SPRINGER
REFERENCE

Bertram O. Fraser-Reid
Kuniaki Tatsuta
Joachim Thiem
Editors

VOLUME 1

Glycoscience

2nd Edition

Chemistry and Chemical Biology

Springer

Glycoscience

Chemistry and Chemical Biology

Bertram O. Fraser-Reid • Kuniaki Tatsuta • Joachim Thiem (Eds.)

Glycoscience

Chemistry and Chemical Biology

Section Editors: G. Coté, S. Flitsch, Y. Ito, H. Kondo, S.-I. Nishimura, B. Yu

With contributions by numerous experts

Second Edition

With 638 Figures and 159 Tables

Springer

Prof. Dr. Bertram O. Fraser-Reid
Natural Products and Glycotechnology
Research Institute Inc.
North Caroline State University
4118 Swarthmore Road
Durham, NC 27707
USA
dglucose@aol.com

Prof. Dr. Kuniaki Tatsuta
University Graduate School
of Science and Engineering
3-4-1 Ohkubo, Shinjuku
Tokyo
Japan
tatsuta@waseda.jp

Prof. Dr. Joachim Thiem
Faculty of Science
Department of Chemistry
University of Hamburg
Martin-Luther-King-Platz 6
20146 Hamburg
Germany
thiem@chemie.uni-hamburg.de

Library of Congress Control Number: 2008921863

ISBN: 978-3-540-30429-6

This publication is available also as:
Print publication under ISBN: 978-3-540-36154-1 and
Print and electronic bundle under ISBN: 978-0-387-36157-2
DOI: 10-1007/978-3-540-30429-6

Springer is part of Springer Science+Business Media

springer.com

Editor: Marion Hertel, Heidelberg, Germany
Development Editor: Sylvia Blago, Heidelberg, Germany
Typesetting and Production: le-tex Jelonek, Schmidt & Vöckler GbR, Leipzig, Germany
Cover Design: Frido Steinen-Broo, Girona, Spain

Printed on acid-free paper SPIN: 11539346 2109 letex — 5 4 3 2 1 0

Preface to the Second Edition

All of the commitments that we expressed in the preface of the 2001 first edition still hold true. The fact that we now introduce a second edition, seven years later, reflects the burgeoning research developments in glycoscience, and the broad impact these developments are having on contemporary biological science. The term "Glycoscience", selected first for a comprehensive collection of reviews a decade ago in Topics in Current Chemistry, has become the generally accepted language to describe the interests of scientists at all levels who are involved in research associated with chemistry, chemical biology and the biology of saccharides.
Readers of Glycoscience 2nd Edition will immediately notice that the editorial board has been supplemented with six younger colleagues of top scientific rank in the glycoscience area. We are happy to have been able to convince:

G. Coté, National Center for Agricultural Utilization and Research, Agricultural Research Service, Peoria, USA;
S. Flitsch, University of Manchester, UK;
Y. Ito, Institute of Physical and Chemical Research (RIKEN), Saitama, Japan;
H. Kondo, Shionogi & Co., Ltd., Osaka, Japan;
S.-I. Nishimura, Hokkaido University, Sapporo, Japan;
B. Yu, Shanghai Institute of Organic Chemistry, Chinese Academy of Sciences, Shanghai, China

to serve as Section Editors.
Their dedicated work in modifying and extending the coverage of the first edition, while maintaining its focus, has resulted in a contemporary appearance. This is particularly true of chapters which cover the bridges that link chemistry to biological chemistry to medicinal chemistry.
On behalf of the Editors and the Section Editors it is our pleasure to extend our gratitude to our dedicated expert colleagues, for their prodigious and scholarly work in updating the chapters. Finally, we are very pleased to note the constant and dedicated support and productive collaboration that we enjoyed with Drs. Marion Hertel and Sylvia Blago, and their highly dedicated editorial staff from Springer, Heidelberg.

January 2008

Bertram O. Fraser-Reid
Kuniaki Tatsuta
Joachim Thiem

Preface to the First Edition

Attempts at understanding of the structure and function of sugars date back to the work of eminent scientists active at the turn of the 19th/20th century. This field of research has developed into today's flourishing area of chemistry and biochemistry of carbohydrates. Over the years major scientific contributions have been associated with this research area and the results have had a strong impact on general chemistry.

At the end of the last century the significance of glycoscience, covering all aspects of the chemistry and chemical biology of carbohydrates and glycoconjugates, increased enormously owing to the challenge of interpreting complex natural processes at the molecular level.

Thus, the purpose of this comprehensive approach is to address and assist a broad readership ranging from graduate students to research scientists involved in glycochemistry and glycobiology, working in areas such as organic chemistry, biochemistry, molecular biology, immunology, and microbiology, as well as pharmaceutical, medicinal, agricultural, food chemistry, etc. The special design of this timely and major modern work on glycoscience is intended to go beyond the collection of specialized reviews. Accordingly, the first general chapters are focussed on principles and methods, giving detailed insight into the current status of structure analysis, synthesis and mechanistic interpretation. Subsequent chapters deal with mono-, oligo-, polysaccharides as well as with the chemistry and chemical biology of glycolipids, glycoproteins, other glycoconjugates and glycomimetics, and they always involve in depth discussion of the biochemical, biological and biomedical background. Because of this presentation, the impact of this collection should be attractive for both specialists, and non-specialists, as well as for newcomers to glycoscience from borderline research areas.

For the editors the task to put together such a comprehensive treatise was feasible only after they could convince the leaders in the fields to contribute and write their chapters with absolute authority. If this work does fulfil the desired purpose of becoming a useful, comprehensive collection in the area of chemistry and chemical biology of glycoscience, many thanks will be due to all of the expert authors.

July 2001

Bertram O. Fraser-Reid
Kuniaki Tatsuta
Joachim Thiem

Table of Contents

Editors-in-Chief

Prof. Dr. Bertram O. Fraser-Reid
Natural Products and Glycotechnology
Research Institute Inc.
North Caroline State University
4118 Swarthmore Road
Durham, NC 27707
USA
dglucose@aol.com

Prof. Dr. Kuniaki Tatsuta
University Graduate School
of Science and Engineering
3-4-1 Ohkubo, Shinjuku
Tokyo
tatsuta@waseda.jp

Prof. Dr. Joachim Thiem
Faculty of Science
Department of Chemistry
University of Hamburg
Martin-Luther-King-Platz 6
20146 Hamburg
Germany
thiem@chemie.uni-hamburg.de

Section Editors

Gregory L. Coté
National Center for Agricultural Utilization
U.S. Department Agriculture
Bioproducts Research Unit
Peoria, IL, USA
greg.cote@ars.usda.gov

Sections:
1. General Principles
4. Monosaccharides
5. Oligosaccharides
6. Complex Polysaccharides

Sabine Flitsch
School of Chemistry
University of Manchester
Manchester, UK
Sabine.Flitsch@manchester.ac.uk

Sections:
11. Biosynthesis and Degradation
12. Glycomedicine

Yukishige Ito
Institute of Physical and Chemical Research (RIKEN)
Saitama, Japan
yukito@riken.jp

Sections:
3. Chemical Glycosylation Reactions
5. Oligosaccharides

Hirosato Kondo
Discovery Research Laboratories
Shionogi & Col, Ltd.
Osaka, Japan
hirosato.kondou@shionogi.co.jp

Sections:
9. Glycomimetics
10. Key Technologies and Tools for Functional Glycobiology

Shin-Ichiro Nishimura
Laboratory of Advanced Chemical Biology
Graduate School of Advanced Life Science
Hokkaido University
Sapporo, Japan
shin@glyco.sci.hokudai.ac.jp

Sections:
7. Glycolipids
8. Glycoproteins

Biao Yu
Chinese Academy of Sciences
Shanghai Institute of Organic Chemistry
Shanghai, China
byu@mail.sioc.ac.cn

Sections:
2. General Synthetic Methods
4. Monosaccharides

List of Contributors

Adachi, Masaatsu
Graduate School of Bioagricultural Sciences
Nagoya University
464-8601 Nagoya
Japan
madachi@agr.nagoya-u.ac.jp

Aich, Udayanath
Department of Biomedical Engineering
The Johns Hopkins University
Baltimore, MD 21218
USA

Allscher, Thorsten
Department Chemie und Biochemie
Ludwig-Maximilians-Universität München
Butenandtstr. 9
81377 München
Germany

Anderson, Kevin
Division of Medicinal and Natural Products
Chemistry, College of Pharmacy
University of Iowa
115 South Grand Avenue
Iowa City, IA 52242-1112
USA

Ando, Hiromune
Division of Instrumental Analysis
Life Science Research Center
Gifu University
Gifu-shi
Gifu 501–1193
Japan
hando@gifu-u.ac.jp

Ariga, Toshio
Institute of Molecular Medicine and
Genetics, Institute of Neuroscience
Medical College of Georgia
Augusta, GA 30912–2697
USA

Asano, Naoki
Faculty of Pharmaceutical Science
Hokuriku University
920-1181 Kanazawa
Japan
n-asano@hokuriku-u.ac.jp

Balbin, Yury Valdés
Center for Synthetic Antigens
Faculty of Chemistry
University of Havana
Ciudad Habana
Cuba 10400
yury@fq.uh.cu

Beau, Jean-Marie
Laboratoire de Synthèse de Biomolécules
Institut de Chimie Moléculaire
et des Matériaux associé au CNRS
Université Paris-Sud
91405 Orsay Cedex
France
jmbeau@icmo.u-psud.fr

BeMiller, James N.
Department of Food Science
Whistler Center for Carbohydrate Research
Purdue University
West Lafayette, IN 47909–2009
USA
bemiller@purdue.edu

Bennett, Clay S.
Chemistry Department
The Scripps Research Institute
La Jolla, CA 92037
USA

Bezouska, Karel
Department of Biochemistry
Faculty of Science
Charles University Prague
Hlavova 8
12840 Praha 2
Czech Republic
bezouska@biomed.cas.cz

Bishop, Lee
Department of Chemistry
Wayne State University
5101 Cass Avenue
Detroit, MI 48202
USA

Blixt, Ola
Department of Cellular and Molecular
Medicine, Faculty of Health Sciences
University of Copenhagen
2200 Copenhagen N
Denmark
olablixt@imbg.ku.dk

Breton, Christelle
Molecular Glycobiology
CERMAV-CNRS
(affiliated with Université Joseph Fourier)
Grenoble 38041
France
christelle.breton@cermav.cnrs.fr

Brown, Stephanie
Department of Microbiology
and Molecular Genetics
Harvard Medical School
Boston, MA 2115
USA
Department of Chemistry
and Chemical Biology
Harvard University
Cambridge, MA 02138
USA

Calderón, Janoi Chang
Center for Synthetic Antigens
Faculty of Chemistry
University of Havana
Ciudad Habana
Cuba 10400
janoi@fq.uh.cu

Ceroni, Alessio
Imperial College London
Division of Molecular Biosciences
Biopolymer Mass Spectrometry Group
London SW7 2AZ
UK
a.ceroni@imperial.ac.uk

Chen, Qi
The State Key Laboratory of Environmental
Chemistry and Ecotoxicology
Research Center for Eco-Environmental
Sciences
Chinese Academy of Sciences
Beijing 100085
China

Collins, Michelle E.
Department of Food Biosciences
The University of Reading
Reading RG6 6AP
UK
m.e.collins@reading.ac.uk

Dell, Anne
Biopolymer Mass Spectrometry Group
Division of Molecular Biosciences
Imperial College London
SW7 2AZ London
UK
a.dell@imperial.ac.uk

Doisneau, Gilles
Laboratoire de Synthèse de Biomolécules
Institut de Chimie Moléculaire
et des Matériaux associé au CNRS
Université Paris-Sud
91405 Orsay Cedex
France
gdoisneau@icmo.u-psud.fr

Du, Yuguo
The State Key Laboratory of Environmental Chemistry and Ecotoxicology
Research Center for Eco-Environmental Sciences
Chinese Academy of Sciences
Beijing 100085
China
duyuguo@rcees.ac.cn

Eggleston, Gillian
United States Department of Agriculture
SRRC-ARS-USDA
1100 Robert E. Lee Boulevard
New Orleans, LA 70124
USA
gillian@srrc.ars.usda.gov

Evans, Philip G.
School of Pharmacy
University of Reading
Whiteknights
Reading, Berkshire RG6 6AD
UK

Evers, David
Division of Medicinal and Natural Products Chemistry, College of Pharmacy
University of Iowa
115 South Grand Avenue
Iowa City, IA 52242-1112
USA

Feizi, Ten
Glycosciences Laboratory
Faculty of Medicine
Imperial College London
Northwick Park and St Mark's Campus
Harrow, Middlesex HA1 3UJ
UK
t.feizi@imeprial.ac.uk

Fernandes, Daryl
Ludger Ltd.
Culham Science Centre
Abingdon, Oxfordshire OX14 3EB
UK
daryl.fernandes@ludger.com

Fokt, Izabela
M. D. Anderson Cancer Center
The University of Texas
Houston, TX 77030
USA

Friedman, Robert B.
Friedmann Associates
6654 North Mozart Street
Chicago, IL 60645
USA
bobbf@juno.com

Fujimoto, Yukari
Department of Chemistry
Graduate School of Science
Osaka University
Toyonaka, Osaka 560–0043
Japan

Fukase, Koichi
Department of Chemistry
Graduate School of Science
Osaka University
Osaka 560-0043
Japan
koichi@chem.sci.osaka-u.ac.jp

Ginsberg, Cynthia
Department of Microbiology and Molecular Genetics
Harvard Medical School
Boston, MA 2115
USA
Department of Chemistry and Chemical Biology
Harvard University
Cambridge, MA 02138
USA

Glasser, Wolfgang G.
Dept. Wood Science and Forest Products
Virginia Tech
Blacksburg, VA 24061
USA
wglasser@vt.edu

Gómez, Ana M.
Instituto de Química Orgánica General
(CSIC)
28006 Madrid
Spain
iqog106@iqog.csic.es

Grindley, T. Bruce
Department of Chemistry
Dalhousie University
Halifax, NS B3H 4J3
Canada
Bruce.Grindley@Dal.Ca

Gruber, Todd D.
Department of Chemistry
University of Wisconsin
1101 University Avenue
Madison, WI 53706
USA

Grynkiewicz, Grzegorz
Pharmaceutical Research Institute
Rydgiera 8
01-793 Warsaw
Poland
g.grynkiewicz@ifarm.waw.pl

Guo, Zhongwu
Department of Chemistry
Wayne State University
5101 Cass Avenue
Detroit, MI 48202
USA
zwguo@chem.wayne.edu

Haslam, Stuart M.
Biopolymer Mass Spectrometry Group
Division of Molecular Biosciences
Imperial College London
SW7 2AZ London
UK

Holst, Otto
Structural Biochemistry
Research Center Borstel
23485 Borstel
Germany
oholst@fz-borstel.de

Icart, Luis Peña
Center for Synthetic Antigens
Faculty of Chemistry
University of Havana
Ciudad Habana
Cuba 10400
luisp@fq.uh.cu

Imberty, Anne
Molecular Glycobiology
CERMAV-CNRS
(affiliated with Université Joseph Fourier)
Grenoble 38041
France
anne.imberty@cermav.cnrs.fr

Ishiwata, Akihiro
RIKEN (The Institute of Physical
and Chemical Research)
Saitama 351–0198
Japan

Isobe, Minoru
Graduate School of Bioagricultural Sciences
Nagoya University
464-8601 Nagoya
Japan
isobem@agr.nagoya-u.ac.jp

Ito, Yukishige
RIKEN (The Institute of Physical
and Chemical Research)
Saitama 351–0198
Japan
yukito@riken.jp

Jang-Lee, Jihye
Biopolymer Mass Spectrometry Group
Division of Molecular Biosciences
Imperial College London
SW7 2AZ London
UK

Jarosz, Sławomir
Institute of Organic Chemistry
Polish Academy of Sciences
01–224 Warsaw
Poland
sljar@icho.edu.pl

Jeon, Heung Bae
Center for Bioactive Molecular Hybrids
and Department of Chemistry
Yonsei University
Seoul 120–749
Korea

Joshi, Hiren J.
German Cancer Research Center
Central Spectroscopy B090
69120 Heidelberg
Germany
h.joshi@dkfz.de

Kiessling, Laura L.
Department of Chemistry
University of Wisconsin
1101 University Avenue
Madison, WI 53706
USA
kiessling@chem.wisc.edu

Kim, Kwan Soo
Center for Bioactive Molecular Hybrids
and Department of Chemistry
Yonsei University
Seoul 120–749
Korea
kwan@yonsei.ac.kr

Kiso, Makoto
Department of Applied
Bioorganic Chemistry
Faculty of Applied Biological Sciences
Gifu University
Gifu-shi
Gifu 501–1193
Japan
kiso@cc.gifu-u.ac.jp

Klüfers, Peter
Department Chemie und Biochemie
Ludwig-Maximilians-Universität München
Butenandtstr. 9
81377 München
Germany
kluef@cup.uni-muenchen.de

Kobayashi, Yoshiyuki
Daiichi Sankyo Research Institute
4250 Executive Square
La Jolla, California 92037
USA
ykobayashi@daiichisankyo-us.com

Koeller, Kathryn M.
Chemistry Department
The Scripps Research Institute
La Jolla, CA 92037
USA

Kondo, Hirosato
Discovery Research Laboratories
Shionogi & Co., Ltd.
12-4, Sagisu 5-chome
Fukushima-ku 553-0002 Osaka
Japan
hirosato.kondou@shionogi.co.jp

Křen, Vladimír
Institute of Microbiology
Academy of Sciences of the Czech Republic
Centre of Biocatalysis and Biotransformation
Vídeňská 1083
142 20 Prague
Czech Republic
kren@biomed.cas.cz

Lieth, Claus-W. von der
(deceased)

Lindhorst, Thisbe K.
Otto Diels-Institut für Organische Chemie
Christian-Albrechts-Universität zu Kiel
24098 Kiel
Germany
tklind@oc.uni-kiel.de

Liu, Jun
The State Key Laboratory of Environmental Chemistry and Ecotoxicology
Research Center for Eco-Environmental Sciences
Chinese Academy of Sciences
Beijing 100085
China

Liu, Yan
Glycosciences Laboratory
Faculty of Medicine
Imperial College London
Northwick Park and St Mark's Campus
Harrow, Middlesex HA1 3UJ
UK

López, J. Cristóbal
Instituto de Química Orgánica General
CSIC
Juan de la Cierva 3
28006 Madrid
Spain
clopez@iqog.csic.es

Lucas, Susana Dias
Pharmaceutical Research
Discovery Chemistry
F. Hoffmann-La Roche Ltd.
4070 Basel
Switzerland

Maaß, Kai
Institute of Biochemistry
Faculty of Medicine
University of Giessen
35392 Giessen
Germany
Kai.maass@biochemie.med.uni-giessen.de

Macmillan, Derek
Department of Chemistry
University College London
20 Gordon Street
WC1H 0AJ London
UK
d.macmillan@ucl.ac.uk

Madsen, Robert
Department of Chemistry
Center for Sustainable and Green Chemistry
Technical University of Denmark
Lyngby 2800
Denmark
rm@kemi.dtu.dk

Mayer, Peter
Department Chemie und Biochemie
Ludwig-Maximilians-Universität München
Butenandtstr. 9
81377 München
Germany

McGill, Nathan W.
School of Chemistry
The University of Melbourne
Parkville, VIC 3052
Australia
n.mcgill@pgrad.unimelb.edu.au

Mori, Yasutaka
Department of Chemistry
Graduate School of Science
Osaka University
Toyonaka, Osaka 560–0043
Japan

Morris, Howard R.
Biopolymer Mass Spectrometry Group
Division of Molecular Biosciences
Imperial College London
SW7 2AZ London
UK

Mortell, Kathleen H.
Department of Chemistry
University of Wisconsin
1101 University Avenue
Madison, WI 53706
USA

Murphy, Paul V.
Centre for Synthesis and Chemical Biology
School of Chemistry
and Chemical Biology
University College Dublin
Belfield
Dublin 4
Ireland
paul.v.murphy@ucd.ie

Nakata, Masaya
Department of Applied Chemistry
Faculty of Science and Technology
Keio University
Yokohama 223-8522
Japan
msynktxa@applc.keio.ac.jp

Nishikawa, Toshio
Graduate School of Bioagricultural Sciences
Nagoya University
464-8601 Nagoya
Japan
nisikawa@agr.nagoya-u.ac.jp

Nishimura, Shin-Ichiro
Laboratory of Advanced Chemical Biology
Graduate School of Advanced Life Science
Frontier Research Center for the
Post-Genome Science and Technology
Hokkaido University and Drug-Seeds
Discovery Research Laboratory
National Institute of Advanced Industrial
Science and Technology
Sapporo
Japan
shin@glyco.sci.hokudai.ac.jp

Nowogródzki, Marcin
Institute of Organic Chemistry
Polish Academy of Sciences
01–224 Warsaw
Poland

Osborn, Helen M. I.
School of Pharmacy
University of Reading
Whiteknights
Reading, Berkshire RG6 6AD
UK
h.m.i.osborn@reading.ac.uk

Oscarson, Stefan
Department of Organic Chemistry
Arrhenius Laboratory
Stockholm University
Stockholm 106 91
Sweden
stefan.oscarson@ucd.ie

Pang, Poh-Choo
Biopolymer Mass Spectrometry Group
Division of Molecular Biosciences
Imperial College London
SW7 2AZ London
UK

Panico, Maria
Biopolymer Mass Spectrometry Group
Division of Molecular Biosciences
Imperial College London
SW7 2AZ London
UK

Parry, Simon
Biopolymer Mass Spectrometry Group
Division of Molecular Biosciences
Imperial College London
SW7 2AZ London
UK

Payne, Richard J.
School of Chemistry
The University of Sydney
Sydney, NSW
Australia

Polt, Robin
University of Arizona
Tucson, AZ 85721-0041
USA
polt@u.arizona.edu

Priebe, Waldemar
M. D. Anderson Cancer Center
The University of Texas
Houston, TX 77030
USA
wpriebe@mdanderson.org, wp@wt.net

Ranzinger, René
German Cancer Research Center
Central Spectroscopy B090
69120 Heidelberg
Germany
r.ranzinger@dkfz.de

Rastall, Robert A.
Department of Food Biosciences
The University of Reading
Reading RG6 6AP
UK
r.a.rastall@reading.ac.uk

Razi, Nahid
Department of Molecular Biology
Glycan Array Synthesis Core D
Consortium for Functional Glycomics
The Scripps Research Institute
La Jolla, CA 92037
USA
nrazi@scripps.edu

Rice, Kevin G.
Division of Medicinal and Natural Products
Chemistry, College of Pharmacy
University of Iowa
115 South Grand Avenue
Iowa City, IA 52242-1112
USA
kevin-rice@uiowa.edu

Richardson, Jonathan P.
Department of Chemistry
University College London
20 Gordon Street
WC1H 0AJ London
UK

Robina, Inmaculada
Departamento de Química Orgánica
Universidad de Sevilla
41071 Sevilla
Spain
robina@us.es

Robyt, John F.
Laboratory of Carbohydrate Chemistry and
Enzymology, Department of Biochemistry,
Biophysics, and Molecular Biology
Iowa State University
Ames, IA 50011
USA
jrobyt@iastate.edu

Sadamoto, Reiko
Graduate School of Advanced Life Science
Hokkaido University
Sapporo 001–0021
Japan
reikosd@glyco.sci.hokudai.ac.jp

Santana, Violeta Fernández
Center for Synthetic Antigens
Faculty of Chemistry
University of Havana
Ciudad Habana
Cuba 10400
violeta@fq.uh.cu

Schmidt, Richard R.
Fachbereich Chemie, Fach M 725
Universität Konstanz
78457 Konstanz
Germany
richard.schmidt@uni-konstanz.de

Spencer, Daniel
Ludger Ltd.
Culham Science Centre
Abingdon, Oxfordshire OX14 3EB
UK
daniel.spencer@ludger.com

Steiner, Andreas
Glycogroup, Institut für Organische Chemie
Technische Universität Graz
8010 Graz
Austria

Stone, Bruce A.
Department of Biochemistry
La Trobe University
Bundoora, VIC 3083
Australia
b.stone@latrobe.edu.au

Stütz, Arnold
Glycogroup, Institut für Organische Chemie
Technische Universität Graz
8010 Graz
Austria
stuetz@tugraz.at

Suda, Yasuo
Department of Nanostructure
and Advanced Materials
Graduate School of Science and Engineering
Kagoshima University
890-0065 Kagoshima
Japan
ysuda@eng.kagoshima-u.ac.jp

Sutton-Smith, Mark
Biopolymer Mass Spectrometry Group
Division of Molecular Biosciences
Imperial College London
SW7 2AZ London
UK

Svensson, Birte
Department of Chemistry
Carlsberg Laboratory
2500 Valby
Denmark

Tadano, Kin-ichi
Department of Applied Chemistry
Keio University
Hiyoshi, Kohoku-ku
223-8522 Yokohama
Japan
tadano@applc.keio.ac.jp

Tanaka, Katsunori
Department of Chemistry
Graduate School of Science
Osaka University
Osaka 560-0043
Japan

Tanaka, Shin-ichi
Department of Chemistry
Graduate School of Science
Osaka University
Toyonaka, Osaka 560–0043
Japan

Thiem, Joachim
Department of Chemistry
Faculty of Science
University of Hamburg
20146 Hamburg
Germany
joachim.thiem@chemie.uni-hamburg.de

Thimm, Julian
Department of Chemistry
Faculty of Science
University of Hamburg
20146 Hamburg
Germany

Tissot, Berangere
Biopolymer Mass Spectrometry Group
Division of Molecular Biosciences
Imperial College London
SW7 2AZ London
UK

Toshima, Kazunobu
Department of Applied Chemistry
Faculty of Science and Technology
Keio University
Yokohama 223–8522
Japan
toshima@applc.keio.ac.jp

Totani, Kiichiro
RIKEN (The Institute of Physical
and Chemical Research)
2-1 Hirosawa, Wako-shi
351-0198 Saitama
Japan
totani@riken.jp

Tvaroška, Igor
Institute of Chemistry, Centre for Glycomics
Slovak Academy of Sciences
845 38 Bratislava
Slovak Republic
chemitsa@savba.sk

Urban, Dominique
Laboratoire de Synthèse de Biomolécules
Institut de Chimie Moléculaire
et des Matériaux associé au CNRS
Université Paris-Sud
91405 Orsay Cedex
France
domurban@icmo.u-psud.fr

Vauzeilles, Boris
Laboratoire de Synthèse de Biomolécules
Institut de Chimie Moléculaire
et des Matériaux associé au CNRS
Université Paris-Sud
91405 Orsay Cedex
France
bvauzeil@icmo.u-psud.fr

Velasco-Torrijos, Trinidad
Centre for Synthesis and Chemical Biology
School of Chemistry
and Chemical Biology
University College Dublin
Belfield
Dublin 4
Ireland
Current address:
Department of Chemistry
National University of Ireland Maynooth
Co. Kildare
Ireland
trinidad.velascotorrijos@nuim.ie

Verez-Bencomo, Vicente
Center for Synthetic Antigens
Faculty of Chemistry
University of Havana
Ciudad Habana
Cuba 10400
vicente@fq.uh.cu

Vogel, Pierre
Laboratoire de glycochimie
et de synthèse asymétrique
Ecole Polytechnique Fédérale de Lausanne
(EPFL), BCH
1015 Lausanne-Dorigny
Switzerland
pierre.vogel@epfl.ch

Wakao, Masahiro
Department of Nanostructure
and Advanced Materials
Graduate School of Science and Engineering
Kagoshima University
890-0065 Kagoshima
Japan
wakao@eng.kagoshima-u.ac.jp

Walker, Suzanne
Department of Microbiology
and Molecular Genetics
Harvard Medical School
Boston, MA 2115
USA
Department of Chemistry
and Chemical Biology
Harvard University
Cambridge, MA 02138
USA
suzanne_walker@hms.harvard.edu

Wang, Yuhang
The State Key Laboratory of Natural
and Biomimetic Drugs
School of Pharmaceutical Sciences
Peking University
Xue Yuan Rd #38
Beijing 100083
China

Wessel, Hans Peter
Pharmaceutical Research
Discovery Chemistry
F. Hoffmann-La Roche Ltd.
4070 Basel
Switzerland
hans_p.wessel@roche.com

Williams, Spencer J.
School of Chemistry
The University of Melbourne
Parkville, VIC 3052
Australia
sjwill@unimelb.edu.au

Wilson, Iain B. H.
Department für Chemie
Universität für Bodenkultur
(University of Natural Resources and
Applied Life Sciences)
Muthgasse 18
1190 Wien
Austria
iain.wilson@boku.ac.at

Witczak, Zbigniew J.
Department of Pharmaceutical Sciences
Nesbitt School of Pharmacy
Wilkes University
Wilkes-Barre, PA 18766
USA
zbigniew.witczak@wilkes.edu

Wittmann, Valentin
Fachbereich Chemie
Universität Konstanz
78457 Konstanz
Germany
mail@valentin-wittmann.de

Wong, Chi-Huey
Chemistry Department
The Scripps Research Institute
La Jolla, CA 92037
USA
wong@scripps.edu

Wrodnigg, Tanja
Glycogroup, Institut für Organische Chemie
Technische Universität Graz
8010 Graz
Austria

Yanagisawa, Makoto
Institute of Molecular Medicine and
Genetics, Institute of Neuroscience
Medical College of Georgia
Augusta, GA 30912–2697
USA

Yarema, Kevin J.
Department of Biomedical Engineering
The Johns Hopkins University
Baltimore, MD 21218
USA
kyarema1@jhu.edu

Ye, Xin-Shan
The State Key Laboratory of Natural and Biomimetic Drugs
School of Pharmaceutical Sciences
Peking University
Xue Yuan Rd #38
Beijing 100083
China
xinshan@bjmu.edu.cn

Young, Travis
Department of Chemistry
University of Wisconsin
1101 University Avenue
Madison, WI 53706
USA

Yu, Robert K.
Institute of Molecular Medicine and Genetics, Institute of Neuroscience
Medical College of Georgia
Augusta, GA 30912–2697
USA
ryu@mcg.edu

Zeng, Guichao
Institute of Molecular Medicine and Genetics, Institute of Neuroscience
Medical College of Georgia
Augusta, GA 30912–2697
USA

Zhang, Jianbo
Department of Chemistry
East China Normal University
200062 Shanghai
China
jbzhang@chem.ecnu.edu.cn

Zhu, Xiangming
School of Chemistry
and Chemical Biology
University College Dublin
Belfield
Dublin 4
Ireland
xiangming@ucd.ie

Part 1
General Principles

1.1 Structure and Conformation of Carbohydrates

T. Bruce Grindley
Department of Chemistry, Dalhousie University,
Halifax, NS B3H 4J3, Canada
Bruce.Grindley@Dal.Ca

Abstract

The conformational analysis of monosaccharides, disaccharides, and oligosaccharides is reviewed. Conformational terms are introduced through examination of the conformations of cyclohexane and cyclopentane then applied to the pyranose, furanose, and septanose rings. Concepts such as the anomeric effect are discussed. Topics of current interest, such as hydroxymethyl group and hydroxyl group rotation and disaccharide conformations are summarized. Physical methods for studying conformation are outlined.

In: *Glycoscience*. Fraser-Reid B, Tatsuta K, Thiem J (eds)
Chapter-DOI 10-1007/978-3-540-30429-6_1:

Keywords

Conformation; Chair; Anomeric effect; Hydroxymethyl; Monosaccharides; Disaccharides; Oligosaccharides; Coupling constants; Molecular mechanics; Molecular dynamics

Abbreviations

A-value	free energy cost for a substituent on a chair conformation of a cyclohexane ring to change from an equatorial to axial orientation
B	boat
B3LYP	the Becke-3 Lee–Yang–Parr density functional
C	chair
CD	circular dichroism
DFT	density functional theory
E	envelope conformation of a five-membered ring
G	Gaussian
gg*, *gt*, *tg	conformations of hydroxymethyl or other side-chains on pyranose or furanose rings
GVB	generalized valence bond
H	half chair
HOMO	highest energy occupied molecular orbital
KDO	3-deoxy-D-*manno*-2-octulosonic acid
MP	Moeller–Plesset
NANA	5-acetamido-3,5-dideoxy-D-glycero-D-galacto-2-nonulosonic acid
nOe	nuclear Overhauser enhancement
ORD	optical rotatory dispersion
ROA	resonant Raman optical activity
S	skew conformation of a six-membered ring
T	twist conformation of a five-membered ring
TB	twist boat
TC	twist chair

1 Conformational Analysis

1.1 Introduction

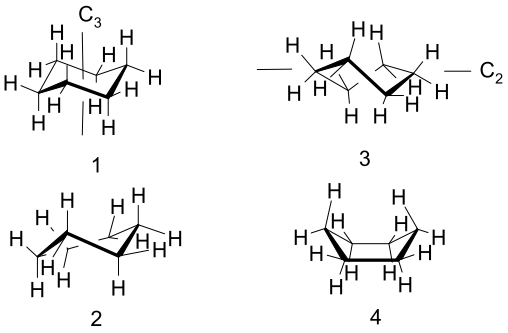
The shapes or conformations of carbohydrates strongly influence both their reactivity and their physical properties. Several reviews of various aspects of the conformational analysis of carbohydrates have been published [1,2,3,4,5,6,7].

Monosaccharides exist chiefly in the six-membered pyranose form and in the five-membered furanose form, both of which adopt puckered shapes to minimize eclipsing interactions. These puckered shapes are most commonly described in terms of the shapes of the symmetrical conformations adopted by the analogous hydrocarbons, cyclohexane and cyclopentane [8,9,10,11], but may also be described in terms of puckering parameters [12,13,14]. Minima on the conformational potential energy surface are called conformers.

1.2 Conformations of Cycloalkanes and Heterocycles

1.2.1 Conformations of Cyclohexanes

The most stable conformer of cyclohexane is the D_{3d} symmetric chair (C) (**1**) conformation, in which orientations parallel to the C_3 axis are termed axial while those roughly perpendicular to the axis are termed equatorial. Chair conformers invert or undergo ring reversal via the C_2 symmetric half-chair (H) saddle point (**2**) and intermediate D_2 twist-boat, or, as used for carbohydrates, skew (S) conformer (**3**). The barrier to ring reversal ($\Delta G^{\neq}$) is about 43 kJ mol^{-1}; values of $\Delta H^{\neq}$ and $\Delta S^{\neq}$ obtained for cyclohexane-d_{11} [15] and cyclohexane-1,1,2,2,3,3,4,4-d_8 [16] were 44.8 and 48.2 kJ mol^{-1} and 9.2 and 19.2 J mol^{-1} K^{-1}, respectively. The skew conformer is flexible, exchanging atoms above and below the plane defined by the two carbons on the C_2 axis and the two adjacent to one of these, via the C_{2v} symmetric boat (B) (**4**) conformation, a saddle point on the conformational potential energy surface. Rapid cooling of a mixture of cyclohexane and argon from 800 °C to 20 K isolated a conformational mixture containing some skew form. The rate of conversion to the chair was used to give a $\Delta H^{\neq}$ value for this process, which, combined with the known chair to chair barrier, gave a value of 23 kJ mol^{-1} for the stability of the S form relative to that of the C [17].



A substituent on a cyclohexane ring can assume either an equatorial or an axial orientation and the free energy preference for the equatorial conformer is termed the *A* value for the substituent (◉ *Fig. 1*). Tables of *A* values are available [8,18]. The values (at 300 K) most relevant for the current topic are those for methyl (7.31 [19] or 7.61 kJ mol^{-1} [20]), for hydroxymethyl (7.36 kJ mol^{-1} [21]), for hydroxyl (2.5 to 4.6 kJ mol^{-1} depending on sol-

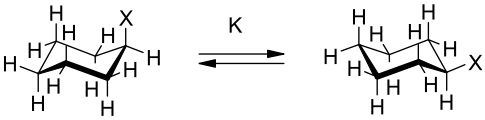


Figure 1
Derivation of *A* values for substituted cyclohexanes: $A = RT \ln K = -\Delta G°$

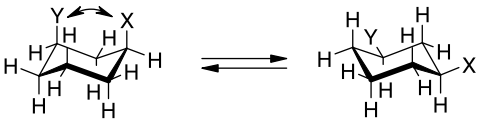
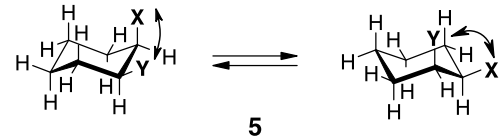


Figure 2
The 1,3-diaxial interaction in *cis*-1,3-disubstituted cyclohexanes

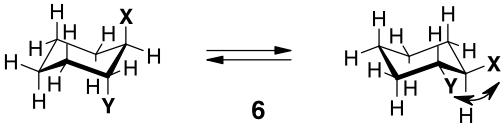
vent [8,22,23]), for methoxyl (2.3–3.1 kJ mol^{-1} [23,24]), for carboxyl (5.7 kJ mol^{-1} [25]), for carboxylate (7.9–8.2 kJ mol^{-1} [25,26]), and for methoxycarbonyl (4.6–5.5 kJ mol^{-1} [24,25,26,27]). On the basis of the lack of change in geometry of the *syn* axial CH bonds on addition of an axial methyl, it was concluded [20] that the cause of the equatorial preference is repulsive steric interactions between the axial group and the ring carbons including the *gauche* torsional interaction rather than the traditional [8] explanation of steric interactions with the 1,3-related axial hydrogens (synaxial interaction). However, see below.

The interactions between substituents are important for the stabilities of the conformations of carbohydrates. Non-geminally disubstituted cyclohexanes can exist as *cis*- and *trans*-isomers. The relative stabilities of *cis*- and *trans*-1,4-disubstituted cyclohexanes or the relative stabilities of the diaxial versus the diequatorial conformers of *trans*-1,4-disubstituted cyclohexanes can be predicted successfully from the *A* values of the two substituents [8]. Another factor arises when *cis*-1,3-disubstituted-1,3-cyclohexanes are considered: the interaction between *syn*-1,3-diaxially related substituents, as shown in *Fig. 2*. Corey and Feiner summarized values for this interaction [28]. For carbohydrates, the most important of those available are: the OH / OH interaction (7.9 kJ mol^{-1} [29]), the CH_3 / OH interaction (7.9–11.3 kJ mol^{-1} [30,31,32]) and the OAc / OAc interaction (8.4 kJ mol^{-1}[33]).

The two chair conformers of *cis*-1,2-disubstituted cyclohexanes both have additional *gauche* interactions between substituents (**5**). One might expect that the stability differences between the diequatorial and diaxial conformers of the *trans*-isomer would be the sum of the *A*-values for the axial substituents minus the value of the *gauche* interaction for the equatorial substituents (**6**). For methyl substituents, where the *gauche* interaction is that of butane, this has been shown to be approximately true [34,35]. Corey and Feiner [28] found that *gauche* interactions between substituents other than methyl could not be estimated by averaging 1/2 the substituent's *A* values but have tabulated these *gauche* interaction values. Those involving polar groups are highly dependent on solvent because the diaxial and diequatorial conformers have very different dipole moments. Direct determination by measurement of the difference in stability between the diequatorial and diaxial conformers requires the subtraction of the substituent *A* values from the total stability difference. These latter values may be uncertain or have large solvent dependencies, leading to large uncertainties in the interaction energies. The most relevant values are the OH:OH interaction (1.5 kJ mol^{-1} in water [29], the OH:OMe interaction (2.7 kJ mol^{-1} in carbon disulfide [36], 1.9 kJ mol^{-1} in pentane [37], 1.9 kJ mol^{-1} in methanol [37]), the OMe:OMe interaction (5.3 kJ mol^{-1} in pentane, 2.3 kJ mol^{-1} in methanol [37]), and the OH:CH_3 interaction (1.6 kJ mol^{-1} [38]).

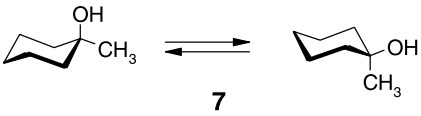


5



6

The relative stabilities of the two chair conformers of 1,1-geminally disubstituted cyclohexanes are additive based on *A* values [39,40] unless the populations of rotamers present for one of the two substituents are different to when in the monosubstituted cyclohexane [8,41] or unless there is differential solvation of one of the substituents in this compound as opposed to the cyclohexane derivative [39]. The latter situation applies with 1-methylcyclohexanol (**7**), where ΔG is 1.3 kJ mol^{-1} in carbon disulfide in favor of the methyl equatorial conformation, whereas subtraction of *A* values gives 3.2 kJ mol^{-1} [39]. In contrast, for 1-methoxy-1-methylcyclohexane, the observed and calculated ΔG values are 3.1 and 3.7 kJ mol^{-1}, respectively [40], a difference that will not have large conformational consequences. The difference between observed and calculated ΔGs for **7** presumably arises because the difference in the aggregation of the equatorial and axial oriented hydroxyl conformers through hydrogen bonding is much larger for the tertiary alcohol **7** than for the corresponding conformers of cyclohexanol. This effect could be important for conformational equilibria of ketoses or branched chain sugars but is probably negligible in hydroxylic solvents.



7

1.2.2 Conformations of Tetrahydropyran Derivatives

The replacement of a CH_2 group by an oxygen atom to give tetrahydropyran or oxane alters the shape of the chair conformation slightly by replacing two 153 pm C–C bonds by two 142 pm C–O bonds [42,43]. This has little effect on the barrier to ring reversal [8,44]. However, the substituent *A*-values now depend on position. The *A*-values for a methyl group at positions 2, 3, and 4 are 12.0, 5.98, and 8.16 kJ mol^{-1}, respectively [45], compared to ~7.5 kJ mol^{-1} for methylcyclohexane. The value at position 2 is significantly increased relative to that for methylcyclohexane because the shorter bond lengths increase steric repulsion, while that at position-3 is decreased because one of the CH_3/H synaxial interactions is replaced by a CH_3/lone pair interaction (*Fig. 3*).

A methoxy group at position-2 of tetrahydropyran is much more stable in the axial orientation than would be expected based on its *A* value, a manifestation of the anomeric effect [46,47,48,49,50,51,52,53], originally observed with carbohydrate derivatives by

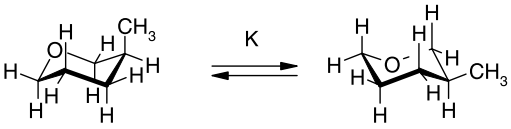


Figure 3
The *A* value for a methyl group at position 3 of tetrahydropyran is less than that of cyclohexane, 6.0 vs 7.5 kJ mol^{-1}

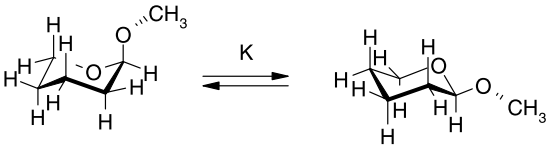


Figure 4
The anomeric effect: the axial conformer is more stable than predicted based on the *A* value of the methoxy group. The exoanomeric effect: the preferred rotamers about the exocyclic C–O bond are those with the methyl *gauche* to the endocyclic oxygen atom and to H-1

Edwards [54] and named by Lemieux [33]. This effect is due to electronic interactions between the exocyclic and endocyclic oxygen atoms and their C–O bonds (see later). It had been recognized that cyclohexane *A* values were not appropriate as reference values for the size of the anomeric effect because of the different geometry in the heterocycle [55]. Franck [56] suggested that the cyclohexane *A* values for polar substituents be scaled to obtain reference heterocyclic *A* values by using a line defined by the relationship of heterocyclic to cyclohexane *A* values for non-polar substituents such as methyl and using hydrogen as the zero point. The resulting equation is $\Delta G(\text{2-thp}) = 1.53 \times \Delta G(\text{Cyhx}) + 0.04\,\text{kJ mol}^{-1}$. This suggestion has now been generally accepted [50,53]. Thus, an observed equatorial to axial free energy difference for a methoxy group at position-2 in tetrahydropyran of -3.3 kJ mol^{-1} in a non-polar solvent [57] and the cyclohexane *A* value of 2.7 kJ mol^{-1} scaled to 4.2 kJ mol^{-1} (see above) gives an anomeric effect for methoxy of 7.5 kJ mol^{-1} in tetrahydropyran (*Fig. 4*).

Two explanations have been advanced to account for the anomeric effect. The first involves the stabilizing effect of bonding interactions between *n* electrons on one oxygen atom and the σ^* orbital of the bond connecting the other oxygen atom and the central anomeric carbon atom [58]. The second involves destabilizing dipole-dipole repulsion between the two oxygen atoms and their lone pairs [54]. The relative importance of these factors has been difficult to establish.

Complicating the discussion are the two different descriptions of the nature of the lone pairs involved. In the traditional view, the lone pairs on oxygen are depicted as sp^3 hybridized. However, these lone pairs would be equivalent for water, which is not compatible with the photoelectron spectra, where the lone pair orbitals are very different in energy [59,60]. In the correct description, the two orbitals differ in their extent of p character; the higher energy orbital of HOMO, n_p is close to being a pure p orbital, while the lower energy orbital, n_σ has much more a s character, resulting in a picture that resembles sp^2 hybridization, with the n_p orbital perpendicular to the plane containing the two substituents bonded to the oxygen

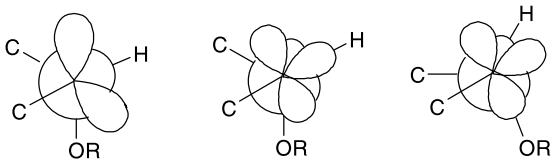
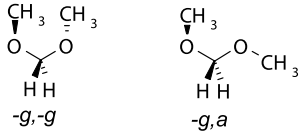


Figure 5
Newman projections of lone pair orbitals on one oxygen atom in a COCOR unit. *Left:* the geometry of an axial OR group in a tetrahydropyran ring with sp^3 orbitals on the endocyclic oxygen. *Center:* the geometry of an axial OR group in a tetrahydropyran ring with n_p and n_σ orbitals on the endocyclic oxygen. *Right:* the geometry has been altered to show the best overlap of the n_p orbital with the σ^* orbital of the OR bond. This results in the COCC torsional angle closing from about staggered to ~30°

atom [48,50,52]. This description has somewhat different stereochemical consequences for $n \rightarrow \sigma^*$ overlap (● *Fig. 5*). In the traditional description, the most favorable overlap from the n orbital of O1 occurs when the C–O1–C–O2 torsional angle is 60°, which matches that present for the axial OR group in a 2-substituted tetrahydropyran. The correct description leads to the most favorable overlap at a torsional angle of 90° where the intra-ring COCC torsional angle has been reduced to 30°. Considerable overlap is still present at the tetrahydropyran intra-ring torsional angle but Dubois et al. have suggested that this type of overlap explains the larger anomeric effect present in furanoses [61]. The recent conclusion that acyclic acetals $RCH(OR_1)(OR_2)$ increasingly adopt conformations with the smaller of R_1 or R_2 eclipsed with the acetal H as the two other R groups increase in size is in accord with this description of the oxygen lone pairs [62].

Evidence for the importance of the first explanation includes changes in bond lengths about the anomeric center. In the axial conformer, a lone pair on the endocyclic oxygen atom is aligned with the exocyclic C–O bond leading to orbital interaction and bond shortening and concomitant bond lengthening of the exocyclic C–O bond, the endo anomeric effect. In both conformers, if the methoxy methyl is *gauche* to the endocyclic C–O bond and *anti* to the C1–C2 bond, a lone pair will interact with the aligned endocyclic C–O bond leading to exocyclic C–O bond shortening. The preference for these conformers is termed the exo anomeric effect. This bond shortening has been clearly observed for the central C–O bonds in the favored *+gauche*,*+gauche* (*+g*,*+g*) or *-gauche*,*-gauche* (*−g*,*−g*) (**8**) conformers of dimethoxymethane, both in the gas phase [63] and recently in the solid [64]. In this conformation, there are two possible $n \rightarrow \sigma^*$ interactions that compete. The C–O bond arrangements in this conformation correspond to those in the axial conformer of 2-methoxytetrahydropyran while those in the *gauche*,*anti* (*g*,*a*) conformer of dimethoxymethane (**9**) are similar to those in the equatorial conformer of 2-methoxytetrahydropyran. Where only one oxygen atom is the source of the n electrons, it is calculated using high level ab initio methods that greater bond shortening of the *a* C–O bond occurs, as in the *g*,*a* conformer [65,66]. These calculations also support the contention that $n \rightarrow \sigma^*$ interactions are very important for the anomeric effect. Similar bond shortening is observed for the appropriate conformers of $ClCH_2OH$ [67]. Statistical analyses of bond length data from X-ray diffraction studies support the bond shortening due to $n \rightarrow \sigma^*$ interactions [61,68,69] and are in agreement with greater bond shortening in the a,p conformer. In addition, electron withdrawing groups in 2-alkoxytetrahydropyrans [70,71] or

2-phenoxytetrahydropyrans [72,73,74], which lower the energy of the σ^* orbital and increase the interaction, were found to increase the preference for the axial conformer. A similar effect was seen in substituted 2,2-diphenyl-1,3-dioxanes where the phenyl rings bearing electron-withdrawing substituents preferred the axial orientation and vice versa [75]. The smaller magnitude of the $^1J_{C,H}$ value at anomeric centers for axial CH bonds than for equatorial CH bonds has been interpreted in terms of $n \rightarrow \sigma^*$ overlap [76,77], but recently it has been shown that the calculated dependence of the size of this value on the HCOC torsional angle is incompatible with this explanation [78]. The effects of replacing hydrogen by deuterium atoms on the positions of anomeric equilibria for the D-glucopyranoses were interpreted in terms of $n \rightarrow \sigma^*$ overlap [79].



The alternative explanation for the anomeric effect involves destabilizing dipole-dipole repulsion between the two oxygen atoms and their lone pairs [33,54] (see *Fig. 6*). This explanation was reexamined by Box [80,81], who suggested that the bond shortening evidence can be explained by n orbital repulsion. Persuasive support for this position has come from ab initio calculations on 2-methoxytetrahydropyran conformers by da Silva and coworkers, performed using the generalized valence bond-perfect wave function at the GVB-PP/6–31G(d,p) level [52]. This approach only uses a localized description of bonding which inherently excludes $n \rightarrow \sigma^*$ overlap. These calculations were able to fully account for the relative conformer energies and bond length shortening and lengthening previously explained by $n \rightarrow \sigma^*$ overlap. The changes in bond lengths with geometry are caused by differences in the % s character in the local orbitals and the contribution of $n \rightarrow \sigma^*$ overlap was considered to be insignificant [52]. In agreement, Perrin et al. showed how dipole-dipole repulsion could also lead to bond length alteration [82]. Wiberg and Marquez demonstrated that there was a significant solvent effect on the axial-equatorial equilibrium of 4,6-dimethyl-2-methoxytetrahydropyran and suggested that the reduction of electrostatic interaction between dipoles with increasing solvent polarity

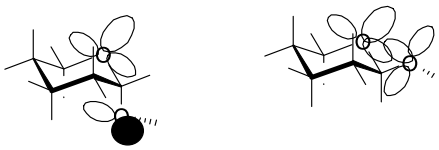


Figure 6
Approximate representation of the geometries of the n_p and n_s orbitals on oxygen atoms at the anomeric center. In the equatorial conformer, the n_s orbital on the exocyclic oxygen atom is aligned with one lobe of the n_p orbital on the endocyclic oxygen atom and one lobe of the n_p orbital on the exocyclic oxygen atom is aligned with the n_p orbital on the endocyclic oxygen atom, leading to repulsive destabilization. In the axial conformer, only one pair of orbitals is aligned, leading to less repulsion

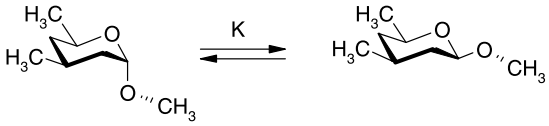


Figure 7
ΔG for the equilibrium ranged from 4.2 kJ/mol in benzene to 1.8 kJ/mol in acetonitrile [83]

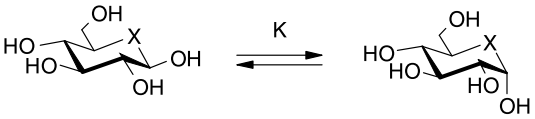
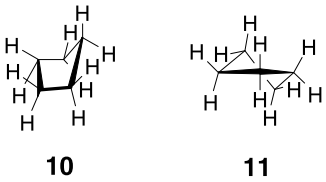


Figure 8
For *X* = O, *K* is 0.56 at 60 °C in water; for *X* = NH, *K* is 1.7 at 50 °C in water [87]

was responsible for this change [83] (*Fig. 7*). Lemieux and coworkers had shown earlier that solvent effects influenced equilibria in 2-alkoxy-substituted tetrahydropyrans [71,84] but they interpreted the solvent dependency in terms of specific solvent molecule interactions. Jorgenson et al. pointed out that the solvent effect reflects the fact that the equatorial conformer has a larger dipole moment and hence is better solvated in more polar solvents [85]. Perrin et al. stated that the fact that the positions of the axial equatorial equilibria were almost identical in 2-methoxy-1,3-dimethylhexahydropyrimidine and 2-methoxy-1,3-dioxane indicated that electrostatic repulsion was dominant in determining the anomeric effect [77]. If $n \rightarrow \sigma^*$ interactions were more important, the nitrogen donor would be expected to have a much larger axial stabilizing effect which was not observed. However, interpretation of the results of this study was complicated by the steric effects of the methyl groups on nitrogen and its conclusion was disputed by Salzner [86]. In addition, for norjirimycin, 5-amino-5-deoxy-D-glucopyranose-glucopyranose, which exists entirely in the piperidine ring form, the anomeric effect is 3.0 kJ mol^{-1} larger than for D-glucopyranose [87], consistent with $n \rightarrow \sigma^*$ interactions (*Fig. 8*). However, on balance, the calculations of da Silva's group indicate that $n \rightarrow \sigma^*$ interactions are of minor importance in explaining the anomeric effect.

1.2.3 Conformations of Cyclopentanes and Tetrahydrofurans

The symmetric puckered conformations of cyclopentane are the C_s symmetric envelope (E) (**10**) with four carbon atoms in a plane and the C_2 symmetric twist (T) (**11**) with three carbon atoms in a plane [88]. Unlike cyclohexane, these conformations are of almost equal energy and are separated by barriers of about RT or less [89]. There are ten envelope conformations, each with one of the five carbon atoms out of the plane in one of the two directions, and ten corresponding twist conformations. The individual conformations freely exchange which atom or atoms are out of the plane, a process termed pseudorotation, and the whole sequence of conformations is called the pseudorotational itinerary (*Fig. 9*).



10 **11**

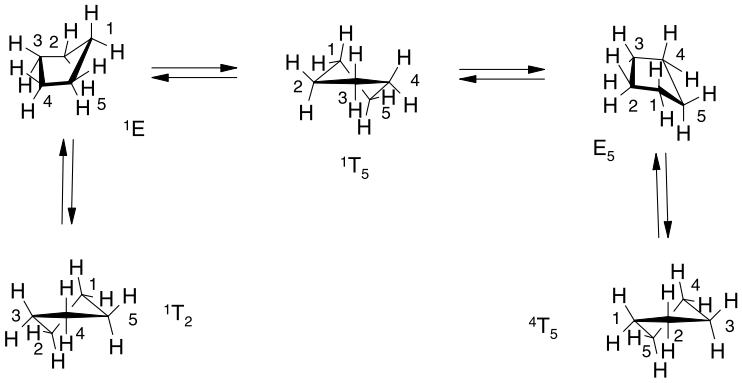
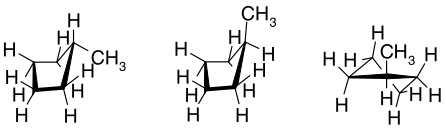


Figure 9
Part of the pseudorotational itinerary of cyclopentane

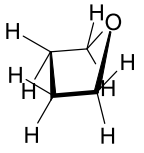
Substituents cause parts of the pseudorotational itinerary to become more populated to avoid eclipsing and 1,3-diaxial interactions. In the envelope conformation, the ring atom out of the plane has pseudoequatorial (**12**) and pseudoaxial positions (**13**) as do the atoms next to it that partially avoid eclipsing interactions. Substituents on the remaining two atoms are eclipsed. In the twist conformation, the atom on the C_2 axis has two identical isoclinal positions that also partially avoid eclipsing interactions (**14**). Thus, the barrier to pseudorotation for methylcyclopentane, which corresponds to having the methyl group in the least stable site, is larger, 14.2 kJ mol^{-1} [88].



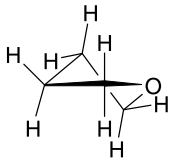
12 **13** **14**

Microwave [90,91,92] and far-IR [93,94] spectra indicate that the barrier to pseudorotation in tetrahydrofuran is small (0.8 to 2.0 kJ mol^{-1}). Modeling the potential energy surface has

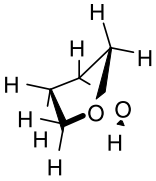
proven to be difficult. The current theoretical view, obtained using calculations at the MP2/cc-pV5ZMP2/cc-pVTZ level, is that there are two minima, the most stable being the C_s symmetric envelope conformer (^{1}E) (**15**), the next the C_2 symmetric twist conformer ($^2T_3$) (**16**), 0.55 kJ mol^{-1} higher, with the saddle point on the pseudorotational itinerary being a C_1 symmetric conformation 0.78 kJ mol^{-1} above the global minimum [95]. Older methods obtain the twist form as more stable [96]. In the solid state, the twist conformer is observed in both 104 and 140 K X-ray [97] and 5 K high-resolution neutron powder [98] diffraction studies. In solid tetrahydrofuran, the pseudorotational motion is a large-amplitude ring deformation vibration with an amplitude of about 140 cm^{-1} [99]. The pseudorotational barrier increases in water by about 1.0 kJ mol^{-1} to about 2.1 ± 0.8 kJ mol^{-1} [100]. When substituents are present, their requirements become more important than the inherent tetrahydrofuran preferences. For instance, the microwave spectrum of 3-hydroxytetrahydrofuran shows that it exists in a ^2E conformation having an axial hydroxyl group hydrogen bonded to the ring oxygen (**17**), with no evidence for pseudorotation both in the gas phase [101] and in aqueous solution [102].



15



16



17

1.3 Conformations of Monosaccharides

1.3.1 Conformations of Acyclic Carbohydrates

In the solid state, most acyclic carbohydrate derivatives adopt a conformation having the carbon atoms in an extended, planar zig-zag arrangement, unless there are parallel, 1,3-steric interactions between oxygen atoms (Hassel–Ottar effect), written in abbreviated form as O//O interactions [103,104,105,106] (*Fig. 10*). In the last few years, a number of examples have been observed where either O//O or CO interactions are present, particularly for compounds with chains longer than five carbons [107,108,109,110,111,112,113,114]. It is clear that the magnitude of the destabilizing effect associated with these interactions is less than in six-membered rings where torsional and bond angle relaxation is more difficult energetically [108,112]. *O*-Acetyl derivatives are more likely to adopt conformations with O//O or CO interactions than unsubstituted compounds [115].

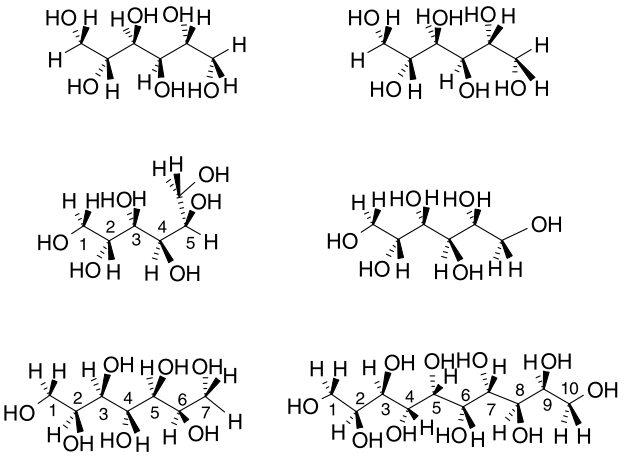


Figure 10
The conformations in the solid state of top left : galactitol, planar zig-zag with extended oxygen atoms [94]; top right, D-mannitol, planar zig-zag with *gauche* oxygen atoms [95]; center left, D-altritol, the $_4G^-$ conformation adopted in the solid-state with a CO interaction [105]; center right, D-altritol, showing the planar zig-zag conformation avoided; bottom left, *meso*-D-*glycero*-L-*altro*-heptitol, with a O//O interaction between O-3 and O-5 [100]; bottom right, L-*galacto*-D-*galacto*-decitol, with a O//O interaction between O-5 and O-7 [98]

In solution, the same considerations apply [116]. The ^{1}H NMR spectra of aqueous solutions of all of the alditols from four to seven carbons have been analyzed [117,118,119] and the coupling constants have been used to determine the most populated conformers, with the assumption that all rotamers have ideal staggered torsional angles. Similar but not as precise information has been obtained from ^{13}C NMR chemical shifts [108,120,121]. It was concluded that the order of the magnitudes of repulsive interactions is the same as for six-membered rings but the actual sizes of the repulsive interactions were smaller. The chain adopts a planar, zig-zag conformation, except where O//O interactions are present, when it twists to replace an O//O by a C//H interaction, or where more than one O//O interactions are present, the chain may twist to give a *gauche* conformation that contains a C//O interaction. Nomenclature for *gauche* conformations was developed by Horton and Wander [116]; a $_2G^-$ conformation is obtained from the planar zig-zag conformation by a 120° clockwise rotation of the remote atom along the C-2–C-3 bond; a $_3G^+$ conformation is obtained from the planar zig-zag conformation by a 120° counterclockwise rotation of the remote atom along the C-3–C-4 bond. Chain preference falls in the order: planar, single twist, double twist. Chain twisting generally results in an oxygen atom extending the chain rather than a hydrogen atom. The preferred rotamer of the hydroxymethyl groups has the oxygen atom extending the chain [119] although significant amounts of all hydroxymethyl rotamers are observed. On the basis of studies of alditols in a variety of solvents [122,123], it was concluded that as solvents change from low polar to protic to very polar aprotic, the preference of C–O bonds in 2,3-butanediol units changes from *gauche* to *anti* [123].

1.3.2 Conformations of Pyranoses

Ring Conformations The two chair conformations of pyranose sugars are named by defining a reference plane that contains four ring atoms and has the lowest numbered carbon atom out of the plane [124]. In the chair, abbreviated *C*, the positions of the atoms above and below the plane are given as superscripts and subscripts, respectively. To remove ambiguity, it is also necessary to stipulate that the atom written as a superscript is on the side of the reference plane from which the numbering of the remaining atoms appears clockwise (*Fig. 11*). The abbreviations for boat, skew or twist-boat, and half-chair conformations are *B*, *S*, and *H*, respectively. Only one reference plane is possible for the boat but the skew is named so that the selected reference plane, which contains three adjacent atoms and one other, has the atom with the lowest possible number exocyclic (*Fig. 12*) [124].

Pyranose derivatives adopt chair conformations unless an unusual combination of destabilizing interactions is present. Angyal developed a set of destabilizing interactions that can be used to estimate the relative stabilities of the two chair conformers in aqueous solution [125,126]. These values were determined before many of the *A*-values discussed above for cyclohexane and tetrahydropyran derivatives were measured and are formulated in terms of 1,3-diaxial

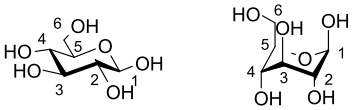


Figure 11
Naming conformers of β-D-glucopyranose. For 18, the more stable chair conformer, the lowest number C, C-1 is below the reference plane and from the atom above the plane, C-4, the order of the remaining atoms appears clockwise. This atom is designated as the superscript, that is, this is the $^4C_1$ conformation. For 19, the less stable chair conformer, C-1 is projecting above the reference plane and the order of the remaining atoms appears clockwise. This atom is designated as the superscript, that is, this is the $^1C_4$ conformation

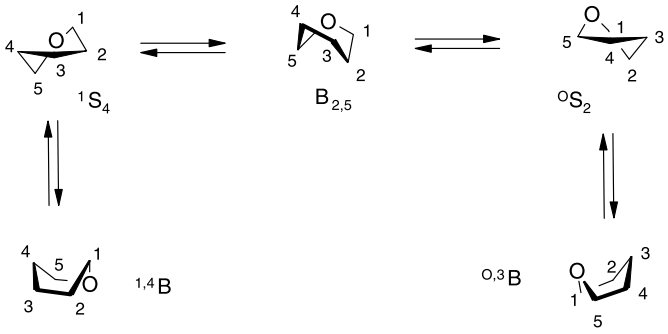


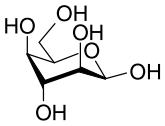
Figure 12
Part of the skew-boat pseudorotational itinerary for aldopyranoses to illustrate conformational nomenclature

Table 1
Destabilizing effects in pyranose and cyclitol rings in aqueous solution at room temperature in kJ/mol (22 °C or 25 °C) [126]

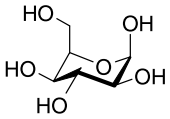
1,3-Diaxial interactions		Gauche interactions			Anomeric effects	
C:O	10.4	C/O	1.9	for OH	if O-2 is equatorial	2.3
O:O	6.3	O/O	1.5		if O-2 is axial	4.2
C:H	3.8	C/C	3.8		if O-2 and O-3 are axial	3.6
O:H	1.9				if 2-deoxy	3.6

and *gauche* interactions and the anomeric effect (see *Table 1*). In general, these terms are remarkably successful for predicting which chair conformer is most populated in aqueous solution and also for predicting anomer populations. For D-aldopyranoses, the group on the pyranose ring with the largest *A* value is the hydroxymethyl group and this group causes most derivatives to adopt the $^4C_1$ conformation. In the $^1C_4$ conformation, this group is axial and has 1,3-diaxial interactions with substituents or hydrogen atoms on C-1 or C-3. Augé and David [127] noted that the value adopted by Angyal for the 1,3-diaxial interaction of a methyl and a hydrogen (3.76 kJ mol^{-1}) gave an equatorial preference for the hydroxymethyl group of 7.5 kJ mol^{-1}, similar to the cyclohexane methyl *A* value of 7.4 kJ mol^{-1} [19,20]. However, Eliel et al. [45] found that the *A* value for a hydroxymethyl group at C-2 of tetrahydropyran was 12.1 kJ mol^{-1}, about 4.6 kJ mol^{-1} greater than its cyclohexane value. Therefore, Augé and David [127] proposed a correction termed the "proximity" correction of 4.6 kJ mol^{-1} for pyranose conformations with axial hydroxymethyl or methyl substituents at C-5. This correction only increases the proportion of the already dominant $^4C_1$ conformation of most sugars and gave improved agreement for data from many substituted idopyranosyl and altropyranosyl derivatives.

The idopyranose anomers will be considered as examples of applications of Angyal's interaction energies. They are particularly interesting because the α-anomer is one of three α-aldohexopyranoses where both chair conformers are populated significantly (α-altropyranose and α-gulopyranose are the others) and also because the conformations of L-iduronic acid have significance for the biological properties of heparin and other glycosylaminoglycans [128]. For β-D-idopyranose, estimation of the destabilizing interactions by Angyal's method [125,126] gives: for the $^4C_1$ conformer, 2 × 1.9 (O:H) + 6.3 (O:O) + 1.9 (C/O) + 1.5 (O/O) + 3.6 (anomeric effect) = 17.1 kJ mol^{-1}; for the $^1C_4$ conformer, 1.9 (O:H) + 3.8 (C:H) + 10.4 (C:O) + 1.9 (C/O) + 3 × 1.5 (O/O) = 22.5 kJ mol^{-1}. Populations of conformers can be estimated from average NMR coupling constants between vicinal hydrogen atoms, $^3J_{H,H}$, if the values for the individual conformers are known or can be estimated or calculated. These coupling constants are known for most aldoses and ketoses [129,130,131,132]. The ΔG value, 5.4 kJ mol^{-1} predicts $K(^1C_4/^4C_1) = 0.1$ at 298 K; the value derived from observed coupling constants and estimated values for $J_{2,3}$ and $J_{3,4}$ in pure conformers [131] is $K = 0.33$, which corresponds to ΔG of 2.7 kJ mol^{-1}. The difference between the observed and calculated equilibrium constant is larger than for most pyranoses but the $^4C_1$ conformer (**20**) is correctly predicted to be more stable than the $^1C_4$ conformer (**21**).

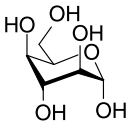


20

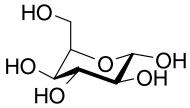


21

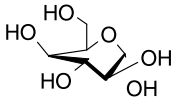
The conformational situation for α-D-idopyranose is more complicated. Estimation of the destabilizing interactions by Angyal's method [125,126] gives: for the $^4C_1$ conformer (**22**), 2×1.9 (O:H) $+ 2 \times 6.3$ (O:O) $+ 1.9$ (C/O) $= 18.3$ kJ mol^{-1}; for the $^1C_4$ conformer (**23**), 2×3.8 (C:H) $+ 1.9$ (C/O) $+ 3 \times 1.5$ (O/O) $+ 2.3$ (anomeric effect) $= 16.3$ kJ mol^{-1}. Considering the uncertainties involved in the method, the difference of 2.0 kJ mol^{-1} is too small to be confident that only the $^1C_4$ conformer is populated significantly [$K(^4C_1/^1C_4) = 0.4$ at 298 K]. Synder and Serianni calculated $K = 0.25$ from the observed values of $J_{2,3}$ and $J_{3,4}$ using standard values for $J_{a,a}$ and $J_{e,e}$ of 9.5 and 2.0 Hz, respectively, which predict the $^1C_4$ conformer being more stable by 3.4 kJ mol^{-1} [131]. A complication is that the $J_{4,5}$ value observed, 5.0 Hz, is too large for an ideal $^1C_4$ conformer, where $J_{a,e}$ should be 3–4 Hz, as it also should be in the $^4C_1$ conformer. It is possible that a skew conformer, the $^OS_2$ conformer (**24**), which would have similar J values to the $^1C_4$ for H-1 to H-4, but a larger $J_{4,5}$, contributes, perhaps significantly [127,133,134]. Alternatively, the axial hydroxymethyl group in the $^1C_4$ conformer may be bent away from the ring resulting in a smaller H-4 H-5 torsional angle than normal and a bigger $J_{4,5}$ value [131]. For both anomers, the $^4C_1$ anomer was observed to be less stable than calculated using Angyal's method. On the basis of these equilibria in aqueous media, it does not appear that Angyal's values need to be corrected as suggested by Augé and David [127]. Recent ab initio calculations for α-D-idopyranose in the gas phase suggest that the conformational mixture present includes the $B_{3,O}$ conformer as well as those mentioned above [135].



22

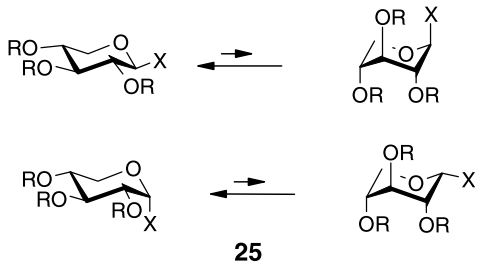


23

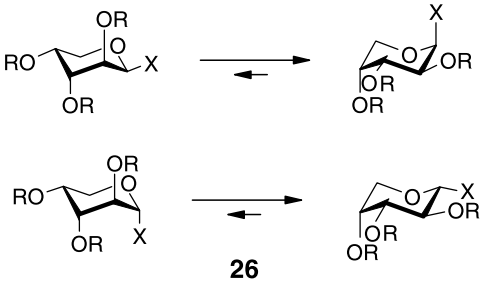


24

The aldopentopyranoses are not constrained to the $^4C_1$ conformation by the presence of a hydroxymethyl group. As a result, the equilibria between the $^4C_1$ and $^1C_4$ conformers are more closely balanced, with the D-xylopyranoses present in aqueous solution mainly as $^4C_1$ conformers (**25**), the D-arabinopyranoses mainly as $^1C_4$ conformers (**26**), and the others as mixtures [136,137,138].



25



26

Comparison of the populated conformers of substituted and non-substituted aldopentopyranoses illustrates how various types of substitution influence the equilibria (see *Table 3*). It should be noted that Angyal's interaction energies [125,126] predict the relative stabilities of the $^4C_1$ and $^1C_4$ conformers of the unsubstituted aldopentopyranoses remarkably well, except for those of arabinose, where the $^1C_4$ conformer is more stable than expected. It may be that the interaction energies overestimate the effect of having an axial hydroxyl group at C-2 or C-4, where there is a *gauche* interaction with the ring oxygen [139]. The acetylated derivatives generally agree with these relative stabilities as well [140], once allowance is made for the greater stabilization of the $^4C_1$ conformer of the α-anomer and the greater destabilization of the $^4C_1$ conformer of the β-anomer by the larger anomeric effects (*Table 2*) associated with methoxy, acetoxy, and chloro groups than with the hydroxyl group. It is interesting that tri-*O*-acetyl-β-D-xylopyranosyl fluoride exists entirely in a conformation with the fluoride group axial [141]. It is likely that 1,3-synaxial repulsive interactions for two acetates are much less than expected based on hydroxyl or methoxyl values and the same is true to a greater extent for 1,3-diaxially related sulfates [142].

Table 2
Other anomeric effects (kJ/mol)

Br[a]	> 13.4	Cl[a]	11.1	for OMe[c]	if O-2 is equatorial	5.6[d,e]
			12.8 if scaled[b]		if O-2 is axial	9.0[e]
OAc[f]	5.9					

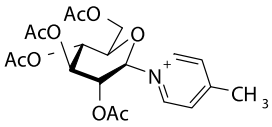
[a]From tetrahydropyran equilibria [143]. [b]By the method of Franck [56]. [c]In methanol [2]. [d][144]. [e][145]. [f][146]

Table 3
Percentages of selected D-pentopyranose derivatives that exist as $^4C_1$ conformers

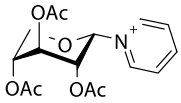
Configuration	Predicted[a]	Unsubstituted in water-d_2	Methyl 2,3,4-tri-*O*-acetyl[b]	1,2,3,4-Tetra-*O*-acetyl[b]	2,3,4-Tri-*O*-acetyl chloride[c]
α-*ribo*	54	41[d]	65[e]	77[e]	
β-*ribo*	73	74[f]	39[e]	43[e]	6[e]
α-*arabino*	13	~0[d]	17[c,e]	21[c,e]	
β-*arabino*	30	~2[d]	3[e]	4[e]	2[e]
α-*xylo*	94	~100[d]	> 98[e]	> 98[e]	> 98[e]
β-*xylo*	98	~100[d]	81[e]	72[e]	21[e]
α-*lyxo*	72	56[d]	83[e]	71[e]	91[e]
β-*lyxo*	85	86[d]	58[e]	39[e]	

[140] [a]Calculated from conformer energy differences using Angyal's method [125]. [b]In acetone-d_6 unless otherwise specified. [c]In chloroform-d. [d]Calculated from data in [130] using $J_{1a,2a} = 7.8$ Hz, $J_{1e,2e} = 1.2$ Hz. If H-1 and H-2 are not *trans*, the values of $^3J_{H,H}$ for all pairs of *trans* non-anomeric vicinal protons were averaged, with $^3J_{a,a} = 9.5$ Hz, $^3J_{e,e} = 1.8$ Hz for $J_{2,3}$ and $J_{3,4}$, but $^3J_{a,a} = 10.3$ Hz, $^3J_{e,e} = 1.8$ Hz for $J_{4,5}$. [e] [140]. [f]Calculated from data in [138]

The origin of the anomeric effect was discussed above in connection with tetrahydropyran conformations. The term "reverse anomeric effect" was coined to describe the tendency of substituents with atoms with formal positive charges attached to C-1, initially pyridinium ions, to adopt equatorial rather than axial orientations [147]. This effect, if real, has great significance for reactions at the anomeric center because many proceed via intermediates that involve increase in positive charge density on atoms attached to the anomeric center. Amongst the pieces of evidence in support of this effect were the surprising observations that in solution the *N*-(2,3,4,6-tetra-*O*-acetyl-α-D-glucopyranosyl)-4-methyl-pyridinium ion adopts a boat conformation (**27**) [148] and the *N*-(2,3,4-tri-*O*-acetyl-α-D-xylopyranosyl)pyridinium ion adopts a $^1C_4$ conformation (**28**) [149]. However, the steric effects of these groups are large, similar to phenyl, which has an *A* value of 9.2 kJ mol^{-1}, magnified to 14.6 kJ mol^{-1} at position-2 of a tetrahydropyran derivative [51]. The strongest evidence for this effect was the observation that the position of the equilibrium between the $^4C_1$ and $^1C_4$ conformers of *N*-(2,3,4-tri-*O*-acetyl-α-D-xylopyranosyl)imidazole moves toward the latter on protonation in chloroform-*d* as observed through the considerable changes in the average vicinal coupling constants [149,150]. Perrin and coworkers have reinvestigated this phenomenon [151,152,153] by studying the titration of glycosylimidazole anomers and other compounds and have concluded that the reverse anomeric effect does not exist. The most significant evidence is that glycosyl imidazole groups are more basic when axial than equatorial, contrary to prediction based on the reverse anomeric effect. Other groups have now come to the same conclusion [154,155,156]. Ammonium ions do not show this effect and it now seems clear that the observations that led to the concept of the reverse anomeric effect were due to unexpectedly large steric effects and perhaps to particular electrostatic interactions with imidazolium ions [157].

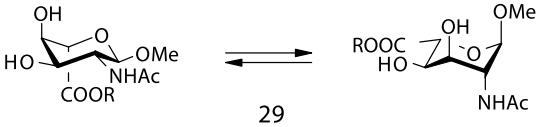


27



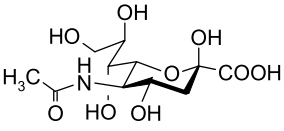
28

Biologically important monosaccharides contain a number of functional groups different to those considered thus far. Aldohexopyranuronic acids favor the $^4C_1$ conformation to about the same extent as do aldopyranoses [1]. Only α-L-iduronic and α-alturonic acids contain significant amounts of other conformations and the former is particularly complicated because the ^2S_O conformation is also populated [158]. Interestingly, the 2-sulfate of methyl α-L-iduronate exists predominantly in the $^1C_4$ conformation [159]. Methyl 2-acetamido-2-deoxy-α-L-altropyranuronic acid exists as a 65:35 $^4C_1$:$^1C_4$ mixture (**29**). The sodium salt and the methyl ester are present mainly as the $^4C_1$ conformer and entirely as the $^1C_4$ conformers, respectively, suggesting that charge on C-5 stabilizes the $^4C_1$ conformer. Interestingly, when this acid is linked at O-1 and O-4 as part of the O-specific polysaccharide of *S. sonnei*, it is only present in the $^4C_1$ conformation [160].



29

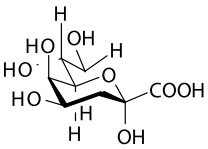
N-Acetyl-D-neuraminic acid (5-acetamido-3,5-dideoxy-D-*glycero*-D-*galacto*-2-nonulosonic acid, NANA) is a critical part of cell surface oligosaccharides. Under different conditions, it crystallizes as a dihydrate and in the anhydrous form. In both crystals, NANA is in the β-pyranose form with the ring in the $^2C_5$ conformation [161,162] that has the side chain and the carboxyl group equatorial and the glycosyl hydroxyl axial (**30**). In aqueous solution, the β-pyranose isomer is accompanied by a small amount (about 8%) of the α-pyranose form [163,164,165]. The $^2C_5$ conformation is maintained in solution for α- and β-pyranose forms, for their alkyl esters and for their glycosides [165,166,167,168]. The conformation of the side chain will be discussed below.



30

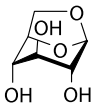
3-Deoxy-D-*manno*-2-octulosonic acid (KDO) is present in the inner core region of gram negative bacterial lipopolysaccharides [169,170] and in a green algae [171,172] as the α-pyranoside. In most bacterial exopolysaccharides, it is in the β-pyranoside form [173]. In X-ray structures, the pyranose rings of ammonium 3-deoxy-α-D-*manno*-2-octulopyranosidonate [174,175], of methyl (methyl 4,5,7,8-tetra-*O*-acetyl-3-deoxy-α-D-*manno*-2-octulopyranosid)onate [176], and of an α-(2→4)-linked disaccharide [177] all adopt the $^5C_2$

conformers that have the side chain and the carboxyl group equatorial (**31**). Solutions of free KDO show a much more complex tautomeric mixture than NANA, containing 60–65% α-pyranose, 2–11% β-pyranose, 20–25% α-furanose, and 8–9% β-furanose [178,179]. The ammonium salts of the α- and β-pyranose methyl glycosides are also present in solution as $^{5}C_{2}$ conformers but these compounds differ in their C-7-C-8 rotameric populations [174].

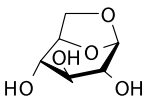


31

Non-chair conformations are often populated for anhydro sugars. Depending on the type of substitution, these sugars can exist entirely in boat conformations as in 2,6-anhydro [180] or 3,6-anhydro derivatives or to a small extent as for 1,6-anhydro-β-D-glucopyranose (❯ *Fig. 13*). 1,6-Anhydro-β-D-hexopyranoses are in equilibrium with the D-hexopyranoses and water when the latter compounds are heated in aqueous acid [181]. On formation of a 1,6-anhydro ring, the normal $^{4}C_{1}$ conformer is forced into a $^{1}C_{4}$ conformation. The amount of the 1,6-anhydro-β-D-hexopyranose present at equilibrium at 100 °C in aqueous acid can be predicted quite successfully from Angyal's interaction energies with a value of $\Delta G°$ for anhydro ring formation of $-11.7\,\mathrm{kJ\,mol^{-1}}$ [181]. The $^{3}J_{H2,H3}$ and $^{3}J_{H3,H4}$ values observed for most 1,6-anhydro-β-D-glucopyranose derivatives are small, 1–2 Hz, consistent with a $^{1}C_{4}$ conformation and all of the derivatives that have been studied by X-ray diffraction adopt this conformer [182,183]. It was therefore surprising that the $^{3}J_{H2,H3}$ and $^{3}J_{H3,H4}$ values for 3-amino-3-deoxy-1,6-anhydro-β-D-glucopyranose were both 5.5 Hz in dimethyl sulfoxide-d_6 [184], even though this compound adopted the $^{1}C_{4}$ conformation in the solid state [185], and $^{3}J_{H2,H3}$ and $^{3}J_{H3,H4}$ values for 3-amino-2-*O*-benzyl-3-deoxy-1,6-anhydro-β-D-glucopyranose were both < 2.0 Hz in chloroform-*d* [183]. It was concluded that the all axial substituent orientations in 1,6-anhydro-β-D-glucopyranose derivatives bring the energy of the $^{1}C_{4}$ conformation (**32**) close to a boat conformation, the $B_{O,3}$ (**33**). The more polar $B_{O,3}$ conformer is favored by more polar solvents and even the parent compound is present in this conformer to an extent of about 20% in water or dimethyl sulfoxide [183]. Two large or very polar groups on C-2 and C-4 also move the equilibrium toward the boat [186,187,188,189] as can unusual hydrogen bonding situations [190].



32



33

Exocyclic Groups The conformations adopted by the exocyclic groups will be considered in three sections: the hydroxymethyl group and longer side chains, the hydroxyl groups, and the anomeric group.

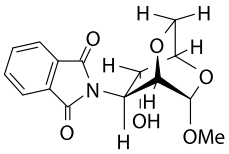


Figure 13
A boat conformation of a 2,6-anhydro derivative recently observed [165]

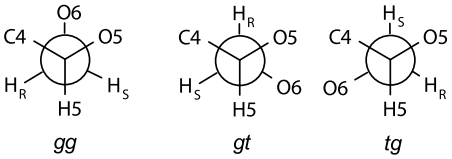


Figure 14
Nomenclature for hydroxymethyl rotamers in a D-aldopyranose

The conformations of the hydroxymethyl group are normally discussed in terms of the three staggered rotameric conformations (*Fig. 14*), termed *gg*, *gt*, and *tg*. The two letters indicate the orientation (*gauche* or *trans*) of O-6 with respect to the ring oxygen, then with respect to C-4. This topic was reviewed by Bock and Duus in 1994 [191]. The *tg* orientation has recently been shown to destabilize anomeric oxacarbenium ions significantly more than the other two orientations giving this topic significance for anomeric reactivity [192,193]. The preferences have been determined mainly from X-ray data [194,195], from NMR studies using average $^{3}J_{H,H}$ values [191], from $^{2}J_{H,H}$, $^{2}J_{C,H}$, $^{3}J_{C,H}$, and $^{4}J_{C,H}$ values [196,197,198], from chiroptical methods [199,200], and recently by resonant two-photon ionization and resonant ion-dip infrared spectroscopy [201,202]. The most important factor in determining the relative populations of the three rotamers is the orientation of the C-4 substituent with respect to the hydroxymethyl group. For the D-pyranoses, sugars with the *trans* orientation (*gluco-* or *manno-* configurations) have approximate populations *gg*:*gt*:*tg* of 6:4:0 while those with the *cis* orientation (*galacto*-configuration) have *gg*:*gt*:*tg* of 2:6:2. The configuration at C-1 [191,203], substitution [204,205], and solvent effects [204,205,206,207,208,209] also influence rotamer populations. For instance, the rotamer populations are *gg*:*gt*:*tg* 40:53:7 for methyl α-D-glucopyranoside but 31:61:8 for methyl β-D-glucopyranoside [197]. In aqueous solutions for glucopyranosides, hydrogen bonding to O-4 in the *tg* rotamer is weakened, leading to the observed predominance of the *gg* and *gt* rotamers [207]. Thus, the major effects influencing these populations are O//O interactions with the substituents on O-4, the *gauche* effect, hydrogen bonding and solvation [191]. Interestingly, the ratio *gg*:*gt*:*tg* only changes to 45:54:1 when the 4-hydroxy group is removed from methyl α-D-glucopyranoside (or galactopyranoside) or to 31:65:4 when removed from the β-anomer [191]. The *gauche* effect obviously has considerable importance in determining the position of these equilibria. Considerable advances in computational methods for the incorporation of solvation have been achieved recently that now allow the position of this equilibrium to be predicted accurately [207,210,211,212,213,214,215,216] and the pathway for rotamer interconversion to be studied [216].

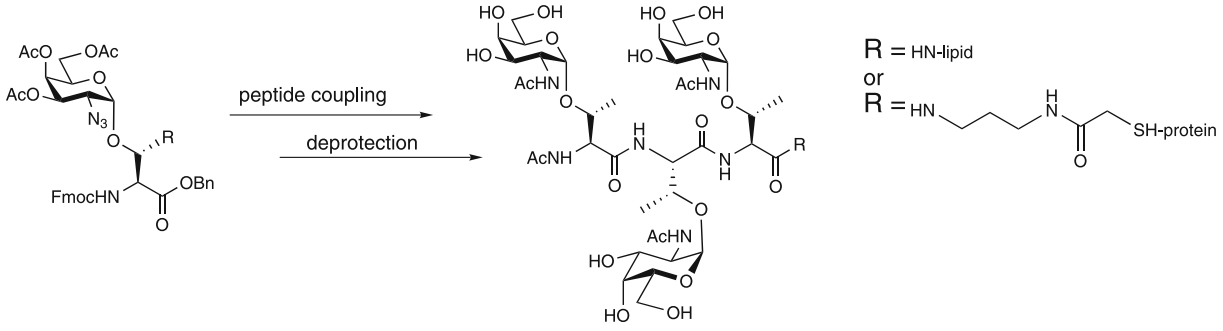
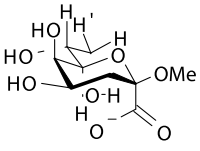


Figure 15
Possible side-chain conformations of *N*-acetyl-D-neuraminic acid

The conformations of the side-chains of sialic acid and KDO can be predicted using the principles outlined above. For *N*-acetyl-D-neuraminic acid (◉ *Fig. 15*), the conformation of the side chain can be termed *gg*, *gt*, or *tg*, using the orientations of C-8 with respect to O-6 and C-5 to define the conformers, as for hydroxymethyl conformers. The *gt* conformer has three minor O//H interactions, an N//H interaction, two C/O and an O/O interaction, the *gg* conformer has C//H and the major C//O interaction plus two C/O and a C/C interaction, while the *tg* conformer has major C//N and C//O interactions plus C/C and O/O interactions. Not surprisingly, the *gt* conformer is preferred as shown by small $J_{6,7}$ values both for the parent compound (0.8Hz [166]) and for α-(268) linked oligomers (0.5 to 1.0Hz [217]). The planar extended conformation about the C-7 – C-8 bond avoids the O//O and C//O interactions present in the other rotamers and is strongly preferred ($J_{7,8}$ 9.2 Hz) for the parent compound [166,218]. The $J_{8,9}$ values are 6.2 and 2.5 Hz, consistent with a mixture of the OH extended conformer (◉ *Fig. 15* bottom center) and the *gauche* conformer which avoids an O//O interaction (bottom right). The NMR relaxation times for C-7 and C-8 are similar to those of the secondary ring suggesting that there is only one populated conformer for the C-7 C-8 portion of the side chain [167,168].

However, in the α-(268) linked sialic acid oligomers, the OH-8 is replaced by OR-8, where R is the bulky quaternary center C-2' of the next sialic acid. Not surprisingly, in the α-(268) linked sialic acid trisaccharide and in the α-(268) linked sialic acid polymer, colominic acid, mixtures of rotamers about the C-7–C-8 bond are adopted [219,220] that have been interpreted for the polymer as giving a helix [221] or a short-lived helix [219], consistent with this polymer being a conformational epitope for serogroup B of *Neisseria meningitidis*. More recent detailed NMR relaxation studies have concluded that the polymer is a random coil with no helix present for any extended period and that the polymer has considerable flexibility about the exocyclic torsional angles [222].

The axial hydroxyl group at C-5 and the *R*-configuration at C-7 of 3-deoxy-α-D-*manno*-2-octulopyranosidonic acid result in the *gt* conformer about the C-6–C-7 bond being strongly preferred to avoid C//O and O//O interactions [174] (◉ *Fig. 16*). Interestingly, the populations of the hydroxymethyl group rotamers change when the anomeric configuration of the methyl glycoside of the ammonium salt is converted from α to β. The *a* conformer (defined by the relationship of C-6 and O-8) should be more stable; it has only a O/O interaction, while the *g*− conformer has an O//O interaction and the *g*+ conformer has O/O and O//O interactions. The *a* conformer fits the *J* values for the α anomer [174] and is present in the crystal structure [174,175]. The β anomer has *J* values closer to that expected for the *g*+ conformer; it seems reasonable that a postulated [223] hydrogen bond from the axial carboxylate anion stabilizes this rotamer (**34**). However, the peracetate of the α-anomer methyl ester also adopts this conformer, both in the crystal and in solution [176], as does the inner KDO in an α(268) linked dimer [177].



34

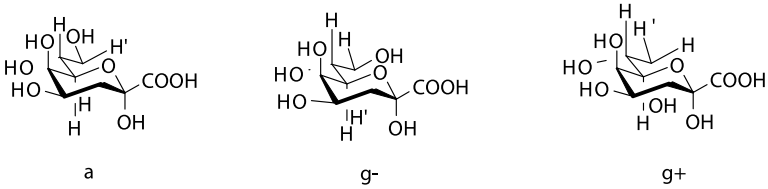


Figure 16
Possible side-chain conformations of 3-deoxy-α-D-*manno*-2-octulopyranosidonic acid

When considering the orientations of the hydroxyl groups in carbohydrates, it should be noted that the rotational barrier for methanol, 4.48 kJ mol^{-1} [224] is about one-third the value in ethane or dimethyl ether [225]. Thus, preferences for staggered conformers in HCOH units are considerably smaller than for hydroxymethyl groups. Information about the orientations of OH groups has come from theoretical studies, from ^{1}H NMR studies in dimethyl sulfoxide, where exchange is slow, in cooled water or water/acetone mixtures, and recently from low temperature gas phase IR spectra. In the gas phase, calculations indicate that the secondary hydroxyl groups are all oriented either clockwise or counterclockwise because this sets up semicircular intramolecular hydrogen bonding networks [210,226,227,228].

A recently developed method, resonance ion-dip infrared spectroscopy, has provided experimental support for this conclusion about the orientations of hydroxyl groups of carbohydrates in the gas phase [229,230]. In this technique, a sample heated in a controlled manner in an oven to between 100 and 230° under an Ar pressure of 4 to 5 bar is allowed to escape through a nozzle. The expansion causes the jet to cool rapidly to 5 to 10 K. IR and UV spectra are recorded using pulsed tunable lasers and the spectra of individual conformers can be selected by UV hole-burning using a high power laser tuned to a band from that conformer and identified by comparison with spectra calculated using DFT theory. The compound being studied must contain a chromaphore and phenyl glycosides have been used. Surprisingly, only one to three conformers are observed for monosaccharides and disaccharides with clockwise or counterclockwise circular hydrogen bonding networks [231,232,233]. Monohydrates can be studied adding water to the Ar atmosphere and identified by mass spectrometry. It was found that molecules of water insert into the H-bonding circuits where the hydrogen bonds are weakest and the hydrated conformers can be the same as the unhydrated ones, minor unhydrated structures or structures that had negligible populations when unhydrated [233] (see *Fig. 17* for examples). Recently, this technique has been combined with resonant Raman optical activity (ROA) [234] spectra of aqueous solutions at room temperature to suggest that in most cases, the same conformations are populated under these conditions as in the hydrated Ar jet at 5 to 10 K [235], although benzyl β-D-lactoside was found to change to the conformation found for lactose by NMR spectroscopy [236].

In dimethyl sulfoxide, exchange of hydroxyl protons is slow enough that hydroxyl protons are observed separately coupled to adjacent CH protons in the ^{1}H NMR spectra of carbohydrates [237,238,239]. Hydroxyl protons involved in intramolecular hydrogen bonds are shielded in comparison to those involved in intermolecular hydrogen bonds to dimethyl sulfoxide [240,241]. The coupling constants can be used in conjunction with Karplus-type relationships [242,243] to identify the orientation of the OH groups. The *J* values observed for

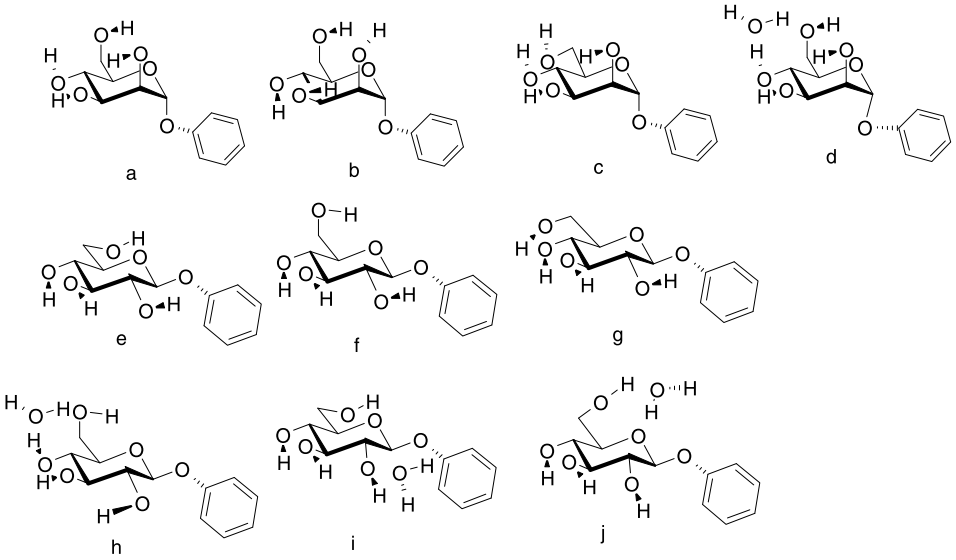


Figure 17
Hydroxy group conformers observed in the gas phase at 5–10 K by ion-dip IR spectroscopy: a–c, phenyl α-D-mannopyranoside, with stability decreasing from a to c; d, the only conformer observed for phenyl α-D-mannopyranoside hydrate; e to f, phenyl β-D-glucopyranoside, with stability decreasing from e to f; g to i, monohydrates of phenyl β-D-glucopyranoside, with stability decreasing from g to i. Note that the most stable conformer, g, was not populated for the anhydrous compound and was calculated to be 10 kJ mol^{-1} less stable than e

fully solvated equatorial hydroxyls are consistent with equal occupancy of all three rotamers (4.5–5.5 Hz) [240,241]. Fully solvated axial hydroxyl groups have slightly smaller J values (4.2–4.4 Hz), consistent with lower occupancy of the rotamer with the hydrogen atom on the interior of the ring. Values outside these ranges indicate that the OH group is an intramolecular H-bond donor [240]. Partial labeling with deuterium can be used to identify which hydroxyl groups are involved in hydrogen bonds by noting which hydroxyl signals show splitting due to the isotopic effect of the hydrogen-bonded hydrogen isotopes (SIMPLE NMR) [244]. The signs of the shift changes provide structural information [245]. Alternatively, the temperature dependence ($\Delta\delta/\Delta T$) of OH chemical shifts can be used; fully solvated OH groups show large $\Delta\delta/\Delta T$ while OH groups that are intramolecularly H-bonded show small $\Delta\delta/\Delta T$ [246]. All the hydroxyl groups in α-D-glucopyranose are involved in weak intramolecular hydrogen bonding. However, since this does not affect the magnitudes of their coupling constants, individual hydroxyl groups must be hydrogen bonded only a small proportion of the time. *Syn*-1,3-diaxial hydroxyl groups form stronger intramolecular hydrogen bonds that do affect J values [238].
In the more polar solvent water, there should be less tendency for ordering through intramolecular hydrogen bonding. Thus, hydroxyl orientations are mostly randomly distributed in water for monosaccharides, as indicated by low temperature NMR experiments [247,248]. As in DMSO solutions, equatorial hydroxyl groups flanked by axial oxygen atoms have slightly larger coupling constants [247]. More recently, information about the orientations of OH

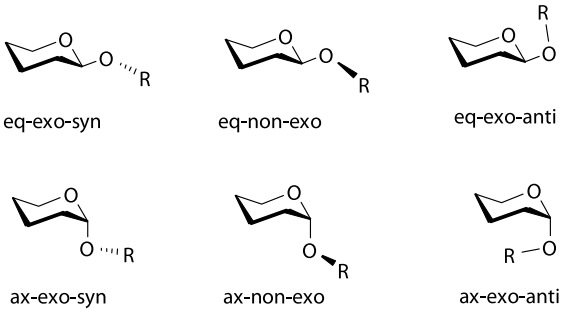


Figure 18
Representation of the three basic conformations about the anomeric CO bond: top, for β-D-pyranoses; bottom, for α-D-pyranoses

groups has been obtained in a wider range of NMR data. Chemical shifts are particular informative [249,250,251], but information can also be obtained from temperature coefficients of chemical shifts, rates of exchange with water, and nOe and NOESY experiments [252,253]. Some strong hydrogen bonds that are present in DMSO are not present in aqueous solutions [254,255] and sometimes different hydrogen bonds form [256]. In other situations, the same hydrogen bonds are present in water as in DMSO [257,258]. A variety of types of evidence has shown that certain intraresidue hydrogen bonds persist in aqueous solutions of oligosaccharides [252,259,260,261,262].

The orientations about the anomeric CO bond can be described as shown in *Fig. 18* [261]. This nomenclature was chosen over the traditional ag^+ nomenclature because the latter changes when the ring chair is inverted. The *exo-syn* and *exo-anti* conformations for both α and β anomers are favored by the exo-anomeric effect [49,52,71,263], which minimizes lone pair repulsions. The *exo-anti* conformations are disfavored by the two *gauche* interactions with O-5 and C-2. As a result, the *exo-syn* conformations are the only conformations populated in crystal structures of carbohydrates [69,264]. Theoretical calculations support the significance of the exo-anomeric effect [52,265,266]. Comparison of the conformations of C-glycosides, such as C-lactose and related compounds, which cannot have an exo-anomeric effect, with those of their *O*-containing analogs, decisively demonstrate the significance of the exo-anomeric effect [6,267,268].

It should also be noted that the non-exo conformations are disfavored by equatorial substituents at C-2, due to the O//O interaction, thus sterically increasing the preference for the eq-*exo-syn* conformation [269] (*Fig. 19*). However, even without the exo-anomeric effect and without

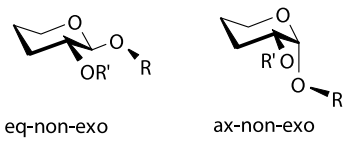


Figure 19
Steric destabilization of non-exo conformations by equatorial substituents on C-2

Figure 20
Steric destabilization of non-exo conformations by H-2

substituents at O-2, there is a substantial preference for the *exo-syn* conformations as observed for 2-deoxy-C-glycosides [270] and calculated for 2-ethyltetrahydropyran [269]. The preference arises because the CH_2R group avoids the RH interaction with the hydrogen on C-2 by being *syn* to the oxygen atom (*Fig. 20*).

1.3.3 Conformations of Furanoses

Because pseudorotation is so facile for cyclopentane and tetrahydrofuran derivatives and because substituent conformational preferences are not large, furanose derivatives are normally present as mixtures of conformations dominated by the interactions between substituents. These can be described as mixtures of ideal twist and envelope conformations but this description is often inadequate for intermediate conformations. The alternative description in terms of the pseudorotational itinerary is more precise [12,271,272,273] but less easy to visualize. Two different formalisms are used. The Altona–Sundaralingam (AS) system [272] described here is related to the more general Cremer–Pople (CP) system [12] by subtracting 90° from the CP phase angle. The CP puckering amplitude can be converted to the AS amplitude by dividing by 100 [7]. In the AS system, the infinite number of conformations on the pseudorotational itinerary are described in terms of the maximum torsion angle, θ_m, and the pseudorotation phase angle P. The pseudorotation phase angle P is calculated from the endocyclic torsional angles, θ_0, θ_1, θ_2, θ_3, and θ_4 according to Eq. (1) [272,273,274]:

$$\tan P = \frac{[(\theta_4 + \theta_1) - (\theta_3 + \theta_0)]}{[2\,\theta_2(\sin 36° + \sin 72°)]}\,. \tag{1}$$

The phase angle P is defined to be 0° when θ_2 has a value that is maximally positive, corresponding to the conformation ${}^3T_2$ (*Fig. 21*) and returns to the same point at $P = 360°$. From the phase angle P, the five torsion angles are related by:

$$\theta_j = \theta_m \cos(P + j\delta) \tag{2}$$

where $j = 0$ to 4 and $\delta = 720°/5 = 144°$. The maximum torsion angle, θ_m, is derived by setting $j = 0$:

$$\theta_m = \frac{\theta_0}{\cos P}\,. \tag{3}$$

In the pseudorotation cycle (*Fig. 21*), a change of P by 180° reverses the signs of all torsion angles. At every phase angle P, the sum of the torsion angles is 0°. Envelope and twist conformations alternate every 18° and T conformations are found at even multiples of 18°. The section of the pseudorotation cycle with phase angles of $0 \pm 90°$ is referred to as the

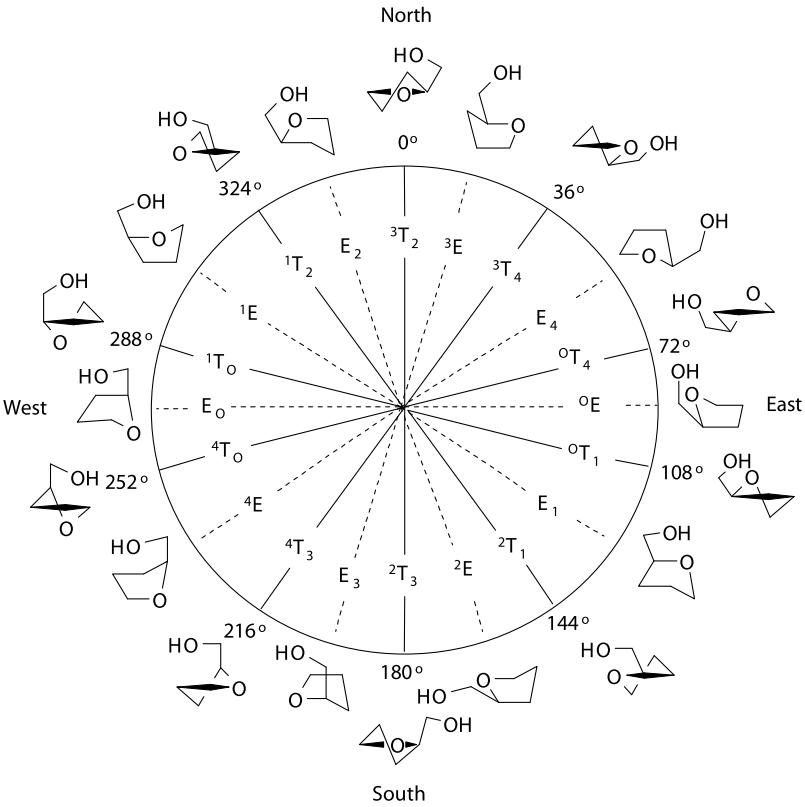


Figure 21
Pseudorotational itinerary for a D-aldopentofuranose

northern half or N while that with phase angles of $180 \pm 90°$ (the lower half of the pseudorotational itinerary) is referred to as the southern half or S. For nucleotides and nucleosides, the preferred conformations lie in part of the N section with *P* angles of 0 to 36° and in the S section with phase angles of 144 to 190° [272,275,276,277,278,279,280]. In the N section, these correspond to the three conformations with C-3 above the plane of the ring (C3-endo) while the favored conformations in the S section have C-2 above the plane of the ring (C2-endo). Interconversion of N and S conformers is rapid, on the same time-scale as molecular reorientation [281]. Alteration of substitution and regiochemistry markedly alters the populated conformations [276,280,282,283]. Conformational studies on nucleosides and nucleotides have shown that the N Ω S equilibrium is biased towards S in 2-deoxyribofuranose derivatives but is approximately 1:1 for ribofuranoses [273,284].

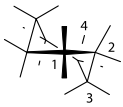
Lowary, Haddad and coworkers studied methyl 3-*O*-methyl-α-D-arabinofuranoside computationally at the B3LYP/6–31G* level unsolvated and in aqueous solution using the B3LYP/6–31+G**SM5.42/BPW91/631G* method and obtained conformer distributions close to those derived from $^3J_{H,H}$ values, mainly a mixture of the $^\circ T_4$ or E_4 conformers (N) and the $^2T_1$ or ^2E_o conformers (S) [285]. The same group studied all of the methyl aldofuranosides using

the same methods except that only idealized envelope conformers were considered [286]. In most cases, the change in the regions of the pseudorotational itineraries that were populated on moving to the solution were calculated to be small. Agreement with results from detailed analyses of $^3J_{H,H}$ and $^3J_{COCH}$ data [287] was good as was agreement with the conformations present in the crystal [288]. For individual compounds, the conformational situations varied widely; methyl α-D-xylofuranoside appears to exist as close to a single E_1 conformer in solution with most conformers $> 12\,kJ\,mol^{-1}$ higher in energy while the conformers of methyl β-D-xylofuranoside lie on a fairly flat potential energy surface. β-Ribofuranose has also been studied using an ab initio molecular dynamics method that concluded that the ring conformer and the hydroxymethyl group rotamer populated were linked as well as emphasizing the importance of solvation [289,290]. The former observation had been made earlier [291,292]. β-D-Galactofuranose derivatives exist preferentially in the southwestern part of the pseudorotational cycle, in the ^4E, ^4T_O, E_O, and ^1T_O range, as shown by analysis of coupling constants and molecular mechanics calculations using the PIMM91 force field [229]. Methyl α-D-mannofuranoside is present in an E_3 conformation in the crystal [293].

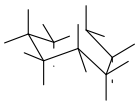
Anomeric effects were first demonstrated in furanosides by means of X-ray diffraction [61,68, 294]. The importance of the anomeric effect for furanosides in solution was shown by studies of fused ring systems [295,296], from studies of nucleosides [297,298], and by comparing *C*-, *N*-, and *O*-furanosyl glycosides in the solid state and solution [299]. Interestingly, the magnitude of the anomeric effect for nucleosides is pH dependent with the largest effect being observed under the most acidic conditions [300].

1.3.4 Conformations of Septanoses

There have been recent indications that septanose sugars can have biological activity [301,302, 303]. Stoddard outlined clearly how the conformational properties of cycloheptanes [304,305] can be applied to oxepanes and septanoses [2] and recent studies [306,307] have used the nomenclature scheme introduced by Stoddart that is analogous to those used for conformations of furanoses and pyranoses [124]. For cycloheptane, the C_2-symmetric twist-chair (TC) (**35**) is the most stable conformation as indicated by molecular mechanics calculations [304,308,309] and high level ab initio studies [310,311], and confirmed by spectroscopic information [310,312,313], electron diffraction [314], and inelastic neutron scattering [315]. The TC conformer pseudorotates through a C_s-symmetric barrier, the chair (C) conformation (**36**), calculated to be about 4 to $6\,kJ\,mol^{-1}$ above the TC [310,311] at room temperature. Another family of pseudorotating conformations includes the boat (B) and twist-boat (TB) where the B is minimum, about $12\,kJ\,mol^{-1}$ above the TC [311]. Virtually all conformational information on septanoses has been interpreted in terms of TC conformers. In the TC conformation, the C_2 axis runs through one carbon atom and the center of the opposing bond giving four different types of carbon positions, numbered 1 to 4 starting with the atom on the axis (**35**). Positions 2–4 have axial and equatorial substituent positions. Hendricksen calculated with an early forcefield that it is much more unfavorable to have methyl groups on axial positions at positions 2 and 3 than in cyclohexane [305]. Position 1 has two identical substituent positions, termed isoclinal that were calculated to be similar to equatorial positions [305].

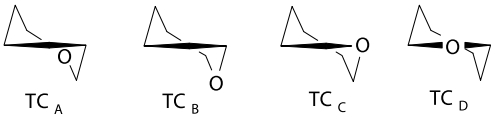


35

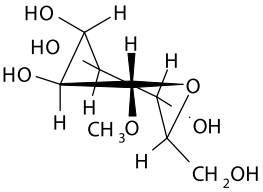


36

For oxepane, there are four different types of TC conformations, with the O atom on the axis of symmetry or at positions 2, 3, or 4 with respect to a pseudoaxis of symmetry, termed TC_D, TC_C, TC_B, and TC_A, respectively [313] (● *Fig. 22*). Molecular mechanics calculations indicate that the C_2-symmetric conformer is less stable [313,316] and this is supported by analysis of IR and Raman spectra [312]. The IR and Raman spectra were interpreted in terms of only two of the remaining TC conformations being populated, TC_C and TC_B [313]. The full TC/C pseudorotational itinerary includes 14 TC conformations, arising from pseudoaxes of symmetry going through each heavy atom and with right or left-handed twists of the remaining heavy atoms. In the TC conformation, the atom on the C_2 axis and the two flanking atoms lie in a plane with two adjacent remaining atoms above the plane and two below. The numbers of these out-of-plane atoms are used for naming TC conformations of septanoses [2], e. g., the most stable conformer of methyl α-D-*glycero*-D-idoseptanoside is $^{3,4}TC_{5,6}$ (**37**).



● Figure 22
The four possible TC conformations of oxepane with the nomenclature of Bocian and Strauss [313]



37

The conformational principles developed earlier appear to control septanose conformations. In both solution and in crystals, TC conformations are preferred that have hydroxymethyl, alkoxy groups, and hydroxyl groups in axial or isoclinal positions [307,317,318,319,320,321,322,323,324]. The favored TC conformations have oxepane conformations of type TC_B or TC_C. The equilibria between conformers are often closely balanced as indicated by the significant change in conformers populated when O-5 of methyl β-D-*glycero*-D-guloseptanoside was methylated [319].

1.4 Conformations of Disaccharides, Trisaccharides, and Oligosaccharides

1.4.1 Conformations of Disaccharides

The conformations adopted by disaccharides are defined in terms of the torsional angles across the glycosidic linkage, starting from the anomeric center. The angle Φ is defined by the torsional angle H1–C1–O1–Ci, where i is the number of the carbon atom in the aglycone. In virtually all structures determined by X-ray crystallography that were not constrained to have other values, Φ lies in the range expected on the basis of the exo-anomeric effect [69]. In solution, some glycosidic linkages have been observed to have minor populations of non-exo conformers [325,326,327] and one branched oligosaccharide was observed to have a considerable population of non-exo conformers about a β-ribofuranosyl linkage [328]. For glycosides of 2-uloses, the definition of this torsional angle is changed to C1–C2–O2–Ci. The second torsional angle, Ψ, is defined as C1–O1–Ci–Hi, where i is defined as above. If the first carbon atom in the aglycone is exocyclic, it is necessary to define a third torsional angle, ω, O1–Ci–Cj–Hj and Ψ now becomes C1–O1–Ci–Cj (see *Fig. 23* and *Fig. 24*). In crystal structures, these torsional angles are often defined in terms of the heavy atoms, that is, C1–C1–O1–Ci or

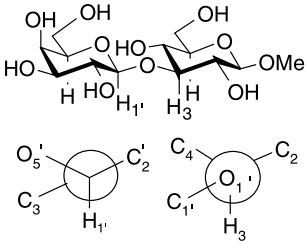


Figure 23
Definition of torsional angles Φ and Ψ: for methyl β-D-galactopyranosyl-(1→3)-β-D-glucopyranoside, Φ is the torsional angle H1′–C1′–O1′–C3, Ψ is the torsional angle C1′–O1′–C3–H3. The Newman projections show how torsional angles involving heavy atoms from X-ray data correspond to angles defined as above

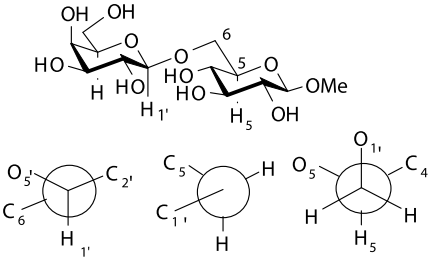


Figure 24
Definition of torsional angles Φ, Ψ, and ω: for methyl β-D-galactopyranosyl-(1→6)-β-D-glucopyranoside, Φ is the torsional angle H1′–C1′–O1′–C6, Ψ is the torsional angle C1′–O1′–C6–C5, ω is the torsional angle O1′–C6–C5–H5

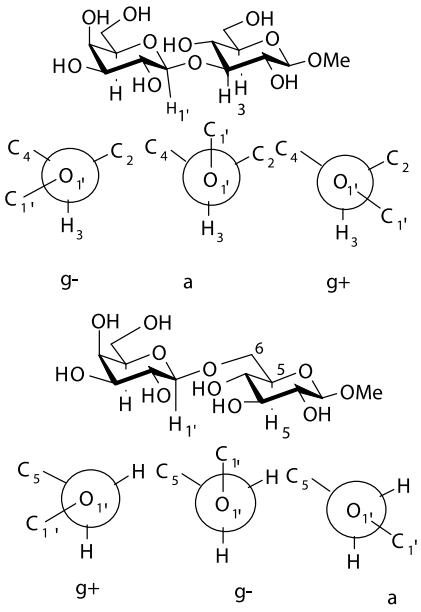


Figure 25
Designations of rotamers about the Ψ angle in disaccharides

O5–C1–O1–Ci, rather than H1–C1–O1–Ci. Approximate conversions can be made by adding + or −120°, as appropriate (see *Fig. 23* and *Fig. 24*) [1]. The book by Rao et al. [1] contains an extensive compilation of X-ray structures of disaccharides.

Unlike the torsional angle Φ, the angle Ψ commonly adopts a variety of values. It is convenient to be able to designate these values in terms of rotamers and a scheme for doing this is shown in *Fig. 25*. Conformational analysis of disaccharides can proceed by using a theoretical approach to estimate the variation in energy with Φ and Ψ. These results are usually presented in a two dimensional plot of Φ against Ψ, representing energy changes as contours, called a Ramachandran plot [329]. Analysis of CC, CO, CH and OH steric interactions in the different rotamers about the Ψ angle can simplify the evaluation of experimental data [330,331]. This procedure is illustrated in *Fig. 26* for β-(1→3) linked disaccharides. It can be seen that the *a* conformer is disfavored because having H-3 *anti* to C-1′ results in the remaining two carbons being *gauche* to C-1′. Indeed, no disaccharides linked through secondary oxygen atoms listed in the Tables in the book by Rao et al. [1] adopt this conformation in the solid state. For the particular linkage shown in *Fig. 26*, the *g*+ conformer is favored over the *g*− conformer (see *Table 4*) , because in the latter, C-4 has a 1,3-diaxial interaction with O-5, whereas in the former, the 1,3-diaxial interaction of the carbon atom that is *gauche* is that of C-2 with H-1′. In the solid state, both *g*− and *g*+ rotamers are observed but the values of the Ψ angle are normally much less than staggered [1], presumably because it is energetically advantageous to decrease the interactions of the *gauche* carbon at the expense of increasing the non-bonded interactions involving the aglycone H at the linkage center. In agreement with this statement, Lemieux and Koto concluded from hard-sphere calculations

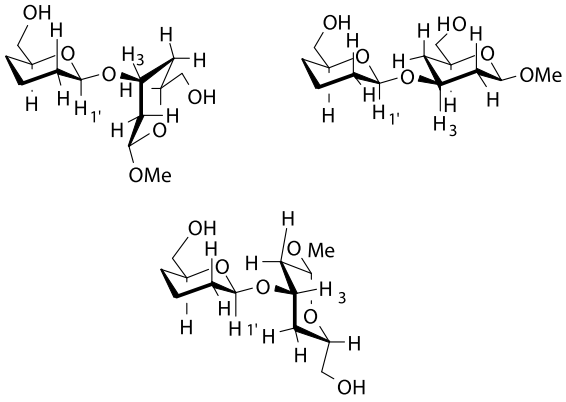


Figure 26
Estimation of steric energies in disaccharides, for methyl 2,3,4-trideoxy-β-D-*glycero*-aldohexopyranosyl-(1→3)-2,4-dideoxy-β-D-*threo*-aldohexopyranoside: top left, *g*+ conformer; top right, *g*−- conformer; bottom, *a* conformer

Table 4
Steric interactions in methyl 2,3,4-trideoxy-β-D-*glycero*-aldohexopyranosyl-(1→3)-2,4-dideoxy-β-D-*threo*-aldohexopyranoside

Conformer	Interaction	Value (kJ mol^{-1})
g+	H-30-5′	1.9
	C-2H-1′	3.8
	H-2axC-1′	3.8
	Total	9.5
g−	C-40-5′	10.4
	H-4eC-1′	3.8
	Total	14.2
	C-4H-1′	3.8
	C-20-5′	10.4
	C-1′H-2ax	3.8
	C-1′H-4ax	3.8
	Total	21.8

on both cyclohexyl D-glucopyranoside anomers that the Ψ angle should have values close to 0° [332], but the calculated preference is much too large based on observed conformer mixtures for disaccharides [326,333,334,335]. Anderson has provided evidence that eclipsed conformations of this type are also important for acetals [62]. High-level ab initio calculations on 2-cyclohexyloxytetrahydropyran indicate that the potential energy surface is fairly flat at Ψ angles of $0 \pm 50°$, and the *anti* conformer is much less stable ($\sim$16 kJ mol^{-1}) [336]. Thus, it is perhaps more appropriate to refer to a conformer with a C1–O1–Ci–Hi torsional angle close to 0° as a *syn* conformer.

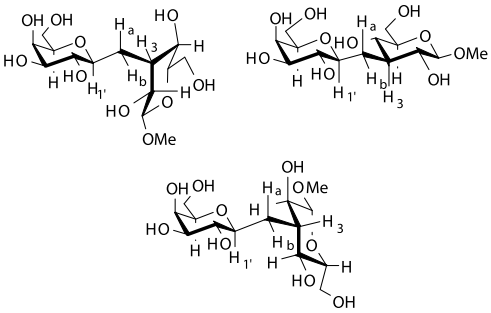


Figure 27
Estimation of steric energies in C- and *O*-disaccharides, for methyl β-D-galactopyranosyl-(1→3)-β-D-glucopyranoside and its C-disaccharide analogue: top left, *g*+ conformer; top right, *g*− conformer; bottom, *a* conformer

Substituents can markedly influence the inherent stabilities as shown in *Fig. 27* where the effects of adding equatorial hydroxyl groups for both *O*- and *C*-glycosides to give a typical 163-linked disaccharide are illustrated, with the values given in *Table 5*. This *O*-linked disaccharide has been observed by nOe measurements to have a small proportion of the *a* conformer along with the major *syn* and *g*− conformers [325]. In β-(164) linked derivatives, such as lactose or cellobiose, the interactions of the adjacent CH_2OH group become important and, in solution, the *a* conformer can become comparable in stability to, or even more favorable than, a *g* conformer [6,331,337,338]. Although a number of disaccharides and oligosaccharides have been found to have intermolecular hydrogen bonds in water, water/acetone, or dimethyl sulfoxide solutions by careful NMR measurements [241,252,259,262,326], no general criteria have yet been developed to predict their influence on disaccharide conformations. Although usually the most populated conformations are those bound by proteins [331,339,340,341], some oligosaccharides bind in conformations that contain a glycosidic linkage in an *a* conformer [338] or other conformations that are not highly populated in the free state [342,343]. A tethered disaccharide has been synthesized that is constrained to this conformation [344].

Calculation of the stabilities and geometries of conformers using the force fields of molecular mechanics or molecular dynamics programs provides an outline of the conformational possibilities that can be used to help interpret the experimental measurements [4,207,333,339,345,346]. It was concluded that most of the current force fields agree on the geometries of the lower energy conformers of disaccharides although there is disagreement about the relative stabilities of these minima [347].

Recently, Almond performed molecular dynamics calculations on a number of disaccharides and oligosaccharides in water [348] using the CHARMm forcefield modified for carbohydrates [349]. He observed a number of persistent intersaccharide hydrogen bonds that influenced the mixture of conformers predicted to be present. On the basis of these calculations, he made the important suggestion that disaccharides linked through α-linkages will be flexible with many hydrogen bonds to water, while disaccharides linked through β-linkages will be involved in intersaccharide hydrogen bonds in water and will be relatively inflexible [348].

Table 5
Steric interactions for methyl β-D-galactopyranosyl-(1→3)-β-D-glucopyranoside and its C-disaccharide analogue

Conformer	Interaction	Value for *C*-glycoside (kJ mol^{-1})	Value for *O*-glycoside (kJ mol^{-1})
g+	O-2C1′	10.4	10.4
	H-3O-5′	1.9	1.9
	C-2H-1′	3.8	3.8
	O-2′Hb	1.9	
	O-4Ha	1.9	
	C-1Hb	3.8	
	Total	23.7	16.1
g−	C-4O-5′	10.4	10.4
	O-4C-1′	10.4	10.4
	O-2′Hb	1.9	
	O-2Hb	1.9	
	Total	24.6	20.8
a	C-4H-1′	3.8	3.8
	C-2O-5′	10.4	10.4
	C-1′H-2	3.8	3.8
	C-1′H-4	3.8	3.8
	O-2Ha	1.9	
	O-2′Hb	1.9	
	O-4Hb	1.9	
	Total	27.5	21.8

1.4.2 Conformations of Trisaccharides and Oligosaccharides

No new factors arise when the conformations of trisaccharides or oligosaccharides are considered but intermonosaccharide interactions can become more important. For reviews, see [350,351,352]. However, discussion of these conformations using the terms employed with disaccharides is much more complicated because of the multiple ΦΨ maps and because of the possibility of coordinated motion. An alternative description uses a generalized order parameter, S^2, which is related to the spatial restriction of internal reorientation, with values ranging from 0 to 1 [353,354,355,356,357]. The dynamics of isotropically tumbling molecules are described in terms of overall and internal correlation times. The generalized order parameter [353] has been extended to include two distinct correlation times [358]. When the oligosaccharide is relatively rigid, S^2 for all segments is close to 1. Terminal monosaccharides in linear oligosaccharides often have lower values of S^2 than the central regions consistent with independent motion [359,360], as do hydroxymethyl groups [361].

2 Physical Methods

2.1 Introduction

The physical methods used for the analysis of carbohydrate conformations are becoming increasingly sophisticated as the molecules studied and the questions asked become larger and more complicated. X-ray crystallography provides a reference point, the conformations of compounds in the solid state. In solution, molecular modeling supplies a framework which can be used to interpret the results of the vast array of NMR spectral observations and those from other methods. For reviews, see [2,354].

2.2 X-Ray Crystallography

The determination of crystal structures of small molecules has now become routine provided that suitable crystals are available [362,363,364,365]. The advent of higher power X-ray sources, more synchronized light, and better recording devices has meant that smaller and smaller crystals can be solved by diffraction methods. At the same time, the analysis of patterns from larger and larger molecules has become easier so that the limit on the use of X-ray crystallography is the production of crystals.

The crystal structures of a large number of carbohydrate derivatives have been solved [1,106,366]. Crystal structures provide critical data on bond lengths, bond angles, torsional angles, intramolecular distances, etc., and this data has played and is continuing to play an important role in the discussion of the factors that influence conformation. Consideration of conformations of individual molecules using X-ray data must take into account intermolecular forces within the crystal lattice. Analysis of large sets of X-ray data from the Cambridge Data File avoids this difficulty and this method has influenced discussion of the anomeric effect [61,62,69]. Although most molecules crystallize in the conformation that is most highly populated in solution, this is not always true. For instance, 3-ammonio-3-deoxy-1,6-anhydro-β-D-glucopyranose is present in solution mostly in a boat conformation [184,367] but crystallizes in a chair [367]. Of increasing use is the determination of crystal structures of carbohydrates bound to proteins [368,369]. This allows analysis of the conformation of the bound carbohydrate and of the binding factors, both of which allow the synthesis of better binding molecules that may have medicinal applications.

2.3 NMR Spectroscopy

NMR spectroscopy is the most important technique for the examination of structure and conformation in solution [370,371,372,373]. Many aspects of NMR spectroscopy can be employed for the study of carbohydrates [130,351,374,375,376,377]. Assignment techniques are discussed in many books and will not be considered here. Chapters by Widmalm [354] and Serianni [376] include sections on assignment of carbohydrate spectra.

2.3.1 Chemical Shifts

^{1}H and ^{13}C NMR chemical shifts for unsubstituted and simple substituted carbohydrate derivatives and the effects of substituents on these shifts have been summarized [130,350,351,378,379,380,381]. The bulk of carbohydrate protons are observed between 3.2 and 4.2 ppm with equatorial protons being less shielded than axial and primary less shielded than secondary [130,375,382,383]. Anomeric protons are more deshielded (~4.3 to 5.5 ppm) and similarly anomeric carbons (~ 90 to 110 ppm) are deshielded with respect to the bulk of the secondary (~ 67 to 82 ppm) and primary carbons (~60–65 ppm). The factors that influence hydroxyl proton chemical shifts have been summarized and evaluated computationally [239,251].

2.3.2 Scalar Coupling

One-Bond Coupling The magnitudes of $^{1}J_{C,H}$ are about 125 Hz in saturated alkanes but increase in the presence of electronegative substituents [384] and are typically 140–150 Hz for non-anomeric carbons. The sizes of anomeric $^{1}J_{C,H}$ are related to the relative orientations of the bonds at the anomeric center of pyranoses, which are in turn related to CH bond lengths [76]; values are ~170–175 Hz if H-1 is equatorial (typically α-D anomers in $^{4}C_{1}$ conformations) but ~160–165 Hz if H-1 is axial (typically β-D anomers in $^{4}C_{1}$ conformations) [385,386,387]. Values from both furanoside anomers are normally similar, > 170 Hz [388,389] although conformational restriction makes the $^{1}J_{C,H}$ values for 2,3-anhydrofuranosides diagnostic of configuration [390]. For the same reasons, $^{1}J_{C,H}$ values for the anomeric CH units of septanoses are unreliable indicators of configuration [307]. Additional electronegative substituents increase the magnitude of $^{1}J_{C,H}$ again to 176–185 Hz in orthoesters [391]. In oligosaccharides, the sizes of anomeric $^{1}J_{C,H}$ are also influenced by the $\Phi\Psi$ angles in predictable ways [387].

Two-Bond Coupling A number of different two-bond couplings can provide useful structural and conformational information. The magnitudes of H–C–H coupling constants in saturated systems range from about 0 to −15 Hz and adjacent electronegative atoms cause them to increase algebraically in an orientationally dependent fashion [392], for instance, $^{2}J_{H,H}$ is ~−6 Hz in 1,3-dioxanes but > −2 Hz in 1,3-dioxolanes. $^{2}J_{H,H}$ has been used for conformational studies [196,393] and its magnitude at C-6 of pyranose sugars is related to both the C-5 C-6 torsional angle as well as the C-6 O-6 torsional angle [196]. It could be utilized more, particularly in the study of 1,6-linked oligosaccharides.
The magnitude of $^{2}J_{C,H}$ depends on substitution, electronegativity, and bond angle and can be positive or negative, increasing with increased numbers of electronegative substituents [394]. Values range from about −8 to +10 Hz in carbohydrates with larger values if one of the carbons is the anomeric carbon [395,396,397]. Rules that relate the size of $^{2}J_{C,H}$ to the orientation of the oxygen atoms have been formulated [395,397] and used to assign the configuration in *N*-acetylneuraminic acid and derivatives [398,399] and in oligosaccharides [400]. More recently, conformational information has been obtained by calculating these values for all possible conformational minima for comparison with experimental values for hydroxymethyl rotamers [197].

The effects on the size of $^{2}J_{C,C}$ across both oxygen and carbon atoms have been investigated experimentally and theoretically [401,402]. Across the anomeric oxygen, the magnitude of $^{2}J_{C,C}$ depends on the size of the Φ torsional angle much more than on the Ψ torsional angle but also depends on the C–O–C bond angle [402], making its use more difficult.

Three-Bond Coupling The magnitude of the coupling constant between protons on adjacent carbons, $^{3}J_{H,H}$, is related to the torsional angle between the protons [403] by the Karplus equation [404]. Several groups have modified the original equation to encode the effects of the orientations and electronegativities of electronegative substituents on the magnitude of $^{3}J_{H,H}$ but the equation of Haasnoot et al. [405], derived using $^{3}J_{H,H}$ values in six-membered rings, has seen the most use. The electronegativity values used with it have been modified and it has been reparameterized [406,407] to be:

$$^{3}J_{H,H} = 14.63\cos^{2}(\phi) - 0.78\cos(\phi) + \sum_{i}\lambda i\left\{0.34 - 2.31\cos^{2}[s_i(\phi) + 18.4|\lambda i|]\right\} \quad (4)$$

where Φ is the torsional angle, 8_i are the modified electronegativities, and s_i is the sign factor, either +1 or −1, defined according to the sign of the torsional angle [406]. Altona et al. have noted that use of this equation is limited to unstrained saturated systems but molecular orbital calculations have suggested that bond angle variation along the coupling path also influence the size of $^{3}J_{H,H}$ [408,409,410,411]. This latter factor may make Eq. (4) less accurate when extended to furanosides or acyclic systems.

$^{3}J_{H,H}$ values have been used extensively to evaluate which conformation is present. When conformational mixtures are present, the coupling constants for contributing conformers can be measured below the temperature of coalescence or estimated from the Haasnoot–Altona equation, Eq. (4). The position of the equilibrium is then evaluated by using the weighted average of the values for the contributors. This technique has been employed by many researchers [119,130,140,191,205,412], largely because it is accepted that the methods available for estimating values for contributors are reasonably accurate. It has been justified theoretically [413].

Karplus-type equations have been developed for many other systems of interest, most notably H–C–O-H [242,243], H–C–O–C [387,414,415], H–C–C-C [197,387], and C–O–C–C [416, 417]. It should be noted that coupling to hydroxyl protons can be observed in water or water/acetone or water/DMSO mixtures if the NMR tubes are rinsed in phosphate buffers and the carbohydrate derivatives are thoroughly deionized [239,249]. If the magnitudes of the coupling constants to carbon are determined on natural abundance material, considerable amounts are required. Alternatively, ^{13}C labeled material can be used [416,417,418,419] and starting materials with varieties of labels are available.

Long-Range Coupling Long range coupling constants ($^{4}J_{H,H}$ and $^{5}J_{H,H}$) were examined in the early literature on NMR of carbohydrates [420,421]. In saturated systems, they are largest over a W coupling pathway (< 2 Hz) and thus have some potential to reveal conformational information. They can be obtained conveniently using gradient-enhanced, two-dimensional homonuclear correlation techniques [422]. Long-range CH couplings ($^{4}J_{COCCH}$ and $^{4}J_{CCCCH}$) to enriched ^{13}C atoms have also proven useful for conformational studies [198].

2.3.3 Dipolar Coupling

In normal liquids, most molecules move isotropically and dipolar couplings are averaged to 0. In static solids and in liquid crystals, molecular motion is non-existent or anisotropic and dipolar couplings are observed. In solids, dipolar couplings are obscured by chemical shift anisotropy but in liquid crystals, they are observed and can be large, up to the kHz range [423]. The size of the direct dipolar coupling (D_{ij}) between nuclei i and j is:

$$D_{ij} = \frac{-K_{ij}S_{ij}}{r_{ij}^3} \tag{5}$$

where K_{ij} is equal to a constant times the product of γ_i and γ_j and S_{ij} is a traceless symmetric tensor, the order tensor, indicating how the spins are oriented with respect to the liquid crystal axis and the magnetic field [424]. Analysis of these spectra is complex since every spin 1/2 nucleus is coupled to every other nucleus. However, Tjandra and Bax showed that partial orientation of small proteins can be achieved in phospholipid bilayers known as bicelles [425]. The value of the order parameter increases with increasing bicelle concentrations allowing separation of dipolar and scalar couplings. This technique has been applied to carbohydrates [236,333,426,427,428,429,430,431]. Measurements of dipolar couplings provide a valuable complement to NOE measurements to define internuclear distances since the distance dependence of this interaction is r^{-3}, rather than r^{-6}. A concern is that interaction of the carbohydrate with the liquid crystal may alter the populations of individual conformers [432].

2.3.4 Nuclear Overhauser Effect

Techniques based on the nuclear Overhauser effect (nOe) are now the most important methods used for defining three-dimensional structures of carbohydrates. There is a recent edition [433] of earlier texts [434,435]. Overhauser originally predicted that saturation of electrons in a metal would polarize the metal nuclear spins [436] and the term nOe refers to the change in intensity of a signal resulting upon irradiation of another signal. For a two-spin system, *IS*, the nOe effect $f_I(S)$ is defined as the fractional change in the intensity of I on saturating S:

$$f_I(S) = \frac{(I - I^\circ)}{I^\circ} \tag{6}$$

where I° is the equilibrium intensity of I [435]. ❯ *Figure 28* shows a diagram of transition probabilities (W_{0IS}, W_{1I}, W_{1S}, and W_{2IS}) and spin states for the two-spin system. Solomon [437] showed that on saturating S:

$$f_I(S) = \frac{\gamma_S}{\gamma_I}\frac{W_{2IS} - W_{0IS}}{W_{0IS} + 2W_{1I} + W_{2IS}} = \frac{\gamma_S}{\gamma_I}\frac{\sigma_{IS}}{\Delta_{IS}} \tag{7}$$

where γ_I and γ_S are the magnetogyric ratios for the two nuclei, σ_{IS} is the cross-relaxation rate, and Δ_{IS} is the longitudinal relaxation rate. Relaxation for ^{1}H and ^{13}C nuclei normally occurs via the dipole-dipole mechanism where the motion of nearby magnets, usually ^{1}H nuclei that move with the molecular motion, creates a magnetic field that has a frequency component that matches the transition frequency. For small molecules in solution at room temperature, the

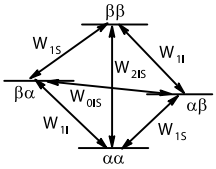


Figure 28
Energy level diagram for a two-spin system, IS, showing definitions of transition probabilities and spin states

average correlation time, τ_C, the time in which a molecule rotates one radian, is on the order of 10^{-11} s, giving a tumbling rate of about 600,000 MHz. For protons in a 500-MHz spectrometer, a tumbling rate of 500 MHz would lead to the most efficient W_1 transitions, and of 1000 MHz would lead to the most efficient W_2 transitions. The W_{0IS} transition corresponds to the frequency difference between the two nuclei which can be 0 to a few kHz for homonuclear cases but will be large for heteronuclear cases, comparable to W_1.

For relaxation in the spin pair *IS* by the dipolar mechanism in a molecule rotating equally in all directions (isotropically), it can be shown [433,438] that:

$$W_{0IS} = \frac{1/10K^2\,\tau_C}{1+(\omega_I-\omega_S)^2\,\tau_C^2}, \qquad W_{1I} = \frac{3/20K^2\tau_C}{1+\omega_I^2\,\tau_C^2},$$
$$W_{1S} = \frac{3/20K^2\,\tau_C}{1+\omega_S^2\tau_C^2}, \qquad W_{2IS} = \frac{3/5K^2\,\tau_C}{1+(\omega_I+\omega_S)^2\tau_C^2}, \tag{8}$$

where ω is the angular frequency, $2\pi\nu$, and $K = (\mu_o/4\pi)(h/2\pi)\gamma_I\gamma_S r_{IS}^{-3}$. When $\omega\tau_C$ is $\ll 1$, $\log W$ increases linearly with $\log\,\tau_C$ (extreme narrowing). Small molecules in non-viscous solvents generally have correlation times in this region. For homonuclear spin pairs, ω_I–ω_S is small and $\log W_{0IS}$ increases linearly with $\log\,\tau_C$ to large values of τ_C. Log W_1 and $\log W_{2IS}$ increase linearly with $\log\,\tau_C$ until $\omega\tau_C \approx 1$ when the slope of the curve changes sign to be linear with $-\log\,\tau_C$. The curves bend when molecular weights are in the 1000–2000 range, the molecular weight range of about pentasaccharides to decasaccharides. For homonuclear pairs, the maximum nOe, η_{max}, is 0.5 under extreme narrowing conditions but becomes -1 when tumbling is slow, typical of polysaccharides. Near $\omega\tau_C = 1$, η_{max} becomes 0 and nOes are small in this tumbling regime. This change over occurs in the tumbling regime where $\omega\tau_C$ changes from about 10^{-1} to 10. The difficulty with employment of nOes in this region can be overcome by using a spin-locking field to lock the spins in the rotating frame. Under these conditions, called a ROESY experiment in the 2D version, extreme narrowing enhancements are observed [439].

When more than two spins are present, the relationships for steady-state nOe become complicated by transfer of magnetization. However, the initial rate of buildup of nOe in multispin systems depends only on σ_{IS}, which for homonuclear relaxation is proportional to r_{IS}^{-6}. It is convenient to compare the initial σ_{IS} for a pair of protons of interest with those for a reference pair where the internuclear distance is known [433,440,441]. Then,

$$\frac{f_I(S)}{f_{ref}} = \frac{\sigma_{IS}}{\sigma_{ref}} = \left(\frac{r_{IS}}{r_{ref}}\right)^{-6}. \tag{9}$$

Interpretation of nOe results must be performed with caution because a number of difficulties, both technical, such as spin diffusion, and theoretical, such as the use of oversimplified or incorrect models can lead to incorrect interpretations.

2.4 Circular Dichroism and Optical Rotatory Dispersion

Only chiral molecules exhibit circular dichroism (CD) or optical rotatory dispersion (ORD) spectra. ORD is circular birefringence spectroscopy, that is, the difference in refractive index for circularly polarized light as a function of wavelength, or simply, the measure of optical rotation as a function of wavelength. Each individual UV absorption appears as an *S*-shaped or mirror-image *S*-shaped curve that is null at 8_{max}. CD is the difference in absorption of right and left circularly polarized light and appears as a Gaussian peak centered on 8_{max} for the UV absorption [442,443]. CD has been employed more for carbohydrates than ORD [444,445]. Since absorption maxima for most carbohydrates occur below 185 nm, vacuum CD must be used [444]. Considerable conformational information can be obtained because the appearance of the curves is very sensitive to the three-dimensional orientations of groups with respect to the chromophores. The most important factors in carbohydrate CD [445] are the anomeric configuration and the orientation about the Φ angle, normally exo-syn. Simple monosaccharides and some disaccharides have been studied [446] and conformational conclusions have been drawn for disaccharides and polymers [447,448,449]. Compounds containing other longer wavelength chromophores have also been studied [450,451].
Alternatively, chromophores can be added to carbohydrates through the formation of *para*-substituted benzoate esters or other functional groups containing chromophores. The high intensities of the CD spectra of these derivatives make this technique quite sensitive [452]. The orientations of the chromophores with respect to the sugar can be estimated and the resulting CD curves are very sensitive to substituent orientations. Thus, this technique is also useful for conformational studies [199,200,203,453].

2.5 Molecular Modeling

Molecular modeling has become an essential technique to outline the possibilities inherent for a particular system. The experimental observations, which are usually weighted averages, can then be interpreted in terms of these possibilities. Information about structure, conformation, and dynamics can be obtained. It is impossible to interpret complex sets of experimental observations, such as the multiple sets of average internuclear distances provided by nOe experiments, without knowledge about the geometries of two, three or more potential energy minima and without having some information about molecular dynamics. Methods for performing molecular modeling studies of carbohydrates are well established [329,347,454,455]. Initial studies used the hard sphere method to estimate non-bonded interactions [456] calculated with the potential of Kitaigorodsky [457]:

$$V = 3.5\left[-0.04(r_o/r)^6 + 8.6 \cdot 10^3 \exp\left[-13\left(\frac{r_o}{r}\right)\right]\right]. \tag{10}$$

Addition of terms to account for the exoanomeric effect in α and β glycosides gave the Hard Sphere Exo-Anomeric (HSEA) force field [458,459], which was very successful in predict-

ing conformations of the blood group oligosaccharides [460]. Force-fields have continued to develop [4,195,347,461,462,463,464,465,466,467]. Most can be summarized as follows:

$$E_{\text{potential}} = E_{\text{bond}} + E_{\text{angle}} + E_{\text{dihedral}} + E_{\text{non-bonded}} + E_{\text{electrostatic}} + \text{cross-terms} \quad (11)$$

where each term describes the increase in energy associated with change in geometry from an equilibrium value for a particular structural feature. For instance, bond stretching can be represented by a simple harmonic [454]:

$$E_{\text{bond}} = \frac{1}{2} K_{\text{b}}(l - l_o)^2 \quad (12)$$

where K_{b} is the force-constant for bond-stretching and l and l_{o} are the actual and the equilibrium bond lengths, respectively. The energy cost of deviations from equilibrium is more severe at greater deviations and higher order terms can be added to represent this [468]:

$$E_{\text{bond}} = \frac{1}{2} K_{\text{b}}(l - l_o)^2[1 - c_1(l - l_o) + c_2(l - l_o)^2] \quad (13)$$

where c_1 and c_2 are additional empirical parameters, or a Morse curve can be used [469]. Force-fields that only include harmonic terms and explicit diagonal elements in the force constant matrix are termed class 1 force fields. Class 2 force-fields add cubic and higher terms and contain off-diagonal elements in the force-constant matrix, that is, terms such as stretch-bend interactions, which detail how the effect of bond-angle bending is changed as the bonds are stretched [470].

Further steps in evaluation of conformational contributors include minimization of the energy from an initial geometry. Techniques for performing this are embedded in all modern force fields. The potential energy surface for carbohydrates contains many minima due to the exocyclic hydroxy, hydroxymethyl, and anomeric groups. Torsional angle driving, where one torsional angle is given an arbitrarily high force-constant, can be used in manual or automatic fashion. Once all minima have been identified, the contribution of individual conformers to the conformational ensemble can be calculated using the Boltzmann distribution.

Dihedral angle driving becomes extremely tedious beyond the monosaccharide stage [471]. Random methods can be used. In stochastic searching, a minimization is performed on the initial geometry, then every Cartesian coordinate of every atom is altered a random amount, then another minimization is performed. The second minimum is compared with the first and saved if it is different. Multiple repetitions allow a library to be built up [472]. Similar methods can be used with torsional angles rather than Cartesian coordinates. Monte Carlo searching imposes an energy test on whether the new conformation is saved and used as a new starting point [473,474,475]. These methods are useful if conformational minima only are required.

Solvent effects can be incorporated in various ways. The simplest is to treat the solvent as a bulk dielectric medium and adjust the equation used to evaluate electrostatic interactions for the dielectric constant [454]. A more sophisticated theory, the reaction-field model, is based on the theory of interactions of a multipole solute molecule in a polarizable continuum. The polar molecule induces a reaction field in the solvent which decreases the energy [476]. An alternative approach is to describe the effect of solvation as follows:

$$\Delta G^{\circ}_{\text{solvation}} = \Delta E_{\text{internal}} + G_{\text{P}} + G_{\text{CDS}} \quad (14)$$

where G_P accounts for the favorable effects of mutual electrostatic polarization of the solute and the solvent minus the cost of distorting the solvent, $\Delta E_{internal}$ accounts for the cost of internal distortion of the solute and G_{CDS}, which depends on the solvent-accessible surface area, includes the effects of interactions dominated by the first solvation shell, such as cavitation [211,477,478]. Another approach is to use Langevin dynamics [333,479,480]. This area is under active development [478,481,482,483,484,485].

Molecular dynamics allows the examination of the time dependence of a system that includes a number of solute and solvent molecules in a cell [486,487,488]. The system of particles is termed the ensemble and the number of particles, the volume and either the energy or the temperature are kept constant. The time evolution of the ensemble is obtained from Newton's second law of motion:

$$\frac{dv}{dt} = \frac{F}{m} \tag{15}$$

where the acceleration is dv/dt, F is the force, and m is the atomic mass. Very small time steps keep the energy constant. The force is obtained from the molecular mechanics force field,

$$\frac{dE_{pot}}{dr} = -F\,. \tag{16}$$

First derivatives of the potential energy with respect to the Cartesian coordinates are determined as part of the minimization process. The total energy of the system is the sum of the potential and kinetic energies. The kinetic energy is proportional to the temperature:

$$E_{kinetic} = \frac{1}{2}mv^2 = \frac{3}{2}kT \tag{17}$$

and this can be used to define the temperature of the ensemble. A difficulty is handling the borders of the cell and various approaches have been used, including periodic border conditions, where the cell is symmetrically replicated in all directions.

The necessity of using small time steps means that many steps must be calculated in order to adequately sample conformational space. Therefore, molecular dynamics calculations are normally performed with class 1 force fields. Molecular dynamics simulations are analyzed by following the trajectory of a variable with time, often $\Phi\Psi$ angles for oligosaccharides. Most modeling programs now can perform molecular dynamics calculations [347,462,466,489].

Quantum mechanical calculations [490,491,492,493] are now fast and accurate enough that they have become the method of choice for studying conformations of individual monosaccharides and disaccharides [231,494,495,496,497]. The use of density functional theory [498] allows the study of much larger systems. High-level methods have to be employed to yield accurate results in terms of energies [499]. The continuing improvements in computational power will result in acceleration of the use of these methods.

Acknowledgement

I would like to thank Al French for comments on this chapter from the previous edition.

References

1. Rao VSR, Qasba PK, Balaji PV, Chandrasekaran R (1998) Conformation of Carbohydrates. Harwood Academic Publishers, Amsterdam
2. Stoddart JF (1971) Stereochemistry of Carbohydrates. Wiley-Interscience, New York
3. Dais P (1995) Adv Carbohydr Chem Biochem 51:63
4. Vliegenthart JFG, Woods RJ (2006) NMR Spectroscopy and Computer Modeling of Carbohydrates: Recent Advances. American Chemical Society, Washington, DC
5. Imberty A, Pérez S (2000) Chem Rev 100:4567
6. Jiménez-Barbero J, Espinosa JF, Asensio JL, Cañada FJ, Poveda A (2001) Adv Carbohydr Chem Biochem 56:235
7. French AD (1998) Carbohydrates. Wiley, New York, p 233
8. Eliel EL, Wilen SH, Mander LN (1994) Stereochemistry of Organic Compounds. Wiley, New York
9. Juaristi E (1995) Conformational Behaviour of Six-Membered Rings. VCH Publishers, New York
10. Moss GP (1996) Pure Appl Chem 68:2193
11. Nasipuri D (1991) Stereochemistry of Organic Compounds: Principles and Applications. Wiley, New York
12. Cremer D, Pople JA (1975) J Am Chem Soc 97:1354
13. Cremer D, Szabo KJ (1995) Ab Initio Studies of Six-membered Rings: Present Status and Future Developments. In: Juaristi E (ed) Conformational Behaviour of Six-Membered Rings. VCH Publishers, New York, p 59
14. Haasnoot CAG (1992) J Am Chem Soc 114:882
15. Anet FAL, Anet R (1975) Conformational Processes in Rings. In: Jackman LM, Cotton FA (eds) Dynamic Nuclear Magnetic Resonance Spectroscopy. Academic Press, New York, p 543
16. Höfner D, Lesko SA, Binsch G (1978) Org Magn Reson 11:179
17. Squillacote M, Sheridan RS, Chapman OL, Anet FAL (1975) J Am Chem Soc 97:3244
18. Bushweller CH (1995) Stereodynamics of Cyclohexane and Substituted Cyclohexanes. Substituent A Values. In: Juaristi E (ed) Conformational Behaviour of Six-Membered Rings. VCH Publishers, New York, p 25
19. Booth H, Everett JR (1980) J Chem Soc, Perkin Trans 2 255
20. Wiberg KB, Hammer JD, Castejon H, Bailey WF, DeLeon EL, Jarret RM (1999) J Org Chem 64:2085
21. Kitching W, Olszowy HA, Adcock W (1981) Org Magn Reson 15:230
22. Eliel EL, Gilbert EC (1969) J Am Chem Soc 91:5487
23. Schneider H-J, Hoppen V (1978) J Org Chem 43:3866
24. Jensen FR, Bushweller CH, Beck BH (1969) J Am Chem Soc 91:344
25. Eliel EL, Reese MC (1968) J Am Chem Soc 90:1560
26. Tichý M, Sipos F, Sicher J (1966) Coll Czech Chem Commun 31:2889
27. Booth H, Dixon JM, Khedhair KA (1992) Tetrahedron 48:6161
28. Corey EJ, Feiner NF (1980) J Org Chem 45:765
29. Angyal SJ, Mchugh DJ (1956) Chem Ind 1147
30. Eliel EL, Haubenstock H (1961) J Org Chem 26:3504
31. Tichý M, Orahovats A, Sicher J (1970) Coll Czech Chem Commun 35:459
32. Tavernier D, De Pessemier F, Anteunis M (1975) Bull Soc Chim Belg 84:333
33. Lemieux RU, Chu P (1958) ACS Meeting, San Francisco, CA
34. Manoharan M, Eliel EL (1983) Tetrahedron Lett 24:453
35. Booth H, Grindley TB (1983) J Chem Soc, Chem Commun 1013
36. Subbotin OA, Sergeev NM, Zefirov NS, Gurvich LG (1975) J Org Chem USSR 11:2233
37. Rockwell GD, Grindley TB (1996) Austral J Chem 49:379
38. Tichý M, Sicher J (1967) Coll Czech Chem Commun 32:3687
39. Subbotin OA, Sergeev NM (1978) J Org Chem USSR 14:1388
40. Jordan EA, Thorne MP (1986) Tetrahedron 42:93
41. Wiberg KB, Castejon H, Bailey WF, Ochterski J (2000) J Org Chem 65:1181
42. Riddell FG (1980) The Conformational Analysis of Heterocyclic Compounds. Academic Press, London
43. Kleinpeter E (1998) Adv Het Chem 69:217
44. Lambert JB, Mixan CE, Johnson DH (1973) J Am Chem Soc 95:4634

45. Eliel EL, Hargrave KD, Pietrusiewicz KM, Manoharan M (1982) J Am Chem Soc 104:3635
46. Szarek WA, Horton D (1979) Anomeric Effect: Origin and Consequences. ACS, Washington, DC
47. Thatcher GRJ (1993) Anomeric Effect and Associated Stereoelectronic Effects. ACS, Washington, DC
48. Kirby AJ (1983) The Anomeric Effect and Related Stereoelectronic Effects at Oxygen. Springer, Berlin Heidelberg New York
49. Tvaroška I, Bleha T (1989) Adv Carbohydr Chem Biochem 47:45
50. Juaristi E, Cuevas G (1994) The Anomeric Effect. CRC Press, Boca Raton, FL
51. Franck RW (1995) Stereoelectronic Effects in Six-membered Rings. In: Juaristi E (ed) Conformational Behaviour of Six-Membered Rings. VCH Publishers, New York, p 159
52. Bitzer RS, Barbosa AGH, da Silva CO, Nascimento MAC (2005) Carbohydr Res 340:2171
53. Graczyk PP, Mikolajczyk M (1994) Topics Stereochem 21:159
54. Edwards JT (1955) Chem Ind 1102
55. Eliel EL, Giza CA (1968) J Org Chem 33:3754
56. Franck RW (1983) Tetrahedron 39:3251
57. Booth H, Dixon JM, Readshaw SA (1992) Tetrahedron 48:6151
58. Romers C, Altona C, Buys HR, Havinga E (1969) Topics Stereochem 4:39
59. Brundle CR, Turner DW (1968) Proc Roy Soc A 307:27
60. Laing M (1987) J Chem Ed 64:124
61. Cossé-Barbi A, Watson DG, Dubois J-E (1989) Tetrahedron Lett 30:163
62. Anderson JE (2000) J Org Chem 65:748
63. Astrup EE (1973) Acta Chem Scand 27:327
64. Yokoyama Y, Ohashi Y (1999) Bull Chem Soc Jpn 72:2183
65. Petillo PA, Lerner LE (1993) Origin and Quantitative Modeling of the Anomeric Effect. In: Thatcher GRJ (ed) Anomeric Effect and Associated Stereoelectronic Effects. ACS, Washington, DC, p 156
66. Salzner U, von Rague Schleyer P (1994) J Org Chem 59:2138
67. Omoto K, Marusaki K, Hirao H, Imade M, Fujimoto H (2000) J Phys Chem A 104:6499
68. Cossé-Barbi A, Dubois J-E (1987) J Am Chem Soc 109:1503
69. Schleifer L, Senderowitz H, Aped P, Tartakovsky E, Fuchs B (1990) Carbohydr Res 206:21
70. Pierson GO, Runquist OA (1968) J Org Chem 33:2572
71. Praly J-P, Lemieux RU (1987) Can J Chem 65:213
72. Ouedrago A, Lessard J (1991) Can J Chem 69:474
73. Kirby AJ, Williams NH (1992) J Chem Soc, Chem Commun 1286
74. Briggs AJ, Glenn R, Jones PG, Kirby AJ, Ramaswamy P (1984) J Am Chem Soc 106:6200
75. Uehara F, Sato M, Kaneko C, Kurihara H (1999) J Org Chem 64:1436
76. Wolfe S, Pinto BM, Varma V, Leung RYN (1990) Can J Chem 68:1051
77. Juaristi E, Cuevas G (1992) Tetrahedron 48:5019
78. Cuevas G, Martínez-Mayorga K, Fernandez-Alonso MD, Jiménez-Barbero J, Perrin CL, Juaristi E, López-Mora NS (2005) Angew Chem Int Ed 44:2360
79. Lewis BE, Schramm VL (2001) J Am Chem Soc 123:1327
80. Box VGS (1990) Heterocycles 31:1157
81. Box VGS (1998) Heterocycles 48:2389
82. Perrin CL, Armstrong KB, Fabian MA (1994) J Am Chem Soc 116:715
83. Wiberg KB, Marquez M (1994) J Am Chem Soc 116:2197
84. Lemieux RU, Pavia AA, Martin JC, Watanabe KA (1969) Can J Chem 47:4427
85. Jorgensen WL, de Tirago PIM, Severance DL (1994) J Am Chem Soc 116:2199
86. Salzner U (1995) J Org Chem 60:986
87. Pinto BM, Wolfe S (1982) Tetrahedron Lett 23:3687
88. Fuchs B (1978) Topics Stereochem 19:1
89. Pitzer KS, Donath WE (1959) J Am Chem Soc 81:3213
90. Engerholm GR, Luntz AC, Gwinn WD, Harris DO (1969) J Chem Phys 50:2446
91. Meyer R, Lopez JC, Alonso JL, Melandri S, Favero PG, Caminati W (1999) J Chem Phys 111:7871
92. Mamleev AH, Gunderova LN, Galeev RV (2001) J Struct Chem 42:365
93. Greenhouse JA, Strauss HL (1969) J Chem Phys 50:124
94. Davidson R, Warsop PA (1972) J Chem Soc, Faraday Trans 2 1875
95. Rayón VM, Sordo JA (2005) J Chem Phys 122:204303
96. Wu AA, Cremer D (2003) Int J Mol Sci 4:158
97. Luger P, Buschmann J (1983) Angew Chem Int Ed Eng 22:410

98. David WIF, Ibberson RM (1992) Acta Crystallogr C48:301
99. Cadioli B, Gallinella E, Coulombeau C, Jobic H, Berthier G (1993) J Phys Chem 97:7844
100. Strajbl M, Baumruk V, Florián J (1998) J Phys Chem B 102:1314
101. Lavrich RJ, Rhea RL, McCargar JW, Tubergen MJ (2000) J Mol Spectrosc 199:138
102. Lavrich RJ, Torok CR, Tubergen MJ (2001) J Phys Chem A 105:8317
103. Berman HM, Rosenstein RD (1968) Acta Crystallogr B24:435
104. Kim HS, Jeffrey GA, Rosenstein RD (1968) Acta Crystallogr B24:1449
105. Jeffrey GA, Kim HS (1970) Carbohydr Res 14:207
106. Jeffrey GA (1990) Acta Crystallogr B46:89
107. Köll P, Kopf J, Morf M, Zimmer B, Brimacombe JS (1992) Carbohydr Res 237:289
108. Angyal SJ, Saunders JK, Grainger CT, Le Fur R, Williams PG (1986) Carbohydr Res 150:7
109. Kopf J, Morf M, Hagen B, Bischoff M, Köll P (1994) Carbohydr Res 262:9
110. Köll P, Bischoff M, Bretzke C, Kopf J (1994) Carbohydr Res 262:1
111. Köll P, Bruns R, Kopf J (1997) Carbohydr Res 305:147
112. Köll P, Kopf J (1996) Austral J Chem 49:391
113. Schouten A, Kanters JA, Kroon J, Comini S, Looten P, Mathlouthi M (1998) Carbohydr Res 312:131
114. Kopf J, Bischoff M, Köll P (1991) Carbohydr Res 217:1
115. Kopf J, Morf M, Zimmer B, Köll P (1992) Carbohydr Res 233:35
116. Horton D, Wander JD (1974) J Org Chem 39:1859
117. Hawkes GE, Lewis D (1984) J Chem Soc, Perkin Trans 2 2073
118. Lewis D (1986) J Chem Soc, Perkin Trans 2 467
119. Lewis D, Angyal SJ (1989) J Chem Soc, Perkin Trans 2 1763
120. Angyal SJ, Le Fur R (1980) Carbohydr Res 84:201
121. Schnarr GW, Vyas DM, Szarek WA (1979) J Chem Soc, Perkin Trans 1 496
122. Franks F, Kay RL, Dadok J (1988) J Chem Soc, Faraday Trans 1 84:2595
123. Gallwey FB, Hawkes JE, Haycock P, Lewis D (1990) J Chem Soc, Perkin Trans 2 1979
124. IUPAC-IUB Joint Commission on Biochemical Nomenclature (1981) Pure Appl Chem 1902
125. Angyal SJ (1968) Austral J Chem 21:2737
126. Angyal SJ (1969) Angew Chem Int Ed Eng 8:157
127. Augé J, David S (1984) Tetrahedron 40:2101
128. Casu B, Petitou M, Provasoli M, Sinaÿ P (1988) TIBS 13:221
129. Angyal SJ, Pickles VA (1972) Austral J Chem 25:1695
130. Bock K, Thøgersen H (1982) Ann Rep NMR Spectrosc 13:1
131. Snyder JR, Serianni AS (1986) J Org Chem 51:2694
132. Snyder JR, Johnston ER, Serianni AS (1989) J Am Chem Soc 111:2681
133. Dowd MK, French AD, Reilly PJ (1994) Carbohydr Res 264:1
134. Tobiason FL, Swank DD, Vergoten G, Legrande P (2000) J Carbohydr Chem 19:959
135. Kurihara Y, Ueda K (2006) Carbohydr Res 341:2565
136. Rudrum M, Shaw DF (1965) J Chem Soc 52
137. Lemieux RU, Stevens JD (1966) Can J Chem 44:249
138. Franks F, Lillford PJ, Robinson G (1989) J Chem Soc, Faraday Trans 1 85:2417
139. Srivastava RM, Pavão AC, Seabra GM, Brown RK (1997) J Mol Struct 412:51
140. Durette PL, Horton D (1971) Adv Carbohydr Chem Biochem 26:49
141. Hall LD, Manville JF (1969) Can J Chem 47:19
142. Probst KC, Wessel HP (2001) J Carbohydr Chem 20:549
143. Anderson CB, Sepp DT (1967) J Org Chem 32:607
144. Bishop CT, Cooper FP (1963) Can J Chem 41:2743
145. Smirnyagin V, Bishop CT (1968) Can J Chem 46:3085
146. Lemieux RU (1964) In: de Mayo P (ed) Molecular Rearrangements, Part 2. Wiley, New York, p 735
147. Lemieux RU, Morgan AR (1965) Can J Chem 43:2205
148. Lemieux RU (1971) Pure Appl Chem 25:527
149. Paulsen H, Györgydeák Z, Friedmann M (1974) Chem Ber 107:1590
150. Vaino AR, Chan SSC, Szarek WA, Thatcher GRJ (1996) J Org Chem 61:4514
151. Perrin CL (1995) Tetrahedron 51:11901
152. Perrin CL, Fabian MA, Brunckova J, Ohta BK (1999) J Am Chem Soc 121:6911
153. Perrin CL, Kuperman J (2003) J Am Chem Soc 125:8846

154. Randell KD, Johnston BD, Green DF, Pinto BM (2000) J Org Chem 65:220
155. Jones PG, Kirby AJ, Komarov IV, Wothers PD (1998) Chem Commun 1695
156. Vaino AR, Szarek WA (2001) J Org Chem 66:1097
157. Chan SSC, Szarek WA, Thatcher GRJ (1995) J Chem Soc, Perkin Trans 2 45
158. Ernst S, Venkataraman G, Sasisekharan V, Langer R, Cooney CL, Sasisekharan R (1998) J Am Chem Soc 120:2099
159. Hricovíni M (2006) Carbohydr Res 341:2575
160. Batta G, Lipták A, Schneerson R, Pozsgay V (1997) Carbohydr Res 305:93
161. Flippen JL (1973) Acta Crystallogr B29:1881
162. Ogura H, Furuhata K, Saitô H, Izumi G, Itoh M, Shitori Y (1984) Chem Lett 1003
163. Jaques LW, Brown EB, Barrett JM, Brey WS Jr, Weltner W Jr (1977) J Biol Chem 252: 4533
164. Dabrowski U, Friebolin H, Brossmer R, Supp M (1979) Tetrahedron Lett 4637
165. Beau J-M, Schauer R, Haverkamp J, Dorland L, Vliegenthart JFG (1980) Carbohydr Res 82:125
166. Brown EB, Brey WS Jr, Weltner W Jr (1975) Biochim Biophys Acta 399:124
167. Czarniecki MF, Thornton ER (1977) J Am Chem Soc 99:8273
168. Batta G, Gervay J (1995) J Am Chem Soc 117:368
169. Unger FM (1981) Adv Carbohydr Chem Biochem 38:324
170. Wilkinson SG (1996) Prog Lipid Res 35:283
171. Becker B, Lommerse JPM, Melkonian M, Kamerling JP, Vliegenthart JFG (1995) Carbohydr Res 267:313
172. Becker B, Melkonian M, Kamerling JP (1998) J Phycol 34:779
173. Jann K (1983) Exopolysaccharides of E. Coli. In: Anderson L, Unger FM (eds) Bacterial Polysaccharides: Structure, Synthesis, and Biological Activities. ACS, Washington, DC, p 171
174. Birnbaum GI, Roy R, Brisson JR, Jennings HJ (1987) J Carbohydr Chem 6:17
175. Sengupta D, van Derveer D (1985) Ind J Chem 24B:1268
176. Kratky C, Stix D, Unger FM (1981) Carbohydr Res 92:299
177. Mikol V, Kosma P, Brade H (1994) Carbohydr Res 263:35
178. Cherniak R, Jones RG, Gupta DS (1979) Carbohydr Res 75:39
179. Brade H, Zähringer U, Rietschel ET, Christian R, Schulz G, Unger FM (1984) Carbohydr Res 134:157
180. Coxon B (1999) Carbohydr Res 322:120
181. Angyal SJ, Dawes K (1968) Austral J Chem 21:2747
182. Cerný M, Stanek J Jr (1977) Adv Carbohydr Chem Biochem 34:24
183. Grindley TB, Cude A, Kralovec J, Thangarasa R (1994) The Chair-Boat Equilibrium of 1,6-Anhydro-β-D-glucopyranose and Derivatives: An NMR and Molecular Mechanics Study. In: Witczak ZJ (ed) Levoglucosenone and Levoglucosans, Chemistry and Applications. ATL Press, Mount Prospect, ILL, p 147
184. Trnka T, Cerný M, Budesínský M, Pacák J (1975) Coll Czech Chem Commun 40:3038
185. Noordik JH, Jeffrey GA (1977) Acta Crystallogr B33:403
186. Paulsen H, Koebernick H (1976) Chem Ber 109:104
187. van Rijsbergen R, Anteunis MJO, De Bruyn A (1982) Bull Soc Chim Belg 91:297
188. Dais P, Shing TKM, Perlin AS (1984) J Am Chem Soc 106:3082
189. Li C, Bernet B, Vasella A (1991) Carbohydr Res 216:149
190. Wessel HP (1992) J Carbohydr Chem 11:1039
191. Bock K, Duus JØ (1994) J Carbohydr Chem 13:513
192. Jensen HH, Nordstrom LU, Bols M (2004) J Am Chem Soc 126:9205
193. Jensen HH, Bols M (2006) Acc Chem Res 39:259
194. Marchessault RH, Pérez S (1979) Biopolymers 18:2369
195. Kouwijzer MLCE, Grootenhuis PDJ (1995) J Phys Chem 99:13426
196. Stenutz R, Carmichael I, Widmalm G, Serianni AS (2002) J Org Chem 67:949
197. Thibaudeau C, Stenutz R, Hertz B, Klepach T, Zhao S, Wu QQ, Carmichael I, Serianni AS (2004) J Am Chem Soc 126:15668
198. Pan QF, Klepach T, Carmichael I, Reed M, Serianni AS (2005) J Org Chem 70:7542
199. Nobrega C, Vázquez JT (2003) Tetrahedron: Asym 14:2793
200. Roen A, Padron JI, Vázquez JT (2003) J Org Chem 68:4615
201. Jockusch RA, Kroemer RT, Talbot FO, Simons JP (2003) J Phys Chem A 107:10725
202. Jockusch RA, Talbot FO, Simons JP (2003) Phys Chem Chem Phys 5:1502

203. Padrón JI, Morales EQ, Vázquez JT (1998) J Org Chem 63:8247
204. de Vries NK, Buck HM (1987) Carbohydr Res 165:1
205. Rockwell GD, Grindley TB (1998) J Am Chem Soc 120:10953
206. de Vries NK, Buck HM (1987) Recl Trav Chim Pays-Bas 106:453
207. Gonzalez-Outeiriño J, Kirschner KN, Thobhani S, Woods RJ (2006) Can J Chem 84:569
208. Kirschner KN, Woods RJ (2001) Proc Nat Acad Sci 98:10541
209. Tvaroska I, Taravel FR, Utille JP, Carver JP (2002) Carbohydr Res 337:353
210. Barrows SE, Storer JW, Cramer CJ, French AD, Truhlar DG (1998) J Comp Chem 19:1111
211. Cramer CJ, Truhlar DG (1999) Chem Rev 99:2161
212. Dolney DM, Hawkins GD, Winget P, Liotard DA, Cramer CJ, Truhlar DG (2000) J Comp Chem 21:340
213. Momany FA, Appell M, Willett JL, Bosma WB (2005) Carbohydr Res 340:1638
214. Momany FA, Appell M, Willett JL, Schnupf U, Bosma WB (2006) Carbohydr Res 341:525
215. Tvaroška I, Taravel FR, Utille JP, Carver JP (2002) Carbohydr Res 337:353
216. Suzuki T, Kawashima H, Sota T (2006) J Phys Chem B 110:2405
217. Baumann H, Brisson JR, Michon F, Pon R, Jennings HJ (1993) Biochemistry 32:4007
218. Sawada T, Hashimoto T, Nakano H, Shigematsu M, Ishida H, Kiso M (2006) J Carbohydr Chem 25:387
219. Brisson JR, Baumann H, Imberty A, Pérez S, Jennings HJ (1992) Biochemistry 31:4996
220. Vasudevan SV, Balaji PV (2002) Biopolymers 63:168
221. Yamasaki R, Bacon B (1991) Biochemistry 30:851
222. Henderson TJ, Venable RM, Egan W (2003) J Am Chem Soc 125:2930
223. Bhattacharjee AK, Jennings HJ, Kenny CP (1978) Biochemistry 17:645
224. Lees RM, Baker JG (1968) J Chem Phys 48:5299
225. Lowe JP (1968) Prog Phys Org Chem 6:1
226. Barrows SE, Dulles FJ, Cramer CJ, French AD, Truhlar DG (1995) Carbohydr Res 276:219
227. Gregurick SK, Kafafi SA (1999) J Carbohydr Chem 18:867
228. Suzuki T, Sota T (2005) J Phys Chem B 109:12603
229. Talbot FO, Simons JP (2002) Phys Chem Chem Phys 4:3562
230. Simons JP, Jockusch RA, Carcabal P, Hung I, Kroemer RT, Macleod NA, Snoek LC (2005) Int Rev Phys Chem 24:489
231. Carcabal P, Patsias T, Hunig I, Liu B, Kaposta C, Snoek LC, Gamblin DP, Davis BG, Simons JP (2006) Phys Chem Chem Phys 8:129
232. Carcabal P, Hunig I, Gamblin DP, Liu B, Jockusch RA, Kroemer RT, Snoek LC, Fairbanks AJ, Davis BG, Simons JP (2006) J Am Chem Soc 128:1976
233. Carcabal P, Jockusch RA, Hunig I, Snoek LC, Kroemer RT, Davis BG, Gamblin DP, Compagnon I, Oomens J, Simons JP (2005) J Am Chem Soc 127:11414
234. Barron LD, Hecht L, Mccoll IH, Blanch EW (2004) Mol Phys 102:731
235. Macleod NA, Johannessen C, Hecht L, Barron LD, Simons JP (2006) Int J Mass Spect 253:193
236. Martin-Pastor M, Canales A, Corzana F, Asensio JL, Jiménez-Barbero J (2005) J Am Chem Soc 127:3589
237. Gillet B, Nicole D, Delpuech J-J, Gross B (1981) Org Magn Reson 17:28
238. Angyal SJ, Christofides JC (1996) J Chem Soc, Perkin Trans 2 1485
239. Sandström C, Kenne L (2006) Hydroxy Protons in Structural Studies of Carbohydrates by NMR Spectroscopy. In: Vliegenthart JFG, Woods RJ (eds) NMR Spectroscopy and Computer Modeling of Carbohydrates: Recent Advances. ACS, Washington, DC, p 114
240. Bernet B, Vasella A (2000) Helv Chim Acta 83:995
241. Bernet B, Vasella A (2000) Helv Chim Acta 83:2055
242. Fraser RR, Kaufman M, Morand P, Govil G (1969) Can J Chem 47:403
243. Rader CP (1969) J Am Chem Soc 91:3248
244. Christofides JC, Davies DB (1987) J Chem Soc, Perkin Trans 2 97
245. Vasquez TE, Bergset JM, Fierman MB, Nelson A, Roth J, Khan SI, O'Leary DJ (2002) J Am Chem Soc 124:2931
246. Muddasani PR, Bozó E, Bernet B, Vasella A (1994) Helv Chim Acta 77:257
247. Adams B, Lerner LE (1994) Magn Reson Chem 32:225
248. Batta G, Kövér KE (1999) Carbohydr Res 320:267
249. Sandström C, Baumann H, Kenne L (1998) J Chem Soc, Perkin Trans 2 809

250. Sandström C, Baumann H, Kenne L (1998) J Chem Soc, Perkin Trans 2 2385
251. Bekiroglu S, Sandström A, Kenne L, Sandström C (2004) Org Biomol Chem 2:200
252. Bekiroglu S, Kenne L, Sandström C (2004) Carbohydr Res 339:2465
253. Siebert HC, Andre S, Vliegenthart JFG, Gabius HJ, Minch MJ (2003) J Biomol NMR 25:197
254. Adams B, Lerner LE (1992) J Am Chem Soc 114:4827
255. Leeflang BR, Vliegenthart JFG, Kroon-Batenburg LMJ, van Eijck BP, Kroon J (1992) Carbohydr Res 230:41
256. Bock K, Frejd T, Kihlberg J, Magnusson G (1988) Carbohydr Res 176:253
257. Poppe L, van Halbeek H (1991) J Am Chem Soc 113:363
258. Sandström C, Magnusson G, Nilsson U, Kenne L (1999) Carbohydr Res 322:46
259. Bekiroglu S, Kenne L, Sandström C (2003) J Org Chem 68:1671
260. Hakkarainen B, Fujita K, Immel S, Kenne L, Sandström C (2005) Carbohydr Res 340:1539
261. Asensio JL, Cañada FJ, Garcia-Herrero A, Murillo MT, Fernández-Mayoralas A, Johns BA, Kozak J, Zhu ZZ, Johnson CR, Jiménez-Barbero J (1999) J Am Chem Soc 121:11318
262. Ivarsson I, Sandström C, Sandström A, Kenne L (2000) J Chem Soc, Perkin Trans 2 2147
263. Lemieux RU, Koto S, Voisin D (1979) The Exoanomeric Effect. In: Szarek WA, Horton D (eds) Anomeric Effect, Origin and Consequences. ACS, Washington, DC, p 17
264. Marchessault RH, Pérez S (1978) Carbohydr Res 65:114–120
265. Cramer CJ, Truhlar DG, French AD (1997) Carbohydr Res 298:1
266. Tvaroška I, Carver JP (1998) Carbohydr Res 309:1
267. Espinosa JF, Cañada FJ, Asensio JL, Martín-Pastor M, Dietrich H, Martín-Lomas M, Schmidt RR, Jiménez-Barbero J (1996) J Am Chem Soc 118:10862
268. Asensio JL, Cañada FJ, Chen XH, Khan N, Mootoo DR, Jiménez-Barbero J (2000) Chem Eur J 6:1035
269. Houk KN, Eksterowicz JE, Wu Y-D, Fuglesang CD, Mitchell DB (1993) J Am Chem Soc 115:4170
270. Wu T-C, Goekjian PG, Kishi Y (1987) J Org Chem 52:4819
271. Hall LD, Steiner PR, Pedersen C (1970) Can J Chem 48:1155
272. Altona C, Sundaralingam M (1972) J Am Chem Soc 94:8205
273. Saenger W (1984) Principles of Nucleic Acid Structure. Springer, Berlin Heidelberg New York
274. Altona C, Geise HJ, Romers C (1968) Tetrahedron 24:13
275. Gelbin A, Schneider B, Clowney L, Hsieh SH, Olson WK, Berman HM (1996) J Am Chem Soc 118:519
276. Plavec J, Thibaudeau C, Chattopadhyaya J (1996) Pure Appl Chem 68:2137
277. Podlasek CA, Stripe WA, Carmichael I, Shang MY, Basu B, Serianni AS (1996) J Am Chem Soc 118:1413
278. Felli IC, Richter C, Griesinger C, Schwalbe H (1999) J Am Chem Soc 121:1956
279. Polak M, Seley KL, Plavec J (2004) J Am Chem Soc 126:8159
280. Plevnik M, Crnugelj M, Stimac A, Kobe J, Plavec J (2001) J Chem Soc, Perkin Trans 2 1433
281. Plavec J, Roselt P, Földesi A, Chattopadhyaya J (1998) Magn Reson Chem 36:732
282. Polak M, Mohar B, Kobe J, Plavec J (1998) J Am Chem Soc 120:2508
283. Verberckmoes F, Esmans EL (1995) Spectrochim Acta A 51:153
284. Olson WK, Sussman JL (1982) J Am Chem Soc 104:270
285. Houseknecht JB, McCarren PR, Lowary TL, Hadad CM (2001) J Am Chem Soc 123:8811
286. Houseknecht JB, Lowary TL, Hadad CM (2003) J Phys Chem A 107:5763
287. Houseknecht JB, Lowary TL, Hadad CM (2003) J Phys Chem A 107:372
288. Evdokimov A, Gilboa AJ, Koetzle TF, Klooster WT, Schultz AJ, Mason SA, Albinati A, Frolow F (2001) Acta Crystallogr B57:213
289. Suzuki T, Sota T (2005) J Phys Chem B 109:12603
290. Suzuki T, Kawashima H, Kotoku H, Sota T (2005) J Phys Chem B 109:12997
291. Gordon MT, Lowary TL, Hadad CM (1999) J Am Chem Soc 121:9682
292. Gordon MT, Lowary TL, Hadad CM (2000) J Org Chem 65:4954
293. Temeriusz A, nulewicz-Ostrowska R, Paradowska K, Wawer I (2003) J Carbohydr Chem 22:593
294. Kopf J, Köll P (1984) Carbohydr Res 135:29
295. Ellervik U, Magnusson G (1994) J Am Chem Soc 116:2340
296. Grundberg H, Eriksson-Bajtner J, Bergquist KE, Sundin A, Ellervik U (2006) J Org Chem 71:5892

297. Thibaudeau C, Földesi A, Chattopadhyaya J (1998) Tetrahedron 54:1867
298. Luyten I, Thibaudeau C, Sandström A, Chattopadhyaya J (1997) Tetrahedron 53:6433
299. Oleary DJ, Kishi Y (1994) J Org Chem 59: 6629
300. Luyten I, Thibaudeau C, Chattopadhyaya J (1997) J Org Chem 62:8800
301. Castro S, Duff M, Snyder NL, Morton M, Kumar CV, Peczuh MW (2005) Org Biomol Chem 3:3869
302. Bozo E, Gati T, Demeter A, Kuszmann H (2002) Carbohydr Res 337:1351
303. Bozo E, Medgyes A, Boros S, Kuszmann J (2000) Carbohydr Res 329:25
304. Hendrickson JB (1967) J Am Chem Soc 89:7036
305. Hendrickson JB (1967) J Am Chem Soc 89:7043
306. DeMatteo MP, Snyder NL, Morton M, Baldisseri DM, Hadad CM, Peczuh MW (2005) J Org Chem 70:24
307. DeMatteo MP, Mei S, Fenton R, Morton M, Baldisseri DM, Hadad CM, Peczuh MW (2006) Carbohydr Res 341:2927
308. Entrena A, Campos JM, Gallo MA, Espinosa A (2005) Arkivoc 88
309. Entrena A, Campos J, Gómez JA, Gallo MA, Espinosa A (1997) J Org Chem 62:337
310. Anconi CPA, Nascimento CS, Dos Santos HF, De Almeida WB (2006) Chem Phys Lett 418:459
311. Wiberg KB (2003) J Org Chem 68:9322
312. Bocian DF, Strauss HL (1977) J Am Chem Soc 99:2866
313. Bocian DF, Strauss HL (1977) J Am Chem Soc 99:2876
314. Dillen J, Geise HJ (1979) J Chem Phys 70:425
315. Verdal N, Wilke JJ, Hudson BS (2006) J Phys Chem A 110:2639
316. Espinosa A, Gallo MA, Entrena A, Gómez JA (1994) J Mol Struct 323:247
317. Pakulski Z (1996) Pol J Chem 70:667
318. Pakulski Z (2006) Pol J Chem 80:1293
319. DeMatteo MP, Snyder NL, Morton M, Baldisseri DM, Hadad CM, Peczuh MW (2005) J Org Chem 70:24
320. Ng CJ, Craig DC, Stevens JD (1996) Carbohydr Res 284:249
321. Tran TQ, Stevens JD (2002) Austral J Chem 55:171
322. Driver GE, Stevens JD (2001) Carbohydr Res 334:81
323. Choong W, McConnell JF, Stephenson NC, Stevens JD (1980) Austral J Chem 33:979
324. James VJ, Stevens JD (1982) Cryst Struct Commun 11:79
325. Dabrowski J, Kozar T, Grosskurth H, Nifant'ev NE (1995) J Am Chem Soc 117:5534
326. Landersjö C, Stenutz R, Widmalm G (1997) J Am Chem Soc 119:8695
327. Eklund R, Lycknert K, Söderman P, Widmalm G (2005) J Phys Chem B 109:19936
328. Asensio JL, Hidalgo A, Cuesta I, Gonzalez C, Cañada J, Vicent C, Chiara JL, Cuevas G, Jiménez-Barbero J (2002) Chem Eur J 8:5228
329. French AD, Brady JW (1990) Computer Modeling of Carbohydrates, An Introduction. In: French AD, Brady JW (eds) Computer Modeling of Carbohydrate Molecules. ACS, Washington, DC, p 1
330. Babirad SA, Wang Y, Goekjian PG, Kishi Y (1987) J Org Chem 52:4825
331. Ravishankar R, Surolia A, Vijayan M, Lim S, Kishi Y (1998) J Am Chem Soc 120:11297
332. Lemieux RU, Koto S (1974) Tetrahedron 30:1933
333. Landersjö C, Stevensson B, Eklund R, Östervall J, Söderman P, Widmalm G, Maliniak A (2006) J Biomol NMR 35:89
334. Garcia-Aparicio V, Fernandez-Alonso MDC, Angulo J, Asensio JL, Cañada FJ, Jiménez-Barbero J, Mootoo DR, Cheng XH (2005) Tetrahedron: Asym 16:519
335. Cheetham NWH, Dasgupta P, Ball GE (2003) Carbohydr Res 338:955
336. Odelius M, Laaksonen A, Widmalm G (1995) J Phys Chem 99:12686
337. Asensio JL, Espinosa JF, Dietrich H, Cañada FJ, Schmidt RR, Martín-Lomas M, Andre S, Gabius HJ, Jiménez-Barbero J (1999) J Am Chem Soc 121:8995
338. Milton MJ, Bundle DR (1998) J Am Chem Soc 120:10547
339. Harris R, Kiddle GR, Field RA, Milton MJ, Ernst B, Magnani JL, Homans SW (1999) J Am Chem Soc 121:2546
340. Poppe L, Brown GS, Philo JS, Nikrad PV, Shah BH (1997) J Am Chem Soc 119:1727
341. Fernandez-Alonso MD, Cañada FJ, Solis D, Cheng XH, Kumaran G, Andre S, Siebert HC, Mootoo DR, Gabius HJ, Jiménez-Barbero J (2004) Eur J Org Chem1604
342. Haselhorst T, Espinosa JF, Jiménez-Barbero J, Sokolowski T, Kosma P, Brade H, Brade L, Peters T (1999) Biochemistry 38:6449
343. Haselhorst T, Weimar T, Peters T (2001) J Am Chem Soc 123:10705

344. Geyer A, Müller M, Schmidt RR (1999) J Am Chem Soc 121:6312
345. Siebert HC, Jiménez-Barbero J, Andre S, Kaltner H, Gabius HJ (2003) Describing Topology of Bound Ligand by Transferred Nuclear Overhauser Effect Spectroscopy and Molecular Modeling. In: Lee YC, Lee RT (eds) Recognition of Carbohydrates in Biological Systems Pt A: General Procedures. Academic Press, San Diego, p 417
346. Eklund R, Widmalm G (2003) Carbohydr Res 338:393
347. Pérez S, Imberty A, Engelsen SB, Gruza J, Mazeau K, Jiménez-Barbero J, Poveda A, Espinosa JF, van Eyck BP, Johnson G, French AD, Louise M, Kouwijzer CE, Grootenuis PDJ, Bernardi A, Raimondi L, Senderowitz H, Durier V, Vergoten G, Rasmussen K (1998) Carbohydr Res 314:141
348. Almond A (2005) Carbohydr Res 340:907
349. Woods RJ, Dwek RA, Edge CJ, Fraser-Reid B (1995) J Phys Chem 99:3832
350. Hounsell EF (1994) Adv Carbohydr Chem Biochem 50:311
351. Hounsell EF (1995) Prog Nuc Magn Reson Spectrosc 27:445
352. Bush CA, Martín-Pastor M, Imberty A (1999) Ann Rev Biophys Biomol Struct 28:269
353. Lipari G, Szabo A (1982) J Am Chem Soc 104:4559
354. Widmalm G (1998) Physical Methods in Carbohydrate Research. In: Boons G-J (ed) Carbohydrate Chemistry. Blackie, London, p 448
355. Lycknert K, Rundlof T, Widmalm G (2002) J Phys Chem B 106:5275
356. Lycknert K, Widmalm G (2004) Biomacromolecules 5:1015
357. Andersson A, Ahl A, Eklund R, Widmalm G, Maler L (2005) J Biomol NMR 31:311
358. Clore GM, Szabo A, Bax A, Kay LE, Driscoll PC, Gronenborn AM (1990) J Am Chem Soc 112:4989
359. Kjellberg A, Rundlöf T, Kowalewski J, Widmalm G (1998) J Phys Chem B 102: 1013
360. Kjellberg A, Widmalm G (1999) Biopolymers 50:391
361. Mäler L, Lang J, Widmalm G, Kowalewski J (1995) Magn Reson Chem 33:541
362. Glusker JP, Lewis M, Rossi M (1994) Crystal Structure Analysis for Chemists and Biologists. Wiley-VCH, New York
363. Hammond C (1997) The Basics of Crystallography and Diffraction. Oxford University Press, New York
364. Ladd MFC, Palmer RA (2003) Structure determination by X-ray crystallography. 4th edn. Kluwer Academic/Plenum Press, New York
365. Hammond C (2001) The Basics of Crystallography and Diffraction. 2nd edn. Oxford University Press, New York
366. Jeffrey GA, Sundaralingam M (1985) Adv Carbohydr Chem Biochem 43:203
367. Maluszynska H, Takagi S, Jeffrey GA (1977) Acta Crystallogr B33:1792
368. Ravishankar R, Suguna K, Surolia A, Vijayan M (1999) Acta Crystallogr D55:1375
369. Alibes R, Bundle DR (1998) J Org Chem 63:6288
370. Friebolin H (2005) Basic One- and Two-Dimensional NMR Spectroscopy. 4th edn. Wiley-VCH, New York
371. Becker ED (2000) High Resolution NMR : Theory and Chemical Applications. 3rd edn. Academic Press, San Diego, CA
372. Braun S, Berger S (2004) 200 and more Basic NMR Experiments: A Practical Course. 3rd edn. Wiley-VCH, New York
373. Homans SW (1992) A Dictionary of Concepts in NMR. Rev. edn. Clarendon, Oxford
374. Vliegenthart JFG, Dorland L, van Halbeek H (1983) Adv Carbohydr Chem Biochem 41:209
375. van Halbeek H (1996) Carbohydrates and Glycoconjugates. In: Grant DM, Harris RK (eds) Encyclopedia of Nuclear Magnetic Resonance. Wiley, New York, p 1107
376. Serianni AS (2000) Carbohydrate Structure, Conformation, and Reactivity: NMR Studies with Stable Isotopes. In: Hecht SM (ed) Biorganic Chemistry: Carbohydrates. Oxford University Press, New York, p 244
377. Jiménez-Barbero, J. and Peters, T.(2003) NMR Spectroscopy of Glycoconjugates. Wiley-VCH, Weinheim
378. Bock K, Pedersen C (1983) Adv Carbohydr Chem Biochem 41:27
379. Bock K, Pedersen C, Pedersen H (1984) Adv Carbohydr Chem Biochem 42:193
380. Shashkov AS, Nifant'ev ÉE, Amochaeva VY, Kochetkov NK (1993) Magn Reson Chem 31:599
381. Hobley P, Howarth O, Ibbett RN (1996) Magn Reson Chem 34:755
382. Lemieux RU, Stevens JD (1965) Can J Chem 43:2059

383. Jansson P-E, Kenne L, Widmalm G (1989) Carbohydr Res 188:169
384. Marshall JL (1983) Carbon-Carbon and Carbon-Proton NMR Couplings : Applications to Organic Stereochemistry and Conformational Analysis. Verlag Chemie International, Deerfield Beach, FL
385. Perlin AS, Casu B (1969) Tetrahedron Lett 2921
386. Bock K, Pedersen C (1974) J Chem Soc, Perkin Trans 2 293
387. Tvaroška I, Taravel FR (1995) Adv Carbohydr Chem Biochem 51:15
388. Cyr N, Perlin AS (1979) Can J Chem 57:2504
389. Mizutani K, Kasai R, Nakamura M, Tanaka O, Matsuura H (1989) Carbohydr Res 185:27
390. Callam CS, Gadikota RR, Lowary TL (2001) J Org Chem 66:4549
391. Hällgren C (1992) J Carbohydr Chem 11:527
392. Cahill R, Cookson RC, Crabb TA (1969) Tetrahedron 25:4681
393. Grindley TB, Szarek WA (1974) Can J Chem 52:4062
394. Hansen PE (1981) Prog Nuc Magn Reson Spectrosc 14:175
395. Cyr N, Hamer GK, Perlin AS (1978) Can J Chem 56:297
396. Schwarcz JA, Cyr N, Perlin AS (1975) Can J Chem 53:1872
397. Bock K, Pedersen C (1977) Acta Chem Scand B 31:354
398. Prytulla S, Lambert J, Lauterwein J, Klessinger M, Thiem J (1990) Magn Reson Chem 28:888
399. Staaf M, Weintraub A, Widmalm G (1999) Eur J Biochem 263:656
400. Pachler KGR (1996) Magn Reson Chem 34:711
401. Church T, Carmichael I, Serianni AS (1996) Carbohydr Res 280:177
402. Cloran F, Carmichael I, Serianni AS (2000) J Am Chem Soc 122:396
403. Lemieux RU, Kullnig RK, Bernstein HJ, Schneider WG (1958) J Am Chem Soc 80:6098
404. Karplus M (1963) J Am Chem Soc 85:2870
405. Haasnoot CAG, DeLeeuw FAAM, Altona C (1980) Tetrahedron 36:2783
406. Altona C, Francke R, de Haan R, Ippel JH, Daalmans GJ, Hoekzema AJAW, van Wijk J (1994) Magn Reson Chem 32:670
407. Altona C (1996) Vicinal Coupling Constants and Conformation of Biomolecules. In: Grant DM, Harris RK (eds) Encyclopedia of Nuclear Magnetic Resonance. Wiley, New York, p 4909
408. Barfield M, Smith WB (1992) J Am Chem Soc 114:1574
409. Barfield M, Smith WB (1993) Magn Reson Chem 31:696
410. Imai K, Osawa E (1990) Magn Reson Chem 28:668
411. Osawa E, Ouchi T, Saito N, Yamoto M, Lee OS, Seo MK (1992) Magn Reson Chem 30:1104
412. Marino JP, Schwalbe H, Griesinger C (1999) Acc Chem Res 32:614
413. Anet FAL, Freedberg DI (1993) Chem Phys Lett 208:187
414. Mulloy B, Frenkiel TA, Davies DB (1988) Carbohydr Res 184:39
415. Tvaroška I, Gajdos J (1995) Carbohydr Res 271:151
416. Milton MJ, Harris R, Probert MA, Field RA, Homans SW (1998) Glycobiology 8:147
417. Bose B, Zhao S, Stenutz R, Cloran F, Bondo PB, Bondo G, Hertz B, Carmichael I, Serianni AS (1998) J Am Chem Soc 120:11158
418. Zhu YP, Pan QF, Thibaudeau C, Zhao SK, Carmichael I, Serianni AS (2006) J Org Chem 71:466
419. Coxon B, Sari N, Batta G, Pozsgay V (2000) Carbohydr Res 324:53
420. Coxon B (1972) Conformational Analysis via Nuclear Magnetic resonance Spectroscopy. In: Whistler RL, BeMiller JN (eds) Methods in Carbohydrate Chemistry. Academic Press, New York, p 513
421. Kotowycz G, Lemieux RU (1973) Chem Rev 73:669
422. Otter A, Bundle DR (1995) J Magn Reson B 109:194
423. Emsley JW, Lindon JC (1975) NMR Spectroscopy using Liquid Crystal Solvents. Pergamon Press, Oxford
424. Diehl P (1996) Structure of Rigid Molecules Dissolved in Liquid Crystalline Solvents. In: Grant DM, Harris RK (eds) Encyclopedia of Nuclear Magnetic Resonance. Wiley, New York, p 4591
425. Tjandra N, Bax A (1997) Science 278:1697
426. Bolon PJ, Prestegard JH (1998) J Am Chem Soc 120:9366
427. Kiddle GR, Homans SW (1998) FEBS Lett 436:128
428. Rundlöf T, Landersjö C, Lycknert K, Maliniak A, Widmalm G (1998) Magn Reson Chem 36:773
429. Venable RM, Delaglio F, Norris SE, Freedberg DI (2005) Carbohydr Res 340:863
430. Zhuang TD, Leffler H, Prestegard JH (2006) Protein Science 15:1780

431. Bush CA (2003) Origins of Flexibility in Complex Polysaccharides. In: Cheng HN, English AD (eds) NMR Spectroscopy of Polymers in Solution and in the Solid State. ACS, Washington, DC, p 272
432. Berthault P, Jeannerat D, Camerel F, Salgado FA, Boulard Y, Gabriel JCP, Desvaux H (2003) Carbohydr Res 338:1771
433. Neuhaus D, Williamson MP (2000) The Nuclear Overhauser Effect in Structural and Conformational Analysis. 2nd edn. Wiley-VCH, New York
434. Neuhaus D, Williamson MP (1989) The Nuclear Overhauser Effect in Structural and Conformational Analysis. VCH Publishers, New York
435. Noggle JH, Schirmer RE (1971) The Nuclear Overhauser Effect. Academic Press, New York
436. Overhauser AW (1953) Phys Rev 92:411
437. Solomon I (1955) Phys Rev 99:959
438. Krishna NR, Agresti DG, Glickson JD, Walter R (1978) Biophys J 24:791
439. Malliavin TE, Desvaux H, Delsuc MA (1998) Magn Reson Chem 36:801
440. Keepers JW, James TL (1984) J Magn Reson 57:404
441. Donati A, Rossi C, Martini S, Ulyanov NB, James TL (1998) App Magn Reson 15:401
442. Nakanishi K (1994) Circular Dichroism : Principles and Applications. VCH, New York
443. Woody RW (1996) Theory of Circular Dichroism of Proteins. In: Fasman GD (ed) Circular Dichroism and the Conformational Analysis of Biomolecules. Plenum Press, New York, p 25
444. Johnson WCJr (1987) Adv Carbohydr Chem Biochem 45:73
445. Stevens ES (1996) Carbohydrates. In: Fasman GD (ed) Circular Dichroism and the Conformational Analysis of Biomolecules. Plenum Press, New York, p 501
446. Arndt ER, Stevens ES (1993) J Am Chem Soc 115:7849
447. Stroyan EP, Stevens ES (2000) Carbohydr Res 327:447
448. Arndt ER, Stevens ES (1997) Carbohydr Res 303:73
449. Schafer SE, Stevens ES (1996) Carbohydr Polym 31:19
450. Szabo L, Smith BL, McReynolds KD, Parrill AL, Morris ER, Gervay J (1998) J Org Chem 63:1074
451. Andersson M, Kenne L, Stenutz R, Widmalm G (1994) Carbohydr Res 254:35
452. Wiesler WT, Vázquez JT, Nakanishi K (1987) J Am Chem Soc 109:5586
453. Morales EQ, Padrón JI, Trujillo M, Vázquez JT (1995) J Org Chem 60:2537
454. Burkert U, Allinger NL (1982) Molecular Mechanics. ACS, Washington
455. Woods RJ (1996) The Application of Molecular Modeling Techniques to the Determination of Oligosaccharide Solution Conformations. In: Lipkowitz KB, Boyd DB (eds) Reviews in Computational Chemistry. VCH Publishers, New York, p 129
456. Rao VSR, Sundararajan PR, Ramakrishnan C, Ramachandran GN (1967) In: Ramachandran GN (ed) Conformations of Biopolymers. Academic Press, New York, p 721
457. Kitaigorodsky AI (1978) Chem Soc Rev 7: 133
458. Lemieux RU, Bock K, Delbaere LTJ, Koto S, Rao VS (1980) Can J Chem 44:631
459. Stuike-Prill R, Meyer B (1990) Eur J Biochem 194:903
460. Thøgersen H, Lemieux RU, Bock K, Meyer B (1982) Can J Chem 44:44
461. Kuttel M, Brady JW, Naidoo KJ (2002) J Comp Chem 23:1236
462. Case DA, Cheatham TE, III, Darden T, Gohlke H, Luo R, Merz KM Jr, Onufriev A, Simmerling C, Wang B, Woods RJ (2005) J Comp Chem 26:1668
463. Lii JH, Chen KH, Johnson GP, French AD, Allinger NL (2005) Carbohydr Res 340:853
464. Kony D, Damm W, Stoll S, Hunenberger PH (2004) J Phys Chem B 108:5815
465. Lins RD, Hunenberger PH (2005) J Comp Chem 26:1400
466. Ewig CS, Berry R, Dinur U, Hill JR, Hwang MJ, Li HY, Liang C, Maple J, Peng ZW, Stockfisch TP, Thacher TS, Yan L, Ni XS, Hagler AT (2001) J Comp Chem 22:1782
467. Stortz CA (2006) Carbohydr Res 341:663
468. Allinger NL, Yuh YH, Lii J-H (1989) J Am Chem Soc 111:8551
469. Engelsen SB, Rasmussen K (1997) J Carbohydr Chem 16:751
470. Hwang MJ, Stockfisch TP, Hagler AT (1994) J Am Chem Soc 116:2515
471. Stortz CA (1999) Carbohydr Res 322:77
472. Saunders M (1987) J Am Chem Soc 109:3150
473. Peters T, Meyer B, Stuike-Prill R, Somorjai R, Brisson JR (1993) Carbohydr Res 238:49
474. von der Lieth CW, Kozar T, Hull WE (1997) Theochem 395:225

475. Brocca P, Bernardi A, Raimondi L, Sonnino S (2000) Glycoconjugate J 17:283
476. Abraham RJ, Bretschneider E (1974) In: Orville-Thomas WF (ed) Internal Rotation in Molecules. Wiley-Interscience, New York, p 481
477. Tomasi J, Mennucci B, Cammi R (2005) Chem Rev 105:2999
478. Chamberlin AC, Cramer CJ, Truhlar DG (2006) J Phys Chem B 110:5665
479. Shen MY, Freed KF (2005) J Comp Chem 26:691
480. Dixon AM, Venable R, Widmalm G, Bull TE, Pastor RW (2003) Biopolymers 69:448
481. Kelly CP, Cramer CJ, Truhlar DG (2006) J Phys Chem A 110:2493
482. Curutchet C, Orozco M, Luque FJ, Mennucci B, Tomasi J (2006) J Comp Chem 27:1769
483. Chen JH, Im WP, Brooks CL (2006) J Am Chem Soc 128:3728
484. Lin ST, Hsieh CM (2006) J Chem Phys 125:
485. Wang ML, Wong CF (2006) J Phys Chem A 110:4873
486. Haille JM (1992) Molecular Dynamics Simulation. Elementary Methods. Wiley, New York
487. Balbuena PB, Seminario JM (1999) Molecular Dynamics: from Classical to Quantum Methods. Elsevier, Amsterdam
488. Allen MP, Tildesley DJ (1987) Computer Simulation of Liquids. Clarendon Press, Oxford
489. Spieser SAH, van Kuik JA, Kroon-Batenburg LMJ, Kroon J (1999) Carbohydr Res 322: 264
490. Levine IN (2000) Quantum Chemistry. 5th edn. Prentice Hall, Upper Saddle River, NJ
491. Lowe JP, Peterson KA (2006) Quantum Chemistry. 3rd edn. Elsevier Academic Press, Burlington, MA
492. Lewars E (2003) Computational Chemistry: Introduction to the Theory and Applications of Molecular and Quantum Mechanics. Kluwer Academic Publishers, Boston, MA
493. Cramer CJ (2004) Essentials of Computational Chemistry: Theories and Models. 2nd edn. Wiley, Chichester, UK
494. Schnupf U, Willett JL, Bosma WB, Momany FA (2007) Carbohydr Res 342:196
495. French AD, Johnson GP (2006) Can J Chem 84:603
496. Hricovíni M (2006) Carbohydr Res 341:2575
497. Miura N, Taniguchi T, Monde K, Nishimura SI (2006) Chem Phys Lett 419:326
498. Becke AD (1993) J Chem Phys 98:5648
499. Lii JH, Ma BY, Allinger NL (1999) J Comp Chem 20:1593

1.2 General Properties, Occurrence, and Preparation of Carbohydrates

John F. Robyt
Laboratory of Carbohydrate Chemistry and Enzymology, Department of Biochemistry, Biophysics, and Molecular Biology, 4252 Molecular Biology Building, Iowa State University, Ames, IA 50011, USA
jrobyt@iastate.edu

In: *Glycoscience*. Fraser-Reid B, Tatsuta K, Thiem J (eds)
Chapter-DOI 10-1007/978-3-540-30429-6_2:

Abstract

D-Glucose and its derivatives and analogues, *N*-acetyl-D-glucosamine, *N*-acetyl-D-muramic acid, D-glucopyranosyl uronic acid, and D-glucitol represent 99.9% of the carbohydrates on the earth. D-Glucose is found in the free state in human blood and in the combined state in disaccharides, sucrose, lactose, and α,α-trehalose, in cyclic dextrins, and in polysaccharides, starch, glycogen, cellulose, dextrans; *N*-acetyl-D-glucosamine and an analogue *N*-acetyl-D-muramic acid are found in bacterial cell wall polysaccharide, murein, along with teichoic acids made up of poly-glycerol or -ribitol phosphodiesters. Other carbohydrates, D-mannose, D-mannuronic acid, D-galactose, *N*-acetyl-D-galactosamine, D-galacturonic acid, L-iduronic acid, L-guluronic acid, L-rhamnose, L-fucose, D-xylose, and *N*-acetyl-D-neuraminic acid are found in glycoproteins, hemicelluloses, glycosaminoglycans, and polysaccharides of plant exudates, bacterial capsules, alginates, and heparin. D-Ribofuranose-5-phosphate is found in many coenzymes and is the backbone of RNAs (ribonucleic acid), and 2-deoxy-D-ribofuranose-5-phosphate is the backbone of DNA (deoxyribonucleic acid). D-Fructofuranose is found in sucrose, inulin, and levan. The general properties and occurrence of these carbohydrates and general methods of isolation and preparation of carbohydrates are presented.

Keywords

D-Glucose; D-Fructose; Sucrose; Lactose; α,α-Trehalose; Starch; Glycogen; Cyclodextrins; Dextrans; Alternan

Abbreviations

ADPGlc	adenosine-diphospho-glucose
ATP	adenosine triphosphate
CGTase	cyclomaltodextrin glucanyltransferase
d.s.	degree of substitution
DNA	deoxyribonucleic acid
FACE	fluorophore-assisted capillary electrophoresis
FAD	oxidized flavin adenine dinucleotide
HPLC	high pressure liquid chromatography
MALDI-TOF MS	matrix-assisted laser desorption ionization–time of flight mass spectrometry
NAD^+	oxidized nicotinamide adenine dinucleotide
NADH	reduced nicotinamide adenine dinucleotide
NADPH	reduced nicotinamide adenine dinucleotide phosphate
ORD	optical rotatory dispersion
RNA	ribonucleic acid
NADPH	nicotinamide adenine dinucleotide-phosphate
NAG	*N*-acetyl-D-glucosamine
NAM	*N*-acetyl-D-muramic acid
TLC	thin-layer chromatography
TFMS	trifluoromethane sulfonic acid
UDPGlc	uridine-diphospho-glucose

1 General Properties and Occurrence of Carbohydrates

Carbohydrates have the following major properties: (1) they are polyhydroxy aldehydes or ketones; (2) they have chiral or asymmetric carbons that are generally manifested by the rotation of plane polarized light; (3) they have the ability to form multiple hydrogen bonds, generally giving them the property of being water-soluble, but they also can be water-insoluble when they form intermolecular hydrogen bonds with each other to give crystals or large, high molecular weight, insoluble crystalline aggregates, granules, or fibers; (4) many have reactivities of aldehydes that can be oxidized to acids by reagents that are thereby reduced (e. g., reducing an oxidizing agent such as an alkaline solution of copper(II) or ferricyanide/cyanide), and they, hence, are considered to be reducing sugars, or they can themselves be reduced by reducing reagents, such as $NaBH_4$, to give sugar alcohols; (5) the aldehyde or ketone groups in carbohydrates with five or more carbons will react with intramolecular alcohol groups to form cyclic structures with hemiacetal and hemiketal hydroxyl groups; (6) the hemiacetal or hemiketal hydroxyls are more reactive than the alcohols and can react intermolecularly with alcohols and amines to give acetals or ketals (glycosidic bonds) that are fairly stable; (7) they have two kinds of alcohol groups, secondary and primary, that can undergo the usual reactions of alcohols to give esters and ethers and can be replaced, for example, by hydrogen, halogens (F, Cl, Br, and I), amino groups, *N*-acetyl amino groups, and sulfhydryl groups; (8) they are generally, although not all of them, sweet-tasting (for example, D-glucose, D-glucitol, D-fructose, D-xylose, D-xylitol and sucrose are sweet-tasting) by forming specific hydrogen and hydrophobic bonds with the sweet-taste receptors on the tongue; and (9) when attached to proteins or cell surfaces, the structural diversity of oligosaccharides mediate a large number of biochemical and biological processes.

In the 19th century, several naturally occurring carbohydrates were known, such as glucose (then called dextrose), fructose (then called levulose), mannose, galactose, sucrose, lactose, starch, and cellulose. Some of these had been known for thousands of years, for example, sucrose, starch, and cellulose. Also in the 19th century, the empirical formula for all of these materials was found to be $C_n(H_2O)_n$ and they were originally thought to be hydrates of carbon, hence the name carbohydrates.

Carbohydrates are now more completely defined as polyhydroxy aldehydes or ketones and compounds that can be derived from them by reduction to give sugar alcohols, oxidation to give sugar acids, substitution of hydroxyl group(s) by hydrogen to give deoxy sugars or by amino or *N*-acetyl amino groups to give deoxy-amino sugars, derivatization of a hydroxyl group by phosphate or sulfate to give sugar phosphates or sugar sulfates, and by condensation reactions of a hydroxyl group of one sugar with the hemiacetal group of another sugar to give disaccharides, trisaccharides, oligosaccharides, and polysaccharides.

2 Carbohydrate Property of Optical Rotation of Plane Polarized Light

An important property of carbohydrates that was recognized in the 19th century was that they generally, but not always, rotated plane polarized light and that this was specific for each carbohydrate. This property is due to the presence of asymmetric or chiral carbons that have four different groups attached to the carbons. Those carbohydrates that rotate plane polarized light

are said to be optically active. It was also recognized that the optical rotation was dependent on several factors: (1) the structure of the substance; (2) the length of the cell; (3) the concentration of the substance; (4) the wavelength of the plane polarized light; and (5) the temperature. The following relationship was derived to encompass these variables:

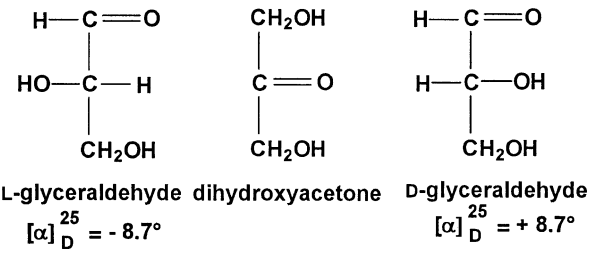
$$\alpha_{obs} = [\alpha]^t_\lambda lc$$

where α_{obs} = the observed optical rotation in degrees, l = the length of the cell holding the compound in dm (decimeter), c = the concentration of the sample in g mL^{-1}, usually in water, $[\alpha]_\lambda{}^t$ = the specific optical rotation constant of the substance at temperature, t, and wavelength, λ. Most polarimetric measurements are made with the D-line from a sodium lamp and each carbohydrate has a characteristic $[\alpha]^t_D$, although the optical rotation can also be measured continuously as a as a function of the wavelength (i. e., optical rotatory dispersion, ORD). Carbohydrate molecules with two-fold symmetry about a central point or plane do not rotate plane polarized light and are said to have a meso-structure.

3 The Structures of Carbohydrates

3.1 The Simplest Carbohydrates

There are three carbohydrates that are the simplest carbohydrates that fulfill the definition given above. They are the following:



Scheme 1
The simplest carbohydrates

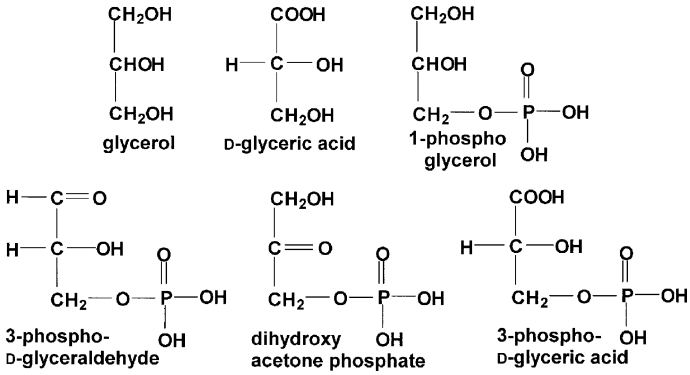
There are two forms for glyceraldehyde that are distinct and cannot be superimposed onto each other. Prof. Fischer defined the one with the chiral hydroxyl group to the right as D-glyceraldehyde, where the "D" indicates that the hydroxyl group is to the right or *dextro* and the one with the chiral hydroxyl group to the left as L-glyceraldehyde where the "L" indicates that the hydroxyl group is to the left or *levo*. It just so happened that for D-glyceraldehyde, plane polarized light was rotated to the right and L-glyceraldehyde rotated plane polarized light to the left. This is not always the case. Some carbohydrates with the D-configuration rotate plane polarized light to the left and some carbohydrates with the L-configuration rotate plane polarized light to the right.

A large majority of the carbohydrates found on the earth belong to the D-family of structural isomers. In the course of evolution, the reason that the D-family of structural isomers was

selected over those of the L-family of structural isomers is not clear. It, however, is not likely that it was a matter of chance. It has been known for many years that irradiation of a racemic mixture of D- and L- isomers with circularly polarized light will selectively destroy one of the two isomers, leaving the other more or less intact [1]. Circularly polarized light has been observed when there is high sunspot activity. High levels of circularly polarized light have also been observed coming from the *Orion nebula* [2]. The selection of D-carbohydrates could have occurred by this type of irradiation when carbohydrates were first being formed on the earth.

3.2 Analogues of D-Glyceraldehyde

D-Glyceraldehyde has some derived analogues, such as the reduced sugar alcohol, glycerol, its oxidized product, D-glyceric acid, and their phosphorylated analogues, 1-phospho-D-glycerol, 3-phospho-D-glyceraldehyde, and 3-phospho-D-glyceric acid, whose structures are shown as:



Scheme 2
Analogues and derivatives of the naturally occurring three-carbon carbohydrates

Glycerol is found as the backbone compound that is esterified by fatty acids to give a class of lipids known as triacyl glycerols (glycerides), and glycerol-1-phosphate is the backbone of a major class of phospholipids. 3-Phospho-D-glyceraldehyde, dihydroxy acetone phosphate, and 3-phospho-D-glyceric acid are all found in both the reactions of photosynthesis and in the degradative reactions of glycolysis.

3.3 The Formation of Carbohydrates Containing More than Three Carbons

From a theoretical stand point, Professor Emil Fischer showed that by adding a new chiral carbon between the aldehyde group and the asymmetric carbon of D-glyceraldehyde, a chiral pair of 4-carbon D-tetraoses would be obtained, namely, D-erythrose and D-threose. Adding another set of similar chiral carbons to each of the two D-tetraoses, gives four 5-carbon D-pentoses, and likewise adding a similar set of chiral carbons to each of the four D-pentoses, gives eight 6-carbon D-hexoses. It should be noted that it is only the configuration of the last asymmetric

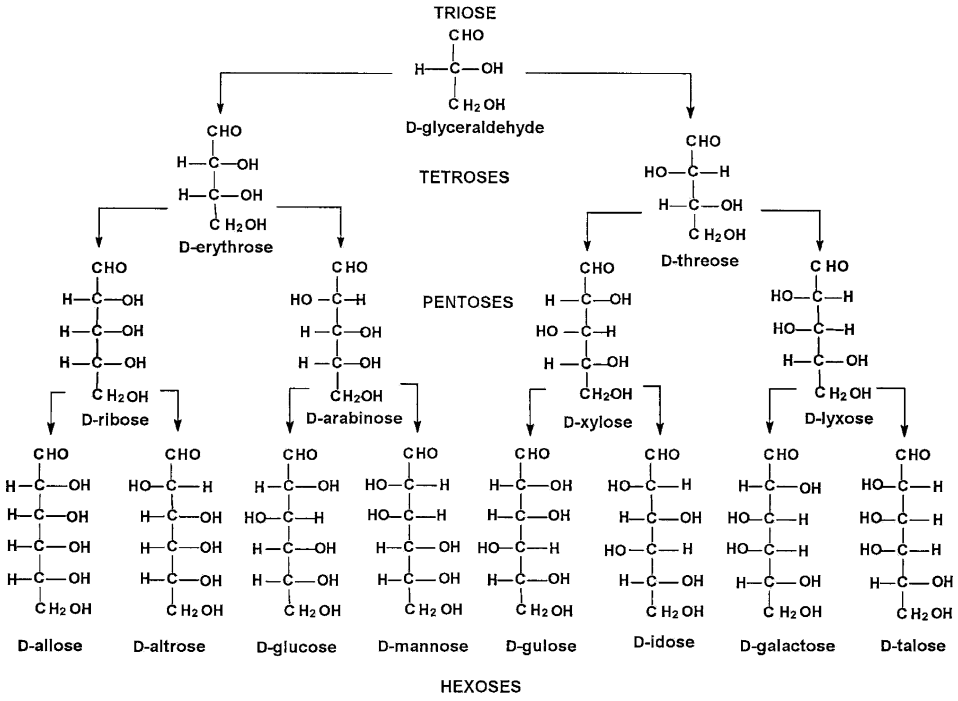


Figure 1
Structural family of D-carbohydrates, from triose to hexoses, with their names

carbon with the hydroxyl group to the right that makes it a "D" carbohydrate and the entire mirror image of the D-carbohydrates gives the L-carbohydrate. See *Fig. 1* for the names and structures of the family of D-carbohydrates.

3.4 Special Properties of Pentoses and Hexoses

The pentoses and hexoses have a propensity for forming six-membered rings in which one of their hydroxyl groups reacts intramolecularly with the aldehyde group to form a cyclic hemiacetal. This reaction creates a new asymmetric center on the aldehyde carbon to give two isomers, called alpha (α) and beta (β). Two five-membered rings are also formed, but in much smaller amounts, as the six-membered cyclic structures are much more thermodynamically stable (i. e., less strained) than are the five-membered rings. The six-membered rings are called pyranoses and the five-membered rings are called furanoses. For D-glucose at 20 °C and equilibrium, there are five compounds: 0.004% the open aldehyde chain, 66% β-D-glucopyranose, 33% α-D-glucopyranose, and 0.498% each of β-D-glucofuranose and α-D-glucofuranose. See *Fig. 2* for the structures of the five forms of D-glucose in equilibrium.
If one starts with α-D-glucopyranose, $[\alpha\text{-}]_D^{25} = +112°$, the optical rotation drops to +52° and if one starts with β-D-glucopyranose, $[\alpha\text{-}]_D^{25} = +19°$, the optical rotation increases and becomes constant at +52°, which is the optical rotation for an equilibrium mixture of the five structural

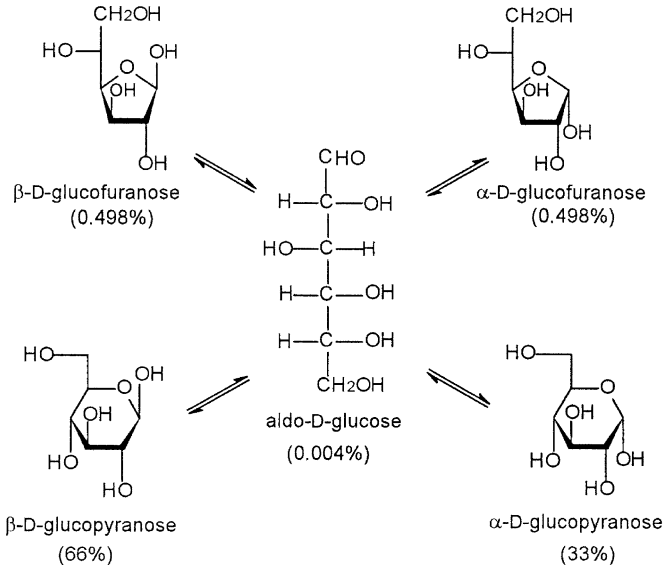


Figure 2
The five structural forms of D-glucose at equilibrium in aqueous solution at 20 °C

forms of D-glucose. The process is known as mutarotation and is relatively slow at pH 7 and 20 °C. It can be accelerated by catalysis with either acid or base or by adding an enzyme, known as mutarotase. Dilute base (pH 10) is a better catalyst by a factor of 5,000 than dilute acid (pH 4). Mutarotase acts as an acid–base catalyst and catalyzes the reactions 4–5 orders of magnitude faster than base.

3.5 D-Glucose: the Most Prominent Carbohydrate on the Earth

Of the 15 possible D-carbohydrates in *Fig. 1*, only a handful occurs in nature to any extent. By far, D-glucose and its analogues are the most prominent and represent 99.9% of the carbohydrates on the earth. Why is this? While D-glucose forms the six-membered cyclic structure and has the Haworth structure as shown in *Fig. 3*, the ring actually has a three-dimensional chair conformation, with two kinds of geometric bonds around the carbons, those that are within the plane of the ring (called equatorial bonds) and those that are perpendicular to the ring (called axial bonds). D-Glucose can exist in two chair conformations, the C1 or $^{4}C_{1}$ chair and the 1C or $_{4}C^{1}$ chair (see *Fig. 3* for the structures). In the C1 or $^{4}C_{1}$ conformation, all of the hydroxyl or bulkiest groups for β-D-glucopyranose are attached to the ring by equatorial bonds that put the hydroxyl or bulkiest groups as far apart as possible from each other, giving the most thermodynamically stable structure possible. If β-D-glucopyranose is in the other chair conformation, 1C or $_{4}C^{1}$, all of the hydroxyl groups are axial and are placed as close together as possible, giving the most thermodynamically unstable structure possible. Thus, β-D-glucopyranose exists primarily in the C1-conformation. α-D-Glucopyranose also exists

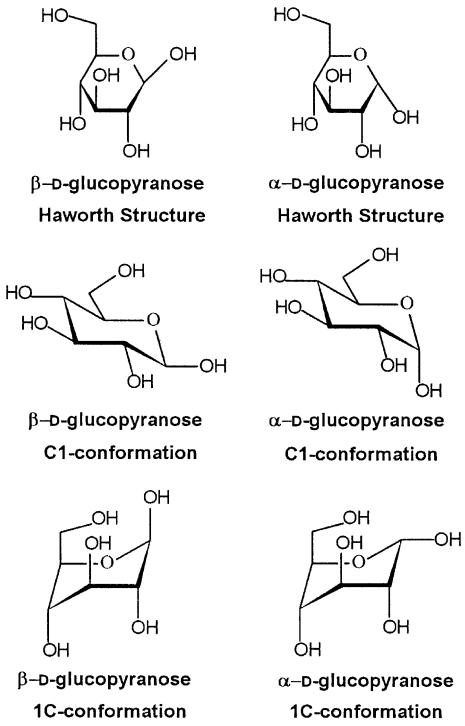


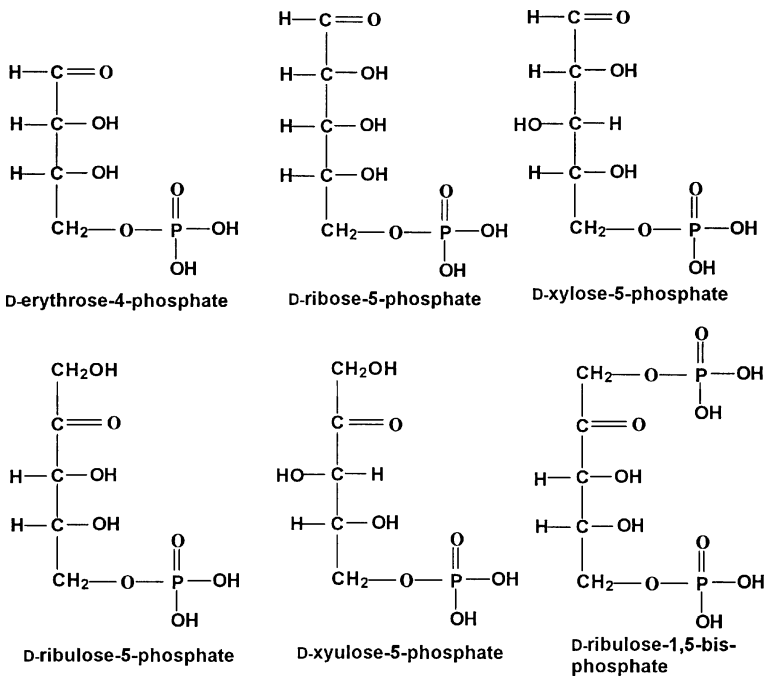
Figure 3
Haworth structures for α- and β-D-glucopyranose and their C1 and 1C conformations

in the C1-conformation, with only the hemiacetal hydroxyl group in the axial position. This is the most likely reason that D-glucose is the predominant carbohydrate on the earth, as it is the only D-hexose that can have all of its hydroxyl groups (exclusive of the hemiacetal hydroxyl group) equatorial, and at equilibrium in solution β-D-glucopyranose has a ratio of $\sim$2:1 to α-D-glucopyranose, which has its hemiacetal hydroxyl group axial. D-Xylose, a pentose, can also form a six-membered ring by its terminal hydroxyl group reacting with the aldehyde group and form a C1 conformation and place all of its bulky hydroxyl groups equatorial, but because it has five carbons, it would have to be split into a 2-carbon fragment and a 3-carbon fragment for metabolism and would require two separate pathways for further metabolism. D-Glucose has 6-carbons and is split into two 3-carbon fragments, D-glyceraldehyde-3-phosphate and dihydroxy acetone-phosphate that are interconvertible and requires only a single metabolic pathway for further metabolism. This is an additional plausible reason that D-glucopyranose is the predominant carbohydrate on the earth.

The D-pentoses will also exist in solution as the six-membered ring structure, but both D-xylose and D-ribose often have their C-5 hydroxyl groups phosphorylated and the C-5 hydroxyl group cannot react with the aldehyde group to form the six-membered ring, and therefore they do the next best thing, with the C-4 hydroxyl group reacting with the aldehyde group to form the five-membered, furanose ring, hemiacetal structure.

3.6 Occurrence of D-Erythrose, D-Ribose, and D-Xylose

Of the two D-tetroses, only 4-phospho-D-erythrose is found in any quantity, as an intermediate in the photosynthetic reactions. Of the four pentoses, only D-ribose and D-xylose occur to any extent. The phospho-*aldo*-D-pentoses, D-ribose-5-phosphate and D-xylose-5-phosphate, are found in the photosynthetic reactions. Three phospho-*keto*-D-pentoses are also found as intermediates in the photosynthetic reactions: D-ribulose-5-phosphate, D-xylulose-5-phosphate, and D-ribulose-1,5-bis-phosphate, the latter carbohydrate being directly involved in the fixing of CO_2in photosynthesis.



Scheme 3
Important naturally occurring four- and five-carbon sugar phosphates

Nicotinamide adenine dinucleotide (NAD^+ and NADH), coenzymes containing D-ribofuranose-diphosphate, are involved in many oxidation and reduction reactions, respectively, of carbohydrate metabolism. D-Ribofuranose is also the main component of the universal energy donor and energy carrier, adenosine triphosphate (ATP), which is one of the primary products of the light reactions of photosynthesis and is responsible for providing the energy for the formation of the carbon–carbon bond in the fixation of CO_2. It is also important in the transfer and utilization of energy in the metabolism of nonphotosynthesizing organisms. Another important D-ribofuranose coenzyme that is formed as a primary product of the light reactions of photosynthesis is the reducing coenzyme, nicotinamide adenine dinucleotide-phosphate,

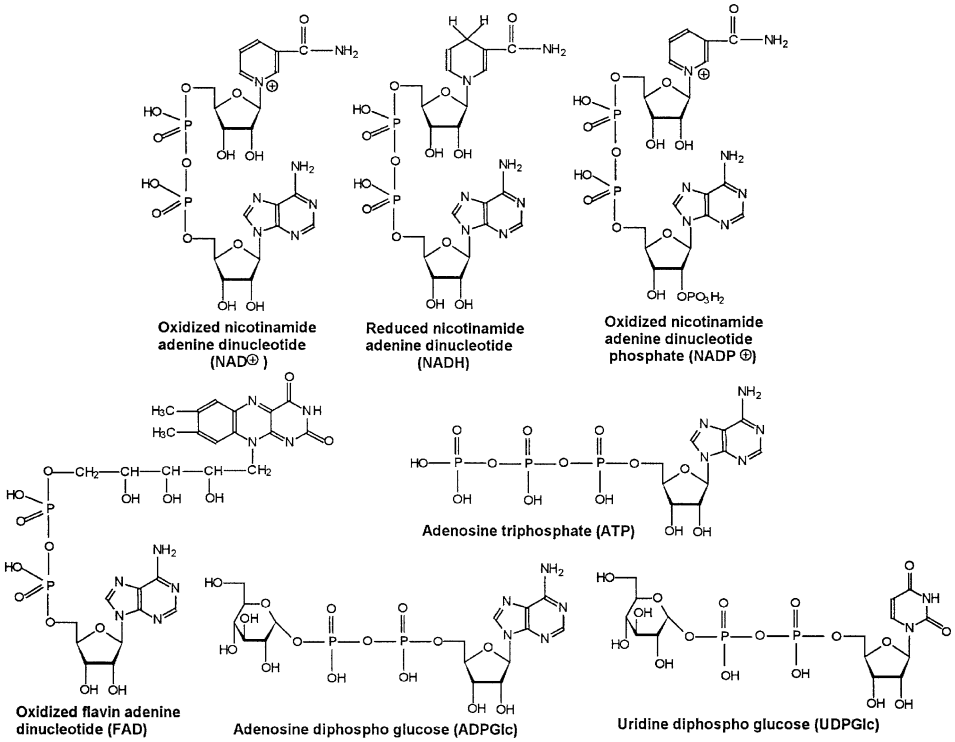


Figure 4
Structures of the D-ribofuranose-5-phosphate nucleotide coenzymes that are important in biochemical metabolism

(NADPH), which is similar to the coenzymes, NAD^+ and NADH, mentioned above, but with an additional phosphate attached to the 2-position of the ribose unit. See *Fig. 4* for the structures of these coenzymes. NADPH is responsible for reducing the carbon–carbon bond that is formed in fixing CO_2in photosynthesis.

ATP is the universal energy carrier and source of energy in biochemical systems; NAD^+ and NADH, oxidation and reduction coenzymes; FAD, oxidative coenzyme, containing ribitol-5-phosphate; ADPGlc, high-energy glucose donor, involved in starch biosynthesis; and UDPGlc, another high-energy glucose donor, involved in cellulose, glycogen, and sucrose biosyntheses, and in the enzymatic conversions of D-glucopyranose to many other sugars, such as, D-galactopyranose, D-glucopyranouronic acid, D-xylopyranose, and L-arabinopyranose.

D-Ribofuranose-5-phosphate occurs as the backbone component of the ribonucleic acids, RNA, that are involved in the biosynthesis of proteins. There are three kinds of RNA's: a small RNA, transfer-RNA that forms a high-energy, amino-acid covalent compound that transfers individual amino acids to the ribosome to be incorporated into proteins; an intermediate sized RNA, messenger-RNA that carries the codon or genetic information of a protein to the ribosome where the code is read and the peptide bonds of the protein are synthesized; and the largest sized RNA, ribosomal-RNA that composes the ribosome, the organelle where pro-

teins are synthesized and assembled. 2-Deoxy-D-ribofuranose-5-phosphates are the backbone components of deoxyribonucleotides (DNA), which primarily act as the carrier of the genetic information, necessary for the formation of proteins involved in life processes. Like RNA, these molecules have two purines (adenine and guanine) and two pyrimidines (uracil and cytosine for RNA and thymidine and cytosine for DNA) that are linked β to carbon-1 of the D-ribofuranose-phosphate units to give *N*-glycosides.

D-Xylopyranose occurs as one of the major components in the hemicelluloses (see ❯ *Sect. 6* on cellulose and ❯ *Sect. 7* on hemicelluloses and ❯ Chap. 6.3 on cellulose).

3.7 Occurrence of Hexoses

Of the eight D-hexoses in ❯ *Fig. 1*, only three occur to any extent: D-glucose, its 2-isomer, D-mannose, and its 4-isomer, D-galactose. Another hexose that is found is D-fructose, which is a keto-sugar that can be derived from dihydroxy acetone, as the aldehyde carbohydrates were derived from D-glyceraldehyde. D-Fructose is formed when D-glucose or D-mannose are treated with alkali, which isomerizes carbons 1 and 2 of the two D-hexoses [3]. The diphosphate of D-fructose, D-fructose-1,6-bis-phosphate is the first hexose that is formed in the photosynthetic process and it is rapidly converted into D-fructose-6-phosphate, and then into D-glucose-6-phosphate.

4 Properties and Occurrence of D-Glucose

4.1 D-Glucose in the Free State

Free D-glucose occurs primarily in the blood of many higher animals, where it serves as an immediate source of energy and as a stabilizer of the osmotic pressure, and a precursor for the formation of glycogen and fat in muscle tissue. In normal humans, the concentration of blood glucose is 80–120 mg/100 mL^{-1} or 5–7 mM. This can increase to 200–300 mg/100 mL^{-1} or 11–20 mM after a high-carbohydrate meal, but then is relatively rapidly decreased by the action of insulin and often goes below 80 mg/mL^{-1} and then goes slowly up to normal levels. In uncontrolled diabetics the glucose often goes much higher, to 140–1100 mg/100 mL^{-1} or 8–60 mM. In controlled diabetics, the glucose will often be slightly higher than normal, for example, 120–140 mg/100 mL^{-1}.

D-Glucose is also found in the free state in honey, grapes, and raisins. It is produced by the action of glucoamylase (amyloglucosidase) on starch, where glucoamylase hydrolyzes the α-(1→4) linkage of the glucose units at the nonreducing-ends of the starch chains, giving inversion of the configuration, forming β-D-glucose. Glucoamylase will also hydrolyze α-(1→6) branch linkages, although at a rate about 0.1 that of the α-(1→4) linkage. Eventually glucoamylase will completely convert all of the starch into D-glucose.

4.2 D-Glucose in the Combined State

D-Glucose is found in a combined form in which its hemiacetal group has reacted with a hydroxyl on another glucose or another carbohydrate to form an acetal or glucosidic link-

age and form such carbohydrate substances as, starch, dextran, cellulose, sucrose, lactose, α,α-trehalose.

4.2.1 Occurrence of D-Glucose Combined with D-Fructose, D-Galactose, and D-Glucose, and High-Energy D-Glucose Donors

There are three major, naturally occurring disaccharides: sucrose, lactose, and α,α-trehalose and D-glucopyranose is found in all three. Sucrose is a nonreducing disaccharide composed of D-glucopyranose joined to D-fructofuranose (1→2) to give α-D-glucopyranosyl-(1→2)-β-D-fructofuranoside. Sucrose is widely distributed in plants, primarily as 6^{Fru}-phosphosucrose, which is the transport form of carbohydrates in plants. When it reaches certain parts of the plant, the glucose moiety is converted into starch by a series of reactions that give adenosine-diphospho-glucose (ADPGlc) or uridine-diphospho-glucose (UDPGlc), high-energy donors of glucose for the biosynthesis of starch and cellulose, respectively. Relatively large amounts (15–20% by weight) of free sucrose are found in the stems and tubers of sugar cane and the tubers of sugar beets. Free sucrose is also found in many other plants but in lower amounts, such as their fruits. Sucrose is the predominant sugar found in honey, produced by bees, in the sap of maple trees, giving maple syrup, and in sorghum, and in dates. See *Fig. 5* for the structures of sucrose, lactose, and α,α-trehalose and *Fig. 4* for the structures of ADPGlc and UDPGlc.

4.2.2 Properties and Occurrence of Sucrose and Sucrose Oligosaccharides Containing D-Galactose

Sucrose has five important properties: (a) the linkage between D-glucopyranose and D-fructofuranose is of high energy being an acetal–ketal linkage of a six-membered ring attached to a five-membered ring, making sucrose a nonreducing sugar. (b) The acetal–ketal linkage is

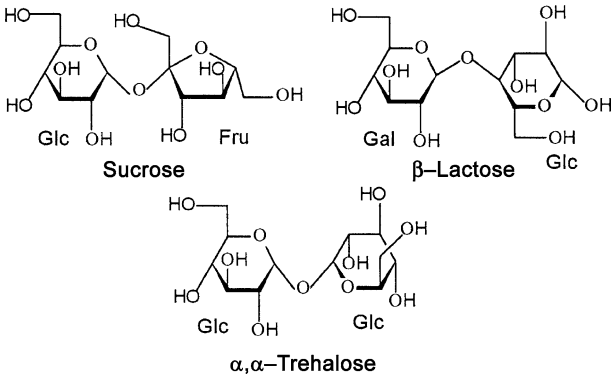


Figure 5
Structures of the three naturally occurring disaccharides: sucrose, lactose, and α-α-trehalose, containing α-D-glucopyranose combined with β-D-fructofuranose, β-D-galactopyranose, and α-D-glucopyranose, respectively

relatively labile and is hydrolyzed by mild acid (pH 4) and is the donor of D-glucopyranose for the biosynthesis of dextrans and related polysaccharides by glucansucrases and the donor of D-fructofruanose for the biosynthesis of levan and inulin (see *Sect. 11*), fructofuranose polysaccharides by levansucrase and inulinsucrase; (c) it is the sugar of commerce because of the ease of obtaining it in large quantities in a pure state from sugar cane and sugar beets; (d) it crystallizes relatively easily; and (e) it has a pleasant sweet taste and has been recognized by humans for over 10,000 years as a sweet food and a natural sweetening agent. See *Fig. 5* for the structure of sucrose.

The origin of sucrose is thought to have been in the Indus Valley, where many woody, wild sugar cane plants that have the fundamental characteristics of the modern cultivated strains can still be found growing today. Sugar cane grows well in a warm, humid, tropical or semi-tropical climate. In the late 18th century on the European continent, the sugar beet was found to be an alternative source of sucrose that did not require a tropical or semi-tropical climate for growth.

Sucrose is hydrolyzed into its component sugars (D-glucose and D-fructose) by the action of the enzyme, invertase, a β-fructofuranosidase, and by mild acid. In this form it is known as invert sugar, due to the fact that the direction of rotation of polarized light is inverted from dextrorotatory to levorotatory on hydrolysis. Honey is usually a mixture of sucrose and invert sugar. Yeasts also have invertase and can hydrolyze sucrose and then ferment the component sugars into ethyl alcohol.

In addition to sucrose, several plants also form a series of sucrose-based oligosaccharides with chains of α-1→6 D-galactopyranose units linked to the D-glucose moiety of sucrose [4]. The first in the series is the trisaccharide, raffinose, in which D-galactopyranose is linked α-(1→6) to sucrose; the second is a tetrasaccharide, stachyose, in which D-galactopyranose is linked α-(1→6) to the D-galactopyranose unit of raffinose. The next is a pentasaccharide, verbascose, with D-galactopyranose linked α-(1→6) to the terminal D-galactose unit of stachyose, and the next is a hexasaccharides, ajugose, with D-galactopyranose linked α-(1→6) to the terminal D-galactose unit of verbascose.

These D-galactopyranosyl sucrose oligosaccharides are particularly found in the tubers and seeds of legumes. Raffinose is found in cottonseeds and in sugar beets. Although sugar beets only contain about 0.05% by weight raffinose as compared with 16–18% sucrose, it has been isolated and crystallized with a purity of better than 99% from sugar beet syrup, where it accumulates during the processing of sucrose.

Soybeans are a good source of stachyose, where it is found to the extent of 2–3% by weight. In general, legume seeds and the mullein root are sources of verbascose. The enzyme invertase and mild acid specifically hydrolyze the oligosaccharides to give D-fructose and the corresponding reducing oligosaccharides that are terminated at the reducing-end with D-glucose. For example, raffinose is hydrolyzed to give D-fructose and the reducing disaccharide, melibiose [α-D-galactopyranosyl-(1→6)-D-glucose].

Another series of galacto-sucrose oligosaccharides involves the attachment to the D-fructofuranose moiety of sucrose [4]. The attachment of α-D-galactopyranose 1→6 to the fructose moiety gives the nonreducing trisaccharide, planteose. It is found primarily in the seeds of the *Plantago* family of plants, for example, the common weed and herb, plantain. Mild acid hydrolysis gives D-glucose and the reducing keto-disaccharide, planteobiose [α-D-galactopyranosyl-(1→6)-D-fructose]. Another nonreducing trisaccharide, melezitose, has α-D-glucopyranosyl linked (1→3) to the D-fructofuranose moiety of sucrose. It is found

in the sweet exudates of many trees, such as larch, Douglas fir, Virginia pine, and poplars. Mild acid hydrolysis of melezitose gives D-glucose and the reducing disaccharide, turanose, [α-D-glucopyranosyl-(1→3)-D-fructose].

Several sucrose analogues have been enzymatically synthesized in the laboratory. Levansucrase can transfer a D-fructofuransoyl unit from raffinose to D-xylose, giving a nonreducing sucrose disaccharide analogue, xylsucrose [α-D-xylopyranosyl-(1→2)-D-fructofuranoside] and the reducing disaccharide, melibiose [α-D-galactopyranosyl-(1→6)-D-glucose [5]. A similar reaction of levansucrase with raffinose and D-galactose gives galactosucrose [α-D-galactopyranosyl-(1→2)-D-fructofuranoside] (also referred to as galsucrose) and melibiose [6]. Reaction of sucrose and lactose with levansucrase gives D-glucose and the nonreducing trisaccharide, lactosucrose [4^{Glc}-β-D-galactopyranosyl sucrose] [7]. Reaction of dextransucrase with sucrose and D-fructose gives an unusual reducing disaccharide, leucrose, containing an α-(1→5) linkage of D-glucopyranosyl linked (1→5) to D-fructopyranose [α-D-glucopyranosyl-(1→5)-D-fructopyranose] [8,9].

A relatively large number of sucrose derivatives have been chemically synthesized [10]. Some notable chloro derivatives have been obtained by the reaction of sucrose with sulfuryl chloride in pyridine/chloroform at low temperatures, for example 4,6,1′,6′-tetrachloro-4,6,1′,6′-tetradeoxy-glactosucrose, 4,6,6′-trichloro-4,6,6′-trideoxy-sucrose, and many others were formed [11,12,13,14,15]. These chloro-compounds of sucrose were 10–100 times sweeter than sucrose. One of them, Sucralose (4,1′,6′-trichloro-4,1′,6′-trideoxy-galactosucrose), was 650-times sweeter than sucrose, with no after-taste, and a sweet-taste identical to sucrose. It is used commercially as a noncariogenic and noncaloric sweetener in soft drinks, candies, cookies, jellies, and many other prepared foods, as well as a general substitute for table sugar. Sucralose is enzymatically inert and passes through the human body without being metabolized or absorbed.

4.2.3 Properties and Occurrence of D-Glucose Combined with D-Galactose to Give Lactose and Higher Oligosaccharides

Lactose is a disaccharide composed of β-D-galactopyranose linked (1→4) to D-glucose and is found in the milk of mammals, where it serves as a source of energy and nourishment for the newborn. Lactose is a reducing disaccharide because the D-glucopyranose residue has a free hemiacetal group at C1. Human milk contains 85 g L^{-1} lactose and cow's milk contains about 50 g L^{-1}.

Human milk also contains lactose oligosaccharides in which various different monosaccharide residues are attached to the D-galactopyranosyl residue. α-L-Fucose (6-deoxy-L-galactose) is attached (1→2) to the galactose moiety [16], α-*N*-acetyl-D-neuraminic acid is attached (2→3) to the galactose moiety [16,17], and the β-*N*-acetyl-D-glucosamine residue is attached either (1→3) or (1→6) to the galactose moiety [18]. The latter serves to produce a core structure that can be further extended by the addition of β-D-galactopyranose residues linked either (1→2) or (1→4) [19]. The β-D-galactopyranosyl-β-*N*-acetyl-D-glucosamine disaccharide is often added in multiples to give a repeated core structure to which α-*N*-acetyl-D-neuraminic acid and α-L-fucose residues are added to the ends of the oligosaccharides. The so called human blood group determinants (see ❯ *Sect. 14*) have structural similarities to the human milk oligosaccharides [19] and it is thought that through this relationship the milk oligosac-

charides impart some form of early immunological protection to the newborn. See *Fig. 5* for the structure of lactose.

4.2.4 Properties and Occurrence of α-D-Glucose Combined with α-D-Glucose to Give α,α-Trehalose

α,α-Trehalose is also one of the three naturally occurring disaccharides. It is a nonreducing disaccharide with two D-glucopyranose residues joined together in an α-1↔α-1 acetal–acetal linkage and is of relatively high energy, like sucrose. Unlike sucrose, however, the acetal–acetal linkage is quite stable and is one of the most difficult linkages to be hydrolyzed by acid. The reason for this is not absolutely clear, but it has been hypothesized to be due to the stabilization of the molecule by intramolecular hydrogen bonds between the D-glucose residues. α,α-Trehalose is found in insect lymph fluid ("insect blood") where it acts as a source of chemical energy [19] and it is also found in mushrooms, honey, yeast, fungi, lobster, and shrimp as a source of energy [20,21] all of which have the enzyme, trehalase, that hydrolyzes α,α-trehalose to give two molecules of D-glucose that can be used for energy.
Besides being used as a source of energy, some plants and animals also use α,α-trehalose as a stabilizing agent during extreme conditions [22]. High concentrations of α,α-trehalose in the tissues of certain insects and in desert plants allows them to survive in a state of suspended animation under conditions of water deficiency. α,α-Trehalose helps frogs to survive in a frozen state and it helps to protect the DNA of salmon sperm from dehydration. α,α-Trehalose has also found applications in the preservation of organs taken for use in organ transplants [23,24]. See *Fig. 5* for the structure of α,α-trehalose.

5 Properties and Occurrence of D-Glucose in Polysaccharides and Cyclodextrins

5.1 Properties and Occurrence of Starch

Starch is an abundant polysaccharide composed of D-glucose residues. It is found in the green leaves, stems, roots, seeds, fruits, tubers, and bulbs of most plants, where it serves as the storage of chemical energy obtained from the energy of the sun light in the process of photosynthesis. Starch also serves as the major source of chemical energy for most nonphotosynthesizing organisms such as bacteria, fungi, insects, and animals. It is found in relatively large amounts in the major food crops of the world. Starch is present 80% by weight in the rice kernel, 78% in the potato tuber, 75% in green bananas, 73% in the maize kernel, 68% in wheat flour, and 60% in rye and lentils. Starch provides about 65% of the dietary calories in the human diet.
Starch occurs in plants as water-insoluble granules produced in plant organelles, plastids (chloroplasts and amyloplasts). The granules have specific shapes and sizes that are characteristic of their botanical source [25]. Most starches are composed of a mixture of two types of polysaccharides, a linear polysaccharide, consisting of α-(1→4) linked D-glucopyranose residues, called amylose, and a branched polysaccharide of α-(1→4) linked D-glucopyranose

residues with 5–6% α-(1→6) branch linkages, called amylopectin. Amylose has an average of 500 to 5,000 D-glucopyranose residues per molecule, depending on the source; amylopectin is much larger and has an average of 100,000 to 1,000,000 D-glucopyranose residues per molecule [26,27]. When at equilibrium with its surroundings, starch granules will contain 10–15% w/w water. The amylose and amylopectin molecules in the granules can be solubilized by heating the granules in water, where they swell and eventually burst, releasing the individual molecules. Starch granules can also be dissolved in 9:1 dimethyl sulfoxide/water solutions [28]. See ❯ Fig. 1 in ❯ Chap. 6.2 for the structures of segments of amylose and amylopectin.

The amounts of amylose and amylopectin differ for starches from different botanical sources. Most so-called normal starches have 20–30% amylose and 80–70% amylopectins, respectively [29,30]. There are mutant varieties, such as waxy maize, waxy rice, and waxy potato, that are composed of 100% amylopectin. There also are the high amylose varieties, such as amylomaize-V that consists of 53% amylose and 47% amylopectin and amylomaize-VII that is 70% amylose and 30% amylopectin, just the reverse of the "normal" starches. Many of the "normal" starches have been found to have an intermediate component that is slightly branched amylose with 0.5–3% α-(1→6) branch linkages [26,27,29,30].

All starches can be completely converted into D-glucose by acid hydrolysis at high temperatures (100 °C) and by the action of the enzyme, glucoamylase, at lower temperatures (20–40 °C), when the granules are solubilized. Humans and other organisms can completely convert solubilized starches into D-glucose by the combined action of several enzymes, such as α-amylases found in saliva and in the small intestine, and α-(1→6)-glucosidase and α-(1→4)-glucosidase that are secreted by special cells in the lining of the small intestines.

Starches have been chemically modified to improve their solution and gelling characteristics for food applications. Common modifications involve the cross linking of the starch chains, formation of esters and ethers, and partial depolymerization. Chemical modifications that have been approved in the United States for food use, involve esterification with acetic anhydride, succinic anhydride, mixed acid anhydrides of acetic and adipic acids, and 1-octenylsuccinic anhydride to give low degrees of substitution (d.s.), such as 0.09 [31]. Phosphate starch esters have been prepared by reaction with phosphorus oxychloride, sodium trimetaphosphate, and sodium tripolyphosphate; the maximum phosphate d.s. permitted in the US is 0.002. Starch ethers, approved for food use, have been prepared by reaction with propylene oxide to give hydroxypropyl derivatives [31].

The solubility of the starch granules has been increased by reaction of starch granules in water with 7% hydrochloric acid for one week at 20 °C to give "Lintner soluble starch". Recent modifications to increase the solubility of starch granules have involved the reaction of the starch granules with hydrochloric acid in anhydrous alcohols, such as methanol, ethanol, 2-propanol, and 1-butanol to give a new class of limit dextrins whose average degree of polymerization can be controlled between 1800 and 30 [32,33,34]. Enzymatic conversions of starches into mixtures of maltodextrins are used in food preparations. Starch is the major source for the commercial preparation of D-glucose and D-fructose. Starches have been modified to give tertiary amino alkyl ethers, quaternary ammonium ethers, amino ethylated ethers, cyanamide ethers, starch anthranilates, cationic dialdehyde starch, carboxymethyl ethers, and carboxy starch for various applications in the sizing of paper, formation of coatings, sizing of textiles, flocculation, and emulsification technologies [35].

5.2 Properties and Occurrence of Glycogen

Glycogen is an α-glucan that is widely distributed in mammals in the liver, muscle, and brain and in fish, insects, and some species of bacteria, fungi, protozoa, and yeasts, as a reserve form of chemical energy. It is a high molecular weight polysaccharide (1×10^6 to 2×10^9 Da) composed of D-glucopyranose residues linked together by α-(1→4) glycosidic linkages with 10–12% α-(1→6) branch linkages [36]. It has been compared with amylopectin and called "animal starch." But, it is quite different from amylopectin in that it has over twice as many α-(1→6) branch linkages per molecule, giving the many chains an average chain length of 8 to 10 D-glucopyranose residues compared to 20 for amylopectin. Further, the branch linkages do not occur in clusters, as they do in amylopectin, and are randomly distributed, giving glycogen different chemical and physical properties from amylopectin. Glycogen does occur in particles or granules of about 25 nm, called β-particles [37]. The β-particles are further combined into a larger mass, called α-particles, which consists of approximately 100 β-particles. Nevertheless, in contrast to starch granules, the glycogen particles are quite water-soluble, because of the relatively high percent of branch linkages and the absence of intermolecular bonding, giving the absence of crystallinity. Glycogen reacts poorly with triiodide, giving a light brown color or no color, but never the blue color given by starch granules and amylose nor the maroon color given by amylopectin.

Glycogen has a specific function in the liver of mammals, where its primary role is to maintain the normal concentration of D-glucose in the blood. In humans, it can provide 100–150 mg of glucose per minute over a sustained period of 12 h, if necessary [38]. In skeletal muscle, its primary function is to provide immediate energy for muscle movement by being converted into α-glucopyranose-1-phosphate and in the human brain, where glycogen normally provides about 100 g of α-glucose-1-phosphate per day for energy used by the brain [38].

Blue-green algae, which are photosynthetic bacteria (cyanobacteria) and not eukaryotic algae, synthesize glycogen as a reserve energy storage polysaccharide instead of synthesizing starch [39]. Glycogen is also synthesized by nonphotosynthetic bacteria, such as *Escherichia coli* that synthesizes it intracellularly from UDPGlc [39] and *Neisseria perflava* that synthesizes it extracellularly from sucrose by the enzyme, amylosucrase [40]. The function of glycogen for these bacteria has been postulated to provide reserve energy in times of the absence of nutrients and as a source of energy for the formation of spores [39].

5.3 Properties and Occurrence of Dextrans, Alternan, Mutan, and Pullulan

Dextrans are a large family of bacterial polysaccharides that have a contiguous series of D-glucopyranose residues linked α-(1→6) to each other [41]. Over 100 strains of *Leuconostoc mesenteroides* [42], *Streptococcus mutans*, *S. sobrinus*, and *S. salivarius* produce specific enzymes, dextransucrases that synthesizes dextrans from sucrose [41]. All of the dextrans are branched, primarily by α-(1→3) glycosidic linkages, but also by α-(1→2) and α-(1→4) linkages in specific *L. mesenteroides* strains. Differences in the number and arrangements of these branches, such as having single glucose branches or long α-(1→6) linked branch chains, and the order and frequencies of the branches, impart differences in the structures and properties [41,42]. The classic prototypical dextran is the commercial product synthesized by dex-

transucrase from *L. mesenteroides* NRRL B-512F. It has 95% α-(1→6) linkages with 5% α-(1→3) branch linkages that have branches that are both single glucose residues and long α-(1→6) linked chains and contains ~10^6 to 10^8 D-glucopyranose residues. Other strains make a wide variety of dextrans with various degrees of branching, not only through α-(1→3) linkages, but through α-(1→2) and/or α-(1→4) linkages as well. In some cases, the degree of branching is as low as 3%, and in other instances, virtually every glucose residue in the backbone may be substituted with a branch linkage [41].

Some of these bacteria also elaborate glucansucrases that synthesize polysaccharides that are not considered dextrans because they do not have contiguous α-(1→6) linked main chains. *L. mesenteroides* NRRL B-1355 secretes a dextransucrase that synthesizes a B-512F-type dextran and another enzyme, alternansucrase, that synthesizes an α-glucan from sucrose that has alternating α-(1→6) and α-(1→3) linked glucose residues in the main chains with 7–11% α-(1→3) branch chains of alternating α-(1→6) and α-(1→3) linked glucose residues [41]. This α-glucan is called alternan. Another glucansucrase, mutansucrase, secreted by *Streptococcus mutans*, synthesizes a linear glucan in which the D-glucopyranose residues are linked α-(1→3). It is particularly characterized by being extremely water-insoluble in contrast to the dextrans and alternan that are highly water-soluble [41].

Pullulan is a polysaccharide that is elaborated by several species of the fungus, *Aureobasidium*, particularly *A. pullulans*. This fungus is typified by the presence of black pigments and is sometimes called "black yeast" [43]. Pullulan is a water-soluble, linear polysaccharide of D-glucopyranose residues joined together by a repeating sequence of two α-(1→4) and one α-(1→6) linkages. The structure is that of a polymer of maltotriose units joined together end to end by α-(1→6) linkages [44,45,46]. In addition to maltotriose units, it also has ~5–7% maltotetraose units located in the interior of the polysaccharide chain [47].

These bacterial polysaccharides have been considered to be "slimes"; they are often in reality loose capsules that are produced extracellularly by the bacteria. It was found that low molecular weight *L. mesenteroides* NRRL B-512F dextran could be used as a blood plasma extender and was produced on a relatively large scale during the "cold war", but also found uses as a gel-filtration material when cross-linked by epichlorohydrin to give a family of cross-linked dextrans [41].

5.4 Properties and Occurrence of D-Glucose in Cyclic Dextrins

A number of different kinds of nonreducing cyclic dextrins containing D-glucopyranose residues occur. The first to be observed were the cyclomaltodextrins (sometimes referred to in the older literature as cyclodextrins or Schardinger dextrins), which have been known for over 100 years. They were first found in rotting vegetables and then in the fermentation of starch by a heat-resistant microorganism called *Bacillus macerans*. The compounds were crystallized from alcohol solutions and shown to be α-(1→4) linked, nonreducing, cyclic dextrins composed of six, seven, and eight D-glucopyranose residues, named cyclomaltohexaose, cyclomaltoheptaose, and cyclomaltooctaose or α-CD, β-CD, and γ-CD [48]. These cyclomaltodextrins are formed from starch by the enzyme, cyclomaltodextrin glucanyltransferase (CGTase). *Bac. macerans* CGTase primarily forms α-CD; other bacteria, for example *Bac. circulans* elaborates a CGTase that primarily forms β-CD and *Brevibacterium sp.* elaborates a CGTase that primarily forms γ-CD [49]. Larger cyclomaltodextrins, having 9, 10, 11, and

12 D-glucopyranose residues were later obtained in relatively small quantities, and even later, cyclomaltodextrins having as many as 25 glucose residues were obtained [50]. The internal cavity of the cyclomaltodextrins is relatively hydrophobic, giving them the property of forming complexes with a wide variety of organic molecules [51].

Cycloisomaltodextrins, linked α-(1→6) containing seven, eight, and nine D-glucopyranose residues have been found to be formed by a bacterial cycloisomaltodextrin dextran-glucanyltransferase, acting on B-512F dextran [52]. A cyclic tetrasaccharide, containing four glucose residues with alternating α-(1→6) and α-(1→3) linkages has been obtained from the reaction of a bacterial enzyme on alternan [53]. This enzyme, 3-α-isomaltosyltransferase, is part of a 2-enzyme system that converts starch to the cyclic tetrasaccharide [54,55]. Also, *Bacillus stearothermophilus* starch branching enzyme catalyzed a reaction with amylose to give macrocyclic dextrins with one α-(1→6) linkage at the site of transglycosylation coupling [56].

Another group of cyclic dextrins is the cyclosophorans, which consist of 17–40 D-glucopyranose residues linked β-(1→2). They are produced by *Rhizobium* species involved in nitrogen-fixing nodules on the roots of legumes [59] and are also found in plant crown galls, produced by *Agrobacterium tumefaciens* [58]. The sizes of the cyclosophorans vary depending on the particular species of *Rhizobium*, which are also specific for the particular type of legume that they associate with to form the nodules. There is some evidence that the cyclosophorans play a role in the formation of the nodules and the crown galls [59].

Bradyrhizobium species synthesize a related cyclic dextrin that contains 12 D-glucopyranose residues linked by a repeating sequence of three contiguous β-(1→6) linkages followed by three contiguous β-(1→3) linkages. One of the β-(1→3) sequences has a single branched D-glucopyranose residue substituted β-(1→6) onto the center D-glucopyranose residue [60]. A cyclodextrin containing only β-(1→3) linkages (cyclolaminarinose) has been found to be elaborated by a recombinant strain of *Rhizobium meliloti* TY7 mutant that is deficient in forming cyclosophoran, but carrying the genetic locus of *Bradyrhizobium japonicum* USDA 110. The cyclolaminarinose dextrin has 10 D-glucopyranose residues, with a single laminaribiose disaccharide substituted β-(1→6) onto the ring [61].

5.5 Properties and Occurrence of Cellulose

Cellulose is usually considered the most abundant carbohydrate on the earth, occurring in all plant cell walls to the extent of approximately 50% by weight; 20–40% of the cell wall is made up of hemicelluloses, and the remaining 10–30% is the noncarbohydrate, lignin, which acts as a cross-linking and cementing agent in the plant cell wall, covalently attached to the hemicelluloses [62]. Hemicelluloses are a family of polysaccharides, with a structure similar to cellulose, but besides D-glucopyranose residues, they contain several other monosaccharide residues, such as D-xylopyranose, D-mannopyranose, D-galactopyranose, D-glucopyranose uronic acid, and L-arabinofuranose residues (see *Sect. 6*).

Cellulose is a very large, linear polysaccharide of $\sim 10^6$ to 10^8 D-glucopyranose residues, linked β-(1→4) to each other. Because of its high water-insolubility, its actual size has never been accurately determined. It is a β-glucan with a very tight helical structure in which the individual glucose residues are oriented 180°to each other [63]. Because of this conformation and the β-linkages, cellulose chains readily form intermolecular hydrogen bonds, giving

multiple chains associated together in 3-dimensional bundles that further associate with other bundles to form micelles or fibers [63,64]. These micelles make a very tough and resistant material that gives shape, strength, and water and substance impermeability to the plant cell

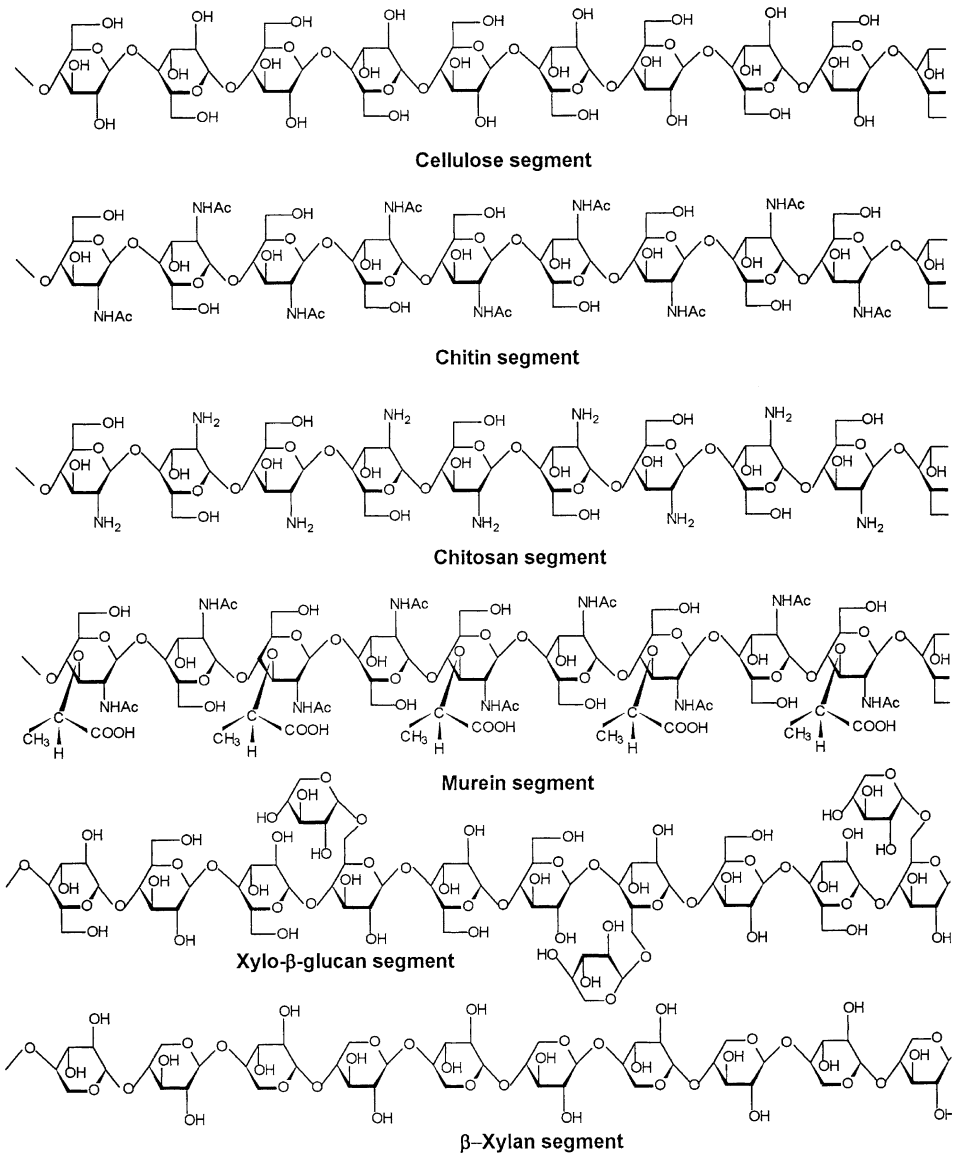


Figure 6
Haworth structures of β-(1→4) linked segments of cellulose, chitin, chitosan, murein, xylo-β-glucan, and xylan

wall. The fibers are also quite resistant to chemical and enzymatic attack. Very pure cellulose is found in the cotton boll and is also synthesized in a relatively pure form by some species of bacteria, such as *Acetobacter xylinum*, *Agrobacterium tumefaciens*, and related bacteria [65]. When one thinks of cellulose, it is usually in connection with trees and wood that come to mind. Cellulose is a major component of flax, comprising 80% (w/w), and jute, comprising 60–70% (w/w). Grasses, such a papyrus and bamboo have been important sources of cellulose going back into ancient times. Papyrus was used as an early form of paper, made from the pith of the papyrus plant, a wetlands sedge that grows to 5 meters (~16 ft) in height and was once abundant in the Nile Delta of Egypt. It was first known to have been used as a writing material in ancient Egypt (at least as far back as the First Dynasty, 3,000 BC) but it was also widely used throughout the Mediterranean, as well as in Europe and Southwest Asia, until about the 11th Century AD. Papyrus was prepared as a thin film from the outer bark that was glued together with starch paste to give it body and the ability to hold ink. Bamboo also served man from very early times, and continues to do so, as a building material to form houses, roofs, furniture, and so forth. Paper today is manufactured from several cellulose sources, such as wood chips and sawdust, the fibers of the sugar cane plant (called bagasse), maize (corn) stalks, and the straws of rye, oats, and rice. The β-(1→4) glycosidic linkage of cellulose is more resistant toward acid hydrolysis than the α-(1→6) linkage of amylopectin. Cellulose is slowly hydrolyzed by 1-M HCl at 100 °C. See ❯ *Fig. 6* for the structure of a segment of the cellulose molecule.

6 Properties and Occurrence of Hemicelluloses

Hemicelluloses are a family of four basic types of polysaccharides, composed of two or more monosaccharide residues. All have structural features similar to cellulose in that they have their main chains that are β-(1→4) linked, with the exception of the arabinoglactans that are β-(1→3) linked. The main chains are homopolysaccharides composed of a single monosaccharide residue, but they are highly branched by one or two different kinds of monosaccharides that are linked for the most part to give single monosaccharide branches.

As previously mentioned, the cell walls of most plants contain 40–60% cellulose. The remaining carbohydrate, representing 40–50% (w/w) of the cell wall is composed of hemicelluloses. The composition of the hemicelluloses varies from one plant type to another [66]. The four basic types are

1. Xyloglucans composed of α-D-xylopyranose linked 1→6 to approximately every third D-glucose residue of cellulose [66,67,68];
2. Xylan composed of D-xylopyranose linked β-(1→4) and glucurono-arabino-xylan, which is composed of β-(1→4) D-xylopyranose chain with 4-*O*-methyl-α-D-glucopyranosyluronic acid linked 1→2 and α-L-arabinofuranosyl linked 1→3 to the xylan chain [66,67,68,69];
3. Mannan is composed of β-(1→4)-D-mannopyranose chains [70]; another type of D-mannan is galactomannan that has D-galactopyranose linked α-(1→6) to the D-mannan chain [71].
4. Arabinogalactan is composed of β-(1→3) linked D-galactopyranose chain with β-(1→6) linked D-galactopyranose branches and to a lesser degree a L-arabinofuranose disaccharide,

β-L-arabinofuranosyl-(1→3)-α-L-arabinopyranosyl linked 1→6 to the D-galactopyranose chain [72,73].

Xyloglucans and xylans are widely distributed, found in most plant cell walls, see *Fig. 6* for the structures of these hemicelluloses. Glucurono-arabino-xylan is also widely distributed, especially in soft-wood trees [74]. Glucuronoxylan is prevalent in hard-wood trees [74]. Hemicellulose composed exclusively of D-mannopyranose is found in palm seed endosperm [70], where it is known as vegetable ivory or ivory nut mannan. Galactomannan are particularly found in soft-wood trees, but also are found in the seedpods of the locust bean. In particular, a galactomannan known as guar gum has been obtained from the seeds of the legume, guar, grown in the semiarid regions of India. It has found widespread use as a thickening agent in food products. Arabinogalactans are prevalent in soft woods such as, larch, black spruce, Douglas fir, cedar, and juniper [74]. Segments of the structures of D-xyloglucan and D-xylan are shown in *Fig. 6*.

6.1 Properties and Occurrence of Pectin

Pectins are related to hemicelluloses and occur in the plant cell wall in low amounts of 1–5% (w/w). They are more prevalent in fruits, for example, apple pulp (10–15%) and in orange and lemon rinds (20–30%). Pectins do have a number of chemical and physical properties that differ from the hemicelluloses. They are composed of D-galactopyranosyl uronic acids linked α-(1→4) and have a relatively wide percentage of the carboxyl groups esterified with methyl groups. There also is a small amount (1 in ~25 uronic acid residues) with L-rhamnopyranosyl (6-deoxy-L-mannopyranosyl) residues linked α-(1→2) to the D-galactopyranosyl uronic acid residues. In addition, pectins also have 2-*O*-acetyl or 3-*O*-acetyl ester groups attached to the uronic acids [75]. The average molecular weight can vary from 20,000 to 400,000 Da, with a typical average molecular weight of 100,000 Da. Pectins act in plants as an intercellular cementing agent that provides body to fruits, and of course, it is used in foods as a gelling agent, especially in the preparation of jellies and confections.

7 Cellulose-like Polysaccharides Containing *N*-Acetyl-D-Glucosamine and D-Glucosamine

There are several polysaccharides containing the β-(1→4) structure, but with monomer residues other than D-glucopyranose, such as *N*-acetyl-D-glucosamine, D-glucosamine, *N*-acetyl-D-muramic acid.

7.1 Properties and Occurrence of Chitin

Chitin is a polysaccharides with the exact same structure as cellulose but containing *N*-acetyl-D-glucosamine or D-glucosamine and are fairly widely distributed [76] (see *Fig. 6* for the structure of a segment of chitin). It is a structural polysaccharide that forms fibers, is water impermeable, and replaces cellulose in the cell walls of many species of lower organisms,

such as fungi, yeasts, green algae, and brown and red seaweeds. It also comprises the major component of the exoskeleton of insects, where it makes a hard shell-like material that is quite strong. Chitin is also found in the cuticles of worms and in the shells of crustaceans, such as mollusks, shrimps, crabs, and lobsters [77].

Chitin, like cellulose, has a highly ordered, crystalline structure in which the chains are intermolecularly hydrogen bonded in an antiparallel arrangement, a parallel arrangement, and a mixed arrangement of two parallel and one antiparallel repeating arrangement [78]. Also like cellulose, it is very insoluble in water and most other solvents. In arthropods, the chitinous shell, or exoskeleton, does not grow, and is periodically cast off or molted. After the old shell is shed, a new, larger shell is produced, providing room for further growth. Chitin is very rigid, except between some body segments and joints, where it is much thinner and allows movement of the various parts.

7.2 Properties and Occurrence of Chitosan

Chitosan is a polysaccharide very similar to chitin, except that the *N*-acetyl-D-glucosamine is replaced by D-glucosamine in which the *N*-acetyl group is removed (see ❯ *Fig. 6* for the structure of a segment of chitosan). Chitosan is found occurring naturally mixed with chitin in the cell walls of some fungi and seaweeds. It is, however, primarily produced chemically by treating chitin with strong alkali to deacetylate the *N*-acetyl-amino group [79]. The degree of deacetylation can range from 60 to 100%, giving a family of chitosans. The free amino group of chitosan has a pK_a value of ~6.5 and it can be protonated in mildly acidic solutions, giving a positive charge to the glucosamine residues. The positive charges on chitosan produce very different physical and chemical properties from chitin. Because of the repulsion of the positive charges, chitosan chains do not line up and associate to form micelles and fibers, as does cellulose and chitin. Chitosan, thus, is water-soluble at acidic pH values.

Because of the positive charges on chitosan, it has found a number of applications. It binds to negatively charged surfaces, such as mucosal membranes, and has been used as a bandage material for wounds that is biocompatible and biodegradable [80,81]. Positively charged chitosan enhances the transport of polar drugs across epithelial tissues and is used to transport drugs in humans [82]. It has been used as an enhancer for plant growth, and as an aid in the defense of plants against fungal infections. Chitosan is used in water purification, as a material in a sand filtration system where it binds fine sediment particles during filtration, greatly aiding the removal of turbidity; it also removes phosphates by ion exchange, heavy metals by chelation, and oils by hydrophobic adsorption from water [81]. Chitosan has also been found useful for the immobilization of enzymes and cells [81,82,83,84].

7.3 Properties and Occurrence of *N*- Acetyl-D-Glucosamine and *N*-Acetyl-D-Muramic Acid in Murein – The Bacterial Cell Wall

The major component of all known bacterial cell walls is a polysaccharide composed of *N*-acetyl-D-glucosamine (NAG) linked together by β-(1→4) glycosidic bonds, as in chitin, but with every other NAG residue substituted at C-3 by an ether linkage to the hydroxyl group of L-lactic acid to give *N*-acetyl-D-muramic acid (NAM) [85,86,87,88]. This results in a nine-

carbon *N*-acetyl-amino-sugar acid, with a repeating β-(1→4)-NAG-NAM sequence of 40–150 residues, giving a polysaccharide, called murein [89].

A pentapeptide is attached to the carboxyl group of the L-lactic acid by an amino group that forms a peptide (amide) bond [90]. Attached to this pentapeptide is a pentaglycine linked to the ε-amino group of an L-lysine by a carboxyl group. The glycine end forms a cross-link to another decapeptide [90]. Using slightly different amino acids in both the pentapeptide and the cross-linking peptide gives different peptides that are genus-dependent.

The murein-peptidoglycan gives rigidity and different specific shapes, such as rods, spheres, or spirals to bacterial cells. Because of the cross-linking of the murein chains, the peptidoglycan is considered one giant, bag-shaped macromolecule [91]. The structures of segments of chitin, chitosan, and murein are shown in *Fig. 6*.

7.4 Properties and Occurrence of Glycosaminoglycans Composed of Amino Sugars and Uronic Acids

Glycosaminoglycans make up a group of polysaccharides that are found in animal tissues. They are composed of repeating disaccharides units of *N*-acetyl-D-glucosamine or *N*-acetyl-D-galactosamine and D-glucopyranosyluronic acid residues. The linkages are primarily β at positions 3 and 4. They are most often attached to protein backbones, forming what is called a proteoglycan [92].

7.4.1 Hyaluronic Acid

Hyaluronic acid consists of repeating disaccharides of β-D-glucopyranosyluronic acid-(1→3)-*N*-acetyl-D-galactosamine linked β-(1→4) to the next disaccharide. This proteoglycan can have between 500 and 50,000 residues per chain [92]. Hyuronic acid is found widely distributed in mammalian cells and tissues, where it is found in synovial fluid that lubricates the joints, in the vitreous humor of the eye, and in connective tissue, such as the umbilical cord, the dermis, and the arterial wall. It also occurs as a capsular polysaccharide around certain bacteria, such as pathogenic streptococci [92].

7.4.2 Chondroitin Sulfate

Chondroitin sulfate also consists of a repeating disaccharide of β-D-glucopyranosyluronic acid-(1→3)-*N*-acetyl-D-galactosamine linked β-(1→4) to the next disaccharide unit, with sulfate groups attached to C-4 or C-6 of the *N*-acetyl-D-galactosamine residue. It occurs as a major component of cartilage found in the cornea of the eye, the aorta, the skin, and lung tissues, where it is located between fibrous protein molecules. Chondroitin sulfate provides a soft and pliable texture to these tissues [92].

7.4.3 Dermatan Sulfate

Dermatan sulfate is derived from chondroitin 4-sulfate by the action of a C-5 epimerase that inverts the carboxyl group of the β-D-glucuronic acid, giving the very rare sugar, α-L-idopy-

ranosyl uronic acid (α-L-iduronic acid). Some of the L-iduronic acids are sulfated at C-2. Dermatan sulfate is found primarily in the skin [93].

7.4.4 Keratan Sulfate

Keratan sulfate consists of the disaccharide, *N*-acetyl-lactosamine, linked β-(1→3) to the D-galactopyranose residue of the next *N*-acetyl-lactosamine unit. Keratan sulfate is the most heterogeneous of the glycosaminoglycans, with variable sulfate content linked to C-4 or C-6 of the D-galactopyranose residue in lactosamine, and small amounts of L-fucose (6-deoxy-L-galactose), D-mannose, and *N*-acetyl-neuraminic acid residues [93] (also see ❯ *Sects. 10.2* and ❯ *14* on the occurrence of *N*-acetyl-D-neuraminic acid and L-fucose in other systems). Keratan sulfate is found in the cornea, on the surfaces of erythrocytes, in cartilage, and in bone.

7.4.5 Heparan Sulfate

Heparan sulfate consists of the repeating disaccharide β-D-glucopyranosyluronic acid-(1→4)-*N*-sulfato-2-amino-2-deoxy-α-D-glucopyranosyl linked (1→4). This polysaccharide is linked to a core protein to give a proteoglycan that is found as a matrix component of arterial wall, lung, heart, liver, and skin [93].

8 Polysaccharides Containing Uronic Acids That Have Some of Their Carboxyl Groups Inverted by a C-5 Epimerase to Give New Polysaccharides with New Properties

8.1 Heparin Sulfate

Heparin sulfate is formed from heparan sulfate by the action of an enzyme, C-5 epimerase that inverts the carboxylate group attached to C-5 of the D-glucopyranosyl uronic acid residues to give the rare and unusual sugar, α-L-idopyranosyl uronic acid. Heparin sulfate is released from the heparin of proteoglycans of mast cells into the blood stream when there is an injury to blood vessels, in the heart, liver, lungs, and skin. The release of heparin near the site of the injury acts as an anti-coagulating agent, preventing massive clotting of the blood and, hence, preventing run-away clot formation [93].

8.2 Alginates

Alginates are found primarily in brown seaweeds in amounts of 18–40% by weight of the plant. The majority is extracellular, being located between the cells [94]. One of the major species of seaweeds that contains alginates is the giant kelp, *Macrocyctis pyrifera*. It grows along the California coast of the US, the northwestern and southwestern coasts of South America, and the southeastern coasts of Australia and New Zealand [95].

Alginates are formed from poly-D-mannopyranosyluronic acid by the action of a C-5 epimerase that inverts the C-5 carboxyl group to give approximately 33% (w/w) of the

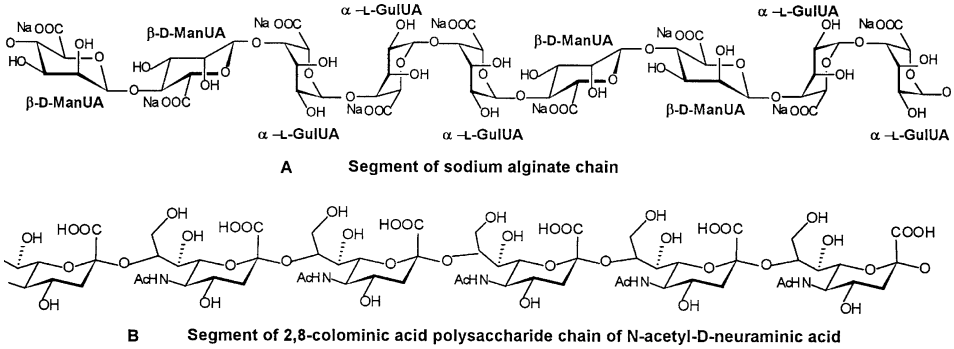


Figure 7
Conformational structures of segments of sodium alginate, containing β-D-mannopyranosyl uronic acid, and α-L-gulopyranosyl uronic acid; and 2,8-α-colominic acid, containing *N*-acetyl-α-D-neuraminic acid, linked (2→8)

rare and unusual sugar α-L-gulopyranosyl uronic acid. The ratio of the two uronic acids varies with different species of seaweed, type of tissue, and age of the plant [96]. The two kinds of uronic acids are combined together in blocks of variable numbers and also as alternating residues. In the brown seaweed, alginates most frequently exists as the calcium salt and are converted to the sodium salt when isolated.

The biological role that alginates play in seaweeds is that of a protective agent against desiccation during low tide. An unusual and useful property of sodium alginate is the ability to instantly form gels when in contact with divalent metal ions, such as calcium, barium, strontium, copper, cobalt, nickel, and so forth [97]. The strength and firmness of the gels are proportional to the amount of α-L-gulopyranosyluronic acid present in the alginate [98]. The strength and firmness is also dependent on the starting concentration of the alginate, the higher the concentration, for example 5% (w/v), gives very strong, firm gels, even though it is only 5% calcium alginate and 95% water.

Sodium alginate is used in food preparations as a thickening agent, a stabilizer, and an emulsifier in ice cream, cream cheese, salad dressings, frozen foods, pharmaceuticals, and so forth. Calcium alginate is the major ingredient in the "pimento" found in stuffed olives. A very important use of calcium alginate is the formation of gels that are used to encapsulate enzymes, hormones, drugs, and whole cells for carrying out various processes while being immobilized [99]. See *Fig. 7* for the conformational structure of a segment of alginate.

Alginates are also produced extracellularly by some bacteria, such as *Pseudomonas aeruginosa* and *Azotobacter vinelandii* [100], where they are believed to play a role in biofilm formation, pathogenesis, and soil aggregation. Their gels, however, are inferior to seaweed alginates, because of the presence of *O*-acetylation, which inhibits gel formation.

9 Occurrence and Properties of Plant Exudate Polysaccharides

Several complex polysaccharides are secreted by plants to seal wounds. Gum arabic is a complex material, containing protein, lipid, and carbohydrate, produced by Acacia trees found in the arid regions of Africa, in Nigeria, Mauritania, Senegal, and the Republic of Sudan. The

structure of the polysaccharide portion has a main chain of D-galactopyranose residues linked β-(1→3) and D-glucopyranosyl uronic acid linked β-(1→6). The main chains have branch chains of two to five residues, consisting of α-L-arabinofuranosyl, α-L-rhamnopyranosyl, β-D-glucopyranosyl uronic acid, and 4-*O*-methyl-β-D-glucopyranosyl uronic acid [101]. The latter two uronic acids occur most frequently at terminal ends of the branched chains.
Another plant exudate is gum ghatti or Indian gum that can be obtained from a large tree grown in the deciduous forests of India and Sri Lanka. Gum ghatti is composed of L-arabinofuranose, D-galactopyranose, D-mannopyranose, D-glucopyranose uronic acid, and D-xylopyranose in approximately the molar ratios of 10:6:2:2:1 [102]. A third exudate gum is gum tragacanth that is primarily obtained from trees growing in Iran, Syria, and Turkey. It is a highly branched arabinogalactan with α-D-xylopyranose and α-L-fucopyranose branch residues [103]. These gums are primarily used to increase viscosity, provide body, stabilize emulsions, and suspend other materials and have been used for thousands of years in confectioneries, cosmetics, textiles, coatings, paints, pastes, and polishes.

10 Occurrence of Carbohydrates in Bacterial Polysaccharides

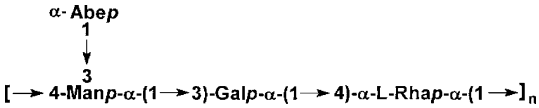
A large number of bacterial polysaccharides are known [104]. The major structural component of the bacterial cell wall is a polysaccharide, known as murein and composed of a repeating unit of one *N*-acetyl-D-glucosamine and an *O*-lactyl substituted *N*-acetyl-D-glucosamine (*N*-acetyl-D-muramic acid) see *Sect. 7.3*.

10.1 Xanthan, a Water-Soluble Bacterial Polysaccharide

In the 1950s, the US Department of Agriculture's Northern Regional Research Laboratories in Peoria, Illinois screened bacterial cultures to obtain a replacement for the plant exudates, which had become rare and expensive. They found that *Xanthomonas campestris*, when grown on D-glucose in an aerobic submerged fermentation, produces xanthan, a water-soluble polysaccharide gum [105,106]. It has a cellulose backbone of β-(1→4) linked D-glucopyranose residues with a trisaccharide of D-mannopyranose linked β-(1→4) to D-glucopyranosyl uronic acid linked β-(1→2) to a D-mannopyranosyl [β-D-Man *p*-(1→3)-β-D-Glc *p*UA-(1→2)-α-D-Man *p*-(1→3)-] attached to every other D-glucose residue in the cellulose chain by an α-(1→3) linkage [107,108]. Some of the nonreducing terminal D-mannopyranose residues of the trisaccharide have a cyclic six-membered pyruvic acid ketal attached to C4 and C6, and some of the inner D-mannopyranose units are acetylated at C6 [109].
The branching of the cellulose chain by the trisaccharide makes the otherwise insoluble cellulose molecule water-soluble. At low concentrations, xanthan produces high viscosities at low temperatures. These properties provide a number of uses as a thickener and bulking agent for prepared foods, such as salad dressings, syrups, toppings, relishes, ice cream, and baked goods. It is also used as a carrier and emulsifying agent in cosmetics and pharmaceuticals [110].
There are other bacterial gel polysaccharides with different properties, composed of D-glucopyranose, L-rhamnopyranose, D-glucuronic acid, and L-mannopyranose that are obtained from *Pseudomonas elodea* (syn. *Sphingomonas elodea*), which produces gellan, and also species of *Alcaligenes* that produce welan and rhamsan [111].

10.2 Pathogenic Bacterial Capsular Polysaccharides

Salmonella species have an O-antigen that is a heteropolysaccharide, imparting pathogenicity to the organism. It is composed of a repeating tetrasaccharide unit, made up of a sequence of D-mannopyranose, L-rhamnopyranose, and D-galactopyranose, with a variable 3,6-dideoxy-D- or L-hexose linked to the D-mannose residue as a branch residue [112]. *Salmonella* easily mutates and over 100 different kinds of capsular polysaccharides have been identified for various species and mutants. The polysaccharides vary according to the linkage positions and the α- or β-configurations [112]. They also vary with the nature of the attachment of four different 3,6-dideoxy carbohydrate residues: D-paratose (3,6-dideoxy-D-glucopyranose), D-tyvelose (3,6-dideoxy-D-mannopyranose), D-abequose (3,6-dideoxy-D-galactopyranose) and L-colitose (3.6-dideoxy-L-galactose) [113]. *Salmonella* readily mutates the structure of this polysaccharide, giving very wide diversity of structures and a basis of avoiding antibody neutralization.



Scheme 4
A typical *O*-antigen *Salmonella* capsule polysaccharide

Streptococcus pneumoniae strains constitute a large group of pathogens, responsible for bacterial pneumonia. All virulent strains have a voluminous capsule that is responsible for their pathogenicity. The capsules are all relatively complex heteropolysaccharides with diverse structures. The monosaccharide residues contain D-glucopyranose, D-glucopyranose uronic acids, L-rhamnopyranose, and *N*-acetyl-D-glucosamine [113]. There are some capsules with unusual carbohydrates, such as sugar alcohols (glycerol, erythritol, D-threitol, and ribitol), amino-sugars (*N*-acetyl-L-fucosamine, *N*-acetyl-D-mannosamine, *N*-acetyl-2-amino-2,6-dideoxy-L-talose [commonly called L-pneumosamine]), as well as D-galactofuranose and phosphodiesters. The structures are repeating tetra-, penta-, and hexa-saccharides. Like *Salmonella O*-antigens, the repeating units have permuted glycosidic linkages at different positions and with either α- or β-configurations to give a wide diversity of structures, with over 120 different known structures [114].

An unusual acidic polysaccharide capsule is produced by the Gram-negative pathogens, *Neisseria meningitides* and *Escherichia coli*. These polysaccharides contain the unusual nine-carbon sugar, *N*-acetyl-D-neuraminic acid, which is formed by an enzyme catalyzed aldo-condensation between the methyl group of pyruvic acid and the aldehyde group of *N*-acetyl-D-mannosamine, followed by the formation of a six-membered ring with a three hydroxy-carbon side chain. The sugar acid is linked α-(2→8) or α-(2→9) with itself to give a linear polysaccharide, called colominic acid [115,116] (see ❯ *Fig. 7* for the structure of α-(2→8) colominic acid). An interesting variation is the colominic acid produced by *E. coli* Bos-2 that has the alternating sequence of α-(2→8) and α-(2→9) linkages [117].

11 Properties and Occurrence of D-Fructose in Polysaccharides

Polysaccharides that exclusively contain D-fructose are known as fructans and there are two known kinds, inulin and levan. Inulin is a polysaccharide containing β-D-fructofuranose linked (2→1) [118]. Inulins are found in the roots and tubers of the family of plants known as the Compositae, which includes asters, dandelions, dahlias, cosmos, burdock, goldenrod, chicory, lettuce, and Jerusalem artichokes. Other sources are from the Liliacae family, which includes lily bulbs, onion, hyacinth, and tulip bulbs. Inulins are also produced by certain species of algae [119]. Several bacterial strains of *Streptococcus mutans* also produce an extracellular inulin from sucrose [120].

Levan is a polysaccharide containing β-D-fructofuranose residues linked (2→6) with (2→1) branch linkages. They are primarily found in grasses [119] and are produced extracellularly by several bacterial strains of *Bacillus subtilis*, *Aerobacter levanicum* (syn. *Erwinia herbicola*) [121], and *Streptococcus salivarius* [122]. They are of higher molecular weight than the inulins, having 100–200 D-fructofuranose residues per molecule. The branch chains are relatively short, containing 2–4 D-fructofuranose residues.

12 Properties and Occurrence of Sugar Alcohols

12.1 Glycerol

When aldoses or ketoses are reduced, sugar alcohols are formed. For example, glycerol is a simple, three carbon sugar alcohol, formed by the reduction of glyceraldehydes. It is found as a major component in two types of lipids, triacylglycerol (triglyceride fats and oils) and phospholipids. In the former, the three hydroxyl groups of glycerol are esterified by fatty acids. In the latter, glycerol is esterified by two fatty acids at the first two carbons and by phosphoric acid at the third carbon. The phosphoric acid is further esterified by the hydroxy groups of ethanolamine, *N*,*N*,*N*-trimethyl ethanolamine (choline), or by the hydroxy group of L-serine. The triglycerides make up the well-known fat deposits found in adipose tissue and the phospholipids are major components found in the lipid bilayers of membranes of cells and organelles and play important roles in nerve transmission. Glycerol is also a common component in the teichoic acids (see *Sect. 12.3*).

Free glycerol is obtained from the saponification of fats and oils. It is a slightly sweet, highly water-soluble liquid. It has the ability to absorb water, making it a valuable humectant and an emollient for skin conditioners. It is also used as a plasticizer in the formation of polymeric materials and is used in the manufacture of pharmaceuticals and the explosive, trinitroglycerine.

12.2 Properties and Occurrence of Free Sugar Alcohols, D-Glucitol, D-Mannitol, Ribitol, Xylitol, and D-Arabinitol

D-glucose can be reduced either chemically or enzymatically to give D-glucitol (frequently called D-sorbitol). It was first obtained from the fresh juice of the berries of the mountain ash [123]. D-Glucitol occurs widely in plants, being found in algae and higher plants. It is especially prevalent in red seaweed, where it occurs to the extent of 10–14% by weight [124]

and is found in relatively large amounts in pears, apples, cherries, prunes, peaches, and apricots, where it imparts a sweet taste to these fruits [125].

D-Mannitol is also widely distributed in plants and was the first crystalline sugar alcohol to be obtained from a natural source, the manna ash [126]. It is also found in large amounts (70–90% w/w) in the exudates of the olive and the plane trees [127]. D-Mannitol is found in relatively large amounts in seaweeds of *Laminaria* and *Mycrocystis* species [128]. Species of the mold *Aspergillus*, produce D-mannitol by fermentation, using D-glucose or acetate as carbon sources [129].

Ribitol, the sugar alcohol from the reduction of D-ribose, is found as a constituent of the vitamin riboflavin (vitamin B_2). It is also a constituent of the teichoic acids, see *Sect. 12.3*.

The reduced product of D-xylose is xylitol, which has a very sweet taste and also imparts an unusual cooling sensation. It is found in several fruits, such as plums, raspberries, and strawberries, where it occurs to the extent of about 1% by weight and gives a distinctive and pleasant taste to these fruits. D-Arabinitol is found in mushrooms in amounts as high as 9–10% by weight [130], in lichens [131], and in avocado seeds [132]. D-Arabinitol is produced by some species of yeast (*Debaryomyces subglobosus* and *Endomycopsis chodati*) through fermentation of D-glucose, D-mannose, and sucrose [133].

12.3 Sugar Alcohols in Teichoic Acids

The teichoic acids are bacterial polymers of sugar alcohols (glycerol or ribitol) and phosphoric acid joined end to end by phosphodiester linkages to the primary alcohol groups. They are found in conjunction with the peptidoglycan of Gram-positive bacterial cell walls [134]. The C2 hydroxy group of glycerol is frequently acylated by D-alanine or glycosylated by *N*-acetyl-D-glucosamine or D-glucopyranose [135]. In some *Bacillus* species phosphodiester linkages join glycerol units between the C1 hydroxy of one unit to the C2 hydroxy of the next unit. Ribitol residues are joined together between the C1 hydroxyl of one unit to the C5 hydroxy of the adjoining unit. The C3 or C4 hydroxy groups can be acylated by D-alanine and the C2 hydroxyl group can be glycosylated by a number of different carbohydrate residues, for example, *N*-acetyl-D-glucosamine, D-glucopyranose, and di- or tri-saccharides of D-glucopyranose. More complex teichoic acids occur that have a repeating sequence of glycerol joined (1→4) to *N*-acetyl-D-glucosamine by a phosphodiester linkage and *N*-acetyl-D-glucosamine joined (1→3) by a phosphodiester linkage to the next glycerol unit [136].

13 Properties and Occurrence of Deoxy Sugars

The most abundant and probably best known deoxy sugar is 2-deoxy-D-ribofuranose, which is found as the carbohydrate component in the genetic polymer, deoxyribonucleic acid, the carrier of genes in the chromosomes of living organisms. Other deoxy sugars include 6-deoxy-L-mannose (L-rhamnose), which is found in glycosides and in *Salmonella* sp. *O*-antigen polysaccharides (see *Sect. 10.2*). The third deoxy sugar is 6-deoxy-L-galactose (L-fucose), found in glycocoproteins, such as the blood group substances (see *Sect. 14*). The fourth deoxy sugar, 6-deoxy-D-glucose (D-quinovose) is found in acarbose, the naturally occurring pseudotetrasaccharide, produced by *Actinoplanes* sp. fermentation. Acarbose is an inhibitor of α-glucosidase [137]. It also occurs in some of its analogues, such as

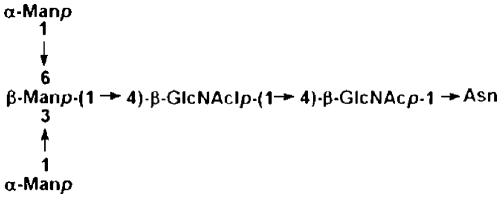
α-acarviosine-(1→6)-cellobiose and α-acarviosine-(1→6)-lactose, which act as inhibitors of β-glucosidases and β-galactosidase [138]; and α-maltohexaose-(1→4)-acarbose and α-maltododecaose-(1→4)-acarbose, which are potent inhibitors of α-amylases in the nM range [139]. Four naturally occurring 3,6-dideoxy sugars appear in the different *Salmonella* sp. *O*-antigen capsular polysaccharides, see *Sect. 10.2*, for their names and structures. Four 2,6-dideoxy sugars (2,6-dideoxy-D-ribo-hexaose [D-digitoxose]; 2,6-dideoxy-3-*O*-methyl-D-ribo-hexose [D-cymarose]; 2,6-dideoxy-D-xylo-hexose [D-boivinose]; and 2,6-dideoxy-3-*O*-methyl-D-xylo-hexose [D-sarmentose]) are found in a number of plants, as the carbohydrate component of the so-called cardiac glycosides [140].

14 Properties and Occurrence of Carbohydrates in Glycoproteins

Glycoproteins make up a large class of important biological compounds. It is estimated that over 75% of the known (~3,000) proteins are glycosylated. The carbohydrates are believed to mediate a number of biological functions: (1) the correct folding of a protein tertiary structure after biosynthesis, (2) establishment and stabilization of protein conformation, (3) secretion of proteins through membranes, (4) control of protein turnover, (5) protection of proteins from proteinase hydrolysis, (6) increase in protein water-solubility, (7) biological recognition involved in growth, cell differentiation, organ formation, fertilization, processes of bacterial and viral infections, formation of tumors, tumor metastasis, allergies, and autoimmune diseases.

Carbohydrates are primarily attached to proteins in two ways: (1) by linkage of C1 to the amide nitrogen of L-asparagine, giving *N*-linked carbohydrate proteins and (2) by formation of acetal linkages with the hydroxyl group of L-serine or L-threonine, giving *O*-linked carbohydrate proteins. The carbohydrate can be a single monosaccharide residue or it can be an oligosaccharide, containing several monosaccharide residues. There are six major carbohydrates involved in glycoproteins; they are *N*-acetyl-D-glucosamine, *N*-acetyl-D-galactosamine, D-mannopyranose, D-galactopyranose, L-fucose, and *N*-acetyl-D-neuraminic acid [141].

The attachment of carbohydrate to nitrogen is invariably by *N*-acetyl-D-glucosamine and the attachment to oxygen is invariably by *N*-acetyl-D-galactosamine. *N*-linked carbohydrates are invariably composed of a "core" pentasaccharide [142] of the following structure:



Scheme 5
Core oligosaccharide for *N*-linked glycoproteins

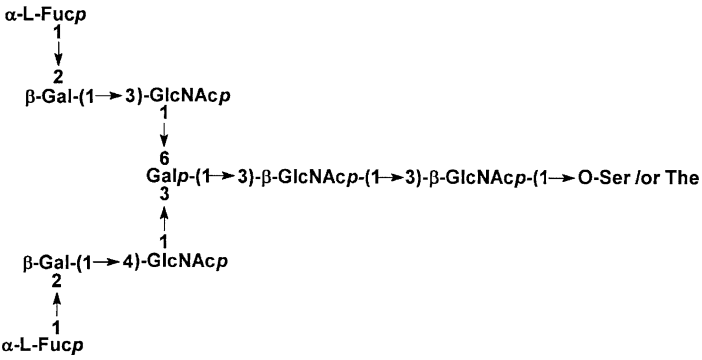
The *N*-linked glycosides can be classified into three families that result from the further attachment of monosaccharides to the two branched D-mannopyranose residues at the nonreducing-end of the core pentasaccharide. These additional residues make up the variable regions of the oligosaccharides. The first family is the "high mannose family" that has additional

α-D-mannopyranose residues attached to the two terminal D-mannopyranose residues of the core pentasaccharide. These residues are linked α-(1→3) and α-(1→6). Substitution is terminated by α-(1→2) linkages of the D-mannopyranose residues. An example of this type of *N*-glycoside is found attached to ovalbumin.

The second family of *N*-linked oligosaccharides are called the "lactosamine family" in which the D-mannose residues of the core pentasaccharide are substituted 1→2 by lactosamine, which is a lactose analogue with *N*-acetyl-D-glucosamine substituted for D-glucopyranose at the reducing-end of lactose. The lactosamine is frequently substituted by *N*-acetyl-D-neuraminic acid (for the structure of *N*-acetyl-D-neuraminic acid, see the monomer residue in colominic acid, ❯ *Sect. 8.2* and ❯ *Fig. 7*) linked 2→3 or 2→6 [142,143]. The third family has a mixed structure of the high mannose and lactosamine families [144,145,146,147].

The *O*-linked saccharides are not as common as the *N*-linked saccharides. A relatively simple saccharide is that attached to L-threonine of the highly glyosylated (1 out of every 3 amino acid residues), antifreeze protein found in the blood sera of fish living in the Arctic and Antarctic waters. This glycoside is a disaccharide, β-D-Galp-(1→3)-α-D-GalNAc *p* [147].

Lactosamine and isolactosamine (lactosamine analogue with D-galactopyranose linked β-(1→6) to *N*-acetyl-D-glucosamine) are a well characterized set of *O*-linked oligosaccharides that make up the ABO human blood group substances [148]. They are on the surface of erythrocytes and divide human blood into four distinct types. The following core structure makes up the *O*-blood type and is found in all four blood types:



Scheme 6
Type O human blood group oligosaccharide

The core is composed of nine monosaccharide residues with two chains terminating in β-D-galactopyranose residues linked (1→4) and (1→3). The other three human blood groups have two additional monosaccharide residues added to the two β-D-galactopyranose residues of the two chains. A-type human blood group has two α-D-galactosamine residues, one each linked (1→3) to the ends of the two chains; B-type has two α-D-galactopyranose residues, one each linked (1→3) to the ends of the two chains; and AB-type has a mixture of α-D-galactosamine and α-D-galactopyranose residues, one each linked (1→3) to the two chains. *O*-Type blood is the universal blood donor and can give blood to all four types, but can accept blood only from *O*-type donors; AB-type blood is the universal blood acceptor and can accept blood from all four types; A-type and B-type can accept blood from *O*-type donors or from donors

with their own blood type, as they make antibodies against either the A- or B-types, and AB-type donors, precipitating the blood [148,149,150].
There are additional blood group variations. A common variation is an isomerization in which α-L-fucopyranose is moved from β-D-Galp to β-D-GlcNAcp and linked (1→4) to give the Lewis-a blood type. A second and related variation is the addition of another α-L-fucopyranose residue to β-D-GlcNAcp linked (1→4) to give two α-L-Fucp residues on the first chain, giving Lewis-b blood type. These kinds of variations can occur for each of the ABO blood types, giving *O*-type-Lewis-a, *O*-type-Lewis-b, A-type-Lewis-a, and so forth: *O*-Le^a, *O*-Le^b, A-Le^a, A-Le^b, B-Le^a, B-Le^b, AB-Le^a, AB-Le^b [150].

15 Separation and Purification of Carbohydrates

The source and the specific physical and chemical properties of carbohydrates determine the methods that are used for their separation and purification. Mono-, di-, tri- and sometimes higher-saccharides, for example maltodextrins, isomaltodextrins, and raffinose-sucrose dextrin series, are usually quite soluble in water. Carbohydrates, thus, are often obtained by the extraction of natural materials with hot water. As many impurities as possible are removed in an extraction mixture, such as salts, proteins, and lipids. Salts can be removed by precipitation and/or the use of ion exchangers. Lipids are removed with organic solvents, such as a 2:1 mixture of chloroform and methanol, and proteins are precipitated with acids and heat. High amounts of alkali and acid, however, should be avoided. Frequently, some of the last impurities in the aqueous extract, especially colored yellow to brown materials, can be removed by adding activated charcoal and filtering it out to give a clear solution before the extract is concentrated. The concentrated carbohydrate extract is obtained at an elevated temperature (50–60 °C) and an organic solvent such as methanol or ethanol is slowly added to the point where the clear solution just becomes cloudy. The solution is then cooled to ≈20 °C to give crystallization of the carbohydrate and then 4 °C to obtain additional crystals. Monosaccharides and disaccharides will often crystallize, while higher oligosaccharides are frequently obtained as amorphous precipitates that can be removed by centrifugation or filtration and dehydrated.
Many different chromatographic methods of separation (on charcoal, BioGel, silica gel, hydroxyapatite, paper) can be used on a preparative scale to give pure materials that can be studied and used even though they are not crystalline. Two typical examples are given for the isolation, purification, and crystallization of a monosaccharide, α-D-xylopyranose, and a disaccharide, lactose, from natural sources.

15.1 Isolation and Purification of α-D-Xylopyranose from Corn Cobs

Coarsely ground corn cobs or crude xylan can be used as starting materials. The xylan in either source is hydrolyzed with 7% (v/v) sulfuric acid by refluxing for 2.5 h. The mixture is filtered through cloth on a Büchner funnel with as much liquid as possible obtained by suction. The residue is washed with an equal volume of water by suspension as thin slurry and then filtered. A few drops of 1-octanol are added to the combined filtrates that are neutralized with barium carbonate. The solids (primarily barium sulfate) in the mixture are filtered and the residue washed by suspension in water and filtered. If corn cobs are used as the starting material,

approximately one-sixth of a cake of baker's yeast is finely suspended in 10–15 mL of water and added to the clear filtrate that is covered with a cotton plug. Fermentation is allowed to go ≈15 h at 37 °C to remove D-glucose (if crude xylan is used, this step can be omitted). After removal of the yeast, activated charcoal is added with an equal weight of Celite 535 and the mixture is filtered by suction. The filtrate is concentrated to a syrup under reduced pressure at 50–60 °C, and three volumes of methanol are added with stirring. The solution is filtered and concentrated to syrup under reduced pressure. The syrup is dissolved in water and passed through a column (5.5 × 50 cm) of equal amounts of activated charcoal and Celite 535. The column is then washed with 6 L of water and the washings and the original filtrate are combined and concentrated under pressure to a syrup (~30 mL) that is filtered through a coarse sintered-glass filter. The filtrate is allowed to stand at 20 °C until crystallization of α-D-xylopyranose is complete. The crystals are removed by filtration and washed with cold (4 °C) 85% (v/v) aqueous methanol. A second crop of crystals are usually obtained by placing the supernatant at 4 °C [151].

15.2 Isolation and Purification of Lactose from Milk

Commercial skimmed (defatted) milk contains ~3% casein, 0.7% albumin, 4–5% lactose, and 1% minerals, along with small amounts of lactosamine and lactosamine oligosaccharides, with the remainder being water. The casein is first precipitated by warning to 40 °C and the addition of 1:10 (v/v) glacial acetic acid and water to 200 mL of milk, with continuous stirring. The dilute acetic acid is added until casein no longer separates. The precipitated casein is removed by centrifugation. Then 5 g of calcium carbonate is immediately added and stirred for ~5 min and then the solution is heated to boiling for ~10 min. This produces almost complete precipitation of the albumin, which is removed by vacuum filtration. The filtrate is concentrated by roto-vacuum evaporation to ~30 mL. Then 166 mL of hot ethanol is added, along with 5 g of activated charcoal; after it has been mixed well, the warm solution is filtered through a bed of Celite. The clear filtrate is allowed to stand 15–25 h at 20 °C or longer for crystallization. When crystallization is complete, the crystals are removed by filtration and a second crop of crystals are obtained by placing the clear solution at 4 °C.

15.3 Analysis, Isolation, and Purification of Monosaccharides and Oligosaccharides

Individual monosaccharides and their reduced sugar alcohols can be separated and analyzed by multiple ascent silica-gel, thin-layer chromatography [152], as well as the more complex mixtures of a series of homologous oligosaccharides, such as maltodextrins, isomaltodextrans, cellodextrins, chitosan- and chito-dextrins, cyclomaltodextrins, and the raffinose-sucrose dextrins can be quantitatively analyzed by multiple ascent silica-gel, thin-layer chromatography (TLC), followed by scanning densitometry [152,153,154]. See ❯ *Fig. 8* for a TLC separation of maltodextrins and cyclomaltodextrins. Pure individual oligosaccharides can be obtained in 50–200 mg amounts by preparative descending paper chromatography, using 70:30 (v/v) propanol-1/water solvent on 23 × 54 cm Whatman 3MM paper for 24–36 h on which the saccharides are separated, and detected by $AgNO_3$/NaOH/$Na_2S_2O_3$ development of a 1-cm strip

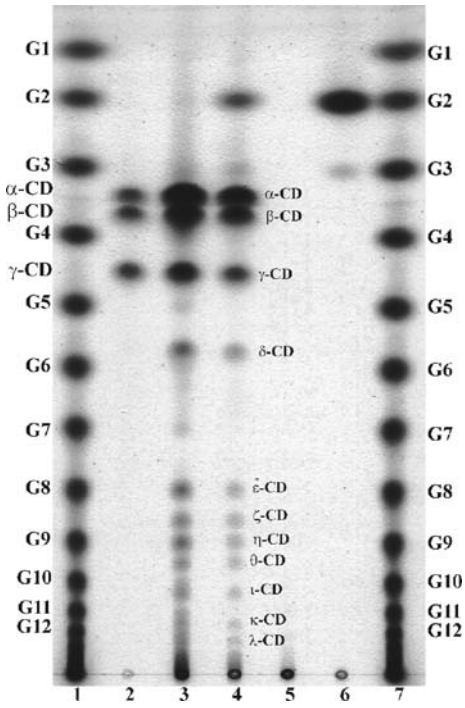


Figure 8

Thin-layer silica-gel chromatographic (TLC) separation of maltodextrins and cyclomaltodextrins and using Whatman K5 silica gel plate, irrigated 18 cm three-times with 85:25:55:50 volume proportions of acetonitrile, ethyl acetate, propanol-1, water solvent. The carbohydrates were visualized on the plate by dipping it into a methanol solution, containing 0.3% (w/v) *N*-(1-naphthyl)ethylene diamine and 5% (v/v) sulfuric acid, dried, and heated at 120 °C for 10 min. The saccharides can be quantitated by scanning densitometry [152,153]. Lanes 1 and 7 are D-glucose (G1) and maltodextrins (G2 to G12); lane 2, cyclomaltohexaose (α-CD), cyclomaltoheptaose (β-CD), and cyclomaltooctaose (γ-CD); lane 3, a mixture of maltodextrins and cyclomaltodextrins; lane 4, cyclomaltodextrins (α-CD) to (λ-CD), with 6 to 16 D-glucopyranose residues; and lane 6 is maltose and maltotriose. From [166], reproduced by permission of the publisher, Elsevier Press

on each side of the paper. The paper is sectioned, and then the individual saccharides are eluted from the sectioned pieces of paper in pure form [155]. They can also be obtained in pure form in larger quantities by charcoal-Celite column chromatography: for example, cellodextrins [156], isomaltodextrins [157], maltodextrins [158], and xylodextrins [159] have been prepared in this way. Sialyl oligosaccharides from human milk have been separated by ion-exchange chromatography [160] and maltodextrins have been separated by high performance liquid chromatography (HPLC) [161,162].

Capillary electrophoresis has been used to separate and analyze synthetically modified carbohydrates in the nanogram to milligram range [163]. Fluorophore-assisted capillary electrophoresis (FACE) has successfully been used to separate nanogram amounts of maltodextrins, containing 4–76 D-glucose residues [164,165,166], see *Fig. 9*. Many carbohydrates can be analytically separated by matrix-assisted, laser desorption, ionization-time of flight,

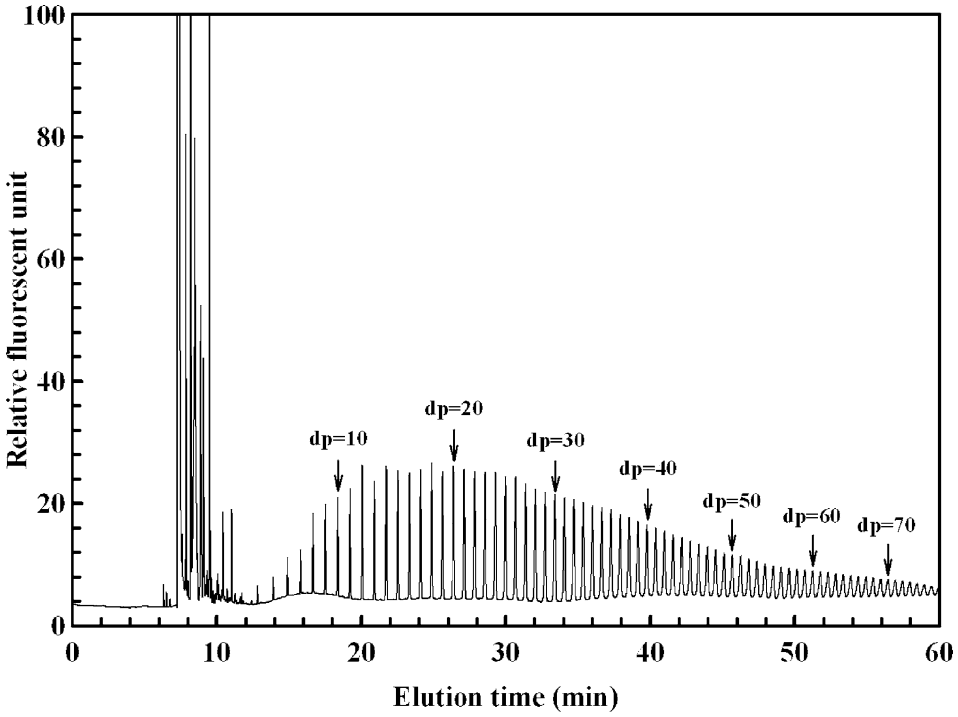


Figure 9
Fluorophore-assisted capillary electrophoresis (FACE) analysis of maltodextrins from G2 to G77. From [166], reproduced by permission of the publisher, Elsevier Press

mass spectrometry (MALDI-TOF MS). For example, cyclomaltodextrins, containing 6–25 D-glucose residues are readily analyzed by this technique [166] (see *Fig. 10*).

15.4 Separation and Purification of Water-Soluble Polysaccharides

Water soluble polysaccharides found in bacterial fermentations or produced enzymatically, such as the dextrans or xanthan, and so forth can be obtained and purified from the aqueous solutions by the addition of two volumes of ethanol, centrifugation, and the resulting pellet redissolved by slowly adding it to boiling water or by suspending it in water and autoclaving at 121 °C for 30 min and then reprecipitating it with two volumes of ethanol. The resulting precipitate can be obtained as a dry powder by treating it several times (5–10) with anhydrous acetone and then once with anhydrous ethanol, and dried in a vacuum oven at 40–50 °C for 12 h. The acetone removes the bulk of the water and the ethanol removes the last traces of water as the 95% azeotrope.

Dextrans with different structures synthesized by distinct dextransucrases that were elaborated by the same strains of *Leuconostoc mesenteroides* have been separated by differential ethanol precipitation, using different concentrations of ethanol, for example strains B-742,

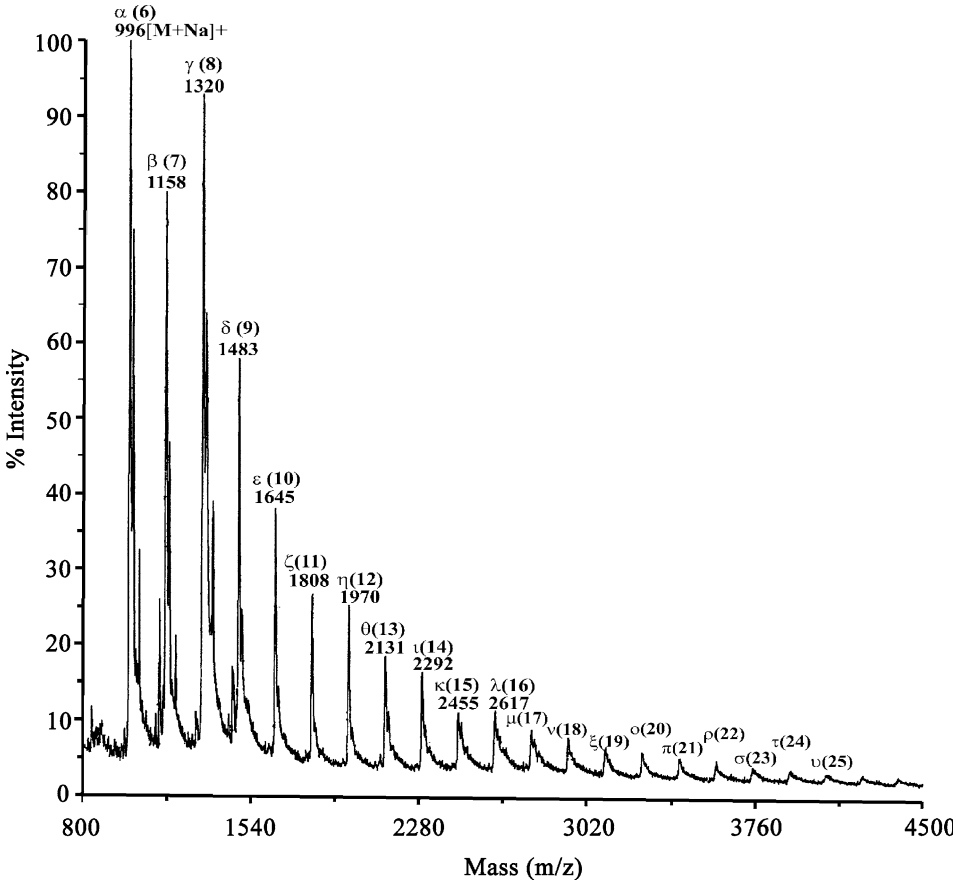


Figure 10
Matrix-assisted laser desorption ionization-time of flight mass spectrometry (MALDI-TOF MS) analysis of cyclomaltodextrins with 6 to 27 D-glucopyranose residues. The numbers in parentheses above the peaks are the number of D-glucopyranose residues in the cyclomaltodextrins. From [166], reproduced by permission of the publisher, Elsevier Press

B-1254, and B-1299 each gave two dextrans with different structures that were separated from each other; dextran and alternan, produced by strain B-1355, were also separated in this way [167].

15.5 Separation and Purification of Water-Insoluble Polysaccharides, Starch and Cellulose

The separation and purification of starch granules from plant extracts (see ❯ Sect. 2 in ❯ Chap. 6.2) and the fractionation of amylose and amylopectin is given in ❯ Sect. 5 in ❯ Chap. 6.2. The separation of cellulose and hemicelluloses from plant materials and from each other is given in ❯ Sect. 3.2 in ❯ Chap. 6.3.

15.6 Separation and Purification of Cyclomaltodextrins

The cyclomaltodextrins (α-CD, β-CD, and γ-CD) can be selectively obtained from a fermentation culture or an enzyme digest of cyclomaltodextrin glucanotransferase reaction with solubilized starch. The majority of the cyclomaltohexaose (α-CD) can be separated from cyclomaltoheptaose (β-CD) and γ-CD by their selective precipitation with p-cymene from the culture supernatant or from an enzyme digest [168]. The α-CD can then be precipitated from the supernatant with cyclohexene, which is extracted with acetone to remove the cyclohexene and the α-CD can be crystallized from water or a propanol-1/water solution [169]. The p-cymene precipitates of β-CD and γ-CD are put into a water solution and β-CD selectively precipitated from γ-CD with fluorobenzene. The γ-CD is then precipitated with anthracene saturated in diethyl ether. After the removal of the fluorobenzene from β-CD with acetone or ethanol extraction, β-CD can be crystallized from water, and after the removal of anthracene with acetone or ethanol extraction from γ-CD, it can also be crystallized from water [170,171]. The selective precipitations of the cyclomaltodextrins with various organic molecules is based on the selective formation of complexes of the organic molecules with the specific sizes of the cyclomaltodextrins and the relatively hydrophobic interior cavities of the cyclomaltodextrins [166,167,168].

15.7 Release of Oligosaccharides from Glycoproteins

The *O*-linked oligosaccharides are relatively easily released from the protein by a β-elimination reaction, using mild alkali (0.05–0.5-M NaOH) and temperatures of 0–45 °C for 15–216 h [172,173,174]. A standard procedure is 0.1-M NaOH at 37 °C for 48 h. Conditions, however, must be determined for each glycoprotein. The asparagine *N*-linked glycosides can be cleaved by hydrazinolysis [175]. The glycoprotein is heated at 100 °C with anhydrous hydrazine for 8–12 h in a sealed tube. Various endoglycosidases, such as endo-β-*N*-acetylglucosaminidase have also been used to release the oligosaccharides from glycoproteins [176,177]. A chemical method that releases both *O*- and *N*-linked oligosaccharides from glycoproteins involves trifluoromethane sulfonic acid (TFMS) [178]. TFMS reactions are performed at 0 °C for 0.5–2 h under nitrogen. After reaction, the mixture is cooled below −20 °C in a dry ice-ethanol bath and slowly neutralized by the addition of 60% (v/v) aqueous pyridine that is previously cooled to −20 °C. More information on the structure and analysis of glycoproteins may be found in Chap. 8.

References

1. Bonner WA (1991) Origins Life Evolut Biosph 21:72
2. Bailey J, Chrysostomou A, Hough JH, Gledhill TM, MeCall A, Clark S, Ménard F, Tamura M (1998) Science 281:672
3. Lobry de Bruyn CA, Alberda van Ekenstein W (1895) Rec Trav Chim 14:156; 203; (1896) 15:92; (1897) 16:241, 262, 274, 282; (1899) 18:147
4. French D (1954) Adv Carbohydr Chem 9:149
5. Avigad G, Feingold DS, Hestrin S (1956) Biochim Biophys Acta 20:129
6. Feingold DS, Avigad G, Hestrin S (1957) J Biol Chem 224:295

7. Avigad G (1957) J Biol Chem 229:121
8. Stodola FH, Sharpe ES, Koepsell HJ (1956) J Am Chem Soc 78:2514
9. Sharpe ES, Stodola FH, Koepsell HJ (1960) J Org Chem 25:1062
10. Khan R (1976) Adv Carbohydr Chem Biochem 33:236
11. Ballard JM, Hough L, Richardson AC, Fairclough PH (1973) J Chem Soc Perkin Trans I 1524
12. Hough L, Phadnis SP, Tarelli E (1975) Carbohydr Res 44:37
13. Parolis H (1976) Carbohydr Res 48:132
14. Hough L, Phadnis SP (1976) Nature 263:800
15. Hough L, Khan R (1978) Trends Biol Sci 3:61
16. Kuhn R, Gauhe A (1962) Chem Ber 95:518
17. Got R, Font J, Bourrillon R, Cornillot P (1963) Biochim Biophys Acta 74:247
18. Kuhn R, Ekong D (1963) Chem Ber 96: 683
19. Kalf GF, Rieder SV, J (1958) Biol Chem 230:691
20. Clegg JS, Filosa MF (1961) Nature 192:1077
21. Stewart LC, Richtmeyer NK, Hudson CS (1950) J Am Chem Soc 72:2059
22. Ingram J, Bartels D (1996) Ann Rev Plant Physiol Plant Molec Biol 47:377
23. Beattie GM, Leibowitz G, Lopez DA, Levine F, Hayek A (2000) Cell Transplant 9:91
24. Han B, Bischof N (2004) Cell Preserv Technol 2:91
25. Jane J-l, Kasemsuwan T, Leas S, Zobel H, Robyt JF (1994) Starch/Stärke 46:121
26. Hizukuri S, Takeda Y, Yasuda M, Szuki A (1981) Carbohydr Res 94:205
27. Hizukuri S (1991) Carbohydr Res 217:251
28. Leach HW, Schoch TJ (1962) Cereal Chem 39:318
29. Banks W, Greenwood CT (1975) Starch and its Components, Edinburgh University Press, Edinburgh, pp 15, 30
30. Whistler RJ, Daniel JR (1984) In: Whistler RJ, BeMiller JN, Paschall EF (eds) Starch: Chemistry and Technology. Academic Press, New York, pp 153–178
31. BeMiller JN, Whistler RL (1996) Starch Derivatives. In: Fennema OR (ed) Food Chemistry. Marcel Dekker, New York, pp 201–203
32. Fox JD, Robyt JF (1992) Carbohydr Res 227:163
33. Robyt JF, Choe J-Y, Hahn RS, Fuchs EB (1996) Carbohydr Res 281:203
34. Robyt JF, Choe J-Y, Fox JD, Hahn RS, Fuchs EB (1996) Carbohydr Res 283:141
35. Ruttenberg MW, Solarek D (1984) Starch Derivatives: Production and Uses. In: Whistler RJ, BeMiller JN, Paschall EF (eds) Starch: Chemistry and Technology, 2nd edn. Academic Press, San Diego, pp 311–388
36. Geddes R, Harvey JD, Wills PR (1977) Biochem J 30:257
37. Wanson J-C, Drochmans P (1968) J Cell Biol 38:130; (1972) 54:206
38. Geddes R (1985) Glycogen. In: Aspinall GO (ed) The Polysaccharides, vol. 3. Academic Press, San Diego, p 316
39. Preiss J, Walsh DA (1981) The Comparative Biochemistry of Glycogen and Starch. In: Ginsburg V, Robbins P (eds) Biology of Carbohydrates, vol. 1. Wiley, New York, p 203
40. Okada G, Hehre E (1973) Carbohydr Res 26:240; Tao BY, Reilly PJ, Robyt JF (1988) Carbohydr Res 181:163
41. Robyt JF (1995) Adv Carbohydr Chem Biochem 51:133
42. Jeanes A, Haynes WC, Wilham CA, Rankin JC, Melvin EH, Austin MJ, Cluskey JE, Fisher BE, Tsuchiya HM, Rist EE (1954) J Am Chem Soc 76:5041
43. Tsiyisaka Y, Mitsuhashi M (1993) Pullulan. In: Whistler RL, BeMiller JN (eds) Industrial Gums: Polysaccharides and Their Derivatives, 3rd edn. Academic Press, San Diego, pp 447–460
44. Bender H, Lehmann J, Wallenfels K (1959) Biochim Biophys Acta 36:309
45. Bouveng HO, Kiessling B, Lindberg B, McKay J (1962) Acta Chem Scand 16:615; (1963) 17:792
46. Wallenfels K, Keilich G, Bechtler G, Freudenberger D (1965) Biochem Z 341: 433
47. Catley BJ, Whelan WJ (1971) Arch Biochem Biophys 143:138
48. French D, Rundle RE (1942) J Am Chem Soc 64:1651; Freudenberg K, Cramer F (1948) Z Naturforsch 36:464; French D, Knapp DV, Pazur JH (1950) J Am Chem Soc 72:5150
49. Penninga D, Strokopytov B, Bozeboom HJ, Lawson CL, Dijkstra BW, de Vries GE, Bergsma J, Dijkhuizen L (1995) Biochemistry 34:3368; Mori S, Hirose, S, Oya T, Kitahata N (1994) Biosci Biotech Biochem 58:1968
50. Pulley AO, French D (1961) Biochem Biophys Res Commun 5:11; Endo T, Ueda H, Kobayashi S, Nagai T (1995) Carbohydr Res 269:369; Endo T, Nagase H, Ueda H, Kobayashi S, Nagai T (1997) Chem Pharm Bull 45:532

51. Szejtli J (1976) Stärke 29: 26; (1978) 30: 427; (1981) 33:387
52. Oguma T, Horiuchi T, Kobayashi M (1993) Biosci Biotech Biochem 57:1225
53. Côté G, Biely P (1994) Eur J Biochem 226:641
54. Aga H, Nishimoto T, Kuniyoshi M, Maruta K, Yamashita H, Higashiyama T, Nakada T, Kubota M, Fukuda S, Krimoto M, Tsujisaka Y (2003) J Biosci Bioeng 95:215
55. Kim Y-K, Kitaoka M, Hayashi K, Kim C-H, Côté GL (2003) Carbohydr Res 338:2213
56. Takaha T, Yanase M, Takata H, Okada S, Smith SM (1996) J Biol Chem 271:2902
57. Takata H, Takaha T, Okada S, Takagi M, Imanka T (1996) J Bacteriol 178:1600; Gorin PAF, Spencer JFT, Westlake DWS (1961) Can J Chem 39:1067; Zevenhuizen LPTM, Scholten-Koerselman HJ (1979) Antonie Leewenhoek 45:165; York WS, McNeil M, Darvill AG, Albersheim P (1980) J Bacteriol 142:243; Da Castro JM, Bruneteau M, Mutaftshiev S, Truchet G, Michel G (1983) FEMS Microbiol Lett 18:269
58. McIntire FC, Peterson WH, Riker AJ (1942) J Biol Chem 143:491
59. Dylan T, Helinski DR, Ditta GS (1990) J Bacteriol 172:1400; Cangelosi GA, Hung L, Pvanesarajah V, Stacey G, Ozga DA, Leigh JA, Nester EW (1987) J Bacteriol 169:2086
60. Miller KJ, Gore RS, Johnson R, Benesi AJ, Reinhold VN (1990) J Bacteriol 172:136; Rolin DB, Pfeffer PE, Osman SF, Szergold BS, Kappler F, Benesi AJ (1992) Biochim Biophys Acta 1116:215
61. Pfeffer PE, Osman SF, Hotchkiss A, Bhagwat AA, Keister DL, Valentine KM (1996) Carbohydr Res 296:23
62. O'Sullivan AC (1997) Cellulose 4:173
63. Marchessault RH, Sundararajan PR (1983) In: Aspinall GO (ed) The Polysaccharides, vol. 2. Academic Press, San Diego, pp 25–65
64. Hess K, Mahl G, Gütter E (1957) Kolloid Z 155:1
65. Ross P, Mayer R, Benziman M (1991) Microbiol Rev 55:35
66. Hus DS, Reeves RE (1967) Carbohydr Res 5:202
67. Aspinall GO, Krishnamurthy TN, Rosell K-G (1977) Carbohydr Res 55:11
68. Aspinall GO (1959) Adv Carbohydr Chem 14:429
69. Timell TE (1964) Adv Carbohydr Chem 19:247; (1965) 20:409
70. Aspinall GO, Molloy A, Craig JWT (1969) Can J Biochem 47:1063; Gould SEB, Rees DA, Wright NJ (1971) Biochem J 124:47
71. Bauer WD, Talmadge KW, Keestra K, Albersheim P (1973) Plant Physiol 51:174
72. Timell TE (1964) Adv Carbohydr Chem 19:247; (1965) 20:409
73. Stephen AM (1983) Other Plant Polysaccharides. In: Aspinall GO (ed) The Polysaccharides, vol. 2. Academic Press, San Diego, pp 166–169
74. Saka S (1990) Xyloglucans and Xylans. In: Hon DN-S, Shiraishi N (eds) Wood Cellulosic Chemistry. Marcel Dekker, New York, pp 59–88
75. Rolin C (1993) Pectin. In: Whistler RL, BeMiller JN (eds) Industrial Gums: Polysaccharides and Their Derivatives, 3rd edn. Academic Press, San Diego, pp 257–282
76. Karrer P, Francois G (1929) Helv Chem Acta 12:986
77. Ward K Jr, Seib PA (1970) Chitin. In: Pigman W, Horton D (eds) The Carbohydrates, vol. IIA. Academic Press, New York, pp 435–437
78. Rudall KM (1963) Adv Insect Physiol 1:257
79. Horton D, Lineback DR (1965) Methods Carbohydr Chem 5:403
80. Ueno H, Mori T, Fujinaga T (2001) Adv Drug Deliv Rev 52:105
81. Sanford PA (1989) Applications of Chitosan. In: Skjåk-Bræk G, Anthonsen T, Sanford PA (eds) Chitin and Chitosan. Elsevier Applied Science, London, pp 51–69
82. Vorlop KD, Klein J (1981) Biotech Lett 3:9
83. Jeon Y-J, Shahid F, Kim S-K (2000) Food Revs Int 16:159
84. Vandenberg GW, De La Noue J (2001) J Microencap 18:433
85. Jeanloz RW, Sharon N, Flowers HM (1963) Biochem Biophys Res Commun 13:20
86. Tipper DJ, Ghuysen J-M, Strominger JL (1965) Biochemistry 4:468
87. Sharon N, Osawa T, Flowers HM, Jeanloz RW (1966) J Biol Chem 242:223
88. Tipper DJ, Strominger JL (1966) Biochem Biophys Res Commun 22:48
89. Krulwich TA, Ensign JC, Tipper DJ, Strominger JL (1967) J Bacteriol 94:734
90. Ghuysen J-M (1968) Bacteriol Rev 32:425
91. Weidel W, Pelzer H (1964) Adv Enzymol 26:193
92. Fransson L-Å (1985) Mammalian Polysaccharides. In: Aspinall GO (ed) The Polysaccharides, vol. 3. Academic Press, San Diego, pp 338–386
93. Casu B (1985) Adv Carbohydr Chem Biochem 43:51

94. Painter TJ (1983) Algal Polysaccharides. In: Aspinall GO (ed) The Polysaccharides, vol. 2. Academic Press, San Diego, pp 263–264
95. Clare K (1993) Algin. In: Whistler RL, BeMiller JN (eds) Industrial Gums, Polysaccharides, and Their Derivatives, 3rd edn. Academic Press, San Diego, pp 108–118
96. Haug A, Larsen B (1962) Acta Chem Scand 16:1908
97. McNeely WH, Pettit DJ (1973) Algin. In: Whistler RJ, BeMiller JN (eds) Industrial Gums 2nd edn. Academic Press, San Diego, pp 74–75
98. Skjåk-Bræk G, Smidsrød O, Larsen B (1986) Int J Biol Macromol 8:330
99. Scott CD (1987) Enzyme Microb Technol 9:66
100. Linker A, Jones RS (1966) J Biol Chem 241:3845; Evans LR, Linker A (1973) J Bacteriol 116:915; Gorin PAJ, Spencer JFT (1966) Can J Chem 44:993; Pindar DF, Bucke CC (1975) Biochem J 152:617
101. Anderson DMW, Gill MCL, Jeffrey AM, McDougal FJ (1985) Phytochemistry 24:71
102. Aspinall GO, Hirst EL, Wickstrom A (1955) J Chem Soc 1160; Aspinall GO, Auret BJ, Hirst EL (1958) J Chem Soc 4408
103. Aspinall GO, Baille J (1963) J Chem Soc 1702
104. Sanford PA (1979) Adv Carbohydr Chem Biochem 36:266; Robyt JF (1998) Essentials of Carbohydrate Chemistry. Springer, Berlin, Heidelberg, New York, pp 193–218
105. Sloneker JH, Jeanes A (1962) Can J Chem 40:2066
106. Sloneker JH, Orentas DG (1962) Can J Chem 40:2188
107. Sloneker JH, Orentas DG, Jeanes A (1964) Can J Chem 42:1261
108. Lindberg B, Lorngren J, Thompson JF (1973) Carbohydr Res 28:351
109. Melton LD, Mindt L, Rees DA, Serson GR (1976) Carbohydr Res 46:245
110. Kang KS, Veeder GT, Mirrasoul PJ, Kaneko T, Cottrell IW (1982) Appl Environ Microbiol 43:1086
111. Jansson PE, Lindberg B, Widmalm G, Sanford PA (1984) Carbohydr Res 139:217; Kuo M-S, Mort AJ, Dell A (1986) Carbohydr Res 156:173; Jansson P-E, Lindberg B, Lindberg J, Maekawa E, Sanford PA (1986) Carbohydr Res 156:157
112. Lüderitz O, Staub AM, Westphal O (1966) Bacteriol Rev 30:193; Robbins PW, Uchida T (1962) Biochemistry 1:323
113. Bagdian G, Lüderitz O, Staub AM (1966) Ann NY Acad Sci 133:849
114. How MJ, Brimacombe JS, Stacey M (1964) Adv Carbohydr Chem 19:303; Larm O, Lindberg B (1976) Adv Carbohydr Chem Biochem 33:295; Lindberg B (1990) Adv Carbohydr Chem Biochem 48:279
115. McGuire EJ, Binkley SB (1964) Biochemistry 3:247
116. Bhattacharjee AK, Jennings HJ, Kenny CP, Martin A, Smith ICP (1975) J Biol Chem 250:1926
117. Egan W, Lui T-Y, Dorow D, Cohen JS, Robbins JD, Gotschlich EC, Robbins JB (1977) Biochemistry 16:3687
118. French AD, Waterhouse AL (1993) Structural Chemistry of Inulin. In: Suzuki M, Chatterton NJ (eds) Science Technology of Fructans. CRC Press, Boca Raton, FL, pp 41–82
119. Hendry GAF, Wallace RK (1993) Occurrence of Inulins and Levans. In: Suzuki M, Chatterton NJ (eds) Science Technology of Fructans. CRC Press, Boca Raton, FL, pp 119–140
120. Baird JK, Longyear VMC, Ellwood DC (1973) Microbios 8:143; Rossel K-G, Birkhead D (1974) Acta Chem Scand Ser B 28:589; Ebisu S, Kato K, Kotani S, Misaki A (1975) J Biochem (Tokyo) 78:879; Corrigan AJ, Robyt JF (1979) Infect Immun 26:387
121. Avigad G (1968) Bacterial Levans. In: Mark HF, Gaylard NG, Bikales NM (eds) Encyclopedia of Polymer Science Engineering, vol. 8. Wiley Interscience, New York, pp 71–78
122. Garzozynski SM, Edwards JR (1973) Arch Oral Biol 18:239; Eshrlich J, Stivala SS, Bahary WS, Garg SK, Long LW, Newbrun E (1975) J Dent Res 54:290; Marshall K, Weigel H (1976) Carbohydr Res 49:351
123. Boussingault J (1972) Compt Rend 74:939
124. Asahina Y, Shimoda H (1930) J Pharm Soc Japan 50:1; Haas P, Hill TG (1932) Biochem J 26:987
125. Strain HH (1937) J Am Chem Soc 56:2264
126. Proust M (1806) Ann Chim Phys 57:144
127. Jrier E (1893) Compt Rend 117:498
128. Bidwell RGS (1958) Can J Botany 36:337
129. Archibald AR, Baddiley J (1966) Adv Carbohydr Chem 21:354
130. Frèryacque M (1939) Compt Rend 208:1123
131. Lindberg B, Misiorny A, Wachtmeister CA (1953) Acta Chem Scand 7:591; Aghoramarthy K, Sarma KG, Seshadri TR (1961) Tetrahedron 12:173
132. Richtmyer NK (1970) Carbohydr Res 12:135
133. Anderson FB, Harris G (1963) J Gen Microbiol 33:137; Hajny GJ (1964) Appl Microbiol 12:87

134. Armstrong JJ, Baddiley J, Buchanan JG, Carss B, Greenberg GR (1958) J Chem Soc 4344
135. Archibald AR, Baddiley J (1966) Adv Carbohydr Chem 21:323
136. Archibald AR, Baddiley J, Burton D (1968) Biochem J 110:543; Archibald AR, Baddiley J, Heckels JE, Heptinstall S (1971) Biochem J 125:353
137. Truscheit E, Frommer W, Junge B, Muller L, Schmidt DD, Wingender W (1981) Angew Chem Int Ed Engl 20:744
138. Lee S-B, Park KH, Robyt JF (2001) Carbohydr Res 33:13
139. Yoon S-H, Robyt JF (2003) Carbohydr Res 338:1969
140. Courtois JÉ, Percheron F (1970) Phenanthrene Glycosides. In: Pigman W, Horton D (eds) The Carbohydrates, vol. IIA. Academic Press, New York, pp 216, 221–222
141. Montreuil J (1980) Adv Carbohydr Chem Biochem 37:158
142. Montreuil J (1975) Pure Appl Chem 42:431
143. Fournet B, Strecker G, Montreuil J, Dorl L, Haverkamp J, Vliegenthart JFG, Schmid K, Binette JP (1978) Biochemistry 17:5206
144. Nilsson B, Nordén NE, Svensson S (1979) J Biol Chem 254:4545
145. Tai T, Yamashita K, Setsuko I, Kobata A (1977) J Biol Chem 252:6687
146. Yamashita K, Tachibana Y, Kobata A (1978) J Biol Chem 253:3862
147. Feeney RE, Yeh Y (1978) Adv Prot Chem 32:191
148. Watkins WM (1966) Science 152:172
149. Lloyd KO, Kabat EA (1968) Proc Natl Acad Sci US 61:1470
150. Ginsburg V (1972) Adv Enzymol 36:131
151. Whistler RL, BeMiller JN (1962) Methods Carbohydr Chem 1:88
152. Robyt JF (2000) Thin-layer Chromatography of Carbohydrates. In: Wilson ID, Cooke M, Poole CF (eds) Encyclopedia of Separation Science, vol. 5. Academic Press, San Francisco, pp 2235–2244
153. Robyt JF, Mukerjea R (1994) Carbohydr Res 251:187
154. Han NS, Robyt JF (1998) Carbohydr Res 313:135
155. Robyt JF, White BJ (1987) Biochemical Techniques: Theory and Practice. Waveland Press, Prospect Heights, IL, pp 82–86
156. Miller GL, Dean J, Blum R (1960) Arch Biochem Biophys 91:21
157. Whelan WJ (1962) Methods Carbohydr Chem 1:321
158. French D, Robyt JF, Weintraub M, Knock P (1966) J Chromatog 24:68
159. Havlicek J, Samuelson O (1972) Carbohydr Res 22:307
160. Smith FD, Zopf DA, Ginsburg V (1978) Anal Biochem 85:602
161. Kainuma K, Nakakuki T, Ogawa T, (1981) J Chromatog 212:126
162. Ammeraal RN, Delgado GA, Tenbarge FL, Friedman RB (1991) Carbohydr Res 215:179
163. Kerns RJ, Vlahov IR, Lindhardt RJ (1995) Carbohydr Res 267:143
164. O'Shea MG, Sammuel MS, Konik CM, Morrell MK (1998) Carbohydr Res 307:1
165. Mukerjea Ru, Robyt JF (2003) Carbohydr Res 338:1811
166. Yoon S-H, Robyt JF (2002) Carbohydr Res 337:2245
167. Wilham CA, Alexander BH, Jeanes A (1955) Arch Biochem Biophys 59:61
168. Cramer F (1958) Chem Ber 91:308
169. French D (1957) Adv Carbohydr Chem 12:189
170. Thoma J, Stewart L (1965) Cycloamyloses. In Whistler RJ, Paschall EF, BeMiller JN, Roberts HJ (eds) Starch: Chemistry and Technology, vol. I. Academic Press, New York, pp 209–249
171. Anderson B, Seno N, Sampson P, Reilly JG, Hoffman P, Meyer K (1964) J Biol Chem 239:2716
172. Spiro RG, Bhoyroo VD (1971) Fed Proc Am Soc Exp Biol 30:1223
173. Spiro RG (1972) Methods Enzymol 28:35
174. Takasaki S, Mizuochi T, Kobata A (1982) Methods Enzymol 83:263
175. Muramatsu T (1978) Methods Enzymol 50:555
176. Kobata A (1978) Methods Enzymol 50:560
177. Tarentino AL, Trimble RB, Maley F (1978) Methods Enzymol 50:574
178. Sojar HT, Bahl OP (1987) Methods Enzymol 138:341

Part 2
General Synthetic Methods

2.1 Reactions at Oxygen Atoms

Ana M. Gómez
Instituto de Química Orgánica General, (CSIC),
28006 Madrid, Spain
iqog106@iqog.csic.es

In: *Glycoscience*. Fraser-Reid B, Tatsuta K, Thiem J (eds)
Chapter-DOI 10-1007/978-3-540-30429-6_3:

Abstract

Synthetic protocols based on carbohydrates require the differentiation of their abundant hydroxyl groups, by and large, in order to expose just one single hydroxyl group to the selected reagent. This differentiation is usually carried out with the assistance of protecting groups that block the rest of the hydroxyl groups while being compatible with the given reaction conditions. By corollary, the knowledge and apt choice of the appropriate protecting groups is a key factor in successful synthetic endeavors. In this chapter, an overview of the most commonly employed protecting groups in carbohydrate chemistry is given. Alkyl ethers, being robust protecting groups, have a long history in synthetic carbohydrate chemistry and in related structural studies of polysaccharides. Acetals and ketals, which are of fundamental importance in carbohydrate chemistry, are then discussed. Acyl and silyl protecting groups, which also play an important role in modern monosaccharide transformations, are also presented. Finally, recent blocking strategies are described, including orthogonal strategies, by which the protecting groups are harmoniously combined in modern carbohydrate chemistry.

Keywords

Protecting groups; Alkylation; Acetalation; Acylation; Carbonylation; Silylation; Phosphorylation

Abbreviations

ADMB	4-acetoxy-2,2-dimethylbutanoyl
All	allyl
Alloc or Aloc	allyloxycarbonyl
APAC	2-(allyloxy) phenyl acetyl
BDA	butane 2,3-diacetals
Bn	benzyl
Bocdene	2-(*tert*-butoxycarbonyl)-ethylidene
BOM	benzyloxymethyl ether
Bz	benzoyl
BzOBT	1-*N*-benzyloxy-1,2,3-benzotriazol

Cac or ClAc	chloroacetyl
CAN	ceric ammonium nitrate
CBz	benzyloxycarbonyl
CDA	cyclo-hexane-1,2-diacetal
CSA	camphorsulfonic acid
CCL	*Candida cylindracea* lipase
CVL	*Chromobacterium viscosum* lipase
DABCO	diazabicyclo[2.2.2]octane
DBMP	ditertbutylmethylpyridine
DBU	1,8-diazabicyclo[5,4,0]undec-7-ene
DCC	dicyclohexylcarbodiimide
DDQ	dichlorodicyanoquinone
DIB	(diacetoxyiodo)benzene
DISAL	3,5-dinitrosalicylate
Dispoke	dispiroketal
DMAP	4-(dimethylamino)pyridine
DME	dimethoxyethane
DMF	*N*,*N*-dimethylformamide
DMSO	dimethylsulfoxide
DMTr	dimethoxytriphenylmethyl
DMTST	dimethyl(methylthio)sulfonium trifluoromethane sulfonate
DTBMP	di-*tert*-butylmethylpyridine
DTBS	ditertbutylsilylene
EDAC	1-[3-(dimethylamino)propyl]-3-ethylcarbodiimide hydrochloride
Fmoc	fluoren-9-ylmethoxycarbonyl
HATU	2-(7-aza-1*H*-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate
IDCP	iodonium di-*sym*-collidine perchlorate
Lev	levulinoyl
MEM	methoxyethoxymethyl
MMTr	monomethoxytriphenylmethyl
Mocdene	2-(methoxycarbonyl)-ethylidene
MOM	methoxymethyl ether
NAP	2-naphthylmethyl
NBS	*N*-bromosuccinimide
NIS	*N*-iodosuccinimide
NMO	*N*-methylmorpholine-*N*-oxide
NMR	nuclear magnetic resonance
NPM	*p*-nitrophenylmethyl
PAB	pivaloylaminobenzyl
PBB	*p*-bromobenzyl
PCB	*p*-chlorobenzyl
PFL	*Pseudomonas fluorescens* lipase
Piv	pivaloyl
PIB	*p*-iodobenzyl

PLE	pig liver esterase
PMB	*p*-methoxybenzyl
PMBM	*p*-methoxybenzyloxymethyl
PN	protease *N*-neutral protease
Poc	propargyloxycarbonyl
PPL	lipase from porcine pancreas
PPTS	pyridinium *p*-toluenesulfonate
PSE	phenylsulfonylethylidene
RJL	*Rhizopus javanicus* lipase
SEE	1-[2-(trimethylsilyl)ethoxy]ethyl
SEM	trimethylsilylethoxymethyl ether
SET	single electron transfer
TBS or TBDMS	*tert*-butyldimethylsilyl
TBDPS	*tert*-butyldiphenylsilyl
TES	triethylsilyl
TFA	trifluoroacetic acid
TFAA	trifluoroacetic anhydride
THF	tetrahydrofurane
TIBAL	triisobutylaluminum
TIPDS	1,1,3,3-tetraisopropyldisiloxane
TIPS	trisopropylsilyl
TMEDA	tetramethylethylenediamine
TMS	trimethylsilyl
TMTr	trimethoxytriphenylmethyl
TPS	triphenylsilyl
Tr	trityl
Troc	2,2,2-trichloroethyloxycarbonyl
TBAF	tetrabutylammonium fluoride

1 Introduction

This chapter describes the chemical reactions at the oxygen atoms of carbohydrates along with some of their fundamental characteristics. The hydroxyl groups of carbohydrates display all the chemical properties associated with simple alcohols. The only difference is that carbohydrates contain many hydroxy groups with similar chemical character. Since the hydroxy groups in carbohydrates play different biological roles depending on their positions, the ability to perform chemical reactions on a particular hydroxy group is highly important. However, the regioselective transformation of one out of several hydroxy groups is far from being trivial. While the differentiation between the primary *versus* the secondary hydroxy groups is in general not too difficult, the discrimination between secondary hydroxy groups is a difficult task.

Usually, partially substituted derivatives are made with the aid of protecting groups. The protecting groups used in carbohydrate chemistry are the same as in any other area in organ-

ic chemistry [1,2,3,4]. In addition to this fact, it is important to point out that in carbohydrate derivatives protecting groups do more than protect; they also confer other effects to the molecule and can alter the course of a reaction. Important examples of such effects are the use of 2-*O*-participating groups in glycosyl donors [5] or the armed/disarmed concept for glycoside coupling [6].

This chapter aims to impart general synthetic strategies for most sugars and oligosaccharide structures through the use of some basic, well-proven protecting groups, coupled with general strategies towards regioselectivity. The discussion outlines frequently used protecting groups in carbohydrate chemistry, briefly surveying conditions for their introduction, stability, and removal. It should be noted at this stage that the hydroxyl group of the anomeric center, is unique in having two attached oxygen atoms and therefore it will be treated in a separated section.

2 Reactions at Non-Anomeric Hydroxyl Groups

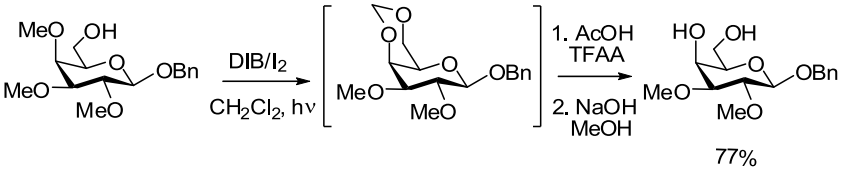
2.1 Alkylation Reactions: Ether-Type Protecting Groups

Alkyl and aryl ethers are relatively stable to acids and bases due to the high C–O bond energy and it is difficult to recover the parent alcohols from them; therefore, most useful ether-type protections utilize resonance stabilization (by delocalization) of the benzylic-type cation or radical to facilitate the cleavage.

2.1.1 Methyl Ethers

Conversion to methyl ethers of non-anomeric hydroxyl groups is a long-established procedure used, in conjunction with ethylation and deutero-methylation, for the analysis of glycosides, oligosaccharides, and polysaccharides.

Methyl ethers are not normally regarded as protecting groups (though they may be considered in special cases [7]) because the removal is difficult requiring conditions not compatible with other functional groups. A recent study has demonstrated a wide range of susceptibilities to methylation of the hydroxyls in various methyl pyranosides using diazomethane together with transition-metal chlorides and boric acid [8]. On the other hand, the selective removal of an

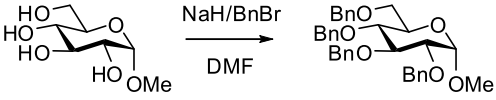


Scheme 1
Selective removal of methoxy protecting groups

ether adjacent to a hydroxyl group in carbohydrate substrates was accomplished with (diacetoxyiodo)benzene (DIB) and I_2 under irradiative conditions (Scheme 1). In this step, the methoxy protecting group was transformed into a mixture of acetals (methylenedioxy acetal or *O*-methyl acetate) which upon basic hydrolysis provides the diol [9,10].

2.1.2 Benzyl (Bn) Ethers

The classical permanent protecting group of carbohydrate hydroxyl functions is probably the benzyl ether. It is very stable and can be readily removed under essentially neutral conditions. For this reason, numerous benzylation and *O*-debenzylation procedures have been described. Benzyl ether formation is usually achieved by the reaction of alcohols and benzyl halides in the presence of a base such as sodium hydride in anhydrous DMF (Scheme 2) [11], or a mild base (Ag_2O) in THF using a phase-transfer catalyst [12]. Benzylation can also be accomplished by the use of an acidic catalyst with benzyltrichloroacetimidate as the reagent [13]. A method using the reductive etherification of TMS ethers under non-basic conditions has also been reported [14].



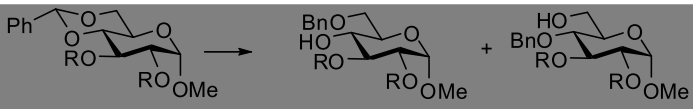
Scheme 2
Benzylation of methyl α-D-glucopyranoside

Benzyl ethers are highly stable to a wide range of reagents but are readily removed through catalytic reductive conditions [15]. Hydrogenolysis is commonly carried out using hydrogen gas with a palladium catalyst absorbed on charcoal although modifications involving hydrogen transfer have been used. A variety of alternative strategies include Na/liquid ammonia [16], anhydrous $FeCl_3$ [17,18], and $CrCl_2$/LiI [19].

Selective Benzylation Selective benzylation of carbohydrate hydroxyl functions by direct one-step protection is difficult to achieve. Therefore, several techniques for the selective protection have been developed over the years and the most common are discussed below.

Reductive Opening of Benzylidene Acetals. An attractive approach for the selective introduction of benzyl groups is provided by the regioselective opening of *O*-benzylidene acetals [20,21,22]. Generally, one of the two C–O bonds in benzylidene acetals can be selectively cleaved, and the direction of the cleavage is dependent on steric and electronic factors as well as, on the nature of the cleavage reagent. Reductive ring-opening of the 1,3-dioxane ring of 4,6-*O*-benzylidene-α-D-glucopyranosides gives the 6-*O*-benzyl and 4-*O*-benzyl ethers respectively in different ratios, depending on the combination of the reagent Lewis acid, solvent, and the substituent at C-3. Some examples are shown in Table 1.
4,6-*O*-Benzylidene-D-galactopyranosides behave in a similar manner to the D-gluco analogs in most cases [29,30]. In the ring opening of the dioxolane rings of 2,3-*O*-benzylidene-α-D-man-

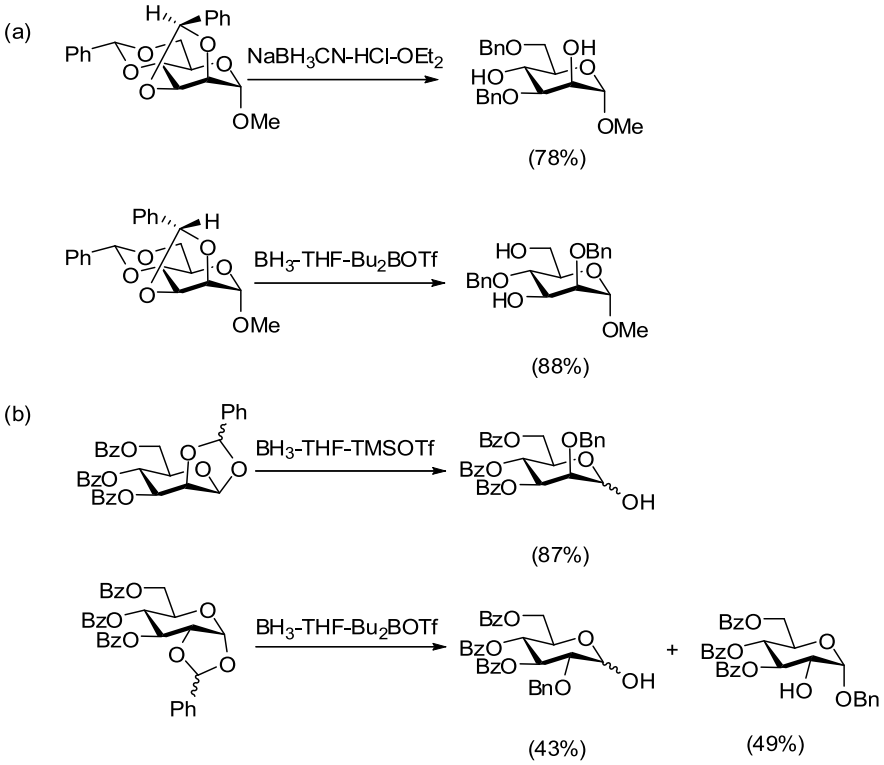
Table 1
Reductive opening of 4,6-*O*-benzylidene-α-D-glucopyranoside



Entry	Reagent	Solvent and temperature	R	Yield (%) 6-*O*-Bn	Yield (%) 4-*O*-Bn	References
1	$NaBH_3CN/HCl$	THF, 0 °C	Bn	81		[23]
2		THF, 0 °C	Bn	95		[23]
3		THF, rt	Bn (1-*O*-Allyl)	80		[24]
4		THF, rt	Bn (1-*O*-Allyl)	79	16	[25]
5	$Me_3N{\cdot}BH_3/AlCl_3$	THF, rt	Bn	71		[26]
6		THF, rt	Bz	74		[26]
7		Toluene, rt	Bn		50	[26]
8		Toluene, rt	Bz		40	[26]
9	$Me_3N{\cdot}BH_3/BF_3{\cdot}Et_2O$	MeCN, 0 °C	Bn	30	55	[27]
10		CH_2Cl_2, 0 °C	Bn	3	73	[27]
11	Et_3SiH/CF_3CO_2H	CH_2Cl_2	Bn	81	55	[28]
12		CH_2Cl_2	Ac	98		[28]
13	$BH_3{\cdot}THF/Bu_2BOTf$	CH_2Cl_2, 0 °C	Bn		87	[30]
14	$BH_3{\cdot}THF/Cu(OTf)_2$	CH_2Cl_2, rt	Bn		94	[31]
15	$Me_2EtSiH/Cu(OTf)_2$	CH_3CN, 0 °C	Bn	84		[31]

noside derivatives, the directions of the reaction are determined by the configuration of the benzylidene carbon (*Scheme 3a*). Regarding the reductive ring-opening of 1,2-*O*-benzylidene derivatives, in the case of *manno*-type derivative only a C–O1 bond was cleaved, whereas both the C–O1 and C–O2 bonds were cleaved in the case of the gluco-type compound (*Scheme 3b*) [32].
Benzylidene acetals can also be opened under oxidative conditions, typically NBS in CCl_4, to give benzoyl ester protected halogen derivatives, thereby providing an entry into deoxycarbohydrate compounds [33].

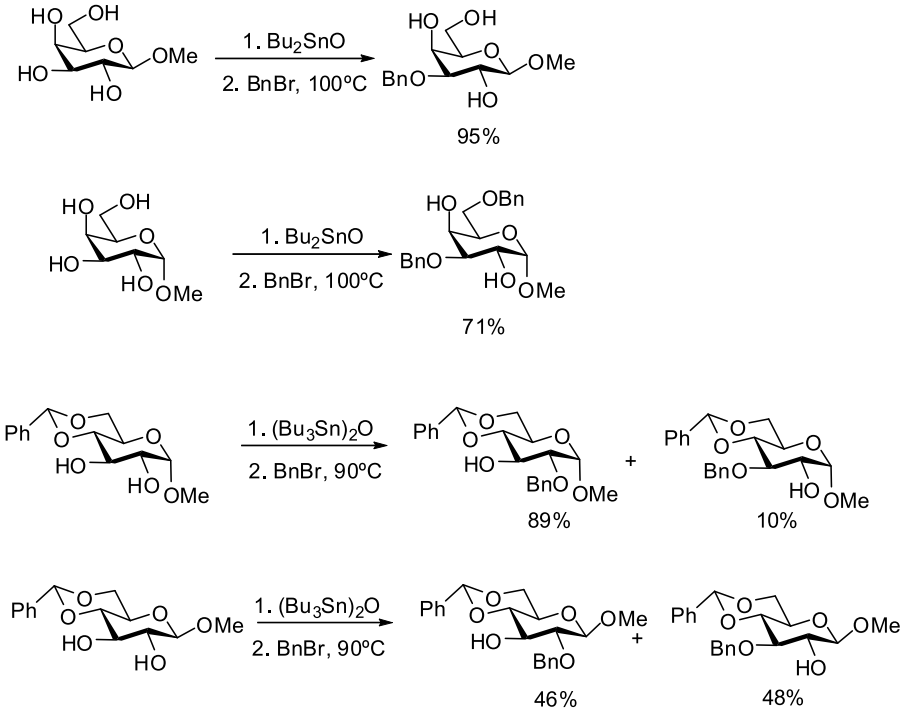
Formation of Organotin Intermediates. Another method for selective benzylation refers to the activation of the hydroxyl groups of saccharides by the formation of organotin intermediates such as trialkylstannyl ethers or dialkylstannylene acetals [34]. When the substrate is treated with the tin reagent, one or two Sn–O bonds are formed, enhancing the nucleophilicity of the oxygen atom in the stannyl ether or stannylene acetal. This effect is not identical for the two oxygen atoms of a Sn-acetal, resulting in a differential increase of their nucleophilicity and an ensuing higher regioselectivity.
The activation is carried out by reaction of the polyol with bis(trialkyltin) oxide or a dialkyltin oxide with heating and can be performed in various solvents, the most common being methanol



Scheme 3
Regioselective benzylation by reductive ring opening of benzylidene acetals

or toluene. Subsequent treatment of the preformed tin intermediates dissolved in a polar aprotic solvent with alkyl halides in the presence of added nucleophiles such as tetrabutylammonium halides [35] or CsF [36] yields the corresponding alkyl (benzyl) ethers. Regarding regioselectivity, this is much the same irrespective of which type of tin derivative is used, the primary hydroxyl group and equatorial hydroxyl group in a vicinal *cis*-dioxygen configuration are preferentially benzylated. As exemplified in *Scheme 4*, this rule is generally correct, but the degree of selectivity is also dependent on structural features and other factors such as the presence of additives.

Regioselective de-*O*-benzylation. An alternative strategy to partially benzylated carbohydrates has been accomplished by selective de-*O*-benzylation of easily available polybenzylated precursors. This has been achieved in limited cases by catalytic hydrogenolysis [37], catalytic hydrogen-transfer cleavage [38], acetolysis [39], hypoiodite fragmentation [40], iodine-mediated addition-elimination sequences [41], or use of Lewis acids [42].
Recently, isobutylalanes [43,44,45,46] or the combination $CrCl_2$/LiI [47] have been shown as efficient agents for the selective deprotection of poly-benzylated carbohydrates. The reaction with isobutylalanes is assumed to proceed through the formation of a penta-coordinated com-



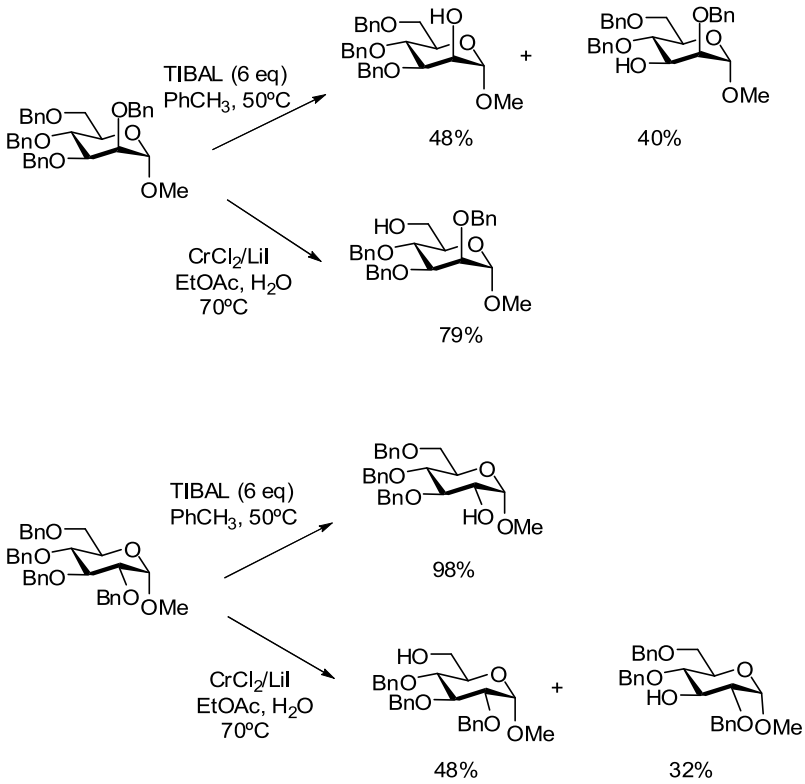
Scheme 4
Examples of stannyl mediated regioselective benzylation

plex between the aluminum reagent and a 1,2-*cis* oxygen pattern of the sugar. A second aluminum atom then selects the less hindered oxygen atom and directs the de-*O*-benzylation [46]. When one oxygen atom is clearly more accessible than the other the reaction is highly regioselective. In contrast, when the system $CrCl_2$/LiI is used, a three-point coordination model of the carbohydrate with Cr(II) or Cr(III) is needed for optimal selectivity [47] (*Scheme 5*).

2.1.3 Substituted Benzyl Ethers

To increase the scope of available hydroxyl protecting groups substituted benzyl ethers, which can be selectively removed in the presence of unsubstituted benzyl ethers have been developed. These substituted benzyl ethers are generally less stable to different reaction conditions than unsubstituted benzyl ethers and therefore are used as temporary protecting groups.

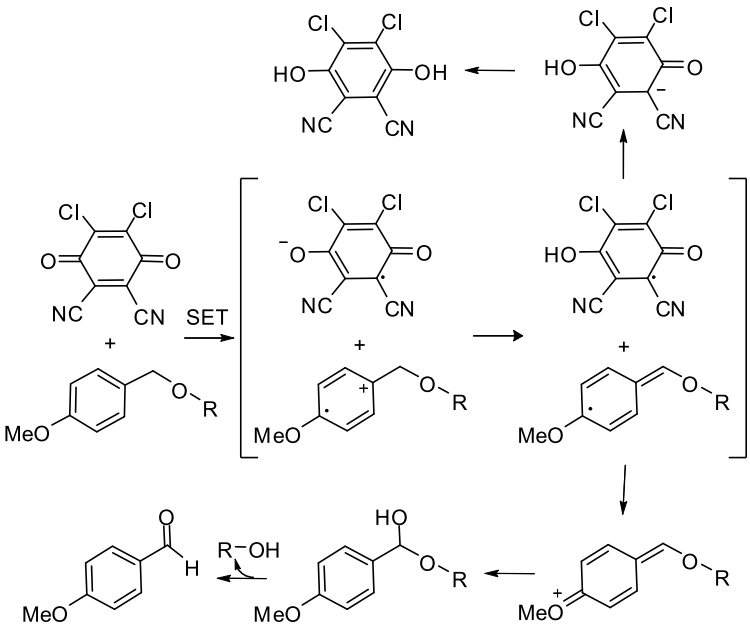
p-Methoxy Benzyl (PMB) Ethers Of the several benzyl ether-type protecting groups reported, *p*-methoxy benzyl (PMB) enjoys a unique position in carbohydrate chemistry due to the ease of its introduction and removal. PMB group demasking, in general, is mediated either by oxidizing agents or by Lewis acids. Thus, hydroxyl moieties protected as PMB ethers can be regenerated easily by oxidation with DDQ [48,49], DDQ-$FeCl_3$ [50], DDQ-$Mn(OAc)_3$ [51], or CAN [52,53]. In the case of DDQ, the *p*-methoxybenzyl group is cleaved selectively with-



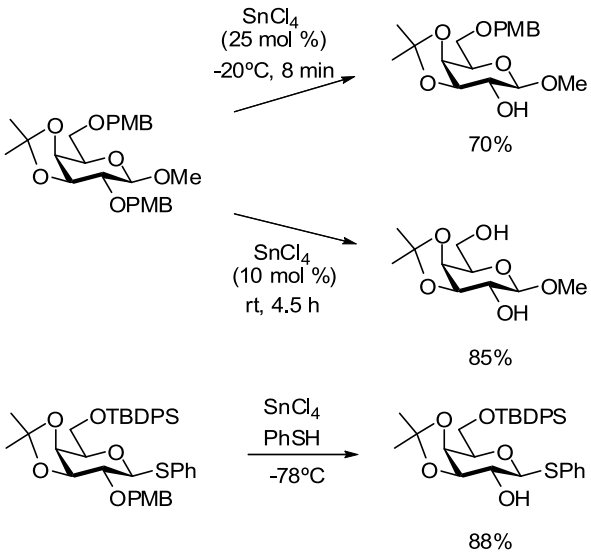
Scheme 5
Examples of regioselective de-*O*-benzylation

out affecting several other protecting groups, including benzyl ether. The reaction is assumed to proceed through an easy single electron transfer (SET) to DDQ to generate an oxonium ion which can be captured by water [48,49] (*Scheme 6*).

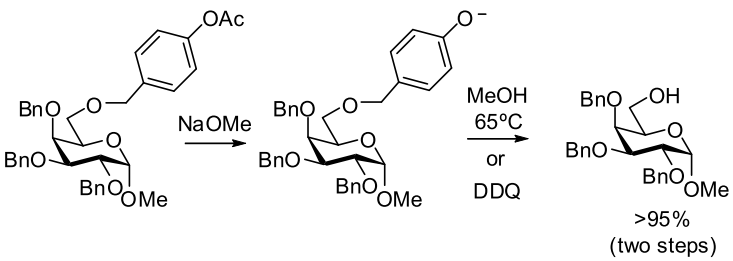
Deprotection of polyhydroxylated carbohydrate PMB ethers can also be accomplished with $SnCl_4$ [54]. The reaction is compatible with benzyl, TMS, isopropylidene acetal, or methoxy protecting groups. Sometimes the reaction results in unusual regioselectivity and partial deprotection is observed. Preferential mono or bis cleavage of PMB ethers was achieved with careful control of the reaction conditions. However, the general conditions fail in the case of thioglycosides, and a combination of $SnCl_4$/PhSH needs to be used [55] (*Scheme 7*). This combination is particularly useful in the cases where oxidative reagents such as DDQ or CAN need to be avoided. PMB ethers can also be cleaved with $ZrCl_4$ [56], $SnCl_2$/TMS-Cl/anisole [57], CF_3CO_2H in CH_2Cl_2 [58], $CeCl_3 \cdot 7H_2O$/NaI [59], or I_2/MeOH [60]. The PMB group has also been used as an in situ-removable protecting group for a reactive hydroxyl group in a one-pot reaction involving two sequential glycosilations. The deprotection was performed with *N*-iodosuccinimide-trifluoromethanesulfonic acid at 0 °C and the procedure was used in the synthesis of the globotetraose (Gb4) tetrasaccharide [61].



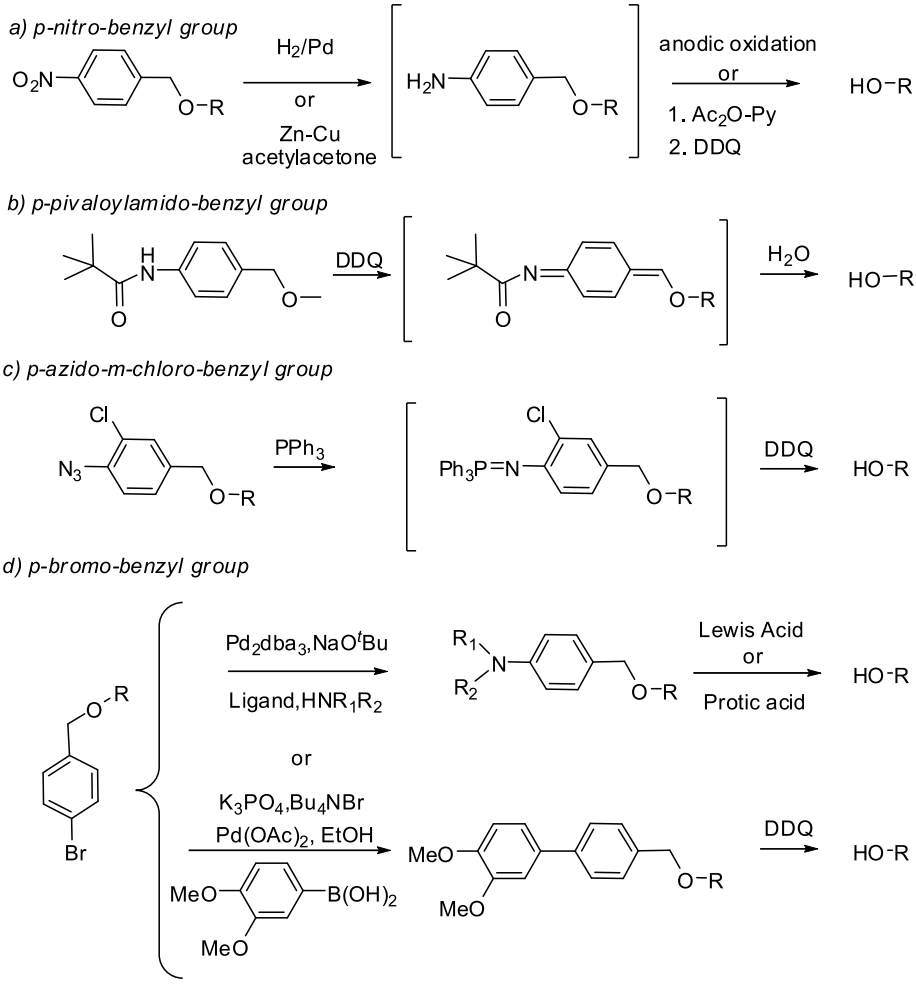
Scheme 6
Removal of *p*-methoxybenzyl ethers with DDQ



Scheme 7
Examples of deprotection of *p*-methoxybenzyl ethers with $SnCl_4$



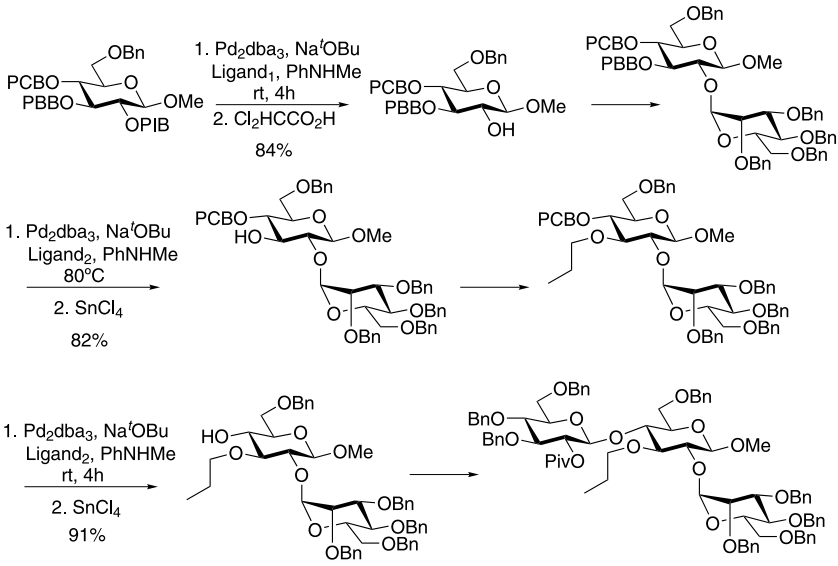
Scheme 8
Two-step deprotection of *p*-acetoxybenzyl protecting group



Scheme 9
Examples of *p*-substituted benzyl-type protecting groups

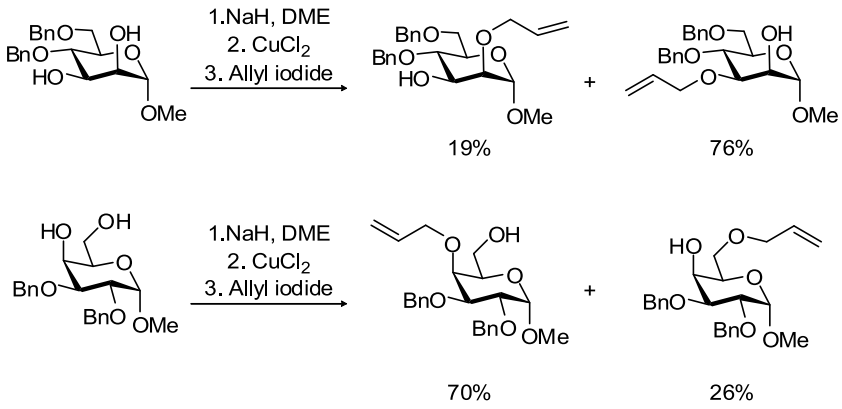
Other *p*-hydroxybenzyl-derived protecting groups include *p*-acetoxybenzyl ether and 2-(trimethylsilyl)ethoxymethoxybenzyl ether [62]. These groups require a two-stage deprotection strategy in which treatment with base or fluoride was followed by thermolysis or mild oxidation of the obtained *p*-phenoxybenzyl intermediate (*Scheme 8*). The conditions are compatible with many of the standard manipulations of oligosaccharide synthesis and with the presence of benzyl or PMB ethers.

Other p-Substituted Benzyl-Type Ethers Although the PMB protecting group has been extensively utilized in oligosaccharide synthesis, the acid sensitivity of this group sometimes restricts its synthetical application especially during glycosylations. Therefore, other *p*-substituted benzyl groups have been developed (*Scheme 9*). The *p*-nitrobenzyl (or *p*-nitrophenylmethyl NPM) group, which is acid-stable, is readily cleaved via a two-stage procedure involving reduction to an *p*-amino-benzyl ether followed by mild anodic oxidation [63]. *p*-Acetamidobenzyl and *p*-pivaloylaminobenzyl (PAB) derivatives are also used as protecting groups for hydroxyl groups [64,65]. These ethers are much more stable under acidic conditions than a PMB ether, can be obtained by direct alkylation of the hydroxyl group or by acylation of the corresponding *p*-aminobenzylether, and are deprotected by treatment with DDQ. The oxidation occurs at a rate comparable to PMB ethers, so that no preferential cleavage could be achieved with DDQ between these two groups. *p*-Azido benzyl groups are also useful as protecting groups of hydroxyl moieties [66,67]. They can be removed much



Ligand$_1$ = 1-(N,N-dimethylamino)-1´-(diciclohexylphosphino)biphenyl
Ligand$_2$ = (*o*-biphenyl)P(*t*Bu)$_2$

Scheme 10
Iterative deprotection of *p*-halobenzyl ether protecting groups



Scheme 11
Example of regioselective allylation from copper complexes

faster than the PMB group by DDQ oxidation after conversion of the azide group into the corresponding iminophosphorane. This group allowed for temporary protection of hydroxyl groups in solid-phase synthesis of oligosaccharides [68].

The chemically stable *p*-halobenzyl ethers (PIB = *p*-iodobenzyl, PBB = *p*-bromobenzyl; PCB = *p*-chlorobenzyl) are converted to labile arylamines via Pd-catalyzed amination [69]. Rapid deprotection of the amine benzyl ethers was observed under very mild Lewis acid conditions. Regarding compatibility of these novel protecting groups with others commonly used, selective cleavage was achieved in the presence of silyl ethers, PMB groups, and glycal double bonds. As shown in *Scheme 10*, the differences in the rates of reaction between aryl chlorides, bromides, and iodides in the Pd-catalyzed amination reactions allows for iterative deprotection. In a related method, the *p*-bromobenzyl group is converted to a DDQ-labile *p*-(3,4-dimethoxyphenyl)benzyl ether by a Suzuki–Miyaura coupling reaction. This protecting group played a key role in the access to the fully lipidated malarial GPI disaccharide [70].

2.1.4 Allyl and Related Ethers

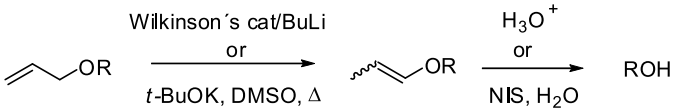
The protection of alcohols with allyl [71] and related (prenyl, methylallyl, cinnamyl, homoallyl) groups is of great importance in carbohydrate synthesis due to their stability under the conditions required for glycoside formation. These groups are moderately stable to acids and bases, and offer the potential for selective dealkylation of differentially protected sites.

The most general method of preparing allyl ethers is to react the alcohol with allyl bromide or iodide in the presence of sodium hydride. The reaction is best carried out in a polar solvent, usually DMF [72]. Alcohols may also be alkylated after conversion to their barium salts. This technique is employed in the case of *N*-acyl derivatives of aminosugars to avoid any risk of alkylation at nitrogen which would accompany the use of sodium hydride as the base [73,74]. Conversion of alcohols to allyl carbonates, followed by palladium-catalyzed extrusion of CO_2, constitutes a milder alternative to the classical Williamson-type pro-

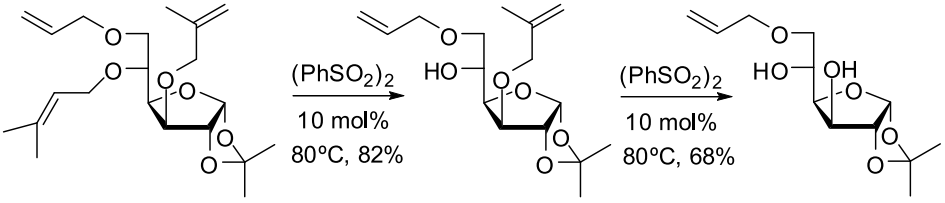
cedure [75,76]. Compounds containing base labile groups can also be allylated using allyl-trichloroacetimidate [77].

As in the case of benzylation, strategies used for selective allylation include prior conversion to organotin derivatives. When comparisons have been made, there does not seem to be significant differences in selectivities between benzylation and allylation [33,78]. Diols have also been selectively alkylated as their copper (II) salts. Under such conditions no disubstitution is observed and regioselectivity towards formation of a preferred monoallylated product (4,6-diols give mainly 4-substitution and 2,3-diols give mainly 3-substitution) is usually high (*Scheme 11*) [79].

Common allyl deprotection methods are two-stage procedures that include isomerization to the more labile 1-propenyl group with a variety of agents (*Scheme 12*). The most frequently employed conditions are treatment of the allyl ether with *t*-BuOK [80], Wilkinson catalyst [81], Pd/C [82], $PdCl_2$ [83], ruthenium(II) [84], and iridium(I) complexes [85] followed usually by acid hydrolysis or oxidation of the resulting enol ether. Also reported are methods including oxidative conditions such as DDQ [86], SeO_2 [87], NBS-*hv* [88], and OsO_4/NMO/$NaIO_4$ [89]. As a general rule substitution of the allylic framework either slows down or even inhibits the transition-metal catalyzed isomerization [71]. The crotyl and the prenyl groups are readily removed in DMSO/*t*-BuOK through γ-hydrogen elimination reactions. These processes are faster than the allyl to prop-1-enyl isomerization but the difference in rates does not appear to be sufficient to allow good selectivities [90]. $Yb(OTf)_3$ [91], I_2/CH_2Cl_2 [92], or DDQ [93] are mild and efficient methods to cleave prenyl ethers. Remarkable selectivity in the order methyl-prenyl > prenyl > mathallyl > allyl has been observed by using diphenyldisulfone in a sealed tube at 80 °C (*Scheme 13*) [94]. These reactions are initiated by the benzene-sulfonyl radical formed from thermal homolysis of $(PhSO_2)_2$. The reaction conditions are compatible with the presence of other protecting groups such as acetals and allyl, benzyl, or silyl ethers.



Scheme 12
Two-stage removal of allyl ether protecting groups



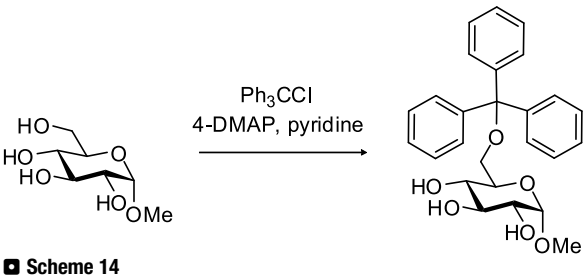
Scheme 13
Selective cleavage of branched allyl ethers

2.1.5 Trityl (Tr) Ethers

The usefulness of triphenylmethyl ethers as protecting groups in organic synthesis, in general, and in carbohydrate and nucleoside chemistry in particular is well documented. Its utility is attributed to the ease in preparing and removing them as well as to the high selectivity for primary positions observed in polyols. Tritylation of primary hydroxyl groups using trityl chloride in pyridine is one of the oldest selective alkylation processes described in carbohydrate chemistry (*Scheme 14*). Forcing conditions (trityl perchlorate and 2,4,6-tri-*tert*-butyl pyridine in dichloromethane) may cause etherification of secondary hydroxyl groups [95]. Alcohols can also be protected as triphenylmethyl ethers by treatment with p-methoxybenzyl trityl ether (PMBOTr) and DDQ under virtually neutral conditions [96].

Trityl ethers are generally cleaved under protic or Lewis acid conditions, such as formic acid [97], trifluoroacetic acid [98], BCl_3 [99], $Yb(OTf)_3$ [100], and $VO(OTf)_2$ [101]. Recently, supported-acids [102,103] or Nafion-H [104] have been found to be useful reagents for the removal of the triphenylmethyl group. Finally, trityl ethers are readily cleaved to the corresponding alcohols by using CBr_4/MeOH [105] or CBr_4-photoirradiation conditions [106].

Substituted trityl groups such as its mono- (MMTr), di- (DMTr) and trimethoxy- (TMTr) derivatives are also used for the protection of primary hydroxyls (*Fig. 1*). The MMTr and DMTr groups can be cleaved [107,108] under much weaker acidic conditions than the parent trityl ether due to the electron-releasing effect of their methoxy groups toward the benzene ring. None of these trityl ethers is stable enough to survive under normal glycosylation conditions, and therefore they are only used as intermediates to construct building blocks in carbohydrate chemistry. However, the use of DMTr as a protecting group is extremely widespread in oligonucleotide chemistry.



Scheme 14
Tritylation of methyl α-D-glucopyranoside

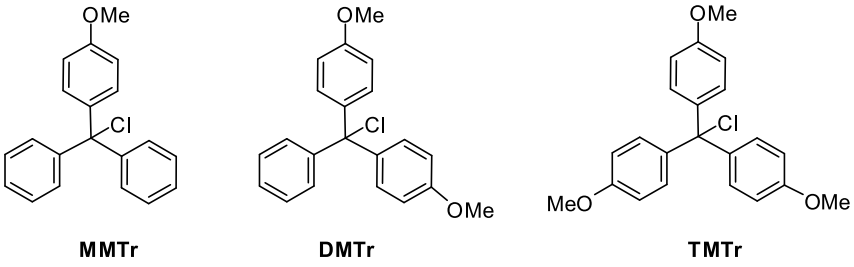
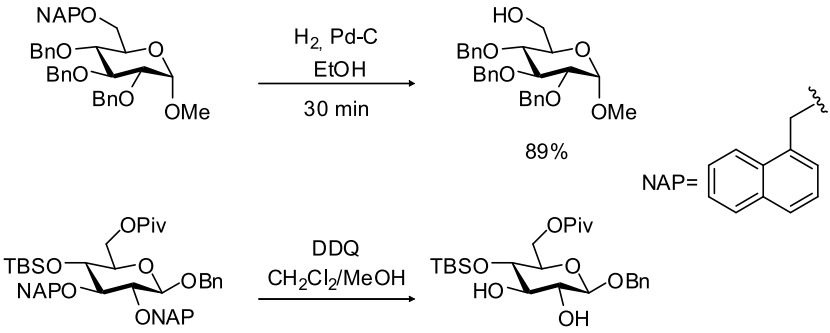


Figure 1
Substituted trityl protecting groups

2.1.6 2-Naphthylmethyl (NAP) Ethers

The 2-naphthylmethyl (NAP) group was introduced by Esko et al. [109] and Spencer et al. [110] as a protecting group for polyhydroxy systems. It is stable under conditions normally used for glycoside formation and offers the potential for selective cleavage by hydrogenolysis even in the presence of benzyl groups [110] (● *Scheme 15*). Standard conditions for introduction of the NAP group are, the alkylation with naphthyl bromide [110] or the hydrogenolysis of dioxolane-type (2-naphthyl)methylene acetals [111,112].



● Scheme 15
Selective removal of NAP ethers

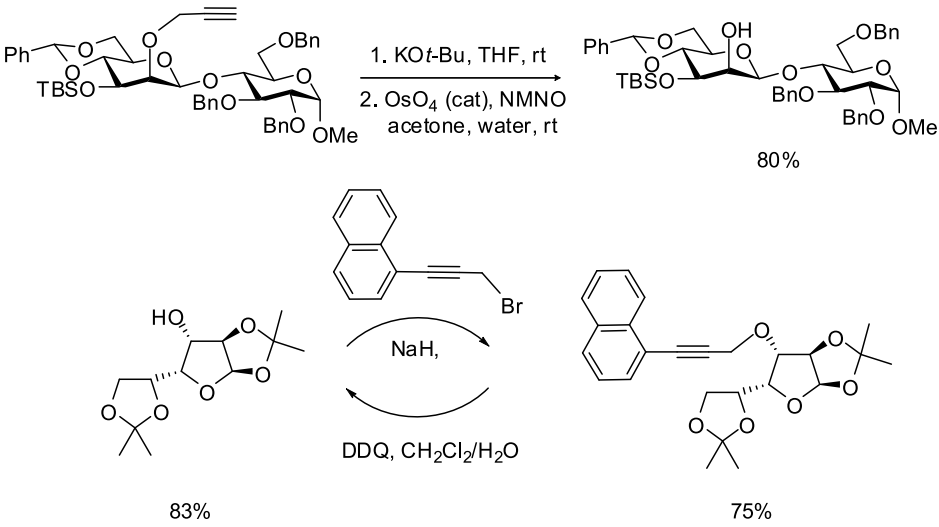
The NAP ethers can easily be removed by hydrogenolysis [110] or by DDQ oxidation under conditions which other usual protecting groups like acetyl, pivaloyl, phthalimido, benzyl, and benzylidene survive [113,114]. Recently some successful applications of sugar NAP ethers in the synthesis of complex oligosaccharides have been reported [115].

2.1.7 Propargyl Ethers

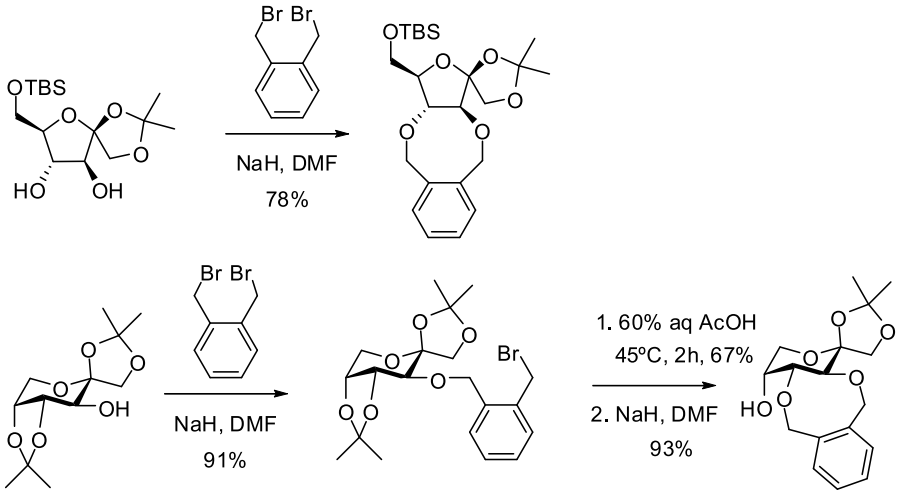
Crich and coworkers, have successfully used propargyl ethers as protecting groups in oligosaccharide synthesis [116,117]. Propargyl ethers are readily introduced under standard techniques and are cleaved by a two-set deprotection protocol: an initial treatment with base followed by catalytic osmylation of the resulting allenyl ether (● *Scheme 16*). The use of the (1-naphthyl)-propargyl group opens the possibility for its one-step cleavage with DDQ [117]. These two protecting groups, because of their minimal steric character, are useful for improving the diastereoselectivity of β-mannosylation reactions [116,117,118].

2.1.8 *o*-Xylylene Ethers

In contrast to all ether-type protecting groups so far mentioned, the *o*-xylylene group is a bifunctional protecting group devised for simultaneous protection of two vicinal hydroxyl groups of a carbohydrate molecule [119]. The *o*-xylylene group can be easily introduced by direct alkylation of the diol with α,α'-dibromoxylene or by a two-step process involving an initial alkylation of one hydroxyl function followed by an intramolecular ring-closing reaction (● *Scheme 17*). The *o*-xylylene protecting group has been successfully used as an element of conformational control of remote stereochemistry in the synthesis of spiroketals [120].



Scheme 16
Propargyl ethers as protecting groups



Scheme 17
The *o*-xylylene protecting group

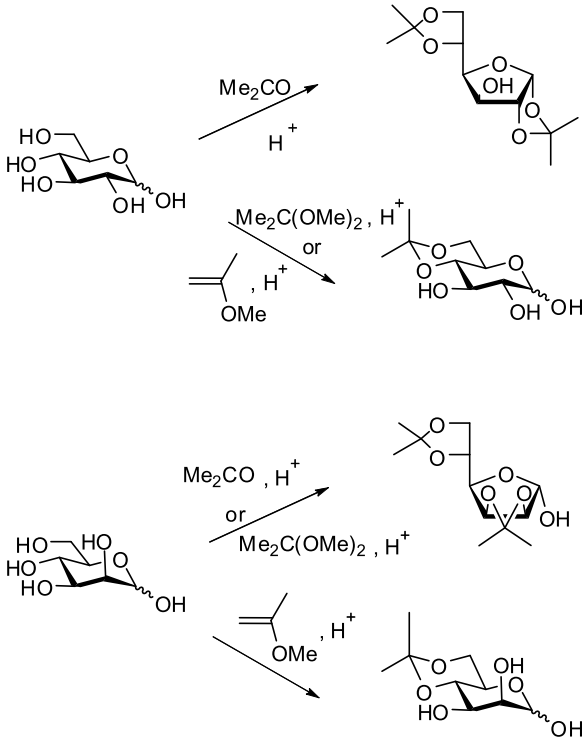
2.2 Acetalation Reactions: Acetal-Type Protecting Groups

Emil Fischer described as early as 1895 the formation of acetals of glycoses [121]. Since then, this type of protecting group has been extensively used in carbohydrate chemistry. Acetal protecting groups are readily available, easily introduced and removed, and stable to a good range of reactions. Standard conditions for the formation of acetals include treatment of a diol with a carbonyl reagent together with some acid catalyst.
Cyclic acetals such as benzylidene, isopropylidene, or 1,2-diacetals have been effectively used in the regioselective protection of diol systems. The ease of formation and the structures of products are a function of the regio- and stereochemistry of the hydroxyl groups and the properties of the employed carbonyl reagent. Acetals have been applied as protecting groups in many sugars including aminosugars and oligosaccharides. Exhaustive lists of catalyst and conditions can be found in reviews devoted to carbohydrates [122].

2.2.1 Cyclic Acetals

Isopropylidene (acetonides) and benzylidene derivatives are the most commonly used acetals for the simultaneous protection of 1,2- and 1,3-diols in carbohydrate and nucleoside chemistry [123]. Cyclohexylidene acetals are occasionally used, most often as an alternative to benzylidene acetals. Protection using cyclohexane-1,2-diacetals or the related butane-2,3-diacetals represents a new approach which has proved its value in complex oligosaccharide synthesis [124].
Besides being useful for selective protection of monosaccharides, cyclic acetals can display a number of interesting reactions, such as reductive or oxidative ring opening, that amplify the synthetic interest of these protecting groups [125].

Isopropylidene, Benzylidene, and Related Acetals Isopropylidene or benzylidene acetals are formed either by direct condensation of the diol with the appropriate carbonyl compound (acetone or benzaldehyde, respectively) or by transacetalation with the corresponding dimethoxy acetal. Both processes are carried out in acidic conditions [123].
One advantage of these acetals is their regioselective introduction. Benzylidene derivatives are formed preferentially with 1,3-diols of which anomeric or primary hydroxyl groups are a part. Therefore, they are generally used for 4,6-*O*-protection of pyranoses forming either *cis*- or *trans*-fused 1,3-dioxane rings. In these six-membered rings only the thermodynamically more stable [126] equatorial phenyl-substituted derivatives are observed. Formation of benzylidene acetals has also been achieved under basic conditions using α,α-dihalotoluenes in refluxing pyridine [127].
In contrast, isopropylidene acetals are more stable as five-membered 1,3-dioxane rings formed on *cis*-1,2-diols. Practically all examples in the literature show that, the use of acetone for the acetonation of sugars, leads to 1,3-dioxane rings, which are thermodynamically favored [128]. If 2-alkoxypropene is used as reagent, a reversal on the regioselectivity is observed, and the kinetic products (4,6-*O*-isopropylidene acetals) are preferentially formed [129]. An intermediate behavior is observed for the transacetalation process involving 2,2-dimethoxypropane which gives results either similar to those obtained with acetone or similar to those obtained with enol ethers (➤ *Scheme 18*) [130].

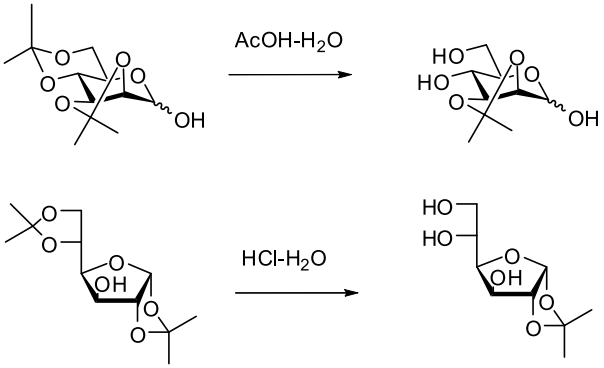


Scheme 18
Examples of isopropylidene formation on hexoses

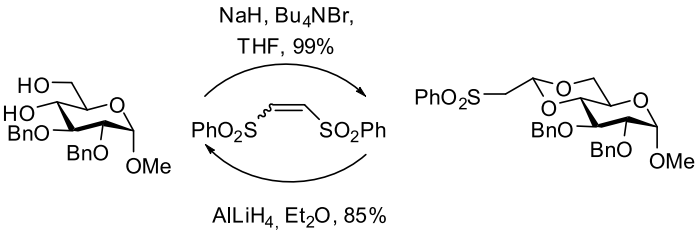
The most extensively adopted method for removal of these protecting groups involves the use of acidic conditions. There are various types of protic and Lewis acids that have been used for this purpose: aq. H_2SO_4 [131], Dowex acidic ion-exchange [132], trifluoroacetic acid [133], $Zn(NO_3)_2$ [134], supported $HClO_4 \cdot SiO_2$ [135] or recoverable $VO(OTf)_2$ [136]. Thiourea can also provoke the cleavage under essentially neutral conditions [137].

Relative to their deprotection, it should be emphasized, that the selective removal of one acetal in the presence of the same (or different) type of acetal, at distinct positions in the same molecule, is possible and has been observed quite often [138]. Several well-established observations can be summarized as follows (i) 1,3-dioxanes are hydrolyzed more easily than the corresponding 1,3-dioxolane, (ii) implication of the anomeric center renders the acetal function more stable, (iii) *cis*-fused 1,3-dioxolanes in a furanose or pyranose ring are more stable than the ones that involve a side chain, and (iv) *trans*-fused benzylidene acetals of hexopyranoses are hydrolyzed faster than the corresponding *cis*-fused acetals [123]. Some examples are shown in *Scheme 19*.

Some other alkyliden acetals with atypical properties have been used as protecting groups in carbohydrate chemistry. Thus, phenylsulfonylethylidene (PSE) acetals can be synthesized from glycosides under basic conditions. These derivatives are suitable for the protection of



Scheme 19
Examples of selective removal of acetal protecting groups



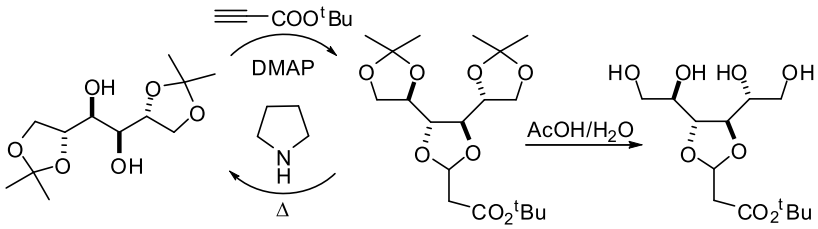
Scheme 20
Phenylsulfonylethylidene acetal as protecting group

1,2- and 1,3-diols. The equatorially configurated cyclic acetals are exclusively formed with 1,3-diols whereas the diastereoselectivity of the dioxolane-type acetals from 1,2-diols is quite poor. PSE acetals are deprotected to the corresponding diols under classical reductive conditions ($LiAlH_4$) (● *Scheme 20*) [139].

Vicinal diols in sugar substrates can also be protected as their 2-(*tert*-butoxycarbonyl)-ethylidene ("Bocdene") or 2-(methoxycarbonyl)-ethylidene ("Mocdene") derivatives in the reaction with *tert*-butyl or methyl propynoate. The acetal-like structures of these protecting groups is of interest because they are stable under acidic conditions, which allows their selective deprotection versus other acetals, and can be removed under basic conditions via an addition-elimination mechanism (● *Scheme 21*) [140]. The procedure is not suitable for 1,3- or 1,4 diols.

Diacetal Protecting Groups The pioneer work of Ley's group concerning the application of 1,2-diacetals such as the dispiroketal (dispoke) [141,142,143,144,145], the cyclo-hexane-1,2-diacetal (CDA) [146], and the butane 2,3-diacetals (BDA) [147] has found widespread application in carbohydrate chemistry (● *Fig. 2*) [148].

1,2-Diacetals are highly selective protecting agents which are able to discriminate di-equatorial diols in many carbohydrate derivatives [148]. These protecting groups are stable to functional group manipulation, glycosidation, and are easily removed at the end of a synthetic sequence



Scheme 21
2-(*tert*-Butoxylcarbonyl)-ethylidene acetal as protecting group

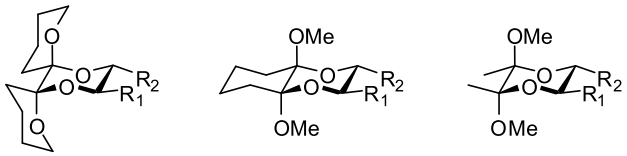


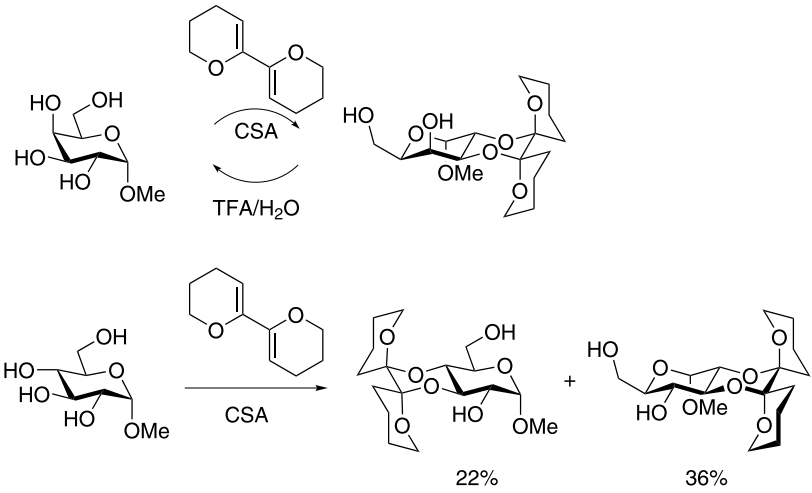
Figure 2
Diacetal protecting groups

by aqueous trifluoroacetic acid [149]. Furthermore, 1,2-diacetal protected substrates present a rigid structural architecture that is able to effect reactivity-tuning during glycosidation reactions. This property has been successfully used in oligosaccharide synthesis [150].

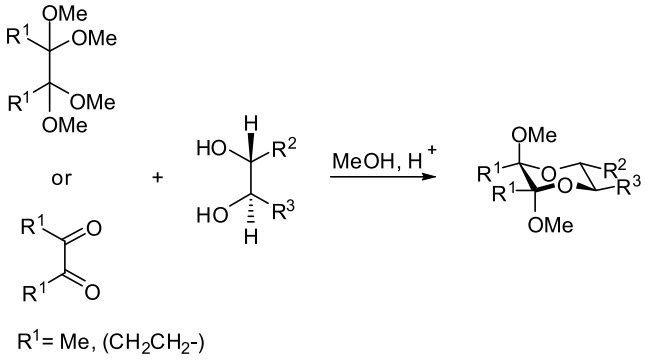
The dispiroketal protection of monosaccharides is controlled by the stabilizing influence of multiple anomeric effects leading to a single diastereomeric derivative. In certain examples, where there is more than one diequatorial diol pair present in the molecule, as for example in D-glucose derivatives, reaction affords a mixture of diacetals (*Scheme 22*). The reaction, often giving crystalline compounds, is carried out by treatment of the polyol with 3,4,3′,4′-tetrahydro-6,6′-bis-2H-pyran in chloroform at reflux in the presence of a catalytic amount of CSA [151].

Cyclohexane-1,2-diacetals (CDA) are, however, a better alternative for application to highly polar derivatives within the carbohydrate area. The cyclohexane-1,2-diacetals are formed by reacting 1,1,2,2-tetramethoxycyclohexane [146] or 1,2-cyclohexanedione [152] in boiling methanol containing a catalytic amount of CSA (*Scheme 23*). The corresponding 1,4-dioxane products are formed with high stereoselective control owing to favorable anomeric effects and equatorial placement of functionality around the periphery of the 1,4-dioxane ring. The CDA derivatives are often highly crystalline and usually do not require chromatography. They are stable but can be deprotected readily. They are also able to withstand a wide variety of reaction types such as iodination, reduction, oxidation, Wittig coupling, silylation, and glycosidation reactions.

Likewise, butane-2,3-diacetals (BDA) are good protecting groups for vicinal diequatorial diols. They are prepared either from butane-2,3-dione [153] or from the tetramethoxy butane-2,3-diacetal [147] (*Scheme 23*). BDA derivatives are usually isolated as solids rather than



Scheme 22
Examples of protection of diequatorial vicinal diols with dispiroketals



Scheme 23
Acid-catalyzed formation of 1,2-diacetals from 1,2–*trans*-diols

highly crystalline materials but they have very desirable NMR features that help analysis of the products since the methyl groups act as useful diagnostic markers. As with CDA, BDA groups are readily removed at the end of the synthetic sequence.

Pyruvate Acetals Pyruvate ketals are present in many lipopolysaccharides of bacterial origin, in capsular polysaccharides (*Fig. 3*) [154], and also in glycolipids isolated from fish nerve fibers [155]. As a result of the unique structural features including the presence of a negative charge on the carboxyl functional group and the chiral center, pyruvate ketals influence immunological specificity and patterns of immunological cross reactivity and therefore play an important role in cell–cell recognition processes [156]. Hexopyranosides containing pyru-

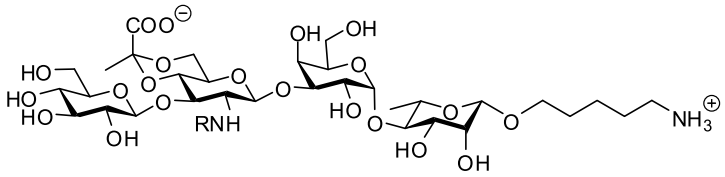
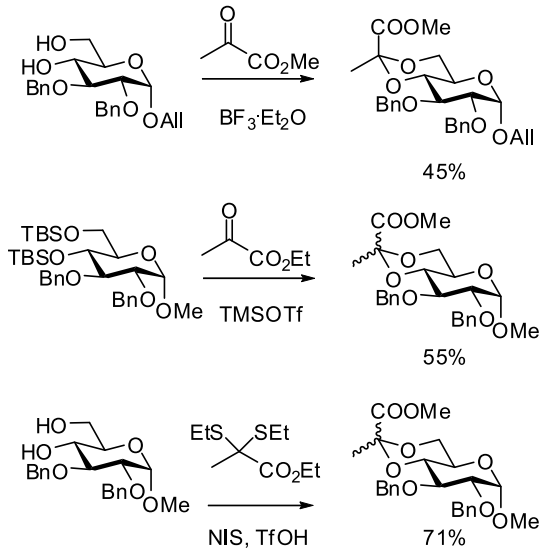


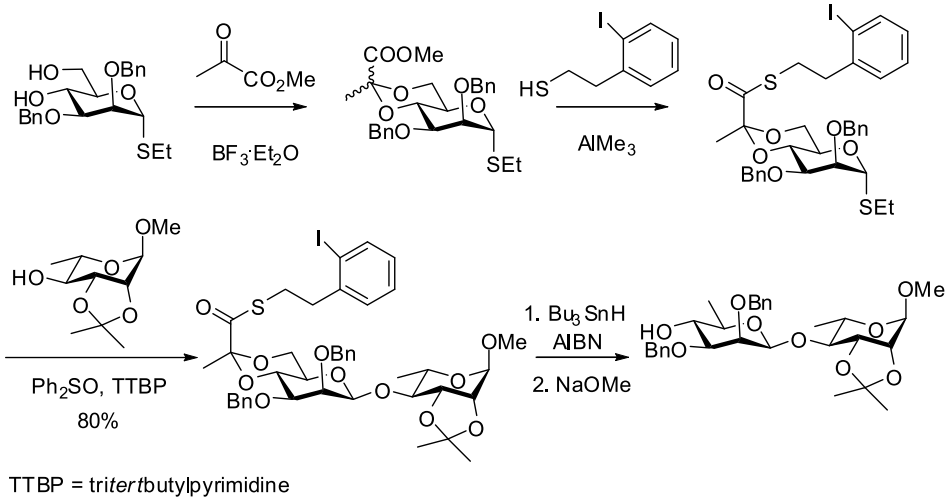
Figure 3
Pyruvated tetrasaccharide related to *Streptococcus pneumoniae* Type 27

vic acid ketals are also useful tools for immunochemical studies of *Klebsiella* polysaccharides [157]. Most commonly, pyruvate ketals are present as 4,6-ketals of hexose residues [158]. However, they also occur as 1,4-dioxolanes formed either from *cis*-axial-equatorial [159] or *trans*-diequatorial hydroxyl groups [160].

Pyruvate ketals can be synthesized [161] by direct condensation of a pyruvate ester with a diol in the presence of a Lewis acid, but this is less preferred because of the electron-withdrawing effect of the adjacent carboxylate group [162,163]. Therefore, several indirect methods for the acetalization have been introduced including condensation with pyruvate derivatives [164,165] or generation of the carboxylate group by oxidation of a suitable precursor [166,167,168,169]. A more efficient route to pyruvic acid acetals starts from silylated diols [170] or by the reaction between diols and methyl pyruvate dialkyl dithioacetal [171,172] activated by methyl triflate, dimethyl(methylthio)sulfonium trifluoromethane sulfonate (DMTST), nitroso tetrafluoroborate ($NOBF_4$), SO_2Cl_2-trifluoromethanesulfonic acid, or *N*-Iodosuccinimide (NIS) and trifluoromethanesulfonic acid [173] (*Scheme 24*).



Scheme 24
Examples of synthesis of pyruvate acetals of carbohydrates



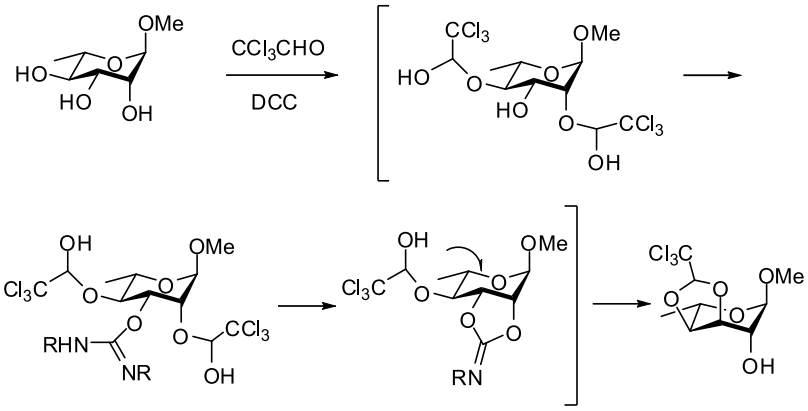
Scheme 25
Pyruvate acetals for stereoselective formation of β-D-mannopyranosides

Pyruvate-related acetals have been introduced by Crich as new protecting groups for carbohydrate thioglycoside donors [174]. The group conveys strong β-selectivity with thiomannoside donors and undergoes a tin-mediated radical fragmentation to provide high yields of the synthetically challenging β-rhamnopyranosides (*Scheme 25*) [174]. Besides this protecting group, it has been shown that this approach can also be applied to other related cyclic acetals [175,176,177].

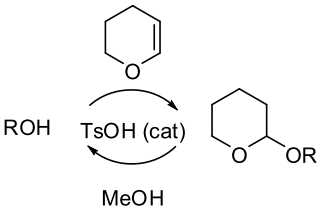
2,2,2-Trihaloethylidene Acetals In 1992 it was found that the reaction of hexafluoroacetone or chloral and dicyclohexylcarbodiimide (DCC) with bis-vicinal triols having a *cis-trans* sequence of hydroxyl groups resulted in the formation of cyclic acetals in which the central carbon of the triol had the inverted configuration [178,179]. In this acetalization the oxygen atom of the carbonyl compound (but not that of the alcoholic component) is inserted into the acetal moiety. As shown in *Scheme 26*, this non-classical pathway involves the in situ formation of a cyclic imidocarbonic ester intermediate, followed by an intramolecular S_N2-attack by a deprotonated neighboring hemiacetal moiety [180,181]. The resulting cyclic acetals are acid-stable but can be converted into the acid-labile ethylidene acetals by treatment with Raney Ni or Bu_3SnH.

2.2.2 Acyclic Acetals

Acyclic *O,O*-acetals are used for the temporary protection of mono-alcohols. Most commonly used are the tetrahydropyranyl (THP), the methoxymethyl (MOM), the benzyloxymethyl (BOM), or the methoxyethoxymethyl (MEM) protecting groups.



Scheme 26
Example of epimerization by non-classical acetalization



Scheme 27
Tetrahydropyranyl protecting group

Tetrahydropyranyl Ether The tetrahydropyranyl ether is one of the most frequently used protecting groups for alcohols during multi-step organic synthesis. It is usually introduced by treatment of the corresponding alcohol with 3,4-dihydro-2H-pyran in the presence of an acid catalyst (*p*-toluenesulfonic acid, PPTS, $BF_3 \cdot OEt_2$, or cation-exchange resins) (*Scheme 27*). Several other methods aiming to introduce the THP group under neutral conditions have been reported [182,183,184,185,186,187]. The resulting tetrahydropyranyl ethers offer stability towards strongly basic reaction conditions, organometallics, hydrides, acylating reagents, and alkylation reagents and the deprotection is usually performed as an acidic hydrolysis or alcoholysis.

In spite of the stability of THP ethers, its use in carbohydrate chemistry is currently limited due to the resulting diastereomeric mixtures obtained by the introduction of an additional stereocenter. This fact makes chromatographic separation and characterization of the products difficult and therefore their use will probably decrease in the future.

Alkoxymethyl Ethers The principal members of this set of protecting groups are: methoxymethyl ether (MOM) [188], methoxyethoxymethyl ether (MEM) [189], benzyloxymethyl ether (BOM) [190], *p*-methoxybenzyloxymethyl ether (PMBM) [191], and trimethylsilylethoxymethyl ether (SEM) [192] (*Fig. 4*). Since these protecting units are devoid of chirality, their use introduces no stereochemical complications.

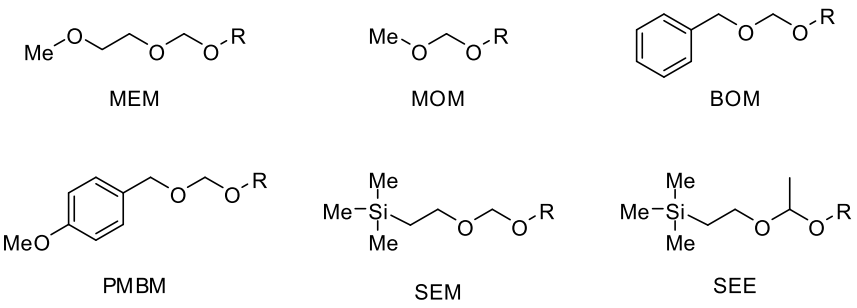


Figure 4
Principal members of the alkoxymethyl ether family

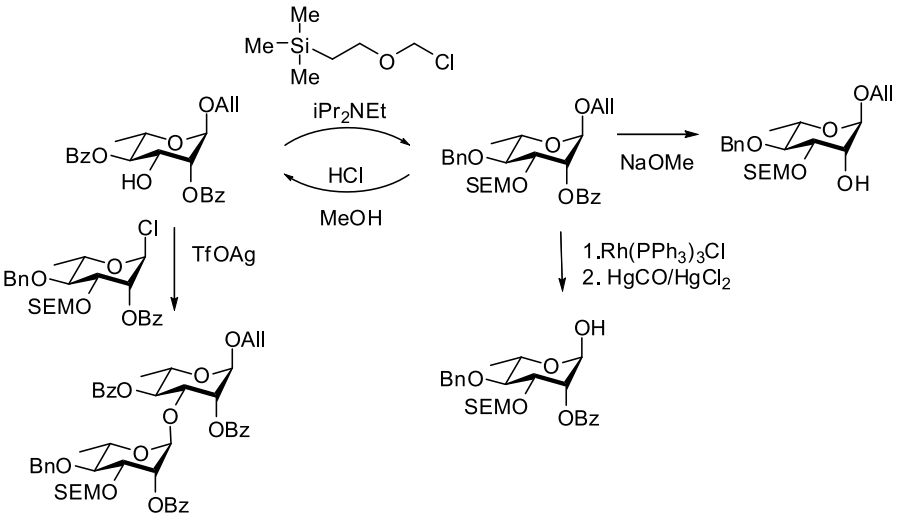
Usually, the formation of the alkoxymethyl ethers is effected by reaction of the corresponding chloride with either the sodium or lithium salt of the alcohol to be protected. The resulting ethers are stable to a wide variety of conditions including many organometallic reagents, reducing conditions, oxidizing agents, and mild acids. Indeed a number of the alkoxymethyl ethers are orthogonal to each other and can be used strategically in multistep syntheses. For example, SEM, BOM, and PMBM ethers are cleaved using a fluoride source, hydrogenolysis and oxidation, respectively. The alkoxymethyl ethers also vary in their degree of lability to Brønsted acids with MOM ether being the most robust.

The SEM group can be removed under milder reaction conditions than the MEM or MOM analogs. As shown in *Scheme 28*, the compatibility of this group with the conditions required for other selective functional group transformation, together with its stability under glycosylation conditions, allowed the preparation of rhamnopyranosyl synthons for the elaboration of higher order oligosaccharides (*Scheme 28*) [193].

The 1-[2-(trimethylsilyl)ethoxy]ethyl (SEE) group closely resembles the SEM group and can be introduced to alcohols with 2-(trimethylsilyl)ethyl vinyl ether in the presence of a catalytic amount of PPTS under neutral or slightly acidic conditions [194]. The SEE group can be removed with TBAF in THF (24 h, rt. to 45 °C).

2.3 Acylation Reactions: Ester-Type Protecting Groups

Acylation of hydroxyl groups of carbohydrates is one of the most commonly used functional group protection techniques in the synthesis of oligosaccharides. Acyl groups are readily introduced with many acylating agents of different reactivity. They are easily removed under basic (aqueous or non-aqueous) conditions, but are fairly stable under acidic conditions. The main drawback of esters as protecting groups is that they have a tendency to migrate (especially acetates), both under acidic and basic conditions. This is a concern in partially protected derivatives, and results in a mixture with the most stable compound preponderant. Thus, in *cis*-hydroxyls, there is normally a preferred migration from the axial position to the equatorial one and in 4,6-diols the migration goes from *O*-4 to *O*-6 preferentially [195].



Scheme 28
Example of use of SEM protecting group

There is a very large number of different ester protecting groups available and only the more common representatives in carbohydrate chemistry will be treated here. This includes acetate and substituted acetates such as chloroacetate, pivaloate, and levulinate groups. Aroyl groups are frequently used, such as benzoyl and substituted benzoyl e. g. p-phenylbenzoyl and 2,4,6-trimethylbenzoyl groups.

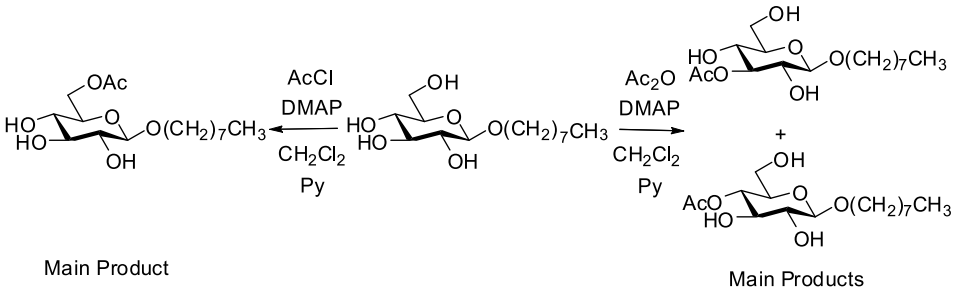
2.3.1 Acetyl (Ac) and Benzoyl (Bz) Esters

General Aspects In carbohydrate chemistry, per-*O*-acetylated sugars are inexpensive and useful intermediates for the synthesis of several natural products containing glycosides, oligosaccharides, and other glycoconjugates [196]. The acetylation reaction has also been employed for structural elucidation of many natural products containing carbohydrates. Acetylation of sugar alcohols is often carried out using a large excess of acetic anhydride or, more rarely, acetyl chloride, in the presence of pyridine (or other tertiary amine). Pyridine derivatives, such as 4-(dimethylamino)pyridine and 4-(pyrrolidino)pyridine have been added to the reaction as co-catalyst to speed up the acetylation reaction [197,198]. Similar considerations are valid for the *O*-benzoylation with the exception that benzoyl chloride rather than the anhydride is used.

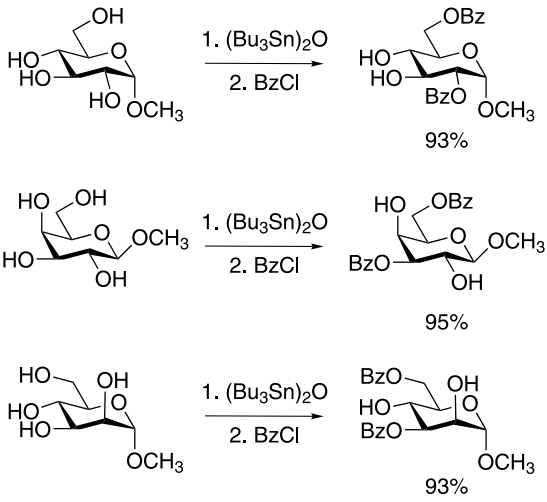
Recently, imidazole has been successfully applied as a catalyst for the acetylation of carbohydrates in acetonitrile [199]. A variety of other catalysts in combination with excess of acetic anhydride and solvent includes sodium acetate [200], sulfuric acid [201], perchloric acid [202], and a number of Lewis acid catalysts such as, iodine [203], $Sc(OTf)_3$ [204], $Cu(OTf)_2$ [205], $CoCl_2$ [206], $BiOCl\text{-}SOCl_2$ [207], $LiClO_4$ [208], $FeCl_3$ [209], $BiCl_3$ [210], and a series of heterogeneous catalysts such as, montmorillonite K-10 [211], zeolites [212], nafion-H [213], $HClO_4\text{-}SiO_2$ [214], or molecular sieves [215]. Recently, a $ZnCl_2$–sodium acetate combina-

tion [216] or $InCl_3$ [217] with acetic anhydride under microwave conditions has been reported for the acetylation of carbohydrates. A few reports have also appeared on the acetylation of carbohydrates using ionic liquids as solvents and catalysts [218,219].
Zemplén deacylation is the most commonly used deblocking reaction for the removal of ester protecting groups [220]. Using this transesterification reaction, OH-functions can be regenerated under mild conditions, in methanol with a catalytic amount of sodium methoxide at room temperature. The difference in the rate of benzoate and acetate solvolysis is sufficient to enable removal of acetates in the presence of benzoates. Typical conditions for this selective cleavage include ammonia in MeOH.

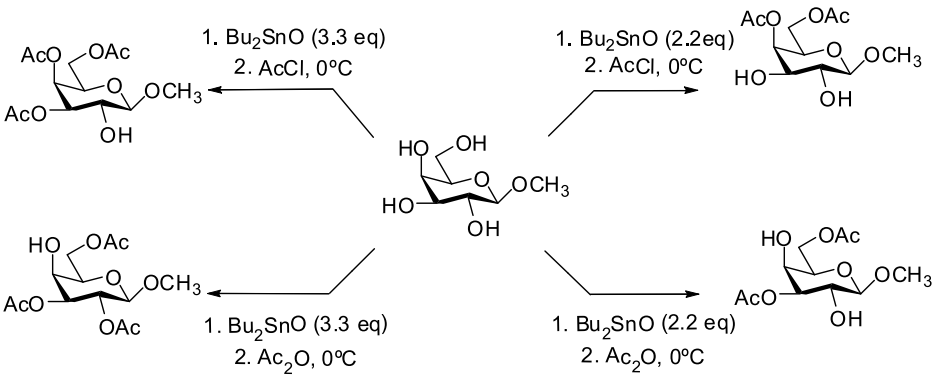
Regioselective Acylation Regioselective esterification of carbohydrates may be achieved in part by making use of the differing reactivities of hydroxyl groups. While the selective protection of primary hydroxy groups with sterically demanding acyl residues (e. g. by pivaloylation) is rather easily achieved, it is more difficult to protect one of a number of secondary hydroxy groups. Factors determining the regioselectivity of the acylation of secondary hydroxy groups in carbohydrates have been studied [221]; the most important ones being steric hindrance, intramolecular hydrogen bonding, and the configuration of the hydroxy groups. For instance, the presence of vicinal axial heteroatoms, such as O and S, enhances the nucleophilicity of the corresponding vicinal equatorial OH.
Regioselective acetylations have been promoted by different reagents such as alumina in refluxing ethyl acetate [222,223], silica gel-supported lanthanide chlorides and methylorthoacetate [224], a hindered base and acetyl chloride at low temperature [225], iminophosphorane bases and vinylacetate [226], NaH and 3-acetyl-thiazolidine-2-thiones [227], PPh_3/CBr_4 in ethyl acetate at high temperatures [228]. Recently, it has been shown that the rate and the selectivity of an acetylation reaction can be controlled by the counterion of the acetylating agent under nucleophilic catalysis. The team play of reagent, catalyst, and auxiliary base is responsible for the outcome of the reaction [229]. Thus, octyl β-D-glucopyranoside can be acetylated with high selectivity either on the primary or on secondary OH groups by using different acetylation agents under otherwise identical conditions (● *Scheme 29*).
The use of organotin reagents (● *Sect. 2.1.2* under ● *"Formation of Organotin Intermediates"*) provides a useful means of efficient regioselective acylations. There does not seem to be significant differences in selectivities between alkylation and acylation although forma-



Scheme 29
Counterion-directed regioselective acetylation of octyl β-D-glucopyranoside



Scheme 30
Examples of stannyl mediated regioselective benzoylation



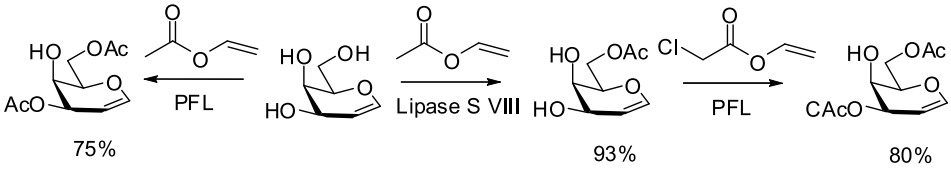
Scheme 31
Reagent-dependent multiprotection of methyl β-D-galactopyranoside

tion of *O*-acyl derivatives is a much faster reaction in any solvent, and it does not require heating or the presence of an additional nucleophile. Thus, treating methyl α-D-glucopyranoside with $(Bu_3Sn)_2O$ and then with BzCl gave the 2,6-di-*O*-benzoate, while the 3,6-di-*O*-benzoates were obtained upon analogous benzoylation of methyl β-D-galactopyranoside and methyl α-D-mannopyranoside, respectively (*Scheme 30*) [230].

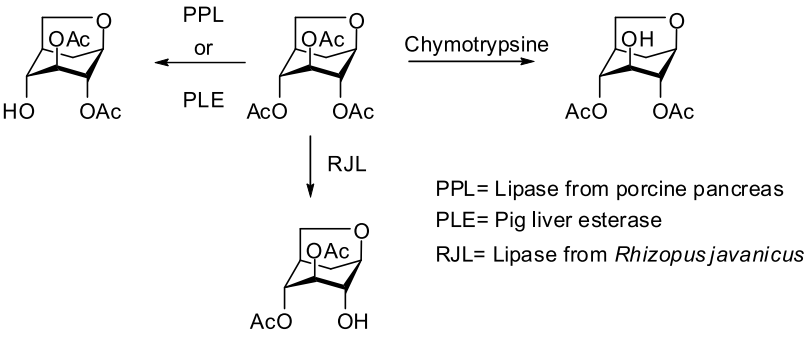
An extension of this tin chemistry to the regioselective acylation of unprotected sugars bound to a resin shows the possibility of using solid-phase techniques for the preparation of *O*-acyl derivatives of carbohydrates [231]. Very recently, it has been reported that organotin-mediated multiple carbohydrate esterifications can be controlled by the acylating reagent and the solvent polarity. When acetyl chloride is used, the reactions are under thermodynamic control, whereas when acetic anhydride is employed, kinetic control takes place (*Scheme 31*) [232].

Enzymatic regioselective protection techniques are also an interesting and useful method for the protection of hydroxyl groups [233]. Such techniques are exclusively directed at regioselective acylation and deacylation, mostly by using different lipases [234] or proteases [235], which can catalyze acyl transfer reactions from activated esters to suitable acceptors. The most frequently used enzymes are *Porcine pancreatic lipase* (PPL), *Protease N-neutral protease* (PN), *Pseudomonas fluorescens lipase* (PFL), *Chromobacterium viscosum lipase* (CVL), and *Candida cylindracea lipase* (CCL). The results of the enzymatic acylation of several pyranoses and furanoses have been reviewed [236]. Almost all combinations of enzymes and substrates lead to acylation of the primary hydroxy group. The regioselectivities are usually higher than 70%, and the conversions between 40 and 100%. However, if the 6-OH groups are protected first or deoxygenated, in the corresponding enzymatic reactions, selectivities on the acylation of secondary hydroxyl groups are observed. An example is shown in *Scheme 32*, where enzymatic acyl transfer reactions turned out to be a viable method for the complete differentiation of the hydroxyl groups of glycal derivatives [237].

Enzymes are not only capable of introducing but also of removing acyl groups into carbohydrates [233]. For example, each of the three OH groups in 1,6-anhydroglucopyranose can be liberated selectively making use of enzymatic reactions (*Scheme 33*) [238,239,240]. The lipase-mediated hydrolysis proceeds with higher velocity and, in many cases with better selectivity, if butanoates or pentanoates are employed as substrates instead of acetates. In all cases the reaction conditions are so mild that the acid sensitive structures remain unaffected.



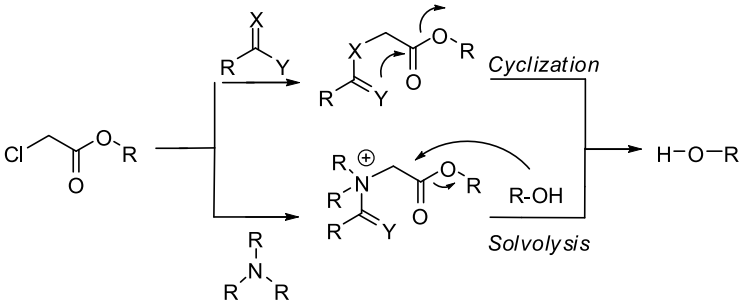
Scheme 32
Examples of enzymic regioselective acylation



Scheme 33
Examples of enzymic regioselective deacylation

2.3.2 Substituted Acetyl Esters

The lability of acetates is enhanced by introducing chlorine atoms in the α-position. Thus, chloroacetates (Cac or ClAc) hydrolyze faster than acetates and trichloroacetates are so reactive that they are rarely used in synthesis. Thus far, thiourea [241], hydrazine dithiocarbonate [242], pyridine [243] and diazabicyclo[2.2.2] octane (DABCO) [244] are representative of dechloroacetylation reagents. 1-Selenocarbamoylpiperidine also deprotects the *O*-chloroacetyl group with high chemoselectivity in the presence of other acyl groups such as acetyl, pivaloyl, and Fmoc without the assistance of a base [245]. Thiourea, hydrazine dithiocarbonate, or 1-selenocarbamoylpiperidine are believed to deprotect the ClAc group by following a cyclization mechanism. This mechanism is illustrated in *Scheme 34*: the nucleophilic atom (X) of the reagent replaces the chlorine atom of the ClAc group, and then another nucleophilic atom (Y) attacks the carbonyl carbon to break the C–O bond, thereby resulting in the production of free hydroxyl. In contrast, tertiary-amine-containing reagents such as pyridine and DABCO presumably attack the α-carbon to form onium salt, which is then solvolyzed by water, MeOH, or EtOH to produce naked hydroxyl group.



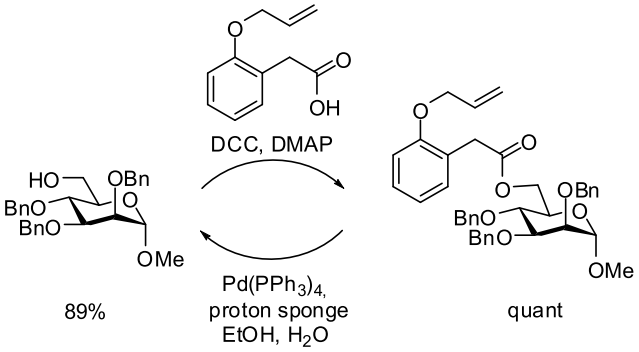
Scheme 34
Plausible mechanisms of removal of chloroacetyl groups

The chemoselective deprotection of the CAc group does not affect other protecting groups such as acyl derivatives (acetyl, benzoyl, or levulinoyl groups), carbonates, *p*-methoxybenzyl or silyl ethers and therefore has been included in sets of orthogonal protecting groups. However, the sensitivity of the chloroacetyl group may impose limitations for its application in the synthesis of complex oligosaccharides.

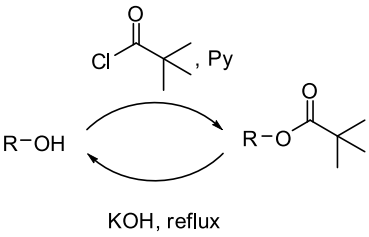
The 2-(allyloxy) phenyl acetyl (APAC) group has been proposed as a new robust acyl-type protecting group for hydroxyl groups [246]. It can be removed under mild conditions by relay deprotection whereby the phenolic allyl ether is cleaved by treatment with a transition metal followed by intramolecular ester cleavage by nucleophilic attack of the revealed hydroxyl (*Scheme 35*). It is compatible with glycosylations and can perform efficiently neighboring group participation leading to the exclusive formation of 1,2-*trans* glycosides.

2.3.3 Pivaloyl (Piv) Esters

The bulky pivaloyl group has been used as a protecting group in the synthesis of acylated nucleosides [247], monosaccharides, and disaccharides [248]. The pivaloyl esters are usually highly crystalline compounds, its position in a molecule is easily detectable by ^{1}H NMR, and it can be



Scheme 35
The 2-(allyloxy)phenyl acetyl protecting group



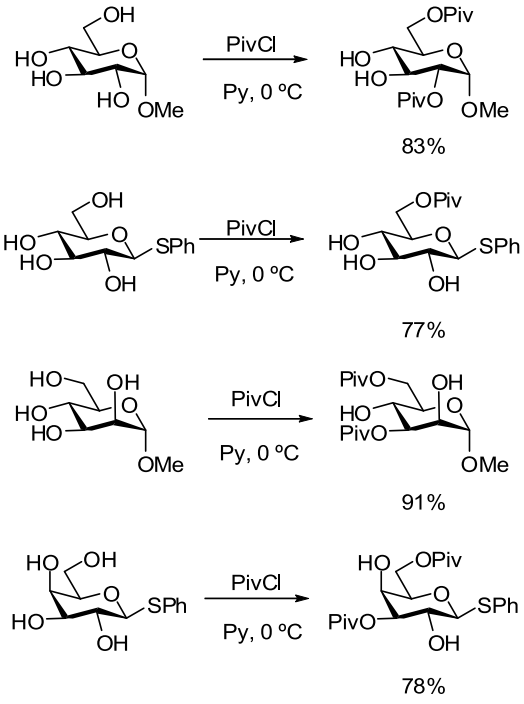
Scheme 36
The pivaloyl protecting group

removed totally or selectively by esterases from mammalian sera [249]. Furthermore, the use of pivaloate esters is advantageous in the stereoselective preparation of β-glycosidic linkages because they inhibit the sometimes competitive process of orthoester formation [250,251].

Pivaloylation of sugar alcohols is carried out using pivaloyl chloride in the presence of pyridine or with *N*-pivaloyl imidazole [252] in DMF. Pivaloyl esters are typically cleaved by base-catalyzed solvolysis (*Scheme 36*). Their greater steric hindrance makes them react slower than other acyl groups and some selectivity in the hydrolysis of different ester protecting groups may be observed.

A systematic study in the selective pivaloylation of various pyranosides and oligosaccharides [253] has shown that, in the absence of adjacent axial alcoxy groups, pivaloylation preferentially occurs at the primary hydroxyl groups. However, in the presence of an adjacent axial function, the reactivity of the vicinal secondary hydroxyl group is as high as that of the primary group towards pivaloylation (*Scheme 37*).

A limitation in the use of pivaloyl esters as protecting groups in a polyfunctional system is the harsh condition required for its cleavage (especially at sterically hindered secondary centers). The 4-acetoxy-2,2-dimethylbutanoyl (ADMB) esters have been proposed as an alternative because they are easily prepared, show similar reactivity in carbohydrate acylations, and are removed under much milder conditions (catalytic quantity of DBU at room temperature) [254].

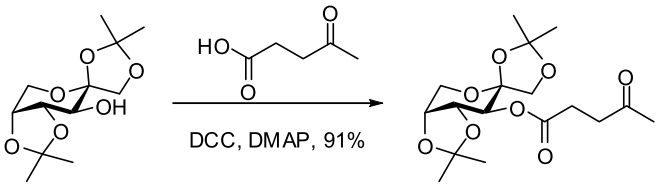


Scheme 37
Regioselective pivaloylation of hexopyranosides

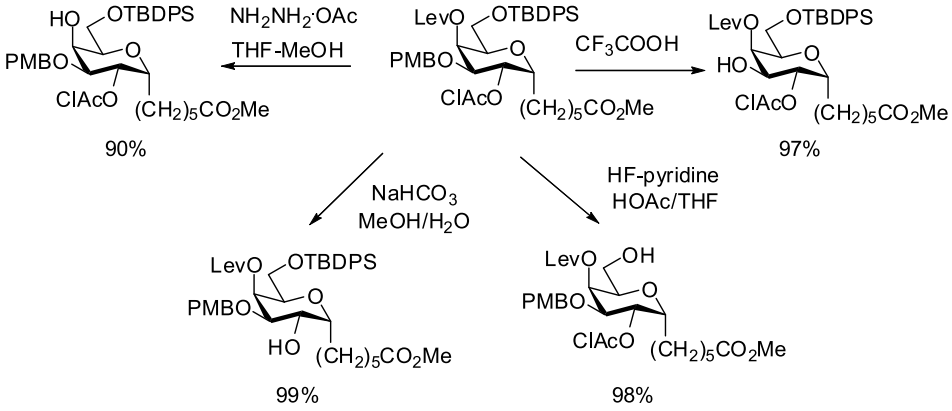
2.3.4 Levulinoyl (Lev) Esters

The levulinoyl (4-oxopentanoyl) moiety is a very useful temporary protecting group in nucleotide, polysaccharide, and glycolipid synthesis.

The levulinoyl esters are prepared from the free hydroxyl group by treatment of levulinic acid with DCC [255,256] or 1-[3-(dimethylamino)propyl]-3-ethylcarbodiimide hydrochloride (EDAC) [257] in the presence of DMAP (*Scheme 38*). Additionally, 3′- and 5′-*O*-levulinyl protected derivatives of 2′-deoxy nucleosides have been prepared by regioselective enzymatic acylation using a variety of lipases and acetonoxime levulinate as acylating agent [258]. In contrast to other ester substituents, the *O*-levulinoyl group is far less prone to migration [259].



Scheme 38
Example of introduction of levulinoyl protecting group



Scheme 39
Example of orthogonal hydroxyl protecting groups

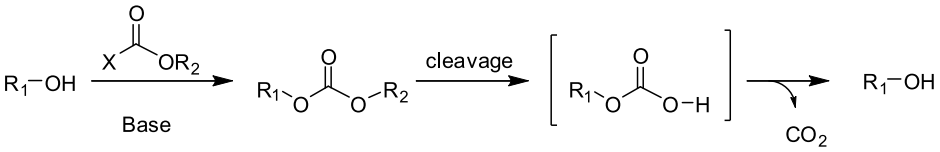
Levulinates are stable during coupling reactions and can be selectively cleaved without affecting other protecting groups in the same molecule. Actually, levulinoyl esters can be removed with hydrazine acetate under conditions that do not harm other acyl groups such as acetates, benzoates, or even chloroacetates [255]. This selectivity has allowed the levulinoyl group to be included in a number of orthogonal sets designed for the synthesis of collections of oligosaccharides. For example, Wong and coworkers showed that chloroacetyl, *p*-methoxybenzyl, levulinoyl, and *tert*-butyldiphenylsilyl groups can each be removed selectively and the freed hydroxyl group employed in glycosylation reactions (*Scheme 39*) [260]. Zhu and Boons showed that Fmoc, Lev, and diethylisopropylsilyl are another attractive set of orthogonal hydroxyl protecting groups for aminosugars [261].

2.4 Carbonylation Reactions: Carbonate-Type Protecting Groups

Carbonates represent an important family of protecting groups of hydroxyl groups. All of the members of the carbonate family are easy to introduce by reaction of the free alcohol with chloroformates or mixed carbonate esters. In general, carbonates are less reactive than esters towards basic hydrolysis owing to the reduced electrophilicity of the carbonyl afforded by the resonance deactivation by two oxygens. However, the conditions that attack esters may also attack carbonates.

Besides simple alkoxycarbonyl groups which are usually removed under basic hydrolysis, more sophisticated groups have been designed which are removed under milder and more specific conditions. In general, all the carbonate protecting groups are close relatives to the most important carboxyl protecting groups. The adaptation works because *O*-alkyl cleavage releases an unstable intermediate which decomposes with loss of carbon dioxide to give the free alcohol (*Scheme 40*).

The most commonly installed carbonates on carbohydrate derivatives include CBz, Troc, Aloc, Poc, and Fmoc groups.



Scheme 40
Formation and cleavage of carbonate-type protecting groups

2.4.1 Benzyl Carbonates (Cbz)

The benzyloxycarbonyl group (Cbz or Z) is useful in carbohydrate synthesis, not only for *N*-protection of amino sugars, but also to protect alcohols [262,263]. The main advantage of this group is that it is cleaved by hydrogenolysis, and when compared to benzyl ethers, benzyl carbonates are not only removed more readily [264] but also allow hydroxyl group protection under softer conditions than those employed for benzylation.

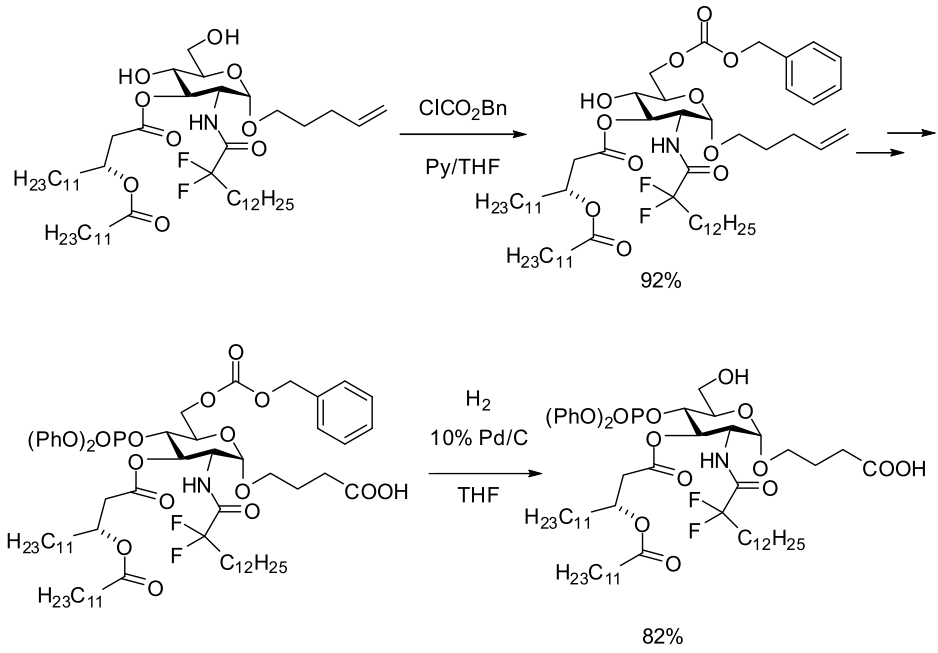
The benzyloxycarbonates are usually prepared by treatment of the alcohol with benzyloxycarbonyl chloride in the presence of a base (DMAP or *N*-ethyldiisopropylamine). Aqueous basic medium has to be avoided in polyol systems since these conditions favor the obtention of cyclic carbonates [265].

The benzyloxycarbonyl group has been used for the selective protection of monosaccharides. For instance, in the synthesis of a 1-*O*-carboxyalkyl GLA-60 analogue, a primary alcohol was selectively protected with benzylchloroformate and pyridine (◉ *Scheme 41*) [266]. Furthermore, Gotor and Pulido showed that the reaction of D-glucose, D-mannose, and D-galactose with acetone *O*-(benzyloxycarbonyl)oxime in dioxane in the presence of a lipase from *Candida antarctica* allowed the selective benzyloxycarbonylation of the primary hydroxyl group [267]. On the other hand, the regioselective protection of secondary alcohols in pyranosides has also been achieved in high yields [268]. Thus, in the α-D-mannopyranoside series the 3-OH is the more reactive secondary alcohol whereas in the α-D-gluco and α-D-galacto series, the 2-OH is the more reactive group.

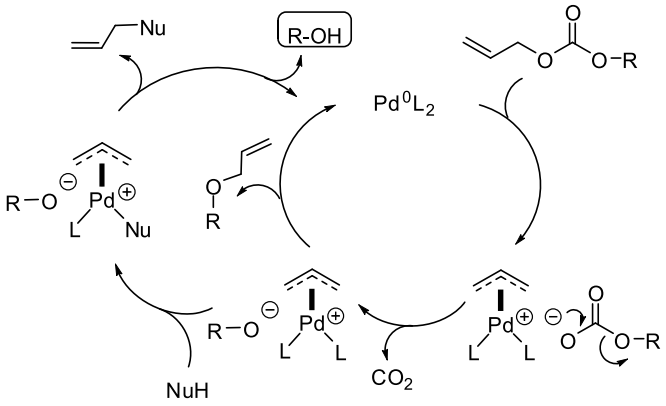
2.4.2 Allyl Carbonates (Aloc or Alloc)

The allyloxycarbonyl group [269] has shown a wide application in organic synthesis, especially in the fields of peptides, nucleotides, and carbohydrates. Allyloxycarbonyl derivatives are more easily prepared than the corresponding allyl ethers and they are more stable than ester protecting groups which find frequent use in carbohydrate chemistry.

Allyloxycarbonyl groups have been conveniently installed on primary and secondary hydroxyl groups of carbohydrate derivatives by reaction with allylchloroformate in the presence of TMEDA [270]. On the other hand, the Alloc group can be cleaved by transition-metal catalysts under conditions that are specific and with a high tolerance of other functional groups. As depicted in ◉ *Scheme 42*, allyl carbonates undergo facile oxidative addition with palladium(0) catalyst to afford π-allyl palladium complexes, which eject CO_2 to give, initially, allylpalladium alkoxides. Depending on the conditions, these intermediates either collapse to the allyl ether (◉ *Sect. 2.1.4*) or are intercepted by an external nucleophile to give the free alcohol [271].

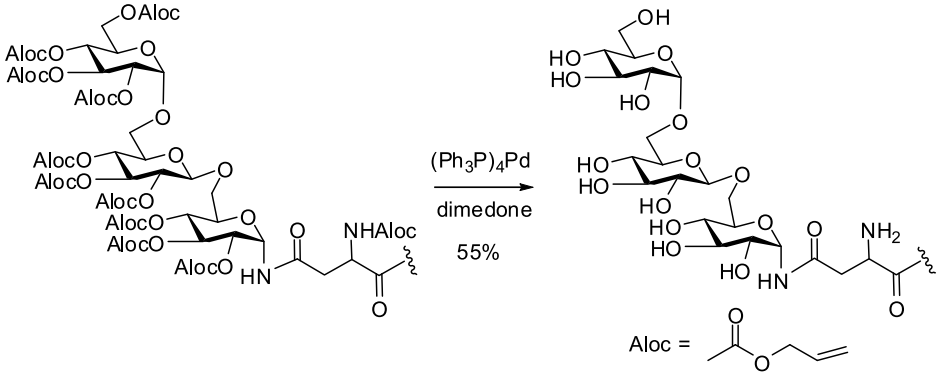


Scheme 41
Use of benzyloxycarbonyl group in the synthesis of a 1-*O*-carboxyalkyl GLA-60 analogue



Scheme 42
Palladium (0)-mediated cleavage of allyloxycarbonyl groups

An impressive example of how this protecting group has become a powerful tool for the construction of glycopeptides and oligosaccharides comes from the first synthesis of the glycopeptide nephritogenoside [272]. Its structure shows a trisaccharide composed of three glucose moieties linked to a peptide of 21 aminoacids. Taking into account the instability of the



Scheme 43
Aloc as a powerful tool for the construction of glycopeptides

molecule under acidic and basic conditions, the Aloc group was chosen as the final protecting group of amino and hydroxyl groups. Removal of a total of 11 allyl carbonates was carried out in one single step by treatment with palladium(0) and dimedone to give free nephritogenoside (*Scheme 43*).

2.4.3 Propargyl Carbonates (POC)

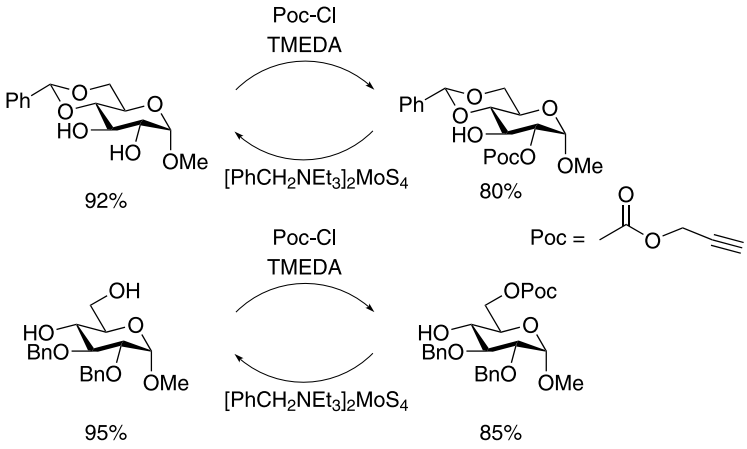
Recently, it has been shown that the propargyloxycarbonyl (Poc) group can be used for the protection of the hydroxyl function in carbohydrates [273]. The protection is achieved by treating the alcohol with propargyloxycarbonyl chloride (PocCl) in the presence of a suitable base. The mild reaction conditions can be modulated to attain regioselective protections (*Scheme 44*). The resulting propargyl carbonates are compatible with acidic, basic, and also glycosylation conditions.

Propargyl esters are deprotected effectively using benzyltriethylammonium tetrathiomolybdate $[PhCH_2Net_3]_2MoS_4$. The deprotected products usually can be isolated by simple filtration. Under the conditions of deprotection benzylidene acetals, benzyl ethers, acetyl and levulinoyl esters, and allyl and benzyl carbonates are left untouched and therefore can be used effectively for orthogonal protection in carbohydrate chemistry.

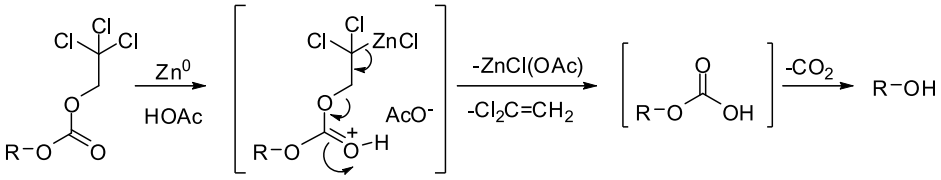
The utility of propargyloxycarbonyl chloride in simultaneous protection of alcohols and amines has been explored and it is possible to deblock propargyl carbonates leaving propargyl carbamates untouched [274].

2.4.4 2,2,2-Trichloroethyl Carbonate (TrOC)

Although very popular as an amino protecting group in peptide and glycopeptide chemistry, only a few reports deal with the trichloroethoxycarbonyl group as a hydroxyl protecting group in carbohydrate derivatives. The selective deprotection is carried out by treatment with zinc in acetic acid to give 1,1-dichloroethylene [275] (*Scheme 45*).

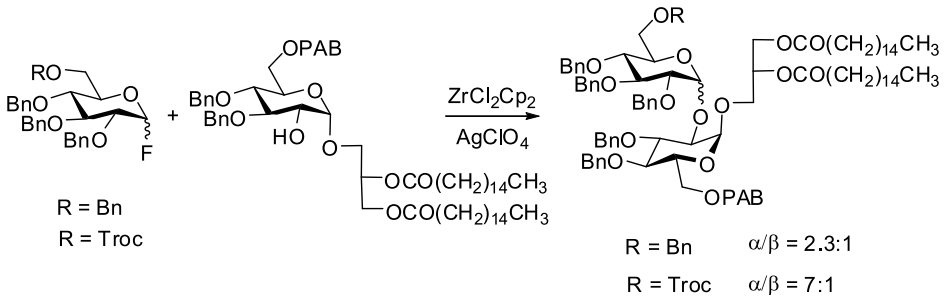


Scheme 44
Propargyloxycarbonyl as a protective group in carbohydrates



Scheme 45
Selective deprotection of 2,2,2-trichloroethoxycarbonyl group

Trichloroethoxycarbonyl groups have been installed on primary and secondary hydroxyl groups of carbohydrate derivatives by standard coupling with 2,2,2-trichloroethyl chloroformate. It has been shown that a Troc group in the primary position of a glycosyl donor reduces its reactivity but enhances α-selectivity in glycosylation couplings (*Scheme 46*) [276].



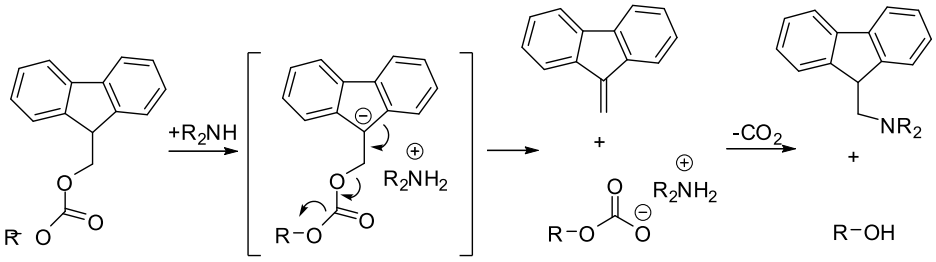
Scheme 46
Example of the α-orienting effect of the 6-*O*-Troc group

2.4.5 Fluoren-9-ylmethoxycarbonyl (Fmoc) Group

The Fmoc group is a well-established amino-protecting group [277] often used in peptide synthesis, but only recently has been recognized as a temporary hydroxyl protecting group for oligosaccharide synthesis [277,278,279,280].

The Fmoc group is readily introduced under standard conditions using FmocCl and a catalytic amount of DMAP in pyridine. The resulting carbonates are exceptionally stable under acidic conditions and therefore survive glycosylation reactions.

The Fmoc group can be removed with mild bases such as ammonia, piperidine, or morpholine [281]. The cleavage goes through a rapid deprotonation of the fluorene group to generate an aromatic dibenzocyclopentadienide anion. In a subsequent slower step, elimination generates dibenzofulvene (itself an unstable species that rapidly adds nucleophiles) and a carbonate residue, which then decomposes with loss of carbon dioxide to release the free alcohol (*Scheme 47*).



Scheme 47
Selective deprotection of fluoren-9-ylmethoxycarbonyl group

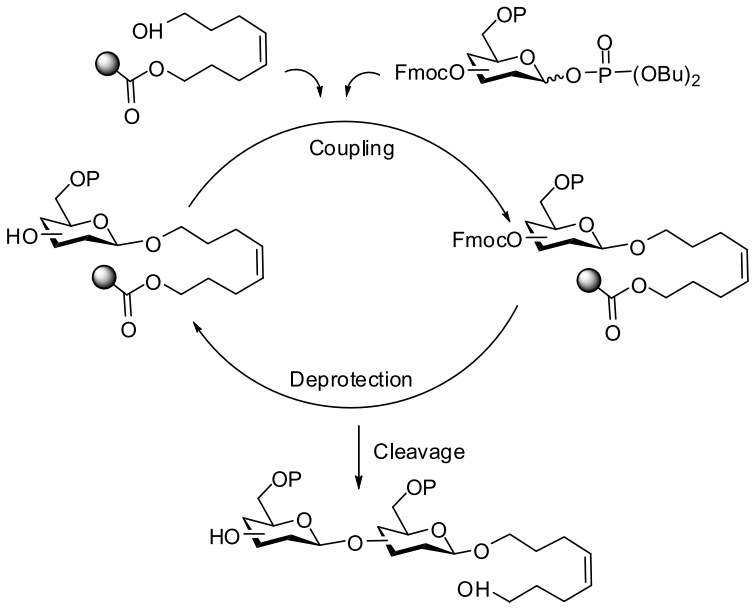
Recently, several groups have reported the use of glycosyl donors bearing Fmoc-protected hydroxyl groups for the solid-phase synthesis of saccharide libraries. Thus, a lactosyl donor, bearing an Fmoc-protected hydroxyl group, has permitted the effective construction of lactose-containing oligosaccharides in a solid-phase system [282,283].

The Fmoc group has also been used as a temporary protecting group in the automated synthesis of Lewis antigens. The UV active dibenzofulvene moiety released after Fmoc cleavage allowed for real-time monitoring of the reaction progress and provided a qualitative assay for the efficiency of each glycosylation and deprotection cycle during automated assembly (*Scheme 48*) [284].

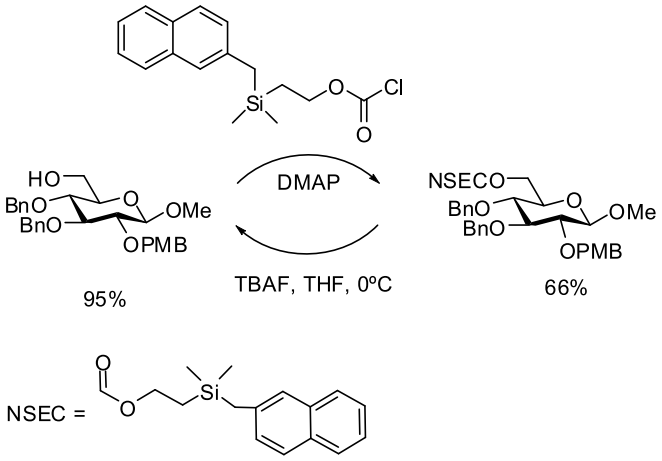
2.4.6 2-[Dimethyl(2-naphthylmethyl)silyl]ethoxycarbonyl (NSEC) Group

The 2-[dimethyl(2-naphthylmethyl)silyl]ethoxycarbonyl group is a novel temporary protecting group to mask hydroxyl groups [285]. In an analogous manner to the Fmoc protection, this group may be particularly useful for the automated assembly of oligosaccharides, as its cleavage can be followed by UV.

The NSEC group can be introduced under standard conditions using NSECCl which is available in three steps from chlorodimethylvinylsilane and 2-(bromomethyl)naphthalene (*Scheme 49*). The NSEC group may be difficult to introduce in sterically demanding positions.



Scheme 48
Fmoc as temporary protecting group in the automated synthesis of oligosaccharides



Scheme 49
NSEC as a protective group in carbohydrates

The NSEC group is stable to glycosylation conditions using glycosyl phosphates, and it is not affected under deprotection conditions that facilitate the removal of Lev, Fmoc, allyl, and PMB groups. For ester-type protecting groups selective deprotection in the presence of NSEC derivatives is not possible.

The removal of NSEC carbonates is carried out by reaction with TBAF in the presence of esters including acetyl, levulinoyl, benzoyl and pivaloyl, but also allyl and PMB ethers are not affected. However, the NSEC group cannot be selectively removed in the presence of Fmoc protecting groups.

2.5 Silylation Reactions: Silyl-Type Protecting Groups

2.5.1 Silyl Ethers. General Aspects

Silyl ethers have become very important alcohol protecting groups, which are used, among other ways, routinely in carbohydrate chemistry [286]. Silyl derivatives are generally stable over a wide variety of reaction conditions and at the same time are selectively removable in the presence of other functional groups including other protecting groups.

Additionally, the ability to vary the organic groups on silicon introduces the potential to alter the R_3Si group in terms of both its steric and electronic characteristics and thereby influence the stability of the silylated species to a wide variety of reaction and deprotection conditions [287].

The synthetic potential of silyl ethers as protecting groups for the hydroxyl groups was appreciated in the early 1970s and now these derivatives are probably used more than any other protecting group in organic synthesis. ❯ *Figure 5* summarizes the structures, names, and abbreviations of the most commonly used silyl ethers.

The usual method for their introduction into sugars is the reaction of one or several hydroxyl functions in the sugar with a trialkylsilyl chloride in the presence of a base, such as pyridine and imidazole. The less sterically hindered the silyl group the easier it is to introduce. The introduction of the sterically unimpeded trimethylsilyl group to a primary, secondary, or tertiary alcohol is a straightforward process taking place with a variety of reagents under mild, high-yield reaction conditions. On the other hand, the introduction of the more steri-

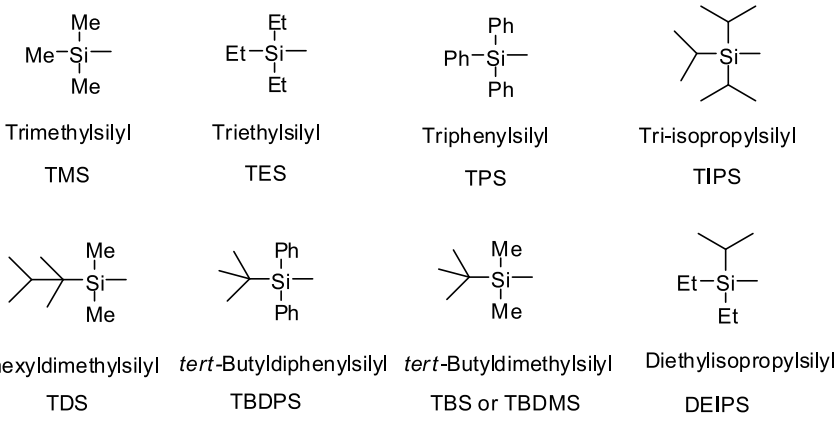


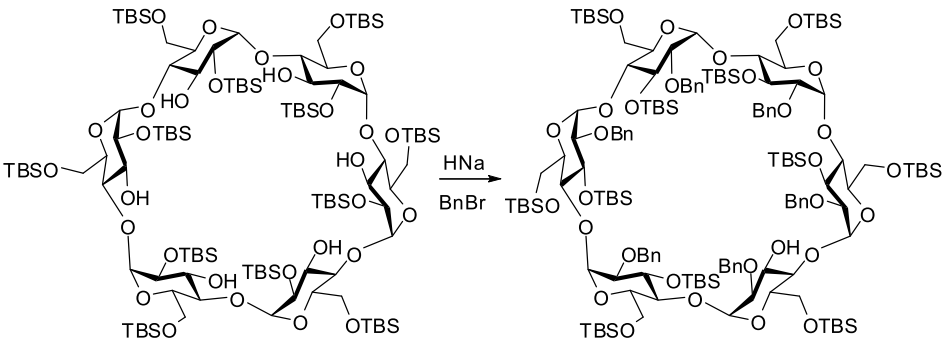
Figure 5
Principal members of the silyl ether family

cally demanding *tert*-butyldimethylsilyl group requires reaction of the alcohol with *tert*-butyldimethylchlorosilane in the presence of imidazole as a catalyst and the formation of the *tert*-butyldimethylsilyl ether of tertiary alcohols is very difficult. Alternatively, a more reactive form such as trialkylsilyl trifluoromethanesulfonates (R_3SiOTf) can be used.

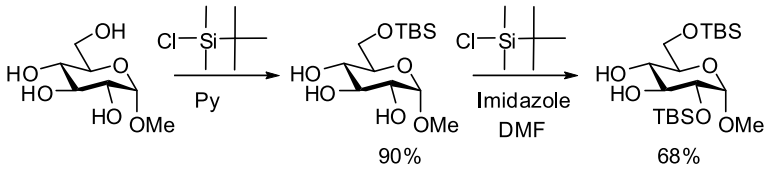
As a rule, the bulkier the substituents, the greater the stability of the resulting silyl derivatives. However, stability is not only a function of steric bulk since electronic effects play a role as well, which can be exploited to differentiate stability under acidic or basic conditions. For example, phenyl-substituted silyl ethers are equal or more reactive than their trimethylsilyl counterparts under alkaline conditions, but less reactive under acidic conditions. In general terms, however, the relative stabilities of the silyl-protected functional groups will follow the order of: $^{i}Pr_3Si$ > $ThMe_2Si$ > $^{t}BuPh_2Si$ > $^{t}BuMe_2Si$ > $^{i}PrMe_2Si$ > Et_3Si > Ph_2MeSi > Me_3Si [288].

On the other hand, silyl groups can migrate between different nucleophilic sites in a molecule under basic conditions. These migrations have to be considered as possible side reactions, and sometimes provide a valuable approach to interesting products that are not directly available [289]. For instance, in the synthesis of chemically modified cyclodextrines, the migration of the TBS groups from the 2-O to the 3-O on all the D-glucopyranose residues was observed during alkylation with sodium hydride in THF (*Scheme 50*) [290].

In general, the size of the substituent on the silicon atom is directly related to the rate of deprotection with smaller silyl substituents being more easily cleaved under acidic conditions. Similarly, if the same protecting group is used to protect two or more hydroxyl groups, the silyl ether derived from the less sterically encumbered alcohol is usually the first to be deprotected [291]. On the other hand, removal of silicon protecting groups occurs under extremely mild and highly specific conditions using a fluorine source. In general, the order of cleavage of silyl ethers with basic fluoride reagents (such as TBAF) parallels the order found for basic hydrolysis; similarly, slightly acidic fluorine-based reagents such as HF-acetonitrile parallel the order found for acid hydrolysis.



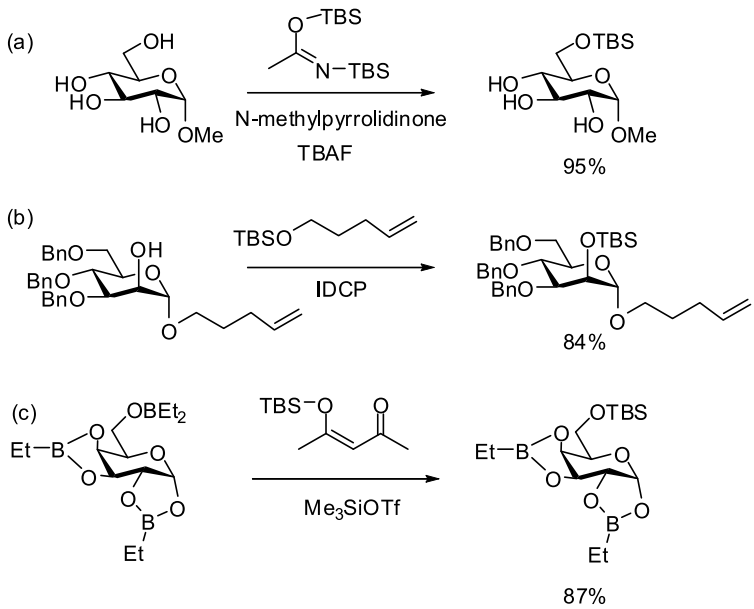
Scheme 50
Migrations of silyl groups in the synthesis of modified cyclodextrines



Scheme 51
Selective silylation of methyl α-D-glucopyranoside

2.5.2 *tert*-Butyldimethylsilyl (TBS or TBDMS) Group

Since its introduction in 1972 [292], the *tert*-butyldimethylsilyl group has become the most popular of the general purpose silicon protecting groups. It can be easily installed in high yields under mild conditions and it is robust to a variety of reaction conditions. The TBS group is commonly introduced via *tert*-butyldimethylchlorosilane, in the presence of basic activators such as DMAP or imidazole in a dipolar aprotic solvent such as DMF (*Scheme 51*). Hindered secondary alcohols can be silylated with TBSOTf using 2,6-lutidine as the base [293]. When the reaction is mediated by equimolecular amounts of dibutyltin oxide, the silylation with TBSCl gives the 6-monosilylated products in excellent yields [294].
N,*O*-Bis(*tert*-butyldimethylsilyl)acetamide silylates tertiary and hindered secondary alcohols in the presence of a catalytic amount of TBAF or another source of fluoride anion. Protection of primary hydroxyl groups in the presence of secondary ones is also possible (*Scheme 52a*) [295].



Scheme 52
Alternative conditions for the *tert*-butyldimethylsilyl ether protection

tert-Butyldimethylsilyl pentenyl ether is also a suitable reagent for efficient silylation of primary and secondary hydroxyl groups (● *Scheme 52b*). Activation is carried out with iodonium di-*sym*-collidine perchlorate (IDCP) and this procedure can be applied even to pentenyl glycosides [296].

An unusual way for the preparation of TBS ethers involves the reaction of diethylboronyl ethers, obtained by the reaction of the corresponding alcohol with BEt_3, with the TBDMS-enolate of pentane-2,4-dione in the presence of a catalytic amount of TMSOTf (● *Scheme 52c*) [297].

The palladium(0) nanoparticle-catalyzed silylation of sugars by silane alcoholysis of *tert*-butyldimethylsilane has been proposed as an attractive alternative to the established silyl chloride method. The methodology gives convenient access to the 3,6-silylated methyl glycopyranosides as the dominant products rather than the 2,6-silylated glycosides typically obtained by the silyl chloride method [298]. Changing to homogeneous cationic catalysts of iridium and rhodium, 2,3,6- and 2,4,6-trisilylated derivatives are obtained in synthetically useful yields [299].

The TBDMS group has also been introduced [300] to alcohols or phenols by the Mitsunobu reaction (DEAD/PPh_3, THF, −78 °C) using *tert*-butyldimethylsilanol.

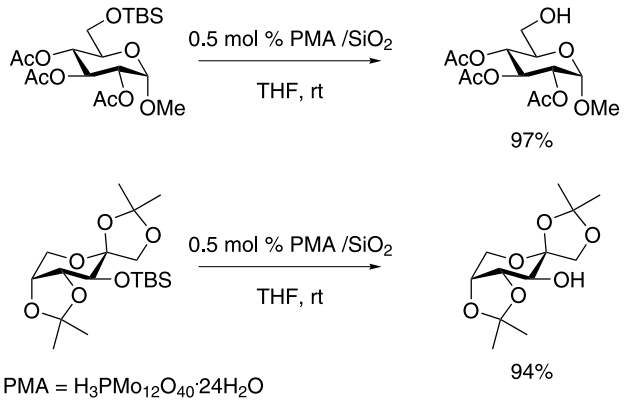
Numerous methods are now available in the literature for the deprotection of TBS ethers under a variety of conditions. One of the most effective ways for the cleavage of silyl ethers is based on the exploitation of the high affinity of silicon towards fluoride ions. Thus, a number of reagents involving one form of fluoride or another, such as tetrabutylammonium fluoride [292], $BF_3{\cdot}Et_2O$ [301], hydrofluoric acid [302], fluorosilicic acid [303], ammonium fluoride [304], silicon fluoride [305], lithium tetrafluoroborate [306], and chlorotrimethylsilane/potassium fluoride dehydrate [307] have been developed for the deprotection of TBDMS ethers. Among these, TBAF is most frequently used but the strong basicity of the fluoride anion makes it inappropriate for base sensitive functionalities.

Similarly, acidic reagents such as HCl [308], H_2SO_4 [309], PPTS [310], TFA [311], TsOH [312] etc., have also been employed for this purpose but cannot be used in the presence of acid-sensitive functionalities. This has led to the development of several Lewis acids and other reagents including $BF_3{\cdot}OEt_2$ [313], BCl_3 [314], $Sc(OTf)_3$ [315], $Ce(OTf)_4$ [316], $InCl_3$ [317], $ZnBr_2$ [318], $Zn(BF_4)_2$ [319], $CeCl_3$–NaI [320], $BiBr_3$ [321], $BiOClO_4$ [322], Cs_2CO_3 [323], CBr_4–MeOH [324], I_2 [325] and CAN [326] for desilylation.

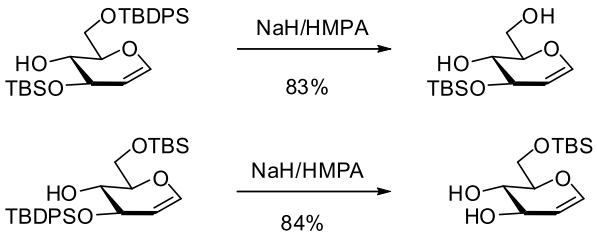
Recently, an environmentally benign phosphomolybdic acid supported on silica gel has been used for the chemoselective deprotection of TBS ethers in carbohydrate derivatives (● *Scheme 53*). The mild conditions are compatible with the presence of other protecting groups such as isopropylidene acetal, OTBDPS, OTHP, OAllyl, OBn, OAc, OBz, *N*-BOc, *N*-CBz, and *N*-Fmoc which are stable under the reaction conditions. Another advantage of this procedure is that the catalyst can be readily recovered and recycled [327].

2.5.3 *tert*-Butyldiphenylsilyl (TBDPS) Group

The TBDPS group was introduced by Hanessian and Lavallee in 1975 [328]. The TBDPS group has greater steric demands than the TBS group and, therefore can result in much more selective protections of hydroxyl groups. The group is also less prone to migrate to proximate hydroxyl groups under neutral or acidic conditions than the TBS group but it may migrate under basic conditions [329].



Scheme 53
The ring oxygen is missing in the fructose derivatives



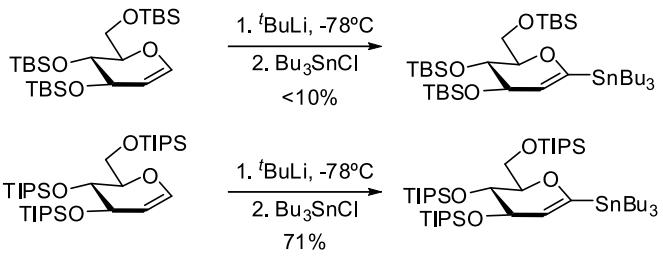
Scheme 54
Selective removal of TBDPS ethers in the presence of TBS ethers

The TBDPS ethers are prepared by treating alcohols with TBDPS-Cl in DMF in the presence of imidazole (the primary hydroxy group reacts faster than the secondary one). Tertiary alcohols do not silylate. The silylation of hindered alcohols is greatly accelerated with the aid of $AgNO_3$, NH_4NO_3, or NH_4ClO_4 [330].

TBDPS ethers are generally cleaved under the same acidic conditions as those used for TBS ethers but longer reaction times are necessary and consequently selective removal of TBS groups in the presence of TBDPS groups is very common [331,332]. The electron-withdrawing effect of the phenyl substituents enhances the electrophilicity of the silicon atom and therefore is more susceptible towards nucleophiles. For this reason it is possible to reverse the tendency of TBS ethers to cleave more easily than TBDPS ethers using ion fluoride or basic hydrolysis [333]. Some examples in glycal derivatives are shown in *Scheme 54*.

2.5.4 Triisopropylsilyl (TIPS) Group

The TIPS group [334] is one of the most sterically hindered silyl protecting groups, being removed only slowly under standard acid- or base-catalyzed hydrolysis conditions. The large steric bulk ensures high selectivity in the protection of primary hydroxyl groups over sec-

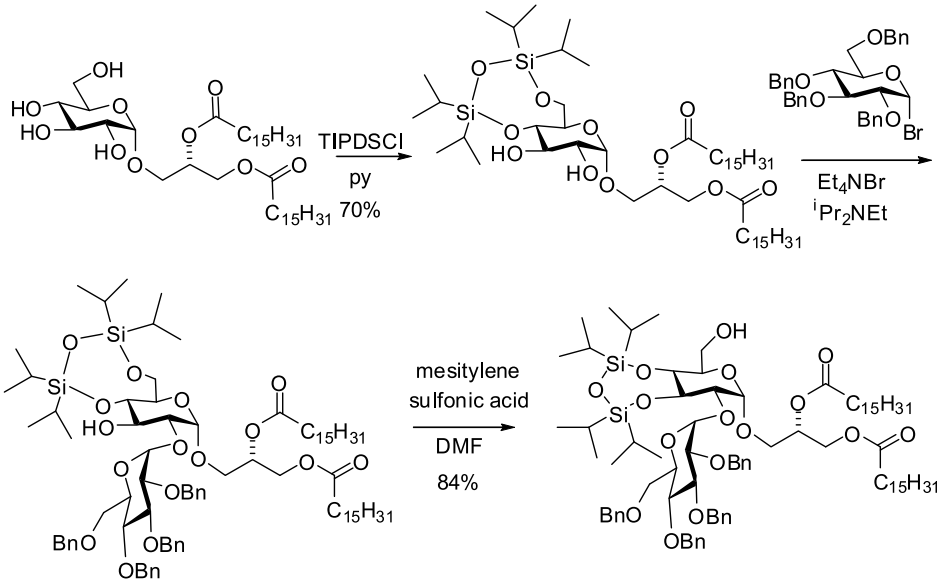


Scheme 55
TBS- vs. TIPS- ethers as protecting groups in the anomeric lithiation of glycals

ondary and valuable stability under a wide range of reaction conditions. It is noteworthy that TIPS groups are inert towards powerful bases such as *tert*-butyllithium, and therefore can be used as protecting groups in the anomeric lithiation of glycals (*Scheme 55*) [335].
The TIPS group is usually introduced from triisopropylchlorosilane [336], but protection of hindered alcohols can be very slow in which case triisopropylsilyl triflate in the presence of 2,6-lutidine is used [293].
TIPS ethers are cleaved under the same conditions as those used for TBS ethers but longer reaction times are frequently necessary; consequently TBS ethers can be removed selectively in many cases.

2.5.5 1,1,3,3-Tetraisopropyldisiloxane (TIPDS) Group

The tetraisopropyldisiloxane-1,3-diyl group was introduced by Markiewicz et al. for simultaneous protection of the 3′- and 5′-hydroxy groups of ribonucleosides [337]. The group is usually introduced by the reaction of the bifunctional reagent 1,3-dichloro-1,1,3,3-tetraisopropyldisiloxane with the substrate in pyridine though imidazole in DMF solution can also be used. When applied to pyranoses, these conditions give the kinetic product, an 8-membered ring, which is formed by rapid reaction first at the least hindered hydroxyl group followed by a second intramolecular silylation with the next proximate hydroxyl at C-4.
Additionally, the eight-membered rings of TIPDS-acetals formed in this way can rearrange under the influence of acidic catalyst to the thermodynamically more stable seven-membered derivatives bridging two vicinal secondary hydroxyl functions [338].
The usefulness of the TIPDS protecting group in carbohydrate chemistry is well illustrated by the synthesis of a glyco(phospho)lipid of *Streptococci* cell membranes (*Scheme 56*). Selective protection of the C-4 and C-6 hydroxyl groups of the pyranose was easily accomplished using the TIPDS group. Then the C-2 hydroxyl participated in a regioselective glycosylation under basic conditions to give the coupling product. At this stage of the synthesis the dynamic properties of the TIPDS group were exploited and the subsequent acid-catalyzed isomerization of the 4,6-O-disilyl-protected product results in the formation of the more stable 3,4-*O*-disiloxane. The freed primary hydroxyl function was then ready to be reacted with stearoyl chloride after which the naturally occurring glycolipid was eventually obtained [339].



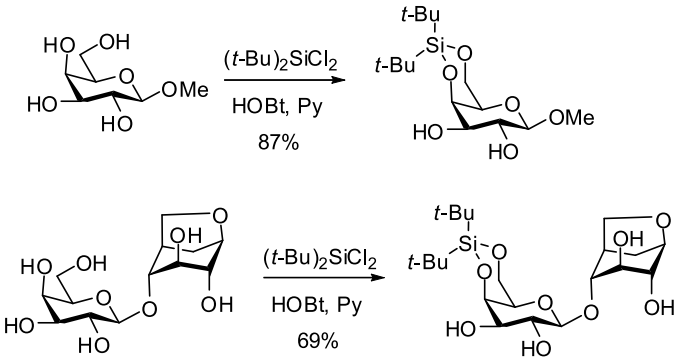
Scheme 56
Application of TIPDS protecting group in the synthesis of a glyco(phospho)lipid

However, despite the attractive properties, which are inherent in the use of the TIPDS protecting group in sugar chemistry, its general applicability is limited, to some extension by the fact that this group can not withstand acidic conditions, which are commonly used in carbohydrate chemistry.

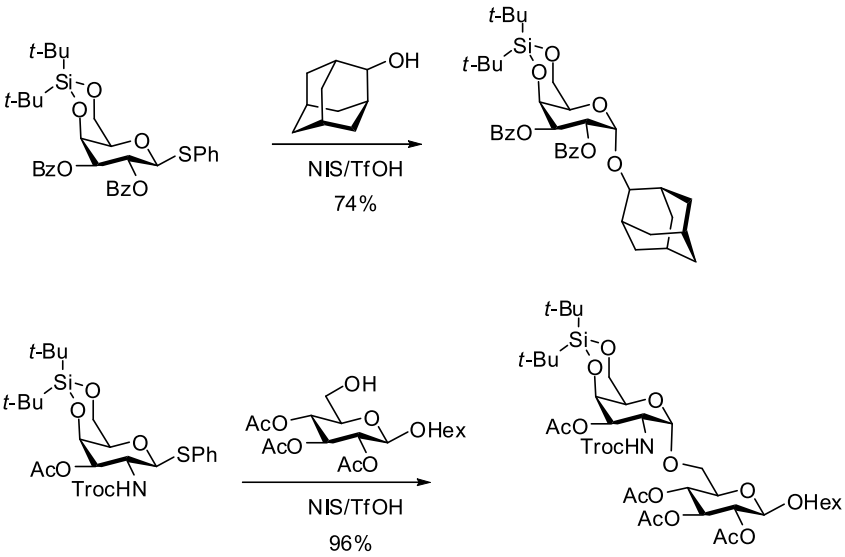
2.5.6 Di-*tert*-butylsilylene (DTBS) Group

The di*tert*butylsilylenediyl is a convenient and versatile protecting group introduced by Trost [340] and often used for the synthesis of anthracyclines [341] and nucleotide [342] derivatives. The DTBS group is not as robust as isopropylidene or benzylidene acetals and therefore its use is appropriate for systems requiring deprotection under very mild conditions. DTBS derivatives survive hydroboration, mild oxidation, Lewis acids, mild protic acids, and strong bases but hydrolysis occurs readily with HF in acetonitrile, HF·pyridine complex, or TBAF.

The formation of the silylene derivatives is effected by treatment of the diol with di*tert*butyldichlorosilane in the presence of 1-hydroxy-benzotriazole (HOBT) (*Scheme 57*) [340]. Di*tert*butylsilyl ditriflate and 2,6-lutidine effects silylene formation faster and under milder conditions than the less-reactive dichloride [343]. When the silylation is carried out in pyranoside derivatives, the reaction proceeds selectively at 1,3-diol groups of C-4 and C-6 positions and the formation of the five-membered DTBS derivatives of 1,2 diols was not observed [344]. It has been shown that a DTBS group at the *O*4-*O*6 position of a galacto-type sugar directs α-predominant selective glycosylation in spite of the presence of a participatory group at C-2 such as benzoyl or Troc groups [345] (*Scheme 58*). This new glycosylation method is a powerful strategy for the synthesis of α-galactosyl and galactosaminyl glycans [346,347].



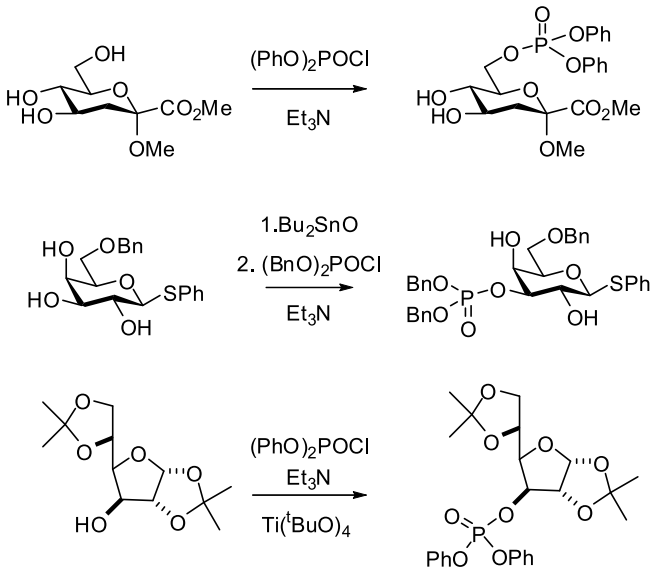
Scheme 57
Cyclic di-t-butysilylene as protecting group



Scheme 58
Examples of di-t-butysilylene-directed α-galactosylation

2.6 Phosphorylation Reactions

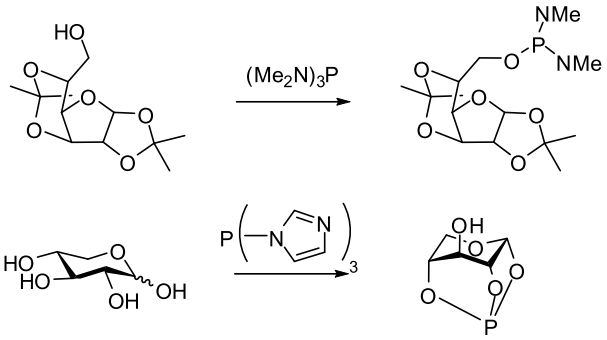
Phosphate esters play an important role in a wide variety of structurally diverse natural and biologically active compounds such as glycolipids, nucleic acids, nucleotides, proteins, coenzymes, steroids, and in particular carbohydrates. Introduction of a phosphate group essentially changes the physical and chemical properties of the parent molecule, resulting in changes to the polarization and intermolecular bonding characteristics of that molecule. Given the importance of this functional group it is not surprising that many methods have been developed for the phosphorylation of alcohol functions [348,349]. Both chemical and enzymic methods are available for the synthesis of specific phosphates.



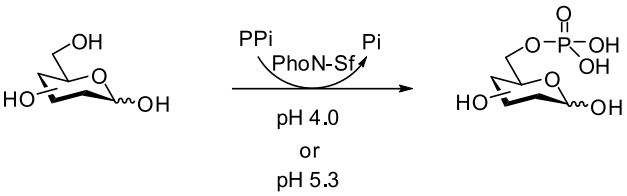
Scheme 59
Examples of chemical phosphorylation with phosphorous (V) reagents

Chemically, the most common phosphorylation reagents used are chlorophosphates [350]. These compounds are generally commercially available and are as stable as their routinely used acyl chloride counterparts to both air and moisture. The problem most commonly encountered with the use of such reagents is the conditions under which they will react. Phosphorylation is usually performed either through formation of the lithium [351] or thalium alkoxide [352], followed by the reaction with the chlorophosphate or simply by use of a proton scavenger such as pyridine [353] or Et_3N [354]. Alternatively, nucleophilic catalysis with DMAP [355] or tin-mediated phosphorylation [356] may be employed. A method for the phosphorylation of hydroxyl groups using a Lewis acid catalyst has been recently reported [357] (*Scheme 59*). Solid-phase phosphorylating reagents have been used for the phosphorylation of unprotected nucleosides and carbohydrates [358,359]. These procedures exhibit high regioselectivity and only one monophosphorylated product is obtained. Carbohydrate and nucleoside diphosphates have also been synthesized by using solid-phase reagents [360].

Aside from P^V reagents, the most widely used and most successful of all chemical phosphorylation techniques is the use of reagents containing trivalent phosphorous which ensure the highest phosphorylation rates and permits one to avoid many side processes [361]. This methodology has been well developed and is used extensively in the construction of oligonucleotides. Phosphorous triamides phosphorylate efficiently monosaccharides whose molecules contain one free alcoholic hydroxyl [362] (*Scheme 60*). Phosphorylation by dialkyl [363] or alkanediyl phosphoramidites [364], phosphonamidites [365] and phosphinamidites [366] follows a similar pathway. Treatment of monosaccharide derivatives whose molecules contain two closely located hydroxy groups with phosphamides results in cyclophosphorylation [367].



Scheme 60
Examples of chemical phosphorylation with phosphorous (III) reagents



Scheme 61
Example of enzymatic phosphorylation

A non-specific bacterial acid phosphatase from *Shigella flexneri* (PhoN-Sf) has been screened for regioselective phosphorylation of primary alcohol(s) of more than 20 different cyclic and acyclic monosaccharides using pyrophosphate as the phosphate donor (*Scheme 61*) [368]. These studies have shown that PhoN-Sf is capable of phosphorylating a range of hexoses (D-glucose epimers, glycosides, and C-2 derivatives), pentoses, heptoses, ketoses, and acyclic carbohydrates.

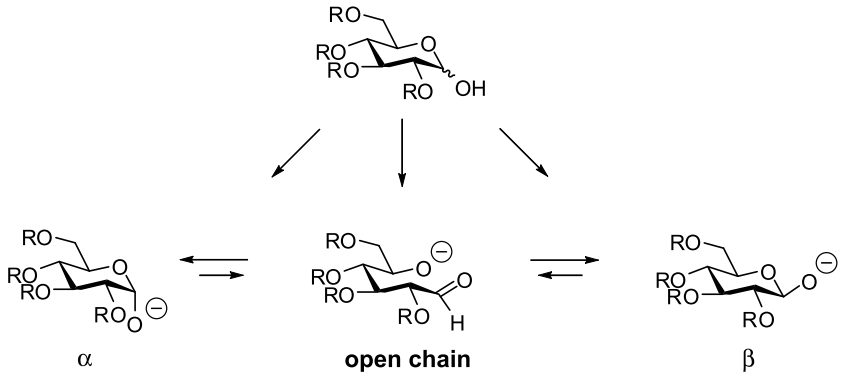
3 Reactions at the Anomeric Hydroxyl

3.1 Alkylation Reactions

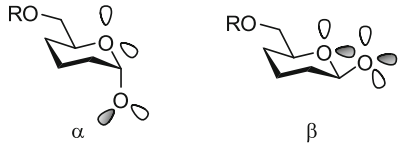
In general, reactions and conditions for the introduction of protecting groups in the anomeric hemiacetal group are the same as those mentioned previously for non-anomeric hydroxyl groups. Probably, the only exception relates to the alkylation reaction since alkyl ethers of the anomeric hydroxyl group, which are acetals rather than ethers, are normally formed under Fischer glycosylation conditions using the alcohol as the aglycon [369]. This process involves cleavage of the C-1–O-1 bond at the anomeric center and, therefore will not be treated here, but it is the method of choice for the preparation of alkyl glycosides.

3.1.1 Anomeric *O*-Alkylation and *O*-Arylation

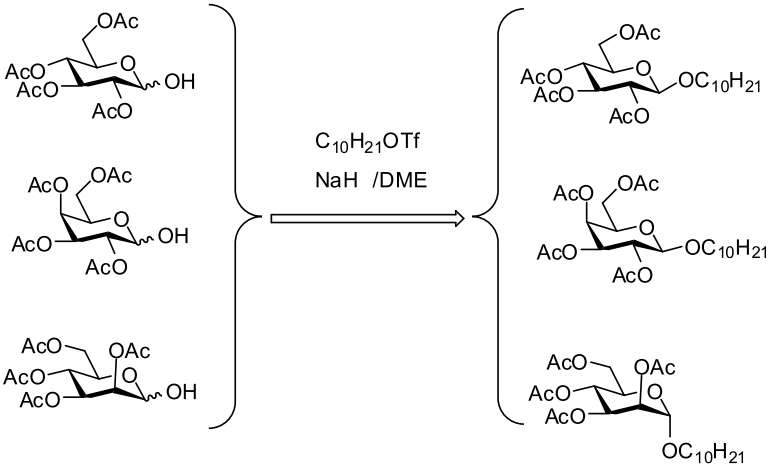
The 1-*O*-alkylation of carbohydrates with simple alkylating agents, particularly methyl iodide and dimethyl sulfate, has long been known [370,371,372]. The reactivity of pyranoses and furanoses deprotonated at *O*-1 is, thus, analogous to that of alkoxides. Alkylation of fully protected pyranoses, due to the ring chain tautomerism between the two anomeric forms α and β and the open chain form (● *Scheme 62*), can take place at three different sites [373].
However, when the alkylation of 2,3,4,6-tetra-*O*-benzyl D-glucose is carried out in dioxane with sodium hydride and methyl triflate, the β-glucoside was obtained practically exclusively [374]. This selectivity has been explained on the basis of an enhanced nucleophilicity of the β-oxide atom which can be attributed to a steric effect in combination with a stereoelectronic effect resulting from repulsion of the lone electron pairs (kinetic anomeric effect) in the β-oxide [375] (● *Fig. 6*). Conversely, if the reaction is carried out at lower temperatures (−40 °C) the formation of α-anomer is preferred. Despite of the use of NaH, neither acyl migration nor orthoester formation occurred during the 1*O*-alkylation of acetyl-protected derivatives (● *Scheme 63*) [376,377]. The stereoelectronic effects in α- and β-furanosyl oxides should differ less for conformational reasons and the stereocontrol results primarily from steric and chelation effects.
The higher acidity of the 1-OH group of the hemiacetal (resulting from the indirect stabilization by the ring oxygen atom) allows for regioselective *O*-alkylation at this position regardless of the presence of other sugar hydroxy groups. Thus, as shown in ● *Scheme 64*, the alkyla-



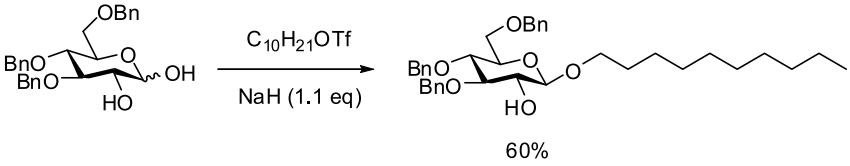
● Scheme 62
Ring chain tautomerism of fully protected pyranoses



● Figure 6
Kinetic anomeric effect in the β-oxide



Scheme 63
Stereoselective anomeric *O*-alkylation of acyl-protected sugars



Scheme 64
Anomeric *O*-alkylation of partially protected pyranoses

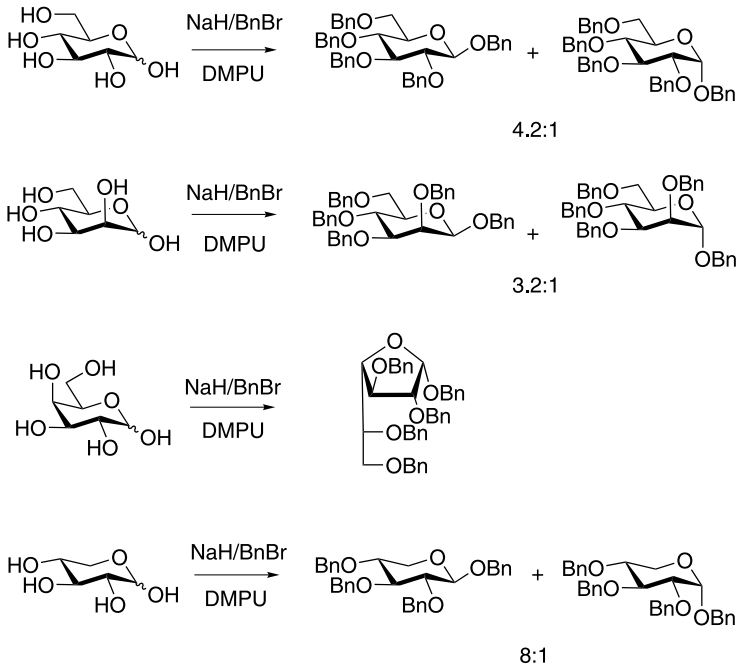
tion of 2-*O*-unprotected tribenzylglucopyranose afforded exclusively the 1-*O*-alkylated product when the reaction was done with one equivalent of NaH, whereas the undesired 1,2-disubstituted isomer was also obtained when two equivalents of NaH were used [378].

Although ring-chain equilibration permits the formation of many products in fully unprotected monosaccharides, the regiocontrol in the per-*O*-benzylation towards uniform glycoside bond formation is generally very high [379,380] (› *Scheme 65*).

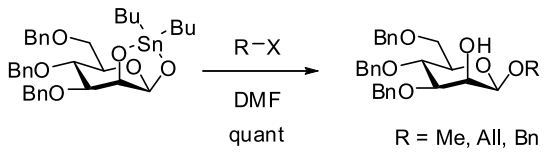
The 1-*O*-alkylation of pyranoses has also been used for glycosidic bond formation [381].

An alternative activation of the anomeric hydroxyl makes use of 1,2-*O*-dibutylstannylene acetals [382]. Thus, for instance condensation of the stannylene acetal of 3,4,6-tribenzylmannose with methyl iodide, allyl or benzyl bromide afforded the corresponding β-mannosides in almost quantitative yields (› *Scheme 66*).

On the other hand, carbohydrates carrying an aromatic aglycon are important natural products and therefore methods for the arylation of anomeric hemiacetals have also been developed. Both Mukaiyama [383] and Smith [384] have synthesized aryl glycosides by nucleophilic aromatic substitution for use as glycosyl donors. The method is quite efficient but requires activation by electron-withdrawing groups in the aromatic counterpart. Thus, direct reaction of 1-fluoro-2,4-dinitrobenzene with the 1-OH group of the hemiacetal gave 2,4-dinitroglycosides in excellent yields (› *Scheme 67a*). In the case of dinitrosalicylic (DISAL) acid



Scheme 65
Two bonds are missing in the first compound



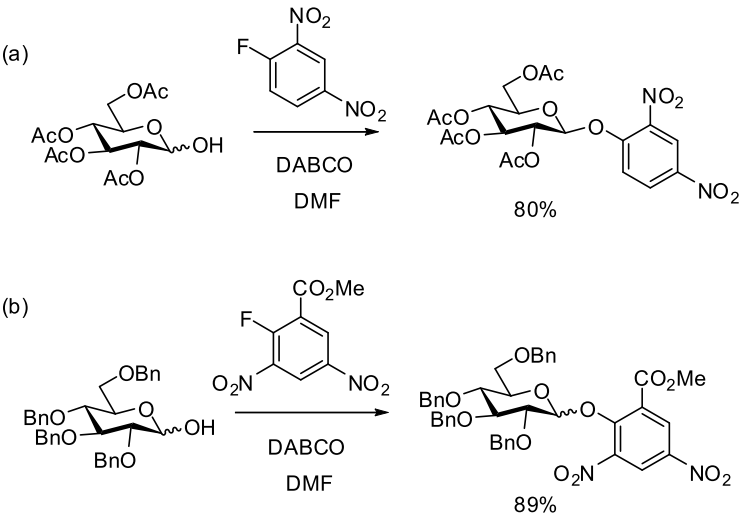
Scheme 66
Anomeric *O*-alkylation via 1,2-*O*-stannylene acetals

derivatives, the use of DMAP as the base gave an α/β ratio similar to the starting 1-OH derivative. In contrast, formation of the β-anomer was favored using 1,4-dimethyl piperazine (*Scheme 67b*) [385,386,387].

The reverse situation, in which the phenol acts as the nucleophile attacking activated carbohydrate hemiacetals, has also found several practical applications preparing *O*-aryl glycosides [388]. However, this situation implies an attack at the anomeric carbon and not at the anomeric oxygen.

3.1.2 Anomeric *O*-Dealkylation

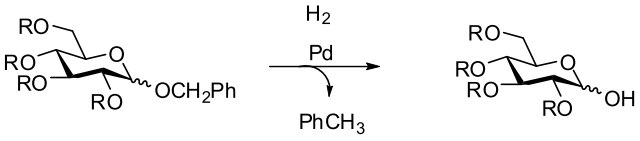
On treatment with aqueous acid, glycosides are hydrolyzed to give the corresponding alcohol and the reducing sugar. Solvolysis of glycosidic bonds is one of the most general and important reactions in carbohydrate chemistry and so the literature on the acid hydrolysis of *O*-glycosidic



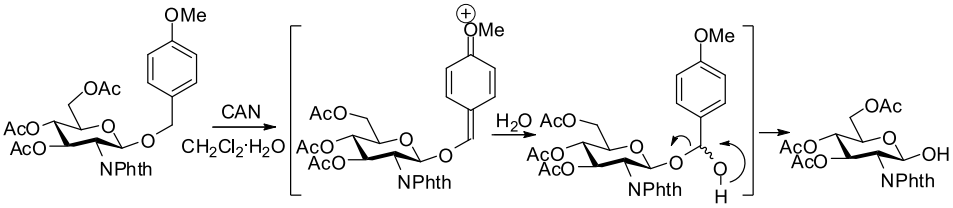
Scheme 67
Anomeric *O*-arylation via nucleophilic aromatic substitution

bonds covers thousands of titles. The data on more detailed studies of this reaction are covered by several reviews [389,390,391,392]. Although this process is formally the reverse direction of the alkylation reaction, the key step involves reaction at the anomeric carbon rather than at the anomeric oxygen and thus is beyond the scope of this chapter.
Nevertheless, a variety of protecting groups have been applied to the anomeric center, which are synthetically useful and provide alternative ways for the selective liberation of the anomeric oxygen. These include the following:

Benzyl Glycosides In 1928, Freudenberg found that benzyl ethers of sugars were cleaved by hydrogenolysis with sodium amalgam and by catalytic hydrogenolysis that could be effected in acetic acid in the presence of platinum metals [393]. On palladium catalysis, hydrogen splits off the benzyl β-D-glycosides, at room temperature and atmospheric pressure to afford toluene and the reducing sugar (*Scheme 68*) [394]. Hydrogenolysis is commonly carried out using hydrogen gas with a palladium catalyst absorbed on charcoal although modifications involving hydrogen transfer have been used.



Scheme 68
Hydrogenolysis of benzyl glycosides



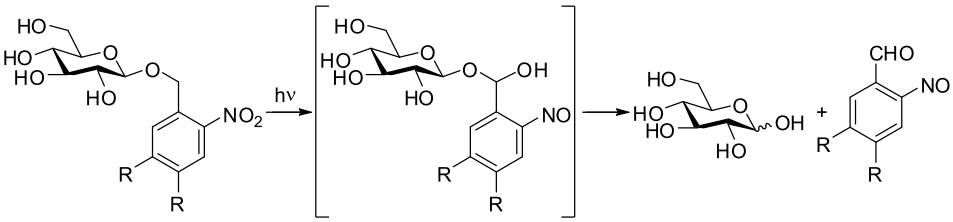
Scheme 69
Oxidative deprotection of p-methoxybenzylglycoside

A remarkable rate difference in the hydrogenolysis of α- and β-benzyl D-gluco-and D-galactopyranosides has been reported, with the β-anomers being more readily cleavable [395]. Ferric chloride has been employed for anomeric debenzylation in oligosaccharides [18].

Methoxy-substituted benzyl glycosides have been used as precursors for reducing sugars [396]. As mentioned in *Sect. 2.1.3* under *"p-Methoxy Benzyl (PMB) Ethers"*, their utility lies in the fact that they are more readily cleaved oxidatively than the unsubstituted benzyl ethers (*Scheme 69*). All these transformations have great synthetic value although the process is not regioselective since it is operational for all the benzyloxy groups present in the sugar.

Particularly useful in protecting the anomeric hydroxyl are those substituted benzyl groups that are light-sensitive. Such groups are stable to a wide variety of chemical treatments and at the same time are sensitive to irradiation under conditions that leave other functional groups in the molecule unaffected. Thus, 2-nitro benzyl and 3,4-dimethoxy-6-nitrobenzyl (6-nitroveratryl) glycosides are more stable to acid hydrolysis than are the corresponding benzyl glycosides but are readily photolyzed at 320 nm to the reducing sugars in high yields (*Scheme 70*) [397]. In this context photocleavable linkers for solid-phase synthesis, based on the lability of 2-nitro benzyl moieties under irradiation, have been applied to the liberation of the anomeric center in oligosaccharides [398].

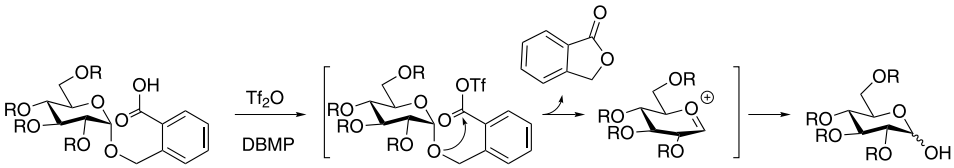
Other substituted-benzyl glycosides that can be selectively removed in the presence of unsubstituted benzyl ethers have been developed. For example, 2-(hydroxycarbonyl)benzyl glycosides are easily sovolyzed by treatment with Tf_2O in the presence of di-*tert*-butylmethylpyridine (DBMP). The reaction implies anomeric C–O bond cleavage since it takes place by lactonization via the mixed anhydride to generate phthalide and the oxocarbenium ion (*Scheme 71*) [399].



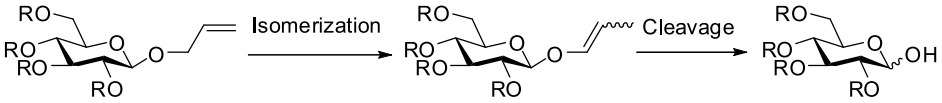
R= H, OMe

R= H, OMe

Scheme 70
Example of light-sensitive glycosides: 2-nitrobenzyl and 6-nitroveratryl glycosides

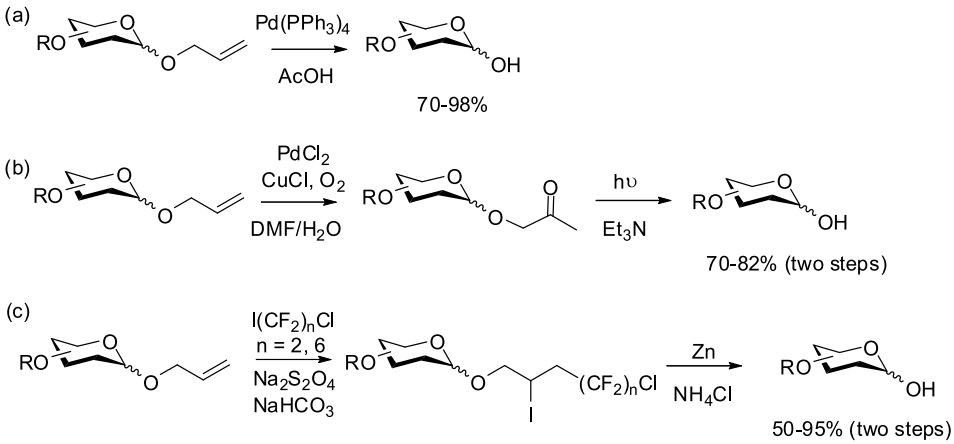


Scheme 71
Solvolysis of 2-(hydroxycarbonyl)benzyl glycosides

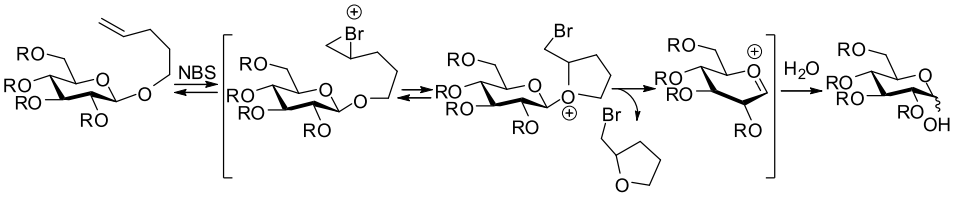


Scheme 72
Usual method for the deprotection of allyl glycosides

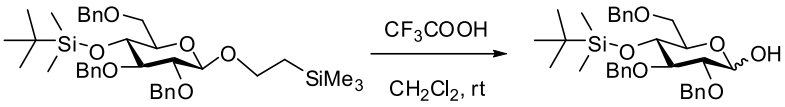
Allyl Glycosides Another commonly used protecting group for the anomeric oxygen is the allyl group [400,401]. The most usual method for deblocking allyl glycosides involves the two-step reaction in which the allyl group is first converted into the more labile propenyl group and then is cleaved under mildly acidic conditions (see ❯ *Sect. 2.1.4*, ❯ *Scheme 72*).
Alternative methods for the deprotection of the allyl group at the anomeric position include $Pd(PPh_3)_4$/AcOH [402], in which the reaction proceeds by the formation of a π-allyl complex, or $PdCl_2/CuCl/O_2$ followed by photolysis in the presence of triethylamine [403]. Perfluoroalkylation with perfluoroalkyl iodide under sodium dithionite and sodium bicarbonate followed by elimination in the presence of zinc powder and ammonium chloride has also been disclosed as an efficient procedure for deprotection of the anomeric allyl group of carbohydrates (❯ *Scheme 73*) [404]. The reaction goes through the intermediacy of a radical addition of a perfluoroalkyl iodide to the double bond followed by Zn-mediated reductive β-elimination.



Scheme 73
Alternative methods for the deprotection of the allyl group at the anomeric position



Scheme 74
Deprotection of n-pentenyl glycosides



Scheme 75
Deprotection of 2-(trimethylsilyl) ethyl glycosides

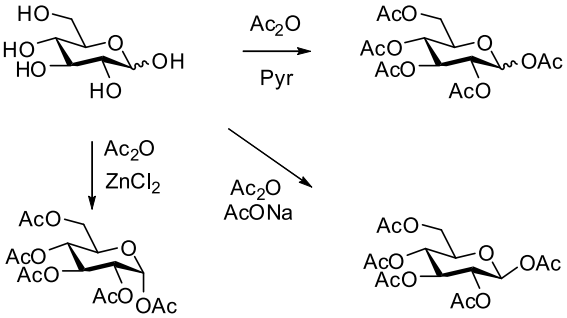
n-Pentenyl Glycosides In 1988 Fraser-Reid and Mootoo reported the NBS-mediated reaction of *n*-pentenyl glycosides, in the presence of water to yield reducing monosaccharides (*Scheme 74*) [405]. This transformation proved to be highly chemoselective leaving a wide variety of other functional groups unaffected.
The reaction takes place with cleavage of the anomeric C–O bond by electrophilic addition to the olefin followed by intramolecular displacement by the ring oxygen and eventual expulsion of the pentenyl chain, in the form of a halomethyltetrahydrofuran, to form an oxonium species (*Scheme 74*). Trapping with water then leads to the reducing sugar. This transformation has also been extended to the use of *n*-pentenoyl esters [406,407].

2-(Trimethylsilyl) Ethyl Glycosides An alternative procedure for protecting the anomeric center is based on the use of 2-(trimethylsilyl) ethyl glycosides [408,409]. Lipshutz et al. first found that $LiBF_4$ in CH_3CN caused the deblocking of the anomeric center [408], although extensive experimentation led Magnusson and coworkers to report on the use of trifluoroacetic acid in dichloromethane as the most effective reagent to carry out the same transformation (*Scheme 75*) [409]. The reaction conditions are fully compatible with most of the normally used protecting groups, including silyl ethers and therefore this protecting group has found wide application in the synthesis of complex oligosaccharides.

3.2 Acylation Reaction

3.2.1 Anomeric *O*-Acylation

Free sugars, since they are polyhydroxy aldehydes or ketones, can be acylated through their hydroxyl groups (including the anomeric hydroxyl group) to give esters. However, the unusual property of the anomeric center to be a mixed function (ester and acetal) confers the glycosyl esters a special reactivity. Acetylation of unprotected sugars is complicated by the fact that they exist in solution as equilibrium mixtures of tautomers. The isomer obtained depends on



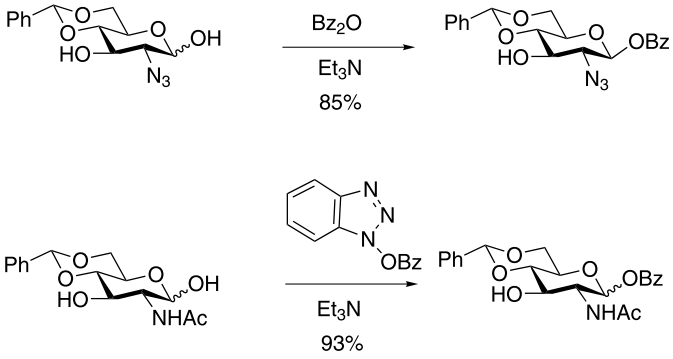
Scheme 76
Acetylation of D-glucopyranose

the catalyst used and on the temperature. For example, acetylation of pure α- and β-D-glucopyranoses with Ac_2O and pyridine at 0 °C occurs with retention of the configuration at the anomeric carbon [410], whereas acetylation of a mixture of α- and β-D-glucopyranoses in the presence of an acid catalyst (Ac_2O, $ZnCl_2$) takes place with predominant formation of the thermodynamically preferred (anomeric effect) α-D-glucopyranose pentaacetate, due to acid-induced anomerization. Conversely, β-D-glucopyranose pentaacetate, is formed preferentially when the acetylation is carried out in the presence of sodium acetate at higher temperatures (Ac_2O, NaOAc, Δ), a fact which has been explained on the basis of a lower rate for acetylation when compared with mutarotation together with a preferential reactivity for the equatorial anomeric hydroxyl groups (*Scheme 76*). For acetylation of ketoses low temperature acidic catalysts are preferred.

The ring size of the cyclic acetates formed under common acetylation procedures is normally pyranoid although sugars that form relatively stable furanose rings give more complex mixtures. D-Galactose, for instance, in the presence of sodium acetate or pyridine at elevated temperatures gives appreciable amounts of furanose acetates [411].

Sugar benzoates have also been widely used since they are easy to prepare and more stable than the corresponding acetates. Benzoyl chloride in pyridine is the reagent of choice to carry out this transformation [412], and benzoylation in hot pyridine may lead to the isolation of glycofuranose benzoates [413]. More recently, a new method for the benzoylation of alcohols has been described using TMEDA as a base, which gave the expected benzoates in excellent yields [414]. In 2-*N*-protected 4,6-*O*-ketal derivatives of D-glucosamine a highly regio- and stereoselective acylation of the anomeric hydroxyl groups is possible using 1*N*-benzyloxy-1,2,3-benzotriazol (BzOBT) or benzoic anhydride and triethylamine as a base [415,416]. Because of the kinetic stereoelectronic effect or 1,3-diaxial repulsion, the *O*- is oriented in the equatorial position and only the β-anomer is formed (*Scheme 77*).

When positions other than 1-OH are fully protected, anomeric acylation can be carried out by the usual methods for the esterification of alcohols (*Scheme 78*). For example, carbodiimide-mediated coupling [417,418] was the method used for the preparation of glycosyl benzyl phthalates [419] or *n*-pentenoyl esters [406,407] from the corresponding 1-OH sugars (*Scheme 78a*). Combination of carbodiimides, active ester-forming reagents, and base catalysts have been studied for the selective acylation of monoprotected glucuronate esters and the uranium reagent HATU [420] has been found to be the reagent of choice [421].



Scheme 77
The 1-OH shall be 1-OBz in the lower right structure

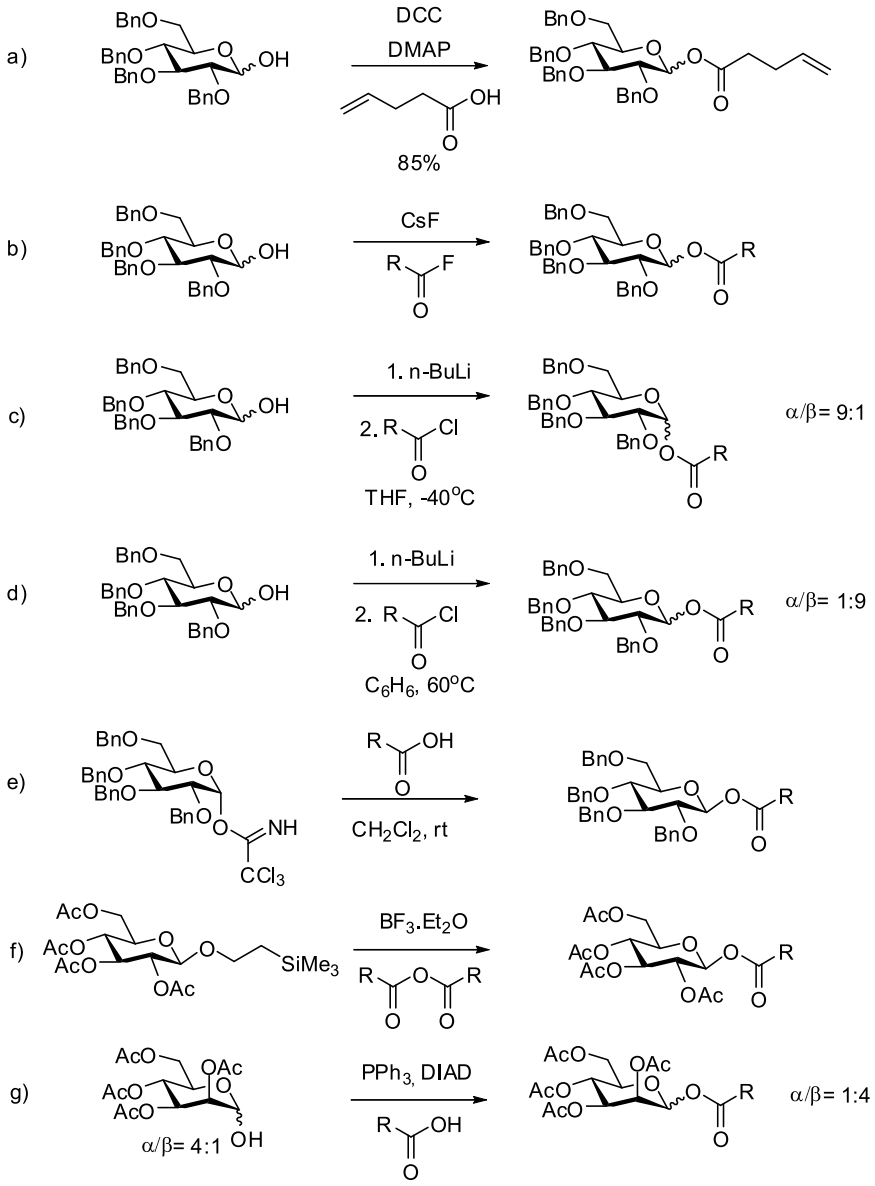
1-OH sugars react with acyl fluorides in the presence of cesium fluoride to furnish the corresponding glycosyl esters under essentially neutral conditions, with the α/β ratio being affected by changes in the order of addition of the reagents (*Scheme 78b*) [422].

Acylation of the lithium salt of 1-OH sugars allows complete stereocontrol in the formation of glycosyl esters [423,424]. Metalation of 2,3,4,6-tetra-*O*-benzyl-α-D-glucopyranose, in tetrahydrofuran at $-40\,°C$ with 1.1 equiv. of *n*-BuLi, followed by acylation with acid chlorides produce mainly or exclusively α-glucosyl esters (*Scheme 78c*). Increasing the reaction temperature and changing the solvent to benzene led to the preferred formation of β-glucosyl esters (*Scheme 78d*). Analogously, tributylstannyl alkoxides can be used in place of the corresponding lithium salts [425].

2-Acylthio-3-nitropyridines, prepared from the corresponding carboxylic acids, have been used as acylating agents when the corresponding acid chlorides are unstable [426].

As an alternative to the direct esterification of the anomeric hydroxyl group, glycosyl esters have been prepared by displacement of a good leaving group at the anomeric position, although in that case the key step involves reaction at the anomeric carbon rather than at the anomeric oxygen. In this context, the direct glycosylation of trichloroacetimidates [427] with carboxylic acids is a particularly advantageous method (*Scheme 78e*) [428]. This reaction, which involves inversion of the configuration at the anomeric center, is the method of choice for the stereoselective preparation of β-acyl-glycosides. The requisite trichloroacetimidates can be selectively produced from the corresponding hemiacetals under thermodynamically controlled conditions [429]. Analogously, 2-(trimethylsilyl) ethyl glycosides have also been transformed into the corresponding 1-*O*-acyl sugars by reaction with the appropriate anhydride in the presence of $BF_3{\cdot}Et_2O$ (*Scheme 78f*) [430].

The Mitsunobu protocol has also been investigated in the stereocontrolled synthesis of glycosyl esters (*Scheme 78g*) [431]. Complete stereochemical inversion at C-1 of the starting sugar is observed when the esterification is conducted with anomerically pure glycosyl hemiacetals. By corollary, complementary ratios of inverted products are formed when an anomeric mixture of sugars is esterified. The stereochemical outcome of the esterification is not affected



Scheme 78
Acylation reactions of fully protected pyranoses

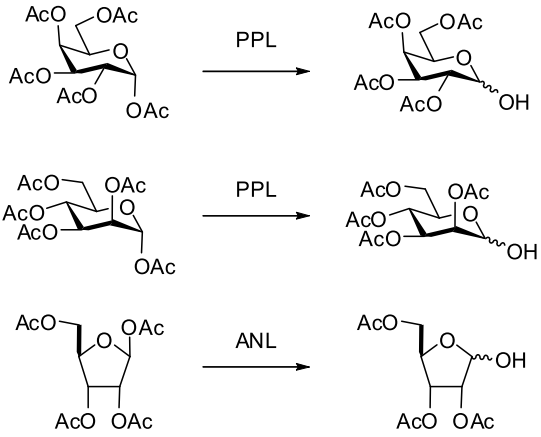
by anchimeric assistance from acyl groups at C-2. Accordingly, 2,3,4,6-tetra-*O*-acetyl-D-mannose furnishes a mixture of 1-*O*-benzoates in which the β-anomer predominates, a result that is especially significant in view of the difficulties generally encountered in obtaining β-glycosides of D-mannose.

3.2.2 Anomeric *O*-Deacylation

Deprotection of the anomeric acyl group in acylated sugars can be effected in a number of manners including chemical and enzymatic methods.

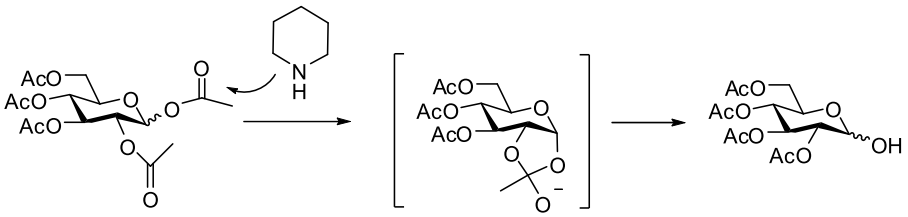
Enzymic Deacylation Both furanose and pyranose sugars can be efficiently deacetylated by suitable lipases under proper reaction conditions. The removal of the 1-*O*-acetyl group of glucose pentaacetate by *Aspergillus niger* lipase was reported after 20% conversion [432]. More recently it was found that the regioselectivity could be enhanced in the pyranose case by the presence of DMF [433]. Porcine pancreatic lipase in 10% DMF exclusively cleaved glucose pentaacetate ester at C-1 (70% isolated yield), and similar selectivities (and yields) were obtained for several peracetylated hexopyranoses. Peracetylated furanoses were deacylated at C-1 by the use of the lipase from *Aspergillus niger*. Finally, peracetylated reducing disaccharides have been specifically hydrolyzed at the anomeric center with a lipase from *Aspergillus niger* (Lipase A Amano 6) in a mixture of organic solvents and phosphate buffer (*Scheme 79*) [434].

Chemical Deacylation Several methods for the regioselective 1-*O*-deacylation of carbohydrates have been reported. Most of them involve regioselective nucleophilic attack upon the carbonyl group at *O*-1 thus liberating the anomeric oxygen. Nitrogen-containing nucleophiles have been widely used in this transformation: piperidine [435], hydrazine acetate [436], and hydrazine hydrate [437], have been reported to selectively hydrolyze anomeric acetates in peracetylated disaccharides (*Scheme 80*). Hydrazine acetate in DMF [436], benzylamine in chloroform [438], hydrazine hydrate in pyridine [439], ammonia in an aprotic solvent [440], and 2-aminoethanol [441], have been used to regioselectively 1-*O*-deacylate per-*O*-



PPL = Porcine pancreas lipase
ANL = *Aspergillus Niger* lipase

Scheme 79
Enzymic deacylation of acyl-glycosides



Scheme 80
Chemical deacylation of acyl-glycosides

acylaldoses. Other reagents used include potassium hydroxide [442], potassium cyanide [442], sodium methoxide [443], bis(tributyltin)oxide [442,444], tributyltin methoxide [444], ammonium carbonate [445], ammonium acetate [446], and mercuric chloride/mercuric oxide [447]. Heterogeneous anomeric deacetylation has also been reported by the use of magnesium oxide in methanol [448], or silica gel in methanol [449].

Finally, acid-catalyzed solvolysis of per-*O*-acyl hexopyranoses ($SnCl_4$, CH_3CN, H_2O) is an efficient method for removal of the anomeric acetyl group [450]. In this case the reaction takes place by cleavage of the C-1–OAc bond [451]. This reaction proceeds in 1 h at room temperature for sugars containing 1,2-trans-acetoxy groups and at 40 °C for 1,2-*cis* acylated pyranoses, and confirms the anchimeric assistance provided by the ester group at C-2.

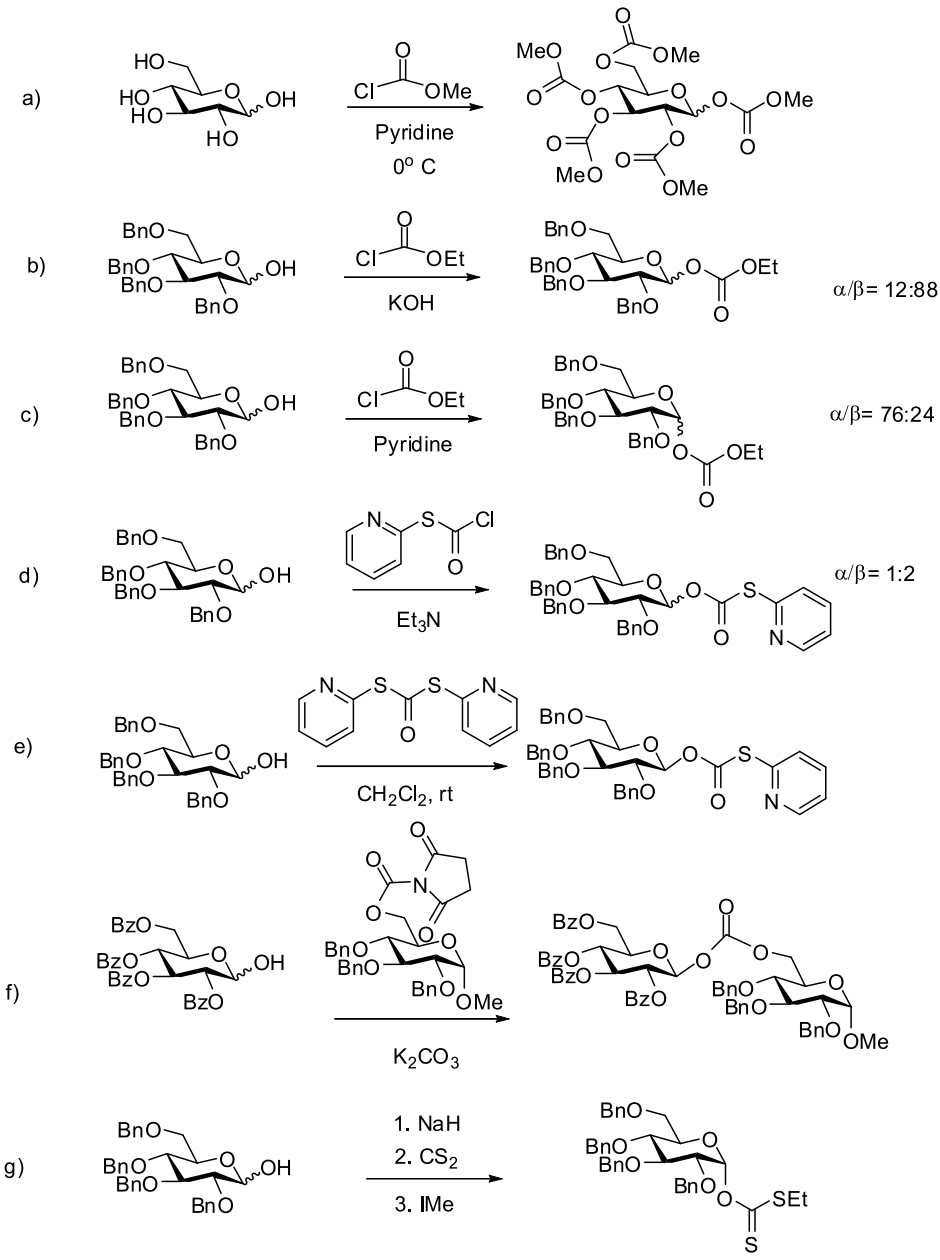
3.3 Carbonylation and Thiocarbonylation of the Anomeric Hydroxyl

The reagents most commonly used for the preparation of sugar carbonates are phosgene, alkyl chloroformates, and diaryl carbonates. Phosgene reacts with free sugars giving rise to cyclic carbonates preferentially having five-membered rings. Depending on the sugar the anomeric position may be involved, for instance when D-glucose is treated with phosgene and pyridine a 1,2:5,6-diester derivative is obtained. Unprotected sugars also react with chloroformic esters in the presence of pyridine, although to yield alkoxycarbonyl compounds (*Scheme 81a*) [452].

In the case of protected sugars, the anomeric hydroxy group reacts with chloroformic esters to give mixed esters (*Scheme 81b*). Usually the coupling reaction is not stereoselective giving rise to an anomeric mixture of carbonates, although the α/β ratio can be influenced by the choice of the proper base (*Scheme 81c*) [453]. Reaction of 2-thiopyridyl chloroformate with a glucose derivative results in an anomeric mixture (α:β, 1:2) (*Scheme 81d*) whereas the use of bis(2-thiopyridyl)carbonate yields exclusively the β-anomer (*Scheme 81e*) [454]. Very recently, a highly regio- and stereoselective reaction of D-glucopyranose 1,2-diols with allyl chloroformate or ethyl chloroformate has been reported [455].

Diaryl carbonates (e. g. carbonyl diimidazol, 4-nitrophenyl carbonate) can react sequentially with carbohydrate derivatives to furnish mixed sugar carbonates (*Scheme 81f*) [456]. Although normally anomeric mixtures are generated the use of a succinimidyl group, in the presence of K_2CO_3, was effective for the synthesis of pure β-carbonates.

Anomeric alkyl xanthates are prepared by treatment of 1-OH sugars with sodium hydride in the presence of a catalytic amount of imidazole, carbon disulfide, and an alkyl halide (*Scheme 81g*) [457].



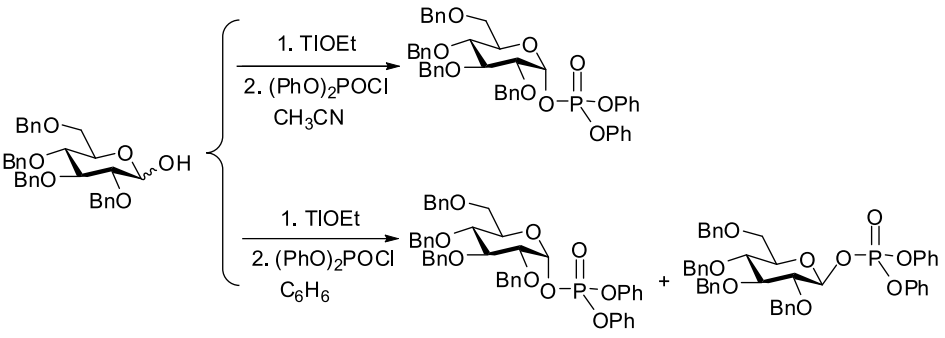
Scheme 81
Preparation of sugar carbonates

3.4 Silylation

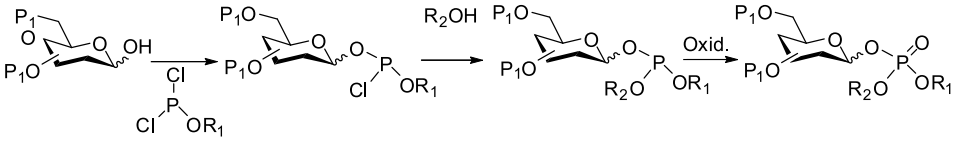
Anomeric silyl ethers have been prepared from 1-OH sugars and the corresponding silyl chloride in the presence of a base. When the hydroxyl group at C-2 is unprotected silyl group migrations away from the anomeric center have been observed [458].

3.5 Phosphorylation and Phosphitylation

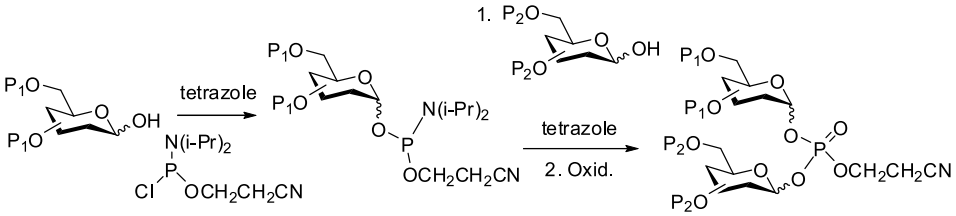
Glycosyl phosphates are intermediates in biological glycosyl transfer and are constituents of cell membranes [459]. Both chemical and enzymic methods are available for the synthesis of specific phosphates. In the preparation of certain glycosyl phosphates, enzymic synthesis with the appropriate phosphorylase provides the simplest preparation. In this fashion, α-D-glucopyranosyl phosphate is readily prepared by the phosphorolysis of starch or glycogen [460].
Chemically, synthesis of glucosyl phosphates also may involve two different approaches based either on activation of the glycosyl oxygen or activation at the anomeric carbon. In the latter, glycosyl acetates [461], orthoesters [462], glycosyl halides [463], trichloroacetimidates [464], vinyl glycosides [465], glycals [466], or 1,2-orthoesters [467] are used as glycosyl donors and they are not the aim of this chapter. The above two-step procedures ensure in most cases the anomeric purity of the final glycosyl phosphates.
On the other hand, several alternative procedures for the synthesis of glycosyl phosphates involving 1-OH activation have been developed. The thallium salt of the anomeric hydroxyl readily undergoes substitution with a phosphochloridate in benzene or acetonitrile (*Scheme 82*) [468]. The configuration of an organo-phosphate moiety introduced at the anomeric position is strongly influenced by the choice of solvent, so that a preponderance of either the α- or β-phosphate may be attained. These differences are reminiscent of solvent effects observed in syntheses of *O*-acyl esters and reflect differences in the anomeric composition in the reducing sugars as well as in the relative reactivities of the two anomers. Similarly, α-phosphates of *N*-acylglucosamine are prepared in high yields via the reaction of the corresponding 1-*O*-lithium salts with phosphorochloridate at low temperatures [469].



Scheme 82
Preparation of aldosyl phosphates



Scheme 83
Preparation of aldosyl phosphates



Scheme 84
Preparation of glycosyl phosphodiesters

A different approach to the synthesis of aldosyl phosphates involves the intermediacy of aldosyl phosphites [470]. The reaction of the anomeric hydroxyl group, with a trivalent phophitylating reagent, furnished an anomeric phosphorochloridite, which is able to react with a hydroxyl-containing compound, to generate a phosphorous triester which, upon oxidation gives the corresponding aldosyl phosphoric triester (*Scheme 83*).

More recently phophitylating reagents which, after oxidative transformation to the corresponding phosphates, allow removal of the protecting group at phosphorous (V) by mild base treatment, have been reported [471,472].

Recent studies have shown that stabilization in *O*-glycosyl phosphites can be achieved with the help of an electron-withdrawing *O*-alkyl group at the phosphite moiety (e. g. trichloroethyl vs. ethyl group) [473].

The phosphytilation approach has also been applied for the preparation of compounds in which two anomeric centers are part of a phosphodiester bond (*Scheme 84*) [471].

Acknowledgement

A.M.G. is grateful to the Ministerio de Educación y Ciencia of Spain [CTQ2006–15279-CO3–02/BQU] for financial support.

References

1. Kocienski PJ (2005) Protecting Groups, 3rd edn. Georg Thieme Verlag, Stuttgart
2. Robertson J (2000) Protecting Group Chemistry. Oxford University Press, New York
3. Greene TWP, Wuts GM (1999) Protective Groups in Organic Synthesis, 3rd edn. Wiley, New York
4. Hanson JR (1999) Protecting Groups in Organic Synthesis. Sheffield Academic Press, New York
5. Goodman L (1967) Adv Chem and Biochem 22:109
6. Mootoo DR, Konradsson P, Udodong U, Fraser-Reid (1988) J Am Chem Soc 110:5583
7. Schürrle K, Beier B, Werbitzky O, Piepersberg W (1991) Carbohydr Res 212:321
8. Evtushenko EV (1999) Carbohydr Res 316:187
9. Boto A, Hernández D, Hernández R, Suárez E (2004) Org Lett 6:3785
10. Boto A, Hernández D, Hernández R, Suárez E (2006) J Org Chem 71:1938
11. Czerncki S, Georgoulis C, Provelenghiou C (1976) Tetrahedron Lett 39:3535
12. Hijfte LV, Little RD (1985) J Org Chem 50:3940
13. Wessel HP, Lersen T, Bundle DR (1985) J Chem Soc Perkin Trans I 2247
14. Hatakeyama S, Mori H, Kitano K, Yamada H, Nishizawa M (1994) Tetrahedron Lett 35:4367
15. Weissman SA, Zewge D (2005) Tetrahedron 61:7833
16. Iseloh U, Dudkin V, Wang Z G, Danishefsky S (2002) Tetrahedron Lett 43:7027
17. Park MH, Takeda R, Nakanishi K (1987) Tetrahedron Lett 28:3823
18. Rodebaugh R, Debenham JS, Fraser-Reid B (1996) Tetrahedron Lett 37:5477
19. Falck JR, Barma DK, Baati R, Mioskowski Ch (2001) Angew Chem Int Ed 40:1281
20. Bhattacharjee SS, Gorin PA (1969) J Can Chem 47:1195
21. Gelas J (1981) Adv Carbohydr Chem 39:71
22. Garegg PJ (1997) In: Hanessian S (ed) Preparative Carbohydrate Chemistry. Marcel Dekker, New York, p 53
23. Garegg PJ, Hultberg H (1981) Carbohydr Res 93:C10
24. Mani NS, Kanakamma PP (1994) Tetrahedron Lett 35:3629
25. Ernst A, Vasella A (1996) Helv Chim Acta 79:1279
26. Ek M, Garegg PJ, Hultberg H, Oscarson S (1983) J Carbohydr Chem 2:305
27. Oikawa M, Liu WC, Nakai Y, Koshida S, Fukase K, Kusumoto S (1996) Synlett 1179
28. DeNinno MP, Etienne JB, Duplantier KC (1995) Tetrahedron Lett 36:669
29. Jiang L, Chan TH (1998) Tetrahedron Lett 39:355
30. Garegg PJ, Hultberg H, Wallin S (1982) Carbohydr Res 108:97
31. Shie ChR, Tzeng ZH, Kulkarni SS, Uang BJ, Hsu ChY, Hung SCh (2005) Angew Chem Int Ed 44:1665
32. Suzuki K, Nonaka H, Yamaura M (2003) Tetrahedron Lett 44:1975
33. Hanessian S (1968) Adv Chem Ser 74:159
34. Grindley TB (1998) Adv Carbohydr Chem Biochem 53:17 and references therein
35. Alais J, Veyrières A (1987) J Chem Soc Perkin Trans I 377
36. Danishefsky SJ, Hungate R (1986) J Am Chem Soc 108:2486
37. Beauepère D, Boutbaiba I, Wadouachi A, Frechou C, Demailly G, Uzan R (1992) New J Chem 16:405
38. Cruzado C, Martín-Lomas M (1986) Tetrahedron Lett 27:2497
39. Yang G, Ding X, Kong F (1997) Tetrahedon Lett 38:6725
40. Madsen J, Viuf C, Bols M (2000) Chem Eur J 6:1140
41. Cipolla L, Lay L, Nicotra F (1997) J Org Chem 62:6678
42. Hori H, Nishida Y, Ohrui H, Meguro H (1989) J Org Chem 54:1346
43. Sollogoub M, Das SK, Mallet JM, Sinaÿ P (1999) C R Acad Sci Paris, t 2, Série IIc, 441
44. Chevalier-du Roizel B, Cabianca E, Rollin P, Sinaÿ P (2002) Tetrahedron 58:9579
45. Jia C, Zhang Y, Zhang LH, Sinaÿ P, Sollogoub M (2006) Carbohydr Res 341:2135
46. Lecourt T, Herault A, Pearce AJ, Sollogoub M, Sinaÿ P (2004) Chem Eur J 10:2960
47. Falck JR, Barma DK, Venkataraman SK, Baati R, Mioskowski C (2002) Tetrahedron Lett 43:963
48. Oikawa Y, Yoshioka T, Yonemitsu O (1982) Tetrahedron Lett 23:885
49. Horita K, Yoshioka T, Tanaka T, Oikawa Y, Yonemitsu O (1986) Tetrahedron 42:3021

50. Chandrasekhar S, Sumithra G, Yadav JS (1996) Tetrahedron Lett 37:1645
51. Sharma GVM, Lavanya B, Mahalingam AK, Radha Krishna P (2000) Tetrahedron Lett 41:10323
52. Johansson R, Samuelsson B (1984) J Chem Soc Perkin Trans 1 2371
53. Classon B, Garegg PJ, Samuelsson B (1984) Acta Chem Scand B 38:419
54. Kartha KPR, Kiso M, Hasegawa A, Jennings H (1998) J Carbohydr Chem 17:811
55. Yu W, Su M, Gao Z, Yang Z, Jin Z (2000) Tetrahedron Lett 41: 4015
56. Sharma GVM, Reddy ChG, Krishna PR (2003) J Org Chem 68:4574
57. Akiyama T, Shima H, Ozaki S (1992) Synlett 415
58. Yan L, Kahne D (1995) Synlett 523
59. Vaino AR, Szarek WA (1995) Synlett 1157
60. Cappa A, Marcantoni E, Torregiani E (1999) J Org Chem 64:5696
61. Bhattacharyya S, Magnusson BG, Wellmar U, Nilsson UJ (2001) J Chem Soc Perkin Trans 1, 2001, 886
62. Jobron L, Hindsgaul O (1999) J Am Chem Soc 121:5835
63. Fukase K, Tanaka H, Torii S, Kusumoto S (1990) Tetrahedron Lett 31:389
64. Fukase K, Yoshimura T, Hashida M, Kusumoto S (1991) Tetrahedron Lett 32: 4019
65. Fukase K, Egusa K, Nakai Y, Kusumoto S (1996) Molecular Diversity 182
66. Fukase K, Hashida M, Kusumoto S Tetrahedron Lett (1991) 32:3557
67. Egusa K, Fukase K, Kusumoto S (1997) Synlett 675
68. Egusa K, Kusumoto S, Fukase K (2003) Eur J Org Chem 3435
69. Plante OJ, Buchwald SL, Seeberger PH (2000) J Am Chem Soc 122:7148
70. Liu X, Seeberger PH (2004) Chem Commun 1708
71. Guibé F (1997) Tetrahedron 40:13509
72. Gill J, Gigg R, Payne S, Conant R (1987) J Chem Soc Perkin Trans I 423
73. Rollin P, Sinaÿ P (1977) J Chem Soc, Perkin Trans I 2513
74. Hindsgaul O, Norberg T, Le Pendu J, Lemieux RU (1982) Carbohydr Res 109:109
75. Oltvoort JJ, Klosterman M, Van Boom JH (1983) Recl Trav Chim Pays-Bas 102:501
76. Lakhmiri R, Lhoste P, Sinou D (1989) Tetrahedron Lett 30:4669
77. Iversen T, Bundle DR (1981) J Chem Soc, Chem Commun 1240
78. David S, Hanessian S (1985) Tetrahedron 41:643
79. Eby R, Schuerch C (1982) Carbohydr Res 100:C41
80. Gent PA, Gigg R (1976) Carbohydr Res 49:325
81. Warren CD, Jeanloz RW (1977) Carbohydr Res 53:67
82. Boss R, Scheffold R (1976) Angew Chem Int Ed Engl 15:558
83. Ogawa T, Nakabayashi S, Kitajima T (1983) Carbohydr Res 114:225
84. Nicolaou KC, Hummel CW, Bockovich NJ, Wong CH (1991) J Chem Soc Chem Commun 870
85. Oltvoort JJ, van Boeckel CAA, de Koning JH; van Boom JH (1981) Synthesis 305
86. Yadav JS, Chandrasekhar S, Sumithra G, Kache R (1996) Tetrahedron Lett 37:6603
87. Kariyone K, Yazawa H (1970) Tetrahedron Lett 11: 2885
88. Diaz RR, Melgarejo CR, Lopez-Espinosa MTP, Cubero II (1994) J Org Chem 59:7928
89. Kitov PI, Bundle DR (2001) Org Lett 3:2835
90. Gent PA, Gigg R, Conant R (1972) J Chem Soc Perkin Trans I 1535
91. Sharma GVM, Ilangovan A, Mahalingam AK (1998) J Org Chem 63:9103
92. Vatèle JM (2002) Synlett 507
93. Vatèle JM (2002) Tetrahedron 58:5689
94. Markovic D, Vogel P (2004) Org Lett 6:2693
95. Helferich B (1948) Adv Carbohydr Chem 3:79
96. Sharma GVM, Mahalingam AK, Prasad TR (2000) Synlett 1479
97. Bessodes M, Komiotis KA (1986) Tetrahedron Lett 27:579
98. MacCoss M, Cameroon DJ (1978) Carbohydr Res 60:206
99. Jones GB, Hynd G, Wright JM, Sharma A (2000) J Org Chem 65:263
100. Lu RJ, Liu D, Giese RW (2000) Tetrahedron Lett 41:2817
101. Yan MCh, Chen YN, Wu HT, Lin ChCh, Chen ChT, Lin ChCh (2007) J Org Chem 72:299
102. Pathak AK, Pathak V, Seitz LE, Tiwari KN, Katar MS, Reynolds RC (2001) Tetrahedron Lett 42:7755
103. Agarwal A, Vankar YD (2005) Carbohydr Res 340:1661
104. Rawal GK, Rani S, Kumar A, Vankar YD (2006) Tetrahedron Lett 47:9117

105. Yadav JS, Subba Reddy BV (2000) Carbohydr Res 329:885
106. Chen MY, Patkar LN, Lu KCh, Lee ASY, Lin ChCh (2004) Tetrahedron 60:11465
107. Schaller H, Weimann G, Lerch B, Khorana HG (1963) J Am Chem Soc 85:3821
108. Caruthers MH (1989) J Chem Ed 66:577
109. Sarkar AK, Rostand KS, Jain RK, Matta KL, Esko JD (1997) J Biol Chem 272:25608
110. Gaunt MJ, Yu J, Spencer JB (1998) J Org Chem 63:4172
111. Liptak A, Borbás A, Jánossy L, Szilágyi L (2000) Tetrahedron Lett 41:4949
112. Borbás A, Szabo ZB, Szilágyi L, Bényei A, Lipták A (2002) Carbohydr Res 337:1941
113. Kia J, Abbas SA, Locke RD, Piskorz CF, Alderfer JL, Matta KL (2000) Tetrahedron Lett 41:169
114. Liao W, Locke RD, Matta KL (2000) Chem Comm 369
115. Xia J, Alderfer JL, Piskorz CF, Matta KL (2001) Chem Eur J 7:356
116. Crich D, Jayalath P (2005) Org Lett 7:2277
117. Crich D, Wu B (2006) Org Lett 8:4879
118. Crich D, Jayalath P, Hutton TK (2006) J Org Chem 71:3064
119. García-Moreno MI, Aguilar M, Ortiz Mellet C, García Fernández JM (2006) Org Lett 8:297
120. Balbuena P, Rubio EM, Ortiz Mellet C, García Fernández JM (2006) Chem Commun 2610
121. Fisher E (1895) Berichte der Deutschen Chemischen Gesellschaft 28:1145
122. de Belder AN (1977) Adv Carbohydr Chem Biochem 34:179
123. Calinaud P, Gelas J (1997) In: Hanessian S (ed) Preparative Carbohydrate Chemistry. Marcel Dekker, New York, p 3
124. Ley SV, Baeschlin DK, Dixon DJ, Foster AC, Ince SJ, Priepke HWM, Reynolds DJ (2001) Chem Rev 101:53
125. Gelas G (1981) Adv Carbohydr Chem Biochem 39:71
126. Eliel EL (1971) Pure Appl Chem 25:509
127. Garegg PJ, Swahn CG (1980) Methods Carbohydr Chem 8:317
128. Foster AB (1972) In: Pigman W, Horton D (eds) The Carbohydrates: Chemistry, Biochemistry. Academic Press, New York, p 391
129. Gelas J, Horton D (1981) Heterocycles 16:1587
130. Clode DM (1979) Chem Rev 79:491
131. Manna S, Jacques YP, Falck JR (1986) Tetrahedron Lett 27:2679
132. Park KH, Yoon YJ, Lee SG (1994) Tetrahedron Lett 35:9737
133. Lablance Y, Fitzsimmons J, Adams EP, Rokacha J (1986) J Org Chem 51:789
134. Vijayasaradhi S, Singh J, Aidhan IS, (2000) Synlett 110
135. Agarwal A, Vankar YD (2005) Carbohydr Res 340:1661
136. Yan MCh, Chen YN, Wu HT, Lin ChCh, Chen ChT, Lin ChCh (2007) J Org Chem 72:299
137. Majumdar S, Bhattacharya A (1999) J Org Chem 64:5682
138. Haines AH (1981) Adv Carbohydr Chem Biochem 39:71
139. Chéry F, Rollin P, De Lucci O, Cossu S (2000) Tetrahedron Lett 41:2357
140. Ariza X, Costa AM, Faja M, Pineda O, Vilarrasa J (2000) Org Lett 2:2809
141. Ley SV, Woods M, Zanotti-Gerosa A (1992) Synthesis 52
142. Ley SV, Leslie R. Tiffin PD, Woods M (1992) Tetrahedron Lett 33:4767
143. Ley SV, Boons GJ, Leslie R, Woods M, Hollinshead DM (1993) Synthesis 689
144. Hughes AB, Ley SV, Priepke HWM, Woods M (1994) Tetrahedron Lett 35:773
145. Ley SV, Downham R, Edwards PJ, Innes JE, Woods M (1995) Contemp Org Synth 2:365
146. Ley SV, Priepke HWM, Warriner SL (1994) Angew Chem Int Ed Engl 33:2290
147. Montchamp JL, Tian F, Hart ME, Frost JW (1996) J Org Chem 61:3897
148. Ley SV, Baeschlin DK, Dixon DJ, Foster AC, Ince SJ, Priepke HWM, Reynolds DJ (2001) Chem Rev 101:53
149. Douglas NL, Ley SV, Osborn HMI, Owen DR, Priepke HWM, Warriner SL (1996) Synlett 793
150. Litjens REJN, van den Bos LJ, Codée JDC, Overkleeft HO, van der Marel GA (2007) Carbohydr Res 342:419
151. Ley SV, Downham R, Edwards PJ, Innes JE, Woods M (1995) Contemp Org Synth 2:365
152. Grice P, Ley SV, Pietruuszka J, Priepke HWM, Warriner SL (1997) J Chem Soc Perkin Trans I 351
153. Hense A, Ley SV, Osborn HMI, Owen DR, Poisson JF, Warriner SL, Wesson KE (1997) J Chem Soc Perkin Trans 1 2023
154. Lindberg B (1990) Adv Carbohydr Chem Biochem 48:279
155. Araki S, Abe S, Yamada S, Satake M, Fujiwara N, Kon K, Ando S (1992) J Biochem 112:461
156. Dudman WF, Heidelberger M (1969) Science 164:954

157. Thayer WR, Bazic CM, Camphausen RT, McNeil M (1990) J Clin Microbiol 28: 714
158. Hirase S (1957) Bull Chem Soc Jpn 30:68
159. Dutton GGS, Karunaratane DN (1984) Carbohydr Res 134:103
160. Gorin PAJ, Mazurek M, Duarte HS, Duarte JH (1981) Carbohydr Res 92:C1
161. Ziegler T (1997) Top Curr Chem 186:203
162. Liptak A, Szabo L (1989) J Carbohydr Chem 8:629
163. Ziegler T, Eckhardt E, Herold G (1992) Tetrahedron Lett 33:4413
164. Ziegler T, Eckhardt E, Herold G (1992) Liebigs Ann Chem 441
165. Ziegler T, Eckhardt E, Neumann K, Birault V (1992) Synthesis 1013
166. Gorin PAJ, Ishikawa T (1967) Can J Chem 45:521
167. Garegg PJ, Lindberg B, Kvarnstrom I (1979) Carbohydr Res 77:71
168. Collins PM, McKinnon C, Manro A (1989) Tetrahedron Lett 30:1399
169. Aspinell GO, Ibrahim IH, Khare NK (1990) Carbohydr Res 200:247
170. Hiruma K, Tamura J, Horito S, Yoshimura J, Hashimoto H (1994) Tetrahedron 50:12143
171. Liptak A, Szabo L (1988) Carbohydr Res 184:C5
172. Liptak A, Bajza I, Kerekgyarto J, Hajko J, Szilagyi L (1994) Carbohydr Res 253:111
173. Agnihotri G, Misra AK (2006) Tetrahedron Lett 47:8493
174. Crich D, Bowers AA (2006) J Org Chem 71:3452
175. Crich D, Banerjee A (2005) Org Lett 7:1395
176. Crich D, Yao Q (2004) J Am Chem Soc 126:8232
177. Crich D, Yao Q (2003) Org Lett 5:2189
178. Miethchen R, Rentsch D, Stroll N (1992) Tetrahedron 48:8393
179. Miethchen R (2003) J Carbohydr Chem 22:801
180. Miethchen R, Rentsch D, Frank M, Lipták A (1996) Carbohydr Res 281:61
181. Miethchen R, Rentsch D, Frank M (1996) J Carbohydr Chem 15:15
182. Kumar P, Dinesh CU, Reddy RS, Pandey B (1993) Synthesis 1069
183. Ranu BC, Saha M (1994) J Org Chem 59:8269
184. Bhalerao UT, Davis KJ, Rao BV (1996) Synth Commun 26:3081
185. Yadav JS, Srinivas D, Reddy GS (1998) Synth Commun 28:1399
186. Tanemura K, Horaguchi T, Suzuki T (1992) Bull Chem Soc Jpn 65:304
187. Maity G, Roy SC (1993) Synth Commun 23:1667
188. Kluge AF, Untch KG, Fried JH (1972) J Am Chem Soc 94:7827
189. Corey EJ, Gras J, Ulrich P (1976) Tetrahedron Lett, 11: 809
190. Stork G, Isobe M (1975) J Am Chem Soc 97:4745
191. Kozikowski AP, Wu J (1987) Tetrahedron Lett 28:5125
192. Lipshutz BH, Pegram JJ (1980) Tetrahedron Lett 21:3343
193. Hosoya T, Takashiro E, Matsumoto T, Suzuki K (1994) J Am Chem Soc 116:1004
194. Pinto BM, Buiting MMW, Reimer KB (1990) J Org Chem 55:2177
195. Haines AH (1976) Adv Carbohydr Chem Biochem 33:11
196. Garegg PJ (1992) Acc Chem Res 25:575
197. Hofle G, Steglich W, Vorbruggen H (1978) Angew Chem Int Ed Engl 17:569
198. Scriven EFV (1983) Chem Soc Rev 12:129
199. Tiwari P, Kumar R, Maulik PR, Misra AK (2005) Eur J Org Chem 4265
200. Wolfrom ML, Thompson A (1963) Methods Carbohydr Chem 2:211
201. Hyatt JA, Tindall GW (1993) Heterocycles 35:227
202. Binch H, Stangier K, Thiem J (1998) Carbohydr Res 306:409
203. Kartha KPR, Field RA (1997) Tetrahedron 53:11753
204. Lee JC, Tai CA, Hung SC (2002) Tetrahedron Lett 43:851
205. Tai AA, Kulkarni SS, Hung SC, (2003) J Org Chem 68:8719
206. Ahmad S, Iqbal J (1987) J Chem Soc Chem Commun 114
207. Ghosh R, Chakraborty A, Maiti S (2004) Tetrahedron Lett 45:9631
208. Lu KC, Hsieh SY, Patkar LN, Chen CT, Lin CC (2004) Tetrahedron 60:8967
209. Dasgupta F, Singh PP, Srivastava HC (1980) Carbohydr Res 80:346
210. Montero JL, Winum JY, Leydet A, Kamal M, Pavia A A, Roque JP (1997) Carbohydr Res 297:175
211. Bhaskar PM, Loganathan D (1998) Tetrahedron Lett 39:2215
212. Bhaskar PM, Loganathan D (1999) Synlett 129

213. Kumareswaran R, Pachamuthu K, Vankar YD (2000) Synlett 1652
214. Tiwari P, Misra AK (2006) Carbohydr Res 341:339
215. Adinolfi M, Barone G, Iadonisi A, Schiattarella M (2003) Tetrahedron Lett 44:4661
216. Limousin C, Cleophax J, Petit A, Loupy A, Lukacs G (1997) J Carbohydr Chem 16:327
217. Das SK, Reddy KA, Krovvidi VLNR, Mukkanti K (2005) Carbohydr Res 340:1387
218. Murugesan S, Karst N, Islam T, Wiencek JM, Linhardt RJ (2003) Synlett 1283
219. Forsyth SA, Macfarlane DR, Thomson RJ, von Itzstein M (2002) Chem Commun 714
220. Zemplén G, Pacsu E (1929) Ber Dtsch Chem Ges 62:1613
221. Kurahashi T, Mizutani T, Yoshida JI (1999) J Chem Soc Perkin Trans I 465
222. Posner GH, Oda M (1981) Tetrahedron Lett 22:5003
223. Rana SS, Barlow JJ, Matta KL (1981) Tetrahedron Lett 22:5007
224. Bianco A, Brufani M, Melchioni C, Romagnoli P (1997) Tetrahedron Lett 38:651
225. Ishihara K, Kurihara H, Yamamoto H (1993) J Org Chem 58:3791
226. Ilankumaran P, Verkade JG (1999) J Org Chem 64:9063
227. Yamada S (1992) J Org Chem 57:1591
228. Hagiwara H, Morohashi K, Sakai H, Suzuki T, Ando M (1998) Tetrahedron 54:5845
229. Kattnig E, Albert M (2004) Org Lett 6:945
230. Ogawa T, Matsui M (1981) Tetrahedron 37:2363
231. Peri F, Cipolla L, Nicotra F (2000) Tetrahedron Lett 41:8587
232. Dong H, Pei Z, Byström S, Ramström O (2007) J Org Chem 72:1499
233. Waldmann H, Sebastian D (1994) Chem Rev 94:911
234. Schmid RD, Verger R (1998) Angew Chem Int Ed Engl 37:1608
235. Bordusa F (2002) Chem Rev 102:4817
236. Drueckhammer DG, Hennen WJ, Pederson RL, Barbas III CF, Gautheron CM, Krach, T, Wong CH (1999) Synthesis 499
237. Holla EW (1989) Angew Chem Int Ed Engl 28:220
238. Kooeterman M, De Nijs MP, Weijnen JG, Schoemaker HE, Meijer EM (1989) J Carbohydr Chem 8:333
239. Zemek J, Kucar S, Anderle D (1987) Collect Czech Chem Commun 52:2347
240. Csuk R, Glhzer BJ (1988) Z Naturforsch 436:1355
241. Naruto M, Ohno K, Naruse N, Takeuchi H (1979) Tetrahedron Lett 20:251
242. van Boeckel CAA, Beetz T (1983) Tetrahedron Lett 24:3775
243. Johnson F, Starkovsky NA, Paton AC, Carlson AA (1964) J Am Chem Soc 86:118
244. Lefeber DJ, Kamerling JP, Vliegenthart FG (2000) Org Lett 2:701
245. Sogabe S, Ando H, Koketsu M, Ishihara H (2006) Tetrahedron Lett 47:6603
246. Arranz E, Boons GJ (2001) Tetrahedron Lett 42:6469
247. Robins MJ, Hawrelak SD, Kanai T, Siefert JM, Mengel R (1979) J Org Chem 44:1317
248. Tomic S, Petrovic V, Matanovic M (2003) Carbohydr Res 338:491
249. Petrovic V, Tomic S, Ljevakovic D, Tomasic J (1997) Carbohydr Res 302:13
250. Kunz H, Harreus A (1982) Liebigs Ann Chem 41
251. Marin J, Blaton MA, Briand JP, Chiocchia G, Fournier C, Guichard G (2005) ChemBioChem 6:1796
252. Santoyo-González F, Uriel C, Calvo-Asín JA (1998) Syntesis 1787
253. Jiang L, Chan TH (1998) J Org Chem 63:6035
254. Yu H, Williams DL, Ensley HE (2005) Tetrahedron Lett 46:3417
255. van Boom JJ, Burgers PMJ (1976) Tetrahedron Lett 52:4875
256. Chery F, Cronin L, O'Brien JL, Murphy PV (2004) Tetrahedron 60:6597
257. Adamo R, Kovác P (2006) Eur J Org Chem 2803
258. García J, Fernández S, Ferrero M, Sanghvi Y, Gotor V (2003) Tetrahedron: Asymmetry 14:3533
259. Rej RN, Glushka JN, Chef W, Perlin AS (1989) Carbohydr Res 189:135
260. Wong CH, Ye X, Zhang Z (1998) J Am Chem Soc 120:7173
261. Zhu T, Boons GJ (2000) J Am Chem Soc 122:10222
262. Morère A, Mouffouk F, Chavis C, Montero JL (1997) Tetrahedron Lett 38:7519
263. Morère A, Menut Ch, Vidil C, Skaanderup P, Thorsen J, Roque JP, Montero JL (1997) Carbohydr Res 3000:175
264. Mouffouk F, Morère A, Vidal S, Leydet A, Montero JL (2004) Synth Commun 34:303
265. Barker GR, Gillam IC, Lord PA, Douglas T, Spoors JW (1960) J Chem Soc 3885

266. Shiozaki M, Deguchi N, Macindoe WM, Arai M, Miyazaki H, Mochizuki T, Tatsuta T, Ogawa J, Maeda H, Kurakata SI (1996) Carbohydr Res 283:27
267. Pulido R, Gotor V (1993) J Chem Soc Perkin Trans I 589
268. Morère A, Mouffouk F, Jeanjean A, Leydet A, Montero JL (2003) Carbohydr Res 338:2409
269. Guibe F, Saint M'Leux Y (1981) Tetrahedron Lett 22:3591
270. Adinolfi M, Barone G, Guariniello L, Iadonisi A (2000) Tetrahedron Lett 41:9305
271. Kunz H, Unverzagt C (1984) Angew Chem Int Ed Engl 23:436
272. Teshima T, Nakajima K, Takahashi M, Shiba T (1992) Tetrahedron Lett 33:363
273. Sridhar PR, Chandrasekaran S (2002) Org Lett 4:4731
274. Ramesh R, Bhat RG, Chandrasekaran S (2005) J Org Chem 70:837
275. Woodward RB, Heusler K, Gosteli J, Naegeli P, Oppolzer W, Ramage R, Ranganathan S, Vorbruggen H (1966) J Am Chem Soc 88:852
276. Fukase K, Yoshimura T, Kotani S, Kusumoto S (1994) Bull Chem Soc Jpn 67:473
277. Carpino LA, Han GY (1970) J Am Chem Soc 92:574
278. Roussel F, Knerr L, Grathwohl M, Schmidt RR (2000) Org Lett 2:3043
279. Zhu T, Boons GJ (2000) Tetrahedron: Asymmetry 11:199
280. Freese SJ, Vann WF (1996) Carbohydr Res 281:313
281. Gioeli C, Chattopadhyaya JB (1982) J Chem Soc Chem Commun 672
282. Roussel F, Takhi M, Schmidt RR (2001) J Org Chem 66:8540
283. Roussel F, Knerr L, Schmidt RR (2001) Eur J Org Chem 2067
284. Love KR, Seeberger PH (2004) Angew Chem Int Ed 43:602
285. Bufali S, Höleman A, Seeberger PH (2005) J Carbohydr Chem 24:441
286. LaLonde M, Chan TH (1985) Synthesis 817
287. Hwu RJR, Tsay SC, Cheng BL (1998) In: Rappoport Z, Apeloig Y (eds) Chemistry of Organic Silicon Compounds, vol 2. Wiley, New York, p 431
288. Larso GL Silicon-Based Blocking Agents. In: Suplement to the Gelest-Catalog (ABCR) Silicon, Germanium & Tin Compounds, Metal Alkoxides and Metal Diketonates; See http://www.gelest.com
289. Arias-Pérez MS, López MS, Santos MJ (2002) J Chem Soc Perkin Trans 2:1459
290. Ashton PR, Boyd SE, Gattuso G, Hartwell EY, Königar R, Spencer N, Stoddart JF (1995) J Org Chem 60:3898
291. Nelson TD, Crouch RD (1996) Synthesis 1031
292. Corey EJ, Venkateswarlu A (1972) J Am Chem Soc 94:6190
293. Corey EJ, Cho H, Rücker C, Hua DH (1981) Tetrahedron Lett 22:3455
294. Bredenkamp MWS (1995) Afr J Chem 48:154
295. Johnson DA, Taubner LM (1996) Tetrahedron Lett 37:605
296. Colombier C, Skrydstrup T, Beau JM (1994) Tetrahedron Lett 44:8167
297. Dahlhoff WV, Taba KM (1986) Synthesis 561
298. Chung MK, Orlova G, Goddard JD, Schlaf M, Harris R, Beveridge TJ, White G, Hallett FR (2002) J Am Chem Soc 124:10508
299. Chung MK, Schlaf M (2005) J Am Chem Soc 127:18085
300. Clive DLJ, Kellner D (1991) Tetrahedron Lett 32:7159
301. Kelly DR, Roberts SM, Newton RF (1979) Synth Commun 9:295
302. Collington EW, Finch H, Smith IJ (1985) Tetrahedron Lett 26:681
303. Shimshock SJ, Waitermire RE, Deshong P (1991) J Am Chem Soc 113:8791
304. Zhang W, Robins MJ (1992) Tetrahedron Lett 33:1177
305. Corey EJ, Yi KY (1992) Tetrahedron Lett 33:2289
306. Metcalf BW, Burkhart JP, Jund K (1980) Tetrahedron Lett 21:35
307. Peng Y, Li WD (2006) Synlett 1165
308. Fukuda Y, Shindo M, Shishido K (2003) Org Lett 5:749
309. Nakamura T, Shiozaki M (2001) Tetrahedron Lett 42:2701
310. Corey EJ, Roberts BE (1997) J Am Chem Soc 119:12425
311. Furstner A, Albert M, Mlynarski J, Metheu M, DeClerq E (2003) J Am Chem Soc 125:13132
312. Shahid KA, Mursheda J, Okazaki M, Shuto Y, Goto F, Kiyooka S (2002) Tetrahedron Lett 43:6377
313. Jackson SR, Johnson MG, Mikami M, Shiokawa S, Carreira EM (2001) Angew Chem Int Ed 40:2694
314. Yang YY, Yang WB, Teo CF, Lin CH (2000) Synlett 1634–1636

315. Oriyama T, Kobayashi Y, Noda K (1998) Synlett 1047
316. Bartoli G, Cupone G, Dalpozzo R, Nino AD, Maiuolo L, Procopio A, Sambri L, Tagarelli A (2002) Tetrahedron Lett 43:5945
317. Jadav JS, Reddy BVS, Madan C (2000) New J Chem 24:853
318. Crouch RD, Polizzi JM, Cleiman RA, Yi J, Romany CA (2002) Tetrahedron Lett 43:7151
319. Ranu BC, Jana U, Majee A (1999) Tetrahedron Lett 40:1985
320. Bartoli G, Bosco M, Marcantoni E, Sambri L, Torregiani E (1998) Synlett 209
321. Bajwa JS, Vivelo J, Slade J, Repic O, Blacklock T (2000) Tetrahedron Lett 41:6021
322. Crouch RD, Romany CA, Kreshock AC, Menkoni KA, Zile JL (2004) Tetrahedron Lett 45:1279
323. Jang JY, Wang YG (2003) Tetrahedron Lett 44:3859
324. Lee A SY, Yeh HC, Shie JJ (1998) Tetrahedron Lett 39:5249
325. Lipshutz BH, Keith J (1998) Tetrahedron Lett 39:2495
326. Hwu JR, Jain ML, Tsai FY, Tsay SC, Balkumar A, Hakimelahi GH (2000) J Org Chem 65:5077
327. Kumar GDK, Baskaran S (2005) J Org Chem 70:4520
328. Hanessian S, Lavallee P (1975) Can J Chem 53:2975
329. Mulzer J, Schöllhorn B (1990) Angew Chem Int Ed Engl 29:431
330. Hardinger SA, Wijaya N (1993) Tetrahedron Lett 34:3821
331. Crouch RD (2004) Tetrahedron 60:5833
332. Prakash C, Saleh S, Blair IA (1989) Tetrahedron Lett 30:19
333. Shekhani MS, Khan KM, Mahmood K, Shah PM, Malik S (1990) Tetrahedron Lett 31:1669
334. Rucker C (1995) Chem Rev 95:1009
335. Friesen RW, Sturino CF, Daljeet AK, Kolaczewska A (1991) J Org Chem 56:1944
336. Bennett F, Knight DW, Fenton G (1991) J Chem Soc Perkin Trans I 1543
337. Markiewicz WT (1979) J Chem Res (S) 24
338. Verdegaal CHM, Jansse PL, de Rooij JFM, van Boom JH (1980) Tetrahedron Lett 21:1571
339. van Boeckel CAA, van Boom JH (1985) Tetrahedron 21:4575
340. Trost BM, Caldwell CG (1981) Tetrahedron Lett 22:4999
341. Trost BM, Caldwell CG, Murayama E, Heissler D (1983) J Org Chem 48:3252
342. Furusawa K, Katsura T (1985) Tetrahedron Lett 26: 887
343. Corey EJ, Hopkins PB (1982) Tetrahedron Lett 23:4871
344. Kumagai D, Miyazaki M, Nishimura SI (2001) Tetrahedron Lett 42:1953
345. Imamura A, Ando H, Korogi S, Tanabe G, Muraoka O, Ishida H, Kiso M (2003) Tetrahedron Lett 44:6725
346. Imamura A, Kimura A, Ando H, Ishida H, Kiso M (2006) Chem Eur J 12:8862
347. Imamura A, Ando H, Ishida H, Kiso M (2005) Org Lett 7:4415
348. Slotin LA (1977) Synthesis 737
349. Lemmen P, Richter W, Werner B, Karl R, Stumpf R, Ugi I (1993) Synthesis 1
350. Edmundson RS (1979) In: Barton D, Ollis WD (eds) Comprehensive Organic Chemistry. Pergamon Press, Oxford, UK, vol 2, p 1267
351. Ireland RE, Muchmore DC, Hengartner U (1972) J Am Chem Soc 94:5098
352. Granata A, Perlin AS (1981) Carbohydr Res 94:165
353. Mora, N; Lacombe JM (1993) Tetrahedron Lett 34:2461
354. Schlimbrene BR , Miller SJ (2001) J Am Chem Soc 123:10125
355. Sabesan S, Neira S (1992) Carbohydr Res 223:169
356. Manning DD, Bertozzi CR, Rosen SD, Kiessling LL (1996) Tetrahedron lett 37:1953
357. Jones S, Seltsianos D, Thompson KJ, Toms SM (2003) J Org Chem 68:5211
358. Parang K (2002) Bioorg Med Chem Lett 12:1863
359. Ahmadibeni Y, Parang K (2005) J Org Chem 70:1100
360. Ahmadibeni Y, Parang K (2005) Org Lett 7:5589
361. Nifantiev EE, Grachev MK, Burmistrov SY (2000) Chem Rev 100:3755
362. Kochetkov NK, Nifantiev EE, Gudkova IP, Ivanona NL, Leskin VA (1972) Zh Obshch Khim 42:450
363. Haines AH, Massy DJR (1996) Synthesis 1422
364. Nifantiev EE, Gudkova IP, Chan DD (1972) Zh Obshch Khim 42:506
365. Nifantiev EE, Tuseev AP, Tarasov VV (1966) Zh Obshch Khim 36:1124
366. Nifantiev EE, Schegolev AA (1965) Vestn Mosk Univ Ser II: Khim 80
367. Kochetkov NK, Nifantiev EE , Koroteev MP, Zhane ZK, Borisenko AA (1976) Carbohydr Res 47:221

368. van Herk T, Hartog AF, van der Burg AM, Wever R (2005) Adv Synth Catal 347:1155
369. Fischer E (1893) Chem Ber 26:2400
370. Purdie T, Irvine JC (1903) J Chem Soc 83:1021
371. Haworth WN (1919) J Chem Soc 107:8
372. Roth D, Pigman W (1960) J Am Chem Soc 82:4608
373. Schmidt RR (1986) Angew Chem Int Ed Engl 25:212
374. Schmidt RR, Reichrath M, Moering V (1984) J Carbohydr Chem 3:67
375. Box VGS (1982) Heterocycles 19:1939
376. Schmidt RR, Reichrath M, Moering U (1980) Tetrahedron Lett 21:3501
377. Klotz W, Schmidt RR (1994) J Carbohydr Chem 13:1093
378. Klotz W, Schmidt RR (1991) Synlett 168
379. Klotz W, Schmidt RR (1993) Liebigs Ann Chem 683
380. Lu W, Navidpour L, Taylor SD (2005) Carbohydr Res 340:1213
381. Esswein A, Rembold H, Schmidt RR (1990) Carbohydr Res 200:287
382. Srivastava VK, Schuerch C (1979) Tetrahedron Lett 35:3269
383. Mukaiyama T, Hashimoto Y, Hayashi Y, Shoda S (1984) Chem Lett 557
384. Huchel U, Schmidt C, Schmidt RR (1998) Eur J Org Chem
385. Sharma SK, Corrales G, Penadés S (1995) Tetrahedron Lett 36:5627
386. Koeners HJ, de Kok AJ, Romers C, van Boom JH (1980) Recl Trav Cim Pays-Bas 99:355
387. Peterson L, Jensen KJ (2001) J Org Chem 66:6268
388. Jacobsson M, Malmberg J, Ullervik U (2006) Carbohydr Res 341:1266
389. Capon B (1969) Chem Rev 69:407
390. BeMiller JM (1967) Adv Carbohydr Chem 22:25
391. Bochkov AF, Zaikov GE (1979) Chemistry of the *O*-glycosidic bond. Pergamon Press, Oxford
392. Garegg PJ (2004) Adv Carbohydr Chem Biochem 59:69
393. Freudenberg K, Dürr W, von Hochstetter H (1928) Chem Ber 61:1735
394. Richtmyer NK (1934) J Am Chem Soc 56:1633
395. Gómez AM, Danelón GO, Valverde S, López JC (1999) Carbohydr Res 320:138
396. Classon B, Garegg PJ, Samuelsson B (1984) Acta Chem Scand B 38:419
397. Zehavi U, Amit B, Patchornik A (1972) J Org Chem 37:2281
398. Rodebaugh R, Fraser-Reid B, Geysen HM (1997) Tetrahedron Lett 38:7653
399. Kim KS, Kim JH, Lee YJ, Lee YJ, Park J (2001) J Am Chem Soc 123 8477
400. Stanek JJr (1990) Topics Curr Chem 154: 209
401. Gigg J, Gigg R (1966) J Chem Soc C 82
402. Nakayama K, Uoto K, Higashi K, Soga T, Kusama T (1992) Chem Pharm Bull 40:1718
403. Lüning J, Möller U, Debski N, Welzel P (1993) Tetrahedron Lett 37:5871
404. Yu B, Zhang J, Lu S, Hui Y (1998) Synlett 29
405. MootooDR, Date V, Fraser-Reid B (1988) J Am Chem Soc 110:266
406. López JC, Fraser-Reid B (1991) J Chem Soc Chem Commun 159
407. Kunz H, Werning P, Schultz M (1990) Synlett 631
408. Lipshutz BH, Pegram JJ, Morey MC (1981) Tetrahedron Lett 22:4463
409. Jansson K, Ahlfors S, Fredj T, Kihlberg J, Magnusson G (1988) J Org Chem 53:5629
410. Behrend R, Roth P (1904) Liebigs Ann Chem 331:359
411. Schlubach HH, Prochownick V (1930) Chem Ber 63:2298
412. Levene PA, Meyer GM (1928) J Biol Chem 76:513
413. Fletcher HGJr (1953) J Am Chem Soc 75:2624
414. Sano T, Ohashi K, Oriyama T (1999) Synthesis 7:1141
415. Hung SCh, Thopate SR, Wang ChCh (2001) Carbohydr Res 330:177
416. Luo SY, Kulkarni SS, Chou ChH, Liao WM, Hung SCh (2006) J Org Chem 71:1226
417. Hassner A, Alexanian V (1978) Tetrahedron Lett 4475
418. Ziegler FE, Berger GD (1979) Synth Commun 9: 539
419. Kim KS, Lee YJ, Kim HY, Kang SS, Kwon SY (2004) Org Biomol Chem 2:2408
420. Carpino LA (1993) J Am Chem Soc 115:4397
421. Perrie JA, Harding JR, Holt DW, Johnston A, Meath P, Stachulski AV (2005) Org Lett 7:2591
422. Shoda S, Mukaiyama T (1982) Chem Lett 6:861
423. Pfeffer PE, Rothman ES, Moore GG (1970) J Org Chem 2925
424. Barrett AGM, Bezuidenhoudt BCB, Gasieki AF, Howell AR, Russel MA (1989) J Am Chem Soc 111:1392
425. Ogawa T, Nozaki M, Matsui M (1978) Carbohydr Res 60:C7
426. Barrett AGM, Bezoindehoudt BCB (1989) Heterocycles 28:209

427. Schmidt RR, Kinzy W (1994) Adv Carbohydr Chem Biochem 50:21
428. Schmidt RR, Michel J (1980) Angew Chem Int Ed Engl 731
429. Schmidt RR, Michel J (1984) Tetrahedron Lett 25:821
430. Ellervik U, Magnusson G (1993) Acta Chem Scand 47:826
431. Smith AB, Hale KJ, Rivero RA (1986) Tetrahedron Lett 27:5813
432. Shaw JF, Klibanov AM (1987) Biotech Bioeng 29:648
433. Hennen WJ, Sweers HM, Wang YF, Wong CH (1988) J Org Chem 53:4939
434. Khan R, Gropen L, Konowicz PA, Matulovà M, Paoletti S (1993) Tetrahedron Lett 34:7767
435. Rowell RM, Feather MS (1967) Carbohydr Res 4:486
436. Excoffier G, Gagnaire D, Utille JP (1975) Carbohydr Res 39:368
437. Khan R, Konowicz PA, Gardossi L, Matulovà M, Paoletti S (1994) 35:4247
438. Ferrer Salat C, Exero Agneseti P, Bemborad Caniato M (1976) Spanish Patent 430 636, Laboratorios Ferrer SL: CA (1977) 87:23683k
439. Ishido Y, Sakairi N, Sekiya M, Nakazaki N (1981) Carbohydr Res 97:51
440. Fiandor J, García López MT, de las Heras FG, Méndez Castrillón PP (1985) Synthesis 1121
441. Grynkiewicz G, Fokt I, Szeja W, Fitak H (1989) J Chem Res S 152
442. Watanabe K, Itoh K, Araki Y, Ishido Y (1986) Carbohydr Res 154:165
443. Itoh T, Takamura H, Watanabe K, Araki Y, Ishido Y (1986) Carbohydr Res 156:241
444. Nudelman A, Herzig J, Gottlieb HE, Keinan E, Sterling J (1987) Carbohydr Res 162:145
445. Mikamo M (1989) Carbohydr Res 191:150
446. Chittaboina S, Hodges B, Wang Q (2006) Let Organ Chem 3:35
447. Sambarah T, Fanwick PE, Cushman M (2001) Synthesis 1450
448. Herzig J, Nudelman A (1986) Carbohydr Res 153:162
449. Avalos M, Babiano R, Cintas P, Jiménez JL, Palacios JC, Valencia C (1993) Tetrahedron Lett 34:1359
450. Banaszek A, Bordas-Cornet X, Zamojski A (1985) Carbohydr Res 144:342
451. Lemieux RU, Brice C (1955) Can J Chem 33:109
452. Zemplén G, Laszlo ED (1915) Chem Ber 48:915
453. Boursier M, Descotes G (1989) C R Acad Sci 308:919
454. Lou B, Huynh HK, Hanessian S (1997) Oligosaccharide synthesis by remote activation: *O*-protected glycosyl 2-thiopyridylcarbonate donors. In Hanessian S (ed) Preparative carbohydrate chemistry. Marcel Dekker, New York, chap 19
455. Zhang J, Liang X, Wang D, Kong F (2007) Carbohydr Res 342:797
456. Iimori T, Shibazaki T, Ikegami S (1997) Tetrahedron Lett 38:2943
457. Pougny JR (1986) J Carbohydr Chem 5:529
458. Lassaletta JM, Schmidt RR (1995) Synlett 925
459. Nikolaev AV, Botvinko IV, Ross AJ (2007) Carbohydr Res 342:297
460. Putman EW (1963) Methods Carbohydr Chem 2:267
461. McDonald DL (1962) J Org Chem 27:1107
462. Volkova LV, Danilov LL, Evstigneeva RP (1974) Carbohydr Res 32:165
463. Putman EW (1963) Methods Carbohydr Chem 2:261
464. Schmidt RR, Stumpp M (1984) Liebigs Ann Chem 680
465. Boons GJ, Burton A, Wyatt P (1996) Synlett 310
466. Plante OJ, Palmacci ER, Seeberger PH (2001) Science 291:1523
467. Ravida A, Liu X, Kovacs L, Seeberger PH (2006) Org Lett 8:1815
468. Granata A, Perlin AS (1981) Carbohydr Res 94:165
469. Inage M, Chaki H, Kusumoto S, Shiba T (1982) Chem Lett 1281
470. Ogawa T, Seta A (1982) Carbohydr Res 110:C1
471. Westerdouin P, Veeneman GH, Marugg JE, van der Marel GA, van Boom JH (1986) Tetrahedron Lett 27:1211
472. Ichikawa Y, Sim MM, Wong CH (1992) J Org Chem 57:2943
473. Müller T, Hummel G, Schmidt RR (1994) Liebigs Ann Chem 325

2.2 Oxidation, Reduction, and Deoxygenation

Robert Madsen
Department of Chemistry, Center for Sustainable and Green Chemistry, Technical University of Denmark, Lyngby 2800, Denmark
rm@kemi.dtu.dk

Abstract

In this chapter, methods for oxidation, reduction, and deoxygenation of carbohydrates are presented. In most cases, the reactions have been used on aldoses and their derivatives including glycosides, uronic acids, glycals, and other unsaturated monosaccharides. A number of reactions have also been applied to aldonolactones. The methods include both chemical and enzymatic procedures and some of these can be applied for regioselective transformation of unprotected or partially protected carbohydrates.

Keywords

Azidonitration; Deoxygenation; Dihydroxylation; Epoxidation; Hydrogenation; Oxidation; Photobromination; Reduction

In: *Glycoscience*. Fraser-Reid B, Tatsuta K, Thiem J (eds)
Chapter-DOI 10-1007/978-3-540-30429-6_4:

Abbreviations

AIBN 2,2′-azobisisobutyronitrile
BSTFA bis(trimethylsilyl)trifluoroacetamide
CAN ceric ammonium nitrate
DCC *N*,*N*′-dicyclohexylcarbodiimide
DIBALH diisobutylaluminum hydride
DMF dimethylformamide
DMSO dimethyl sulfoxide
HMPA hexamethylphosphoric triamide
MS molecular sieves
***m*-CPBA** *m*-chloroperoxybenzoic acid
NBS *N*-bromosuccinimide
NIS *N*-iodosuccinimide
NMO *N*-methyl-morpholine *N*-oxide
PCC pyridinium chlorochromate
PDC pyridinium dichromate
TFAA trifluoroacetic anhydride
TEMPO 2,2,6,6-tetramethyl-1-piperidinyloxy
TPAP tetrapropylammonium perruthenate

1 Introduction

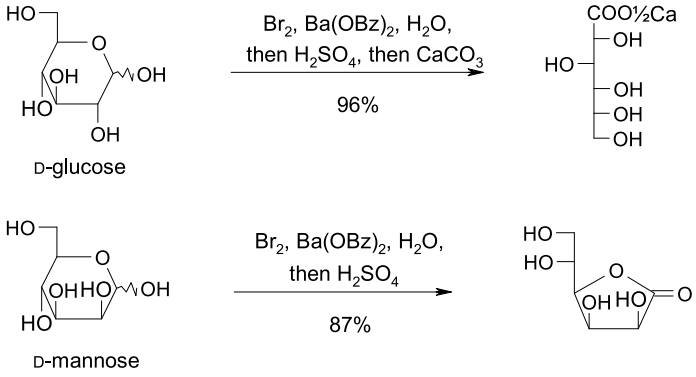
Carbohydrates have found valuable applications as enantiomerically pure starting materials for synthesis of non-carbohydrate compounds. Through chemical synthesis carbohydrates can be transformed into versatile synthetic intermediates that have functional groups and stereogenic centers structured in a framework found in many natural products, and the entire repertoire of common reactions in organic chemistry can be performed on carbohydrates [1]. However, carbohydrates are also densely functionalized molecules and as such constitute a special synthetic challenge. Regio- and chemoselectivity problems are frequently encountered which often makes it necessary to protect functionalities not involved in the desired transformation. In the following, oxidations, reductions, and deoxygenations will be discussed. An eight-volume comprehensive encyclopedia is recommended for further information on the different reagents for these reactions [2].

2 Oxidations

A large variety of oxidation procedures are available to the synthetic organic chemist [3]. As presented in this chapter, many of these methods can be applied to protected carbohydrates and some also for regioselective oxidation of unprotected or partially protected sugars.

2.1 Oxidation at the Anomeric Center

Unprotected aldoses can be selectively oxidized at the anomeric center to afford aldonic acids/aldonolactones. This oxidation can be achieved by chemical as well as by biochemical

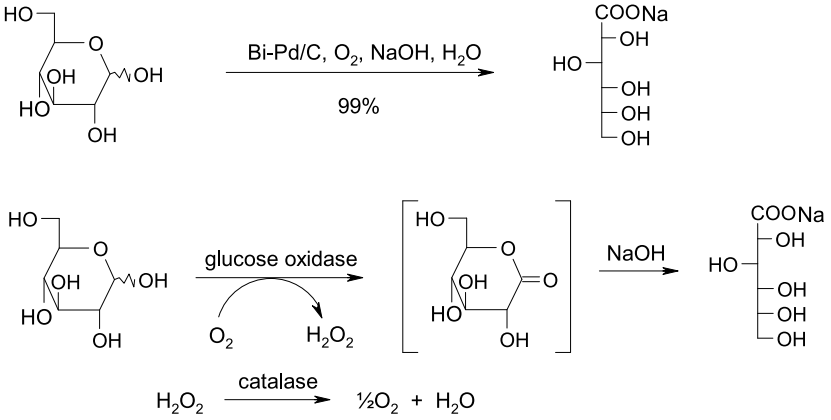


Scheme 1

methods. A common laboratory procedure for oxidation of aldoses uses 1.1 equiv. of bromine in an aqueous solution in the presence of an acid scavenger such as barium benzoate, barium carbonate, or calcium carbonate [4]. The scavenger is necessary since the liberated hydrobromic acid lowers the rate of the oxidation reaction. These conditions are very selective for the anomeric center and will not oxidize other hydroxy groups in the aldose. The product is typically isolated by crystallization either as the aldonolactone or as a salt of the aldonic acid [5]. It should be noted that aldonolactones usually exist as the five-membered 1,4-lactone contrary to aldoses which prefer the six-membered pyranose form. Gluconolactone is an important exception from the rule since it crystallizes as the 1,5-lactone. The oxidation with bromine takes place on the cyclic form of the aldose and not with the free aldehyde. Furthermore, the β-pyranose is oxidized faster than the α-pyranose for all the common aldoses [6]. Therefore, the initial product is the 1,5-lactone which will either ring-open to form a salt of the aldonic acid or rearrange to the thermodynamically more stable 1,4-lactone (*Scheme 1*) [4]. The process can also be turned into a catalytic procedure by using an electrochemical oxidation of calcium bromide to generate bromine in a solution with the aldose and calcium carbonate [7].

Another catalytic method makes use of a homogeneous dehydrogenation catalyst in the presence of a hydrogen acceptor. The complex $RhH(PPh_3)_4$ catalyzes a clean dehydrogenation of unprotected aldoses into aldono-1,4-lactones in DMF [8]. Benzalacetone ($PhCH{=}CHCOCH_3$) serves as the hydrogen acceptor and is converted into 4-phenylbutan-2-one during the course of the reaction.

Although homogeneous catalysts are often used at the laboratory scale industrial applications usually prefer a heterogeneous catalyst due to the easy separation from the product and the recovery of the catalyst. Many heterogeneous catalysts have been studied for aerobic oxidation of unprotected aldoses. The favored catalysts are Pd/C and Au/C, which show very high selectivity for the hemiacetal function [9,10]. A drawback with Pd/C, however, is catalyst deactivation. This can be circumvented by promoting the catalyst with bismuth which seems to improve the catalyst performance by coordinating with the substrate [11]. Thus, aerobic oxidation of glucose over a Bi-Pd/C catalyst at pH 9 with continuous addition of sodium hydroxide gives rise to sodium gluconate in 99% yield (*Scheme 2*) [9]. The catalyst can be



Scheme 2

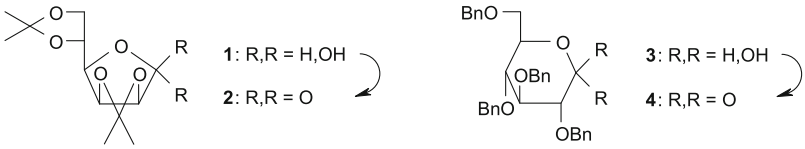
recovered and reused five times without affecting the yield. The same oxidation can also be performed with an Au/C catalyst to afford sodium gluconate in a near quantitative yield, but in this case the reused catalyst reacts more slowly in the following runs [10]. When the oxidation of glucose is carried out with the parent Pd/C catalyst the yield drops to 78% due to incomplete conversion [9].

Despite the advances in heterogeneous catalysis the preferred industrial procedure for oxidation of glucose is still aerobic fermentation. The microbial process is usually performed with *Aspergillus niger* by the use of submerged fermentation [12]. The obtained gluconolactone/gluconic acid will inhibit the fungal growth and is therefore neutralized with sodium hydroxide in order to maintain the pH of the growth medium around 6. The process is highly efficient and is able to oxidize glucose at a rate of 15 g/L per hour [12].

The oxidation can also be carried out by using an enzymatic reaction with the two enzymes that are responsible for the microbial oxidation: glucose oxidase and catalase (*Scheme 2*) [13]. Glucose oxidase dehydrogenates glucose to gluconolactone by simultaneous reduction of dioxygen to hydrogen peroxide. The liberated peroxide is an inhibitor of glucose oxidase and must be removed in order to obtain a good conversion. This is achieved by decomposition with catalase and the conversion of glucose to gluconate is nearly quantitative under these conditions [12]. Glucose oxidase is very specific for D-glucose, but with a prolonged reaction time similar aldoses can also be oxidized in good yield including D-xylose, D-mannose, D-galactose, and D-glucosamine [14].

Protection of aldoses at the non-anomeric positions makes it possible to use many of the common procedures in organic chemistry for oxidizing lactols as shown with mannofuranose **1** and glucopyranose **3** (*Table 1*). The reactions can be divided into three main categories: oxidations mediated by activated dimethyl sulfoxide (DMSO), oxidations with chromium(VI) oxides, and oxidations catalyzed by ruthenium oxides. The DMSO-mediated oxidations of alcohols can be promoted by several activators [27]. With the partially protected aldoses the activation has mainly been achieved with acetic anhydride and oxalyl chloride. Competing β-elimination does usually not occur unless the eliminating group is an ester, e. g., an acetate or a benzoate [27].

Table 1
Oxidation of diisopropylidenemannofuranose 1 and tetrabenzylglucopyranose 3 to lactones 2 and 4



Substrate	Reagent	Solvent	Yield (%)	Reference
1	DMSO, Ac_2O	DMSO	79–96	[15,16]
3	DMSO, Ac_2O	DMSO	84	[17]
1	DMSO, $(COCl)_2$; Et_3N	CH_2Cl_2	82	[18]
3	DMSO, $(COCl)_2$; Et_3N	CH_2Cl_2	60	[19]
1	$CrO_3{\cdot}2C_5H_5N$	CH_2Cl_2	80	[16]
1	$CrO_3{\cdot}2C_5H_5N$, Ac_2O	CH_2Cl_2	97	[20]
3	PDC, 3 Å MS	CH_2Cl_2	86	[19]
1	PCC, 3 Å MS	CH_2Cl_2	86	[21]
3	PCC, 3 Å MS	CH_2Cl_2	95	[22]
1	RuO_2, $NaIO_4$	$H_2O/CHCl_3/CCl_4$	79	[23]
1	TPAP, NMO	MeCN	88	[24]
3	TPAP, NMO	CH_2Cl_2	94	[19]
1	Dess–Martin periodinane	CH_2Cl_2	95	[25]
1	$RhH(PPh_3)_4$, benzalacetone	DMF	93	[8]
1	NiO(OH)[a], K_2CO_3	H_2O	80	[26]

[a]Generated electrochemically

The chromium-mediated oxidations can be performed with a number of chromium(VI) reagents [28]. The Collins reagent ($CrO_3.2C_5H_5N$) is not very effective at oxidizing aldoses and 6 equiv. are needed in order to oxidize **1** in 30 min [16]. The reactivity of the reagent can be enhanced by adding acetic anhydride which makes it possible to oxidize **1** with 4 equiv. of the reagent in 5–10 min [20]. The same number of equivalents is usually required with pyridinium dichromate (PDC) and pyridinium chlorochromate (PCC) which are both commercially available [28]. In all reactions with chromium reagents the work-up is rather tedious and a significant amount of toxic waste is produced. Upon completion of the reaction most of the chromium byproducts are precipitated as a tarry mass which is then followed by purification of the products by silica gel flash chromatography.

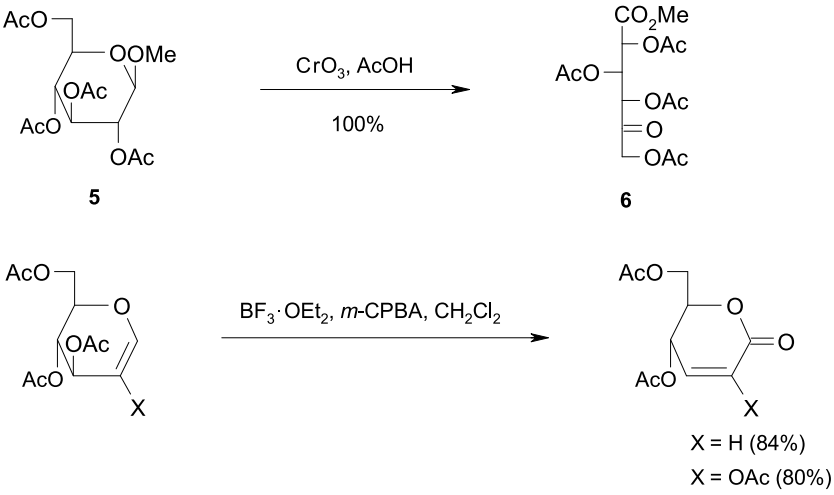
In view of the purification and waste disposal problems with the chromium oxidations catalytic methods with ruthenium catalysts are more attractive. Ruthenium(VIII) oxide is a strong oxidant that will also oxidize alkenes, alkynes, sulfides, and in some cases benzyl ethers. The method is compatible with glycosidic linkages, esters and acetals, and is usually carried out in a biphasic solvent system consisting of water and a chlorinated solvent. Acetonitrile or a phase-transfer catalyst has been shown to further promote the oxidation [29,30]. Normally, a periodate or a hypochlorite salt serve as the stoichiometric oxidant generating ruthenium(VIII) oxide from either ruthenium(IV) oxide or ruthenium(III) chloride [30].

Another ruthenium-catalyzed oxidation uses tetrapropylammonium perruthenate (TPAP) [24]. Being a ruthenium(VII) oxide, the perruthenate ion is a less powerful oxidant than ruthenium(VIII) oxide and more functional groups are stable to the oxidation conditions, including alkenes, alkynes, amines, amides, benzyl, trityl and silyl ethers [24]. However, alcohols and lactols still undergo oxidations in high yield with *N*-methyl-morpholine *N*-oxide (NMO) as the stoichiometric oxidant. The reactions are usually carried out in dichloromethane, acetonitrile, or mixtures of both in the presence of molecular sieves [24].

A number of special oxidation methods have also been applied to partially protected aldoses. The Dess–Martin periodinane, 1,1,1-triacetoxy-1,1-dihydro-1,2-benziodoxol-3(1*H*)-one [31], is a mild and efficient oxidant compatible with carbohydrate protecting groups [25]. However, it is also quite expensive and should only be used if the other procedures fail. The hydrogen transfer reactions catalyzed by 5% of $RhH(PPh_3)_4$ can also be applied to partially protected aldoses [8].

In all the above methods for oxidizing carbohydrates a stoichiometric oxidant is added to the reaction mixture. This can be avoided by using an electrochemical oxidation. A nickel hydroxide electrode has been applied for oxidizing isopropylidene-protected carbohydrates in aqueous base [26]. While secondary hydroxy groups fail to react under these conditions, the hemiacetal at the anomeric center is oxidized to the lactone in good yield [26].

Besides aldoses methyl glycosides and glycals can also be oxidized at the anomeric center. Peracetylated methyl β-D-glucopyranoside **5** reacts with ozone at the anomeric center to give the corresponding open chain methyl ester [32]. When the reaction is performed with 2 equiv. of chromium(VI) oxide further oxidation occurs to give keto ester **6** in quantitative yield (*Scheme 3*) [33]. Interestingly, the α-anomer of **5** does not react under these conditions. Ester-protected glycals or 2-hydroxyglycals react selectively with *m*-chloroperoxybenzoic acid (*m*-CPBA) in the presence of borontrifluoride etherate to afford α,β-unsaturated lactones (*Scheme 3*) [34]. The Lewis acid mediates an allylic rearrangement which is followed by



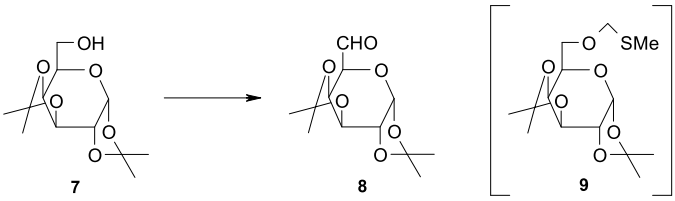
Scheme 3

oxidation at the anomeric center by the peracid. The same transformation can be achieved with indium(III) chloride and 2-iodoxybenzoic acid [35].

2.2 Oxidation of Primary Alcohols to Aldehydes

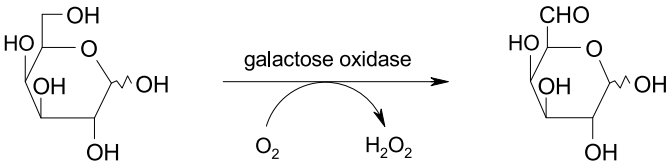
An easily available protected carbohydrate containing a primary hydroxy group is diisopropylidenegalactopyranose **7**. Oxidation of **7** to the corresponding aldehyde **8** illustrates very well the different reagents available for this transformation (*Table 2*). Activated DMSO-mediated oxidations are usually the method of choice for converting a primary alcohol into an aldehyde [27]. The conditions are mild and no overoxidation to the carboxylic acid occurs. The original Moffatt procedure uses *N*,*N*′-dicyclohexylcarbodiimide (DCC) in the presence of a proton source as activator. A good yield is obtained with alcohol **7**, but with some alcohols a significant amount of the corresponding methylthiomethyl ether is formed [27]. The urea byproduct can also be difficult to separate from a protected carbohydrate aldehyde. Activation with sulfur trioxide or oxalyl chloride gives a more straightforward work-up and these procedures also give rise to very little of the methylthiomethyl ether byproduct **9** [27]. The Swern procedure (oxalyl chloride) is the most widely used protocol due to its combination of high reactivity with inexpensive reagents. To avoid the methylthiomethyl byproduct, the Swern oxidation is normally carried out at low temperature. A tertiary amine is added in the last step of the oxidation process. Triethylamine is most commonly used, but more hindered amines have been reported to give better yields in some cases [27]. These basic conditions can cause epimerization of the aldehyde or β-elimination. Although these side reactions have not been observed in the oxidation of **7**, a similar substrate without the 1,2-isopropylidene group gave exclusively the β-elimination product when using the Swern procedure [42]. Acetic anhydride can also be used to activate DMSO, and this procedure has been used for a number of carbohydrate alcohols [27,42]. However, the reaction is slow and yields often moderate due to signifi-

Table 2
Oxidation of diisopropylidenegalactopyranose 7 to aldehyde 8



7 8 9

Reagent	Solvent	Yield (%)	Reference
DMSO, DCC, pyridine·HCl	DMSO	83–87	[36]
DMSO, SO_3·pyridine; Et_3N	DMSO	85	[37]
DMSO, $(COCl)_2$; Et_3N	CH_2Cl_2	82	[38]
CrO_3, pyridine, Ac_2O	CH_2Cl_2	84–93	[39,20]
PDC, Ac_2O	CH_2Cl_2/DMF	71	[40]
$Pb(OAc)_4$	Pyridine	74	[41]



Scheme 4

cant formation of the methylthiomethyl ether and the acetate of the starting alcohol [27,42]. In fact, oxidation of **7** with DMSO/acetic anhydride gave methylthiomethyl ether **9** as the major product together with smaller amounts of the desired aldehyde **8** and the acetate of **7** [43].

Heavy metal reagents can also be used for oxidation of primary alcohols to aldehydes. Although experimentally less attractive than the Swern procedure, good yields can be obtained. Particularly, the chromium(VI)-based oxidants when activated by acetic anhydride have found use for carbohydrate alcohol oxidations even on large scale [20,39]. The major side reaction is overoxidation to the corresponding carboxylic acid. Contrary to these chromium reagents, lead(IV) acetate is generally not very reactive for oxidation of carbohydrate alcohols although oxidation of **7** to **8** has been achieved in a satisfactory yield [41].

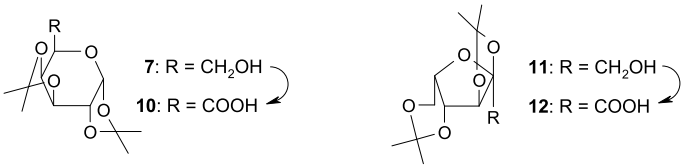
The TPAP/NMO system [24] and the Dess–Martin periodinane [31] have been widely applied for oxidizing alcohols in complex natural product synthesis. Although both reagents are commercially available, they have so far found relatively little use in carbohydrate chemistry for oxidation of primary alcohols to aldehydes [44,45].

Unprotected or partially protected carbohydrates cannot generally be oxidized to the aldehyde using the chemical methods described above. Instead, an enzymatic oxidation can be used in some cases. Galactose oxidase catalyzes the oxidation of certain primary alcohols to aldehydes (*Scheme 4*) [46]. The enzyme is commercially available and is usually isolated from the fungus *Dactylium dendroides*. It is very specific for D-galactose and its derivatives including D-galactopyranosides, 2-deoxy-D-galactose, *N*-acetyl-D-galactosamine, D-talose, and a few alditols [46,47]. Other hexoses or pentoses do not undergo oxidation in the presence of galactose oxidase. The enzyme catalase is usually also added in the oxidation to decompose the hydrogen peroxide liberated during the reaction.

2.3 Oxidation of Primary Alcohols to Carboxylic Acids

Oxidation of primary alcohols with stronger oxidizing agents gives carboxylic acids. Several methods are illustrated for the oxidation of diisopropylidenegalactopyranose **7** and -sorbofuranose **11** (*Table 3*). Oxidation of the latter is an important step in the synthesis of vitamin C (ascorbic acid) from D-glucose [49]. The electrochemical oxidation of **11** with nickel(III) oxide hydroxide has been applied on an industrial scale using various nickel electrodes and chemical reactors [52]. On a laboratory scale, however, the same oxidation can be conveniently accomplished with a catalytic amount of nickel(II) chloride and sodium hypochlorite as the stoichiometric oxidant [53]. Potassium permanganate and ruthenium(VIII) oxide are also strong oxidants for converting a primary alcohol to the carboxylic acid [2]. Both reagents are rather non-selective and will also oxidize olefins, sulfides, and in some cases benzyl

Table 3
Oxidation of diisopropylidenegalactopyranose 7 and -sorbofuranose 11 to carboxylic acids 10 and 12



Substrate	Reagent	Solvent	Yield (%)	Reference
7	$KMnO_4$, NaOH, Bu_4NBr	H_2O/CH_2Cl_2	85	[48]
11	$KMnO_4$, KOH	H_2O	91	[49]
7	$RuCl_3$, $NaIO_4$	H_2O/MeCN/$CHCl_3$	82	[50]
7	RuO_2, NaOCl[a]	H_2O/CCl_4	86	[51]
11	RuO_2, NaOCl[a]	H_2O/CCl_4	83	[51]
7	NiO(OH)[a], KOH	H_2O	93	[26]
11	NiO(OH)[a], KOH	H_2O	90–96	[52]
11	$NiCl_2$, NaOCl, NaOH	H_2O	90	[53]

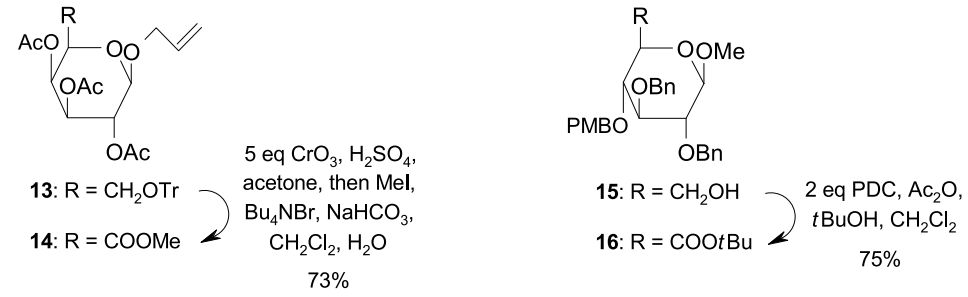
[a]Generated electrochemically

ethers. Ruthenium(VIII) oxide is used catalytically and a biphasic solvent system of carbon tetrachloride, acetonitrile, and water has proven to be beneficial for the oxidation [29,54] although successful ruthenium(VIII) oxide oxidations have also been accomplished in aqueous acetone [55].

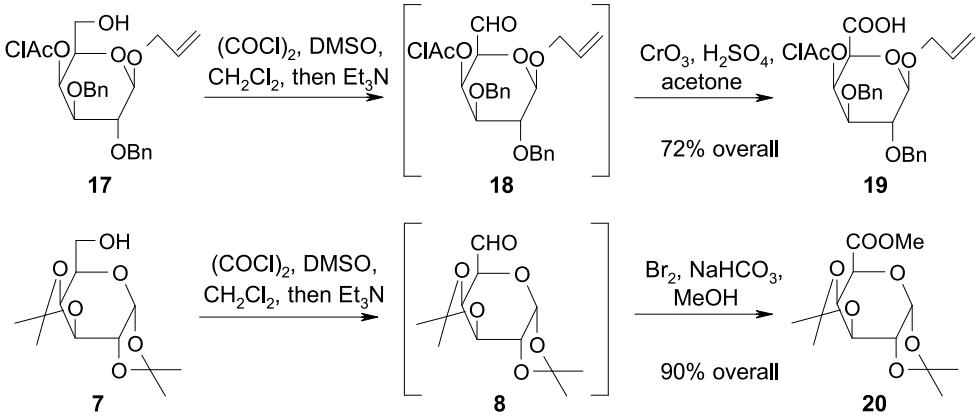
When acid labile protecting groups are not present in the substrate, the Jones oxidation [chromium(VI) oxide, sulfuric acid] can be applied for preparation of uronic acids. Isopropylidene acetals are normally cleaved to some extent under these conditions [56]. The method usually requires an excess reagent (2–5 equiv.) to drive the oxidation to completion. For example, the Jones oxidation of methyl and allyl 2,3,4-tri-*O*-benzyl-α-D-glucopyranoside occurs with 2 equiv. of reagent to give the uronic acids in good yields [57]. Cleavage of acid labile protecting groups during the reaction can in some cases be an advantage. Worthy of note is the direct oxidation of trityl ether **13** to the uronic acid isolated as methyl ester **14** (*Scheme 5*) [56]. The Jones oxidation can also be used on thioglycosides without concomitant oxidation at sulfur [58].

Another chromium(VI) oxidant for the preparation of uronic acids is PDC [28]. The oxidation is carried out in an aprotic solvent like dimethylformamide (DMF) or dichloromethane. Acetals and sulfides are stable under these conditions [59]. Although PDC is also used for oxidation of primary alcohols to aldehydes, use of a larger excess and/or a longer reaction time will give the carboxylic acid [60]. PDC can be further activated by addition of acetic anhydride which will shorten the reaction time [61]. If *tert*-butanol is added to the reaction, the *tert*-butyl ester can be obtained directly as shown by the conversion of **15** into **16** (*Scheme 5*) [62]. Presumably, the intermediate aldehyde forms a hemiacetal with *tert*-butanol which is then further oxidized to the ester.

In some cases a two-step protocol is a milder procedure for oxidation of a primary alcohol to a carboxylic acid. The first step is then usually a Swern oxidation of the alcohol to the



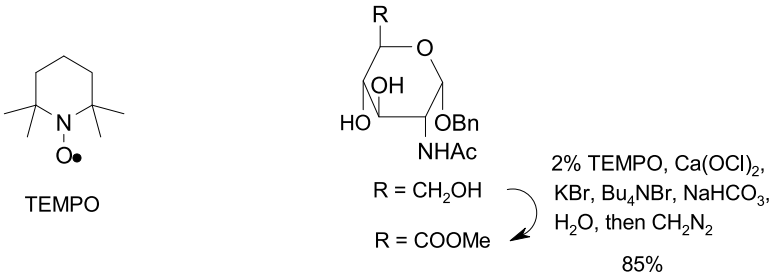
Scheme 5



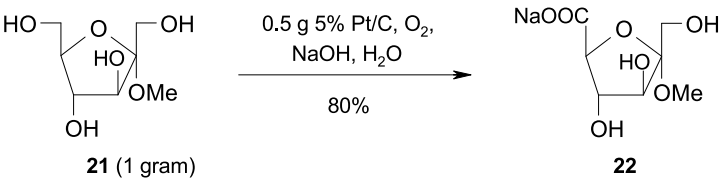
Scheme 6

aldehyde. This is not purified, but taken on directly to the next step. Oxidation of the aldehyde to the uronic acid is now an easier task than the direct oxidation of the starting alcohol. For example, Jones oxidation of alcohol **17** is a sluggish reaction accompanied by significant chloroacetyl migration [63]. However, Jones oxidation of aldehyde **18** proceeds readily to give uronic acid **19** in good overall yield (*Scheme 6*) [63]. The latter aldehyde oxidation can also be achieved effectively with sodium chlorite [64] which has been applied in the oxidation of complex oligosaccharides [65]. By addition of bromine and methanol to the aldehyde, the methyl ester is obtained directly, e. g., **7** → **20** (*Scheme 6*) [66]. The reaction proceeds through the hemiacetal which is more readily oxidized than the starting aldehyde or methanol. In this reaction bromine can be replaced with PDC which has been used for oxidation of thioglycosides that do not tolerate treatment with bromine [67].

Unprotected or partially protected glycosides cannot generally be oxidized to uronic acid by the above-described methods. However, for unprotected carbohydrates milder and more selective oxidants have been developed that take advantage of the primary alcohol function being more sterically accessible. An important reagent for this transformation is 2,2,6,6-tetramethyl-1-piperidinyloxy (TEMPO) which is a shelf-stable and commercially available nitrosyl radical



Scheme 7



Scheme 8

(*Scheme 7*) [68,69]. It is soluble in water and used catalytically with sodium or calcium hypochlorite as the stoichiometric oxidant in the presence of bromide ions. Benzyl ethers, ester groups, olefins, and azides are stable to the TEMPO oxidation conditions [69,70] while thioglycosides may undergo oxidation at sulfur [58]. TEMPO is very regioselective for oxidation of the primary hydroxy group and has been successfully applied for oxidation of naturally occurring polysaccharides [71]. A drawback with the TEMPO procedure is the need for several inorganic salts in the reaction mixture that can be difficult to remove in the work-up. This can be circumvented by using (diacetoxyiodo)benzene as the stoichiometric oxidant in a dichloromethane/water mixture [72]. Under these conditions, thioglycosides can be converted into uronic acids without simultaneous oxidation to the sulfoxide/sulfone.

Another method for selective oxidation of carbohydrate primary alcohols involves platinum-catalyzed oxidation with oxygen in water [73]. Typically, Pt/C is used as the platinum source. The hemiacetal group at the anomeric center is preferentially oxidized, but when this is blocked the primary hydroxy group will undergo oxidation with very high selectivity over the secondary hydroxy groups [74]. Even if two primary hydroxy groups are present, as in methyl α-D-fructofuranoside **21**, the sterically most accessible group will undergo oxidation in high yield (*Scheme 8*) [75]. Olefins are stable to the reaction conditions while amines and sulfides are catalyst poisons [73]. The oxidation is best carried out around neutral pH, and a base is usually added during the reaction to neutralize the acid as it is formed. The reaction is easily worked up as the catalyst is removed by filtration and no other inorganic salts are needed. However, the oxidation does require a large amount of platinum catalyst as is also evident in the oxidation of **21**. Even more problematic is the oxidation of more hindered hydroxy groups. Oxidation of L-sorbopyranose to L-*xylo*-hexulosonic acid (2-keto-L-gulonic acid) requires an average 1 g of platinum metal (i. e., 10 g of 10% Pt/C) to convert 0.3–8 g of L-sorbopyranose

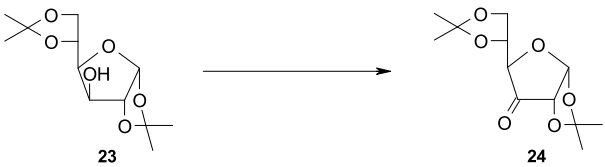
in 1 h [76]. Increased activity can be obtained by using bismuth- or lead-promoted platinum catalysts. The presence of the promoter seems to suppress poisoning of the catalyst caused by accumulation of oxygen on the metal surface. However, the promoters also change the selectivity profoundly [76], hence a general solution to the deactivation problem remains to be found.

When an aldose is not protected at C1, oxidation can occur both at C1 and at the primary position to give an aldaric acid. Strong nitric acid is the classical reagent for this oxidation. For example, treatment of D-glucose with concentrated nitric acid at 60 °C for 1 h affords D-glucaric acid isolated as the crystalline monopotassium salt in 41% yield [77]. Aldoses can also be oxidized to aldaric acids by the platinum-catalyzed oxidation with oxygen, but the need for relatively large amounts of platinum generally makes this procedure less attractive [78]. More recently, a modification of TEMPO was introduced for oxidation of aldoses to aldaric acids [79]. By this procedure glucose is converted into the monopotassium salt of glucaric acid in 85% yield [79].

2.4 Oxidation of Secondary Alcohols to Ketones

A commonly used, protected carbohydrate containing a secondary hydroxy group is diisopropylideneglucofuranose **23**. Oxidation to the corresponding ketone **24** illustrates some of the most widely applied methods for oxidation of secondary alcohols (*Table 4*). Again, the reactions can be divided into three main categories: oxidations mediated by activated DMSO, oxidations with chromium(VI) oxides, and oxidations catalyzed by ruthenium oxides. For oxidations with activated DMSO the Swern procedure is the most widely used [27].

Table 4
Oxidation of diisopropylideneglucofuranose 23 to ketone 24



Reagent	Solvent	Yield (%)	Reference
DMSO, $(COCl)_2$; Et_3N	CH_2Cl_2	92	[80]
DMSO, TFAA; Et_3N	CH_2Cl_2	85	[81]
DMSO, Ac_2O	DMSO	81	[82]
PCC, 3 Å MS	CH_2Cl_2	89	[83]
PDC, AcOH, 4 Å MS	CH_2Cl_2	98	[84]
PDC, Ac_2O	CH_2Cl_2	94	[85]
$CrO_3 \cdot 2C_5H_5N$, Ac_2O	CH_2Cl_2	90	[20]
RuO_2, KIO_4	$H_2O/CHCl_3$	86	[86]
TEMPO, NaBr, NaOCl	H_2O/EtOAc	> 85	[87]
Dess–Martin periodinane	$ClCH_2CH_2Cl$	83	[88]

Trifluoroacetic anhydride (TFAA) is also a very potent activator for DMSO and concomitant trifluoroacetylation of the starting alcohol is usually not observed [27]. Both the Swern and the TFAA procedure are carried out at low temperature to prevent undesired side reactions, particularly formation of the methylthiomethyl ether. Before these two methods became developed, acetic anhydride was often used for DMSO activation. However, the oxidation under these conditions is slower and the methylthiomethyl ether byproduct is often observed [27].

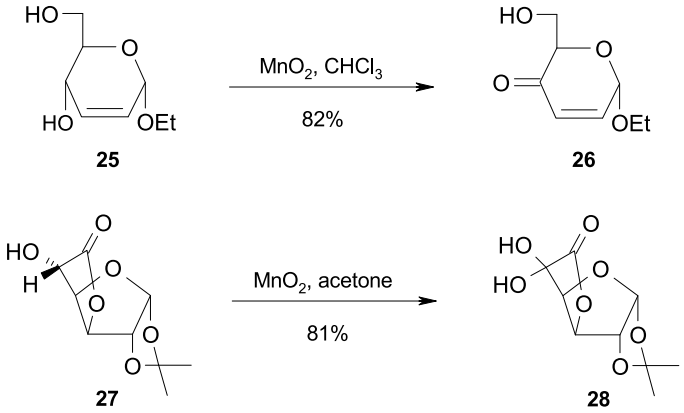
Among the chromium(VI) oxides PCC [28,89] and PDC [28,40] are preferred for oxidation of carbohydrate secondary alcohols. The reaction is in both cases accelerated by molecular sieves [90] and anhydrous acetic acid [91]. Activation of PDC can also be achieved with acetic anhydride [40]. The most widely used procedure, however, seems to be PCC and 3 Å molecular sieves in dichloromethane [28,92]. PCC is mildly acidic, but acetal protecting groups remain stable to the oxidation conditions. Chromium(VI) oxide-pyridine complex is usually not a satisfactory oxidant for carbohydrate secondary alcohols. However, further activation by acetic anhydride gives good yields of ketones with 4 equiv. of reagent [20]. Secondary alcohols in thioglycosides can be oxidized under these conditions without accompanying oxidation at sulfur [93].

If a secondary alcohol is not easily oxidized by other methods the ruthenium(VIII) oxide catalyzed procedure is often recommended. As mentioned previously, this is a strong oxidation method which is not compatible with a number of functional groups. Sodium periodate usually serves as the stoichiometric oxidant, but sodium hypochlorite has also been used in the oxidation of secondary alcohols [94]. Because of the cheap oxidants and a straightforward work-up this reaction is well suited for large-scale oxidations [95]. The TEMPO procedure also employs a cheap stoichiometric oxidant and has been applied in the oxidation of **23** on a kilogram scale [87]. The TPAP-catalyzed method is a milder procedure and many functional groups are stable to these conditions. However, secondary alcohols are still oxidized to ketones in high yield with NMO as the co-oxidant [24].

The Dess–Martin periodinane [31] has also been used for oxidation of carbohydrate secondary alcohols [88,96]. Oxidations are usually carried out under neutral conditions in dichloromethane, chloroform, or acetonitrile. However, the Dess–Martin periodinane is used stoichiometrically and as such becomes a rather expensive oxidant. As a result, it is mostly recommended for special cases where the above-described procedures are insufficient.

Manganese(IV) oxide is very slow at oxidizing isolated secondary alcohols. However, if the alcohol is allylic or alpha to a lactone manganese(IV) oxide is the reagent of choice [3,97]. The reagent has to be activated for the oxidation and a commercial sample is usually not sufficient. Activated manganese(IV) oxide is prepared as a solid from potassium permanganate and manganese(II) sulfate [98]. Oxidations can be carried out in a variety of solvents, but ether, chloroform, or acetone are usually good choices. Hereby, allylic alcohol **25** and α-hydroxylactone **27** undergo oxidation in high yield (*Scheme 9*) [99,100].

Manganese(IV) oxide will also oxidize the C3 hydroxy group in glycals [101]. However, a variety of other oxidants has also been applied for this special case (*Table 5*). Silver carbonate on Celite is a mild and neutral oxidant that also gives good yields for allylic oxidations. Because of the heterogeneous reaction conditions, an excess (5 equiv. or more) of this reagent is needed [102]. Fully protected glycals can be oxidized directly with *N*-bromosuccinimide (NBS)/benzoyl peroxide [105] or with the Koser reagent (PhI(OH)OTs) [106].



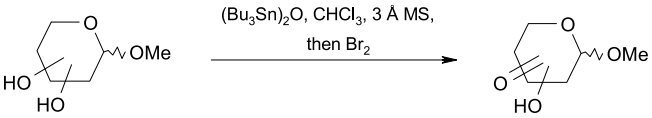
Scheme 9

Table 5
Oxidation of unprotected and protected glucals at C3

R	R′	Reagent	Solvent	Yield (%)	Reference
H	TBS	MnO_2	CH_2Cl_2	72	[101]
H	TBS	Ag_2CO_3, Celite	Benzene	86	[102]
H	H	Ag_2CO_3, Celite	Benzene	70	[103]
H	H	$(Bu_3Sn)_2O$; NIS	Benzene	60	[104]
Ac	Ac	NBS, $(BzO)_2$, K_2CO_3	CCl_4	72	[105]
Ac	Ac	PhI(OH)OTs, 3 Å MS	MeCN	52	[106]

Secondary hydroxy groups in unprotected or partially protected carbohydrates are difficult to oxidize regioselectively by the above-described procedures. Instead, some special conditions have been developed in these cases [107]. Particularly noteworthy is the brominolysis of stannyl ethers and stannylene acetals. Because of the sensitivity of the Sn–O linkage, secondary hydroxy groups can be activated with tin for a variety of regioselective reactions including oxidation to the ketone [108]. The reaction is carried out by first forming the tin compound with either bis(tributyltin) oxide or dibutyltin oxide which is then subsequently treated with bromine in situ to affect the oxidation. Bis(tributyltin) oxide was found to be superior in an extensive study on oxidation of unprotected and partially protected methyl glycopyranosides [109,110]. Oxidation of methyl pento- and hexopyranosides with this reagent and bromine gives ketoglycosides in high yield (*Table 6*). The regioselectivity is surprisingly high and is determined by the configuration of the hydroxy group being oxidized. An axial hydroxy group is always oxidized preferentially as compared to an equatorial hydroxy group. During the oxidation

Table 6
Regioselective oxidation of unprotected and partially protected methyl glycopyranosides



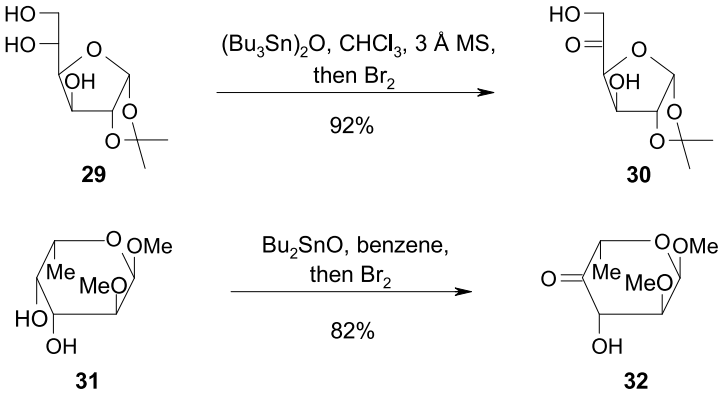
Substrate	Oxidized position	Yield (%)	Reference
Methyl α-D-glucopyranoside	C4	65[b]	[109]
Methyl β-D-glucopyranoside	C3	97	[109]
Methyl α-D-mannopyranoside	C2	53[b]	[110]
Methyl α-D-galactopyranoside	C4	70[b]	[109]
Methyl β-D-galactopyranoside	C3/C4[a]	67	[109]
Methyl α-D-allopyranoside	C3	91	[110]
Methyl β-D-allopyranoside	C3	84	[110]
Methyl α-D-altropyranoside	C3	84	[110]
Methyl β-D-altropyranoside	C3	82	[110]
Methyl α-D-xylopyranoside	C4	92	[109]
Methyl β-D-xylopyranoside	C3	93	[109]
Methyl β-L-arabinopyranoside	C4	93	[109]
Methyl 4,6-*O*-benzylidene-α-D-glucopyranoside	C2	95	[110]
Methyl 4,6-*O*-benzylidene-β-D-glucopyranoside	C3	98	[109]
Methyl 4,6-*O*-benzylidene-α-D-galactopyranoside	C3	59[b]	[110]
Methyl 4,6-*O*-benzylidene-β-D-galactopyranoside	C3	89	[110]
Methyl 4,6-*O*-benzylidene-α-D-altropyranoside	C3	76	[110]

[a] C3/C4 ratio 2/5
[b] Some unreacted starting material is also recovered

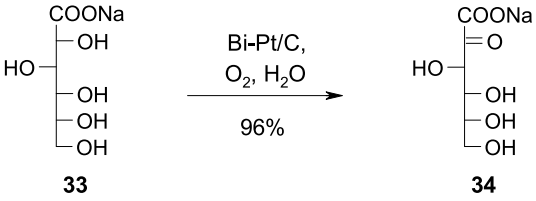
a proton is removed from the carbon bearing the secondary hydroxy group. In general, the more easily available this proton is, the more easily the secondary hydroxy group is oxidized to the corresponding ketone. This rule also applies to protected methyl glycosides, e. g., the 4,6-benzylidenehexopyranosides [109,110]. The products from these regioselective oxidations are hydroxyketones that often dimerize fairly rapidly [110].

Other partially protected carbohydrates also undergo very regioselective oxidation. Noteworthy is the oxidation of isopropylideneglucofuranose **29** to 5-ketofuranose **30** (*Scheme 10*) [109]. For oxidation of the axial hydroxy group in *cis*-1,2 diols, the dibutylstannylene acetal method is often employed. Oxidation of methyl fucoside **31** with this procedure gives ketone **32** in good yield (*Scheme 10*) [111].

The platinum-catalyzed oxidation with oxygen can also be applied for selective oxidation of secondary alcohols if no primary alcohol is present [73]. Like the tin-bromine method, axial secondary hydroxy groups will undergo preferential oxidation over equatorial hydroxy groups. However, as described above large amounts of platinum metal are required for these oxidations. Some improvement in catalyst activity has been achieved by promotion of platinum with bismuth or lead [76]. This also causes a change in selectivity and makes it possible in



Scheme 10

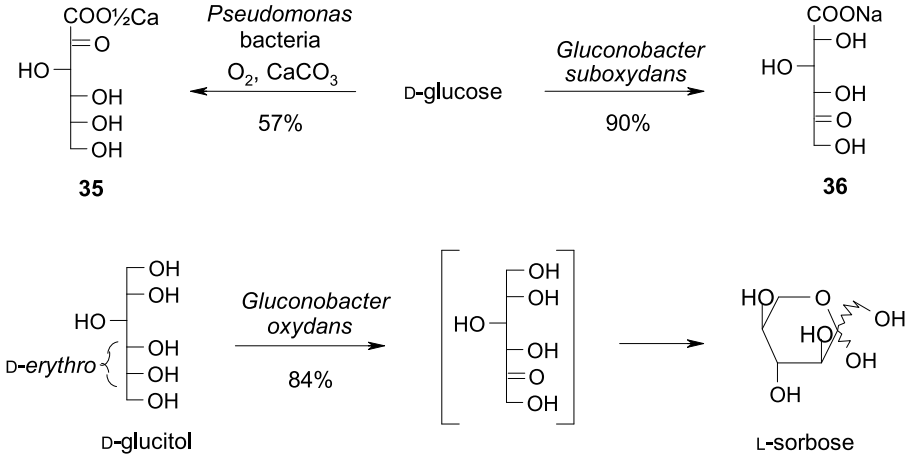


Scheme 11

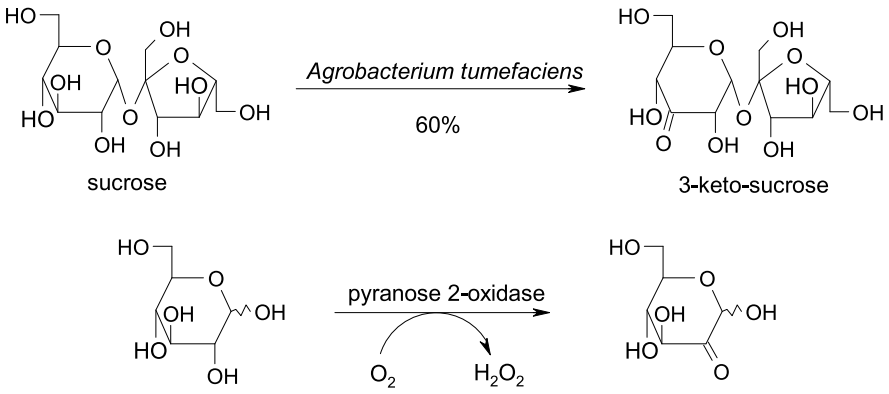
some cases to oxidize a secondary alcohol in the presence of a primary alcohol. A remarkable example is the very regioselective oxidation at C2 in aldonic acids as shown with the conversion of sodium D-gluconate **33** into 2-keto-D-gluconate **34** (*Scheme 11*) [112].

Efficient preparation of keto D-gluconates can also be achieved by fermentation. Microbial oxidation of D-glucose with various bacterial strains of the genus *Pseudomonas* produces 2-keto-D-gluconate which can be isolated by direct crystallization of the calcium salt **35** (*Scheme 12*) [113]. The same product can be obtained by fermentation with *Gluconobacter* species [114]. In fact, with this genus both 2-keto-, 5-keto-, and 2,5-diketo-D-gluconates can be formed and, depending on the strain, good selectivity for either one of the three ketogluconates can be obtained [115]. For example, 5-ketogluconate **36** can be formed in yields up to 90% with *Gluconobacter suboxydans* (*Scheme 12*) [115]. Besides oxidizing aldoses and aldonic acids *Gluconobacter* species are also known to mediate the oxidation of alditols [114]. Only alditols containing a D-*erythro* grouping adjacent to a primary alcohol will react with a reasonable growth rate. The oxidation occurs selectively at the secondary hydroxy group next to the primary alcohol (*Scheme 12*) [116]. An example is the oxidation of D-glucitol (D-sorbitol) to L-sorbose [117] which is the first step in the classical route for production of vitamin C [49].

Some pyranosides can be oxidized at C3 using the bacterium *Agrobacterium tumefaciens*. This method has been particularly successful for oxidation of disaccharides. The conversion of sucrose into 3-keto-sucrose has been studied in detail (*Scheme 13*) [118]. Lactose, maltose,



Scheme 12



Scheme 13

and cellobiose also undergo this selective C3 oxidation [119]. The active enzyme, a 3-dehydrogenase, has been isolated and purified. In addition to disaccharides this enzyme also oxidizes D-glucose, D-galactose and their methyl glycosides at C3 [120].

Selective oxidation at C2 in pyranoses can in some cases be carried out with the enzyme pyranose 2-oxidase, which can be isolated from several wood-degrading fungi [121]. The natural substrate for the enzyme is D-glucose, but significant activity towards C2 oxidation in D-galactose, D-xylose, and D-allose has also been observed (*Scheme 13*) [121]. In contrast, L-sorbose undergoes very selective oxidation at C5 [121]. In some cases, the formed 2-keto-D-aldose is not the oxidation end-product, but is accumulated in the mixture and then further converted into another product. In this way, D-glucose, D-galactose, and D-xylose have been converted into the corresponding 2,3-diketo-aldoses by oxidation with pyranose 2-oxidase from various sources [122].

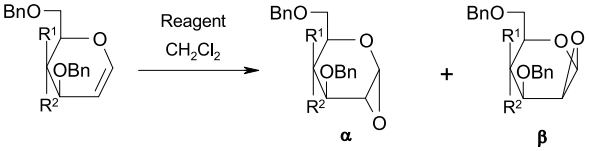
2.5 Epoxidation, Dihydroxylation, and Azidonitration of Olefins

Some unsaturated carbohydrates are easily available especially glycals and products derived from Wittig-type olefinations. The epoxidation and dihydroxylation of these substrates constitutes an important reaction in the synthesis of more complex sugars. Stereocontrolled epoxidation of glycals gives 1,2-epoxy sugars that are useful glycosyl donors [123]. Peroxy acids cannot generally be used for the epoxidation because of the high reactivity of the 1,2-epoxide which will undergo ring-opening by the acid formed during the course of the epoxidation. An exception is the *m*-CPBA-KF mixture (*Table 7*). The addition of potassium fluoride sufficiently reduces the solubility of *m*-chlorobenzoic acid to prevent the formation of products arising from epoxide ring-opening [124]. Strictly anhydrous conditions are necessary and commercial *m*-CPBA samples have to be further dried. Hereby, high yields of the desired 1,2-epoxy sugars can be obtained. The epoxidation occurs predominately from the α-face of the olefin placing the oxygen *anti* to the C3 substituent. Because of the axial C4 substituent, D-galactal derivatives show an even higher α-selectivity in the epoxidation than the D-glucal derivatives. Another reagent for the epoxidation is dimethyldioxirane [123], which has to be prepared from acetone and oxone. The byproduct of the epoxidation in this case is acetone which does not react with the product epoxide. As a result, near quantitative yields are obtained and dimethyldioxirane is often the reagent of choice for epoxidizing glycals [125].

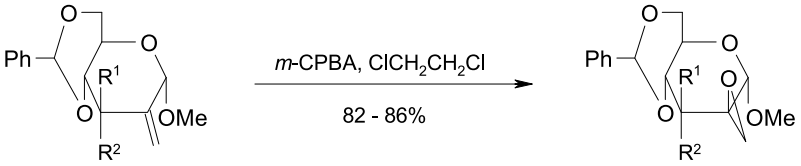
Protected *C*-methylene pyranosides like **37a**,**b** can be obtained by Wittig methylenation from the corresponding ketones (*Scheme 14*) [126]. Epoxidation to give the epoxy-branched sugars can now be carried out with a peroxy acid due to the increased stability of the formed epoxide. The electrophilic attack of the peroxy acid occurs from the less hindered face of the olefin. In this way, 2-*C*-methylene pyranosides **37a** and **37b** undergo epoxidation to give **38a** and **38b**, respectively, controlled by the axial anomeric methoxy group [126].

Hex-5-enopyranosides are available by various elimination reactions and can be epoxidized with *m*-CPBA and with a dioxirane. Like in the case with glycals the epoxides are quite sensitive to ring-opening reactions. Prolonged treatment with the epoxidizing agent in the presence of water gives hexos-5-uloses, which can be isolated in moderate to good yields [127]. For example, hex-5-enopyranoside **39** gives a 7:3 mixture of epoxides **40a**,**b** upon treatment with

Table 7
Epoxidation of 3,4,6-tri-*O*-benzyl-D-glucal and -D-galactal

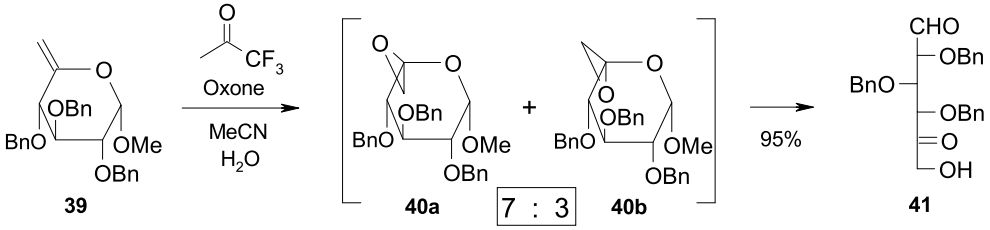


R^1	R^2	Reagent	Yield (%)	α/β-ratio	Reference
H	OBn	*m*-CPBA, KF	95	9/1	[124]
OBn	H	*m*-CPBA, KF	95	20/1	[124]
H	OBn	Dimethyldioxirane	99	20/1	[123]
OBn	H	Dimethyldioxirane	99	only α	[125]



37a: R^1 = H, R^2 = OMe
37b: R^1 = OMe, R^2 = H

38a: R^1 = H, R^2 = OMe
38b: R^1 = OMe, R^2 = H



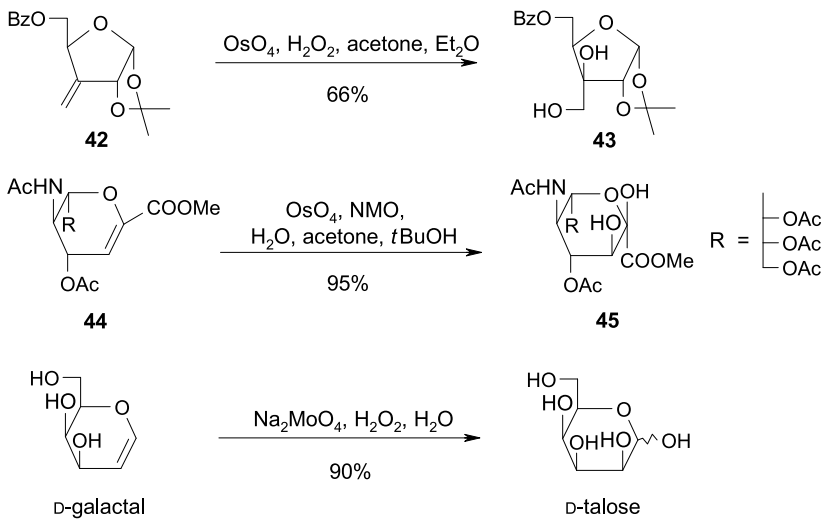
39 **40a** **40b** **41**

Scheme 14

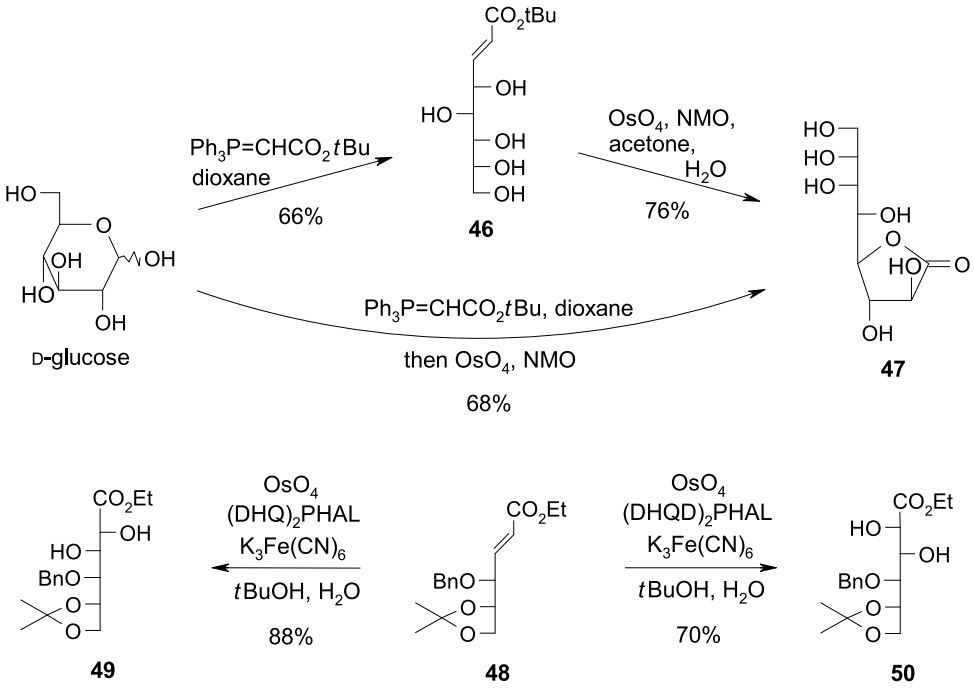
in situ generated methyl(trifluoromethyl)dioxirane while the reaction with *m*-CPBA furnishes **40a**,**b** in a ratio of 3:7 (*Scheme 14*) [127]. Further reaction with water leads to the corresponding hexos-5-ulose which after loss of methanol affords **41** in 95% overall yield.

Dihydroxylation of olefins is typically carried out with a catalytic amount of osmium(VIII) oxide in the presence of NMO as co-oxidant. If the reaction proceeds slowly a tertiary amine is sometimes added to further accelerate the dihydroxylation [128]. In cyclic systems the dihydroxylation takes place from the sterically least encumbered face of the olefin. Sugars containing an exocyclic methylene group are thus converted into branched sugars, e. g., **42** → **43** (*Scheme 15*) [129]. Aldonolactones containing a 2,3-double bond undergo dihydroxylation from the face opposite the side chain [55]. Glycals are dihydroxylated to give the 2,3-*trans* compound as the major product. For example, dihydroxylation of D-galactal with osmium(VIII) oxide gives a 4:1 mixture of D-galactose and D-talose [130]. A more interesting application is the dihydroxylation of sialic acid glycal **44** to give diol **45** in high yield (*Scheme 15*) [131]. This diol can be converted into a special glycosyl donor for α-selective sialylation using neighboring group participation [131]. Dihydroxylation of unprotected glycals from the opposite face to give 2,3-*cis* products can be carried out with catalytic molybdenum(VI) oxide and hydrogen peroxide in water [130]. The reaction presumably involves epoxidation of the double bond directed by the C3 hydroxy group followed by epoxide ring-opening with water. Particularly attractive is the conversion of D-galactal into crystalline D-talose in high yield (*Scheme 15*) [132].

Dihydroxylation of chain-extended unsaturated carbohydrates gives rise to higher sugars. Unsaturated ester **46** is available from glucose by a Wittig reaction in dioxane (*Scheme 16*) [133]. Dihydroxylation of **46** affords a 5:1 mixture of two diastereomers and the major isomer is isolated by crystallization as the octonolactone **47** [134]. Interestingly, the two-step procedure can be converted into a one-pot transformation in dioxane by performing the



Scheme 15



Scheme 16

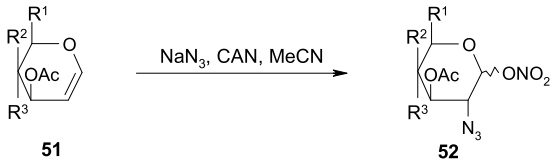
dihydroxylation immediately after the Wittig reaction. In this way, lactone **47** is obtained in a better overall yield from glucose than by isolating the intermediate ester **46**. Similar yields and diastereoselectivities are observed when galactose, arabinose, and xylose are subjected to the one-pot Wittig-dihydroxylation reaction [134]. Mannose, ribose, and lyxose, on the other hand, give a lower diastereoselectivity in the dihydroxylation [134]. In all cases, the stereochemistry of the major product is in accordance with Kishi's empirical rule [135]. This predicts that for dihydroxylation of acyclic allylic alcohols (or protected alcohols), the relative stereochemistry between the preexisting hydroxy group (or protected hydroxy group) and the adjacent, newly introduced hydroxy group in the major product is *erythro*. Kishi's rule applies to dihydroxylation of a variety of carbohydrate allylic systems [136]. The selectivity in the dihydroxylation can be improved or inverted by using a chiral ligand for osmium [137]. This is demonstrated very effectively in a recent study where all six L-hexoses are prepared from L-ascorbic acid [138]. The latter is converted into L-erythrose and L-threose which are both subjected to a Wittig reaction to afford either the (*E*)- or the (*Z*)-unsaturated ester. Each Wittig adduct is then dihydroxylated into either the 3,4-*erythro* or the 3,4-*threo* product depending on the chiral ligand for osmium. Thus, dihydroxylation of (*E*)-olefin **48** in the presence of hydroquinine 1,4-phthalazinediyl diether ($(DHQ)_2PHAL$) gives exclusively the L-altro product **49** while the reaction with hydroquinidine 1,4-phthalazinediyl diether ($(DHQD)_2PHAL$) affords only the opposite L-gluco isomer **50** (*Scheme 16*). Similar results are obtained for the corresponding (*Z*)-olefin [138].

Although osmium(VIII) oxide continues to be the most popular reagent for dihydroxylation of olefins, it does have two major drawbacks: it is very expensive and very toxic. As a result, other reagents have been investigated. Potassium permanganate has been successfully applied for dihydroxylation of the electron-deficient double bond in 2,3-unsaturated aldonolactones and –lactams [139]. Ruthenium(VIII) oxide generated from catalytic ruthenium(III) chloride and stoichiometric sodium periodate has been applied for the dihydroxylation of various protected carbohydrates with very short reaction times [140]. Although good yields can be obtained with both potassium permanganate and ruthenium(VIII) oxide, these reagents are more powerful oxidants than osmium(VIII) oxide and less chemoselective. As a result, byproducts arising from overoxidation and oxidative fission are more common.

Osmium(VIII) oxide also catalyzes the aminohydroxylation of olefins using chloramine-T as a stoichiometric oxidant and nitrogen source. Aminohydroxylation has been performed on several unsaturated carbohydrates including glycals, hex-2-eno- and hex-3-enopyranosides [141]. The yields are typically in the 60–80% range, the major byproduct being the diol. However, the reaction suffers badly from poor regioselectivity and often large amounts of both isomers are obtained.

Azidonitration represents a special aminohydroxylation reaction because the azido group often serves as a masked amino group. Azidonitration is possible only on glycals where the azido group is introduced at C2 and the nitrate at C1. The reaction is typically carried out with 1.5 equiv. of sodium azide and 3 equiv. of ceric ammonium nitrate (CAN) in acetonitrile at about −15 °C [142,143,144]. In this way, a number of acetylated glycals **51a–e** undergo fairly selective azidonitration to give azidonitrates **52a–e** (*Table 8*). The reaction works particularly well on derivatives of D-galactal [142,145,146]. In contrast, azidonitration of triacetyl D-glucal is not very selective under these standard conditions, but gives an almost equal amount of the corresponding 2-azido-D-glucose and –D-mannose products [142,147]. Howev-

Table 8
Azidonitration of acetylated glycals



51 **52**

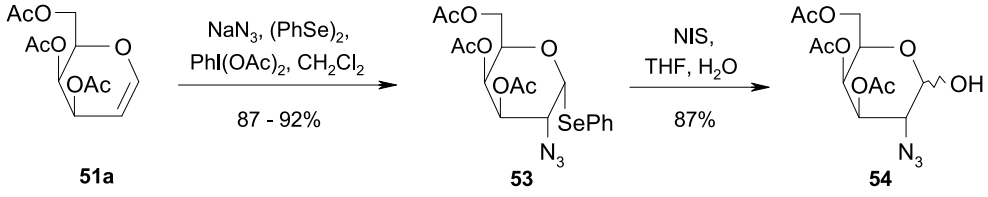
	R^1	R^2	R^3	Yield (%)	Reference
a	CH_2OAc	OAc	H	75	[142]
b	H	OAc	H	55	[143]
c	H	H	OAc	58[a]	[143]
d	COOMe	H	OAc	42[a]	[144]
e	COOMe	OAc	H	44[a]	[144]

[a]Product isolated as glycosyl acetate after treatment with NaOAc

er, increased selectivity for the 2-azido-D-mannose product can be obtained by lowering the reaction temperature to -40 °C and diluting the mixture with ethyl acetate [147]. On the other hand, use of the corresponding 4,6-*O*-benzylidene- or 4,6-*O*-isopropylidene-D-glucal show increased α-selectivity to give mainly the 2-azido-D-glucose product [148].

Although the yields obtained in the azidonitration reaction are sometimes moderate, the reaction continues to be quite important for preparation of 2-azidoglycosyl donors used in 2-aminoglycoside synthesis. In order to obtain a glycosyl donor the glycosyl nitrate can be directly converted into several glycosyl derivatives with the appropriate reagent. This includes conversion into the hemiacetal with hydrazine acetate [145], the glycosyl acetate with sodium acetate [142,143,144], the glycosyl chloride with tetraethylammonium chloride [142], or the glycosyl bromide with lithium bromide [142,146].

A closely related reaction to azidonitration is the azidophenylselenylation reaction that gives 2-azidophenylselenoglycosides from glycals. The reaction is carried out with sodium azide, diphenyl diselenide, and (diacetoxyiodo)benzene in dichloromethane at room temperature [149]. Typically, the yields for azidophenylselenylation of acetylated glycals are slightly higher than in the corresponding azidonitration reaction, e. g., **51a** → **53** (*Scheme 17*) [149]. In addition, only the α-selenoglycoside is formed. Selenoglycosides can serve directly as glycosyl donors in the preparation of *O*- and *C*-glycosides [150] or be hydrolyzed to the corresponding hemiacetal as shown for the conversion of **53** into **54** [149].



51a **53** **54**

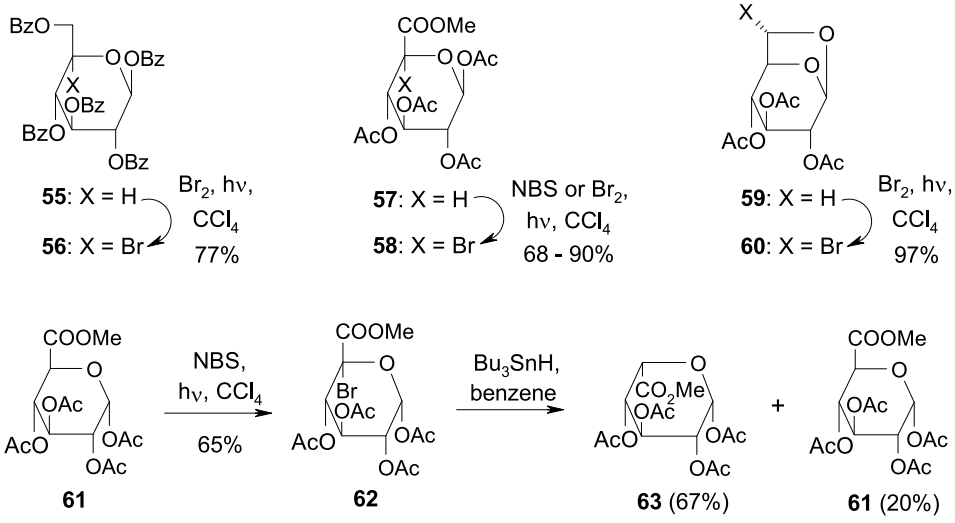
Scheme 17

2.6 Bromination at Ring Positions

In tetrahydrofuran and –pyran systems homolytic hydrogen abstraction will occur at the ether carbon atoms due to the stabilization of the developing radical by one of the oxygen lone-pairs. In aldose derivatives this hydrogen abstraction can be directed towards the non-anomeric carbon adjacent to the ring oxygen if all hydroxy groups are protected with radical destabilizing protecting groups. In this way, bromine atoms can be substituted directly onto these ring positions by free radical photobromination [151]. The reaction is carried out with bromine or NBS in refluxing carbon tetrachloride under a tungsten or heat lamp. Bromine is the most reactive brominating agent, while NBS is more selective in some cases. In addition, bromine produces hydrogen bromide during the course of the reaction, which can lead to side reactions if not trapped with an acid scavenger. The hydroxy groups are most efficiently protected with acetyl or benzoyl groups. Sometimes side reactions can occur with acetyl groups, which can undergo α-bromination to the corresponding bromoacetate [151].

Using these photobromination conditions a number of ester-protected aldose derivatives undergo very selective bromination (*Scheme 18*) [152]. Aldopyranosides (e. g., **55**) and methyl esters of uronic acids (e. g., **57**) are brominated at C5. The bromination is controlled by the anomeric effect to place bromine axial. In these D-glucopyranose derivatives the bromination occurs most readily for the β-anomers. In the corresponding α-anomers the axial anomeric protecting group makes the C5 proton less accessible, thus causing a more sluggish reaction. In 1,6-anhydropyranoses, however, the bromination occurs very selectively at C6 giving rise to the *exo*-bromide adduct, e. g., **60**.

The brominated sugars are usually quite stable and can be useful for a variety of purposes [151]. Substitution of bromine with deuterium can be used for preparation of labeled carbohydrates, while substitution with hydrogen sometimes can be used for inverting the stereochemistry. Noteworthy is the conversion of D-glucuronic acid derivative **61** into the corre-



Scheme 18

sponding L-iduronic acid **63** by bromination and subsequent debromination (*Scheme 18*) [153]. It is important that the reaction is carried out with the α-anomer. If the corresponding β-anomer **58** is subjected to the same debromination conditions, the starting D-glucuronic acid **57** is obtained as the major product [151].

3 Reductions

Numerous reduction procedures are available in organic synthesis [154] and many of these can be applied for the reduction of carbohydrates.

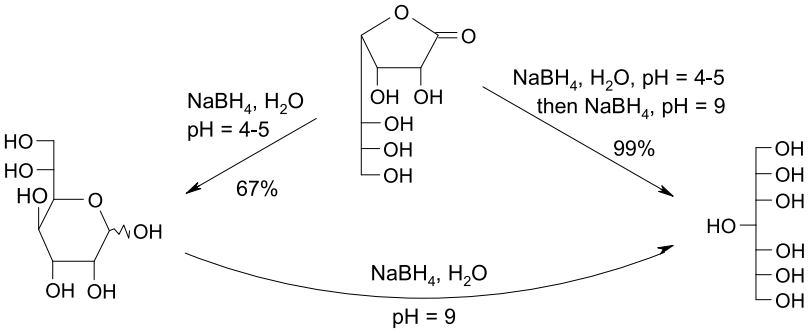
3.1 Reduction at the Anomeric Center

Unprotected aldoses and ketoses can be reduced to afford alditols while aldonolactones can be reduced to give either aldoses or alditols. The reagent of choice for reduction to alditols is sodium borohydride since it is both cheap and convenient to use. The reduction is carried out under mild conditions at room temperature in an aqueous solution. Sodium borohydride is stable in water at pH 14 while it reacts with the solvent at neutral or slightly acidic pH, but at a slower rate than the rate of carbonyl reduction. In some cases, the product will form esters with the generated boric acid. These borate complexes can be decomposed by treatment with hydrochloric acid or a strongly acidic ion-exchange resin and the boric acid can be removed in the work-up as the low boiling trimethyl borate by repeated co-evaporation with methanol at acidic pH [155].

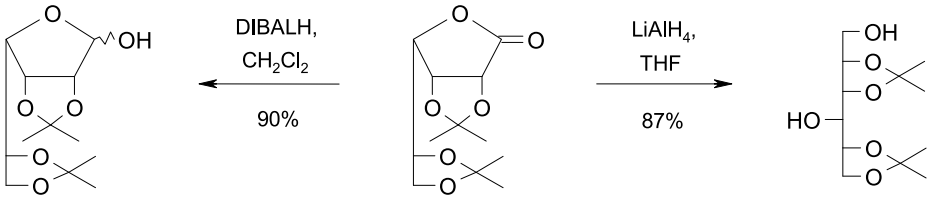
The reduction of aldoses/ketoses occurs readily with sodium borohydride and during the reaction the pH increases to about 9 (*Scheme 19*) [155]. For the reduction of aldonolactones in water the first step of the reduction has to be carried out at a pH around 5 in order to avoid ring-opening of the lactone to the corresponding sodium salt which will not react with sodium borohydride. The pH control can be achieved by performing the reduction in the presence of an acidic ion-exchange resin, e. g., Amberlite IR-120 [156]. In this way, it is possible to stop the reduction at the aldose step. Alternatively, more sodium borohydride can be added and thereby increasing the pH to 9 by which the alditol is obtained (*Scheme 19*). The reduction of aldonolactones to alditols can also be performed in anhydrous methanol or ethanol where hydrolysis of the lactone is not a side reaction [156].

Sodium borohydride only reduces aldonolactones when there is an electronegative substituent at C2. As a result, 2-deoxylactones are not reduced with this reagent, but can instead be reduced with disiamylborane in THF to the corresponding 2-deoxyaldose [157] or with calcium borohydride in ethanol to the alditol [158]. Disiamylborane is easily generated in situ by reacting borane-dimethylsulfide complex with 2-methyl-2-butene prior to addition of the lactone [157].

Sodium borohydride is not a useful reagent for large-scale industrial applications. In this case, catalytic hydrogenation in an aqueous solution over a heterogeneous catalyst is the preferred method for reducing aldoses and ketoses. The favored catalyst is Raney nickel [159] or a promoted Raney nickel [160]. The hydrogenations are typically carried out at high pressure and at temperatures around 120 °C. Lately, Ru/C and modifications thereof are gaining more attention due to a higher activity and fewer problems with metal leaching [161]. A particularly challenging example is the hydrogenation of D-fructose to afford D-mannitol where the best catalysts give D-mannitol:D-glucitol in a ratio of about 3:2 [162].



Scheme 19



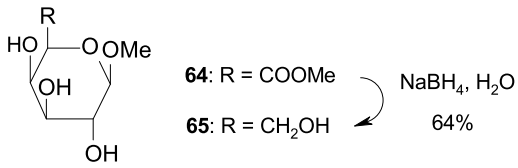
Scheme 20

Partially protected aldoses and ketoses can also be reduced with sodium borohydride to the corresponding alditol. The reduction is typically carried out in an alcoholic solvent or in a mixture of THF and water [163]. The same reduction can be achieved in a non-protic solvent with lithium aluminum hydride and diisobutylaluminum hydride (DIBALH) which are more powerful reducing agents [164]. Protected aldonolactones can be reduced with sodium borohydride in a similar manner as described above in *Scheme 19* [165]. In addition, protected aldonolactones can be reduced with DIBALH [166] and disiamylborane [167] to the corresponding aldose or with lithium aluminum hydride to the alditol (*Scheme 20*) [168]. It should be noticed that ester-protected aldonolactones can be reduced with disiamylborane in high yield without reducing the ester groups [167] while the other reducing agents only tolerate ether and acetal protecting groups. The reduction of protected aldonolactones with sodium borohydride, DIBALH, or disiamylborane gives aldoses with the same ring size as the starting lactone and this method is particularly effective for synthesizing protected aldoses in the furanose form.

3.2 Reduction of Carboxylic Acids to Primary Alcohols

Uronic acids are important components in many naturally occurring polysaccharides. By chemical or enzymatic degradation of these polysaccharides, smaller uronic acid units ranging from monosaccharides to smaller oligosaccharides can be prepared. In this connection, reduction of the carboxyl group can serve both synthetic and analytical purposes.

Reduction of the free carboxylic acid in an otherwise fully protected uronic acid can be accomplished with borane in THF [169]. For example, the borane-THF complex reduces diisopropy-

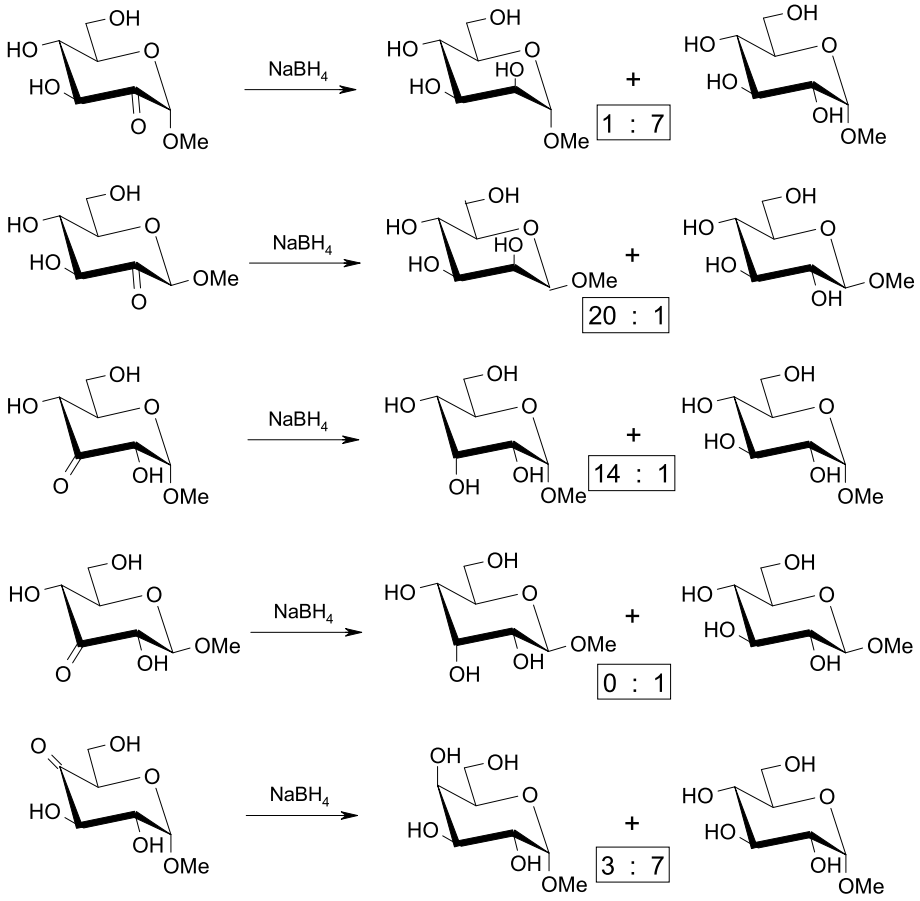


Scheme 21

lidenegalacturonic acid **10** to diisopropylidenegalactose **7** (*Table 3*) [170]. Esters of protected uronic acids are normally reduced to the corresponding alcohols with lithium aluminum hydride in ether or THF [171]. Unprotected glycosides of uronic acids are reduced with sodium borohydride in water. In this way, methyl galacturonate **64** is reduced to galactoside **65** which is isolated by crystallization after work-up with an ion-exchange resin (*Scheme 21*) [171]. Free carboxylic acids, however, do not undergo direct reduction with sodium borohydride. Instead, an initial activation of the acid is necessary. This can be conveniently done in water with the water soluble carbodiimide, 1-(3-dimethylaminopropyl)-3-ethylcarbodiimide, which reacts selectively with the free carboxylic acid [172], which is thus sufficiently activated to be reduced with sodium borohydride. The entire procedure is carried out as a one-pot process in water and the method is well suited for analysis of uronic acids in polysaccharides [172].

3.3 Reduction of Ketones to Secondary Alcohols

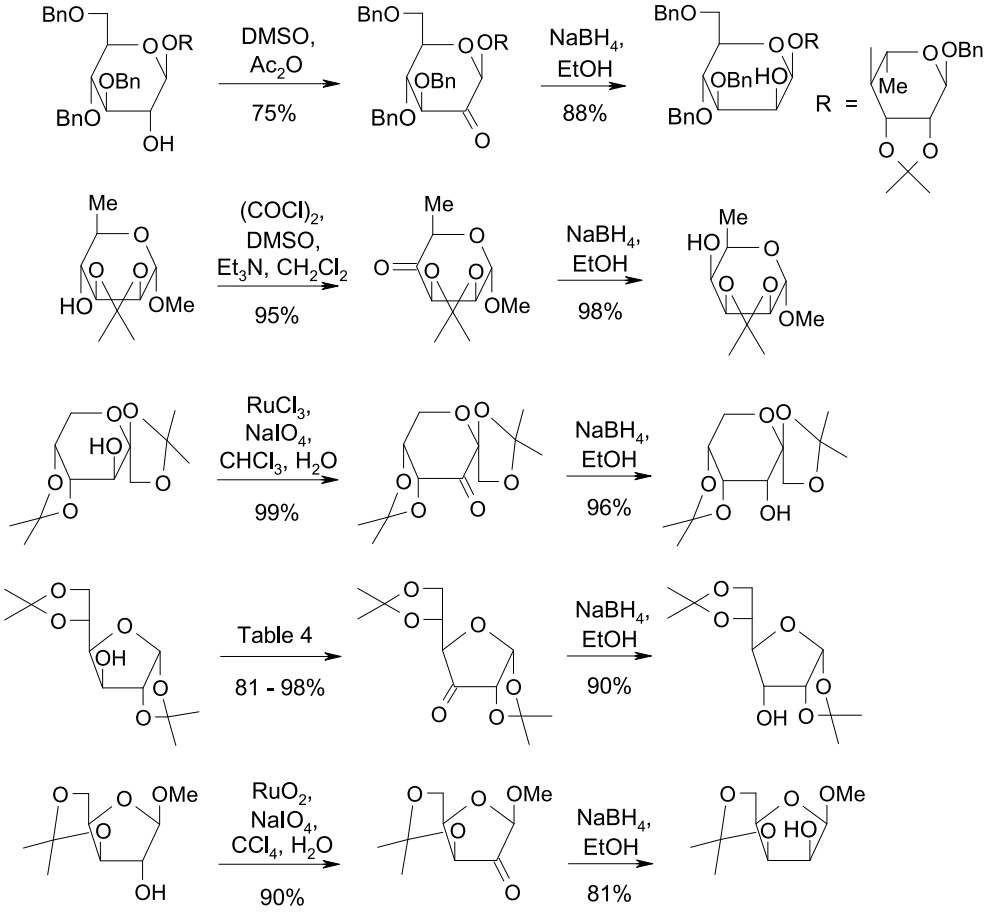
Sodium borohydride is often the reagent of choice for the reduction of carbohydrate keto groups. The reduction is typically carried out in ethanol and the stereochemical outcome depends on steric and electronic factors in the substrate. It is important to note that sodium borohydride is a sterically undemanding reagent that is perfectly capable of approaching the ketone along the seemingly more hindered axial trajectory, thus leading to the equatorial alcohol. In fact, in simple cyclohexanones this axial hydride attack is favored for electronic reasons [154]. Sodium borohydride reductions of several unprotected methyl keto-glucopyranosides is shown in *Scheme 22* [109,110]. The ratio between the axial and the equatorial product alcohol is influenced by the stereochemistry at the anomeric center. For the methyl keto-α-glucosides the reduction always occurs from the face of the ketone opposite to the axial methoxy group at C1. For the keto-β-glucosides the stereochemical outcome is less predictable since the 2-ketoglucoside gives the axial alcohol while the 3-ketoglucoside affords the equatorial alcohol as the major product. With the 2-keto-β-glucoside the reduction has paved the way for one of the more reliable procedures for the preparation of β-mannopyranosides. The glycosylation is performed with a glucosyl donor to give a β-glucoside which is then oxidized at C2 and subsequently reduced to give the β-mannoside [173,174]. For keto-hexopyranosides derived from other aldoses the stereochemical outcome is highly dependent on the substrate and is often difficult to predict. In some cases, the reduction gives complete selectivity for one product while in other cases equal amounts of both diastereomers are obtained [109,110].
Protected ketosugars undergo reduction with sodium borohydride in a similar way. However, in some cases the steric or electronic nature of the protecting groups can have an additional influence on the selectivity for the reduction [175]. Protected ketosugars are normally prepared by direct oxidation of the corresponding hydroxysugar by one of the procedures mentioned



Scheme 22

previously. Subsequent reduction of the keto group then gives the possibility for inverting the stereochemistry. Several examples are illustrated in *Scheme 23* [80,95,176,177,178]. In the β-mannoside synthesis in the first example, the protective group at C3 plays a major role [176]. Additional studies have shown that high selectivity for the β-mannoside is obtained with a benzyl ether at C3 while an ester protective group gives more of the β-glucoside [179]. Inversion by oxidation-reduction can also be an efficient protocol for large-scale synthesis of some rare sugars. This is illustrated here by the preparation of diisopropylidene-protected psicose and allose from the corresponding fructose and glucose compounds [80,95].

Several other reagents have also been used for carbohydrate ketone reductions. Some of these can cause a dramatic change in the selectivity as compared to reduction with sodium borohydride (*Table 9*) [180]. Borane is similar to sodium borohydride in the sense that it is sterically undemanding and also capable of approaching the ketone along the axial trajectory to give the equatorial alcohol. If sodium borohydride does not give a satisfactory yield of an equatorial alcohol, borane can in some cases be a better choice of reducing agent [181].



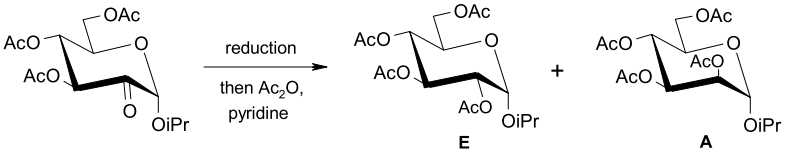
Scheme 23

Preference for the axial alcohol, however, can be obtained by heterogeneous catalytic hydrogenation or by using a sterically hindered borohydride. Catalytic hydrogenation can also be performed with the Pt/C catalyst that has been used for reduction of 2-keto-β-glucosides to β-mannosides [174]. Benzyl ethers are not affected under these conditions. Lithium tri-(*sec*-butyl)borohydride (L-selectride) is a sterically very demanding borohydride which has also found successful use for reduction to β-mannosides [182]. Other ketoglycosides have also been shown to give more of the axial alcohol when reduced with L-selectride instead of sodium borohydride [183].

3.4 Reduction of Oximes to Primary Amines

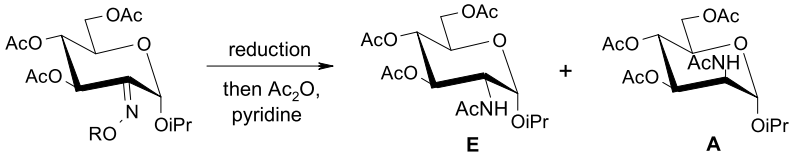
Reduction of a ketoxime is a very useful procedure for introducing a primary amine. Many reducing agents will perform this reduction, including borane, lithium aluminum hydride, and

Table 9
Reduction of 2-keto-α-D-glucoside



Reagent	Solvent	E:A
$NaBH_4$	H_2O/dioxane	96:4
$(BH_3)_2$	THF	94:6
H_2, Pd/C	EtOH	32:68
H_2, Rh/C	EtOH	20:80

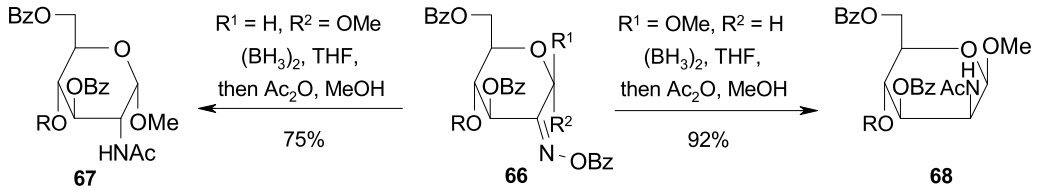
Table 10
Reduction of 2-ketoximes



R	Reagent	Solvent	E:A
H	H_2, Pd/C, HCl	MeOH	1:1
H	$LiAlH_4$	THF	2:3
H	$(BH_3)_2$	THF	7:3
Ac	$(BH_3)_2$	THF	95:5

catalytic hydrogenation over various metals [154]. The oxime carbon is not as electrophilic as a ketone carbon, and as a result sodium borohydride alone does not reduce an oxime. However, in the presence of an additive such as nickel(II) chloride, titanium(III) chloride, or titanium(IV) chloride, borohydrides will also perform oxime reductions [2]. Like other oximes, carbohydrate oximes are also usually prepared from the corresponding ketone. Oximes at C2, however, can also be prepared by nitrosochlorination of glycals followed by reaction with alcohols [184]. The stereoselectivity in the oxime to amine reduction is often similar to the one obtained in the reduction of the corresponding ketone, which has been described in the previous section.

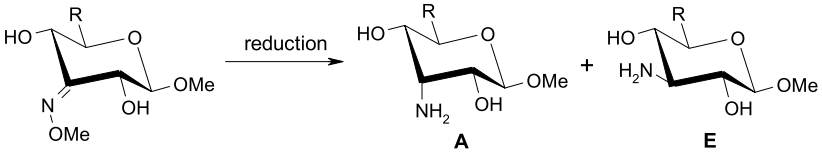
For reduction of 2-ketoximes the stereochemical outcome is determined by the reducing agent, the substituent on the oxime and the stereochemistry at the anomeric center (*Table 10*) [185]. Borane gives the best selectivity for the 1,2-*cis* product and this selectivity can be dramatically enhanced by using the acetylated oxime. This turns out to be general for reduction of glucose derived 2-ketoxime esters (*Scheme 24*) [186]. α-Glucosides of these 2-oxime esters give high yields of *N*-acetylglucosaminide derivatives (e. g., **67**) when reduced with borane followed by *N*-acetylation. On the other hand, β-glucosides give *N*-acetylmannosaminide



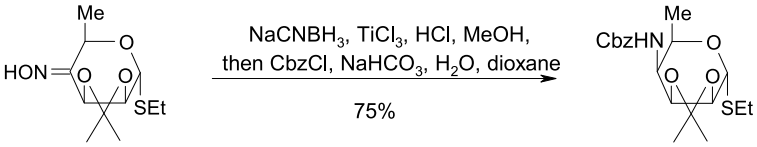
R = 2,3,4,6-tetra-*O*-benzoyl-β-D-galactopyranosyl

Scheme 24

Table 11
Reduction of 3-ketoximes



R	Reagent	Solvent	A:E	Yield (%)
H	AlH_3	THF	0:1	52
H	H_2, PtO_2	AcOH	7:3	87
CH_2OH	AlH_3	THF	1:2	61
CH_2OH	H_2, PtO_2	AcOH	7:1	67



Scheme 25

residues (e. g., **68**) very stereoselectively [186]. Although borane-THF complex is normally the reagent of choice for this reduction of 2-ketoximes, sodium borohydride/nickel(II) chloride [187] and lithium borohydride/trimethylsilyl chloride [188] have also been shown to work well. The latter mixture presumably generates borane in situ.

A study on the reduction of 3-ketoximes shows that the stereochemical outcome depends on the reagent (*Table 11*) [189]. Catalytic hydrogenation favors the axial amine while the sterically less demanding alane gives more of the equatorial amine. For reduction of 4-ketoximes the same study shows that the axial product dominates regardless of the reducing agent [189]. This has also been observed in the reduction of **69** which gives the axial product **70** almost exclusively (*Scheme 25*) [190].

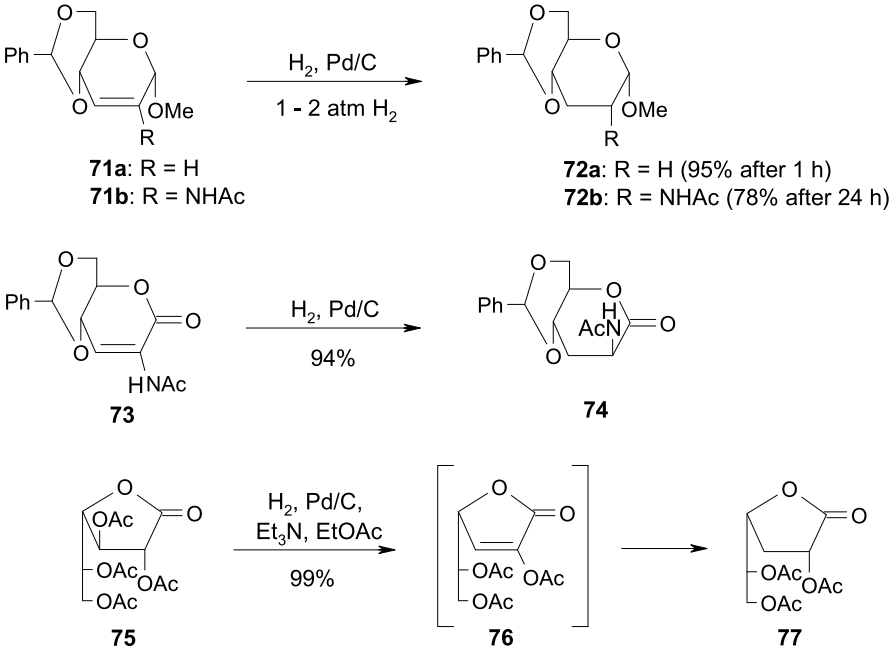
3.5 Hydrogenation of Olefins

Saturation of a carbohydrate double bond is almost always carried out by catalytic hydrogenation over a noble metal. The reaction takes place at the surface of the metal catalyst that absorbs both hydrogen and the organic molecule. The metal is often deposited onto a support, typically charcoal. Palladium is by far the most commonly used metal for catalytic hydrogenation of olefins. In special cases, more active (and more expensive) platinum and rhodium catalysts can also be used [154]. All these noble metal catalysts are deactivated by sulfur, except when sulfur is in the highest oxidation state (sulfuric and sulfonic acids/esters). The lower oxidation state sulfur compounds are almost always catalytic poisons for the metal catalyst and even minute traces may inhibit the hydrogenation very strongly [154]. Sometimes Raney nickel can

be used to remove traces of sulfur impurities prior to the hydrogenation. Raney nickel can also be used on its own as a hydrogenation catalyst although it is less reactive than the noble metals and more catalyst is required [191].

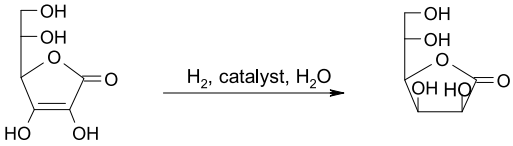
Hydrogenation over noble metals is usually performed at room temperature and at 1–3 atmospheres of hydrogen. The progress of the reaction can be monitored by measuring the hydrogen uptake. Methanol, ethanol, or ethyl acetate are normally the solvents of choice. Catalytic hydrogenation is sensitive to the steric environment around the olefin, but less sensitive to electronic factors. The hydrogenation generally occurs from the sterically less demanding face of the olefin [154].

A number of carbohydrates contain an endocyclic olefin. If this is not further substituted, hydrogenation occurs readily as shown for hex-2-enopyranoside **71a** (*Scheme 26*) [192]. A benzylidene protecting group is unaffected under these conditions. If the double bond is further substituted, as in hex-2-enopyranoside **71b**, the hydrogenation takes longer [193]. However, it occurs very selectively from the face of the olefin opposite to the axial anomeric methoxy group. Interestingly, when this methoxy group is absent, the facial selectivity reverses, as seen for hydrogenation of lactone **73** [194]. This is, however, in accordance with the general trend observed for 2,3-unsaturated aldonolactones, which preferentially undergo hydrogenation from the face of the olefin opposite to the side chain [195]. The elimination to form the unsaturated aldonolactone and the subsequent hydrogenation can be carried out as a one-pot procedure if a base is added to the reaction mixture. In this way, peracetylated aldonolactones can be hydrogenated in the presence of triethylamine to give 3-deoxylactones



Scheme 26

Table 12
Hydrogenation of ascorbic acid



78 79

78	Catalyst	H_2 Pressure (atm)	Temperature (°C)	Time	Yield (%)	Reference
23.1 g	2.2 g of 10% Pd/C	3.4	50	24 h	99	[198]
250.0 g	5 g of $PdCl_2$ on C	50.0	50	72 h	85	[199]
10.0 g	1 g of 5% Rh/C	3.7	rt	1.5 h	75–90	[200]

(*Scheme 26*) [196]. Tetra-*O*-acetyl-D-galactono-1,4-lactone **75** yields 3-deoxylactone **77** through the unsaturated lactone **76**. Because of the facial selectivity in the hydrogenation, the 2-acetate group in the product is always *cis* to the side chain. 2,3-Unsaturated aldonolactones can also be saturated by 1,4-reduction with tributyltin hydride, copper(I) iodide, and trimethylsilyl chloride [197]. Under these conditions other isolated olefins are not affected.

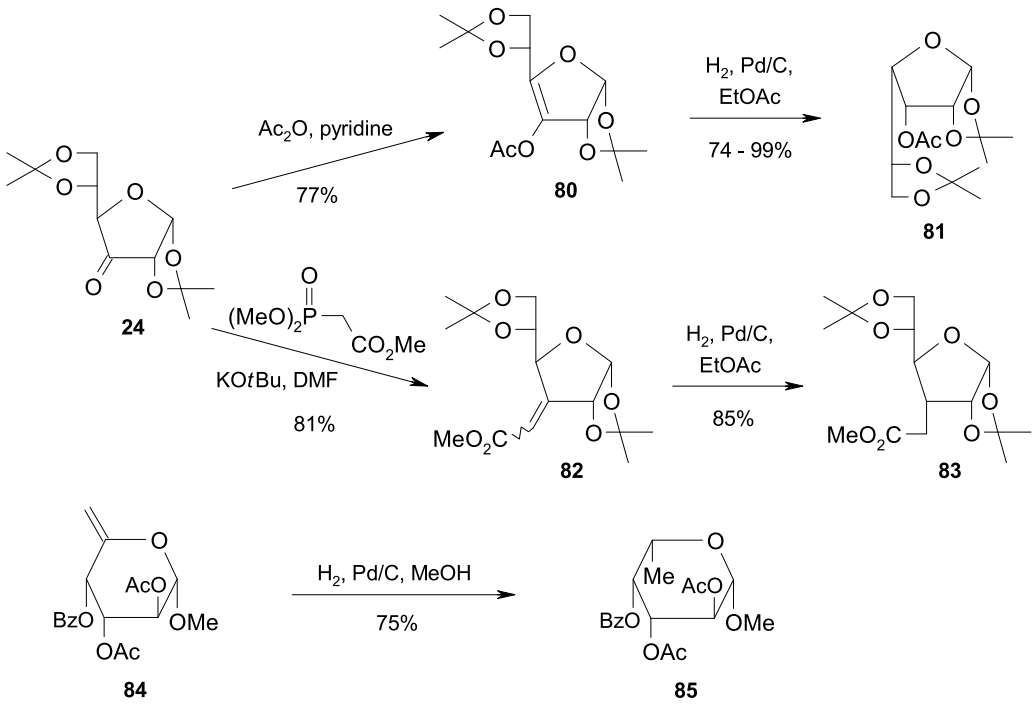
A special, but cheap, 2,3-unsaturated aldonolactone is ascorbic acid **78** which undergoes very selective hydrogenation to L-gulono-1,4-lactone **79** (*Table 12*). However, ascorbic acid is very unreactive towards hydrogenation and a relatively large amount of catalyst is needed. If Pd/C is used as the catalyst, 100 g of ascorbic acid requires about 1 g of palladium metal for the hydrogenation [198,199]. If the more reactive and also more expensive Rh/C is used, about half of that amount is needed to hydrogenate the same amount of ascorbic acid [200]. As a result, the catalyst is the most expensive reactant for hydrogenation of ascorbic acid.

Another special substrate containing an endocyclic double bond is enol acetate **80** prepared from glucose isopropylidene ketone **24** (*Scheme 27*) [201]. Hydrogenation occurs selectively from the face opposite to the 1,2-*O*-isopropylidene group to give gulofuranose **81** [85,202]. This reduction combined with the oxidation in *Table 4* can be used for conversion of D-glucose into D-gulose [201].

For carbohydrates containing an exocyclic double bond, hydrogenation will introduce a stereocenter in the ring. In this way, branched sugars can be obtained by hydrogenation of products derived from Wittig-type olefinations. For example, hydrogenation of **82**, also prepared from ketone **24**, gives the branched furanose **83** [203]. Again, the hydrogenation takes place from the face opposite to the 1,2-*O*-isopropylidene group. For hex-5-enopyranosides the hydrogenation gives 6-deoxypyranosides, e. g., hex-5-enopyranoside **84** gives rise to L-fucoside **85** as the major product (*Scheme 27*) [204]. A smaller amount (17%) of the epimeric 6-deoxy-D-altro compound is also obtained in this reaction.

4 Deoxygenations

A hydroxy group can be removed by a number of methods that usually involve a two-step process where the hydroxy group is first converted into another functional group and this group is then subsequently replaced by hydrogen [205].

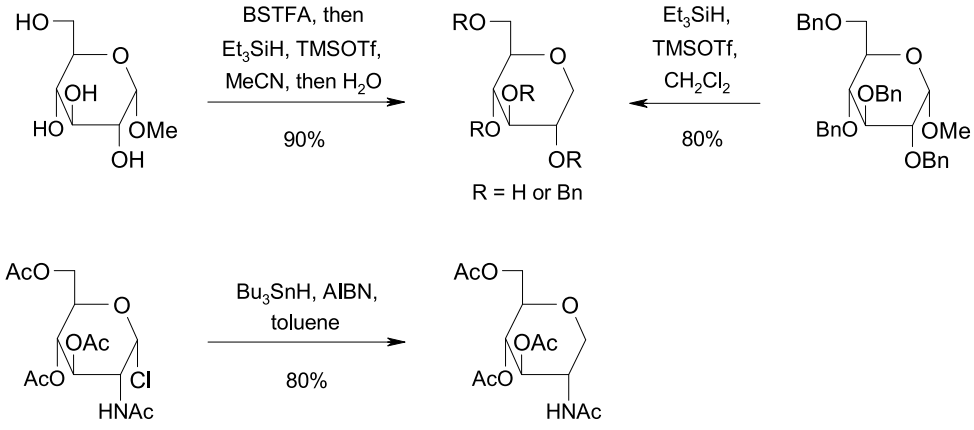

Ac2O, pyridine
77%
80
H2, Pd/C,
EtOAc
74 - 99%
81
24
(MeO)2P
CO2Me
KOtBu, DMF
81%
82
H2, Pd/C,
EtOAc
85%
83
AcO
BzO
OMe
OAc
84
H2, Pd/C, MeOH
75%
Me
AcO
BzO
OMe
OAc
85

Scheme 27

4.1 Deoxygenation at the Anomeric Center

Deoxygenation of aldoses at C1 gives rise to 1,4- and 1,5-anhydroalditols which can be used as chiral scaffolds for further synthesis. One of the preferred methods for deoxygenating aldoses and their glycosides at C1 employs triethylsilane in the presence of trimethylsilyl triflate [206]. The strong Lewis acid mediates the formation of an oxocarbenium ion at C1 which is then reduced by the silane. Under these conditions several unprotected methyl aldopyranosides and aldofuranosides are converted into 1,5- and 1,4-anhydroalditols by persilylation with BSTFA [bis(trimethylsilyl)trifluoroacetamide] followed by treatment with triethylsilane and trimethylsilyl triflate in the same pot (*Scheme 28*) [206]. Methyl hexopyranosides and methyl pentofuranosides give good yield of the corresponding 1,5-anhydrohexitols and 1,4-anhydropentitols, respectively. Some methyl pentopyranosides, on the other hand, undergo rearrangement into 1,4-anhydropentitols under the reaction conditions [206]. This rearrangement can be avoided by using the peracetylated methyl pentopyranosides or the peracetylated pentopyranoses as the starting material [207]. Besides ester-protected substrates the reaction can also be applied to ether-protected aldoses (*Scheme 28*) [208]. Furthermore, a methoxy group and an acetate are not the only groups that can be reductively cleaved from C1. A hydroxy group can be removed in aldoses that are fully protected at all positions but the anomeric center [209]. In addition, 1,2-*O*-isopropylidenefuranoses undergo reductive cleavage of the acetal to afford 1,4-anhydroalditols if borontrifluoride etherate is used as the Lewis acid instead of trimethylsilyl triflate [210].

Another well-adapted method for removing the oxygen functionality at C1 uses a radical reduction of a protected pyranosyl halide (*Scheme 28*) [211]. Formally, this is not a deoxygenation since the hydroxy group at C1 has already been replaced by a halide. However, glycosyl halides are easily available from aldoses by a one-pot procedure [212] and combined with the radical reduction this two-step route gives easy access to a number of 1,5-anhydroalditols. 1,4-Anhydroalditols, on the other hand, are more difficult to obtain by this method since the corresponding furanosyl halides require more steps for their preparation. The rad-



Scheme 28

ical reduction is usually performed with tributyltin hydride and a catalytic amount of 2,2′-azobisisobutyronitrile (AIBN) and goes through the formation of a glycosyl radical which is stabilized by the endocyclic oxygen [213]. The radical reduction at C1 can also be carried out with titanocene borohydride which can be prepared from titanocene dichloride and sodium borohydride [214]. Furthermore, radical reductions can be achieved with protected aldoses containing a phenyl thionocarbonate at C1 [215]. Because of the toxicity of organotin compounds this reduction has been performed with a catalytic amount of tributyltin hydride in the presence of polymethylhydrosiloxane as the stoichiometric reductant [215].

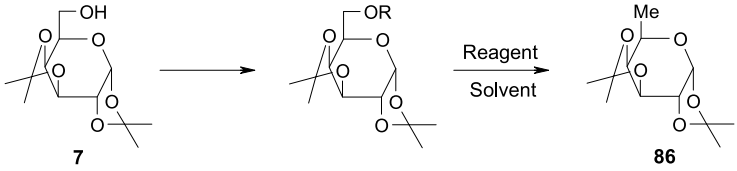
Other methods for preparing 1,5-anhydroalditols employ lithium aluminum hydride reduction of protected 1,2-epoxy pyranoses and glycosyl halides. For example, the reductive ring-opening of the glucose-derived epoxide in *Table 7* affords 1,5-anhydro-3,4,6-tri-*O*-benzyl-D-glucitol in 74% yield [216] while the reduction of 2,3,4-tri-*O*-acetyl-α-L-rhamnopyranosyl bromide gives 1,5-anhydro-L-rhamnitol in 87% yield [217]. Additionally, anhydroalditols can also be prepared by Raney nickel-mediated desulfurization of thioglycosides [218] while the same reduction on aldose dialkyl dithioacetals gives rise to 1-deoxyalditols [219].

4.2 Deoxygenation of Primary Alcohols

Diisopropylidenegalactopyranose **7** has been deoxygenated at C6 under various conditions which illustrates some of the methods that are available for removing a primary alcohol (*Table 13*). Sulfonates can be selectively introduced at the primary position in many carbohydrates and can be displaced by a hydride from either lithium aluminum hydride or sodium borohydride. Besides galactose the reduction has also been applied for removing primary sulfonates in glucose [230], mannose [231], and ribose [232]. Carboxylates, on the other hand, are not displaced by hydride, but can be removed by photolysis at 254 nm in an aqueous hexamethylphosphoric triamide (HMPA) solution [223]. The reaction goes through a radical mechanism and does not tolerate halides and other carbonyl groups in the substrate [233].

A more common radical reaction for deoxygenating alcohols is the Barton–McCombie reaction [234]. In this transformation the alcohol is converted into a thiocarbonyl derivative (xanthate, thionocarbonate, or thionocarbamate) which undergoes homolytic C-OCS cleavage upon treatment with tributyltin hydride and a radical initiator. In the original Barton–McCombie procedure secondary alcohols are treated with *N*,*N*′-thiocarbonyldiimidazole or carbon disulfide/methyl iodide/sodium hydride and the resulting thionocarbamate or xanthate is then reductively cleaved with tributyltin hydride [234]. However, these procedures proved inefficient for deoxygenating primary alcohols due to the slightly lower stability of a primary radical as compared to a secondary radical [234,235]. Instead, improved conditions for primary alcohols have been developed by acylation with 2,4,6-trichlorophenyl or 4-fluorophenyl chlorothionoformate followed by deoxygenation of the resulting thionocarbonate [224]. Although tributyltin hydride is an effective reducing agent which is compatible with esters, ethers, acetals, and olefins, organotin compounds are toxic and difficult to remove completely from the desired products. Furthermore, tributyltin hydride is rather expensive and has a limited shelf-stability. Therefore, alternative reagents have been investigated and particular attention has been given to hydrogen donors containing Si–H or P–H bonds [236]. So far, the most effective tin hydride

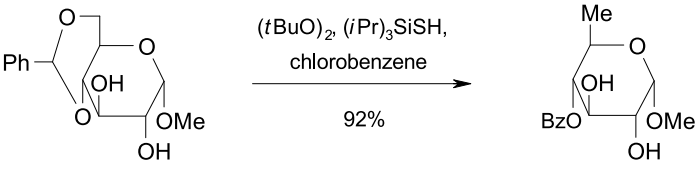
Table 13
Deoxygenation of diisopropylidenegalactopyranose 7 to diisopropylidenefucopyranose 86



R	Reagent	Solvent	Yield (%)	Reference
Ts	$LiAlH_4$	Et_2O	59	[220]
Ts	$NaBH_4$	DMSO	88	[221]
Tf	$NaBH_4$	MeCN	92	[222]
Ac	$h\nu$	H_2O/HMPA	85	[223]
C=S(O-2,4,6-Cl_3Ph)	Bu_3SnH, AIBN	Toluene	91	[224]
C=S(O-4-FPh)	$PhSiH_3$, $(BzO)_2$	Toluene	88	[224]
C=S(O-4-FPh)	Ph_3SiH, $(BzO)_2$	Toluene	88	[225]
C=S(NHPh)	TMS_3SiH, AIBN	Benzene	85	[226]
C=S(O-4-FPh)	H_3PO_2, Et_3N, AIBN	Dioxane	91	[227]
C=S(O-4-FPh)	$(MeO)_2PHO$, $(BzO)_2$	Dioxane	90	[228]
C=S(SMe)	$(Bu_4N)_2S_2O_8$, HCO_2Na	DMF	86	[229]

substitute has been tris(trimethylsilyl)silane, but unfortunately this reagent is also quite expensive. Arylsilanes, hypophosphorous acid, and dialkyl phosphites are significantly less reactive than tributyltin hydride and require longer reaction times and larger amounts of the radical initiator. However, work-up and product purification with these alternative hydrogen donors is easy and particularly the P–H reagents are much cheaper than tributyltin hydride. More recently an interesting new procedure based on tetrabutylammonium peroxodisulfate and sodium formate has been published where a range of alcohols are deoxygenated in excellent yield in less than 1 h [229].

Another radical reaction for deoxygenating the C6 position in hexoses employs the corresponding 4,6-*O*-benzylidene derivative. These acetals undergo a thiol-catalyzed radical redox rearrangement to afford 6-deoxyhexoses with a benzoate at C4 [237]. The rearrangement is initiated by thermal decomposition of a peroxide which then reacts with the thiol to generate the reactive thiyl radical. The reaction works very well with *trans*-fused acetals, as in glucose, while the *cis*-fused acetals, as in galactose, give a poor regioselectivity resulting in deoxygenation at C4 and at C6. The rearrangement tolerates a range of functional groups and can even be



Scheme 29

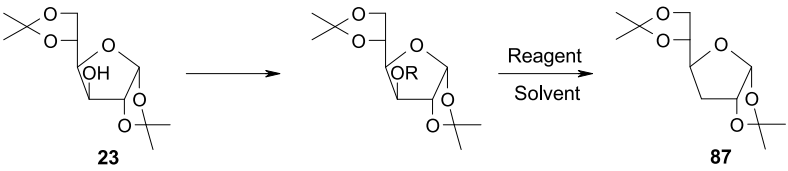
applied to a partially protected substrate (Scheme 29) [237]. A similar radical deoxygenation at C6 can be achieved by using the more complex 4,6-*O*-[1-cyano-2-(2-iodophenyl)ethylidene] acetal [238]. In this case, the redox rearrangement is mediated by tributyltin hydride and AIBN. Again, the *trans*-fused acetals in glucose and mannose undergo a very regioselective fragmentation to yield the 6-deoxy compounds while the regioselectivity with the *cis*-fused acetal in galactose is poor [238].

4.3 Deoxygenation of Secondary Alcohols

The reduction of secondary sulfonates with lithium aluminum hydride or sodium borohydride is usually a poor reaction for deoxygenating secondary alcohols [220,222]. In most cases, the hydride attack will occur at sulfur and result in cleavage of the S–O bond to afford the starting secondary alcohol as the main product. An exception from this rule is observed when tetrabutylammonium borohydride is used for reduction of secondary triflates in refluxing benzene [239]. Under these conditions clean displacement with hydride occurs to give the corresponding deoxy compounds in good yield (Table 14).

A much more common transformation for deoxygenating secondary alcohols is the Barton–McCombie reaction and various modifications of this method. Diisopropylideneglucofuranose **23** has served as a model compound for many of these deoxygenation reactions

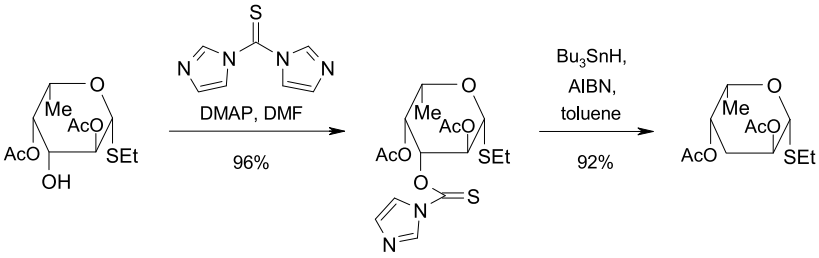
Table 14
Deoxygenation of diisopropylideneglucofuranose 23 to 3-deoxyfuranose 87



23 87

R	Reagent	Solvent	Yield (%)	Reference
Tf	Bu_4NBH_4	Benzene	84	[239]
C=S(SMe)	Bu_3SnH	Toluene	75	[240]
C=S(imidazolide)	Bu_3SnH	Toluene	74	[235]
C=S(O-2,4,6-Cl_3Ph)	Bu_3SnH, AIBN	Benzene	100	[241]
C=S(O-4-FPh)	$PhSiH_3$, $(BzO)_2$	Toluene	100	[225]
C=S(SMe)	Ph_2SiH_2, AIBN	Toluene	92	[242]
C=S(SMe)	Ph_3SiH, $(BzO)_2$	Toluene	95	[225]
C=S(OPh)	TMS_3SiH, AIBN	Toluene	81	[243]
C=S(NHPh)	TMS_3SiH, AIBN	Benzene	99	[226]
C=S(SMe)	H_3PO_2, Et_3N, AIBN	Dioxane	91	[227]
C=S(SMe)	$(MeO)_2PHO$, $(BzO)_2$	Dioxane	97	[227]
C=S(SMe)	$(Bu_4N)_2S_2O_8$, HCO_2Na	DMF	98	[229]
Bz	$Mg(ClO_4)_2$, hν[a]	H_2O/*i*PrOH	86	[244]

[a] 9-Ethyl-3,6-dimethylcarbazole is used as the photosensitizer



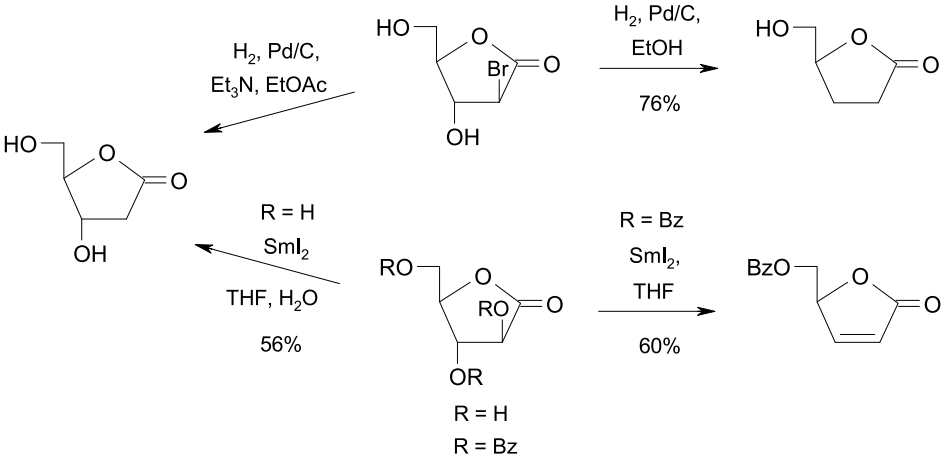
Scheme 30

(*Table 14*). The early procedures used tribultyltin hydride on the corresponding *S*-methyl xanthate or imidazoylthiocarbonyl derivative [234,235]. Later on it was shown that faster conversion and better yields could be obtained by using a radical initiator and a thionocarbonate containing an electron-withdrawing aryl group [241]. AIBN is usually employed as the initiator and together with tributyltin hydride constitute the most widely employed reagent mixture for deoxygenating secondary alcohols (*Scheme 30*) [245]. However, the toxicity of tin hydrides and the problems associated with the work-up has prompted a search for alternative reducing agents. In some applications catalytic amounts of tributyltin hydride or solid-supported tin hydrides have been used for deoxygenation of **23** [215,246]. In most cases, however, the tin hydride has been replaced by another reducing agent containing either a Si–H or a P–H bond. Several examples of these reagents are illustrated in *Table 14*. In addition, photoinduced deoxygenation of benzoates has been used for the removal of secondary alcohols [244]. The reaction is selective for benzoyl esters of secondary alcohols and cannot be used for deoxygenation of primary alcohols [247]. Recently, the photoinduced deoxygenation reaction has been applied for selective deoxygenation at C2 in aldonolactones [248].

The synthesis of 2-deoxyaldonolactones can also be achieved from the parent lactones if they contain a triflate or a tosylate at C2. The treatment of these sulfonated lactones with iodide [249], hydrazine [250], or by catalytic hydrogenolysis [251] leads to the corresponding 2-deoxyaldonolactones in good yields. When 2-bromo-2-deoxyaldonolactones are subjected to catalytic hydrogenolysis the reaction can give either the 2-deoxy- or the 2,3-dideoxylactone depending on the presence or absence of an acid scavenger (*Scheme 31*) [252]. The debromination by hydrogenolysis in the presence of triethylamine is a well-established method for the synthesis of deoxysugars [253]. The formation of the dideoxylactone, however, is an unusual transformation that seems to proceed through the 2,3-unsaturated lactone.

2-Deoxyaldonolactones can also be prepared by a samarium(II) iodide-mediated deoxygenation reaction. By this procedure a range of protected and unprotected aldonolactones undergo selective reduction at C2 with 3 equiv. of the reagent (*Scheme 31*) [254]. If the starting lactone contains an ester at C3 the reduction is accompanied by elimination to afford a 2,3-unsaturated aldonolactone [254].

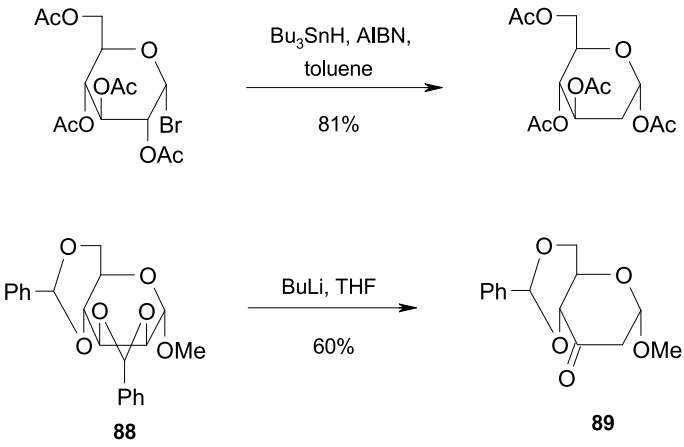
Another radical reaction for preparation of 2-deoxysugars utilizes acylated glycosyl halides as the starting material. As mentioned previously, when these glycosyl halides are treated with tributyltin hydride and AIBN the initially formed glycosyl radical is reduced to give anhydroalditols [211]. However, if tributyltin hydride and AIBN are added very slowly over



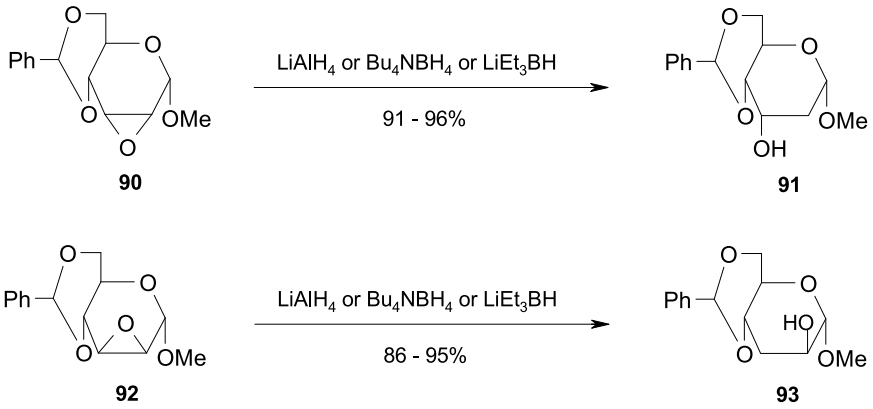
Scheme 31

several hours to a solution of the glycosyl halide, the concentration of the hydrogen donor will be low enough to ensure rearrangement of the glycosyl radical before the reduction takes place [213]. Under these conditions the acyloxy group at C2 will migrate to the anomeric center and the resulting C2 centered radical will then abstract hydrogen to afford a peracylated 2-deoxyaldose (*Scheme 32*) [255]. The slow addition can be avoided if tributyltin hydride is replaced by tris(trimethylsilyl)silane since the latter is a less effective hydrogen donor and therefore allows the rearrangement to take place before the hydrogen abstraction [256].

2-Deoxyaldoses can also be prepared from partially protected glycals by addition of water or acetic acid [257]. A special reaction for synthesis of a 2-deoxysugar employs dibenzylidene-mannoside **88** (*Scheme 32*) [258]. Reaction of **88** with butyl lithium leads to a selective



Scheme 32



Scheme 33

deprotonation at C3 and subsequent elimination of benzaldehyde to give 2-deoxy-3-keto sugar **89**.

A more general procedure for synthesis of deoxysugars uses ring-opening of an epoxide with hydride. Monosubstituted epoxides will react at the primary position while 1,2-disubstituted epoxides can react at both secondary positions. In six-membered rings the epoxide opening is controlled by the Fürst–Plattner rule and gives rise to products with a *trans*-diaxial orientation between the secondary alcohol and the incorporated hydride. Lithium aluminum hydride is often used as the reducing agent [259], but the ring-opening can also be achieved with tetrabutylammonium borohydride [239], lithium triethylborohydride [260], and in situ generated borane [128]. The ring-opening is shown in *Scheme 33* with three different reagents on 2,3-epoxides **90** and **92** [239,259,260]. For both compounds, very regioselective ring-opening is observed to give 2-deoxyglycoside **91** from *allo* epoxide **90** and 3-deoxyglycoside **93** from *manno* epoxide **92**. It should be noticed that epoxides **90** and **92** are both easily prepared from tosylates of methyl 4,6-*O*-benzylidene-α-D-glucopyranoside [261]. The regioselective reduction can also be performed with 3,4-epoxides to give deoxysugars with an axial hydroxy group at C3 or C4 [262]. Furthermore, cyclic sulfates undergo regioselective ring-opening with tetrabutylammonium borohydride in the same way as the corresponding epoxides [263].

Acknowledgement

The author thanks the Lundbeck Foundation for financial support. The Center for Sustainable and Green Chemistry is sponsored by the Danish National Research Foundation.

References

1. Collins PM, Ferrier RJ (1995) Monosaccharides – Their Chemistry and Their Roles in Natural Products. Wiley, Chichester
2. Paquette LA (ed) (1995) Encyclopedia of Reagents for Organic Synthesis. Wiley, Chichester

3. Hudlický M (1990) Oxidations in Organic Chemistry, ACS Monograph Ser No 186. ACS, Washington, DC
4. Hudson CS, Isbell HS (1929) J Am Chem Soc 51:2225; Nelson WL, Cretcher LH (1930) J Am Chem Soc 52:403; Isbell HS (1963) Meth Carbohydr Chem 2:13
5. Isbell HS, Frush HL (1993) J Res Natl Bur Stand 11:649
6. Isbell HS, Pigman WW (1937) J Org Chem 1:505
7. Isbell HS, Frush HL (1931) J Res Natl Bur Stand 6:1145; Frush HL, Isbell HS (1963) Meth Carbohydr Chem 2:14; Zinner H, Voigt H, Voigt J (1968) Carbohydr Res 7:38
8. Isaac I, Stasik I, Beaupère D, Uzan R (1995) Tetrahedron Lett 36:383
9. Besson M, Lahmer F, Gallezot P, Fuertes P, Flèche G (1995) J Catal 152:116
10. Biella S, Prati L, Rossi M (2002) J Catal 206:242
11. Wenkin M, Ruiz P, Delmon B, Devillers M (2002) J Mol Catal A: Chem 180:141
12. Ramachandran S, Fontanille P, Pandey A, Larroche C (2006) Food Technol Biotechnol 44: 185
13. Beltrame P, Comotti M, Della Pina C, Rossi M (2004) J Catal 228:282
14. Pezzotti F, Therisod M (2006) Carbohydr Res 341:2290; Pezzotti F, Therisod H, Therisod M (2005) Carbohydr Res 340:139
15. Horton D, Jewell JS (1966) Carbohydr Res 2:251
16. Manna S, McAnalley BH, Ammon HL (1993) Carbohydr Res 243:11
17. Kuzuhara H, Fletcher HG (1967) J Org Chem 32:2531
18. Buchanan JG, Smith D, Wightman RH (1984) Tetrahedron 40:119
19. Benhaddou R, Czernecki S, Farid W, Ville G, Xie J, Zegar A (1994) Carbohydr Res 260:243
20. Garegg PJ, Samuelsson B (1978) Carbohydr Res 67:267
21. Shing TKM, Gillhouley JG (1994) Tetrahedron 50:8685
22. Hanessian S, Ugolini A (1984) Carbohydr Res 130:261
23. Hall RH, Bischofberger K (1978) Carbohydr Res 65:139
24. Ley SV, Norman J, Griffith WP, Marsden SP (1994) Synthesis 639
25. Csuk R, Dörr P (1995) J Carbohydr Chem 14:35
26. Schäfer HJ, Schneider R (1991) Tetrahedron 47:715
27. Tidwell TT (1990) Org React 39:297; Tidwell TT (1990) Synthesis 857
28. Luzzio FA (1998) Org React 53:1
29. Carlsen PHJ, Katsuki T, Martin VS, Sharpless KB (1981) J Org Chem 46:3936
30. Morris Jr PE, Kiely DE (1987) J Org Chem 52:1149
31. Dess DB, Martin JC (1991) J Am Chem Soc 113:7277; Meyer SD, Schreiber SL (1994) J Org Chem 59:7549
32. Deslongchamps P, Moreau C (1971) Can J Chem 49:2465
33. Pistia G, Hollingsworth RI (2000) Carbohydr Res 328:467
34. Lichtenthaler FW, Rönninger S, Jarglis P (1989) Liebigs Ann Chem 1153
35. Yadav JS, Subba Reddy BV, Suresh Reddy C (2004) Tetrahedron Lett 45:4583
36. Lee HH, Hodgson PG, Bernacki RJ, Korytnyk W, Sharma M (1988) Carbohydr Res 176:59; Howarth GB, Lance DG, Szarek WA, Jones JKN (1969) Can J Chem 47:75
37. Cree GM, Mackie DW, Perlin AS (1969) Can J Chem 47:511
38. Midland MM, Asirwatham G, Cheng JC, Miller JA, Morell LA (1994) J Org Chem 59:4438
39. Lough C, Hindsgaul O, Lemieux RU (1983) Carbohydr Res 120:43
40. Andersson F, Samuelsson B (1984) Carbohydr Res 129:C1
41. Ward DJ, Szarek WA, Jones JKN (1972) Carbohydr Res 21:305
42. Barili PL, Berti G, D'Andrea F, Bussolo VD, Granucci I (1992) Tetrahedron 48:6273
43. Goodman JL, Horton D (1968) Carbohydr Res 6:229
44. Cronin L, Murphy PV (2005) Org Lett 7:2691
45. Abe H, Terauchi M, Matsuda A, Shuto S (2003) J Org Chem 68:7439
46. Mazur AW (1991) In: Bednarski MD, Simon ES (eds) Enzymes in Carbohydrate Synthesis, ACS Symp Ser 466. ACS, Washington, DC, p 99
47. Andreana PR, Sanders T, Janczuk A, Warrick JI, Wang PG (2002) Tetrahedron Lett 43:6525; Root RL, Durrwachter JR, Wong C-H (1985) J Am Chem Soc 107:2997; Avigad G, Amaral D, Asensio C, Horecker BL (1962) J Biol Chem 237:2736
48. Vogel C, Jeschke U, Vill V, Fischer H (1992) Liebigs Ann Chem 1171
49. Reichstein T, Grüssner A (1934) Helv Chim Acta 17:311

50. Godage HY, Fairbanks AJ (2000) Tetrahedron Lett 41:7589
51. Torii S, Inokuchi T, Sugiura T (1986) J Org Chem 51:155
52. Lyazidi HA, Benabdallah MZ, Berlan J, Kot C, Fabre P-L, Mestre M, Fauvarque J-F (1996) Can J Chem Eng 74:405; Nanzer J, Langlois S, Cœuret F (1993) J Appl Electrochem 23:477; Robertson PM, Berg P, Reimann H, Schleich K, Seiler P (1983) J Electrochem Soc 130: 591
53. Weijlard J (1945) J Am Chem Soc 67:1031
54. Hecker SJ, Minich ML (1990) J Org Chem 55:6051
55. Garner P, Park JM (1990) J Org Chem 55:3772
56. Steffan W, Vogel C, Kristen H (1990) Carbohydr Res 204:109; Betaneli VI, Ott AY, Brukhanova OV, Kochetkov NK (1988) Carbohydr Res 179:37
57. Jarosz S (1988) Carbohydr Res 183:201; van Boeckel CAA, Belbressine LPC, Kaspersen FM (1985) Recl Trav Chim Pays-Bas 104:259
58. Allanson NM, Liu D, Chi F, Jain RK, Chen A, Ghosh M, Hong L, Sofia MJ (1998) Tetrahedron Lett 39:1889
59. Westman J, Nilsson M (1995) J Carbohydr Chem 14:949
60. Vogel C, Gries P (1994) J Carbohydr Chem 13:37
61. Halkes KM, Slaghek TM, Hyppönen TK, Kruiskamp PH, Ogawa T, Kamerling JP, Vliegenthart JFG (1998) Carbohydr Res 309:161
62. Nilsson M, Svahn C-M, Westman J (1993) Carbohydr Res 246:161
63. Nakahara Y, Ogawa T (1988) Carbohydr Res 173:306
64. Chambers DJ, Evans GR, Fairbanks AJ (2005) Tetrahedron 61:7184
65. Clausen MH, Madsen R (2003) Chem Eur J 9:3821
66. Lichtenthaler FW, Jarglis P, Lorenz K (1988) Synthesis 790
67. Garegg PJ, Olsson L, Oscarson S (1995) J Org Chem 60:2200
68. De Souza MVN (2006) Mini-Rev Org Chem 3:155
69. Lin F, Peng W, Xu W, Han X, Yu B (2004) Carbohydr Res 339:1219
70. Davis NJ, Flitsch SL (1994) J Chem Soc Perkin Trans 1 359
71. Bragd PL, van Bekkum H, Besemer AC (2004) Top Catal 27:49
72. van den Bos LJ, Codée JDC, van der Toorn JC, Boltje TJ, van Boom JH, Overkleeft HS, van der Marel GA (2004) Org Lett 6:2165
73. Heyns K, Paulsen H (1962) Adv Carbohydr Chem 17:169
74. Fabre J, Betbeder D, Paul F, Monsan P (1993) Synth Commun 23:1357; Vleeming JH, Kuster BFM, Marin GB (1997) Carbohydr Res 303: 175
75. Johnson L, Verraest DL, van Haveren J, Hakala K, Peters JA, van Bekkum H (1994) Tetrahedron Asymmetry 5:2475
76. Mallat T, Brönnimann C, Baiker A (1997) J Mol Cat A: Chem 117:425
77. Mehltretter CL (1963) Meth Carbohydr Chem 2:46
78. Venema FR, Peters JA, van Bekkum H (1992) J Mol Cat 77:75
79. Merbouh N, Bobbitt JM, Brückner C (2002) J Carbohydr Chem 21:65; Merbouh N, Thaburet JF, Ibert M, Marsais F, Bobbitt JM (2001) Carbohydr Res 336:75
80. Yoshikawa M, Okaichi Y, Cha BC, Kitagawa I (1990) Tetrahedron 46:7459
81. Yoshimura J, Sato K, Hashimoto H (1977) Chem Lett 1327
82. Mazur A, Tropp BE, Engel R (1984) Tetrahedron 40:3949
83. Lankin DC, Nugent ST, Rao SN (1993) Carbohydr Res 244:49
84. Shing TKM, Wong C-H, Yip T (1996) Tetrahedron Asymmetry 7:1323
85. Legler G, Pohl S (1986) Carbohydr Res 155:119
86. Baker DC, Horton D, Tindall Jr CG (1972) Carbohydr Res 24:192
87. Bio MM, Xu F, Waters M, Williams JM, Savary KA, Cowden CJ, Yang C, Buck E, Song ZJ, Tschaen DM, Volante RP, Reamer RA, Grabowski EJJ (2004) J Org Chem 69:6257
88. Rao HSP, Muralidharan P, Pria S (1997) Indian J Chem 36B:816
89. Hollenberg DH, Klein RS, Fox JJ (1978) Carbohydr Res 67:491
90. Herscovici J, Egron M-J, Antonakis K (1982) J Chem Soc Perkin Trans 1 1967
91. Agarwal S, Tiwari HP, Sharma JP (1990) Tetrahedron 46:4417; Czernecki S, Georgoulis C, Stevens CL, Vijayakumaran K (1985) Tetrahedron Lett 26:1699
92. Wang Z-X, Tu Y, Frohn M, Zhang J-R, Shi Y (1997) J Am Chem Soc 119:11224
93. Ichikawa S, Shuto S, Matsuda A (1998) Tetrahedron Lett 39:4525

94. Gonsalvi L, Arends IWCE, Sheldon RA (2002) Org Lett 4:1659
95. Mio S, Kumagawa Y, Sugai S (1991) Tetrahedron 47:2133
96. Skaanderup PR, Madsen R (2003) J Org Chem 68:2115; Cook GP, Greenberg MM (1994) J Org Chem 59:4704; Alonso RA, Burgey CS, Rao BV, Vite GD, Vollerthun R, Zottola MA, Fraser-Reid B (1993) J Am Chem Soc 115:6666
97. Gan L-X, Seib PA (1991) Carbohydr Res 220:117
98. Henbest HB, Jones ERH, Owen TC (1957) J Chem Soc 4909; Mancera O, Rosenkranz G, Sondheimer F (1953) J Chem Soc 2189
99. Fraser-Reid B, McLean A, Usherwood EW, Yunker M (1970) Can J Chem 48:2877
100. Kitahara T, Ogawa T, Naganuma T, Matsui M (1974) Agric Biol Chem 38:2189
101. Goodwin TE, Crowder CM, White RB, Swanson JS, Evans FE, Meyer WL (1983) J Org Chem 48:376
102. Dinh TN, Khac DD, Gandolfi I, Memoria Y, Fétizon M, Prangé T (1993) Bull Soc Chim Fr 130:287
103. Tronchet JMJ, Tronchet J, Birkhäuser A (1970) Helv Chim Acta 53:1489
104. Czernecki S, Leteux C, Veyrières A (1992) Tetrahedron Lett 33:221
105. Bouillot A, Khac DD, Fétizon M, Guir F, Memoria Y (1993) Synth Commun 23:2071
106. Kirschning A (1995) J Org Chem 60:1228; Kirschning A (1998) Eur J Org Chem 2267
107. Arterburn JB (2001) Tetrahedron 57:9765
108. Grindley TB (1994) In: Kovác P (ed) Synthetic Oligosaccharides – Indispensable Probes for the Life Sciences, ACS Symp Ser 560. ACS, Washington, DC, p 51
109. Tsuda Y, Hanajima M, Matsuhira N, Okuno Y, Kanemitsu K (1989) Chem Pharm Bull 37:2344
110. Liu H-M, Sato Y, Tsuda Y (1993) Chem Pharm Bull 41:491
111. Aspinall GO, Gammon DW, Sood RK, Chatterjee D, Rivoire B, Brennan PJ (1992) Carbohydr Res 237:57
112. Abbadi A, van Bekkum H (1995) Appl Catal A: General 124:409
113. Lockwood LB (1963) Meth Carbohydr Chem 2:51
114. Deppenmeier U, Hoffmeister M, Prust C (2002) Appl Microbiol Biotechnol 60:233
115. Silverbach M, Maier B, Zimmermann M, Büchs J (2003) Appl Microbiol Biotechnol 62:92; Weenk G, Olijve W, Harder W (1984) Appl Microbiol Biotechnol 20:400; Shinagawa E, Matsushita K, Adachi O, Ameyama M (1983) J Ferment Technol 61:359
116. Hann RM, Tilden EB, Hudson CS (1938) J Am Chem Soc 60:1201
117. Lockwood LB (1962) Meth Carbohydr Chem 1:151
118. Stoppok E, Walter J, Buchholz K (1995) Appl Microbiol Biotechnol 43:706; Stoppok E, Matalla K, Buchholz K (1992) Appl Microbiol Biotechnol 36:604
119. Maeda A, Adachi S, Matsuno R (2001) Biochem Eng J 8:217; Klekner V, Löbl V, Šímová E, Novák M (1989) Folia Microbiol 34:286
120. Van Beeumen J, De Ley J (1968) Eur J Biochem 6:331
121. Giffhorn F (2000) Appl Microbiol Biotechnol 54:727
122. Volc J, Sedmera P, Halada P, Dwivedi P, Costa-Ferreira M (2003) J Carbohydr Chem 22:207; Volc J, Sedmera P, Halada P, Prikrylová V, Haltrich D (2000) Carbohydr Res 329:219; Sedmera P, Volc J, Havlícek V, Pakhomova S, Jegorov A (1997) Carbohydr Res 297:375
123. Halcomb RL, Danishefsky SJ (1989) J Am Chem Soc 111:6661
124. Bellucci G, Catelani G, Chiappe C, D'Andrea F (1994) Tetrahedron Lett 35:8433
125. Cheshev P, Marra A, Dondoni A (2006) Carbohydr Res 341:2714
126. Yoshimura J, Sato K-i, Funabashi M (1979) Bull Chem Soc Jpn 52:2630
127. Enright PM, Tosin M, Nieuwenhuyzen M, Cronin L, Murphy PV (2002) J Org Chem 67:3733
128. Andresen TL, Skytte DM, Madsen R (2004) Org Biomol Chem 2:2951; Murphy PV, O'Brien JL, Smith III AB (2001) Carbohydr Res 334:327
129. Tronchet JMJ, Tronchet J (1977) Helv Chim Acta 60:1984
130. Bílik V, Kucár Š (1970) Carbohydr Res 13:311
131. Castro-Palomino JC, Tsvetkov YE, Schmidt RR (1998) J Am Chem Soc 120:5434
132. Bílik V (1972) Chem Zvesti 26:76
133. Railton CJ, Clive DLJ (1996) Carbohydr Res 281:69
134. Jørgensen M, Iversen EH, Madsen R (2001) J Org Chem 66:4625
135. Cha JK, Christ WJ, Kishi Y (1984) Tetrahedron 40:2247
136. Prenner RH, Binder WH, Schmid W (1994) Liebigs Ann Chem 73; Brimacombe JS, Kabir AKMS (1988) Carbohydr Res 179:21

137. Marshall JA, Beaudoin S (1994) J Org Chem 59:6614
138. Ermolenko L, Sasaki NA (2006) J Org Chem 71:693
139. López-Herrera FJ, Sarabia-García F, Pino-González MS, García-Aranda JF (1994) J Carbohydr Chem 13:767; Rassu G, Casiraghi G, Spannu P, Pinna L, Fava GG, Ferrari MB, Pelosi G (1992) Tetrahedron Asymmetry 3:1035
140. Tiwari P, Misra AK (2006) J Org Chem 71:2911; Shing TKM, Tam EKW, Tai VW-F, Chung IHF, Jiang Q (1996) Chem Eur J 2:50
141. Matsumoto K, Ebata T, Matsushita H (1995) Carbohydr Res 267:187; Schulte G, Meyer W, Starkloff A, Dyong I (1981) Chem Ber 114:1809; Dyong I, Schulte G, Lam-Chi Q, Friege H (1979) Carbohydr Res 68:257
142. Lemieux RU, Ratcliffe RM (1979) Can J Chem 57:1244
143. Hashimoto H, Araki K, Saito Y, Kawa M, Yoshimura J (1986) Bull Chem Soc Jpn 59:3131
144. Darakas E, Hultberg H, Leontein K, Lönngren J (1982) Carbohydr Res 103:176
145. Toyokuni T, Cai S, Dean B (1992) Synthesis 1236
146. Broddefalk J, Nilsson U, Kihlberg J (1994) J Carbohydr Chem 13:129
147. Paulsen H, Lorentzen JP, Kutschker W (1985) Carbohydr Res 136:153
148. Seeberger PH, Roehrig S, Schell P, Wang Y, Christ WJ (2000) Carbohydr Res 328:61
149. Mironov YV, Sherman AA, Nifantiev NE (2004) Tetrahedron Lett 45:9107; Czernecki S, Ayadi E (1995) Can J Chem 73:343; Santoyo-González F, Calvo-Flores FG, García-Mendoza P, Hernández-Mateo F, Isac-García J, Robles-Diás R (1993) J Org Chem 58:6122
150. Jiaang W-T, Chang M-Y, Tseng P-H, Chen S-T (2000) Tetrahedron Lett 41:3127; SanMartin R, Tavassoli B, Walsh KE, Walter DS, Gallagher T (2000) Org Lett 2:4051
151. Somsák L, Ferrier RJ (1991) Adv Carbohydr Chem Biochem 49:37
152. Vogel C, Liebelt B, Steffan W, Kristen H (1992) J Carbohydr Chem 11:287; Ferrier RJ, Tyler PC (1980) J Chem Soc Perkin Trans 1 1528; Blattner R, Ferrier RJ (1980) J Chem Soc Perkin Trans 1 1523; Ferrier RJ, Furneaux RH (1977) J Chem Soc Perkin Trans 1 1996
153. Medakovic D (1994) Carbohydr Res 253:299
154. Hudlický M (1996) Reductions in Organic Chemistry, 2nd edn, ACS Monograph Ser No 188. ACS, Washington, DC
155. Wolfrom ML, Thompson A (1963) Meth Carbohydr Chem 2:65; Hayward LD, Wright IG (1963) Meth Carbohydr Chem 2:258
156. Frush HL, Isbell HS (1956) J Am Chem Soc 78:2844
157. Bock K, Lundt I, Pedersen C (1981) Carbohydr Res 90:7
158. Johansen SK, Kornø HT, Lundt I (1999) Synthesis 171
159. Hough L, Theobald RS (1962) Meth Carbohydr Chem 1:94
160. Gallezot P, Cerino PJ, Blanc B, Fléche G, Fuertes P (1994) J Catal 146:93
161. Hoffer BW, Crezee E, Mooijman PRM, van Langeveld AD, Kapteijn F, Moulijn JA (2003) Catal Today 79–80:35; Kusserow B, Schimpf S, Claus P (2003) Adv Synth Catal 345:289
162. Heinen AW, Peters JA, van Bekkum H (2000) Carbohydr Res 328:449
163. Fleet GWJ, Son JC, Green DSC, di Bello IC, Winchester B (1988) Tetrahedron 44:2649; Helleur R, Rao VS, Perlin AS (1981) Carbohydr Res 89:83
164. Jiang S, Singh G, Wightman RH (1996) Chem Lett 67; Doane WM, Shasha BS, Russell CR, Rist CE (1967) J Org Chem 32:1080
165. Lerner LM, Kohn BD, Kohn P (1968) J Org Chem 33:1780
166. García-Moreno MI, Díaz-Pérez P, Mellet CO, Fernández JMG (2003) J Org Chem 68:8890; Rosen T, Taschner MJ, Heathcock CH (1984) J Org Chem 49:3994
167. Pedersen C, Jensen HS (1994) Acta Chem Scand 48:222; Kohn P, Samaritano RH, Lerner LM (1965) J Am Chem Soc 87:5475
168. Fleet GWJ, Son JC (1988) Tetrahedron 44:2637
169. Baer HH, Breton RL, Shen Y (1990) Carbohydr Res 200:377
170. Danishefsky SJ, Maring CJ (1989) J Am Chem Soc 111:2193
171. Lewis BA, Smith F, Stephen AM (1963) Meth Carbohydr Chem 2:68
172. Taylor RL, Shively JE, Conrad HE (1976) Meth Carbohydr Chem 7:149
173. Liu KK-C, Danishefsky SJ (1994) J Org Chem 59:1892
174. Garegg PJ (1992) Acc Chem Res 25:575
175. Chang C-WT, Hui Y, Elchert B (2001) Tetrahedron Lett 42:7019
176. Kochetkov NK, Dmitriev BA, Malysheva NN, Chernyak AY, Klimov EM, Bayramova NE, Torgov VI (1975) Carbohydr Res 45:283

177. Eis MJ, Ganem B (1988) Carbohydr Res 176:316
178. Bischofberger K, Brink AJ, de Villiers OG, Hall RH, Jordaan A (1977) J Chem Soc Perkin Trans 1 1472; Boeyens JCA, Brink AJ, Jordaan A (1978) J Chem Research (S) 187
179. Lichtenthaler FW, Schneider-Adams T (1994) J Org Chem 59:6728
180. Lemieux RU, James K, Nagabhushan TL (1973) Can J Chem 51:27
181. Borén HB, Ekborg G, Eklind K, Garegg PJ, Pilotti Å, Swahn C-G (1973) Acta Chem Scand 27:2639
182. Lichtenthaler FW, Lergenmüller M, Peters S, Varga Z (2003) Tetrahedron Asymmetry 14:727
183. Pelyvás I, Hasegawa A, Whistler RL (1986) Carbohydr Res 146:193
184. Lemieux RU, Ito Y, James K, Nagabhushan TL (1973) Can J Chem 51:7
185. Lemieux RU, James K, Nagabhushan TL, Ito Y (1973) Can J Chem 51:33; Lemieux RU, Gunner SW (1968) Can J Chem 46:397
186. Lichtenthaler FW, Kaji E, Weprek S (1985) J Org Chem 50:3505; Kaji E, Lichtenthaler FW (1995) J Carbohydr Chem 14:791
187. Smiatacz Z, Paszkiewicz E, Chrzczanowicz I (1991) J Carbohydr Chem 10:315
188. Karpiesiuk W, Babaszek A (1990) J Carbohydr Chem 9:909
189. Tsuda Y, Okuno Y, Iwaki M, Kanemitsu K (1989) Chem Pharm Bull 37:2673
190. Smid P, Jörning WPA, van Duuren AMG, Boons GJPH, van der Marel GA, van Boom JH (1992) J Carbohydr Chem 11:849
191. Thomas SS, Plenkiewicz J, Ison ER, Bols M, Zou W, Szarek WA, Kisilevsky R (1995) Biochim Biophys Acta 1272:37
192. Albano EL, Horton D (1969) J Org Chem 34:3519
193. Calvo-Mateo A, Camarasa MJ, De las Heras FG (1984) J Carbohydr Chem 3:461
194. Varela O, Nin AP, de Lederkremer RM (1994) Tetrahedron Lett 35:9359
195. Choquet-Farnier C, Stasik I, Beaupère D (1997) Carbohydr Res 303:185; Varela OJ, Cirelli AF, de Lederkremer RM (1979) Carbohydr Res 70:27
196. Bock K, Lundt I, Pedersen C (1981) Acta Chem Scand B35:155
197. Song J, Hollingsworth RI (2001) Tetrahedron Asymmetry 12:387
198. Andrews GC, Crawford TC, Bacon BE (1981) J Org Chem 46:2976
199. Czarnocki Z, Mieczkowski JB, Ziólkowski M (1996) Tetrahedron Asymmetry 7:2711
200. Soriano DS, Meserole CA, Mulcahy FM (1995) Synth Commun 25:3263
201. Meyer zu Reckendorf W (1972) Meth Carbohydr Chem 6:129
202. Lemieux RU, Stick RV (1975) Aust J Chem 28:1799; Tronchet JMJ, Bourgeois JM (1970) Helv Chim Acta 53:1463
203. Rosenthal A, Nguyen L (1969) J Org Chem 34:1029
204. Chiba T, Tejima S (1979) Chem Pharm Bull 27:2838
205. Hartwig W (1983) Tetrahedron 39:2609
206. Bennek JA, Gray GR (1987) J Org Chem 52:892
207. Jeffery A, Nair V (1995) Tetrahedron Lett 36:3627
208. Guo Z-W, Hui Y-Z (1996) Synth Commun 26:2067
209. Nicotra F, Panza L, Russo G, Zucchelli L (1992) J Org Chem 57:2154
210. Ewing GJ, Robins MJ (1999) Org Lett 1:635
211. Bamford MJ, Pichel JC, Husman W, Patel B, Storer R, Weir NG (1995) J Chem Soc Perkin Trans 1 1181; Kocienski P, Pant C (1982) Carbohydr Res 110:330
212. Kartha KPR, Jennings HJ (1990) J Carbohydr Chem 9:777
213. Praly J-P (2001) Adv Carbohydr Chem Biochem 56:65
214. Cavallaro CL, Schwartz J (1996) J Org Chem 61:3863
215. Tormo J, Fu GC (2002) Org Synth 78:239
216. Flaherty TM, Gervay J (1996) Tetrahedron Lett 37:961
217. Ness RK, Fletcher Jr HG, Hudson CS (1950) J Am Chem Soc 72:4547
218. Lemieux RU (1951) Can J Chem 29:1079
219. Sarbajna S, Das SK, Roy N (1995) Carbohydr Res 270:93
220. Schmid H, Karrer P (1949) Helv Chim Acta 32:1371
221. Thiem J, Meyer B (1980) Chem Ber 113:3067
222. Barrette E-P, Goodman L (1984) J Org Chem 49:176
223. Pete J-P, Portella C, Monneret C, Florent J-C, Khuong-Huu Q (1977) Synthesis 774
224. Barton DHR, Blundell P, Dorchak J, Jang DO, Jaszberenyi JC (1991) Tetrahedron 47:8969
225. Barton DHR, Jang DO, Jaszberenyi JC (1993) Tetrahedron 49:2793
226. Oba M, Nishiyama K (1994) Tetrahedron 50:10193

227. Barton DHR, Jang DO, Jaszberenyi JC (1993) J Org Chem 58:6838
228. Barton DHR, Jang DO, Jaszberenyi JC (1992) Tetrahedron Lett 33:2311
229. Park HS, Lee HY, Kim YH (2005) Org Lett 7:3187
230. Hanessian S (1972) Meth Carbohydr Chem 6:190; Schmidt OT (1962) Meth Carbohydr Chem 1:198
231. Nishio T, Miyake Y, Kubota K, Yamai M, Miki S, Ito T, Oku T (1996) Carbohydr Res 280:357
232. Sairam P, Puranik R, Rao BS, Swamy PV, Chandra S (2003) Carbohydr Res 338:303
233. Portella C, Deshayes H, Pete JP, Scholler D (1984) Tetrahedron 40:3635
234. Barton DHR, McCombie SW (1975) J Chem Soc Perkin Trans 1 1574
235. Rasmussen JR, Slinger CJ, Kordish RJ, Newman-Evans DD (1981) J Org Chem 46:4843
236. Studer A, Amrein S (2002) Synthesis 835
237. Dang H-S, Roberts BP, Sekhon J, Smits TM (2003) Org Biomol Chem 1:1330
238. Crich D, Bowers AA (2006) J Org Chem 71:3452
239. Sato K-i, Hoshi T, Kajihara Y (1992) Chem Lett 1469
240. Iacono S, Rasmussen JR (1986) Org Synth 64:57
241. Barton DHR, Jaszberenyi JC (1989) Tetrahedron Lett 30:2619
242. Barton DHR, Jang DO, Jaszberenyi JC (1993) Tetrahedron 49:7193
243. Schummer D, Höfle G (1990) Synlett 705
244. Buck JR, Park M, Wang Z, Prudhomme DR, Rizzo CJ (2000) Org Synth 77:153
245. Ruttens B, Kovác P (2004) Synthesis 2505
246. Boussaguet P, Delmond B, Dumartin G, Pereyre M (2000) Tetrahedron Lett 41:3377; Neumann WP, Peterseim M (1992) Synlett 801
247. Wang Z, Prudhomme DR, Buck JR, Park M, Rizzo CJ (2000) J Org Chem 65:5969
248. Bordoni A, de Lederkremer RM, Marino C (2006) Carbohydr Res 341:1788
249. Stewart AJ, Evans RM, Weymouth-Wilson AC, Cowley AR, Watkin DJ, Fleet GWJ (2002) Tetrahedron Asymmetry 13:2667
250. Malle BM, Lundt I, Furneaux RH (2000) J Carbohydr Chem 19:573
251. Kalwinsh I, Metten K-H, Brückner R (1995) Heterocycles 40:939
252. Lundt I, Pedersen C (1986) Synthesis 1052; Bock K, Lundt I, Pedersen C (1981) Carbohydr Res 90:17
253. Horton D, Cheung T-M, Weckerle W (1980) Meth Carbohydr Chem 8:195
254. Hanessian S, Girard C, Chiara JL (1992) Tetrahedron Lett 33:573; Hanessian S, Girard C (1994) Synlett 861
255. Giese B, Gröninger KS (1990) Org Synth 69:66
256. Giese B, Kopping B, Chatgilialoglu C (1989) Tetrahedron Lett 30:681
257. Lam SN, Gervay-Hague J (2003) Org Lett 5:4219; Costantino V, Imperatore C, Fattorusso E, Mangoni A (2000) Tetrahedron Lett 41:9177; Dupradeau F-Y, Hakomori S-i, Toyokuni T (1995) J Chem Soc Chem Commun 221
258. Testero SA, Spanevello RA (2006) Carbohydr Res 341:1057
259. Jones K, Wood WW (1988) J Chem Soc Perkin Trans 1 999; Prins DA (1948) J Am Chem Soc 70:3955
260. Baer HH, Hanna HR (1982) Carbohydr Res 110:19
261. Wiggins LF (1963) Meth Carbohydr Chem 2:188
262. Hedgley EJ, Overend WG, Rennie RAC (1963) J Chem Soc 4701
263. Zegelaar-Jaarsveld K, van der Plas SC, van der Marel GA, van Boom JH (1996) J Carbohydr Chem 15:665

2.3 Heteroatom Exchange

Yuhang Wang, *Xin-Shan Ye*
The State Key Laboratory of Natural and Biomimetic Drugs, School of Pharmaceutical Sciences, Peking University, Xue Yuan Rd #38, Beijing 100083, China
xinshan@bjmu.edu.cn

In: *Glycoscience*. Fraser-Reid B, Tatsuta K, Thiem J (eds)
Chapter-DOI 10-1007/978-3-540-30429-6_5:

Abstract

Saccharides derived with a variety of functional groups are usually of great importance in many biological and chemical aspects. The chemical modifications of saccharides with heteroatoms at nonanomeric positions are reviewed in this chapter.

Keywords

Halogens; Nitrogen; Sulfur; Phosphorus; Nucleophilic substitution; Ring-opening; Addition reaction

Abbreviations

AIBN azo-bis-isobutyronitrile
CAN ceric ammonium nitrate
DAST diethylaminosulfur trifluoride
DBU 1,8-diazabicyclo[5.4.0]undec-7-ene
DEAD diethyl azodicarboxylate
DFMBA *N*,*N*-diethyl-α,α-difluoro-(*m*-methylbenzyl)amine
DIAD diisopropyl azodicarboxylate
dppb 1,4-bis(diphenylphosphino)butane
IDCP iodonium di-*sym*-collidine perchlorate
IBX 1-hydroxy-1,2-benziodoxol-3(1*H*)-one 1-oxide
NBS *N*-bromosuccinimide
PMB *p*-methoxybenzyl
SMDA sodium dihydrobis(2-methoxyethoxy)aluminate
TASF tri(dimethylamino)-sulfonium difluorotrimethylsilicate
TFAA trifluoroacetic anhydride

1 Introduction

The sugar molecule can be considered as a playground for exploratory organic synthesis, and many functional group manipulations can be carried out on the nonanomeric carbons of carbo-

hydrates. Among these reactions replacements of nonanomeric oxygen atoms by heteroatoms (e. g. halogen, nitrogen, sulfur, phosphorus, etc.) afford valuable compounds having biological activities as well as synthetic usefulness. Halogeno sugars can serve as tools for studying carbohydrate-protein interaction, and useful intermediates in the synthesis of other important sugar derivatives. Amino sugars are widely distributed in living organisms and are essential units of various antibiotics such as amino glycosides. Thiosugars exhibit biologically important properties as the substrates for many carbohydrate-related enzymes. Phosphorus derivatives have been widely used as chiral ligands for asymmetric synthesis and also as biological intermediates. Here synthetic methodologies for the preparation of these specific classes of carbohydrates will be reviewed and discussed in this chapter, as the revised version of the corresponding chapter written by Boullanger and Descotes in the previous edition of this book.

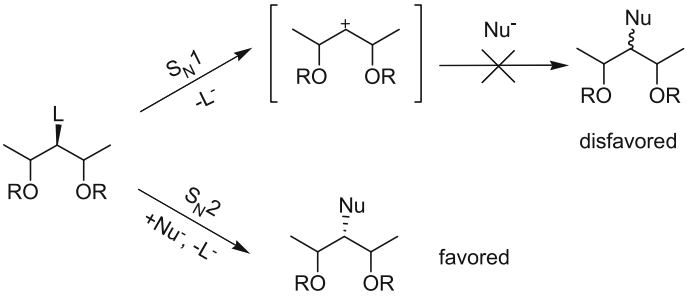
1.1 Nucleophilic Substitution

The manipulation of appropriately activated hydroxyl groups by nucleophilic displacement reactions is an indispensable tool for the introduction of functionalities directly attached to the sugar framework.

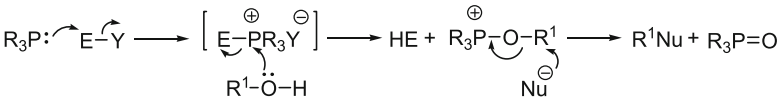
Carbohydrates are excellent substrates to test nucleophilic displacement reactions with a variety of heteroatom-based nucleophiles. These reactions can be tried on primary and secondary hydroxyl groups at different sites of the sugar ring, and with different steric or stereoelectronic dispositions [1].

The substitution reactions at the anomeric carbon usually proceed easily and via a S_N1 mechanism in most cases. However, it is not the case in the nucleophilic substitutions at nonanomeric sites. Because of the presence of vicinal electron-withdrawing substituents (OR or NHR) which strongly destabilize the intermediate carbocations, S_N2 displacement reactions instead of S_N1 reactions are favored (*Scheme 1*).

Besides displacement reactions, direct replacement of hydroxyl groups by heteroatoms is also an attractive general approach to afford heteroatom-substituted sugar derivatives. A general and important method is based on the reaction of an activated triphenyl phosphine derivative with a sugar alcohol to give an alkylphosphonium ion, which in turn is attacked by a nucleophile to give the substituted product with the configuration inversed. The plausible mechanism is depicted in *Scheme 2* [2].



Scheme 1



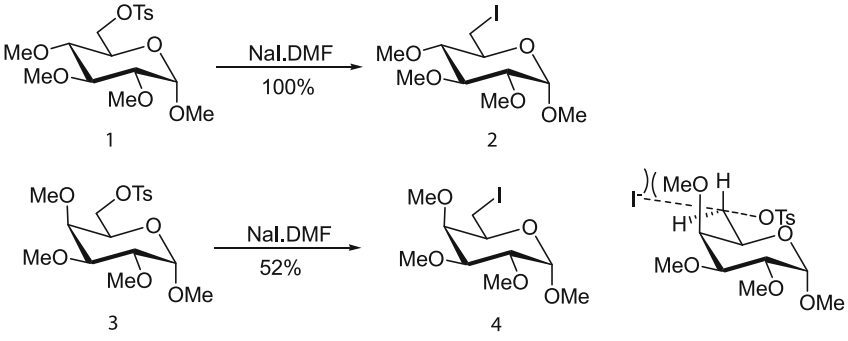
Scheme 2

In addition to the phosphorus-containing reagents, other reagents, such as the Vilsmeier–Haack reagent [$(Me_2N^+{=}CHX)X^-$] [3], sulfuryl chloride [4] and diethylaminosulfur trifluoride (DAST) [5] have also been employed for direct replacement of hydroxyl groups in carbohydrates.

Classical substitution reactions usually involve the introduction of leaving groups, among which sulfonic esters provide a versatile and simple method for activating hydroxyl groups of the carbohydrate for a bimolecular displacement reaction. The most common sulfonate leaving groups consist of mesylates and tosylates. The use of *p*-bromobenzenesulfonates (brosylates), which are ten times more reactive than tosylates, has been occasionally reported [6]. Triflates and imidazole-1-sulfonates (imidazylates) [7] are used in the substitutions at positions where other sulfonates are known to be ineffective. Halogens can also be used as leaving groups.

The efficiency of a displacement critically depends on the position of the leaving group and the chemical environment of the sugar ring (the steric- or stereoelectronic dispositions of substituents). Generally displacement at C-6 (primary position) can proceed under relatively milder conditions. Primary sulfonates of hexopyranosides are readily displaced by nucleophiles provided that the C-4 oxygen is in an equatorial orientation (e. g. 6-*O*-sulfonates of D-glucosides). And the analogous reactions in the D-galactopyranose series are particularly sluggish, presumably because of the polar, repulsive forces in the transition state involving lone pairs of electrons on the axial O-4 and the ring-oxygen atom. The examples are illustrated in *Scheme 3* by the conversion of **1** and **3** to **2** and **4**, respectively [8,9].

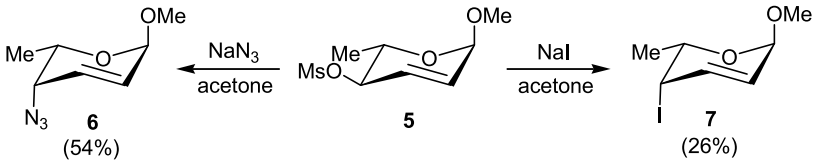
Displacements at C-3 or C-4 (secondary position) proceed with more difficulty than those at C-6. Much more drastic conditions (long reaction time, high temperatures, and polar aprotic solvents) are necessary. This mainly results from the influence of steric and polar factors from other groups in the carbohydrate ring. The difference of substitution reactivity may probably rely on the orientations of leaving groups. Axial leaving groups are substituted faster (about three times) than equatorial ones.



Scheme 3

The displacements at C-2 are very difficult owing to the vicinal C-1 bearing two electron-withdrawing oxygens, which retard the departing of the leaving group. This substitution depends strongly on the orientations of the anomeric substituents. When C-3 also bears an electron-withdrawing group, the displacement at C-2 of α-glycosides can hardly proceed, while the corresponding reaction at C-2 of β-glycosides is much more facile.

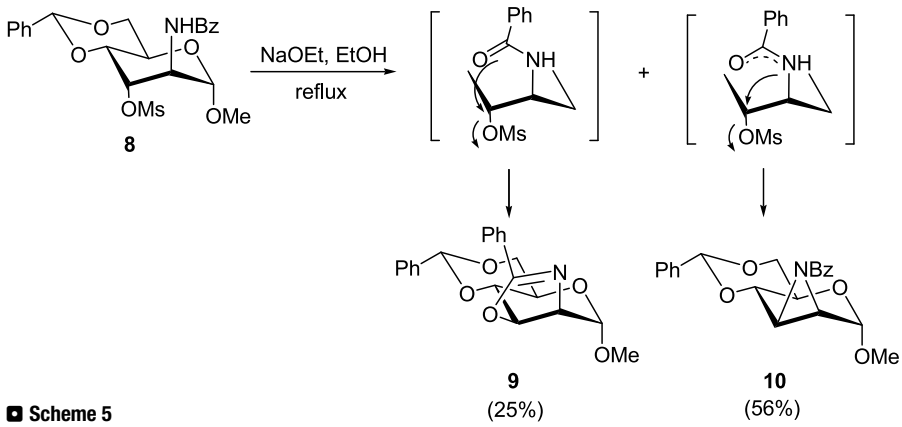
Furthermore, S_N2 displacements in furanoid rings are easier in comparison with pyranoid rings, since the more flexible five-membered rings have a smaller increasing strain, favoring the formation of the transition state. In unsaturated rings, substituents at the allylic positions are particularly susceptible to nucleophilic displacement due to the presence of double bonds. For example, mesylate compound **5** can be converted to the corresponding azide **6** and iodide **7**, when treated with NaN_3 and NaI, respectively (*Scheme 4*). Both reactions are easily realized in acetone at room temperature [10,11].



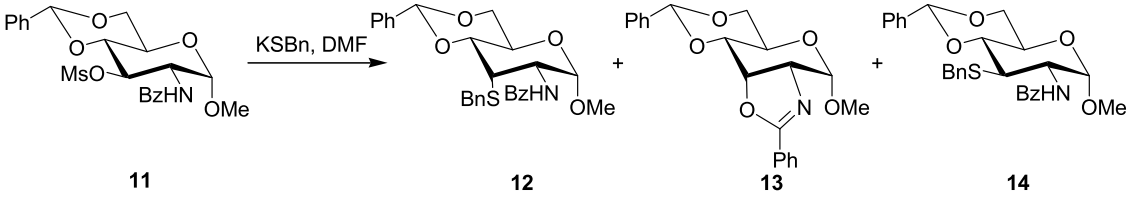
Scheme 4

Effective S_N2 substitution reactions are also greatly effected by anion nucleophilicity. An anion will be less nucleophilic when it is effectively solvated and when it is restricted by its counterion. This can be circumvented by selection of a tetraalkylammonium counterion, addition of a crown ether, and use of a solvent that effectively solvate cations.

Neighboring group participation involving acylamino or acyloxy groups is common in nucleophilic substitution. For example, in the reaction of methyl 4,6-*O*-benzylidene-2-deoxy-2-benzoylamino-3-*O*-mesyl-α-D-altro-pyranoside **8** with NaOEt, no 3-*O*-ethoxy-mannoside derivative was obtained. Instead, oxazoline **9** and epimine **10** were identified in this reaction (*Scheme 5*) [12]. This results from the 1,2-*trans*-diaxial relationship between the leaving



Scheme 5



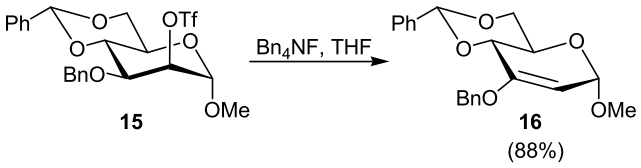
11 **12** **13** **14**

Scheme 6

group and the acylamino group, which is the ideal position for neighboring group participation. On the other hand, when the substituents in methyl 4,6-*O*-benzylidene-2-deoxy-2-benzoylamino-3-*O*-mesyl-α-D-glucopyranoside **11** are in a 1,2-*trans*-diequatorial orientation, the nucleophile competes with the acylamino group to form substituted products **12**, **14**, and oxazoline **13** (● *Scheme 6*) [13]. The participation of acyloxy groups is not very often encountered in pyranoid rings, even when the substituents are in suitable orientations. It seems that acyloxy group participation is more facile in furanoid rings.

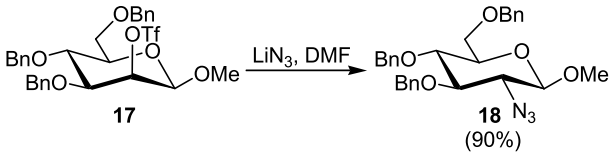
It has to be realized that an elimination reaction is the most frequent competing side reaction of nucleophilic substitution. Strong bases and weak nucleophiles favor this reaction. And elimination reactions are more likely when the leaving group is in a *trans*-diaxial relationship with a α-hydrogen atom at a sterically hindered position. Among halide ions, the fluoride ion is the strongest base and the weakest nucleophile and, as a result, substitution by this anion is often accompanied by elimination; sometimes, elimination products are the only ones formed (e. g. **15** to **16** in ● *Scheme 7*) [14]. To overcome elimination reactions, fluorides have been introduced by nucleophilic ring opening of epoxides, which is discussed in detail in ● *Sect. 1.2*. The azide ion also is sufficiently basic to promote elimination reactions; however, because the azide ion is an effective nucleophile, substitution is usually the dominant or exclusive reaction (e. g. **17** to **18** in ● *Scheme 8*) [15].

In addition, another competing side reaction is molecular rearrangement of carbohydrate. For example, attempted displacement of the 2-*O*-imidazyl ester of methyl 3,4,6-tri-*O*-methyl-β-D-



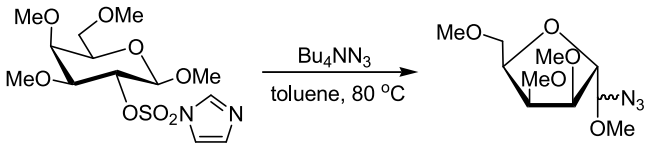
15 **16** (88%)

● **Scheme 7**



17 **18** (90%)

● **Scheme 8**



19 **20** (59%)

● **Scheme 9**

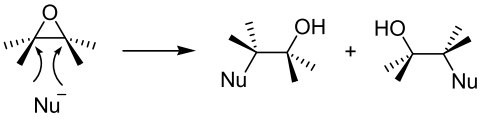
galactopyranoside **19** with an azide ion led to the formation of the corresponding 2,5-anhydro sugar **20** (◉ *Scheme 9*) [16]. The predominance of a ring-contraction reaction over a substitution can be explained by a steric interaction between the incoming nucleophile and the axial C-4 substituent, and by the favored antiparallel disposition of the C-1-O-5 bond and the equatorial leaving groups at C-2.

1.2 Ring-Opening Reaction

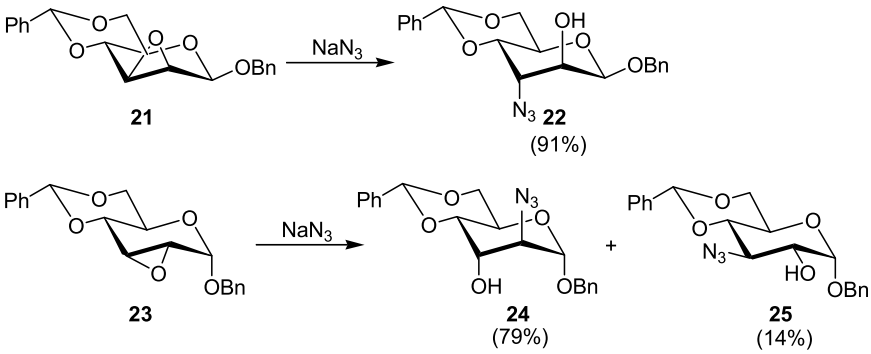
1.2.1 Nucleophilic Ring Opening of Epoxides

Nucleophilic ring opening of epoxides with heteroatom nucleophiles is another valuable method for the synthesis of many heteroatom-modified carbohydrate derivatives. The cyclic nature of epoxides renders the competing elimination process stereoelectronically unfavorable. Analogous to the above-discussed S_N2 nucleophilic mechanism, nucleophiles can open epoxide rings, and give rise to Walden inversion at the attacked carbon, furnishing α-hydroxy derivatives as illustrated in ◉ *Scheme 10*.

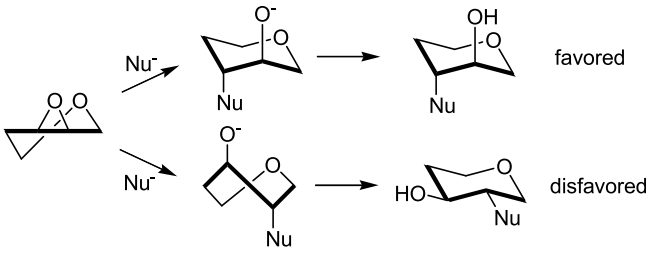
In theory, nucleophiles can attack either of the two carbons on epoxide rings. Therefore, for asymmetric epoxides, two regioisomeric products will be formed. However, in many cases, the epoxide-opening reactions in carbohydrate rings can be highly regio- and stereoselective, even specific. For example, the *trans*-fused benzylidene D-manno compound **21**, when treated with NaN_3 gives the D-altro adduct **22** as the only product, while the same reaction of its D-allo epoxide **23** gives the D-altro adduct **24** preferentially (◉ *Scheme 11*) [17]. The propensity of the reaction to give *trans*-diaxial products possibly results from the transition state, in which the diaxial product adopts a chair conformation, whereas the diequatorial product



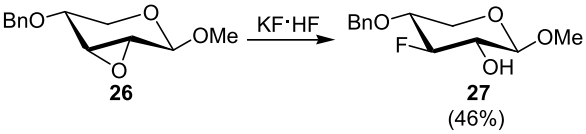
Scheme 10



Scheme 11



Scheme 12

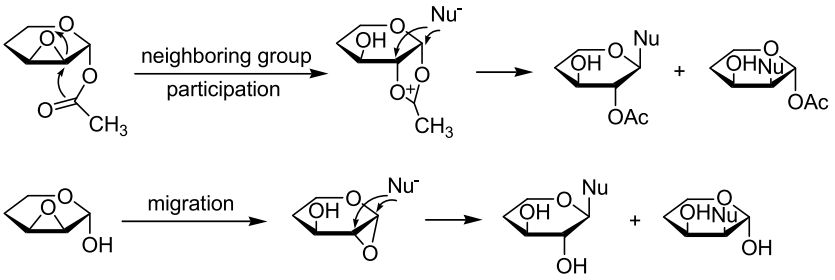


Scheme 13

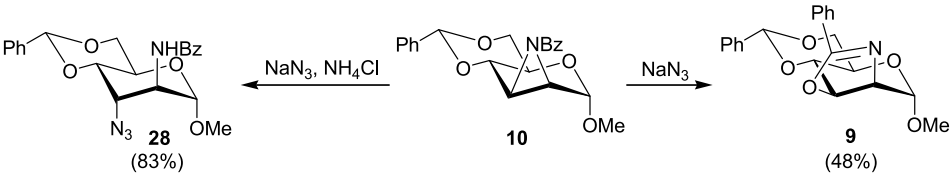
adopts a twist-boat (*Scheme 12*). Usually, the formation of a highly strained twist boat is energetically unfavorable and hence the diequatorial product is often not formed. In principle, the regioselectivity and stereoselectivity of carbohydrate epoxide rings can be improved by the introduction of additional rigidity in the starting material forcing the pyranose ring into one particular conformation. And for the less rigid conformations, it's difficult to predict the products.

Besides anionic nucleophiles (e. g. N_3^-), poor nucleophiles such as F^- can also be used in the ring-opening reactions in the presence of acids or Lewis acids as catalysts. For example, the reaction of methyl 2,3-anhydro-β-D-ribopyranoside **26** with F^- does not proceed except when acidic KF·HF is used to provide the nucleophile, yielding 3-deoxy-3-fluoroxylopyranoside **27** (*Scheme 13*) [18].

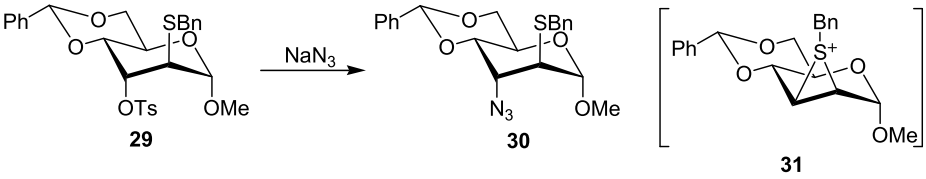
Sometimes, intramolecular rearrangements involving neighboring group participation or epoxide-ring migration if free hydroxyl groups are suitably situated in the vicinity of the epoxide ring, can complicate the reaction, giving rearranged products together with expected epoxide ring-opening products (*Scheme 14*) [19].



Scheme 14



Scheme 15



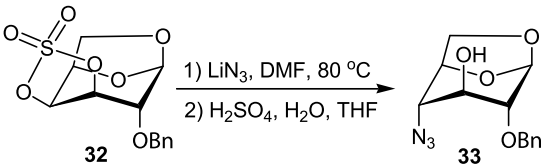
Scheme 16

Furthermore, the opening reactions of other three-membered rings such as aziridines and episulfides are similar to epoxides, except that they are more complicated. For example, the reaction of benzoylaziridine **10** with NaN_3 gives the *trans*-diaxial adduct **28** in the presence of NH_4Cl, whereas the oxazoline **9** is obtained as the major product in the absence of NH_4Cl (*Scheme 15*) [20]. In another instance, the 2-benzylthio-2-deoxy-3-*O*-tosyl-α-D-altropyranoside **29** reacts with NaN_3 to afford the configuration-retained substitution derivative **30** (*Scheme 16*), indicating the reaction occurs via the sulfonium salt intermediate **31** formed by intramolecular displacement from compound **29** [21].

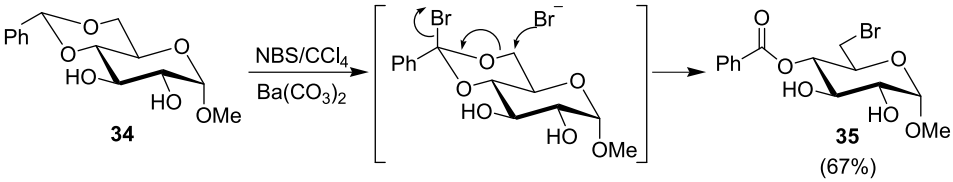
1.2.2 Ring-Opening Reactions of Five/Six-Membered Rings

Five-membered ring compounds, such as cyclic sulfites, sulfates, sulfamidates, and thionocarbonates, can also be opened by treatment with nucleophiles in an analogous fashion to epoxide ring-opening. Nucleophilic ring-opening of cyclic sulfate **32** with NaN_3, for example, gives the azido-substituted product **33** (*Scheme 17*) [22]. Noteworthy, these cyclic compounds are usually prepared from corresponding *cis* diols, whereas epoxides are synthesized from *trans* diols.

In some cases, six-membered rings can be involved in ring-opening reactions. For example, the reaction of 4,6-*O*-benzylidene glucopyranoside **34** with *N*-bromosuccinimide (NBS) in the presence of barium carbonate leads to the corresponding 4-*O*-benzoyl-6-bromo-6-deoxy



Scheme 17

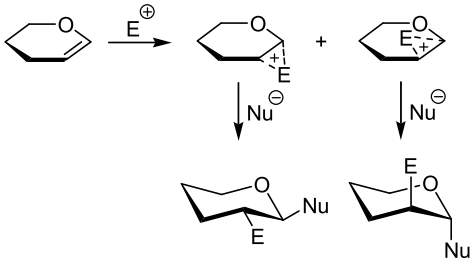


Scheme 18

glucopyranoside **35** (see *Scheme 18*) [23]. The reaction may probably undergo the radical bromination of the benzylic carbon atom followed by rearrangement to the 6-deoxy-6-bromo derivative.

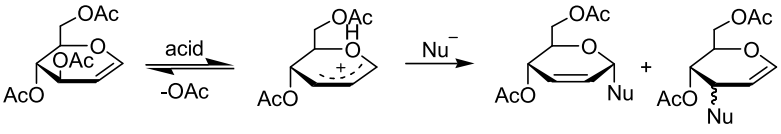
1.3 Addition Reaction

Unlike the carbohydrates with double bonds at positions other than between C-1 and C-2 ('isolated alkenes'), which exhibit normal alkene chemistry, glycals are vinyl ethers and therefore undergo a number of highly selective addition reactions due to the strongly polarized double bonds and the presence of bulky substituents at the C-3 allylic centers. Straightforward addition reaction includes initial electrophilic addition at the double bond, followed by the addition of a nucleophile at C-1 to give the 1,2-*trans* adduct (*Scheme 19*).

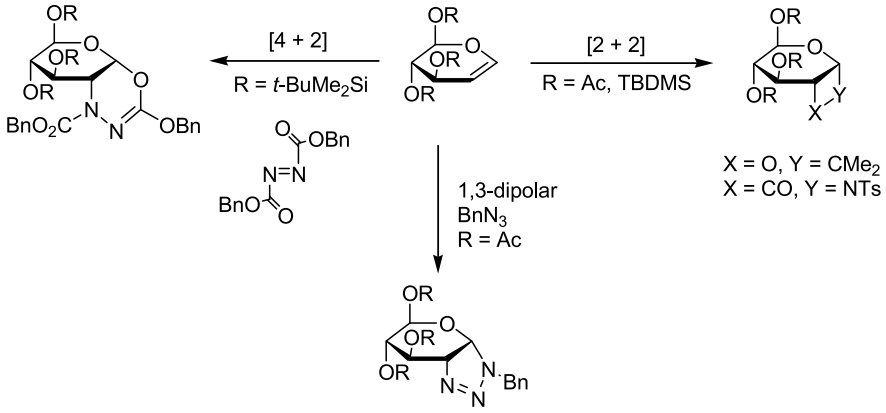


Scheme 19

The classical Ferrier rearrangements of acetyl glycals lead to the introduction of a nucleophilic group at C-1 (preferentially α-anomers) and the migration of the double bond to C-2 in the presence of acid catalysts (*Scheme 20*) [24]. In some particular conditions (high temperature or stronger acid), 3-substituted isomeric products can be formed (*Scheme 20*) [25].



Scheme 20



Scheme 21

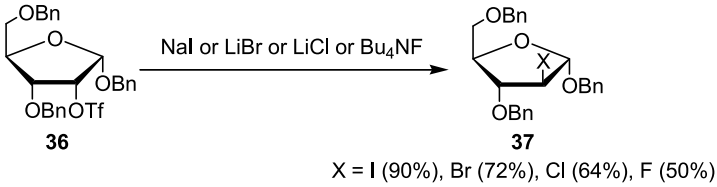
Glycals can also undergo cycloaddition reactions and their derivatives are of interest for synthetic purposes. Diels–Alder [4+2], Paterno–Büchi [2+2], and 1,3-dipolar additions can be applied to the construction of fused cycloadducts (*Scheme 21*) [26,27,28].
Radical additions can also be performed on the double bonds of unsaturated carbohydrates and regio- and stereoselectivities are high in some particular cases [29].

2 Introduction of Halogens

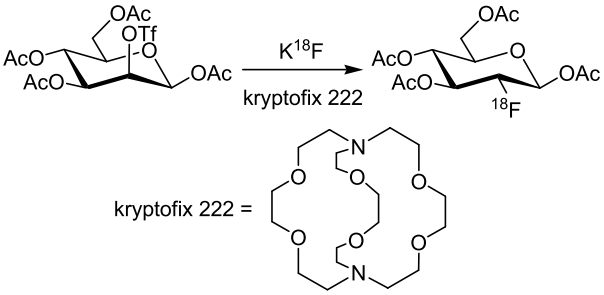
Deoxyhalogeno sugars are carbohydrate derivatives in which hydroxyl groups at positions other than the anomeric center have been replaced by halogen atoms. These compounds are susceptible to nucleophilic attack, leading to displacement, elimination, or anhydro-ring formation, therefore are useful intermediates in the synthesis of other uncommon sugars, such as deoxy and aminodeoxy sugars. Deoxyfluoro sugars are extensively applied to the study of carbohydrate metabolism and transport [30]. ^{18}F-labeled carbohydrates are employed for medical imaging [31]. Many chlorinated sugars have antibiotic activities and the trichloro sugar 'sucralose', 4,1′,6′-trichloro-4,1′,6′-trideoxy-*galacto*-sucrose, is known for its sweet profile 650 times stronger than sucrose [32,33]. The bromo and iodo derivatives have also been widely used, and have undergone a variety of nucleophilic substitution reactions [34].

2.1 Displacement of Sulfonic Esters

Sulfonate displacement has been well established as a common method for introducing halogen atoms into carbohydrates. The high reactivity and easy preparation of sulfonic esters contribute to their use in carbohydrate chemistry. Triflate displacements are popular for the preparation of secondary halogenated carbohydrates and often give satisfactory results when mesylates and tosylates fail to give products.
The efficiency of the displacement is to a large degree dependent upon the nucleophilicity of the halide. As illustrated in *Scheme 22*, for example, all four 2-deoxy-2-halo-α-D-arabino-



Scheme 22



Scheme 23

furanosides **37** (X = F, Cl, Br, I) have been obtained from the corresponding 2-*O*-triflyl-α-D-ribofuranoside **36**, and the yields decrease in the order: I > Br > Cl > F [35].

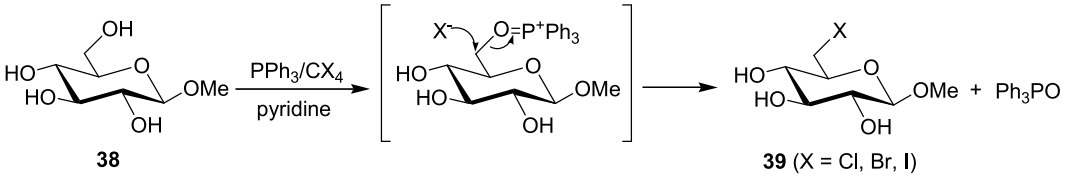
Alkalimetal halides or the more highly soluble tetraalkylammonium salts are frequently used as the halide sources. The weakly nucleophilic character of the fluoride ion strongly affects its displacement efficiency. In some cases, a cryptand is used to increase its effective nucleophilicity (*Scheme 23*) [36].

A variety of fluoride ion sources have been used in an effort to improve product yields in deoxyfluoro sugar synthesis. These salts include cesium fluoride, tetrabutylammonium fluoride, and tetrabutylammonium difluoride (Bu_4NHF_2) [37,38]. And currently, the most successful fluoride source is tri(dimethylamino)-sulfonium difluorotrimethylsilicate (TASF), which is soluble in many organic solvents and produces an anhydrous fluoride ion [39].

2.2 Direct Displacement of Hydroxyl Groups

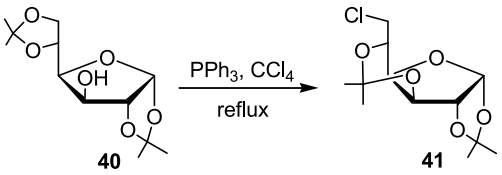
2.2.1 Application of Alkoxyphosphonium Salts

Several reagents of this type are now available for replacing hydroxyl groups in carbohydrates by halogen atoms. The reaction of triphenylphosphine and tetrahalomethanes produces a halogenophosphonium ion in situ; this ion, when used in pyridine with unprotected sugars, can react with the hydroxymethyl groups to selectively give primary halo sugars. As shown in *Scheme 24*, treatment of methyl β-D-glucopyranoside **38** with triphenylphosphine and carbon tetrachloride, tetrabromide, or tetraiodide produces the corresponding 6-deoxy-6-haloglucosides **39** in almost quantitative yields [40,41].

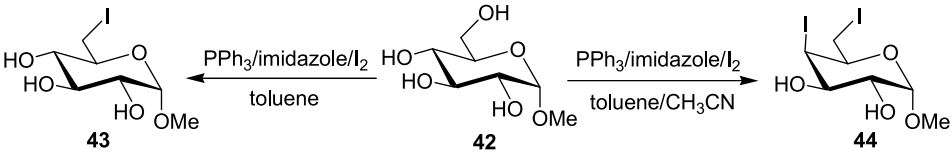


38 **39** (X = Cl, Br, I)

Scheme 24



Scheme 25



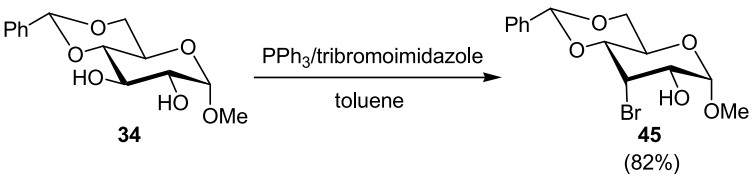
Scheme 26

However, some limitations in the application of this halogenation procedure have been noted. For example, 1,2;5,6-di-*O*-isopropylidene-α-D-glucofuranose **40**, in which C-3 is sterically hindered, gives the 5,6-acetal rearranged product **41** (*Scheme 25*) [42].

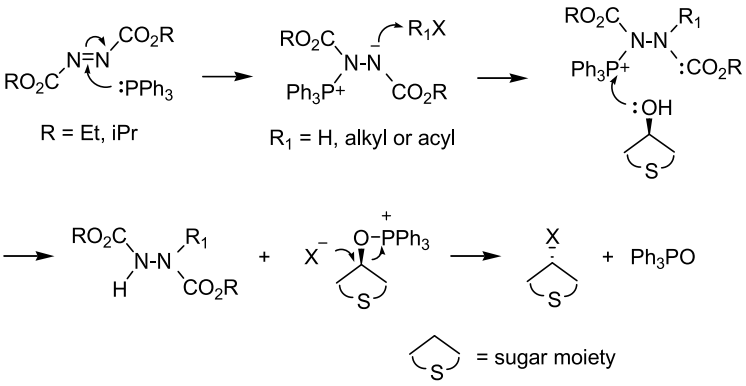
Other combinations of reagents, such as triphenylphosphine/*N*-halosuccinimides [43], triphenylphosphine/imidazole/iodine, and triphenylphosphine/2,4,5-trihaloimidazole, react by a similar mechanistic pathway, providing a degree of regioselectivity. For example, methyl α-D-glucopyranoside **42**, when heated with PPh_3/imidazole/I_2 mixture in toluene, gives methyl 6-deoxy-6-iodo-α-D-glucoside **43**, whereas in toluene/acetonitrile, a solvent of increased polarity in which reactants are more soluble, displacements occur at both the 4- and 6-position to give methyl 4,6-dideoxy-4,6-diiodo-α-D-galactoside **44** (*Scheme 26*) [44,45,46].

In another instance, when 2,4,5-tribromoimidazole is used, methyl 4,6-*O*-benzylidene-3-bromo-3-deoxy-α-D-allopyranoside **45** is selectively formed from the methyl glucoside **34** (*Scheme 27*) [47]. This results from the greater nucleophilicity of C-3 hydroxyl, which can form alkoxyphosphonium ion easily.

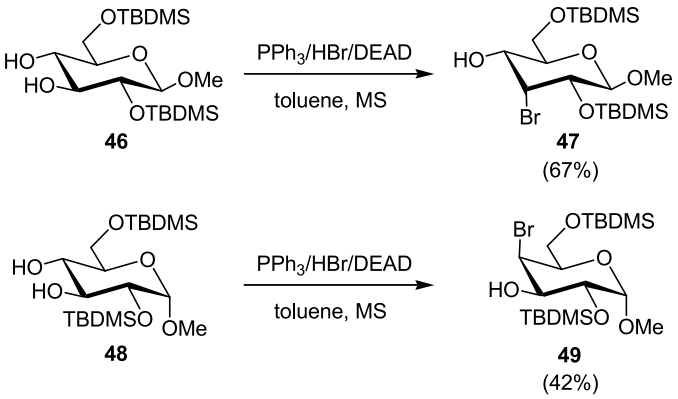
The Mitsunobu procedure, which uses diethyl azodicarboxylate (DEAD) or diisopropyl azodicarboxylate (DIAD) to react with PPh_3 and alcohols providing configuration-inversed substitution products via alkoxytriphenylphosphonium ion intermediates (*Scheme 28*), presents another route for introducing halogen atoms into carbohydrates [48].



Scheme 27



Scheme 28



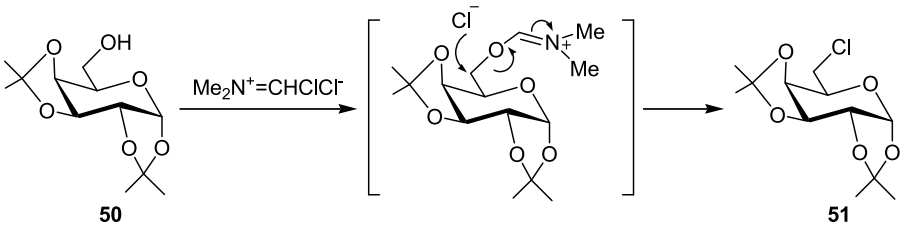
Scheme 29

For example, methyl 2,6-di-*O*-*t*-butyldimethylsilyl-β-D-glucopyranoside **46** can be converted mainly to 3-deoxy-3-bromoalloside **47** when treated with PPh_3/HBr/DEAD, whereas the corresponding α-glucoside **48** reacts to give C-4 brominated product **49** preferentially under the same conditions (*Scheme 29*) [49]

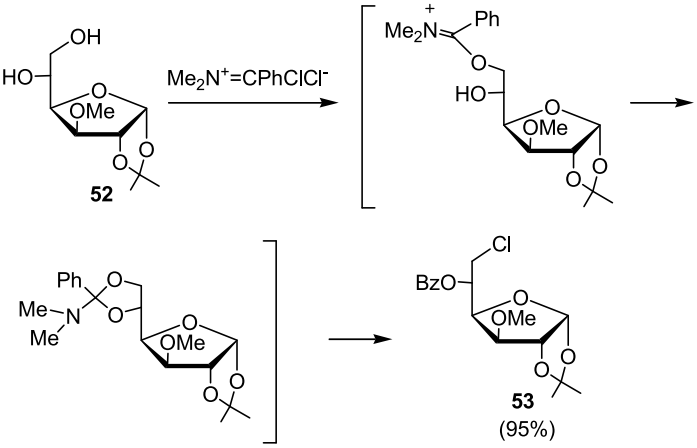
Triphenylphosphite methiodide $[(PhO)_3P^+MeI^-]$ and dihalides $[(PhO)_3P^+XX^-]$ (Rydon reagents), which are closely related to the reagents described above, have also been successfully applied to the synthesis of halogeno sugars [50].

2.2.2 Application of Iminium Salts

N,*N*-dimethylformamide reacts with chlorides of inorganic acids (phosgene, phosphoryl chloride, phosphorus trichloride, and thionyl chloride) to form an active salt: (chloromethylene)dimethyliminium chloride, which is also called Vilsmeier's imidoyl chloride reagent. This reagent, as shown in *Scheme 30*, reacts with the free hydroxyl group of the sugar derivative **50** to give an activated intermediate that can undergo nucleophilic substitution by



Scheme 30



Scheme 31

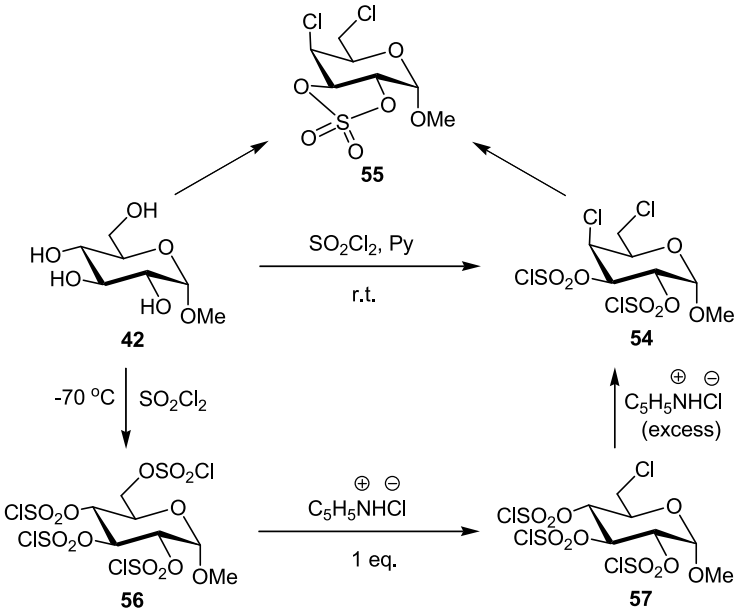
chloride ion with loss of dimethylformamide yielding the chlorinated derivative **51** in an excellent yield [51].

α-Chlorobenzylidene-*N*,*N*-dimethyliminium chloride reacts with the diol of sugar derivative **52** to form an addition product, which is then transformed into the 5-*O*-benzoyl-6-chloro-6-deoxy-glucose derivative **53** (*Scheme 31*) [52].

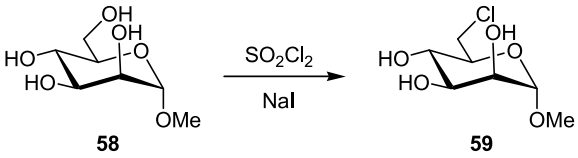
2.2.3 Application of Chlorosulfates

The reaction of carbohydrates with sulfuryl chloride can form chlorosulfate groups initially, followed by S_N2 displacement of the liberated chloride ion. This provides another effective method for the preparation of chlorodeoxy sugars [34,53]. Those positions, where the steric and polar factors are favorable for a S_N2 reaction, are more susceptible to displacement. The chlorosulfates that have been substituted with a chloride can easily be cleaved by treatment with sodium iodide.

Thus, the representative reaction manner of this transformation is presented by the reactions of methyl α-D-glucopyranoside **42**, as outlined in *Scheme 32*. When a minimum proportion of pyridine is employed, the reaction of sulfuryl chloride with **42** produces 4,6-dichlorodeoxy sugar **54**. While in the presence of an excess of pyridine, 2,3-cyclic sulfate **55** can be formed from either compound **42** or **54**. And methyl α-D-glucopyranoside 2,3,4,6-tetra(chlorosulfate)



Scheme 32



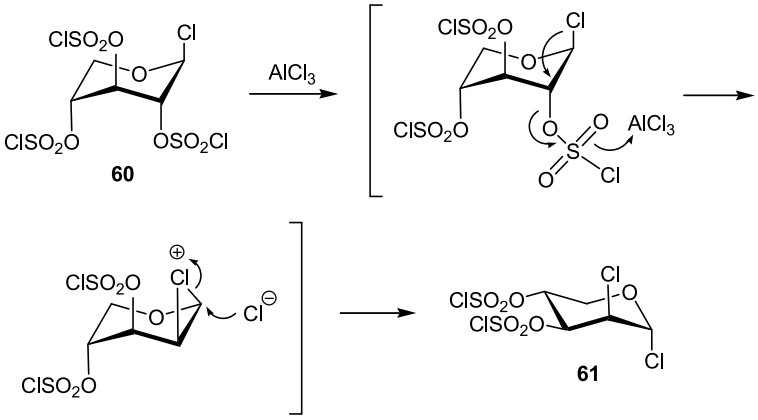
Scheme 33

56 is isolated when the reaction proceeds at −70 °C. Compound **56** is further converted into the sugar derivative **54** via the formation of 6-chlorodeoxy sugar **57** on treatment with pyridine chloride [54,55,56,57].

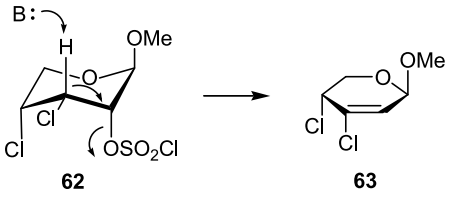
In the transformation of **42** to **54**, the lack of substitution of the chlorosulfonyloxy group at C-3 is attributed to the presence of the *β-trans* axial methoxyl group at C-1; also, the chlorosulfonyloxy group at C-2 is deactivated to displacement by chloride ion. This '*β-trans*-axial substituent effect' can also interpret the regioselectivity of the reaction of the α-D-mannopyranoside **58** to **59** (*Scheme 33*) [57].

The chlorosulfated glycosyl chloride **60** is converted into 2-chloro-2-deoxy-α-D-lyxopyranosyl chloride **61** when treated with aluminum chloride (*Scheme 34*) [58]. This reaction proceeds via an initial intramolecular displacement of chlorosulfate at C-2 by the anomeric chlorine atom, followed by the chloride ion attack at the more highly reactive center (C-1).

Pyranoid derivatives having a chlorosulfonyloxy group in a *trans*-diaxial relation with a ring proton, such as **62**, may undergo an elimination reaction to yield the unsaturated compound **63** (*Scheme 35*) [59].



Scheme 34

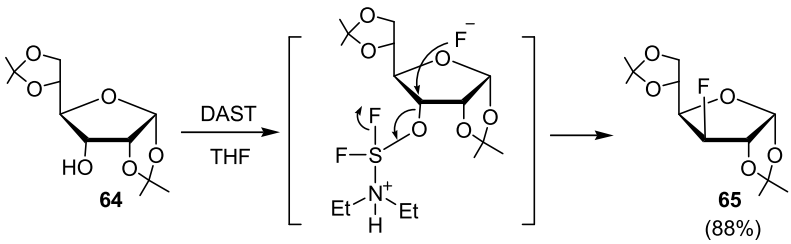


Scheme 35

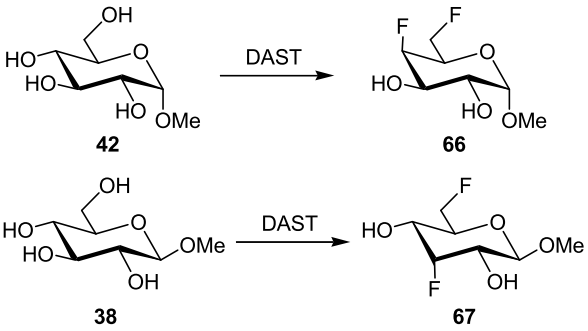
2.2.4 Application of Diethylaminosulfur Trifluoride (DAST)

To date, diethylaminosulfur trifluoride (Et_2NSF_3) may be the most commonly applied reagent for direct fluorination. An alcohol displaces a fluoride of DAST resulting in an activated intermediate, which in turn is substituted by the liberated fluoride, as illustrated in the formation of **65** from **64** (*Scheme 36*) [60].

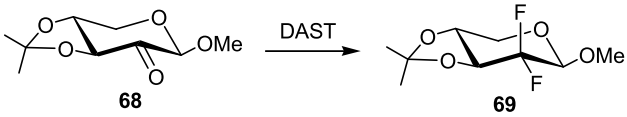
Some selectivity has been observed with this reagent, and is similar to chlorosulfate displacement. Thus, due to the '*β-trans*-axial substituent effect', methyl α-D-glucopyranoside **42** gives 4,6-difluorodeoxy sugar **66**, and methyl β-D-glucopyranoside **38** gives 3,6-difluorodeoxy sugar **67** (*Scheme 37*) [61,62].



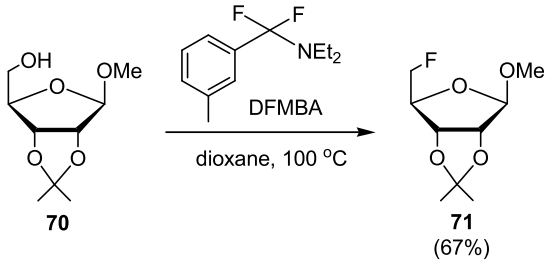
Scheme 36



Scheme 37



Scheme 38



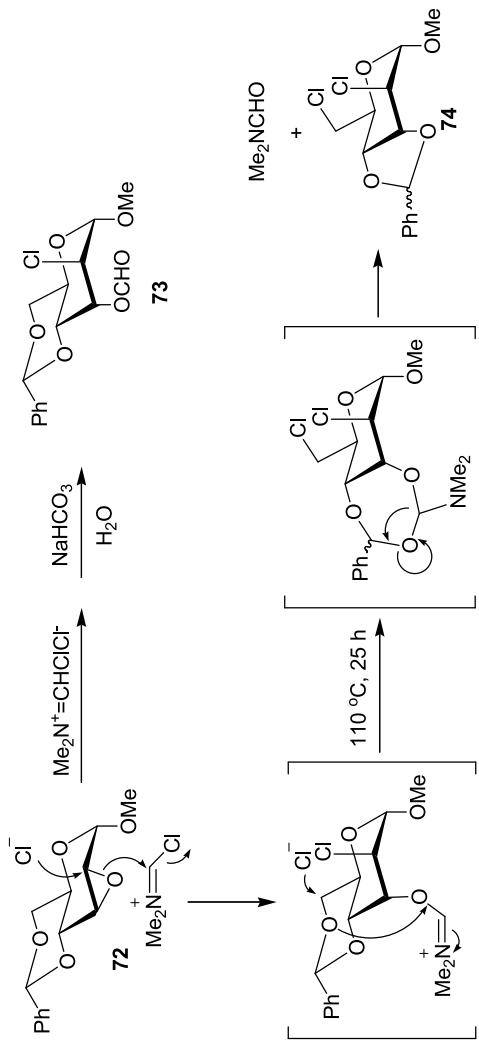
Scheme 39

A gem-difluoride is formed when a ketone or aldehyde is treated with DAST, as outlined in the reaction of **68** to **69** (*Scheme 38*) [60].
For the synthesis of **71** from **70**, *N*,*N*-diethyl-α,α-difluoro-(*m*-methylbenzyl)amine (DFMBA) is used to avoid the unexpected migration of the methoxy group from the 1- to 5-position when DAST is used (*Scheme 39*) [63].

2.3 Ring-Opening Reactions

2.3.1 Opening of Epoxide Rings

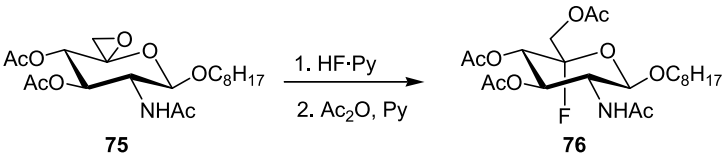
As an alternative practical method, the opening of epoxides by halide ions has often been used for introducing halogens, especially fluorine, into sugar molecules. Sodium iodide, magnesium bromide, and hydrogen halides (HCl, HF) are frequently used as halogen sources [64,65].



Scheme 40

A special example of epoxide opening reaction is outlined in ❯ *Scheme 40*. Methyl 2,3-anhydro-4,6-*O*-benzylidene-α-D-allopyranoside **72** gives the 2-chloro-2-deoxy-3-*O*-formyl-altroside **73** on treatment of (chloromethylene)dimethyliminium chloride at room temperature. Alternatively, when the reaction mixture is heated for a longer time, methyl 3,4-*O*-benzylidene-2,6-dichloro-2,6-dideoxy-α-D-altroside **74** is obtained by means of a rearrangement of the benzylidene group [51].

An interesting regioselectivity of epoxide opening occurs in the reaction of sugar **75** with HF · pyridine (❯ *Scheme 41*) [66]. Excellent selectivity of fluoride ion substitution at C-5 is obtained giving compound **76** after acetylation.

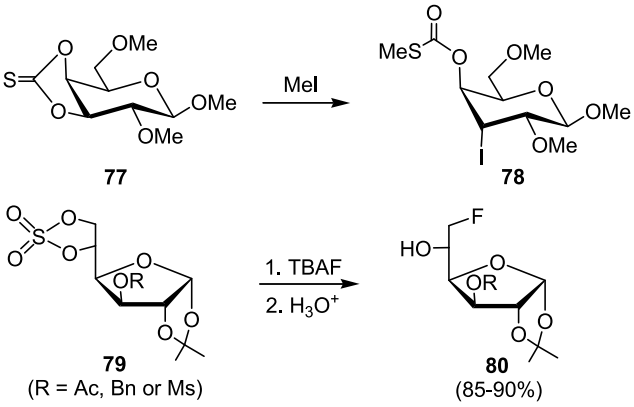


75 **76**

Scheme 41

2.3.2 Opening of Sulfur-Containing Rings

Five-membered ring compounds such as cyclic thionocarbonates can also be opened on treatment of methyl iodide. For example, thionocarbonate **77** gives regioselectively the ring-opening product **78** (❯ *Scheme 42*) [67]. Another regioselective ring opening has been observed in the transformation of fluorosugar **80** from the 5,6-cyclic sulfates **79** (❯ *Scheme 42*) [68,69].



77 **78**

79 (R = Ac, Bn or Ms)

80 (85-90%)

Scheme 42

2.3.3 Opening of Benzylidene Acetals

N-Bromosuccinimide converts benzylidene acetals into bromodeoxy benzoates, and offers an excellent method for preparing 6-bromodeoxyhexoses (see ❯ *Sect. 1.2.2*). The reaction can

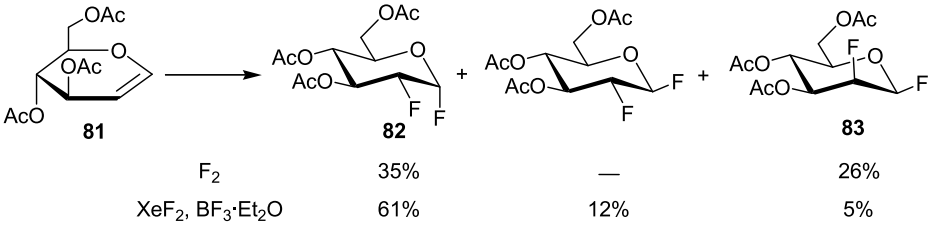
also be applied to 2,3- and 3,4-benzylidene acetals of pyranosides and furanosides, and proceeds regioselectively [23]. When there are no stereochemical constrains to effect regioselectivity, mixtures of isomers are formed. And similar transformations can be performed using bromotrichloromethane as the bromine source [70].

2.4 Addition to Unsaturated Sugar Derivatives

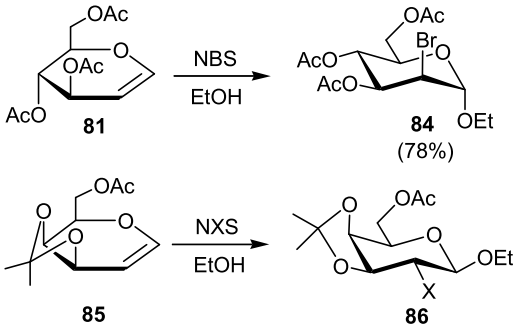
Addition reactions of electrophilic halogenation reagents to unsaturated sugars constitute a valuable synthetic alternative to the S_N-strategy. 3,4,6-tri-*O*-acetyl-D-glucal **81** reacts with fluorine diluted with argon to give mixtures of 3,4,6-tri-*O*-acetyl-2-deoxy-2-fluoro-α-D-glucopyranosyl fluoride **82** and the corresponding β-D-*manno*-difluoride **83** [71]. An efficient variant of this reaction utilizes xenon difluoride additions catalyzed by boron trifluoride etherate, giving compound **82** in an increased selectivity (● *Scheme 43*) [72].

High regio- and stereoselectivity are obtained in the reaction given in ● *Scheme 44*. The reaction of glucal **81** with NBS mainly leads to 2-bromo-2-deoxy-α-D-mannopyranoside **84**, whereas 3,4-*O*-isopropylidene-D-galactal **85** affords only the β-D-galactopyranoside **86**. Thus, this reaction has made glycals very significant starting materials for the preparation of some important glycosides, especially 2-deoxy oligosaccharides after reduction of the halide groups [73,74].

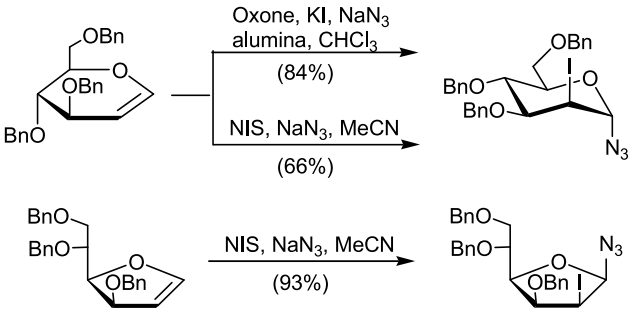
Addition of iodine azide to glycals gives mostly the 1,2-*trans*-2-deoxy-2-iodoglycosyl azides [75]. In view of the hazardous nature of iodine azide, two mild methods were devel-



● **Scheme 43**



● **Scheme 44**

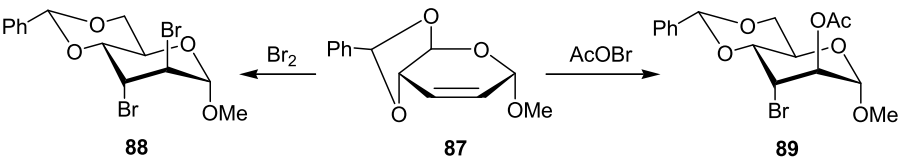


Scheme 45

oped for the preparation of 1,2-*trans*-2-deoxy-2-iodoglycosyl azides from glycals in good yields. The first method involves reaction of a glycal with oxone, potassium iodide, sodium azide, and neutral alumina in chloroform at room temperature (*Scheme 45*). Whereas the second method involves its reaction with *N*-iodosuccinimide and sodium azide in acetonitrile at 0 °C; however, it is interesting to note that while pyran glycals give 1,2-*trans*-2-deoxy-2-iodoglycosyl azides, the furan glycals give exclusively the 1,2-*cis*-2-deoxy-2-iodoglycosyl azides (*Scheme 45*) [76].

Under the conditions of the addition reaction glycals such as 2-fluoro-glycals can be converted into corresponding 2-deoxy-2-fluoro-2-iodo derivatives [77].

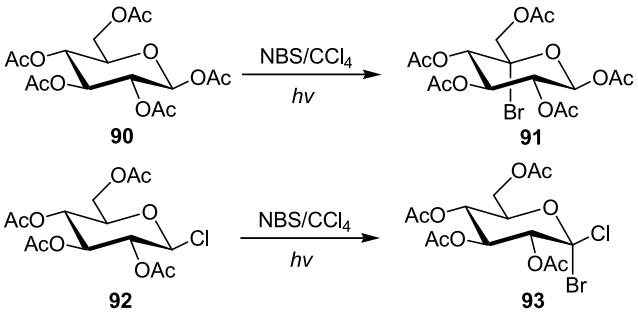
The isolated alkenes undergo normal additions, frequently with high regio- and stereoselectivity. For example, methyl 4,6-*O*-benzylidene-2,3-dideoxy-α-D-*erythro*-hex-2-enopyranoside **87** gives the 2,3-dibromo-D-*altro*-adduct **88** on treatment with bromine, and the 2-*O*-acetyl-3-bromo-3-deoxy adduct **89** is obtained when treated with acetyl hypobromite (*Scheme 46*), indicating that the reactive bromonium ion intermediate is formed on the lower (α) face of the ring [78].



Scheme 46

2.5 Radical Bromination Reactions

Bromine can be introduced directly into particular ring positions of certain sugar derivatives by free radical processes which can show high selectivity and efficiency. For example, peracetylated glucose **90** is selectively brominated at the C-5 position by NBS yielding **91** however, in the case of glycosyl halides, such as tetra-*O*-acetyl-β-D-glucopyranosyl chloride **92**, anomeric substituted product **93** is mainly formed (*Scheme 47*) [79,80].



90 91 92 93

Scheme 47

3 Introduction of Nitrogen

Amino sugars are carbohydrate derivatives in which a hydroxyl group is replaced by an amino group at nonanomeric positions. 2-Amino-2-deoxy-D-glucose (D-glucosamine) is abundant in nature, appearing in particular in the polysaccharide chitin. 2-Amino-2-deoxy-D-galactose (D-galactosamine) is a constituent monosaccharide unit of dermatan and chondroitin sulfate found in mammalian tissue and cartilage. Nonulosaminic acids are 5-amino-5-deoxynonose derivatives usually found in combined form in mucolipids or mucopolysaccharides.

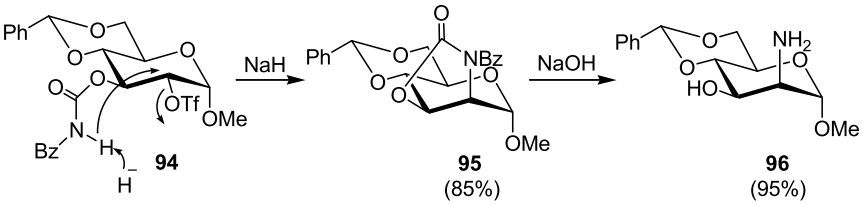
3.1 Nucleophilic Displacement Reactions

This substitution reaction can be performed using nucleophilic nitrogen-containing reagents such as ammonia, hydrazine, or sodium azide. Ammonia and hydrazine can overcome the dipolar repulsion against charged nucleophiles, but their products are still nucleophilic and can perform a second displacement. Azide ions are stable under many reaction conditions and can be converted into amines by a wide range of reducing agents. Sometimes, the addition of crown ether is required to increase the nucleophilicity of azide ions [7]. Phthalimide ions have also successfully been applied in substitution reactions to yield a protected amino sugar derivative [81,82].

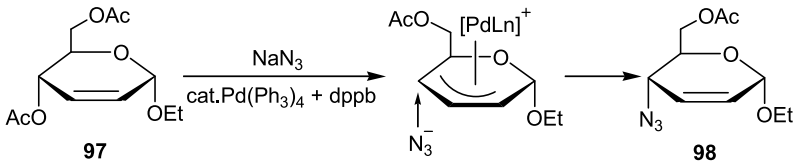
Misunobu substitution conditions (PPh_3, DIAD) have also been successfully applied to the direct displacement of hydroxyl using zinc azide-pyridine complex [83].

Intramolecular substitutions provide a convenient and stereoselective approach for the preparation of vicinal *cis*-hydroxy amino derivatives. Thus, the sodium hydride-mediated cyclization of the D-*gluco* compound **94** leads to the formation of the oxazolidinone **95**, which is hydrolyzed by base affording the D-mannosamine **96** (*Scheme 48*) [84].

Palladium-catalyzed azidation reaction is a novel regio- and stereoselective method for introducing azide groups into 2,3-dideoxyhex-2-enopyranosides. For example, ethyl 4,6-di-*O*-acetyl-α-D-*erythro*-hex-2-enopyranoside **97**, when treated with NaN_3 in the presence of $Pd(PPh_3)_4$ and dppb [1,4-bis(diphenylphosphino)butane], gives predominantly the 4-substituted product **98** with retention of configuration (*Scheme 49*). The (π-allyl)palladium intermediate and the crowded C-2 by the aglycon moiety may contribute to the *trans* attack of N_3^- at C-4 only [85].



Scheme 48



Scheme 49

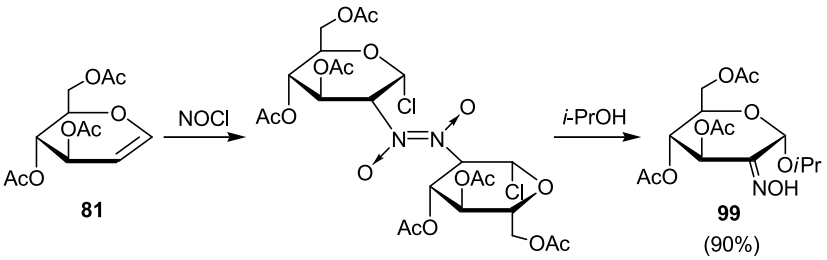
3.2 Opening of Epoxide Rings

Nucleophilic ring-opening of epoxides with nitrogen nucleophiles offers another route to amino sugars. Similar to the above-mentioned nucleophilic substitutions, epoxide openings can be conducted with ammonia, primary amines, guanidine, or azide ions [17,86]. Ammonium chloride is often used to neutralize the alkoxide produced during the opening of an epoxide.

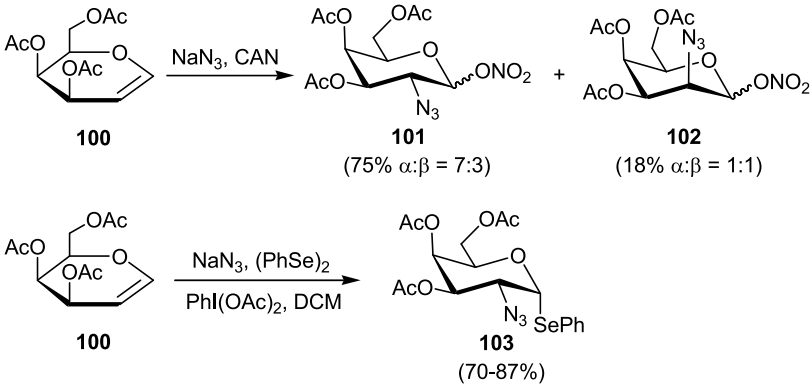
3.3 Addition of Nitrogenous Reagents to Double Bonds

3.3.1 Addition to Glycals

Chloronitroxylation of glycal double bonds has been used as a means for synthesizing 2-amino-2-deoxy sugar derivatives. As illustrated in ❯ *Scheme 50*, the addition of nitrosyl chloride to tri-*O*-acetyl-D-glucal **81** gives a dimeric adduct which can be converted into the oxime **99** in high stereoselectivity with alcohols [87]. Reduction of the product provides corresponding amino sugar derivatives (see ❯ *Sect. 3.5*).



Scheme 50

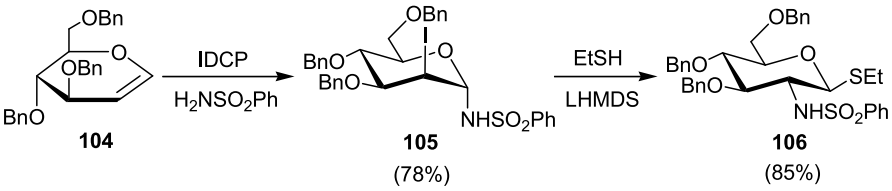


Scheme 51

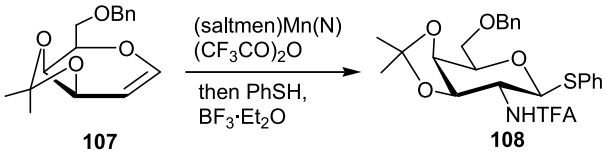
Addition to the double bond can also be conducted by the so-called azidonitration reaction. This reaction occurs with sodium azide and ceric ammonium nitrate (CAN) resulting in a 2-azido-2-deoxyglycosyl nitrate via a radical azido addition, as illustrated in *Scheme 51* by the conversion of tri-*O*-acetylgalactal **100** into the D-*galacto* adduct **101** and a small amount of the *talo*-isomer **102**. The anomeric nitrate can be readily replaced by a halide, acetyl, or hydroxyl functionality [88]. In another similar reaction (*Scheme 51*), sodium azide and diphenyl diselenide in the presence of (diacetoxy)iodobenzene react with the galactal **100** to give stereoselectively the α-phenylselenyl galactoside **103** (azidophenylselenylation) [89].

Halosulfonamidation of hexose-derived glycals followed by sulfonamide migration reaction provides also a useful approach for the synthesis of 2-amino sugar derivatives. For example, the reaction of tri-*O*-benzyl glucal **104** with iodonium di-*sym*-collidine perchlorate (IDCP) and benzenesulfonamide gives the *trans*-diaxial iodosulfonamide **105**, which undergoes sulfonamide migration in the presence of lithium ethanethiolate yielding the 2-amino thioglycoside **106** (*Scheme 52*) [90].

A transition metal-mediated approach to amidation of glycal substrates can lead to the formation of 2-deoxy-2-trifluoroacetyl-amido derivatives. Thus, when a solution of (saltmen)Mn(N) is added to a mixture of trifluoroacetic anhydride (TFAA) and glycal **107** followed by sequential treatment with thiophenol and $BF_3 \cdot Et_2O$, the 2-*N*-trifluoroacetamido thioglycoside **108** can be obtained in good yield and excellent diastereoselectivity (*Scheme 53*) [91].



Scheme 52



107 108

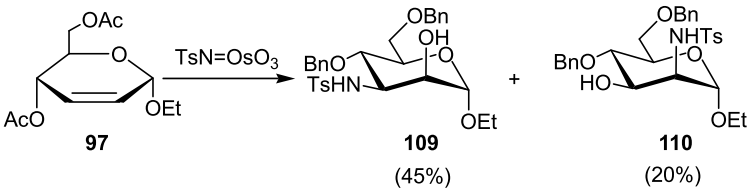
Scheme 53

3.3.2 Addition to Isolated Alkenes

The *p*-toluenesulfonylimidoosmium reagent, which may be generated from Chloramine-T and osmium tetraoxide, leads to *cis*-addition of OH and HNTs to the least-hindered face of double bonds. With the hex-2-enopyranoside **97**, the addition occurs to the top face of the double bond to give the regioisomers **109** and **110** (*Scheme 54*) [92].

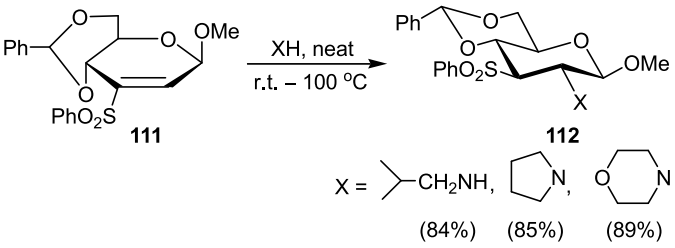
Vinyl sulfone-modified carbohydrates, such as **111**, can be subjected to Michael addition reaction with primary or secondary amines to construct **112** bearing amino groups at the C-2 carbon in equatorial configurations (*Scheme 55*) [93].

An 1-hydroxy-1,2-benziodoxol-3(1*H*)-one 1-oxide (IBX)-mediated process provides another efficient and stereoselective preparative route for the synthesis of amino sugars. For example, reaction of the D-glucal derivative **113** with *p*-methoxybenzene isocyanate in the presence of a catalytic amount of 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU) followed by treatment with IBX furnishes the cyclic carbonate **115**, through urethane **114**. Subsequent removal of the *p*-methoxybenzyl (PMB) protecting group by CAN affords the protected amino sugar **116** (*Scheme 56*) [94].



97 109 (45%) 110 (20%)

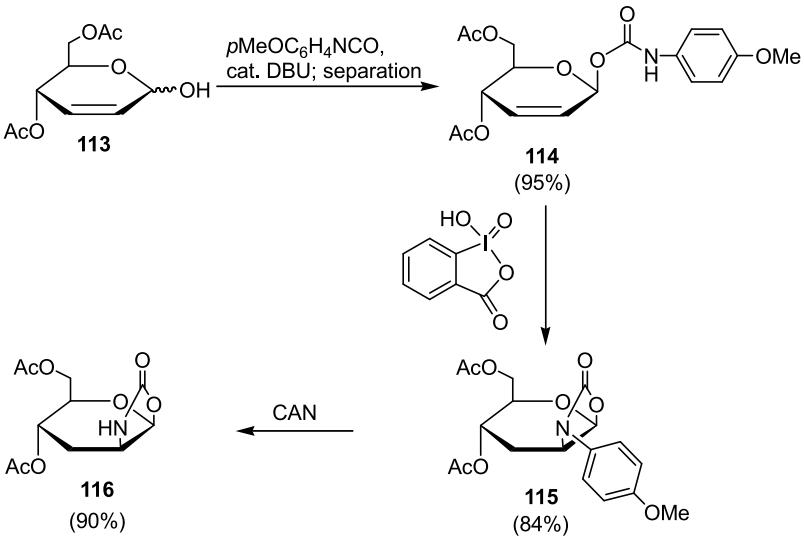
Scheme 54



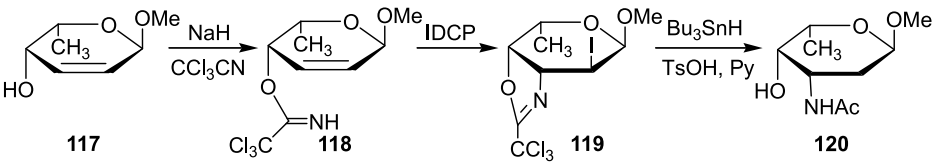
111 112

(84%) (85%) (89%)

Scheme 55



Scheme 56



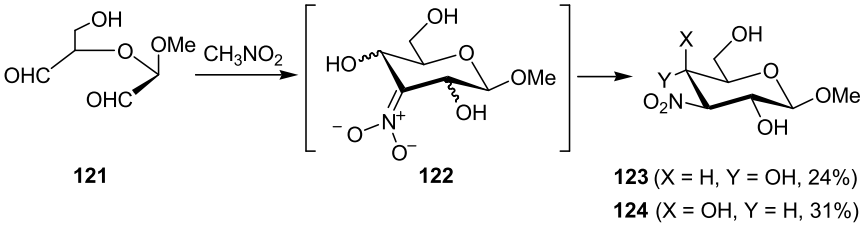
Scheme 57

Regio- and stereo-controlled iodocyclizations of allylic trichloroacetimidates provide a route to *cis*-hydroxyamino sugar from hexenopyranosides. For example, conversion of the hydroxyl of compound **117** into the trichloroacetimidate **118** followed by IDCP-mediated intramolecular cyclization, gives iodo-oxazoline derivative **119**, which is reduced (Bu_3SnH) and hydrolyzed (pyridine, TsOH) to afford *N*-acetyldaunosamine methyl glucoside **120** (*Scheme 57*) [95].

3.4 Cyclization of Dialdehydes

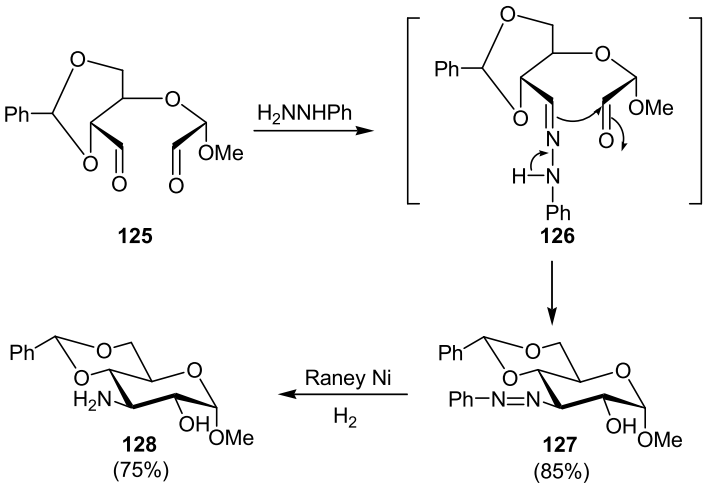
Dialdehydes, obtained by periodate oxidation of appropriate cyclic polyhydroxy precursors, condensing with some nitrogen-containing reagents in basic solution has been developed into a useful synthesis of amino sugars. Thus, the dialdehyde **121**, on treatment with nitromethane in the presence of sodium methoxide produces a mixture of *aci*-nitro salt isomers **122**, which mainly give 3-amino-3-deoxyhexopyranosides **123** and **124** after reduction (*Scheme 58*) [96].

In another instance, when the dialdehyde **125** is treated with phenylhydrazine, a cyclization occurs, probably via the monophenylhydrazone **126** to give the phenylazo derivative **127**,



121 **122** **123** (X = H, Y = OH, 24%)
124 (X = OH, Y = H, 31%)

Scheme 58



125 **126** **127** (85%) **128** (75%)

Scheme 59

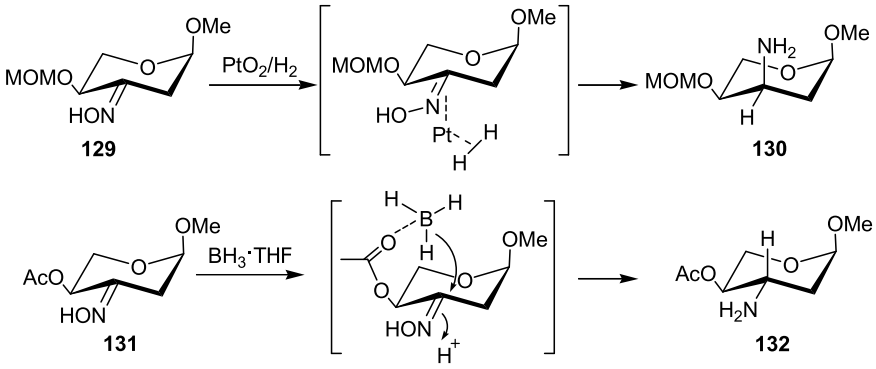
which can be hydrogenolyzed to the 3-amino-3-deoxy-α-D-glucopyranoside **128** with a good stereoselectivity (*Scheme 59*) [97].

3.5 Reduction of Ulose Oximes

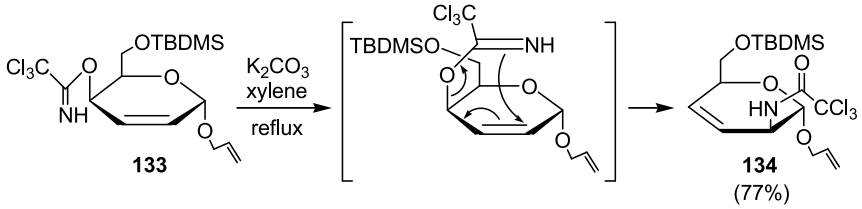
Oxime formation from uloses and subsequent reduction provides yet another route to amino sugars. The stereochemical outcome is dependent on the reducing agent, solvent, and protecting groups [98]. For example, PtO_2 reduction of the oxime **129** gives only the L-*ribo* derivative **130** with the amino group in an axial orientation, whereas borane reduction of acylated oxime **131** produces mainly the L-*arabino* derivative **132** with an equatorial amino substituent (*Scheme 60*) [99].

3.6 Rearrangement Reactions

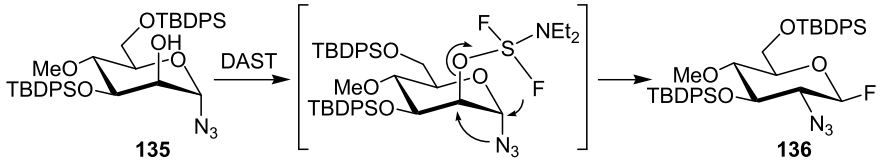
The rearrangement of allyl trichloroacetimidate into allyl trichloroacetamide (Overman rearrangement) has been used for the synthesis of amino sugars. For example, the trichloroace-



Scheme 60



Scheme 61

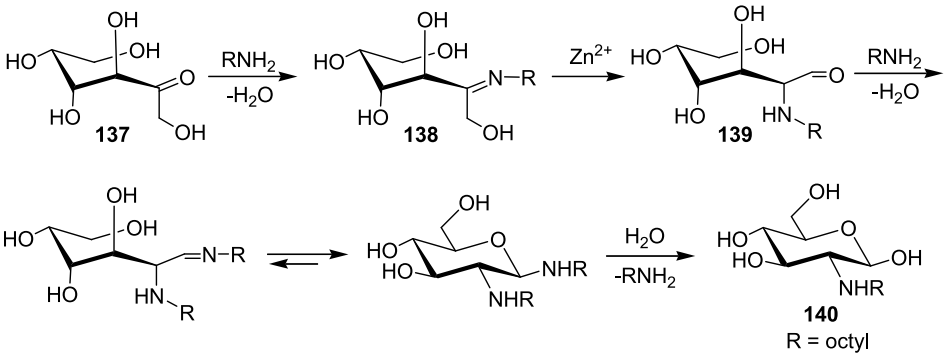


Scheme 62

timidate **133** undergoes reflux in xylene in the presence of potassium carbonate, giving the corresponding trichloroacetamide **134** with total stereoselectivity (*Scheme 61*) [100].
In another interesting reaction, 2-azido-2-deoxypyranosyl fluoride **136** is obtained from the 2-OH unprotected mannosyl azide **135**, by DAST-mediated displacement involving neighboring group participation and migration of the anomeric azide group (*Scheme 62*) [101].

3.7 Miscellaneous Methods

Under dehydrating conditions, D-fructose **137** will effectively react with amines to produce the corresponding imine product **138**, which can be transformed into the aldohexose **139** by using zinc halide as catalyst. Condensation of **139** with another equiv. of amine followed by ring-closure and hydrolysis gives the 2-alkylamino-2-deoxy-D-glucopyranose **140** (*Scheme 63*) [102].



Scheme 63

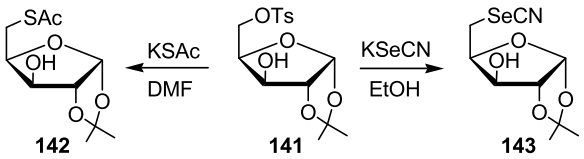
A modification of the Kiliani–Fischer method for amino functionality introduction involves the condensation of free sugars with arylamines to give imines, which are treated directly with hydrogen cyanide to produce aminonitriles. Hydrogenolysis and hydrolysis of the resulting aminonitriles gives 2-amino-2-deoxyaldoses with one carbon extended [103].

4 Introduction of Sulfur and Selenium

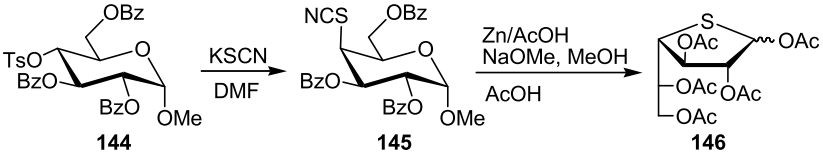
4.1 Nucleophilic Substitutions

Thiosugars discussed here are the carbohydrates with oxygen replaced by sulfur at the nonanomeric positions. Normally these compounds can be obtained by the displacement of suitably positioned leaving groups using reagents containing nucleophilic sulfur such as thiocyanate ion ($^{-}$SCN), thiolacetate [Me(C=O)S^{-}], benzylthiolate (PhCH$_2$S^{-}) and ethyl xanthate [EtO(C=S)S^{-}] [104].

As expected, these substitutions occur readily at primary positions, as outlined by the thioacetate displacement applied to the 5-*O*-tosylxylose derivative **141** giving 5-thio-D-xylose **142** (*Scheme 64*). Similarly, the selenosugar derivative **143** is obtained from compound **141** by substitution using potassium selenocyanate (*Scheme 64*) [104]. And 6-*S*-phenyl-gluco/galactopyranosides are readily prepared from the corresponding 6-hydroxy-glycosides by direct substitution of the hydroxyl group using phenyl disulfide and tri-*n*-butylphosphine [105].



Scheme 64

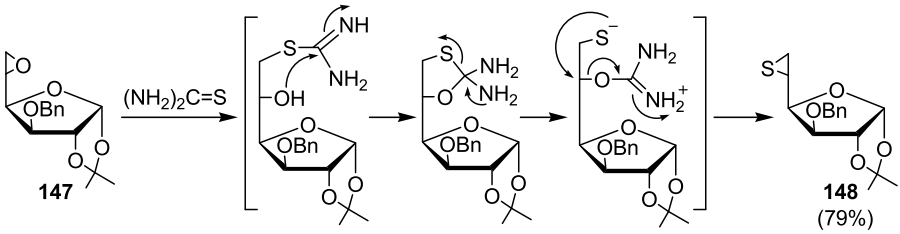


Scheme 65

A sluggish displacement by the thiocyanate ion occurs at C-4 in methyl 2,3,6-tri-*O*-benzoyl-4-*O*-tosyl-α-D-glucopyranoside **144** to give the 4-thiocyanogalactose derivative **145**. Subsequent reduction, deprotection and acetolysis yield preferentially the five-membered form **146** in which sulfur is in the ring (*Scheme 65*) [106].
Phosphorus dithioacids [RR′P(S)SH] are effective sulfur nucleophiles for the synthesis of 3′-*S*- or 5′-*S*-nucleosides [107]. Furthermore, Mitsunobu reactions have also been used to convert primary sugar alcohols into corresponding thiosugar derivatives in one-pot manner [108].

4.2 Opening of Epoxide Rings

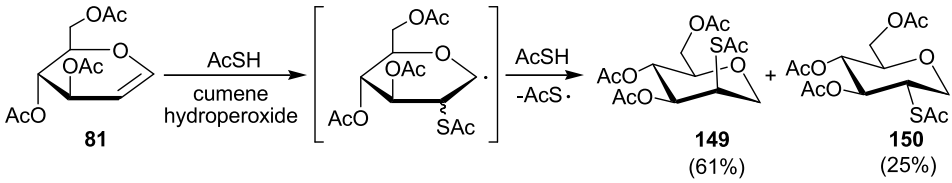
By the means of epoxide ring opening, thiosugars can also be obtained. Sometimes with nucleophiles such as thiocyanate or thiourea a more complex reaction occurs to convert epoxides directly into episulfides, as illustrated by the conversion of the 5,6-anhydro-L-*ido* compound **147** to the D-*gluco*-5,6-episulfide **148** (*Scheme 66*). Also, episulfides can be ring opened with nucleophiles [109].



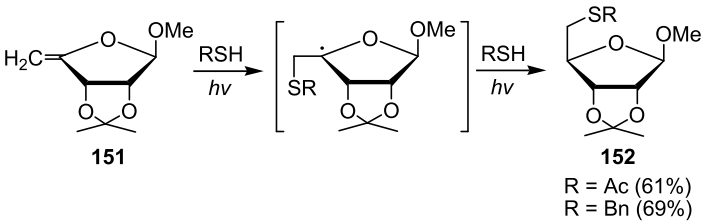
Scheme 66

4.3 Addition to Unsaturated Carbohydrates

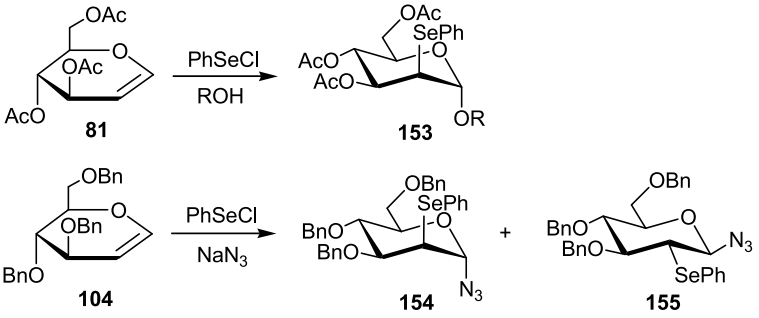
Many known addition reactions of sulfur-containing reagents to unsaturated sugar derivatives involve free radical processes. For example, when triacetyl-D-glucal **81** is treated with thiolacetic acid in the presence of free radical initiators, the mixture of 2-thio-D-mannitol **149** and glucitol **150** is formed (*Scheme 67*) [110].
In an analogous fashion, radical additions to exocyclic unsaturated sugars have also successfully been performed with high regio- and stereoselectivity. Thus, treatment of the compound **151** with benzylthio or thiolacetate radicals gives mainly **152** (*Scheme 68*) [111].



Scheme 67



Scheme 68



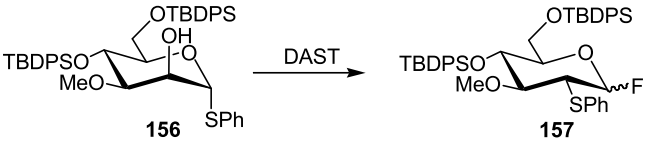
Scheme 69

As for the introduction of a selenium atom into carbohydrates, addition reactions to glycals are also effective. Thus, the glucal **81** can be treated with phenylselenyl chloride and gives mainly the *trans*-diaxial product **153** (*Scheme 69*) [112]. Similarly, azidophenylselenylation of perbenzylated glucal **104** proceeds smoothly to give 2-*Se*-phenyl-2-selenoglycosylazides **154** and **155** (*Scheme 69*) [113].

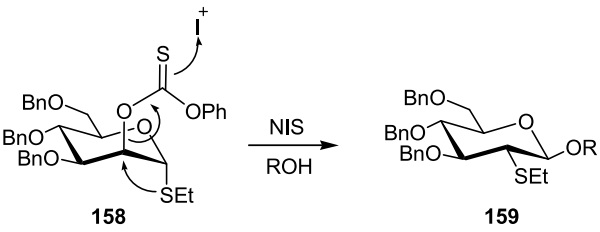
4.4 Rearrangement Reactions

4.4.1 1,2-Migrations

In 1,2-*trans*-thioglycosides, DAST can mediate a 1,2-migration of the anomeric thio functionality through an episulfonium ion to give the 1,2-*trans*-product, as illustrated by the conversion of **156** to **157** (*Scheme 70*) [101]. And under the same conditions 2-hydroxy-1-seleno gly-



Scheme 70



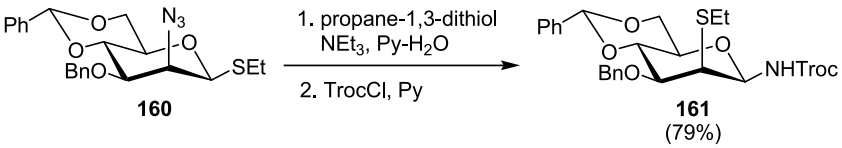
Scheme 71

cosides result in a similar 1,2-migration of the selenium group, with simultaneous installation of a fluoride group at C-1 [113].

In another similar reaction, activation of the phenoxythiocarbonyl ester on C-2 of the thiomannoside **158** by iodonium ions leads to the anomeric thioethyl group migration to the C-2 position, yielding the 1,2-*trans*-glucoside **159** (Scheme 71) [114].

1,2-Migration and concurrent glycosidation of phenyl 2,3-*O*-thionocarbonyl-1-thio-α-L-rhamnopyranosides under the action of methyl trifluoromethanesulfonate (MeOTf) also give in high yields the 3-*O*-(methylthio) carbonyl-2-*S*-phenyl-2,6-dideoxy-β-L-glucopyranosides, which are ready precursors to the corresponding 2-deoxy-β-glycosides [115].

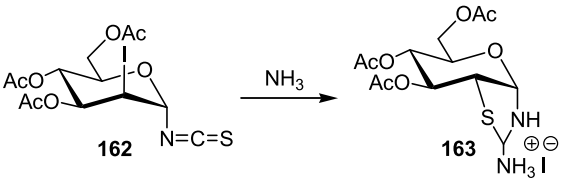
In the reduction of 2-azido-3-*O*-benzyl-4,6-*O*-bezylidene-2-deoxy-1-thio-β-D-mannopyranoside **160** with propane-1,3-dithiol and NEt_3 followed by the protection of the formed amino group, migration of the ethylthio group at C-1 and the amino group at C-2 occurs with the retention of configuration in both positions, to give 3-*O*-benzyl-4,6-*O*-bezylidene-2-*S*-ethyl-2-thio-β-D-mannopyranosyl-(2,2,2-trichloroethoxycarbonyl)amine **161** (Scheme 72) [116].



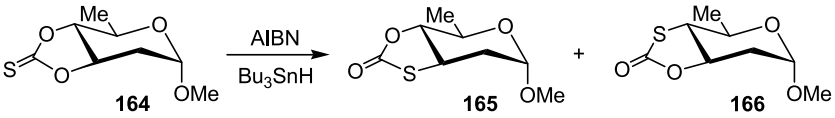
Scheme 72

4.4.2 Rearrangements

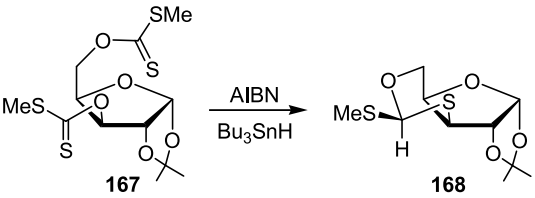
When 2-deoxy-2-iodoglycosylisocyanate **162** is exposed to ammonia, a long-distance rearrangement occurs to yield 1,2-fused aminothiazoline **163** (Scheme 73) [117].



Scheme 73



Scheme 74



Scheme 75

Thionocarbonates may proceed a radical-induced rearrangement to form thiosugars. Thus, treatment of 3,4-thionocarbonate **164** with azo-bis-isobutyronitrile (AIBN) and Bu_3SnH gives the 3- and 4-thio sugar derivatives **165** and **166** with retention of configuration (*Scheme 74*) [118].

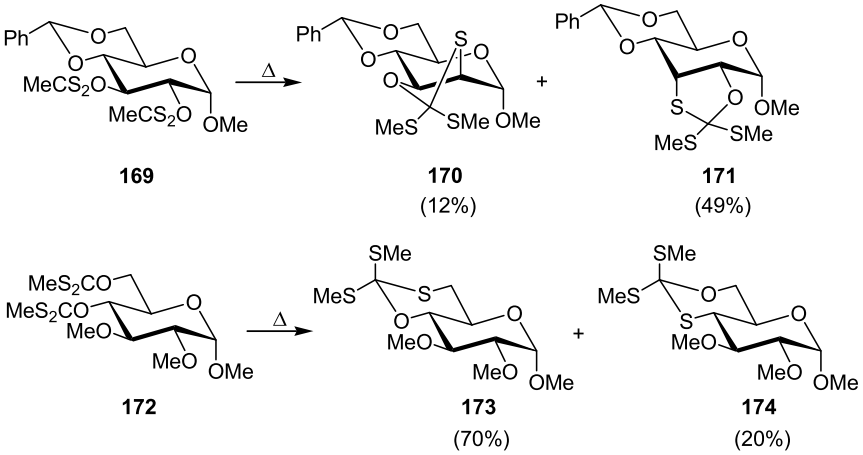
Dixanthates such as **167** also undergo rearrangement reactions under similar radical conditions to afford 3,5-cyclic dithioorthoformate **168** instead of the expected deoxygenation products (*Scheme 75*) [119].

In the case of heating dixanthate **169** at 200 °C, orthotrithiocarbonates **170** and **171** are formed via a regioselective cyclization and with inversion of configuration. A similar procedure can also be applied to the 4,6-dixanthate **172** giving **173** and **174** (*Scheme 76*) [120].

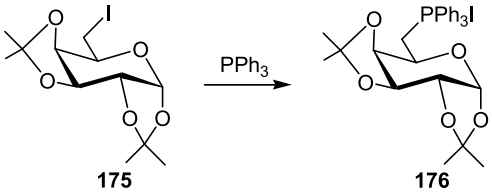
5 Introduction of Phosphorus

5.1 Nucleophilic Substitutions

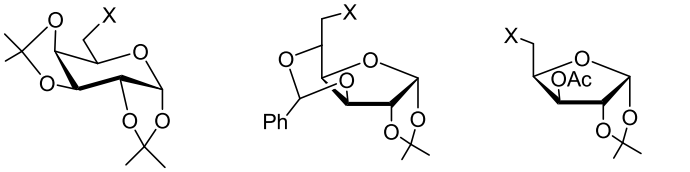
The trivalent phosphorus atom bears a lone pair of electrons and therefore can be used as a nucleophilic reagent for substitution. Triphenylphosphine displacements on alkyl halides give phosphonium salts which, after the conversion into phosphorus ylides by strong bases,



Scheme 76



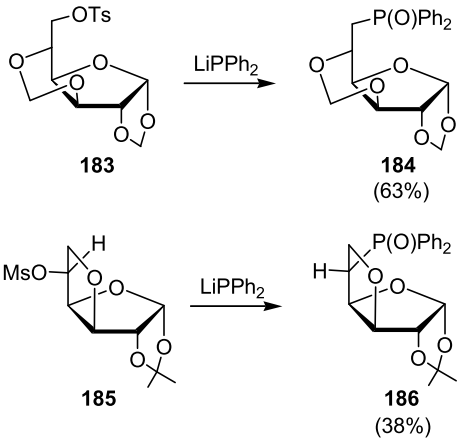
Scheme 77



Scheme 78

are useful in Wittig alkene synthesis. A limited number of reactions of this type have been performed on primary sugar halides. For example, the reaction of 6-iodogalactose derivative **175** with PPh_3 gives the phosphonium iodide **176** (*Scheme 77*) [121].

Other trivalent phosphorus-containing nucleophiles, such as trialkyl phosphites [$P(OR_3)$], dialkyl alkylphosphonites [$R'P(OR)_2$], or alkyl dialkylphosphinites [$R'_2P(OR)$], can convert primary carbohydrate halides **177**, **178**, **179** to corresponding phosphorus derivatives **180**, **181**, **182** via the Michaelis–Arbusov reaction (*Scheme 78*) [122,123].

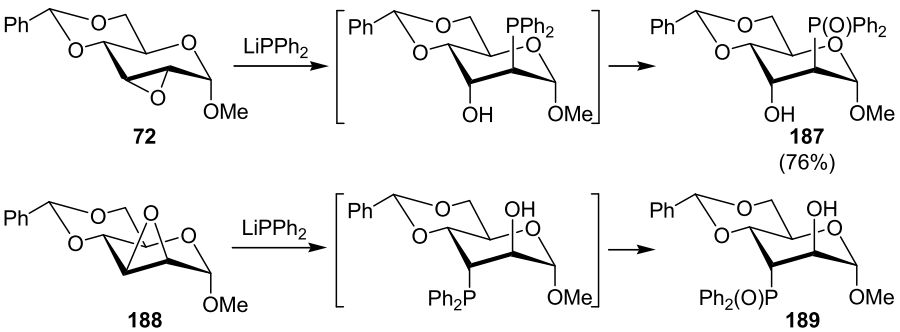


Scheme 79

Lithium diphenylphosphine ($LiPR_2$) has also been used to substitute sulfonates on sugar derivatives such as **183** and **185** to yield, after spontaneous oxidation, **184** and **186** (*Scheme 79*) [124].

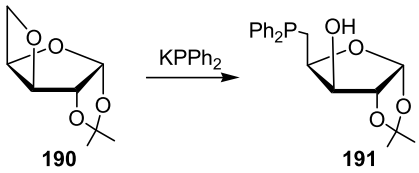
5.2 Ring-Opening Reactions

Epoxide ring opening is also an alternative route for the introduction of a phosphorus atom into carbohydrates. Thus, the D-*allo* epoxide **72** reacts with $LiPPh_2$, giving an intermediate phosphine which is spontaneously oxidized to the 2-phosphonate **187** [125], whereas the isomeric D-*manno* epoxide **188** affords the regioisomeric 3-phosphonate **189** (*Scheme 80*) [124].



Scheme 80

A slight excess of potassium diphenylphosphine can open the oxetane ring of compound **190** to afford phosphine **191** (*Scheme 81*) [126].



Scheme 81

The 1,6-anhydro ring of 1,6-anhydro-D-glucopyranose can also be opened by using phosphinic acid (H_3PO_2) and sodium phosphinate (H_2PO_2Na) to give the 6-phosphate derivative [127].

5.3 Addition Reactions

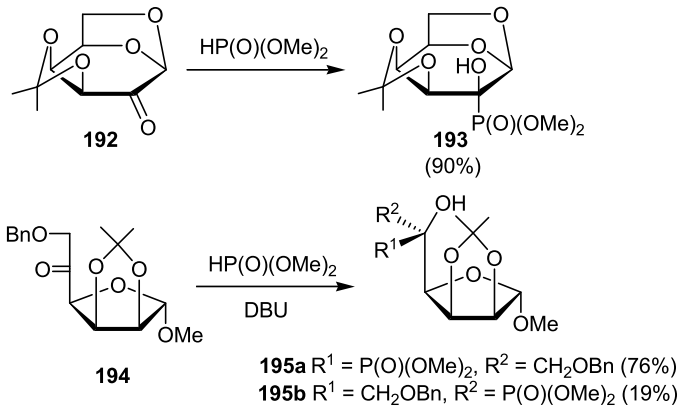
5.3.1 Addition to Carbonyl Compounds

The addition of dialkyl phosphonates [$HP(O)(OR)_2$] to the carbonyl group of oxosugars gives rise to geminal phosphorus adducts. And the introduction orientation of phosphorus is effected by steric hindrance on sugar rings.

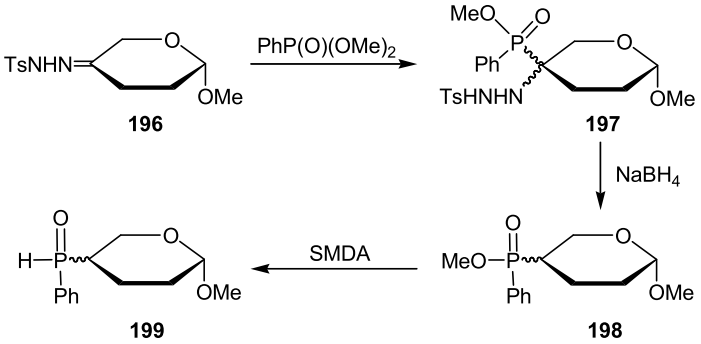
Thus, for example, the reaction of 1,6-anhydro-2-oxo-carbohydrate **192** with dimethyl phosphonate gives rise to the 2-phosphonate derivative **193** with a high stereoselectivity (*Scheme 82*) [128]. And the addition of dimethyl phosphonate to the ulose **194** in the presence of DBU affords the (5*R*)-5-(dimethylphosphinyl)-D-*lyso*-hexofuranoside derivative **195a** and its (5*S*)-epimer **195b** (*Scheme 82*) [129].

Tosylhydrazones can also undergo additions of pentavalent phosphorus derivatives, such as alkyl phenylphosphinate [Ph(H)P(O)OR]; dimethyl phosphonate [$HP(O)(OMe)_2$], or dimethyl phenylphosphonate [$PhP(O)(OMe)_2$], to introduce phosphorus atoms.

For example, compound **196** reacts with dimethyl phenylphosphonate to afford the adduct **197** and reductive removal of the tosylhydrazino group by $NaBH_4$ leads to the phosphi-



Scheme 82

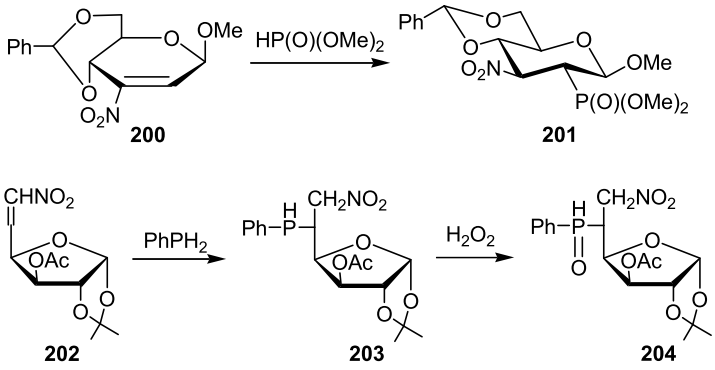


Scheme 83

nate **198**, which can be further reduced to the phosphine oxide **199** by reaction with sodium dihydrobis(2-methoxy-ethoxy)aluminate (SMDA) (*Scheme 83*) [130].

5.3.2 Addition to Isolated Double Bond

Dialkyl phosphonates [e. g. $HP(O)(OMe)_2$] can be added to isolated double bonds of the 3-nitro derivative **200** to give the 3-nitro-2-phosphonate **201**. Also, phenylphosphine can react with the exocyclic double bond of compound **202** to afford the phosphine **203** in 63% yield together with a dimeric compound [131]. And compound **203** can be further oxidized to produce the phosphine oxide **204** quantitatively (*Scheme 84*) [132].

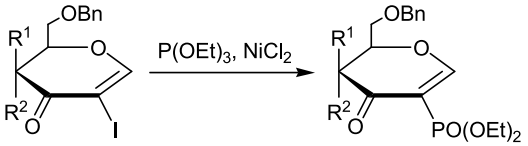


Scheme 84

5.4 Coupling Reaction

A modified Arbuzov reaction has been applied to the coupling of the D-glucal/D-galactal derived vinyl halides **205** and triethyl phosphite at 150 °C in the presence of $NiCl_2$

as a catalyst, and 2-(diethoxyphosphoryl)hex-1-en-3-uloses **206** are obtained in fair yields (◉ *Scheme 85*) [133].



205a, 205b **206a, 206b**

a: R^1 = H, R^2 = OBn (60%)
b: R^1 = OBn, R^2 = H (60%)

Scheme 85

References

1. Hanessian S (1997) Selected reactions in carbohydrate chemistry. In: Hanessian S (Ed) Preparative carbohydrate chemistry. Dekker, New York, p 85
2. Boons GJ (2000) Functionalised saccharides. In: Boons GJ, Hale KJ (Eds) Organic synthesis with carbohydrates. Sheffield Academic Press, Sheffield, p 56
3. Szarek WA (1973) Adv Carbohydr Chem Biochem 28:250
4. Szarek WA (1973) Adv Carbohydr Chem Biochem 28:230
5. Card PJ (1985) J Carbohydr Chem 4:451
6. Wu MC, Anderson L, Slife TW, Jensen LJ (1974) J Org Chem 39:3014
7. Hanessian S, Vatèle JM (1981) Tetrahedron Lett 22:3579
8. Nadkarni S, Williams NR (1965) J Chem Soc 3496
9. Capon B (1969) Chem Rev 69:407
10. Brimacombe JS, Doner LW, Rollins AJ (1972) J Chem Soc Perkin Trans 1 2977
11. Brimacombe JS, Doner LW, Rollins AJ (1973) J Chem Soc Perkin Trans 1 1295
12. Buss DH, Hough L, Richardson AC (1963) J Chem Soc 5295
13. Zu Reckendorf WM, Bonner WA (1963) Tetrahedron 19:1711
14. Binkley RW, Hehemann DG (1979) Adv Carbohydr Chem 24:139
15. Vox JN, Van Boom JH, Van Boechel CAA, Beetz T (1984) 3:117
16. Richardson AC (1969) Carbohydr Res 10:395
17. Gnichtel H, Rebentisch D, Tompkins TC, Gross PH (1982) J Org Chem 47:2691
18. Wright JA, Taylor NF (1966–1967) Carbohydr Res 3:333
19. Buchanan JG, Schwarz JCP (1962) J Chem Soc 4770
20. Guthrie RD, Williams GJ (1976) J Chem Soc Perkin Trans 1 801
21. Christensen JE, Goodman L (1961) J Am Chem Soc 83:3827
22. Van Der Klein PAM, Filemon W, Veeneman GH, Van Der Marel GA, Van Boom JH (1992) J Carbohydr Chem 11:837
23. Hanessian S, Pleassas NR (1969) J Org Chem 34:1035, 1045 and 1053
24. Ferier RJ, Prasad N (1969) J Chem Soc 570
25. Ferier RJ, Ponpipom MJ (1971) J Chem Soc Chem Commun 553
26. Leblanc Y, Fitzsimmons BJ, Springer JP, Rokach J (1989) J Am Chem Soc 111:2995
27. Dios A, Geer A, Marzabadi CH, Franck RM (1998) J Org Chem 63:6673
28. Dahl RS, Finney NS (2004) J Am Chem Soc 126:8356
29. Giese B (1986) Radical in organic synthesis: Formation of carbon-carbon bonds. Pergamon Press, Oxford
30. Taylor NF (1988) Fluorinated carbohydrate. Chemical and biological aspects, ACS Symp Ser 374
31. Beuthien-Baumann B, Hamacher K, Oberdorfer F, Steinbach J (2000) Carbohydr Res 327:107

32. Tomas SS, Plenkiewicz J, Ison ER, Bols M, Zou W, Szarek WA, Kisilevsky R (1995) Biochim Biophys Acta 1272:37
33. Suami T, Hough L, Tsuboi M, Machinami T, Watanabe N (1994) J Carbohydr Chem 13:1079
34. Szarek WA (1973) Adv Carbohydr Chem Biochem 28:225
35. Su TL, Klein RS, Fox JJ (1982) J Org Chem 47:1506
36. Hamacher K, Coenen HH, Stocklin G (1986) J Nucl Med 27:235
37. Tsuchiya T, Takahashi Y, Endo M, Umezawa S, Umezawa H (1985) J Carbohydr Chem 4:587
38. Haradahira T, Maeda M, Kai Y, Omae H. Kojima M (1985) Chem Pharm Bull 33:165
39. Doboszewski B, Hay GW, Szarek WA (1987) Can J Chem 65:412
40. Anisuzzaman AKM, Whistler RL (1978) Carbohydr Res 61:511
41. Whistler RL, Anisuzzaman AKM (1980) Methods Carbohydr Chem 8:227
42. Haylock CR, Melton LD, Slessor KN, Tracey AS (1971) Carbohydr Res 16:375
43. Hanessian S, Ponpidom MM, Lavallee P (1972) Carbohydr Res 24:45
44. Garegg PJ, Samuelsson B (1980) J Chem Soc Perkin Trans 1 2866
45. Abram TS, Baker R, Exon CM (1982) J Chem Soc Perkin Trans 1 285
46. Garegg PJ (1984) Pure Appl Chem 56:845
47. Kielberg J, Frejd T, Jansson K, Sundin A, Magnusson G (1988) Carbohydr Res 176:271
48. Mitsunobu S (1981) Synthesis 1
49. Hannelore V, Brandstetter H, Zbiral E (1980) Helv Chim Acta 63:327
50. Verheyden JPH, Moffatt JG (1970) J Org Chem 35:2319
51. Hanessian S, Plessas NR (1969) J Org Chem 34:2163
52. Back TG, Barton DHR, Rao BL (1977) J Chem Soc Perkin Trans 1 1715
53. Hanessian S (1968) Adv Chem Ser 74:159
54. Bragg PD, Jones JKN, Turner JC (1959) Can J Chem 37:1412
55. Jones JKN, Perry MB, Turner JC (1960) Can J Chem 38:1122
56. Jennings HJ, Jones JKN (1963) Can J Chem 41:1151
57. Jennings HJ, Jones JKN (1965) Can J Chem 43:2372
58. Jennings HJ (1970) Can J Chem 48:1834
59. Jennings HJ, Jones JKN (1965) Can J Chem 43:3018
60. An SH, Bobek M (1986) Tetrahedron lett 27:3219
61. Card PJ (1983) J Org Chem 48:393
62. Card PJ (1985) J Carbohydr Chem 4:453
63. Kobayashi S, Yoneda A, Fukuhara T, Hara S (2004) Tetrahedron lett 45:1287
64. Penglis AAE (1981) Adv Carbohydr Chem Biochem 38:195
65. Tsuchiya T (1990) Adv Carbohydr Chem Biochem 48:91
66. Hartman MCT, Coward JK (2002) J Am Chem Soc 124:10036
67. Patroni JJ, Stick RV (1987) Aust J Chem 40: 795
68. Fuentes J, Andulo M, Pradera MA (1998) Tetrahedron lett 39:7149
69. Fuentes J, Andulo M, Pradera MA (1999) Carbohydr Res 319:192
70. Chana JS, Collins PN, Farnia F, Peacock DJ (1988) J Chem Soc Chem Commun 94
71. Ido T, Wan CN, Fowler JS, Wolf AP (1977) J Org Chem 42:2341
72. Korytnyk W, Petrieó CR (1982) Tetrahedron 38:2547
73. Thiem J, Klaffke W (1992) Top Curr Chem 154:285
74. Horton D, Priebe W, Snaidman M (1990) Carbohydr Res 205:71
75. Lafont D, Descotes G (1987) Carbohydr Res 166:195
76. Rawal GK, Rani S, Madhusudanan KP, Vankar YD (2007) Synthesis 2:294
77. Mc Carter JD, Adam MJ, Wihters SG (1995) Carbohydr Res 266:273
78. Ansell MF (1983) Rodd's chemistry of carbon compounds, vol 1F, G Supplement. Elsevier, Amsterdam
79. Blattner R, Ferrier RJ (1980) J Chem Soc Perkin Trans 1 1523
80. Ferrier RJ, Haines SR, Gainsford GJ, Gabe EJ (1984) J Chem Soc Perkin Trans 1 1683
81. Karpeisiuk W, Banaszek A, Zamojski A (1989) Carbohydr Res 186:156
82. Kloosterman M, Westerduin P, Van Boom JH (1986) Recl Trav Chim Pays-Bas 105:136
83. Viand MC, Rollin P (1990) Synthesis 130
84. Knapp S, Kukkola PJ, Sharma S, Murali Dhar TG, Naughton ABJ (1990) J Org Chem 55:5700
85. De Oliveira RN, Cottier L, Sinou D, Srivastava RM (2005) Tetrahedron 61:8271
86. Williams NR (1970) Adv Carbohydr Chem Biochem 25:109

87. Lemieux RU, James K, Nagabhushan TL (1973) Can J Chem 51:48
88. Lemieux RU, Ratcliffe M (1979) Can J Chem 57:1244
89. Czernecki S, Ayadi E, Randriamandimby D (1994) J Org Chem 59:8256
90. Griffith DA, Danishefsky SJ (1990) J Am Chem Soc 112:5811
91. Du Bois J, Tomooka CS, Hong J, Carreira EM (1997) J Am Chem Soc 119:3179
92. Dyong I, Schilte G, Lam-Chi Q, Friege H (1979) Carbohydr Res 68:257
93. Ravindran B, Sakthivel K, Suresh CG, Pathak T (2000) J Org Chem 65:2637
94. Nicolaou KC, Baran PS, Zhong YL, Vega JA (2000) Angew Chem 112:2625
95. Paul HW, Fraser-Reid B (1986) Carbohydr Res 150:111
96. Baer HH (1972) Methods Carbohydr Chem 6:245
97. Guthrie RD, Johnson LF (1961) J Chem Soc 4166
98. Lemieux RU, James K, Nagabhushan TL, Ito Y (1973) Can J Chem 51:33
99. Pelyvas I, Hasegawa A, Whistler RL (1986) Carbohydr Res 146:193
100. Montero A, Mann E, Herradón B (2005) Tetrahedron Lett 46:401
101. Nicolaou KC, Ladduwahetty T, Randall JL, Chucholowski A (1986) J Am Chem Soc 108:2466
102. Piispanen PS, Norin T (2003) J Org Chem 68:628
103. Brossmer R (1962) Methods Carbohydr Chem 1:216
104. Trimnell D, Stout EI, Doane WM, Russel CR (1975) J Org Chem 40:1337
105. Yu B, Zhu X, Hui Y (2001) Tetrahedron 57:9403
106. Arela O, Cicero D, De Lederkremer RM (1989) J Org Chem 54:1884
107. Dabbkowski W, Michalska, M, Tworowska I (1998) J Chem Soc Chem Commun 427
108. von Itzstein M, Jenkins MJ, Mocerino M (1990) Carbohydr Res 208:287
109. Whistler RL, Lake WC (1972) Methods Carbohydr Chem 6:286
110. Igarashi K, Honma T (1970) J Org Chem 35:606
111. Matsuura K, Maeda S, Araki Y, Ishido Y (1970) Tetrahedron Lett 2869
112. Kaye A, Neidle S, Reese CB (1988) Tetrahedron Lett 29:2711
113. Nicolaou KC, Mitchell HJ, Fylaktakidou KC, Suzuki H, Rodríguez RM (2000) Angew Chem Int Ed 39:1089
114. Zuurmond HM, Van Deklein PAM, Van Der Marel GA, Van Boom JH (1993) Tetrahedron 49:6501
115. Yu B, Yang Z (2001) Org Lett 3:377
116. Veselý J, Rohlenová A, Džoganová M, Trnka T, Tišlerová I, Šaman D, Ledvina M (2006) Synthesis 4:699
117. Santoyo-Gonzales F, Garcia-Calvo-Flores F, Isaac-Garcia J, Hernandez-mateo P, Garcia-mendoza P, Robles-Diaz R (1994) Tetrahedron 50:2877
118. Somask L, Ferrier RJ (1991) Adv Carbohydr Chem Biochem 49:37
119. Herdewijn PAM, Van Aerschot A, Jie L, Esmans E, Peneau-Dupont J, Declerc JP (1991) J Chem Soc Perkin Trans 1 1729
120. Faure A, Kryczka B, Descotes G (1979) Carbohydr Res 74:127
121. Karpiesiuk W, Banaszek A (1994) Tetrahedron 50:2965
122. Seo K, Inokawa S (1970) Bull Chem Soc Jpn 43:3224
123. Seo K, Inokawa S (1975) Bull Chem Soc Jpn 48:1237
124. Hall LD, Steiner PR (1971) J Chem Soc Chem Commun 84
125. Brown MA, Cox PJ, Howie RA, Melvin OA, Taylor OJ, Wardell JL (1995) J Organomet Chem 498:275
126. Pàmies O, Diéguez M, Net G, Ruiz A, Claver C (2001) J Org Chem 66:8364
127. Nifant'ev EE, Gudkova IP, Kochetkov NK (1970) J Gen Chem USSR 40:425
128. Paulson H, Greve W (1973) Chem Ber 106:2124
129. Hanaya T, Yamamoto H (2002) Helv Chim Acta 85:2608
130. Yamashita M, Yoshikane M, Ogawa T, Inokawa S (1979) Tetrahedron 35:741
131. Paulson H, Greve W (1973) Chem Ber 106:2114
132. Takayanagi H, Yamashita M, Seo K, Yoshida H, Ogata T, Inokawa S (1974) Carbohydr Res 38:C19
133. Leonelli F, Capuzzi M, Calcagno V, Passacantilli P, Piancatelli G (2005) Eur J Org Chem 2671

2.4 Anhydrosugars

Sławomir Jarosz, Marcin Nowogródzki
Institute of Organic Chemistry, Polish Academy of Sciences,
01–224 Warsaw, Poland
sljar@icho.edu.pl

Abstract

The anhydrosugars are reviewed in this chapter. The emphasis is placed on the general methodology of their preparation, as well as, their application in stereocontrolled organic synthesis. The material is divided into two main parts: anomeric anhydrosugars and non-anomeric ones. In the first class, 1,2-sugar epoxides and 1,6-anhydrosugars are the most important since they offer significant synthetic potential in targeted synthesis of important compounds. From the second class, sugar epoxides (2,3- and 3,4-oxiranes) are particularly useful. The synthesis of both classes of anhydrosugars (anomeric and non-anomeric) is illustrated by selected examples including older examples (i. e. those described also in the previous edition of this monograph) and those published more recently. Various methods for the preparation of anhydrosugars are described in order to give the reader a general impression of the importance of such derivatives both as biologically active targets (illustrated by selected examples of anhydronucleosides) and optically pure building blocks in targeted synthesis. The material included in this chapter

In: *Glycoscience*. Fraser-Reid B, Tatsuta K, Thiem J (eds)
Chapter-DOI 10-1007/978-3-540-30429-6_6:

should help the reader solve any problems faced in their laboratories connected with planning and execution of the synthesis of optically pure targets.

Keywords

Anhydrosugars; Synthesis; Sugar oxiranes; 1,2-Sugar epoxides; 1,6-Anhydrosugars; Rearrangement

Abbreviations

AZT	3-azidothymidine
DET	diethyl tartrate
DIAD	di-isopropyl azodicarboxylate
DMDO	dimethyldioxirane
IDCP	iodonium dicollidine perchlorate
LNA	locked nucleic acids
MCPBA	*m*-chloroperbenzoic acid
OTs	toluene-*p*-sulfonate
RCM	ring closing metathesis
RNA	ribonucleic acid
THF	tetrahydrofuran
THP	tetrahydropyran
TPP	triphenylphosphine

1 Introduction

The material presented in this chapter describes the concise methodology of the preparation of anhydrosugars and their application in synthesis, updating the chapters published in the last edition [1,2]. However, basic information from the earlier edition is also included here allowing the reader to follow the present text more easily. Therefore, citation to this data refers (mostly) to the previously published chapters and not to original papers.

The anhydrosugars are an important class of saccharides, the synthesis and properties of which are described in many monographs [3,4,5,6,7]. Such derivatives are used for the preparation of modified carbohydrates (including *C*-glycosides) and complex enantiomerically pure products in which the chirality of the parent sugar is transferred to the target ('chiron approach' [8]). Many compounds from this class are also components of biologically active products. Anhydrosugars (also called 'intramolecular anhydrides') are derivatives that formally arise from elimination of the molecule of water from the parent carbohydrate. Various carbon atoms may be engaged in this formal process, thus providing different classes of anhydrosugars [3]. Generally, they belong to two main groups in which: (1) the anomeric carbon atom is involved in the anhydro structure and (2) the anhydro linkage is built between other carbon atoms of the sugar (◉ *Fig. 1*). These two classes of compounds will be described separately.

To construct the anhydro skeleton in the sugar molecule standard synthetic methods, used for the formation of 'normal' heterocyclic derivatives, are applicable [4,5]. Generally, one of the hydroxyl groups of the diol (from which the anhydro ring is formed) is activated (by

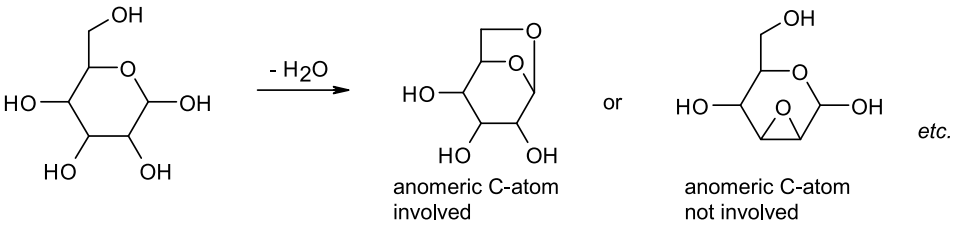


Figure 1
Examples of anhydrosugars derived from parent monosaccharides by formal elimination of the molecule of water

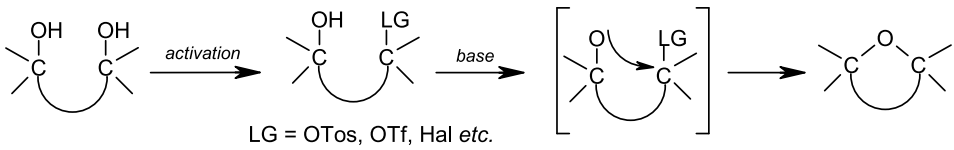


Figure 2
General method leading to anhydrosugars

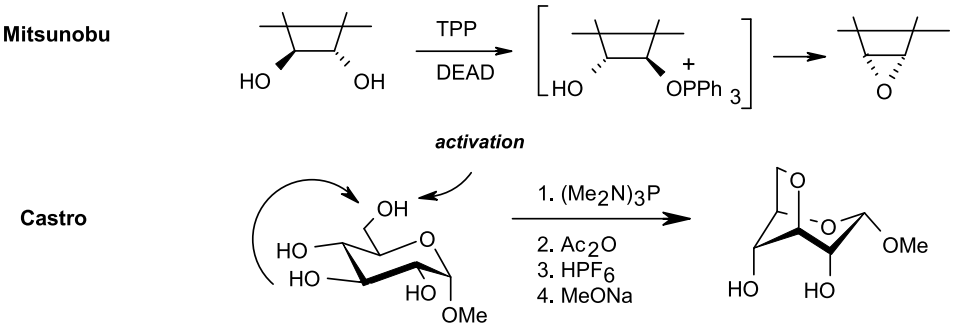


Figure 3
Synthesis of anhydrosugars by the Mitsunobu and Castro reaction

conversion into the leaving group such as tosylate, triflate, halogen etc.), while the second one acts as a nucleophile in basic media, which results in an intramolecular S_N2 closure of the anhydro ring (*Fig. 2*).

Milder conditions, such as activation of one of the hydroxyl groups according either to Mitsunobu [9] or Castro [10] protocol, are also used for the preparation of anhydrosugars, which may be illustrated by examples shown in *Fig. 3*.

Compounds with the oxirane rings are special cases which, beside this general methodology, are also prepared by a number of other methods (direct epoxidation, reaction with sulfur ylides, Darzens' reaction, etc.) [1].

Anhydrosugars are also classified on the basis of the size of the heterocyclic anhydro ring. Thus, oxiranes (sugar epoxides), oxetanes, tetrahydrofuran, and tetrahydropyran derivatives are known. Reactivity of anhydrosugars is determined by the size of the heterocyclic anhy-

dro ring. The most reactive are oxiranes and oxetanes, with oxolane (THF) and oxane (THP) derivatives being much less reactive. Selected examples of the representative derivatives: their synthesis and reactivity, from each group will be presented.

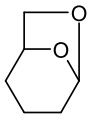
2 Anomeric Anhydrosugars: Synthesis and Reactions

Anomeric anhydrosugars represent a class of molecules that can be regarded as intramolecular glycosides. The glycosydic bond may engage the terminal alcohol of a sugar molecule (1,6-anhydropyranoses, 1,6-anhydrofuranoses, etc.) or any other atom (1,2-, 1,3-anhydrosugars etc.). The most representative are 1,6- and 1,2-anhydro-sugars; others are less common and have rather limited synthetic potential [3].

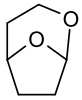
2.1 1,6-Anhydrosugars

2.1.1 1,6-Anhydrohexopyranoses

The most common 1,6-anhydrosugars are anhydro-aldopyranoses, which are formal derivatives of 6,8-dioxabicyclo[3.2.1]octane; less common are anhydro-aldofuranoses.

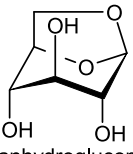


1,6-anhydrohexopyranose
6,8-dioxabicyclo[3.2.1]octane

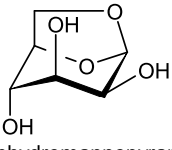


1,6-anhydrohexofuranose
2,8-dioxabicyclo[3.2.1]octane

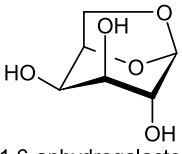
1,6-Anhydro-β-D-glucopyranose (levoglucosan), the most representative example of this class of compounds, was first isolated in pure form in 1894 by Tanret upon treatment of naturally occurring phenolic glycoside with barium hydroxide [2]. This compound is now conveniently obtained by pyrolysis of starch or cellulose and its production is covered by many patents. Other 1,6-anhydrosugars with different configurations, such as mannosan or galactosan, may be prepared by pyrolysis of mannan ivory nut meal or α-lactose [2,3].



1,6-anhydroglucopyranose

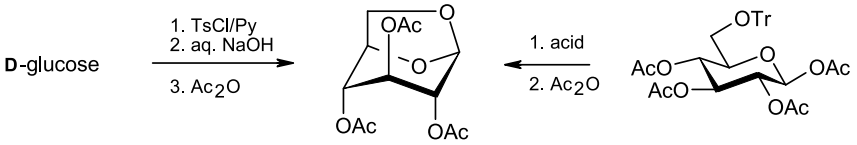


1,6-anhydromannopyranose

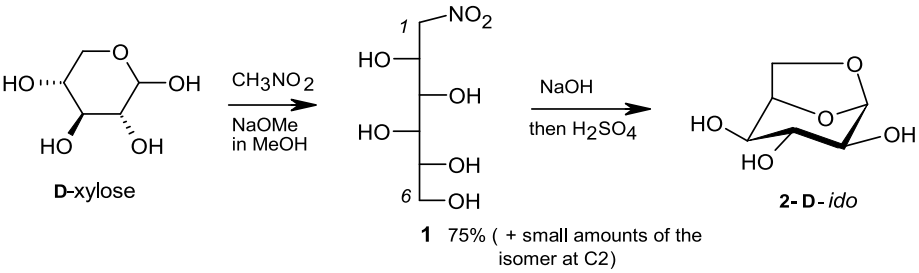


1,6-anhydrogalactopyranose

Generally, treatment of a sugar with a free hydroxyl group at the C-6 position possessing a good leaving group at the anomeric position (halogen, azide, tosylate etc.) with a strong base affords 1,6-anhydropyranoses [3]. Another method consists of a selective activation of a terminal hydroxyl group; this procedure is applicable for most 1,6-anhydro-pyranoses. Fraser-Reid and co-workers described a large-scale synthesis of 1,6-anhydro-D-*gluco*- and D-*manno*-pyranoses from free sugars by selective activation of the terminal position with tosyl chloride



Scheme 1

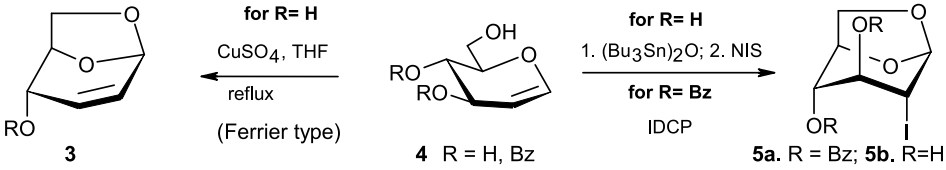


Scheme 2

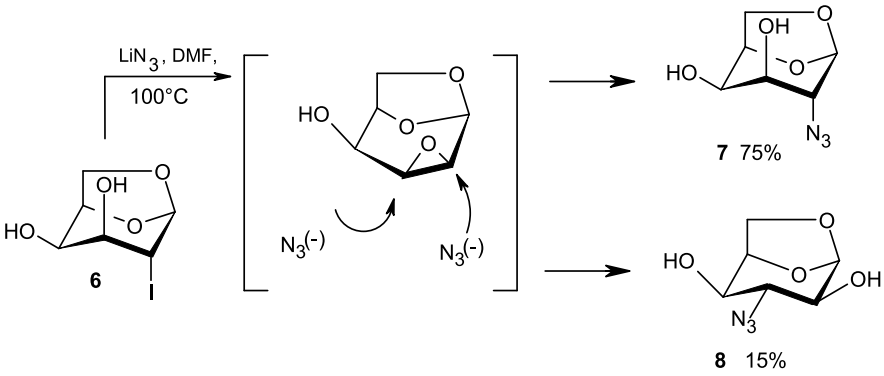
followed by cyclization of the intermediate in basic media [11]. This process can be facilitated by microwave irradiation [12]. Alternatively, the labile ethers (such as trityl) at the C6 position, while activated with Lewis acid, provide the anhydrosugars in excellent yield (*Scheme 1*) [2].

In some cases special activation is not necessary. In acidic media, D-idose is easily converted into 1,6-anhydro-D-*ido*-pyranose, because the equatorial arrangement of all hydroxyl groups in the anhydro molecule facilitates this process. For example, starting from D-xylose, Köll prepared 1,6-anhydro-D-idose in good yield in a sequence of reaction shown in *Scheme 2* [13]. The free idose, obtained from 1-deoxy-1-nitro-D-iditol (**1**) by the Nef reaction, cyclized readily to the desired anhydrosugar [13]. During the synthesis of L-idose from the D-glucose [14] one may encounter the problem of the isolation of the free sugar, since it readily undergoes cyclization to the anhydro form.

Other methods of preparation of these valuable compounds are shown in *Scheme 3*. All of them are based on the intramolecular glycosylation of glycals by a terminal hydroxyl group. Treatment of the 3,4-di-*O*-benzoyl-D-glucal (**4**; R = Bz) with iodonium dicollidine perchlorate (IDCP) results in formation of the anhydro derivative **5a** (R = Bz), while activation of the hydroxyl group in a free glycal (**4**; R = H) with $(Bu_3Sn)_2O$ followed by iodocyclization pro-



Scheme 3

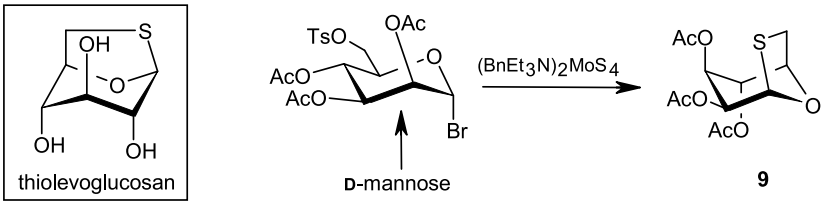


Scheme 4

vides the free compound **5b** (R = H). Internal Ferrier-type glycosylation leading to compound **3** was achieved by treatment of the free glycal with Lewis or protic acid (*Scheme 3*) [2].

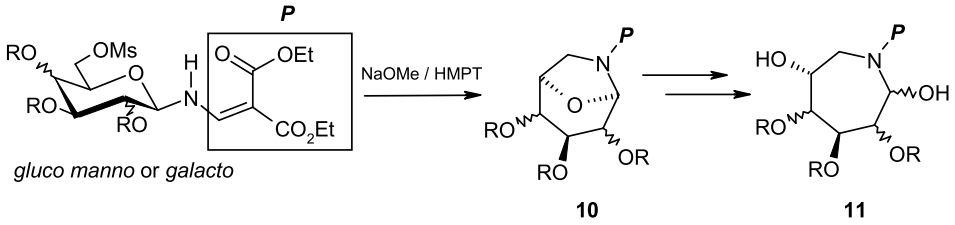
The presence of a leaving group in the anhydro skeleton opens up an easy path to modified anhydrosugars. For example, reaction of 2-iodo-2-deoxy-1,6-anhydrogalactose (**6**) with lithium azide provides the 2-azido-galactose derivative **7** together with small amounts of 3-azido-idose derivative **8** (*Scheme 4*).

1,6-Anhydro-pyranoses with other than an oxygen heteroatom in the ring are known (although not common). Thiolevoglucosan has been known for many years [15]. An interesting example of the synthesis of 1,6-anhydrothiomannose (**9**) was presented recently [16].

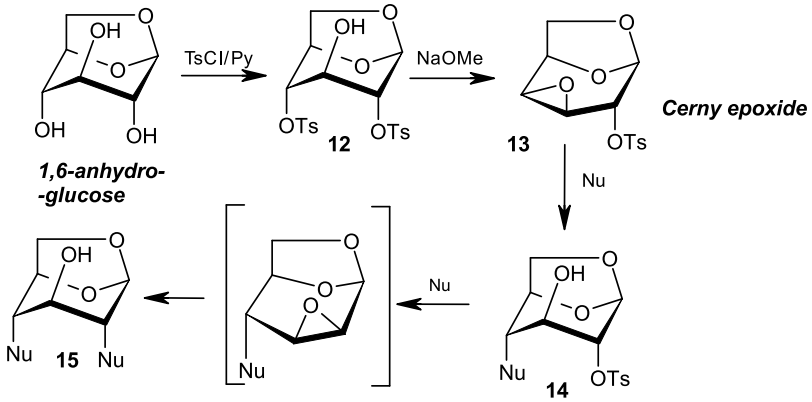


Such 'hetero' anhydrosugars may be used as convenient synthetic intermediates. For example, septanose iminosugars **11** were prepared via the aza-anhydrosugars **10** by Fuentes (*Scheme 5*) [17].

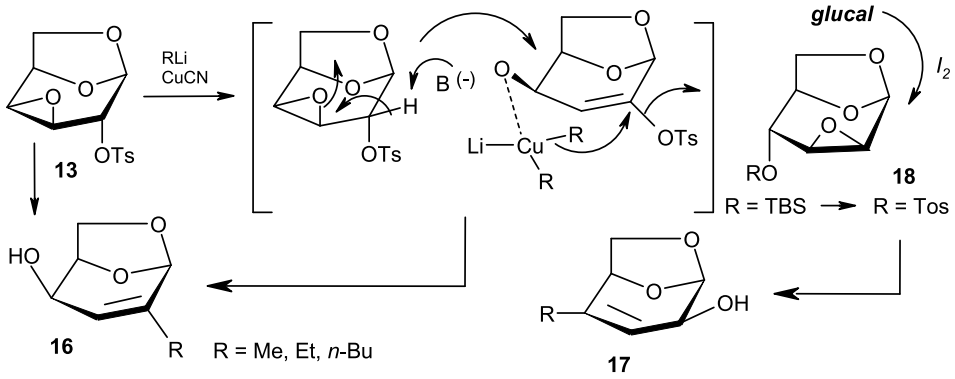
The hydroxyl groups in anhydrosugars differ in steric orientation and, hence, their reactivity is also different [3]. This feature may be well illustrated by selective transformations of the hydroxyl groups in 1,6-anhydroglucose. Because of the $^1C_4$ conformation of its skeleton, all hydroxyl groups in 1,6-anhydro-D-glucose are placed at the axial positions. The hydroxyl group at the C-3 position is most hindered, thus less reactive than those at the C-2 and C-4 positions. This offers a great advantage for the preparation of many useful building blocks. For example, di-tosylation of 1,6-anhydro-D-glucopyranose leads to 2,4-di-*O*-tosyl derivative **12**, which readily cyclizes to the 'Cerny' epoxide **13** (*Scheme 6*) [3,6].



Scheme 5



Scheme 6



Scheme 7

Opening of the three-membered ring in **13** with various nucleophiles provides the corresponding mono- (**14**; at the C-2 position) and di-substituted (**15** at the C-2 and C-4 positions) derivatives (*Scheme 6*) [3].

However, when organocuprate is used as a nucleophile, another product is formed. Instead of opening the oxirane ring, cuprate acts as a base which abstracts a proton from the C-2

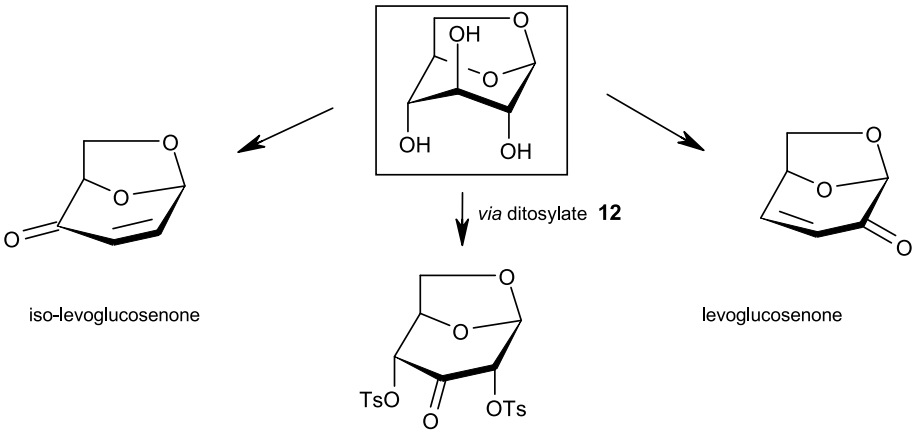
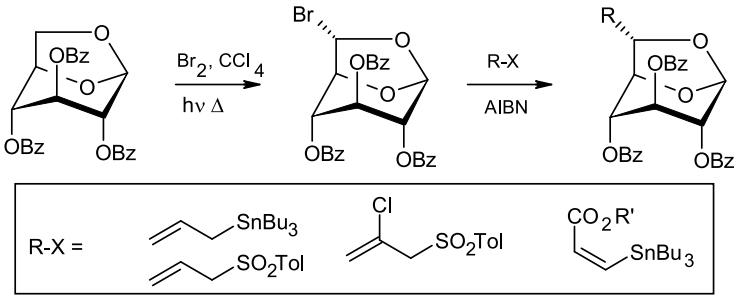


Figure 4
Examples of keto-anhydrosugars prepared from 1,6-anhydroglucose



Scheme 8

position inducing a cascade process leading finally to 2,3-unsaturated anhydrosugar **16** with a nucleophile placed at the C-2 position [18].

The alternative compound **17**, with the nucleophile placed at the C-4 position, was prepared by the same method [19] from regioisomeric (to Cerny epoxide) oxirane **18**, which was obtained from D-glucal by iodocyclization followed by standard reactions (*Scheme 7*).

Many useful building blocks (besides those already presented) with preservation of the 1,6-anhydro skeleton such as levoglucosenone and its regioisomer *iso*-levoglucosenone can be prepared from 1,6-anhydroglucose [20]. The carbonyl group can be also introduced at the C-3 position by oxidation of the 2,4-di*O*-tosyl-1,6-anhydroglucose (**12**) [3] (*Fig. 4*).

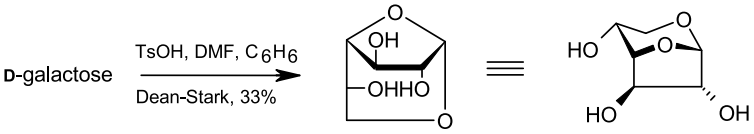
Besides the reactions involving the oxygen atoms, other processes directed to functionalization of the carbon skeleton are known. One of the most useful strategies is based on bromination of the C-6 position and further replacement of the bromine atom with a functional group under the radical conditions (*Scheme 8*) [21].

Such derivatives, functionalized at the C-6 position, are convenient synthons for a variety of interesting compounds.

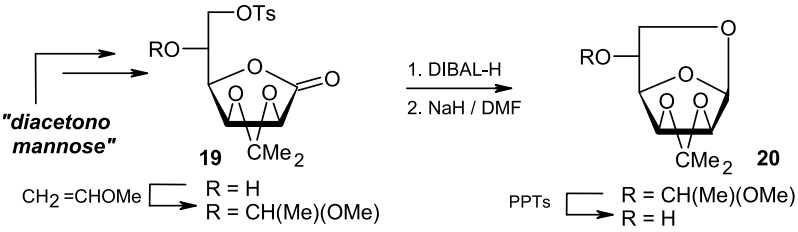
2.1.2 1,6-Anhydrofuranoses

1,6-Anhydrofuranoses derived from all eight diastereoisomeric aldohexoses were described in the literature [6]. These compounds are formed as side products in the synthesis of 1,6-anhydropyranoses from free (or partially protected) sugars. Treatment of free sugars with toluene-*p*-sulfonic acid in DMF solution affords the furanose and pyranose 1,6-anhydrides with the furanose form up to 33% for the *galacto*-, *allo*-, and *talo*- isomers [3]; the example is shown in ❯ *Scheme 9*.

A common procedure for the efficient synthesis of 1,6-anhydrofuranoses is based on cyclization of the 6-*O*-tosyl derivative of furanose with the free anomeric position and protected hydroxyl function at the C-5 position. Synthesis of 1,6-anhydro-D-mannofuranose is a good example of such a strategy (❯ *Scheme 10*) [22].



Scheme 9



Scheme 10

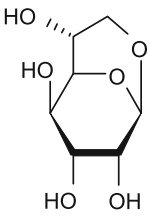
Standard transformation of the "diacetonomannose" provided 6-*O*-tosyl-5-*O*-protected lactone **19**, which was reduced to lactole with diisobutyl hydride. Cyclization under basic conditions afforded 1,6-anhydromannofuranose **20**, which can be deprotected easily at the C-5 position.

2.2 Higher Anhydroaldoses and Anhydroketoses

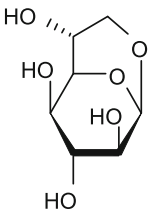
Higher sugars also can form the anhydrides in which the anomeric position and the terminal carbon atom are linked via a heterocyclic ring. For example, heptoses can form 1,7-anhydro derivatives although they are very rare. Only a few examples of such derivatives have been described in the literature (❯ *Fig. 5*) [3].

Anhydroketoses are derived mostly from 2-ketosugars and have similar properties to anhydroaldoses. Several such derivatives (e. g. **21**) were isolated from residues after pyrolysis of

parent ketohexoses such as, for example, fructose [3,23]. The 3,8-anhydro derivative **24**, was isolated during the synthesis of D-*glycero*-D-*manno*-oct-3-ulose (**23**) by self-aldol condensation of D-erythrose (**22**) (*Fig. 6*).



D-*glycero* -**D**-*gulo* -
1,7-anhydroheptopyranose



D-*glycero* -**D**-*ido* -
1,7-anhydroheptopyranose

Figure 5
Examples of higher anomeric anhydrosugars

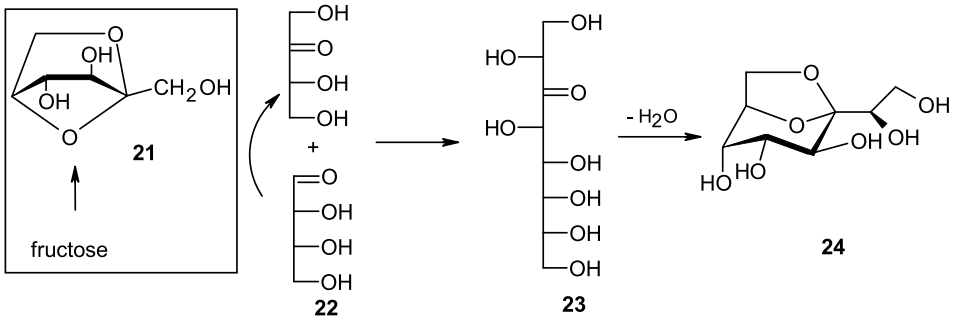
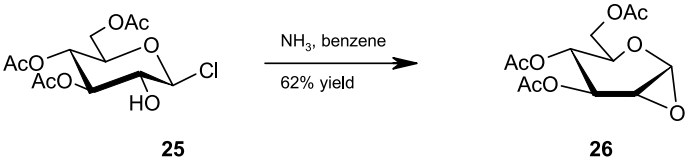


Figure 6
Selected examples of anhydroketoses

2.3 1,2-Anhydrosugars

1,2-Anhydrosugars are an important class of compounds in which the three-membered oxirane ring is connected to the anomeric and the C-2 atoms of a sugar. The first compound of this class, the so-called Brigl's anhydride (**26**), was prepared by treatment of β-chloro-3,4,6-tri-*O*-acetyl-D-glucose (**25**) with ammonia (*Scheme 11*) [3].



Scheme 11

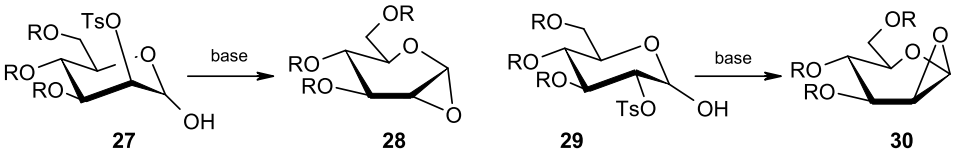
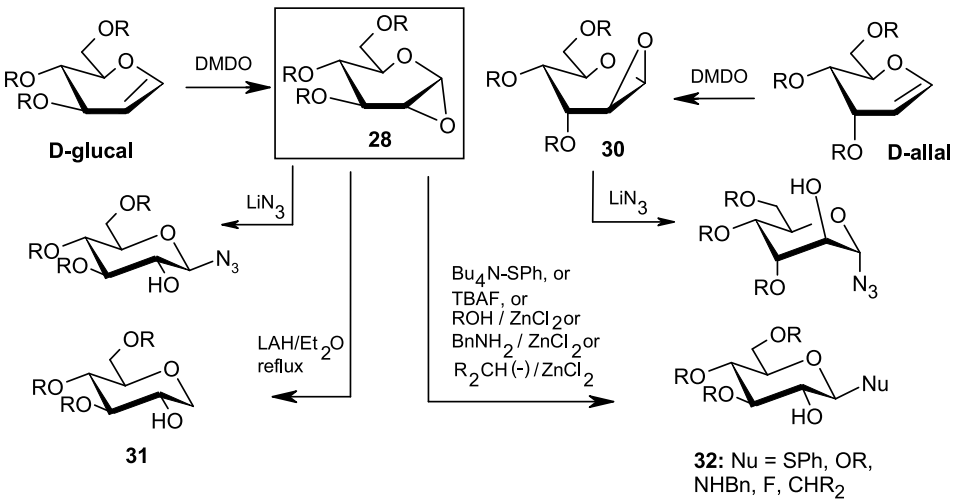


Figure 7
Synthesis of 1,2-anhydrosugars via activation of the 2-OH group

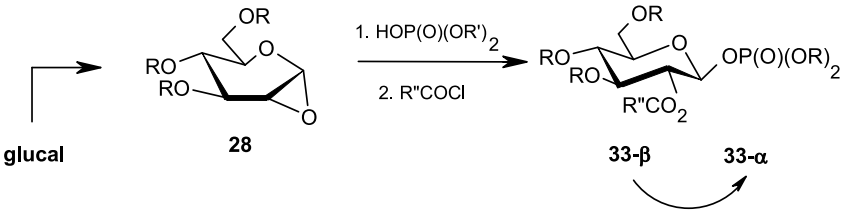


Scheme 12

There are three main routes leading to 1,2-anhydrosugars. The first one (similar to Brigl's original preparation) involves the attack of the hydroxyl group from the C-2 position on the anomeric center activated with a leaving group; both groups have to be in the *trans* arrangement. The second route is just the opposite: activated hydroxyl at the C-2 position (e. g. OTs) is attacked by the free anomeric hydroxyl with the inversion of the configuration at the C-2 stereogenic center [2,3] (*Fig. 7*).

Treatment of the 2-OTs D-mannose derivative **27** with a base affords the 1,2-anhydro-glucose derivative **28**, while a similar reaction of the 2-OTs glucopyranose **29** yields the 1,2-anhydro-D-mannopyranose **30**.

The third method of synthesis of such epoxides, proposed by Danishefsky [24], involves a direct epoxidation of glycals with dimethyldioxirane (DMDO). The stereoselectivity of this reaction depends on the configuration at the C-3 position. If D-allal is used instead of the D-glucal the opposite epoxide is formed (*Scheme 12*) [25]. Opening of the epoxide with azide (*Scheme 12*) or oxygen nucleophiles proceeds with the inversion of the configuration at the anomeric center [25,26]. When such 1,2-epoxides are treated with a 'hydride' reagent smooth reduction is observed, which provides 1-deoxy-sugars in good yield (e. g. compound **31**) [27].



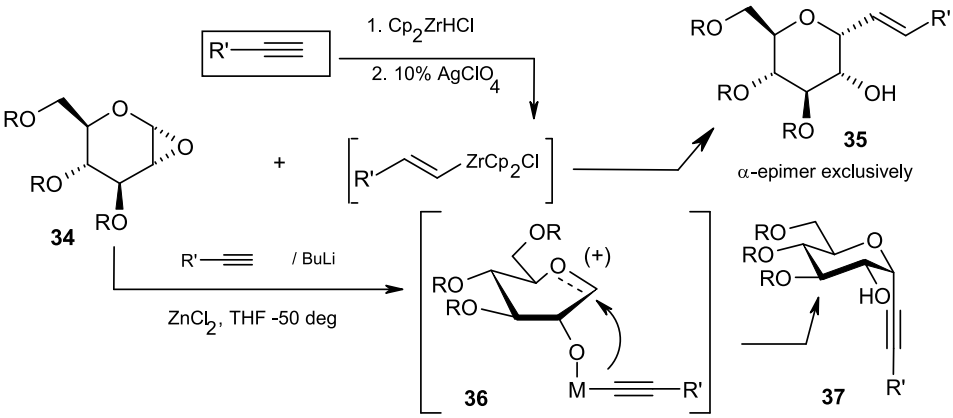
Scheme 13

The 1,2-anhydrosugars are convenient synthons for the preparation of a variety of important derivatives. Generally, opening of the oxirane ring with various nucleophiles (alcoholates, amines, sulfides, carboanions, etc.) proceeds with the inversion of the configuration at the anomeric center [28,29,30].

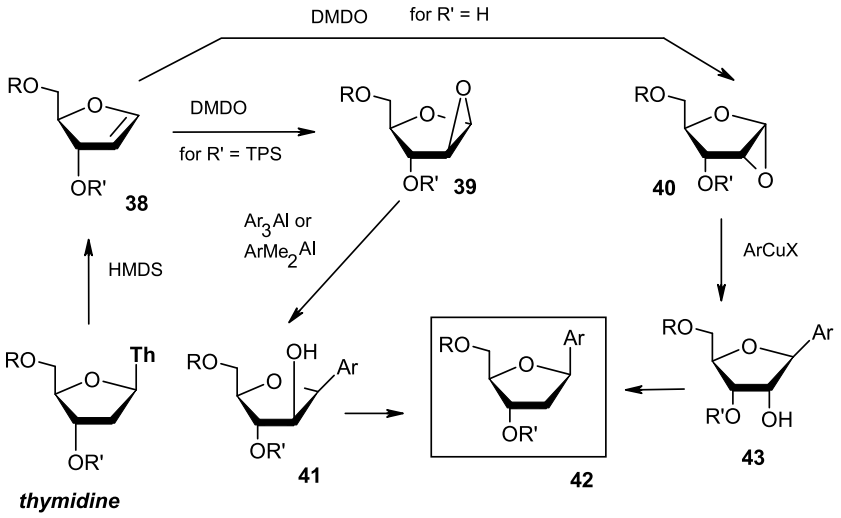
Glycosyl phosphates, building blocks in automated solid-phase synthesis of complex oligosaccharides, were prepared via a 1,2-epoxide (*Scheme 13*). Oxidation of the appropriately substituted D-glucal **28** with DMDO followed by reaction with dialkyl phosphate provided the β-phosphate **33**, which anomerized further to the α-analog [31].

Reaction of 1,2-anhydro derivatives such as **34** with acetylene organometallics provides acetylenic *C*-glycoside **37** with the *retention* of the configuration at the anomeric center [29,32]. This can be explained by a complexation of the organometallic to the oxirane oxygen atom with a cleavage of the C1–O bond (**36**) and subsequent attack of the electrophile from the same (α) side. An interesting variation of this procedure with zirconium organometallics was proposed recently by Wipf [33]. This reaction provides also the α-anomers (**35**) substituted at the anomeric center with the *E*-olefin (*Scheme 14*).

Diastereoselective synthesis of aryl *C*-glycosides of furanoses was realized in two ways. Oxidation of 1,2-unsaturated furanose **38** (obtained from thymidine according to the Danishefsky procedure) provides either D-*ribo*- (when the 2-OH is unprotected) or D-*arabino*- (the OH is protected with a bulky substituent) epoxides (**40** and **39**, respectively). The *trans* opening



Scheme 14



Scheme 15

of the oxirane ring [34] afforded the D-*ribo*-*C*-glycoside **43**, while the *cis*-opening with aluminum compounds provided the D-*arabino*-*C*-glycoside **41**. Both were further converted into the same 2-deoxy derivative **42** (*Scheme 15*) [35].

2.4 1,3- and 1,4-Anhydrosugars

Such derivatives are obtained by reaction of partially protected glycosyl chlorides with a strong base (NaH, ^{t}BuOK) in THF. The alternative version consists of treatment of the 4-*O*- or 5-*O*-

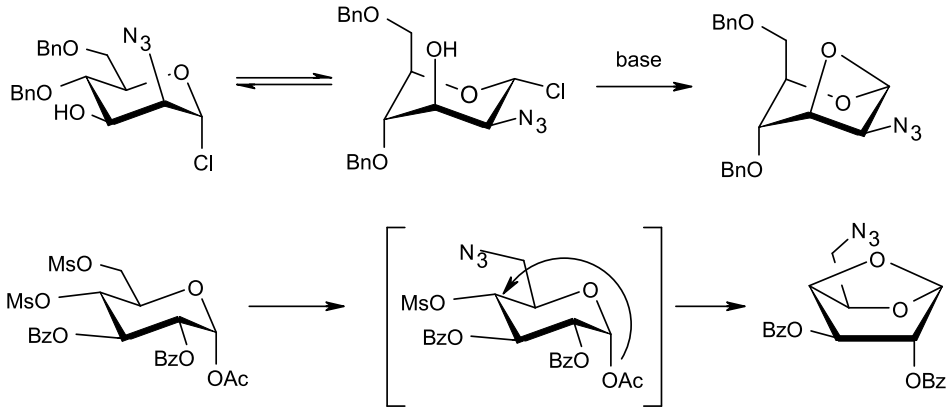


Figure 8
Examples of the 1,3- and 1,4-anhydropyranoses

sulfonylated partially protected sugars with a base or azide anion (NaN_3). Examples of the preparation of 1,3-anhydropyranose and 1,4-anhydropyranose are shown in ❯ *Fig. 8* [3]. Substituted 1,3-anhydroglucopyranoses, in the presence of the acidic catalysts, undergo regioselective ring-opening polymerization providing (1→3)-*β*-D-glucopyranans. The 1,4-anhydrosugars are stable in basic media, but are readily hydrolyzed with acid. 1,4-Anhydro-galactose is so unstable that it is already decomposed on silica gel.

3 Non-anomeric Anhydrosugars: Synthesis and Reactions

This group of compounds is characterized by the presence of a free (or protected as a *intermolecular* glycoside) anomeric position. The anhydro function may be created between various (except anomeric) carbon atoms, thus a large number of such derivatives are possible. The most convenient method of classification of such anhydrosugars is based on the size of the anhydro ring: sugar oxiranes, oxetanes, THF and THP derivatives are known. The latter are rather rare, however, interest in them has increased recently. Discussing the anhydrosugars one has to consider also two different positions of the anhydro ring, namely the *exo*- and *endo*-cyclic rings (❯ *Fig. 9*).

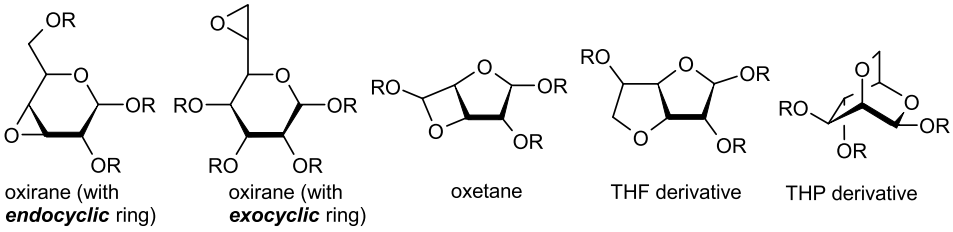
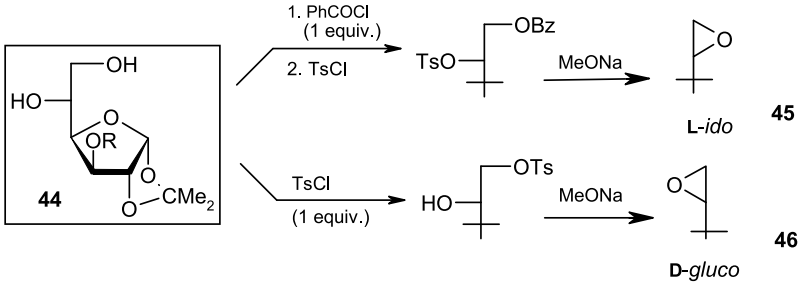


Figure 9
Different types of anhydrosugars with an anomeric carbon atom not involved in the anhydro ring

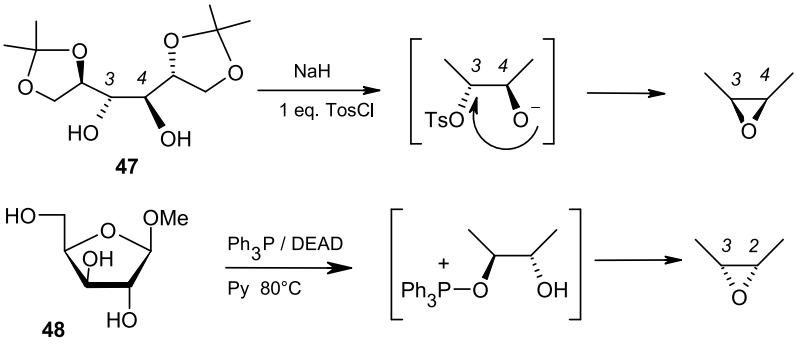
3.1 Sugar Oxiranes

Because of the well-pronounced differentiation of the reactivity of the hydroxyl groups in the sugar molecule, it is possible to prepare anhydrosugars differing in configuration from the same precursor. This strategy may be exemplified by preparation of either 5,6-anhydro-*α*-D-*gluco*- or 5,6-anhydro-*β*-L-*ido*-hexofuranoses (**46** and **45**, respectively) from 3-*O*-benzyl-1,2-*O*-isopropylidene-*α*-D-glucofuranose (**44**). Activation of the primary hydroxyl group (by selective tosylation) followed by cyclization in basic media affords the L-*ido*-isomer **46**. Alternatively, protection of the most reactive group (6-OH) as benzoate followed by activation of the secondary one at the C5 position affords the 6-*O*-benzoyl-5-*O*-tosyl derivative; in basic media hydrolysis of the benzoate occurs readily and the anion generated at the oxygen atom from the C-6 position attacks the C-5 center with the inversion of the configuration providing the D-*gluco*-isomer **45** (❯ *Scheme 16*) [1].

A synthetically useful procedure for the one-pot conversion [1] of vicinal diols into epoxides involves selective mono-activation of one hydroxyl group (by reaction with 1 equiv. of tosyl



Scheme 16



Scheme 17

chloride) followed by the attack of a second hydroxyl group with the closure of the oxirane ring. Thus, treatment of the diol (e. g. diacetonomannitol **47**) with two equivalents of sodium hydride generates the di-anion which reacts further with one equivalent of toluene-*p*-sulfonyl chloride. The intermediate thus formed, undergoes intramolecular S_N2 reaction leading to the appropriate oxirane (*Scheme 16*). The activation can be performed also under much milder Mitsunobu conditions (e. g. compound **48** in *Scheme 17*) [36].

Synthesis of the smallest heterocyclic ring (the oxirane) is achieved conveniently also by other methods. The oxidation of a carbon–carbon double bond with hydrogen peroxide (alone or with organic nitriles: MeCN or PhCN), peracids (e. g. *m*-chloroperbenzoic), or organic peroxides\transition-metal catalysts [37] is applied to form sugar epoxides. Alternatively, transformation of a carbon–oxygen double bond into epoxides could be done by reaction of sugar aldehydes with: α-halogeno-acids (the Darzens' reaction), diazo-acids, or sulfonium ylides (*Fig. 10*) [1].

This methodology may be illustrated by the synthesis of exocyclic epoxide **50** (a key compound for the preparation of sugar–lysine chimeras), which was realized by reaction of lactone **49** with methyl bromoacetate (*Scheme 18*) [38].

The oxirane ring can be formed between C2-C3, or C3-C4, or other carbon atoms; the most common are 2,3-epoxides. Such compounds are important building blocks in the preparation of, for example, di- and oligosaccharides as well as modified monosaccharides. The sugars

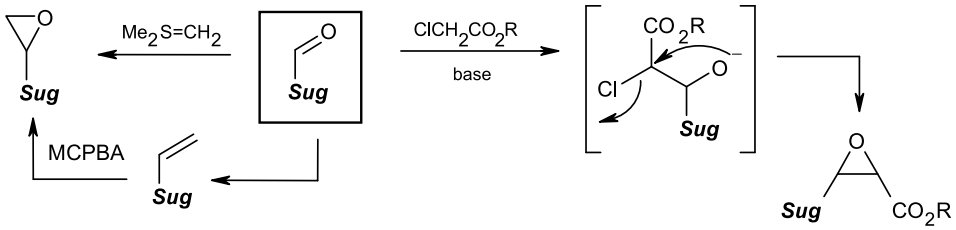
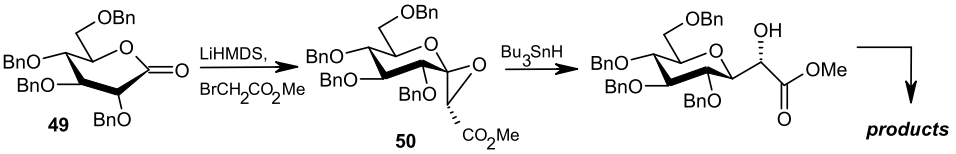


Figure 10
Methods for the preparation of sugar epoxides



Scheme 18

epoxides can be opened with a high degree of regiochemical control [39], which is governed by stereoelectronic factors. In some cases the regioselectivity of the opening of the oxirane ring may be inverted, if this process is done under chelating conditions. A detailed analysis of this process for several 6-deoxy-2,3- and 3,4-epoxysygars (prepared by the general methods already described within this chapter) was performed by Crotti [40]. The representative examples of this analysis are shown in *Fig. 11*. For the *cis*-relationship between the 3,4-epoxide ring and the methyl group at the C-6 position the opening of the ring was highly regioselective providing the C-3 product regardless of the reaction conditions (chelating or standard).
However, for the *trans*-relationship the regiochemistry of this process is slightly different. Opening of the 3,4-epoxide **51**-*trans* provides significant amounts of the C-4 product (*Fig. 11*).

3.1.1 2,3-Anhydrosugars

These compounds are formed either by a nucleophilic displacement of appropriate leaving groups by a hydroxyl group placed at the C-α-position or by direct epoxidation of the C2–C3 double bond. Synthesis of 2,3-anhydro-derivatives of α-D-*manno*- and α-D-allopyranosides (**54** and **53**, respectively) from methyl 4,6-*O*-benzylidene-α-D-glucopyranoside (**52**) is a good example of the first method. Activation of the more acidic 2-OH group in α-methyl glucoside **52** with 1 equiv. of tosyl chloride affords the 2-*O*-Tos derivative which, upon treatment with base (NaOMe), cyclizes readily to the *manno*-epoxide **54** via nucleophilic attack of the anion generated from the hydroxyl group at the C-3 position. When, however, ditosylate is used under the same conditions, the stereoisomeric *allo*-epoxide **53** is formed, which results from the preferential hydrolysis of the tosylate from C-2 and subsequent attack of the 2-O anion on the C-3 position (*Scheme 17*) [1,41]. Such sugar epoxides (e. g. **54**) were recently

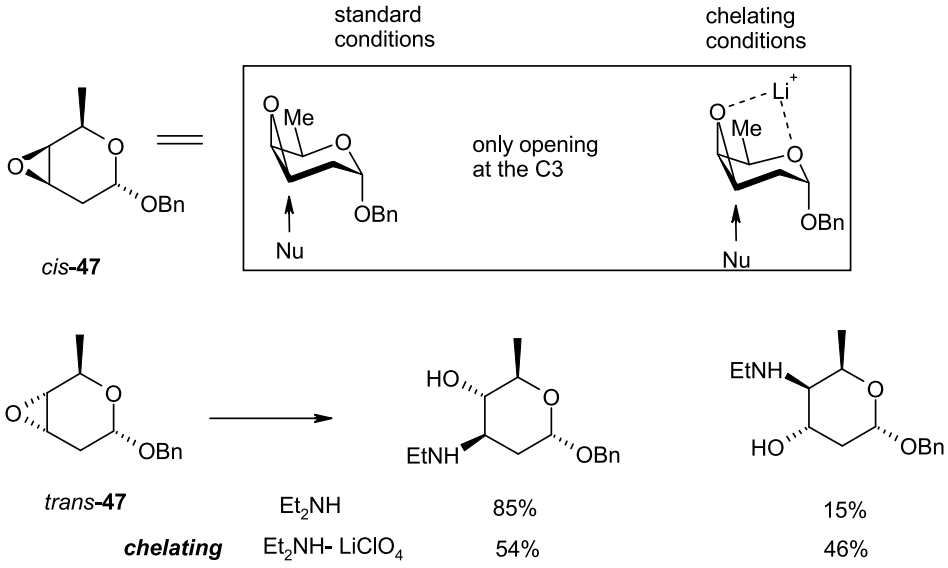
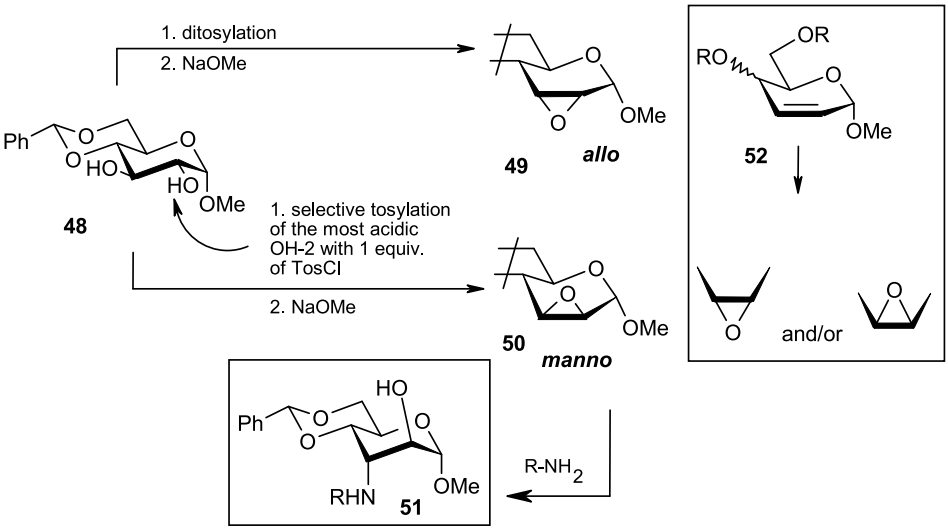


Figure 11
Examples of the opening of the epoxides with nucleophiles under chelating and standard conditions



Scheme 19

used for the preparation of aminosugars (e. g. **55**) (*Scheme 19*) [42]. Epoxidation reaction of the double bond in monosaccharides (e. g. **56**) is used mainly for the preparation of 2,3-anhydrosugars or saccharides with the exocyclic oxirane ring.

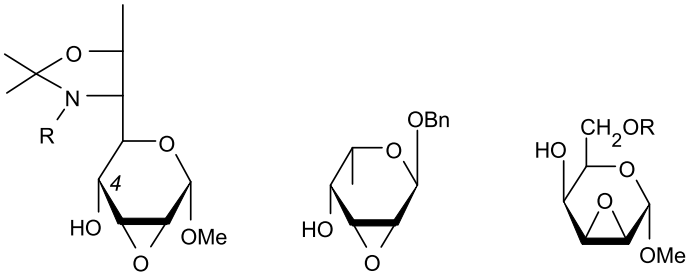
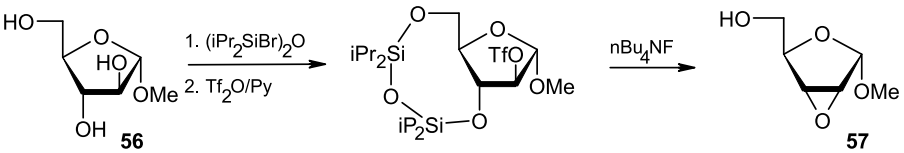


Figure 12
Examples of 2,3-anhydrosugars obtained by epoxidation of the parent unsaturated derivatives



Scheme 20

Stereoselectivity of the direct epoxidation of 2,3-unsaturated sugars with carboxylic or imino-peroxo acids (H_2O_2 + RCN) depends on the relative configuration of substituents at the adjacent centers, for example the configuration at C-1(OR) and C-4(OH). Since both such substituents may complex the epoxidizing agent by hydrogen bonding [43] they exhibit a directing effect in the oxidation process. Therefore, the highest selectivity of the epoxidation of 2,3-unsaturated sugars should be expected when both (oxy) groups at the C-1 and C-4 positions are in relative *cis* configuration. No wonder, therefore, that compound **57** was obtained as the only product by epoxidation of the corresponding unsaturated sugar, since both oxygen functions are in the *cis*-relation (*Fig. 12*) [1].

The effect of a free OH group is usually much more pronounced than that of the protected one. For example, oxiranes **58** and **59** were obtained with high stereoselectivity (directed by the free OH), although both oxy-substituents act in the opposite directions [1,44] (*Fig. 12*). When the epoxidation is performed not on allylic alcohols but allylic esters the selectivity of this process is low; both possible oxiranes are obtained in comparable amounts. The stereoselective formation of the 2–3-epoxides can be achieved also by the S_N2 process. Recently, Lowary proposed an efficient synthesis of 2,3-epoxy-*arabino*-furanoside **61** from the parent glycoside **60** in a sequence of reactions presented in *Scheme 20* [45]. Such anhydrosugars are convenient precursors for further functionalization at either the C-2 or C-3 position.

When the sulfur function is placed at the anomeric center (e. g. **62**), this molecule serves as a glycosyl donor in glycosylation reactions, thus it becomes a precursor of oligosaccharides (e. g. **63**). The epoxy ring located at the C2-C3 atoms may be properly functionalized, allowing us to prepare disaccharides of the desired stereochemistry (e. g. **64**, **65**; *Fig. 13*) [46,47,48]. By this methodology a number of di- and trisaccharides were obtained [49].

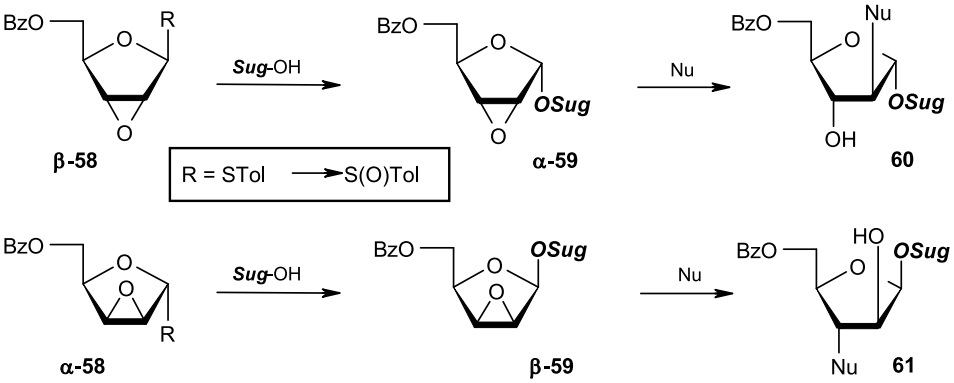


Figure 13
The concise methodology of the synthesis of di-saccharides from anhydrosugars

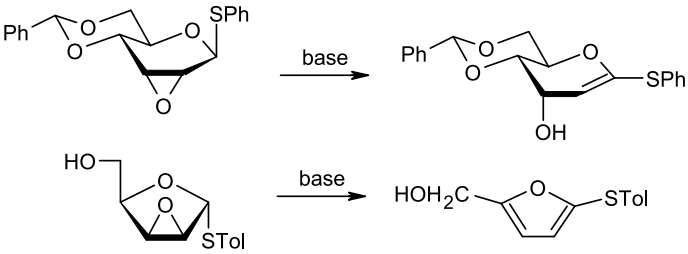
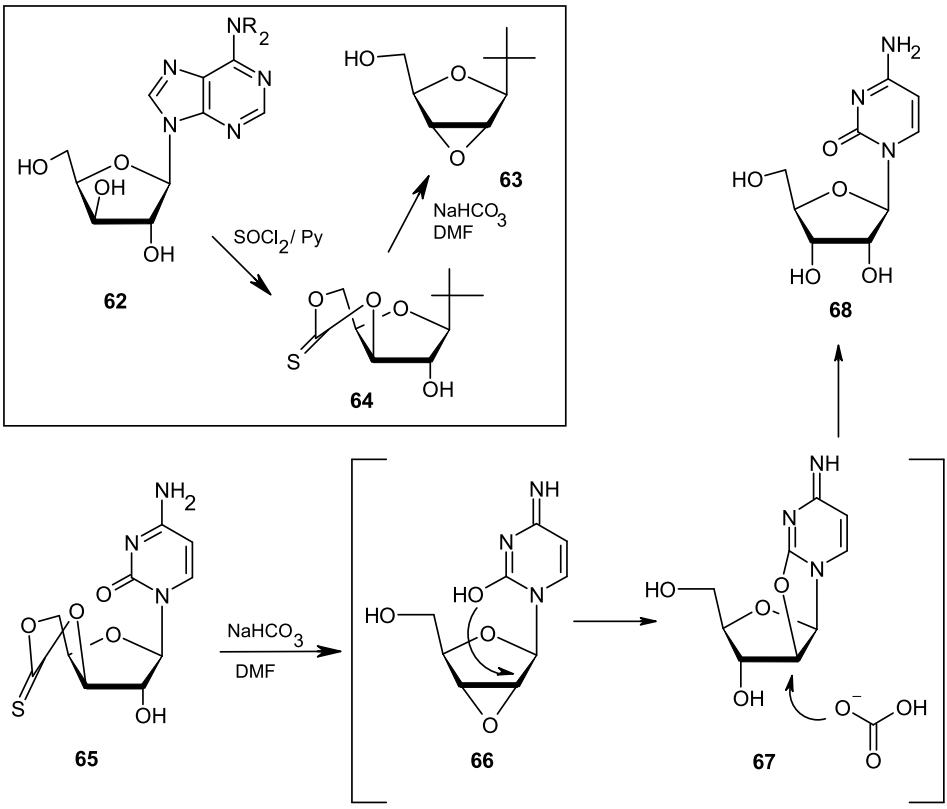


Figure 14
Behavior of 2,3-anhydro-thioglycosides under basic conditions

However, 2,3-anhydro-thioglycosides undergo—under basic conditions—other undesired reactions. If the anion is generated at the α-position to the anhydro ring, the β-elimination occurs readily providing the unsaturated carbohydrates [41]. Furanoses, under the same conditions undergo aromatization (*Fig. 14*)

The carbon atoms in cyclic sulfates and sulfites are highly reactive towards nucleophilic reagents [50], which allows us to use such functionalities as the epoxide substitutes in many reactions. The 3,5-sulfites having the neighboring free hydroxyl group were used for preparation of anhydronucleosides. Thus, D-*xylo*-furanoside nucleoside **66** reacted with $SOCl_2$ in pyridine to afford the cyclic sulfite **67**, which upon treatment with sodium bicarbonate furnished the 2,3-anhydro-nucleoside **68** (*Scheme 21*).

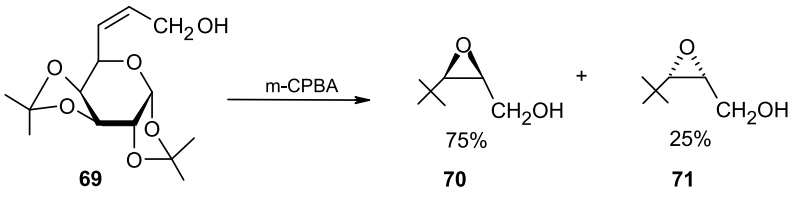
However, synthesis of 2′,3′-anhydro-cytidine **70** could not be realized in this way, since treatment of the corresponding cyclic sulfite **69** with bicarbonate afforded the D-*ribo*-derivative **72**. The initially formed epoxide **70** was opened with the oxygen atom from the heterocyclic moiety and this anhydro derivative **71** upon another ring-opening process with a bicarbonate anion provided **72** (*Scheme 21*) [51].



Scheme 21

3.1.2 Exocyclic Epoxides

These compounds are conveniently obtained by epoxidation of the exocyclic double bond in the sugar skeleton. Such oxidation with *m*-chloroperbenzoic acid (MCPBA) usually proceeds with moderate selectivity. Treatment of homologated galactose **73** with MCPBA afforded both epoxides **74** and **75** in a 3:1 ratio; these compounds are synthons for the preparation of antibiotic olguinine (*Scheme 22*) [1].



Scheme 22

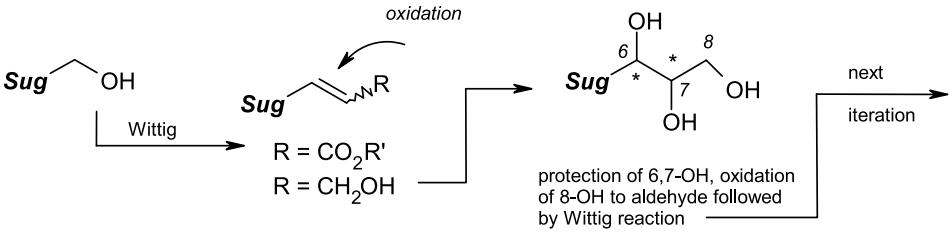
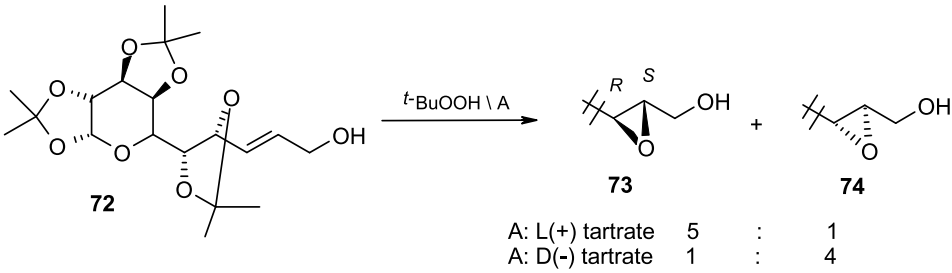


Figure 15
Brimacombe methodology of the synthesis of higher carbon sugars



Scheme 23

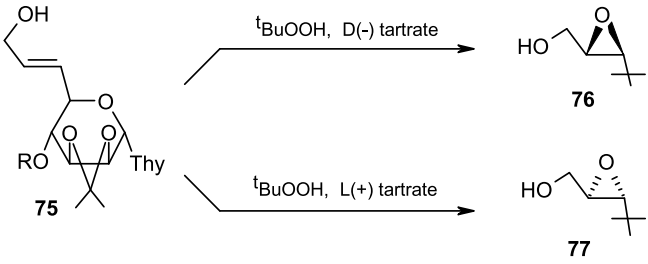
One of the first applications of carbohydrate derivatives with an exocyclic oxirane ring in stereocontrolled synthesis was proposed by Brimacombe in his synthesis of higher sugars [52]. The general idea was based on the elongation of the parent monosaccharide at the terminal position by two carbon atoms using the Wittig methodology followed by functionalization of the resulting double bond, which was achieved either by osmylation or epoxidation. By this iterative elongation Brimacombe was able to prepare decoses (*Fig. 15*.)

One of the final steps of this synthesis involved epoxidation of the unsaturated decose (e. g. **76**), which was performed under the Sharpless conditions. When L-tartrate was used as chiral catalyst the 8(*R*),9(*S*) epoxide **77** was obtained as the main product, while with D-tartrate the opposite isomer **78** was formed (*Scheme 23*) [52].

The asymmetric Sharpless epoxidation allowed us to obtain the epoxide of desired stereochemistry by use of the proper catalyst, however, the selectivity was not high. Much more selective was the epoxidation process of a precursor of higher sugar pyranosidic nucleosides **79**, which provided only epoxide **80** with (–)-DET, while (+)-DET afforded exclusively the opposite stereoisomer **81** (*Scheme 24*) [1].

3.1.3 Rearrangement of Sugar Epoxides

In compounds with the free hydroxyl group placed at the α-position with respect to the oxirane ring the interconversion between epoxides may be noted. The first such rearrangement was observed by Lake and Peat already in 1939 and later by Buchanan [1]. Treatment of methyl 2,3-di-*O*-benzyl-4-*O*-tosyl-6-*O*-trityl-α-D-glucopyranoside with alkali resulted in formation of methyl 3,4-anhydro-α-D-galactoside together with the 2,3-anhydro-D-guloside. When the



Scheme 24

hydroxyl group is placed at the β- or γ-positions with respect to the oxirane ring, formation of other anhydrosugars with a different size of heterocyclic ring might be observed [1,3]. The classical examples of such transformations are shown in *Fig. 16*.

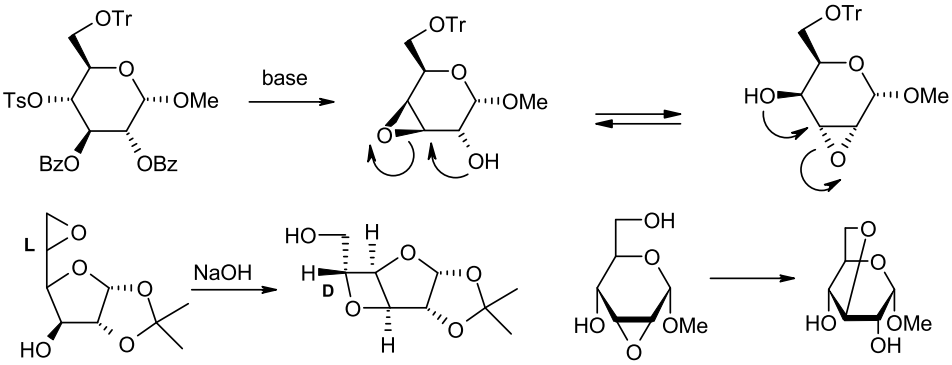


Figure 16
Selected examples of the rearrangement of sugar epoxides

3.2 Sugar Oxetanes, Oxolanes (THF), and THP Derivatives

These compounds are rather rare. One of the first examples was presented by Helferich who prepared 3,5-anhydro-1,2-*O*-isopropylidene-α-D-xylofuranose (**82**) from the appropriate 3,5-dimesyl derivative by treatment with ethanolic potassium hydroxide [1]. Another convenient approach to such derivatives utilized the reaction of 1,2-*O*-isopropylidene-5-*O*-triflate-α-D-glucofuranos-3,6-lactone (**79**) with base (K_2CO_3) in methanol, which produced the corresponding L-*ido*-anhydrosugar **80**. This was done by opening of the lactone ring with formation of the alcoholate at 3-OH, which substituted the triflate with inversion of the configuration at the C-5 atom (*Fig. 17*).

The general methodology for the preparation of anhydropyranoses containing tetrahydrofuran and tetrahydropyran anhydro rings is described in the leading book for preparative carbohydrate chemistry: "*Methods in Carbohydrate Chemistry*" [53]. 3,6-Anhydrofuranoses

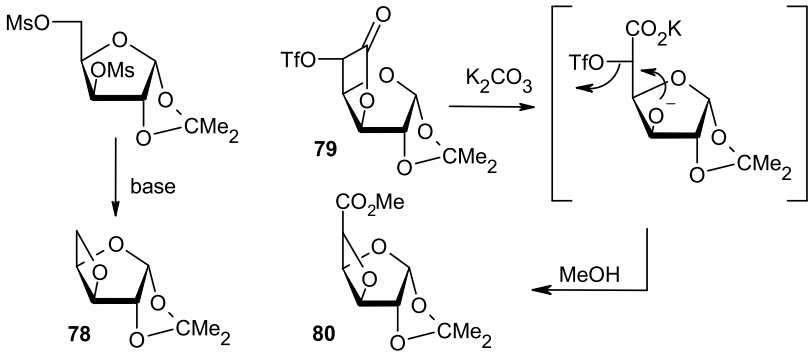


Figure 17
Formation of sugar oxetane derivatives

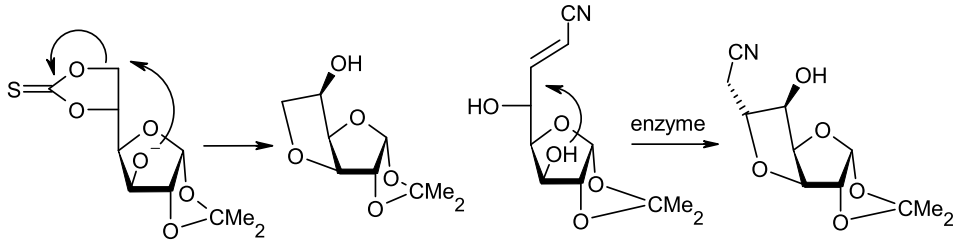


Figure 18
Selected examples of the preparation of oxolane anhydrosugars

are prepared similarly. The thiocarbonate, sulfate, or sulfite functionalities are used as leaving groups. The 1,4 addition of a sugar hydroxyl group to an activated electrophile is also used for the preparation of anhydrosugars. For example, treatment of 1,2-*O*-isopropylidene-α-D-glucofuranos-5,6-thiocarbonate with base caused the intramolecular attack of the 3-OH anion onto the C-6 atom with formation of 3,6-anhydro-glucofuranose.

No attack at C-5 leading to the oxetane (3,5-anhydro) has been noted [1]. Enzymatic cyclization of the α,β-unsaturated sugar nitrile provided the corresponding 3,6-anhydro-glucofuranose derivative in good yield [1] (*Fig. 18*).

The most important compounds from this class are undoubtedly 3,6-anhydrofuranoses. Furanodictine A and B (produced by cellular slime mold *Dictyostelium discoideum*) showing neuronal differentiation activity [54] are good examples. These interesting derivatives may be conveniently obtained from the open-chain sugars. For example, synthesis of furanodictine A was realized from compound **85** obtained from D-arabinose in a few well-defined steps as shown in *Fig. 19* [55].

The tin methodology is particularly useful in organic chemistry. The tin moiety activates the allylic fragments; besides, the stannyl unit can be replaced with metal cations (generally lithium), thus organotin derivatives can be regarded as stable precursors of carbanions [56]. A highly oxygenated heterocycle isolated from natural sources—goniofurone, representative

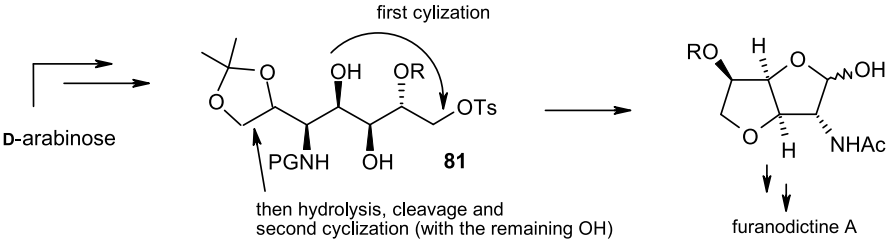
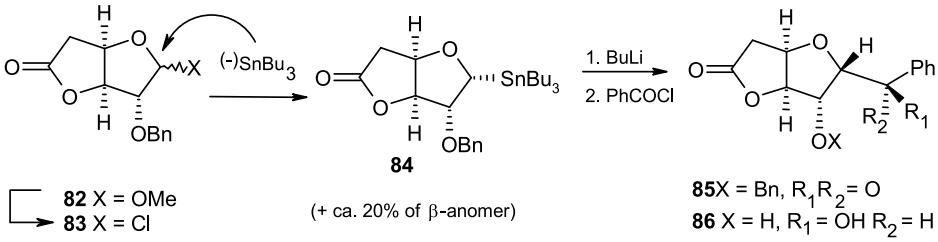


Figure 19
Synthesis of a representative oxolane anhydrosugar



Scheme 25

of 3,6-anhydrofuranoses—has been synthesized from the stannyl intermediate. Compound **86**, readily obtained in a few steps from D-glucurono-3,6-lactone, was converted into chloride **87**. The S_N2 displacement of the chlorine with the tin anion [$Bu_3Sn^{(-)}$ generated from $Bu_3SnSiMe_3$ by action of fluorides] afforded the stannyl derivative **88**, which reacted with benzoyl chloride to give ketone **89** finally transformed into **90** (*Scheme 25*) [57].

Among tetrahydropyran anhydrosugar derivatives the most common are 3,6-anhydro-pyranoses. An example illustrating the synthesis of such types of compounds is the preparation of methyl 3,6-anhydro-β-D-glucoside from methyl 2,3,4-tri-*O*-acetyl-6-bromo-6-deoxy-β-D-gluco-pyranoside upon treatment with barium hydroxide. 3,6-Ahydro-D-galactose and 3,6-ahydro-D-mannose are prepared in a similar way (*Fig. 20*).

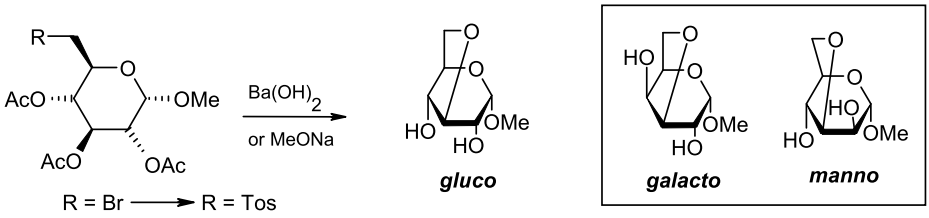


Figure 20
Examples of 3,6-anhydropyranoses

4 Anhydronucleosides

Several anhydronucleosides such as oxiranes **91** [58] or oxetanes **92** [59] (Fig. 21) inhibit HIV replication.

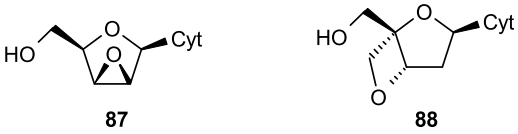


Figure 21
Examples of anhydronucleosides

They are also useful intermediates in the preparation of biologically important compounds [60] including unsaturated nucleosides. Several representatives of the latter compounds such as 3′-azidothymidine (AZT), 2′,3′-dideoxyinosine (ddI), 2′,3′-dideoxycytosine (ddC), and 2′,3′-dideoxy-2′,3′-didehydro-thymidine (d4T) (see Scheme 25) are approved as *anti*-HIV drugs by the US Food and Drug Administration [1].

Nucleosides with conformationally restricted carbohydrate rings (locked nucleic acids—LNA) are RNA mimics containing the 2′-*O*,4′-*C*-methylene linkage. The bicyclic structure locks

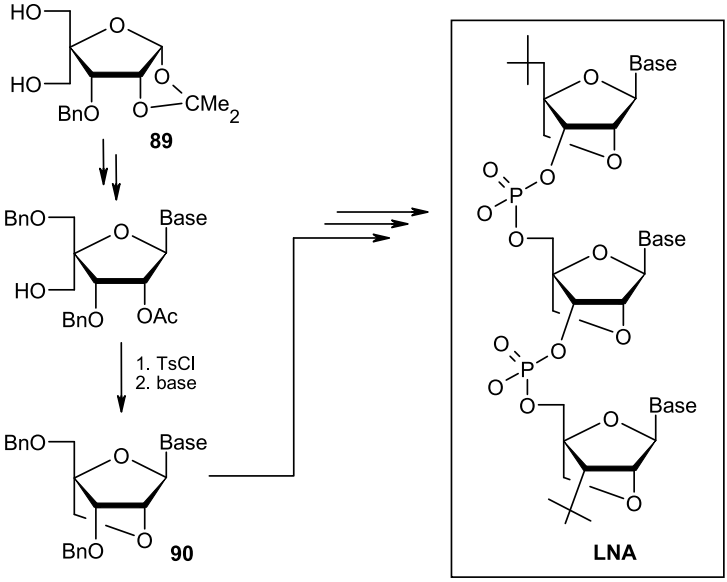


Figure 22
Synthesis of locked nucleic acids (LNA)

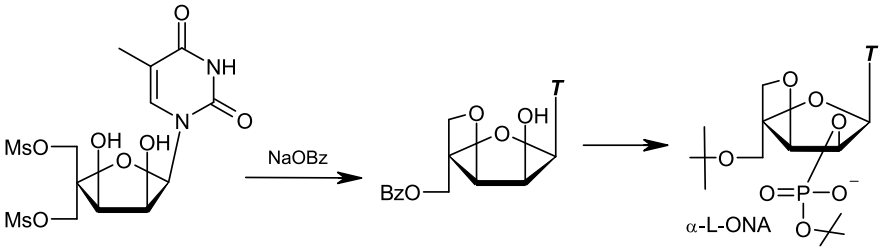


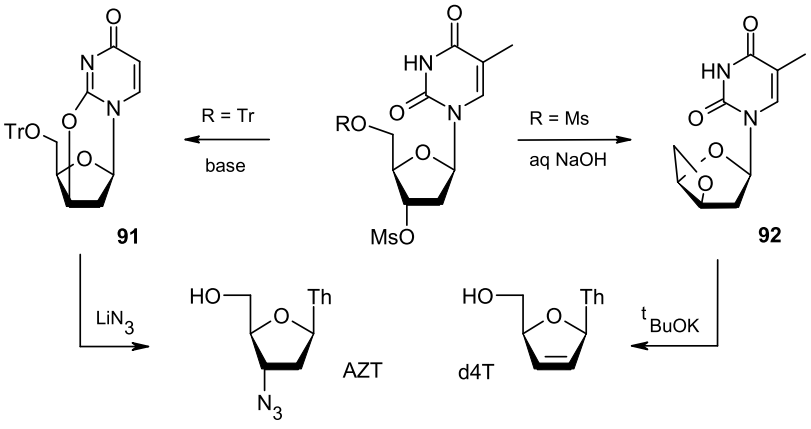
Figure 23
Synthesis of 3,4-anhydronucleosides

nucleosides in the 3′-endo conformation, favorable for thermal stability of A-type duplexes. The key-compound in their synthesis is oxetane **94**, being prepared by standard reactions from **93** (*Fig. 22*) [61].

Another example of monomers of this type, which have significant potential as building blocks for antigene and antisense strategies, was proposed recently by Wengel (*Fig. 23*) [62].

Strained carbohydrates such as 2,3- and 3,5-anhydrofuranoses are important compounds, since the heterocyclic oxirane and oxetane rings can be cleaved by various reagents, thus giving rise to either synthetically useful intermediates or products possessing biological activity [63]. Recently, the synthesis of such derivatives from partially protected furanosides (or furanoses) has been reported [36].

2′,3′-Anhydronucleosides are usually prepared from appropriate mono- or di-sulfonates [1]. 1-(2-Deoxy-3,5-anhydro)-*β*-D-*threo*-pentosylthymine (**96**), an example of the oxetane-type of anhydronucleoside, was prepared from dimesyl thymidine by the action of an alkali. The first synthesis of 2′,3′-dideoxy-2′,3′-didehydro-thymidine (d4T) was accomplished in 1966 by Horwitz et al. from this compound, which can be now prepared in quite large scale according to a slightly modified procedure proposed by the group from Bristol-Myers. The 5′-tritylated derivative of 3′-*O*-mesyl-thymidine was converted into 3′,2-anhydroucleoside **95**, from which



Scheme 26

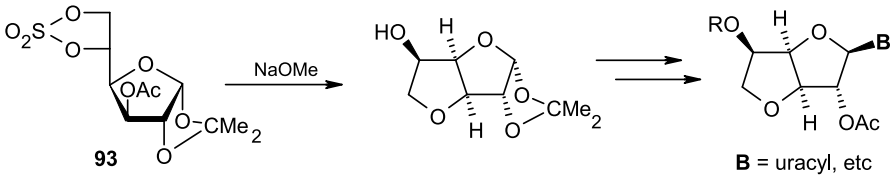
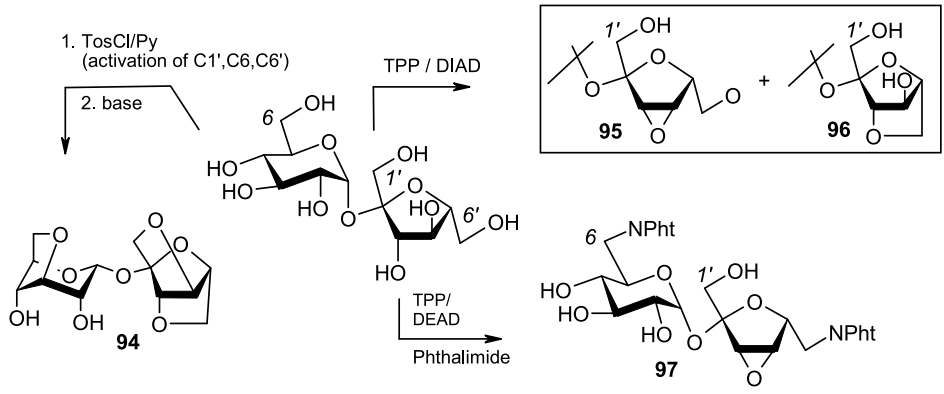


Figure 24
Synthesis of 3,6-anhydronucleosides

the 3-azido-thymidine (AZT) was prepared by an opening of the anhydro ring with lithium azide (*Scheme 26*) [1]. Both compounds are used in the treatment of AIDS.
Oxolane anhydronucleosides were prepared recently from cyclic 5,6-sulfate **97** derived from diacetonoglucose. Under the basic conditions the hydroxyl group at the C-3 position was liberated and the anion thus formed attacked the C-6 position of the sugar providing the target anhydro derivative (*Fig. 24*) [64].
This compound is a precursor of the corresponding anhydronucleosides.

5 Miscellaneous

Selected examples of the preparation of different types of anhydrosugars, which were not included in previous paragraphs, are presented in this Section. Many examples of such types of compounds are noted in the chemistry of sucrose. Sucrose is the most common disaccharide occurring in nature and 150 mln tons per year is produced. Intensive work is carried out to utilize this sugar also in markets other than food. The first example of anhydrosucrose was presented by Khan, who obtained the trianhydro-derivative **98** by treatment of tri-*O*-tosyl-sucrose with base [65]. Anhydro derivatives of sucrose are often obtained during Mitsunobu-type reactions [66,67], which are aimed at the synthesis of, for example, sucrose fatty esters. The esters at the terminal positions C-6 and C-6′ may be conveniently prepared under the Mitsunobu conditions (DIAD, TPP), when rather reactive acids are used [68]. However, when the acids are not reactive enough (or not present [67,69]), intramolecular etherification can compete with the desired intermolecular esterification, leading to anhydro-derivatives either at positions 3′,4′ or 3′,6′ (**99** and **100**, respectively) [70]. Treatment of free sucrose with phthalimide under the Mitsunobu conditions affords modified derivatives in which the primary 6-OH and 6′-OH groups are replaced with phthalimide while the secondary ones at the C-3′ and C-4′-positions were converted into the epoxide (**101**, *Scheme 27*) [71].
Reaction of the artificial sweetener sucralose (being 650 times sweeter than sucrose) with triphenylphosphine and diethyl azodicarboxylate afforded epoxide **102** from which the tetrachloro-derivative **103** was obtained (*Fig. 25*) [72].
Triflation at *O*-2 in hepta-*O*-acetylsucrose **104**, followed by S_N2 displacement with amines led to the *C*-2 epimer of its deoxyamino analog, as well as 2,3-epoxide (**105**) and its ring-opening products (*Scheme 28*) [73].
Sugar stannanes, convenient intermediates for the preparation of interesting enantiomerically pure compounds, are stable precursors of highly reactive carbanions. Such organostannanes may be easily obtained from anhydrosugars, as shown for example in *Fig. 26* [53].



Scheme 27

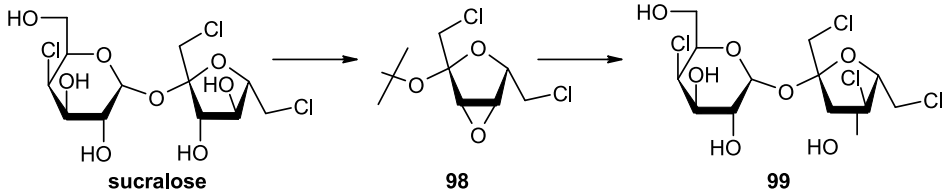
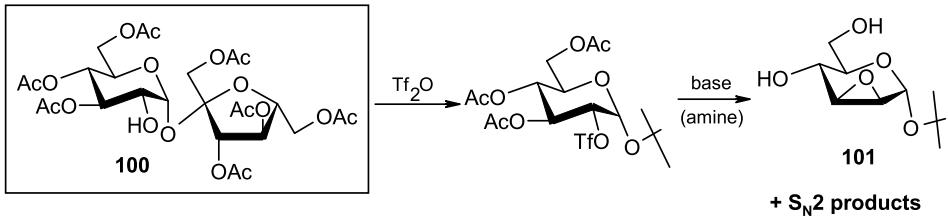


Figure 25
Selective transformation of sucralose



Scheme 28

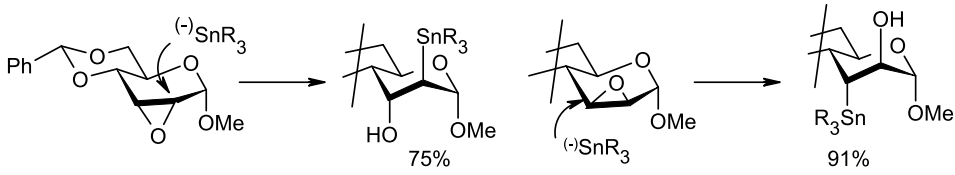
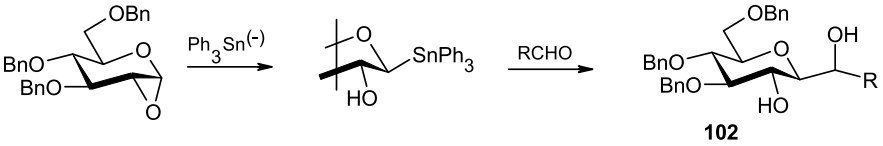
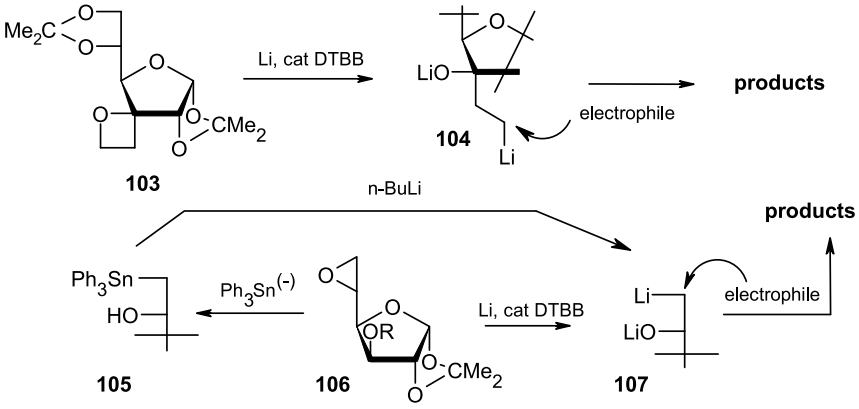


Figure 26
Opening of the ring of sugar epoxides with tin nucleophiles



Scheme 29



Scheme 30

If the 1,2-epoxide is used, the stannyl derivative obtained in this way may be applied in the synthesis of *C*-glycosides (e. g. **106**) with the retention of the configuration at the anomeric center (*Scheme 29*) [53].

Reductive lithiation [74] of epoxides (**110**) or oxiranes (**107**) is a good method for the preparation of highly reactive carbohydrate anions (**111** or **108**). Such species may also be obtained in a two-step procedure involving the opening of the epoxide ring with a tin nucleophile (to **109**), which is replaced finally with lithium (*Scheme 30*) [75].

Sugars with the 3-membered anhydro ring were used as convenient intermediates in the synthesis of biologically important compounds such as azasugars, which represent an important class of transition-state analogue inhibitors of glycosidases and glycosyl transferases. Compernolle applied sugar aziridines (e. g. **112**) for the preparation of heterocycles with a 6-membered ring [76], while sugar oxiranes (e. g. **113**) were used as substrates for the septanose-type azasugars [77] (*Fig. 27*).

An interesting approach to azasugars (e. g. **116** in *Scheme 31*) from the epoxsides built at the anomeric centers of unprotected carbohydrates was presented recently [78]. Such epoxides (e. g. **115**) were obtained with very high selectivity by reaction of the corresponding hemiacetals (e. g. **114**) with sulfur ylid (*Scheme 31*). This reaction is applicable for various sugar aldehydes [79].

The cascade reaction of symmetrical bis-epoxide **117** with azide provided the cyclopentane derivative **119** which was applied in the preparation of peptide-based drugs (*Scheme 32*) [80].

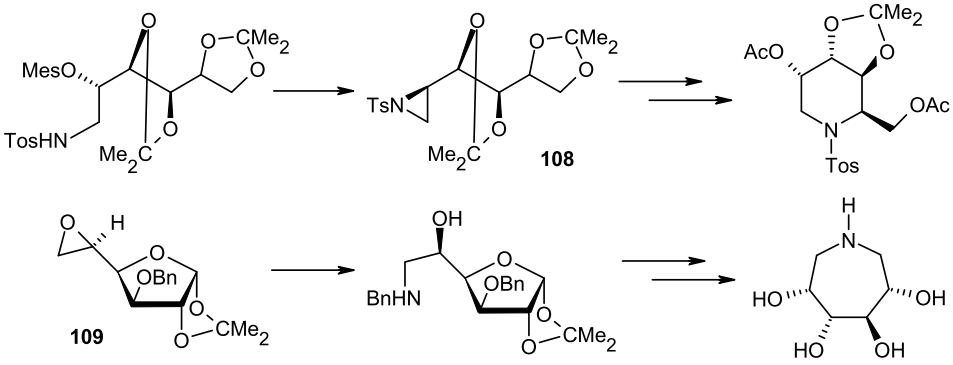
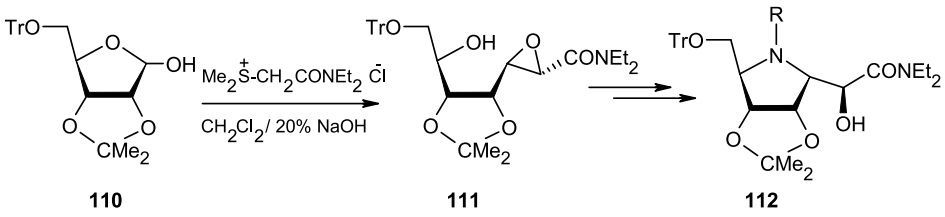
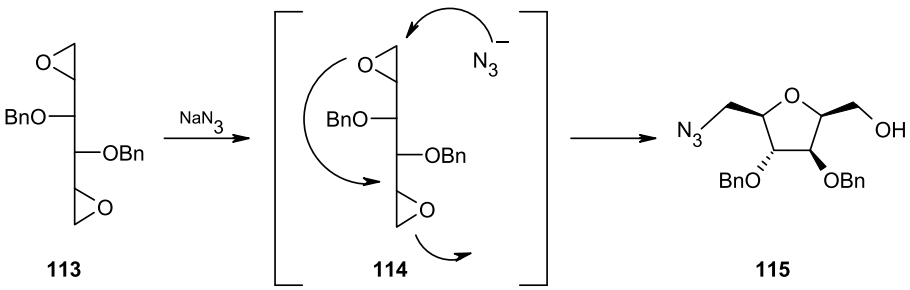


Figure 27
Examples of the synthesis of azasugars from anhydrosugar intermediates



Scheme 31



Scheme 32

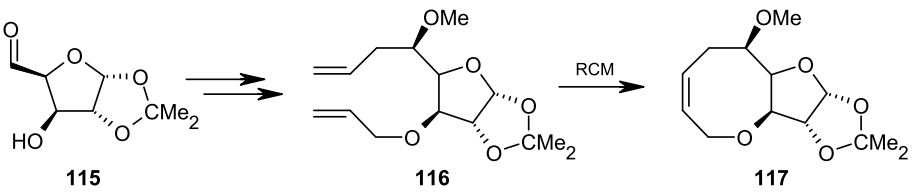


Figure 28
Methodology of the synthesis of anhydrosugars via the RCM process

Application of the ring-closing olefin metathesis reaction (RCM) for the preparation of carbocyclic rings from the appropriately functionalized sugars is now well documented [81]. Recently it was applied for the synthesis of compound **122**, which can be regarded as a (n,m)-anhydrosugar [81]. The intermediate diolefin **121** was prepared in a few standard steps from the corresponding dialdose **120** (*Fig. 28*) [82].

6 Conclusion

This chapter briefly describes different classes of anhydrosugars, both anomeric and non-anomeric. Attention is paid mostly to anhydrosugars of great importance such as 1,6- and 1,2-anhydrosugars, and non-anomeric sugar epoxides. Several anhydrosugars (mostly the modified ones with the nucleoside structure) show interesting biological activity, the others are used as optically pure synthons in stereocontrolled organic synthesis. The material presented in this chapter comprises the basic chemistry of such derivatives, which is already described in many monographs and reviews (including the previous chapters from this series) and provides new applications of such important derivatives. The new data (updating previously published work) are not comprehensive, however, they illustrate the important aspects of the chemistry in which anhydrosugars are engaged. Therefore, only selected examples from the new literature are included here. This material should give the reader a general impression of the chemistry of such compounds and elucidate the modern aspects of anhydrosugar chemistry.

References

1. Jarosz S (2001) Glycoscience—Chemistry and Chemical Biology 1:291
2. Stick RV, Williams SJ (2001) Glycoscience—Chemistry and Chemical Biology 1:627
3. Cerny M (2003) Adv Carbohydr Chem Biochem 58:121
4. Ball DH, Parish FW (1969) Adv Carbohydr Chem Biochem 24:139
5. Williams NR (1970) Adv Carbohydr Chem Biochem 25:109
6. Cerny M, Stanek J Jr (1977) Adv Carbohydr Chem Biochem 34:23
7. Ferrier RJ, Middleton S (1993) Chem Rev 93:2779
8. Hanessian S (1983) Total Synthesis of Natural Products: The Chiron Approach. Pergamon Press, New York; Inch TD (1984) Tetrahedron 40:3161; Fraser-Reid B, Tsang R (1989) Strategies and Tactics in Organic Synthesis. Academic Press, New York
9. Mitsunobu O (1981) Synthesis 1
10. Castro BR (1983) Org React 29:1
11. Zottola MA, Alonso R, Vite GD, Fraser-Reid B (1989) J Org Chem 54:6123
12. Bailliez V, de Figuiredo RM, Olesker A, Cleophax J (2003) Synthesis 1015
13. Dromowicz M, Köll P (1998) Carbohydr Res 308:169
14. Wiggins LF (1963) Methods Carbohydr Chem 2:188; Whistler RL, Lake WC (1972) Methods Carbohydr Chem 6:286
15. Lundt I, Skelbaek-Pedrsen B Acta Chim Scand Ser B (1981) 35:637; Skelto B, Stick RV, Matthew D (2000) Austr J Chem 53:389; Sridhar P, Saravanan V, Chandrasekarn S (2005) Pure Appl Chem 77:145
16. Sivapriya K, Chandrasekaran S (2006) Carbohydr Res 341:2204
17. Fuenets J, Olano D, Pradera MA Tetrahedron Lett (1999) 40:4063
18. Krohn K, Gehle D, Flörke U (2005) Eur J Org Chem 2841
19. Krohn K, Gehle D, Flörke U (2005) Eur J Org Chem 4557
20. Witczak ZJ, Kaplon P, Kolodziej M (2002) J Carbohydr Chem 21:143
21. Kelly DR, Mahdi JG (2002) Tetrahedron Lett 43:511

22. Manna S, McAnalley BH, Ammon HL Carbohydr Res (1993) 243:11
23. Goursaud F, Peyrane F, Veyrieres A (2002) Tetrahedron 58:3629
24. Halcomb RL, Danishefsky SJ (1989) J Am Chem Soc 111:6661; Danishefsky SJ, Bilodeu MT (1996) Angew Chem Int Ed Engl 35:1380
25. Lee GS, Min HK, Chung BY (1999) Tetrahedron Lett 40:543
26. Timmers CM, van Starten NCR, van der Marel GA, van Boom JH J Carbohydr Chem (1998) 17:471
27. Flaherty TM, Gervay J (1996) Tetrahedron Lett 37:961
28. Du Y, Linhard RJ (1998) Tetrahedron 54:9913
29. Allwein SP, Cox JM, Howard BE, Johnson HWB, Rainier JD (2002) Tetrahedron 58:1997
30. Jacobsson M, Malmberg J, Ellervik U (2006) Carbohydr Res 341:1266
31. Seeberger PH, Haase W-CH (2000) Chem Rev 100:4349; Plante OJ, Palmacci ER, Andrade RB, Seeberger PH (2001) J Am Chem Soc 123:9545
32. Leeuwenburgh MA, van der Marel GA, Overkleeft HS, van Boom JH (2003) J Carbohydr Chem 22:549; Riseeuw DP, Grotenberg GM, Witte MD, Tuin AW, Leeuwenburgh MA, van der Marel GA, Overkleeft HS, Overhand M Eur (2006) J Org Chem 3877
33. Wipf P, Pierce JG, Zhuang N (2005) Org Lett 7:483
34. Chow K, Danishefsky S (1990) J Org Chem 55:4211
35. Singh I, Seitz O (2006) Org Lett 8:4319
36. Schultze O, Voss J, Adiwidjaja G (2001) Synthesis 229
37. Sharpless KB, Behrens CH, Katsuki T, Lee AWM, Martin VS, Takatani M, Viti SM, Walker FJ, Woodard SS (1983) Pure Appl Chem 55:589; Katsuki T, Martin VS (1996) Org Reactions 48:1
38. Zhang K, Wang J, Sun Z, Nguyen D.-H, Schweizer F (2007) Synlett 239; Knorr R, Treciak A, Bannwarth E, Gillesen D (1989) Tetrahedron Lett 30:1927
39. Zamojski A, Banaszek B, Grynkiewicz G (1982) Adv Carbohydr Chem Biochem 40:1
40. Crotti P, Di Bussolo V, Favero L, Macchia F, Pineschi M (2002) Tetrahedron 58:6069
41. Wang Y, Li Q, Cheng S, Wu Y, Guo D, Fan Q-H, Wang X, Zhang L-H, Ye X-S (2005) Org Lett 7:5577
42. Maxwell VL, Evinson EL, Emmerson DPG, Jenkins PR (2006) Org Biomol Chem 4:2724
43. Berti G (1973) Topics Stereochem 7:83
44. Coleman RS, Jones AB, Danishefsky SJ (1990) J Org Chem 55:2771; Horita K, Sakurai Y, Nagasawa M, Yonemitsu O (1997) Chem Pharm Bull 45:1558
45. Callam CS, Gadikota RR, Lowary TL Carbohydr Res (2001) 330:267; Bai Y, Lowary TL (2006) J Org Chem 71:9658
46. Gadikota RR, Callam CS, Wagner T, Del Fraino B, Lowary TL (2003) J Am Chem Soc 125:4155
47. Callam CS, Gadikota RR, Krein DM, Lowary TL (2003) J Am Chem Soc 125:13112
48. Tilekar JN, Lowary TL (2004) Carbohydr Res 339:2895
49. Cociorva OM, Lowary TL (2004) Tetrahedron 60:1481
50. Byun H-S, He L, Bittman R (2000) Tetrahedron 56:7051
51. Takatsuki K-I, Yamamoto M, Ohgushi S, Kohmoto S, Kishikawa K, Yamashita H (2004) Tetrahedron Lett 45:137
52. Brimacombe JS (1989) In: Atta-ur Rahman (ed) Studies in Natural Product Chemistry. Elsevier, Amsterdam, 4C:157
53. Lewis BA, Smith F, Stephen AM (1963) Methods Carbohydr Chem 2:172
54. Kikuchi H, Saito Y, Komiya J, Takaya Y, Honma S, Nahakata N, Ito A, Oshima Y (2001) J Org Chem 66:6982
55. Yoda H, Suzuki Y, Takabe K (2004) Tetrahedron Lett 45:1599
56. Jarosz S, Zamojski A (2003) Curr Org Chem 7:13
57. Ye J, Bhatt RK, Falck JR (1993) Tetrahedron Lett 34:8007
58. Webb TR, Mitsuya H, Broder S (1988) J Med Chem 31:1475
59. Counde O-Y, Kurz W, Eugui EM, Mc Roberts MJ, Verheyden JHP, Kurz LJ, Walker KAM (1992) Tetrahedron Lett 33:41
60. Huryn DM, Okabe M (1992) Chem Rev 92:1745
61. Rozners E (2006) Current Org Chem 10:675
62. Madsen AS, Hrdlicka PJ, Kumar TS, Wengel J (2006) Carbohydr Res 341:1398.
63. Unger FM, Christian R, Waldstätten P (1978) Carbohydr Res 67:257
64. Molas MP, Matheu MI, Castillon S, Isac-Garcia J, Hernandez-Mateo F, Calvo-Flores FG, Santoyo-Gonzalez F (1999) Tetrahedron 55:14649
65. Khan R (1972) Carbohydr Res 22:441
66. Descotes G, Mentech J, Veesler S (1989) Carbohydr Res 190:309.

67. Guthrie RD, Jenkins ID, Thang S, Yamasaki R (1983) Carbohydr Res 121:109; ibid. item. (1988) 176:306
68. Bottle S, Jenkins IA (1984) J Chem Soc Chem Commun 385; Abouhilale S, Greiner J, Riess JG (1991) Carbohydr Res 212:55; Baczco K, Nugier-Chauvin C, Banoub J, Thilbault P, Plusquellec D (1995) Carbohydr Res 269:79
69. Buchanan JG, Cummerson DA (1972) Carbohydr Res 21:293
70. Molinier V, Fitremann J, Bouchu A, Queneau Y (2004) Tetrahedron: Asymmetr 15:1753
71. Amariutei L, Descotes G, Kugel C, Maitre JP, Mentech J (1998) J Carbohydr Chem 7:21
72. Hough L (1991) In: Lichtenthaler FW (ed) Carbohydrates as Organic Raw Materials I. VCH, Weinheim, p 32
73. Lichtenthaler FW, Mondel S (1997) Carbohydr Res 303:293
74. Soler T, Bachki A, Falvello LR, Foubelo F, Yus M (2000) Tetrahedron: Asymmetr 11:493
75. Taylor OJ, Wardell JL (1988) Recl Trav Chem Pays-Bas 107:267; Cox PhJ, Doidge-Harrison SMSV, Howie RA, Nowell IW, Taylor OJ, Wardell JL (1989) J Chem Soc Perkin Trans 1 2017
76. Compernolle F, Joly GJ, Peeters K, Toppet S, Hoornaert GJ, Kilonda A, Babady-Bila (1997) Tetrahedron 53:12739; Pearson MSN, Mathe-Allainmat M, Ergeas V, Lebreton J (2005) Eur J Org Chem 2159
77. Tilekar JN, Patil NT, Jadhav HS, Dhavale DD (2003) Tetrahedron 59:1873
78. Pino-Gonzales MS, Assiego C (2005) Tetrahedron: Asymmetr 16:199
79. Lopez-Herrera FJ, Pino-Gonzalez MS, Sarabia Garcia FA, Heras Lopez, Ortega-Alcantara JJ, Pedraza-Cebrian MG (1996) Tetrahedron: Asymmetr 7:2065;
80. Poitout L, Le Merrer Y, Depezay JC (1995) Tetrahedron Lett 36:6887; Gruner SAW, Locardi E, Lohof E, Kessler H (2002) Chem Rev 102:491
81. Hansen FG, Bundgaard E, Mdsen R (2005) J Org Chem 70:10139
82. Kaliappan KP, Kumar N (2005) Tetrahedron 61:7461

2.5 C–C Bond Formation

*Yuguo Du**, *Qi Chen*, *Jun Liu*
The State Key Laboratory of Environmental Chemistry and Ecotoxicology, Research Center for Eco-Environmental Sciences, Chinese Academy of Sciences, Beijing 100085, China
duyuguo@rcees.ac.cn

Abstract

In this chapter, synthetic methodologies for the preparation of *C*-branched carbohydrates, and nucleosides will be summarized and discussed.

Keywords

Chain extension; Aldol reactions; Branched-chain sugars; Radical reactions

Abbreviations

AIBN azobisisobutyronitrile
CAN ceric(IV) ammonium nitrate
CAMC 3′-β-carbamoylmethylcytidine

In: *Glycoscience*. Fraser-Reid B, Tatsuta K, Thiem J (eds)
Chapter-DOI 10-1007/978-3-540-30429-6_7:

COT cyclooctatetraene
DLP dilauoryl peroxide
DMF dimethylformamide
HDA hetero-Diels–Alder
HMDS 1,1,1,3,3,3-hexamethyldisilazane
MBH Morita–Baylis–Hillman
MDA methyl diazoacetate
MMTr monomethoxytrityl
TBS *tert*-butyldimethylsilyl
TBAF tetra-*n*-butylammonium fluoride
TBAI tetra-*n*-butylammonium iodide
TBDPS tertbutyldiphenylsilyl
TEA triethylamine
TEDMS 2-(trimethylsilyl)ethynyl]dimethylsilyl
THF tetrahydrofuran
THP tetrahydropyran

1 Introduction

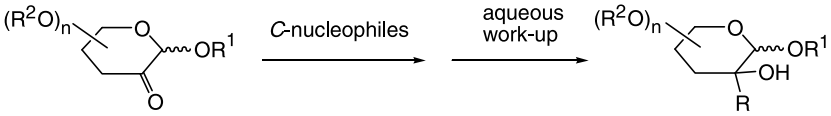
The carbohydrates represent a structurally diverse group of compounds, which are usually derivatized with a variety of functional group modifications. *C*-branched sugar derivatives, in which a new C–C bond was formed at the non-anomeric center, can provide a versatile compound pool for drug screening. Generally speaking, any chemistry method (traditional or new) applied to C–C bond formation could be useful in *C*-branching sugar preparation. Among these methods are intramolecular alkylation and intramolecular condensation of aldehyde with enolates, phosphonates, and nitro-stabilized anions. Metal-mediated radical reactions, cycloadditions, and rearrangements have also been applied frequently.

2 C–C Bond Formation by Means of Nucleophilic Additions

2.1 Addition of Organometallics to Carbonyl Sugar

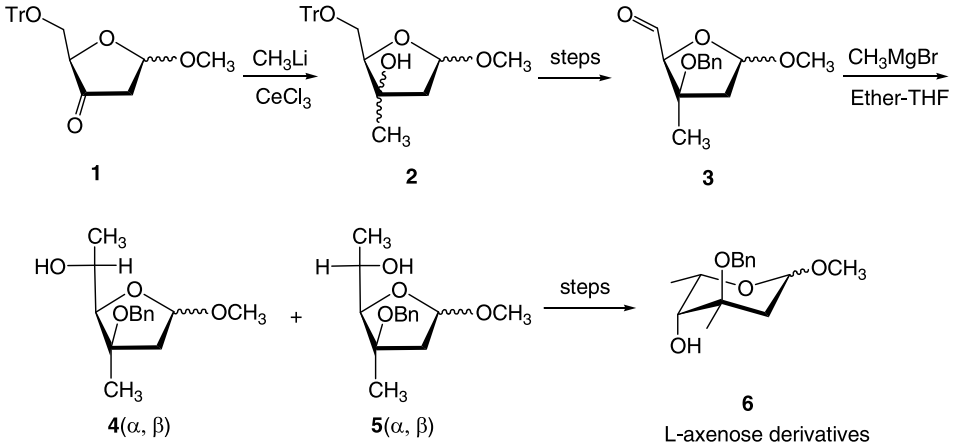
Organometallics, such as Grignard reagents and alkyllithium, are the most commonly used *C*-nucleophiles in C–C bond formation by addition to a carbonyl-functionalized sugar (❯ *Scheme 1*). Occasionally, organocopper [1] and alkylcerium reagents [2], prepared from alkyllithium and copper or cerium salts, are applied in this type of reaction. These nucleophilic additions, especially for Grignard reagents, are usually of high stereoselectivity due to the possible chelate-complexation between the metal ions and oxygenate group in the molecule [3]. The diastereofacial selectivity can be predicted and explained by different models [4]. The hydroxyl group generated in situ can also be further deoxylated or transformed into other functional groups.

For synthesizing the *C*-methyl-branched sugars [5] such as L-axenose, L-cladinose, L-arcannose, and L-vinelose, methylmagnesium bromide and methyllithium are often used to intro-



C-nucleophiles: RMgX, RLi, R_2CuLi, etc.

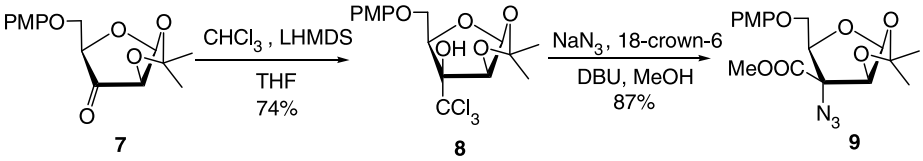
Scheme 1



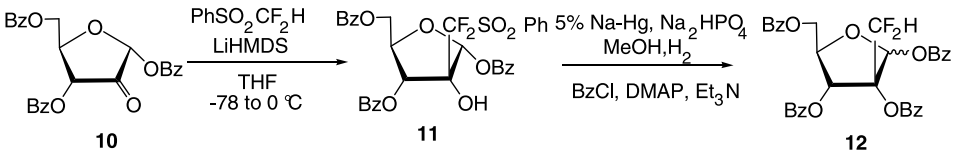
Scheme 2

duce a branching methyl group as well as a hydroxyl group with a certain configuration. Giuliano et al. [6] reacted 2-deoxy-D-glycero-pentofuranosid-3-ulose derivative **1** with methylcerium to give a 5:1 ratio of *threo* to *erythro* products **2** in the case of the α-anomer, and an exclusive *threo* product for the corresponding β-anomer. However, the same reaction gave unsaturated products using methyllithium as the nucleophile. When aldehyde **3** was treated with methylmagnesium bromide, a 1.3:1 ratio of products **4**β and **5**α favoring the D-arabino isomer **4**β in ether/tetrahydrofuran (THF) co-solvent were obtained. Equilibration of **5**β in methanolic hydrogen chloride gave pyranosides **6** as L-axenose derivatives (*Scheme 2*).

Nielsen et al. [7] reported a synthesis of the [3.2.0] bicyclic β-nucleoside mimicking the anti-HIV drug AZT. Ketone **7** reacted with trichloromethyllithium to give the alcohol **8** with absolute stereoselectivity in a reasonable yield. Following the modified Corey–Link reaction [8], **8** was converted into the α-azido methyl ester **9**, taking advantage of in situ formation of a dichloroepoxide, ring opening by the azide ion, and methanolysis of the acyl chloride intermediate. The reaction was absolutely stereoselective, and only one compound was obtained (*Scheme 3*). 2′-*C*-branched nucleosides have served as valuable probes to explore biomolecular structure and functions [9,10]. Piccirilli et al. [11] synthesized a series of ribonucleotides bearing substituents with increasing electron-withdrawing power (CH_3, CH_2F, CHF_2, CF_3) (*Scheme 4*). Addition of difluoromethyl phenyl sulfone to ketoribose **10** in the presence of lithium hexamethyldisilazane in THF/HMPA gave sulfone **11** in good yield.



Scheme 3



Scheme 4

^{19}F-^{1}H NOE experiments indicated that the phenylsulfonyl difluoromethyl group attached stereoselectively to the β-face of the sugar. Single-electron-transfer reduction of **11** assisted by Na-Hg-MeOH-Na_2HPO_4 under hydrogen pressure, followed by benzoylation with BzCl, afforded perbenzoylated difluoromethylribose **12**.

A number of nucleoside analogues, either used clinically as anticancer drugs or evaluated in clinical studies, are of *C*-branched structures which can be prepared by introducing methyl [12], allyl [13], ethynyl [12a,14], trifluoromethyl [15], and the other groups [16] through addition of organometallics to sugar moieties. This will not be discussed in detail here.

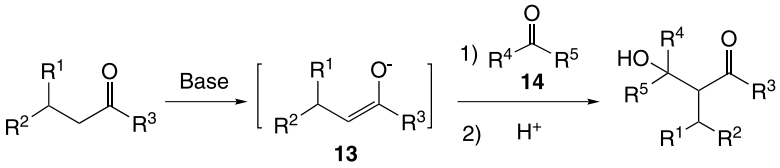
2.2 Aldol-Type Condensations

The aldol condensation, as the most extensively studied reaction in C–C bond formation, is also commonly applied to the synthesis of *C*-branched sugars. In general, it involves the base-catalyzed addition of one molecule of carbonyl compound to a second molecule in such a way that the α-carbon of the first attached to the carbonyl carbon of the second to form a β-hydroxyl compound.

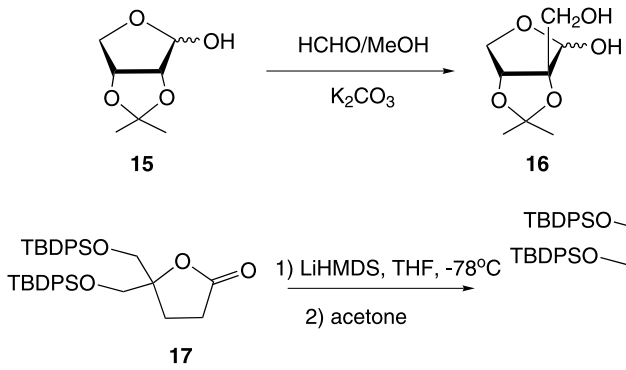
2.2.1 Aldol Reactions

Here, the name "aldol" refers to a β-hydroxycarbonyl compound which is derived from the nucleophilic additions between enol (or enolate) **13** and ketone (or aldehyde) **14** (*Scheme 5*). Sugars, as the chiral polyhydroxy aldehyde or ketone compounds, are natural substrates for these reactions.

Serianni et al. [17] reported the synthesis of branched-chain aldotetrose **16** from lactol **15** and CH_2O under mild basic conditions through aldol reaction. In the same way, condensation of lactone **17** and acetone gave C-2 branched **18** using LiHMDS as a base at low temperature [18] (*Scheme 6*).



Scheme 5



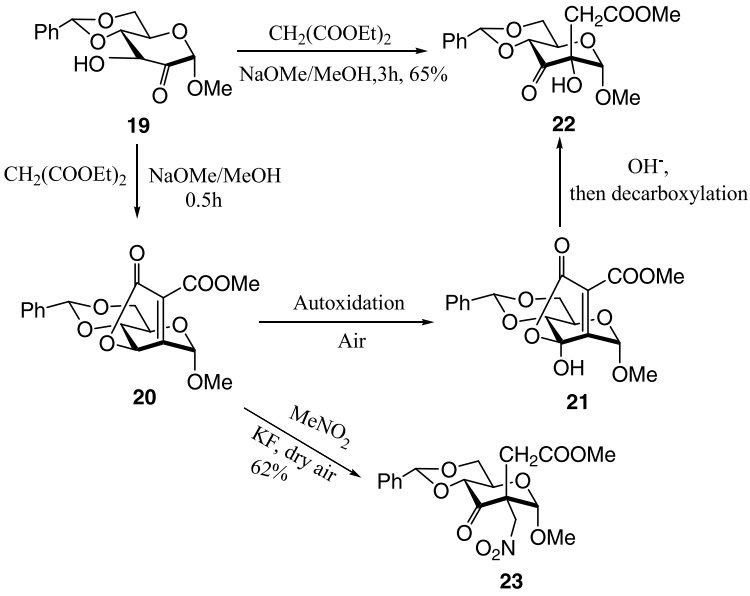
Scheme 6

An interesting aldol condensation between 2-oxoglucopyranoside **19** and diethyl malonate has been carefully investigated and the reaction mechanism was illustrated in *Scheme 7* [19]. A butenolide-containing sugar **20**, available from the aldol condensation of methyl 4, 6-*O*-benzylidene-α-D-glucopyranosid-2-ulose **19**, and diethyl malonate, was autoxidized by air at the C-3 position affording α,β-unsaturated γ-lactone sugar **21**, which subsequently underwent 1,4-conjugate (Michael) addition of hydroxide ion (or water) leading to 2-*C*-branched-chain glycopyranosid-3-ulose **22**. The autoxidations could be performed in either weak basic, neutral, or weak acidic medium, respectively. When active methylene compound was introduced, a new type of *C*-branch sugar **23** was obtained [20].

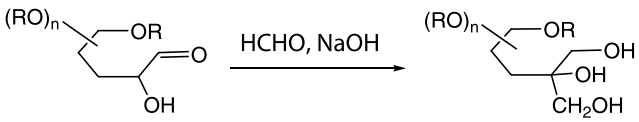
2.2.2 Aldol-Cannizzaro Reactions

When aldehyde, carrying an active α-hydrogen, is coupled with formaldehyde, the product becomes a suitable substrate for Cannizzaro reaction, which can react with formaldehyde subsequently in one pot to give the corresponding alcohol and sodium formate. Sodium hydroxide is believed to be a good catalyst for this reaction (*Scheme 8*).

An efficient method for large-scale preparation of apiose was developed by Kóóš et al. [21], using an Aldol-Cannizzaro reaction as a key step. 2,3-*O*-isopropylidene-L-*threo*-tetrodialdose acetal **24** was reacted with excessive formaldehyde and gave 3-*C*-(hydroxymethyl)-2,3-*O*-isopropylidene-D-*glycero*-tetrose acetal **25**, which was deprotected to afford apios **26**. Besides, treatment of **27** with formaldehyde in the presence of sodium hydroxide yielded branched furanoside **28** under the same conditions [22] (*Scheme 9*). It is worth noting that the prod-



Scheme 7



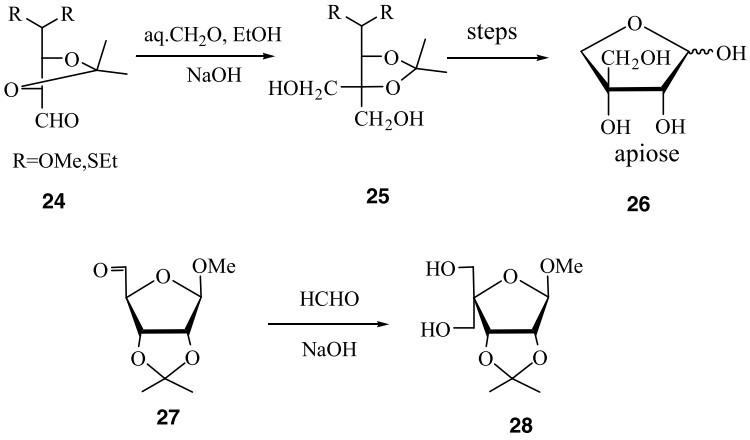
Scheme 8

uct of the Aldol-Cannizzaro reaction can be further transformed into locked nucleic acid (LNC) [23] containing other functional groups, such as ethenyl [24], ethynyl [25], and so on [26] (*Scheme 10*).

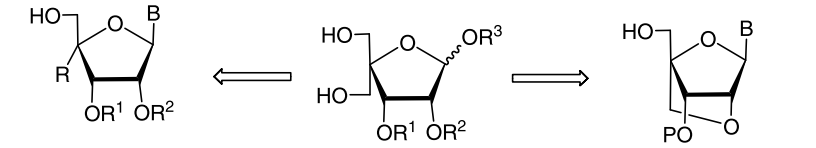
2.2.3 Nitroaldol Condensations (Henry Reaction)

Condensation of aldehyde or ketone with nitroalkane, bearing an α-hydrogen, stereoselectively enables the formation of a new carbon–carbon bond under mild basic conditions with the generation of a β-nitroalcohol, which is called nitroaldol condensation or the Henry reaction (*Scheme 11*). This is a good method for preparing the branch-chain sugars containing nitrogen, since it provides a polyhydroxylated carbon framework with multiple stereogenic centers as well as the possibility for transformation of the nitro group to other functional groups. Many natural or unnatural azasugars have been obtained through this reaction for the development of new glycosidase inhibitors in recent years.

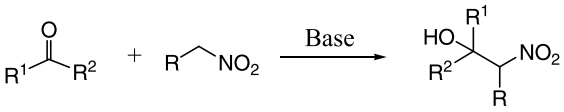
Bols [27] reported an improved synthesis of isofagomin **32** and noeuromycin **34** from D-arabinose in six and seven steps, respectively. As shown in *Scheme 12*, Henry reaction of ketone sugar **29** gave the nitromethane adduct benzyl 4-deoxy-4-*C*-nitromethylene-D-arabi-



Scheme 9



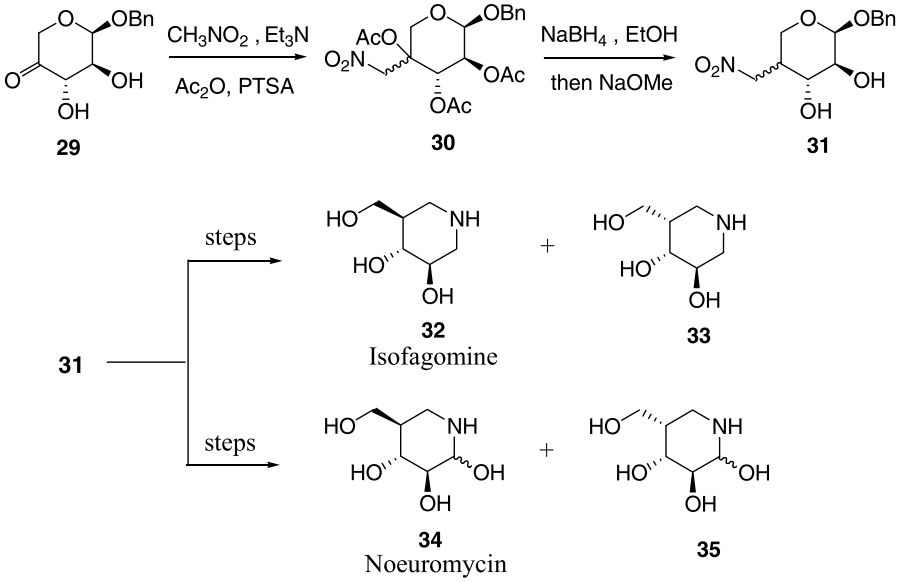
Scheme 10



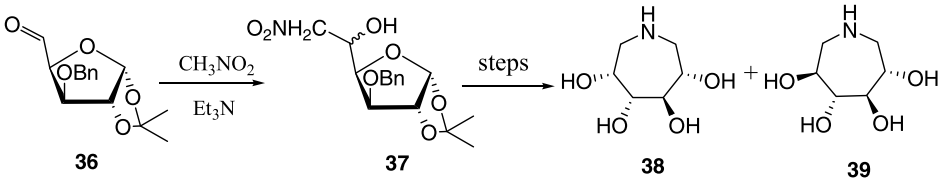
Scheme 11

no-pyranoside. After acetylation, fully protected **30** was obtained as a mixture of epimers which could be separated easily. Reductive elimination of **30** with $NaBH_4$ gave a 4*R*,4*S* mixture of benzyl 4-deoxy-4-*C*-nitromethylene-D-arabinopyranoside **31** in which the desired isofagomine/noeuromycin precursor (4*S*)-**31** was slightly favored (<2:1). Mixture **31** can be transformed into isofagomin **32** and noeuromycin **34**, along with their isomers **33** and **35**, respectively.

In the preparation of polyhydroxylated azepane as potential glycosidase inhibitors, Dhavale [28] described a short synthetic route utilizing the Henry approach. The nitroaldol reaction of 1,2-*O*-isopropylidene-3-*O*-benzyl-α-D-xylo-pentodialdose **36** and nitromethane in the presence of triethylamine at room temperature afforded α-D-gluco- and β-L-ido- nitroaldose **37**, the precursors to (2*S*, 3*R*, 4*R*, 5*R*) and (2*S*, 3*R*, 4*R*, 5*S*) tetrahydroxyazepanes **38** and **39**, in a 88:12 ratio in 95% yield (*Scheme 13*).



Scheme 12

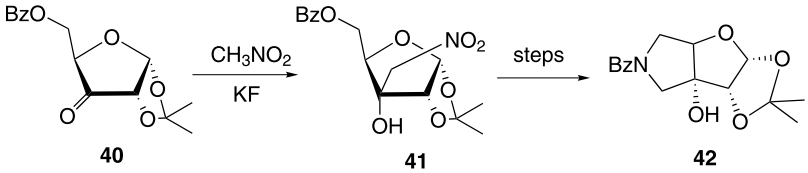


Scheme 13

Using the Henry reaction [29], addition of nitromethane to the carbonyl group of 5-*O*-benzoyl-1,2-*O*-isopropylidene-α-D-erythro-ketofuranos-3-ulose **40** took place stereoselectively to give ribo isomer **41** in almost qunatitative yield (› *Scheme 14*). KF was found to be better than any other bases for this reaction. The absolute configuration of all asymmetric carbon atoms in **41** was confirmed by single-crystal X-ray crystallographic analysis. The stereoselectivity probably resulted from the steric hindrance of the 1,2-*O*-isopropylidene group. Several new *C*-branched iminosugar derivatives bearing a pyrrole ring, such as **42**, were synthesized from **41**.

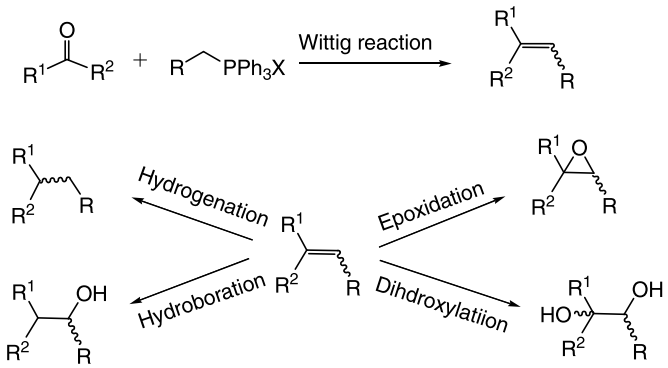
2.3 Wittig Reaction

Wittig reaction and its variants (e. g., Wadsworth–Emmons, Wittig–Horner, etc) are widely applied in the synthesis of *C*-branched sugar derivatives. The Wittig reagent reacts with aldehyde or ketone sugar affording the C=C bond as a potent functional group which could

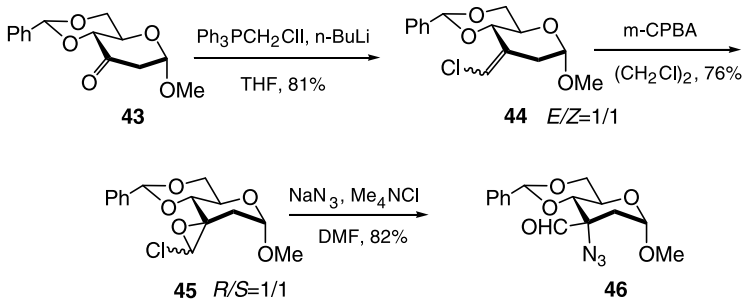


40 41 42

Scheme 14



Scheme 15

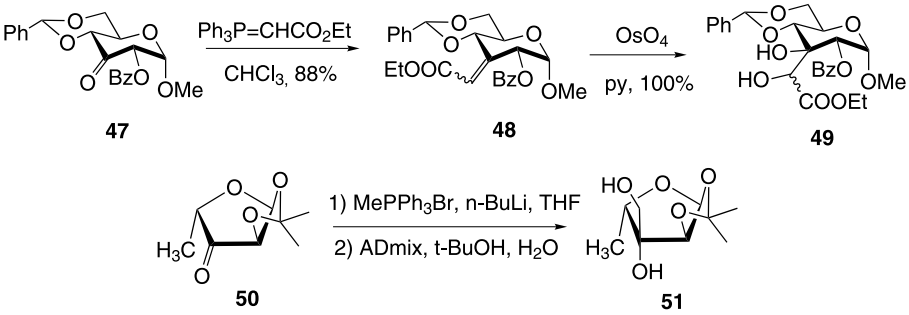


43 **44** *E*/*Z*=1/1 **45** *R*/*S*=1/1 **46**

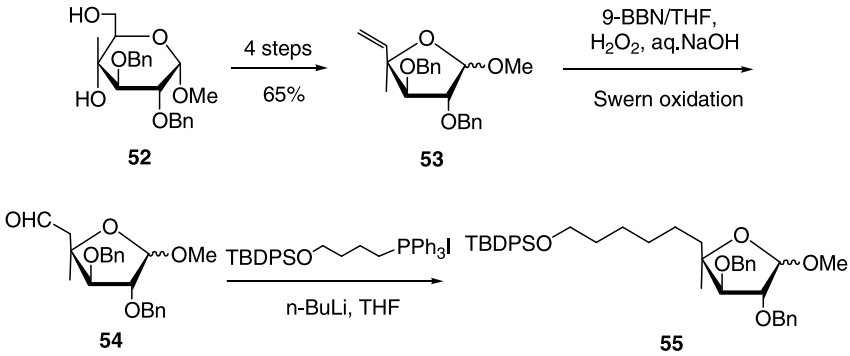
Scheme 16

be converted into epoxides, alcohols, and other compounds via epoxidation, dihydroxylation, hydroboration, and catalytic hydrogenation, etc (*Scheme 15*).

A new synthetic method towards sugar aldehyde via α-chloroolefine and α-chloroepoxides was exploited by Sato et al. [30]. As shown in *Scheme 16*, Wittig reaction of **43** with $Ph_3P(I)CH_2Cl$ and n-BuLi in THF gave the corresponding chloroolefins **44** (*E*/*Z* = 1/1). The configurations of **44***E* and **44***Z* were determined by NMR spectrum (NOESY: H-4 and olefin proton). A mixture of these diastereomers was oxidized with *m*-chloroperbenzoic acid in 1,2-dichloroethane at 70 °C to give a *R*,*S*-mixture (1:1) of spiro α-chloroepoxide **45**, which was then treated with NaN_3 and Me_4NCl in dimethylformamide (DMF) at 80 °C furnishing the corresponding α-azidoaldehyde derivative **46** regioselectively.



Scheme 17

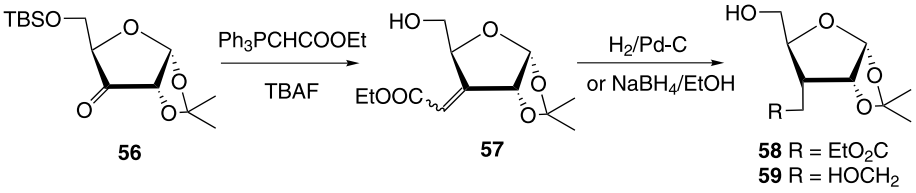


Scheme 18

In the synthesis of a miharamycin sugar moiety [31], a C-3-branched hexopyranoside was constructed through Wittig reaction and dihydroxylation. Olefination of the ketosugar **47** was performed with [(ethoxycarbonyl)-methylene]triphenylphosphorane affording **48** as a mixture of isomers (*Z*:*E* = 7:3). Oxidation of **48** with OsO_4in pyridine led to a mixture of stereoisomers **49** in a quantitative yield, whereas a catalytic amount of OsO_4 and 4-methylmorpholine *N*-oxide in acetone/water gave only a 66% yield of **49**. Similarly, ketone sugar **50** was converted into the *exo*-methylene derivative by reaction with $H_2C{=}CHPPh_3$, a further Sharpless dihydroxylation using AD-mix (ether α or β) afforded the single stereoisomer [32] (*Scheme 17*).

Asymmetric synthesis of macrolide antibiotic, ossamycin, was started from known **52**, which was further converted into furanoside **53** via zinc reduction and subsequent cyclization (*Scheme 18*). Compound **53** was subjected to hydroboration, Swern oxidation, Wittig olefination, and hydrogenation furnished key intermediate **55** towards ossamycin [33].

Treatment of 5-*O*-TBS-1,2-*O*-isopropylidene-α-D-*erythro*-pentofuranos-3-ulose **56** with [(ethoxycarbonyl)-methylene]triphenylphosphorane gave (*E*/*Z*)-**57** (7:1; 90%) (*Scheme 19*). Desilylation and hydrogenation of **57** at 25 psi H_2/Pd-C gave **58** with a trace amount of the over-reduced diol **59**. Formation of **59** was minimized at lower hydrogen pressure (5 psi). Reduction of **57** with $NaBH_4$/EtOH gave worse results. The 1,2-*O*-isopropylidene group is known to direct incoming reagents from the β-face of furanosyl derivatives and thus *ribo* diastereomers **58** (or **59**) were obtained with either catalytic hydrogenation or chemical reduction [34].

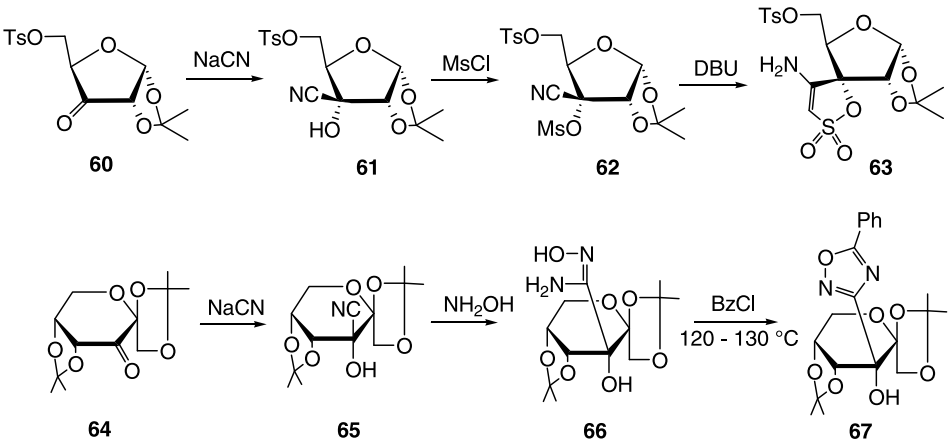


Scheme 19

2.4 Other Nucleophilic Additions

One carbon branch unit could be introduced into a sugar compound by addition of cyanide to sugar aldehyde or sugar ketone. As reported by San-Félix et al. [35], treatment of 3-ulose **60** with sodium cyanide afforded cyanohydrin **61**. The high stereoselectivity in the formation of **61** could be explained by the presence of conformationally rigid 1,2-*O*-isopropylidene functionality, which dictates the approach of the cyanide ion from the sterically less-hindered β-face of the ulose. Mesylation (→**62**) and subsequent aldol-type cyclo-condensation afforded *C*-branched-3-spiro derivative **63**. A psicopyranose derivative containing 1,2,4-oxadiazole could be achieved [36] with the same method, and a potent functionality, such as cyanide, could also be introduced at the same time. Reaction of 1,2:4,5-di-*O*-isopropylidene-β-D-erythro-2-hexulopyranose-3-ulose **64** with NaCN under phase-transfer conditions yielded 1,2:4,5-di-*O*-isopropylidene-3-*C*-cyano-β-D-psicopyranose **65** in nearly quantitative yield and complete stereoselectivity. Refluxing of **65** with hydroxylamine in anhydrous methanol (→**66**) followed by benzoyl chloride treatment expectantly gave 1,2,4-oxadiazole derivatives **67** (*Scheme 20*).

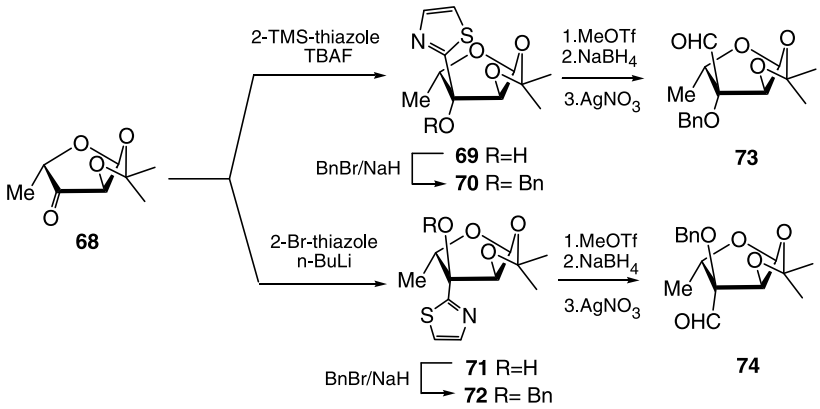
Thiazole-based one carbon extension has been proved to be a useful tool in the synthesis of *C*-branched sugar derivatives [37]. To investigate an approach towards the synthesis of



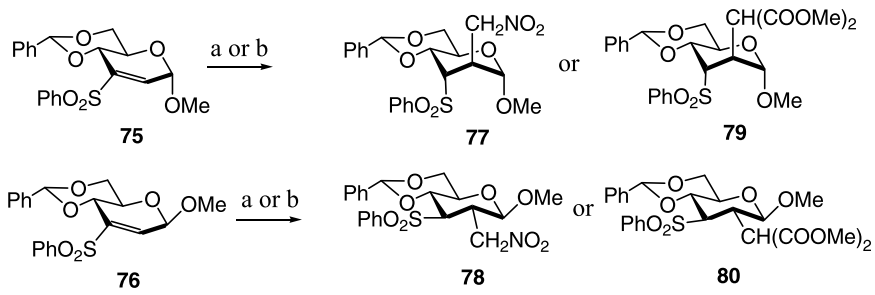
Scheme 20

C-branched sugar aceric acid [32] (● *Scheme 21*), 3-ulose **68** was treated with 2-(trimethylsilyl) thiazole (2-TST) in THF under thermodynamic control, followed by desilylation with tetra-*n*-butylammonium fluoride (TBAF), led to a 5:1 mixture of *endo*- and *exo*-adducts **69** and **71**, respectively. In contrast, addition of 2-thiazolyl lithium to compound **68** gave only the adduct **71**, arising from sterically controlled *exo*-addition to the trioxabicylo[3.3.0]octane ring system in an unoptimized 40% yield. Benzylation of isomeric adducts **69** and **71** afforded 3-*O*-benzyl derivatives **70** and **72**, respectively. The stereochemistry of thiazole addition to 3-ulose **68** was established by X-ray structural analyses of compounds **69** and **72**, which revealed that **72** has the required C-3 configuration for aceric acid synthesis.

The stereoselectivity in the reaction of 3-ketofuranose **68** and 2-thiazolyl lithium can be attributed to the kinetic control and steric hindrance from *endo*-face addition. The thiazolyl group of compounds **70** and **72** was converted into the formyl functionalized sugars **73** and **74** using essentially a one-pot, three-step literature method [38,39].



Scheme 21



Condition a: CH_3NO_2, NaOMe, MeOH, rt
Condition b: $CH_2(COOMe)_2$, NaH, THF, rt

Scheme 22

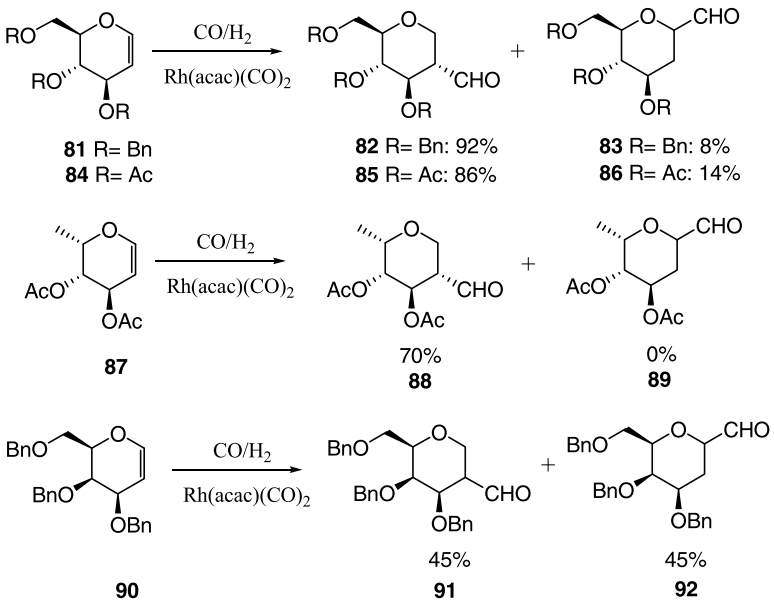
Suresh et al. [40] reported a diastereoselective addition of carbon nucleophiles to vinyl sulfone-modified carbohydrates, in which nucleophiles added to the C-2 position from a direction opposite to that of the anomeric methoxy group (Scheme 22). The nucleophile, generated from CH_3NO_2 and NaOMe, reacted with **75** (or **76**) producing unique **77** (or **78**). Similarly, the nucleophile generated from dimethyl malonate and sodium hydride produced exclusively **79** and **80**, respectively. Interestingly, α-methyl glycoside **75** gave **77** and **79** having α-D-*altro* configuration, while β-methyl glycoside **76** produced β-D-gluco analogues **78** and **80** under the same reaction conditions. Crystal structures of **77**/**79** and **78**/**80** unambiguously established the absolute configurations at positions C-2 and C-3 of these compounds. Pentofuranosides presented the same results.

3 C–C Bond Formation by Metal or Metal Complex Mediated Reactions

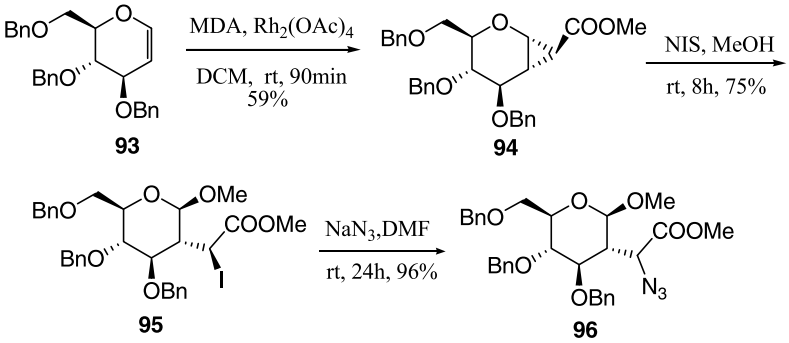
Metal or metal complex-mediated reactions have been widely applied to *C*-glycoside synthesis [41]. Accordingly, various metal or metal complex catalysts such as rhodium, indium, samarium, palladium, and so forth have also been developed for the preparation of *C*-branched-chain sugars.

3.1 Rhodium Complex Catalyzed Reactions

Hydroformylation of alkenes by homogenous rhodium catalysts has attracted the attention of researchers for functionalization of complex molecules [42,43]. Al-Abed et al. [44] recently



Scheme 23



Scheme 24

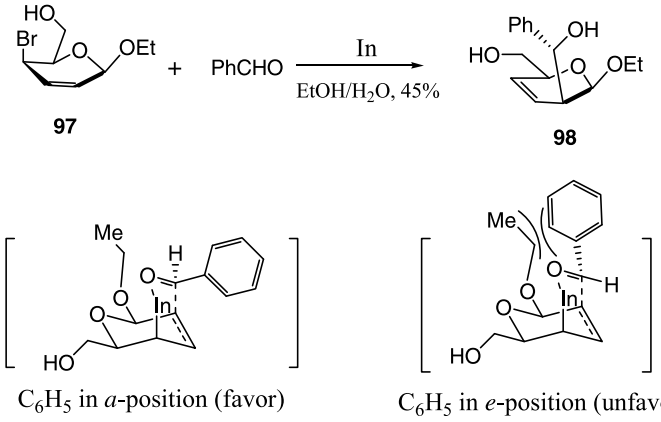
reported hydroformylation of glycals utilizing the catalyst $Rh(acac)(CO)_2$ to give a mixture of the C-1 formyl 2-deoxy-*C*-glycoside and/or the C-2 formylpyran counterpart (*Scheme 23*). Thus, hydroformylation of tri-*O*-benzyl-D-glucal **81** gave formyl adducts **82**:**83** in a ratio of 92:8, and the corresponding acetyl glucal **84** obtained similar results under these conditions. Unintelligibly, the 3,4-di-*O*-acetyl-6-deoxy idopyranose derivative **87** gave only the C-2 formyl product **88**, while tri-*O*-benzyl-D-galactal **90** gave a 1:1 mixture of formyl compounds **91** and **92**. The regioselectivity of formyl addition depends on polarization of the olefin, relative stability of the alkyl-metal complexes, the difficulty of β-elimination for conformationally rigid substrates, and the ratio of acyl-metal intermediates [45].

An efficient method for the synthesis of 2-*C*-branched glyco-amino acid derivatives by diastereoselective ring opening of carboxylated 1,2-cyclopropane sugars has been achieved using the stereocontrolled cyclopropanation of glycals mediated by rhodium acetate [46]. As shown in *Scheme 24*, tri-*O*-benzyl-D-glucal **93** was treated with methyl diazoacetate (MDA) in the presence of a catalytic amount of rhodium acetate furnishing 1,5-anhydro-2-deoxy-1,2-*C*-(*exo*-carbomethoxy methylene)-3,4,6-tri-*O*-benzyl-α-D-glucitol **94** in 59% yield. Treatment of **94** with NIS/MeOH afforded methyl-3,4,6-tri-*O*-benzyl-2-deoxy-2-*C*-(iodomethyl acetate)-β-D-glucopyranoside **95** as a single diastereomer in which two new stereocenters were introduced in a single reaction. Further reaction of **95** with NaN_3/DMF afforded azide **96** which could be converted into a glyco-amino acid ester after reduction.

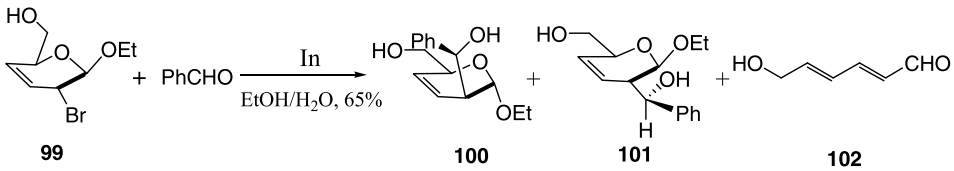
3.2 Indium-Promoted Reactions

Organometallic reactions in aqueous media have been developed and their application in organic synthesis has been increasingly explored [47]. From this perspective, indium chemistry has captured much recent attention due to the comparable catalyzing abilities in aqueous media [48].

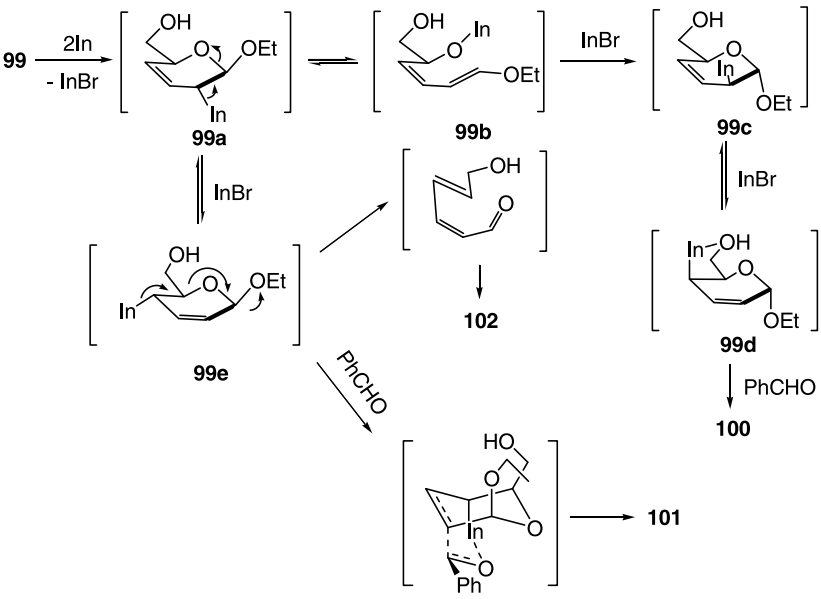
Lubineau reported a series of research works on the preparation of *C*-branched monosaccharides and *C*-disaccharides under indium promoted Barbier-type allylation in aqueous media [49]. In the case of substrate **97**, the reaction, which took place in H_2O/EtOH (1:2) at 50 °C, gave unique stereoisomer **98** with complete regio- and diastereoselectivity. From



Scheme 25



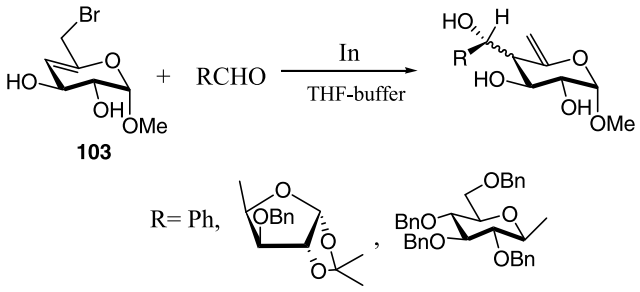
Scheme 26



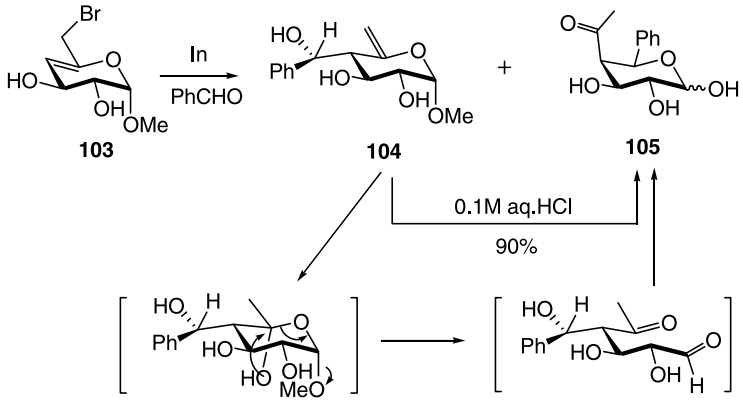
Scheme 27

a mechanistic point of view, the stereochemistry at C-7 can be explained by the formation of a six-membered cyclic transition state between the carbonyl and the allyl-indium complex moiety in which the phenyl group is in the axial position as depicted in *Scheme 25*. Indeed, molecular modeling showed clearly an unfavorable steric interaction between the equatorial phenyl substituent and the β-ethoxy group. When 2-bromo-4-enopyranoside **99** was employed in this reaction, a mixture of the C-2 axial product **100**, the C-2 equatorial adduct **101**, and the known aldehyde **102** was obtained in a 6/6/1 ratio (*Scheme 26*), which could be rationalized by the mechanism given in *Scheme 27*. Firstly, an allylindium(I) species **99a** is formed which may undergo a ring opening through β-elimination. Secondly, indium(I) bromide formed in the reaction may act as a Lewis acid and cause the cyclization of the enol ether derivative **99b** to give **99c**, which will take a more favorable state **99d** via a stereospecific 1,3-allylindium migration. In the same manner, the species **99a** exists in equilibrium with its regioisomer **99e** in the presence of InBr. The desired products **100–102** were derived from these intermediates, respectively.

Starting from methyl 6-bromo-4,6-dideoxy-α-D-*threo*-4-enopyranoside **103**, 4-*C*-branched sugars have been prepared from various aldehydes with the same manner in a THF–phosphate buffer (0.11 M, pH 7.0) as the solvent (*Scheme 28*). If the reaction is conducted in pure



Scheme 28



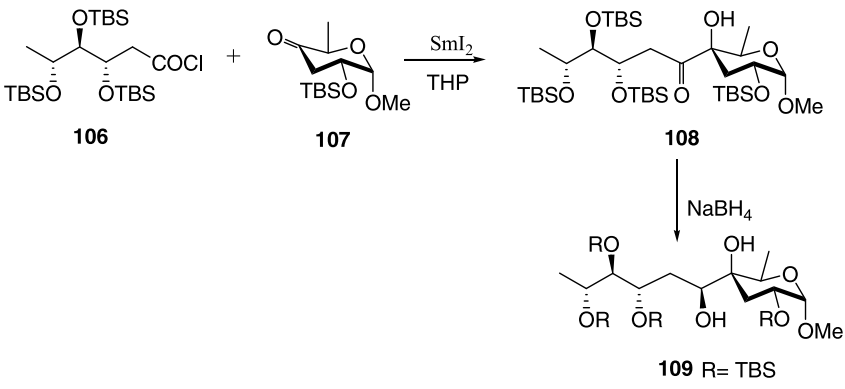
Scheme 29

water, in the absence of the phosphate buffer, the 4-*C*-adduct **104** is slowly transformed (30% after 1.5 h) into an α,β mixture of **105**. In fact, **105** α, β can also be formed from the acid hydrolysis of the labile enol ether derivative **104** (*Scheme 29*). After electrophilic addition of water, followed by pyran ring opening and elimination of the anomeric methoxy group, the cyclization of the C-7 hydroxyl group with the aldehyde led to the keto-derivatives **105** α, β. In order to confirm these results, compound **104** was treated with 0.1 M aq HCl affording **105** as a mixture of anomers (α/β, 1:3) in 90% yield.

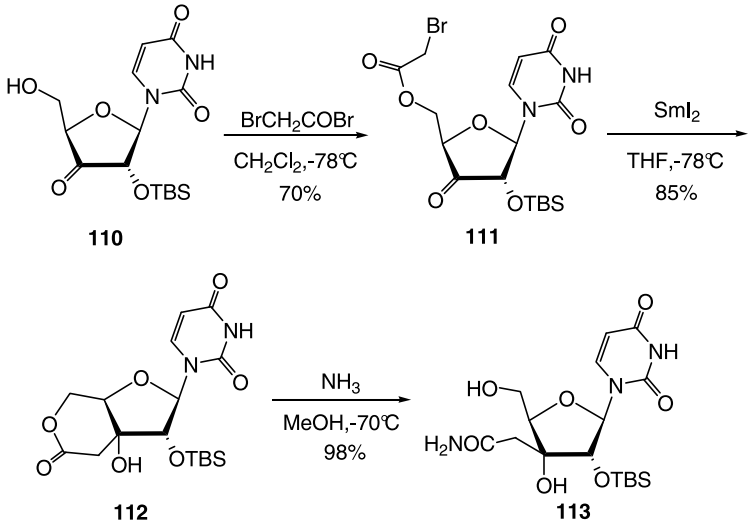
3.3 Diiodosamarium-Mediated Reactions

Prandi [50] reported the preparation of methyl α-D-caryophylloside, a natural 4-*C*-branched sugar, in which the key step was diiodosamarium-promoted coupling reaction. As illustrated in *Scheme 30*, C–C bond formation between the crude acid chloride **106** and ketone **107** was mediated smoothly by SmI_2 in tetrahydropyran (THP). Expected products were isolated in 63% yield and in a 8:1diastereoisomeric ratio. Reduction of the major diastereomer **108** with sodium borohydride in methanol at 0 °C was very slow, but the expected **109** was eventually obtained in 73% yield after 24 h at room temperature.

3′-β-Carbamoylmethylcytidine (CAMC) exhibits potent cytotoxicity against various human tumor cell lines and was synthesized using an intramolecular Reformatsky-type reaction promoted by SmI_2 as the key step [51]. The synthesis of the 3′-β-branched-chain sugar pyrimidine nucleosides is shown in *Scheme 31*. 2′-*O*-TBS-3′-ketouridine **110** was acylated with a bromoacetyl group to give the 5′-*O*-bromoacetyl derivative **111**, the precursor to the intramolecular Reformatsky-type reaction. Treatment of **111** with 2 equiv. of SmI_2 in THF at −78°C afforded the desired lactone **112** in good yield, while the Zn-promoted Reformatsky reaction did not obtain **112** under standard conditions. This is attributed to the strong chelating ability of the samarium enolate to the 3′-carbonyl oxygen to form a six-membered transition state. Ammonolysis of the lactone **112** at −70 °C gave the 3′-carbamoylmethyluridine derivative **113**.



Scheme 30



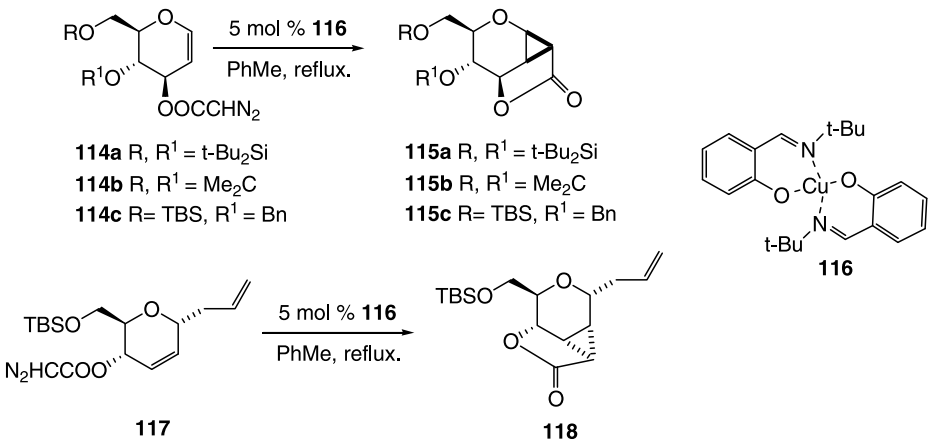
Scheme 31

3.4 Other Metal-Catalyzed Reactions

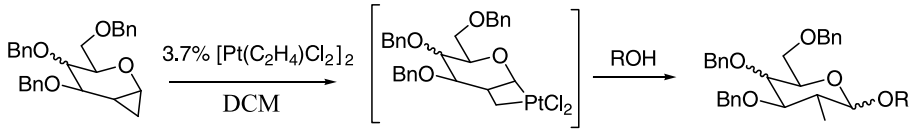
Copper-catalyzed intramolecular cyclopropanation of glycal-derived diazoacetates has been well investigated by Pagenkopf [52]. The cyclopropanation was achieved by the addition of the diazoester to a refluxing solution of 5 mol% bis(*N*-*tert*-butylsalicylaldiminato)copper(II) **116** in either dichloromethane or toluene. Slow addition of the diazoester over 8–12 h was necessary to minimize the formation of dimeric fumarates and maleates. Catalyst **116** performed admirably in the cyclopropanation reactions in this study, whereas other catalysts, including $Rh_2(OAc)_4$, gave lower yields. The intramolecular cyclopropanation of glycals is compatible with a variety of protecting groups at the C-4 and C-6 positions, including cyclic silylene (**114a**), acetonide group (**114b**), and acyclic benzyl and TBS groups (**114c**). Cyclopropane **115a**, **115b**, or **115c** shown in *Scheme 32* was formed as an exclusive stereoisomer in each case. As expected, the cyclopropanation is not limited to electron-rich olefins, and reaction of **117** proceeded with equal efficiency. In addition to the high diastereoselectivity inherent to this intramolecular reaction, the products present a parallel selectivity regardless of the protecting groups employed at C-4 and C-6.

Zeise's dimer $[Pt(C_2H_4)Cl_2]_2$ catalyzed ring opening of 1,2-cyclopropanate of sugars with *O*-nucleophiles generated 2-*C*-branched carbohydrates [53]. A number of *O*-nucleophiles can participate in the ring opening including alcohols, phenols, and water. A wide range of alcohols has been employed to give 2-*C*-branched glycosides ranging from simple methyl glycosides to complex disaccharides. A very high diastereoselectivity is obtained at the newly formed C-1 stereocenter. The α-glycoside, favored by the anomeric effect, is always the major product regardless of the stereochemistry of the starting cyclopropane (*Scheme 33*).

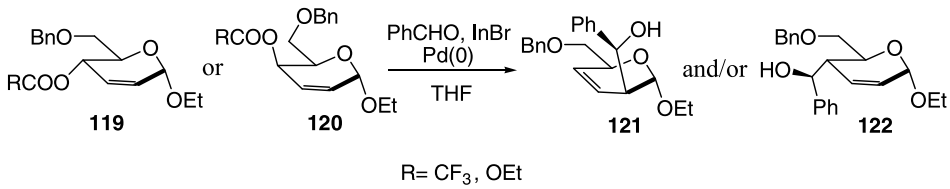
Palladium-catalyzed carbonyl allylation [54] can be effectively applied to the regio- and diastereoselective synthesis of 2-*C*- and 4-*C*-branched sugars **121** and **122** from allylic esters



Scheme 32

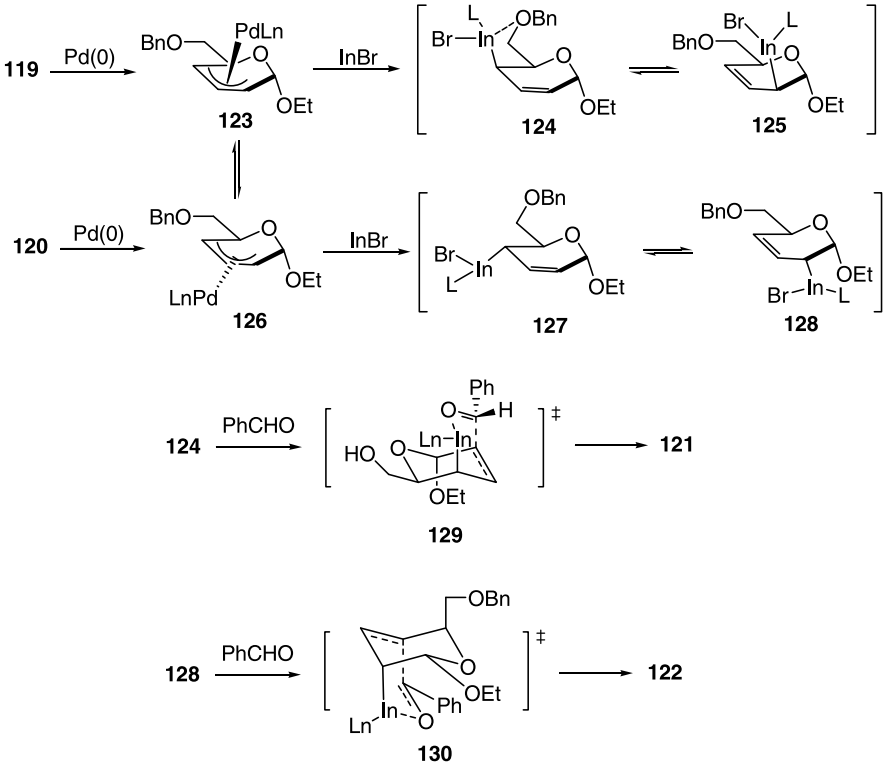


Scheme 33



Scheme 34

or carbonates, **119** and **120**, via formation of π-allylpalladium(II) intermediates and reductive transmetalation with indium(I) bromide (*Scheme 34*). A postulated mechanism is depicted in *Scheme 35*. In the catalytic process, Pd(0) complexes may react with alkenes **119** or **120** to form π-allyl species with inversion of configuration of the carbon bearing the leaving group. The resulting π-allyl palladium(II) complexes **123** or **126** are then reductively transmetalated with indium(I) to give allylindium(III) species **124** and **125** (or **127** and **128**), followed by reacting with benzaldehyde to afford target compounds **121** and **122**, respectively. As was recently confirmed by Lubineau's group [55], InBr approaches the substrate from the same face as the palladium leading to a reductive transmetalation with retention of configuration.



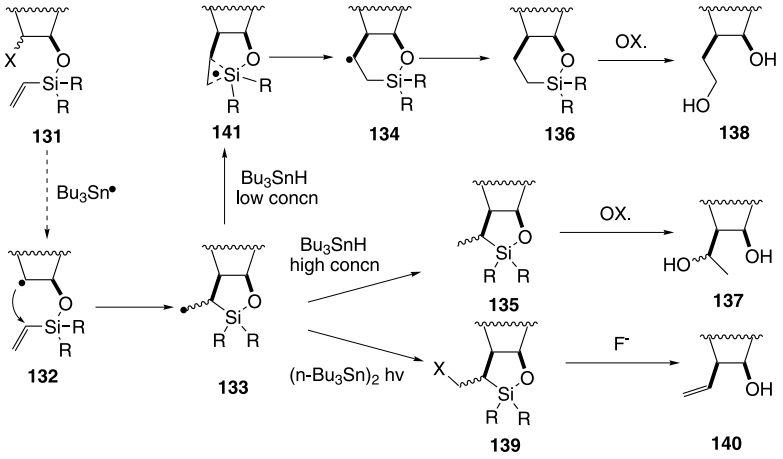
Scheme 35

4 Radical Cyclization

Over the past decades, radical chemistry has been developed into an important and integral part of organic chemistry. Radical cyclization becomes a facile and useful strategy for stereo- and regioselective C–C bond formation, affording useful chiral synthons for the synthesis of *C*-branched sugar derivatives [56,57]. The reactions in this section are divided into intramolecular and intermolecular free radical cyclization.

4.1 Intramolecular Free Radical Cyclization

The most representative examples of intramolecular free radical cyclization in carbohydrate chemistry are the syntheses of *C*-branched nucleosides derivatives. The key step in *C*-branched nucleoside preparation is the regio- and stereo-controlled formation of a new C–C bond at the branching point of the ribofuranose ring [58]. Among published reports, a temporary silicon connection is becoming a growing interest in the syntheses of *C*-branched nucleosides by intramolecular radical cyclization.



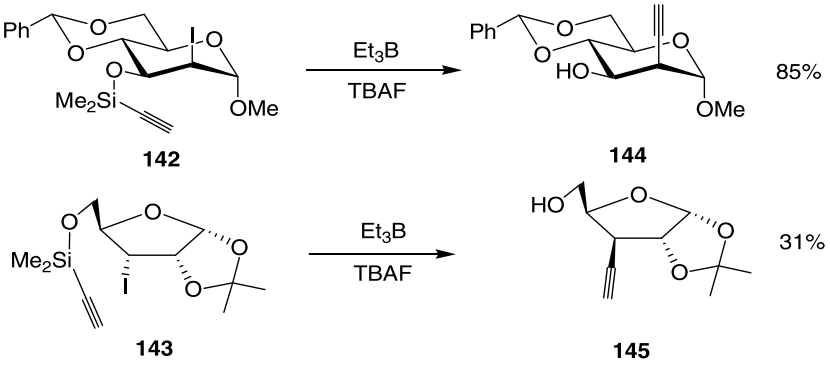
Scheme 36

Shuto's group has developed a highly versatile regio- and stereoselective method for introducing the C_2 substituent via an intramolecular radical cyclization reaction, in which a silicon-containing group was applied as a temporary radical acceptor tether [59,60], as summarized in *Scheme 36*.

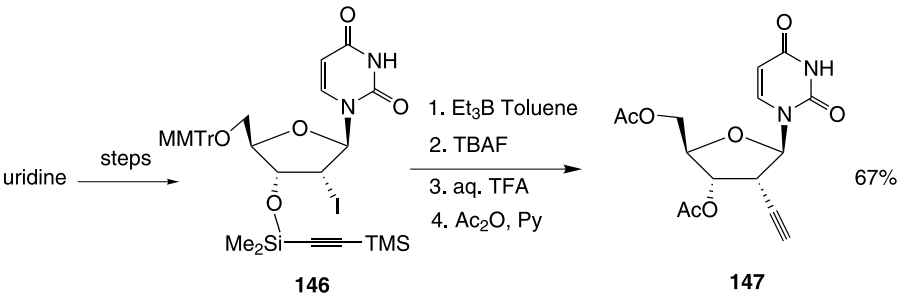
The selective introduction of both 2-hydroxyethyl and 1-hydroxyethyl groups can be achieved, depending on the concentration of *n*Bu_3SnH in the reaction system, via a 5-*exo*-cyclization intermediate **135** or a 6-*endo*-cyclization intermediate **136**, respectively. A vinyl group can also be introduced, to give **140**, by irradiation of the vinylsilyl ether in the presence of (n-$Bu_3Sn)_2$, followed by treatment of the resulting 5-*exo*-cyclization product **139** with a fluoride ion. The mechanism of radical cyclization is that the kinetically favored 5-*exo*-cyclized radical **133** was trapped by high concentration of n-Bu_3SnH to give **135**. At lower concentrations of n-Bu_3SnH and at higher reaction temperatures, the radical **133** rearranged into a more stable, ring-enlarged 4-oxa-3-silacyclohexyl radical **134** via a pentavalent-like silicon radical transition state **141**, which was then trapped with n-Bu_3SnH to give **136**. This method has been applied to the synthesis of many *C*-branched chain sugar nucleoside analogues.

The stereoselective introduction of an ethynyl group in various five- and six-membered iodohydrins has also been developed by the same group (*Scheme 37*) [60]. Treatment of 3′-ethynyldimethylsilyl 2-deoxy-2-iodo-D-mannopyranoside **142** with Et_3B and TBAF in toluene furnished the methyl 2-deoxy-2-*C*-ethynyl-4,6-*O*-benzylidene-α-D-mannopyranoside **144** in 85% yield. However, a similar reaction of 3-iodo-D-ribosyl substrate **143** having the [2-(trimethylsilyl) ethynyl]dimethylsilyl at C-5 gave the desired product 5-*O*-acetyl-3-deoxy-3-*C*-ethynyl-1,2-*O*-(1-methyl ethylidene)-α-D-xylofuranose, in which the ethynyl group was introduced at the γ-*cis* position to the 5′-hydroxyl, in only 31% yield. This may be explained by an unfavored 6-*exo* radical cyclization.

2′-Deoxy-2′-iodo-3′-*O*-TEDMS uridine derivative **146**, readily available from uridine via 2,2′-anhydrouridine, was subjected to the above-mentioned procedure. After removal of the monomethoxytrityl (MMTr) group and subsequent acetylation with Ac_2O in pyridine,



Scheme 37



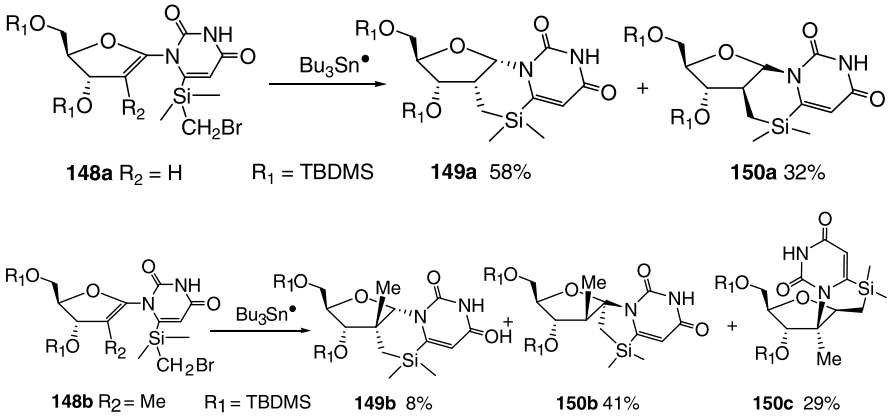
Scheme 38

3,′5′-di-*O*-acetyl-2′-deoxy-2′-*C*-ethynyluridine **147**, a key substrate designed for potential antimetabolites (*Scheme 38*), was isolated in 67% yield.

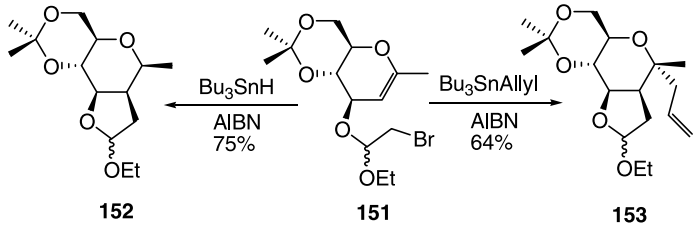
The cyclization (5-*exo* and 6-*endo*) of radicals derived from 6-(bromomethyl)dimethylsilylated glycal uridine, was extensively investigated by Tanaka and coworkers [61] (*Scheme 39*). Without the 2-substituent, 6-(bromomethyl)dimethylsilyl-1-[3,5-bis-*O*-TBDMS-2-deoxy-D-*erythro*-pent-1-enofuranosyl] uracil **148a** was subjected to a radical reaction with azobisisobutyronitrile (AIBN) and Bu_3SnH in refluxing benzene producing specifically 6-*endo*-cyclized products **149a** (58%) and **150a** (32%). No 5-*exo*-cyclized products were formed in this case. In the presence of the 2-methyl substituent, **148b** predominantly afforded 5-*exo* product **150b** under the same reaction conditions. However, the C2 radical intermediate was not stable and a substantial amount (29%) of rearranged product **150c** was formed through anomeric radical species. Cyclization of the other 2-substituted (R_2= CO_2Me, OBz, Cl) derivatives furnished exclusively 5-*exo* products.

Bromo-acetal **151** was subjected to tributyltin hydride promoted radical cyclization and afforded a stereoselective β-*C*-methyl derivative **152**. (*Scheme 40*) Similarly, treatment of **151** under Keck's conditions [62] with allyltributyltin and AIBN resulted in the formation of a bis-*C*,*C*-glycoside **153** [63].

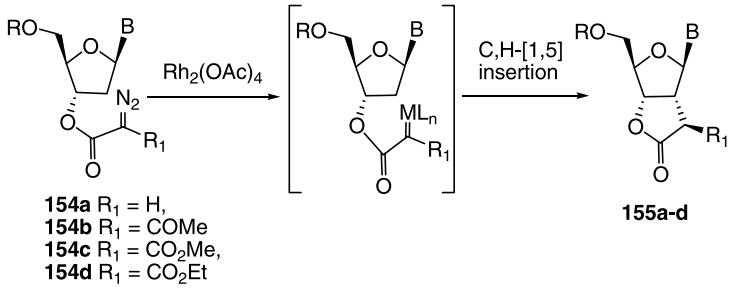
Kim et al. (*Scheme 41*) developed a new method for the stereoselective syntheses of fused α-substituted γ-butyrolactone nucleosides via [1,5]-*C*,*H* insertion of α-diazo-γ- butyrolac-



Scheme 39

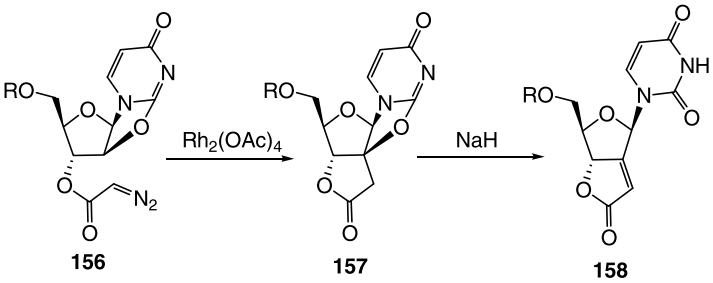


Scheme 40

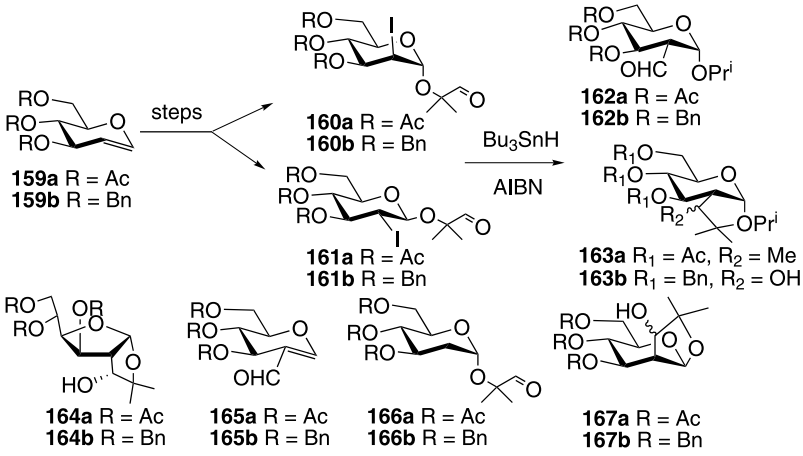


Scheme 41

tone nucleosides [58,64]. 2′-Deoxy-3′-diazoacetates of nucleosides **154a–d** (R_1=H, COMe, CO_2Me, CO_2Et) were chosen as templates for the insertion reactions. Initial studies on stereocontrolled *C,H*-insertion of 2′-deoxy-3′-α-diazoacetate nucleosides were performed in the presence of dirhodium tetraacetate (1.0 mol%) in dichloromethane at room temperature and only a trace amount of product was obtained. However, when the reaction mixture was refluxed, γ-butyrolactones **155a–d** were obtained with high diastereoselectivities in 58–80% yields.



Scheme 42

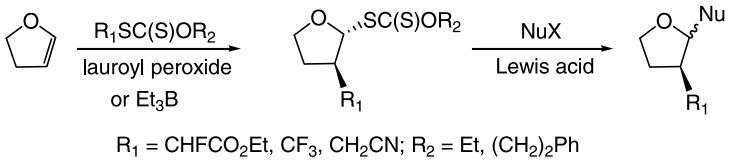


Scheme 43

The same method was applied to the synthesis of biologically active nucleosides (*E*)-2′-deoxy-2′-(carboxymethylene)-5′-*O*-trityluridine-3′,2′-γ-lactone **158** (*Scheme 42*). Exposure of 2′,5′-cyclouridine derivative to the House–Blankey protocol afforded the corresponding diazo compound **156**, which could be converted to the γ-butyrolactone of uridine **157** in 65% yield via [1,5]-*C*,*H* insertion. Lactone **158** could be smoothly obtained by an elimination reaction with sodium hydride in 85% yield.

A new three-step procedure of iodoetherification, ozonolysis, and radical cyclization-fragmentation, converting glucals into C-2 formyl pyranosides, was reported by Choe et al. [65]. The free-radical promoted cyclization of 1,1-dimethyl-2-oxoethyl-3,4,6-tri-*O*-acetyl-2-deoxy-2-iodo-α-D-mannopyranoside **160a**, prepared in 72% yield from commercially available tri-*O*-acetyl-D-glucal **159a** in two steps, produced a mixture of products in varying yields depending on the reaction conditions (*Scheme 43*).

When the reaction was conducted in benzene with 1.0–1.5 equiv. of Bu_3SnH and 0.01–0.5 equiv. of AIBN under reflux, formyl-transfer product isopropyl 3,4,6-tri-*O*-acetyl-2-deoxy-2-*C*-formyl-α-D-glucopyranoside **162a** was isolated in 40% yield. The bicyclic alcohols



Scheme 44

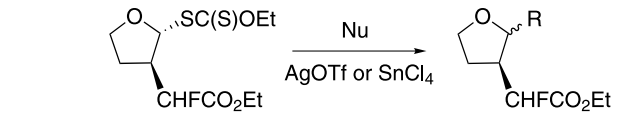
163a (a 1:1 mixture of diastereomers), the C-2 formyl glycal **165a**, C-2 deoxysugar **166a**, and an unexpected furanoside product [1,2-bis-(acetyloxy)ethyl]hexahydro-2,2-dimethyl-furo-[2,3-b]furan-3,4-diol,4-acetate **164a** were also isolated from this reaction with different concentrations of **160a** and the reaction time. If the above radical reactions were promoted by tris(trimethylsilyl)silane, it did not form any of the furanoside **164a**. However, a 65% yield of the C-2 formyl glycal **165a** was isolated and the desired formyl transfer product **162a** was isolated in 41% yield under optimal conditions. This method can also be extended to synthesize C-2 branched galactosides and C-2 cyano glucopyranosides.

4.2 Intermolecular Free Radical Cyclization

A two-step preparation of 2-*C*-branched nucleoside derivatives through highly diastereoselective tandem xanthate radical addition–substitution reactions was investigated by Lequeux and coworkers (*Scheme 44*) [66]. The addition was regioselective and the fluorocarboxylic ester functional group was unambiguously added onto C-2. Attempts to trap the radicals with ethyl acrylate were unsuccessful, and no addition to electron-poor alkenes was observed for this reaction because of the high electrophilic character of the carboxyfluoromethyl radical. The displacement of the resulting anomeric xanthates with various nucleophiles in the presence of Lewis acid allowed the formation of 2,3-disubstituted tetrahydrofuran derivatives.

In the above-mentioned reaction, the corresponding acetals or glycofuranoside derivatives can be prepared using *O*-nucleophiles. Treatment of a single diastereomer, ethyl (2-ethoxythio carbonylsulfanyl-tetrahydrofuran-3-yl)-fluoroacetate (**168**), with ethanol or neopentanol in the presence of silver triflate in toluene led to the formation of a *trans*/*cis* mixture of the corresponding acetals (*Scheme 45*). The reactions gave low stereoselectivity, and the mixtures of 2,3-*trans* and -*cis* isomers ethyl (2-ethoxy-tetrahydrofuran-3-yl)fluoroacetate **169a** and ethyl[2-(2,2-dimethyl-propoxy)-tetrahydrofuran-3-yl] fluoroacetate **169b** were obtained in a 7:3 and 3:2 ratios, respectively.

Introduction of a *C*-nucleophile has been attempted by using organomagnesium reagents (PhMgBr, EtMgBr) in the presence of Lewis acid (AgOTf, $BF_3{\cdot}Et_2O$, $SnCl_4$). Substitution of the anomeric xanthate to form *C*-glycoside was not observed in these cases. However, $SnCl_4$ promoted carbon–carbon bond formation can be performed from the *trans* ethyl (2-ethoxythio carbonylsulfanyl tetrahydrofuran-3-yl)-fluoroacetate **168** and Me_3SiCN at −78 °C, resulting in 2-cyanotetrahydrofurans **169c** with a high yield (83%). This reaction was also attempted with nucleophile TMSAll at 0 °C or −78 °C in the presence of AgOTf, $Cu(OTf)_2$, or $SnCl_4$, but the corresponding alkylated tetrahydrofuran derivatives were only detected as a minor component of the products.

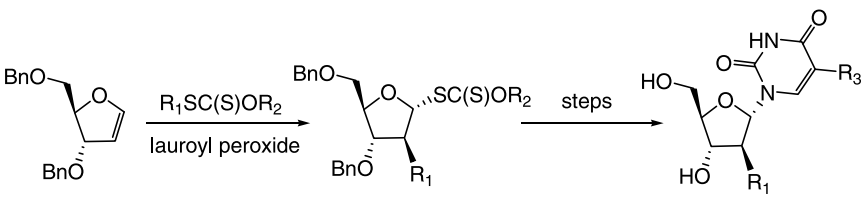


168

Nu = EtOH, $^{t}BuCH_2OH$, TMSCN, $TMSN_3$

169a R = EtO
169b R = $^{t}BuCH_2O$
169c R = CN
169d R = N_3

Scheme 45



170 **171** **172**

R_1 = $CHFCO_2Et$; R_2 = Et; R_3 = CH_3

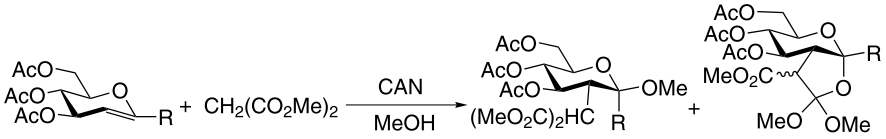
Scheme 46

A two-step synthesis of modified 2'-*C*-nucleoside precursor, ethyl [2-(5-methyl-2,4-dioxo-3,4-dihydro-2*H*-pyrimidin-1-yl)-4-hydroxyl-5-hydroxymethyltetra-hydrofuran-3-yl]fluoro-acetate **172**, from protected glycal **170** and xanthate has been developed following the same idea, and a diastereomeric 1:1 mixture of 2,3-*trans* product **171** was obtained in 57% yield (*Scheme 46*). The use of triethylborane as a free-radical initiator was less successful and a longer reaction time was also required. Interestingly, introducing thymine at C-1 in the presence of silver triflate at 0 °C was highly stereoselective, and only a *C*1,*C*2-*trans* linked product was detected.

Dimethyl malonate in combination with ceric(IV) ammonium nitrate (CAN) has been used to generate electrophilic malonyl radicals [67]. Linker and coworkers extended this methodology with the addition of such radicals to substituted glycals [68] (*Scheme 47*). The reactions proceed smoothly to afford the C-2 branched carbohydrates **174** and **175** with good yields and excellent regioselectivities. For unsubstituted galactal **173a** and carboxamide **173b**, only methyl glycosides **174a** and **174b** were obtained, respectively, whereas the nitrile **173d** afforded exclusively the ortho esters **175d**. On the other hand, the ester **173c** gave a mixture of both products **174c** and **175c**.

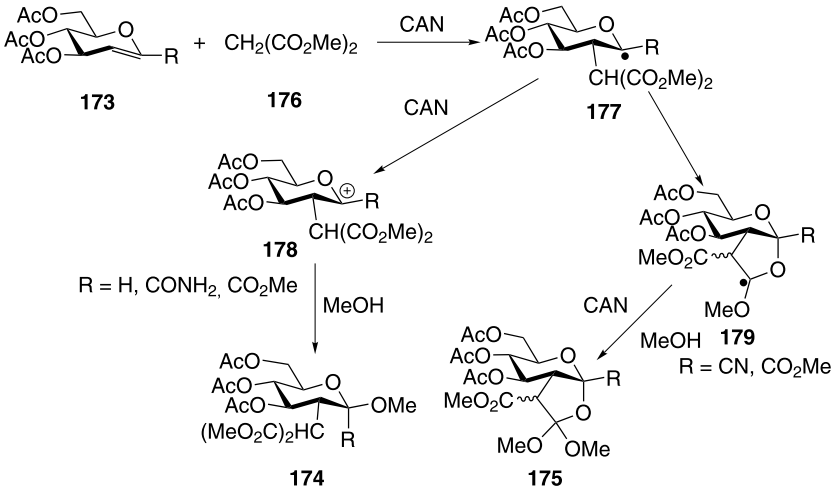
This result can be rationalized by the interaction between the SOMO of the electrophilic radical and the HOMO of the double bond. Furthermore, due to the steric shielding of the pseudo axial *O*-acetyl group, the radicals attack the double bond selectively from the α-face. The reaction mechanism is depicted in *Scheme 48*.

In an approach towards the synthesis of *C*-oligosaccharide containing α-D-Man-(1–4)-D-Man repeating units, C-4 allylated building block **183** was designed [69]. However, the allylation of iodide **181** at C-4 with allyltributyltin and AIBN or dilauoryl peroxide (DLP) slow-

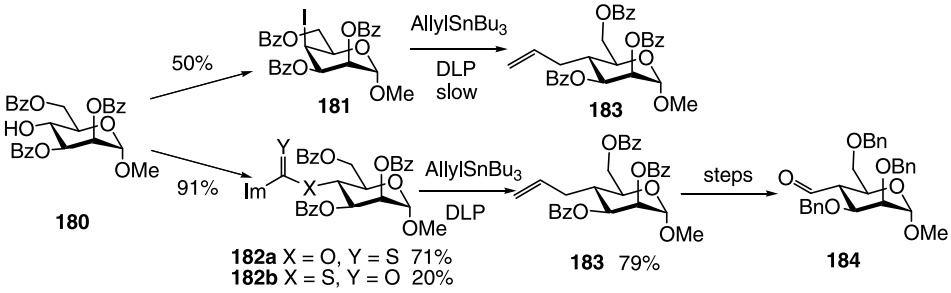


173a R = H — **174a** 86% — 0%
173b R = $CONH_2$ — **174b** 71% — 0%
173c R = CO_2Me — **174c** 7% — **175c** 71%
173d R = CN — 0% — **175d** 80%

Scheme 47



Scheme 48



Scheme 49

ly gave desired **183**, together with considerable 4-deoxy byproduct due to 1,3-diaxial repulsion between iodide and the C_2-benzoate (*Scheme 49*). When thionocarbamate **182** was subjected to allyltributyltin and dilauoryl peroxide in refluxing benzene, only the equatorially C_4-allylated methyl 2,3,6-tri-*O*-benzoyl-4-deoxy-4-*C*-allyl-D-mannopyranoside **183** was

obtained with 79% yield. After the transformation of protecting groups, double bond migration and olefin ozonolysis, formyl-branched monosaccharide **184** was achieved in high yield.

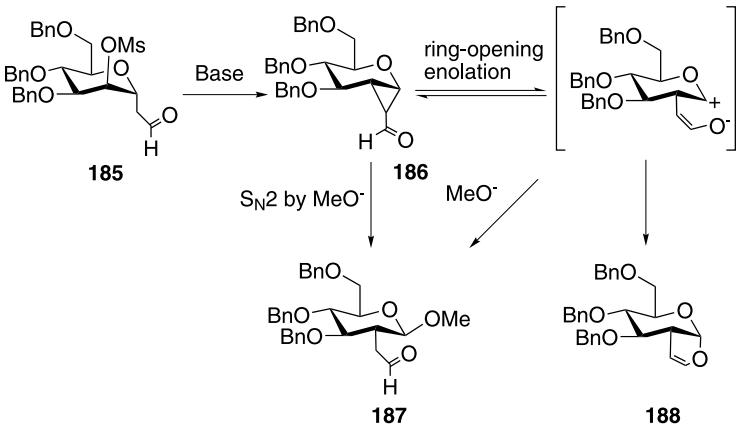
5 Rearrangement and Cycle Additions

Many of the 2-*C*-branched sugars are synthesized from glycals through 1,2-cyclopropanation and a subsequent selective ring opening via solvolysis in the presence of a stoichiometric amount of mercury(II) salts, strong acid, or halonium ions. In all cases an anomeric mixture of glycosides were often provided through an oxocarbonium-like intermediate, although α-glycosides are being favored due to the anomeric effect.

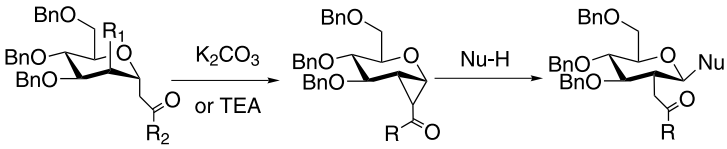
Zou recently reported a novel 1,2-migration of a 2′-oxoalkyl group via 1,2-cyclopropanate sugar derivatives [70]. Treatment of 2′-oxoalkyl 2-*O*-Ms-α-*C*-mannosides **185** with sodium methoxide in methanol produced 2-*C*-branched methyl β-glycoside **187** (40–50%) and a bicyclic derivative **188** (10–15%) (*Scheme 50*). The proposed mechanism is that base-catalyzed enolate of 2′-oxoalkyl 2-*O*-Ms-α-*C*-mannosides **185** afforded 1,2-cyclopropanated intermediate **186** smoothly, which in turn underwent ring-opening with methoxide at the anomeric carbon to give 2-*C*-formylmethyl-2-deoxy-β-glycoside **187**. Presumably, an intramolecular rearrangement afforded bicyclic **188**. When **185** was treated with weaker bases, such as TEA in methanol or K_2CO_3 in acetonitrile-methanol co-solvent, **187** was obtained as the sole product, since the ring-opening enolation becomes less likely with a weaker base.

The ring opening of 1,2-cyclopropanated sugars, formed from 2′-aldehydo(acetonyl)-2-*O*-Ms(Ts)-α-*C*-glycosides, by nucleophiles such as alcohols, thiols, and azide under weak basic conditions, resulted in the formation of 2-*C*-branched β-glycosides and glycosyl azides in good to excellent yields. The best results were obtained with thiol nucleophiles and 1,2-*trans* 2-*C*-branched β-glycosides were always given (*Scheme 51*).

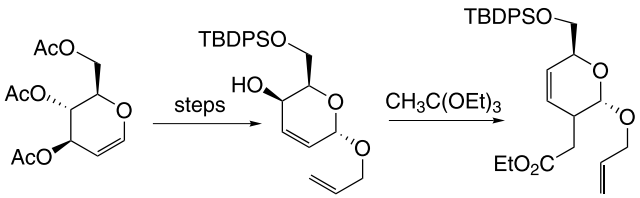
In the development of new sugar amino acids and peptidomimetics, 6-allyloxy-2-[(*tert*-butyl diphenylsilanyloxy)methyl]-3,6-dihydro-2*H*-pyran-3-ol **189** was subjected to the refluxing



Scheme 50



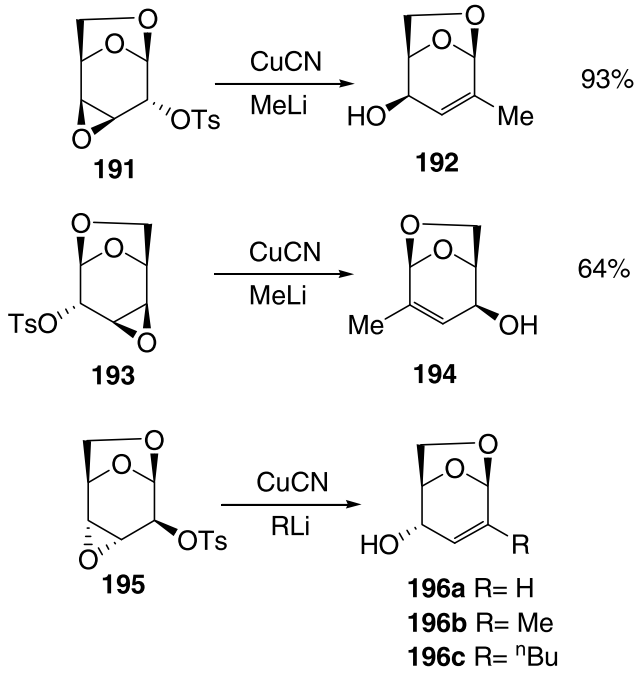
Scheme 51



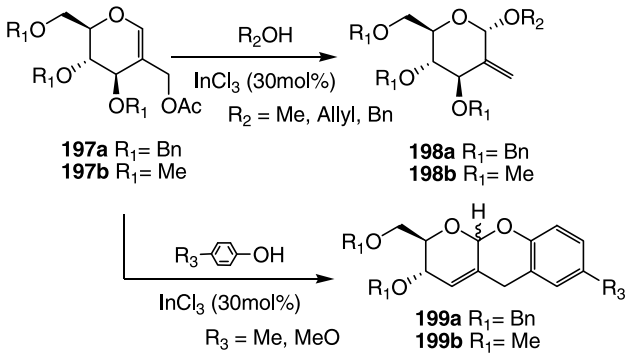
Scheme 52

triethyl orthoacetate in the presence of propanoic acid and hydroquinone, obtaining ethyl 2-{(2*S*,3*S*,6*S*)-2-allyloxy-6-[(*tert*-butyldiphenylsilanyloxy)-methyl]-3,6-dihydro-2*H*-pyran-3-yl} acetate **190** as the only product via Claisen–Johnson rearrangement [71] (*Scheme 52*). Krohn reported a base-catalyzed vinyl tosylate-cuprate cross coupling to form regio- and stereodiverse *C*-branched 1,6-anhydrosugar building blocks [72]. As shown in *Scheme 53*, treatment of epoxide 1,6:3,4-dianhydro-2-*O*-tolylsulfonyl-β-D-galactopyranose **191** in THF with Gilman methyl cuprate afforded 1,6-anhydro-2,3-dideoxy-2-methyl-β-D-threo-hex-2-enopyranose **192** as the sole product. The reaction was initiated from the rearrangement of epoxide to allyl alcohol under basic conditions, and followed by a subsequent cross-coupling between vinyl tosylate and the cuprate. Under the same reaction conditions, 1,6:2,3-dianhydro-4-*O*-*p*-tolylsulfonyl-β-D-mannopyranose **193** and 1,6:3,4-dianhydro-2-*O*-*p*-toluolsulfonyl-β-D-altropyranose **195** was converted into 1,6-anhydro-3,4-dideoxy-4-methyl-β-D-threohex-3-enopyranose **194**, and 1,6-anhydro-2,3-dideoxy-2-methyl-β-D-erythrohex-2-enopyranose **196b**, respectively. Compound **195** was also reacted with the corresponding *n*-butyl Gilman cuprate to afford 1,6-anhydro-2-butyl-2,3-dideoxy-β-D-erythrohex-2-enopyranose **196c** in 68% yield, in addition to 22% of 1,6-anhydro-2,3-dideoxy-β-D-erythrohex-2-enopyranose **196a**. The product **196a** may be formed by hydrolysis of intermediates, such as metalated vinyl species.

An efficient route towards the syntheses of chiral 2-*C*-methylene-*O*-glycosides and chiral pyrano[2,3-b][1] benzopyrans using $InCl_3$ as the catalyst was reported [73]. Accordingly, reaction of 2-*C*-acetoxymethyl glycal derivatives and aliphatic (or aromatic) hydroxyl compounds in the presence of $InCl_3$ via Ferrier rearrangement furnished the corresponding 2-*C*-methylene glycosides (*Scheme 54*). $InCl_3$, $In(OTf)_3$, and $Yb(OTf)_3$ were tested to catalyze the Ferrier rearrangement of 2-*C*-acetoxymethyl-3,4,6-tri-*O*-benzyl-D-glucal **197a** and the best result was obtained with 30 mol% $InCl_3$. The substrates **197a** and **197b**, on treatment with benzyl



Scheme 53



Scheme 54

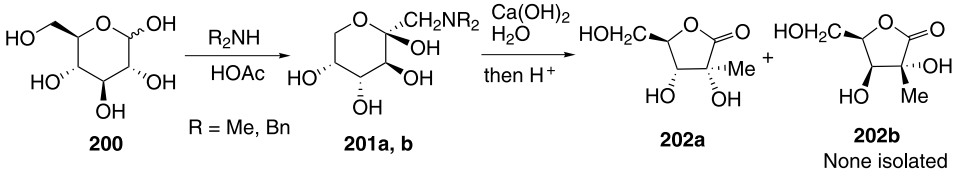
alcohol, obtained the corresponding 2-*C*-methylene glycosides **198a** and **198b**, respectively, in good yields and exclusive α-selectivity. The excellent yields and exclusive α-selectivity were also obtained in the reactions of 2-*C*-acetoxymethyl-3,4,6-tri-*O*-methyl-D-galactal with benzyl, allyl, and *tert*-butyl alcohols. Interestingly, the reaction of 2-*C*-acetoxyglycals **197a** (or **197b**) with phenols formed the corresponding chiral carbohydrate-pyranobenzopyran derivatives **199a** (or **199b**) via Ferrier rearrangement and tandem cyclization in excellent yields and moderate to high stereoselectivities.

The reaction of dibenzylamine with glucose **200** in ethanol, in the presence of acetic acid, gave a high yield of Amadori ketose **201**, which was subsequently treated with calcium hydroxide in water and afforded, after acidic work-up, the branched ribono-1,4-lactone **202a**. No epimeric arabinonolactone **202b** was formed under these conditions [74] (*Scheme 55*).

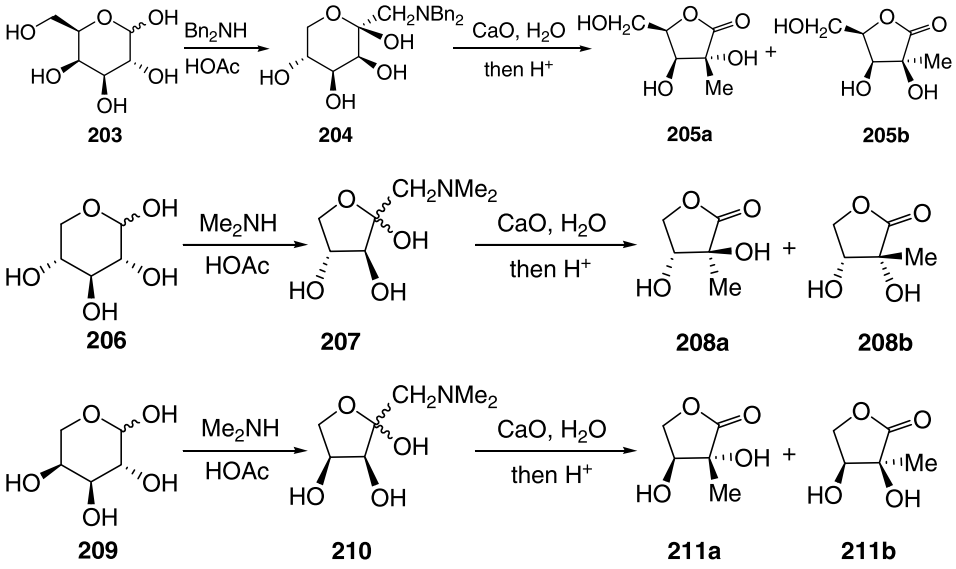
D-galactose **203** underwent the same Amadori rearrangement [75,76] to give crystal α-anomer of tagatosamine **204** (88% yield), which was subjected to CaO in H_2O and an acidic work-up, affording separable 2-*C*-methyl-branched lyxono-**205a** and xylono-**205b** [77] (*Scheme 56*).

Treatment of D-xylose **206** or L-arabinose **209** with the same procedure obtained 2-*C*-methyl-D-erythrono-**208a** and D-threono-**208b**, or L-threono-**211a** and L-erythrono-**211b**, respectively.

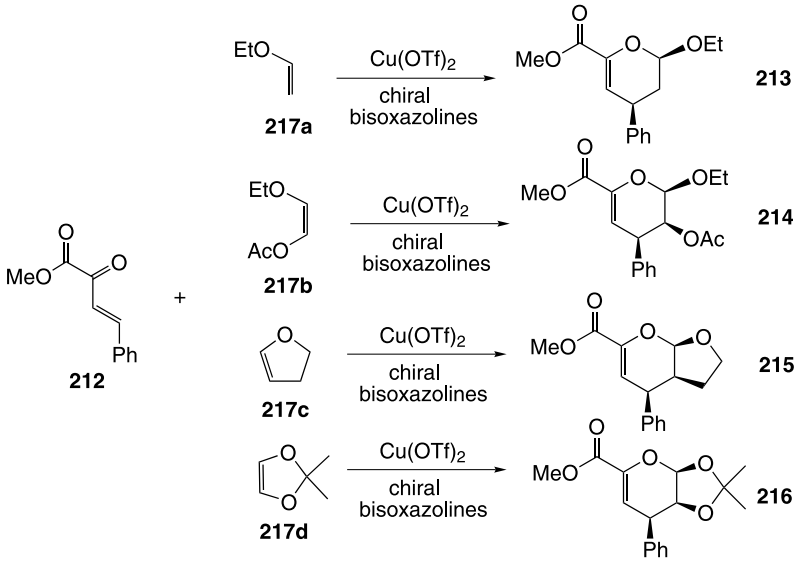
Several diastereoselective HDA (hetero-Diels–Alder) reactions of α,β-unsaturated carbonyl compounds and electron-rich alkenes have been exploited to gain carbohydrate derivatives with good diastereomeric excess. In HDA reaction, up to three chiral centers are formed with high stereoselectivity at each chiral carbon [78].



Scheme 55



Scheme 56



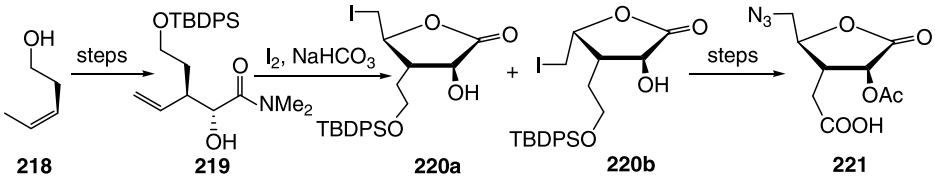
Scheme 57

Two independent groups demonstrated that the reaction between β,γ-unsaturated γ-keto esters and ethyl vinyl ether, in the presence of chiral bisoxazoline copper(II) complexes, led to enantiomerically enriched dihydropyrans which could be converted into attractive carbohydrate derivatives [79,80]. The representative reactions of (*E*)-2-oxo-4-phenylbut-3-enoic acid methyl ester **212** with ethyl vinyl ether **217a** in the presence of different C_2-bisoxazoline ligands and copper(II) salts are presented in *Scheme 57*. The HDA reactions gave the dihydropyran **213** in very high yield (93–99%) with predominantly one diastereomer (de > 98% and ee > 99.5%). The same reaction with **212** and **217b–d** was carried out smoothly to give dihydropyran moieties **214–216** with high diastereoselectivity (de > 95%) and enantioselectivity (up to 99.5% ee), respectively. These HDA products **213–216** are good synthons in the preparation of spirosugars and *C*-branched sugars.

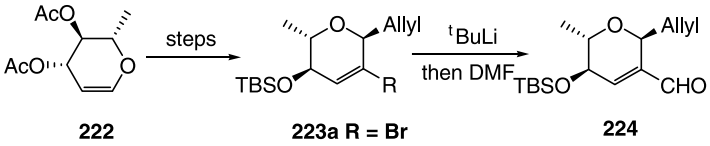
6 Other Methods

Traditional syntheses of *C*-branched sugar analogs usually start from readily available chiral pool compounds, such as nucleosides or carbohydrates, by taking advantage of the already set or easily adjusted stereochemical relationships. However, such routes are frequently lengthy and laborious due to the multifunctional group sensitivities and extensive protecting group manipulations. In contrast, the asymmetric synthesis is extremely flexible for optimization. Because the upstream starting materials are usually small molecules, it is much easier to find a new reaction that provides a better way to a key intermediate [81].

Stereoselective iodolactonization of small achiral molecules is a very useful methodology to create a tetrahydrofuran framework leading to 3,′5′-*C*-branched carbohydrates [82]. Starting



Scheme 58



Scheme 59

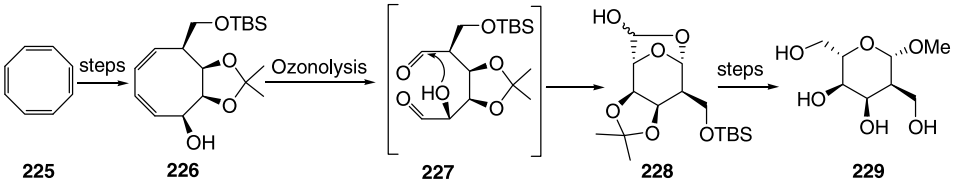
from the commercially available (*Z*)-3-penten-1-ol **218**, iodolactonization of dimethyl amide **219** in THF in the presence of $NaHCO_3$ gave a mixture of 3,5-*trans* and *cis* iodolactones **220a** and **220b** in a ratio of 4:1. Attempts to optimize the reaction conditions, such as temperature, solvent, base, and amide substituents, did not improve the *trans*/*cis* ratio. Azide substitution of **220a** produced the enantiomerically pure ribonucleoside analog **221** (*Scheme 58*).

Fraser-Reid [83] reported a two-step preparation of 2-*C*-branched sugar derivatives through halogen/metal exchange (*Scheme 59*). 5-Bromo-3-(*tert*-butyldimethylsilyl)oxy-2,3-dihydro-2-methyl-6-(prop-2-en-1-yl)-6*H*-pyran **223a** was prepared in four steps from commercially available 3,4-di-*O*-acetyl-l-rhamnal **222**. All attempts to trap the vinyl lithium with a variety of halomethyl alkoxy electrophiles were unsuccessful. However, when a THF solution of **223a** was mixed with 5 equiv. of DMF and 6 equiv. of *tert*-butyllithium at −78 °C, the enal **224** was furnished quantitatively. It is worth mentioning that adding *tert*-butyllithium to the mixture, prior to mixing with DMF, led to low yields (33–69%) of enal **224**, with substantial formation of **223b**.

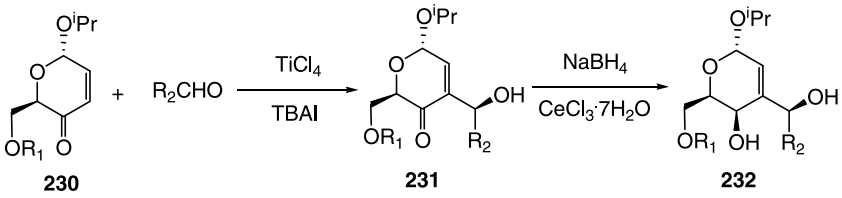
The employing of cyclic polyene, like cyclooctatetraene (COT) **225**, seems to be an esoteric and interesting synthetic approach aiming at the synthesis of hexoses and their branched analogues. Mehta reported the transformation of **225** into a rare sugar (DL)-β-allose and its C_2-branched sibling. Acetonide and TBS-protected cyclooctadienediol **226**, readily available from **225** in steps, was subjected to ozonolysis leading to the bicyclic hemiacetal **228** through the intermediacy **227**. Further modification of hydroxyl groups led to (DL)-methyl-2-deoxy-C_2-hydroxymethyl-β-allopyranoside **229** [84] (*Scheme 60*).

Highly diastereoselective synthesis of C-3 branched deoxysugars has been studied by Shaw's group using Morita–Baylis–Hillman (MBH) reactions. The three-component reactions of aldehyde, sugarenone **230** (prepared from 3,4,6-tri-*O*-acetyl-D-glucal) and $TiCl_4$ were investigated under various reaction conditions. Gradually increasing reaction temperature from −78 °C to −30 °C, together with the adding of TBAI or Me_2S, obtained satisfactory yields and stereo outcomes [85].

Both aromatic and aliphatic aldehydes formed the products in good to excellent yields with almost complete diastereoselectivity (diastereoselectivity >99%). The adduct **231** underwent

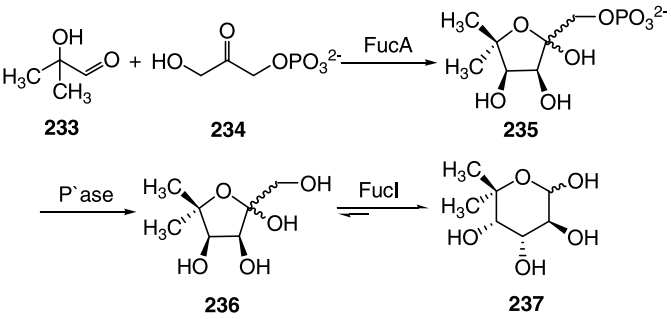


Scheme 60



R$_1$ = H, Ac, *p*-NBz, TBDPS, TBDMS R$_2$ = aromatic and aliphatic groups

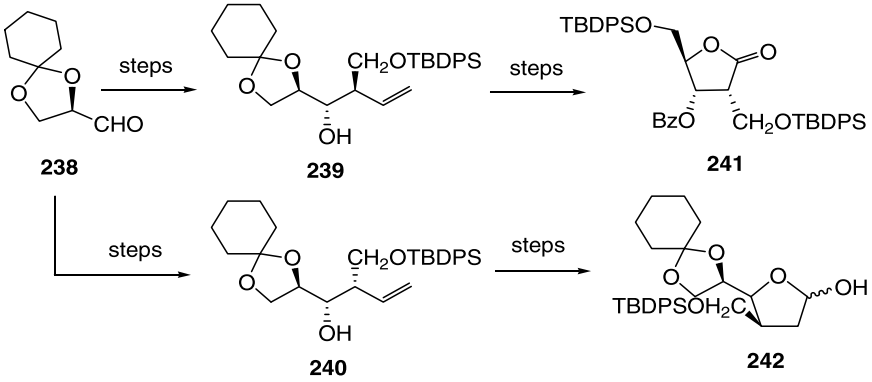
Scheme 61



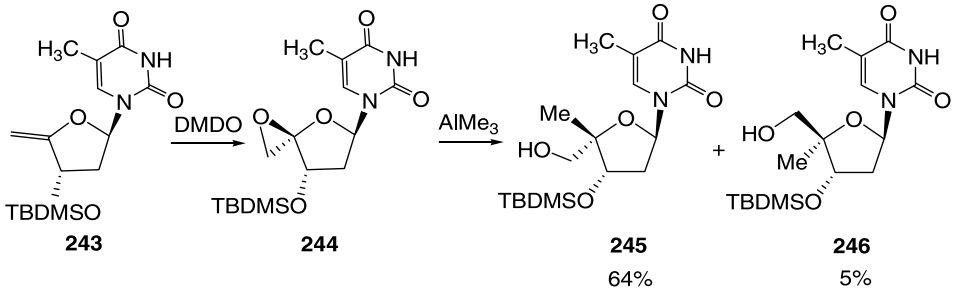
Scheme 62

$NaBH_4$ reduction with the help of $CeCl_3 \cdot 7H_2O$ to obtain *threo* derivatives **232** (Scheme 61). An enzymatic route for the synthesis of L-fucose analogs modified at the non-reducing end is reported by Fessner et al. [86]. Using 2-Hydroxy-2-methylpropanal **233** and dihydroxyacetone phosphate **234** as substrates, branched fucose derivative **237** has been prepared via recombinant L-fuculose 1-phosphate aldolase (FucA) and L-fucose ketol isomerase (FucI) in *E. coli* (Scheme 62).

Zinc-mediated Barbier-type addition of **238**, followed by Luche's procedure obtained a mixture of the homoallylic alcohol diastereomers **239** and **240**. Alcohol **239** was carried through benzoylation, deketalization, silylation, and ozonolysis, to produce *C*-branched γ-lactone **241**. Benzoylation of **240** followed by hydroboration, PCC oxidation, debenzoylation, and alkaline-promoted cyclization directly formed *C*3-branching 2-deoxyfuranose **242** [87] (Scheme 63). A novel method for stereoselective synthesis of 4′-α-carbon-substituted nucleosides, through epoxidation of 4,′5′-unsaturated nucleosides and $SnCl_4$-promoted epoxy ring opening, was



Scheme 63



Scheme 64

reported by Haraguchi [88]. When 3-*O*-TBDMS-4,5-unsaturated thymidine **243** was treated with an acetone solution of DMDO (1.5 equiv.) at −30 °C, 4,5-epoxythymidine **244** was obtained as a single isomer. Ring-opening of **244** with Me_3Al (3 equiv.) afforded 4′-β-methyl isomer **245** (64%) and its isomer **246** (5%) (*Scheme 64*). However, reaction of **244** with allyltrimethylsilane (3 equiv.) and $SnCl_4$ (3 equiv.) gave a mixture of expected 4′-β-allylthymidine and byproduct 5′-*O*-trimethylsilyl derivative.

In conclusion, C–C bond formation on carbohydrates has absorbed lots of widely used methods from general synthetic organic chemistry. Typical among these methods are intramolecular alkylation and intramolecular condensation of aldehyde with enolates, phosphonates, and nitro-stabilized anions. Metal-mediated radical reactions, cycloadditions, and rearrangements have also been applied frequently.

Acknowledgement

Financial support from NNSFC (Project 20621703) is gratefully acknowledged.

References

1. Lipshutz BH, Sengupta S (1992) Org React 41:135
2. Giuliano RM, Villani FJ (1995) J Org Chem 60:202
3. Danishefsky SJ, Pearson WH, Harvey DF, Maring CJ, Springer JP (1985) J Am Chem Soc 107: 1256
4. Gawley RE, Aubé J (1996) Principles of Asymmetric Synthesis. Pergamom Press, Oxford
5. Sato K, Yoshimura J (1982) Carbohydr Res 103:221
6. Smith GR, Villani FJ, Failli L, Giuliano RM (2000) Tetrahedron Asymmetry 11:139
7. Sørensen MH, Nielsen C, Nielsen P (2001) J Org Chem 66:4878
8. Dominguez C, Ezquerra J, Baker SR, Borrelly S, Prieto L, Espada M, Pedregal C (1998)Tetrahedron Lett 39:9305
9. Baker CH, Banzon J, Bollinger JM, Stubbe J, Samano V, Robins MJ, Lippert B, Jarvi E, Resvick R (1991) J Med Chem 34:1879
10. Ong SP, McFarlan SC, Hogenkamp HP (1993) Biochemistry 32:11397
11. Ye JD, Liao XG, Piccirilli JA (2005) J Org Chem 70:7902
12. (a) Girardet JL, Gunic E, Esler C, Cieslak D, Pietrzkowski Z, Wang G (2000) J Med Chem 43:3704;
    (b) Eldrup AB, Prhavc M, Brooks J, Bhat B, Prakash TP, Song Q, Bera S, Bhat N, Dande P, Cook PD, Bennett CF, Carroll SS (2004) J Med Chem 47:5284;
    (c) Franchetti P, Cappellacci L, Pasqualini M, Petrelli R, Vita P, Jayaram HN, Horvath Z, Szekeres T, Grifantini M, (2005) J Med Chem 48:4983
13. (a) Babu BR, Keinicke L, Petersen M, Nielsen C, Wengel J (2003) Org Biomol Chem 1:3514
    (b) Garcia I, Feist H, Cao R, Michalik M, Peseke K (2001) J Carbohydr Chem 20:681;
    (c) Herrera L, Feist H, Quincoces J, Michalik M, Peseke K (2003) J Carbohydr Chem 22:171;
    (d) Feist IOH, Herrera L, Michalik M, Quincocesd J, Pesekea K (2005) Carbohydr Res 340:547;
    (e) Li X, Uchiyama T, Raetz CRH, Hindsgaul O (2003) Org Lett 5:539
14. Nomura M, Sato T, Washinosu M, Tanaka M, Asao T, Shuto S, Matsuda A (2002) Tetrahedron 58:1279
15. Lena C, Mackenzie G (2006) Tetrahedron 62:9085
16. González Z, González A (2000) Carbohydr Res 329:901
17. Zhao S, Petrus L, Serianni AS (2001) Org Lett 3:3819
18. (a) Sigano DM, Peach ML, Nacro K, Choi Y, Lewin NE, Nicklaus MC, Blumberg PM, Marquez VE (2003) J Med Chem 46:1571;
    (b) Tamamura H, Bienfait B, Nacro K, Lewin NE, Blumberg PM, Marquez VE (2000) J Med Chem 43:3209
19. (a) Liu HM, Zhang F, Zhang J (2001) Carbohydr Res 334:323;
    (b) Liu HM, Zhang F, Zhang J, Li S (2003) Carbohydr Res 338:1737
20. Liu HM, Zhang F, Zou DP (2003) Chem Commun 2044
21. Koóš M, Mičová J, Steiner B, Alföldi J (2002) Tetrahedron Lett 43:5405
22. Gunic E, Girardet JL, Pietrzkowski Z, Eslerb C, Wang G (2001) Bioorgan Med Chem 9:163
23. (a) Hăkansson AE, Koshkin AA, Sørensen MD, Wengel J (2000) J Org Chem 65:5161;
    (b) Koshkin AA, Fensholdt J, Pfundheller HM, Lomholt C (2001) J Org Chem 66:8504;
    (c) Koshkin AA (2004) J Org Chem 69:3711;
    (d) Meldgaard M, Hansen FG, Wengel J (2004) J Org Chem 69:6310
24. Montembault M, Bourgougnon N, Lebreton J (2002) Tetrahedron Lett 43:8091
25. Ohrui H, Kohgo S, Kitano K, Sakata S, Kodama E, Yoshimura K, Matsuoka M, Shigeta S, Mitsuya H (2000) J Med Chem 43:4516
26. Wu T, Nauwelaerts K, Aerschot AV, Froeyen M, Lescrinier E, Herdewijn P (2006) J Org Chem 71:5423
27. Andersch J, Bols M (2001) Chem Eur J 7:3744
28. Chakraborty C, Dhavale DD (2006) Carbohydr Res 341:912
29. Ji XM, Mo J, Liu HM, Sun HP (2006) Carbohydr Res 341:2312
30. (a) Sato K, Sekiguchi T, Hozumi T, Yamazaki T, Akai S (2002) Tetrahedron Lett 43:3087;
    (b) Sato K, Miyama D, Akai S (2004) Tetrahedron Lett 45:1523
31. Rauter A, Ferreira M, Borges C, Duarte T, Piedade F, Silva M, Santos H (2000) Carbohydr Res 325:1
32. Jones NA, Nepogodiev SA, MacDonald CJ, Hughes DL, Field RA (2005) J Org Chem 70:8556

33. Kutsumura N, Nishiyama S (2005) Tetrahedron Lett 46:5707
34. Robins MJ, Doboszewski B, Timoshchuk VA, Peterson MA (2000) J Org Chem 65:2939
35. Cordeiro A, Quesada E, Bonache MC, Velázquez S, Camarasa MJ, San-Félix A (2006) J Org Chem 71:7224
36. Yu J, Zhang S, Li Z, Lu W, Cai M (2005) Bioorg Med Chem 13:353
37. Hanessian S (1997) Preparative Carbohydrate Chemistry. Marcel Dekker, New York
38. Carcano M, Vasella A (1998) Helv Chim Acta 81:889
39. Dondoni A, Fogagnolo M, Medici A, Pedrini P (1985) Tetrahedron Lett 26:5477
40. Sanki AK, Suresh CG, Falgune UD, Pathak T (2003) Org Lett 5:1285
41. Du Y, Linhardt RJ, Vlahov IR (1998) Tetrahedron 54:9913
42. Horiuchi T, Ohta T, Shirakawa E, Nozaki K, Takaya H (1997) J Org Chem 62:4285
43. Kollar L, Sandor P (1993) J Organomet Chem 445:257 and references therein
44. Seepersaud M, Kettunen M, Abu-Surrah AS, Voelter W, Al-Abed Y (2002)Tetrahedron Lett 43:8607
45. Fernandez E, Ruiz A, Claver C, Castillon S (1998) Organometallics 17:2857
46. Sridhar PR, Ashalu KC, Chandrasekaran S (2004) Org Lett 6:1777
47. Li CJ (1993) Chem Rev, 93:2023 and references therein
48. (a) Li CJ, Chan TH (1999) Tetrahedron 55:11149; (b) Chan TH, Isaac MB (1996) Pure Appl Chem 68:919
49. (a) Canac Y, Levoirier E, Lubineau A (2001) J Org Chem 66:3206
(b) Lubineau A, Canac Y, Goff N (2002) Adv Synth Catal 344:319;
(c) Levoirier E, Canac Y, Norsikian S, Lubineau A (2004) Carbohydr Res 339:2737
50. Prandi J (2001) Carbohydr Res 332:241
51. Ichikawa S, Minakawa N, Shuto S, Tanaka M, Sasaki T, Matsuda A (2006) Org Biomol Chem 4:1284
52. Yu M, Lynch V, Pagenkopf BL (2001) Org Lett 3:2563
53. Beyer J, Skaanderup PR, Madsen R (2000) J Am Chem Soc 122:9575
54. Norsikian S, Lubineau A (2005) Org Biomol Chem 3:4089
55. Fontana G, Lubineau A, Scherrmann MC (2005) Org Biomol Chem 3:1375
56. Lawrence AJ, Pavey JBJ, Chan MY, Fairhurst RA, Collingwood SP, Fisher J, Cosstick R, O'Neil IA (1997) J Chem Soc Perkin Trans 1:2761
57. Kittaka A, Tanaka H, Odanaka Y, Ohnuki K, Yamaguchi K, Miyasaka T (1994) J Org Chem 59:3636
58. Lim J, Choo DJ, Kim YH (2000) Chem Commun 553
59. (a) Kodama T, Shuto S, Nomura M, Matsuda A (2001) Chem Eur J 7:2332;
(b) Sukeda M, Shuto S, Sugimoto I, Ichikawa S, Matsuda A (2000) J Org Chem 65:8988
60. Sukeda M, Ichikawa S, Matsuda A, Shuto S (2003) J Org Chem 68:3465
61. (a) Ogalino J, Mizunuma H, Kumamoto H, Takeda S, Haraguchi K, Nakamura KT, Sugiyama H, Tanaka H (2005) J Org Chem 70:1684;
(b) Kumamoto H, Shindoh S, Tanaka H, Itoh Y, Haraguchi K, Gen E, Kittaka A, Miyasaka T, Kondo M, Nakamura KT (2000) Tetrahedron 56:5363
62. Keck GE, Enholm EJ, Yates BE, Wiley MR (1985) Tetrahedron 41:4079
63. Gomez AM, Casillas M, Valverde S, Lopez JC (2001) Tetrahedron Asymmetry 12:217
64. (a) Padwa A, Weingarten MD (1996) Chem Rev 96:223;
(b) Doyle MP, Forbes DC (1998) Chem Rev 98:911
65. Choe SWT, Jung ME (2000) Carbohydr Res 329:731
66. Baptiste LJ, Yemets S, Legay R, Lequeux T (2006) J Org Chem 71:2352
67. (a) Linker T, Hartmann K, Sommermann T, Scheutzow D, Ruckdeschel E (1996) Angew Chem Int Ed Engl 35:1730;
(b) Linker T, Sommermann T, Kahlenberg F (1997) J Am Chem Soc 119:9377
68. (a) Gyollai V, Schanzenbach D, Somsak L, Linker T (2002) Chem Commun 1294;
(b) Sommermann T, Kim BG, Peters K, Peters EM, Linker T (2004) Chem Commun 2624
69. Mikkelsen LM, Skrydstrup T (2002) J Org Chem 68:2123
70. (a) Shao HW, Ekthawatchai S, Chen CS, Wu SH, Zou W (2005) J Org Chem 70:4726;
(b) Shao HW, Ekthawatchai S, Wu SH, Zou W (2004) Org Lett 6:3497

71. Ana M, Enrique M, Bernardo H (2004) Eur J Org Chem 3063
72. (a) Krohn K, Gehle D, Flörke U (2005) Eur J Org Chem 2841;
(b) Krohn K, Gehle D, Flörke U (2005) Eur J Org Chem 4557
73. Ghosh R, Chakraborty A, Maitia DK, Puranikb VG (2005) Tetrahedron Lett 46:8047
74. Hotchkiss DJ, Jenkinson SF, Storer R, Heinz T, Fleet GWJ (2006) Tetrahedron Lett 47:315
75. Hodge JE, Fisher BE (1963) Methods in Carbohydrate Chemistry, vol II. Academic Press, New York
76. Hou Y, Wu X, Xie W, Braunschweiger PG, Wang PG (2001) Tetrahedron Lett 42:825
77. Hotchkiss DJ, Soengas R, Booth KV, Weymouth-Wilson AC, Eastwich-Field V, Fleet GWJ (2007) Tetrahedron Lett 48:517
78. (a) Schmidt RR, Maier M, Tetrahedron Lett (1985) 26:2065;
(b) Schmidt RR, Haag ZB, Hoch M (1988) Liebigs Ann Chem 885;
(c) Dujardin G, Molato S, Brown E (1993) Tetrahedron Asymmetr 4:193
(d) Tietze LF, Montenbruck A, Schneider C (1994) Synlett 509;
(e) Tietze LF, Schneider C, Grote A (1994) Chem Eur J 2:139;
(f) Dujardin G, Rossignol S, Brown E (1998) Synthesis 763
79. (a) Evans DA, Johnson JS (1998) J Am Chem Soc 120:4895;
(b) Thorhauge J, Johannsen M, Jørgensen KA (1998) Angew Chem Int Ed 37:2404;
(c) Evans DA, Olhava EJ, Johnson JS, Janey JM (1998) Angew Chem Int Ed Engl 37:3372;
(d) Evans DA, Johnson JS, Olhava EJ, Janey JM (2000) J Am Chem Soc 122:1635
80. Audrain H, Thorhauge J, Hazell RG, Jørgensen KA (2000) J Org Chem 65:4487
81. (a) Robins MJ, Sarker S, Xie M, Zhang W, Peterson MA (1996) Tetrahedron Lett 37:3921;
(b) Peterson MA, Nilsson BL, Sarker S, Doboszewski B, Zhang W, Robins MJ (1999) J Org Chem 64:8183;
(c) Robins MJ, Doboszewski B, Timoshchuk VA, Peterson MA (2000) J Org Chem 65:2939
82. (a) Rozners E, Qun X (2003) Org Lett 5:3999;
(b) Rozners E, Liu Y (2003) Org Lett 5:181;
(c) Rozners E, Liu Y (2003) J Org Chem 70: 9841;
(d) Qun X, Rozners E (2005) Org Lett 7:2821
83. Fraser-Reid B, Chen XT, Haag D, Henry KJ, McPhail AT (2000) Chirality 12:488
84. Mehta G, Pallavi K (2004) Tetrahedron Lett 45:3865
85. Sagar R, Pant CS, Pathak R, Shaw AK (2004) Tetrahedron 60:11399
86. Fessner W-D, Goße C, Jaeschke G, Eyrisch O (2000) Eur J Org Chem 125
87. Chattopadhyay A, Goswami D, Dhotare B (2006) Tetrahedron Lett 47:4701
88. Haraguchi K, Takeda S, Tanaka H (2003) Org Lett 5:1399

2.6 C=C Bond Formation

Sławomir Jarosz, Marcin Nowogródzki
Institute of Organic Chemistry, Polish Academy of Sciences,
01–224 Warsaw, Poland
sljar@icho.edu.pl

Abstract

The material presented in this chapter describes the general methodology used for the preparation of unsaturated sugars. The 'older' methods (i. e. those being developed since at least the 1950s) which are still very useful and have general application are also presented but they are illustrated by newer examples. The direct formation of the double bond(s) is emphasized, but the methodology based on the rearrangement of unsaturated sugars into other olefinic carbohydrates is also reviewed.
The emphasis is placed on the general methods rather than synthesis of individuals, since this should give the reader a general view of the importance of unsaturated sugars and their application to stereocontrolled organic synthesis. All main classes of unsaturated sugars are described briefly. The anomeric unsaturated derivatives are excluded from this review unless they serve as starting materials in the preparation of other unsaturated carbohydrates via rearrangement reactions.

Keywords

Synthesis; Rearrangement; Unsaturated sugars; Sugar chirons; Targeted synthesis

In: *Glycoscience*. Fraser-Reid B, Tatsuta K, Thiem J (eds)
Chapter-DOI 10-1007/978-3-540-30429-6_8:

Abbreviations

DBU 1,8-diazabicyclo[5.4.0]undec-7-ene
DMSO dimethyl sulfoxide
DMF dimethylformamide
HCMV human cytomegalovirus
NHDF normal human fibroblasts
PTC phase transfer catalysis
RCM ring-closing metathesis
TIBAL tri-*iso*-butyl aluminum

1 Introduction

Sugar derivatives containing a double bond are an important class of compounds often used in stereocontrolled synthesis of optically pure targets. Their chemistry: preparation, reactions, and application in targeted synthesis is well developed and described in a number of monographs [1,2,3,4,5]. Therefore the classical methods, which have been introduced to the synthetic chemistry of sugars since (at least) the 1950s and are still very useful for their preparation will be presented in this report, which is an update of an earlier report [6]. The general methodology presented here will refer to the previous report rather than to original literature, but will be illustrated by more recent examples.

Only occasionally are unsaturated sugar derivatives such as blasticidin A (**1**), which inhibits blast disease in rice, found in nature. However, the non-natural unsaturated sugars often possess potent biological activity. The most representative example is, undoubtedly, 2,′3′-dideoxy-2,′3′-didehydro-thymidine d4T (Stavudine, Cerit®)—approved as the fourth *anti*-HIV drug by the US Food and Drug Administration—which exhibits an effect comparable to AZT in HIV-infected CEM cells (in vivo), is less toxic than AZT (the first drug used in an *anti*-HIV treatment) for bone marrow stem cells and was found to be less inhibitory to mitochondrial DNA replication [6,7]. Another example, neuraminic acid derivative **2**, is the antisialyl-inhibitory anti-influenza drug [1]. It is also known that the L-nucleosides are important components of antiviral agents; for example, L-2′3′-dideoxy-2′3′-didehydro-5-fluorocytydine (L-Fd4C) exhibits potent anti-HBV (antihepatitis B virus) activity (*Fig. 1*) [8,9].

Generally two main classes of unsaturated sugars in which (1) the anomeric carbon atom is involved in a double bond (glycals) and (2) the unsaturation is placed between other carbon atoms are known. One has to discuss also derivatives with the *exo*- and *endo*-unsaturated bonds (either olefins/dienes or acetylenes). Another class of unsaturated monosaccharides is represented by the open-chain sugars with the double bond(s) and/or triple bond present in the molecule. The examples of different types of unsaturated sugars are shown in *Fig. 2*.

Compounds such as glycals will not be discussed here, unless they are used as substrates for the preparation of other, non-anomeric, unsaturated derivatives.

Unsaturated sugar chirons are versatile starting materials in organic synthesis. Many complex natural products whose structure incorporates five- and six-membered carbocyclic rings

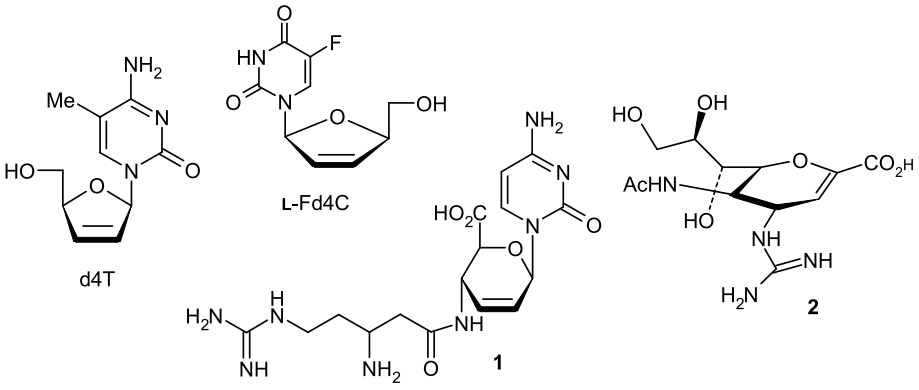


Figure 1
Examples of unsaturated sugars with potent biological activity

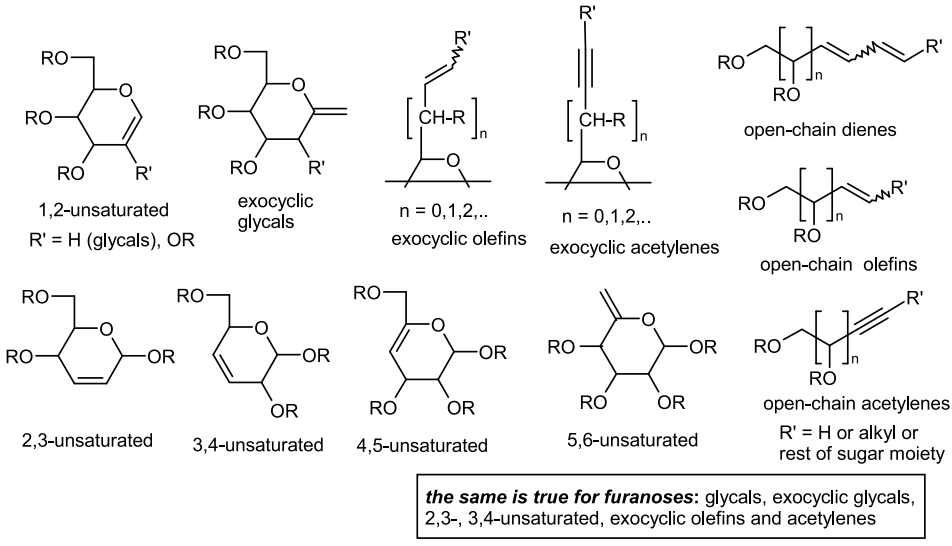


Figure 2
Different types of unsaturated sugars

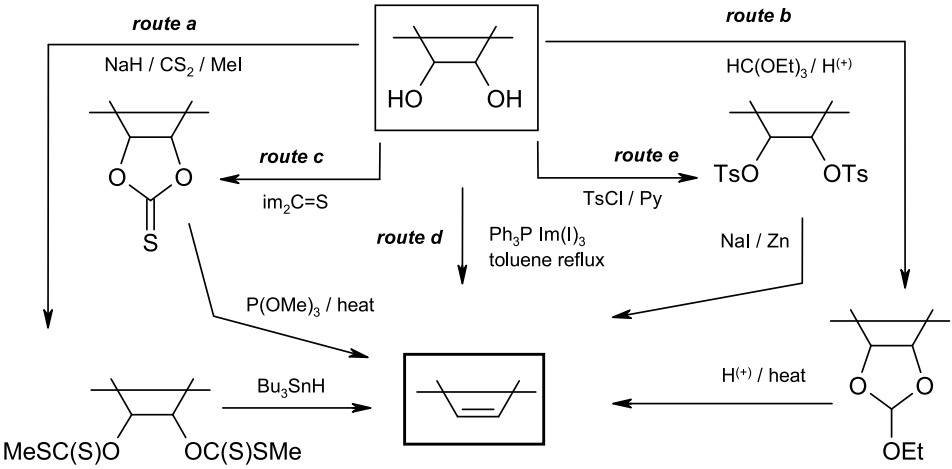
have been synthesized from such readily available substrates [5,10]. The older methods, still used with great success in the synthesis of complex modified unsaturated carbohydrates, may be found in an excellent review by Ferrier [11]. For the newer ones see the recent reviews [1,2,4,6].

2 General Methods of the Formation of the Double Bond

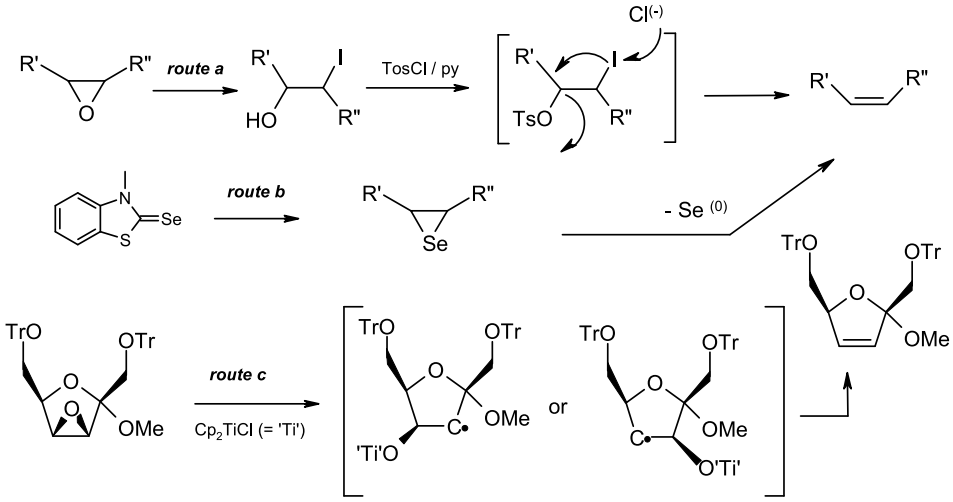
The double bond may be introduced into the carbohydrate skeleton in several different ways. The methodology of the preparation of unsaturated sugars is well developed and almost any derivative can be prepared using classical methods described in a number of monographs and text-books. These methods include reduction of oxygen functions, elimination reactions, reactions of carbonyl compounds with the Wittig-type reagents, and rearrangements of a sugar skeleton (e. g. preparation of 2,3-unsaturated sugars from glycals). Carbohydrates are polyhydroxylated compounds and conversion of the *vic*-diol function into an olefin is the simplest and most obvious choice. Several different methods shown in *Scheme 1* were elaborated for such purposes. Direct conversion is achieved by treatment of a diol with triphenyl-phosphine and tri-iodoimidazole in refluxing toluene (Garegg's procedure, *route d*) [12]. Other common preparations require first activation of a diol followed by further transformation of such active intermediate(s) into olefins.

The di-*O*-tosylates (prepared by action of tosyl chloride in pyridine) are reduced with zinc (NaI/Zn; *route e*; Tipson–Cohen reaction) [13]. Cyclic ortho-esters (prepared by reaction of the diol with ethyl orthoformate) are transformed into olefins by simple heating in the presence of acids (Eastwood reaction, *route b*) [14]. Cyclic thiocarbonates (obtained by reaction of a diol with thiophosgene or *N*,*N*′-thiocarbonyl-di-imidazole) are reduced to olefin with trimethyl phosphite (Corey–Winter method, *route c*) [15]. Finally, reduction of vicinal di-xanthates with tri-*n*-butyltin hydride according to the Barton procedure [16] affords olefins via a reductive elimination process (*route a*). The Corey–Winter, Garegg, and Tipson–Cohen methods are most commonly applied for deoxygenation of sugar diols.

Epoxides and episulfides are also convenient precursors of unsaturated sugars. The first step involves nucleophilic opening of the three-membered ring to afford the intermediate from which the desired unsaturated derivative is obtained [1]. For example, conversion of the



Scheme 1



Scheme 2

epoxide into iodohydrin and subsequent reaction with pyridine\tosyl chloride affords appropriate olefinic sugars (*route a* in *Scheme 2*). Reduction of oxiranes with selenium reagents, such as 3-methyl-2-selenoxo-1,3-benzothiazole, provides olefins in good yields (*route b*). Bis(cyclopentadienyl) titanium chloride induces a radical deoxygenation of epoxides; this titanium species is an extremely mild reducing agent, which may be illustrated by the fact that the olefin (obtained from the corresponding epoxide according to a mechanism shown in *route c*) could be isolated in good yield, although it is so unstable that even traces of acidic impurities present in $CDCl_3$ may cause its aromatization [6].

All these already presented general methods may be used for the construction of either *exo*- or *endo*-cyclic double bonds in a carbohydrate skeleton.

There are also special procedures which allow the preparation of sugars with only *endo*- or only *exo*-cyclic double bonds. The first group of compounds with *endo*-cyclic double bond may be prepared by total synthesis from non-carbohydrate precursors. The particularly useful hetero Diels–Alder reaction (*Fig. 3*) allows one to obtain the dihydropyran skeleton either by reaction of a diene with a heterodienophile [3,17] or by reaction of a heterodiene with a 'normal' dienophile [18].

The approach proposed first by Zamojski was based on the thermal reaction of 1-methoxy-1,3-butadiene with a highly active heterodienophile (butyl glyoxylate), which provided a dihydropyran derivative—precursor of racemic monosaccharides [3,17].

Several years later Danishefsky introduced 1-methoxy-3-trimethylsilyloxy-1,3-butadiene, a highly reactive diene, which upon reaction with (not activated) aldehydes catalyzed with mild Lewis acids [$(Eu(fod)_3$] afforded cyclic α,β-unsaturated ketones [19]. Another method involved reaction of 1,4-di-alkoxy(acyloxy)-butadiene with an activated heterodienophile, which led to more functionalized derivatives (Schmidt) [20]. This methodology may be illustrated by the 'classical' synthesis of the precursor of purpurosamine B (**3**) and higher sugar

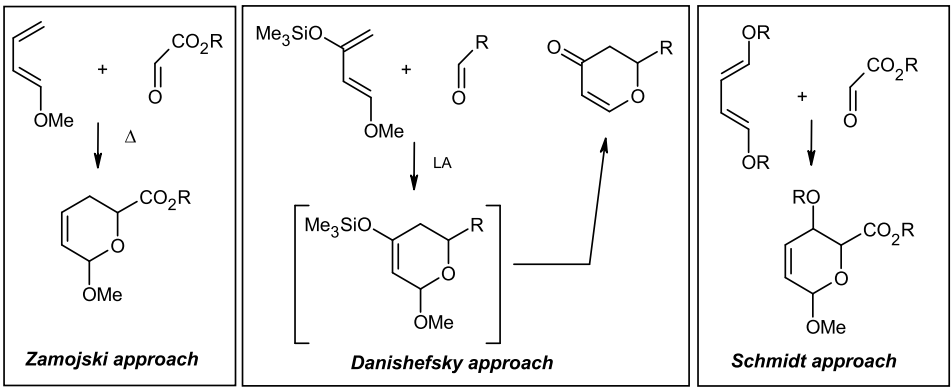
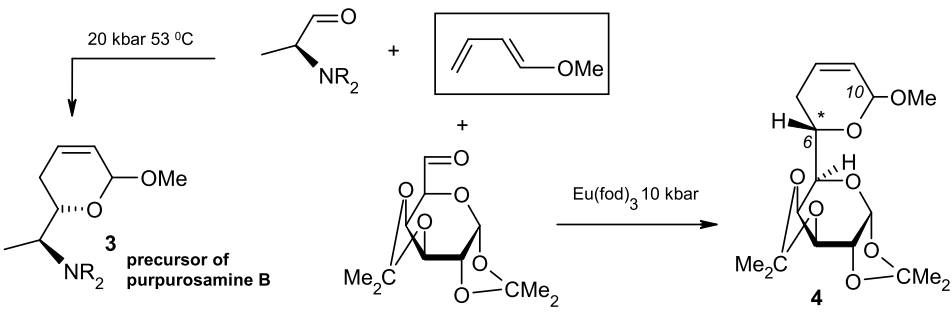


Figure 3
Synthesis of unsaturated sugars by a hetero Diels–Alder approach



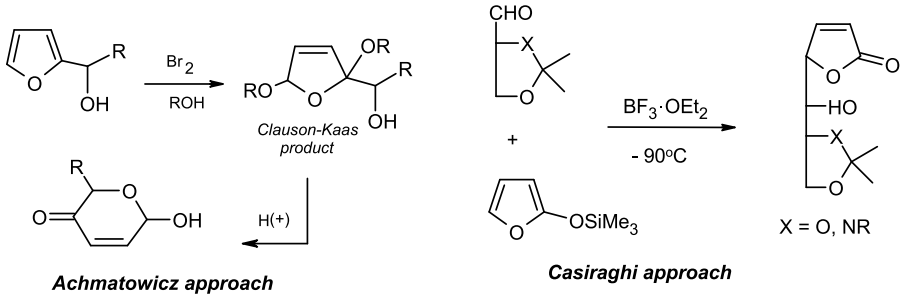
Scheme 3

dialdose **4** (*Scheme 3*). The second reaction performed under 20 kbar pressure afforded the target with almost 100% diastereoselectivity. It was possible to lower the pressure (to 10 kbar) if the reaction was additionally catalyzed with mild Lewis acid [$Eu(fod)_3$] [6].

This process (hetero Diels–Alder reaction leading to a dihydropyran system) may be also conducted in an asymmetric version; application of chiral transition-metal catalysts based on BINOL, BINAP, bisoxazolines, etc. provides adducts in very high optical purity (*ee* up to 99%) [1,6]. In a series of papers Jurczak reported recently a highly enantioselective cycloaddition of 1-methoxy-1,3-butadiene and butyl glyoxylate catalyzed with chiral salen complexes [21].

The convenient precursors of unsaturated sugars, dihydropyran or dihydrofuran derivatives, can be also obtained from an appropriately functionalized furan molecule. Two main approaches are used. The first is based on the oxidative rearrangement of furfuryl alcohols (Achmatowicz reaction) [22], the second one involves an acid-catalyzed reaction of aldehydes with 2-trimethyl-silyloxyfuran (Casiraghi reaction) [23] (*Scheme 4*).

In the original Achmatowicz approach, the (racemic) furfuryl carbinol is oxidized with bromine in the presence of methanol under weakly basic conditions to provide the Clauson-



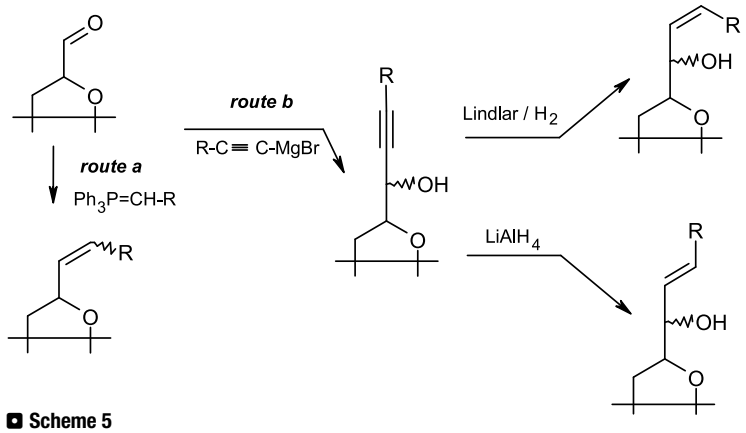
Scheme 4

Kaas product, which is further hydrolyzed under rigorously controlled acidic conditions into the ald-2-enos-4-ulose derivative. Many other modifications [24] of the original Achmatowicz procedure such as oxidation of the furan ring with: *m*-CPBA, PCC, NBS, ^{t}BuOOH\VO(OAc)$_2$, or singlet oxygen are also used for this transformation [6].

Application of the Sharpless procedure for the kinetic resolution of racemic allyl alcohols [Ti(O*i*Pr)$_4$\$^{t}$BuOOH and L- or D-DET] to such a process provided an optically active dihydropyran together with enantiomerically pure unreacted furylcarbinol [6]. The Casiraghi approach leads to 2,3-unsaturated furanoses (or amino furanoses) by an acid-catalyzed reaction of 2-(trimethylsilyloxy)furan with sugar aldehydes or aminoaldehydes.

For introduction of the exocyclic double bond two methodologies are particularly useful. The first one is based on the reaction of the aldehydes (ketones) with the Wittig reagents (*route a*) [6]. In the second one (*route b*), an acetylenic functionality is introduced [25], which may be further converted either into the *Z*- or *E*-olefins by proper reducing reagents (H_2/Lindlar catalyst or e. g. $LiAlH_4$, respectively) (*Scheme 5*) [6].

All these general methods presented here will be illustrated by proper examples in the next sections of this chapter.



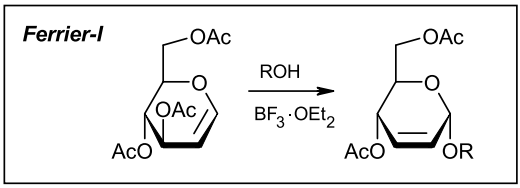
Scheme 5

3 Monosaccharides with the Endocyclic Double Bond

In this section the selected and most representative examples (including also several older ones) of the preparation of unsaturated monosaccharides with the double bond present in the sugar ring will be described. The material will be divided into three main classes: 2,3-, 3,4-, and 4,5-unsaturated sugars in both pyranose and furanose form.

3.1 2,3-Unsaturated Monosaccharides

One of the most convenient methods for the synthesis of 2,3-unsaturated sugars is the so-called Ferrier-I rearrangement of glycals (*Scheme 6*). This reaction allows one to obtain the unsaturated targets via the process in which the double bond migrates from the C1–C2 to the C2–C3 position [1,2,4,5,11].

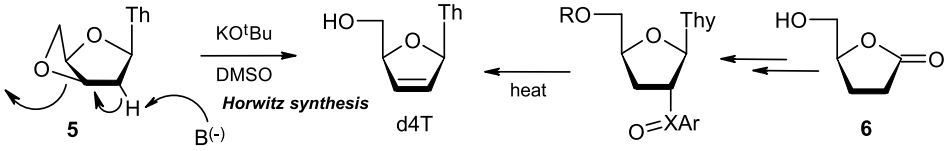


Scheme 6

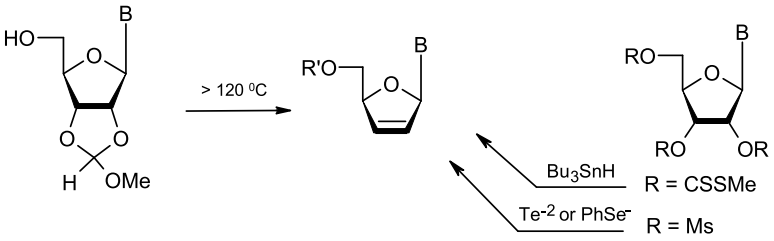
It is not concerned, therefore, with the formation of the double bond and this process will be described in more detail in *Sect. 3.4* (*rearrangement*).

In this section only the direct synthesis of 2,3-unsaturated sugars, in which the double bond is created will be presented. Numerous syntheses of these valuable compounds have been described in the literature. One of the most important classes of such derivatives are 2,′3′-dideoxy-2,′3′-didehydronucleosides which possess a broad spectrum of biological activity. Two different approaches to 2,′3′-dideoxy-2,′3′-didehydro-thymidine (d4T) illustrate the general methodology of their preparation.

The first synthesis of d4T was accomplished in 1966 by Horwitz from 1-(2-deoxy-3,5-epoxy-β-D-*threo*-pentosyl)thymine; abstraction of a proton from the 2′-position by potassium ^{t}butoxide in DMSO resulted in an opening of the oxetane ring and formation of d4T. Other syntheses were based on the elimination of sulfoxide or the selenoxide moiety placed at the C-2′ position in the nucleoside, which was prepared from cheap lactone **6** (*Scheme 7*) [6].



Scheme 7



Scheme 8

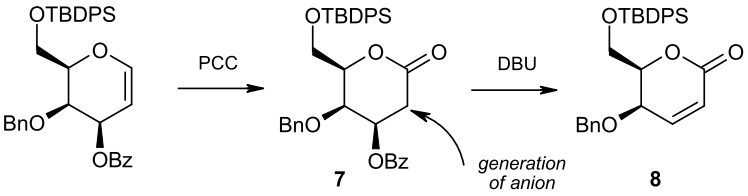
The 2,′3′-double bond in unsaturated nucleosides may be created also by thermal degradation of *ortho*-esters, reduction of appropriate di-xanthates with tri-*n*-butyltin hydride [26], or reductive elimination of dimesylates caused by telluride or selenide anions (*Scheme 8*) [6,27].

The β-elimination reaction is a useful method for the preparation of unsaturated sugars. For example, treatment of the sugar lactone **7** (readily available from the corresponding glycal by the Ferrier-I rearrangement followed by standard transformations) with a strong base generates the anion from which the benzoate is eliminated providing the 2,3-unsaturated derivative **8** (*Scheme 9*) [28].

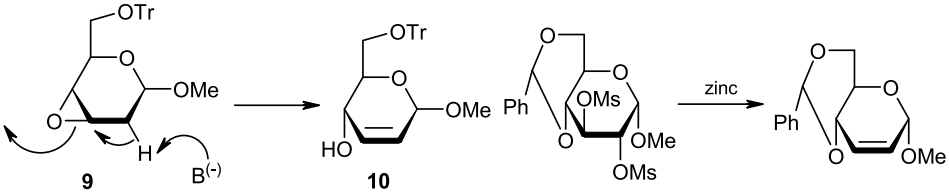
Base-catalyzed isomerization of epoxides into allylic alcohols is a method of choice for the preparation of unsaturated sugars; olefin **10** was prepared by treatment of oxirane **9** with a strong base. This process is similar to the Horwitz synthesis of d4T presented in *Scheme 7*.

The Tipson–Cohen methodology is frequently applied for the synthesis of 2,3-unsaturated sugars as shown in *Scheme 10*; such reduction might be performed also under microwave irradiation [6].

Formation of 2,3-unsaturated sugars was illustrated by the representative examples. Application of any other general methods which are described in *Sect. 2* should be considered when planning the synthesis of such derivatives.



Scheme 9



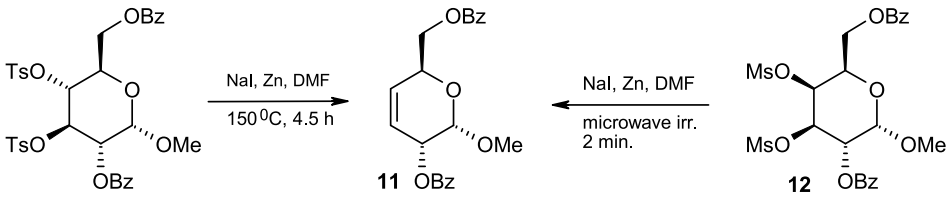
Scheme 10

3.2 3,4-Unsaturated Monosaccharides

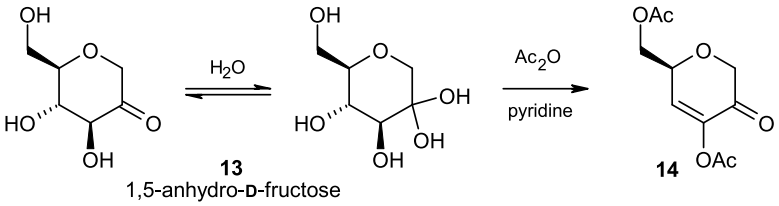
Synthesis of 3,4-dideoxy-3,4-didehydro-sugars can be conveniently performed by the Cohen–Tipson reduction of appropriate sulfonates. For example, methyl 3,4-dideoxy-3,4-didehydro-α-D-*erythro*-hex-3-enopyranoside is obtained from methyl α-D-glucopyranoside by selective protection of the 2- and 6-OH groups as benzoyl esters, subsequent tosylation (or mesylation) of the remaining 3,4-hydroxyl groups, and reduction of such di-sulfonic ester with zinc in refluxing DMF. The same target can also be obtained from methyl 2,6-di-*O*-benzoyl-3,4-di-*O*-mesyl-β-D-galactopyranoside; the latter reaction was performed under microwave irradiation to facilitate the process (*Scheme 11*) [6].

Another example of application of the elimination process in the synthesis of these useful derivatives is shown in *Scheme 12*. Derivative **14** was obtained in high yield from 1,5-anhydro-D-fructose **13** via regioselective elimination of acetic acid [29].

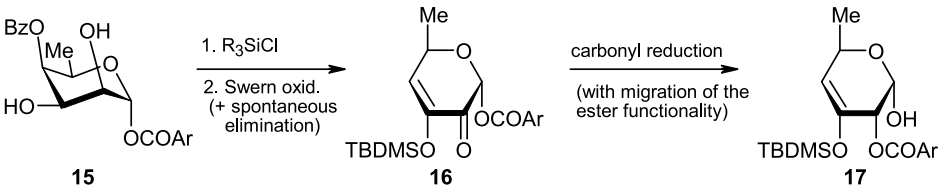
Further illustration of this useful process is provided in *Scheme 13*. It was based on a different reactivity of the hydroxyl group in carbohydrates. Selective protection of the 3-OH in **15** leaves the 2-OH free. Oxidation of this hydroxyl with the Swern reagent affords a ketone from which the benzoic acid is eliminated leading to the 2,3-unsaturated compound **16**. Reduction of the carbonyl group is followed by spontaneous migration of the ester group from C-1, thus affording a free sugar **17** [1].



Scheme 11



Scheme 12



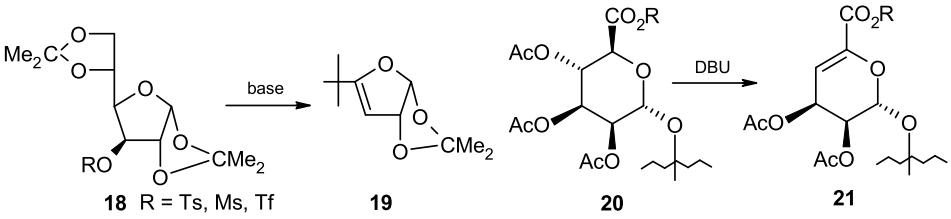
Scheme 13

3.3 4,5-Unsaturated Monosaccharides

The 4,5-unsaturated pyranoses or 3,4-unsaturated furanoses are usually prepared by a base-catalyzed elimination of a leaving group such as halogen, sulfonyloxy etc. from appropriate sugar derivatives. The classical example is represented by the synthesis of compound **19** from 1,2:5,6-di-*O*-isopropylidene-α-D-glucofuranose **18** [1]. Analogous elimination performed for **20** led to derivative **21** from which capuramycin—a complex nucleoside antibiotic could be prepared readily (● *Scheme 14*) [30].

Elimination of the substituent from the β-position with respect to the electron-withdrawing group is often observed and leads to appropriate unsaturated products. Some natural glycopyranuronate conjugates with the characteristic 4-*O*-glycosidic bond are degraded to 4,5-unsaturated carbohydrates with liberation of sugar(s) sub-unit(s) by enzymes or in a proton-poor media (● *Fig. 4*) [6].

An example of chemically induced β-elimination is shown in ● *Scheme 15*. Treatment of the 3,4-*O*-isopropylidene-D-galactoside with a strong base induced the elimination of acetone with formation of the 3,4-unsaturated sugar [31].



Scheme 14

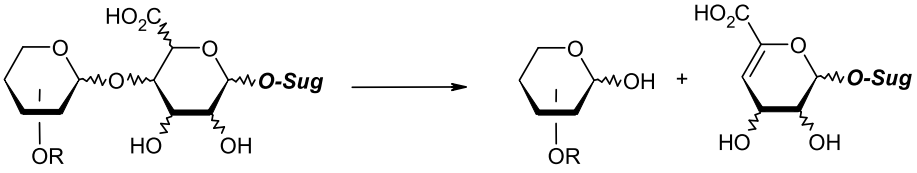
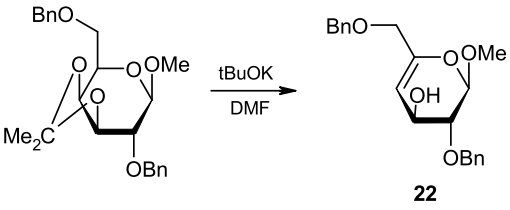


Figure 4
Degradation of oligo- and polysaccharides by enzymes in proton-poor media



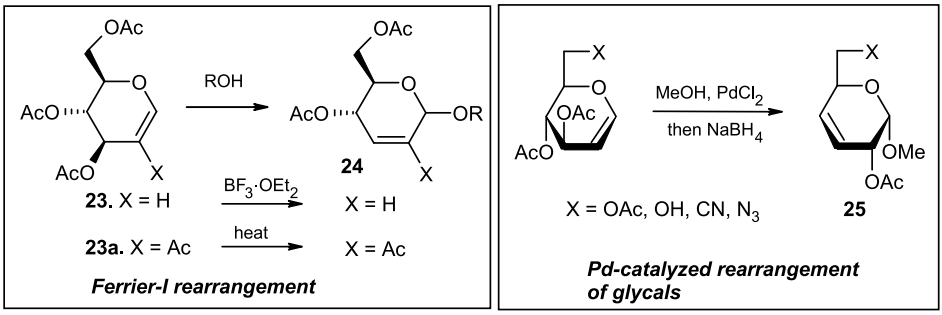
Scheme 15

3.4 Rearrangement Reactions of Unsaturated Sugars

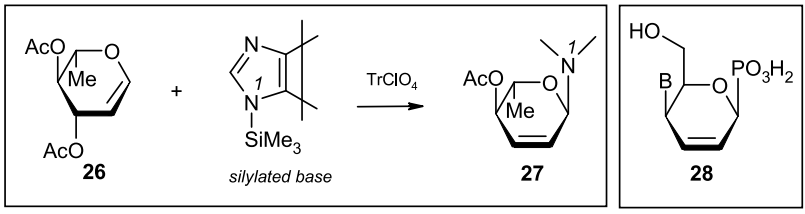
Rearrangements of unsaturated sugars are not concerned with a *direct formation* of a C=C bond, but with a migration of a double bond from one position to another. Surely, the most useful process of this type is the so-called Ferrier rearrangement [3,32], i. e. reaction of 1,2-unsaturated pyranoses or furanoses with nucleophiles in the presence of Lewis acids leading to 2,3-unsaturated sugars. In the original procedure, tri-*O*-acetyl-D-glucal (**23**; X = H) was treated with ethanol in the presence of boron trifluoride etherate which effected the addition of an alcohol at the C1 center with a simultaneous shift of the C1–C2 double bond to the C2–C3 position and elimination of acetic acid providing **24** (X = H). Other variants of this reaction, such as the Pd-catalyzed process [33] or microwave irradiation [34] allow one to obtain the desired products under milder reaction conditions. The 2-acyloxyglycals **23a** react with alcohols at elevated temperature to afford 2-substituted products **24** (X = Ac) [35]. When glycals are treated with alcohols in the presence of palladium chloride the corresponding 3,4-unsaturated sugars **25** are formed in good yield (› *Scheme 16*) [36].

The Ferrier-I rearrangement represents now a classical methodology which should be considered in planning the syntheses of optically pure targets from sugar chirons. Such a conclusion may be illustrated by the preparation of pyranosyl nucleosides obtained by this procedure.

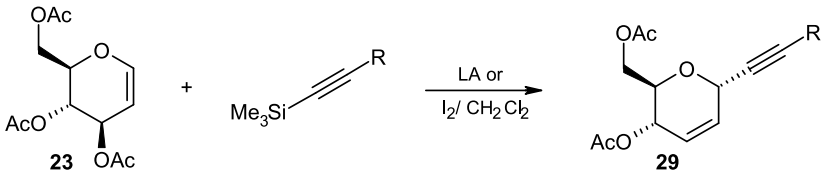
For example, treatment of di-*O*-acetyl-L-rhamnal **26** with silylated pyrimidine or purine bases in the presence of a mild acidic catalyst (trityl perchlorate) afforded 2,3-unsaturated pyranosyl nucleosides, as a mixture of α,β-anomers **27** (with the β-anomer predominating) [37]



Scheme 16



Scheme 17

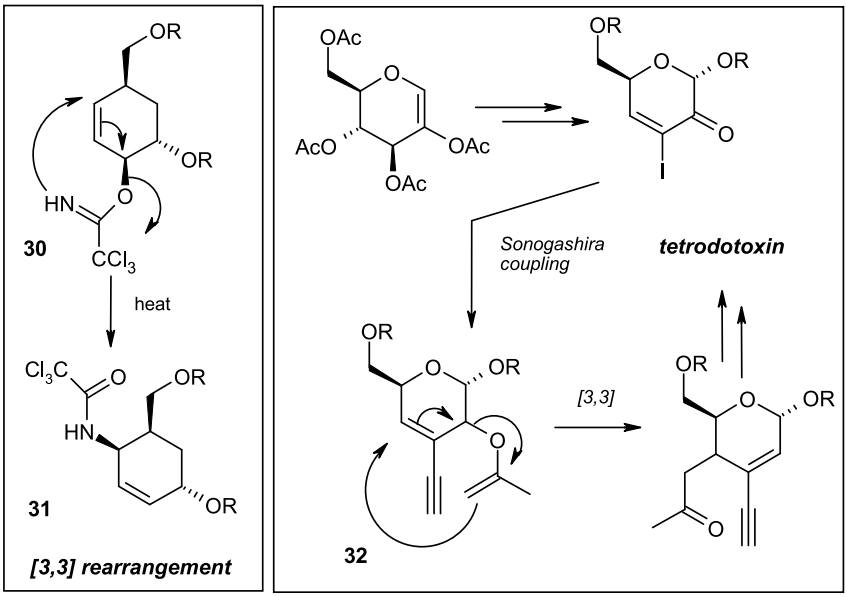


Scheme 18

(*Scheme 17*). In the synthesis of phosphorylated pyranosyl nucleosides (which inhibit replication of HCMV in NHDF cells) such as **28**, the key step [38] consisted of the Ferrier-type coupling of glycals with alkyl phosphites catalyzed with boron trifluoride etherate.

Ferrier reaction is applicable also to many other nucleophiles (thiols, amines), which allow one to prepare a wide variety of useful 2,3-unsaturated products, many of which—especially 2,3-unsaturated furanoses—possess interesting biological properties. Introduction of the diphenylphosphine oxide group by Ferrier rearrangement of glycals was reported [39]. A particularly interesting process is represented by reaction of glycals with *C*-nucleophiles (such as e. g. allylsilanes, silylacetylenes, etc.), which directly provides the 2,3-unsaturated *C*-glycosides [1,2]. For example, addition of acetylenic species to glycal **23** (with its rearrangement) provides the 2,3-unsaturated-1-*C*-sugar acetylenes **29** in good yields (*Scheme 18*) [40].

Another useful process applied in the synthesis of target unsaturated carbohydrates are sigmatropic rearrangements as for example the [3,3] rearrangement of **30** (prepared from appropriate alcohol and trichloroacetonitrile) leading to the amine **31** [41]. The [3,3] rearrangement of



Scheme 19

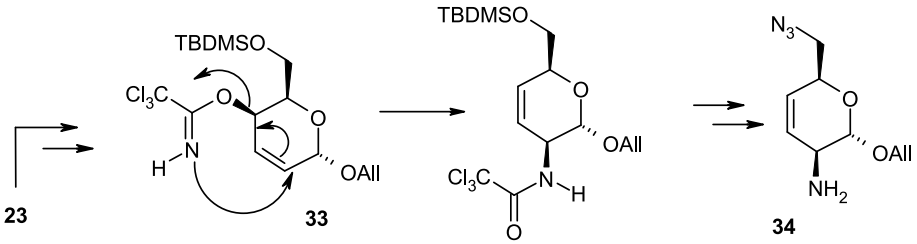
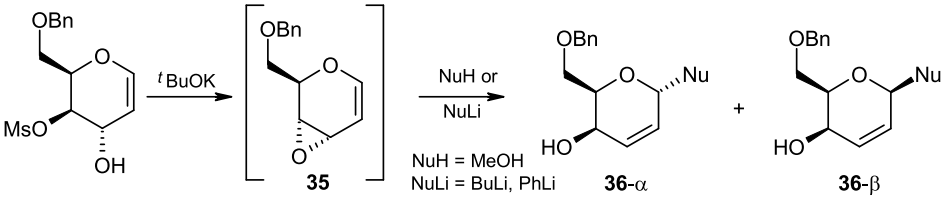


Figure 5
Synthesis of 3,4-unsaturated aminosugars via rearrangement of 2,3-unsaturated sugars



Scheme 20

unsaturated compound **32** was a key step in the synthesis of tertrodotoxin (a toxic principle of puffer fish poisoning) realized recently by Isobe (*Scheme 19*) [42].

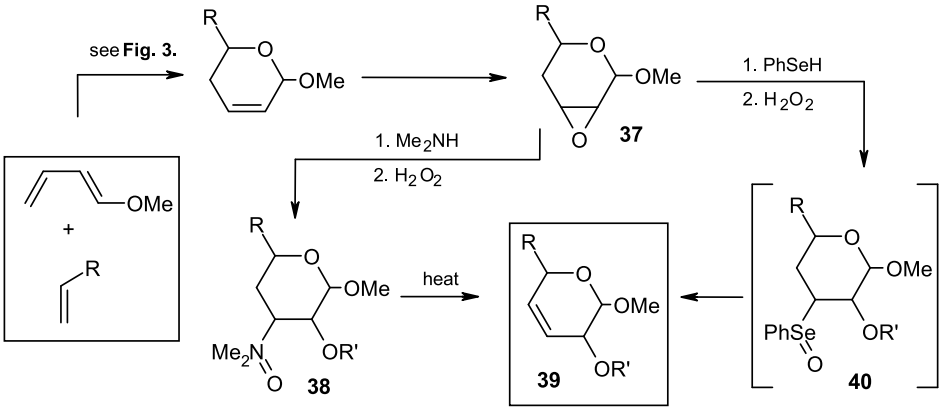
The key-step in the synthesis of a unsaturated sugar **34** bearing two amino groups (a scaffold which were used for the preparation of peptides with aromatic rings) was realized by the rearrangement of the adduct **33** readily prepared from D-glucal **23** (*Fig. 5*) [43].

Another interesting example of the preparation of unsaturated *O*- and *C*-glycosides **36** was presented recently [44]. This was based on the S_N2' attack of the nucleophile on unsaturated epoxide **35**. Substitution at C-1 proceeded with simultaneous migration of the double bond and opening of the three-member ring providing 2,3-unsaturated derivative **36** (*Scheme 20*).

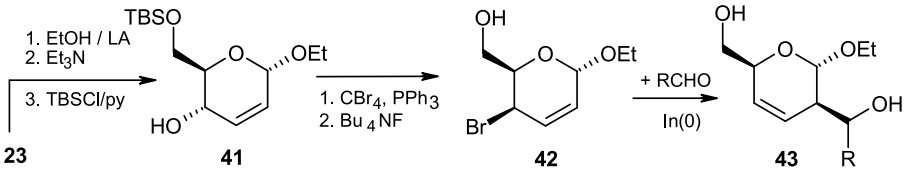
Conversion of 4-deoxy-2,3-unsaturated sugars into 3,4-unsaturated ones reported by Banaszek and Zamojski already in 1972 [3] was one of the key-steps in total synthesis of monosaccharides by a hetero Diels–Alder approach.

Epoxide **37** (either of possible stereoisomers) was converted into the appropriate aminosugar (by reaction with aqueous dimethylamine), which was further oxidized with hydrogen peroxide to the *N*-oxide **38** and transformed into the 3,4-unsaturated sugar **39** via the thermal Cope-type rearrangement. A milder procedure for the preparation of derivative **39** was proposed by David [45]; after opening of the oxirane ring with a phenylselenide anion, the resulting sugar selenide was oxidized in situ with H_2O_2 to the selenoxide **40**, which underwent a smooth conversion into allylic alcohol **39** ($R' = H$) (*Scheme 21*).

Another example of the transposition of the C2–C3 double bond into the C3–C4 position was reported by Lubineau [46]. Tri-*O*-acetyl-D-glucal **23** was converted in three standard steps into 2,3-unsaturated sugar **41**. The modified Appel reaction provided bromide **42** which reacted with aldehyde in the presence of metallic indium affording the 3,4-unsaturated product **43** (*Scheme 22*).



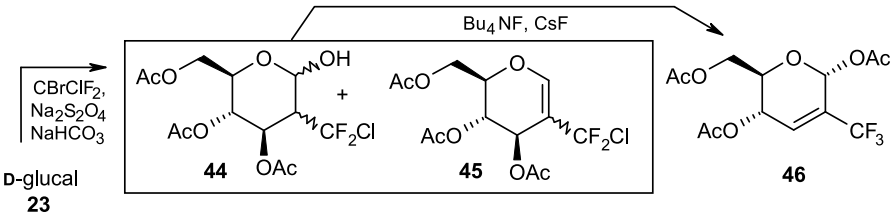
Scheme 21



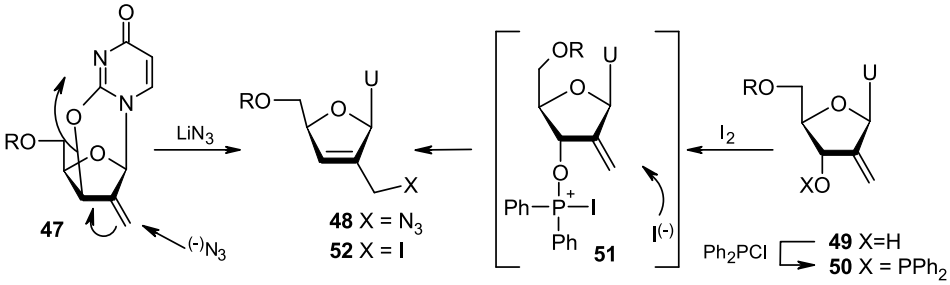
Scheme 22

Synthesis of unsaturated sugars with the strongly electron-withdrawing trifluoromethyl group was initiated also from the tri-*O*-acetyl-D-glucal (**23**). Radical addition of the difluorochloromethyl fragment to a double bond afforded two products **44** and **45**; reaction of the former with cesium fluoride furnished 2-trifluromethyl-2-eno-saccharide **46** (*Scheme 23*) [47]. The transposition of the exocyclic double bond into the C2′–C3′ position leading to unsaturated nucleosides is presented also in *Scheme 24*. The S_N2' reaction of **47** with lithium azide led to 2′-azidomethylene derivative **48**, while rearrangement of phosphinate **50** (prepared from **49** and Ph_2PCl) caused by iodine afforded **52** (via transition state **51**) [6].

The selected examples (both older and newer ones) shown in this part should emphasize to the reader the problem of planning and execution of the synthesis of unsaturated sugars via



Scheme 23

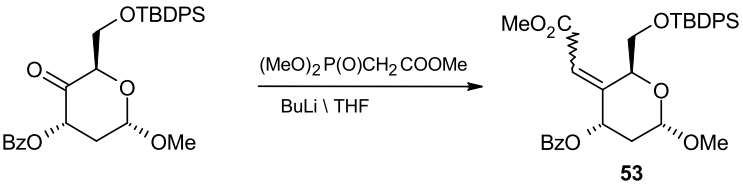


Scheme 24

rearrangement reactions. This may also be illustrated by other syntheses and the reader can find more examples in the literature.

4 Monosaccharides with the Exocyclic Double Bond(s)

The Wittig (or Wittig-type) reaction is one of the most frequently used methods for the formation of a new carbon–carbon double bond. The terminal olefinic bond is usually constructed by reaction of aldehydes with methylenetriphenylphosphorane ($Ph_3P=CH_2$) or the dialkyl methyl-phosphonate anion. Both methods are applied for the preparation of unsaturated sugars [6]. However, serious problems are sometimes encountered using $Ph_3P=CH_2$, because of the rather high basicity of this reagent (which may cause undesired rearrangements of the parent carbonyl compounds) and moderate nucleophilicity. The ylids may be replaced with more reactive and less basic reagents such as Tebbe's reagent [48] or $CH_2Br_2/Zn/TiCl_4$ [49] (these reagents convert also esters into alkoxy-olefins). For the elongation of sugars by a C_2-unit appropriate aldehydes are treated with stabilized phosphoranes ($Ph_3P=CHCOR$), which leads to the α,β-unsaturated derivatives. Sugar ketones, being less reactive than aldehydes very often do not react with stabilized phosphoranes. More nucleophilic phosphonate anions might be, therefore, applied to overcome this problem as illustrated by synthesis of *exo*cyclic unsaturated compound **53**, an intermediate in the Hanessian synthesis of thromboxane (*Scheme 25*) [6]. α-Alkoxy-α,β-unsaturated sugars are conveniently obtained by reaction of the corresponding aldehydes (ketones) with alkoxyphosphoranes [50] [$Ph_3P=C(OR)(CO_2R)$] or phosphonates [51] [$(MeO)_2P(O)CH(OR)CO_2R$].



Scheme 25

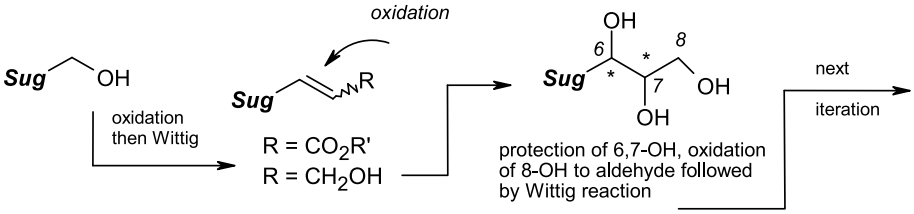
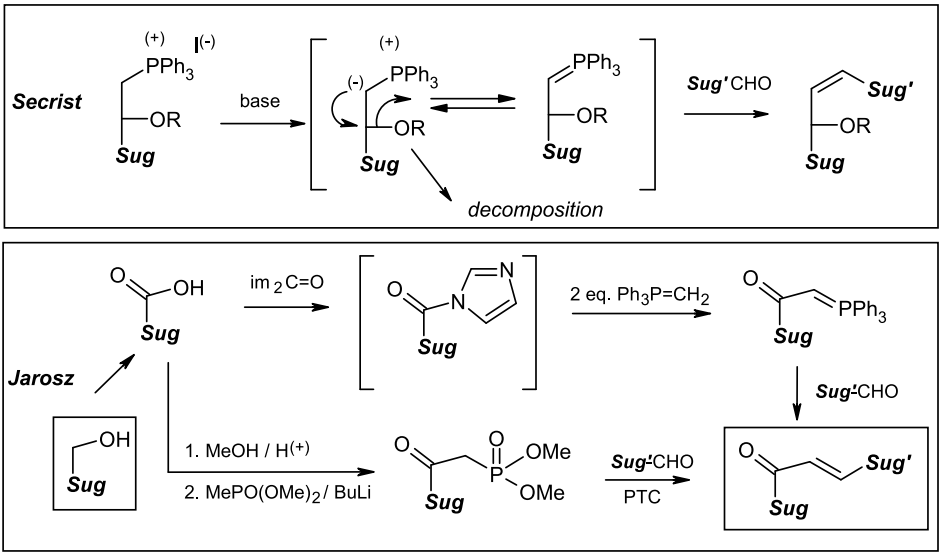


Figure 6
Brimacombe methodology of the synthesis of higher carbon sugars

The Wittig reaction is a common methodology applied in the preparation of higher carbon sugars. In the mid 1980s Brimacombe proposed a general synthesis of such compounds by a two carbon atom elongation of the parent monosaccharide, followed by functionalization of the resulting double bond [52]. By this iterative elongation Brimacombe was able to prepare decoses (*Fig. 6*).

A more convenient method for higher sugar synthesis consists, undoubtedly, of the coupling of two properly activated monosaccharide subunits. Application of unstabilized sugar-derived phosphoranes was proposed by Secrist [53]. The phosphonium iodide upon treatment with a base should generate the ylid, which might either decompose in a β-elimination process or react with an aldehyde partner to afford the *cis*-Wittig adduct. Indeed, the precursor of a higher sugar antibiotic hikozamine was obtained according to this general idea under rigorously controlled conditions (*Scheme 26*). This methodology was also applied for the synthesis of tunicamine [54], a backbone of another higher sugar antibiotic—tunicamycin.



Scheme 26

A general methodology for the preparation of higher carbon sugars was proposed by Jarosz in the mid 1980s [6,55,56]. Protected monosaccharide with the terminal OH free was oxidized to an acid, which was further converted either into phosphorane or phosphonate. Upon reaction with another sugar synthon (aldehyde), α,β-unsaturated higher sugar enones were formed in good yields (Scheme 26). The second approach is more versatile, since the phosphonates are more reactive and can be prepared in much higher yields than the corresponding phosphoranes [6,55].

This methodology may be also successfully applied to complex sugars. For example, selectively protected sucrose **54** with the 6-OH free was converted into the corresponding phosphonate, which upon reaction with '*diacetonogalactose*' aldehyde provided the higher sucrose precursor **55** in good yield (Fig. 7) [57].

Another general approach leading to higher carbon sugars utilized acetylenic precursors [55, 58]. Conversion of an aldehyde into sugar acetylene was achieved using Corey's methodology [59] (Ph_3P/CBr_4, then a base). Treatment of the acetylene **56** with tributyltin hydride under the radical conditions afforded the *E*-vinyltin **57** (Scheme 27).

Replacement of the tin moiety with lithium (this reaction proceeds with the retention of the configuration at the double bond) followed by reaction with sugar aldehyde afforded two diastereoisomeric higher sugar allylic alcohols **58** with the *E*-geometry across the double bond.

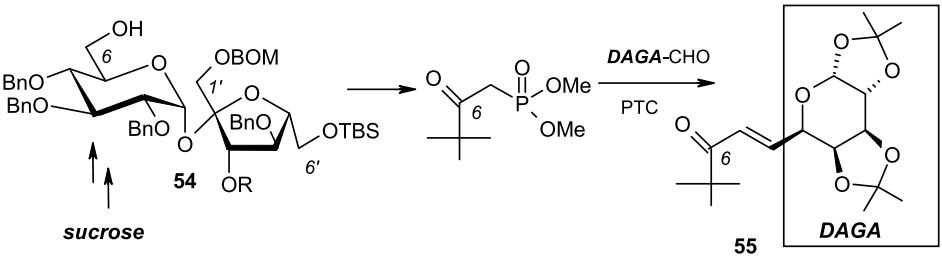
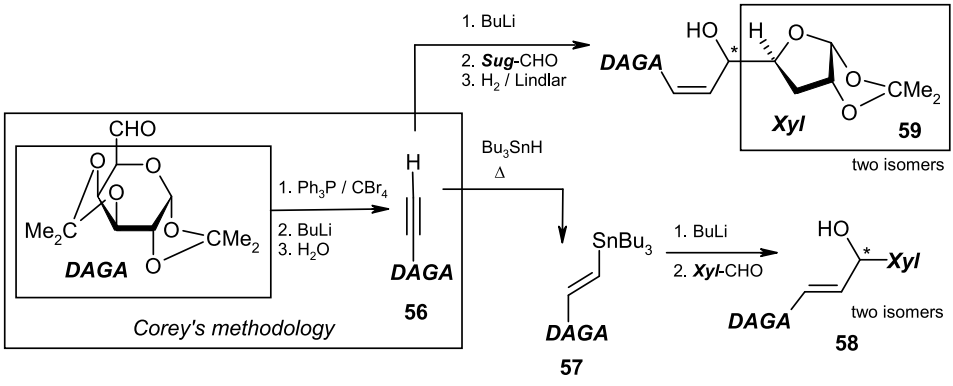
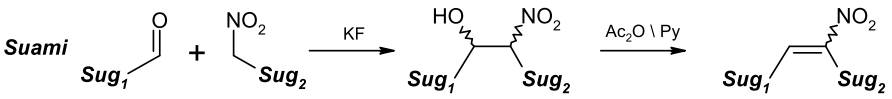


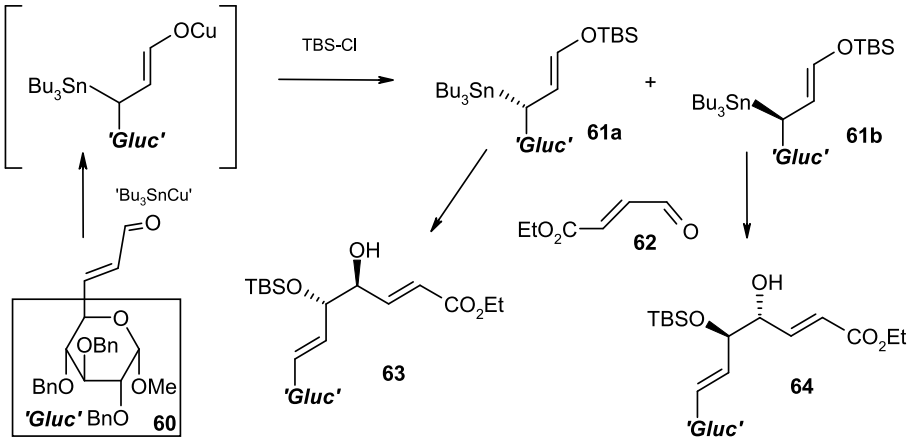
Figure 7
Concise approach to "higher sucrose" via the phosphonate method



Scheme 27



Scheme 28



Scheme 29

Alternatively, the anion generated from the acetylene **56** reacted with sugar aldehydes providing two isomeric higher sugar propargylic alcohols, which were reduced (H_2/Lindlar catalyst) into the *Z*-olefins **59** (*Scheme 27*) [58].

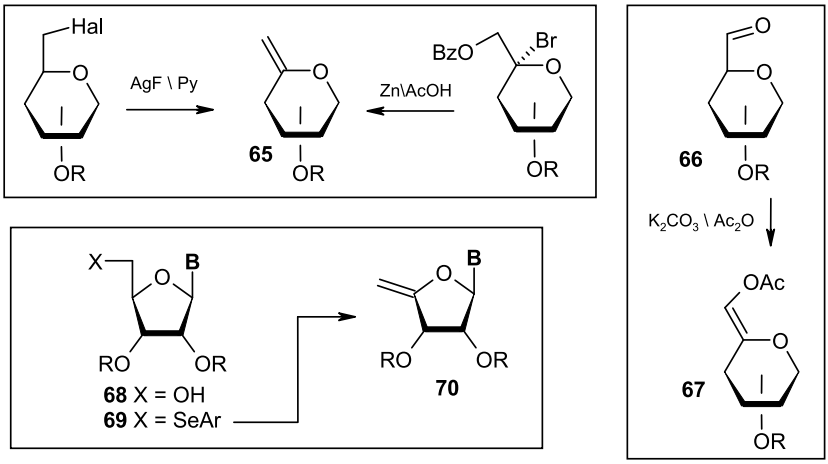
Other methods for coupling of monosaccharide sub-units leading finally to unsaturated higher sugars are also used, as for example the Henry reaction applied by Suami in his marvelous synthesis of tunicamine (*Scheme 28*) [60].

Another interesting example consists of application of reactive (but at the same time stable) stannyl derivatives. Marshall prepared unsaturated precursors of dodecoses according to *Scheme 29* [61].

The synthesis was initiated from the easily available unsaturated aldehyde **60**, which reacted with tributyltin cuprate affording a 1,4-adduct (as a mixture of stereoisomers), trapped as trimethylsilyl ethers **61a** and **61b**. Reaction of these reactive species with synthon **62** afforded the unsaturated precursors of higher carbon sugars **63** and **64**.

4.1 5,6-Unsaturated Pyranosides (and 4,5-Unsaturated Furanosides)

This class of compounds needs special interest. Such derivatives are especially important, since they serve as extremely useful intermediates in the preparation of a wide variety of carbocyclic compounds with strong biological activity (e. g. inositols, carbasugars, etc.) [5,62].



Scheme 30

5,6-Unsaturated pyranosides are easily converted into cyclohexane derivatives by a $HgCl_2$-catalyzed rearrangement discovered by Ferrier in 1979 (The Ferrier-II rearrangement) [5].

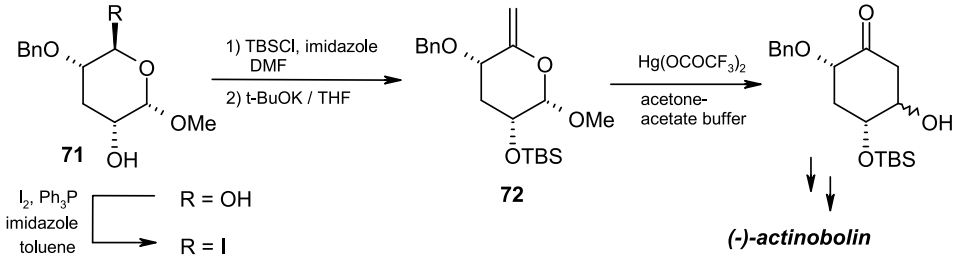
The first compound with the 5,6-unsaturation was synthesized already in 1930; the action of silver fluoride in pyridine on 6-halogenosugars caused elimination of a HX molecule with formation of compound **65** (*Scheme 30*). This reaction soon became a standard method for the preparation of 5,6-unsaturated pyranosides [63]. Much cheaper than silver fluoride reagents such as DBU [64] or sodium hydride [65] may also be applied in such a process. Reduction of 5-bromo-sugars with zinc in acetic acid also affords the exocylic derivatives **65**, however, small amounts of the endocyclic (4,5-unsaturated) isomer are formed (*Scheme 30*) [66].

Treatment of aldehydosugars **66** with potassium carbonate and acetic anhydride provides appropriate 6-acyloxy-5,6-unsaturated sugars **67** (*Scheme 30*) [67]. An example of the preparation of 4,5-unsaturated furanose **70** (used for the synthesis of analogs of antibiotic sinefungin) is presented in *Scheme 30*. Alcohol **68** was converted into selenide **69** which was oxidized to selenoxide with $NaIO_4$ and further eliminated in the presence of triethylamine [68].

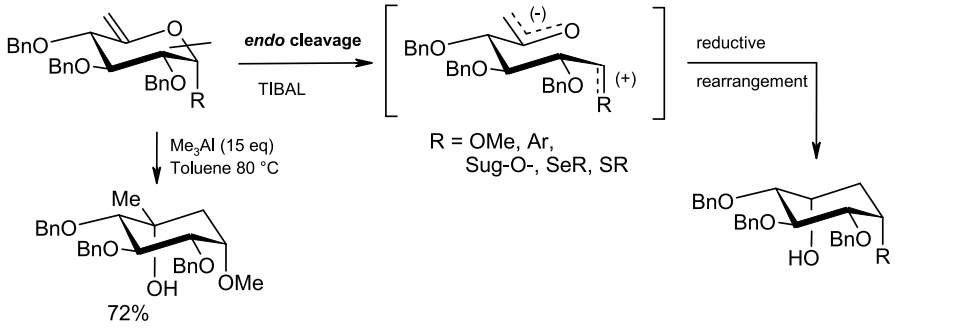
Recently the Ferrier-II rearrangement was used as a key step in the synthesis of unnatural (–) actinobolin. The 5,6-unsaturated compound **72** was prepared by a base-induced elimination of HI from the 6-deoxy-6-iodo-derivative **71** (*Scheme 31*) [69].

The alternative reaction to the Ferrier-II rearrangement was proposed by Sinaÿ. Treatment of the 5,6-unasturated glycoside with tri-*iso*-butyl aluminum (TIBAL) provides the corresponding carbocycles in high yields. The configuration at the anomeric center in the glycoside is preserved during the cyclization [70]. This reaction can be performed for *O*- and *C*-glycosides, disaccharides, thio- and selenyl-glycosides, etc. (*Scheme 32*).

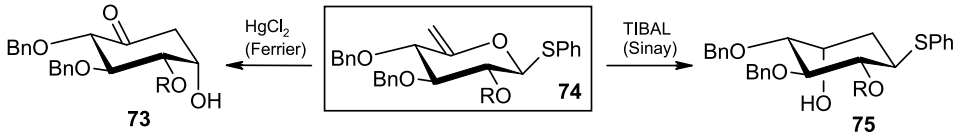
This reaction is complementary to Ferrier rearrangement. For example treatment of the thioglycoside **74** under the Ferrier conditions ($HgCl_2$) affords the ketone **73** with elimination of the sulfur moiety, while application of the Sinaÿ procedure leads to the cyclohexane derivative **75** with the sulfur atom present in the molecule (*Scheme 33*).



Scheme 31



Scheme 32

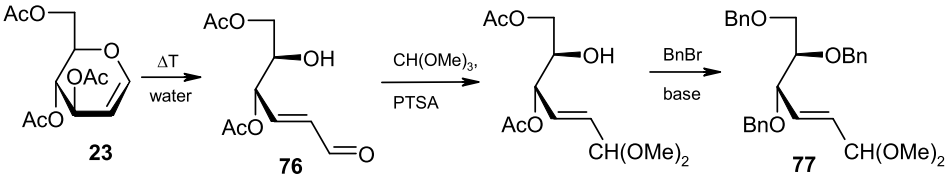


Scheme 33

5 Acyclic Unsaturated Sugars

One of the first unsaturated open-chain sugars (observed for the first time by Emil Fischer) was obtained by hydrolysis of tri-*O*-acetyl-D-glucal (**23**). Heating of this compound in water results in smooth formation of the *E*-α,β-unsaturated open-chain aldose **76** (*Scheme 34*) [1]. The carbonyl group can be protected, which allows one to replace the acetate blocks into more stable benzyl groupings thus providing compound **77** (*Scheme 34*) [71].

In 1979 Vasella introduced a convenient procedure for the preparation of unsaturated open-chain sugars by reductive fragmentation of 6-deoxy-6-halogenopyranosides (and 5-deoxy-5-halogeno-furanosides) with zinc in ethanol [72]; the same fragmentation is induced with *n*-butyl-lithium. Fürstner proposed modification of the original procedure and in his approach activated zinc on graphite was used [73]. The conditions of the Fürstner process are milder and, for example, unsaturated acid **79** can be prepared in very high yield from iodolactone **78** (*Fig. 8*).



Scheme 34

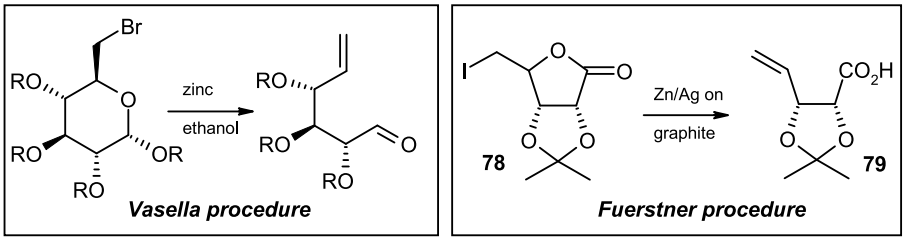
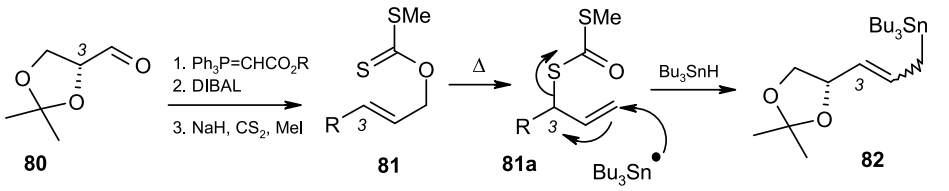


Figure 8
Synthesis of unsaturated open chain sugars by fragmentation reactions

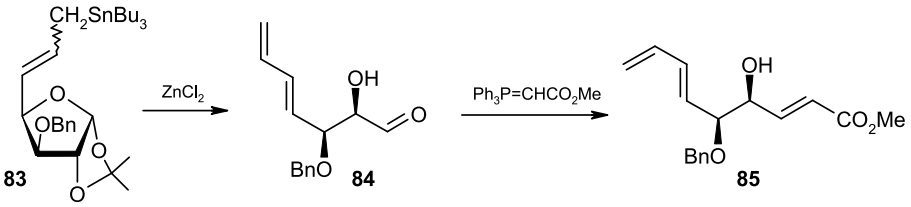
Sugar allyltin derivatives are very useful synthetic intermediates. The most convenient and reliable method for their preparation is a so-called 'xanthate' procedure. The first compound of this class was prepared in 1988 by Mortlock and Thomas [74]. 1,2-*O*-isopropylidene-D-glyceraldehyde (**80**) was converted into allylic alcohol and further transformed into the corresponding xanthate **81**. This compound underwent thermal [3,3] rearrangement into the dithiocarbonate **81a**, which upon reaction with tri-*n*-butyltin hydride provided the final sugar allyltin **82** as a mixture of isomers with the *E*-one strongly predominating (*Scheme 35*).

Application of other sugars (pentoses and hexoses) as starting materials allowed Jarosz to prepare a number of such useful compounds (such as **83** or **86** in Schemes 36 and 37) using the xanthate method [75,76,77].

It was observed that treatment of such allyltins with a Lewis acid (the best being zinc chloride) induced a controlled fragmentation leading to highly oxygenated dienoaldehydes with the *E*-configuration across the internal double bond. This procedure may be, therefore, a method of choice for the synthesis of unsaturated sugar dienes. For example sugar allyltin **83** (obtained in a few steps from 'diacetonoglucose') upon treatment with zinc chloride provided a dienoaldehyde **84**, which can be further elongated into the triene **85** by simple reaction with a stabilized Wittig reagent (*Scheme 36*) [76].



Scheme 35

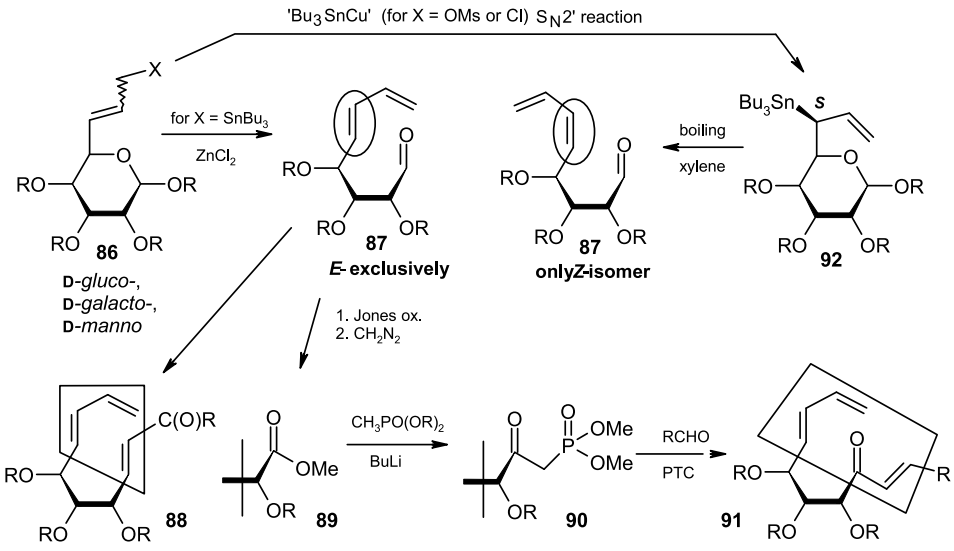


Scheme 36

It was observed, that the hexose-derived allytins **86** are convenient precursors of highly oxygenated carbobicyclic derivatives. When such a compound was treated with $ZnCl_2$ the *E*-dienoaldehydes **87** were formed exclusively, which could be further transformed into the corresponding trienes (*Scheme 37*) [76,77].

Reaction of dienoaldehyde with the stabilized Wittig reagent(s) provided a triene **88**, which upon high-pressure Diels–Alder cyclization furnished bicyclo[4.3.0]nonene [78]. Alternatively the aldehyde **87-*E*** was converted into the methyl ester **89**, which reacted with the dimethyl methylphosphonate anion affording the phosphonate **90**. Reaction of the latter with aldehydes under mild PTC conditions yielded a (regioisomeric to the previous one) triene **91**, which cyclized spontaneously to bicyclo[4.4.0]decene (*Scheme 37*) [79].

The controlled fragmentation of the primary sugar allyltins is a convenient method for the preparation of dienoaldehydes with the *E*-configuration across the internal double bond. The *Z*-dienes are also available from sugar allyltins, but the secondary ones. When the sugar allylic mesylates (or chlorides) **86** reacted with a tin nucleophile, the S_N2' product **92** was obtained.



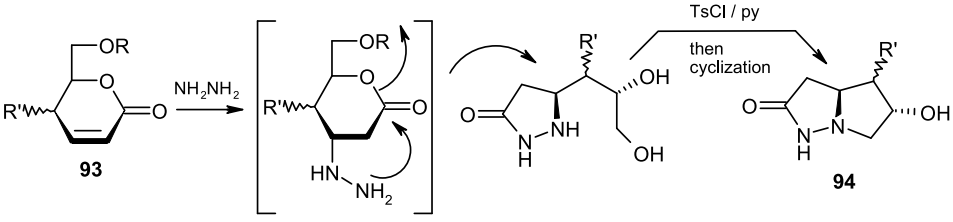
Scheme 37

Moreover, only one stereoisomer was formed regardless of the geometry (*E*- or *Z*-) of the substrate, to which the *S*-configuration at the newly created stereogenic center was assigned [80]. It was found that such secondary allyltins decomposed thermally (the primary ones are stable up to at least 220 °C) into the dienoaldehydes **87** with the *Z*-geometry across the internal double bond. Therefore, both dienes are available with high stereoselectivity (❯ *Scheme 37*) [77].

6 Application of Unsaturated Sugars as Chirons

Azasugars [81], carbasugars [82], and *C*-glycosides [83] are important sugar mimics, which may be prepared by a variety of methods. In this section the selected methods for the preparation of such derivatives (and also other important compounds) from unsaturated sugars will be presented. Although this is not connected directly with the main subject of this chapter (which deals with the formation of a double bond in sugars), these important targets are prepared from unsaturated sugars, which are in turn synthesized by a number of methods described within this chapter. This would also show the very high synthetic potential of unsaturated sugars.

An interesting methodology to highly oxygenated pyrazolidines and indolizidines from 2,3-unsaturated sugar lactones was presented by Chmielewski [84]. The synthesis was initiated by a 1,4-addition of nitrogen nucleophiles to unsaturated lactones **93**, which resulted in formation of appropriate heterocyclic derivative **94** (❯ *Scheme 38*).

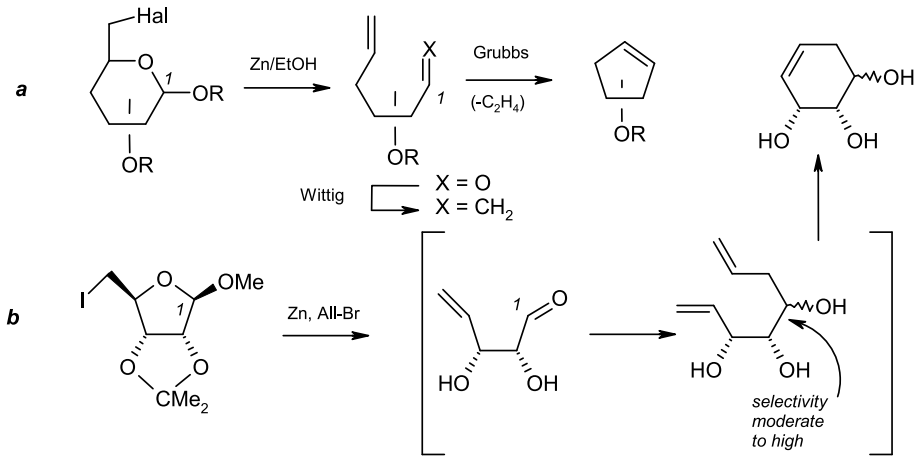


Scheme 38

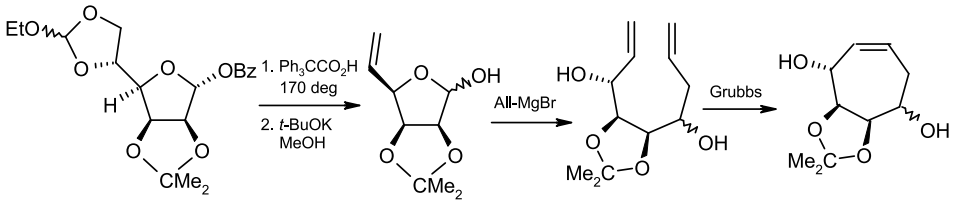
Further transformations allowed one to obtain a variety of natural and unnatural azasugars.

One of the most useful routes to highly oxygenated carbocycles consists of the transformation of monosaccharide into diolefin, which is further subjected to ring-closing metathesis reaction (*RCM*) with formation of an unsaturated carbocycle. The most convenient way to introduce a terminal double bond is provided by reductive dehalogenation of a terminal halogenosugar (the Vasella reaction, which was already presented in ❯ *Fig. 8*, ❯ *Sect. 5*) leading to the corresponding eno-aldehyde in good yield.

The unsaturation at C-1 may be introduced either by a simple Wittig (or Wittig-type) reaction (C_1 homologation; ❯ *Scheme 39*; *route a*) or by reaction with, for example, allylic building blocks (C_3 homologation); these two steps (reductive dehalogenation followed by a C_3 homologation) can be performed simultaneously (❯ *Scheme 39*; *route b*). Such prepared diolefins are then transformed into carbocycles [85] with the Grubbs' (or similar) catalysts. This versatile methodology allows one to prepare carbocycles with different sizes (5–8) of the ring, as shown in ❯ *Scheme 39* and ❯ *Scheme 40*.



Scheme 39



Scheme 40

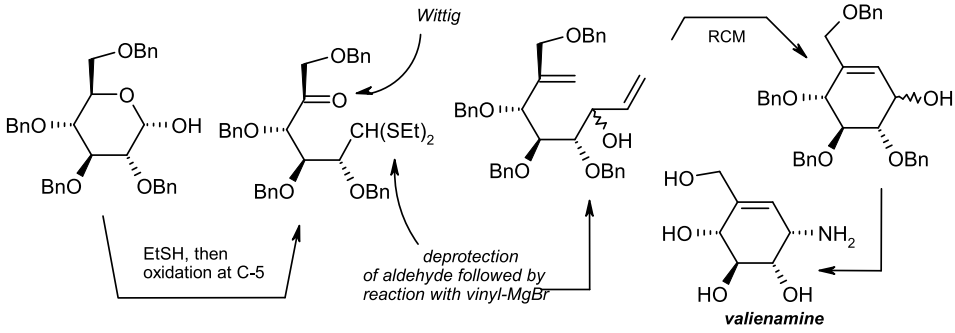
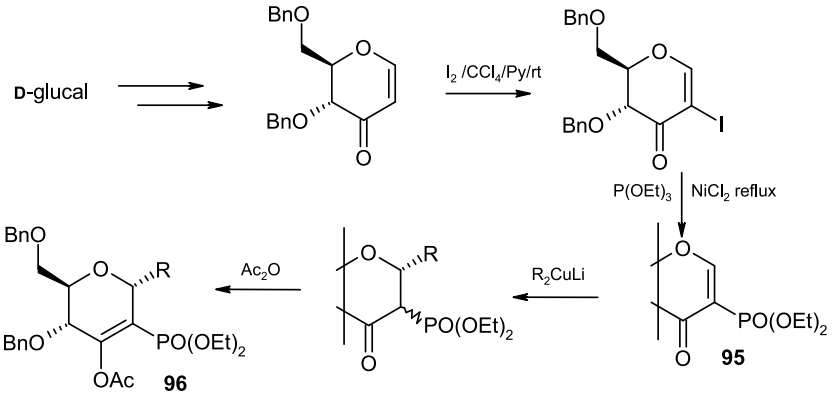


Figure 9
Application of unsaturated sugars in the synthesis of valienamine

The carbocycles substituted at the double bond (with for example, an alkoxymethylene group) can also be prepared from the appropriately functionalized unsaturated sugars. Synthesis of α-glucosidase inhibitor—valienamine—was realized from unsaturated sugar chirons as shown in *Fig. 9* [86].



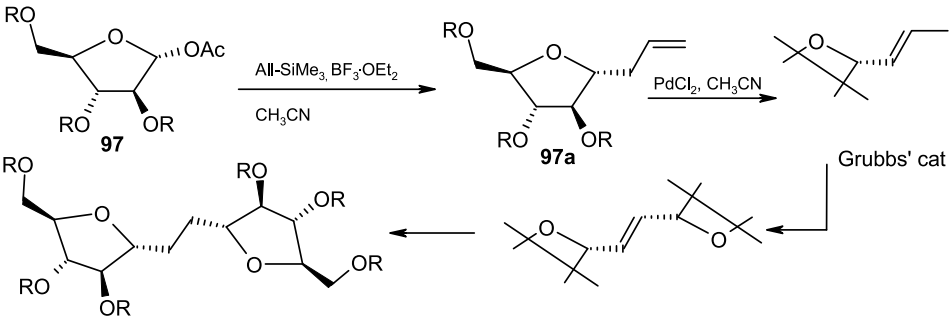
Scheme 41

C-glycosyl phosphonates are important inhibitors of enzymes [87]. An interesting approach to such derivatives was proposed recently [88].

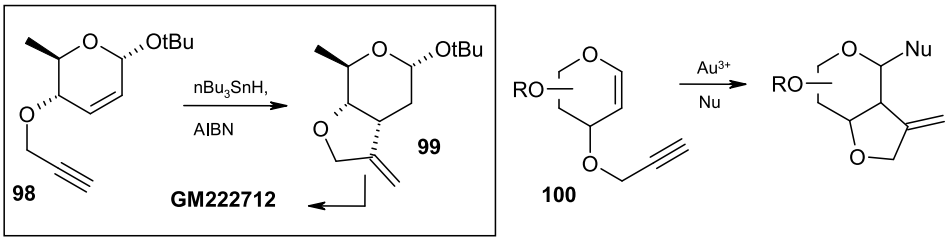
The synthesis was initiated from the unsaturated phosphonate **95** prepared in several steps from the D-glucal as shown in *Scheme 41*. The 1,4-addition of a soft nucleophile provided the adduct which, upon acetylation, was easily converted into the target **96**.

An interesting approach to symmetrical *C*-disaccharides was proposed recently by Lowary. 1-*O*-Acetyl-sugar **97** was converted into 1-*C*-allyl derivative **97a** by classical reaction with allyltrimethylsilane, and this derivative was 'dimerized' under the metathesis conditions with Grubbs' catalyst (*Scheme 42*) [89].

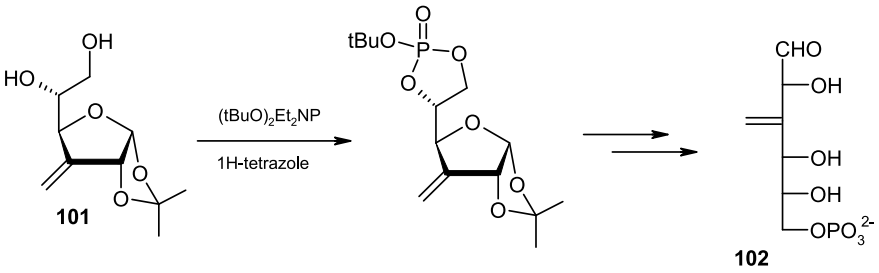
Unsaturated sugars are used also as starting materials for the preparation of complex natural products and their analogs. For example, radical cyclization of **98** (readily prepared in well defined steps from D-glucal) led to synthon **99**, a precursor of antifungal agent GM222712 (a close analog of sordarin) [90]. The cascade reaction of **100** with the nucleophile catalyzed with gold-III provided also the bicyclic skeleton with a different arrangement of unsaturated bonds (*Scheme 43*) [91].



Scheme 42



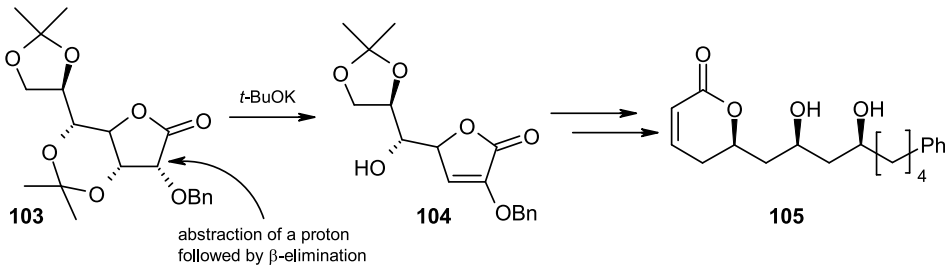
Scheme 43



Scheme 44

The glucose analog, 3-deoxy-3-*C*-methylene-D-*ribo*-hexose (**102**) is a substrate of a D-xylose isomerase. The efficient synthesis from 3-deoxy-3-*C*-methylene-glucofuranose (**101** prepared by the classical method according to [92]) was described recently (*Scheme 44*) [93].

The key step in the preparation of 5,6-dihydro-δ-pyrones with an incorporated 1,3-polyol system (an example, compound **105** possessing inhibitory activity against *C. cucumerinum* is shown in *Scheme 45*) consisted of formation of a 2,3-unsaturated furanose derivative **104**, which was prepared from **103**. Abstraction of an acidic proton from the C-2 induced β-elimination, providing the desired target **104** [94].



Scheme 45

7 Miscellaneous

α,β-Unsaturated γ-lactones occur widely in nature as simple metabolites of a broad class of natural products. Many of these compounds exhibit interesting biological activity such as, for example, antifungal, antibacterial, cytotoxic properties, etc. Because of their biological importance a number of methods for preparation of such compounds were described in the literature. An interesting method based on the elimination reaction was proposed by Rauter (*Scheme 46*) [95].

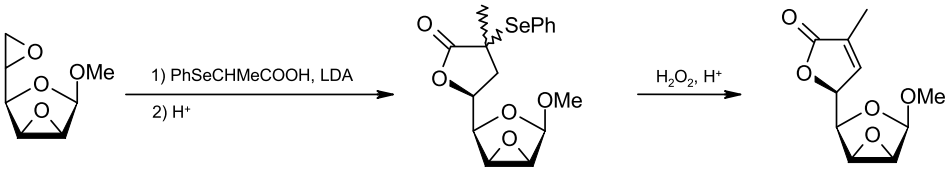
The synthesis of higher carbon sugars by an Achmatowicz approach was realized by Indian scientists [96]. This example should have been presented in *Sect. 2* (describing general methods); however, since it illustrates a concise approach to the synthesis of unsaturated sugars using different methods it is included here.

Addition of allyl species to the free sugar **106** affords an open-chain unsaturated carbohydrate **107**. In a few well-defined steps the furyl derivative **108** is obtained, which upon oxidation with peroxide provides finally the higher carbon sugar **109** (*Scheme 47*).

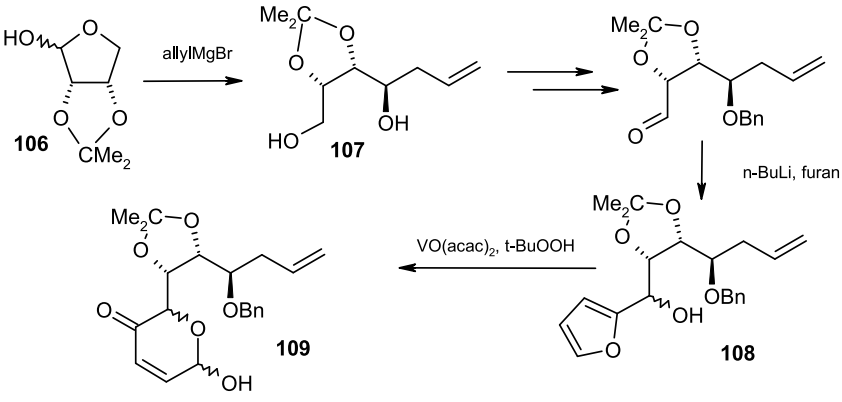
Another example of the usefulness of tri-*O*-acetyl-D-glucal in the preparation of a complex target in optically pure form is illustrated by the synthesis of iridoids.

Compound **110** (obtained by reaction of D-glucal with propargylic alcohol) was subjected to cyclization with $Co_2(CO)_8$, which led to a fused tricyclic derivative **111** (*Scheme 48*) [97].

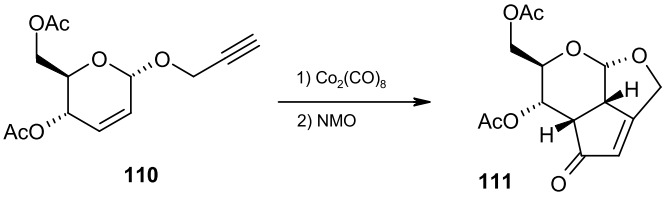
[2+2] Cycloaddition of carbenes to glycals is an efficient route leading to cyclopropanated sugars with high stereoselectivity [1]. Such cyclopropanes can be transformed into a number



Scheme 46



Scheme 47



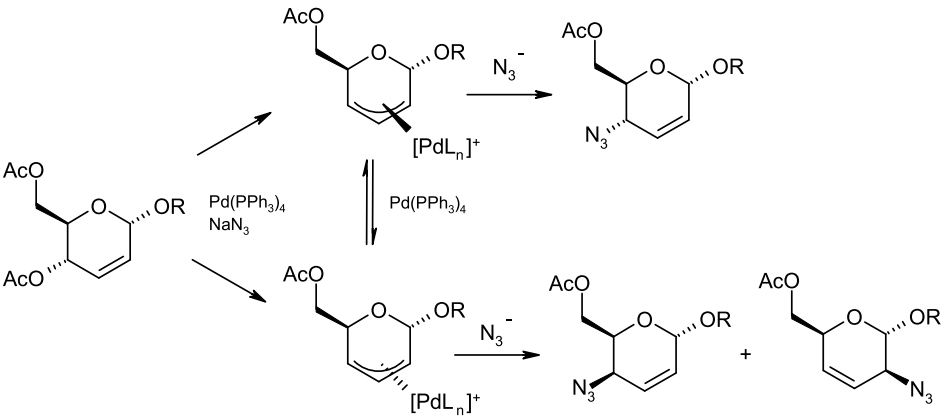
Scheme 48

Scheme 49

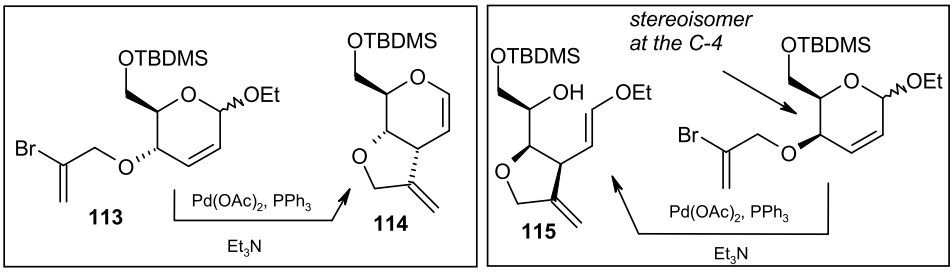
of interesting compounds. The 3,4-unsaturated septanose was prepared from cyclopropane **112** upon reaction with trimethylsilyl azide (*Scheme 49*) [1].

The palladium-catalyzed reaction of unsaturated sugars (already mentioned in *Sect. 3.4, rearrangements*) is a methodology that may also lead to 3,4-unsaturated sugars. A complex obtained from the 2,3-unsaturated derivative and Pd(0) species reacted with an azide anion providing either substituted 2,3- or 3,4-unsaturated sugar azides as shown in *Scheme 50* [98].

As shown in *Scheme 51* the outcome of the intramolecular reaction of unsaturated sugars depends on the configuration of the starting material. Cyclization of the 4(*S*) stereoisomer **113** induced by palladium species led to bicyclic unsaturated sugar **114**, while the same reaction performed for the 4(*R*) stereoisomer led to compound **115** with an opened sugar ring (*Scheme 51*) [99].



Scheme 50



Scheme 51

8 Conclusions

Material presented in this chapter describes the methodology of the synthesis of unsaturated sugars (more precisely the formation of the C=C bond) by various methods. This is an updated version of the chapter that appeared in the 2001 edition of this book presenting the concise approach leading to such derivatives. The chemistry of simple (and also more complex) monosaccharides is very well developed, which allows one to apply this well-established methodology for the preparation of complex targets. Therefore, the older methods already described in the previous chapter are also included here, since they represent the basic knowledge of modern carbohydrate chemistry. The material is illustrated with examples of the synthesis of new targets, which have been prepared using methods developed in the last few decades.

Of course, this chapter does not pretend to be comprehensive. The idea behind this chapter was for the authors to show how to prepare the desired unsaturated sugars using proper method(s) and (eventually) how to apply these synthons in the synthesis of important targets.

In modern organic synthesis, monosaccharides have a very well-developed chemistry that allows us to prepare almost any derivative and they are convenient sources of chirality. We hope that the material presented in this chapter will help the reader to plan the synthesis of optically pure targets using the known and convenient carbohydrate approach.

References

1. Ferrier RJ, Hoberg JO (2003) Adv Carbohydr Chem Biochem 58:55
2. Ferrier RJ, Zubkov OA (2003) Org React 62:569
3. Zamojski A, Banaszek B, Grynkiewicz G (1982) Adv Carbohydr Chem Biochem 40:1
4. Ferrier RJ (2001) Topics Curr Chem 215:277
5. Ferrier RJ, Middleton S (1993) Chem Rev 93:2779
6. Jarosz S (2001) Glycoscience—Chemistry and Chemical Biology 1:365
7. Meier Ch, Lorey M, De Clercq E, Balzarini J (1997) Bioorg Med Chem Lett 7:99
8. Jeong L, Schinazi RF, Beach JW, Kim HO, Nampalli S, Shanmuganathan K, Alves AJ, McMillan A, Chu CK (1993) J Med Chem 36:181
9. Yun M, Moon HR, Kim HO, Choi WJ, Kim YCL, Park CCS, Jeong LS (2005) Tetrahedron Lett 46:5903 and references therein
10. Hanessian S (1983) Total Synthesis of Natural Products: The Chiron Approach. Pergamon Press, New York; Hanessian S, Franco J, Larouche B (1987) Pure Appl Chem 62:1990; Inch TD (1984) Tetrahedron 40:3161; Fraser-Reid B, Tsang R (1989) Strategies and Tactics in Organic Synthesis. Academic Press, New York:2; Fraser-Reid B (1996) Acc Chem Res 29:7
11. Ferrier RJ (1969) Adv Carbohydr Chem 24:199
12. Garegg PJ, Samuelsson B (1979) Synthesis 469; Liu Z, Classon B, Samuelsson B (1990) J Org Chem 55:4273

13. Tipson RS, Cohen A (1965–1966) Carbohydr Res 1:338
14. Josan JS, Eastwood FW (1968) Carbohydr Res 7:161
15. Corey EJ, Winter RAE (1963) J Am Chem Soc 85:2677; Horton D, Thomson JK, Tindall ChG Jr (1972) Methods Carbohydr Chem 6:297
16. Barrett AGM, Barton DHR, Bielski R (1979) J Chem Soc Perkin Trans 1 2378
17. Konowal A, Jurczak, J, Zamojski A (1968) Rocz Chem 42:2045
18. Schmidt RR (1986) Acc Chem Res 19:250
19. Danishefsky SJ, Pearson WH, Segmuller BE (1985) J Am Chem Soc 107:1281; Danishefsky SJ, DeNinno MP (1987) Angew Chem Int Ed Engl 26:15
20. Angerbauer R, Schmidt RR (1981) Carbohydr Res 89:173; Schmidt RR (1986) Acc Chem Res 19:250; Bataille C, Begin G, Guillam A, Lemiegre L, Lys C, Maddaluno J, Toupet L (2002) J Org Chem 67:8054
21. Kwiatkowski P, Asztemborska M, Caille JC, Jurczak J (2003) Adv Synth Catal 345:506; Kwiatkowski P, Asztemborska M, Jurczak J (2004) Synlett 1755; Kwiatkowski P, Asztemborska M, Jurczak J (2004) Tetrahedron: Asymmetr 15:3189; Chaladaj W, Kwiatkowski P, Jurczak J (2006) Synlett 3263
22. Achmatowicz O, Bukowski P, Zwierzchowska Z, Zamojski A (1971) Tetrahedron 27: 1973; Achmatowicz O (1981) In: Trost BM and Hutchison CR (eds) Organic Synthesis Today and Tomorrow 307. Pergamon Press, Oxford
23. Casiraghi G, Colombo L, Rassu G, Spanu P (1991) J Org Chem 56:2135 and 6523
24. Raczko J, Jurczak J (1995) In: Atta-ur Rahman (ed) Studies in Natural Product Chemistry. Elsevier, Amsterdam, 15:639 and literature cited therein
25. Horton D, Tsai JH (1979) Carbohydr Res 75:151
26. Doboszewski B, Chu CK, Van Halbeek H (1988) J Org Chem 53:2777
27. Clive DLJ, Wickens PL, Sgarbi PWM (1996) J Org Chem 61:7426; Clive DLJ, Sgarbi PWM, Wickens PL (1997) J Org Chem 62:3751
28. Baba T, Huang G, Isobe M (2003) Tetrahedron 59:6851
29. Deppe O, Glumer A, Yu S, Buchholz K (2004) Carbohydr Res 339:2077
30. Knapp S, Nandan SR (1994) J Org Chem 59:281
31. Capozzi G, Catelani G, D'Andrea F, Menichetti S, Nativi C (2003) Carbohydr Res 338:123
32. Ferrier RJ, Prasad N (1969) J Chem Soc C 570
33. Babu Satheesh R, O'Doherty GA (2003) J Am Chem Soc 125:12406; Babu Satheesh R, Zhou M, O'Doherty GA (2004) J Am Chem Soc 126:3428
34. Shanmugasundaram B, Bose AK, Balasubramian KK (2002) Tetrahedron Lett 43:6795
35. Ferrier RJ (1972) Methods Carbohydr Chem 6:307
36. Baer HH (1989) Pure Appl Chem 61:1217
37. Herscovici J, Montserret R, Antonakis K (1988) Carbohydr Res 176:219
38. Paulsen H, Thiem J (1973) Chem Ber 106:3850; Alexander P, Krishnamamurthy VV, Prisbe EJ (1996) J Med Chem 39:1321
39. Takano A, Fukuhara H, Ohno T, Kutsuma M, Fujimoto T, Shirai H, Iriye R, Kakehi A, yamamoto I (2003) J Carbohydr Chem 22:443
40. Hosokawa S, Kirschbaum B, Isobe M (1998) Tetrahedron Lett 39:1917; Saeeng R, Sirion U, Sahakitpichan P, Isobe M (2003) Tetrahedron Lett 44:6211
41. Ischikawa Y, Kobayashi Ch, Isobe M (1996) J Chem Soc Perkin Trans 1 377
42. Ohyabu N, Nishikawa T, Isobe M (2003) J Am Chem Soc 125:8798
43. Montero A, Mann E, Herradon B (2004) Tetrahedron Lett 46:401
44. Di Bussolo V, Caselli M, Romano MR, Pineschi M, Crotti P (2004) J Org Chem 69:7383
45. David S, Lubineau A, Vatele JM (1982) Carbohydr Res 104:41
46. Canac Y, Levoirier E, Lubineau A (2001) J Org Chem 66:3206
47. Wegert A, Reinke H, Miethchen R (2004) Carbohydrate Res 339:1833
48. Ali MH, Collins PM, Overend V G (1990) Carbohydr Res 205:428
49. Sharma M, Bobek M (1990) Tetrahedron Lett 31:5839
50. Frick W, Kruelle Th, Schmidt RR (1991) Liebigs Ann Chem 435
51. Itoh H, Kaneko T, Tanami K, Yoda T (1988) Bull Chem Soc Jpn 61:3356
52. Brimacombe JS (1989) In: Atta-ur Rahman (ed) Studies in Natural Product Chemistry. Elsevier, Amsterdam, 4C:157
53. Secrist JA III, Wu SR (1979) J Org Chem 44:1434; Secrist JA III, Barnes KD, Wu SR (1989) In: Horton ED, Hawking DL, McCorrey GJ (eds) Trends in Synthetic Carbohydrate Chemistry. ACS Symposium Series, 386:93
54. Karpiesiuk W, Banaszek A (1994) Bioorg Med Chem Lett 4:879
55. Jarosz S (2001) J Carbohydr Chem 20:93

56. Jarosz S, Skora S, Kosciolowska I (2003) Carbohydr Res 338:407
57. Mach M, Jarosz S (2001) J Carbohydr Chem 20:411; Jarosz S, Mach M (2002) Eur J Org Chem 769
58. Jarosz S (1993) J Carbohydr Chem 12:1149
59. Corey EJ, Hua DH, Pan BCh, Steitz SP (1982) J Am Chem Soc 104:6816
60. Suami T, Sasai H, Matsuno K (1983) Chem Lett 819
61. Marshall JA, Elliott, LM (1996) J Org Chem, 61:4611; Jarosz S (2003) Curr Org Chem 7:13
62. Ferrier RJ (1979) J Chem Soc Perkin Trans 1 1455
63. Blair MG (1963) Methods Carbohydr Chem 2:415
64. Mirza S, Molleyres LP, Vasella A (1985) Helv Chim Acta 68:988
65. Barton DHR, Auge-Dorey S, Camara J, Dalko P, Delaumeny JM, Gero SD, Quiclet-Sire B, Stuetz P (1990) Tetrahedron 46:215
66. Blattner R, Ferrier RJ, Tyler CP (1980) J Chem Soc Perkin Trans 1 1535
67. Prestwich GD (1996) Acc Chem Res 29:503
68. Meade EA, Krawczyk SH, Towsend LB (1988) Tetrahedron Lett 29:4073
69. Imuta S, Ochiai S, Kuribayashi M, Chida N (2003) Tetrahedron Lett 44:5047
70. Sollogub M, Millet JM, Sinaÿ P (2000) Angew Chem Int Ed 39:362; Pearce AJ, Sollogub M, Mallet JM, Sinaÿ P (1999) Eur J Org Chem 2103; Sollogoub M, Pearce AJ, Herault A, Sinaÿ P (2000) Tetrahedron: Asymmetr 11:283; Jia C, Pearce AJ, Bleriot Y, Zhang Y, Zhang LH, Sollogoub M, Sinaÿ P (2004) Tetrahedron: Asymmetr 15:699
71. Pathak R, Shaw AK, Bhaduri AP (2002) Tetrahedron 58:3535
72. Bernet B, Vasella A (1979) Helv Chim Acta 62:1990 and 2400
73. Fuerstner A (1993) Angew Chem Int Ed Engl 32:164
74. Mortlock SV, Thomas EJ (1988) Tetrahedron Lett 29:2479
75. Jarosz S, Fraser-Reid B (1989) J Org Chem 54:4011; Kozlowska E, Jarosz S (1994) J Carbohydr Chem 13:889
76. Jarosz S, Skóra S, Szewczyk K (2000) Tetrahedron: Asymmetr 11:1997
77. Jarosz S, Gawel A (2005) Eur J Org Chem 3415
78. Jarosz S, Skora S (2000) Tetrahedron: Asymmetr 11:1425
79. Jarosz S, Skora S (2000) Tetrahedron: Asymmetr 11:1433
80. Jarosz S, Szewczyk K, Zawisza A (2003) Tetrahedron: Asymmetr 14:1715
81. Afarinkia K, Bahar A (2005) Tetrahedron: Asymmetr 16:1239
82. Kobayashi Y (2001) Glycoscience: Chemistry and Chemical Biology III:2595; Ogawa S (1998) Carbohydrate Mimics: Concepts and Methods 87; Suami T (1990) Top Curr Chem 154:257; Suami T, Ogawa S (1990) Adv Carbohydr Chem Biochem 48:21.
83. Postema NHD, Piper JL, Betts RL (2005) Synlett 1345
84. Rabiczko J, Chmielewski M (1999) J Org Chem 64:1347; Panfil I, Lipkowska-Urbanczyk Z, Suwinska K, Solecka J, Chmielewski M (2002) Tetrahedron 58:1199
85. Jorgensen M, Hadwiger P, Madsen R, Stütz AE, Wrodnigg TM (2000) Curr Org Chem 4:565; Madsen R (2007) Eur J Org Chem 399
86. Chang YK, Lee BY, Kim DJ, Lee GS, Jeon HB, Kim KS (2005) J Org Chem 70:3299
87. Engel R (1977) Chem Rev 77:349; Dondoni A, Marra A (2000) Tetrahedron: Asymmetr 11:305; Orsini F, Caselli A (2002) Tetrahedron Lett 43:7259
88. Leonelli F, Capuzzi M, Calcano V, Passacantilli P, Piancatelli G (2005) Eur J Org Chem 2671
89. Chang GX, Lowary TL (2006) Tetrahedron Lett 47:4561
90. Bueno JM, Coteron JM, Chiara JL, Fernandez-Mayoralas A, Fiandor JM, Valle N (2000) Tetrahedron Lett 41:4379
91. Kashyap S, Hotha S (2006) Tetrahedron Lett 47:2021
92. Brimacombe JS, Miller JA, Zakir U (1976) Carbohydr Res 49:233
93. Burger A, Tritsch D, Biellmann JF (2001) Carbohydr Res 332:141
94. Ramana CV, Srinivas B, Puranik VG, Gurjar MK (2005) J Org Chem 70:8216
95. Rauter AP, Figueiredo J, Ismael M, Canda T, Font J, Figueredo M (2001) Tetrahedron: Asymmetr 12:1131
96. Krishna UM, Deodhar KD, Trivedi GK (2004) Tetrahedron 60:4829
97. Marco-Contelles J, Ruiz-Caro J (1999) J Org Chem 64:8302
98. de Oliveira RN, Cottier L, Sinou D, Srivastava RM (2005) Tetrahedron 61:8271
99. Bedjeguelal K, Joseph L, Bolitt V, Sinou D (1999) Tetrahedron Lett 40:87

2.7 Degradations and Rearrangement Reactions

Jianbo Zhang
Department of Chemistry, East China Normal University,
200062 Shanghai, China
jbzhang@chem.ecnu.edu.cn

In: *Glycoscience*. Fraser-Reid B, Tatsuta K, Thiem J (eds)
Chapter-DOI 10-1007/978-3-540-30429-6_9:

Abstract

This section deals with recent reports concerning degradation and rearrangement reactions of free sugars as well as some glycosides. The transformations are classified in chemical and enzymatic ways. In addition, the Maillard reaction will be discussed as an example of degradation and rearrangement transformation and its application in current research in the fields of chemistry and biology.

Keywords

Degradation; Rearrangement; Hydrolysis; Double bond shift; Ring transformation; Ring-contraction; Ring-expansion; Ferrier carbocyclization; Anomerization; Aromatization; Maillard reaction; Amadori reaction

Abbreviations

AGEs	advanced glycation end products
DAST	diethylaminosulfur trifluoride
DMDO	dimethyldioxirane
DMF	dimethylformamide
EFC	ethanol-from-cellulose
HFIP	1,1,1,3,3,3-hexafluoro-2-propanol
HMF	5-(hydroxymethyl)-2-furaldehyde, 5-hydroxymethylfuraldehyde
IDCP	iodonium dicollidine perchlorate
LTMP	lithium 2,2,6,6-tetramethylpiperidide
m-CPBA	3-chloroperoxybenzoic acid, *meta*-chloroperoxybenzoic acid
PTC	phase transfer catalysis
TASF	tris(dimethylamino)sulfonium difluorotrimethylsilicate
TMSCN	trimethylsilyl cyanide
TMSOTf	trimethylsilyl triflate
TMU	*N*,*N*-tetramethylurea

1 Overview

Degradation and rearrangement reactions in carbohydrate chemistry are described in the standard organic chemistry textbooks [1,2]. Recent books [3,4,5] contain informative surveys of developments in this area. This chapter dealing with degradation and rearrangement reactions of carbohydrate systems covers literature published in the last few years. Degradation reactions are classified into two main categories: hydrolysis from glycosides or polysaccharides into free sugars, and the degradations from free sugars into useful building blocks or chiral synthons for organic synthesis. The rearrangement reactions described herein are classified into four groups: double bond shifts, ring rearrangements associating with a double bond, ring isomerizations (contraction and expansion), and other processes. Current results on the Maillard reaction initiated by the Amadori reaction, which are so intimately associated with degradation and rearrangement reactions, are also discussed in this chapter.

2 Hydrolysis of Glycosides and Polysaccharides

Among all the degradation patterns for glycosides and polysaccharides, hydrolysis is the most important process in carbohydrate chemistry, either in nature or in biological systems. Before long, it became a process important in the food industry for production of free sugars and is now gaining more and more attention because of the present energy crisis. This is because petroleum is not an ideal chemical feedstock for industry, due to its intractability, while glycosides and polysaccharides—which are abundant and recyclable—can be utilized in the production of fuel and chiral synthons to be used instead of traditional petroleum [6].

2.1 Chemical Hydrolysis

Chemical hydrolysis is a very familiar reaction for the sugar industry. However, it may generate an array of possible degradation products. For example, very low rate constants for the spontaneous hydrolysis of nonactivated methyl β-D-glucopyranoside **1** have been determined at 220 °C [7], (*Fig. 1*). At pH > 7, the rate constants approach a constant value. On hydrolysis at pH 10 in the presence of $H_2{}^{18}O$, the results show that the reaction occurs almost exclusively by cleavage of the C1/O1 bond. The β-anomer **1** is roughly twice as reactive as the α-anomer **2**, as are also the anomeric pair of methyl D-ribofuranosides **3** and **4**. Unlike the hydrolysis at pH < 7, the hydrolysis of **1** without catalysts proceeds with a negative entropy of activation. This is consistent with bimolecular attack of water on **1**.

Acid hydrolysis of isopropenyl α-D-glucopyranoside **5** at pH 3.0 and 25 °C occurs by *C*-protonation followed by cleavage of the alkenyl ether C/O bond. The α-anomer **5** is hydrolyzed 4.5-times faster that its β-anomer **6**. Spectroscopic evidence indicates greater conjugation of O1 with the double bond, and hence a greater basicity of the β-carbon of the double bond, in **5** compared to **6** [8].

An accelerating effect by the intramolecular nucleophilic catalysis of a phosphate anion upon hydrolysis of the phosphate **7** at 80TT °C and pH 6–9, in comparison with the unsubstituted **8** and its 2-*O*-methyl derivatives **9**, has been observed [9]. The reaction of **8** is base-catalyzed down almost to pH 7, while that of **9** is pH-independent up to pH 9–10 The hydrolysis of **7** proceeds about 100-times faster than that of **9** at pH ~9 and 80 °C. In comparison, the hydrolysis of **10** is pH-independent down to pH 7 and ~20-times slower than that of **7** at pH 9 and 80 °C [10].

A kinetic study of the acyl migration reaction of the 1-*O*-acyl β-D-glucopyranuronic acid **11**, a model drug ester glucuronide, employing a directly coupled stop-flow HPLC/600 MHz ^{1}H-NMR system at pH 7.4 and 25 °C, has been carried out [11]. The acyl migration rate of the β-1-*O*-acyl group of **11** is greater than any other regio-isomers. The simulating mutarotation rates for the 4-*O*-acyl isomers **12** are in good accord with the experimental values.

The fructofuranosyl cation **13** is the first formed product of the acid-catalyzed melt thermolyses of sucrose **14** (*Scheme 1*). This reacts with hydroxy nucleophiles co-existing in the melt to give fructose-grafted products. Rigorous thermolysis of **14** itself at 170 °C furnishes a fructosylglucan with an average dp ~25 together with the known sucrose thermal oligosaccharides from **14**, such as **15** (3.9%) and **16** (4.1%) [12].

Mechanistic studies on acid hydrolysis of glycosides often encounter the *endo*/*exo*-cyclic cleavage problem [13]. For instance, the sulfuric acid (1%)-catalyzed acetolysis of the

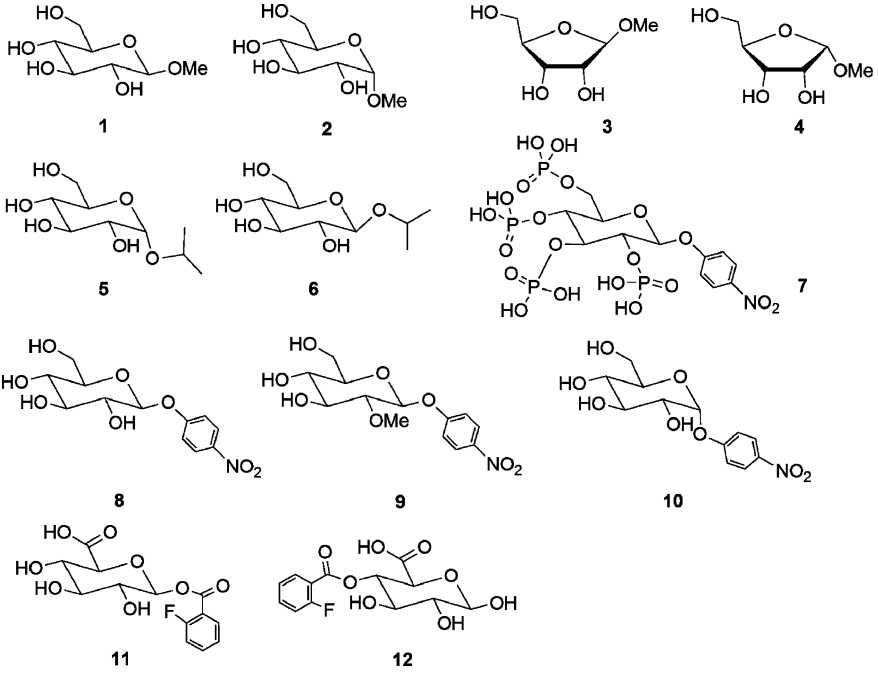
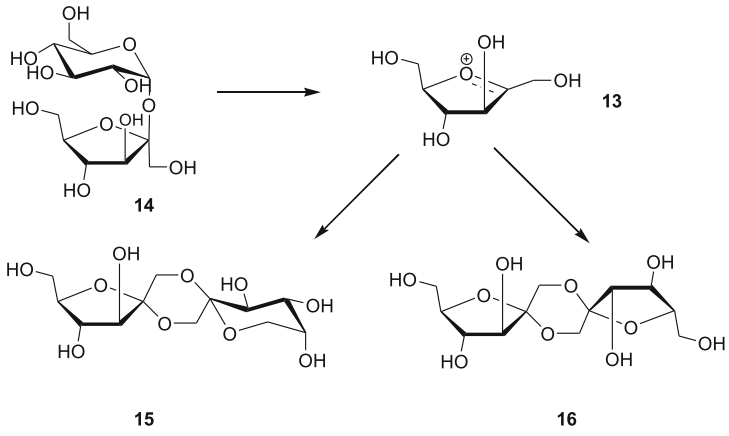


Figure 1
Glycosidesubstrates for nucleophilic hydrolysis



Scheme 1

anomeric ethyl glycoside derivatives **17** and **18** as well as the diastereoisomeric acetal **19** and **20** have been studied kinetically. The time-dependent distribution of the acetolysis products from **17** and **18** shows that their rapid mutual anomerization precedes their acetolysis to **21** and **22**, undoubtedly by way of the *endo*-cyclic cleavage product **23**, the precursor of **19** and **20** [14] (*Fig. 2*).

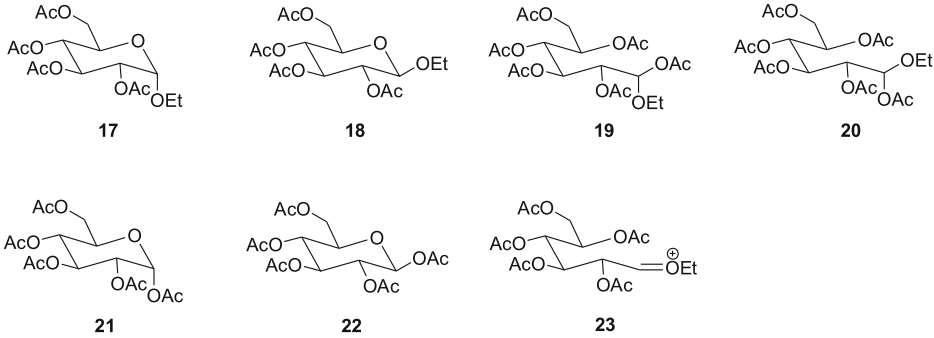


Figure 2
Substrates and intermediates in acetolysis of ethyl glycoside

Sialosides have a distinct mechanism of hydrolysis for its unusual sugar structure of sialic acid. For example, the large β-dideuterium and small primary ^{14}C kinetic isotope effects observed at the anomeric carbon and the large secondary ^{14}C kinetic isotope effect observed at the carboxylate carbon in the acid-catalyzed solvolysis of CMP-*N*-acetyl neuraminate **24** support an oxocarbenium ion-like transition state **25** having the $_5S$ conformation without nucleophilic participation of carboxylate and with the carboxylate anion in a looser environment than in the ground state [15] (*Fig. 3*). Such a zwitterion structure is consistent with the results from calculations using the COSMO-AM1 method for aqueous solutions [16].

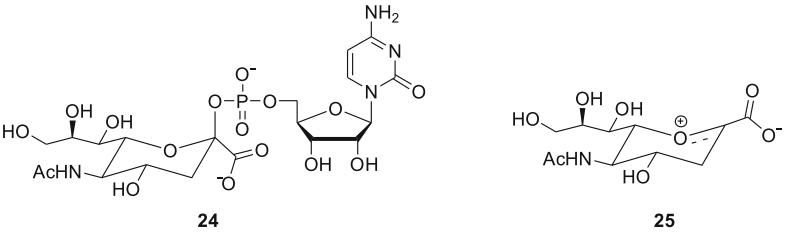
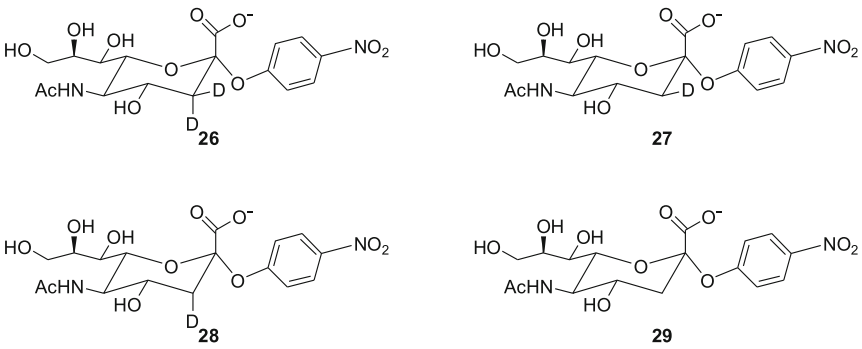


Figure 3
CMP-*N*-acetyl neuraminic acid and oxocarbenium ion-like transition state in sialoside hydrolysis

2.2 Enzymatic Hydrolysis

Glycoside hydrolase is one of the main categories of hydrolases in nature. Many references have suggested a distorted conformation for the substrate, which accelerated the hydrolytic process dramatically [17,18]. The influenza A sialidase hydrolyzes sialyl glycosides with retention of the anomeric configuration [16], whereas the *Salmomella typhimurium* sialidase works with inversion, although their protein folds and presumed active site residues are very similar [19,20]. Comparative studies using deuterium-labeled*p*-nitrophenyl *N*-acetyl-α-neuraminides **26**–**28** have postulated that the reactive substrate adopts a B_{25} conformation with

significant proton donation to the leaving group for the influenza virus enzyme, whereas the *S. typhimurium* enzyme works through a single chemical transition state derived from the ground state $^{2}C_{5}$ conformation with little proton donation to the leaving group [13]. The leaving group ^{18}O isotope effects are higher at pH 6.67 and 60 °C than at pH 2.69 and 50 °C in the nonenzymatic hydrolysis of **29** (◉ *Fig. 4*) [21,22]. This indicates that the C/O bond dissociation is complete at the transition state [21].

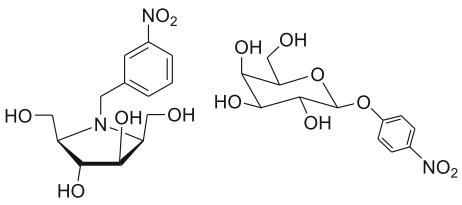


◘ Figure 4
Deuterium-labeled *p*-nitrophenyl *N*-acetyl-α-neuraminides

New analytic tools can be helpful in the research of enzymatic processes. For instance, time-course examination of enzymatic hydrolysis has recently been studied with ^{1}H-NMR spectroscopy. Thus, α-L-rhamnosyl and α-D-galactosyl hydrolysates from *Aspergillus* fungi have recently been found to be inverting hydrolyses [23,24].

With newly gained knowledge of hydrolytic mechanisms a novel artificial enzyme, the antibody enzyme AbZyme, was designed using the known mimics of the transition states [25]. Antibody Ab24, produced by in vitro immunization using the carrier-free hapten **30** and spleen cells in culture, catalyzes the hydrolysis of **8** with a k_{cat} of 0.02 h^{-1} and K_m of 160 µM ($k_{cat}/k_{uncat} = 2.2 \times 10^{4}$). Similarly, antibody Ab21 can catalyze the hydrolysis of galactoside **31** with a k_{cat} of 0.035 h^{-1} and K_m of 310 µM ($k_{cat}/k_{uncat} = 2.5 \times 10^{4}$) (◉ *Fig. 5*) [26].

At the same time, with the progress in development of new separation techniques and biotechnology, more and more enzymes have been found with interesting properties. For instance,



◘ Figure 5
Hapten and substrate for catalytic antibody Ab24

the α-1,4 glucan lyase (EC 4.2.2.-) from *Gracilariopsis lemeneiformis* is a new class of starch/glycogen degrading enzyme that digests the substrate from the nonreducing end while releasing 1,5-anhydro-D-fructose **32** successively, instead of the usual D-glucose **33** (*Fig. 6*) [27,28,29].

32 **33**

Figure 6
1,5-anhydro-D-fructose and D-glucose

Several recent books review the enzymes used in the conversion of renewable feedstocks such as starch and cellulose [30,31,32]. They provide many examples of the use of enzymes in the resource sector, specifically addressing their use in agriculture, forest products, and pulp and paper; they also address the greater use of agriculture and forestry residues and possible enzymatic modification. One recent example is the use of crude α-galactosidase from *Gibberella fujikuroi* to reduce the flatulence-inducing raffinose family sugars in chickpea flour. Crude enzyme treatment of chickpea flour resulted in complete hydrolysis of sugars of the raffinose family [33,34].

3 Degradation of Free Sugars

Although the free sugars are important industrial starting materials, they have not been focused upon until the recent findings of their degradation into useful organic resources [35,36].

3.1 Thermal Degradations

Thermal degradations in aqueous carbohydrate solutions are well documented [37,38,39]. Innovation of technology for the degradation of phytomass has focused upon the production of pyrolysis oil with high H/C and C/O ratios. Hydrothermal degradation appears to be attractive from this point of view [36,40,41,42]. Early in 1964, Qua and Fagerson tentatively identified furfural **34**, dihydroxyacetone **35**, glycolic acid **36**, glycolaldehyde **37**, and 5-(hydroxymethyl)-2-furaldehyde (HMF) **38**, and noted the presence of six additional volatile products from glucose **33** heated at 250 °C for 1 min in air. Recently, the scientists found that HMF **38** can be utilized as a very important intermediate for the petroleum industry [43,44,45].
For example, a problem is the formation of **35** during autoclave sterilization of various solutions for parenteral injection containing **33** as an excipient or a nutritional carbohydrate [46]. Similar degradations occur during food processing, especially in soft-drink production, with compound L-ascorbic acid **39** degrading into **34** and **35** [47,48]. Another example includes the roasting of coffee, during which there are many aphilic acids formed by carbohydrate degradation, which contribute to the smell and taste impact for coffee beans [49,50,51].

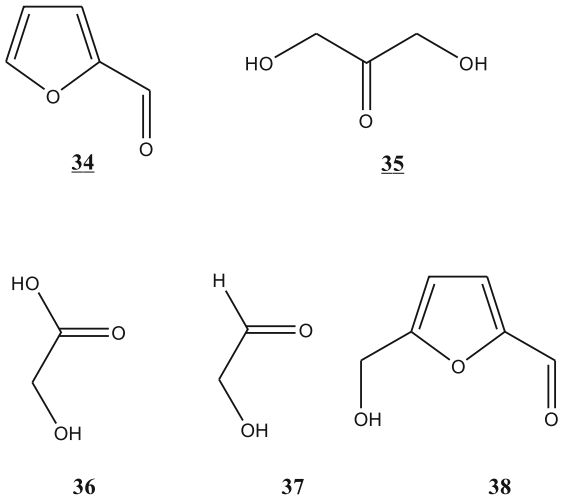
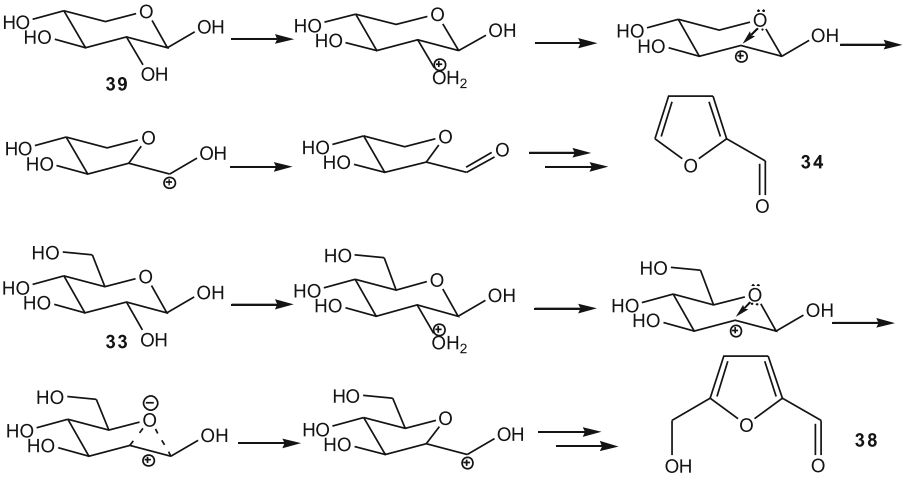


Figure 7
Main thermal degradation products from D-glucose

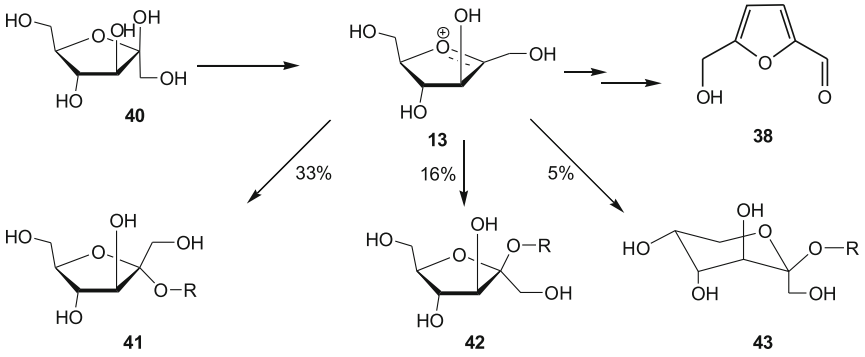
Ab initio molecular dynamics (MD) simulations were also applied in elucidation of xylose and glucose degradation pathways (*Scheme 2*). In the case of D-xylose **39**, a 2,5-anhydride intermediate was observed leading to the formation of furfural **34** through elimination of water. This pathway agrees with one of the mechanisms proposed in the literature in that no open chain intermediates were found. In the case of D-glucose **33**, a series of intermediates were observed before forming the 2,5-anhydride intermediate that eventually leads to HMF **38** (*Fig. 7*). One of these intermediates was a very short-lived open-chain form. Furthermore, two novel side-reaction pathways were identified, which lead to degradation products other than **38** [52].



Scheme 2

3.2 Acidic Degradations

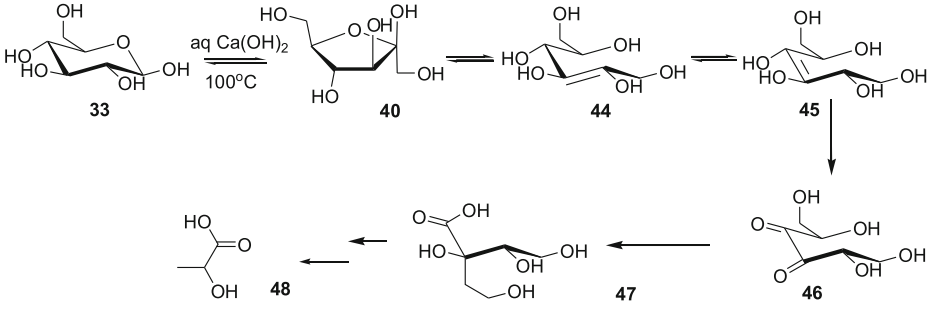
Alkyl glycosides are environmentally benign biosurfactants due to their biodegradability and low toxicity [53]. Usually they are produced through Fisher glycosylation using hydrophobic alcohols in acidic media. A practical problem is the control of the degradation reactions of starting free sugars in the acidic Fischer reaction. The situation is more serious in the case of D-Fructose **40**, which degrades into **38** [54,55,56] (*Scheme 3*). In Fischer reactions of this type, silica-alumina cracking catalysts effectively catalyze reactions to give the glycosides **41**, **42**, and **43**, without formation of **38** [57]. These results convincingly indicate that both the glycosylation giving the furanosides **41** and **42** and the degradation to **38** proceed via the common cyclic intermediates **13** [54].



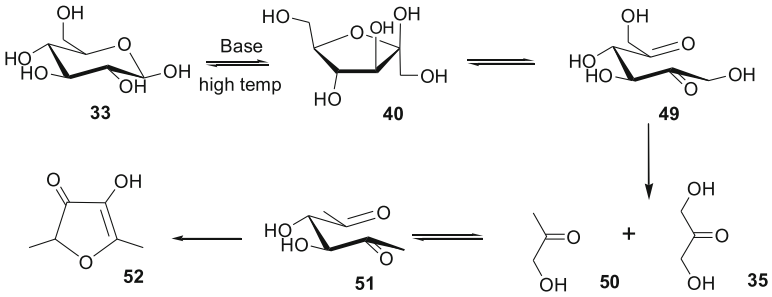
Scheme 3

3.3 Alkaline Degradations

The alkaline degradation of reducing monosaccharides involves a series of consecutive reactions and gives many kinds of products [58]. For example, the alkaline degradation of **33** in aqueous calcium hydroxide at 100 °C results in a complex mixture of more than 50 compounds (*Scheme 4*). Products obtained by the same degradation of **40** are similar to those from the



Scheme 4



Scheme 5

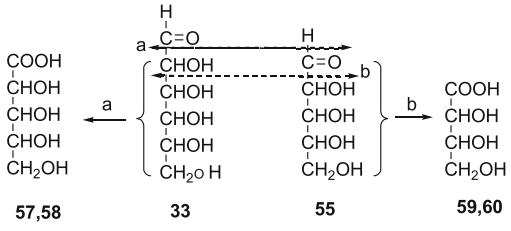
reaction of **33**. Among the degradation products, lactic acid **44** is almost the sole major product in each case [59].

High-temperature alkaline degradation of **33** forms furaneol (**52**), an aroma compound, probably because of fragmentation of **49** into the C_3-fragments **35** and **50** (*Scheme 5*). Fragment **50** dimerizes into the diketone **51**, the precursor of **52** [60].

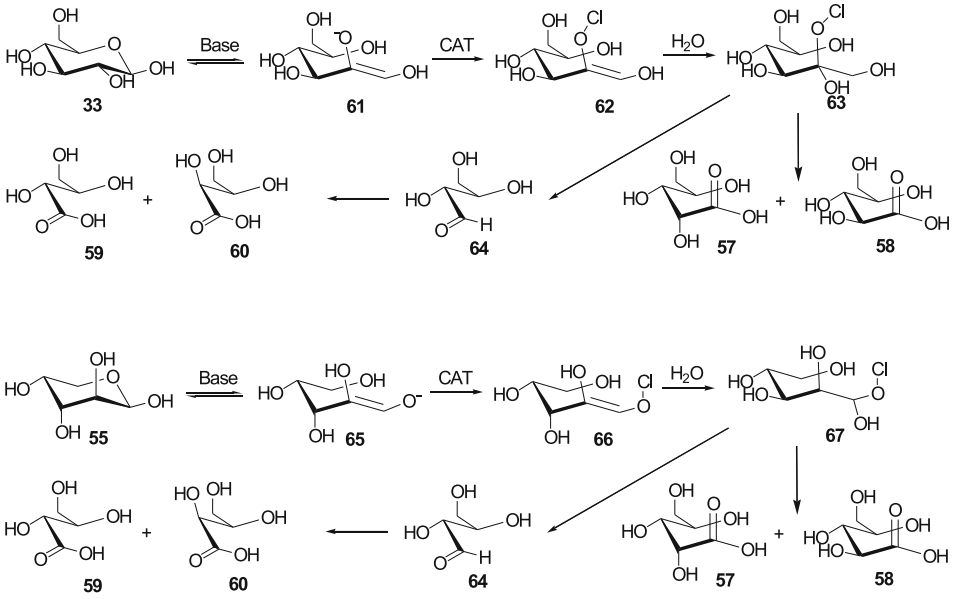
3.4 Oxidative Degradations

Oxidative degradation reactions involving the anomeric center are classic processes and are well documented [61,62]. For example, lactose, maltose, cellobiose, and galactose can be degraded selectively in one step and in high yield into the corresponding next lower aldose and formic acid by H_2O_2 in the presence of borate. The selectivity further improves when a small amount of EDTA is added, in order to suppress the influence of transition metal ions, which catalyze the decomposition of H_2O_2 via radical pathways, leading to nonselective oxidative degradation of aldoses. The function of borate in the selective oxidative degradation of aldoses is two-fold: catalysis of the degradation of the starting aldose and protection of the next lower aldose against oxidation [63].

On alkaline oxidation of aldoses with (*N*-chloro-*p*-toluenesulfonamido) sodium (CAT), the monosaccharides **33**, **40**, D-mannose **54**, D-arabinose **55**, and D-ribose **56**, belonging to the 4,5- or 3,4-*ethythro*-series, afford the C_4-acids **59** and **60** in 35 to 49% yields while the yields of glyceric acid are low [64]. Thus, as illustrated in *Scheme 6*, hexoses are cleaved at the C1/C2 (a) and C2/C3 (b) bonds, whereas pentoses break at the C1/H1 (a) and C1/C2 (b) bonds.



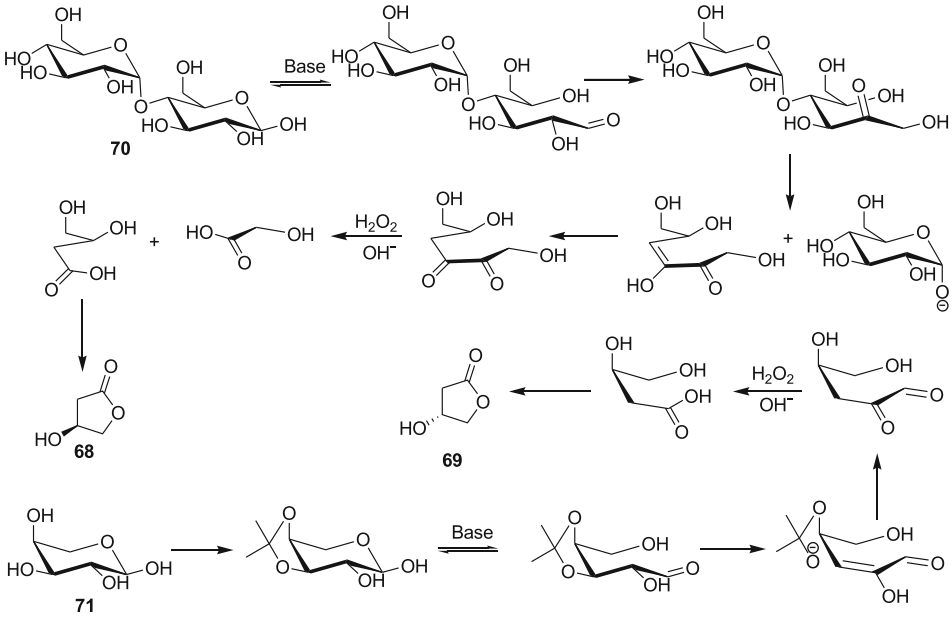
Scheme 6



Scheme 7

These reactions are governed by the alkaline-induced slow equilibrium between hexoses and enediol anions and the irreversible, rate-determining formation of the intermediate **63** (*Scheme 7*). The latter is transformed into **57** and **58** or **64**. In the case of pentose **55**, the intermediate **67** gives out **59** and **60** from the intermediate **64**.

(S)- and (R)-3-hydroxy-γ-butyrolactone (**68** and **69**, respectively) are two extremely flexible chiral synthons. They can be converted to an extremely large number of useful and important intermediates with a wide range of applications. Earlier synthetic routes to these compounds all relied on structural transformations or selective reductions of malic acid. They can now be obtained in high yield from several carbohydrate raw materials. For example, the (S)-lactone **68** can readily be prepared by the oxidation of 4-linked D-hexose sources such as cellobiose, lactose, maltose, maltodextrins, starch, etc., with hydrogen peroxide and an alkaline or alkaline-earth hydroxide. Treatment of a 4-linked hexose **70** with base leads to an isomerization to the 4-linked ketose, which readily undergoes β-elimination to form enone, which then tautomerizes to the diketone. The diketone is readily cleaved with hydrogen peroxide to give the salt of (S)-3,4-dihydroxybutyric acid and glycolic acid. Acidification and concentration yields the lactone **68** [65]. Similarly, the (R)-lactone **69** can be synthesized using a 4-linked L-hexose source since the chiral center in the product is derived from the 5-carbon of the hexose. The (R)-lactone was obtained in high yield from L-arabinose **71** by the simple strategy of functionalizing the 3-position by forming a 3,4-acetal and oxidizing it under similar conditions as those used for the preparation of the other isomer. This oxidation yields the dihydroxy acid and formic acid via the unsaturated aldehyde which tautomerizes to the *R*-dicarbonyl compound. The dihydroxy acid is then converted to the lactone **69** by acidification and concentration (*Scheme 8*) [66].



Scheme 8

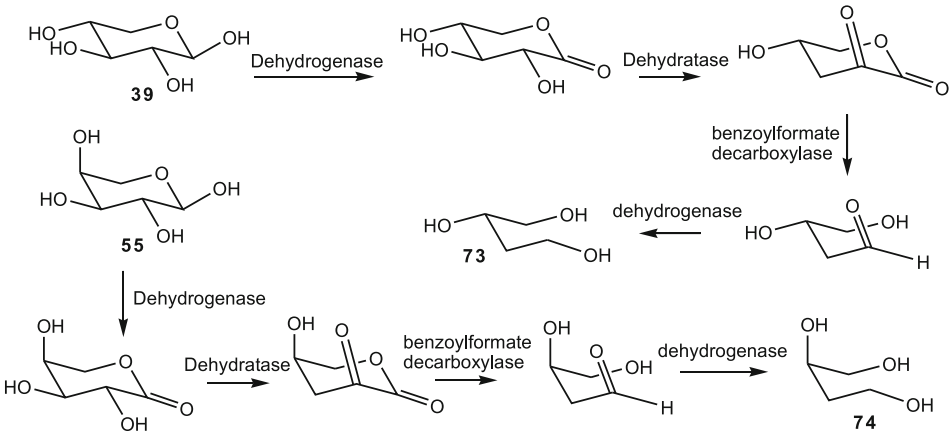
Titanium-containing zeolites, such as Ti-BEA, Ti-FAU, and TS-1 have been tested as catalysts for the Ruff oxidative degradation of calcium D-gluconate **72** to D-arabinose **55** using diluted hydrogen peroxide as the oxidant. Only large-pore zeolites Ti-BEA and Ti-FAU were found to be active. It was shown, in particular, that a very rapid leaching of titanium occurred and that the titanium species present in the solution were responsible for the catalytic activity observed [67,68].

Applying H_2O_2/CuO in alkaline solution, degradation of the carbohydrate-rich biomass residues results with formic, acetic and threonic acids as the main products. Gluconic acid was formed instead of glucaric acid throughout. Reaction of a 10% H_2O_2 solution with sugar beet molasses generated mainly formic and lactic acids. Important advantages of the microwave application were lower reaction times and reduced reagent demands [69].

3.5 Enzymatic Degradations

1,2,4-Butanetriol is an important intermediate in organic synthesis, for instance in the production of D,L-1,2,4-butanetriol trinitrate. Commercial synthesis of D,L-1,2,4-butanetriol employs $NaBH_4$ reduction of esterified D,L-malic acid. For every ton of 1,2,4-butanetriol synthesized, multiple tons of byproduct borates are generated. D,L-malic acid can also be hydrogenated over various catalysts (Cu–Cr, Cu–Al, Ru–Re) at 2900–5000 psi of H_2 and 60–160 °C reaction temperatures. Yields of 1,2,4-butanetriol range from 60 to 80%. A variety of byproducts are also formed during high-pressure hydrogenation. These byproducts are not generated when esterified malic acid is reduced using $NaBH_4$. D,L-malic acid is synthesized from the *n*-butane

component of liquefiable petroleum gas via the intermediacy of maleic anhydride. The new synthesis of 1,2,4-butanetriol has been established with microbes. Enzymes from three different microbes are recruited to create biosynthetic pathways by which D-1,2,4-butanetriol **73** and L-1,2,4-butanetriol **74** are derived from D-xylose **39** and L-arabinose **55**, respectively [70] (*Scheme 9*).



Scheme 9

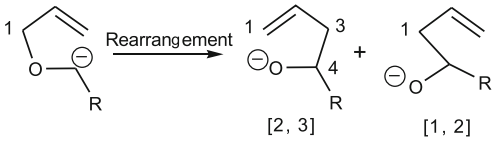
The use of ethanol as an alternative automobile fuel has been steadily increasing around the world for a number of reasons [71]. Domestic production and use of ethanol for fuel can decrease dependence on foreign oil, reduce trade deficits, create jobs in rural areas, reduce air pollution, and reduce global climate change carbon dioxide buildup. Ethanol, unlike gasoline, is an oxygenated fuel that contains 35% oxygen, which reduces particulate and NOx emissions from combustion. Ethanol can be made synthetically from petroleum or by microbial conversion of biomass materials through fermentation. In 1995, about 93% of the ethanol in the world was produced by the fermentation method and about 7% by the synthetic method. The fermentation method generally uses three steps: (1) the formation of a solution of fermentable sugars, (2) the fermentation of these sugars to ethanol, and (3) the separation and purification of the ethanol, usually by distillation. Ethanol-from-cellulose (EFC) holds great potential due to the widespread availability, abundance, and relatively low cost of cellulosic materials. However, although several EFC processes are technically feasible, only recently have cost-effective EFC technologies begun to emerge, which are quite important for rapidly developing countries such as China and Canada [72].

4 Rearrangement with Double Bond Shifts

4.1 [2,3]-Sigmatropic Rearrangements

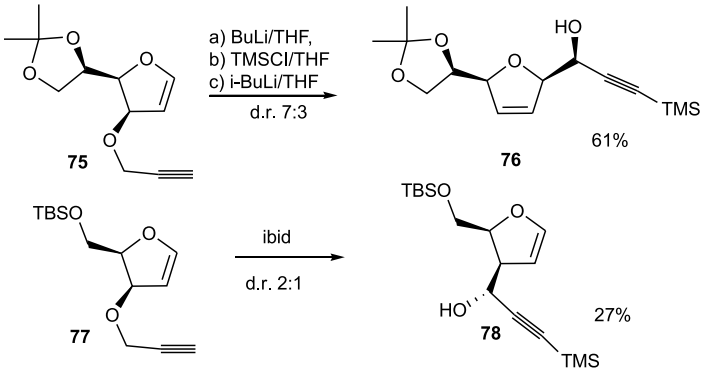
Double bond rearrangements in carbohydrate systems lead to various kinds of sugar transformations. The [2,3]-Wittig rearrangements [73] initiated by deprotonation and followed by

migration of an anionic substituent are illustrated in Scheme 10. The stereochemistry at C1 is well transformed to C3 and/or C4 [74,75]. The [1,2]-Wittig rearrangement without a double bond shift occurs in dependence on the conditions.



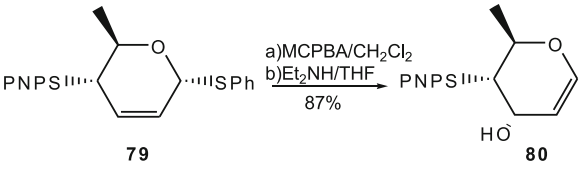
Scheme 10

The [2,3]-Witting rearrangement has been employed in synthetic work on the tetrahydrofuran acetogenins from *Annonaceous* species starting from furanoid glycals [76]. The rearrangement of **75** is induced by a base to generate an anionic species, which rearranges into **76** and its epimer. In this case, erythro-2 predominates. Under the same conditions, **77** with a silyl-protecting group mainly gives the [1,2]-Witting rearrangement product **78** (Scheme 11).



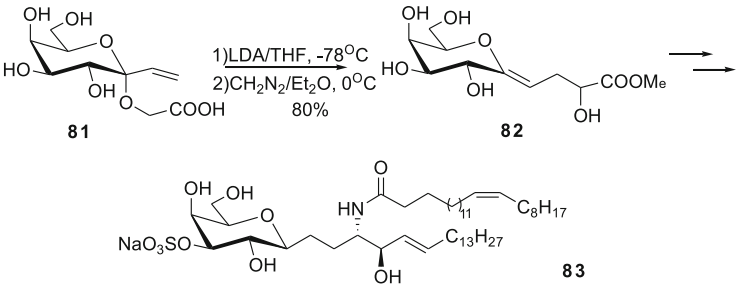
Scheme 11

A [2,3]-sigmatropic rearrangement of a sulfoxide has been employed in the total synthesis of calicheamicin g [77]. The thioglycoside **79** is oxidized to give the sulfoxide intermediate, which spontaneously undergoes a suprafacial sigmatropic shift to the β-position to move the double bond towards the anomeric center. The resulting sulfinate is treated with a secondary amine to afford the desired rearranged glycal derivative **80** (Scheme 12).



Scheme 12

Recently, the *C*-analogue of sulfatide **83** was synthesized through a [2,3]-Wittig sigmatropic rearrangement (*Scheme 13*) [78].

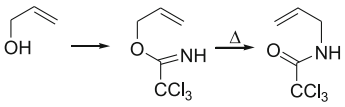


Scheme 13

4.2 [3,3]-Sigmatropic Rearrangements

4.2.1 Overman Rearrangement and Related Reactions

An allylic system that is easy to rearrange is a useful tool in a variety of synthetic methods. Allylic alcohol is readily converted into the corresponding trichloroacetimidate by brief treatment with trichloroacetonitrile in the presence of an appropriate base and usually results in high yields. Simple heating of the imidate of the allylic alcohol system induces the rearrangement reaction [79,80] (*Scheme 14*).

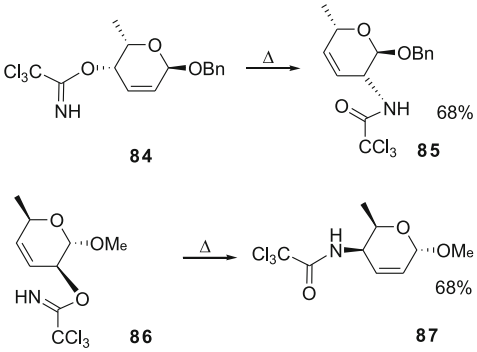


Scheme 14

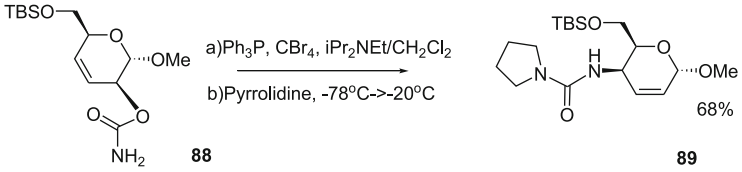
For example, on heating at 160 °C in 1,2-dichlorobenzene, the allylic trichloroacetimidate **84** smoothly rearranges into the corresponding 2-amino-2-deoxy sugar **85** [81]. The suprafacial rearrangement from C2 to C4 is similarly performed to obtain the 4-amino-4-deoxy sugar derivative **87** [82] (*Scheme 15*).

Under dehydrating conditions, the allylic carbamate **88** generates the allyl cyanate which, in turn, rearranges into the reactive allyl isocyanate and then reacts with nucleophiles [83] (*Scheme 16*). The allylic carbamate is prepared by treatment of the allylic alcohol with trichloroacetyl isocyanate in dichloromethane at 0 °C and chemoselective removal of the alkali-labile trichloroacetyl group by mild reaction with cold methanolic potassium carbonate without affecting the carbamate linkage at all. The obtained carbamate **88** is dehydrated by the triphenyl phosphine/tetrabromomethane system under very mild conditions. This leads to the reactive isocyanate via spontaneous rearrangement. The isocyanate thus generated is trapped with a nucleophile such as pyrrolidine to furnish the aminosugar derivative **89**.

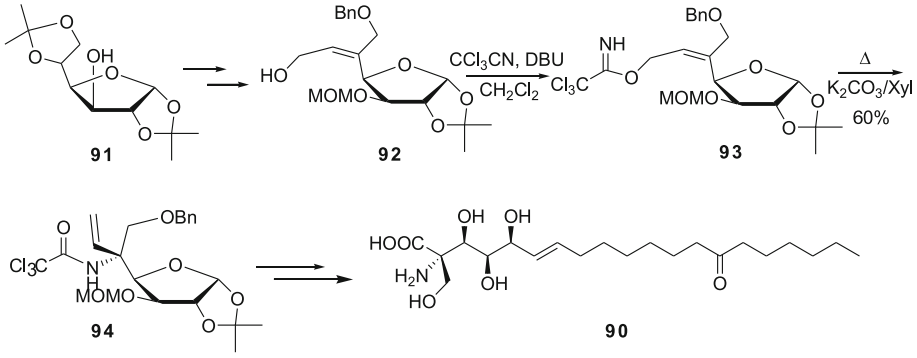
In the recent total synthesis of sphingofungin E (**90**), Overman rearrangement of an allylic trichloroacetimidate derived from diacetone-D-glucose **91** generated tetra-substituted carbon



Scheme 15



Scheme 16



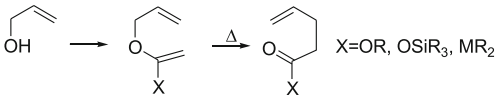
Scheme 17

with nitrogen (**94**), and subsequent Wittig olefination afforded the highly functionalized part in sphingofungin E stereoselectively [84] (*Scheme 17*).

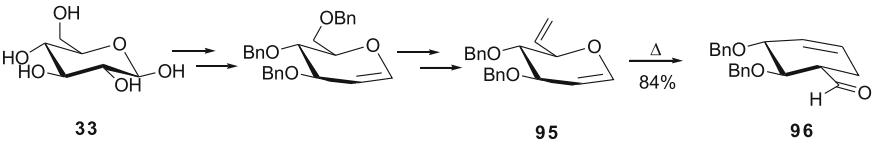
4.2.2 Modified Claisen Rearrangements

Modifications of the Claisen rearrangement have been widely used in a variety of synthetic chemistry reactions [85] (*Scheme 18*).

The simple Claisen rearrangement itself has been employed in the transformation of the vinylglycal **95** into carbocyclic compounds [86] (*Scheme 19*). On heating at 240 °C in o-dichlorobenzene in a sealed tube for 1 h, the desired rearrangement of **95** proceeds in the



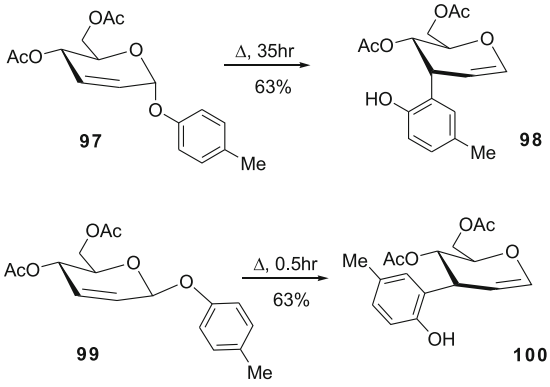
Scheme 18



Scheme 19

expected direction to give the unsaturated carbocyclic system bearing an aldehyde function in 84% yield. This is a useful synthetic intermediate for a variety of pseudosugars.

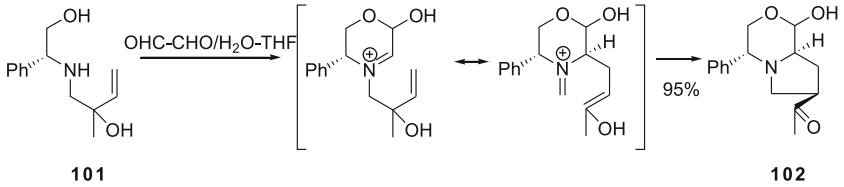
Aromatic Claisen rearrangements in 2,3-unsaturated sugar systems are useful for the stereocontrolled synthesis of aryl-branched sugars [87] (*Scheme 20*). The α-anomer **97** is much less reactive in comparison to the β-anomer **99**. This thermal rearrangement is carried out by refluxing in *N*,*N*-diethylaniline. The efficiency of the reaction is almost independent of the nature of the p-substituent in the phenyl group.



Scheme 20

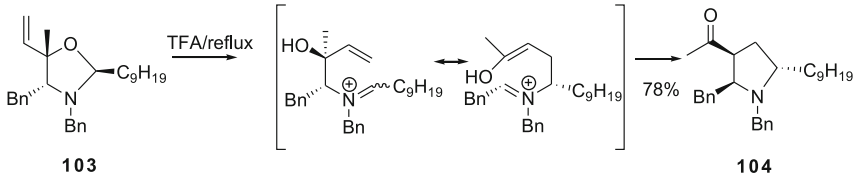
4.2.3 Hetero-Cope Rearrangements

Cationic aza-Cope/Mannich tandem reactions [88,89,90,91] have been applied to the asymmetric synthesis of homochiral proline derivatives (azafuranosides) [92]. The β-amino alcohol **101** reacts with glyoxal at room temperature to generate a cyclic aminoacetal, which undergoes spontaneous dehydration to give rise to the ene-iminium intermediate (*Scheme 21*). Through an aza-Cope reaction, this cationic species transforms into the bond-rearranged exomethylene intermediate. Then a Mannich-type cyclization takes place to give the homochiral proline derivative **102** quantitatively.



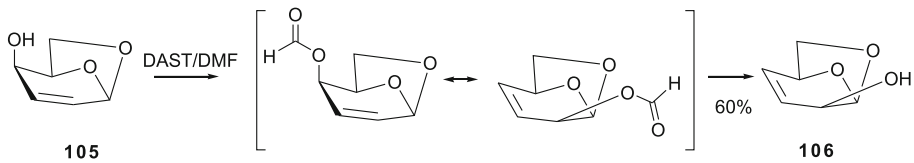
Scheme 21

This protocol was later employed for the synthesis of (−)-α-allokainic acid [93]. Tandem aza-Cope/Mannich reactions of this type have also been employed to construct the framework of (−)-preussin [90]. On refluxing in trifluoroacetic acid, the protonated and functionalized oxazolidine **103** changes into the ene-iminium intermediate, which equilibrates with the bond-rearranged enol compound through an aza-Cope process (*Scheme 22*). This is followed by cyclization through the Mannich reaction to give the functionalized pyrrole **104** in 78% yield with 86% ee. To proceed to (−)-preussin, a retro-Mannich fragmentation-Mannich cyclization of **104** is needed to establish the desired configuration of the pendants on the pyrrolidine ring.



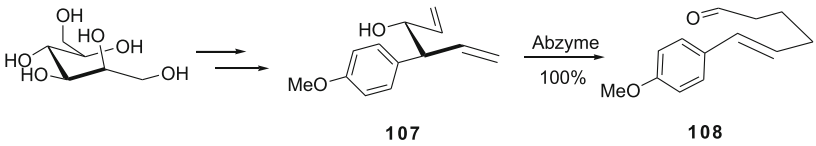
Scheme 22

A neutral, metal-free rearrangement, formally a suprafacial [1,3]-sigmatropic migration, of the hydroxy group has been reported [94] (*Scheme 23*). The direct migration of the hydroxy group is thermally forbidden. This rearrangement reaction probably proceeds by way of an intermediate formate, which undergoes an oxy-Cope rearrangement. Diethylaminosulfur trifluoride (DAST) is considered to react with the solvent DMF to generate the reactive quaternary amine salt, which rapidly converts **105** into the corresponding formate. The hypothetical 4-*O*-formate would then undergo acyloxy group migration accompanied by a double bond shift through an oxy-Cope rearrangement to give the 2-*O*-formate. Thus, the net results are suprafacial 1,3-shifts of the hydroxy group from the C4 to the C2 position. No substitution reaction of the hydroxy group with the fluoride ion seems to occur during this reaction.



Scheme 23

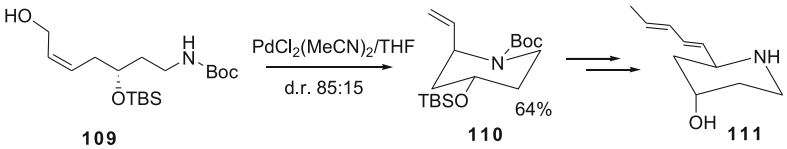
As an artificial enzyme, AbZyme was applied in a similar [3,3]-sigmatropic rearrangement [95,96]. The substrate **107** is prepared from a diacetone-D-mannitol through conventional synthetic transformations. On exposure to the artificial polychronal antibody in the presence of 2-(*N*-morpholino)ethanesulfonic acid and sodium chloride at 37 °C, with a molar ratio of 100:1 of the hexadiene **107** to the antibody, **107** is completely converted into the product **108** in 20 h (*Scheme 24*).



Scheme 24

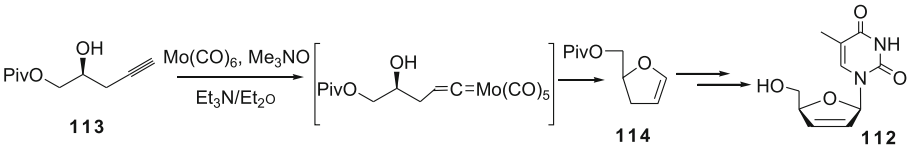
4.3 Double Bond Inducing Ring-Closing Rearrangements

There are a variety of examples of ring-closing rearrangements with exhausting double and/or triple bonds and some recent examples are shown here. Pd(0)-complexes catalyze reactions of the unsaturated amine **109** to give the azasugar **110**, an intermediate in the synthesis of SS20846A **111** [97,98,99,100] (*Scheme 25*).



Scheme 25

Besides the conventional methods, the metallo-carbene route to access cyclic compounds has become a versatile tool in sugar chemistry. Synthesis of stavudine **112**, an antiviral nucleoside, from an allyl alcohol [101] is realized by a $Mo(CO)_5$-mediated cyclization reaction (*Scheme 26*). Molybdenum hexacarbonyl smoothly reacts with the triple bond of **113** to generate the intermediate Mo-carbene, which undergoes a clean cyclorearrangement to yield the furanoid glycal **114**. Alkynol isomerization is effected by group-6 transition metal carbonyl complexes [102].



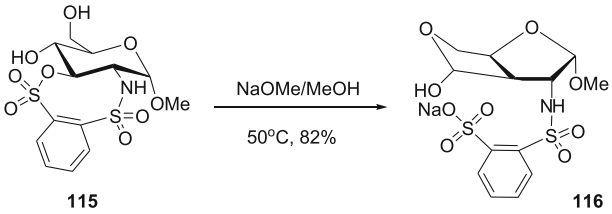
Scheme 26

5 Ring Isomerizations

Ring transformations are useful reactions in synthetic carbohydrate chemistry [103].

5.1 Ring Contractions

Nucleophilic displacement reactions of the sulfonyloxy group or its equivalents in the sugar ring are known to induce unexpected ring-contraction reactions [104]. However, the first target-oriented ring-contraction reaction of the sulfonate **115** (Scheme 27) in the stereoselective total synthesis of (−)-rosmarinecine from D-glucosamine impressively demonstrated the novel utility of this kind of reaction [105]. Ring-contraction reactions of carbohydrates have now become a useful tool for syntheses of various types of compounds [103].

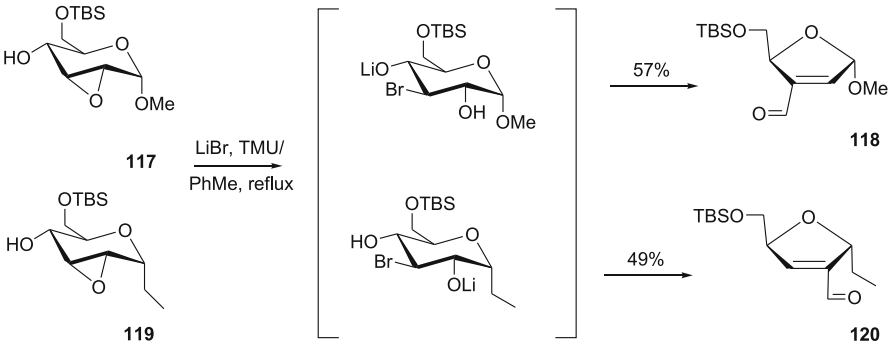


115 **116**

Scheme 27

Epoxy sugars are good starting materials for the preparation of ring-contracted products. Various 2,3-epoxypyranosides such as **117** can be converted into furanosides directly by simple heating under reflux in toluene containing lithium bromide and *N*,*N*-tetramethylurea (TMU) [106] (Scheme 28). Usually, the more stable 3-*C*-formyl derivatives are formed. In the case of the *C*-glycoside **119**, however, the mode of reaction changes to yield mainly the 2-*C*-formyl compound **120**.

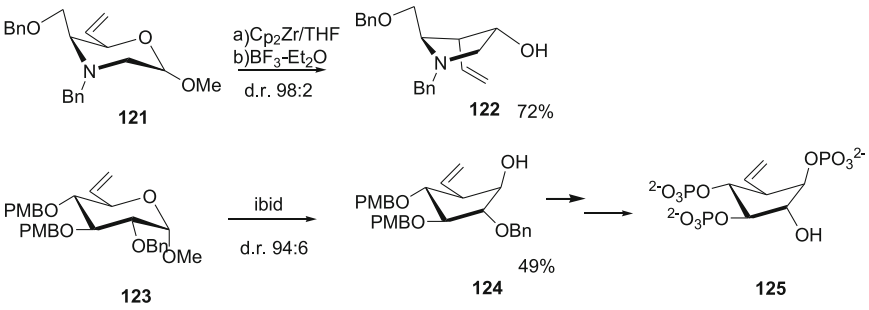
Zr-mediated ring-contraction reactions using vinyl sugars are useful to synthesize carbocycles [107,108]. This method was later employed successfully for aza-sugar synthesis. The



117 **119** **118** **120**

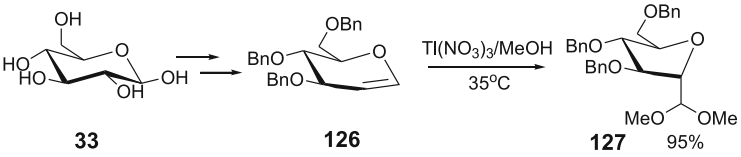
Scheme 28

functionalized morpholine **121** is transformed into the pyrrolidine **122** with excellent stereoselectivity. The stereochemistry at the junction of the main product is *cis*. This protocol has been applied to the synthesis of inositol phosphate analogs using the 5-*C*-vinyl glycoside derivative **123** [109] (● *Scheme 29*).



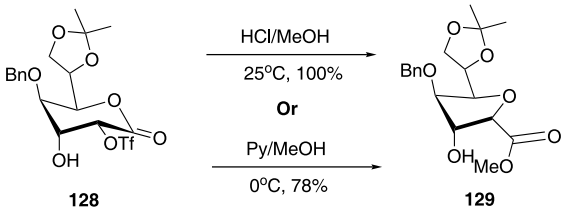
Scheme 29

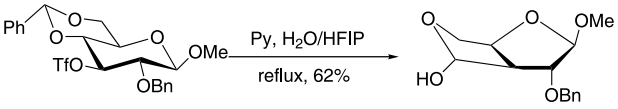
The *O*-benzyl derivative of the glycal **126** undergoes stereoselective ring contractions on treatment with thallium(III) nitrate [110] (● *Scheme 30*).

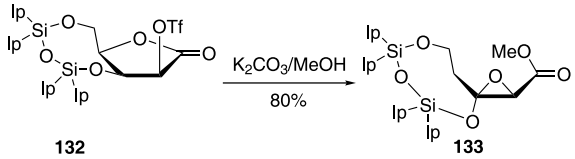


Scheme 30

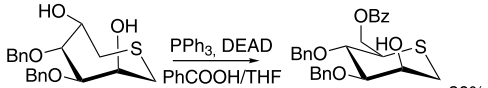
Triflates of aldonolactones are a productive source of ring-contraction reactions. Compound **128** contracts its ring under acidic and basic conditions to give *C*-furanosides [111,112] (● *Scheme 31*). Triflates of glycosides occasionally yield ring-contracted products [113]. Another paper has provided an additional example of a ring-contraction reaction of a sugar triflate on reaction with tetrabutylammonium nitrite in moist toluene; the triflate yields a ring-contracted byproduct [114]. It has been found that the ring-contraction reactions of the triflate in the presence of pyridine depend on the acidity of the solvent; in the acidic solvent 1,1,1,3,3,3-hexafluoro-2-propanol (HFIP), ring-contraction of **130** proceeds smoothly. Oxygenophilic silyl protection of the 2-*O*-triflate of aldonolactone **131** prohibits ring contraction so that the reaction results in the formation of the silyl-migrated epoxide **132** [115].
Mitsunobu conditions smoothly effect clean ring-contraction reactions of thiosugars [116] (● *Scheme 32*). From the thioheptanoid **134**, the thiopyranoid **135** is obtained. Mild sulfonation induces a spontaneous ring-contraction of the azaheptanoid **136** to afford the azapyranoside **137** [117].
Some rare, four-membered sugar rings have been synthesized by ring-contraction reactions [103]. Similar to 2-*O*-triflates, which easily undergo ring-contraction reactions [118]. DAST-treatment of the thiopentofuranose derivative **138** affords the ring-contracted product **139** having a thietane framework (● *Scheme 33*). DAST-assisted ring-contraction has

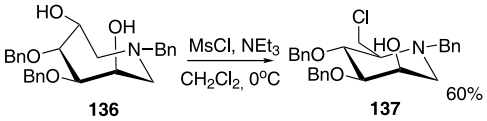






Scheme 31



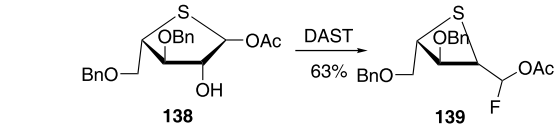


Scheme 32

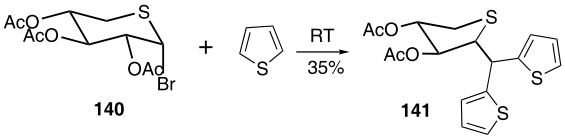
been found in the fluorination reaction of the thiosugar furanose **138** [119]. It is known that the sulfur(IV) fluoride/hydrogen fluoride system also promotes such ring-contraction reactions [120]. On Friedel–Craft reaction of the thiopentosyl bromide **140**, a ring-contraction process occurs [121]. The per-*O*-alkylated glycoside **142** is converted into the δ-lactone **143** with concomitant ring-contraction to furnish product **144** [122].

The Chan rearrangement was effectively used to build up the furanoid structure on the way to taxol [123,124] (*Scheme 34*).

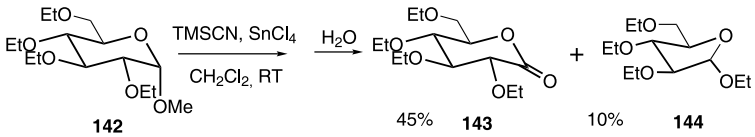
In the syntheses of staurosporin congeners, ring-contraction reactions have been used effectively [125,126]. Novel stereoselective Beckmann-type rearrangement of TAN-1030A **147** produces the K-252 analog **148** via a hypothetical hemiacetal intermediate [127]. Oxidation of the model compound **149**, the staurosporin analog, results in ring contractive benzilic acid rearrangement to give a furanoid **150** possessing the framework of K252a [128] (*Scheme 35*).



138 139

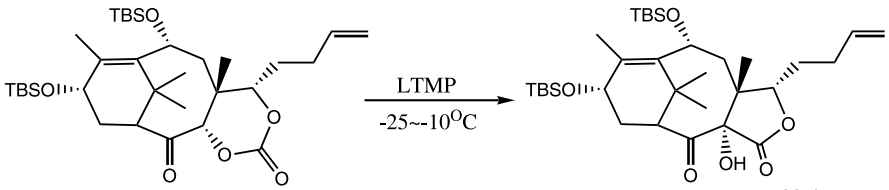


140 141



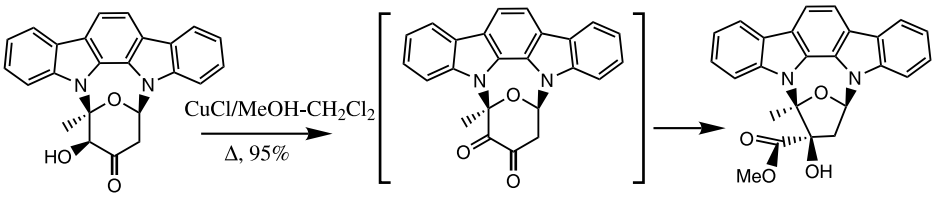
142 45% 143 10% 144

Scheme 33



145 146

Scheme 34



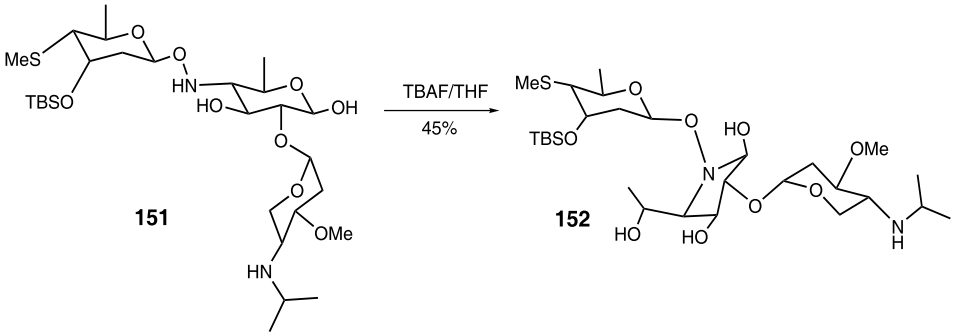
149 150

Scheme 35

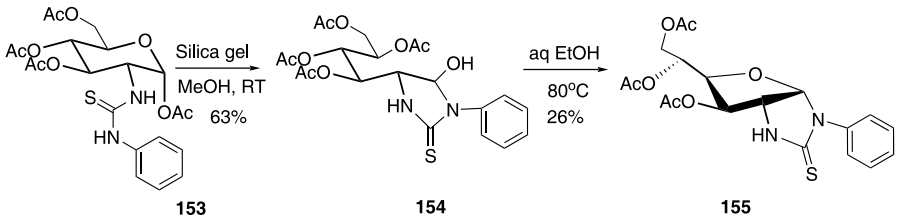
The protected oxyaminoglucoside **151** rearranges to the azafuranose form under deprotecting conditions [77] (*Scheme 36*).

The mild reaction of the thioureido derivative **153** with methanol produces the compound **154** with migration of the acetyl group. On heating, this compound isomerizes into the *cis*-fused cyclic **155** [129] (*Scheme 37*).

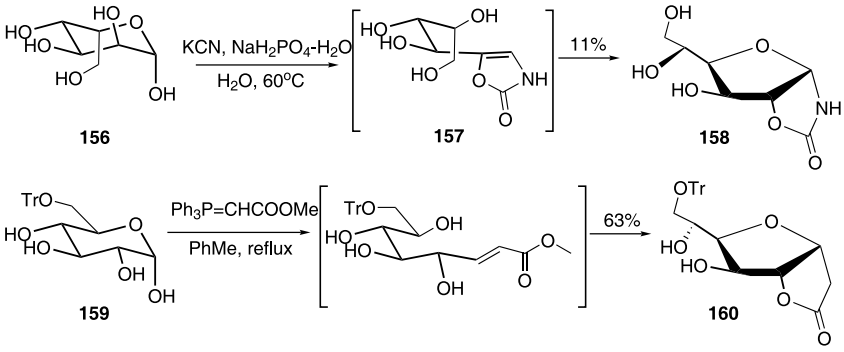
Epimerization at C2 of L-gulose **156** on reaction with KCN in buffered aqueous solution is thought to proceed by way of the open-chain intermediate. Free sugars produce cyclic products directly on reaction with the Wittig reagent. Thus, **159** is converted into **160** on prolonged heating with the reagent (*Scheme 38*).



Scheme 36



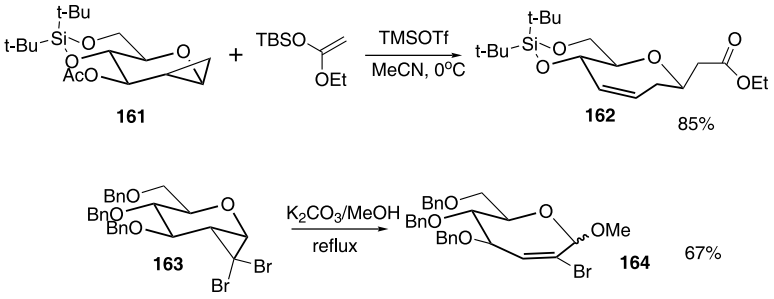
Scheme 37



Scheme 38

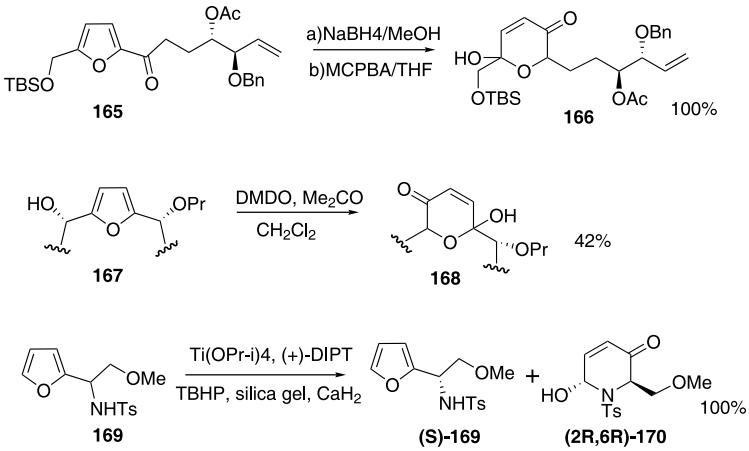
5.2 Ring Expansions

The cyclopropane system is a tool for inserting a methylene unit into a ring system to form a larger ring structure. Even densely functionalized pyranoids such as **161** [130,131] and cyclohexanes [132] expand into heptanoids and cycloheptanes, respectively. 1,2-*C*-dibromomethylene sugar **163** expands its pyranose ring to give oxepine **164** [133] (*Scheme 39*).



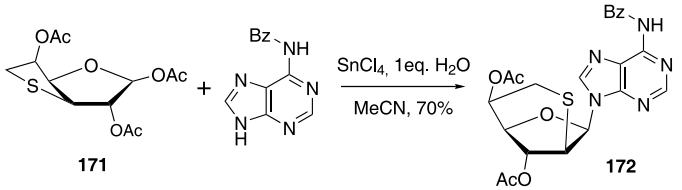
Scheme 39

α-Hydroxyfurans expand to pyranoids via sequential epoxyalcohol rearrangements. Epoxidation of the α-hydroxyfuran **165** with *meta*-chloroperoxybenzoic acid (m-CPBA) induces a cationic rearrangement, followed by dehydration, to form a pyranoid on the way to (+)-resineferatonin [134] (*Scheme 40*). Dimethyldioxirane (DMDO), apparently more sensitive to the steric circumstances than m-CPBA, has been used for the selective epoxidation of the furan **167** [135]. The monoepoxide rearranges followed by hemiacetalization to afford the pyranoid intermediate **165** of the total synthesis of the eleuthesides. α-Aminofurans similarly expand into azapyranoids [136]. Racemic **169** is kinetically resolved to give (S)-**169** and the rearranged (2R, 6R)-**170** with modified Sharpless epoxidation. Compound (S)-**169** is transformed into (2S, 6S)-**170** on treatment with m-CPBA.

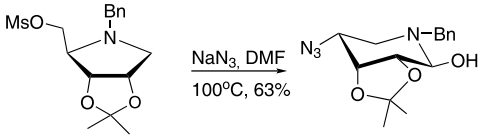


Scheme 40

The reaction of the bicyclic thiosugar **171** with N6-benzoyladenine in the presence of moist tin(IV) chloride furnishes the ring-rearranged nucleoside product **172** instead of the normal glycosylation product [137]. On heating mesylate **173** under reflux in the presence of a nucleophile, the thermal ring-expanding reaction occurs [138] (*Scheme 41*).

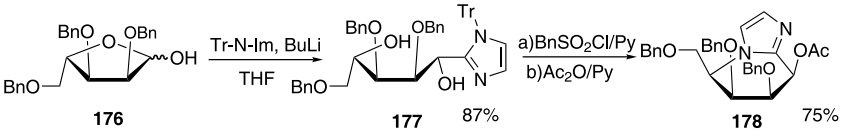


171 172



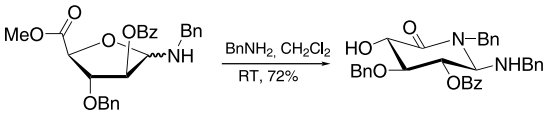
173 174

Scheme 41

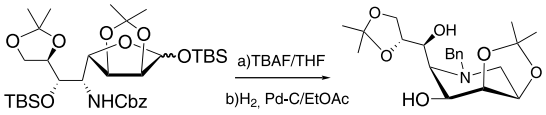


176 177 87% 178 75%

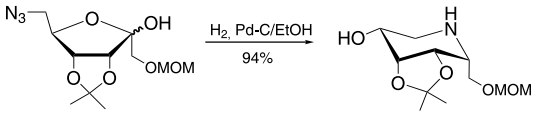
Scheme 42



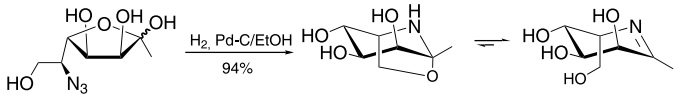
179 180



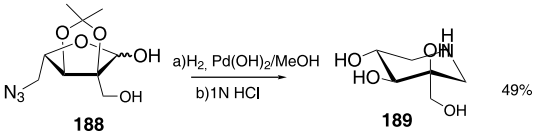
181 182



183 184



185 186 187



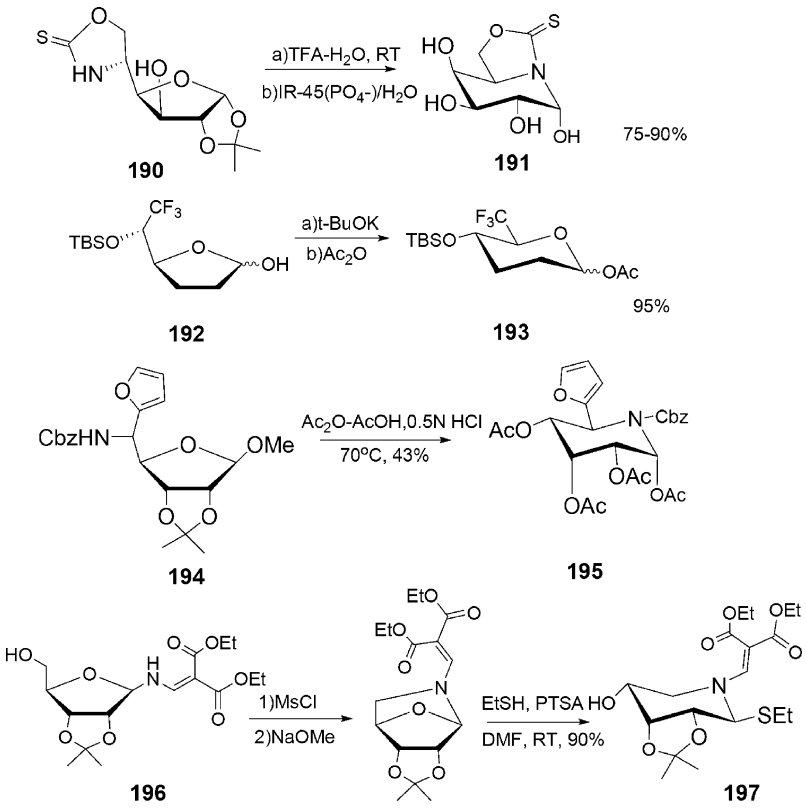
188 189

Scheme 43

The key compound on the way to the debranched nagstatin **175** has been synthesized from the L-ribose derivative **176** by employing ring-chain interconversion involving addition of trityl imidazole, selective sulfonylation, and warm acetylation which causes detritylation with cyclization [139] (*Scheme 42*).

The glycosylamine **179** transforms to the piperidinone **180** on reductive amination [140]. Sequential deprotection to regenerate hemiacetal OH and amino groups from **181** induces ring interconversion and reduction to give the azasugar **182** [141]. Similarly, the azidodeoxyketose derivative **183** can be converted to the piperidine derivative **184** by reductive aminocyclization [142]. Reduction of **185** affords the bicyclic azasugar **186** on intramolecular reductive cyclization, which is not a stable system and forms an equilibrium mixture with the monocyclic imine **187** [143]. Reactions of this type are also of use for the synthesis of the branched-chain 1-*N*-iminosugars such as **189**, which have been the subject of continuous attention as glycosidase inhibitors [144,145,146,147,148,149,150,151,152] (*Scheme 43*).

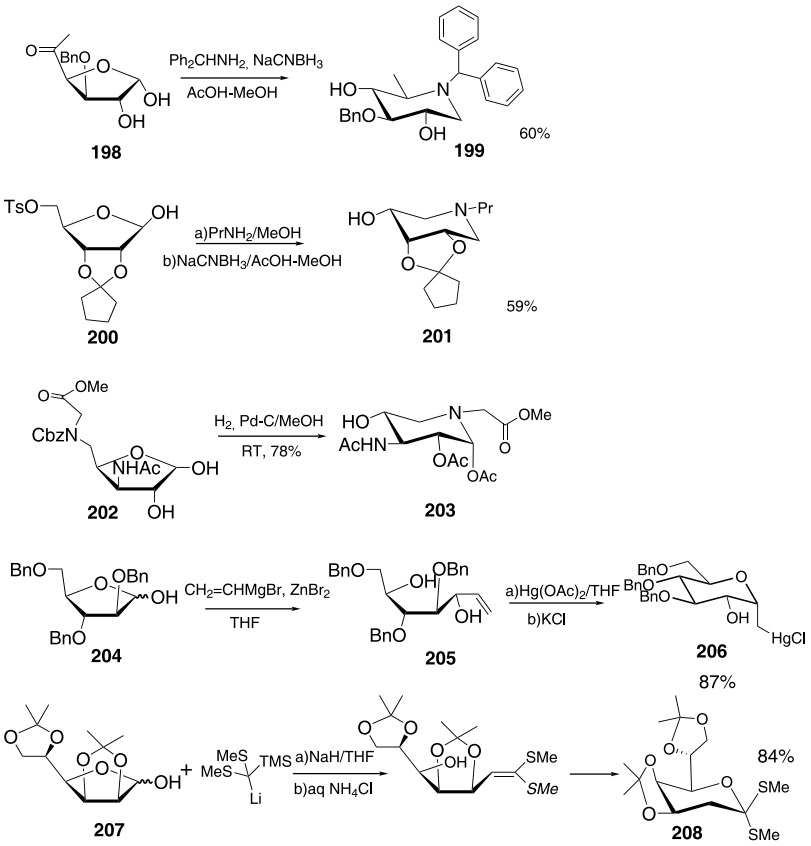
On deprotection, followed by neutralization, the acetal **190** rearranges spontaneously to the azasugar **191** [153], an analog of the indolizine alkaloids for which synthetic approaches starting from carbohydrates [154] have recently been described [155] employing the olefin



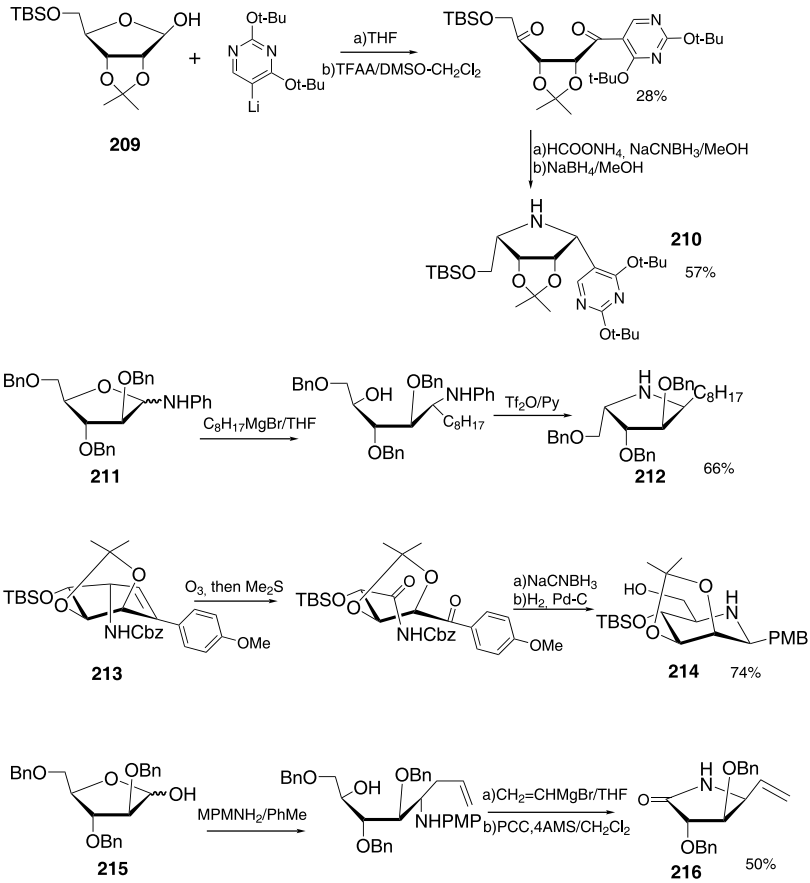
Scheme 44

metathesis protocol [156,157,158,159]. Under basic conditions, silyl-group shifts occur in a 6-deoxy-6,6,6-trifluorosugar **192** and form a pyranoside derivative [160]. Acetolysis of the methyl glycoside **194** mainly affords the piperidine **195** [161]. Azasugar ethyl thioglycoside **197**, a new type of azasugar derivative, can be stereoselectively prepared from suitable glycosylenamine **196**, through anhydroazasugar derivatives. The thioethoxy group is introduced through a highly stereoselective substitution. The attack of EtSH was 100% stereoselective [162,163] (*Scheme 44*).

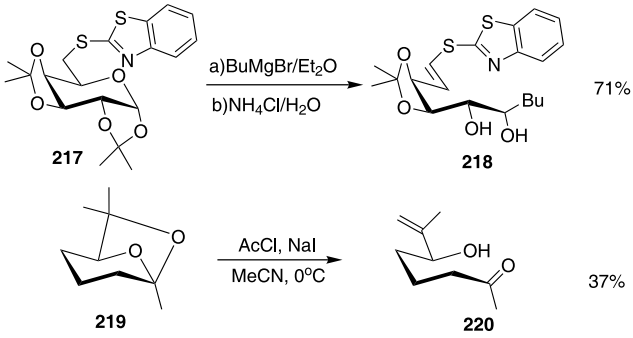
The protected 5-ulose derivative **198** can be converted into the piperidine **199** by reductive amination [164]. The 5-*O*-sulfonyllactol **200** is reductively transformed into the azasugar **201** by way of oxime formation [165]. Reductive deprotection of the aminodeoxylactol derivative **202** affords the *N*-substituted piperidine **203** [166]. The unsaturated alcohol, readily obtained from the lactol **204** and Grignard reagent, cyclizes into the *C*-glycoside **206** [167]. The lactol **207** is converted into the unsaturated dithioacetal, which cyclizes slowly to give **208** on storage [168] (*Scheme 45*).



Scheme 45



Scheme 46



Scheme 47

5.3 Ring Transformation

Additive ring-opening of **209**, followed by Swern oxidation and aminocyclization, affords the aza-*C*-nucleoside **210** [169], belonging to an attractive class of *C*-glycosides [170,171,172,173,174,175]. The glycosylamine **211** is converted to the azasugar **212** via an alkylative ring-opening reaction [176,177]. The aminoaldehyde derivative generated from the unsaturated aminocyclitol **213** cyclizes to give **214** [178] (● *Scheme 46*). Descending oxidative aminocyclization of **215** affords the lactam **216** [179].

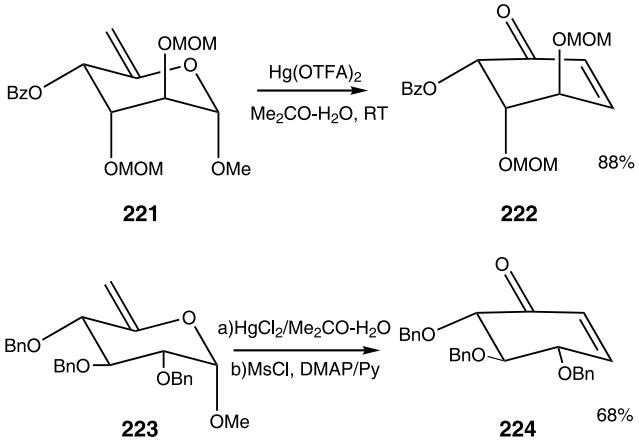
5.4 Ring-Opening Rearrangements

A fragmentation ring-opening rearrangement reaction of **217** using a Grignard reagent has been reported [180]. The combined reagent acetyl chloride/sodium iodide induces a ring-opening rearrangement of the bicyclic ketal **219** [181]. The iodide ion serves to promote the reaction [182] (● *Scheme 47*).

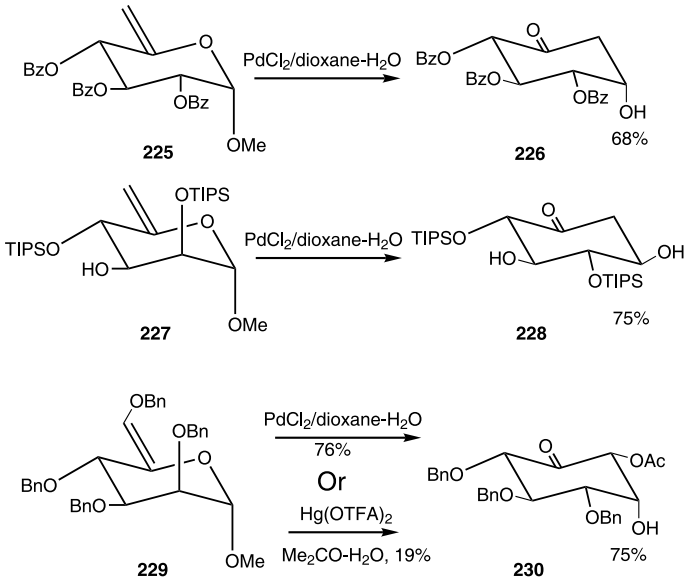
6 Miscellaneous Reactions

6.1 Ferrier Carbocyclization and Related Reactions

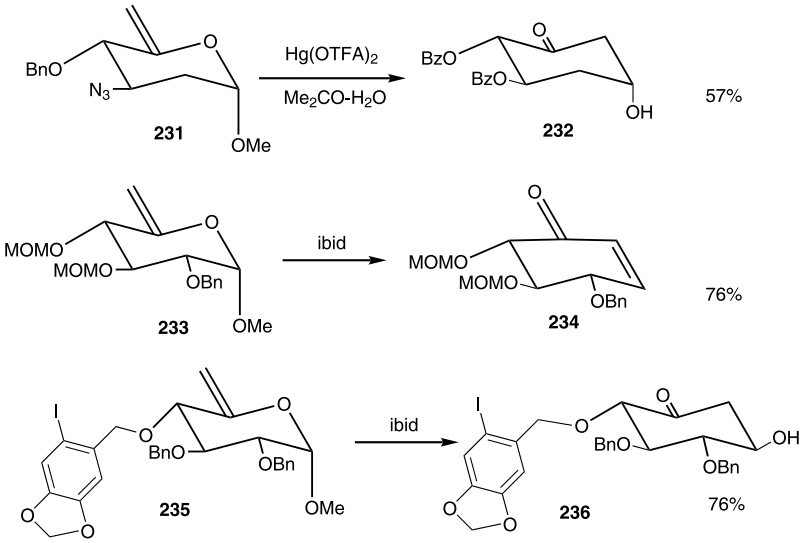
The Ferrier II reaction, a carbocyclization reaction, is of widespread use as a tool for the conversion of glycosides into cyclitols [183,184,185]. Newer examples for the utilization of the reaction conducted under catalytic conditions [186,187] have appeared in the recent literature. Compound **221** is converted into cyclohexanone **222** on the way to (−)-mesembranol [188] (● *Scheme 48*). Compound **223** is transformed to the enone **224**, the precursor of several new cyclitol derivatives [189,190,191].



Scheme 48



Scheme 49

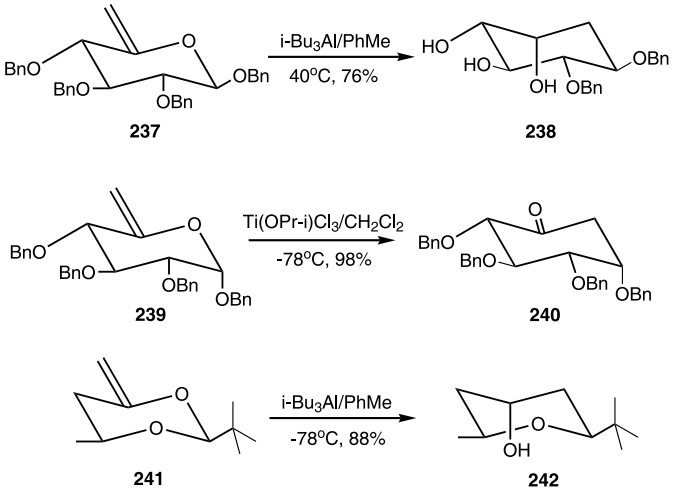


Scheme 50

The Pd-catalyzed carbocyclization affects good control on the orientation of the newly formed OH group [192] (*Scheme 49*). Thus, **225** and **227** afford the corresponding cyclitols with almost complete selectivity. In the rearrangement of **227**, the stereoselectivity is controlled by the bulky silyl ether-protecting group, which effects the conformational change.

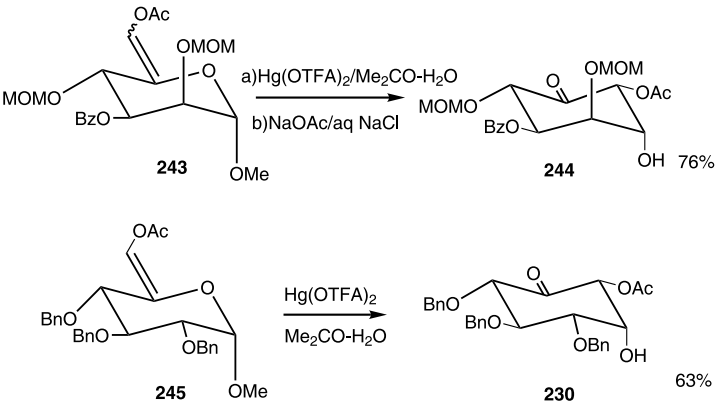
This protocol can be applied to the 6-*O*-acetyl-5-enopyranoside **229** with good efficiency and the utility is well demonstrated by the synthesis of the D-myo-inositol phosphate, IP3 [193,194].

Compound **231** is converted to the Ferrier product **232**, the precursor of novel aminoglucosides [183] (*Scheme 50*). Carbocyclization of the glycoside **233** gives the cyclohexane **234**, from which tetrazoline analogs can be synthesized [195]. The Ferrier cyclization found new utility in the synthetic chemistry of *Amaryllidaceae* alkaloids [196]. Thus, the glycoside **235** is transformed to the Ferrier-II product **236**, the logical intermediate to 7-deoxypancratistatin. A novel reductive carbocyclization of hex-5-enopyranosides retains the substituent at the anomeric center and the ring oxygen remains as the new hydroxy group [197]. The stereo-



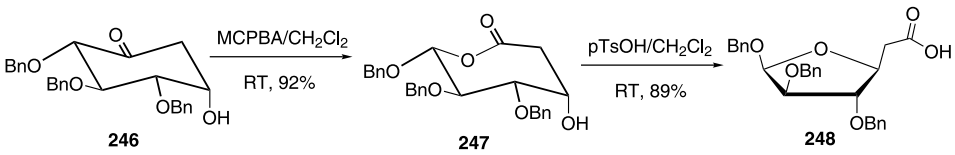
237 **238** **239** **240** **241** **242**

Scheme 51

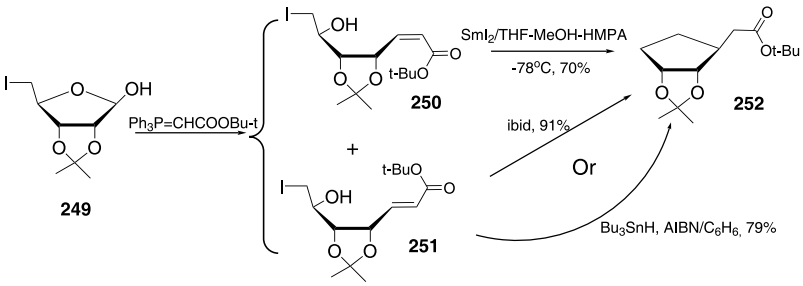


243 **244** **245** **230**

Scheme 52



Scheme 53



Scheme 54

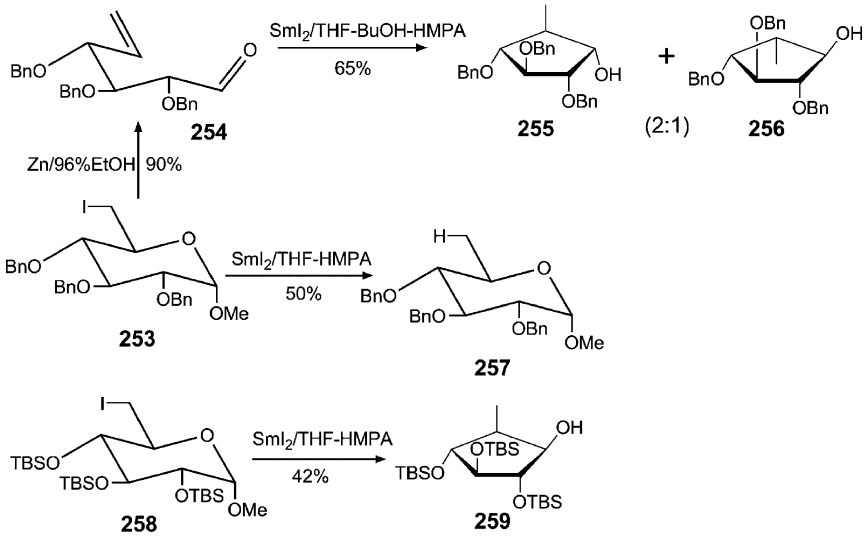
chemistry at the anomeric centers is retained as exemplified by the conversion of **237** to **238** (*Scheme 51*). A more efficient cyclization also retains the aglycone; the glycoside **239** affords the cyclohexanone **240**. The cyclic acetal **241** is converted to the pyran **242** reductively [198,199,200].

The enol acetate **243** affords the Ferrier product **244**, a key compound to L-chiro-inositol polyphosphates [201]. The Ferrier cyclization of **245** is useful for the preparation of the key intermediate to glycosylphophetidylinositols [202] (*Scheme 52*).

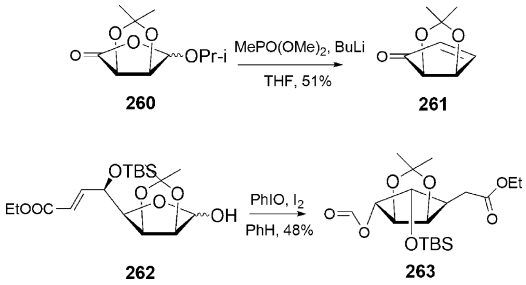
Combination of the Ferrier-II and the Baeyer–Villiger reactions leads to the stereoselective synthesis of rare 5-deoxyfuranosiduronic acids [203]. As exemplified, the oxidation of the Ferrier II product **246**, followed by hydrolysis, gives the acid **248** [204] (*Scheme 53*).

The evolution of SmI_2 as a reagent in synthesis has been one of the exciting recent developments in organic chemistry. The construction of highly functionalized carbocycles from carbohydrates promoted by SmI_2 is currently receiving significant interest and a series of carbocyclization strategies have been described in the literature. Treatment of the lactol **249** with the Wittig reagent readily gives the olefins, which undergo radical-induced cyclization [205]. Cyclization of the (Z)-isomer **250** under the action of SmI_2 is more stereoselective than that of the (E)-isomer **251** [206]. In the case of **252**, the diastereomeric excess of the products significantly depends on the choice of the reducing agents (*Scheme 54*).

Stepwise conversion of the iodoglycoside **253** via Grob–Vasella fragmentation and cyclorearrangement induced by SmI_2 furnishes the carbocycles **255** and **256** with a trans-junction [207,208]. This reaction can be carried out in a one-pot manner whereby SmI_2 induces the fragmentation of the iodoglycoside [209]. While the iodoglycoside **258** mainly affords the carbocycle **259** with a *cis*-junction, the reaction of **253** only gives the quinovoside **257** (*Scheme 55*).



Scheme 55

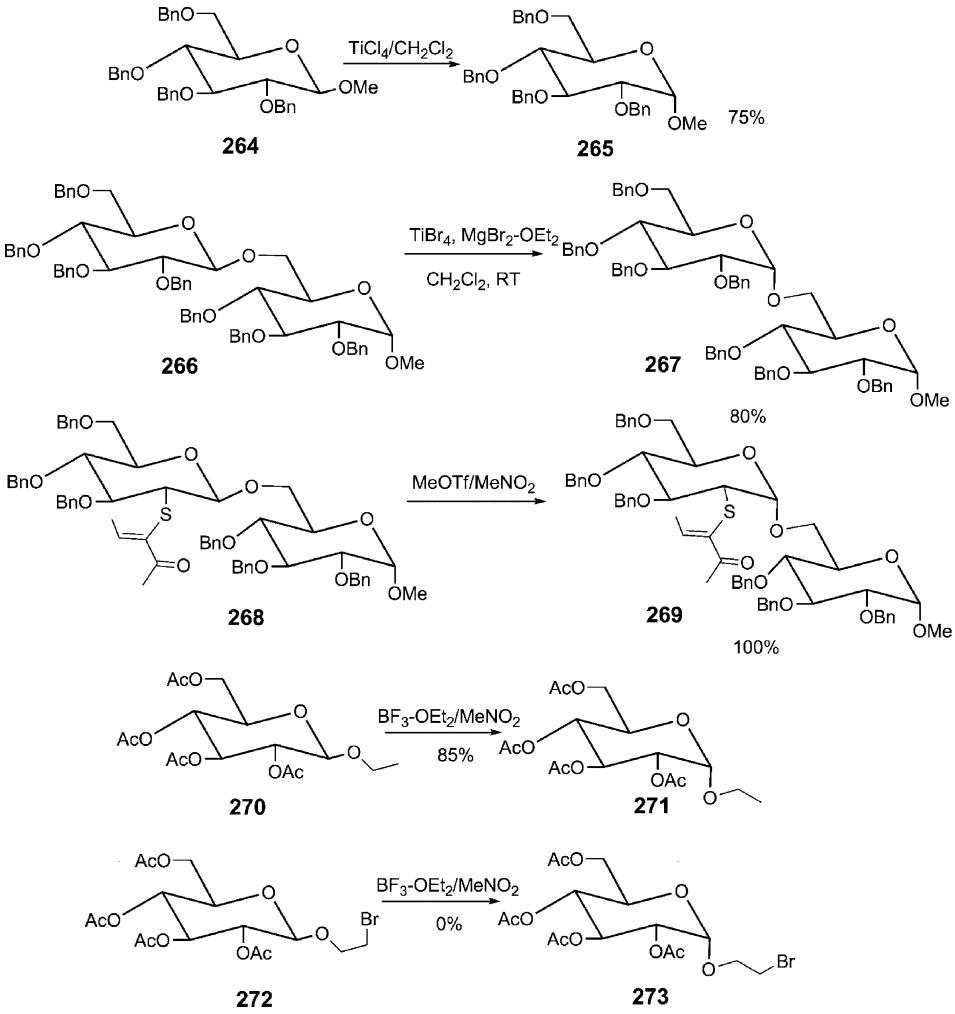


Scheme 56

The glycoside **260** is converted into the cyclopentenone **261** on reaction with dimethyl methanephosphonate and base [210] (*Scheme 56*). The tandem β-fragmentation-cycloisomerization of the unsaturated lactol **262** gives the carbocycle **263** [211].

6.2 Anomerization and Related Rearrangements

Anomerization is a characteristic reaction of sugar [212,213]. The well-known reagent, Pascu's $TiCl_4$ for the anomerization of acetylated glycoside, rapidly anomerizes the benzyl-protected glucoside **264** [214,215] (*Scheme 57*). The results from inhibition experiments indicate that $TiCl_4$ might coordinate with O5 and O6 to form a ring-opened intermediate. The use of catalytic amounts of $TiBr_4$ combined with $MgBr_2 \cdot OEt_2$ allows us to carry out longer reactions: the disaccharide **266** anomerizes to **267** completely. It has been reported that β-glycosides such as **268** anomerizes quantitatively [216]. Although the acetylated glycoside **270** is anomer-

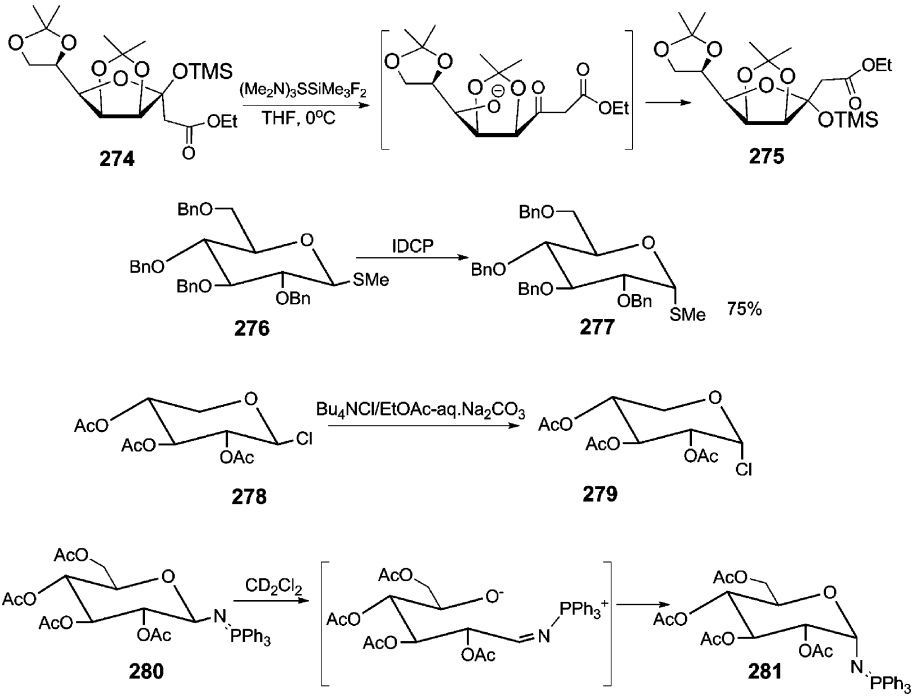


Scheme 57

ized in polar nitromethane containing $BF_3{\cdot}OEt_2$, the bromide **272** is practically inert [217] (*Scheme 57*).

The silylated ketose **274** slowly anomerizes in the presence of TASF by way of the keto-form [218]. The thioglycoside **276** anomerizes in the presence of a catalytic amount of IDCP [219]. Under PTC conditions, the β-chloride **278** also anomerizes [220]. Anomerization of **280** to the α-form **281** via an open-chain zwitterionic intermediate has been suggested [221] (*Scheme 58*).

An investigation of the time course of the anomerization of β-iodide **282** has been carried out using ^{1}H-NMR spectroscopy [222]. The NMR titration method to measure the shift of



Scheme 58

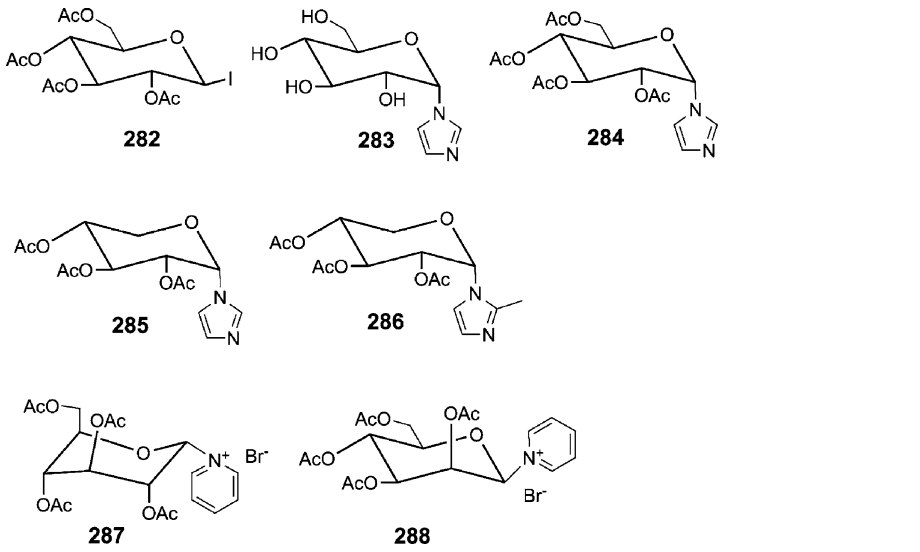
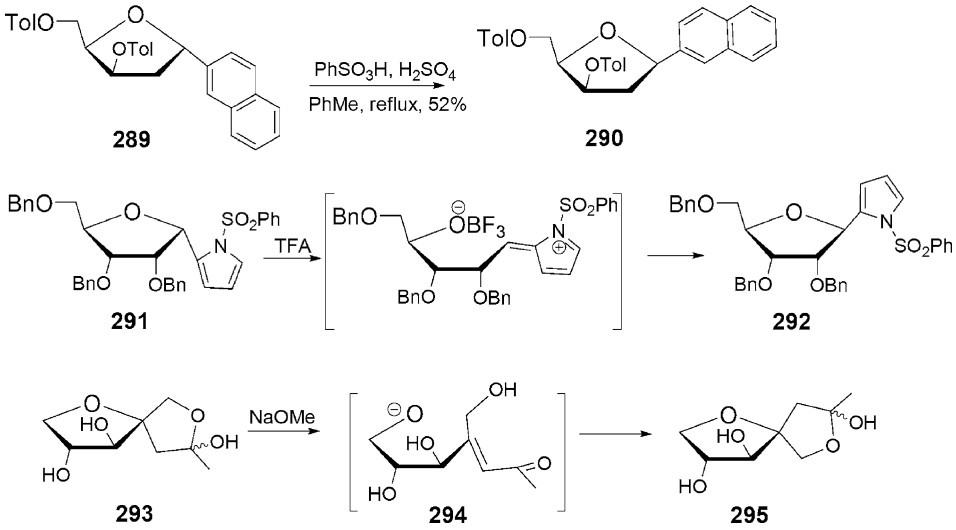


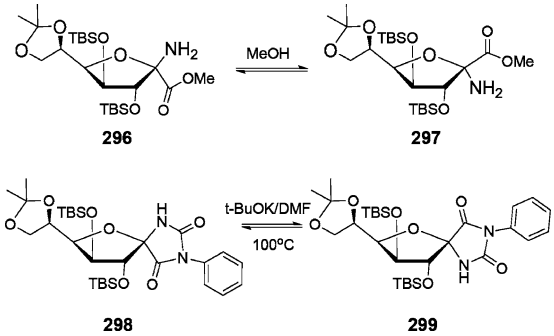
Figure 8
Anomeric equilibrium of glycosides and reverse anomeric effect

the anomeric equilibrium on protonation of **283** and **284** reveals that the protonated imidazoyl group has a small but distinct preference for the axial disposition than does the unprotonated group; which is the opposite of what the reverse anomeric effect predicts [223,224,225]. Compounds **285** and **286** increase the proportion of their $^1C_4$ conformers on *N*-protonation but not when the polarity of the solvent is increased as predicted by the reverse anomeric effect [226]. In solution, the α-glycosylpyridinium salt **287** adopts the $^1C_4$ conformation and the β-mannosyl compound **288** has the $^4C_1$ form; both of them have positively charged groups in an equatorial position at the anomeric center, indicating the manifestation of the reverse anomeric effect [227] (● *Fig. 8*).

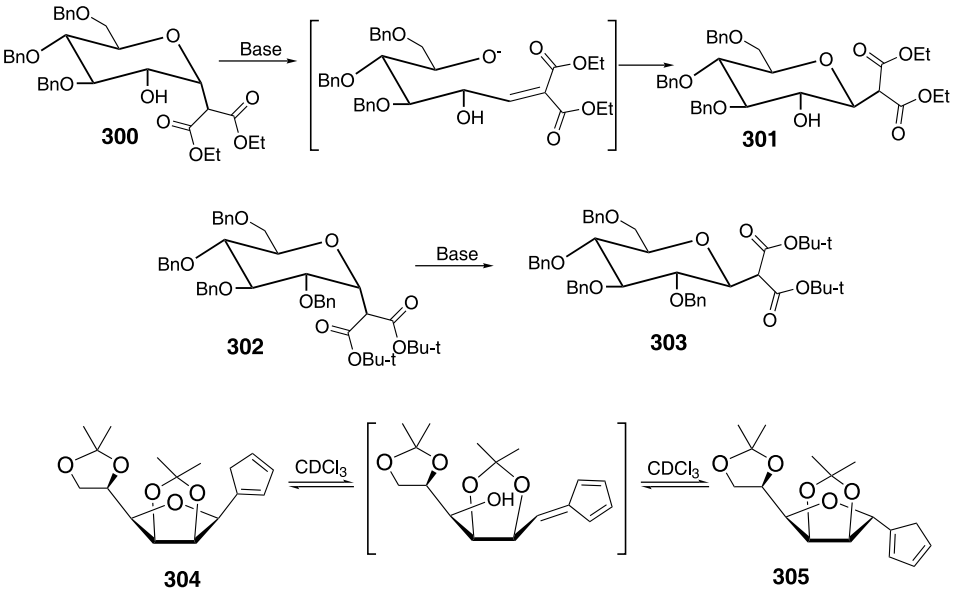
Aryl *C*-glycosides such as **289** and **291** undergo α- to β-anomerization in the presence of an acid by way of open-chain intermediates [228,229]. Under basic conditions, **293** isomerizes into **295** through **294** [230] (● *Scheme 59*).



Scheme 59



Scheme 60

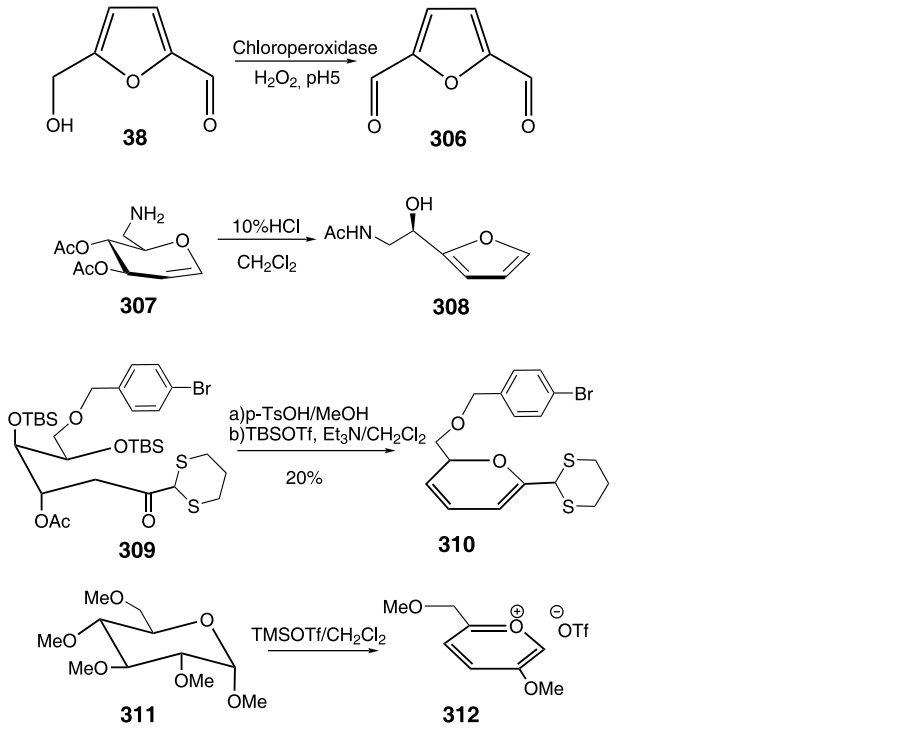


Scheme 61

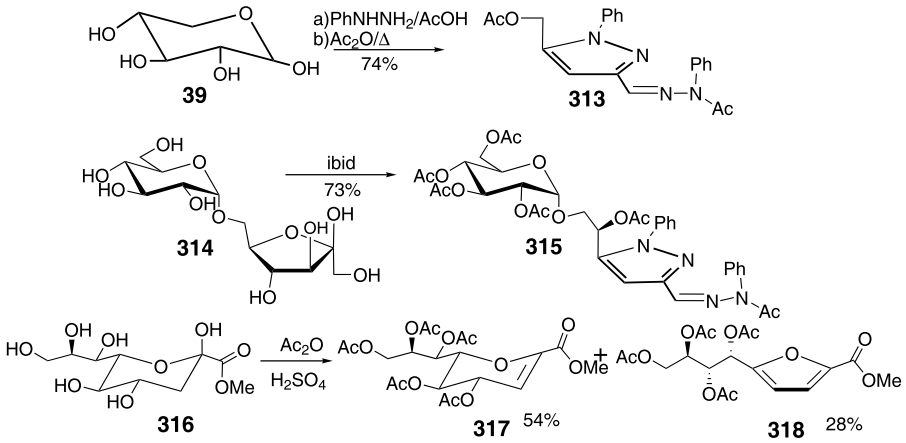
The glycosylamine **296** anomerizes in methanolic solution. The spirohydantoins **298** and **299** form an equilibrium mixture under basic conditions [231] (*Scheme 60*.
A possibility for the α- to β-anomerization of *C*-glycosides **300** and **302** by way of open-chain intermediates generated under basic conditions has been discussed [232,233]. Isomerization of pure **303** to **304** and vice versa probably occurs by way of open-chain isomerization through the linear intermediate [234] (*Scheme 61*).

6.3 Aromatization of Sugars

Sequential elimination reactions, most of them being dehydration, involving the reaction at the anomeric center often produce various aromatic compounds [235] especially furans which have diverse use [236,237]. Explorations have been continued to open a new route to aromatics based on renewable biomass in place of fossilized material.
Oxidation of 5-hydroxymethylfuraldehyde **38** with hydrogen peroxide catalyzed by chloroperoxidase, a hemeperoxidase from *Caldariomyces fumago*, proceeds with good selectivity to furnish **306** [238] (*Scheme 62*). The 6-aminodeoxyglycal derivative **307** is similarly converted into the furan **308** [239]. The dithiane **309** gives the oxacyclohexadiene **310**, on acid treatment [240]. Treatment of **311** with TMSOTf produces the pyrylium salt **312** [241].
The phenylosazone from D-xylose **39** can be converted into the pyrazoles **313** [242] . Isomaltulose **314** affords the glycosylated aromatic compound **315** [243]. On acidic acetylation the ulosonic acid ester **316** forms concomitantly the glycal derivative **317** and the furanoic acid derivative **318** [244] (*Scheme 63*).

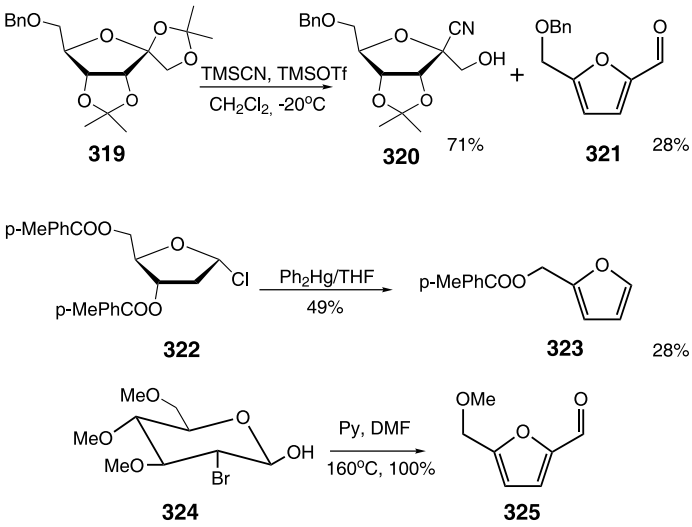


Scheme 62



Scheme 63

Even on mild *C*-glycosylation using the TMSCN/TMSOTf system, a notable amount of the D-psicofuranose derivative **319** degrades to the furan **321** [245]. On reaction with Ph_2Hg, the chloride **322** gives the furan **323** exclusively [246]. The bromohydrin **324** degrades into furan **325** on heating with a base [247] (*Scheme 64*).



Scheme 64

7 The Maillard Reaction

The Maillard reaction is a complex group of degradation/rearrangement reactions initiated by reactions of free sugars and amines [248,249,250,251]. The reaction is of major interest for food processing [252,253,254,255,256,257] and life sciences [258,259,260,261,262,263,264, 265]. Degeneration of amine drugs in the presence of reducing sugars as excipients and deterioration of sugar artifacts are also related to the reaction [266,267].

The nonenzymatic reaction between reducing sugars and long-lived proteins *in vivo* results in the formation of glycation and advanced glycation end products, which alter the properties of proteins including charge, helicity, and their tendency to aggregate. Such protein modifications are linked with various pathologies associated with the general aging process such as Alzheimer disease and the long-term complications of diabetes. Although it has been suggested that glycation and advanced glycation end products altered protein structure and conformation, little structural data and information currently exist on whether or not glycation does indeed influence or change local protein secondary structure [268]. For example, in the blood, D-glucose can react with an NH_2 group of hemoglobin to form an imine that subsequently undergoes an irreversible rearrangement to a more stable a-aminoketone known as hemoglobin-AIc \index{hemoglobin-AIc}% [269].

Diabetes results when the body does not produce sufficient insulin or when the insulin it produces does not properly stimulate its target cells. Because insulin is the hormone that maintains the proper level of glucose in the blood, diabetics have increased blood glucose levels. The amount of hemoglobin-AIc formed is proportional to the concentration of glucose in the blood, so diabetics have a higher concentration of hemoglobin-AIc than nondiabetics. Thus, measuring the hemoglobin-AIc level is a way to determine whether the blood glucose level of a diabetic is being controlled [270,271].

Cataracts, a common complication in diabetics, are caused by the reaction of glucose with the group of proteins in the lens of the eye. It is thought that the arterial rigidity common in old age may be attributable to a similar reaction of glucose with the NH_2 group of proteins [250,272].

7.1 Mechanism of the Maillard Reaction

As illustrated in *Scheme 65*, **33** reacts with an amine to give an imine **326** that isomerizes into an aminoketose **327** (Amadori product), existing as an equilibrium mixture of cyclic hemiacetals, whereas **40** affords, by way of **328**, the hexosamine derivatives **329** and **330** (Heyns products), also in cyclic form. The Amadori–Heyns compounds \index{Amadori–Heyns compounds}% are at the head of the complex sequences of the Maillard reaction. The crystal structure of the Amadori product **331** between **33** and glycine has been determined more than three decades after the first proposal of its structure. Alternative preparations and X-ray analyses of Heyns products **332** and **333** have been reported [273,274].

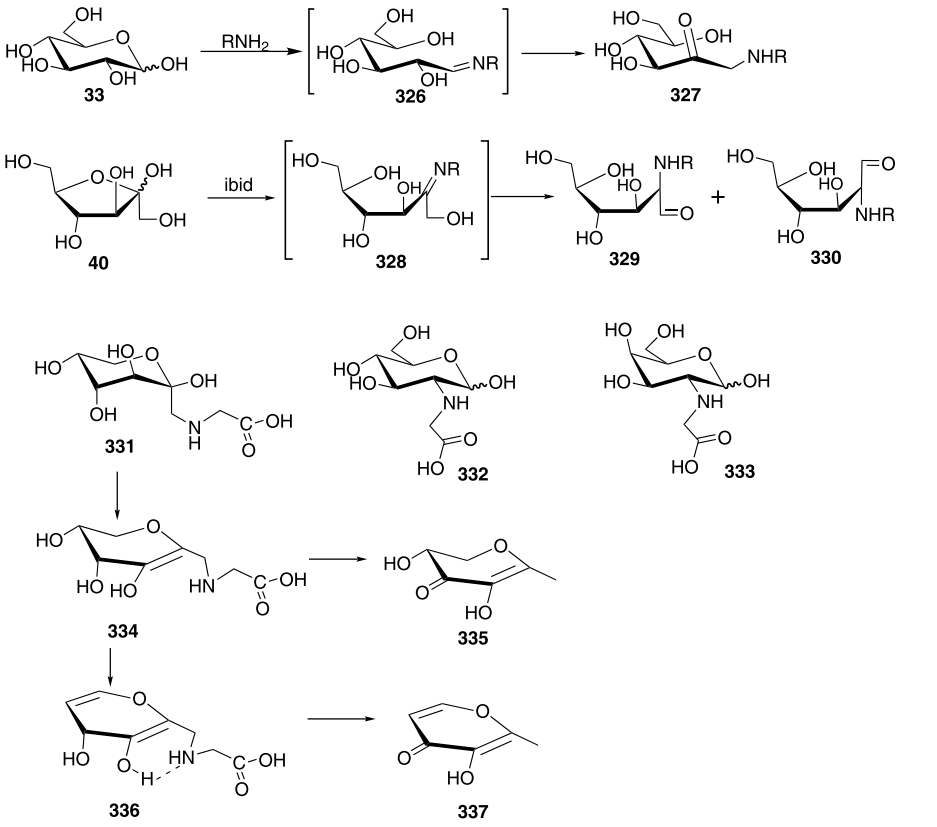
The Amadori product from D-glucose **33** and L-proline decomposes at 130 °C in DMF to afford **33** and D-Mannose **54**, indicating the reversibility of the Amadori reaction. A kinetic study using **33** and phenylalanine indicates that the Schiff's base formation is the rate-determining step of the Maillard reaction [275,276].

Maltol **337** is one of the degradation products in monosaccharide solutions with amino acids forming Amadori compounds but not in the solution of monosaccharides alone. Heated solutions of monosaccharides yield **335**, the logical precursor of **337**, but not **337** itself. On the basis of the molecular mechanics calculation indicating that **335** adopts the conformation unfavorable for dehydration into **337**, a possible route via the dehydrated product **336**, an *ortho*-elimination product, has been postulated as a more favorable alternate reaction pathway [277].

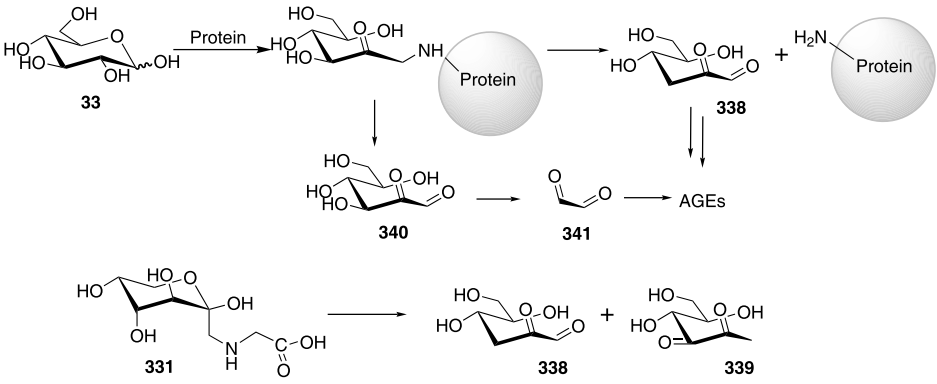
7.2 Chemistry of Biologically Significant Maillard Products

In biological systems, Amadori products formed from aldoses and the amino group in peptides, decompose to release reactive sugar derivatives that are irreversibly consumed in the production of the advanced glycation end products (AGEs). In this sense, **338** is one of the key substances in the Maillard reaction [278]. A new specific assay of **338** has been developed using diaminonaphelene [279]. The dicarbonyl compounds **338** and **339**, the suggested intermediates in the degradation of the Amadori compound **331**, had been trapped with aminoguanidine [280,281]. The role of **338** generated in the Maillard cascade as a cross-linker of proteins has been emphasized [282]. Oxygen and metal cations accelerate the degradation of Amadori products to D-glucosone (**340**), a precursor of glyoxal **341** [283,284] (*Scheme 66*).

It is known that some Maillard products have strand-breaking activities to DNA. Many compounds found in foodstuffs are α,β-unsaturated ketones [285,286]. Compounds **53**, **342**, **343**, and **335** (*Fig. 9*) cleave DNA single strands by generating hydroxyl radicals and other active oxygen radicals in the presence of Fe^{3+} and oxygen [287,288,289,290,291,292]. For example, the key hydrolyzate **344** generates hydroxyl radicals and the oxidation products **345** and **346**. An organic hydroperoxide **347**, presumably formed via direct oxidation of **339** or stepwise from **348**, the precursor of **335**, has been isolated [293,294,295,296,297] (*Scheme 67*).



Scheme 65



Scheme 66

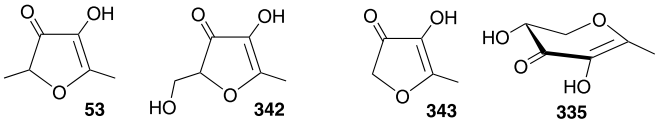
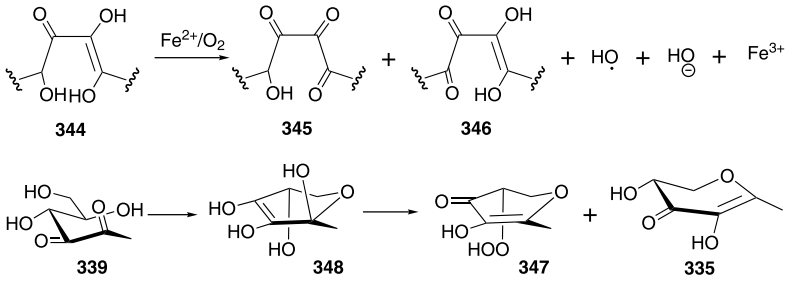


Figure 9
Maillard products having strand-breaking activities to DNA

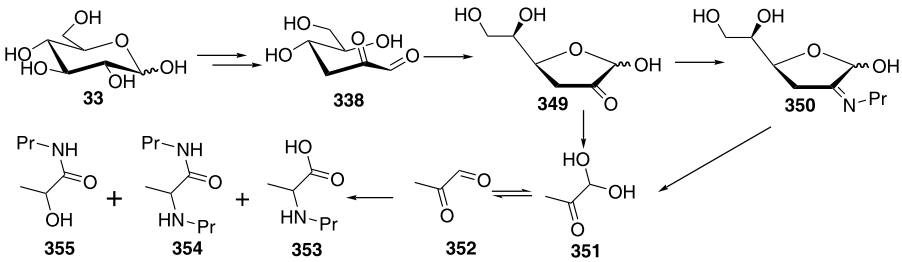


Scheme 67

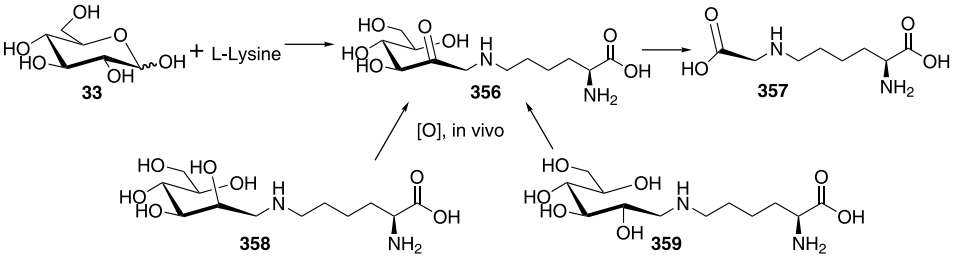
The major intermediate of the Maillard reaction **38**, having an allylic system, seems to furnish a cytotoxic ester on metabolic sulfonation [298]. In contrast to the above findings, some Amadori products, such as pyrazines have antimutagenicity [299,300,301]. Enkastines, the Amadori products of **33** and dipeptides, beneficially prolong the action of enkephaline by inhibiting enkephalinase [302].

Reactive small sugars and related acids appear to play a role in forming AGEs including cross-linked proteins in the aged body as well as inactivation of human Cu,Zn-superoxide dismutase [282,303,304]. Reaction of **33** with *n*-propylamine in phosphate-buffered, neutral solution generates several derivatives of small sugars [305,306], namely, C_2 and C_3 sugar derivatives. The 3-deoxyulose **349**, a hemiacetal form of **338**, yields **351**, the hydrate of methylglyoxal (**352**), as well as the Schiff's base **350** which is thought to be the precursor of the C_3-products [307,308,309] (*Scheme 68*).

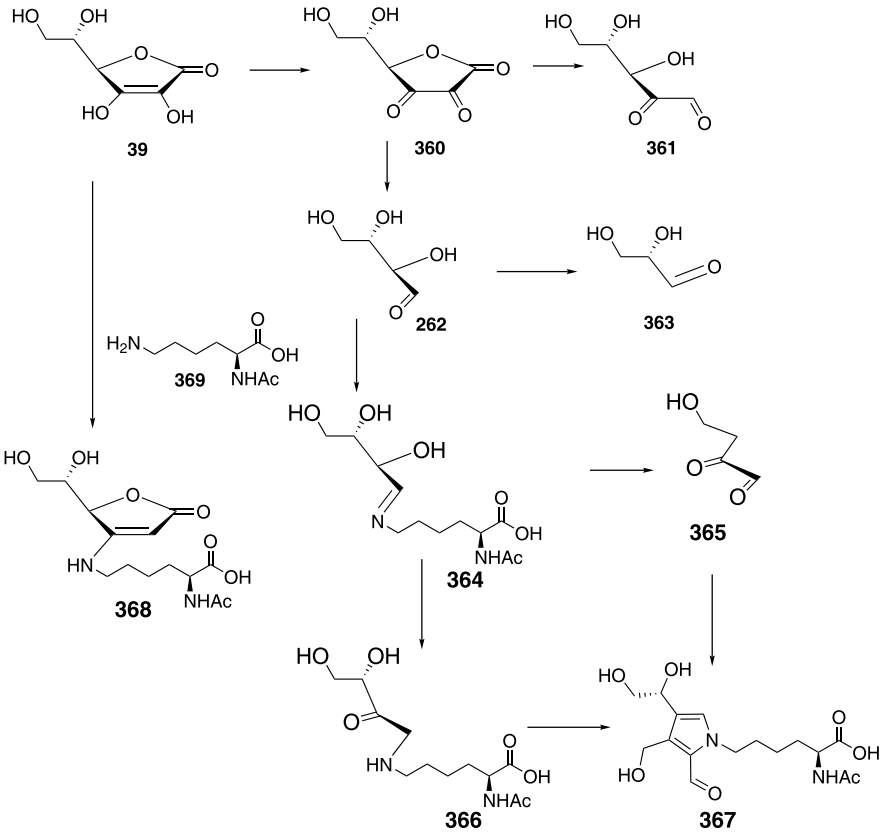
N^{ε}-(carboxymethyl)lysine (**357**), is a main AGE product found in vivo [310,311]. About 50% of **357** seems to be formed via oxidative degradation of the Amadori product **356**. The reduced



Scheme 68



Scheme 69



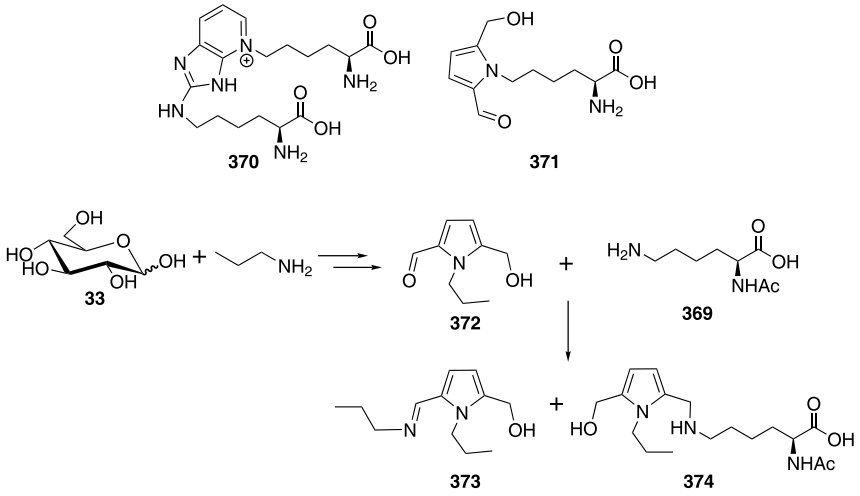
Scheme 70

compounds **358** and **359** also form **356** under aerobic physiological conditions [282]. Reactive **352** combines reversibly with lysine and cysteine residues and irreversibly with arginine residue [312] (*Scheme 69*).

L-threose (**362**), the degradation product of **360** degrades in the presence of N^{α}-acetyl-L-lysine (**369**) at pH 7 into 3-deoxy-tetros-2-ulose (**365**) [313,314]. Only at pH 7 does retro-

aldolization of **362** occur to give glyceraldehyde (**363**). Under physiological conditions, the AGE product **364** is formed from **362** and **369**, apparently via condensation of the Amadori compounds **365** and **366** [315,316,317]. On heating at 100 °C, a hood-processing temperature, in methanol in a sealed vessel, **39** and **369** form an amine **368** [318] (*Scheme 70*).

Some of the heterocyclic compounds among the Maillard products, for example, pentosidine **370** and pyrraline **371**, are AGEs in the skin of diabetic patients as well as in the brain of Alzheimer patients [319,320,321,322,323,324,325,326,327,328,329]. The observation that the aldehyde **372**, the Maillard product of **33** and *n*-propylamine, reacts with the amine and **369** to give **373** and **374**, respectively, led to the assumption that pyrrole aldehydes might also be precursors of the lysine side chain of proteins [320,330,331,332] (*Scheme 71*).



Scheme 71

References

1. Solomons TWG, Fryhle CB (2002) Organic chemistry. Wiley, New York
2. Bruice PY (2004) Organic chemistry. Pearson/Prentice Hall, Upper Saddle River, NJ
3. Stütz AE (2001) Glycoscience: epimerisation, isomerisation and rearrangement reactions of carbohydrates. Springer, Berlin, Heidelberg, New York
4. Lindhorst TK (2003) Essentials of carbohydrate chemistry and biochemistry. Wiley-VCH, Weinheim
5. Levy DE, Fügedi P (2006) The organic chemistry of sugars. Taylor & Francis, Boca Raton, FL
6. Sun Y, Cheng J (2002) Bioresource Technology 83:1
7. Wolfenden R, Lu X, Young G (1998) J Am Chem Soc 120:6814
8. Chenault HK, Chafin LF (1998) J Org Chem 63:833
9. Kirby AJ, Stromberg R (1994) J Chem Soc-Chem Comm: 709
10. Camilleri P, Jones RFD, Kirby AJ, Stromberg R (1994) J Chem Soc-Perkin Trans 2:2085
11. Sidelmann UG, Hansen SH, Gavaghan C, Carless HAJ, Lindon J, Farrant R, Wilson ID, Nicholson JK (1996) Anal Chem 68:2564

12. ManleyHarris M, Richards GN (1996) Carbohydr Res 287:183
13. Sinnott ML (1990) Chem Rev 90:1171
14. Kaczmarek J, Kaczynski Z, Trumpakaj Z, Szafranek J, Bogalecka M, Lonnberg H (2000) Carbohydr Res 325:16
15. Horenstein BA, Bruner M (1996) J Am Chem Soc 118:10371
16. Chong AK, Pegg MS, Taylor NR, von Itzstein M (1992) Eur J Biochem 207:335
17. Davies GJ, Mackenzie L, Varrot A, Dauter M, Brzozowski AM, Schulein M, Withers SG (1998) Biochemistry 37:11707
18. Nagano N, Noguchi T, Akiyama Y (2007) Proteins-Structure Function and Bioinformatics 66:147
19. Crennell SJ, Garman EF, Laver WG, Vimr ER, Taylor GL (1993) Proc Natl Acad Sci USA 90:9852
20. Guo X, Sinnott ML (1993) Biochem J 296 (Pt 2):291
21. Ashwell M, Sinnott ML, Zhang Y (1994) J Org Chem 59:7539
22. Ashwell M, Guo X, Sinnott ML (1992) J Am Chem Soc 114:10158
23. Pitson SM, Mutter M, van den Broek LA, Voragen AG, Beldman G (1998) Biochem Biophys Res Commun 242:552
24. Biely P, Benen J, Heinrichova K, Kester HC, Visser J (1996) FEBS Lett 382:249
25. Schultz PG, Lerner RA (1995) Science 269:1835
26. Yu J, Choi SY, Moon KD, Chung HH, Youn HJ, Jeong S, Park H, Schultz PG (1998) Proc Natl Acad Sci USA 95:2880
27. Yu S, Ahmad T, Kenne L, Pedersen M (1995) Biochim Biophys Acta 1244:1
28. Richard GT, Yu S, Monsan P, Remaud-Simeon M, Morel S (2005) Carbohydr Res 340:395
29. Yu S, Kenne L, Pedersen M (1993) Biochim Biophys Acta 1156:313
30. Saddler JN, Penner MH (1995) Am Chem Soc Meeting. Enzymatic degradation of insoluble carbohydrates. American Chemical Society, Washington, DC
31. Horn SJ, Sikorski P, Cederkvist JB, Vaaje-Kolstad G, Sorlie M, Synstad B, Vriend G, Varum KM, Eijsink VG (2006) Proc Natl Acad Sci USA 103:18089
32. National Research Council (US) Committee on Bioprocess Engineering (1992) Putting biotechnology to work: bioprocess engineering. National Academy of Sciences, Washington, DC
33. Mulimani VH, Ramalingam R (1995) Biochem Mol Biol Int 36:897
34. Kulkarni DS, Kapanoor SS, Girigouda K, Kote NV, Mulimani VH (2006) Biotechnol Appl Biochem 45:51
35. Hollingsworth RI, Wang G (2000) Chem Rev 100:4267
36. Nolasco J, De Massaguer PR (2006) J Food Proc Eng 29:462
37. Hollnagel A, Kroh LW (2000) J Agricult Food Chem 48:6219
38. Miyazawa T, Ohtsu S, Nakagawa Y, Funazukuri T (2006) J Mat Sci 41:1489
39. Srokol Z, Bouche AG, van Estrik A, Strik RCJ, Maschmeyer T, Peters JA (2004) Carbohydr Res 339:1717
40. Griebl A, Lange T, Weber H, Milacher W, Sixta H (2006) Macromolecular Symposia 232:107
41. Kruse A, Gawlik A (2003) Ind Eng Chem Res 42:267
42. Hashaikeh R, Fang Z, Butler IS, Kozinski JA (2005) Proc Combustion Institute 30:2231
43. Chheda JN, Roman-Leshkov Y, Dumesic JA (2007) Green Chem 9:342
44. Sanderson K (2006) Nature 444:673
45. Hayes MH (2006) Nature 443:144
46. Kruse A, Maniam P, Spieler F (2007) Ind Eng Chem Res 46:87
47. Mark J, Pollien P, Lindinger C, Blank I, Mark T (2006) J Agricult Food Chem 54:2786
48. Fan XT (2005) J Agricult Food Chem 53:7826
49. Ginz M, Balzer HH, Bradbury AGW, Maier HG (2000) Eur Food Res Technol 211:404
50. Arya M, Rao LJM (2007) Crit Rev Food Sci Nutrit 47:51
51. Burde RDL, Crayton F, Bavley A (1962) Nature 196:166
52. Qian X, Nimlos MR, Johnson DK, Himmel ME (2005) Appl Biochem Biotechnol 121–124:989
53. Buskas T, Konradsson P (2000) J Carbohydr Chem 19:25
54. Antal MJ, Jr., Mok WS, Richards GN (1990) Carbohydr Res 199:91
55. Halliday GA, Young RJ, Jr., Grushin VV (2003) Org Lett 5:2003
56. Hauck T, Landmann C, Bruhlmann F, Schwab W (2003) J Agric Food Chem 51:1410
57. deGoede ATJW, vanDeurzen MPJ, vanderLeij IG, vanderHeijden AM, Baas JMA, vanRantwijk F, vanBekkum H (1996) J Carbohydr Chem 15:331
58. Cancilla MT, Penn SG, Lebrilla CB (1998) Anal Chem 70:663

59. Yang BY, Montgomery R (2007) Bioresour Technol 98:3084
60. T'i VTU, Pisarnitskii AF (1998) Appl Biochem Microbiol 34:106
61. Manini P, La Pietra P, Panzella L, Napolitano A, d'Ischia M (2006) Carbohydr Res 341:1828
62. Hourdin GL, Germain A, Moreau C, Fajula F (2002) J Catalysis 209:217
63. van den Berg R, Peters JA, van Bekkum H (1995) Carbohydr Res 267:65
64. Rangappa KS, Raghavendra MP, Mahadevappa DS, Gowda DC (1998) Carbohydr Res 306:57
65. Hollingsworth RI (1996) Biotechnol Annu Rev 2:281
66. Hollingsworth RI (1999) J Org Chem 64:7633
67. Stapley JA, BeMiller JN (2007) Carbohyd Res 342:407
68. Hourdin G, Germain A, Moreau C, Fajula F (2000) Catal Lett 69:241
69. Fischer K, Bipp HP (2005) Bioresour Technol 96:831
70. Niu W, Molefe MN, Frost JW (2003) J Am Chem Soc 125:12998
71. Service RF (2007) Science 315:1488
72. Champagne P (2007) Resour Conserv Recy 50:211
73. Kakinuma K, Li HY (1989) Tetrahedron Lett 30:4157
74. Taillefumier C, Chapleur Y (2004) Chem Rev 104:263
75. Sasaki M, Higashi M, Masu H, Yamaguchi K, Takeda K (2005) Org Lett 7:5913
76. Bertrand P, Gesson JP, Renoux B, Tranoy I (1995) Tetrahedron Lett 36:4073
77. Halcomb RL, Boyer SH, Wittman MD, Olson SH, Denhart DJ, Liu KKC, Danishefsky SJ (1995) J Am Chem Soc 117:5720
78. Modica E, Compostella F, Colombo D, Franchini L, Cavallari M, Mori L, De Libero G, Panza L, Ronchetti F (2006) Org Lett 8:3255
79. Jaunzeme I, Jirgensons A (2005) Synlett: 2984
80. Montero A, Mann E, Herradon B (2005) Tetrahedron Lett 46:401
81. Banaszek A, Pakulski Z, Zamojski A (1995) Carbohyd Res 279:173
82. Ichikawa Y, Kobayashi C, Isobe M (1996) J Chem Soc-Perkin Trans 1:377
83. Ichikawa Y, Kobayashi C, Isobe M (1994) Synlett: 919
84. Oishi T, Ando K, Inomiya K, Sato H, Iida M, Chida N (2002) Org Lett 4:151
85. Werschkun B, Thiem J (2001) Glycoscience: Epimerisation, Isomerisation and Rearrangement Reactions of Carbohydrates. Topics in Current Chemistry Vol 215, Springer-Verlag, Berlin, Heidelberg 215:293
86. Sudha AVRL, Nagarajan M (1998) Chem Comm 925
87. Balasubramanian KK, Ramesh NG, Pramanik A, Chandrasekhar J (1994) J Chem Soc-Perkin Trans 2:1399
88. Kuhn C, Legouadec G, Skaltsounis AL, Florent JC (1995) Tetrahedron Lett 36:3137
89. Overman LE, Shim J (1993) J Org Chem 58:4662
90. Deng W, Overman LE (1994) J Am Chem Soc 116:11241
91. Knight SD, Overman LE, Pairaudeau G (1995) J Am Chem Soc 117:5776
92. Agami C, Couty F, Lin J, Mikaeloff A (1993) Synlett: 349
93. Agami C, Couty F, Puchot-Kadouri C (1998) Synlett: 449
94. Oberdorfer F, Haeckel R, Lauer G (1998) Synthesis-Stuttgart: 201
95. Mundorff EC, Hanson MA, Varvak A, Ulrich H, Schultz PG, Stevens RC (2000) Biochemistry 39:627
96. Black KA, Leach AG, Kalani MY, Houk KN (2004) J Am Chem Soc 126:9695
97. Yokoyama H, Otaya K, Yamaguchi S, Hirai Y (1998) Tetrahedron Lett 39:5971
98. Zamojski A (2002) Polish J Chem 76:1053
99. Yokoyama H, Ejiri H, Miyazawa M, Yamaguchi S, Hirai Y (2007) Tetrahedron: Asymmetry 18:852
100. Muzart J (2005) Tetrahedron 61:4179
101. Mcdonald FE, Gleason MM (1995) Angew Chem Int Ed Engl 34:350
102. Castro S, Peczuh MW (2005) J Org Chem 70:3312
103. Redlich H (1994) Angew Chem 106:1407
104. Collins P, Ferrier RJ (1995) Monosaccharides: their chemistry and their roles in natural products. Wiley, New York
105. Tatsuta K, Hosokawa S (2006) Science and Technology of Advanced Materials 7:397
106. Ponten F, Magnusson G (1994) Acta Chemica Scandinavica 48:566
107. Ito H (2003) Yakugaku Zasshi-J Pharmaceut Soc Japan 123:933
108. Ito H, Motoki Y, Taguchi T, Hanzawa Y (1993) J Am Chem Soc 115:8835

109. Jenkins DJ, Riley AM, Potter BVL (1996) J Org Chem 61:7719
110. Bettelli E, D'Andrea P, Mascanzoni S, Passacantilli P, Piancatelli G (1998) Carbohydr Res 306:221
111. Estevez JC, Saunders J, Besra GS, Brennan PJ, Bash RJ, Fleet GWJ (1996) Tetrahedron: Asymmetry 7:383
112. Bichard CJF, Brandstetter TW, Estevez JC, Fleet GWJ, Hughes DJ, Wheatley JR (1996) J Chem Soc, Perkin Trans 1: Organic and Bio-Organic Chemistry: 2151
113. Binkley RW, Ambrose MG (1984) J Carbohydr Chem 3:1
114. Binkley RW (1992) J Org Chem 57:2353
115. Wang YF, Fleet GWJ, Zhao LX (1998) Carbohydr Res 307:159
116. Fuzier M, Le Merrer Y, Depezay J-C (1995) Tetrahedron Lett 36:6443
117. Poitout L, LeMerrer Y, Depezay JC (1996) Tetrahedron Lett 37:1613
118. Charette AB, Cote B (1993) J Org Chem 58:933
119. Jeong LS, Moon HR, Yoo SJ, Lee SN, Chun MW, Lim YH (1998) Tetrahedron Lett 39:5201
120. Welch JT, Svahn B-M, Eswarakrishnan S, Hutchinson JP, Zubieta J (1984) Carbohydrate Research 132:221
121. Baudry M, Barberousse V, Descotes G, Faure R, Pires J, Praly JP (1998) Tetrahedron 54:7431
122. Goebel M, Nothofer HG, Ross G, Ugi I (1997) Tetrahedron 53:3123
123. Holton RA, Somoza C, Kim HB, Liang F, Biediger RJ, Boatman PD, Shindo M, Smith CC, Kim SC, Nadizadeh H, Suzuki Y, Tao CL, Vu P, Tang SH, Zhang PS, Murthi KK, Gentile LN, Liu JH (1994) J Am Chem Soc 116:1597
124. Lee SD, Chan TH, Kwon KS (1984) Tetrahedron Lett 25:3399
125. Csuk R, Fuerstner A, Weidmann H (1986) J Carbohydr Chem 5:271
126. Curran DP, Suh Y-G (1987) Carbohydr Res 171:161
127. Fredenhagen A, Peter HH (1996) Tetrahedron 52:1235
128. Stoltz BM, Wood JL (1996) Tetrahedron Lett 37:3929
129. Avalos M, Babiano R, Cabanillas A, Cintas P, Higes FJ, Jimenez JL, Palacios JC (1996) J Org Chem 61:3738
130. Hoberg JO, Bozell JJ (1995) Tetrahedron Lett 36:6831
131. Hoberg JO (1997) J Org Chem 62:6615
132. Boyer FD, Lallemand JY (1994) Tetrahedron 50: 10443
133. Ramana CV, Murali R, Nagarajan M (1997) J Org Chem 62:7694
134. Wender PA, Jesudason CD, Nakahira H, Tamura N, Tebbe AL, Ueno Y (1997) J Am Chem Soc 119:12976
135. Chen XT, Gutteridge CE, Bhattacharya SK, Zhou BS, Pettus TRR, Hascall T, Danishefsky SJ (1998) Angew Chem Int Ed 37:185
136. Xu YM, Zhou WS (1997) J Chem Soc-Perkin Trans 1:741
137. Marshall JA, Tang Y (1994) J Org Chem 59:1457
138. Kim DK, Kim GH, Kim YW (1996) J Chem Soc-Perkin Trans 1:803
139. Tatsuta K (1998) J Synth Org Chem Japan 56:714
140. Boyer FD, Pancrazi A, Lallemand JY (1995) Synth Comm 25:1099
141. Chen YW, Vogel P (1994) J Org Chem 59:2487
142. Shilvock JP, Fleet GWJ (1998) Synlett: 554
143. Beacham AR, Smelt KH, Biggadike K, Britten CJ, Hackett L, Winchester BG, Nash RJ, Griffiths RC, Fleet GWJ (1998) Tetrahedron Lett 39:151
144. Poitout L, LeMerrer Y, Depezay JC (1996) Tetrahedron Lett 37:1609
145. Ernholt BV, Thomsen IB, Jensen KB, Bols M (1999) Synlett: 701
146. Lohse A, Jensen KB, Bols M (1999) Tetrahedron Lett 40:3033
147. Andreassen V, Svensson B, Bols M (2001) Synthesis-Stuttgart: 339
148. Moriyama H, Tsukida T, Inoue Y, Kondo H, Yoshino K, Nishimura SI (2003) Bioorg Med Chem Lett 13:2737
149. Tschamber T, Gessier F, Neuburger M, Gurcha SS, Besra GS, Streith J (2003) Eur J Org Chem: 2792
150. Fuentes J, Sayago FJ, Illangua JM, Gasch C, Angulo M, Pradera MA (2004) Tetrahedron-Asymmetry 15:603
151. Tsukida T, Moriyama H, Inoue Y, Kondo H, Yoshino K, Nishimura SI (2004) Bioorg Med Chem Lett 14:1569
152. Ichikawa Y, Osada M, Ohtani II, Isobe M (1997) J Chem Soc-Perkin Trans 1:1449
153. Blanco JLJ, Diaz Prez VM, Mellet CO, Fuentes J, Garcia Fernandez JM, Diaz Arribas JC, Canada FJ (1997) Chem Comm (Cambridge): 1969

154. Goti A, Cacciarini M, Cardona F, Cordero FM, Brandi A (2001) Org Lett 3:1367
155. Overkleeft HS, Bruggeman P, Pandit UK (1998) Tetrahedron Lett 39:3869
156. Haakansson AE, Palmelund A, Holm H, Madsen R (2006) Chemistry-A Eur J 12:3243
157. Kotha S, Mandal K, Tiwari A, Mobin SM (2006) Chemistry 12: 8024
158. Andresen TL, Skytte DM, Madsen R (2004) Org Biomol Chem 2: 2951
159. Ovaa H, Lastdrager B, Codee JDC, van der Marel GA, Overkleeft HS, van Boom JH (2002) J Chem Soc-Perkin Trans 1:2370
160. Yamazaki T, Mizutani K, Kitazume T (1996) ACS Symp Series 639:105
161. Dondoni A, Catozzi N, Marra A (2004) J Org Chem 69:5023
162. Pradera MA, Sayago FJ, Illangua JM, Gasch C, Fuentes J (2003) Tetrahedron Lett 44:6605
163. Fuentes J, Illangua JM, Sayago FJ, Angulo M, Gasch C, Pradera MA (2004) Tetrahedron-Asymmetry 15:3783
164. Dhavale DD, Jachak SM, Karche NP, Trombini C (2004) Tetrahedron 60:3009
165. Sun H, Abboud KA, Horenstein NA (2005) Tetrahedron 61:10462
166. Parr IB, Horenstein BA (1997) J Org Chem 62:7489
167. Castro S, Peczuh Mark W (2005) J Org Chem 70:3312
168. Foulard G, Brigaud T, Portella C (1997) J Org Chem 62:9107
169. Momotake A, Mito J, Yamaguchi K, Togo H, Yokoyama M (1998) J Org Chem 63: 7207
170. Zou W (2005) Curr Topic Med Chem 5:1363
171. Lin C-H, Lin H-C, Yang W-B (2005) Curr Topic Med Chem (Sharjah, United Arab Emirates) 5:1431
172. Lee DYW, He MS (2005) Curr Topic Med Chem 5:1333
173. Bililign T, Griffith BR, Thorson JS (2005) Natural Product Reports 22:742
174. Sharma GVM, Krishna PR (2004) Curr Org Chem 8:1187
175. Gascon-Lopez M, Motevalli M, Paloumbis G, Bladon P, Wyatt PB (2003) Tetrahedron 59:9349
176. Cipolla L, La Ferla B, Nicotra F (1998) Carbohydr Polym 37:291
177. Cipolla L, Peri F, La Ferla B, Redaelli C, Nicotra F (2005) Curr Org Synth 2:153
178. Johnson CR, Johns BA (1997) J Org Chem 62:6046
179. Hashimoto M, Terashima S (1994) Chem Lett: 1001
180. Brochard L, Lorin C, Spiess N, Rollin P (1998) Tetrahedron Lett 39:4267
181. Jun JG, Lee DW, Mundy BP (1998) Synth Comm 28:2499
182. Wei BG, Chen J, Huang PQ (2006) Tetrahedron 62:190
183. Pelyvas IF, Madi-Puskas M, Toth ZG, Varga Z, Hornyak M, Batta G, Sztaricskai F (1995) J Antibiotics 48:683
184. Zhou J, Wang G, Zhang L-H, Ye X-S (2006) Curr Org Chem 10:625
185. Ferrier RJ, Hoberg JO (2003) Advances in Carbohydrate Chemistry and Biochemistry, Vol 58, Academic Press, San Diego, London, 58:55
186. Chida N, Ohtsuka M, Ogura K, Ogawa S (1991) Bull Chem Soc Japan 64:2118
187. Wang A, Auzanneau FI (2007) J Org Chem 72:3585
188. Chida N, Takeoka J, Ando K, Tsutsumi N, Ogawa S (1996) Tennen Yuki Kagobutsu Toronkai Koen Yoshishu 38th: 259
189. Letellier P, Ralainirina R, Beaupere D, Uzan R (1994) Tetrahedron Lett 35:4555
190. Letellier P, Ralainairina R, Beaupere D, Uzan R (1997) Synthesis-Stuttgart: 925
191. Letellier P, ElMeslouti A, Beaupere D, Uzan R (1996) Synthesis-Stuttgart: 1435
192. Iimori T, Takahashi H, Ikegami S (1996) Tetrahedron Lett 37:649
193. Takahashi H (2001) Yuki Gosei Kagaku Kyokaishi 59:484
194. Takahashi H (2002) Yakugaku Zasshi-J Pharmaceut Soc Japan 122:755
195. Miyazaki H, Kobayashi Y, Shiozaki M, Ando O, Nakajima M, Hanzawa H, Haruyama H (1995) J Org Chem 60:6103
196. Friestad GK, Branchaud BP (1997) Tetrahedron Lett 38:5933
197. Das SK, Mallet JM, Sinay P (1997) Angew Chem Int Ed Engl 36:493
198. Sollogoub M, Mallet JM, Sinay P (2000) Angew Chem Int Ed Engl 39:362
199. Sollogoub M, Sinay P (2006) Organic Chemistry of Sugars: 349
200. Petasis NA, Yao X (2001) Abstracts of Papers of the American Chemical Society 222: U99
201. Chung S-K, Yu S-H (1996) Bioorg Med Chem Lett 6:1461
202. Jia C, Pearce AJ, Bleriot Y, Zhang Y, Zhang L-H, Sollogoub M, Sinay P (2004) Tetrahedron: Asymmetry 15:699

203. Strukul G (1998) Angew Chem Int Ed 37:1199
204. Mereyala HB, Guntha S (1995) Tetrahedron 51:1741
205. Zhou H, Wang GN, Zhang LH, Ye XS (2006) Curr Org Chem 10:625
206. Bennett SM, Biboutou RK, Zhou ZH, Pion R (1998) Tetrahedron 54:4761
207. Grove JJC, Holzapfel CW (1997) Tetrahedron Lett 38:7429
208. CronjeGrove JJ, Holzapfel CW, Williams DBG (1996) Tetrahedron Lett 37:1305
209. Chiara JL, Martinez S, Bernabe M (1996) J Org Chem 61:6488
210. Sano H, Sugai S (1995) Tetrahedron-Asymmetry 6:1143
211. deArmas P, GarciaTellado F, MarreroTellado JJ, Robles J (1997) Tetrahedron Lett 38:8081
212. Lewis BE, Choytun N, Schramm VL, Bennet AJ (2006) J Am Chem Soc 128:5049
213. Liu FW, Zhang YB, Liu HM, Song XP (2005) Carbohydr Res 340:489
214. Pacsu E (1930) J Am Chem Soc 52:2563
215. Pacsu J (1928) Berichte der Deutschen Chemischen Gesellschaft [Abteilung] B: Abhandlungen 61B:1508
216. Capozzi G, Mannocci F, Menichetti S, Nativi C, Paoletti S (1997) Chem Comm: 2291
217. Ellervik U, Jansson K, Magnusson G (1998) J Carbohydr Chem 17:777
218. Csuk R, Schaade M (1994) Tetrahedron 50:3333
219. Boons GJ, Stauch T (1996) Synlett: 906
220. Kim JM, Roy R (1997) J Carbohydr Chem 16:1281
221. Maunier V, Boullanger P, Lafont D (1997) J Carbohydr Chem 16:231
222. Gervay J, Nguyen TN, Hadd MJ (1997) Carbohydr Res 300:119
223. Fabian MA, Perrin CL, Sinnott ML (1994) J Am Chem Soc 116:8398
224. Vaino AR, Szarek WA (2001) J Org Chem 66:1097
225. Grundberg H, Eriksson-Bajtner J, Bergquist KE, Sundin A, Ellervik U (2006) J Org Chem 71:5892
226. Vaino AR, Chan SSC, Szarek WA, Thatcher GRJ (1996) J Org Chem 61:4514
227. Skorupowa E, Dmochowska B, Madaj J, Kasprzykowski F, Sokolowski J, Wisniewski A (1998) J Carbohydr Chem 17:49
228. Ren RXF, Chaudhuri NC, Paris PL, Rumney S, Kool ET (1996) J Am Chem Soc 118:7671
229. Yokoyama M, Nomura M, Togo H, Seki H (1996) J Chem Soc-Perkin Trans 1:2145
230. Di Florio R, Rizzacasa MA (1998) J Org Chem 63:8595
231. Brandstetter TW, Kim YH, Son JC, Taylor HM, Lilley PMD, Watkin DJ, Johnson LN, Oikonomakos NG, Fleet GWJ (1995) Tetrahedron Lett 36:2149
232. Leeuwenburgh MA, Timmers CM, vanderMarel GA, vanBoom JH, Mallet JM, Sinay PG (1997) Tetrahedron Lett 38:6251
233. Gervay J, Hadd MJ (1997) J Org Chem 62:6961
234. Vedso P, Chauvin R, Li Z, Bernet B, Vasella A (1994) Helvetica Chimica Acta 77:1631
235. Barker JL, Frost JW (2001) Biotechnol Bioeng 76:376
236. Zhu LZ, Talukdar A, Zhang GS, Kedenburg JP, Wang PG (2005) Synlett: 1547
237. Rowe D (2004) Chem Biodiversity 1:2034
238. vanDeurzen MPJ, vanRantwijk F, Sheldon RA (1997) J Carbohydr Chem 16:299
239. Mathews WB, Zajac WW (1995) J Carbohydr Chem 14:287
240. Devianne G, Escudier JM, Baltas M, Gorrichon L (1995) J Org Chem 60:7343
241. Lee CK, Kim EJ, Lee ISH (1998) Carbohydr Res 309:243
242. Diehl V, Cuny E, Lichtenthaler FW (1998) Heterocycles 48:1193
243. Oikawa N, Muller C, Kunz M, Lichtenthaler FW (1998) Carbohydr Res 309:269
244. Sun XL, Kai T, Takayanagi H, Furuhata K (1997) J Carbohydr Chem 16:541
245. Sano H, Mio S, Kitagawa J, Sugai S (1994) Tetrahedron-Asymmetry 5:2233
246. Chaudhari VD, Kumar KSA, Dhavale DD (2006) Tetrahedron 62:4349
247. Kozlowski JS, Marzabadi CH, Rath NP, Spilling CD (1997) Carbohydr Res 300:301
248. Mauron J (1981) Prog Food Nutr Sci 5:5
249. Ledl F, Beck J, Sengl M, Osiander H, Estendorfer S, Severin T, Huber B (1989) Prog Clin Biol Res 304:23
250. John WG, Lamb EJ (1993) Eye 7 (Pt 2): 230
251. Gerrard J (2005) The Maillard reaction: chemistry, biochemistry and implications by Harry Nursten. Royal Society of Chemistry
252. Erbersdobler HF, Somoza V (2007) Mol Nutr Food Res 51:423
253. Charissou A, Ait-Ameur L, Birlouez-Aragon I (2007) J Agric Food Chem 55:4532

254. van Boekel MA (2006) Biotechnol Adv 24:230
255. Kato A (2002) Food Sci Technol Res 8:193
256. Chen CQ, Robbins E (2000) ACS Symp Series 754:286
257. Gerrard JA (2006) Trends Food Sci Technol 17:324
258. Tuohy KM, Hinton DJ, Davies SJ, Crabbe MJ, Gibson GR, Ames JM (2006) Mol Nutr Food Res 50:847
259. Sun Y, Hayakawa S, Chuamanochan M, Fujimoto M, Innun A, Izumori K (2006) Biosci Biotechnol Biochem 70:598
260. Saraiva MA, Borges CM, Florencio MH (2006) J Mass Spectrom 41:755
261. Robert L, Labat-Robert J (2006) Pathologie Biologie 54:371
262. Reddy VP, Beyaz A (2006) Drug Discovery Today 11:646
263. Peyroux J, Sternberg M (2006) Pathol Biol (Paris) 54:405
264. Somoza V (2005) Mol Nutrit Food Res 49:663
265. Baynes JW (2000) Biogerontology 1:235
266. Colaco CA (1993) J R Soc Med 86:243
267. Wirth DD, Baertschi SW, Johnson RA, Maple SR, Miller MS, Hallenbeck DK, Gregg SM (1998) J Pharm Sci 87:31
268. Howard MJ, Smales CM (2005) J Biol Chem 280:22582
269. Rahbar S (2005) Ann N Y Acad Sci 1043:9
270. Monnier VM, Stevens VJ, Cerami A (1981) Prog Food Nutr Sci 5:315
271. Gottlieb Sheldon H (2002) Diabetes Forecast 55:34
272. Stitt Alan W (2005) Ann N Y Acad Sci 1043:582
273. Mossine VV, Glinsky GV, Barnes CL, Feather MS (1995) Carbohydr Res 266:5
274. Mossine VV, Barnes CL, Mawhinney TP (2007) Carbohydr Res 342:131
275. Yaylayan VA, Huyghues-Despointes A (1996) Carbohydr Res 286:179
276. Ge SJ, Lee TC (1996) J Agric Food Chem 44:1053
277. Yaylayan VA, Huyghues-Despointes A (1994) Crit Rev Food Sci Nutrit 34:321
278. Niwa T, Takeda N, Miyazaki T, Yoshizumi H, Tatematsu A, Maeda K, Ohara M, Tomiyama S, Niimura K (1995) Nephron 69:438
279. Yamada H, Miyata S, Igaki N, Yatabe H, Miyauchi Y, Ohara T, Sakai M, Shoda H, Oimomi M, Kasuga M (1994) J Biol Chem 269:20275
280. Hirsch J, Petrakova E, Feather MS (1992) Carbohydr Res 232:125
281. Hirsch J, Petrakova E, Feather MS, Barnes CL (1995) Carbohydr Res 267:17
282. Zyzak DV, Richardson JM, Thorpe SR, Baynes JW (1995) Arch Biochem Biophys 316:547
283. Hayase F, Shibuya T, Sato J, Yamamoto M (1996) Biosci Biotechnol Biochem 60: 1820
284. Hayase F, Nagaraj RH, Miyata S, Njoroge FG, Monnier VM (1989) J Biol Chem 264:3758
285. Yaylayan VA, Keyhani A (1999) J Agric Food Chem 47:3280
286. Kim SW, Rogers QR, Morris JG (1996) J Nutr 126:195
287. Hiramoto K, Aso-o R, Ni-iyama H, Hikage S, Kato T, Kikugawa K (1996) Mutat Res 359:17
288. Hiramoto K, Ishihara A, Sakui N, Daishima S, Kikugawa K (1998) Biol Pharm Bull 21:102
289. Hiramoto K, Kato T, Kikugawa K (1993) Mutat Res 285:191
290. Hiramoto K, Nasuhara A, Michikoshi K, Kato T, Kikugawa K (1997) Mutat Res 395:47
291. Hiramoto K, Sekiguchi K, Aso OR, Ayuha K, Ni-Iyama H, Kato T, Kikugawa K (1995) Food Chem Toxicol 33:803
292. Hiramoto S, Itoh K, Shizuuchi S, Kawachi Y, Morishita Y, Nagase M, Suzuki Y, Nobuta Y, Sudou Y, Nakamura O, Kagaya I, Goshima H, Kodama Y, Icatro Faustino C, Koizumi W, Saigenji K, Miura S, Sugiyama T, Kimura N (2004) Helicobacter 9:429
293. Lertsiri S, Fujimoto K, Miyazawa T (1995) Biochim Biophys Acta 1245:278
294. Lertsiri S, Maungma R, Assavanig A, Bhumiratana A (2001) J Food Process Preserv 25:149
295. Lertsiri S, Oak JH, Nakagawa K, Miyazawa T (2002) Biochim Biophys Acta 1573:48
296. Lertsiri S, Oak J-H, Nakagawa K, Miyazawa T (2002) Biochimica et Biophysica Acta, General Subjects 1573:48
297. Lertsiri S, Shiraishi M, Miyazawa T (1998) Biosci Biotechnol Biochem 62:893
298. Surh YJ, Liem A, Miller JA, Tannenbaum SR (1994) Carcinogenesis 15:2375
299. Tressi R, Piechotta CT, Rewicki D, Krause E (2002) Int Congr Series 1245:203
300. Jenq SN, Tsai SJ, Lee H (1994) Mutagenesis 9:483
301. Friedman M (1996) J Agric Food Chem 44:631

302. Vertesy L, Fehlhaber HW, Kogler H, Schindler PW (1996) Liebigs Annalen: 121
303. Ukeda H, Hasegawa Y, Ishi T, Sawamura M (1997) Biosci Biotechnol Biochem 61:2039
304. Ukeda H, Shimamura T, Tsubouchi M, Harada Y, Nakai Y, Sawamura M (2002) Anal Sci 18:1151
305. Buttner U, Gerum F, Severin T (1997) Carbohydr Res 300:265
306. Buttner U, Ochs S, Severin T (1996) Carbohydr Res 291:175
307. Ahmed MU, Dunn JA, Walla MD, Thorpe SR, Baynes JW (1988) J Biol Chem 263:8816
308. Ahmed MU, Thorpe SR, Baynes JW (1986) J Biol Chem 261:4889
309. Ahmed N, Babaei-Jadidi R, Howell SK, Thornalley PJ, Beisswenger PJ (2005) Diabetes Care 28:2465
310. Glomb MA, Pfahler C (2001) J Biol Chem 276:41638
311. Glomb MA, Tschirnich R (2001) J Agricult Food Chem 49:5543
312. Lo TW, Westwood ME, McLellan AC, Selwood T, Thornalley PJ (1994) J Biol Chem 269:32299
313. Lopez MG, Mancilla-Margalli NA (2000) Frontiers of Flavour Science. Proceedings of the 9th Weurman Flavour Research Symposium, Freising, Germany, June 22–25
314. Li EY, Feather MS (1994) Carbohydr Res 256:41
315. Nagaraj RH, Monnier VM (1995) Biochim Biophys Acta 1253:75
316. Nagaraj RH, Sady C (1996) FEBS Lett 382:234
317. Nagaraj RH, Sarkar P, Mally A, Biemel KM, Lederer MO, Padayatti PS (2002) Arch Biochem Biophys 402:110
318. Pischetsrieder M, Larisch B, Seidel W (1997) J Agric Food Chem 45:2070
319. Dyer DG, Blackledge JA, Katz BM, Hull CJ, Adkisson HD, Thorpe SR, Lyons TJ, Baynes JW (1991) Z Ernahrungswiss 30:29
320. Dyer DG, Blackledge JA, Thorpe SR, Baynes JW (1991) J Biol Chem 266:11654
321. Smith PR, Somani HH, Thornalley PJ, Benn J, Sonksen PH (1993) Clin Sci (Lond) 84:87
322. Smith PR, Thornalley PJ (1992) Eur J Biochem 210:729
323. Nissimov J, Elchalal U, Bakala H, Brownlee M, Berry E, Phillip M, Milner Y (2007) J Immunolog Method 320:1
324. Meerwaldt R, Lutgers HL, Links TP, Graaff R, Baynes JW, Gans ROB, Smit AJ (2007) Diabetes Care 30:107
325. Yu Y, Thorpe SR, Jenkins AJ, Shaw JN, Sochaski MA, Mcgee D, Aston CE, Orchard TJ, Silvers N, Peng YG, McKnight JA, Baynes JW, Lyons TJ, Grp DER (2006) Diabetologia 49:2488
326. Lutgers HL, Graaff R, Links TP, Ubink-Veltmaat LJ, Bilo HJ, Gans RO, Smit AJ (2006) Diabetes Care 29:2654
327. Stitt AW, Jenkins AJ, Cooper ME (2002) Expert Opin Inv Drug 11:1205
328. Raj DSC, Choudhury D, Welbourne TC, Levi M (2000) Am J Kidney Dis 35:365
329. Chiarelli F, Catino M, Tumini S, Cipollone F, Mezzetti A, Vanelli M, Verrotti A (2000) Pediatr Nephrol 14:841
330. DeGroot J (2004) Curr Opin Pharmacol 4:301
331. Chellan P, Nagaraj RH (2001) J Biol Chem 276:3895
332. Biemel KM, Reihl O, Conrad J, Lederer MO (2001) J Biol Chem 276:23405

Part 3
Chemical Glycosylation Reactions

3.1 Glycosyl Halides

Kazunobu Toshima
Department of Applied Chemistry, Faculty of Science and Technology,
Keio University, Yokohama 223–8522, Japan
toshima@applc.keio.ac.jp

Abstract

This chapter describes the preparations and chemical glycosylation reactions of glycosyl halides as glycosyl donors including glycosyl bromides, chlorides, iodides, and fluorides. For a survey on the general current methodological advances, glycosyl halide donors are classified into four groups based on the type of anomeric functional group and their activating methods. Among them, glycosyl fluorides, which are frequently and widely used in current glycosylation reactions, are particularly emphasized and discussed in this chapter.

Keywords

Glycosylation; Glycosyl donor; Glycosyl halide; Glycosyl bromide; Glycosyl chloride; Glycosyl iodide; Glycosyl fluoride

Abbreviations

Ac acetyl
acac acetylacetone
Ar aryl
Bn benzyl
BTF benzotrifluoride
***t*-Bu** *tert*-butyl
Bz benzoyl
Cp cyclopentadienyl
DAST diethylaminosulfur trifluoride
DDQ 2,3-dichloro-5,6-dicyano-1,4-benzoquinone
DEAD diethyl azodicarboxylate
DIEA *N*,*N*-diisopropylethylamine

In: *Glycoscience*. Fraser-Reid B, Tatsuta K, Thiem J (eds)
Chapter-DOI 10-1007/978-3-540-30429-6_10:

Et ethyl
Me methyl
MS molecular sieves
NBS *N*-bromosuccinimide
NIS *N*-iodosuccinimide
Ph phenyl
***i*-Pr** isopropyl
PPTS pyridinium *p*-toluenesulfonate
Pv pivaloyl
TBAF tetrabutylammonium fluoride
TBAI tetrabutylammonium iodide
TEA triethylamine
Tf trifluoromethanesulfonyl
THF tetrahydrofuran
TMS trimethylsilyl
TMU 1,1,3,3-tetramethylurea
Tr triphenylmethyl
Ts toluenesulfonyl

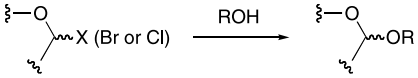
1 Chemical Glycosylation of Glycosyl Bromide and Chloride

The use of glycosyl bromide or chloride as an effective glycosyl donor in the glycosylation reaction was first introduced by Koenigs and Knorr in 1901 [1]. In relation to the anomeric stereochemistry of the glycosylation reaction, three significant basic methods, (1) the neighboring group assisted method for construction of 1,2-*trans* glycosides such as β-gluco- or α-manno type glycoside, (2) the in-situ anomerization method for synthesis of α-gluco or α-manno type glycoside [2], and (3) the heterogenic catalyst method for preparation of β-mannoglycoside [3], were developed in this area [4d, f]. The well-known classical Koenigs–Knorr method used heavy metal salts (mainly silver and mercury salts) as activating reagents as summarized in ❯ *Table 1*. A variety of heavy metal salts such as $AgClO_4$, AgOTf, $AgNO_3$, Ag_2CO_3, Ag_2O, Ag-silicate, $Hg(CN)_2$, $HgBr_2$, $HgCl_2$, and HgI_2 [5] and their combined use were employed in this area [4a–h, j, l]. The order of reactivity of some representative catalysts was generally confirmed [4d, f]. Furthermore, Ag_2CO_3, Ag_2O, HgO, $CdCO_3$, *S*-collidine, and TMU were frequently used as an acid scavenger, and water was generally removed by Drierite® and molecular sieves during these glycosylation reactions [4a–h, j, l].

On the other hand, other glycosylation methods using glycosyl bromide or chloride in the absence of any metal were also widely studied. Lemieux et al. [6] introduced a mild glycosylation method in the presence of Bu_4NBr. Also, the glycosylation protocols which involved a transformation of glycosyl bromide into the corresponding onium salts by Et_3N, Ph_3P, and Me_2S were developed by Schuerch et al. [7]. Furthermore, the uses of other activating reagents without heavy metals and including several Lewis acids such as $SnCl_4$ [8], BF_3Et_2O [8], $Sn(OTf)_2$-collidine [9a], $Sn(OTf)_2$-TMU [9b], TrCl-$ZnCl_2$ [10], $LiClO_4$ [11], I_2-DDQ [12],

Table 1

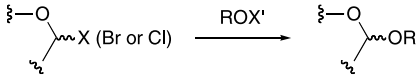
Glycosylation of glycosyl bromide or chloride using heavy metal



Activator	Acid scavenger	Drying agent	Ref.
$AgClO_4$	Ag_2CO_3	Drierite	[4a–h, j, l,5]
AgOTf	Ag_2O	molecular sieves	
$AgNO_3$	HgO		
Ag_2CO_3	$CdCO_3$		
Ag_2O	*s*-collidine		
Ag-silicate	TMU		
Hg(CN)			
$HgBr_2$			
$HgCl_2$			
HgI_2			

Table 2

Glycosylation of glycosyl bromide or chloride using Lewis acid



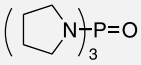
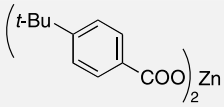
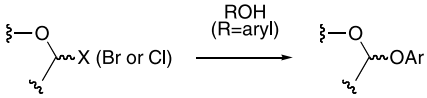
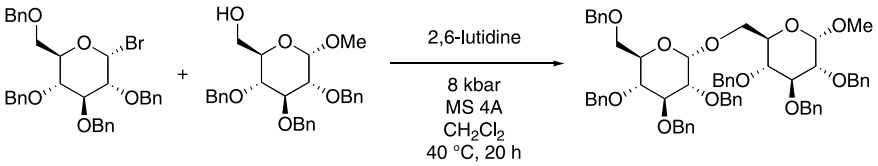
Activator	X′	Ref.
$SnCl_4$	$SnBu_3$	[8]
$BF_3{\cdot}OEt_2$	$SnBu_3$	[8]
$Sn(OTf)_2$-collidine	H	[9a]
$Sn(OTf)_2$-TMU	H	[9b]
TrCl-$ZnCl_2$	H	[10]
$LiClO_4$	H	[11]
I_2-DDQ	H	[12]
IBr	H	[13]
NIS	H	[14]
In	H	[15]
$InCl_3$	H	[16]
TMSOTf	$SnBu_3$	[17]
		
$Zn(acac)_2$	H	[18,19]
$Zn(acac)_2$	H	[20]
		
$Cu(OTf)_2$-BTF	H	[21]
$Cu(OTf)_2$-BTF	H	[22]

Table 3

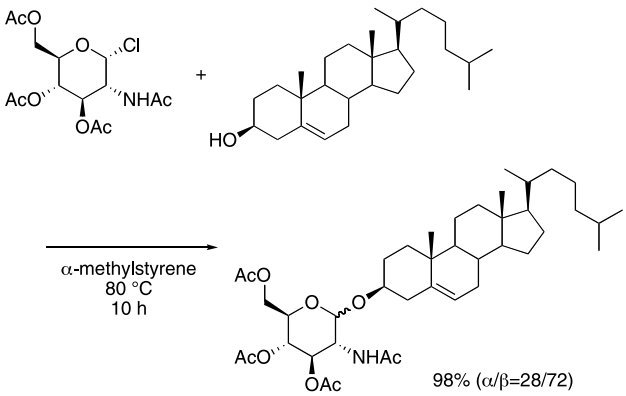
Glycosylation of glycosyl bromide or chloride using phase transfer catalyst



Catalyst	Conditions	Ref.
$Et_3N^+BnBr^-$	$CHCl_3/H_2O/NaOH$	[23]
$Et_3N^+BnCl^-$	$CH_2Cl_2/H_2O/NaOH$ or KOH	[24]
$Me(CH_2)_{15}N^+Me_3Br^-$	$CH_2Cl_2/H_2O/NaOH$	[25,26]
$Bu_4N^+Br^-$	$CH_2Cl_2/H_2O/NaOH$	[27]
$Bu_4N^+HSO_4^-$	$CH_2Cl_2/H_2O/NaOH$ or $NaHCO_3$	[28]
Aliquat®336 (R = alkyl)	$CH_2Cl_2/H_2O/K_2CO_3$	[29]
$BnN^+Bu_3Cl^-$	$CHCl_3/H_2O/K_2CO_3$	[30]



Scheme 1



Scheme 2

IBr [13], NIS [14], In [15], $InCl_3$ [16], TMSOTf [17], $(C_4H_8N)_3P=O$ [18,19], $Zn(acac)_2$ [20], $(t\text{-}BuC_6H_4CO_2)_2Zn$ [21], and $Cu(OTf)_2$-BTF [22] were reported in this field (*Table 2*).

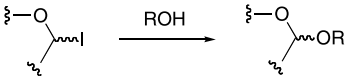
The glycosylations of aryl alcohols using a phase transfer catalyst such as $Et_3N^+BnBr^-$ [23a, b], $Et_3N^+BnCl^-$ [24], $Me(CH_2)_{15}N^+Me_3Br^-$ [25,26], $Bu_4N^+Br^-$ [27], $Bu_4NH^+SO_4^-$ [28], Aliquat®336 [29], and $BnN^+Bu_3Cl^-$ [30] were also developed (*Table 3*).

Alternatively, Sasaki et al. [31] provided a glycosylation method using glycosyl bromide in the presence of hindered amines such as 2,6-lutidine or TMU under high-pressure conditions (Scheme 1). In addition, Nishizawa et al. [32] developed a thermal glycosylation of glycosyl chloride in the presence of α-methylstyrene or TMU as an acid scavenger without any metal salts (Scheme 2).

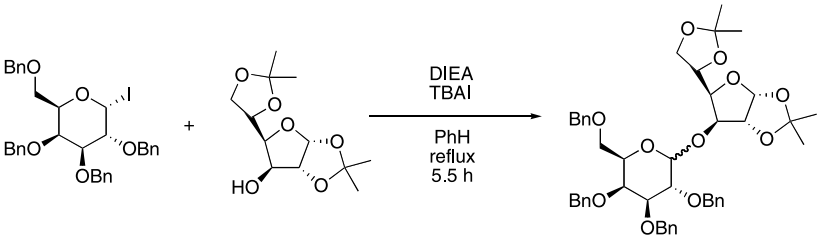
2 Chemical Glycosylation of Glycosyl Iodide

Among the glycosyl halides, only a few examples of the use of glycosyl iodide as an isolatable glycosyl donor are demonstrated due to their instability (Table 4). Thus, $LiClO_4$ [33], TBAI-DIEA [34], $FeCl_3$-I_2 [35], CuCl-I_2 [35], and NIS-I_2-TMSOTf [35] were reported as the activators of the glycosyl iodides (Scheme 3).

Table 4
Glycosylation of glycosyl iodide



Activator	X	Ref.
$LiClO_4$	H	[33]
TBAI-DIPEA	H	[34]
NIS-I_2-TMSOTf	H	[35]
$FeCl_3$-I_2	H	[35]
CuCl-I_2	H	[35]



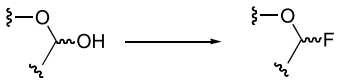
Scheme 3

3 Preparation and Chemical Glycosylation of Glycosyl Fluoride

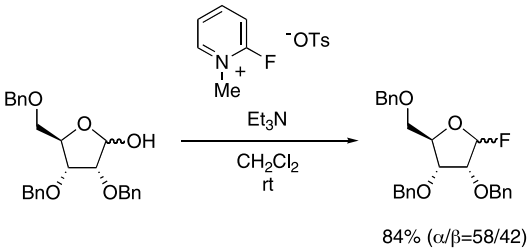
3.1 Preparation of Glycosyl Fluoride

Glycosyl fluorides have been widely and effectively used for glycosylation reactions. One of the notable advantages of the glycosyl fluoride as a glycosyl donor is its higher thermal and chemical stability as compared to other glycosyl halides such as glycosyl chlorides, bromides,

Table 5
Preparation of glycosyl fluoride from aldose



Reagent	Ref.
2-F-C_5H_4NMe·OTs-Et_3N	[37]
DAST	[38,39]
HF/Py	[40,41]
TiF_4	[42]
CF_3ZnBr·2MeCN-TiF_4	[42]
$M_2C=C(F)NMe_2$	[43]
DEAD-PPh_3-Et_3OBF_3	[44]



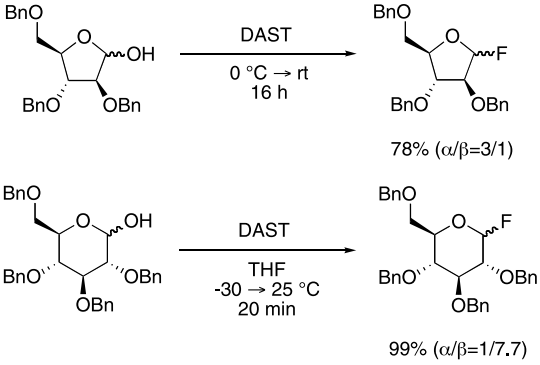
Scheme 4

and iodides. Therefore, the glycosyl fluoride can be generally purified by the appropriate distillation and even by column chromatography with silica gel. The practical use of glycosyl fluoride as a glycosyl donor was first introduced by Mukaiyama et al. in 1981 [36]. After the significant advance in this field, a number of effective methods for fluorinating the anomeric center of several types of sugars have been developed. The representative methods for the conversion of a free anomeric hydroxyl group of sugars into the fluoride are summarized in *Table 5*.

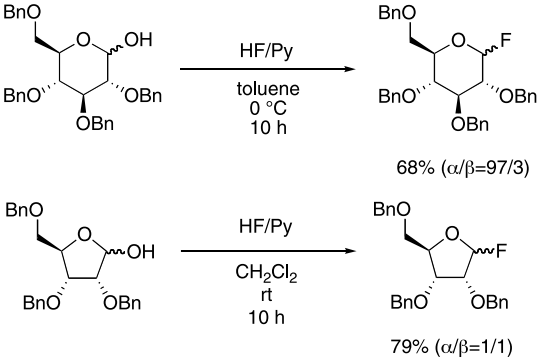
Mukaiyama and coworkers reported that totally benzylated ribofuranose smoothly reacted with 2-fluoro-1-methylpyridinium tosylate in the presence of triethylamine at room temperature to afford an anomeric mixture of the corresponding ribofuranosyl fluorides in high yield (*Scheme 4*) [37].

Diethylaminosulfur trifluoride (DAST) is now widely used as the most preferable fluorinating reagent of 1-glycoses (*Scheme 5*) [38,39]. Pyridinium poly(hydrogen fluoride) (HF/Py) was also found to be an effective reagent for the conversion of various furanose and pyranose hemiacetals into the corresponding glycosyl fluorides (*Scheme 6*) [40,41].

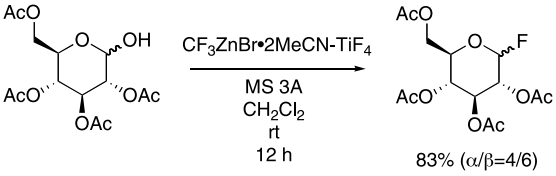
On the other hand, use of metal fluorides such as TiF_4 and the complex CF_3ZnBr·2MeCN-TiF_4 also afforded the glycosyl fluorides from the corresponding glycoses (*Scheme 7*) [42].



Scheme 5



Scheme 6

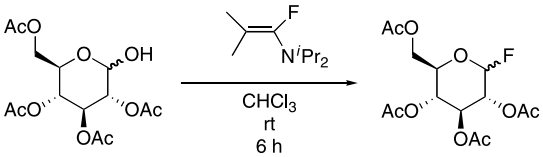


Scheme 7

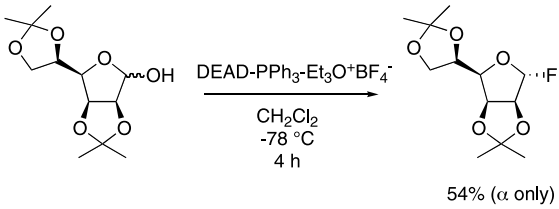
Furthermore, the anomeric hydroxyl group of various furanose and pyranose hemiacetals can be replaced by a fluorine under neutral conditions using an α-fluoroenamine (*Scheme 8*) [43].

Glycosyl fluorides can also be obtained by the modified Mitsunobu reaction using diethyl azodicarboxylate (DEAD)-PPh_3-$Et_3O^+BF_4^-$ (*Scheme 9*) [44]. By this method, the acid sensitive isopropylidene-protected mannofuranosyl fluoride was prepared in moderate yield.

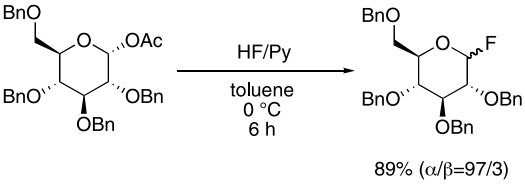
Alternatively, 1-*O*-acylated sugars can be easily transformed into the corresponding glycosyl fluorides. Thus, treatment of 1-*O*-acetylated sugars with HF/Py gave the corresponding glyco-



Scheme 8



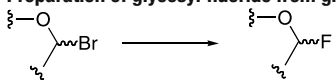
Scheme 9



Scheme 10

Table 6

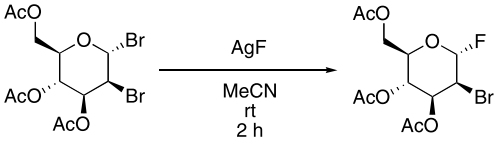
Preparation of glycosyl fluoride from glycosyl bromide



Reagent	Ref.
AgF	[45,46]
ZnF_2-2,2′-bipyridine	[47]
CF_3ZnBr·2MeCN	[42]

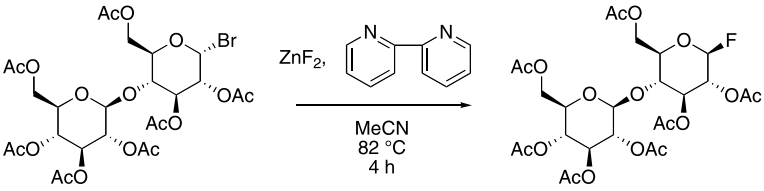
syl fluorides in high yields. Compared to the unprotected analogue, the 1-*O*-acetylated derivatives preferentially underwent fluorination by this reagent (*Scheme 10*) [40]. Furthermore, it was found that the reaction of penta-*O*-acetyl-β-D-glucopyranose, which had a participating acetoxy group at the C-2 position, gave a mixture of the α-and β-fluorides in which the latter predominated.

The preparation of glycosyl fluorides from other glycosyl halides is summarized in *Table 6*. Glycosyl bromides were effectively reacted with metal fluorides such as AgF (*Scheme 11*)



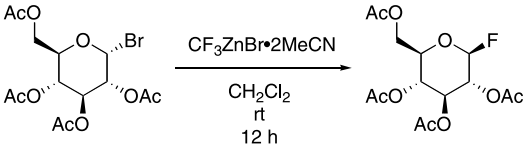
88% (α only)

Scheme 11



69% (β only)

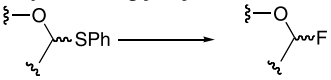
Scheme 12



83% (β only)

Scheme 13

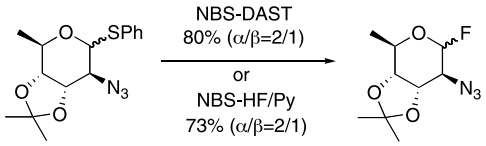
Table 7
Preparation of glycosyl fluoride from thioglycoside



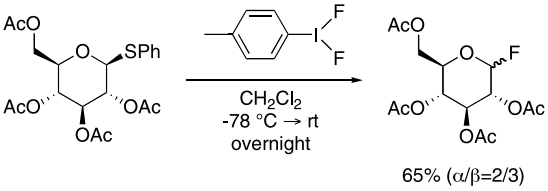
Reagent	Ref.
NBS-DAST	[48]
NBS-HF/Py	[48]
4-$MeC_6H_4IF_2$	[49]

[45,46], ZnF_2−2,2′-bipyridine (*Scheme 12*) [47] or $CF_3ZnBr{\cdot}2MeCN$ (*Scheme 13*) [42] to afford the corresponding glycosyl fluorides in high yields via the nucleophilic halide exchange reaction. Therefore, in general, the glycosyl fluorides were obtained with stereo-inversion of the anomeric center of the glycosyl bromides.

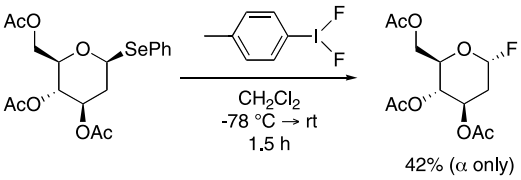
On the other hand, the glycosyl fluorides can be prepared from the corresponding thioglycosides (*Table 7*). Thus, the reactions of thioglycosides with DAST-NBS or HF/Py-NBS gave the corresponding glycosyl fluorides in high yields (*Scheme 14*) [48]. Furthermore, the thio-



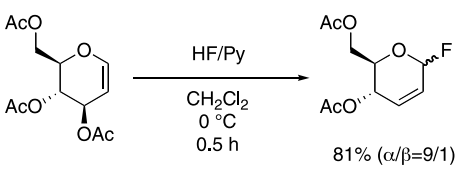
Scheme 14



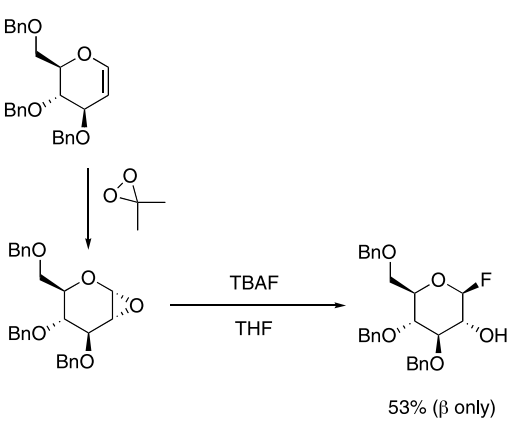
Scheme 15



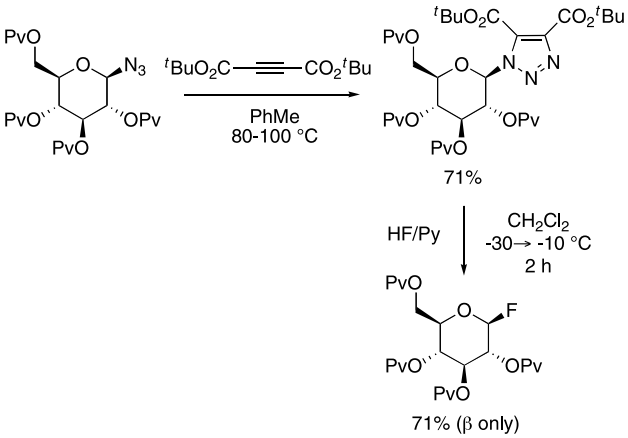
Scheme 16



Scheme 17



Scheme 18



Scheme 19

glycosides may also react with the hypervalent iodoarene, 4-methyl(difluoroiodo)benzene, to yield the corresponding glycosyl fluorides (*Scheme 15*) [49,50]. It is noteworthy that the yields of the produced glycosyl fluorides are significantly improved using the corresponding *p*-chloro-phenylthioglycosides as glycosyl donors in these reactions.

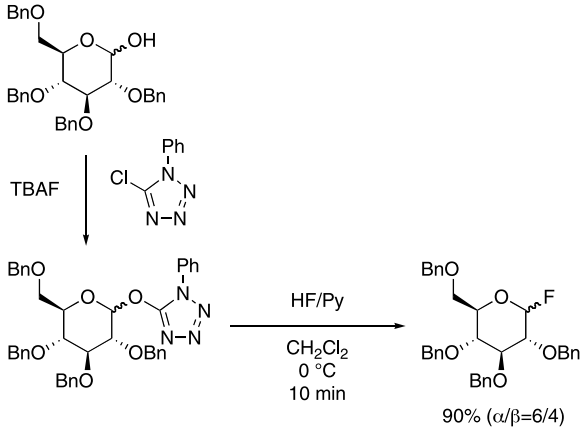
The selenoglycosides were also found to be transformed into the glycosyl fluorides using 4-methyl(difluoroiodo)benzene via an S_N2 inversion of the anomeric centers (*Scheme 16*) [50].

Glycals and 1,2-anhydro sugars, the latter of which are readily obtained from glycals, were prepared and then transformed into the corresponding glycosyl fluorides using HF/Py (*Scheme 17*) [51] and tetrabutylammonium fluoride (TBAF) (*Scheme 18*) [52], respectively. In the former case, the corresponding 2,3-unsaturated glycosyl fluorides are exclusively produced via the Ferrier rearrangement [53] while 1,2-trans glycosyl fluorides are stereoselectively obtained in the latter case.

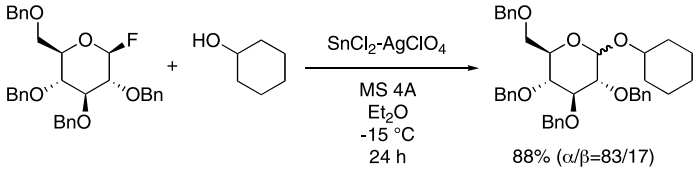
Furthermore, *N*-glycosyl triazole derivatives (*Scheme 19*) [54], prepared from the corresponding glycosyl azides, and (1-phenyl-1*H*-tetrazol-5-yl)glycosides synthesized from the corresponding glycoses and 5-chloro-1-phenyl-1*H*-tetrazole (*Scheme 20*) [55] both smoothly react with HF/Py to furnish the corresponding glycosyl fluorides in good yields.

3.2 Chemical Glycosylation of Glycosyl Fluoride

Since 1981 when Mukaiyama introduced the practical use of glycosyl fluoride with a fluorophilic activator, $SnCl_2$-$AgClO_4$, as a glycosyl donor (*Scheme 21*) [36], a number of specific fluorophilic reagents have been developed for the effective glycosylation reactions (*Table 8*). Following the $SnCl_2$-$AgClO_4$ promoter, the combined use of $SnCl_2$-$TrClO_4$ (*Scheme 22*) [37] and $SnCl_2$-AgOTf (*Scheme 23*) [56] were reported by Mukaiyama and Ogawa, respectively. In these cases, 1,2-*cis*-α-glycosides were obtained predominantly in high yields.



Scheme 20



Scheme 21

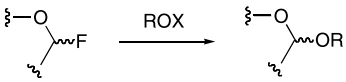
In 1984, Noyori and coworkers announced that the silyl compounds, both SiF_4 and trimethylsilyl trifluoromethanesulfonate (TMSOTf) were very effective for activation of the glycosyl fluorides (*Scheme 24*) [57]. Furthermore, it was found that the stereoselectivity of the glycosylation was highly dependent on the reaction solvent. Thus, the glycosylation in MeCN exclusively gave the β-glycoside while the glycosylation in Et_2O predominately afforded the α-glycoside.

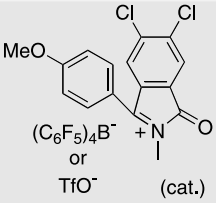
Furthermore, Nicolaou, Kunz, and Vozny independently reported that glycosyl fluorides effectively reacted with a variety of free alcohols and silyl ethers using BF_3Et_2O as an activator to give the corresponding *O*-glycosides in good yields (*Scheme 25*) [58,59,60,61].

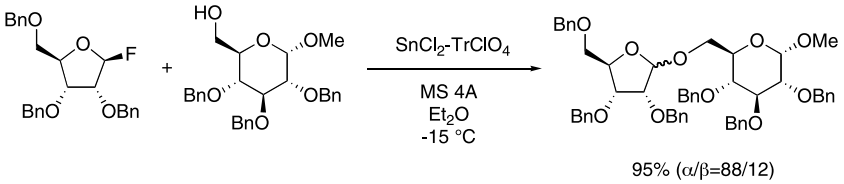
On the other hand, metal fluorides such as TiF_4 and SnF_4 were also used as effective promoters of glycosyl fluorides by Thiem et al., and the stereoselective glycosylation of 2-deoxy glycosyl fluoride was carried out with use of TiF_4. In the case of 2-deoxy-β-glycosyl fluoride, when hexane was used as the solvent, the α-glycoside was selectively produced with the inversion of the anomeric center. On the other hand, when the reaction was performed in Et_2O, the β-glycoside was obtained as the major product via "double S_N2" mechanism which involved the formation of the oxionium cation-ether complex (*Scheme 26*) [62].

Suzuki and coworkers developed new and quite effective methods in which the combined activators including the group IV_B metallocenes such as Cp_2MCl_2-$AgClO_4$ (M = Zr, Hf) (*Scheme 27*) [63], Cp_2ZrCl_2-$AgBF_4$ (*Scheme 28*) [64], and Cp_2HfCl_2-AgOTf [64,65] were used as milder reagents for promoting the glycosylations of glycosyl fluorides. Fur-

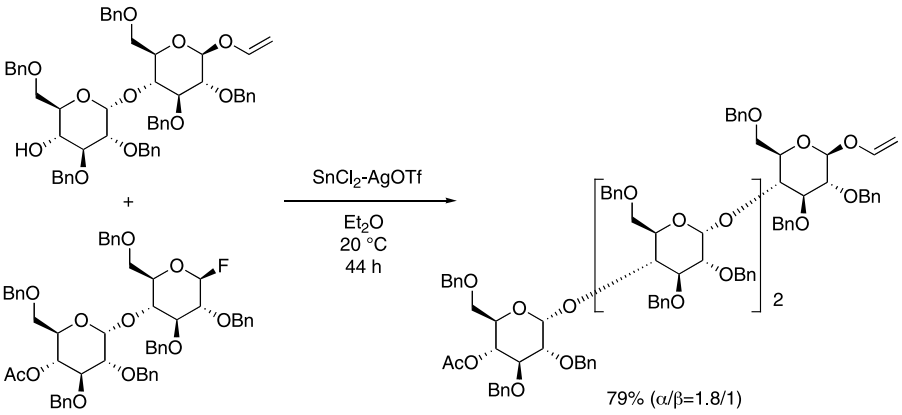
Table 8
Glycosylation of glycosyl fluoride



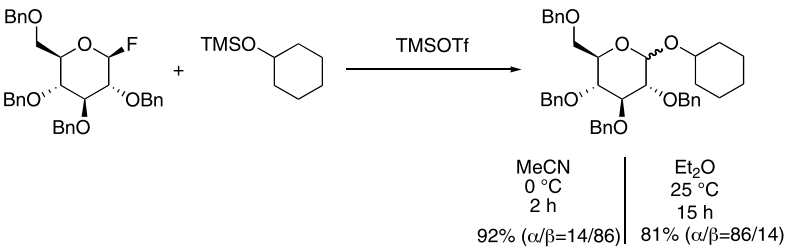
Activator	X	Ref.
$SnCl_2$-$AgClO_4$	H	[36]
$SnCl_2$-$TrClO_4$	H	[37]
$SnCl_2$-AgOTf	H	[56]
TMSOTf (cat.)	TMS	[57]
SiF_4 (cat.)	TMS	[58]
$BF_3{\cdot}Et_2O$	H	[58,59,60,61]
TiF_4	H	[62]
SnF_4	H	[62]
Cp_2MCl_2-$AgClO_4$ (M = Zr or Hf)	H	[63]
Cp_2ZrCl_2-$AgBF_4$	H	[64]
Cp_2HfCl_2-AgOTf	H	[64,65]
$Bu_2Sn(ClO_4)_2$	H	[66]
Me_2GaCl	H	[67]
Tf_2O	H	[68]
$LiClO_4$	H	[69]
$Yb(OTf)_3$	H	[70]
$La(ClO_4)_3{\cdot}nH_2O$ (cat.)	TMS	[71]
$La(ClO_4)_3{\cdot}nH_2O$-$Sn(OTf)_2$	H	[72]
Yb-Amberlyst 15	H	[73]
SO_4/ZrO_2	H	[74]
Nafion-H	H	[74]
montmorillonite K-10	H	[74]
$TrB(C_6F_5)_4$ (cat.)	H	[75]
 (cat.)	H	[75]
$SnCl_4$-$AgB(C_6F_5)_4$ (cat.)	H	[76]
TfOH (cat.)	H	[77]
$HB(C_6F_5)_4$ (cat.)	H	[78]
$HClO_4$ (cat.)	H	[78]
$HNTf_2$ (cat.)	H	[78]
$Yb(NTf_2)_3$	H	[79]
$ZrCl_4$	TMS	[80]
$Cu(OTf)_2$	H	[22]



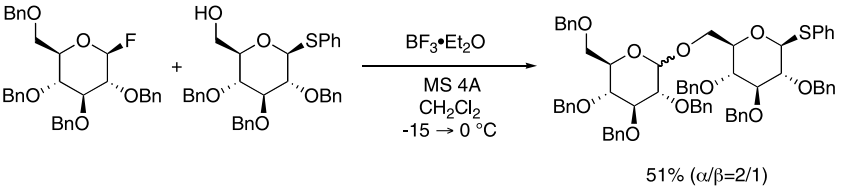
Scheme 22



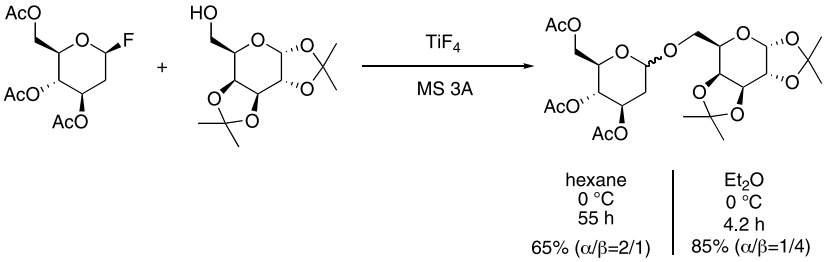
Scheme 23



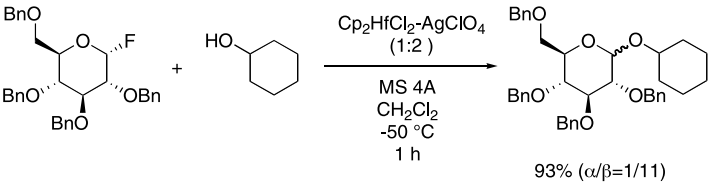
Scheme 24



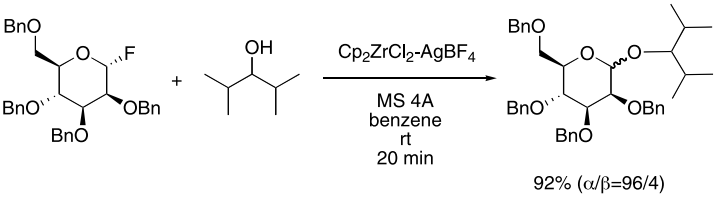
Scheme 25



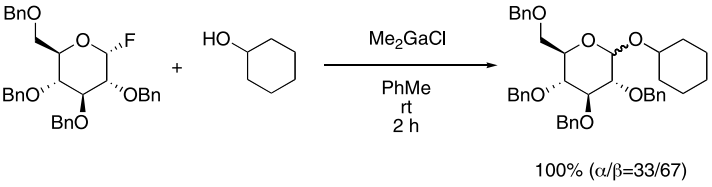
Scheme 26



Scheme 27



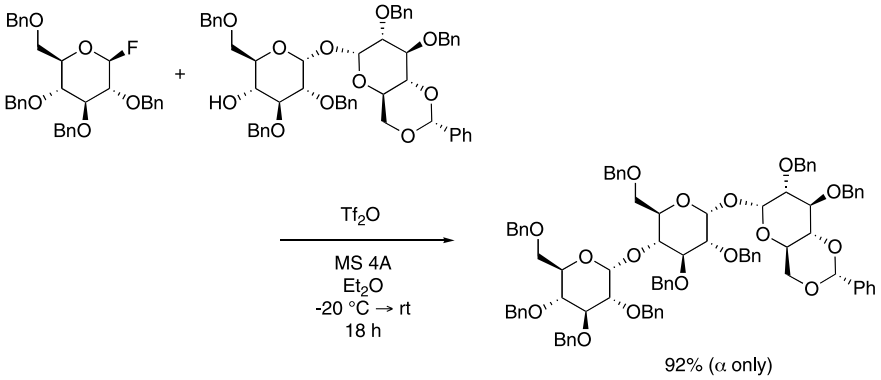
Scheme 28



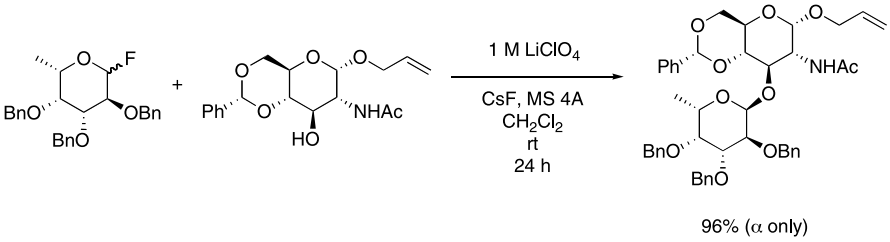
Scheme 29

thermore, the novel combined use of Bu_2SnCl_2–$2AgClO_4$ was found to show promise as an effective promoter of the glycosyl fluoride [66].

Me_2GaCl and Me_2GaOTf were introduced as new promoters of the glycosyl fluorides by Kobayashi due to their strong affinity for fluoride. In this study, it was found that the readily available Me_2GaCl was more effective for the glycosylation reactions of glycosyl fluorides (*Scheme 29*) [67].



Scheme 30



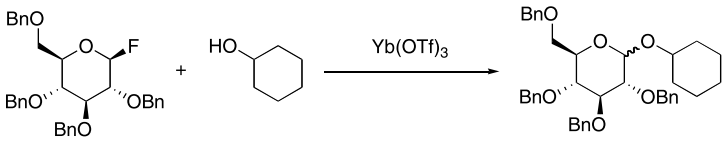
Scheme 31

Furthermore, Wessel et al. announced that Tf_2O was a highly reactive activator during the glycosylation of glycosyl fluorides (*Scheme 30*) [68]. In this report, it was suggested that a sequence of relative reactivity of the examined catalysts was TMSOTf < $SnCl_2$-AgOTf < TiF_4 < Tf_2O.

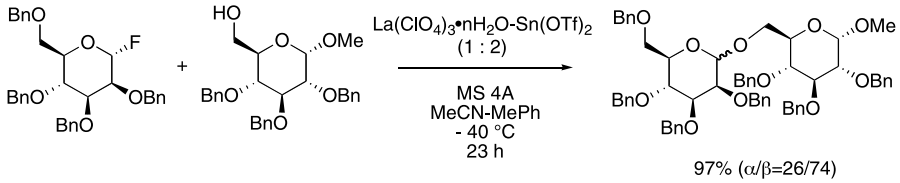
On the other hand, Waldmann et al. reported the use of $LiClO_4$, a milder Lewis acid, for the glycosylation of fucosyl fluoride under neutral conditions (*Scheme 31*) [69].

Shibasaki and coworkers developed the rare earth metal salts, such as $La(ClO_4)_3{\cdot}nH_2O$ and $Yb(OTf)_3$, catalyzed glycosylations of glycosyl fluorides [70,71,72]. The use of either $Yb(OTf)_3$ or $YbCl_3$ in the presence of $CaCO_3$ and MS 4A in Et_2O was found to be effective for α-stereoselective glucosylation. On the other hand, for β-stereoselective glucosylation, the utilization of $Yb(OTf)_3$ containing K_2CO_3 and MS 4A in MeCN gave the best result (*Scheme 32*) [70]. Furthermore, it was found that the combined use of $La(ClO_4)_3{\cdot}nH_2O$ and $Sn(OTf)_2$ was very useful for β-stereoselective mannosylation (*Scheme 33*) [72]. Along this line, Wang et al. reported the use of lanthanide (III) catalysts supported on ion exchange resins for simple glycosylations of glycosyl fluorides with methanol [73].

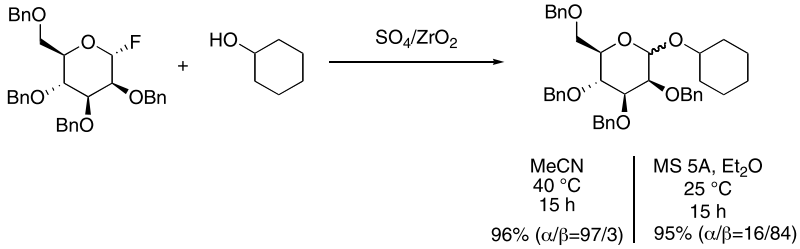
Toshima and coworkers demonstrated that environmentally friendly heterogeneous catalysts such as montmorillonite K-10, Nafion-H®, and SO_4/ZrO_2 were very effective for the glycosylations of glycosyl fluorides [74]. This is the first report on the use of a protic acid for activation of glycosyl fluoride. Among them, SO_4/ZrO_2 was shown to be superior to the



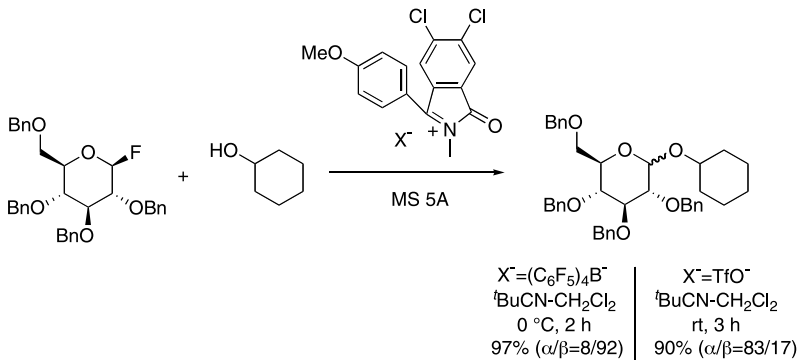
Scheme 32



Scheme 33



Scheme 34



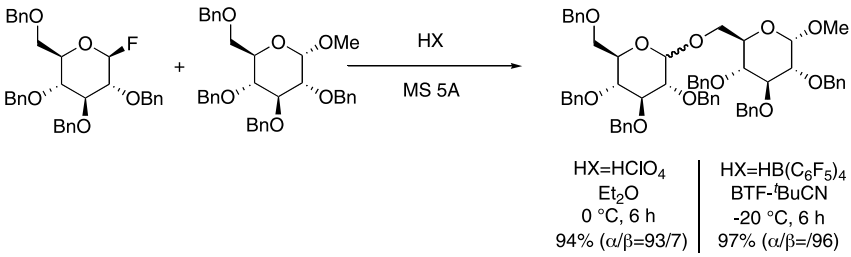
Scheme 35

others for the stereocontrolled glycosylation with α-mannopyranosyl fluoride. Thus, the glycosylations of perbenzylated α-mannopyranosyl and and alcohols using SO_4/ZrO_2in MeCN exclusively gave the corresponding α-glycosides. On the other hand, the corresponding β-glycosides were selectively obtained by the glycosylations employing SO_4/ZrO_2in the presence of MS 5A in Et_2O (Scheme 34). Furthermore, this method could be applied to the stereocontrolled glycosylation of 2-deoxy-α-glucopyranosyl fluoride.

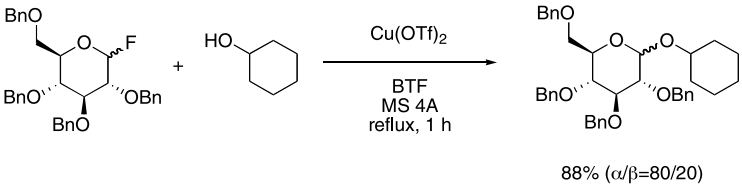
Mukaiyama et al. developed the glycosylations of glycosyl fluorides using carbocationic species paired with tetrakis(pentafluorophenyl) borate or trifluoromethanesulfonate [75]. In this study, it was found that the glycosylation using the former catalyst in CH_2Cl_2 containing *t*-BuCN gave the β-glycoside selectively while the α-glycoside was predominately obtained by the glycosylation using the latter catalyst in CH_2Cl_2 containing Et_2O (Scheme 35). In addition, they found that the use of a 1:2 combination of $SnCl_4$ and $AgB(C_6F_5)_4$ was very effective for the glycosylations of glycosyl fluorides possessing phthaloyl or dichlorophthaloyl protecting group for the C-2 amino function [76].

In contrast to Toshima's glycosylations of glycosyl fluorides using heterogeneous protic acids such as montmorillonite K-10, Nafion-H®, and SO_4/ZrO_2 [74], Mukaiyama and coworkers demonstrated the catalytic glycosylations using various homogeneous protic acids such as TfOH [77], $HClO_4$, $HB(C_6F_5)_4$, and $HNTf_2$ [78]. Among them, when the glycosylation was performed using $HClO_4$ in Et_2O, the α-glycoside was major product while the β-stereoselectivity was observed when $HB(C_6F_5)_4$ was employed in a mixture of BTF and *t*-BuCN (Scheme 36).

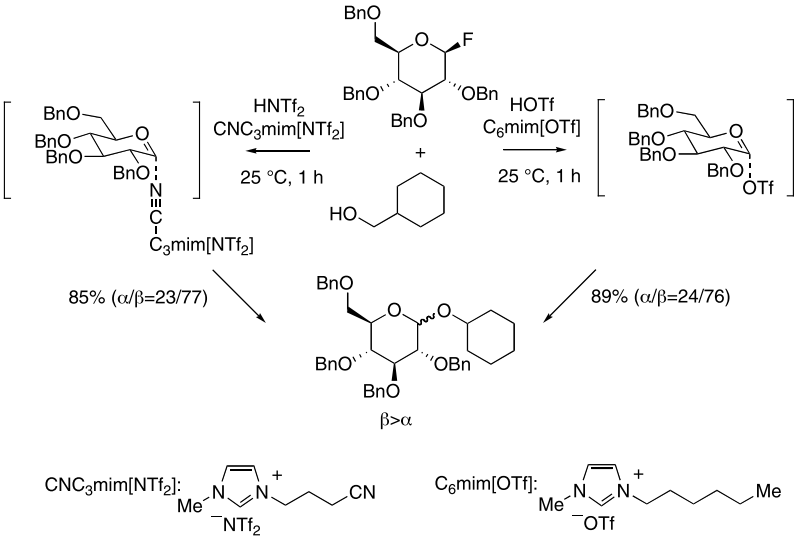
$Yb(NTf_2)_3$ [79] and $ZrCl_4$ [80] were also employed by Switzer et al. and Inazu et al., respectively. Furthermore, Yamada reported that $Cu(OTf)_2$ was a promoter not only for glycosyl bromide but also for glycosyl fluoride [22] (Scheme 37).



Scheme 36



Scheme 37



Scheme 38

Following the extensive studies using a heterogeneous and reusable solid catalyst in a glycosylation reaction by Toshima et al. for greening the chemical glycosylation method [74], the use of an ionic liquid as an environmentally benign reaction media for the glycosylation of glycosyl fluoride was reported by Toshima et al. [81]. In this study, it was demonstrated not only the reusability of the ionic liquid in the glycosylation reaction but also the potency of the ionic liquid for stereocontrol of the glycosylation reaction (*Scheme 38*).

References

1. Koenigs W, Knorr E (1901) Chem Ber 34:957
2. Lemieux RU, Hayami JI (1965) Can J Chem 43:2162
3. (a) Paulsen H, Lockhoff O (1981) Chem Ber 114:3102; (b) Paulsen H, Kutschker W, Lockhoff O (1981) Chem Ber 114:3233
4. (a) Wulff G, Röhle G (1974) Angew Chem Int Ed Engl 13:157; (b) Bochkov A-F, Zaikov GE (1979) (eds) Chemistry of the *O*-Glycosidic Bond: Formation and Cleavage. Pergamon Press, Oxford; (c) Tsutsumi H, Ishido Y (1980) J Synth Org Chem Jpn 38:473; (d) Paulsen H (1982) Angew Chem Int Ed Engl 21:155; (e) Koto S, Morishima N, Zen S (1983) J Synth Org Chem Jpn 41:701; (f) Paulsen H (1984) Chem Soc Rev 13:15; (g) Schmidt RR (1986) Angew Chem Int Ed Engl 25:212; (h) Krohn K (1987) Nachr Chem Tech Lab 35:930; (i) Kunz H (1987) Angew Chem Int Ed Engl 26:294; (j) Schmidt RR (1989) Pure Appl Chem 61:1257; (k) Hashimoto S, Ikegami S (1991) Farmacia 27:50; (l) Schmidt RR (1991) In: Trost BM (ed) Comprehensive Organic Synthesis. Pergamon Press, Oxford, vol 6, p 33; (m) Sinaÿ P (1991) Pure Appl Chem 63:519; (n) Suzuki K, Nagasawa T (1992) J Synth Org Chem Jpn 50:378; (o) Ito Y, Ogawa T (1992) In: Iguchi Y (ed) Jikkenkagakukoza. Maruzen, Tokyo, vol 26, p 267; (p) Banoub J (1992) Chem Rev 92:1167; (q) Toshima K, Tatsuta K (1993) Chem Rev 93:1503; (r) Boons G-J (1996) Tetrahedron 52:1095; (s) Davis BG (2000) J Chem Soc Perkin Trans 1 2137; (t) Pellissier H (2005) Tetrahedron 61:2947; (u) Fügedi P (2006) In: Levy DE, Fügedi P (eds) The Organic Chemistry of Sugars. CRC Press, Boca Raton, Fl, p 89

5. Bock K, Meldal M (1983) Acta Chem Scand Ser B B37:775
6. Lemieux RU, Hendriks KB, Stick RV, James K (1975) J Am Chem Soc 97:4056
7. (a) West AC, Schuerch C (1973) J Am Chem Soc 95:1333; (b) Kronzer FJ, Schuerch C (1974) Carbohydr Res 33:273; (c) Eby R, Schuerch C (1975) Carbohydr Res 39:33
8. Ogawa T, Matsui M (1976) Carbohydr Res 51:C13
9. (a) Lubineau A, Malleron A (1985) Tetrahedron Lett 26:1713; (b) Lubineau A, Le Gallic J, Malleron A (1987) Tetrahedron Lett 28:5041
10. Higashi K, Nakayama K, Soga T, Shioya E, Uoto K, Kusama T (1990) Chem Pharm Bull 38:3280
11. Waldmann H, Böehm G, Schmid U, Röettele H (1994) Angew Chem Int Ed Engl 33:1994
12. Kartha KPR, Aloui M, Field RA (1996) Tetrahedron Lett 37:8807
13. Kartha KPR, Aloui M, Field RA (1997) Tetrahedron Lett 87:8233
14. Stachulski AV (2001) Tetrahedron Lett 42:6611
15. Banik BK, Samajdar S, Banik I, Zegrocka O, Becker FF (2001) Heterocycles 55:227
16. Mukherjee D, Kumar RP, Sankar Chowdhury U (2001) Tetrahedron 57:7701
17. Yamago S, Yamada T, Nishimura R, Ito H, Mino Y, Yoshida J-I (2002) Chem Lett 152
18. Mukaiyama T, Kobashi Y (2004) Chem Lett 33:10
19. Kobashi Y, Mukaiyama T (2005) Bull Chem Soc Jpn 78:910
20. Nishizawa M, Garcia DM, Yamada H (1992) Synlett 797
21. Nishizawa M, Garcia DM, Shin T, Yamada H (1993) Chem Pharm Bull 41:784
22. Yamada H, Hayashi T (2002) Carbohydr Res 337:581
23. (a) Kleine HP, Weinberg DV, Kaufman RJ, Sidhu RS (1985) Carbohydr Res 142:333; (b) Dess D, Kleine HP, Weinberg DV, Kaufman RJ, Sidhu RS (1981) Synthesis 883
24. Brewster K, Harrison JM, Inch TD (1979) Tetrahedron Lett 20:5051
25. Loganathan D, Trivedi GK (1987) Carbohydr Res 162:117
26. Wang Y, Li L, Wang Q, Li Y (2001) Synth Commun 31:3423
27. Roy R, Tropper F (1990) Synth Commun 20:2097
28. Roy R, Tropper FD, Grand-Maître C (1991) Can J Chem 69:1462
29. Bliard C, Massiot G, Nazabadioko S (1994) Tetrahedron Lett 35:6107
30. Hongu M, Saito K, Tsujihara K (1999) Synth Commun 29:2775
31. Sasaki M, Gama Y, Yasumoto M, Ishigami Y (1990) Tetrahedron Lett 31:6549
32. (a) Nishizawa M, Kan Y, Yamada H (1988) Tetrahedron Lett 29:4597; (b) Nishizawa M, Kan Y, Yamada H (1989) Chem Pharm Bull 37:565; (c) Nishizawa M, Kan Y, Shimomoto W, Yamada H (1990) Tetrahedron Lett 31:2431; (d) Nishizawa M, Imagawa H, Kan Y, Yamada H (1991) Tetrahedron Lett 32:5551; (e) Nishizawa M, Shimomoto W, Momii F, Yamada H (1992) Tetrahedron Lett 33:1907
33. Schmid U, Waldmann H (1997) Liebigs Ann Chem 2573
34. (a) Hadd MJ, Gervay J (1999) Carbohydr Res 320:61; (b) Lam SN, Gervay-Hague J (2002) Org Lett 4:2039; (c) Lam SN, Gervay-Hague J (2002) Carbohydr Res 337:1953; (d) Dabideen DR, Gervay-Hague J (2004) Org Lett 6:973; (e) Lam SN, Gervay-Hague J (2005) J Org Chem 70:2387; Du W, Gervay-Hague J (2005) Org Lett 7:2063
35. (a) Perrie JA, Harding JR, King C, Sinnott D, Stachulski AV (2003) Org Lett 5:4545; (b) Harding JR, King CD, Perrie JA, Sinnott D, Stachulski AV (2005) Org Biomol Chem 3:1501
36. Mukaiyama T, Murai Y, Shoda S (1981) Chem Lett 431
37. Mukaiyama T, Hashimoto Y, Shoda S (1983) Chem Lett 935
38. Rosenbrook JrW, Riley DA, Lartey, PA (1985) Tetrahedron Lett 26:3
39. Posner GH, Haines SR (1985) Tetrahedron Lett 26:5
40. Hayashi M, Hashimoto S, Noyori R (1984) Chem Lett 1747
41. Szarek WA, Grynkiewicz G, Doboszewski B, Hay GW (1984) Chem Lett 1751
42. Miethchen R, Hager C, Hein M (1997) Synthesis 159
43. Ernst B, Winkler T (1989) Tetrahedron Lett 30:3081
44. Kunz H, Sanger W (1985) Helv Chim Acta 68:283
45. Teichmann M, Descotes G, Lafont D (1993) Synthesis 889
46. Lichtenthaler FW, Kläres U, Lergenmüller M, Schwidetzky S (1992) Synthesis 179

47. Goggin KD, Lambert JF, Walinsky SW (1994) Synlett 162
48. Nicolaou KC, Dolle RE, Papahatjis DP, Randall JL (1984) J Am Chem Soc 106:4189
49. Caddick S, Motherwell WB, Wilkinson JA (1991) J Chem Soc Chem Commun 674
50. Caddick S, Gazzard L, Montherwell WB, Wilkinson JA (1996) Tetrahedron 52:149
51. MacDonald SJF, McKenzie TC (1988) Tetrahedron Lett 29:1363
52. Gordon DM, Danishefsky SJ (1990) Carbohydr Res 206:361
53. Ferrier RJ, Prasad N (1969) J Chem Soc C:570
54. Bröder W, Kunz H (1993) Carbohydr Res 249:221
55. Palme M, Vasella A (1995) Helv Chim Acta 78:959
56. Ogawa T, Takahashi Y (1985) Carbohydr Res 138:C5; Takahashi Y, Ogawa T (1987) Carbohydr Res 164:277
57. Hashimoto S, Hayashi M, Noyori R (1984) Tetrahedron Lett 25:1379
58. Nicolaou KC, Chucholowski A, Dolle RE, Randall JL (1984) J Chem Soc Chem Commun 1155
59. Kunz H, Sager W (1985) Helv Chim Acta 68:283
60. Kunz H, Waldmann H (1985) J Chem Soc Chem Commun 638
61. Vozny YA, Galoyan AA, Chizhov OS (1985) Bioorg Khim 11:276
62. (a) Kreuzer M, Thiem J (1986) Carbohydr Res 149:347; (b) Jünnemann J, Lundt I, Thiem J (1991) Liebigs Ann Chem 759
63. (a) Matsumoto T, Maeta H, Suzuki K, Tsuchihashi G (1988) Tetrahedron Lett 29:3567; (b) Suzuki K, Maeta H, Matsumoto T, Tsuchihashi G (1988) Tetrahedron Lett 29:3571; (c) Matsumoto T, Maeta H, Suzuki K, Tsuchihashi G (1988) Tetrahedron Lett 29:3575; (d) Matsumoto T, Katsuki M, Suzuki K (1989) Chem Lett 437
64. Suzuki K, Maeta H, Suzuki T, Matsumoto T (1989) Tetrahedron Lett 30:6879
65. (a) Nicolaou KC, Caulfield TJ, Kataoka H, Stylianides NA (1990) J Am Chem Soc 112:3693; (b) Nicolaou KC, Hummel CW, Iwabuchi Y (1992) J Am Chem Soc 114: 3126
66. Maeta H, Matsumot T, Suzuki K (1993) Carbohydr Res 249:49
67. Kobayashi S, Koide K, Ohno M (1990) Tetrahedron Lett 31:2435
68. (a) Wessel HP (1990) Tetrahedron Lett 31: 6863; (b) Wessel HP, Ruiz N (1991) J Carbohydr Chem 10:901
69. Böhm G, Waldmann H (1995) Tetrahedron Lett 36:3843
70. Hosono S, Kim W-S, Sasai H, Shibasaki M (1995) J Org Chem 60:4
71. Kim W-S, Hosono S, Sasai H, Shibasaki M (1995) Tetrahedron Lett 36:4443
72. Kim W-S, Sasai H, Shibasaki M (1996) Tetrahedron Lett 37:7797
73. Yu L, Chen D, Li J, Wang PG (1997) J Org Chem 62:3575
74. (a) Toshima K, Kasumi K, Matsumura S (1998) Synlett 643; (b) Toshima K, Kasumi K, Matsumura S (1999) Synlett 1332
75. (a) Takeuchi K, Mukaiyama T (1998) Chem Lett 555; (b) Yanagisawa M, Mukaiyama T (2001) Chem Lett 224
76. Jona H, Maeshima H, Mukaiyama T (2001) Chem Lett 726–727
77. (a) Mukaiyama T, Jona H, Takeuchi K (2000) Chem Lett 696; (b) Jona H, Takeuchi K, Mukaiyama T (2000) Chem Lett 1278
78. (a) Jona H, Mandai H, Mukaiyama T (2001) Chem Lett 426; (b) Jona H, Mandai H, Chavasiri W, Takeuchi K, Mukaiyama T (2002) Bull Chem Soc Jpn 291
79. Yamanoi T, Nagayama S, Ishida H-K, Nishikido J, Inazu T (2001) Synth Commun 31:899
80. Pikul S, Switzer AG (1997) Tetrahedron: Asymmetr 8:1165
81. Sasaki K, Matsumura S, Toshima K (2004) Tetrahedron Lett 45:7043

3.2 Glycosyl Trichloroacetimidates

*Richard R. Schmidt**[1], *Xiangming Zhu*[2]
[1] Fachbereich Chemie, Fach M 725, Universität Konstanz, 78457 Konstanz, Germany
[2] School of Chemistry and Chemical Biology, University College Dublin, Belfield, Dublin 4, Ireland
richard.schmidt@uni-konstanz.de, xiangming@ucd.ie

Abstract

In the first part of this review, the basic principles of chemical glycosylation reactions are discussed; this way the advantages of *O*-glycosyl trichloroacetimidates and related systems as glycosyl donors become obvious. Many new methods for the generation of *O*-glycosyl trichloroacetimidates and for their use as glycosyl donors have been introduced which are

In: *Glycoscience*. Fraser-Reid B, Tatsuta K, Thiem J (eds)
Chapter-DOI 10-1007/978-3-540-30429-6_11:

compiled as well as their use in solid-phase oligosaccharide synthesis. The power of these glycosyl donors is demonstrated by their application in complex oligosaccharide and glycoconjugate synthesis, as outlined in the second part of this review. Recent applications in glycolipid, glycosyl amino acid and glycopeptide, nucleoside and nucleotide glycosidation, glycosamino glycan, cell wall constituent, and GPI anchor synthesis, glycosylation of various natural products and their metabolites, and finally cyclooligosaccharide generation are compiled. In the last section, related glycosyl donors are briefly discussed.

Abbreviations

ADMB 4-acetoxy-2,2-dimethylbutanoyl
CAc chloroacetyl
CDs cyclodextrins
CSs chondroitin sulfates
DBU 1,8-diazabicyclo[5.4.0]undec-7-ene
DDQ 2,3-dichloro-5,6-dicyno-benzoquinone
DMM *N*-dimethylmaleoyl
DPM diphenylmethyl
FGFs fibroblast growth factors
Fmoc 9-fluorenylmethyl carbamate
GAGs glycosaminoglycans
GPI glycosylphosphatidyl inositol
HS heparan sulfate
ILs ionic liquids
IP inverse procedure
IPG inositolphosphoglycan
LTAs lipoteichoic acids
MPEG mono-methyl polyethyleneglycol
MPLC medium pressure liquid chromatography
NAP naphthylmethyl
PLL photolabile linker
PNBP *p*-nitrobenzylpyridine
PPTS pyridinium *p*-toluenesulfonate
PTBD polymer-supported 1,5,7-triazabicyclo[4.4.0]dec-5-ene
PS-DBU polystyrene-supported DBU
SPOS solid-phase oligosaccharide syntheses
TBS *tert*-butyl-dimethylsilyl
TMSOTf trimethylsilyl trifluoromethanesulfonate

1 Introduction

The biological significance of oligosaccharides and glycoconjugates has stimulated many activities in carbohydrate chemistry. Most of these activities have been devoted to the development of methods for glycoside bond formation as the assembly of monosaccharide building

blocks to complex oligosaccharides and glycoconjugates is the most difficult task in this endeavor.

2 Glycosyl Donor Generation Through Anomeric-Oxygen Exchange Reactions

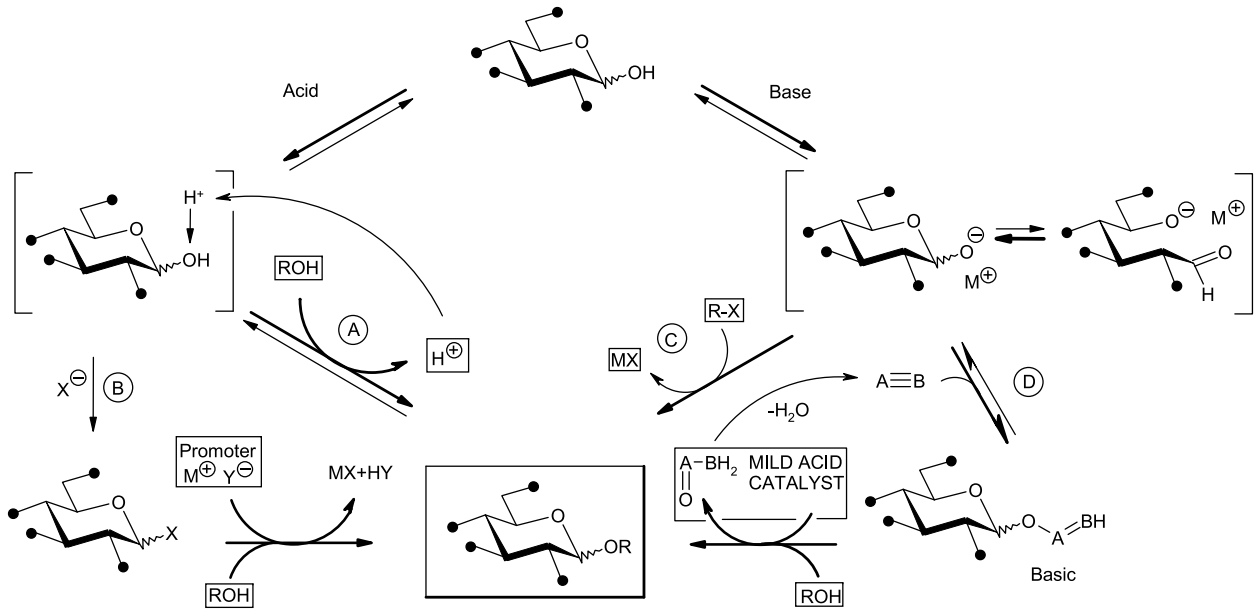
For a long time the methods developed for glycosylation have essentially favored approaches which require for glycosyl donor generation an *anomeric oxygen exchange reaction* on the half-acetal moiety of pyranoses and furanoses [Fig. 1, methods (A) and (B)] [1,2,3,4]. The *Fischer–Helferich method* (A), as a direct acid-catalyzed anomeric-oxygen replacement reaction has been very successfully applied to the synthesis of simple alkyl glycosides. However, because of its reversibility, it is of limited usefulness in the synthesis of complex oligosaccharides and glycoconjugates. For this endeavor, irreversible methods are required, which can be attained by preactivation of the anomeric center by introducing good leaving groups which can be released by leaving group specific promoters or—even better—by catalytic amounts of an activator.

The best-known of these methods is the *Koenigs–Knorr method* (B) in which the anomeric hydroxy group is replaced by chlorine or bromine. Thus, an α-halogen-ether is generated which in the glycosylation step can be readily activated by halophilic promoters frequently incorrectly termed "catalyst." Generally, from one up to four equivalents of heavy metal salts are employed resulting in an irreversible transfer of the glycosyl donor moiety to the acceptor. On the basis of this general method many valuable techniques for the synthesis of complex oligosaccharides and glycoconjugates have been introduced which have been extensively reviewed. However, the obvious limitations of this method were the reason for the search for alternative methods [1,2,3,4,5,6,7].

Therefore, other anomeric oxygen exchange reactions closely related to the *Koenigs–Knorr method* have been extensively investigated. Particularly, the introduction of fluorine and alkyl- and arylthio groups as leaving groups gained great interest because these groups also tolerate manipulations of orthogonal protecting groups. Differences in thio leaving group tendencies (in combination with the "armed"/"disarmed" principle) could be even employed for one-pot consecutive glycosylations leading to more or less pure oligosaccharide products. However, the basic drawbacks of the *Koenigs–Knorr method* are also associated with these promoter systems. For instance, the large amounts of promoter required (and often additional reagents) limit their usefulness particularly in large-scale oligosaccharide and glycoconjugate synthesis.

3 Direct Anomeric Oxygen Alkylation

For more than a century glycosylations were essentially based on methods where the anomeric carbon of the sugar residue to be coupled served as the electrophile (the glycosyl donor) and the alcohol (the glycosyl acceptor) as the nucleophile [Fig. 1, (A), (B)]. Alternatively, base-mediated deprotonation of the anomeric hydroxy group generating at first an anomeric



A **Fischer-Helferich**: (Acid catalyzed Act.)
B **Koenigs-Knorr**: X = Cl, Br, (I) Activation
X = F-Activation
X = S-R-Activation (+ Hetarylthio)

Glycosylation: "X-Philic" Promoter
(Generally: Heavy metal salts: Ag^{+}, Hg^{2+}, etc.)

C **Anomeric O-Alkylation**: (Base Activation)
D **Trichloroacetimidate Activation**: A≡B = CCl_3–CN
Related Systems (?):
$PO(OR)_2$-, $P(OR)_2$-Activation
$SO_2(OR)$-, $SO(OR)$-, SO_2R-Activation

Glycosylation: Mild Acid catalyst
(Generally: $BF_3 \cdot OEt_2$, TMSOTf, etc.)

Figure 1
Synthesis of glycosides and saccharides

oxide structure from a pyranose or a furanose and then anomeric *O*-alkylation leading directly and irreversibly to glycosides should be also available [◗ *Fig. 1*, (C)]. Surprisingly, no studies employing this simple '*anomeric O-alkylation method*' as termed by us [1,2,3,4], for the synthesis of complex glycosides and glycoconjugates had been reported before our work. Only a few scattered examples with simple alkylating agents, for instance, excess methyl iodide or dimethyl sulfate, have been found in the literature [1]. However, in our hands, direct *anomeric O-alkylation* of variously protected and totally unprotected sugars in the presence of a base and triflates or Michael acceptors, respectively, as alkylating agents has become a very convenient method for glycoside bond formation [8,9,10,11]. Base-promoted decomposition reactions, particularly of the acyclic form, were practically not observed. Often even high anomeric diastereocontrol was available. The high diastereocontrol in pyranoses is based on the enhanced nucleophilicity of equatorial oxygen atoms (due to steric effects and the stereoelectronic *kinetic anomeric effect*) [1,2,3,4] and on the higher stability of axial oxygen atom-derived products (due to the *thermodynamic anomeric effect*). Chelation effects can be also employed to design anomeric stereocontrol. The availability and to some extent the stability of the carbohydrate-derived alkylating agents precluded the general applicability of this simple method to the synthesis of complex oligosaccharides and glycoconjugates.

4 Glycosyl Donor Generation Through Retention of the Anomeric Oxygen

The requirements for an efficient glycosylation method are the following:

- high chemical and stereochemical yield,
- applicability to large-scale preparations, and
- avoidance of large amounts of waste materials, i. e. activation of the donor by catalytic amounts of reagent.

These demands were not met by any of the above-described methods for the synthesis of complex oligosaccharides and glycoconjugates. However, the general strategy for glycoside bond formation seems to be correct:

- The first step (*activation step*) should consist of an activation of the anomeric center under formation of a stable glycosyl donor—best by a catalyzed attachment of a leaving group to the anomeric hydroxy group.
- The second step (*glycosylation step*) should consist of a sterically uniform high-yielding glycosyl transfer to the glycosyl acceptor based on glycosyl donor activation with catalytic amounts of promoter, i. e. a catalyst. Obviously, this catalytic procedure has to be orthogonal to the glycosyl donor preparation procedure. Diastereocontrol in the glycosylation step may be derived from the anomeric configuration of the glycosyl donor (by inversion or retention), by anchimeric assistance, by the solvent influence, by thermodynamic and/or stereoelectronic effects, or by any other effects.

The experience with the direct anomeric *O*-alkylation exhibited that these demands can be fulfilled by a simple base-catalyzed anomeric *O*-transformation into a leaving group and

its acid-catalyzed activation in the glycosylation step. This approach should also satisfy the demand for simplicity in combination with efficiency which is decisive for general acceptance.

Obviously, for achieving stereocontrolled activation of the anomeric oxygen atom, the anomerization of the anomeric hydroxy group or the anomeric oxide ion, respectively, has to be considered. Thus, in a reversible activation process and with the help of kinetic and thermodynamic reaction control, possibly both activated anomers should be accessible. From these considerations it was concluded that suitable triple-bond systems A≡B (or compounds containing cumulative double bond systems A=B=C) might be found that add pyranoses and furanoses under base catalysis directly and reversibly in a stereocontrolled manner [◉ *Fig. 1*, (D)].

Electron-deficient nitriles, such as for instance trichloroacetonitrile (and trifluoroacetonitrile [11]), are known to undergo direct and reversible, base-catalyzed addition of alcohols providing *O*-alkyl trichloroacetimidates (or *O*-alkyl trifluoroacetimidate, respectively) [1,12,13] [◉ *Fig. 2*, (1)]. This imidate synthesis has the advantage that the imidates can be isolated as stable adducts which are less sensitive to hydrolysis than the corresponding salts. On acid addition leading to imidate activation [◉ *Fig. 2*, (2)], hydrolysis with water (R^2–OH = H_2O) is a fast reaction furnishing amide CCl_3–$CONH_2$ and alcohol R^1–OH [◉ *Fig. 2*, (3A)]; mechanistically this is an acylative attack of the imidate at the nucleophile. With other nucleophiles, after acylation by the imidate, other transformations are possible, as for instance ortho-ester formation. However, the basic question is: are these activated imidates also good alkylating agents [◉ *Fig. 2*, (3B)] as required for being effective in glycosidation reactions? On consider-

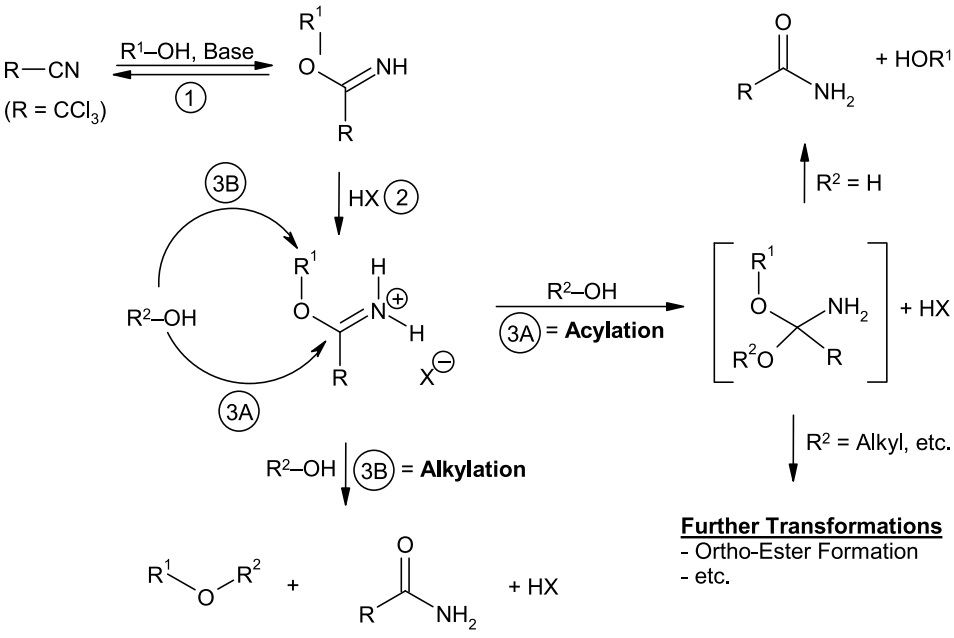


Figure 2
Trichloroacetimidate formation and its acid-catalyzed transformations

ation of the influence of the substituents R and R^1 on these two competing reactions an attack of R^2–OH leads to the following expectations:

- Acylation [reaction (3A)] is supported by R being a small electron-withdrawing group, R^1 destabilizing carbenium ion formation.
- Alkylation [reaction (3B)] is supported by R being a sterically demanding electron-withdrawing group, R^1supporting carbenium ion formation.

From these considerations it can be deduced that the bulky and strongly electron-withdrawing trichloromethyl group as R and the glycosyl group as R^1, which through the α-oxygen atom supports oxocarbenium ion formation at the anomeric center, should provide excellent alkylating agents; hence, *O*-glycosyl trichloroacetimidates should on acid activation exhibit excellent glycosyl donor properties.

As expected, on base-catalyzed addition of the anomeric hydroxy group to trichloroacetonitrile the *O*-glycosyl trichloroacetimidate is formed for which, due to the reversibility of the addition, the different nucleophilicities of the anomeric oxides, and the different thermodynamic stabilities of the *O*-glycosyl trichloroacetimidates, anomeric stereocontrol is possible. Thus, the weak base potassium carbonate could be employed for preferential or exclusive formation of the β-anomer and the strong base sodium hydride could be employed for the thermodynamically more stable α-anomer (❯ *Fig. 3*). This stereocontrol could be successfully extended to S_N2-type glycosidation reactions in solvents of low donicity and under low temperature conditions and particularly to glycosylations of O=X–OH nucleophiles with X being RC or R_2P (see below). However, for many cases anomeric stereocontrol is based on other effects such as for instance neighboring group participation and/or steric effects, stereoelectronic effects, solvent participation effects, etc.

The experience with the *trichloroacetimidate method* exhibited that the demands on a new glycosylation methodology are fulfilled:

(i) The *O*-glycosyl trichloroacetimidates are readily formed and generally stable under room temperature conditions. However, on acid catalysis they exhibit extraordinary high glycosyl donor properties.

(ii) The release of nonbasic trichloroacetamide fulfills the criteria for acid catalysis: The acid is not consumed by the leaving group, therefore generally only catalytic amounts of (Lewis) acid are required (~0.001 to 0.1 equivalents).

(iii) The released trichloroacetamide is also not acidic, therefore the acidity provided by the catalyst amount is maintained throughout the reaction course. Hence, negative effects of increasing acidity in the reaction mixture (found for instance for phosphate, sulfate, sulfonate leaving groups, ❯ *Fig. 1*, D) are avoided.

(iv) Glycosidation is basically a condensation reaction. In this procedure, water is bound to trichloroacetonitrile under trichloroacetamide formation. Hence, drying agents are not required. Often molecular sieves are used in glycosidation reactions. Because they may affect the acidity of the reaction in an unpredictable fashion, their use is not even recommended in this method.

(v) Trichloroacetamide can be removed from the reaction mixture and transformed back to trichloroacetonitrile, thus exhibiting the cost-effectiveness and ecofriendliness of this method in large-scale preparations.

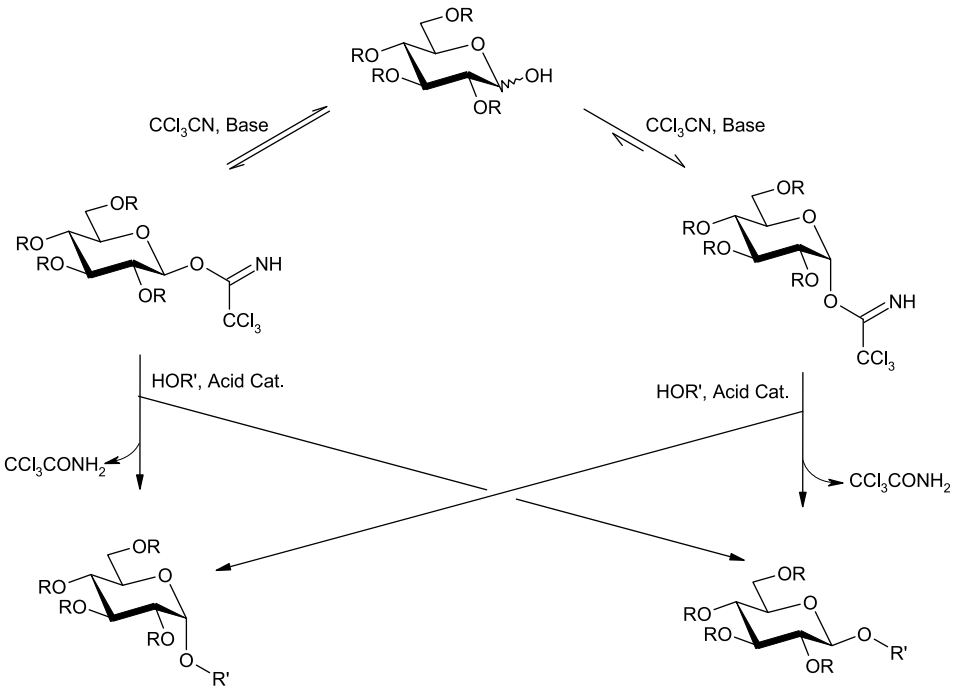


Figure 3
***O*-Glycosyl trichloroacetimidate formation and acid-catalyzed glycosylation of acceptors HOR′**

(vi) Neither in the formation of the *O*-glycosyl trichloroacetimidates nor in the glycosidation reactions are equivalent or even higher amounts of salts produced—a disadvantage of most of the above-mentioned methods, particularly in large-scale preparations. Also, highly expensive sterically hindered bases are not required.

This discussion exhibits that *O*-glycosyl *N*-methyl-acetimidates, introduced by the Sinaÿ group [14,15], are poor glycosyl donors: they lack the strongly electron-withdrawing sterically demanding trichloromethyl group. Hence, they are rather better acylating agents than alkylating agents. In addition, their formation via *O*-glycosylation of *N*-methyl-acetamide with excess amounts of silver oxide is quite cumbersome.

These favorable aspects of *O*-glycosyl trichloroacetimidates led to their frequent use as glycosyl donors for various types of glycosyl acceptors such as nucleophiles (*Fig. 4*). Besides hydroxy groups (of standard alcohols, phenols, sugars, etc.), also carboxylic acids, phosphorous acids, and sulfonic and sulfuric acids, respectively, were successfully employed as acceptors; due to the acidity of these acidic acceptors, generally catalysts are not required for the activation of the glycosyl donors and—presumably via an eight-membered transition state—often the inversion product at the anomeric center is generated. However, also various *C*-, *N*-, *S*-, and *P*-nucleophiles have been successfully glycosylated by *O*-glycosyl trichloroacetimidates [1,2].

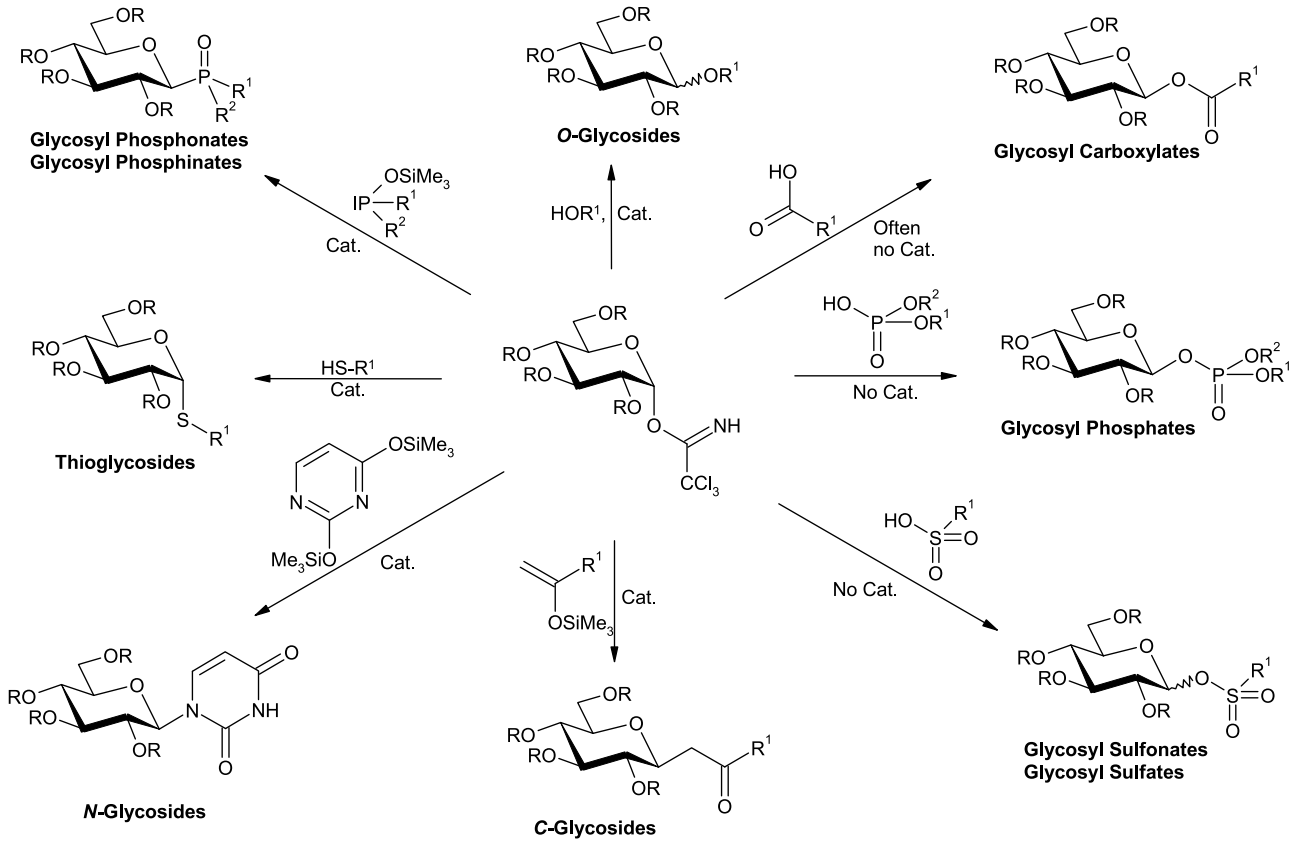


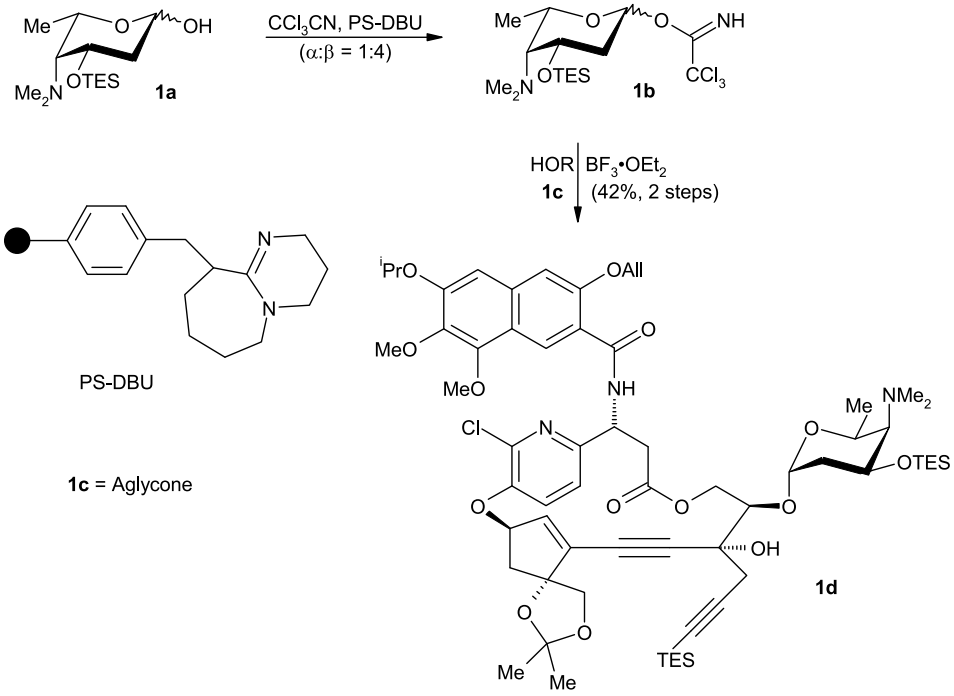
Figure 4
Reaction of *O*-glycosyl trichloroacetimidates with *O*-, *C*-, *N*-, *S*-, and *P*-nucleophiles—activation by (Lewis)acid catalysis

4.1 Methodological Aspects

As outlined above, the anomeric hydroxy groups add under base catalysis readily to the electron-deficient nitrile group of trichloroacetonitrile to furnish generally via equilibration at the anomeric center *O*-glycosyl trichloroacetimidates. As bases mainly potassium carbonate, sodium hydride, DBU, and cesium carbonate are employed. Weak bases, such as for instance potassium carbonate, permit kinetic product control (formation of the equatorial product) whereas strong bases, as for instance sodium hydride, lead to the thermodynamically more stable product (generally the axially oriented trichloroacetimidate group). Because of convenience, DBU has become the most popular base, which also favors axial product formation. Therefore, polystyrene-supported DBU (PS-DBU) [16,17] and various other solid-supported nitrogen bases were successfully employed in this reaction as shown for the particularly difficult trichloroacetimidate formation of 4-dimethylamino-2,4,6-trideoxy sugar kedarosamine (**1a** → **1b**) and its use in the glycosylation of the ansamacrolide substructure (**1b** + **1c** → **1d**) of the kedarcidin chromophore (*Scheme 1*) [17]. Also Dowex 1-X8 OH^- is a good catalyst for the trichloroacetimidate formation of a glucosamine derivative [18,19]. Polymer-supported 1,5,7-triazabicyclo[4.4.0]dec-5-ene (PTBD) turned out to be a powerful catalyst for trichloroacetimidate formation as well [17,19]; this base catalyst in combination with Nafion®-SAC resin (a nanocomposite of Nafion® resin with porous silica) could be successfully employed in one-pot trichloroacetimidate formation and following glycosylation reactions [19].

O-Glycosyl trichloroacetimidates are quite stable under neutral but also under basic conditions. Hence, trichloroacetimidate formation is tolerated by standard *O*- and *N*-protecting groups, such as *O*-acyl, *O*-benzyl, *O*,*O*-alkylidene, *O*-silyl, *N*-acyl and *N*,*N*-diacyl, *N*-phthaloyl (Phth), *N*-dimethylmaleoyl (DMM), and the latent amino functionality azido. It is particularly worth mentioning that trichloroacetimidate formation is also tolerated by Fmoc-protected hydroxy groups; this group has become an important temporary protecting group which is orthogonal to other temporary protecting groups, thus permitting in oligosaccharide synthesis not only regioselective chain extension but also branching [20,21]. The useful 3-*O*-Fmoc-protected galactosyl donor **2d** can be readily prepared from compound **2a** via intermediates **2b** and **2c** (*Scheme 2*) [20]. Additional protecting groups have been recently probed particularly at 2-*O* in order to influence the stereochemical outcome, such as for instance the diphenylmethyl (DPM), the 9-fluorenyl (Fl) group [22], and the 4-acetoxy-2,2-dimethylbutanoyl (ADMB) group [23]. The ADMB group is a useful alternative to the pivaloyl group; it combines the strong 1,2-*trans*-selection in glycosylation reactions of the pivaloyl group with the ease of removal of more reactive acyl groups. Other important new groups are the *O*-trifluoroethylsulfonate group which could be successfully employed to release the sulfate group and the replacement of the 6-hydroxy group by a phenylthio group which is employed for uronate formation [25].

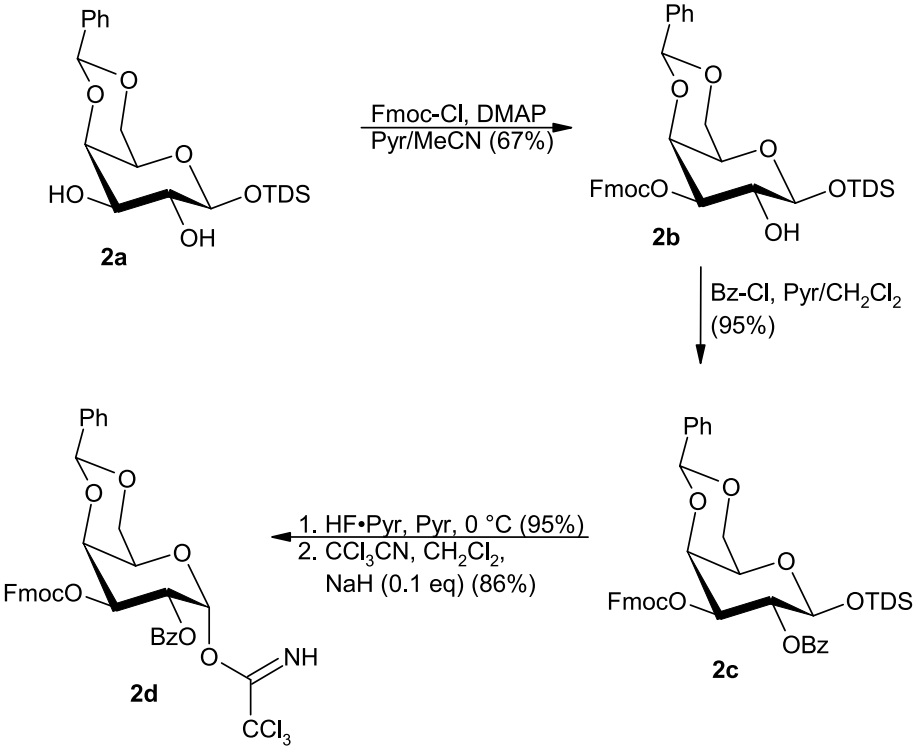
O-Glycosyl trichloroacetimidates are under acidic conditions very powerful glycosyl donors. Reaction with nonacidic nucleophiles is generally performed with catalytic amounts of Brønsted or Lewis acids. TMSOTf and BF_3OEt_2 are the most frequently employed catalysts. TMSOTf is the catalyst of first choice and it is generally used in 0.001 to 0.1 equivalents based on the glycosyl donor. Commonly 1.0 to 1.5 equivalents of glycosyl acceptor are added. Dichloromethane is the solvent of first choice. Because of the high reactivity of *O*-glycosyl



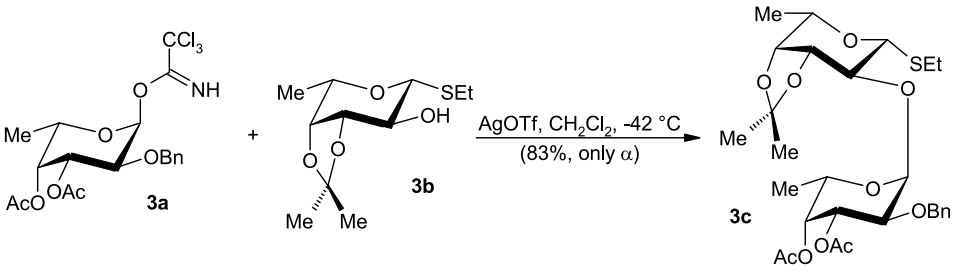
Scheme 1

trichloroacetimidates often reaction temperatures between 0 and −40 °C are selected. For highly sensitive glycosyl donors and for the anomeric stereocontrol many variations have been probed which are shortly summarized.

Numerous acid catalysts have been investigated. Brønsted acids, such as for instance p-TsOH and TfOH, can lead—as previously discussed [1,2,3,4,26]—to *O*-glycosyl sulfonates as reactive intermediates which are in most cases not the glycosyl donors but rather some type of glycosyl oxocarbenium ion intermediates (S_N1-type reactions). Similarly, pyridinium p-toluenesulfonate (PPTS) [27,28] and perchloric acid [29,30] are working as catalysts. Metal triflates such as AgOTf [31,32], $Cu(OTf)_2$ [33], $Sn(OTf)_2$ [34,35], $Sm(OTf)_3$ [36], and $Yb(OTf)_3$ [37], are particularly valuable in glycosylations with acid sensitive glycosyl donors and/or acceptors. On $Cu(OTf)_2$-activation of various donors the trichloroacetimidates have proven to be the most potent glycosyl donors [33]. For instance, AgOTf has proven to be useful, as a catalyst for highly reactive deoxysugars such as glycosyl donors, as shown in *Scheme 3* for the reaction of fucosyl donor **3a** with fucosyl acceptor **3b** giving the α-oligosaccharide **3c** [32]. $Sn(OTf)_2$ permitted successful glycosylations of acid sensitive glycal derivatives as acceptors as shown for the synthesis of **4c** from **4a** and **4b** (*Scheme 4*) [35]. Various other catalyst systems have proven to be useful such as for instance $HB(C_6F_5)_4$ [29,30], *N*-acyl-sulfonamides, phenols [38], I_2 [39], I_2/Et_3SiH [40], and electrophilic carbonyl compounds [41], such as for instance chloral. Silica-supported perchloric acid [42,43], acid-washed molecular sieves (MS



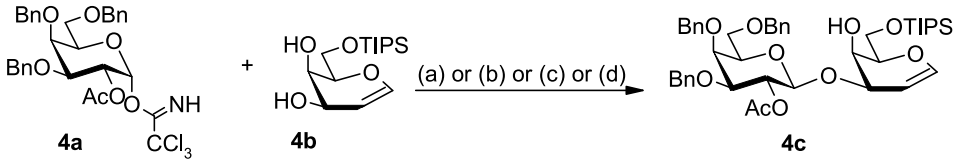
Scheme 2



Scheme 3

AW 300) [44], and Nafion®-SAC resin [19] (see above) are highly valuable solid-supported acid catalysts. This is nicely demonstrated in an avermectin B_{1a} analog synthesis [43]: Rhamnosylation of intermediate **5b** with donor **5a** gave target molecule **5c** in quantitative yield (*Scheme 5*).

The solvent of choice in the glycosylation reactions is dichloromethane. In combination with a solvent of low polarity and/or of low donicity (such as for instance cyclohexane, petroleum ether, etc.), BF_3OEt_2 as a mild catalyst and at low temperature, S_N2-type reactions could be



(a) 0.01 eq TMSOTf, -78°C	–	(**4c**)
(b) 0.01 eq BF_3•OEt_2, -78 °C	–	(**4c**)
(c) 0.5 eq $ZnCl_2$, 20 °C	44%	(**4c**)
(d) 0.02 eq $Sn(OTf)_2$, -20 °C	75%	(**4c**)

Scheme 4

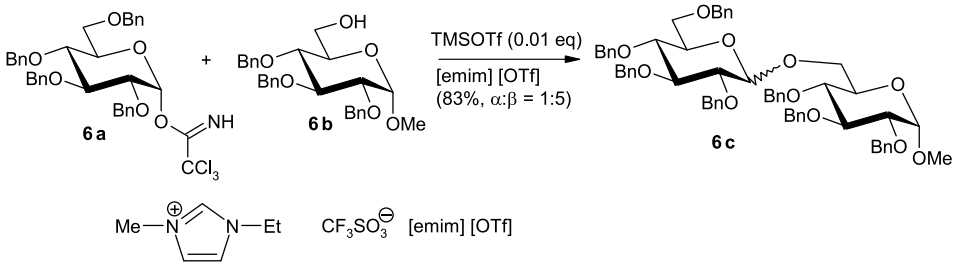
carried out [1,2,3,4]. Solvents of high donicity (such as for instance ethers and nitriles) permit—based on the exo-anomeric effect—a different anomeric stereocontrol: As found and previously explained, in ethers at room temperature generally the axial product (for most cases the α-product) is favored whereas in nitriles at low temperature, based on the nitrile effect, the equatorial product (for most cases the β-product) is preferentially or exclusively obtained [3,4,45,46]. Hence, solvent and temperature selection play also an important role in anomeric stereocontrol.

Another interesting solvent effect is associated with the application of the 'inverse procedure' (IP) which consists of the addition of the glycosyl donor to a mixture of acceptor and acid catalyst [47,48]. In this way, particularly with highly reactive glycosyl donors often dramatic glycosylation yield improvements could be obtained [49]. This result seems to be based on a cluster effect between the catalyst and acceptor molecules; hence, on penetration of the donor into this cluster, activation of the donor and following glycoside bond formation takes place within the cluster in a practically intramolecular fashion. Obviously, this effect is critically dependent on the acceptor type, the solvent, and the temperature; therefore it requires often some experimentation for its successful application.

A dramatic polarity increase of the solvent can be reached by adding $LiClO_4$ to organic solvents (for instance, ether, dichloromethane, etc.) which enabled *O*-glycosyl trichloroacetimidate activation under essentially neutral conditions [50,51]. High polarity solvents are also ionic liquids (ILs) which have gained a lot of interest as solvents for organic reactions [52]. ILs have been probed in glycosylation reactions with *O*-glycosyl trichloroacetimidates and excellent glycosylation yields have been obtained (*Scheme 6*; **6a** + **6b** → **6c**) [53]; for reactive systems again no acid activation is required [53]. Microwave heating has been employed for *O*-glycosyl trichloroacetimidate activation as well, furnishing glycosides in high yields [54]; however, presumably due to the temperature effects only modest anomeric stereocontrol was available.

Besides solvent, temperature, and (kinetic and thermodynamic) stereoelectronic effects, anchimeric assistance by neighboring acyl groups and/or by steric shielding is decisive for the anomeric stereocontrol [1,2,3,4,5,6,7,22]. This way 1,2-*trans*- (β-*gluco*, α-*manno*) and 1,2-*cis*-type (α-*gluco*) glycosides have been generally obtained at wish in good to excellent stereochemical yields. However, a 2-*O*-acyl group does not guarantee β-glucopyranoside or α-mannopyranoside formation. Besides undesired ortho-ester formation, which can often be

Scheme 5



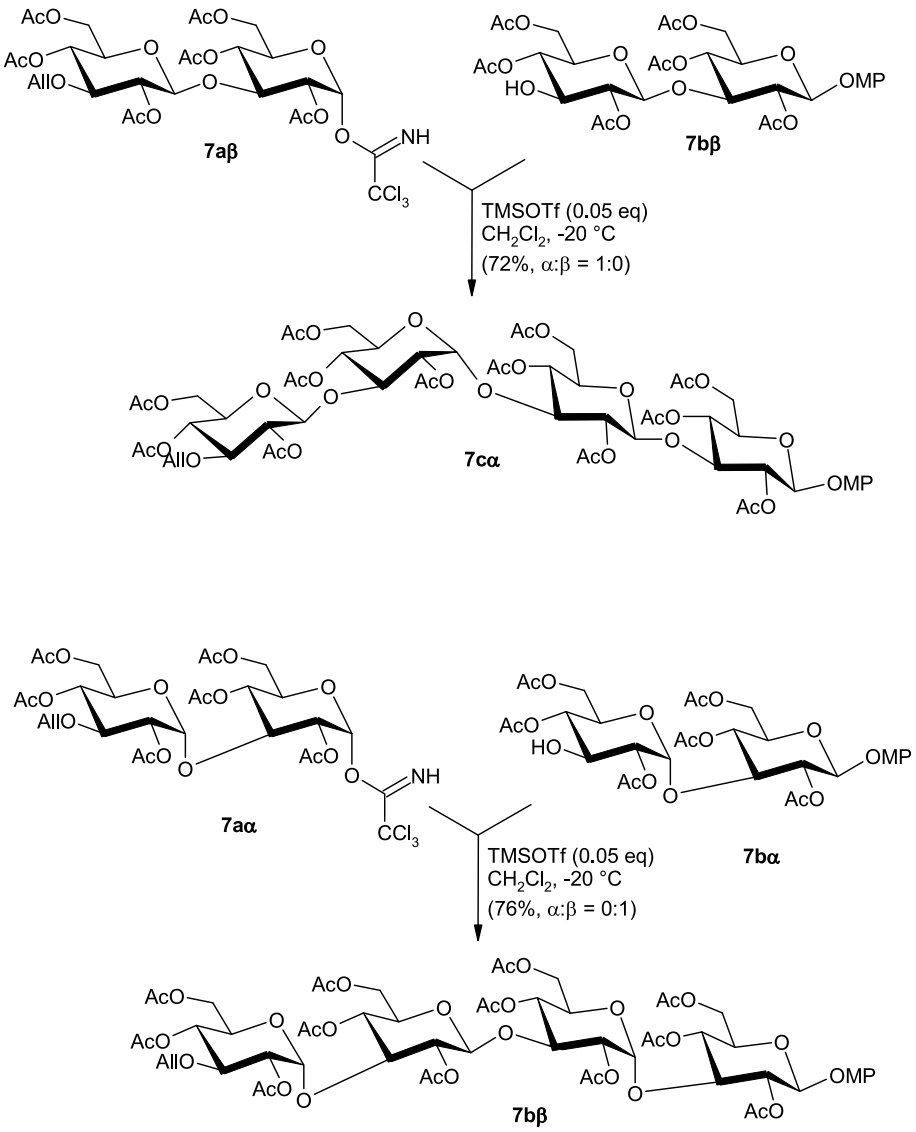
Scheme 6

overcome by using more catalyst or by varying the acyl group, once in a while still the α-glucopyranoside or even some β-mannopyranoside, respectively, is obtained. Interesting cases of α-linkage on attempted 2-*O*-acyl supported β(1–3)-glucan synthesis were recently reported which were explained by 'remote control' [55,56,57]. **7aβ + 7bβ** gave **7cα** whereas **7aα + 7bα** gave **7cβ** (*Scheme 7*). However, this problem could be readily overcome by employing the ADMB group [23]. The influence of the structure of the glycosyl acceptors on the anomeric stereocontrol has often been discussed and since the work of van Boeckel et al. [58,59] the potential importance of matched and mismatched donor-acceptor pairs on the stereochemical outcome of glycosylation reactions has become evident. Recent work on 2-azido-2-deoxy-glucopyranosyl trichloroacetimidates exhibited that the structure of the glycosyl donor has generally the major influence [60].

Other anchimerically assisting groups at C-2, such as–SPh, -SePh, -Br, and -I, have been investigated [61,62]. Particularly worth mentioning is the work on equatorial iodo substitution of glucopyranosyl trichloroacetimidates which strongly favored 1,2-*trans*- i. e. β-glycosidation. This result could be due to generation of an iodonium intermediate or due to the steric demand of the iodo group, thus favoring a twist-boat type oxocarbenium ion intermediate which—as previously discussed in β-mannopyranoside synthesis (see below) [26]—for steric and stereoelectronic reasons favors nucleophilic attack from the β-side [63,64].

An interesting addition to the repertoire of anchimerically assisting groups is the chiral 1-phenyl-2-(phenylsulfanyl)ethyl group [65]. Coupled as (*S*)-isomer **8a** to the 2-hydroxy group of the glucopyranosyl trichloroacetimidates with acceptor **8b** 1,2-*cis*- i. e. α-glucopyranosides **8c** have been obtained in excellent yields and stereoselectivities (*Scheme 8*). The formation of a cyclic β-linked sulfonium ion intermediate **8a***, having a *trans*-decalin type structure has been confirmed by NMR experiments. This intermediate seems to be preferentially or exclusively attacked by the nucleophile from the α-side.

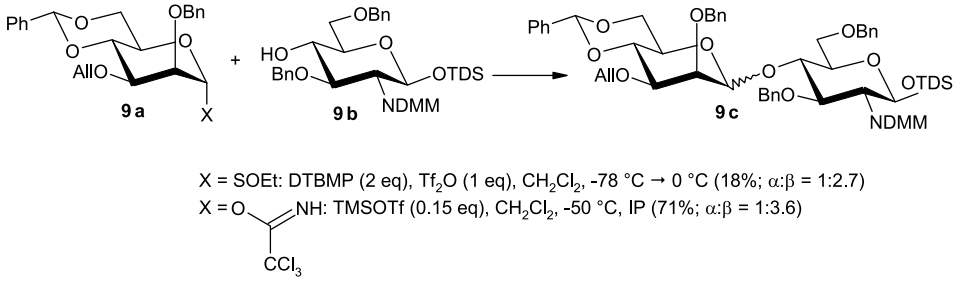
A particularly difficult problem is 1,2-*cis*- i. e. β-linkage formation in β-mannopyranoside synthesis. The presence of β-linked mannopyranosides in various natural products [66], particularly in the *N*-glycan core structure of glycoproteins [66,67], led to the search for efficient methodologies for generating this target structure. Investigation of mannopyranosyl donors with nonparticipating protecting groups and different leaving groups led in general mainly or exclusively to α-products [1,66,67,68]. Several specific methods led to some success in this endeavor [69,70,71,72,73,74], however finally epimerization of β-glucopyranosides at 2-*O* [74,75,76,77,78,79] and intramolecular aglycone delivery [80,81,82,83,84] have led



Scheme 7

to very successful results. Mannopyranosyl donors with diol *O*-protecting groups leading to ring annelation had already been investigated but again with limited success [85]. Therefore, it was a big surprise that 2,3-di-*O*-alkyl-4,6-*O*-benzylidene-protected mannopyranosyl sulfoxides as glycosyl donors gave preferentially β-products with various acceptors at low temperatures [86,87]. The same result is more conveniently obtained with the corresponding trichloroacetimidates as shown in *Scheme 9* with glycosyl donor **9a** furnishing with accep-

Scheme 8

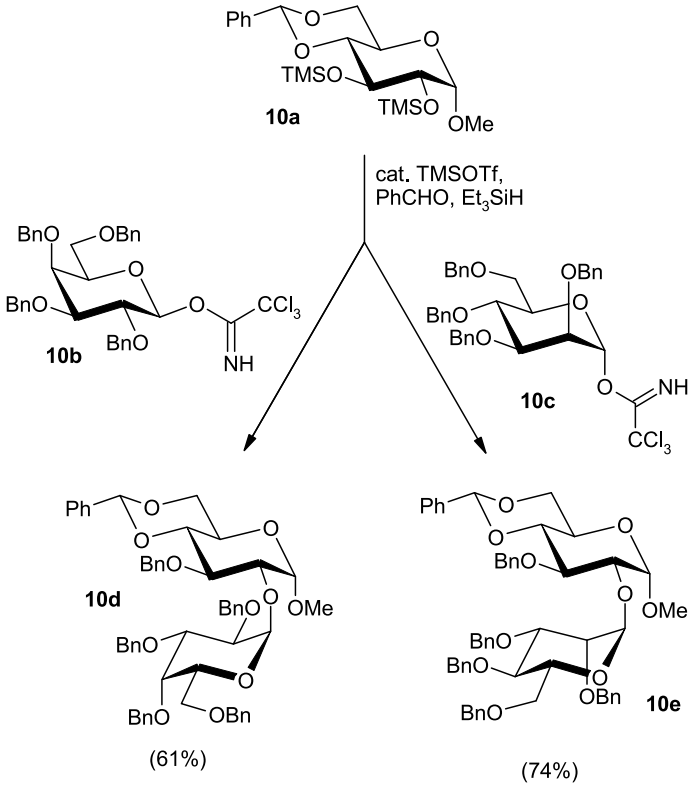


Scheme 9

tor **9b** preferentially the β-disaccharide **9c** [26]. The results obtained on varying the reaction parameters were not compatible with the reaction mechanism proposed for sulfoxide activation, in which an α-triflate intermediate is thought to play the decisive role [87]. Rather the anomeric stereocontrol is caused by a conformational effect enforced by the 4,6-*O*-benzylidene group on the pyranosyl ring, which favors the generation of a flattened twist-boat conformation as the intermediate. For stereoelectronic and steric reasons, this twist-boat intermediate will be preferentially attacked from the β-side, thus forming after equilibration the $^4C_1$-conformer of the β-mannopyranoside [26]. This mechanistic proposal is also confirmed by L-rhamnosylation reactions with bulky 3-*O*- and 4-*O*-protecting groups which enforce the $^4C_1$-conformation and this way favor β-L-rhamnoside formation [88]. On the basis of these mechanistic considerations, β-mannopyranoside formation should be facilitated by nonparticipating, strongly electron-withdrawing groups at the 2-*O* atom because generation of the twist-boat intermediate would gain from a strong dipole effect. This expectation could be confirmed [89].

Regioselectivity in glycosylation reactions is generally based on a sequence consisting first of regioselective functional group protection, making use of various principles, second glycosylation of the *O*-unprotected hydroxy group, and finally of *O*-deprotection. Direct regioselective glycosylation using the difference in reactivity of sugar hydroxy groups is of great interest as it often avoids cumbersome protection and deprotection steps. Many regioselective glycosylations of partly unprotected carbohydrate acceptors have been reported in the literature [1,2,3,4,5,6,7]. Generally, they make use of the higher reactivity of primary hydroxy groups over secondary hydroxy groups (particularly those in axial orientation) and the higher reactivity of equatorial 3-hydroxy groups, for instance in galacto- and glucopyranosides. This was again confirmed in the glycosylation of 4,6-*O*-benzylidene-glucopyranosides which gave with various glycosyl trichloroacetimidates as donors the (1–3)-linked disaccharides in very good yields [90]. Another interesting alternative is one-pot regioselective protection and following *O*-glycosylation with *O*-glycosyl trichloroacetimidates which, as shown in *Scheme 10* in the glycosylation of **10a** with donors **10b** and **10c** affording disaccharides **10d** and **10e**, respectively, makes use of this reactivity difference [91].

Separation of the target glycosides from the reaction mixture generally requires column chromatography purification. Alternatively, highly fluorinated compounds are readily separated from nonfluorinated compounds by a simple phase separation. Therefore, organic synthesis

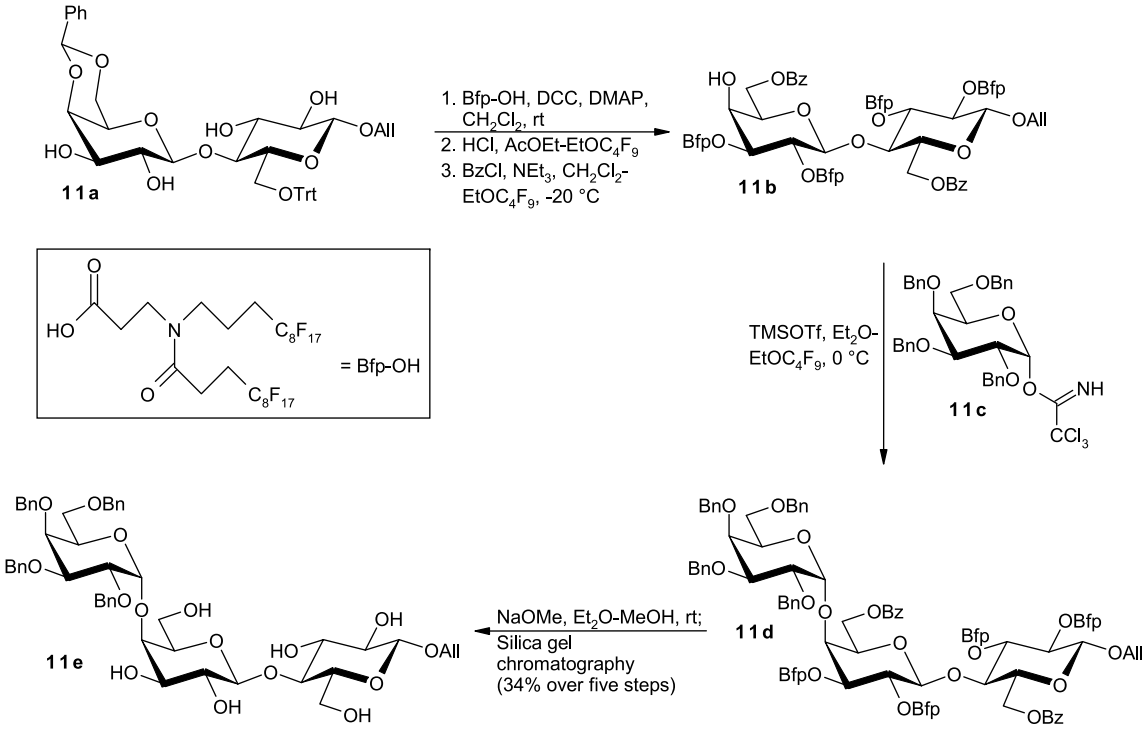


Scheme 10

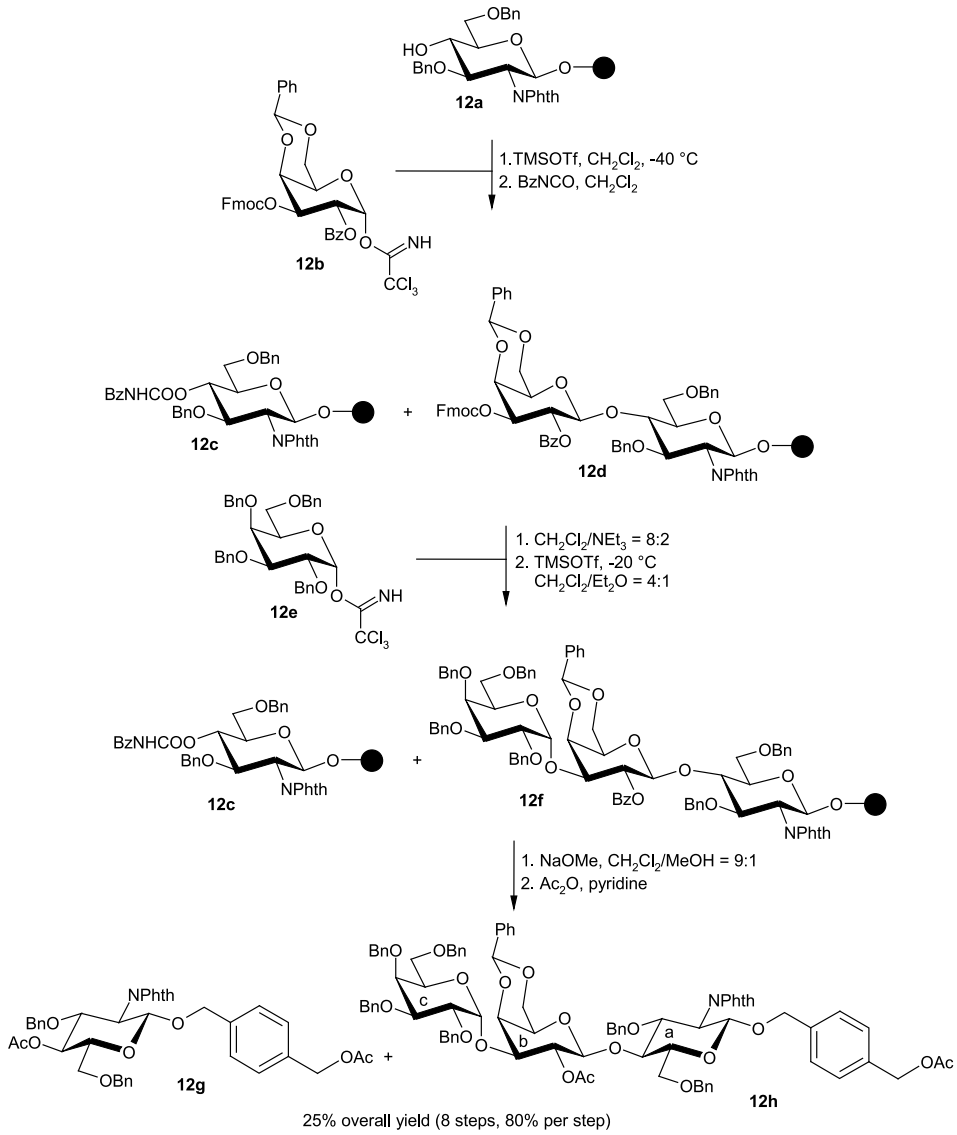
with substrates having fluorinated tags has become a valuable alternative to solid-phase synthesis (see below). This concept has been already applied to oligosaccharide synthesis [92]. On the basis of the novel fluorous acyl protecting group Bfp (**b**is **f**luorous chain type **p**ropanoyl) highly successful oligosaccharide syntheses could be performed [93], which were extended to the synthesis of globo-triaosyl ceramide **11e** from lactoside **11a** as shown in *Scheme 11* [94]. All fluorous intermediates (**11b** and **11d**) were extracted with FC-72 (perfluorohexane isomers) and an organic solvent (toluene or methanol) and were purified without silica gel chromatography. After removal of the Bfp groups by simple base-catalyzed methanolysis the pure allyl trisaccharide **11e** was obtained by only one silica gel column chromatography in 34% total yield after five steps (~81% per step).

4.2 Glycoside Synthesis on Polymer Supports

Successful solid-phase oligosaccharide syntheses (SPOS) have been developed by several research groups [95,96,97,98,99,100,101,102,103,104,105], which exhibit the inherent advantages over solution phase synthesis, such as (i) higher reaction yields due to the use


Ph
HO
OH
HO
OAll
OH
OTrt
11a
1. Bfp-OH, DCC, DMAP, CH2Cl2, rt
2. HCl, AcOEt-EtOC4F9
3. BzCl, NEt3, CH2Cl2-EtOC4F9, -20 °C
HO
OBz
BfpO
OBfp
Bfp
OBfp
OAll
OBz
11b
HO
C8F17
C8F17
= Bfp-OH
TMSOTf, Et2O-EtOC4F9, 0 °C
BnO
OBn
BnO
BnO
NH
CCl3
11c
NaOMe, Et2O-MeOH, rt; Silica gel chromatography (34% over five steps)
11d
11e

Scheme 11



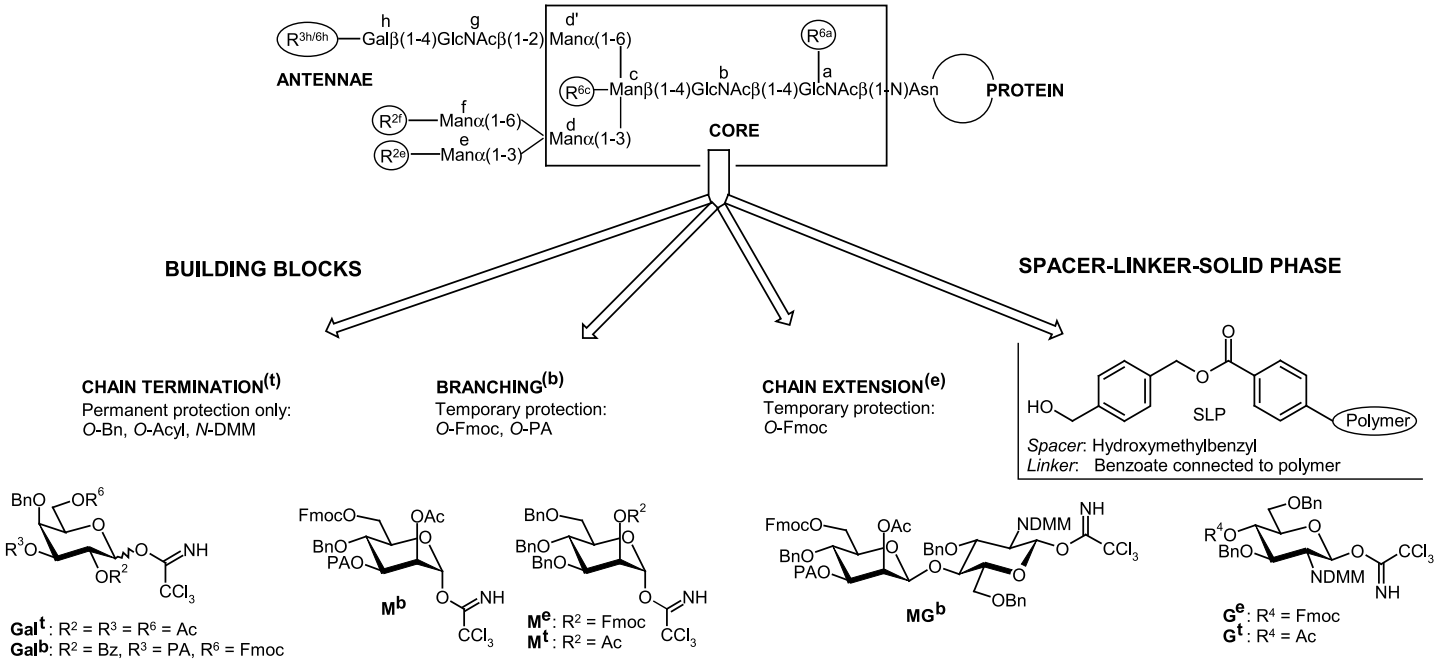
Scheme 12

of excess building blocks, (ii) shorter reaction times for the completion of total syntheses, and (iii) convenient purification procedures by just washing the resin. In addition, methods to avoid undesired byproduct formation in the synthesis of the target molecules have been introduced [106,107,108,109], such as for instance capping procedures of unreacted intermediates (*Scheme 12*) [108]. Low reactive acceptor **12a** after glycosylation with donor **12b** is capped with benzoyl isoyanate yielding **12c** and **12d**. Chain extension with **12e** (→ **12c** + **12f**) and cleavage led to readily separable **12g** and **12h**. So far, no generally accepted strategy has yet

appeared for the efficient construction of various complex oligosaccharides on polymer supports, thus limiting the commercialization of automated synthesizers. However, it seems that *O*-glycosyl trichloroacetimidates and the less reactive *O*-glycosyl phosphates have the potential to become the glycosyl donors of choice because the requirement of catalytic amounts of just one activator is a major advantage over the other glycosylation methods.
After various attempts of the Schmidt group with different linkers between solid support and the carbohydrate groups, such as for instance

- thioglycosides and their cleavage by thiophilic reagents,
- silyl glycosides and their cleavage by fluorides,
- pentenyl-type glycosides and their cleavage by electrophilic reagents, and
- allyl-type glycosides and their cleavage by cross metathesis or by ring-closing metathesis, respectively [21,110,111],

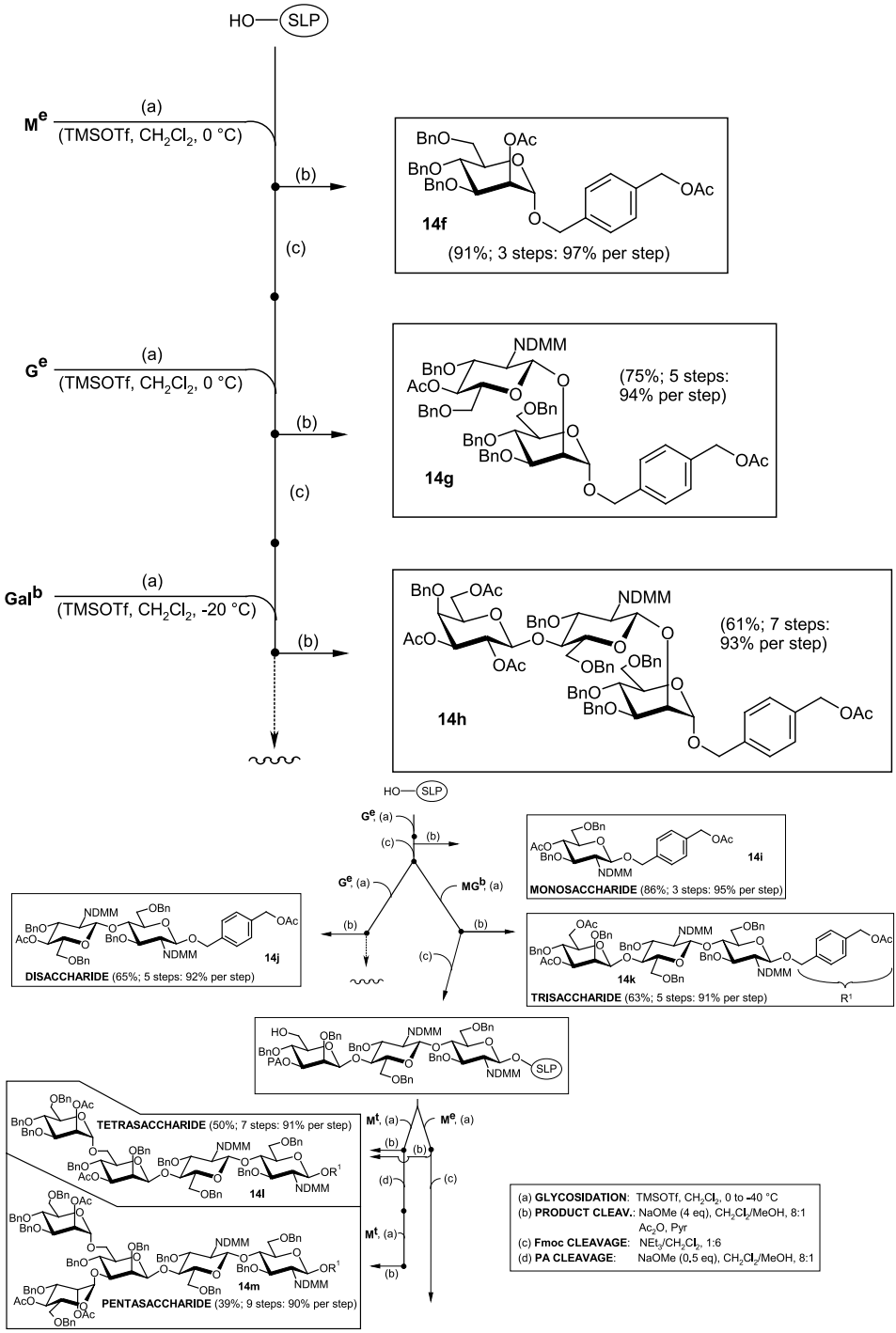
finally the ester-based SPOS design was introduced which gave excellent results in chain extension and product cleavage from the resin [110,111]. However, to cope with complex oligosaccharide synthesis, besides the ester-linker and -spacer system, three types of glycosyl donor building blocks for controlled chain extension (suffix **e**), branching (suffix **b**), and chain termination (suffix **t**) are required. An efficient solution to these requirements for SPOS of a small library of *N*-glycans is exhibited in ❯ *Scheme 13* which shows the retrosynthesis to the required building blocks and the required reaction steps. A particularly important role played the selection of the *N*-protecting group for the glucosamine residue.
The ester-based SPOS methodology (❯ *Scheme 13*) comprises (i) different types of esters, that is, the benzoate group as a linker and for chain termination and the Fmoc and PA (phenoxyacetyl) group as temporary protecting groups for chain extension and branching which can be chemoselectively cleaved (in the sequence Fmoc and then PA), (ii) the benzyl group for permanent *O*-protection and for the spacer between the anomeric center at the reducing end sugar, thus providing after final product cleavage from the resin a structurally defined target molecule, (iii) *O*-glycosyl trichloroacetimidates of type **e**, **b**, or **t** (for chain extension, branching, or termination) as powerful glycosyl donors, which can be readily activated by catalytic amounts of (Lewis) acid, and (iv) benzoic acid residues on the Merrifield resin for the linkage of the hydroxymethylbenzyl spacer. Hence, retrosynthesis of a typical *N*-glycan molecule containing the core pentasaccharide and some antennae leads to spacer-linker connected Merrifield resin SLP and to glycosyl donors $\mathbf{G}^{e}$, $\mathbf{G}^{t}$, $\mathbf{Gal}^{t}$, $\mathbf{Gal}^{b}$, $\mathbf{M}^{b}$, $\mathbf{M}^{t}$, and $\mathbf{MG}^{b}$ which can be selectively converted into acceptors on resin (**e** and **b**-type donor building blocks). Thus, as indicated in ❯ *Scheme 13*, only four simple procedures are required for successful SPOS: (a) glycosidation under TMSOTf catalysis; (b) product cleavage under transesterification conditions; (c) selective Fmoc cleavage under basic conditions; and (d) selective PA cleavage under milder transesterification conditions.
On the basis of this concept, as shown in ❯ *Scheme 14*a–c, a small library of *N*-glycans (compounds **14a** to **14q**) was successfully synthesized: (i) The Merrifield resin as solid support exhibited excellent results during all stages of the assembly; (ii) all glycosylations, including those with *N*-DMM protected glycosyl donors, gave high yields; (iii) the methodology presented herein shows the desired versatility in terms of efficient chain extension and branching requiring only two standard (in one direction) orthogonal protecting

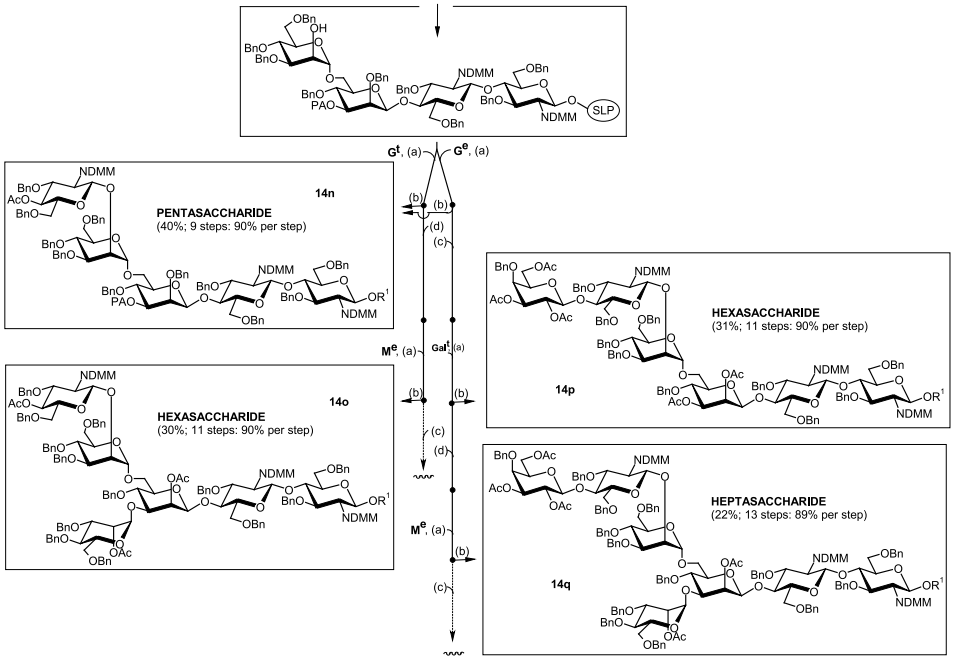


REACTIONS ON SOLID PHASE

(a) **Glycosidation**: *O*-Glycosyl Trichloroacetimidates, TMSOTf (cat.), CH_2Cl_2, 0 to 40 °C
(b) **Product cleavage**: NaOMe (4 eq), CH_2Cl_2/MeOH (8:1); Ac_2O, Pyr
(c) **Fmoc cleavage**: NEt_3/CH_2Cl_2 (1:6)
(d) **PA cleavage**: NaOMe (0.5 eq), CH_2Cl_2/MeOH (8:1)
} UV-Monitoring possible

Scheme 13
Solid phase synthesis of high mannose, complex, and hybrid type *N*-glycans: increased donor and acceptor reactivity

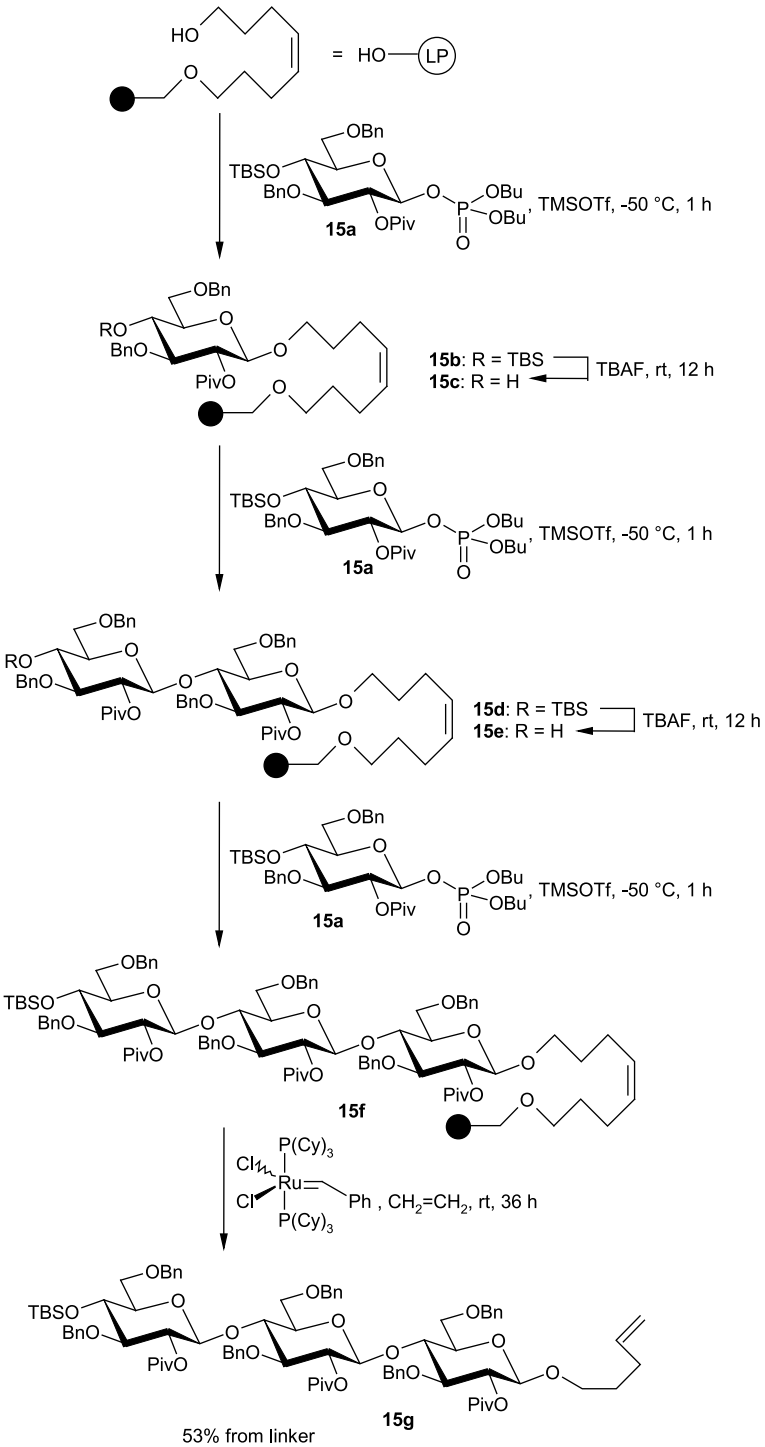

HO—SLP
M^e
(a)
(TMSOTf, CH2Cl2, 0 °C)
(b)
(c)
14f
(91%; 3 steps: 97% per step)
G^e
(a)
(TMSOTf, CH2Cl2, 0 °C)
(b)
(c)
14g
(75%; 5 steps: 94% per step)
Gal^b
(a)
(TMSOTf, CH2Cl2, -20 °C)
(b)
14h
(61%; 7 steps: 93% per step)
14j
DISACCHARIDE (65%; 5 steps: 92% per step)
14i
MONOSACCHARIDE (86%; 3 steps: 95% per step)
14k
TRISACCHARIDE (63%; 5 steps: 91% per step)
TETRASACCHARIDE (50%; 7 steps: 91% per step)
14l
14m
PENTASACCHARIDE (39%; 9 steps: 90% per step)
(a) GLYCOSIDATION: TMSOTf, CH2Cl2, 0 to -40 °C
(b) PRODUCT CLEAV.: NaOMe (4 eq), CH2Cl2/MeOH, 8:1
Ac2O, Pyr
(c) Fmoc CLEAVAGE: NEt2/CH2Cl2, 1:6
(d) PA CLEAVAGE: NaOMe (0.5 eq), CH2Cl2/MeOH, 8:1



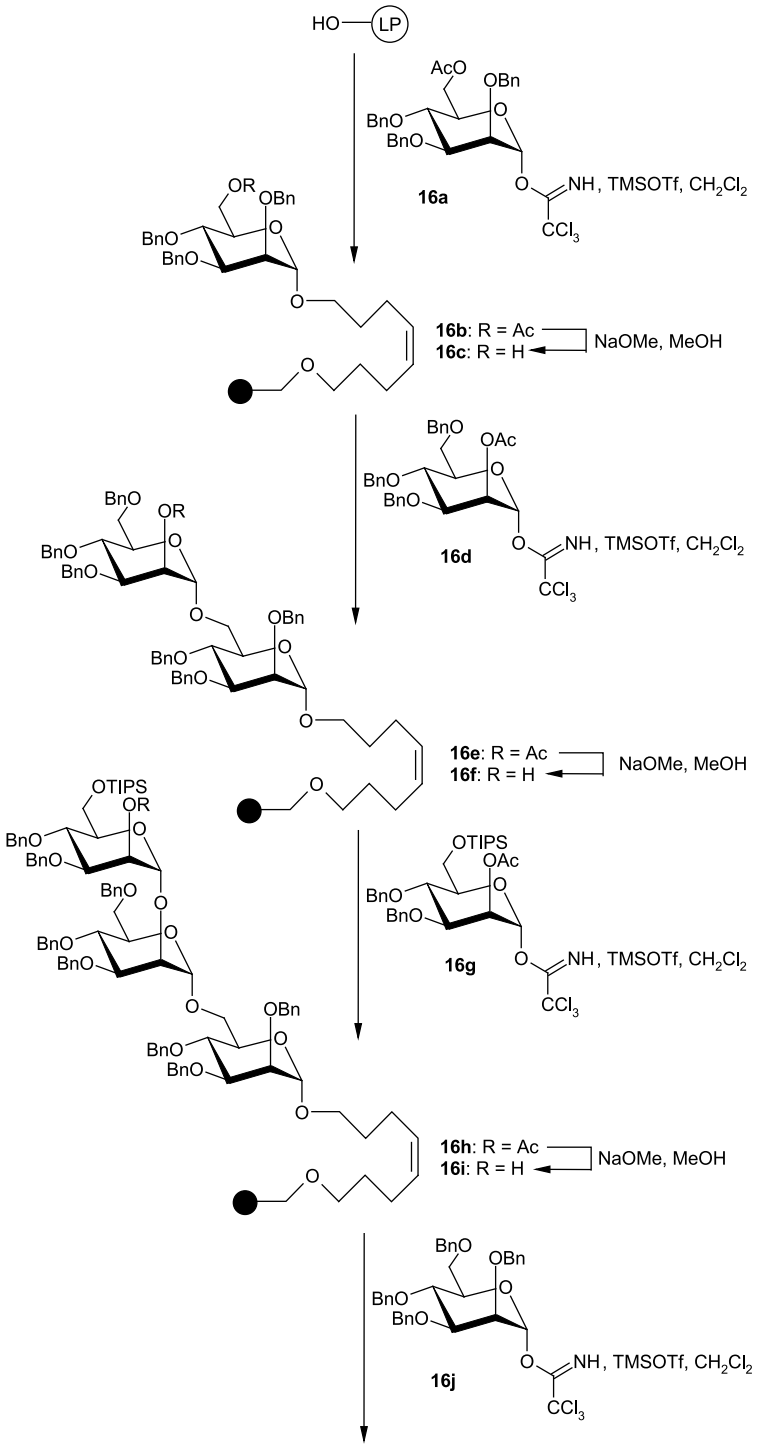
Scheme 14 (left page and above)
Solid phase synthesis of a small library of *N*-glycans: typical constituents

groups; (iv) cleavage of the product from the resin was feasible leading to stable 1-*O*-benzyl type products with only benzyl, DMM and, after acetylation, acetyl protection; (v) the crude products were already of high purity; therefore, standard silica gel chromatography and MPLC were sufficient for purification; (vi) yields of isolated products were high, ranging from 97% per step (after three steps) to 89% per step (after 13 steps) on solid-phase; (vii) the methodology is technically simple, thus lending itself available to automation.

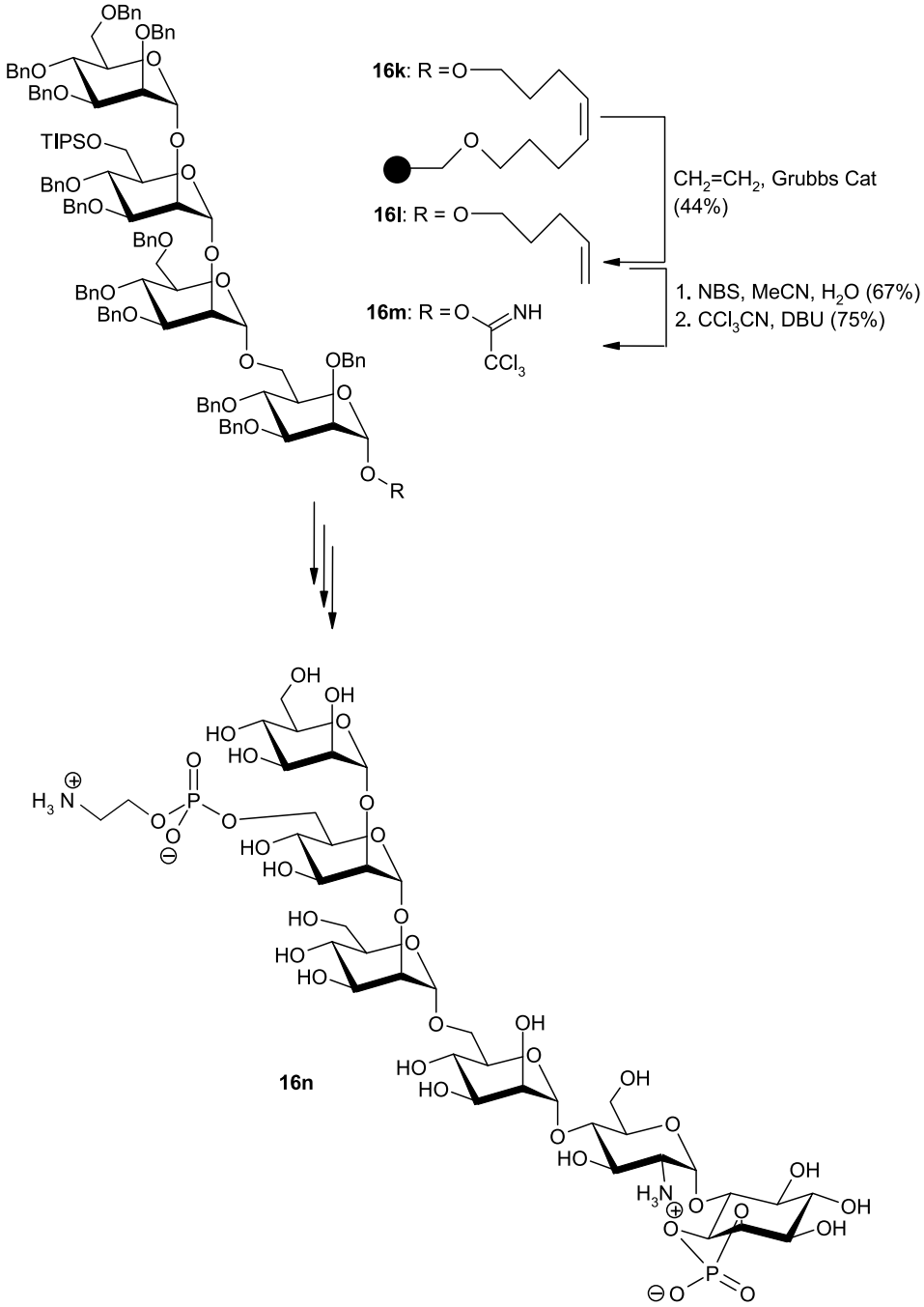
Very good SPOS results have also been obtained by the Seeberger group with *O*-glycosyl phosphates and *O*-glycosyl trichloroacetimidates, respectively, as glycosyl donors and (*Z*)-4-octene-1,8-diol as the linker between the Merrifield resin and the sugar residue. The linker was cleaved by cross-metathesis with ethylene, thus providing pentenyl glycosides as cleavage products (Schemes 15 and 16) [112]. *Scheme 15* shows the synthesis of pentenyl cellotrioside **15g** with **15a** as the donor over seven steps with **15b–f** as intermediates in 53% overall yield. As a temporary protecting group in the glycosyl donor the TBS group in the 4-position was employed. The β-selectivity was supported by a 2-*O*-pivaloyl group. This method was also employed in an automated SPOS which was based on a modified peptide synthesizer and *O*-glycosyl trichloroacetimidates as glycosyl donors (*Scheme 16*) [96,113,114]. This way a glycosylphosphatidyl inositol (GPI) (**16n**) was synthesized which functions as a malarial toxin. Starting from three different *O*-glycosyl trichloroacetimidates (**16a**, **d**, **f**)


HO
= HO—LP
OBn
TBSO
BnO
O
OBu
OBu
P
O
15a
OPiv
, TMSOTf, -50 °C, 1 h
RO
PivO
15b: R = TBS
15c: R = H
TBAF, rt, 12 h
15d: R = TBS
15e: R = H
15f
P(Cy)3
Cl
Ru
Ph , CH2=CH2, rt, 36 h
15g
53% from linker

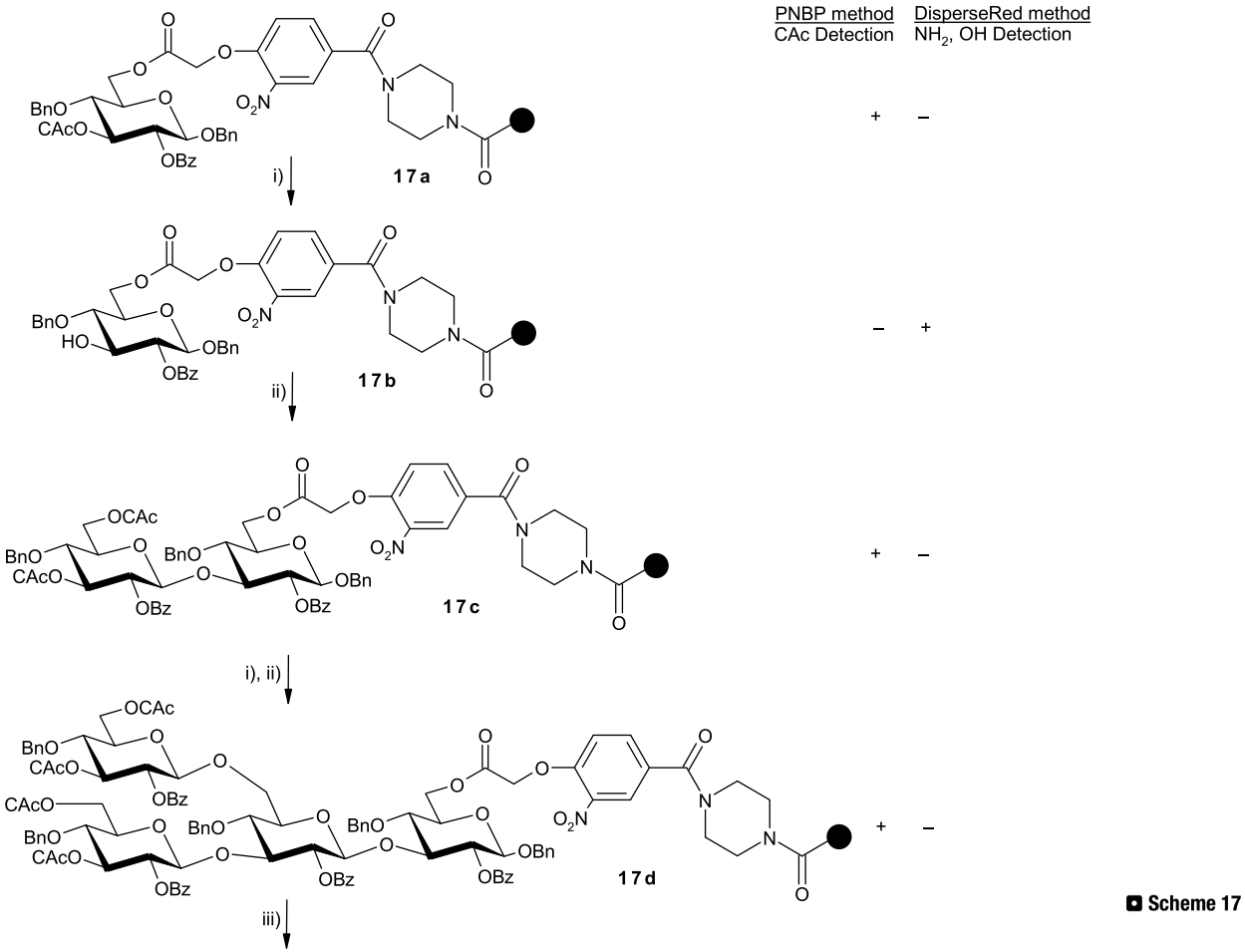
Scheme 15


HO
LP
AcO
OBn
BnO
BnO
O
NH, TMSOTf, CH_2Cl_2
CCl_3
16a
OR
OBn
16b: R = Ac
16c: R = H
NaOMe, MeOH
BnO
OAc
16d
OTIPS
16e: R = Ac
16f: R = H
16g
16h: R = Ac
16i: R = H
16j

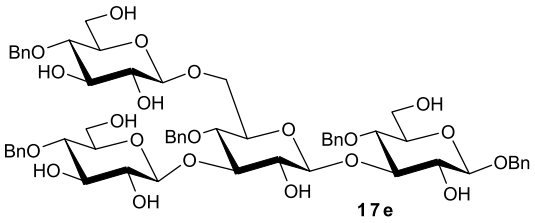
Scheme 16



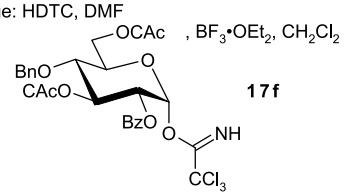
Scheme 16
(continued)


PNBP method
CAc Detection
DisperseRed method
NH2, OH Detection
17a
17b
17c
17d
i)
ii)
i), ii)
iii)
+ –
– +
+ –
+ –

Scheme 17



i) (a) Capping: Ac_2O, $^iPr_2NEt$, CH_2Cl_2
(b) CAc-Cleavage: HDTC, DMF
ii) Glycosylation:



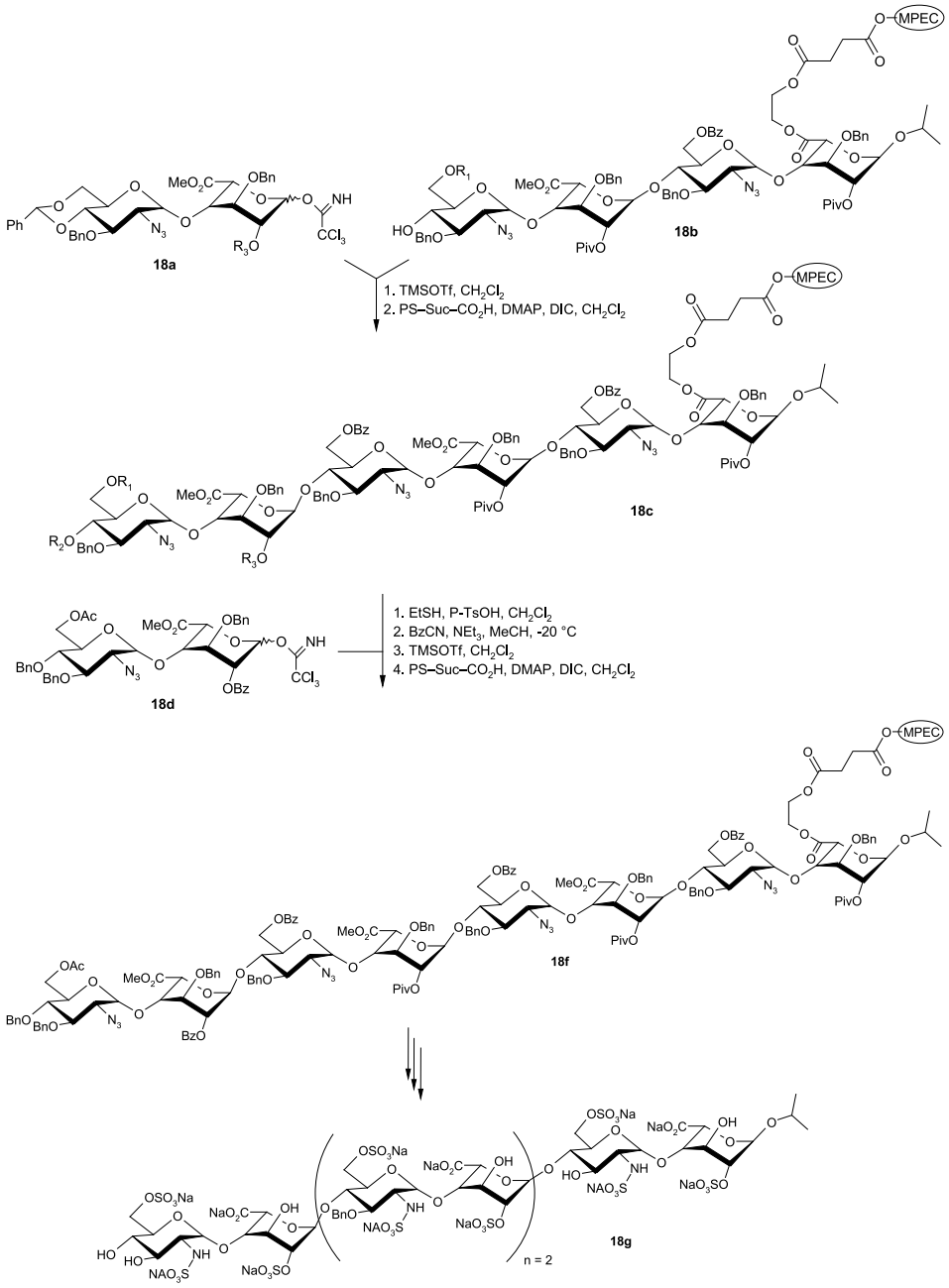
iii) Cleavage from resin: NaOMe, MeOH, THF (30% overall yield)

Scheme 17 (continued)

having *O*-acetyl groups as temporary protection for chain extension and one per-*O*-benzylated *O*-mannopyranosyl trichloroacetimidate (**16j**) for chain termination a tetrasaccharide **16k** was obtained which was cleaved from the resin (→ **16l**) by cross metathesis. Transformation into a trichloroacetimidate **16m** and glycosylation of a pseudodisaccharide led after deprotection to the target molecule **16n**.

For the real-time monitoring of the glycosylation result by the Ito group sensitive color tests were introduced with p-nitrobenzylpyridine (PNBP), reacting with chloroacetyl (CAc) as the temporary protecting group under red color formation, and with cyanuric chloride-Disperse Red conjugate, which readily reacts with hydroxy or amino groups to give a red color on the resin [115]. This way, as shown in *Scheme 17*, the repeating unit **17e** of the immuno-active oligosaccharide schizophyllan was synthesized based on a partly CAc-protected *O*-glucosyl trichloroacetimidate donor **17f** with **17a** as starting material and **17b** to **17d** as intermediates; the presence of the CAc groups could be readily monitored with PNBP and the presence or absence of hydroxy groups was monitored by the Disperse-Red method. As a linker to piperazine-modified TentaGel, a 2-nitrophenyloxyacetyl group [99] was employed that can be readily cleaved under transesterification conditions. This led also to loss of other ester protecting groups furnishing the target molecule **17f** in very good overall yield.

The synthesis of oligosaccharides on soluble supports has been further investigated as well. Particularly worth mentioning is the synthesis of heparin-like oligosaccharides on polyethylene glycol ω-monomethyl ether (MPEG) resin with a succinoyl ester linker and *O*-glycosyl trichloroacetimidates as glycosyl donors [116,117]. A typical example is shown in *Scheme 18* [117]. After the glycosylation steps of **18a** with polymer bound **18b** and **18c** with **18d**, respectively, catalyzed by TMSOTf, a capping step was included in which unreacted MPEG-bound acceptor was captured onto the solid-phase of succinate-functionalized Merri-


18a
18b
1. TMSOTf, CH2Cl2
2. PS–Suc–CO2H, DMAP, DIC, CH2Cl2
18c
18d
1. EtSH, P-TsOH, CH2Cl2
2. BzCN, NEt3, MeCH, -20 °C
3. TMSOTf, CH2Cl2
4. PS–Suc–CO2H, DMAP, DIC, CH2Cl2
18f
18g
n = 2

Scheme 18

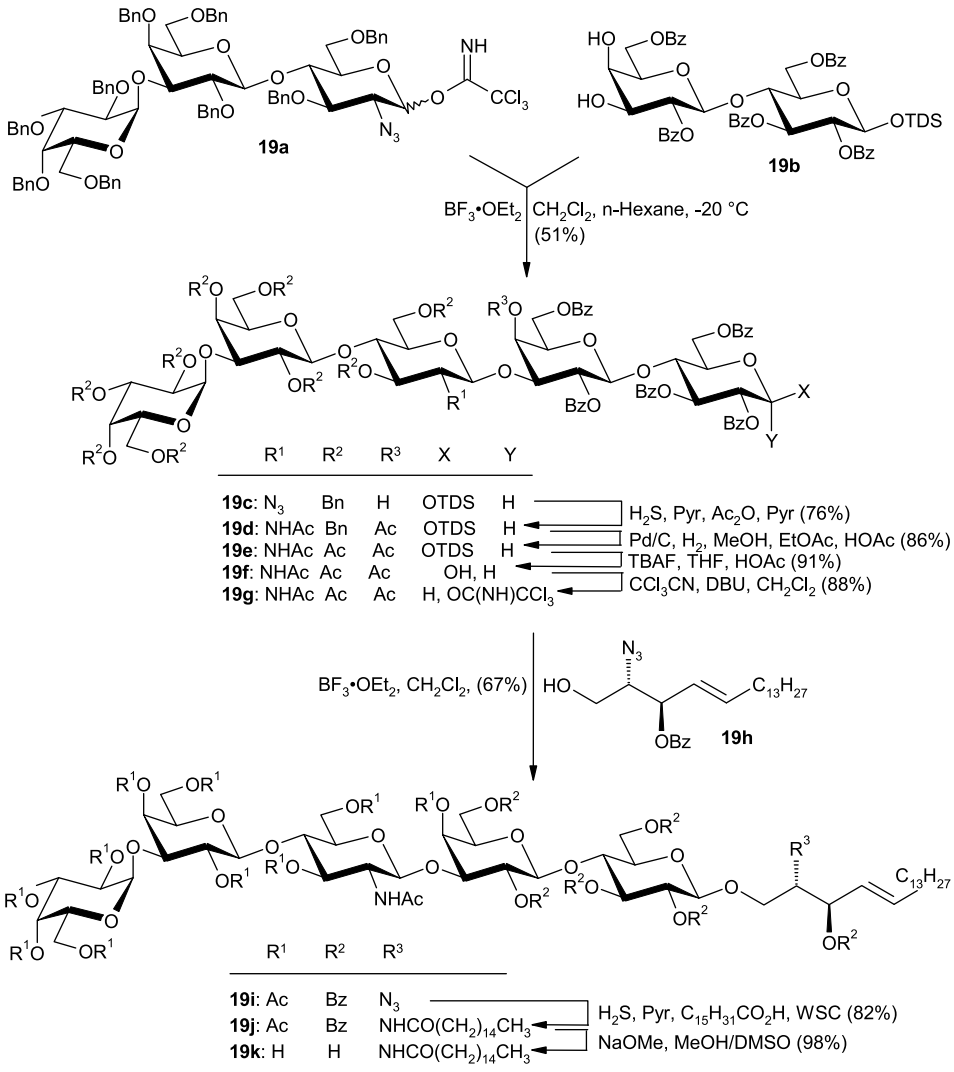
field resin (PS–Suc-CO_2H) by a chemoselective ester formation (→ **18c** and **18f**). Alkaline hydrolysis released the octasaccharide from the resin and also all other ester groups were cleaved. By further known transformations [118] the octasaccharide fragment **18g**, containing the structural motif of the regular region of heparin, was obtained.
An interesting comparative study between solution phase synthesis and synthesis of the same oligosaccharides on MPEG as soluble support was carried out by the Furneaux group [119]. This study clearly reflected the incompatibility of MPEG as a soluble support with some common reagents, reaction conditions, and choice of standard protecting groups. Hence, the shortcomings associated with this support seem to outbalance the merits [119,120].
To avoid some limitations of PEGs, as for instance low loading, hyperbranched polyester such as Boltorn H40 and H50 have been successfully investigated in glycosylation reactions with *O*-glycosyl trichloroacetimidates [121]. For simple mono- and disaccharide synthesis optimization of the reaction conditions led to practically quantitative transformations.

5 Recent Applications of *O*-Glycosyl Trichloroacetimidates in Complex Oligosaccharide and Glycoconjugate Synthesis

Several comprehensive reviews have been devoted to the use of *O*-glycosyl trichloroacetimidates in glycosylation reactions. In this overview, so far the basic principles of the glycosylation methods, the strengths of the *trichloroacetimidate method*, differences to other methods, and methodological variations have been discussed. In this chapter recent applications of this method in complex glycoconjugate synthesis will be highlighted.

5.1 Glycolipids

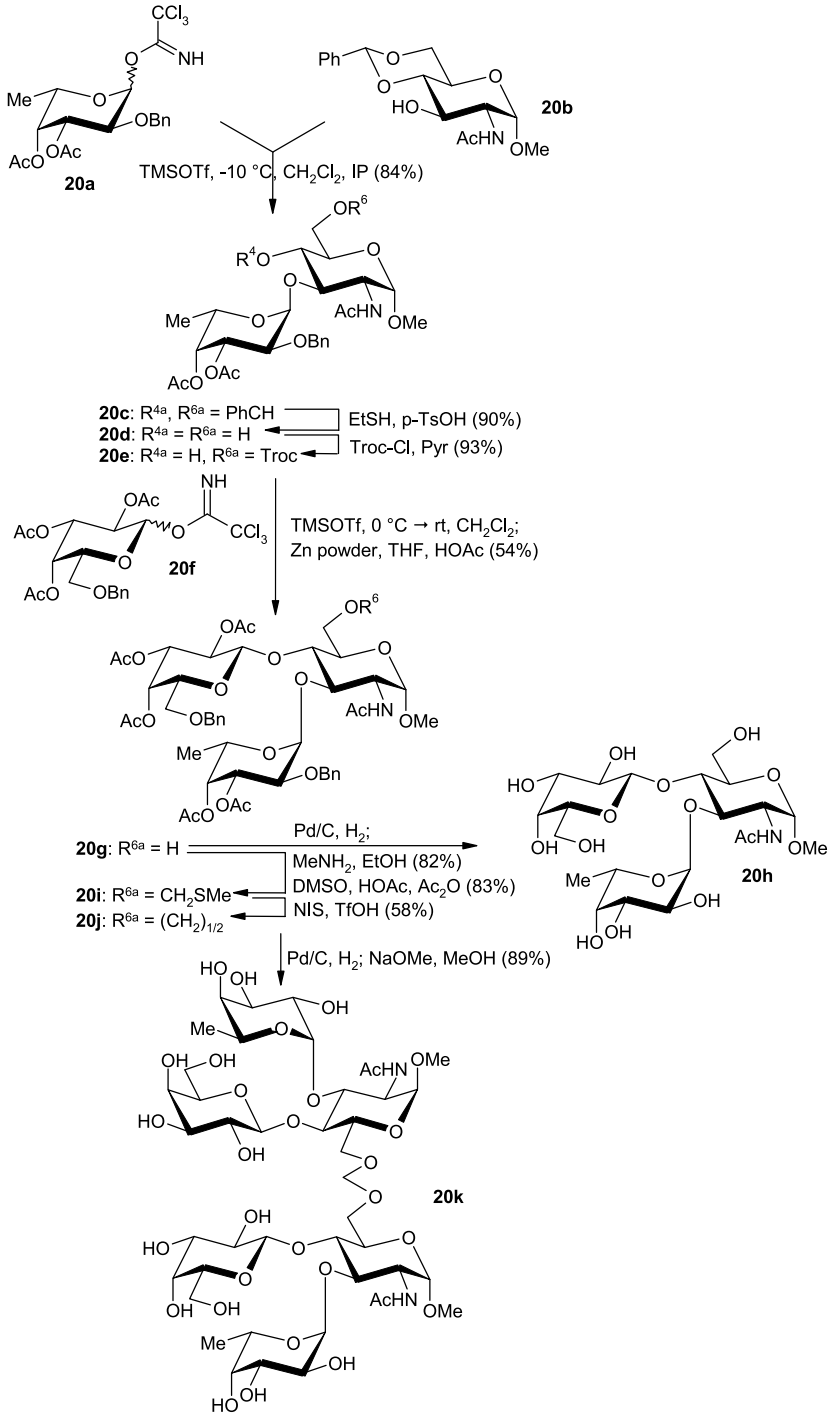
Glycosphingolipid synthesis has remained a target of great interest because of the biological importance of these compounds. Assembly of the sugar residues with *O*-glycosyl trichloroacetimidates as donors and attachment of the ceramide residue via the azidosphingosine *glycosylation procedure* [122,123] has become the method of choice. On the basis of this approach, aminodeoxy analogs of globotriosyl ceramide [124], globo- and isoglobotriosides bearing cinnamoylphenyl tags [125], and a quite practical globotriose synthesis [126], which is essentially based on *O*-acyl protection, have been carried out. An efficient glycosylation protocol for the attachment of an α-linked galactosamine residue for asialo GM2 synthesis has been reported [127]. A novel ether-bridged GM3 lactone analog has been successfully prepared and used in antibody-based cancer therapy studies [128]. Also ganglioside GD3 was synthesized [129]; comparison with bovine-brain derived GD3 showed that the effects in GD3-triggered uncoupling of mitochondrial respiration and induction of apoptosis in oligodendrocytes are very similar. Also ganglioside mimics for binding studies with myelin-associated glycoprotein were prepared [130].
The *lacto*- and *neolacto*-series of glycosphingolipids also attracted further attention. The synthesis of the lacto-*N*-neotetraose and lacto-*N*-tetraose building blocks has been improved by using the *N*-dimethylmaleoyl protecting group [131]. A total synthesis of the natural antigen **19k** involved in a hyperacute rejection response to xenotransplants has been carried out as



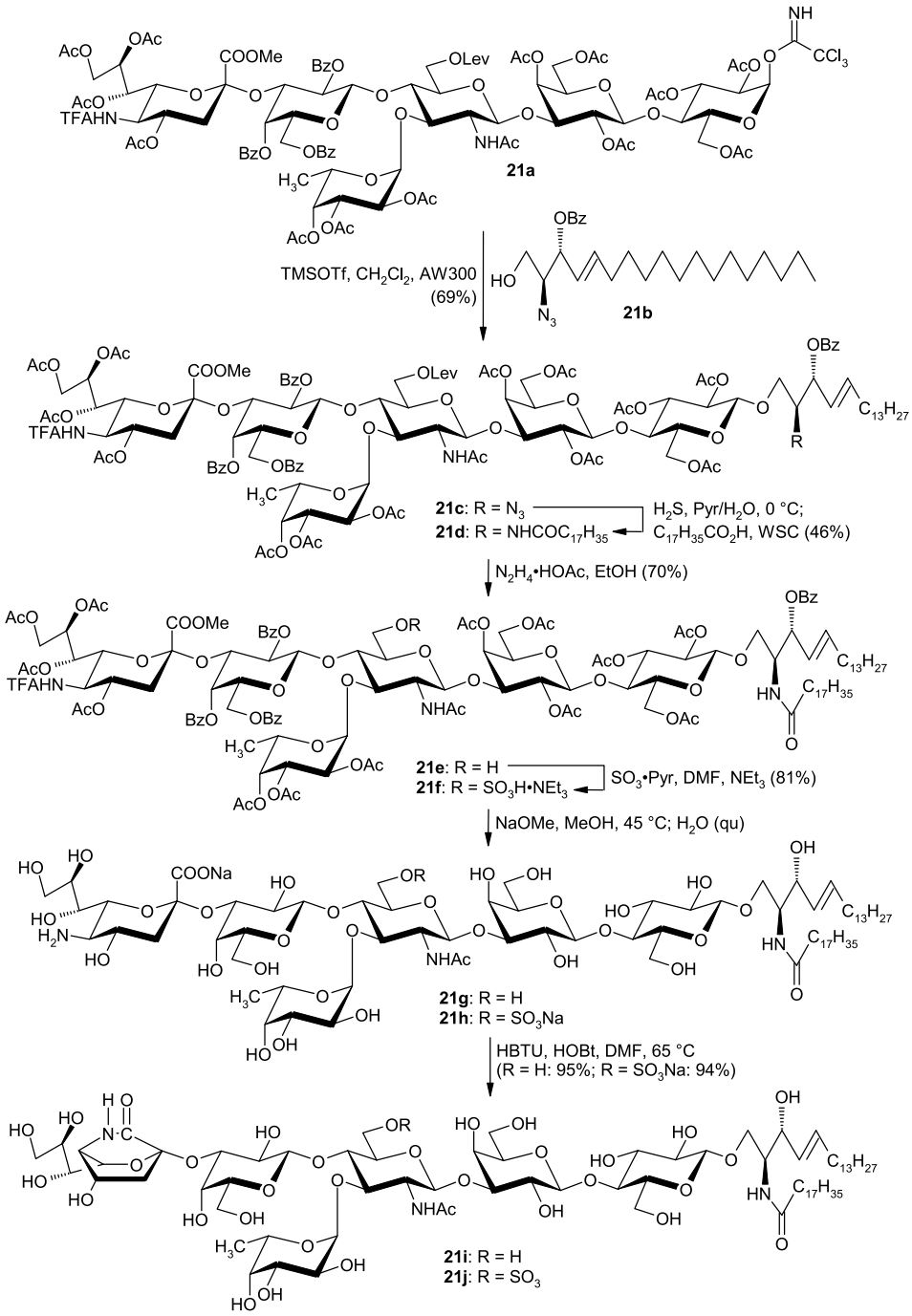
Scheme 19

shown in *Scheme 19* [132]. Glycosyl donor **19a**, also obtained via the *trichloroacetimidate method*, gave with acceptor **19b** pentasaccharide **19c**. Standard transformation via **19d**, **19e**, and **19f** led to pentasaccharide donor **19g**. Application of the *azidosphingosine glycosylation procedure* with **19h** as acceptor furnished **19i** which was transformed via **19j** into the target molecule **19k**.

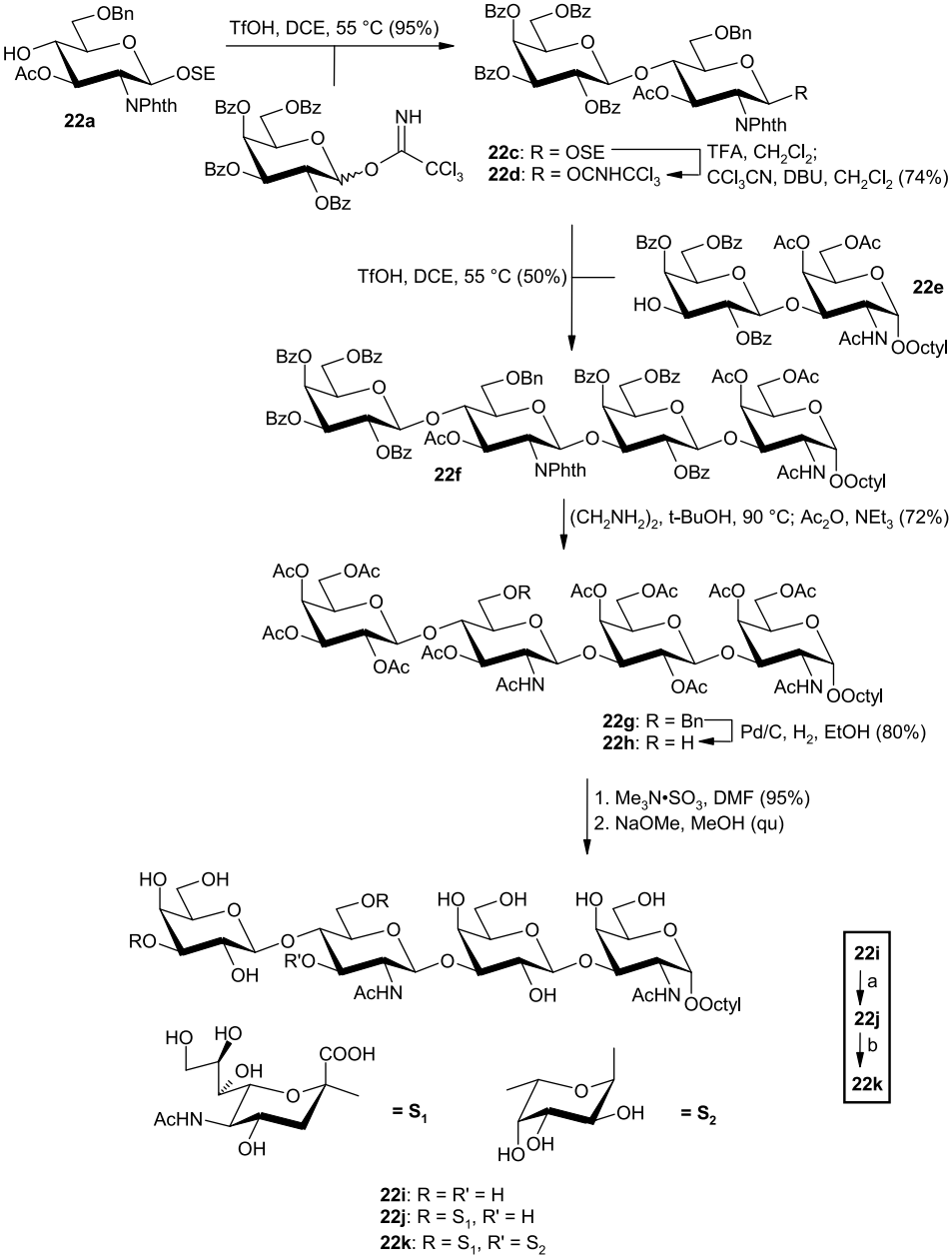
Further, for microdomain formation studies fluorescence labeled sialyl Lewis X [133], for cluster effect studies dimeric sialyl Lewis X [134], and for carbohydrate–carbohydrate recognition studies a dimer of Lewis X [135] were synthesized (*Scheme 20*, **20k**). To this aim,


20a
20b
TMSOTf, -10 °C, CH2Cl2, IP (84%)
20c: R4a, R6a = PhCH
20d: R4a = R6a = H
20e: R4a = H, R6a = Troc
EtSH, p-TsOH (90%)
Troc-Cl, Pyr (93%)
20f
TMSOTf, 0 °C → rt, CH2Cl2;
Zn powder, THF, HOAc (54%)
20g: R6a = H
Pd/C, H2;
MeNH2, EtOH (82%)
20i: R6a = CH2SMe
DMSO, HOAc, Ac2O (83%)
20j: R6a = (CH2)1/2
NIS, TfOH (58%)
20h
Pd/C, H2; NaOMe, MeOH (89%)
20k

Scheme 20


21a
TMSOTf, CH2Cl2, AW300
(69%)
21b
21c: R = N3
21d: R = NHCOC17H35
H2S, Pyr/H2O, 0 °C;
C17H35CO2H, WSC (46%)
N2H4•HOAc, EtOH (70%)
21e: R = H
21f: R = SO3H•NEt3
SO3•Pyr, DMF, NEt3 (81%)
NaOMe, MeOH, 45 °C; H2O (qu)
21g: R = H
21h: R = SO3Na
HBTU, HOBt, DMF, 65 °C
(R = H: 95%; R = SO3Na: 94%)
21i: R = H
21j: R = SO3

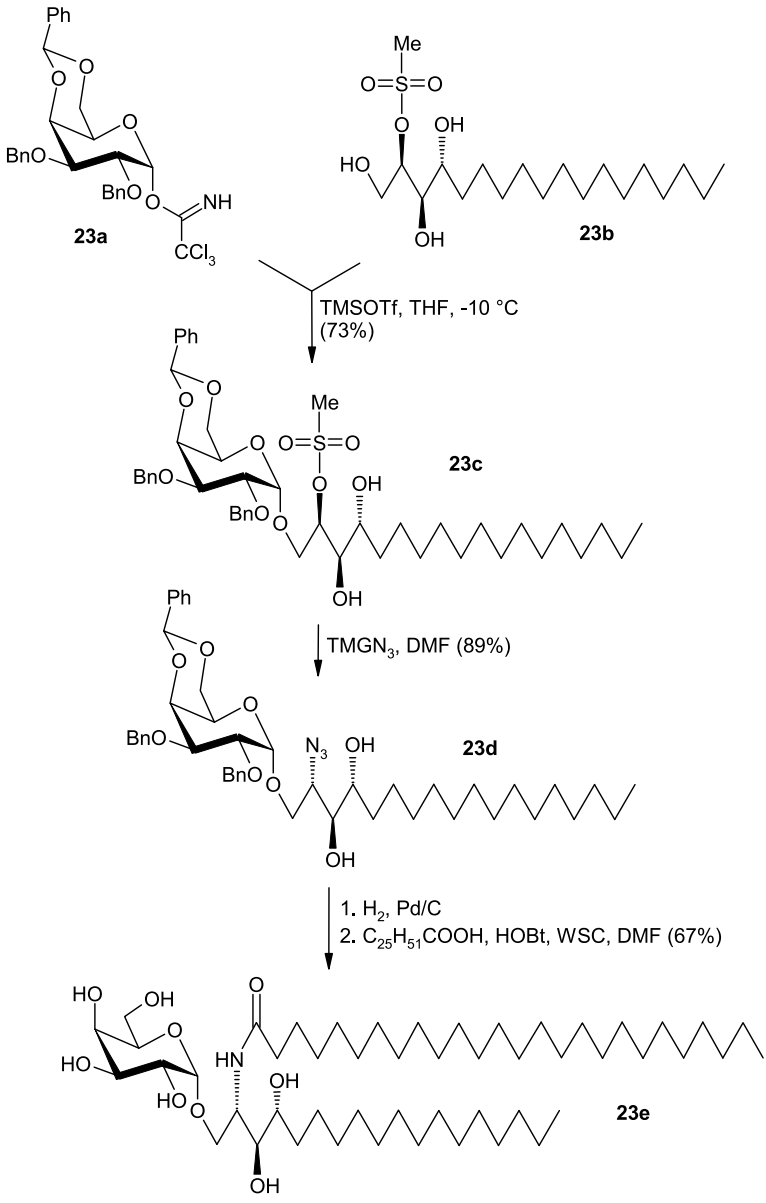
Scheme 21



Scheme 22

Reagents and conditions: (a) SAT(N), CMP-NANA, MOPS (pH 7.4), Triton S4, BSA, H_2O, CIP, 55%; (b) FucT(V), GDP-Fuc, $MnCl_2$, DTT, Tris-HCl (pH 7.4), qu

from building blocks **20a**, **20b**, and **20f** via intermediates **20c**, **d**, **e** protected Lewis X intermediate **20g** was obtained which gave monomer **20h**. For the dimerization, via **20i**, **j**, the target molecule **20k** was obtained. The synthesis of sialyl Lewis X containing glycolipids with different core structures [136] exhibited the importance of the spacer in selectin-binding



Scheme 23

studies [137]. Sulfated sialyl Lewis X variants, containing a lactamized sialyl residue, were obtained by the Kiso-Ishida group (● *Scheme 21*) [138]; these compounds were found to be potent antigenic determinants. Following the standard procedure **21a** and **21b** gave **21c**, which was transformed via **21d–h** into lactamized target molecules **21i** and **21j**. Also chemoenzymatic synthesis was successfully engaged in the synthesis of sulfated sialyl Lewis X connected to a core 1 mucin (T antigen) structure (● *Scheme 22*) [139]. From **22a**, **b** disaccharide **22c** was obtained, which was transformed into donor **22d** giving with acceptor **22e** tetrasaccharide **22f**. Partial deprotection (→ **22g**, **h**), regioselective sulfation and final deprotection gave **22i**; enzymatic sialylation and fucosylation led to target molecules **22j** and **22k**.

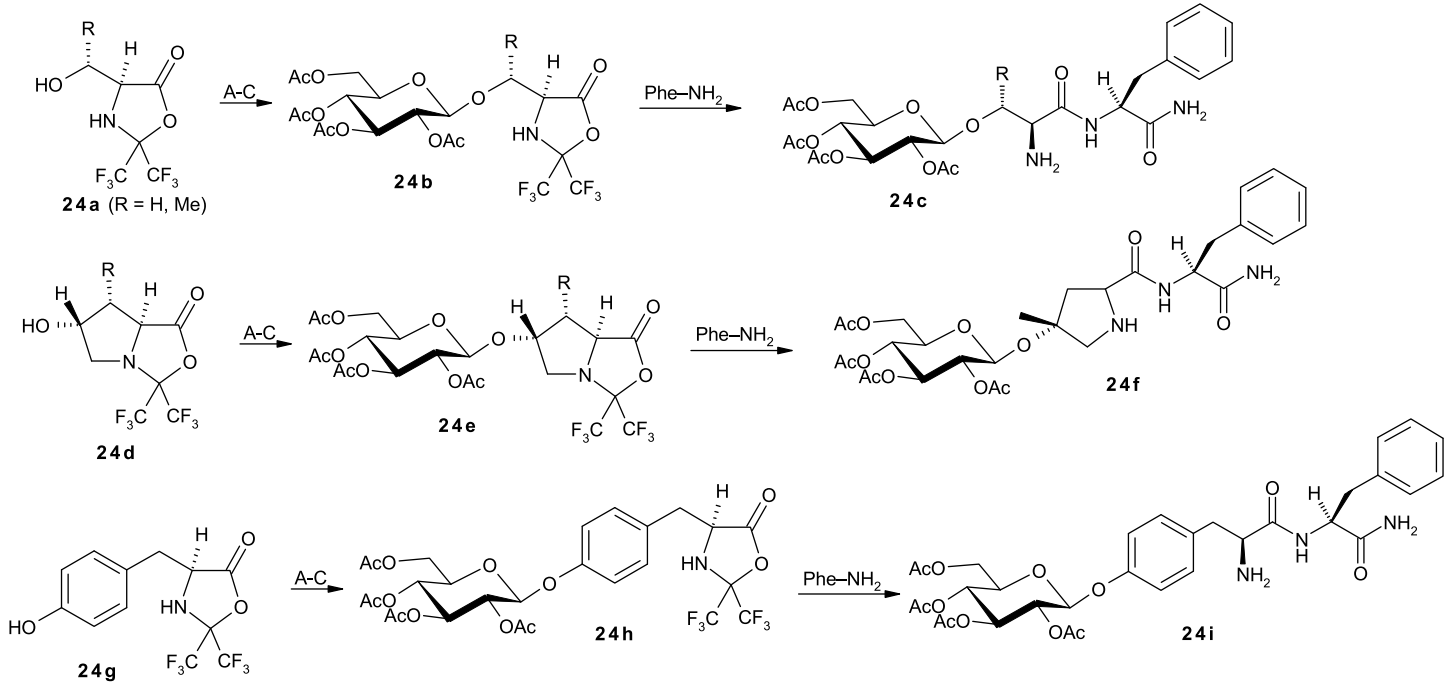
Some α-galactosyl ceramides with a phytosphingosine residue, isolated from the marine sponge *Agelas mauritianus*, were found to possess strong in vivo activities against several murine tumor cells [140,141]. The synthesis of analogs exhibited that compound **23e** (● *Scheme 23*) is a potent candidate for clinical trials [142]; **23e** was also found to have immunostimulatory activity [143]. These findings promoted a great demand for this compound, therefore, several syntheses of **23e** and analogs have been reported [144]. The Schmidt group reported an efficient synthesis which is based on galactosyl trichloroacetimidate **23a** as donor and phytosphingosine derivative **23b** as acceptor which led to the α-linked intermediate **23c**. Azide introduction (→ **23d**), hydrogenolysis of protecting groups and *N*-acylation furnished target molecule **23e** very efficiently.

Lipid-linked T and T_n antigens [145] and glycosidated phosphoglycerolipids [146] were also synthesized based on *O*-glycosyl trichloroacetimidates donors. In a comparative study glucopyranosylation of methyl ω-hydroxy-hexadecanoate with different glucosyl donors was carried out which demonstrated the advantageous properties of the trichloroacetimidate donor; this way the desired β-linked target molecule for biological studies was readily obtained [147]. 3,4,5-Tris(alkyloxy)benzyl glycosides were prepared with standard *O*-glycosyl trichloroacetimidates as donors [148]. The three lipid chains permitted the immobilization of these compounds on a hydrophobic surface and lectin affinity studies by surface plasmon resonance.

5.2 Glycosyl Amino Acids and Glycopeptides

The availability of *O*- or *N*-glycosyl amino acids as building blocks for glycopeptide or eventually glycoprotein synthesis is of great importance. Therefore, various synthetic methods have been reported [149,150]. Direct *O*-glycosylations of serine and threonine derivatives with *O*-glycosyl trichloroacetimidates as donors have been reported [151,152]. The straightforward hexafluoroacetone *O*,*N*-protection of Ser, Thr, Pyp, and Tyr gave generally the best results with trichloroacetimidate donors as shown in ● *Scheme 24* [153] on glucosylation of **24a**, **d**, **g** with per-*O*-acetylated glucopyranosyl donors affording **24b**, **e**, **h**. As this protection also leads to an activated carboxylate group, protecting group cleavage and peptide chain extension can be combined furnishing dipeptides **24c**, **f**, **i**.

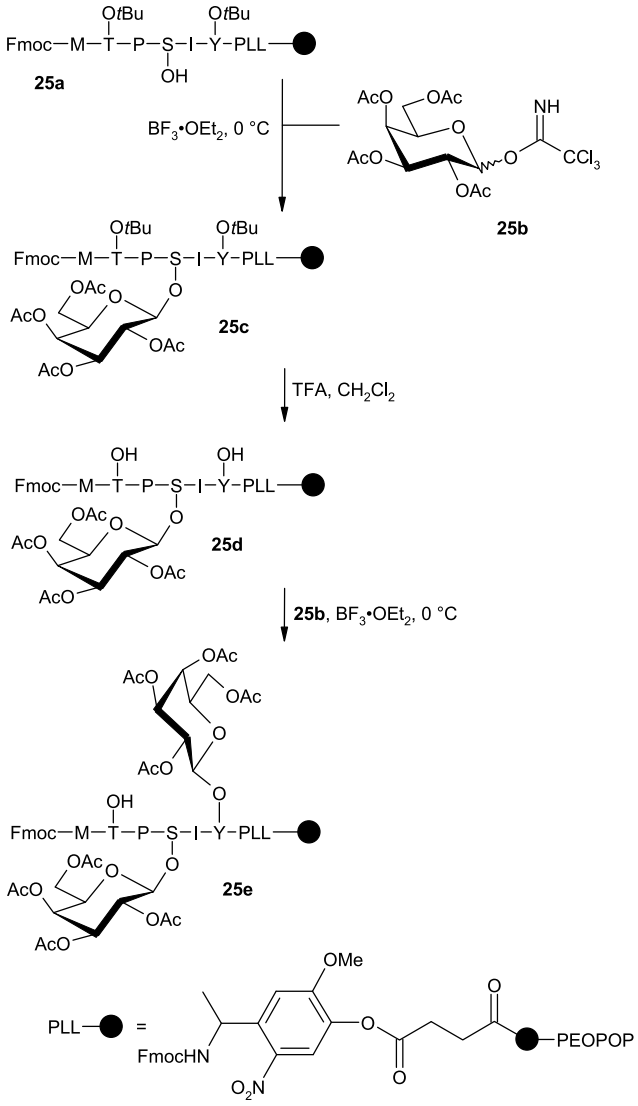
The presence of the Glcα(1–2)Galβ(1-O)Hyl moiety in collagen was reason to synthesize this building block [154]. Replacement of the ω-amino group of Hyl by an azido group and protection of the α-amino group by the Z group and the carboxylate group by a *t*-butyl group led cleanly to the target molecule after two glycosylation steps. Glycosylation of Fmoc-protected Ser, Hse, and Thr with T_N, T, and ST-derived glycosyl donors were also success-



Compound		Method[a]/Yield [%] A	B	C
24a (R = H)	HFA-Ser(Glc)	57	61	78
24a (R = Me)	HFA-Thr(Glc)	86	70	82
24d:	HFA-Hyp(Glc)	21	40	27
24g:	HFA-Tyr(Glc)	-	83	93

[a] Method A: Helferich variant of the Koenigs-Knorr reaction;
method B: Procedure according to Paulsen;
method C: Procedure according to Schmidt

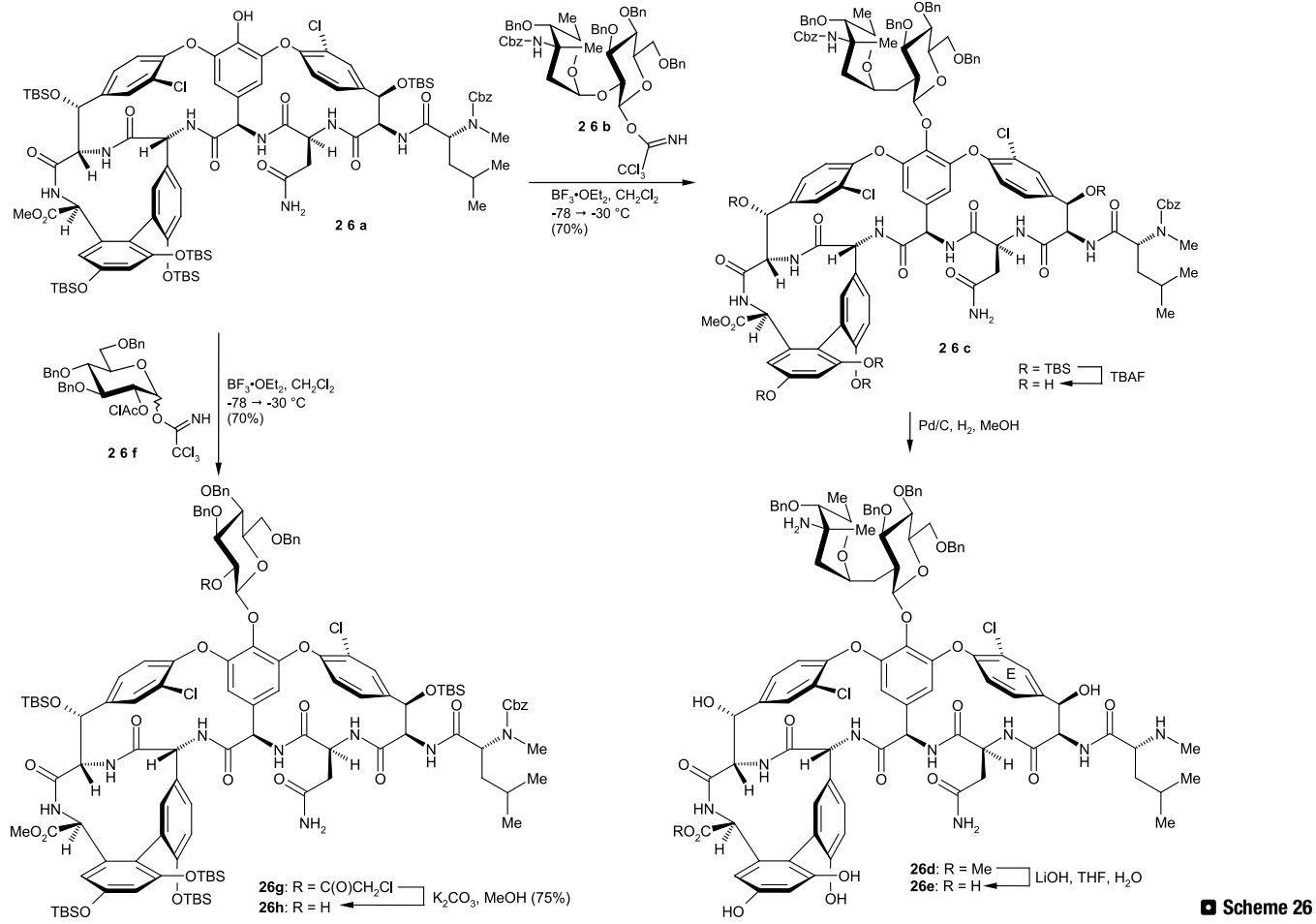
Scheme 24



Scheme 25

fully carried out, partly by following known procedures [155,156]. Also an important Galβ(1–4)GlcNAcβ(1–3)-L-Fuc moiety was prepared, which is part of *O*-linked chains of human clotting factor IX [157].

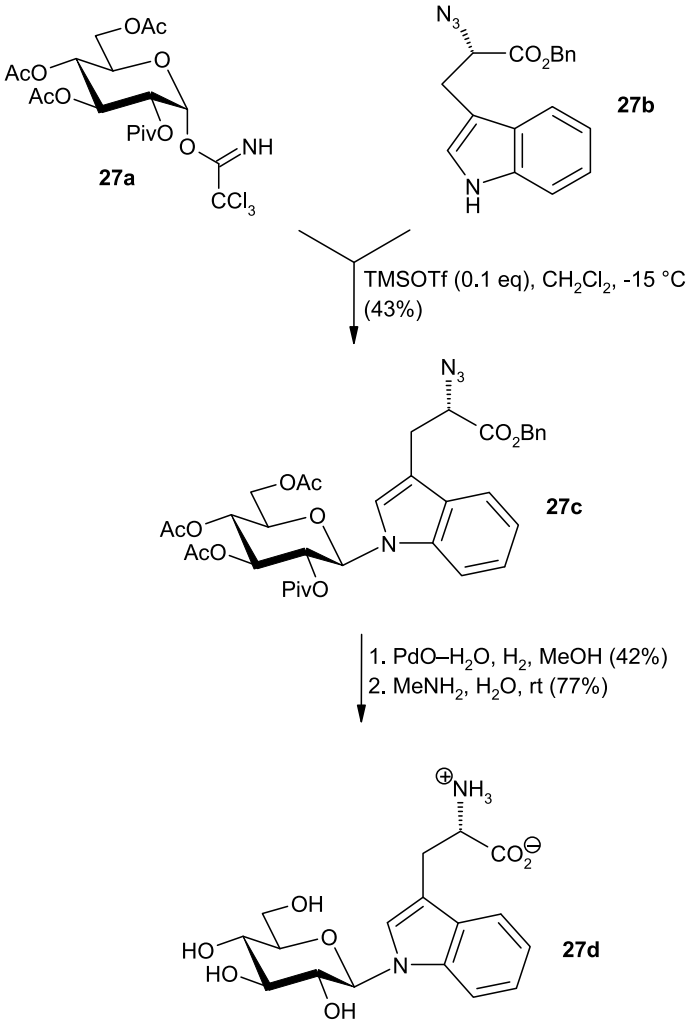
More demanding is direct glycosylation of peptides, which was so far not very successful because of solubility problems and side reactions with the functional groups [158,159,160]. The Meldal group expected that a solid-phase approach would suffer less from these drawbacks, therefore, they undertook a study with Ser, Thr, and Tyr containing hexapeptides, hav-


26 a
26 b
BF3•OEt2, CH2Cl2
-78 → -30 °C
(70%)
26 c
R = TBS
R = H
TBAF
Pd/C, H2, MeOH
26 f
BF3•OEt2, CH2Cl2
-78 → -30 °C
(70%)
26g: R = C(O)CH2Cl
26h: R = H
K2CO3, MeOH (75%)
26d: R = Me
26e: R = H
LiOH, THF, H2O

Scheme 26

ing no other functional side chains (!), leading to direct reaction in the order Tyr > Ser > Thr, as shown in Scheme 25 [161]. The peptide was linked via a photolabile linker (PLL) to the resin (→ **25a**). Galactosylation with **25b** (→ **25c**), then *O*-deprotection (→ **25d**), and galactosylation with **25b** gave **25e**; threonine did not undergo reaction in this case.

A particularly interesting case is the glycosylation of the vancomycin aglycone to achieve the vancomycin total synthesis as successfully investigated by the Nicolaou group [162,163]. They obtained very good results on glycosylation of aglycone **26a** with glycosyl donors **26b** and **26g** affording intermediates **26c** and **26g** (Scheme 26), which were further transformed, thus furnishing via **26d** target molecule **26e**. Similarly, vancomycin analogs were prepared by



Scheme 27

the Wong group [164]; biological studies exhibited growth inhibition of vancomycin sensitive bacteria by several of these compounds.

N-Glycopeptides are generally obtained by treatment of reducing sugars with ammonium bicarbonate [165] or reduction of glycosyl azides [166] with activated aspartic acid. This way a heptasaccharide synthesized with *O*-glycosyl trichloroacetimidates as donors was linked to Asp [167,168,169,170]; further work along these lines was carried out [171]. Direct formation of the *N*-glycosidic linkage with Asn by chemical glycosylation is still an important task. An interesting method was investigated by Ito et al. [172] based on hydroxyamination of the Asp side chain and then glycosylation with a glycosyl donor. The results show, for standard glycosyl donors direct *N*- glycosylation is possible, however so far yields and anomeric stereoselection are not yet satisfactory.

A number of *C*- and *N*-linked tryptophan glycoconjugates were discovered as constituents of natural products [173,174,175,176]. Also tryptophan *N*-glucoside **27d** has been detected for which Unverzagt et al. [177] reported a successful synthesis (*Scheme 27*). With the help of 2-*O*-pivaloyl protection in glucosyl donor **27a** the undesired acetal formation could be overcome; thus with acceptor **27b** *N*-glucoside **27c** could be obtained and then transformed into target molecule **27d**.

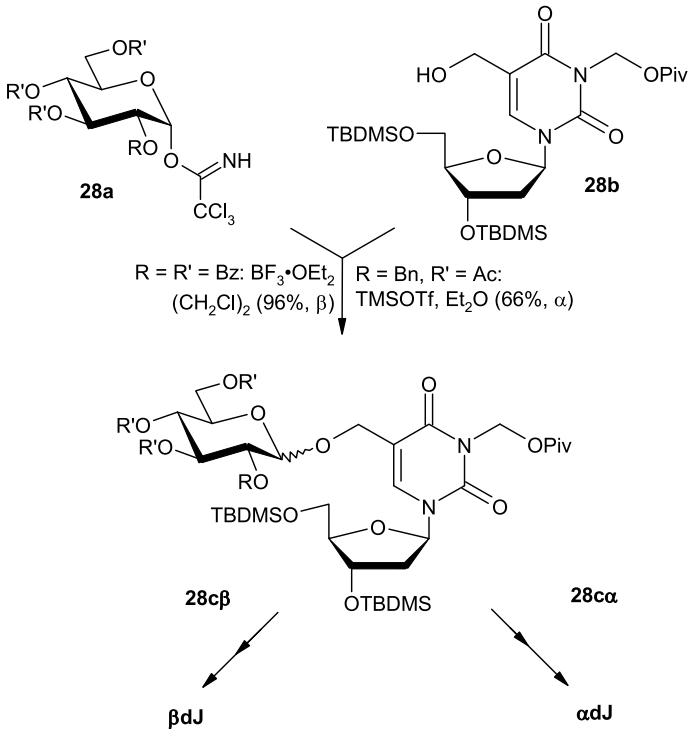
5.3 Nucleoside and Nucleotide Glycosidation

The presence of carbohydrate moieties in biomolecules influences many biological functions, thus also modulation of the functions of the aglycone. Therefore, studies have been undertaken to glycosylate nucleosides [178] and nucleotides [179]. Particularly interesting is the direct glycosylation of CPG-bound oligo-deoxynucleotides containing pyrimidine residues leading to reaction preferentially at the 5′-end.

The occurrence at 5-(β-D-glucopyranosyloxymethyl-2′-deoxyuridine (βdJ) in DNA, for instance of *Trypanosoma brucei* [180,181,182], led to the development of an efficient route for the synthesis of βdJ (*Scheme 28*) and its phosphoramidite as building blocks for DNA synthesis [183]. The synthesis was based on glucosyl donor **28a** (R = R′ = Bz) which gave with acceptor **28b** intermediate **28cβ** and after deprotection βdJ. Similarly, from **28a** (R = Bn, R′ = Ac) and **28b** α-linked **28cα** was obtained which led to αdJ.

5.4 Synthesis of Glycosaminoglycans

Glycosaminoglycans (GAGs) are bioactive oligosaccharides that are highly functionalized, linear, and anionically charged. Because of their complexity their chemical synthesis is a demanding task. The chondroitin sulfates (CSs) are a member of the GAG family. They are found in the extracellular matrix of connective tissues at the surface of many cells and in intracellular secretory granules [184,185]. They are linear heteroglycans built mainly from D-GlcA and GalNAc dimers which are β(1–3)-linked; chain extension employs a β(1–4)-linkage. In addition, the sugar residues are sulfated at various positions (→ A- E- and K-type CSs). Stereocontrolled shark cartilage CS (D-type) synthesis has been successfully carried out by Jacquinet et al. (*Scheme 29*) [186,187]. The GlcA building block **29a** possesses as temporary protecting group for chain extension at 4-*O* the chloroacetyl group, and for sulfation at 2-*O* the benzoyl

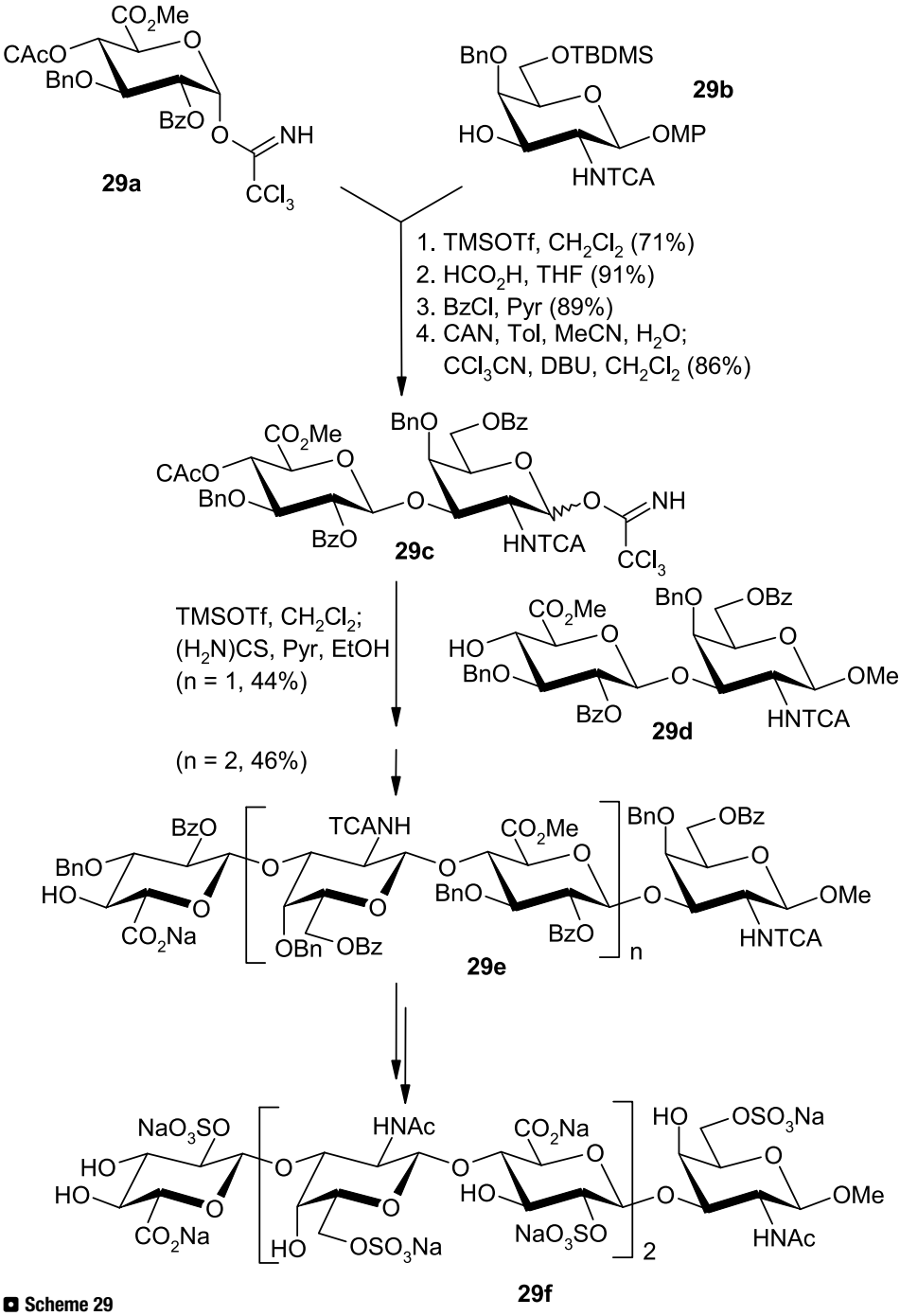


Scheme 28

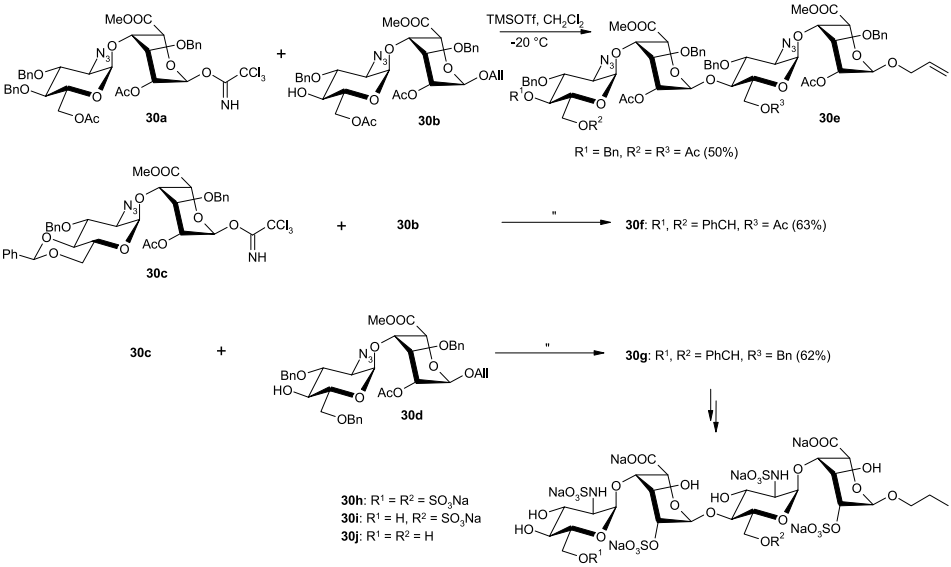
group. Similarly, the galactosamine residue **29b** carries for chain extension at 1-*O* an MP group and for sulfation at *O*-6 the TBDMS group; the amino group is protected by the trichloroacetyl group. Glycosylation and further transformations furnish donor **29c** which on chain extension with **29d** led to the tetrasaccharide **29e** (n = 1) and higher oligomers which could be deprotected to yield, for instance, target molecule **29f**. On the basis of a similar strategy E-type CS was obtained which was found to stimulate neuronal outgrowth [188]. A hexameric E-type CS was also synthesized by Tamura et al. [189]; in this synthetic approach the amino group of the galactosamine residue was replaced by an azido group.

Hyaluronic acids possess a repeating unit consisting of GlcA β(1–3)-linked to GlcNAc; chain extension of this dimer employs a β(1–4)-linkage. The dimeric unit has been successfully synthesized based on GlcA trichloroacetimidate as the donor [190]. A trimer consisting of two GlcNAc residues and one GlcA residue was also successfully obtained [191]. In this synthesis design the carboxylate function was introduced at a late stage, after construction of the trisaccharide backbone.

Fibroblast growth factors (FGFs) display high binding affinities for GAGs [185,192,193,194,195] such as heparan sulfate (HS) and heparin. Heparin is a linear, heterogeneously sulfated, anionic polysaccharide composed of alternating L-iduronic acid and D-glucosamine residues which is found in almost all animal tissues. The biological roles of several members of the FGFs


CO2Me
CAcO
BnO
BzO
O
NH
CCl3
29a
BnO
OTBDMS
O
29b
HO
OMP
HNTCA
1. TMSOTf, CH2Cl2 (71%)
2. HCO2H, THF (91%)
3. BzCl, Pyr (89%)
4. CAN, Tol, MeCN, H2O;
CCl3CN, DBU, CH2Cl2 (86%)
BnO
OBz
29c
TMSOTf, CH2Cl2;
(H2N)CS, Pyr, EtOH
(n = 1, 44%)
(n = 2, 46%)
29d
OMe
HNTCA
BzO
TCANH
OBn
29e
n
NaO3SO
HNAc
CO2Na
OSO3Na
29f
2

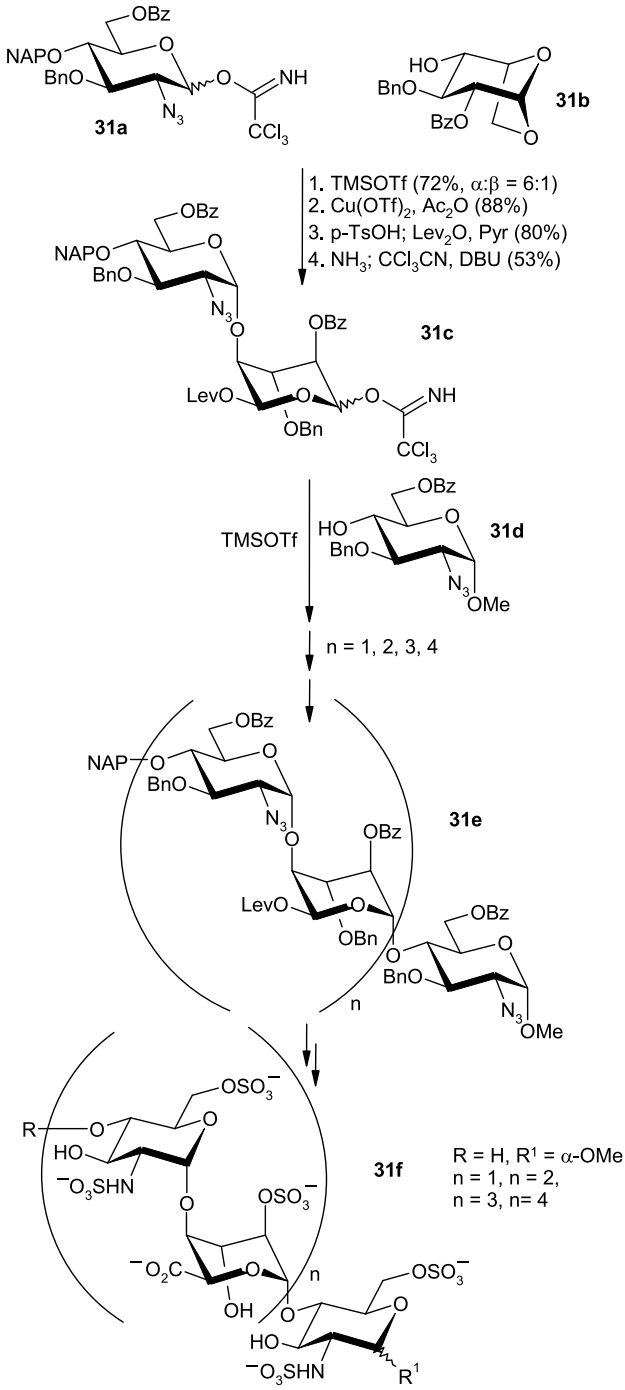
Scheme 29



Scheme 30

have been extensively investigated [196,197], which also called for synthetic endeavors; they are still ongoing. The synthesis of a highly sulfated tetrasaccharide has been reported by Lay et al. (*Scheme 30*) [198]. It is based on a versatile α(1–4)-linked GalN_3-IdoA disaccharide building block from which four differently protected disaccharide acceptors and donors (**30a–d**) are obtained, thus leading to **30e**, **f**, **g**. The deprotection scheme is shown for **30g**, leading to target molecules **30h**, **i**, **j**. An interesting study was also reported by Bonnaffé et al. [199]; for successful glycosylation a remote *N*-acetyl group had to be replaced by an azido group. Tri-, penta-, hepta-, and nonasaccharides of heparin have been synthesized by Hung et al. (*Scheme 31*) [200]. In this synthetic approach, the carboxylate function is again introduced at a late stage by oxidation of the hydroxymethyl group of the idose moiety. Glycosylation of Ido derivative **31b** with donor **31a** led via further transformations to disaccharide donor **31c** which was chain extended with glucosamine acceptor **31d** to yield **31e** with n = 1. Further chain extension is based on **31c** having naphthylmethyl (NAP) temporary protection; chemoselective NAP cleavage is readily performed with DDQ. Thus via the higher oligomers of **31e** after deprotection target molecules **31f** (n = 1, 2, 3, 4) are obtained. Heparin-like oligosaccharide syntheses were also reported by Martin-Lomas et al. [201], by Seeberger et al. [202], and by Yu et al. [203]. The low reactivity of the axial 4-hydroxy group in the iduronate residue in the Seeberger approach led to trichloroacetimidate rearrangement because the donor and acceptor reactivity did not match.

Proteoglycans are other important glycoconjugates with various roles [204], such as lubrication and blood anticoagulation. Several proteoglycans possess a highly conserved tetrasaccharide linkage region joining a GAG to a core protein. This GlcAβ(1–3)Galβ(1–3)Galβ(1–4)Xylβ(1-O)Me tetrasaccharide was also successfully synthesized based on the *trichloroacetimidate method* [205].


OBz
NAPO
BnO
O
O
NH
N3
CCl3
31a
HO
BnO
BzO
O
O
31b
1. TMSOTf (72%, α:β = 6:1)
2. Cu(OTf)2, Ac2O (88%)
3. p-TsOH; Lev2O, Pyr (80%)
4. NH3; CCl3CN, DBU (53%)
31c
LevO
OBn
TMSOTf
31d
OMe
n = 1, 2, 3, 4
NAP
31e
OSO3−
R
HO
−O3SHN
31f
R = H, R1 = α-OMe
n = 1, n = 2,
n = 3, n= 4
−O2C
OH
R1

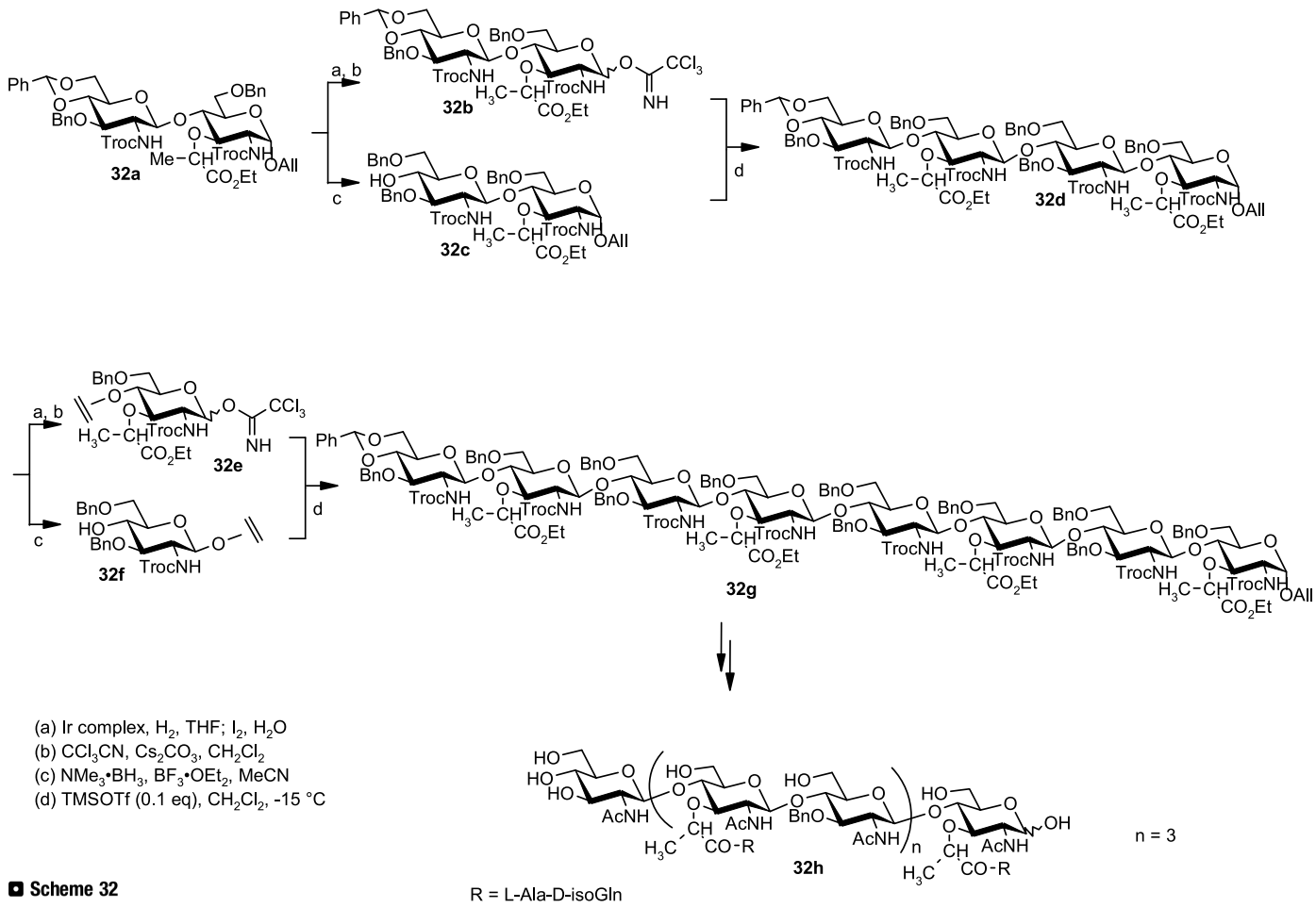
Scheme 31

5.5 Cell Wall Constituents

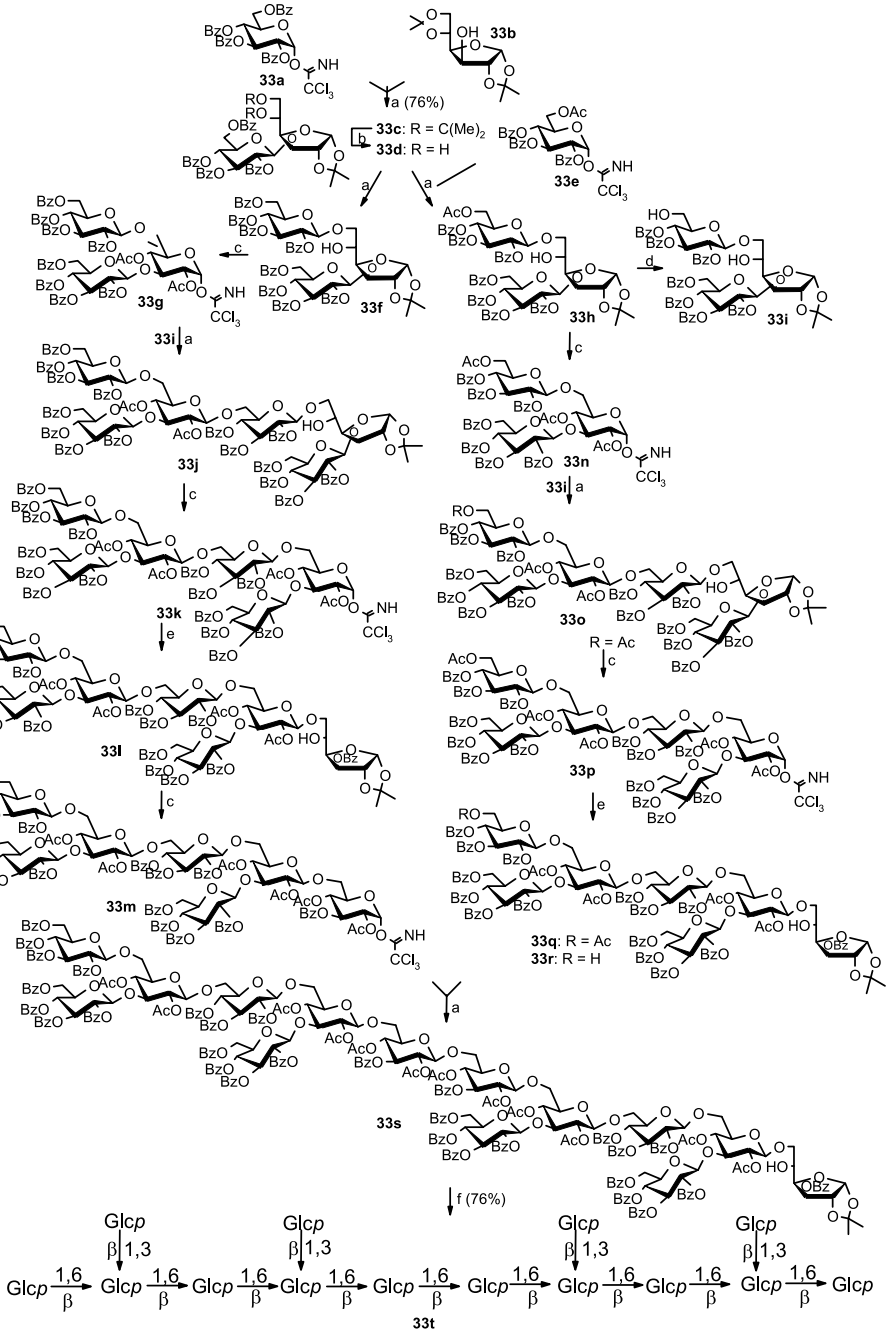
Bacterial cell wall peptidoglycan is known as a strong immunopotentiator which induces various mediators such as cytokines, prostaglandins, platelet activation factor, and NO, thus stimulating the immune system [206]. The receptor for peptidoglycans was shown to be TLR2 [207], the same as for lipoteichoic acids (LTAs) and lipoproteins. Therefore, the synthesis of peptidoglycan fragments is of continuing interest. Recently up to an octasaccharide fragment was obtained as shown in ❯ *Scheme 32* [208,209]. Starting material **32a** was transformed into donor **32b** and acceptor **32c** based on the temporary protection at *O*-1 with allyl and at *O*-4′ with benzylidene. Glycosylation afforded tetrasaccharide **32d** which by similar transformation to **32e** and **32f** led to octasaccharide **32g** which could be deprotected to yield target molecule **32h**. This procedure of chain extension could be even extended to the hexadecasaccharide, however the Troc deprotection of this compound failed so far. To the octasaccharide **32h** also small peptides were attached via the lactate residue and the biological activity of the derived compounds was tested.

Many bacterial and fungal cell walls contain homo- and heteroglycans. The synthesis of fragments of these cell wall constituents has gained increasing interest in the last years. The Manα(1–3)Manα(1–2)Manα(1–6)Manα(1–2)Man heptasaccharide obtained by mild acetolysis from *C. glabrata* IFO 0622 has been successfully synthesized by Kong et al. [210]; they synthesized also a related sulfated pentasaccharide [211]. Also mannan repeating units of *Trichophytan mentagrophytes*, *T. rubrum* [212], *Saccharomyces cerevisiae* X 2180–1A [213], and *Candida kefyr* IFO 0586 [214] have been prepared. Another important class of glycans are β(1–3)- and β(1–6)-linked glucans which are frequently branched. Glucans consisting of β(1–3)-linked glucose residues having β(1–6)-linked branches have been found in many plants and fungi [215,216]. Since these glucans show antitumor activity, the synthesis of at least minimal structural units for biological activity studies has gained great interest [217]. Recently, several successful syntheses of such molecules have been reported mainly by Kong et al. [218,219,220,221,222,223]. Also a 2-branched β(1–3)-glucan was found and dimers of the trisaccharide repeating unit were prepared [224]. 3-Branched β(1–6)-linked glucans were found to be phytoalexin elicitors; they possess antitumor activities as well. A highly efficient synthesis of tetradecasaccharide **33t** was reported again by the Kong group [225] (❯ *Scheme 33*). The synthesis is based on readily available glucose-derived building blocks **33a**, **33b**, and **33e** and leads via disaccharides **33c** and **33d**, trisaccharides **33f**–**33i**, and **33n**, hexasaccharides **33j**, **k**, **o**, **p**, heptasaccharides **33l**, **m**, **q**, **r** to tetradecasaccharide **33s** which was deprotected to furnish target molecule **33t** in good overall yield. Such compounds were also prepared in large scale and with different aglycones [226,227,228].

Mannoglucan from *Microellabosperia grisea* possesses also high antitumor activity, therefore the repeating unit $\{[\text{Glc}\alpha(1\text{–}3)][\text{Glc}\alpha(1\text{–}6)]\text{Glc}\beta(1\text{–}4)\text{Glc}\beta(1\text{–}4)\}_n$ has been synthesized via an efficient route [229]. β(1–6)-Linked galactofuranosyl oligosaccharides are constituents of the cell wall of bacteria and fungi including some clinically significant pathogens [230,231,232]. The highly immunogenic arabinogalactans contain arabinofuranosyl and galactofuranosyl residues. β(1–6)-Linked galactofuranose oligomers, also found in the cell wall of the fungus *Fusarium*, exhibited in plants elicitor activity. Therefore, a β(1–6)-linked *O*-galactofuranosyl hexasaccharide was prepared based on the *trichloroacetimidate method* in high yield [233,234]. The Kong group also synthesized α(1–5)-linked L-arabinofu-

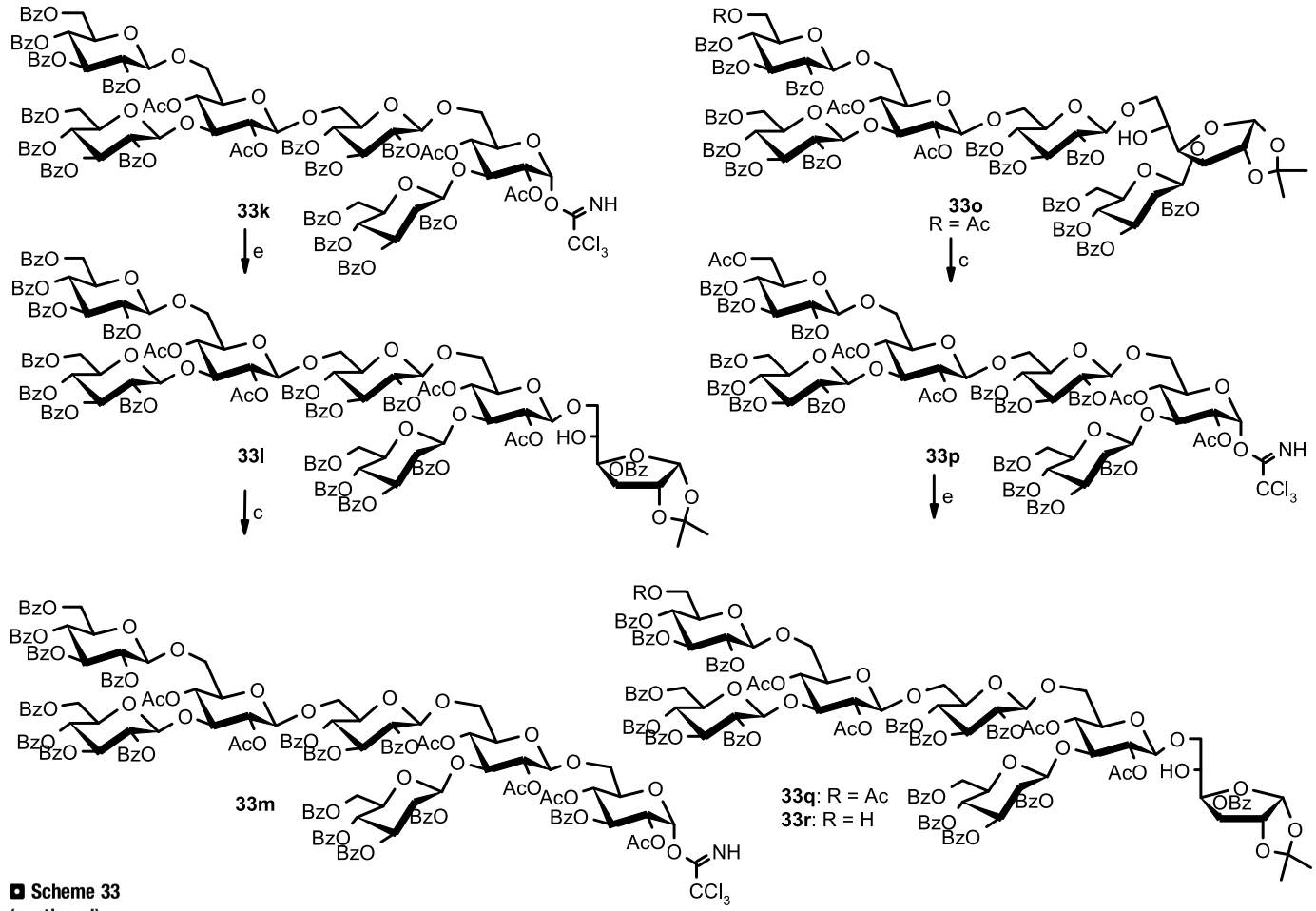

32a
32b
32c
32d
32e
32f
32g
32h
a, b
c
d
(a) Ir complex, H_2, THF; I_2, H_2O
(b) CCl_3CN, Cs_2CO_3, CH_2Cl_2
(c) $NMe_3{\bullet}BH_3$, $BF_3{\bullet}OEt_2$, MeCN
(d) TMSOTf (0.1 eq), CH_2Cl_2, -15 °C
R = L-Ala-D-isoGln
n = 3

Scheme 32

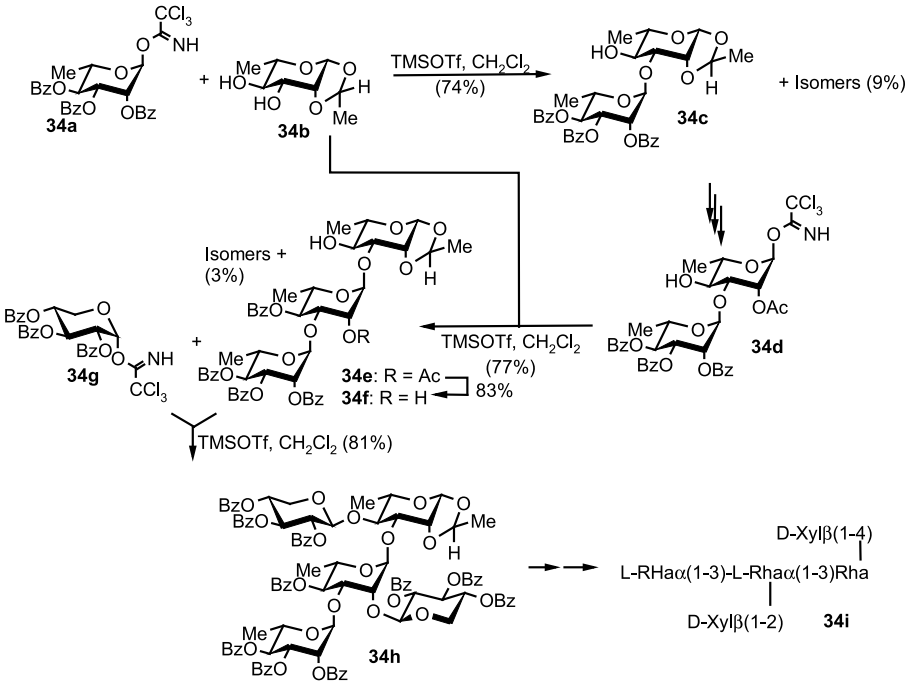


(a) TMSOTf (0.01 eq), CH_2Cl_2, rt; (b) HOAc, H_2O, 40 °C (qu)
(c) HOAc, H_2O, refl.; Ac_2O, Pyr, rt; NH_3, THF, MeOH; CCl_3CN, K_2CO_3, CH_2Cl_2;
(d) HCl, CH_2Cl_2, MeOH; (e) 3-O-Bz-1,2-O-Ip-glucofuranose, TMSOTf, CH_2Cl_2, rt;
(f) HOAc, H_2O, refl.; NH_3, MeOH, CH_2Cl_2.

Scheme 33



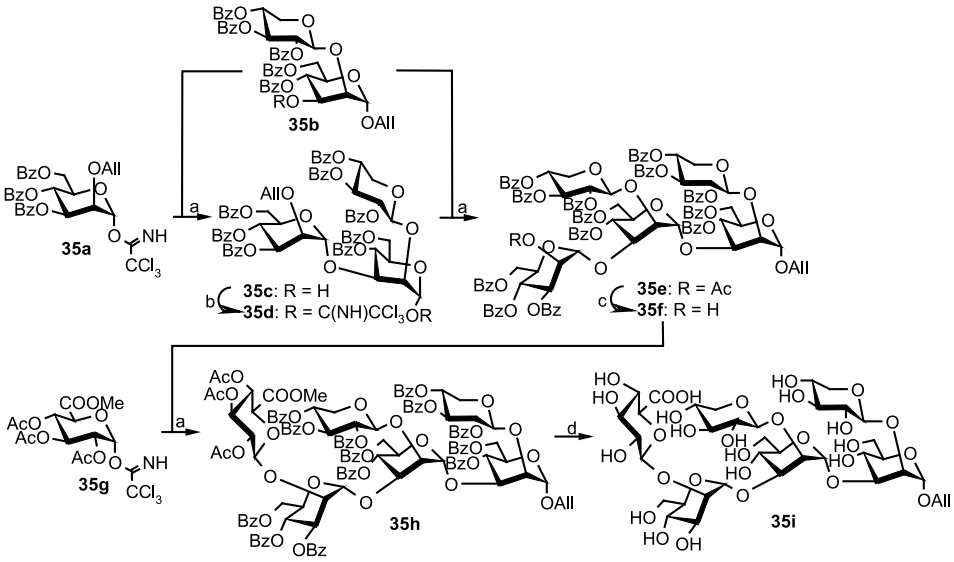
Scheme 33
(continued)



Scheme 34

ranosyl oligosaccharides up to the octamer based on related methodologies [235]. 2-*O*-Arabinofuranosylated β(1–6)-linked galactopyranosyl oligosaccharides, found in arabiogalactans, could be obtained also by Kong et al. [236] who also reported the successful synthesis of other arabinogalactan linkage types based on the *trichloroacetimidate method* [237,238,239,240]. Lipoarabinomannans have attracted great interest as well, because they are part of the cell surface oligosaccharide of mycobacterial species that cause many diseases, including tuberculosis and leprosy. The largest heteroglycan synthesis for this type of compound was recently reported in which the *trichloroacetimidate method* plays an important role [241].

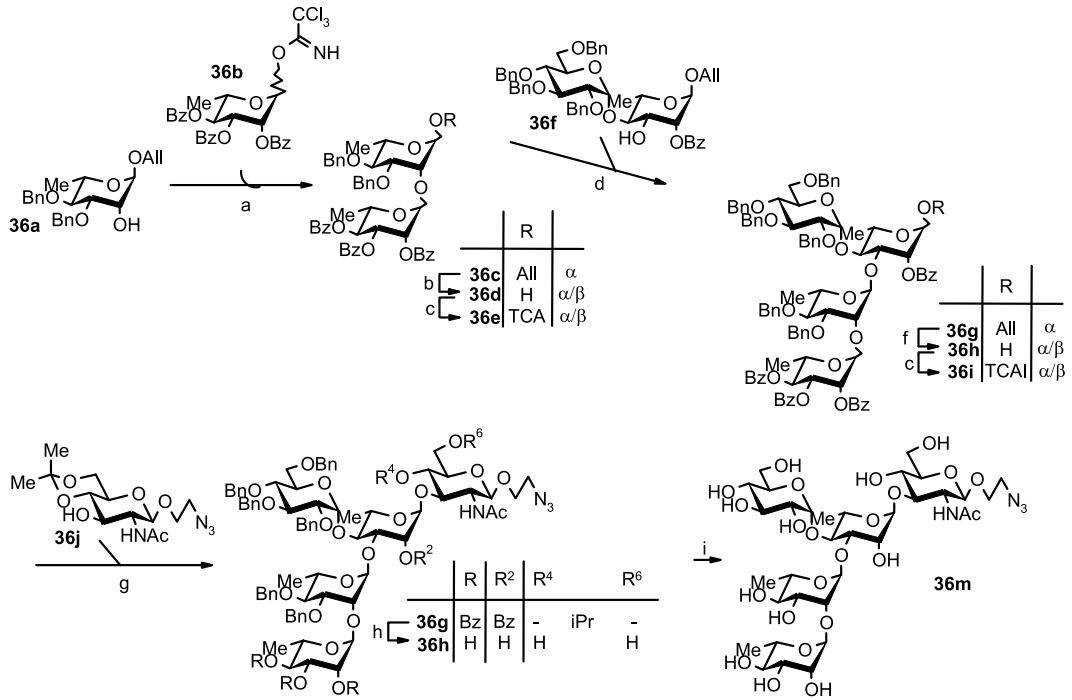
In algal cell wall polysaccharides β(1–3)-linked xylans were found. Fragments up to the hexamer were successfully synthesized [242]. The specific *O*-chain of the lipopolysaccharide of many bacteria contains repeating units composed of various sugar residues. Several of these repeating units or fragments thereof have been synthesized based on the *trichloroacetimidate method*. Rhamnosylated rhamnan [243], the D-Rha*p*α(1–3)[L-Xyl*p*β(1–2)]D-Rha*p*α(1–3)[L-Xyl*p*β(1–2)]D-Rha unit [244], (*Scheme 34*, **34i**), a more complex xylorhamnan [245], a xylosylated GlcNAc rhamnan [246], a rhamnan with GlcNAc in the branch [247,248], the pentasaccharide D-Glcβ(1–2)-[D-Rib*f*β(1–3)]L-Rha*p*α(1–3)L-Rha*p*α(1–3)L-Rha*p*α(1–2)L-Rha*p* [249], the glucurono xylomannan hexasaccharide repeating unit of *C. neoformans* serotype A (*Scheme 35*) [250], a heptasaccharide fragment of *C. neoformans* serotype C [251], a repeating unit of *lactosillan* [252], the *Shigella flexneri* serotype 2a [253,254,255,256,257], and serotype 5a [258] have been efficiently obtained (*Scheme 36*). The straight forward syn-



(a) TMSOTf, CH_2Cl_2, -10 °C → rt (dry); (b) $PdCl_2$, 90% acetic acid-NaOAc, rt, 12 h; then GI_3CN, DBU, CH_2Cl_2, 2 h; (c) 2% CH_3COCl–CH_3OH, 0 °C → rt; (d) sat. NH_3–MeOH, rt, 36 h; then H_2O, rt, 5 h.

Scheme 35

thesis of heptasaccharide **34i** is outlined in *Scheme 34* [244]. Rhamnosyl donor **34a** readily reacts with rhamnose acceptor **34b** to give mainly disaccharide **34c** which is transformed into donor **34d**. Chain extension with **34b** leads to **34e** and then to acceptor **34f** which on reaction with xylosyl donor **34g** furnishes pentasaccharide **34h**; deprotection leads to target molecule **34i**. The hexasaccharide repeating unit synthesis of *C. neoformans* type A is outlined in *Scheme 35* [29]. Mannosyl donor **35a** reacts with acceptor **35b** to afford trisaccharide **35c** leading to donor **35d** which on reaction with **35b** gives pentasaccharide **35e** and acceptor **35f**. Glucuronidation with donor **35g** furnished hexasaccharide **35h** which gave on deprotection target molecule **35i**. The *Shigella flexneri* serotype 2a pentasaccharide fragment is outlined in *Scheme 36* [36]. Rhamnosyl donor **36b** and acceptor **36a** yield disaccharide **36c** which is transformed via **36d** into donor **36e**. On reaction with acceptor **36f** tetrasaccharide **36g** is obtained which is again transformed via **36h** into donor **36i**. Reaction with acceptor **36j** pentasaccharide **36k** is obtained which led to target molecule **36m** having an aminoethyl group at *O*-1 for glycoconjugate synthesis. Linkage to the PADRE sequence, acting as a universal T-cell epitope, was successfully performed and immunogenicity studies were carried out. Also successful investigations towards the synthesis of a tetrasaccharide rhamnogalacturonan related to an antiulcer pectic polysaccharide have been reported [259]. The *Neisseria* lipooligosaccharide contains two heptose residues within the conserved core structure; one heptose residue is 3,4-branched. On the basis of the *trichloroacetimidate method* a successful synthesis of a branched tetrasaccharide unit was successfully performed by Yamasaki et al. [260,261].



(a) cat. TMSOTf, Et_2O, -70 °C → rt, 8 h; (b) i. cat. $[Ir(COD)\{PCH_3(C_6H_5)_2\}_2]^+PF6^-$, THF, rt, 16 h, ii. HgO, $HgBr_2$, acetone/water, rt, 1h; (c) CCl_3CN, DBU, CH_2Cl_2, rt, 2 h; (d) cat. TMSOTf, Et_2O, -60 - -30 °C, 2 h; (f) i. cat $[Ir(COD)\{PCH_3(C_6H_5)_2\}_2]^+PF6^-$, THF, rt, 16 h, ii. HgO, $HgCl_2$, acetone/water, rt, 1h; (g) cat. TMSOTf, 4 Å-MS, CH_2Cl_2, rt, 3 h; (h) i. 50% aq. TFA, CH2Cl2, 0 °C, 2 h; ii. cat. MeONa, MeOH, 55 °C, 2 h; (i) 10% Pd/C, 1 M aq. HCl, EtOH/EtOAc, rt, 2 h.

Scheme 36

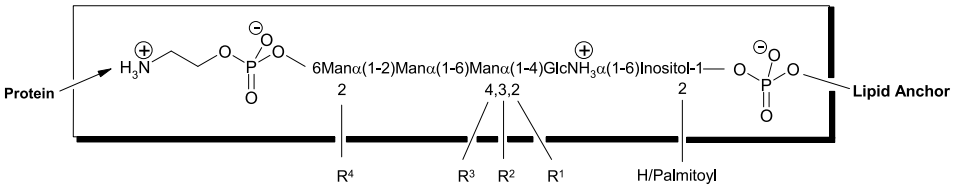
5.6 Synthesis of Glycosylphosphatidyl Inositol Anchors

Glycosylphosphatidyl inositol anchors constitute a class of glycolipids that link proteins and glycoproteins via their *C*-terminus to eukaryotic cell membranes. The first structure of a GPI anchor, that of *Trypanosoma brucei*, was published by Ferguson et al. [262]. Since then quite a few examples of GPI anchors were described, allowing the definition of the core structure depicted in ❯ *Scheme 37* [263].

The diversity within GPI anchors is mainly reflected in the location and nature of the branching groups of the glycan residue (R2, R3, R4). Additional ethanolamine phosphates (R1) seem to be specific for higher eukaryotes [264]. Concerning the lipid residue, many of the structures of GPI anchors contain a diacylglycerol moiety but alkylacylglycerol residues are not uncommon and ceramide structures have also been identified [263]. These modifications of the evolutionary conserved structure give rise to species-, stage-, and tissue-specific GPI structures.

The function of GPI anchors has been extensively discussed. A controversial aspect of GPI anchors is their ability to mediate signaling mechanisms or to function as second messengers, e. g. in insulin-mediated signal transduction processes [265]. Therefore, to perform biological studies elucidating the functions of GPI anchors, it seems to be an important objective to have access to structurally homogeneous GPI anchors and their derivatives. For the total synthesis of GPI anchors, a combination of lipid, phosphate, and oligosaccharide chemistry is required. A highly versatile strategy has been successfully followed for a ceramide-containing GPI anchor of yeast [266,267]. Similarly obtained were the acylglycerol-containing GPI anchors of *Trypanosoma brucei* [268,269] and *rat brain* Thy-1 [270].

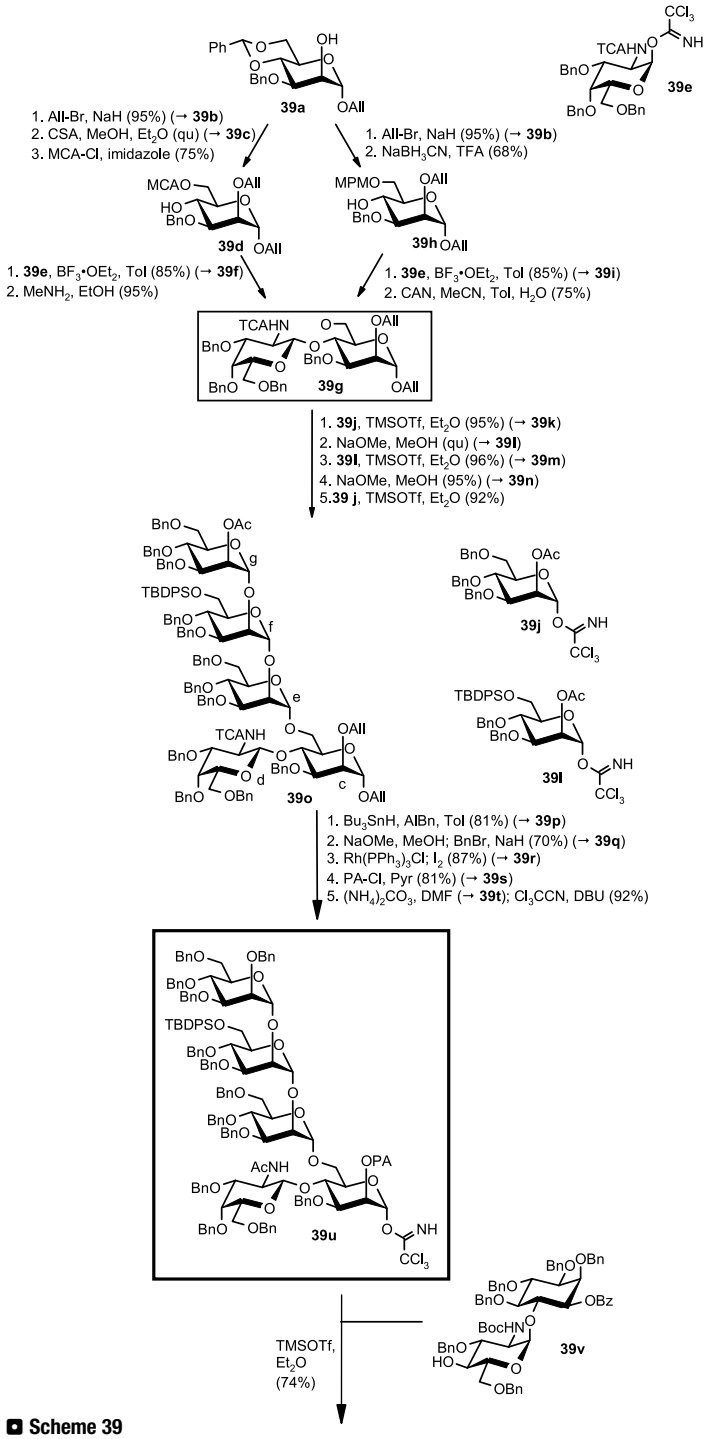
On the basis of earlier work [266,267], the development of a highly variable synthetic strategy, which is also applicable to the preparation of branched GPI anchors, was reported by Schmidt et al. [271,272]. This strategy allows also for the attachment of peptide or protein residues to the GPI anchor. The focus was on 4,6-branched mannose residues as there are sev-



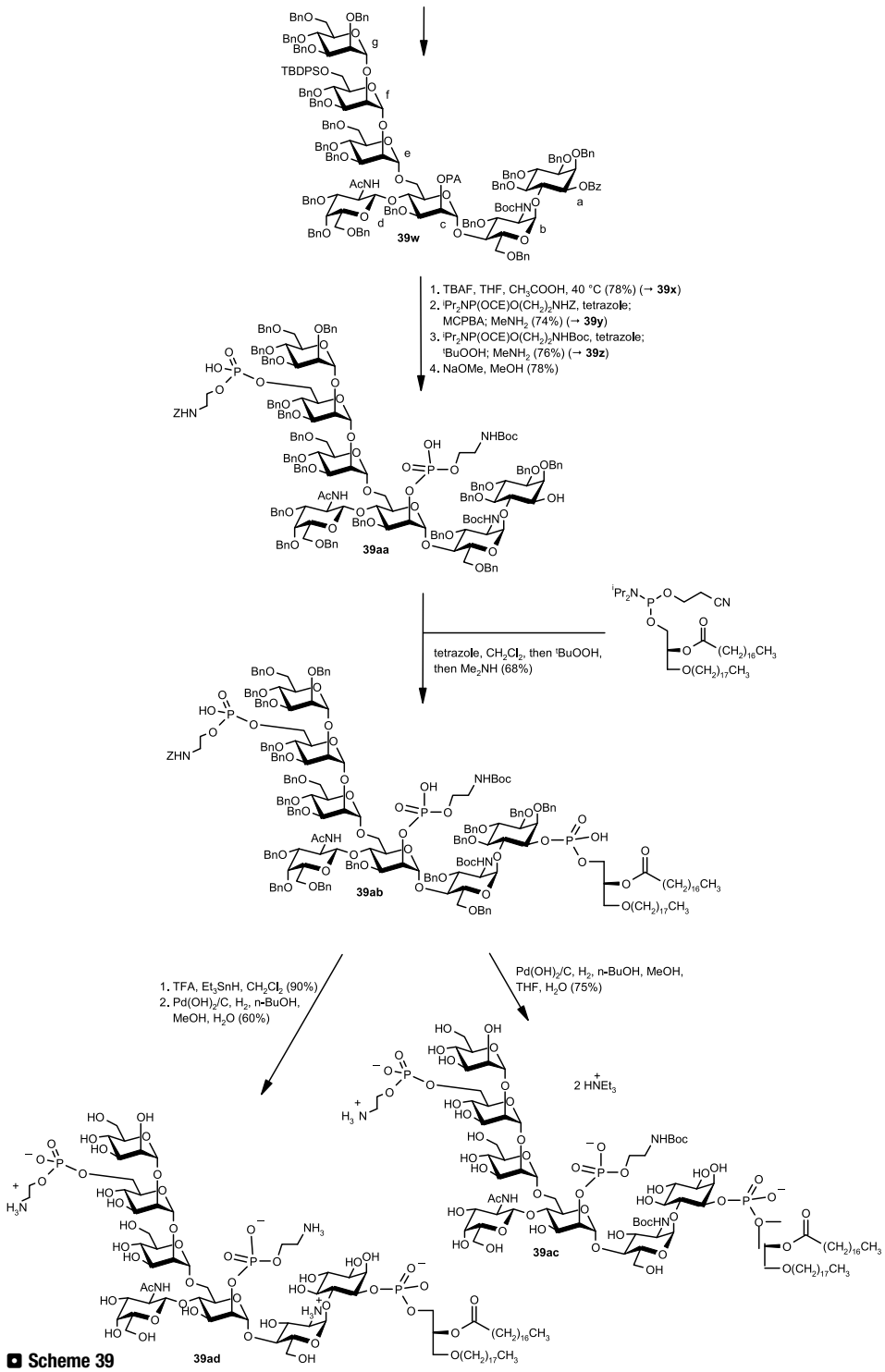
Natural Source	R^4	R^3	R^2	R^1	Lipid
S. cerevisiae	Manα(1-2)	H	H	H	Ceramide/DAG
T. brucei VSG	H	H	$Gal_{2\text{-}4}$α(1-3)	H	DAG
T. gondii A	H	GalNACβ(1-4)	H	H	DAG
B	H	Glcα(1-4)GalNAcβ(1-4)	H	H	DAG
Rat brain *Thy-1*	Manα(1-2)	GalNAcβ(1-4)	H	EA-P	Acylalkylglycerol

Scheme 37
General structure of GPI anchors (EA, ethanolamine; P, phosphate; DAG, diacylglycerol)

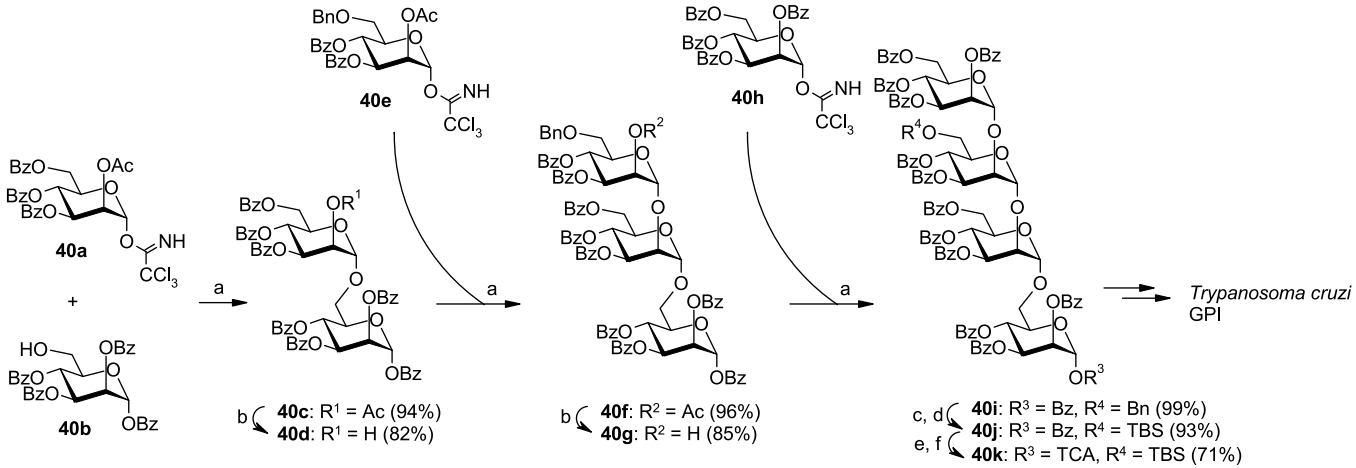
Scheme 38


39a
39e
1. All-Br, NaH (95%) (→ 39b)
2. CSA, MeOH, Et2O (qu) (→ 39c)
3. MCA-Cl, imidazole (75%)
1. All-Br, NaH (95%) (→ 39b)
2. NaBH3CN, TFA (68%)
39d
39h
1. 39e, BF3•OEt2, Tol (85%) (→ 39f)
2. MeNH2, EtOH (95%)
1. 39e, BF3•OEt2, Tol (85%) (→ 39i)
2. CAN, MeCN, Tol, H2O (75%)
39g
1. 39j, TMSOTf, Et2O (95%) (→ 39k)
2. NaOMe, MeOH (qu) (→ 39l)
3. 39l, TMSOTf, Et2O (96%) (→ 39m)
4. NaOMe, MeOH (95%) (→ 39n)
5.39 j, TMSOTf, Et2O (92%)
39j
39l
39o
1. Bu3SnH, AIBn, Tol (81%) (→ 39p)
2. NaOMe, MeOH; BnBr, NaH (70%) (→ 39q)
3. Rh(PPh3)3Cl; I2 (87%) (→ 39r)
4. PA-Cl, Pyr (81%) (→ 39s)
5. (NH4)2CO3, DMF (→ 39t); Cl3CCN, DBU (92%)
39u
TMSOTf,
Et2O
(74%)
39v

Scheme 39



Scheme 39 (continued)



Reagents: (a) $TMSOSO_2CF_3$, CH_2Cl_2; (b) 2% HCl, $MeOH/CH_2Cl_2$; (c) H_2, $Pd(OH)_2/C$, THF; (d) $TBSOSO_2CF_3$, Et_3N, CH_2Cl_2; (e) $NH_2CH_2CH_2NH_2 \cdot HOAc$, THF; (f) CCl_3CN, Cs_2CO_3, CH_2Cl_2. Bn = benzyl, THF = tetrahydrofuran, TMS = trimethylsilyl.

Scheme 40

eral prominent examples in nature (◉ *Scheme 37*). Therefore, the fully phosphorylated GPI anchor of *Toxoplasma gondii* was prepared [271]. Also a new route to rat-brain Thy-1 [272] was developed which is outlined in ◉ *Scheme 38*, showing important disconnections and building blocks; it is totally based on trichloroacetimidate donors. Two efficient routes starting from mannosyl acceptor **39a** and glucosamine-derived donor **39e** leading to disaccharide **39g** were developed (◉ *Scheme 39*). With variously protected mannosyl donors **39j** and **39l**, chain extension to pentasaccharide **39o** could be efficiently performed. Transformation into donor **39n** and reaction with disaccharide acceptor **39v** led to pseudoheptasaccharide **39w**. Introduction of two orthogonally protected aminoethyl phosphate residues led to **39aa** to which the phospholipid was attached (→ **39ab**). Debenzylation afforded *N*-Boc-protected **39ac** which permits regioselective attachment of peptide residues to one aminoethylphosphate moiety and complete deprotection led to target molecule **39ad**. This way a highly variable concept for the synthesis of branched GPI anchors could be established. It is based on versatile building blocks which are readily accessible and provide the desired regio- and stereocontrol. The concept further allows for the regioselective attachment of peptides or proteins.

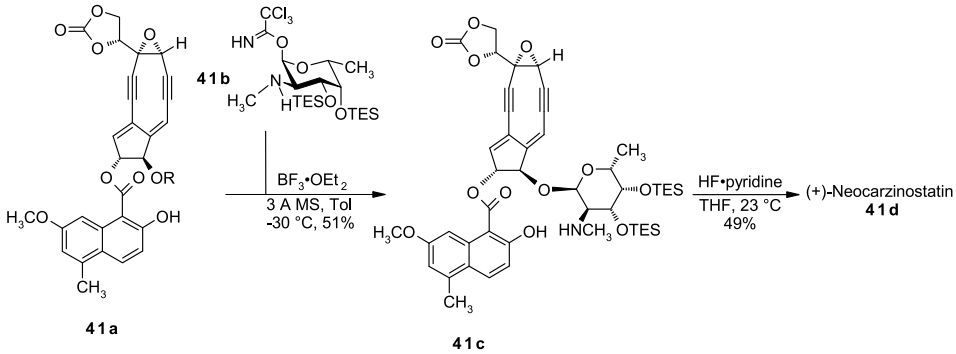
The synthesis of a partial structure of a branched pseudohexasaccharide of an inositolphosphoglycan (IPG) has been reported by Martin-Lomas et al. [273]. Linear oligosaccharides of GPIs were also synthesized based on the *trichloroacetimidate method* [274,275]. The GPI anchor of *Trypanosoma cruzi* containing an unsaturated fatty acid was obtained without any *O*-benzyl protection in the decisive intermediate (◉ *Scheme 40*) [276]. Acyl-protected donor **40a** and acyl-protected acceptor **40b** gave disaccharide **40c** which was transformed into acceptor **40d** which on reaction with donor **40e** gave trisaccharide **40f** and then acceptor **40g**. Glycosylation with **40h** led to tetrasaccharide **40i** which was transformed via **40j** into donor **40k**. From this compound the desired target molecule was prepared.

5.7 Glycosylation of Various Natural Products and Their Metabolites

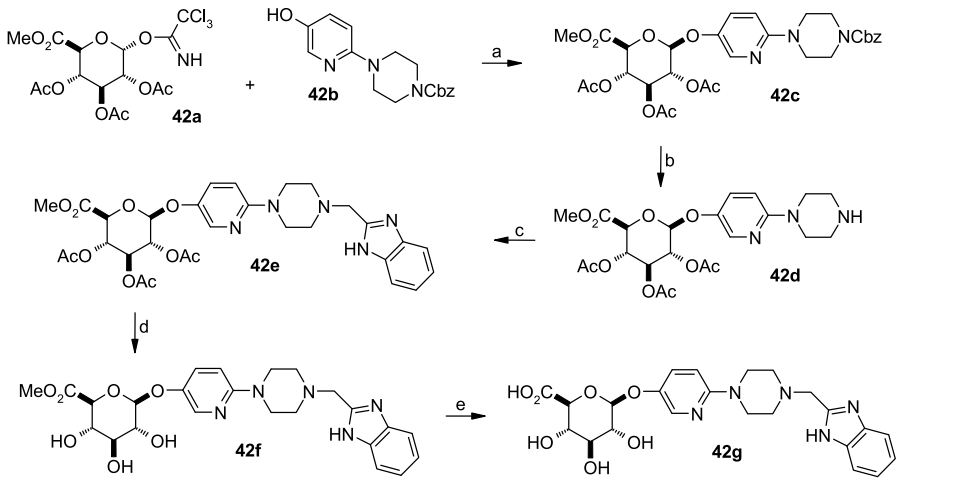
Many antibiotics are glycosylated, however the structural complexity of many antibiotics and their glycosylation is a demanding task. On the basis of the *trichloroacetimidate method* a successful synthesis of landamycin A hexasaccharide was reported [277]. Novobiocin was glycosylated with three different *O*-glycosyl trichloroacetimidates as donors at the 7-hydroxy group [278]. Also the quite sensitive (+)-Neo-carzinostatin aglycone **41a** (◉ *Scheme 41*) could be successfully glycosylated [279]. Reaction of **41a** with glycosyl donor **41b** furnished the desired glycoside **41c** in good yield which on deprotection afforded target molecule **41d**. Hydroquinone glycosylation leading to α-arbutin was also performed [280]; this compound is an inhibitor of human tyrosinase and therefore of interest in the cosmetic industry.

Aminoglycoside antibiotics are of special interest as glycosylation targets. Interesting neomycin trisaccharide mimetics have been prepared by Boons et al. [281]. Glycosidase inhibitors have also been combined with additional carbohydrate residues in order to fine tune their biological properties. Thus, deoxynojirimycin has been combined with β(1–3)- and β(1–6)-linked glucan residues [282].

Many metabolites are glucuronidated, therefore the introduction of the glucuronic acid residue is a major task because of the low reactivity of the derived donors and their tendency to 4,5-elimination. Therefore, often compounds are first glycosidated and at a later stage of the



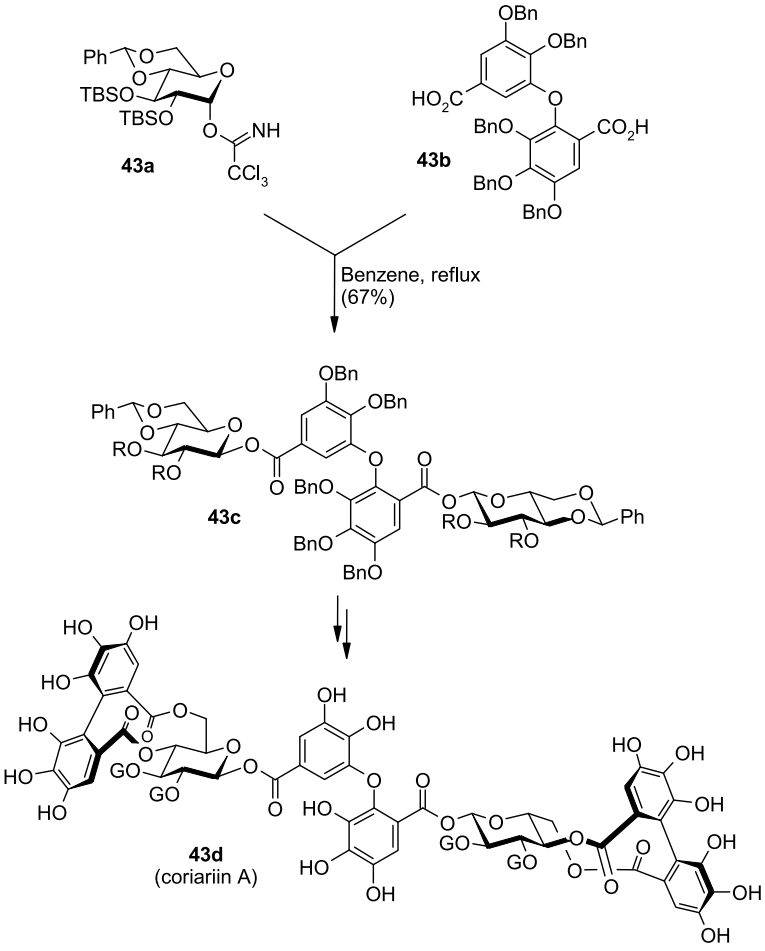
Scheme 41



(a) 2.0 eq **42a**, 1.0 eq **42b**, BF_3•OEt_2, CH_2Cl_2, -60 °C → rt, 41 h, 75%; (b) 5% Pd/C, H_2, MeOH, rt, 83%;
(c) (1) benzimidazolylmethyl chloride, NEt_3, NMP, 0 °C → rt, (2) EtOAc, rt, 62%; (d) 0.1 eq LiOH, MeOH, -10 °C → rt;
(e) 1.9 eq LiOH, MeOH, rt; 2.0 eq AcOH, H_2O, $CHCl_3$ wash; Dowex hydroxide resin, filtration, wash, then $AcOH_{aq}$, 87%.

Scheme 42

synthesis the hydroxymethyl group is selectively oxidized to the carboxylate group. Excellent results in direct glucuronidation based on *O*-isobutyryl-protected trichloroacetimidate have been obtained by Scheinmann and Stachulski et al., who glucuronidated for instance an anti-estrogenic steroid [283,284] and morphine [285]. Also MS 209, a quinoline derivative exhibiting multidrug resistance in cancer therapy was directly glucuronidated based on an *O*-pivaloyl-protected donor [286]. 5-Hydroxypyridine derivative **42b** (*Scheme 42*) which is part of ABT-724, a potent D4 dopamine receptor agonist, could be directly glucuronidated [287] leading to **42c**. Further transformations via intermediates **42e** and **42f** gave the desired glycosylated metabolite **42g**.

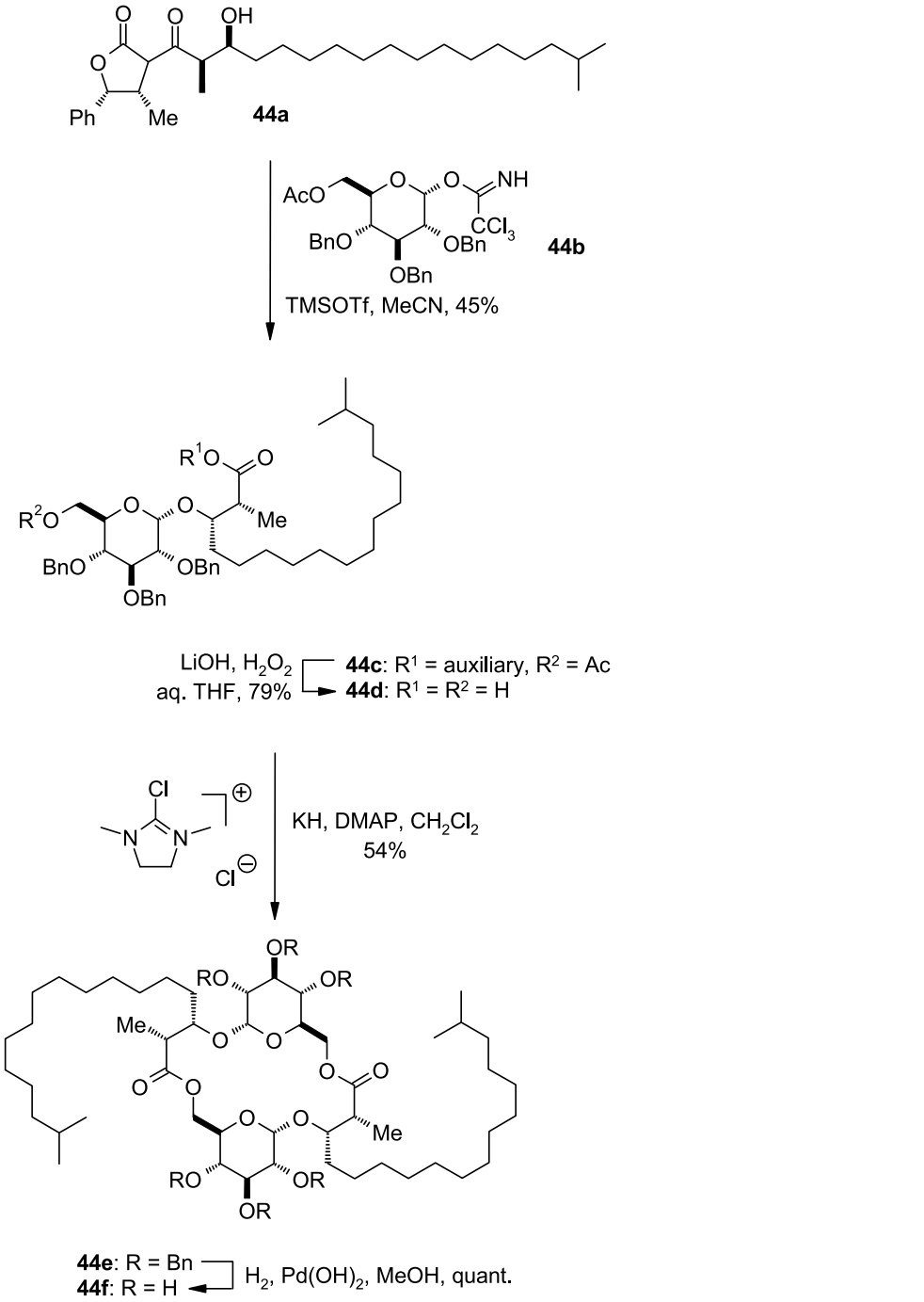


Scheme 43

Dihydroxyphenyl glycosides are a family of plant components. Some representatives have also been synthesized based on the *trichloroacetimidate method*, such as for instance the conandroside [288]. Other plant metabolites, such as for instance the ellagitannins, contain benzoyl groups at the anomeric oxygen. The derived coriariin A (**43d**, *Scheme 43*) was obtained as discussed above just by heating **43b** with glycosyl donor **43a**, leading to **43c**, which was transformed into target molecule **43d** [289]. Macrophylloside D [290] and buprestin A [291] were similarly prepared.

N-Glycosylation of nitrogen-containing heterocycles has also been investigated. An interesting synthesis of *N*-indigoglycosides was undertaken which, as analogs of akashines, exhibit activity against various human tumor cell lines [292].

Unique glycolipids produced by plants are the resin glycosides for which the first synthesis was based on the *trichloroacetimidate method* [293,294]. Other members of this family have been


44a
44b
TMSOTf, MeCN, 45%
LiOH, H2O2
aq. THF, 79%
44c: R1 = auxiliary, R2 = Ac
44d: R1 = R2 = H
KH, DMAP, CH2Cl2
54%
44e: R = Bn
44f: R = H
H2, Pd(OH)2, MeOH, quant.

Scheme 44

prepared with the help of closely related trichloroacetimidate building blocks and a similar strategy, such as for instance for woodrosin I [295] and tricolorin F [296]. Structurally related, highly bioactive antiviral compounds, such as for instance cycloviracin B_1 or glucolipsin A were isolated and successfully synthesized [297,298,299]. The approach to the synthesis of glucolipsin A [299] is exhibited in *Scheme 44*. Glucosylation of fatty acid derivative **44a** with *O*-glycosyl trichloroacetimidate **44b** gave the desired glycoside **44c** in good yield. Ester hydrolysis (→ **44d**) and dimerization to **44e** and then hydrogenolysis of the *O*-benzyl groups led to target molecule **44f** which had physical data in accordance with the natural product (see [299]).

Saponins are steroid- or triterpenoid-based glycolipids which are found in terrestrial and marine plants. They possess various biological activities [300]. Several papers reported on successful glycosylations at the 3-hydroxy group [301,302,303,304,305,306]. When this hydroxy group is sterically hindered due to disubstitution at C-4, this glycosylation is a difficult task. However, with *O*-glycosyl trichloroacetimidate donors good yields and anomeric selectivities have also been obtained for this glycosylation step [307,308,309,310,311,312]. Particularly worth mentioning is the total synthesis of QS 21 A published by Gin et al. [25] where the decisive step is carried out with a trisaccharide donor.

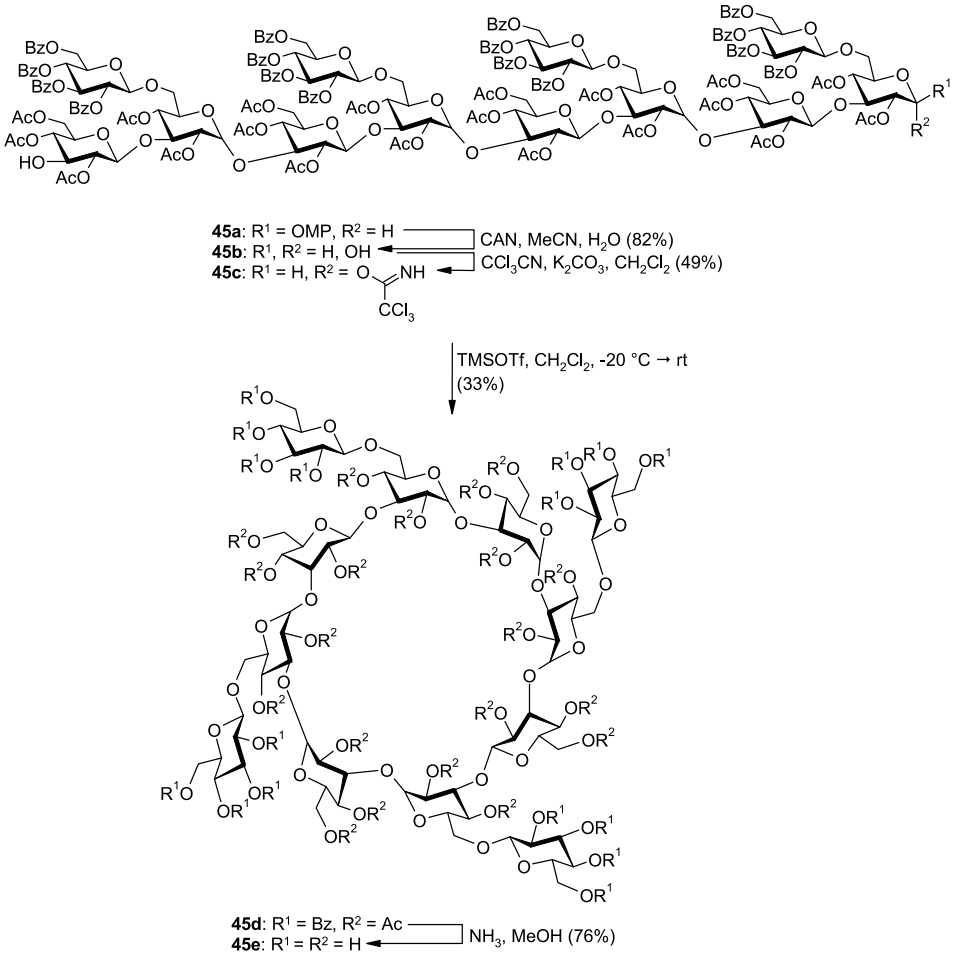
Most carbohydrates contain ethane-1,2-diol fragments, therefore they were also successfully included into crown ether macrocycles [313]. Polyvalency of carbohydrates was studied based on dendrimeric structures which were generated on successful glycosylations of highly branched polyethylene glycol units [314].

5.8 Cyclooligosaccharides

Cyclodextrins (CDs) have gained a lot of interest because they provide useful cavities for the generation of inclusion complexes. Therefore, branched CDs and CDs with various carbohydrate residues have been synthetic targets. The *trichloroacetimidate method* was successfully employed to reach such goals. Mannosyl, galactosyl, lactosyl, and the Galβ(1–4′)-lactosyl residues were selectively introduced at *O*-6 of β- and γ-CD, respectively [315,316]. Also building blocks for oligomannoside attachment to CDs were prepared [317]. Most remarkable is the synthesis of a *O*-6 branched cyclo(1–3)-glucohexaose and octaose [318,319]. The ring closure of the prepared nona- and dodecasaccharide worked extremely well based on *O*-glycosyl trichloroacetimidate donors (*Scheme 45*) [5]. Dodecasaccharide **45a** was transformed via **45b** into trichloroacetimidate **45c** having a 3-hydroxy group at the nonreducing end. Acid-catalyzed activation led to CD derivative **45d** which gave after deacylation target molecule **45e**.

5.9 *C*-Glycoside Synthesis

The synthesis of "*C*-glycosides" has been of great interest [320]. *C*-Glycosylation of electron-rich compounds with *O*-glycosyl trichloroacetimidates continues to be a method of choice. This work is not discussed here in detail. A few references give an entry to the work in this field [321,322,323,324].



Scheme 45

5.10 *O*-Glycosyl Trichloroacetimidates of *N*,*O*- and *S*,*O*-Halfacetals

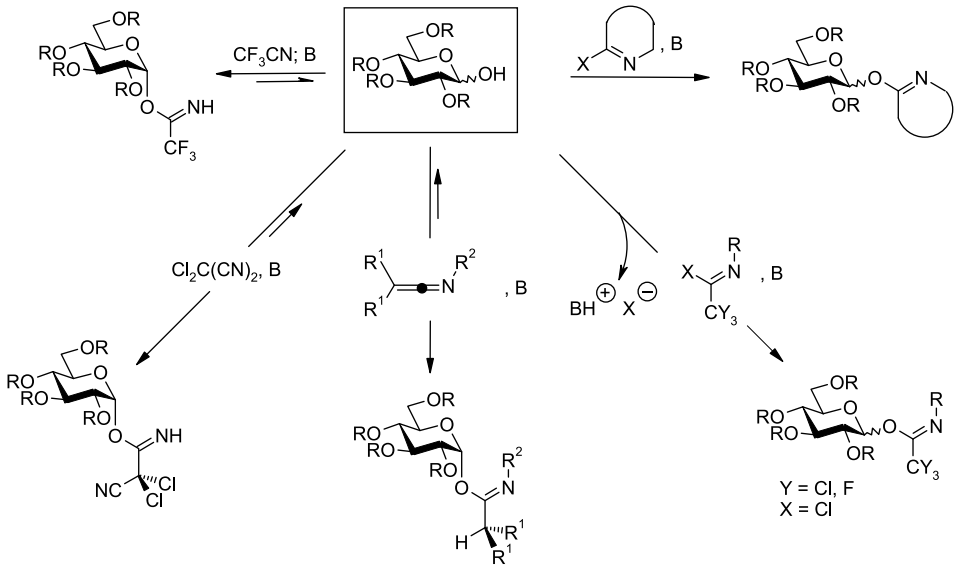
Not only *O*,*O*-halfacetals but also *N*,*O*-, and *S*,*O*-halfacetals have successfully been transformed into trichloroacetimidate derivatives [325,326]. Thus carbohydrates with sulfur in the ring have been successfully transformed into *O*-glycosyl trichloroacetimidates and employed in glycosylation reactions by M. Hashimoto et al. [327,328] and by H. Hashimoto et al. [329,330]. Thy glycosyl donors could be readily obtained by base-catalyzed trichloroacetonitrile addition to the anomeric hydroxy group and glycosylation yields under standard conditions were generally very good.

6 Related Activation Systems

O-Glycosyl trichloroacetimidates have become popular because they are easily available, powerful glycosyl donors. Therefore, investigations were undertaken to provide glycosyl donors which follow the principle of *O*-glycosyl trichloroacetimidate formation and activation. A short outline of this work is attached here.

Already previously trifluoroacetonitrile has been investigated for the generation of glycosyl donors [12,331]. However, because trifluoroacetonitrile is a gas, its use is not as convenient as the use of trichloroacetonitrile. In addition, preliminary glycosylation results were inferior to those obtained with *O*-glycosyl trichloroacetimidates. Much more promising were studies with dichloromalonitrile which is a suitable reagent for the base-catalyzed generation of *O*-glycosyl dichloro-cyanoacetimidates [332,333]. These compounds exhibited glycosyl donor properties closely related to those of *O*-glycosyl trichloroacetimidates (*Scheme 46*). A further important class of compounds is ketene imines, which should readily lead to *O*-glycosyl imidates. However, so far only a few examples were investigated [12,13,330], therefore, the potential of these compounds has not been elucidated yet.

Another interesting class of compounds are imide halides having electron-withdrawing carbon substituents and their heterocyclic equivalents. After some earlier work [12,334,335,336,337,338], recently imide halides have gained increased interest and excellent glycosylation results have been reported [45,339,340]. However, these systems have the disadvantage of furnishing equimolar amounts of salt on glycosyl donor generation. Additionally, glycosyl donor generation is not reversible, therefore α/β-stereocontrol is more difficult or even impossible. Hence, matching or even surpassing the properties of *O*-glycosyl trichloroacetimidates remains a demanding task.



Scheme 46

7 Conclusions

The requirement for efficient glycosylation methods, as outlined at the beginning of this chapter, namely convenient diastereocontrolled anomeric *O*-activation (first step) and subsequent efficient diastereocontrolled glycosylation promoted by catalytic amounts of a promoter (second step) are essentially completely fulfilled by the *trichloroacetimidate method*. In terms of reactivity and applicability toward different acceptors, the *O*-glycosyl trichloroacetimidates have generally proven to be outstanding glycosyl donors, which resemble in various respects the nucleoside diphosphate sugar derivatives used by nature as glycosyl donors. Thus, base-catalyzed generation of *O*-glycosyl trichloroacetimidates followed by acid-catalyzed glycosylation have become a very competitive alternative to other methods mainly requiring anomeric oxygen exchange reactions for glycosyl donor generation and at least equimolar amounts of a promoter system for the glycosylation step. Hence, the *trichloroacetimidate method* can readily be adapted for large-scale preparations. Recently, related activation systems have attracted increasing interest, thus widening the scope of this glycosylation method.

Acknowledgement

Our work reported in this chapter was supported by the Deutsche Forschungsgemeinschaft and the Fonds der Chemischen Industrie.

References

1. Schmidt RR (1986) Angew Chem Int Ed Engl 25:89
2. Schmidt RR, Kinzy W (1994) Adv Carbohydr Chem Biochem 50:21
3. Schmidt RR, Jung KH (1997) Oligosaccharide Synthesis via Trichloroacetimidates. In: Hannessian S (ed) Preparative Carbohydrate Chemistry. Marcel Dekker, p 283
4. Schmidt RR, Jung KH (2000) Trichloroacetimidates. In: Ernst B, Hart GW, Sinaÿ P (eds) Carbohydrates in Chemistry and Biology, Part I: Chemistry of Saccharides, vol 1. Wiley-VCH, Weinheim, p 5
5. Paulsen H (1982) Angew Chem Int Ed Engl 21:155
6. Kunz H (1987) Angew Chem Int Ed Engl 26:294; (1993) Pure Appl Chem 65:1223
7. Toshima K, Tatsuta K (1993) Chem Rev 93:1503
8. Schmidt RR, Reichrath M (1979) Angew Chem Int Ed Engl 18:466
9. Klotz W, Schmidt RR (1994) J Carbohydr Chem 13:1093
10. Schmidt RR, Klotz W (1991) Synlett 168
11. Klotz W, Schmidt RR (1993) Liebigs Ann Chem 683
12. Michel J (1983) PhD thesis, Universität Konstanz
13. Schmidt RR, Michel J (1980) Angew Chem Int Ed Engl 19:731
14. Pougny JR, Sinaÿ P (1976) Tetrahedron Lett 4073
15. Sinaÿ P (1978) Pure Appl Chem 50:1437
16. Ohashi I, Lear MJ, Yoshimura F, Dirama M (2004) Org Lett 6:719
17. Chiara JL, Encinas L, Díaz B (2005) Tetrahedron Lett 46:2445
18. Yoshizaki H, Fukuda N, Sato K, Oikawa M, Fukase K, Suda Y, Kusumoto S (2001) Angew Chem Int Ed 40:1475
19. Oikawa M, Tanaka T, Fukuda, N, Kusumoto S (2004) Tetrahedron Lett 45:4039
20. Roussel F, Knerr L, Grathwohl M, Schmidt RR (2000) Org Lett 2:3043
21. Knerr L, Schmidt RR (2001) The Use of *O*-Glycosyl Trichloroacetimidates for the Polymer Supported Synthesis of Oligosaccharides. In: Seeberger PH (ed) Solid Support Oligosaccha-

ride Synthesis and Combinatorial Carbohydrate Libraries. Wiley, New York, pp 67
22. Ali IAI, El-Ashry ESH, Schmidt RR (2003) Eur J Org Chem 4121
23. Yu H, Williams DL, Ensley HE (2005) Tetrahedron Lett 46:3417
24. Karst NA, Islam TF, Avci FY, Linhardt RJ (2004) Tetrahedron Lett 45:6433
25. Yu B, Zhu X, Hui Y (2001) Tetrahedron 57:9403
26. Weingart R, Schmidt RR (2000) Tetrahedron Lett 41:8753
27. Nicolaou KC, Daines RA, Ogawa Y, Chakraborty TK (1987) J Am Chem Soc 109:2821
28. Nicolaou KC, Daines RA, Ogawa Y, Chakraborty TK (1988) J Am Chem Soc 110:4696
29. Jona H, Mandai H, Chavasiri W, Takeuchi K, Mukaiyama T (2002) Bull Chem Soc Jpn 75:291
30. Jona H, Mandai H, Mukaiyama T (2001) Chem Lett 426
31. Douglas SP, Whitfield DM, Krepinsky JJ (1993) J Carbohydr Chem 12:131
32. Wei G, Gu G, Du Y (2003) J Carbohydr Chem 22:385
33. Yamada H, Hayashi T (2002) Carbohydr Res 337:581
34. Castro-Palomino JC, Schmidt RR (1995) Tetrahedron Lett 36:5343
35. Geiger J, Barroca N, Schmidt RR (2004) Synlett 836
36. Adinolfi M, Barone G, Guariniello L, Iadonisi A (2000) Tetrahedron Lett 41:9005
37. Adinolfi M, Barone G, Iadonisi A, Mangoni L, Schiattarella M (2001) Tetrahedron Lett 42:5967
38. Griswold KS, Horstmann TE, Miller SJ (2003) Synlett 1923
39. Kartha KPR, Karkkainen TS, March SJ, Field RA (2001) Synlett 260
40. Adinolfi M, Barone G, Iadonisi A, Mangoni L, Schiattarella M (2002) Synlett 269
41. Schmidt RR, Gaden H, Jatzke H (1990) Tetrahedron Lett 31:327
42. Mukhopadyay B, Maurer SV, Rudolph N, van Well RM, Russell DA, Field RA (2005) J Org Chem 70:9059
43. Du Y, Wei, G, Cheng S, Hua Y, Linhardt RJ (2006) Tetrahedron Lett 47:307
44. Adinolfi M, Barone G, Iadonisi A, Mangoni L, Schiattarella M (2003) Org Lett 5:987
45. Schmidt RR, Rücker E (1980) Tetrahedron Lett 21:1421
46. Schmidt RR, Behrendt M, Toepfer A (1990) Synlett 694
47. Schmidt RR (1992) New Aspects of Glycosylation Reactions. In: Ogura H, Hasegawa A, Suami T (eds) Carbohydrates—Synthetic Methods and Applications in Medicinal Chemistry. Kodanasha Ltd., Tokyo, 68
48. Toepfer A, Schmidt RR (1991) Tetrahedron Lett 32:3353
49. Bommer R, Kinzy W, Schmidt RR (1991) Liebigs Ann Chem 425
50. Böhm G, Waldmann H (1995) Tetrahedron Lett 36:3843
51. Schmid U, Waldmann H (1997) Liebigs Ann Chem 2573
52. Jain N, Kumar A, Chauhan S, Chauhan SMS (2005) Tetrahedron 61:1015
53. Rencurosi A, Lay L, Russo G, Caneva E, Poletti L (2005) J Org Chem 70:7765
54. Larsen K, Worm-Leonhard k, Olsen P, Hoel A, Jensen KJ (2005) Org Biomol Chem 3:3966
55. Zeng Y, Ning J, Kong F (2002) Tetrahedron Lett 43:3729
56. Yang F, He H, Du Y, Lü M (2002) Carbohydr Res 337:1165
57. Zeng Y, Ning J, Kong F (2003) Carbohydr Res 338:307
58. Spijker NM, van Boeckel CAA (1991) Angew Chem Int Ed Engl 30:180
59. Masamuni S, Choy W, Petersen JC, Sita LR (1985) Angew Chem Int Ed Engl 24:1
60. Cid MB, Alfonso F, Martín-Lomas M (2005) Chem Eur J 11:928
61. Castro-Palomino JC, Schmidt RR (1998) Synlett 501
62. Marzabadi CH, Franck RW (2000) Tetrahedron 56:8385
63. Roush WR, Gung BW, Bennett CE (1999) Org Lett 1:891
64. Chong PY, Roush WR (2002) Org Lett 4:4523
65. Kim JH, Yang H, Park J, Boons GJ (2005) J Am Chem Soc 127:12090
66. Gridley JJ, Osborn HMI (200) J Chem Soc Perkin Trans 1:1471
67. Dwek RA (1996) Chem Rev 96:683
68. Barresi F, Hindsgaul O (1996) In: Khan SH, O'Neill RA (eds) Modern Methods in Carbohydrate Synthesis. Harwood Academic Publishers, Amsterdam, p 251
69. Paulsen H, Lockhoff O (1981) Chem Ber 114:3102
70. Garegg PJ, Ossowski P (1983) Acta Chem Scand 337:249

71. Danishefsky SJ, Hu S, Cirilo PF, Eckhardt M, Seeberger PH (1997) Chem Eur J 3:1617
72. Lichtenthaler FW, Schneider-Adams T (1994) J Org Chem 59:6728
73. Schmidt RR, Moering U, Reichrath M (1982) Chem Ber 115:39
74. Hodosi G, Kovác P (1997) J Am Chem Soc 119:2335
75. Ekborg G, Lindberg B, Lonngren J (1972) Acta Chem Scand B 26:3287
76. Matsuo I, Isomura M, Walton R, Ajisaka K (1996) Tetrahedron Lett 37:8795
77. Kunz H, Günther W (1988) Angew Chem Int Ed Engl 27:1086
78. Weiler S, Schmidt RR (1998) Tetrahedron Lett 39:2299
79. Twaddle GWJ, Yashunsky DV, Nikolaev AV (2003) Org Biomol Chem 1:623
80. Barresi F, Hindsgaul O (1991) J Am Chem Soc 113:9376
81. Stork G, LaChair JJ (1996) J Am Chem Soc 118:247
82. Ito Y, Ogawa T (1994) Angew Chem Int Ed Engl 33:1765
83. Jung KH, Müller M, Schmidt RR (2000) Chem Rev 100:4423
84. Ziegler T, Lemanski G, Rakoczy A (1995) Tetrahedron Lett 36:8973
85. Caregg PJ, Iversen T, Johansson R (1980) Acta Chem Scand B 34:505
86. Crich D, Sun S (1996) J Org Chem 61:4506
87. Crich D, Smith M (2001) J Am Chem Soc 123:9015
88. Ikeda T, Yamada H (2000) Carbohydr Res 329:889
89. Abdel-Rahman AAH, Jonke S, El Ashry ESH, Schmidt RR (2002) Angew Chem Int Ed 41:2972
90. Zhou FY, Huang JY, Yuan Q, Wang YG (2005) Chem Lett 34:878
91. Wang CC, Lee JC, Luo SY, Fan HF, Pai CL, Yang WC Lu LD Hung SC (2002) Angew Chem Int Ed 41:2360
92. Curran DP, Ferritto R, Hua Y (1998) Tetrahedron Lett 39:4937
93. Miura T, Hirose Y, Ohmae M, Inazu T (2001) Org Lett 3:3947
94. Miura T, Inazu T (2003) Tetrahedron Lett 44:1819
95. Haase WC, Seeberger PH (2000) Chem Rev 100:4349
96. Plante OJ, Palmacci ER, Seeberger PH (2001) Science 291:1523
97. Roussel F, Takhi M, Schmidt RR (2001) J Org Chem 66:8540
98. Roussel F, Knerr L, Schmidt RR (2001) Eur J Org Chem 2066
99. Wu X, Grathwohl M, Schmidt RR (2001) Org Lett 3:747
100. Rademann J, Geyer A, Schmidt RR (1998) Angew Chem Int Ed 37:1241
101. Rademann J, Schmidt RR (1997) J Org Chem 62:3650
102. Zhu T, Boons GJ (2001) Chem Eur J 7:2382
103. Zhu T, Boons GJ (2000) J Am Chem Soc 122:10222
104. Nicolaou KC, Watanabe N, Li J, Pastor J, Winssinger N (1998) Angew Chem Int Ed 37:1559
105. Mogemark M, Elofsson M, Kihlberg J (2003) J Org Chem 68:7281
106. Ando H, Manabe S, Nakahara Y, Ito Y (2001) Angew Chem Int Ed 40:4725
107. Pamacci ER, Hewitt MC, Seeberger PH (2001) Angew Chem Int Ed 40:4433
108. Wu X, Schmidt RR (2004) J Org Chem 69:1853
109. Bauer J, Rademann (2005) J Am Chem Soc 127:7296
110. Wu X, Grathwohl M, Schmidt RR (2002) Angew Chem Int Ed 41:4489
111. Jonke S, Liu K, Schmidt RR (2006) Chem Eur J 12:1274
112. Andrade RB, Plante OJ, Melean LG, Seeberger PH (1999) Org Lett 1:1811
113. Ratner DM, Swanson ER, Seeberger PH (2003) Org Lett 5:4717
114. Hewitt MC, Snyder DA, Seeberger PH (2002) J Am Chem Soc 124:13434
115. Manabe S, Ito Y (2002) J Am Chem Soc 124:12638
116. Ojeda R, de Paz JL, Martín-Lomas M (2003) Chem Commun 2486
117. Ojeda R, Terentí O, de Paz JL, Martín-Lomas M (2004) Glycoconjugate J 21:179
118. de Paz JL, Angulo J, Lassaletta JM, Nieto PM, Ridondo-Horcajo M, Lozano RM, Gimenez-Gallego G, Martín-Lomas M (2001) ChemBioChem 2:673
119. Blattner R, Furneaux RH, Ludewig M (2006) Carbohydr Res 341:299
120. Knust B (1996) Diploma thesis, Universität Konstanz
121. Kantchev EAB, Parquette JR (2005) Synlett 1567
122. Schmidt RR, Zimmermann P (1986) Angew Chem Int Ed Engl 25:725

123. Zimmermann P, Schmidt RR (1988) Liebigs Ann Chem 663
124. Hansen HC, Magnusson G (1999) Carbohydr Res 322:190
125. Aly MRE, Rochaix P, Amessou M, Johannes L, Florent JC (2006) Carbohydr Res 341:2026
126. Chen L, Zhao XE, Lai D, Song Z, Kong F (2006) Carbohydr Res 341:1174
127. Sun B, Pukin AV, Visser GM, Zuilhof H (2006) Tetrahedron Lett 14:7371
128. Tietze LF, Keim K, Janßen CO, Tappertzhofen C, Olschimke J (2000) Chem Eur J 6:2801
129. Castro-Palomino JC, Simon B, Speer O, Leist M, Schmidt RR (2001) Chem Eur J 7:2178
130. Janssen S, Schmidt RR (2005) J Carbohydr Chem 24:611
131. Aly MRE, Ibrahim ESI, El Ashry ESH, Schmidt RR (1999) Carbohydr Res 316:121
132. Gege C, Kinzy W, Schmidt RR (2000) Carbohydr Res 328:459
133. Gege C, Oscarson S, Schmidt RR (2001) Tetrahedron Lett 42:377
134. Gege C, Schmidt RR (2002) Carbohydr Res 337:1089
135. Gege C, Geyer A, Schmidt RR (2002) Eur J Org Chem 2475
136. Gege C, Vogel J, Bendas G, Rothe U, Schmidt RR (2000) Chem Eur J 6:111
137. Gege C, Geyer A, Schmidt RR (2002) Chem Eur J 8:2454
138. Yamaguchi M, Ishida H, Kanamori A, Kannagi R, Kiso M (2003) Carbohydr Res 338: 2793
139. Bélot F, Rabuka D, Fukuda M, Hindsgaul O (2002) Tetrahedron Lett 43:7743
140. Natori T, Morita M, Akimoto K, Koezuka Y (1994) Tetrahedron 50:2771
141. Natori T, Koezuka Y, Higa T (1993) Tetrahedron Lett 34:5591
142. Morita M, Motoki K, Akimoto K, Natori T, Sakai T, Sawa E, Yamaji K, Koezuka Y Kobayashi E, Fukushima H (1995) J Med Chem 38:2176
143. Kawano T, Cui J, Koezuka Y, Toura I, Kaneko Y, Motoki K, Ueno H, Nakagawa R, Sato H, Kondo E, Koseki H, Taniguchi M (1997) Science 278:1626
144. Figueroa-Pérez S, Schmidt RR (2000) Carbohydr Res 328:95
145. Laurent N, lafont D, Buillange P (2006) Carbohydr Res 341:823
146. Bartolmäs T, Heyn T, Mickeleit M, Fischer A, Reutter W, Danker K (2005) J Med Chem 48:6750
147. Gouin SG, Pilgrim W, Porter RK, Murphy PV (2005) Carbohydr Res 340:1547
148. Sato R, Toma K, Nomura K, Takagi M, Yoshida T, Azefu Y, Tamiaki H (2004) J Carbohydr Chem 23:375
149. Herzner H, Reipen T, Schultz M, Kunz H (2000) Chem Rev 100:4495
150. Röhrig CH, Retz OA, Hareng L, Hartung T, Schmidt RR (2005) ChemBioChem 6:1805
151. Saha UK, Schmidt RR (1997) J Chem Soc Perkin Trans 1:1855
152. Saha UK, Griffith LS, Rademann J, Geyer A, Schmidt RR (1997) Carbohydr Res 304:21
153. Burger K, Kluge M, Fehn S, Koksch B, Hennig L, Müller G (1999) Angew Chem Int Ed 38:1414
154. Allevi P, Anastasia M, Paroni R, Ragusa A (2004) Bioorg Med Chem Lett 14:3319
155. Komba S, Meldal M, Werdelin O, Jensen T, Bock K (1999) J Chem Soc Perkin Trans 1:415
156. St. Hilaire PM, Cipolla L, Franco A, Tedebark U, Tilly DA, Meldal M (1999) J Chem Soc Perkin Trans 1:3559
157. Xue J, Khaja SK, Locke RD, Matta KL (2004) Synlett 861
158. Paulsen H, Paal M, Schultz M (1983) Tetrahedron Lett 24:1759
159. Paulsen H, Paal M, (1984) Carbohydr Res 135:53
160. Kessler H, Kottenhahn M, Kling A, Kolar C (1987) Angew Chem Int Ed Engl 99:919
161. Halkes KM, Gotfredsen CH, Grøtly M, Miranda LP, Duus JØ, Meldal M (2001) Chem Eur J 7:3584
162. Nicolaou KC, Mitchell HJ, Jain NF, Bando T, Hughes R, Winssinger N, Natarajan S, Koumbis AE (1999) Chem Eur J 5:2648
163. Nicolaou KC, Cho SY, Hughes R, Winssinger N, Smethurst C, Labischinski H, Endermann R (2001) Chem Eur J 7:3798
164. Ritter TK, Mong KKT, Liu H, Nakatani T, Wong CH (2003) Angew Chem Int Ed 42:4657
165. Meinjohanns E, Meldal M, Paulsen H, Dwek RA, Bock K (1998) J Chem Soc Perkin Trans 1: 549
166. Matsuo I, Nakahara Y, Ito Y, Nukada T, Nakahara Y, Ogawa T (1995) Bioorg Med Chem 3:1455
167. Chiesa MV, Schmidt RR (2000) Eur J Org Chem 3541

168. Aly MRE, Ibrahim ESI, El Ashry ESH, Schmidt RR (2001) Carbohydr Res 331:129
169. Zhu Y, Chen L, Kong F (2002) Carbohydr Res 337:207
170. Ratner DM, Plante OJ, Seeberger PH (2002) Eur J Org Chem 826
171. Mendoza VM, Agusti R, Gallo-Rodriguez C, de Lederkremer RM (2006) Carbohydr Res 341:1488
172. Nakano J, Ichiyanagi T, Ohta H, Ito Y (2003) Tetrahedron Lett 44:1742
173. Hofsteenge J, Müller DR, de Beer T, Lüffler A, Richter WJ, Vliegenthart JFG (1994) Biochemistry 33:13524
174. Gäde G, Kellner R, Rinehart KL, Proefke ML (1992) Biochem Biophys Res Commun 189:1303
175. Hofsteenge J, Blommers M, Hess D, Furmanek A, Miroshnichenko O (1999) J Biol Chem 274:32786
176. Diem S, Albert J, Herderich M (2001) Eur Food Res Technol 213:439
177. Schnabel M, Römpp B, Ruckdeschel D, Unverzagt C (2004) Tetrahedron Lett 45:295
178. Adinolfi M, Barone G, De Napoli L, Guariniello L, Iadonisi A, Piccialli G (1999) Tetrahedron Lett 40:2607
179. Adinolfi M, De Napoli L, di Fabio G, Guariniello L, Iadonisi A, Messere A, Montesachio D, Piccialli G (2001) Synlett 745
180. Bernards A, van Harten-Loosbroek N, Borst P (1984) Nucleic Acids Res 12:4153
181. Gommers-Ampt JH, Lugterink J, Borst P (1991) Nucleic Acids Res 19:1745
182. Gommers-Ampt JH, van Leeuwen F, de Beer ALJ, Vliegenthart JFG, Dizdaroglu M, Kowalak JA, Crain PF, Borst P (1993) Cell 75:1129
183. de Kort M, Ebrahimi E, Wijsman ER, van der Marel GA, van Boom JH (1999) Eur J Org Chem 2337
184. Fransson LÅ (1987) Trends Biochem Sci 12:406
185. Kjellen L, Lindahl U (1991) Annu Rev Biochem 60:443
186. Karst N, Jacquinet JC (2000) J Chem Soc Perkin Trans 1:2709
187. Karst N, Jacquinet JC (2002) Eur J Org Chem 815
188. Tully SE, Mabon R, Gama CI, Tsai SM, Liu X, Hsieh-Wilson LC (2004) J Am Chem Soc 126:7736
189. Tamura J, Tokuyoshi M (2004) Biosci Biotechnol Biochem 68:2436
190. Soliman Se, Bassily RW, El-Sokkary RI, Nashed MA (2003) Carbohydr Res 338: 2337
191. Yeung BKS, Hill DC, Janicka M, Petillo PA (2000) Org Lett 2:1279
192. Casu B (1985) Adv Carbohydr Chem 43:51 and references cited therein
193. Faham S, Hileman RE, Fromm JR, Linhardt RJ, Rees DC (1996) Science 271:1116
194. Hileman RE, Fromm JR, Wiler JM, Linhardt RJ (1998) BioEssays 20:156
195. Pellegrine L, Burke DF, von Delft F, Mulloy B, Blundell TL (2000) Nature 107:1029
196. Mach H, Volkin DB, Burke CJ, Middaugh CR, Linhardt RJ, Fromm JR, Loganathan D, Mattsson L (1993) Biochemistry 32:5480
197. Guimond S, Maccarana M, Olwin BB, Lindahl U, Rapraeger AC (1993) J Biol Chem 268:23906
198. Poletti L, Fleischer M, Vogel C, Guerrini M, Torri G, Lay L (2001) Eur J Org Chem 2727
199. Lucas R, Hamza D, Lubineau A, Bonnaffé D (2004) Eur J Org Chem 2107
200. Lee JC, Lu XA, Kulkarni SS, Wen YS, Hung SC (2004) J Am Chem Soc 126:476
201. de Paz JL, Ojeda R, Reichardt N, Martín-Lomas M (2003) Eur J Org Chem 3308
202. Lohman GJS, Seeberger PH (2004) J Org Chem 69:4081
203. Zhou Y, Lin F, Chen J, Yu B (2006) Carbohydr Res 341:1619
204. Bernfield M, Gotte M, Park PW, Reizes O, Fitzgerald ML, Lincecum J, Zako M (1999) Annu Rev Biochem 68:729
205. Chen L, Kong F (2002) Carbohydr Res 337:1373
206. Rietschel ET, Schletter J, Weidemann B, El-Samalouti V, Mattern T, Zähringer U, Seydel U, Brade H, Flad HD, Kusumoto S, Gupta D, Dziarski R, Ulmer AJ (1998) Microb Drug Resist 4:37
207. Takeuchi O, Hoshino K, Kawai T, Sanjo H, Takada H, Ogawa T, Takeda K, Akira S (1999) Immunity 11:443
208. Inamura S, Fukase K, Kusumoto S (2001) Tetrahedron Lett 42:7613
209. Inamura S, Fujimoto Y, Kawasake A, Shiokawa Z, Woelk V, Heine H, Lindner B, Inohara N, Kusumoto S, Fukase K (2006) Org Biomol Chem 4:232
210. Zeng Y, Zhang J, Kong F (2002) Carbohydr Res 337:1367
211. Gu G, Wie G, Du Y (2004) Carbohydr Res 339:1155

212. Ning J, Heng L, Kong F (2002) Carbohydr Res 337:1159
213. Zeng Y, Zhang J, Ning J, Kong F (2003) Carbohydr Res 338:5
214. Xing Y, Ning J (2003) Tetrahedron Asymmetry 14:1275
215. Bruneteau M, Fabre I, Perret J, Michel G, Ricci P, Joseleau J, Kraus J, Schneider M, Blaschek W, Franz G (1988) Carbohydr Res 175:137
216. Johnson J, Kirkwood S, Misaki A, Nelson TE, Scalleti JD, Smith F (1963) Chem Ind (London) 820
217. He H, Yang,F, Du Y (2002) Carbohydr Res 337:1673
218. Zhu Y, Kong F (2000) Synlett 663
219. Zhang G, Fu M, Ning J (2005) Carbohydr Res 340:597
220. Zeng Y, Kong F (2003) Carbohydr Res 338:2359
221. Ning J, Zhang W, Yi Y, Yang G, Wu Z, Yi J, Kong F (2003) Bioorg Med Chem 11:2193
222. Wu Z, Kong F (2004) Carbohydr Res 339:2761
223. Wu Z, Kong F (2004) Carbohydr Res 339:377
224. Li A, Kong F (2004) Carbohydr Res 339:2499
225. Ning J, Yi Y, Kong F (2002) Tetrahedron Lett 43:5545
226. Ning J, Kong F, Lin B, Lei H (2003) J Agric Food Chem 51:987
227. Yi Y, Zhou Z, Ning J, Kong F, Li J (2003) Synthesis 491
228. Huang GL, Mei XY, Liu MX (2005) Carbohydr Res 340:603
229. Zhu Y, Kong F (2000) Carbohydr Res 329:199
230. McNeil M, Wallner SJ, Hunter SW, Brennan PJ (1987) Carbohydr Res 166:299
231. Nita-Lazar M, Chevolot L, Iwahara S, Takegawa K, Furmanek A, Lienart Y (2002) Acta Biochim Pol 49:1019
232. Ramli N, Shinohara H, Takegawa K, Iwahara SJ (1994) Ferment Bioeng 78:341
233. Zhang G, Fu M, Ning J (2005) Carbohydr Res 340:155
234. Gandolfi-Donadio L, Gallo-Rodriguez C, de Lederkremer RM (2003) J Org Chem 68:6928
235. Du Y, Pan Q, Kong F (2000) Carbohydr Res 329:17
236. Ning J, Yi Y, Yao Z (2003) Synlett 2208
237. Li A, Zeng Y, Kong F (2004) Carbohydr Res 339:673
238. Ma Z, Zhang J, Kong F (2004) Carbohydr Res 339:1761
239. Li A, Kong F (2004) Carbohydr Res 339:1847
240. Li A, Kong F (2005) Carbohydr Res 340:1949
241. Borman S (2006) C&EN 84:80
242. Chen L, Kong F (2002) Carbohydr Res 337:2335
243. Bedini E, Carabellese A, Corsaro MM, De Castro C, Parrilli M (2004) Carbohydr Res 339:1907
244. Zhang J, Kong F (2002) Carbohydr Res 337:391
245. Zhang J, Ning J, Kong F (2003) Carbohydr Res 338:1023
246. Zhang J, Kong F (2003) Carbohydr Res 338:19
247. Zhang J, Kong F (2003) Tetrahedron 59:1429
248. Ma Z, Zhang J, Kong F (2004) Carbohydr Res 339:43
249. Zhang J, Kong F (2002) J Carbohydr Chem 21:579
250. Zhang J, Kong F (2003) Carbohydr Res 338:1719
251. Zhao W, Kong F (2005) Carbohydr Res 340:1673
252. Hua Y, Xiao J, Huang Y, Du Y (2006) Carbohydr Res 341:191
253. Bélot F, Costachel C, Wright K, Phalipon A, Mulard LA (2002) Tetrahedron Lett 43:8215
254. Segat-Dioury F, Mulard LA (2002) Tetrahedron Asymmetry 13:2211
255. Mulard LA, Guerreiro C (2004) Tetrahedron 60:2475
256. Bélot F, Wright K, Costachel C, Phalipon A, Mulard LA (2004) J Org Chem 69:1060
257. Wright K, Guerreiro C, Laurent I, Baleux F, Mulard LA (2004) Org Biomol Chem 2:1518
258. Mulard LA, Clément MJ, Imberty A, Delepierre M (2002) Eur J Org Chem 2486
259. Maruyama M, Takeda T, Shimizu N, Hada N, Yamada H (2000) Carbohydr Res 325:83
260. Ishii K, Kubo H, Yamasaki R (2002) Carbohydr Res 337:11
261. Kubo H, Ishii K, Koshino H, Toubetto K, Naruchi K, Yamasaki R (2004) Eur J Org Chem 1202
262. Ferguson MAJ, Homans SW, Dwek RA, Rademacher TW (1988) Science 239:753
263. Frankhauser C, Homas SW, Oates JE, McConville MJ, Desponds C, Conzelmann A, Ferguson MAJ (1993) J Biol Chem 268: 26365
264. Homans SW, Ferguson MAJ, Dwek RA, Rademacher TW, Anad R, Williams AF (1988) Nature 333:269
265. Caro HH, Martín-Lomas M, Bernabé M (1993) Carbohydr Res 240:119
266. Mayer TG, Kratzer B, Schmidt RR (1994) Angew Chem Int Ed Engl 33:2177
267. Mayer TG, Schmidt RR (1999) Eur J Org Chem 1153

268. Murakata C, Ogawa T (1992) Carbohydr Res 234:75 and 235:95
269. Baeschlin DK, Chaperon AR, Green LG, Hahn ME, Ince SJ, Ley SV (2000) Chem Eur J 6:172
270. Campbell AS, Fraser-Reid B (1995) J Am Chem Soc 117:10387
271. Pekari K, Tailler D, Weingart R, Schmidt RR (2001) J Org Chem 66:7432
272. Pekari K, Schmidt RR (2003) J Org Chem 68:1295
273. Martín-Lomas M, Flores-Mosquera M, Chiara JL (2000) Eur J Org Chem 1547
274. Ma Z, Zhang J, Kong F (2004) Carbohydr Res 339:29
275. Gandolfi-Donadio L, Gallo-Rodriguez C, de Lederkremer RM (2002) J Org Chem 67: 4430
276. Yashunsky DV, Borodkin VS, Ferguson MAJ, Nikolaev AV (2006) Angew Chem 118:482
277. Roush WR, Bennett CE (2000) J Am Chem Soc 122:6124
278. Ješelnik M, Plavec J, Polanc S, Kočevar M (2000) Carbohydr Res 328:591
279. Myers AG, Liang J, Hammond M, Harrington PM, Wu Y, Kuo EY (1998) J Am Chem Soc 120:5319
280. Wang ZX, Shi XX, Chen GR, Ren ZH, Luo L, Yan J (2006) Carbohydr Res 341:1945
281. Rao Y, Venot A, Swayze EE, Griffey RH, Boons GJ (2006) Org Biomol Chem 4:1328
282. Blattner R, Furneaux RH, Pakulski Z (2006) Carbohydr Res 341:2115
283. Ferguson JR, Harding JR, Lumbard KW, Scheinmann F, Stachulski AV (2000) Tetrahedron Lett 41:389
284. Ferguson JR, Harding JR, Kilick DA, Lumbard KW, Scheinmann F, Stachulski AV (2001) J Chem: Soc Perkin Trans 1:3037
285. Brown RT, Carter NE, Mayalarp SP, Scheinmann F (2000) Tetrahedron 56:7591
286. Suzuki T, Mabuchi K, Fukazawa N (1999) Bioorg Med Chem Lett 9:659
287. Engstrom KM, Daanen JF, Wagaw S, Stewart AO (2006) J Org Chem 71:8378
288. Kawada T, Asano R, Makino K, Sakuno T (2000) Eur J Org Chem 2723
289. Feldman KS, Lawlor MD (2000) J Am Chem Soc 122:7396
290. Nicolaou KC, Pfefferkorn JA, Cao GQ (2000) Angew Chem Int Ed 39:734
291. Schramm S, Dettner K, Unverzagt C (2006) Tetrahedron Lett 47:7741
292. Hein M, Michalik D, Langer P (2005) Synthesis 3531
293. Jiang ZH, Geyer A, Schmidt RR (1995) Angew Chem Int Ed Engl 34:2520
294. Jiang ZH, Schmidt RR (1994) Liebigs Ann Chem 645
295. Fürstner A, Jeanjean F, Razon P (2002) Angew Chem Int Ed 41:2097
296. Brito-Arias M, Pereda-Miranda R, Heathcock CH (2004) J Org Chem 69:4567
297. Fürstner A, Albert M, Mlynarski J, Matheu M (2002) J Am Chem Soc 124:1168
298. Fürstner A, Mlynarski J, Albert M (2002) J Am Chem Soc 124:10274
299. Fürstner A, Ruiz-Caro J, Prinz H, Waldmann H (2004) J Org Chem 69:459
300. Hostettmann K, Marston A (1995) Saponins, Chemistry and Pharmacology of Natural Products, University Press, Cambridge
301. Chwalek M, Plé K, Voutquenne-Nazabadioko L (2004) Chem Pharm Bull 52:965
302. Deng S, Yu B, Xie J, Hui Y (1999) J Org Chem 36:7265
303. Plé K, Chwalek M, Voutquenne-Nazabadioki L (2004) Eur J Org Chem 1588
304. Thompson MJ, Hutchinson EJ, Stratford TH, Bowler WB, Blackburn GM (2004) Tetrahedron Lett 45:1207
305. Ikeda T, Yamauchi K, Nakano D, Naknishi K, Miyashita H, Ito S, Nohara T (2006) Tetrahedron Lett 47:4355
306. Ikeda T, Miyashita H, Kajimoto T, Nohara T (2001) Tetrahedron Lett 42:2353
307. Zhu X, Yu B, Hui Y, Schmidt RR (2004) Eur J Org Chem 965
308. Schimmel J, Passos Eleutério MI, Ritter G, Schmidt RR (2006) Eur J Org Chem 1701
309. Passos Eleutério MI, Schimmel J, Ritter G, do Céu Costa M, Schmidt RR (2006) Eur J Org Chem 5293
310. Wang P, Kim YJ, Navarro-Villalobos M, Rohde BD, Gin DY (2005) J Am Chem Soc 127:3256
311. Kim YJ, Wang P, Navarro-Villalobos M, Rohde BD, Derryberry JM, Gin DY (2006) J Am Chem Soc 128:11906
312. Sawada N (2004) Synthesis of Deoxy-QS-21A, Universität Konstanz, to be published
313. Dumont-Hornebeck B, Joly JP, Coulon J, Chapleur Y (1999) Carbohydr Res 320:147
314. Rele SM, Cui W, Wang L, Hou S, Barr-Zarse G, Tatton D, Gnanou Y, Esko JD, Chaikof EL (2005) J Am Chem Soc 127:10132

315. Ikuta A, Tanimoto T, Koizumi K (2003) J Carbohydr Chem 22:297
316. Ikuta A, Mizuta N, Kitahata S, Murata T, Usui T, Koizumi K, Tanimoto T (2004) Chem Pharm Bull 52:51
317. Smiljanic N, Halila S, Moreau V, Djedaïni-Pilard F (2003) Tetrahedron Lett 44:8999
318. Damager I, Olsen CE, Møller BL, Motawia MS (1999) Carbohydr Res 320:19
319. Wu Z, Kong F (2004) Synlett 2594
320. Levy DE, Tang C (1995) In: The Chemistry of *C*-Glycosides, Pergamon, Elmsford NY
321. Dondoni A, Marra A, Massi A (1999) J Org Chem 64:933
322. Armitt DJ, Banwell MG, Freeman C, Parish CR (2002) J Chem Soc Perkin Trans 1:1743
323. Herzner H, Palmacci ER, Seeberger PH (2002) Org Lett 4:2965
324. Furuta T, Kumura T, Kondo S, Mihara H, Wakimoto T, Nukaya H, Tsuji K, Tanaka K (2004) Tetrahedron 60:9375
325. Ali IAI, El-Ashry ESH, Schmidt RR (2004) Tetrahedron 40:4773; and references therein
326. Fuchss T, Schmidt RR (2000) J Carbohydr Chem 19:677; and references therein
327. Tsuruta O, Yuasa H, Hashimoto H, Kuromo S, Yazawa S (1999) Bioorg Med Chem Lett 9:1019
328. Izumi M, Tsuruta O, Kajihara Y, Yazawa S, Yuasa H, Hashimoto H (2005) Chem Eur J 11:3032
329. Ohara K, Matsuda H, Hashimoto M, Miyairi K, Okuno T (2002) Chem Lett 626
330. Morii Y, Matsuda H, Ohara K, Hashimoto M, Miyairi K, Okuno T (2005) Bioorg Med Chem 13:5113
331. Schmidt RR, Michel J, Roos M (1984) Liebigs Ann Chem 1343
332. Schmelzer U (1997) PhD thesis, Universität Konstanz
333. Schmelzer U, Zhang Z, Schmidt RR (2007) J Carbohydr Chem (submitted)
334. Huchel U (1998) PhD thesis, Universität Konstanz
335. Huchel U, Schmidt C, Schmidt RR (1995) Tetrahedron Lett 36:9457
336. Huchel U, Schmidt C, Schmidt RR (1998) Eur J Org Chem 1353
337. Hanessian S, Condé JJ, Lion B (1995) Tetrahedron Lett 36:5865
338. Hanessian S (1997) In: Hanessian S (ed) Preparative Carbohydrate Chemistry. Marcel Dekker, New York, p 381–388; and references therein
339. Yu B, Tao H (2001) Tetrahedron Lett 42:2405
340. Yu B, Tao H (2002) J Org Chem 67:9099

3.3 Further Anomeric Esters

Kwan Soo Kim, *Heung Bae Jeon*
Center for Bioactive Molecular Hybrids and Department of Chemistry, Yonsei University, Seoul 120–749, Korea
kwan@yonsei.ac.kr

Abstract

The most representative anomeric ester glycosyl donor is glycosyl acetate, which seems to be a promising and useful glycosylating reagent due to a variety of advantageous properties. Several other anomeric acyl and carbonate-type leaving groups have been devised as new glycosyl donors. In particular, the glycosyl acetate can be the precursor for glycosylating agents of other types, for example, glycosyl halides, 1-thioglycosides, and glycosyl trichloroacetimidates. Therefore, use of glycosyl acetate can make glycosylation reactions easier and less expensive. Furthermore, the intramolecular glycosylation method through the 1-*O*-acyl linkage and the glycosylation with glycosyl donors having remote acyl groups from the anomeric center have been developed to increase the efficiency and the stereoselectivity.

Keywords

Glycosylation; Glycosyl donor; Anomeric leaving group; Anomeric ester; Glycosyl acetate; Glycosyl carbonate; 1-*O*-acyl linkage; Remote acyl group

Abbreviations

CB	2′-carboxybenzyl
DCC	1,3-dicyclohexylcarbodiimide
DIC	diisopropylcarbodiimide
Dmob	*N*-2,4-dimethoxybenzyl
IDCP	iodonium dicollidine perchlorate
LPS	lipopolysaccharide
NIS	*N*-iodosuccinimide
TBSOTf	*tert*-butyldimethylsilyl trifluoromethanesulfonate
TMSOTf	trimethylsilyl trifluoromethanesulfonate

In: *Glycoscience*. Fraser-Reid B, Tatsuta K, Thiem J (eds)
Chapter-DOI 10-1007/978-3-540-30429-6_12:

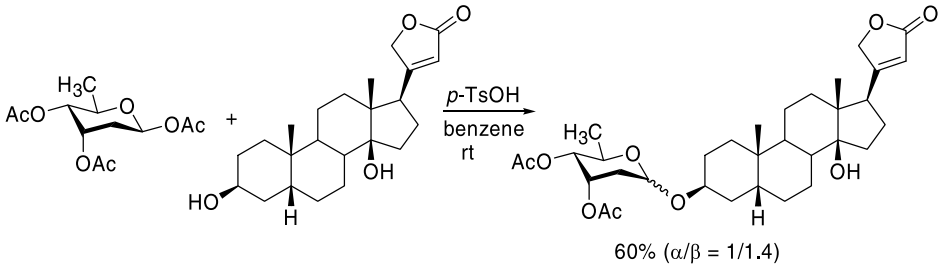
1 Introduction

An advantage of the 1-*O*-acylated glycosyl donors in the glycosylation method is undoubtedly the ease of their preparation. The most representative anomeric functional group in this area is the acetyl group. Glycosyl acetates seem to be promising and useful glycosylating reagents due to a variety of advantageous properties. Among these, the most important are their availability and chemical stability, which allows one to obtain almost any desired amounts of glycosyl acetates which can be stored for unlimited time without special precautions. Furthermore, glycosyl acetates can be precursors for glycosylating agents of other types, for example, glycosyl halides, 1-thioglycosides, and glycosyl trichloroacetimidates. Therefore, the use of glycosyl acetates can make glycosylation reactions easier and less expensive.

Because of the relatively low reactivity of glycosyl acetates they used to occasionally bring about difficulty in controlling stereoselectivity. To overcome the problem associated with glycosyl acetates, several other anomeric acyl leaving groups, for example, ester-type leaving groups such as haloacetyl, (2-methoxyethoxy)acetyl, benzoyl, *p*-nitrobenzoyl, 2-pyridinecarbonyl, pivaloyl, 4-pentenoyl, and 5-hexynoyl, and carbonate-type leaving groups such as phenoxycarbonyl, isopropenyloxycarbonyl, 1-imidazolylcarbonyl, *N*-allyl carbamoyl, sulfonylcarbamoyl, and 2-pyridyl thiocarbonyl, have been devised as new glycosyl donors. Furthermore, the methods of intramolecular glycosylation through the 1-*O*-acyl linkage and glycosylation with glycosyl donors having acyl groups remote from the anomeric center have been developed to increase efficiency and stereoselectivity.

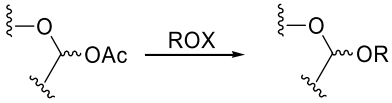
2 Glycosylation with Glycosyl Acetates

Generally the activation of glycosyl acetates is achieved in the presence of a Lewis acid. Since Helferich et al. introduced the use of glycosyl acetates as glycosyl donors in the reaction with phenol in the presence of *p*-toluenesulfonic acid (TsOH) or $ZnCl_2$ in 1933 [1], several Lewis acids have been utilized as effective promoters in the glycosylation with glycosyl acetates. The representative promoters used for the activation of glycosyl acetate donors for the glycosylation are summarized in ❯ *Table 1*. TsOH was also used as a promoter of peracetylated 2-deoxy-L-fucose for the glycosylation of digitoxigenin (❯ *Scheme 1*) [2].

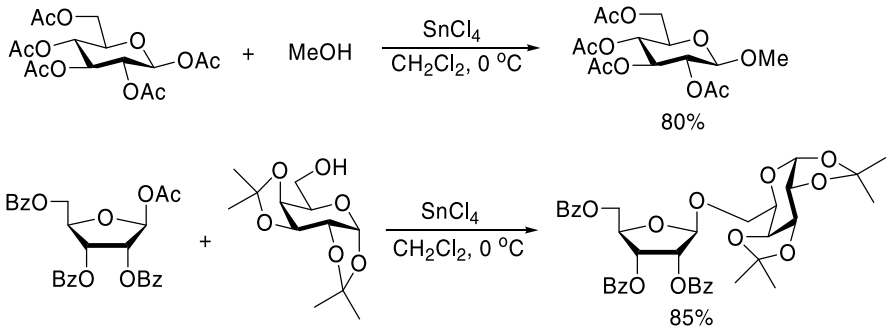


Scheme 1
(Ref. [2])

Table 1
Glycosylation with 1-*O*-acetyl sugar



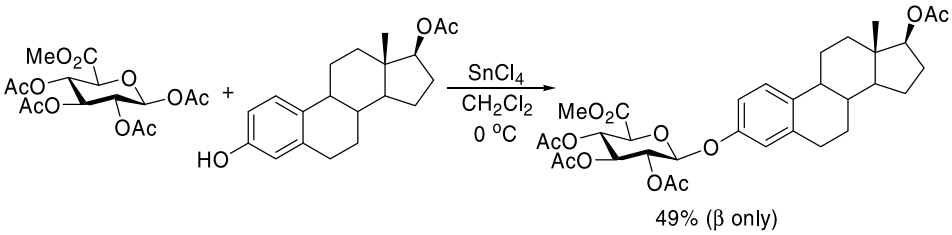
Promoters	X	References
TsOH	H	[1,2]
$ZnCl_2$	H	[1]
$SnCl_4$	H	[3,4,5,6,7,8]
$FeCl_3$	H	[9,10,11,12]
$BF_3{\cdot}Et_2O$	H	[13,14,15,16,17,18,19,20,21]
$BF_3{\cdot}Et_2O/Bi(OTf)_3$	H	[22]
TMSOTf	H	[23,24,25,26,27,28,29,30,31,32,33,34,35,36,37,40,41,42]
TBSOTf	H	[35,38,39]
$TrClO_4$	H	[43,44]
$SnCl_4/Sn(OTf)_2/LiClO_4$	TMS	[45]
$SnCl_4/AgClO_4$	TMS	[46,47,49]
$GaCl_3/AgClO_4$	TMS	[47]
$Me_2SiCl_2/AgClO_4$	H	[48]
K-10 montmorillonite	H	[50]
$Yb[N(O_2SC_4F_9)_2]_3$	H	[51]
$TMSCl/Zn(OTf)_2$	H	[52,53]



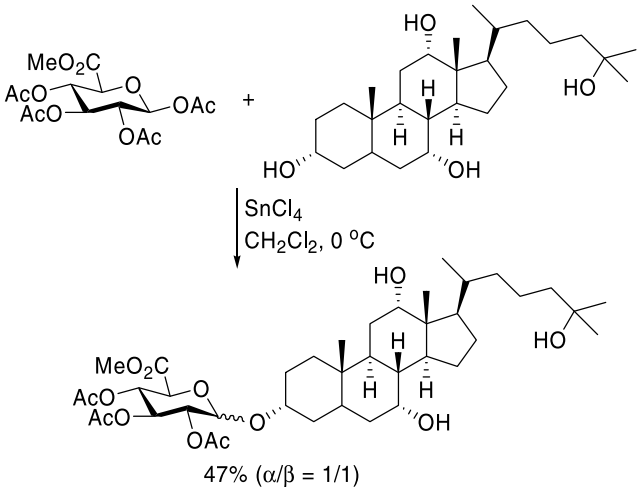
Scheme 2
(Ref. [3,4])

Lemieux and Shyluk reported that the reaction of peracetylated glucose and methanol in the presence of $SnCl_4$ gave methyl tetraacetyl-β-glucopyranoside in moderate yield [3]. Later, Hanessian and Banoub reported that the glycosylation using $SnCl_4$ for the activation of glycosyl acetate donors at low temperature afforded the corresponding β-*O*-glycosides in high yield (*Scheme 2*) [4]. Furthermore, $SnCl_4$ was employed as a promoter for *O*-glucuronylation of some phenols and alcohols including estradiol 17-acetate and estradiol 16,17-diacetate with methyl 1,2,3,4-tetra-*O*-acetyl-β-D-glucopyranuronate (*Scheme 3*) [5]. And, Williams et al. developed a mild and regiocontrolled method for the *O*-glucuronylation of the highly reac-

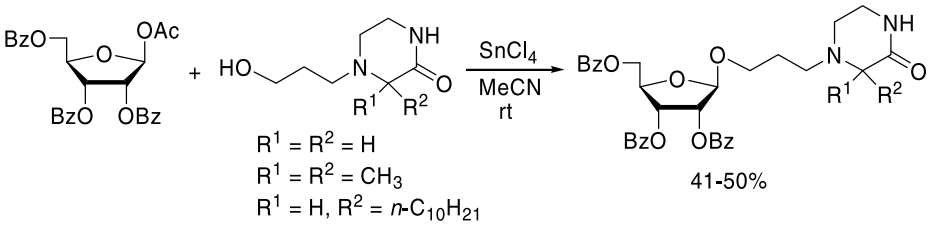
tive 3-hydroxyl group of 5β-cholestane-3α,7α,12α,25-tetrol using $SnCl_4$ in order to study the transformations of bile alcohols and evaluate the biological effects on the blood-brain barrier (*Scheme 4*) [6]. Krausz et al. reported the synthesis of bioactive nucleoside analogues, which have a spacer arm between the sugar and the base moieties, by $SnCl_4$-mediated glycosylation of 3-alkyl N^4-(3-hydroxylpropyl) 2-piperazinones with protected 1-*O*-acetyl ribofura-



Scheme 3
(Ref. [5])



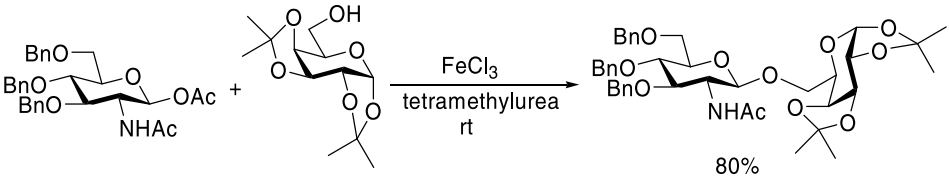
Scheme 4
(Ref. [6])



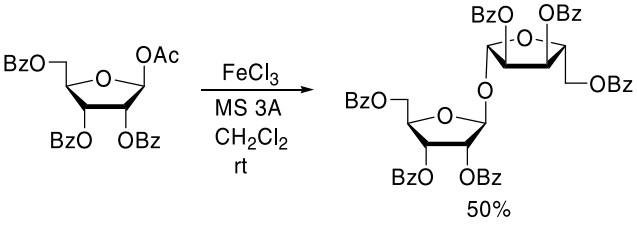
Scheme 5
(Ref. [7])

noses (Scheme 5) [7] and the synthesis of novel *meso*-glycosylarylporphyrins by the similar methodology [8].

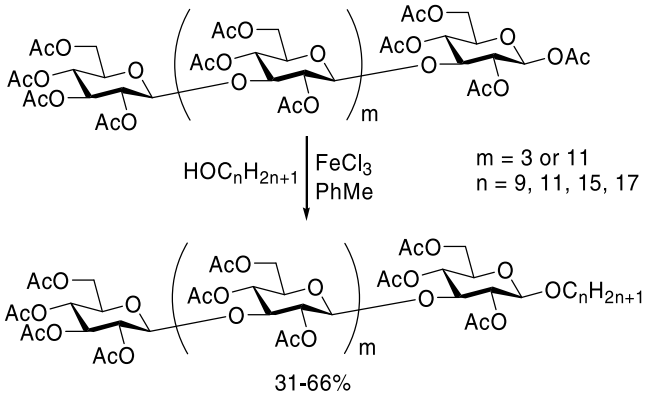
Iron(III) chloride, $FeCl_3$, was introduced for glycosylation with 1-*O*-acetyl-2-acylamido-2-deoxy-β-D-glucopyranose by Kiso and Enderson (Scheme 6) [9], and later, Lerner used $FeCl_3$ in combination with molecular sieves (3Å) for the dimerization of 1-*O*-acetyl-2,3,5-tri-*O*-benzoyl-β-D-ribofuranose to afford a nonreducing disaccharide derivative (Scheme 7) [10]. Uryu and coworkers reported the use of $FeCl_3$ as a promoter for the glycosylation of aliphatic alcohols with malto-hexaoside and malto-heptaoside peracetates (Scheme 8) [11]. On the other hand, while peracetylated sugar donors are typically



Scheme 6
(Ref. [9])



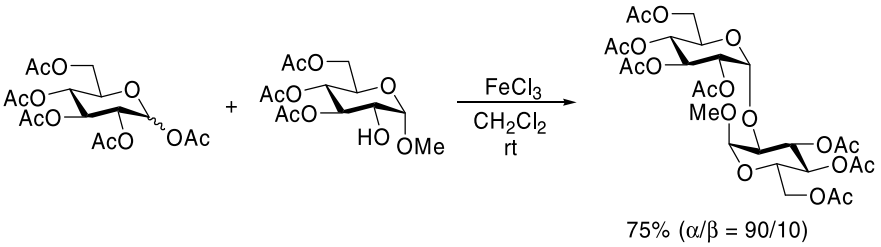
Scheme 7
(Ref. [10])



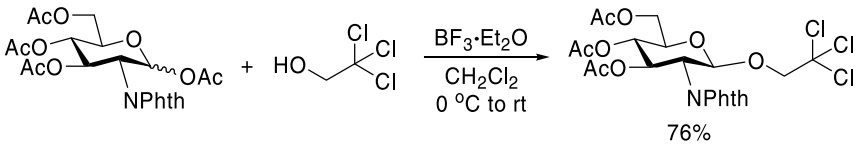
Scheme 8
(Ref. [11])

employed for the 1,2-*trans* glycosylation, it was reported that the glycosylation of a series of alcohol acceptors with peracetylated sugars in the presence of $FeCl_3$ provided stereoselectively 1,2-*cis*-α-glycosides, which might be generated by the anomerization of the 1,2-*trans*-β-glycopyranoside by $FeCl_3$ (Scheme 9) [12].

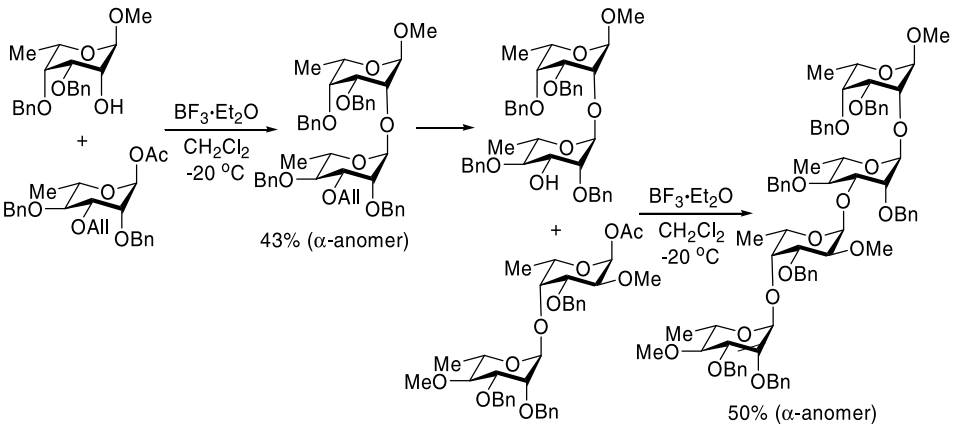
Magnusson et al. introduced $BF_3{\cdot}Et_2O$ as a promoter for the formation of alkyl glycosides from 1-*O*-acetyl-2-deoxy-2-phthalimidoglucopyranose (Scheme 10) [13]. And then, $BF_3{\cdot}Et_2O$ was applied for the synthesis of a number of oligosaccharides as an activator of glycosyl acetate donors. Gurjar et al. applied $BF_3{\cdot}Et_2O$ for the synthesis of the oligo-



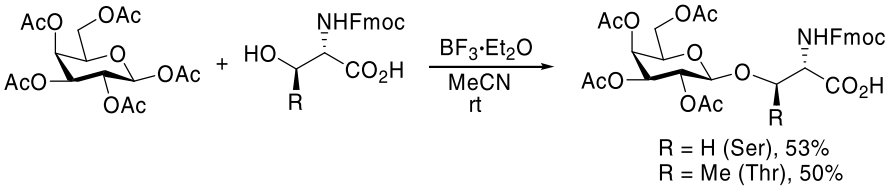
Scheme 9
(Ref. [12])



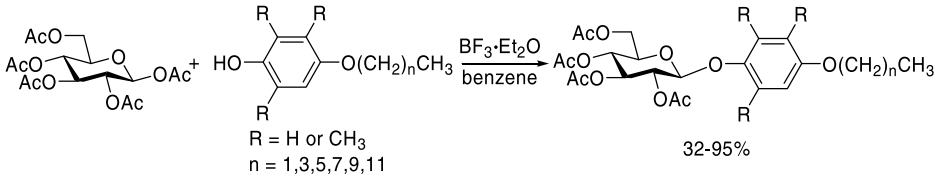
Scheme 10
(Ref. [13])



Scheme 11
(Ref. [14])



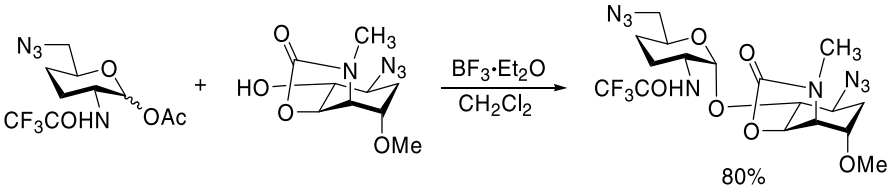
Scheme 12
(Ref. [15])



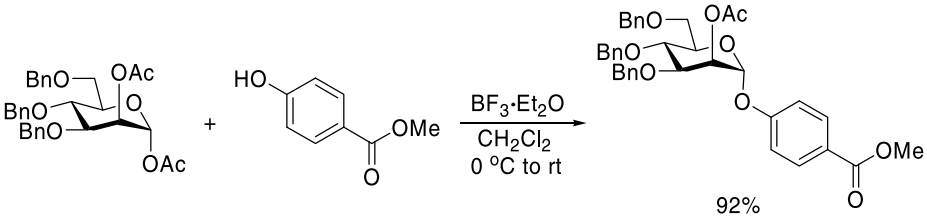
Scheme 13
(Ref. [18])

saccharide segment of the glycopeptidolipid antigen of mycobacterium avium serotype 4 (Scheme 11) [14]. Elofsson et al. discovered that glycosyl acetate donors could be used for the glycosylation of amino acids, in which the carboxylic acid was unprotected, in the presence of $BF_3{\cdot}Et_2O$ (Scheme 12) [15], while Steffan et al. showed that coupling of peracetylated glucose with *N*-Fmoc-protected serine by the same methodology provided the corresponding glycosylated amino acid [16]. This was a major breakthrough, because the approach circumvents the protection-deprotection procedure of the carboxylic acid prior to incorporation into a peptide. Peracetylated glucose and $BF_3{\cdot}Et_2O$ were also utilized for the glycosylation of protected serine and tyrosine to make *O*-glycopeptides [17]. On the other hand, Satoh et al. used $BF_3{\cdot}Et_2O$ as an activator of peracetylated glucose in order to prepare 4-alkoxyphenyl β-D-glucopyranosides, showing strong inhibitory effects on the histamine release from rat peritoneal mast cells induced by concanavalin A (Scheme 13) [18]. A similar methodology was applied for the preparation of sannamycin-type aminoglycoside antibiotics from the purpurosamine *C*-type glycosyl acetate donor and the sannaime-type acceptor by Prinzbach et al. (Scheme 14) [19]. Gervay-Hague and Lam also reported the use of $BF_3{\cdot}Et_2O$ for the glycosylation of methyl paraben with protected mannose 1-*O*-acetate to obtain the corresponding α-mannopyranoside (Scheme 15) [20]. And, $BF_3{\cdot}Et_2O$ was used for the glycosylation with the *N*-2,4-dimethoxybenzyl (Dmob) protected 2-acetamido-1-*O*-acetate glucosyl donor, which gave higher glycosylation yields than the corresponding 2-acetamido glucosyl donor without Dmob protection (Scheme 16) [21].

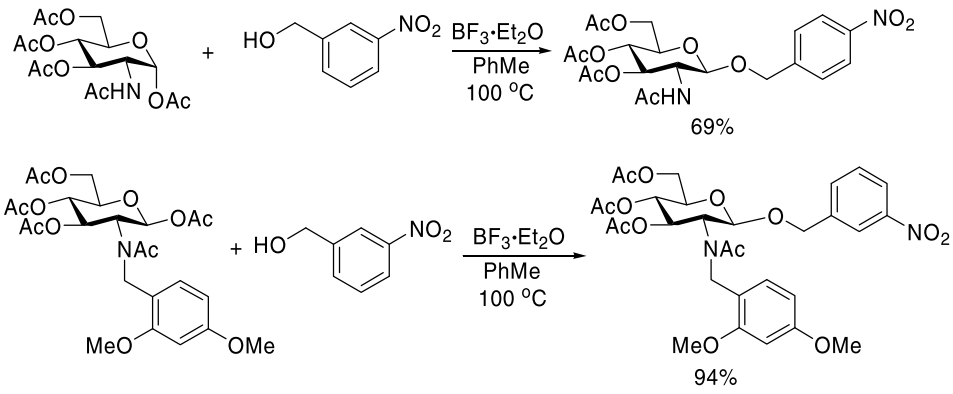
Ikeda et al. reported that a combined system of bismuth triflate [$Bi(OTf)_3$] and $BF_3{\cdot}Et_2O$ in dichloromethane is an efficient promoter for the glycosylation of sialyl acetates among various Lewis acids, for example, $BiCl_3$, $Sc(OTf)_2$, $Yb(OTf)_3$, $Zn(OTf)_2$, and TMSCl, and their combinations, which showed the first example of utilization of $Bi(OTf)_3$ in the glycosylation of common sialic acid derivatives (Scheme 17) [22].



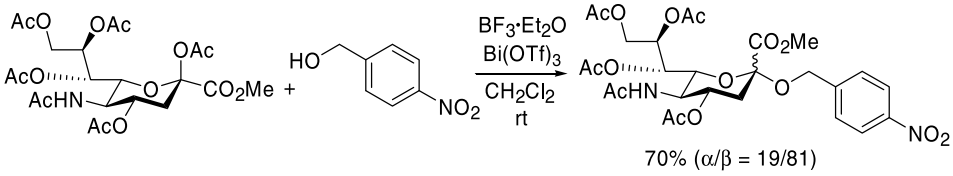
Scheme 14
(Ref. [19])



Scheme 15
(Ref. [20])

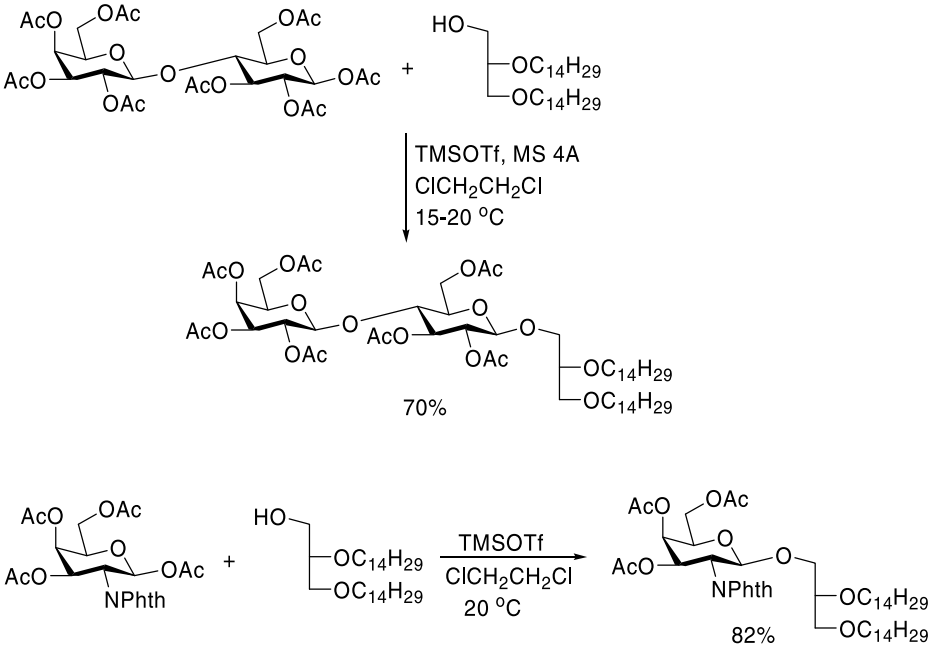


Scheme 16
(Ref. [21])

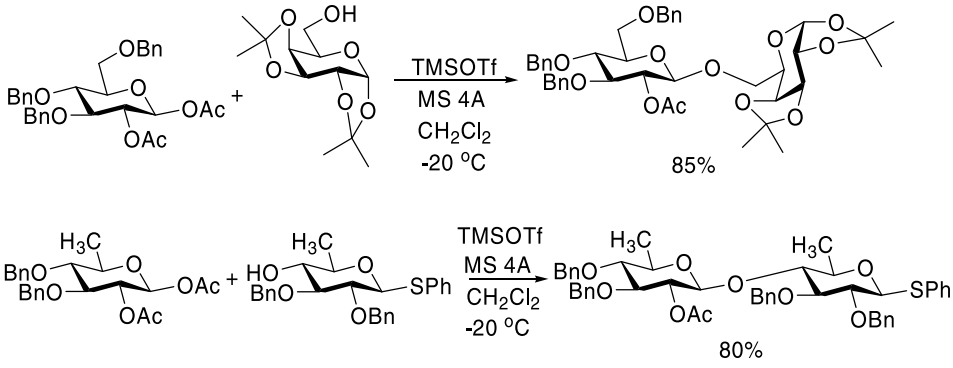


Scheme 17
(Ref. [22])

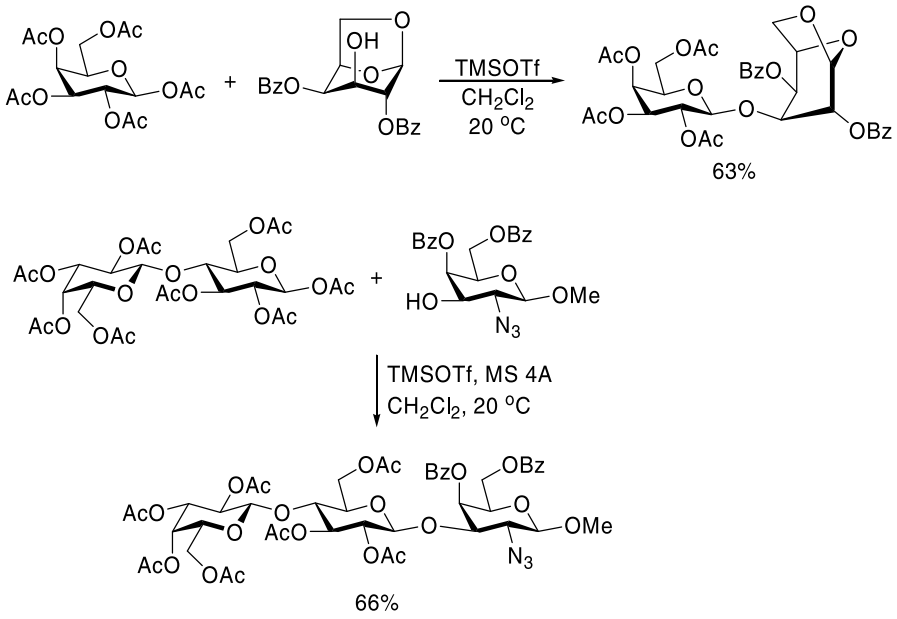
Ogawa and coworkers reported that trimethylsilyl trifluoromethanesulfonate (TMSOTf) was a very effective promoter for the glycosylation of a glycerolipid alcohol with octa-*O*-acetyl-β-lactose [23] and with 1,3,4,6-tetra-*O*-acetyl-2-deoxy-2-phthalimido-β-D-galactopyranose [24] to synthesize model glycoglycerolipids (*Scheme 18*). However, glycosylations with glycosyl acetates in the presence of TMSOTf sometimes accompany side reactions such as transfer of an acyl group from the glycosyl donor to the glycosyl acceptor [25],



Scheme 18
(Ref. [23,24])

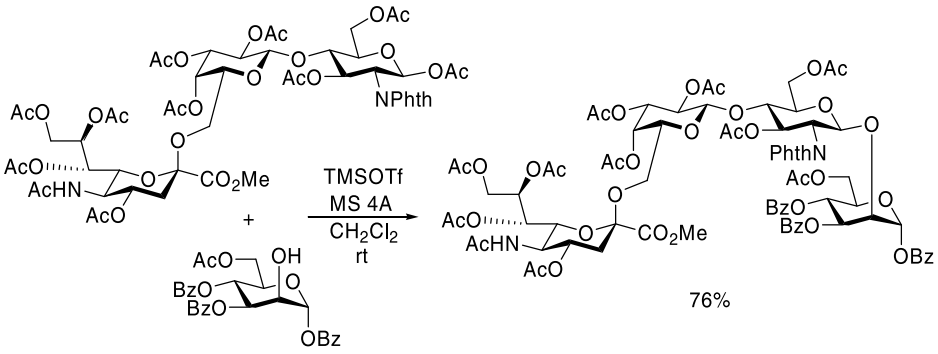


Scheme 19
(Ref. [28])

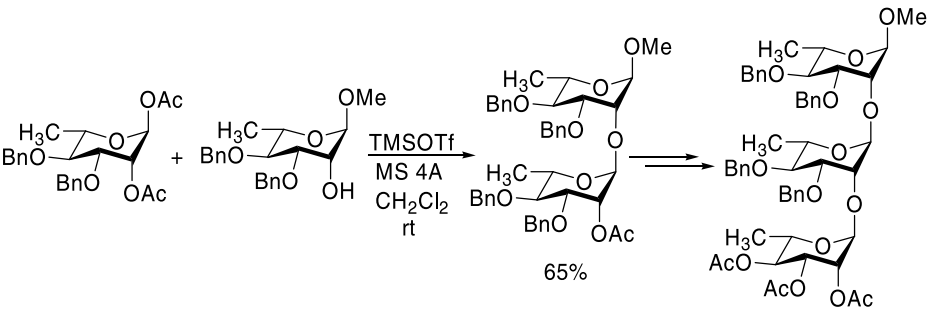


Scheme 20
(Ref. [29])

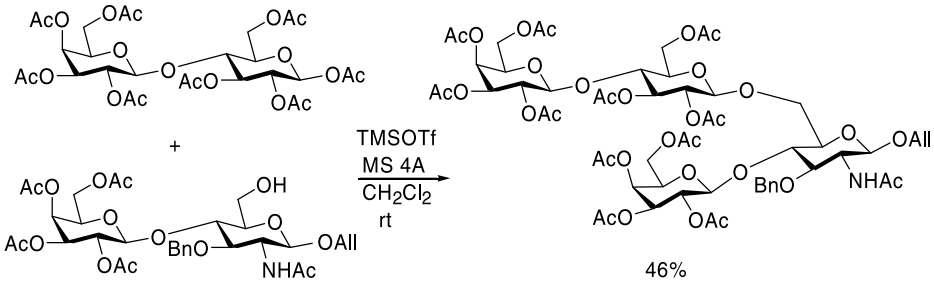
deacetylation [25b], and silylation of the glycosyl acceptor [26]. Thus, Nifant'ev et al. studied the scope and limitation of the use of β-glycosyl acetates as glycosyl donors in the presence of TMSOTf [27]. In spite of these problems, since its first introduction by Ogawa, TMSOTf has been widely used as the most preferred activator of glycosyl acetates for glycoside synthesis. Sinaÿ et al. reported that 1,2-*trans*-diacetates as glycosyl donors with various nonacylated glycosyl acceptors in the presence of TMSOTf gave the corresponding β-disaccharides in high yields (*Scheme 19*) [28]. And, Paulsen and Paal reported the use of TMSOTf in reactions of various glycosyl acetate donors with secondary hydroxyl groups of saccharides having low reactivity to obtain the corresponding di- and oligosaccharide in good yields (*Scheme 20*) [29]. They also used TMSOTf in the final coupling reaction for the synthesis of *N*-acetylneuraminic acid-containing trisaccharide (*Scheme 21*) [30]. TMSOTf was also employed as an efficient activator for the rhamnosylation reaction with 1,2-di-*O*-acetyl-3,4-di-*O*-benzyl-α-L-rhamnopyranose in synthesizing the rhamnotriose moiety of the common antigen among oligosaccharides related to Group B Streptococcal polysaccharides by Jennings et al. (*Scheme 22*) [25b]. They also reported the synthesis of a tri- and a tetrasaccharide fragment of the capsular polysaccharide of type III Group B *Streptococcus* by a similar methodology (*Scheme 23*) [25c], [31]. On the other hand, Scharf et al. showed another example of the TMSOTf-mediated rhamnosylation with 1,4-di-*O*-acetyl-3-*O*-methyl-2-*O*-pivaloyl-α-L-rhamnopyranose for the synthesis of a tetradeoxydisaccharide found in the avermectin family (*Scheme 24*) [32], and Lafont et al. reported the TMSOTf-mediated glycosylation of 1,6-anhydro-β-D-mannopyranose derivatives with peracetylated *N*-allyloxycarbonyl-β-D-glucosamine for the synthesis of glycan fragments of glycoproteins (*Scheme 25*) [33].



Scheme 21
(Ref. [30])

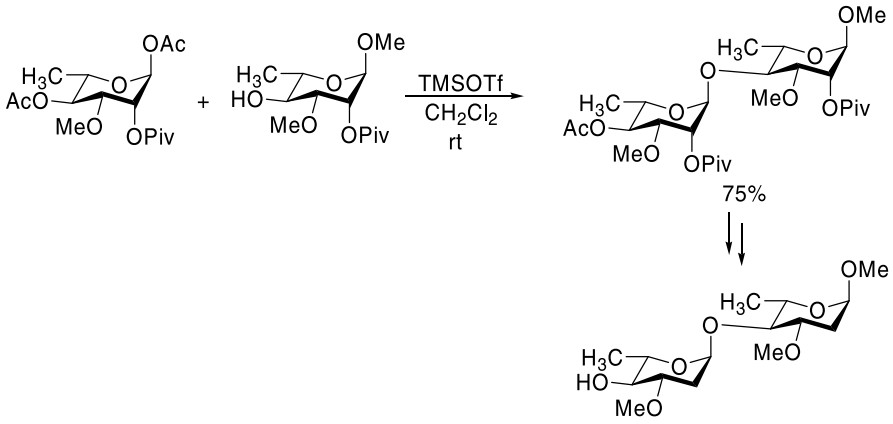


Scheme 22
(Ref. [25b])

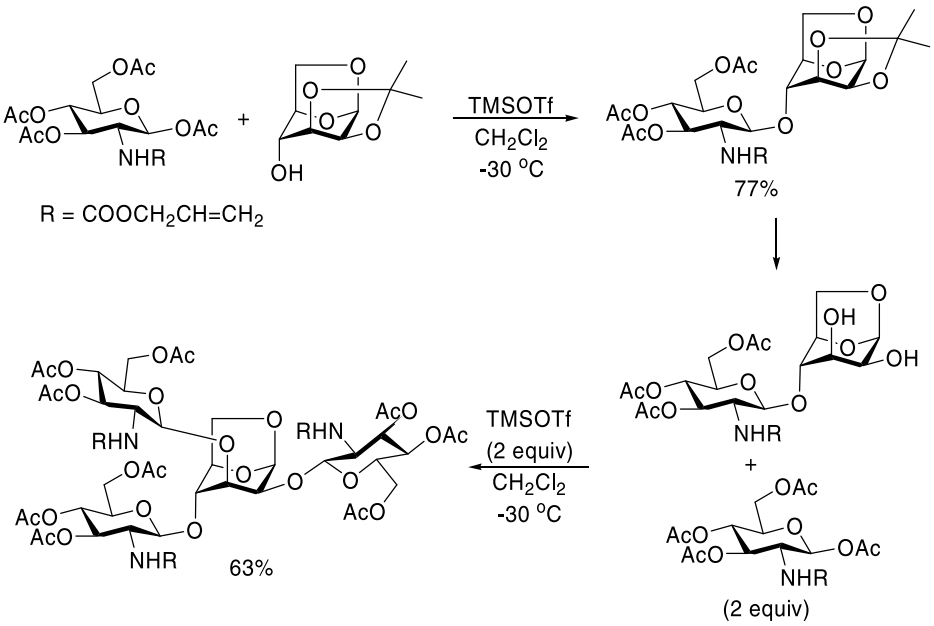


Scheme 23
(Ref. [25c], [31])

On the other hand, Roush and coworkers reported the synthesis of the functionalized C-D-E trisaccharide precursor of Olivomycin A, in which, particularly, the 2-deoxy-2-iodo-3-methyl-α-L-mannopyranosyl linkage was stereoselectively introduced by using the glycosyl acetate donor (*Scheme 26*) [34]. This new α-glycosylation protocol was applied to the synthesis

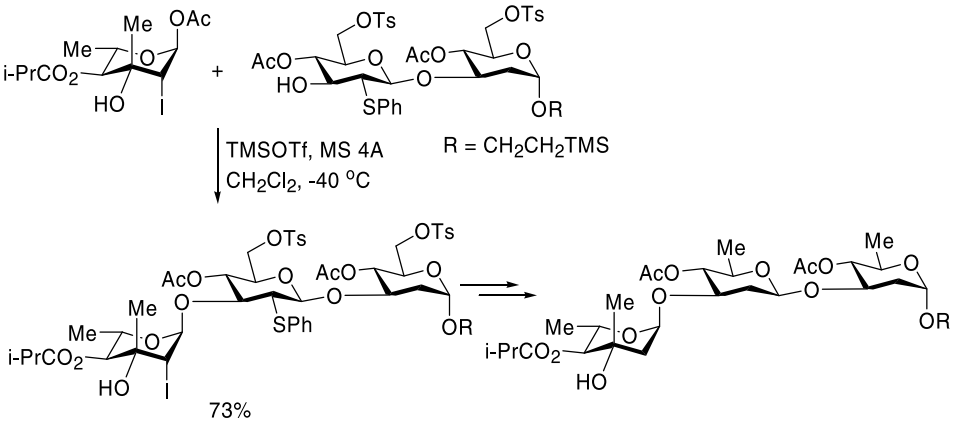


Scheme 24
(Ref. [32])

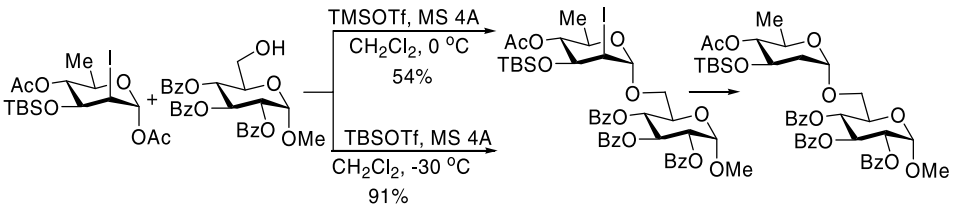


Scheme 25
(Ref. [33])

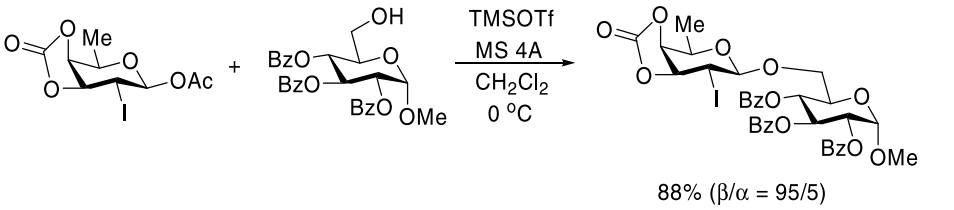
of 2-deoxy-α-D-mannosides and -talosides by reduction of 2-deoxy-2-iodo-α-glycosides generated from the α-stereoselective glycosylation with 2-deoxy-2-iodo glycosyl acetates in the presence of TMSOTf or TBSOTf as the promoter (*Scheme 27*) [35]. They also demonstrated the stereoselective synthesis of 2-deoxy-β-galactosides using 2-deoxy-2-iodo-galactopyra-



Scheme 26
(Ref. [34])

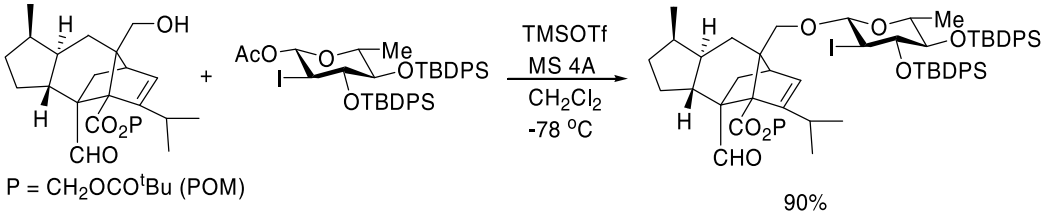


Scheme 27
(Ref. [35])

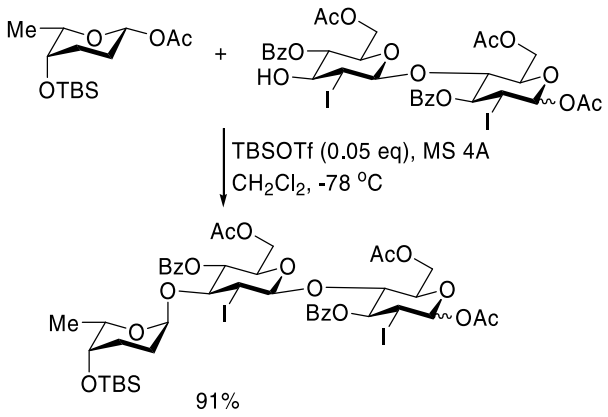


Scheme 28
(Ref. [36])

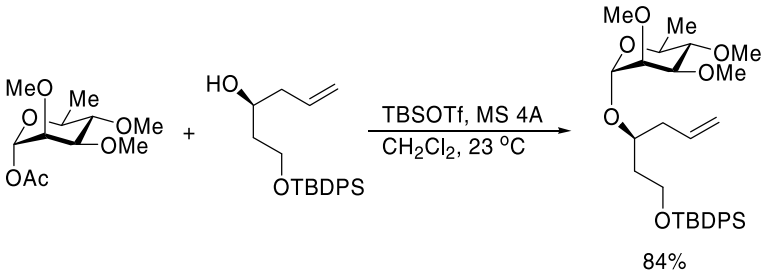
nosyl acetate donors by a similar methodology (*Scheme 28*) [36]. Kaneko et al. reported an approach to the stereoselective semi-synthesis of GM-237354, a potent inhibitor of fungal elongation factor 2 (EF-2), by employing a highly β-selective glycosylation using 2-deoxy-2-iodo-glycopyranosyl acetate donor developed by the Roush group (*Scheme 29*) [37]. Roush et al. also reported the use of TBSOTf as an activator for the glycosylation reactions with L-rhodinosyl (2,3,6-trideoxy-L-hexosyl) acetate derivatives in synthesizing the repeat A-B-C trisaccharide unit of Landomycin A (*Scheme 30*) [38], and for the glycosylation reaction with α-L-rhamnopyranosyl acetate in the early stage of the synthesis of (–)-spinosyn A (*Scheme 31*) [39].



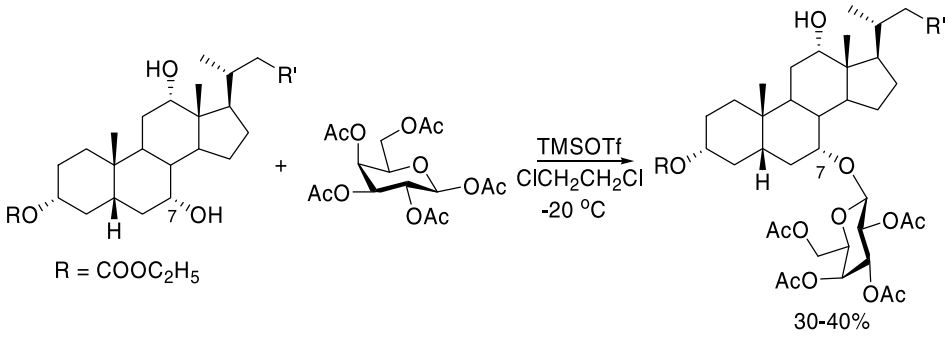
Scheme 29
(Ref. [37])



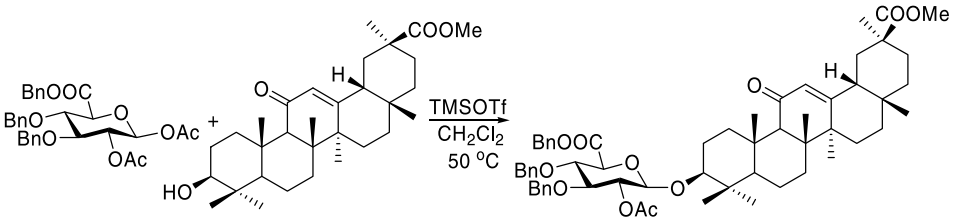
Scheme 30
(Ref. [38])



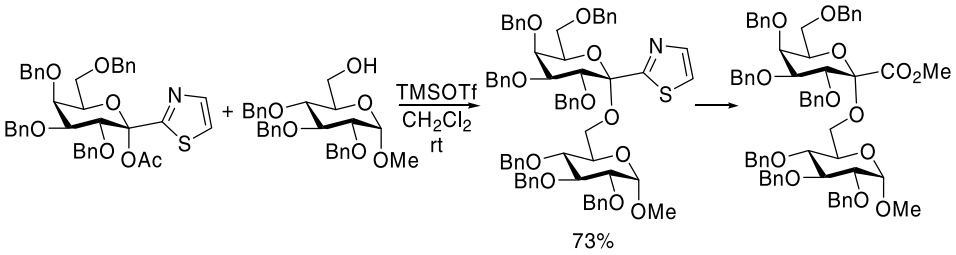
Scheme 31
(Ref. [39])



Scheme 32
(Ref. [40])



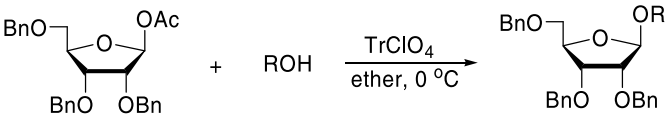
Scheme 33
(Ref. [41])



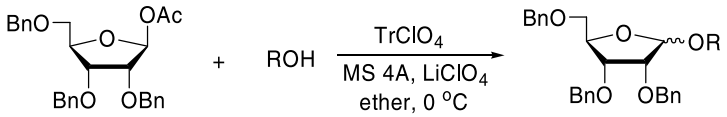
Scheme 34
(Ref. [42])

Nakanishi and coworkers also reported the TMSOTf-mediated glycosylation of a cholic acid derivative with galactose pentaacetate for the synthesis of mosesin-4, a naturally occurring steroid saponin with shark repellent activity, in which the severely hindered 7α-position was selectively glycosylated (*Scheme 32*) [40]. Satto et al. reported that the glycosylation of methyl glycyrrhetinate with methyl D-glucuronatopyranose 1-*O*-acetate in the presence of TMSOTf afforded the corresponding glycoside, a precursor of the glycyrrhetic acid glycoside, although the stereoselectivity was not satisfactory (*Scheme 33*) [41]. On the other hand, Dondoni et al. introduced the TMSOTf-promoted glycosylation with thiazolyketol acetates as glycosyl donors for the stereoselective synthesis of α-linked ketodisaccharides (*Scheme 34*) [42].

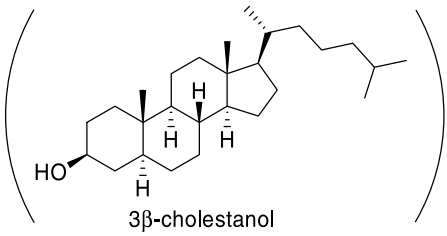
Trityl perchlorate ($TrClO_4$) was found to be an effective promoter for the anomeric acetates by Mukaiyama and coworkers. They reported that the $TrClO_4$-promoted glycosylations of alcohols with 1-*O*-acetyl-β-D-ribose in ether afforded the corresponding β-ribosides exclusively, while the same reaction in the presence of molecular sieves (4Å) and additives such as $LiClO_4$ provided α-ribosides predominantly in good yields (*Scheme 35*) [43]. Evans and coworkers modified the Mukaiyama's stoichiometric $TrClO_4$ glycosylation by employing a catalytic amount of $TrClO_4$ [44] and applied the modified method to the glycosylation with glycosyl acetates for the total synthesis of Cytovaricin [44a] and Lepicidin A [44b]. Thus, the catalytic $TrClO_4$ (ca. 5%)-promoted glycosylation with 2,3,4-tri-*O*-methyl-α-D-rhamnosyl acetate in toluene provided the α-glycoside in 87% yield along with 5% of the undesired β-anomer, and the α-anomer was further transformed into the complex macrolide insecticide, Lepicidin A (*Scheme 36*) [44b].



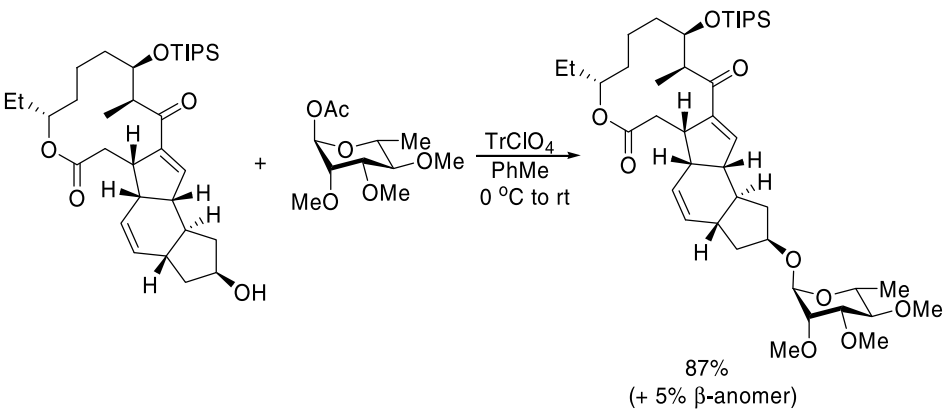
ROH = 1-octadecanol, 95% (β only)
cyclohexanol, 83% (β only)
3β-cholestanol, 88% (β only)



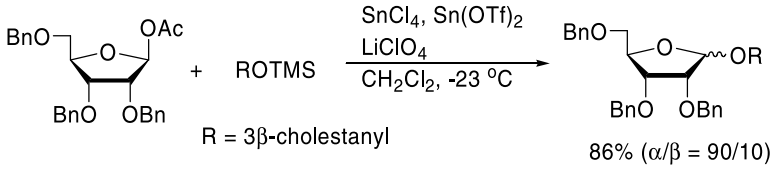
ROH = 2-propanol, 91% (α/β = 73/27)
cyclohexanol, 86% (α/β = 70/30)
3β-cholestanol, 75% (α/β = 79/21)



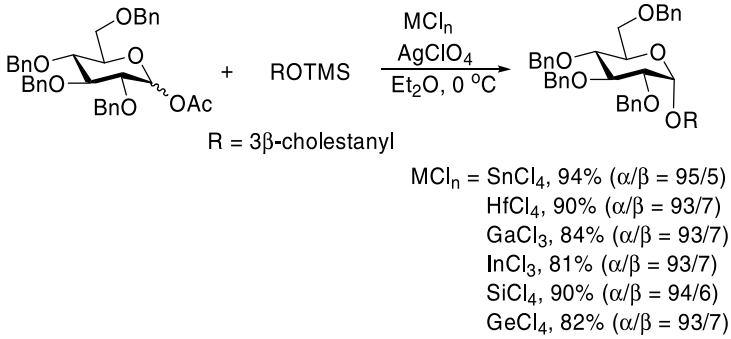
Scheme 35
(Ref. [43])



Scheme 36
(Ref. [44b])

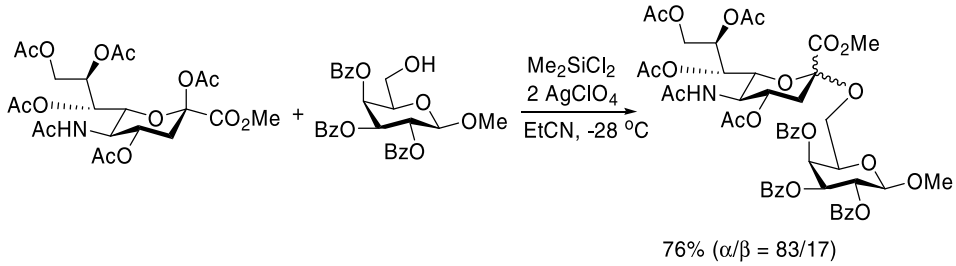


Scheme 37
(Ref. [45])



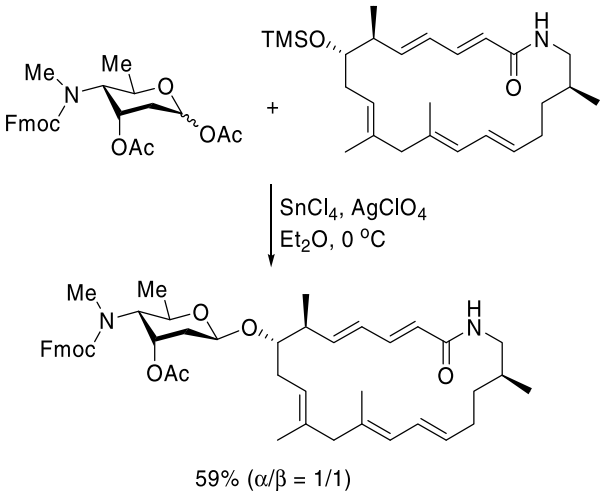
Scheme 38
(Ref. [46,47])

Mukaiyama and coworkers also introduced tin-based Lewis acids such as $SnCl_4/Sn(OTf)_2/LiClO_4$ and $SnCl_4/AgClO_4$ for the glycosylation of trimethylsilylated alcohols with the 1-*O*-acetyl sugar. For example, 1,2-*cis*-α-ribosides were predominantly obtained from 1-*O*-acetyl-riboses and silylated 3β-cholestanol (*Scheme 37*) [45]. On the basis of the effectiveness of the $SnCl_4/AgClO_4$ combination as a promoter [46], they also investigated the glycosylation of silylated alcohols with 1-*O*-acetyl-2,3,4,6-tetra-*O*-benzyl-D-glucopyranose using a series of combinations of other Lewis acids and $AgClO_4$. Among them, the catalysts generated from $HfCl_4$, $GaCl_3$, $InCl_3$, $SiCl_4$, or $GeCl_4$, with $AgClO_4$ led to the formation of the corresponding glycosides effectively in high diastereomeric ratios (*Scheme 38*) [47].

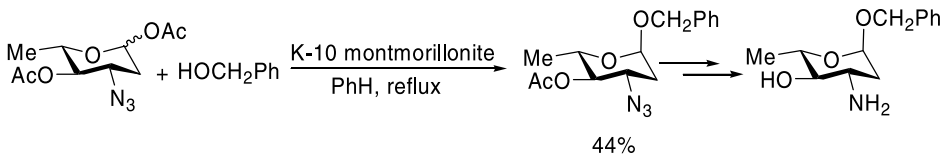


Scheme 39
(Ref. [48])

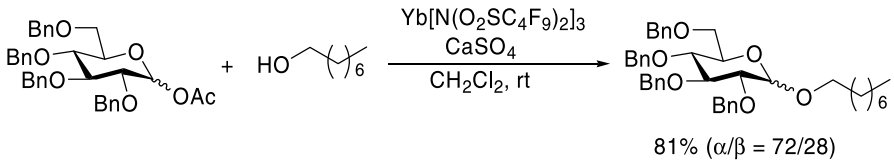
Furthermore, Mukaiyama et al. investigated the glycosylation with *N*-acetyl-α-neuraminosyl acetate for the stereoselective synthesis of *N*-acetyl-α-neuraminosyl-galactose disaccharide by using a variety of combinations of Lewis acids, for example, $SiCl_4$, $TiCl_4$, $SnCl_4$, $MeSiCl_3$, Me_2SiCl_2, or $PhSiCl_3$, with $AgClO_4$, in which the combination of Me_2SiCl_2 and $AgClO_4$ proved to be the most suitable promoter system (*Scheme 39*) [48]. Kakinuma et al. applied the $SnCl_4/AgClO_4$ combination as the promoter to the glycosylation of an *O*-TMS-aglycone with the appropriately protected 1-*O*-acetyl aminosugar for the total synthesis of Vicenistatin, a novel 20-membered macrocyclic lactam antitumor antibiotic (*Scheme 40*) [49].



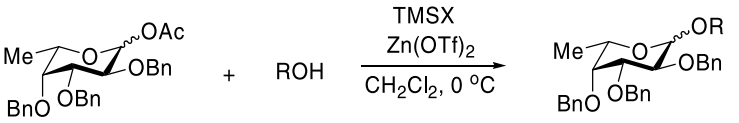
Scheme 40
(Ref. [49])



Scheme 41
(Ref. [50])



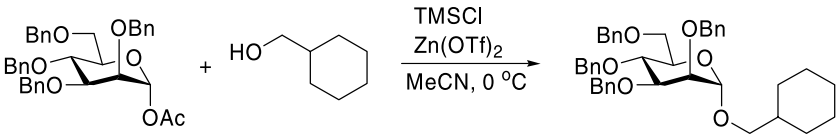
Scheme 42
(Ref. [51])



ROH = 3β-cholestanol

X = Cl, 70% (α/β = 73/27)
Br, 65% (α/β = 85/15)

Scheme 43
(Ref. [52])



92% (α/β = 97/3)

Scheme 44
(Ref. [53])

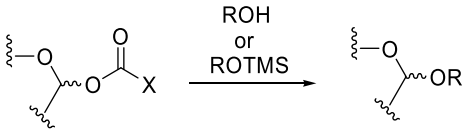
Besides these promoters for the activation of glycosyl acetate donors, K-10 montmorillonite was used as a new inexpensive catalyst in the glycosylation of the simple alcohol such as benzyl alcohol or methanol to give benzyl or methyl acosaminide with high stereoselectivity (*Scheme 41*) [50], and a catalytic amount of ytterbium(III) tris[bis(perfluorobutylsulfonyl)amide] ($Yb[N(O_2SC_4F_9)_2]_3$) was introduced as a new promoter for the glycosylation with glucosyl 1-acetate (*Scheme 42*) [51]. And, the combination of trimethylsilyl halides and $Zn(OTf)_2$ was also employed to give the corresponding glycoside in good yield from benzyl-protected fucosyl acetate (*Scheme 43*) [52]. In particular, a combination of TMSCl and $Zn(OTf)_2$ afforded α-mannosides stereoselectively from benzyl-protected mannosyl acetate without the participating group at C-2 (*Scheme 44*) [53].

3 Glycosylation with Donors Having Other Anomeric Ester Groups

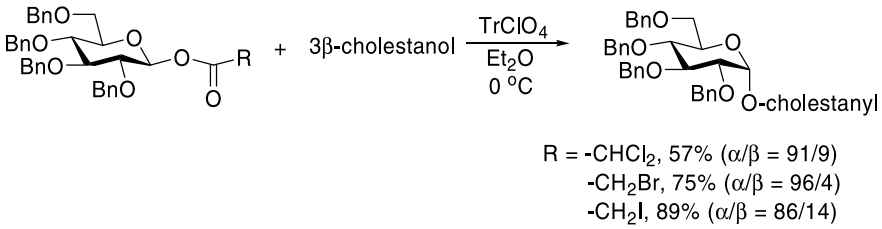
Besides glycosyl acetate, a variety of other acyl groups were employed as good anomeric leaving groups. The representative acyl leaving groups used in glycosyl donors for the glycosylation are summarized in *Table 2*.

Mukaiyama and coworkers prepared some 1-*O*-haloacetyl-β-D-glucopyranoses by treating 2,3,4,6-tetra-*O*-benzyl-D-glucopyranose with haloacetyl chloride in the presence of cesium fluoride or potassium fluoride and performed the glycosylation reactions with the 1-*O*-haloacetyl glucoses in the presence of $TrClO_4$ to obtain the corresponding glycosides in good yields (*Scheme 45*) [43]. Furthermore, they investigated the reactivity of various glucoses having substituted acetyl groups at the anomeric center, such as $-COCH_2OMe$, $-COCH_2CH_2OMe$, $-COCH_2SMe$, and $-COCH_2OCH_2CH_2OMe$, in the glycosylation reaction in the presence of $SnCl_4$ and $AgClO_4$ as a promoter system and they reported that the glucopyranose having the (2-methoxyethoxy)acetyl ($-COCH_2OCH_2CH_2OMe$) group at the anomeric

Table 2
Glycosylation with 1-*O*-ester sugars



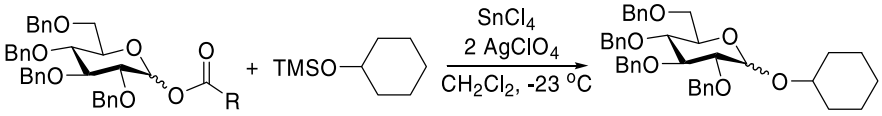
X	Activator	References
$-CHCl_2$, $-CH_2Br$, $-CH_2I$	$TrClO_4$, $SnCl_4/AgClO_4$	[43,54]
$-CH_2OCH_2CH_2OMe$	$SnCl_4/AgClO_4$	[54]
$-CF_3$, $-CCl_3$	TMSOTf, $BF_3{\cdot}OEt_2$	[55,56,57,58]
–Ph	$FeCl_3$, TMSOTf	[10,59]
–Ph-*p*-NO_2	TMSOTf, $BF_3{\cdot}OEt_2$, TMSCl/$Zn(OTf)_2$	[52,53,60,61,62]
–2-pyridine	$Cu(OTf)_2$, $Sn(OTf)_2$	[63]
$-C(CH_3)_3$	$SnCl_4$, TMSOTf	[64,65,66]
$-CH_2CH_2CH{=}CH_2$	IDCP, NIS, PhSeOTf	[67,68,69,70,71]
$-CH_2CH_2CH_2C{\equiv}CH$	$Hg(OTf)_2$	[72]



Scheme 45
(Ref. [43])

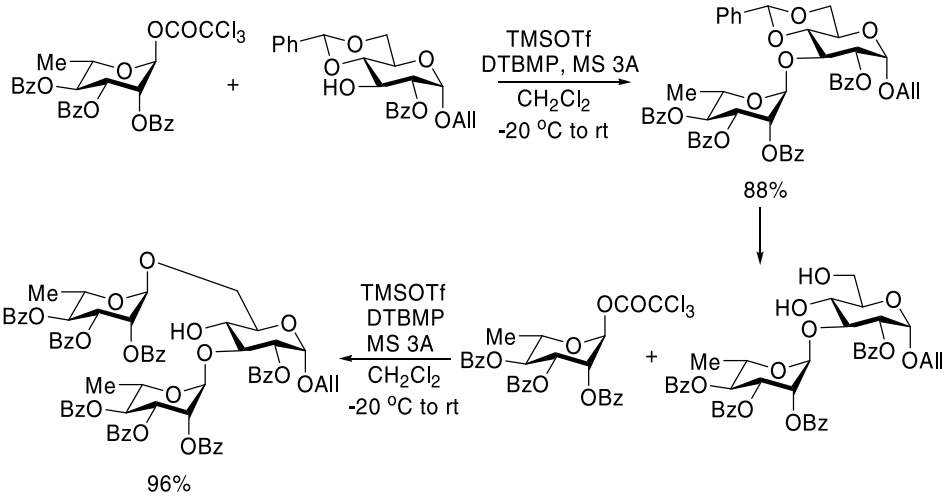
center was the most efficient donor among them (Scheme 46) [54]. The trihaloacetyl group was introduced as a new leaving group at the anomeric center in the glycosylation reaction by Cai and coworkers. 2,3,4,6-Tetra-*O*-acetyl-1-*O*-trifluoroacetyl-α-D-glucopyranose was found to be a good glycosyl donor in the presence of TMSOTf [55] and 1-*O*-trichloroacetyl 2,3,4-tri-*O*-benzoyl-α-L-rhamnopyranose was employed as an efficient donor to prepare disaccharide and trisaccharide residues, the sugar cores of phenylpropanoid glycosides (Scheme 47) [56]. They also reported the glycosylation with 2,3,4,6-*O*-tetrabenzyl-1-α-D-galactopyranosyl trichloroacetate [57] and with 2,3,4,6-tetra-*O*-benzoyl-α-D-mannopyranosyl trichloroacetate [58] in the presence of $BF_3{\cdot}OEt_2$ or TMSOTf.

On the other hand, benzoyl and *p*-nitrobenzoyl groups were employed by Lerner as good anomeric leaving groups and could be activated by $FeCl_3$, TMSOTf, or BF_3–OEt_2 like the acetyl group [10]. Charette et al. reported that the catalytic use of TMSOTf promoted the glycosylation of the trimethylsilyl ether of acceptor alcohols with 1-*O*-benzoyl sugar (Scheme 48) [59]. Terashima and coworkers used the 1-*O*-(*p*-nitrobenzoyl)glycosyl donor and TMSOTf for the synthesis of anthracycline antibiotics (Scheme 49) [60] while Scharf

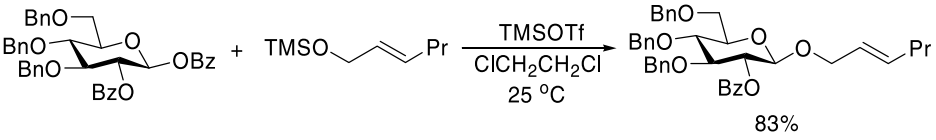


R = -CH_3, 85% (α/β = 24/76)
-CH_2Br, 79% (α/β = 22/78)
-CH_2I, 91% (α/β = 24/76)
-CH_2OMe, 82% (α/β = 18/82)
-CH_2CH_2OMe, 66% (α/β = 19/81)
-CH_2SMe, 40% (α/β = 15/85)
-$CH_2OCH_2CH_2OMe$, 90% (α/β = 16/84)
-Ph-2-OMe, 90% (α/β = 19/81)

Scheme 46
(Ref. [54])

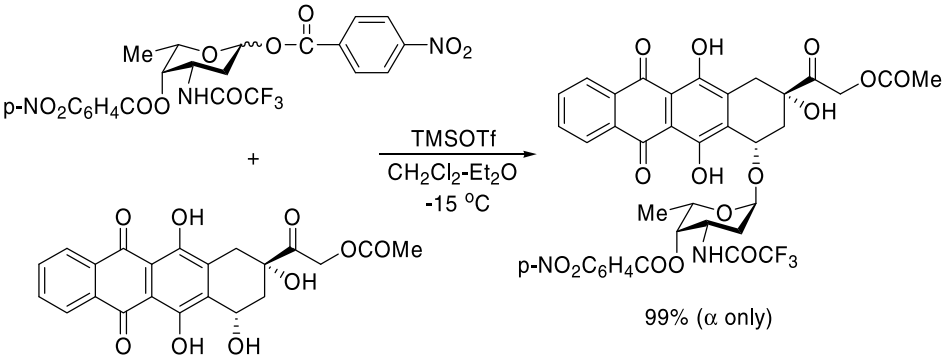


Scheme 47
(Ref. [56])

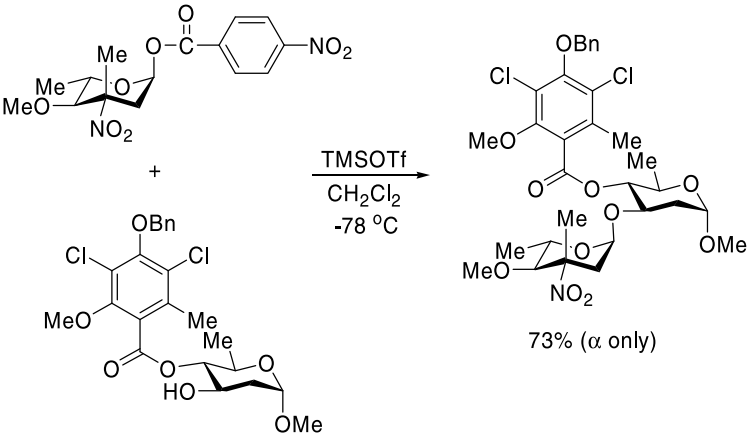


Scheme 48
(Ref. [59])

and coworkers applied Terashima's method to their synthetic study on everninomicin antibiotics (*Scheme 50*) [61]. Rosso et al. reported the use of the 1-*O*-(*p*-nitrobenzoyl)glycosyl donor for the practical synthesis of Disaccharide H, 2-*O*-(α-L-fucopyranosyl)-D-galactopyranose—which is one of the biologically important L-fucose-containing oligosaccharides—in



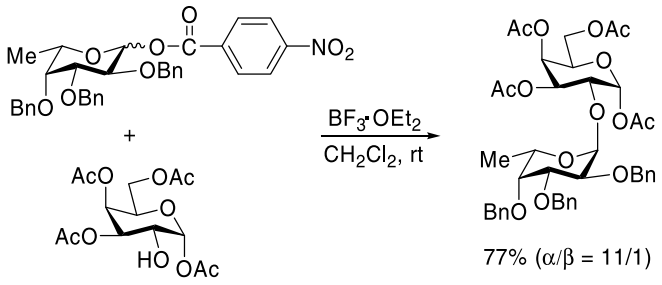
Scheme 49
(Ref. [60])



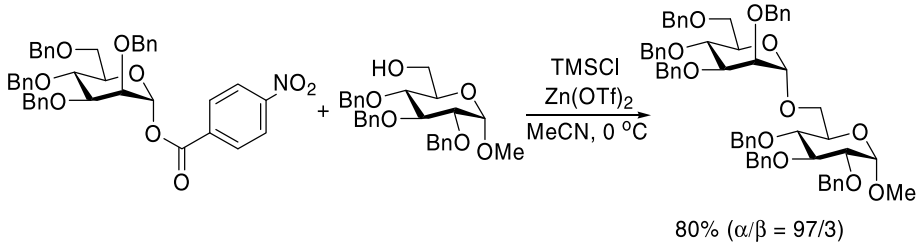
Scheme 50
(Ref. [61])

the presence of $BF_3{\cdot}OEt_2$ (Scheme 51) [62]. The *p*-nitrobenzoate as a leaving group was also activated by the combination of TMSCl and $Zn(OTf)_2$ and thus α-mannosides were obtained from benzyl-protected 1-*O*-(*p*-nitrobenzoyl) mannopyranose (Scheme 52) [52,53]. Kobayashi et al. reported that glycosyl 2-pyridinecarboxylate, as a glycosyl donor, could be activated by $Cu(OTf)_2$ in Et_2O or $Sn(OTf)_2$ in MeCN to predominantly produce the corresponding α- or β-glucoside, respectively (Scheme 53) [63].

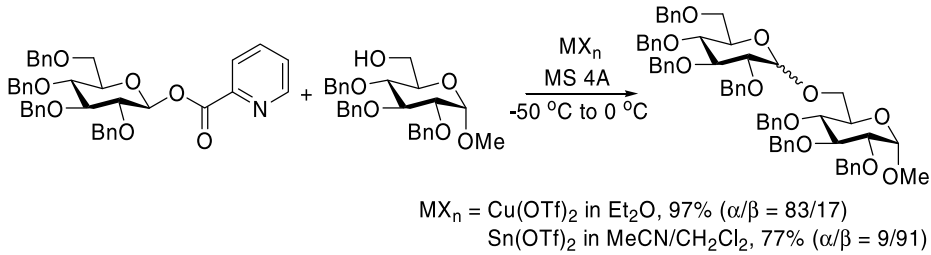
In addition, the pivaloyl group has been employed as an anomeric leaving group. Thus, the fully *O*-pivaloyl-protected galactofuranose glycosyl donor reacted with alkyl primary alcohol acceptors in the presence of $SnCl_4$ as the promoter (Scheme 54) [64]. Kunz et al. also reported that the fully *O*-pivaloyl-protected galactopyranosyl donor reacted with 6-*O*-protected 2-azido-galactose in the presence of TMSOTf to give the precursor structure of the Thomsen–Friedenreich antigen disaccharide, β(1–3)-linked galactosyl galactosamine, with high regioselectivity, although in low yield (Scheme 55) [65]. Very recently, pivaloylated glucosamin-



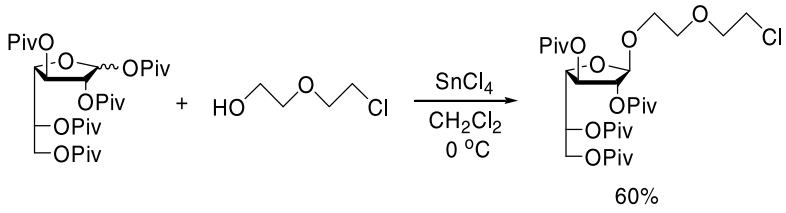
Scheme 51
(Ref. [62])



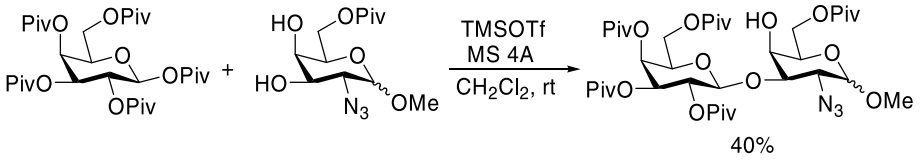
Scheme 52
(Ref. [52,53])



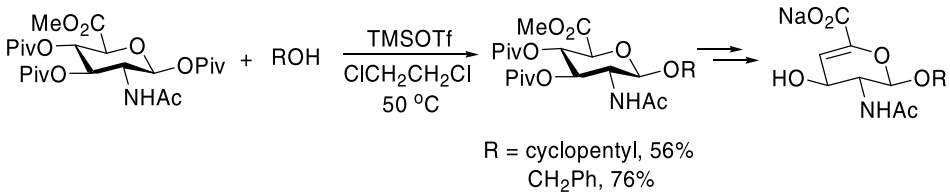
Scheme 53
(Ref. [63])



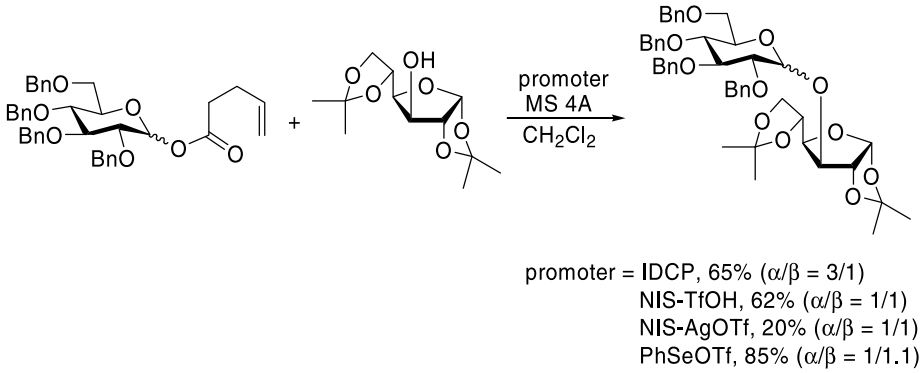
Scheme 54
(Ref. [64])



Scheme 55
(Ref. [65])



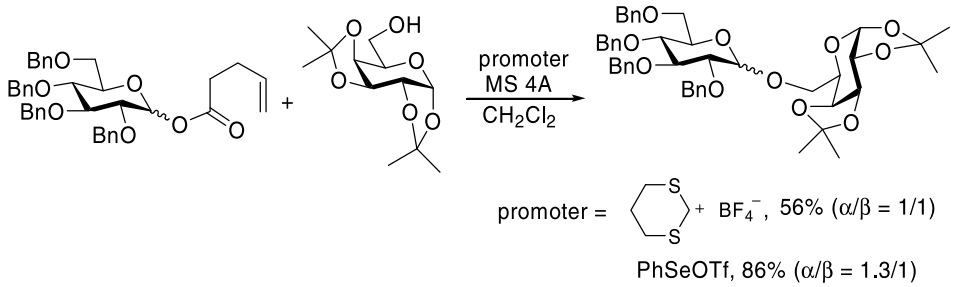
Scheme 56
(Ref. [66])



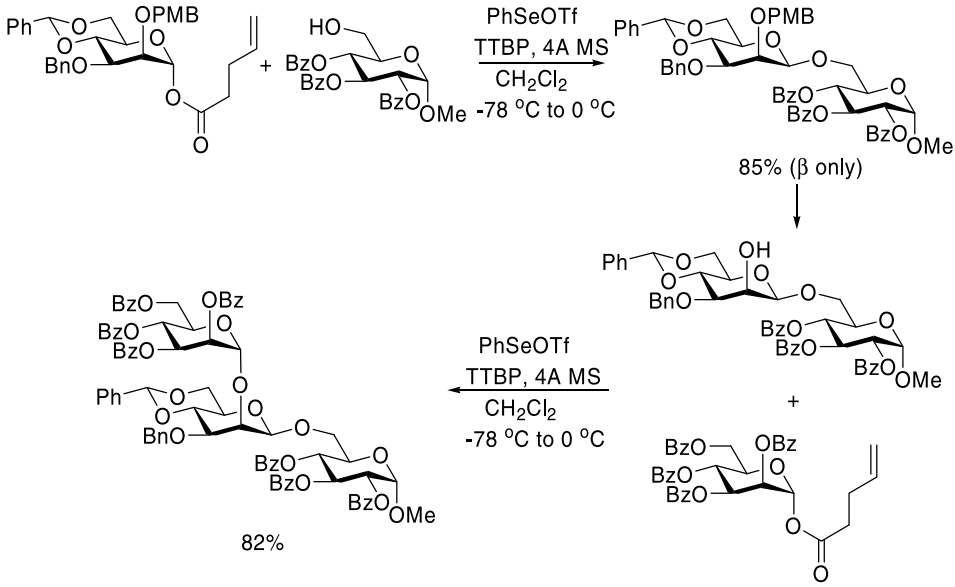
Scheme 57
(Ref. [67,69])

uronates as glycosyl donors were employed with TMSOTf to provide a series of novel *β-O*-glycosides, which are the precursors of unsaturated *N*-acetyl-D-glucosaminuronic acid glycosides as inhibitors of the influenza virus sialidase (*Scheme 56*) [66].

Kunz [67] and Fraser-Reid [68] et al. introduced glucosyl 4-pentenoates as new glycosyl donors. These glucosyl 4-pentenoates were prepared by the condensation of the corresponding glucopyranoses with 4-pentenoic acid using 1,3-dicyclohexylcarbodiimide (DCC) or diisopropylcarbodiimide (DIC) and activated with soft electrophiles, for example, iodonium compounds, iodonium dicollidine perchlorate (IDCP) or *N*-iodosuccinimide (NIS) (*Scheme 57*), or 1,3-dithian-2-ylium tetrafluoroborate (*Scheme 58*). However, the glycosylations with the pentenoates using these promoters were not very efficient. Very recently, Kim and coworkers reported that phenylselenyl triflate (PhSeOTf) was found to be a much more efficient promoter than IDCP, NIS-TfOH, and 1,3-dithian-2-yl tetrafluoroborate for glycosylations with glyco-

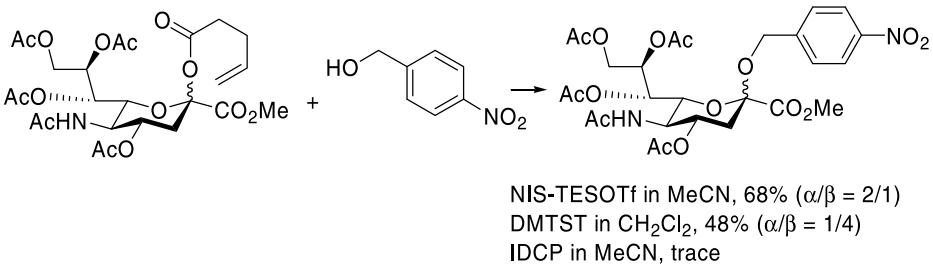


Scheme 58
(Ref. [68,69])

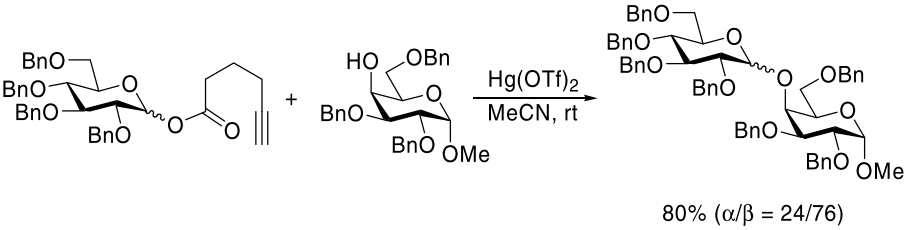


Scheme 59
(Ref. [70])

syl 4-pentenoates as donors (Scheme 57 and Scheme 58) [69]. Moreover, they described a highly reactive and stereoselective procedure for the β-mannopyranosylation employing 4,6-*O*-benzylidene mannopyranosyl pentenoate and PhSeOTf (Scheme 59), and this method was proved to be comparable to and even better than other methods for the β-mannosylation of the simple reactive primary alcohols [70]. On the other hand, the reactivity and stereoselectivity of 4-pentenoic acid ester of *N*-acetylneuraminic acid as a sialyl donor were also investigated with a series of promoters, such as NIS-TESOTf, dimethyl(methylthio)sulfonium triflate (DMTST), and IDCP, but the glycosylation with the sialyl pentenoate gave the *O*-sialosides in low yield and low stereoselectivity (Scheme 60) [71].



Scheme 60
(Ref. [71])



Scheme 61
(Ref. [72])

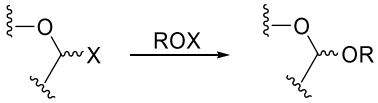
Recently, Nishizawa et al. reported a mercuric triflate [$Hg(OTf)_2$]-catalyzed glycosylation using the alkynoate as the leaving group (*Scheme 61*) [72].

4 Glycosylation with Glycosyl Carbonates and Related Donors

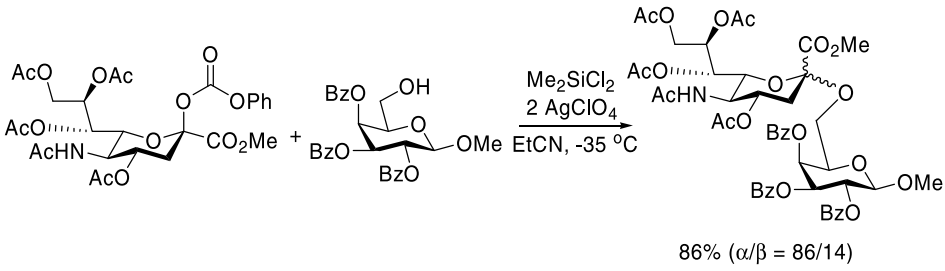
Some representative methods using 1-*O*-carbonate derivatives as glycosyl donors are summarized in *Table 3*.

Mukaiyama et al. reported the glycosylation with the *N*-acetylneuraminic acid donor having an easily accessible phenoxycarbonyloxy leaving group at the anomeric center using a combination of Me_2SiCl_2 and $AgClO_4$ as the promoter for the stereoselective synthesis of *N*-acetyl-α-neuraminosyl-galactose disaccharide (*Scheme 62*) [48]. They also reported the highly stereoselective synthesis of 1,2-*trans*-glycosides using *p*-chlorobenzylated glycosyl carbonates as glycosyl donors in the presence of trityl tetrakis(pentafluorophenyl)borate [$TrB(C_6F_5)_4$], in which several β-D-glucopyranosides were prepared in good yields with high stereoselectivities even in the absence of the anchimeric assistance by the neighboring group at C-2 position (*Scheme 63*) [73,74]. Furthermore, they applied the same method using a galactosyl phenylcarbonate as the glycosyl donor to the convergent total synthesis of the F1α antigen, a member of the tumor-associated *O*-linked mucin glycosyl amino acids (*Scheme 64*) [75]. On the other hand, Sinaÿ and Descotes independently introduced the isopropenyloxycarbonyloxy leaving group at the anomeric center for the glycosylation reaction with TMSOTf. Sinaÿ

Table 3
Glycosylation with 1-*O*-carbonates and related sugars



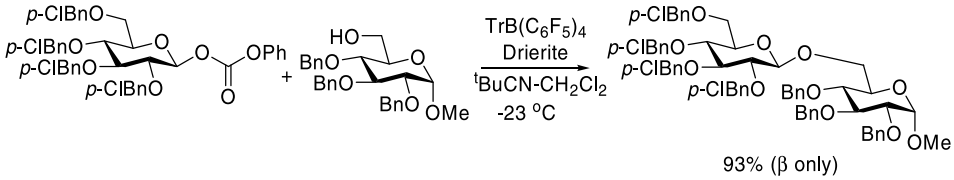
References	X	Activator
—O—C(=O)—OPh	$Me_2SiCl_2/AgClO_4$, $TrB(C_6F_5)_4$	[48,73,74,75]
—O—C(=O)—O—C(=CH_2)—CH_3	TMSOTf	[76]
—O—C(=O)—N(imidazolyl)	$ZnBr_2$	[77]
—O—C(=O)—NH—$CH_2CH{=}CH_2$	IDCP	[78]
—O—C(=O)—NH—Ts	TMSOTf	[79]
—O—C(=O)—S—(2-pyridyl)	AgOTf	[80,81]



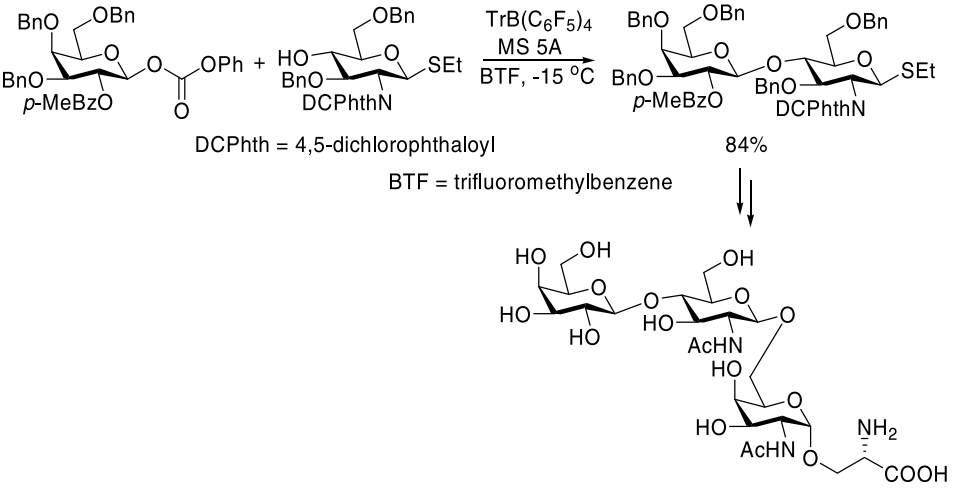
Scheme 62
(Ref. [48])

and coworkers showed the high-yielding β-stereoselective glycosylation employing glucosyl isopropenylcarbonate as the glycosyl donor (*Scheme 65*) [76].

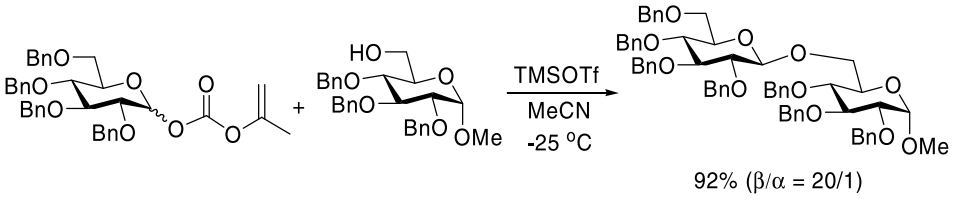
Ley and Ford reported the C-1 (1-imidazolylcarbonyl) glycosides as new glycosyl donors, which reacted with a series of alcohol acceptors in the presence of $ZnBr_2$ to give the corresponding *O*-glycosides (*Scheme 66*) [77]. Kunz and Zimmer reported the *N*-allyl carbamate as an anomeric leaving group with soft electrophiles, for example, IDCP, DMTST, and methyl bis-methylthiosulfonium hexachloroantimonate, as promoters for the *O*-glycoside synthesis (*Scheme 67*) [78]. Kiessling and Hinklin reported glycosyl sulfonylcarbamates



Scheme 63
(Ref. [73,74])

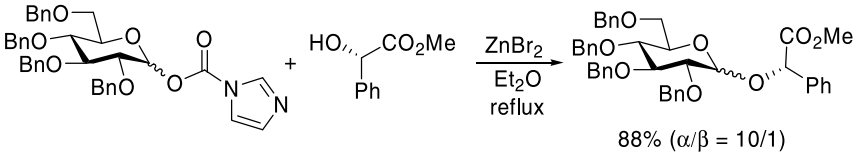


Scheme 64
(Ref. [75])

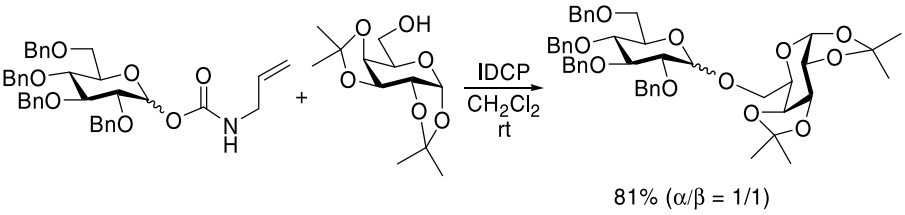


Scheme 65
(Ref. [76])

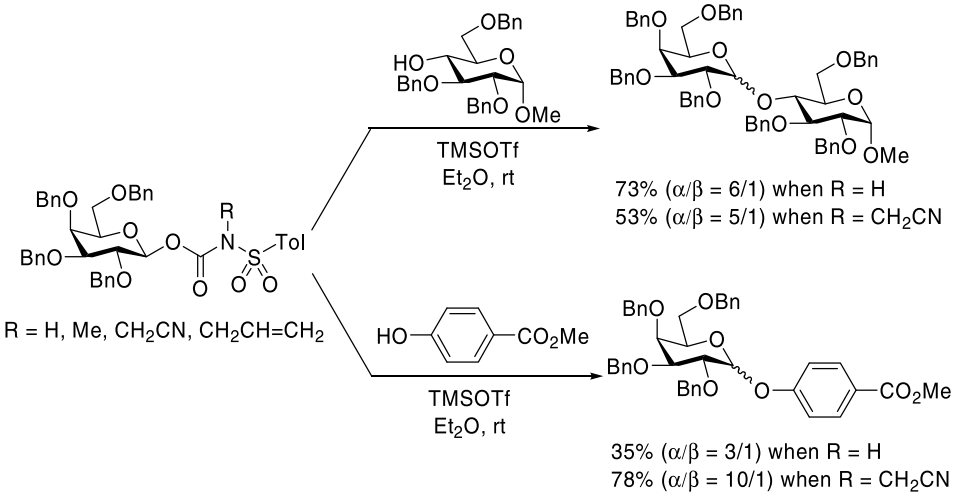
as new glycosyl donors with tunable reactivity by introducing different *N*-alkyl substituents on the sulfonylcarbamate group. For example, the glycosylation of a glucose with tetrabenzylgalactosyl *N*-tosylcarbamate was more efficient than that with tetrabenzylgalactosyl *N*,*N*-(cyanomethyl)tosylcarbamate while the glycosylation of a poor nucleophilic phenol acceptor with tetrabenzylgalactosyl *N*,*N*-(cyanomethyl)tosylcarbamate turned out to be more efficient and stereoselective than that with tetrabenzylgalactosyl *N*-tosylcarbamate (*Scheme 68*) [79].



Scheme 66
(Ref. [77])

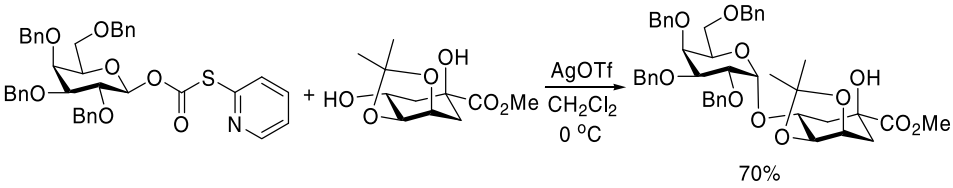


Scheme 67
(Ref. [78])

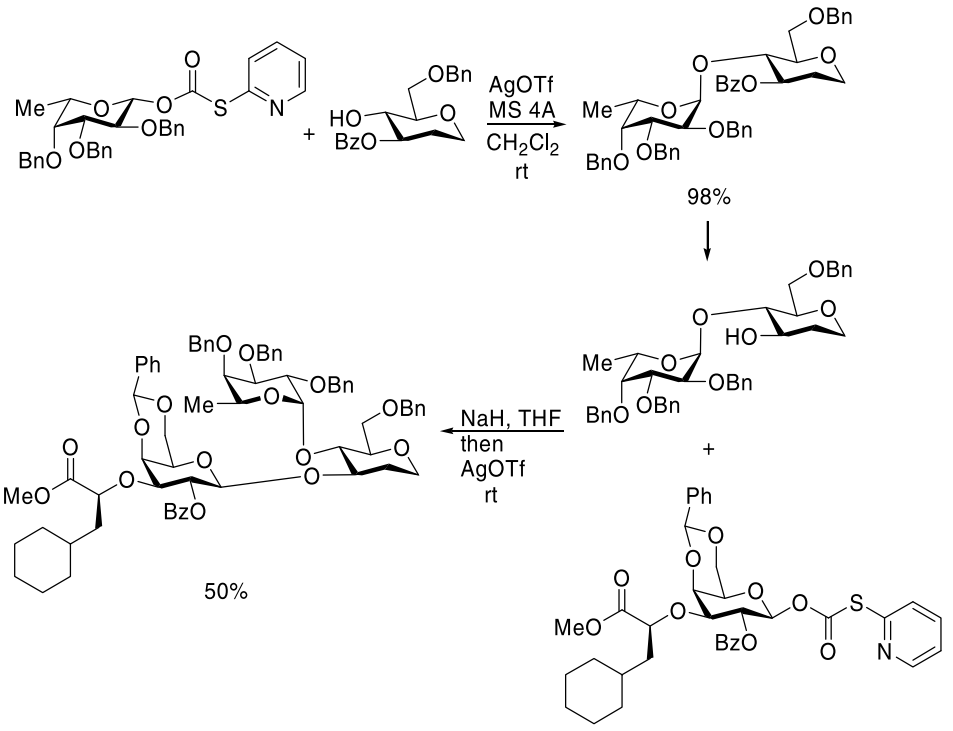


Scheme 68
(Ref. [79])

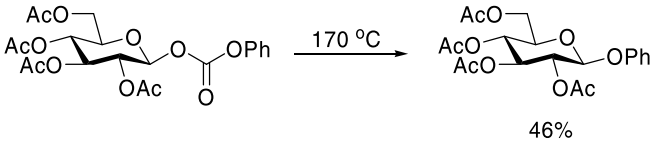
On the other hand, in order to explore a new efficient α-galactopyranosylation method, Hanessian et al. introduced a crystalline glycosyl 2-pyridyl thiocarbonate donor, which gave the pseudo-disaccharide in good yield using AgOTf as the promoter (*Scheme 69*) [80]. Furthermore, they successfully utilized the same method to prepare a carbohydrate-based antagonist of E-selectin, and they found that pretreatment of the hindered disaccharide acceptor with sodium hydride, followed by addition of AgOTf, and finally adding the glycosyl donor led to the target trisaccharide (*Scheme 70*) [81].



Scheme 69
(Ref. [80])



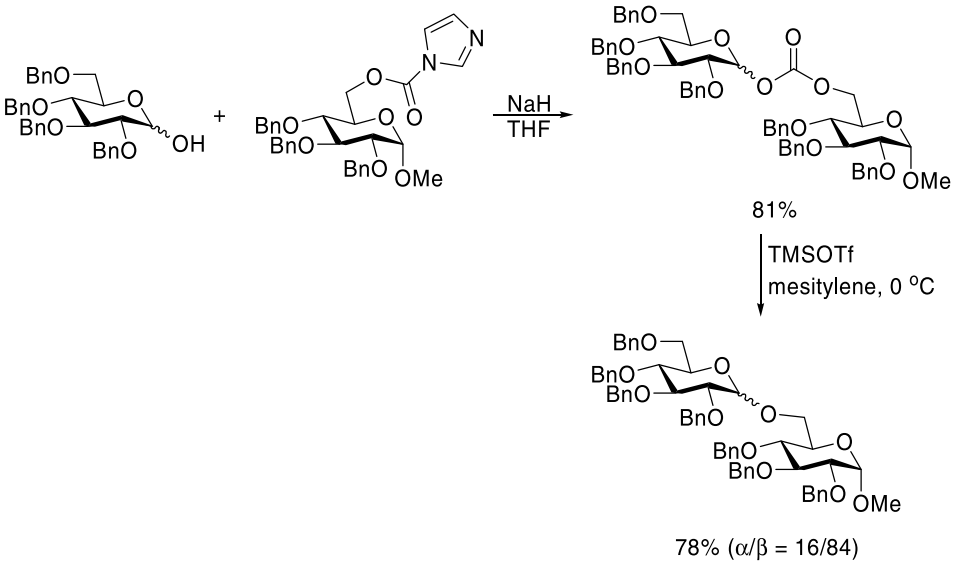
Scheme 70
(Ref. [81])



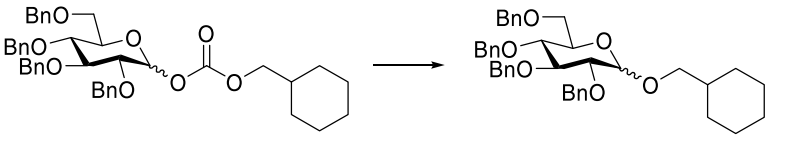
Scheme 71
(Ref. [82,83])

5 Intramolecular Glycosylation Through 1-*O*-Acyl Linkages

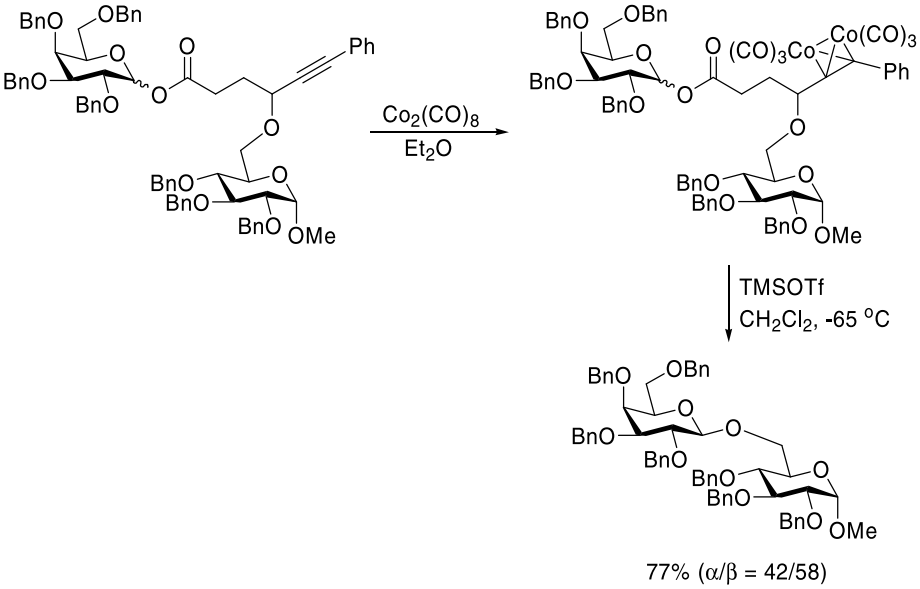
Ishido and coworkers reported a new method for the synthesis of *O*-glycosides by the pyrolysis of 1-*O*-aryloxycarbonyl sugar derivatives: for example, the pyrolysis of 2,3,4,6-tetra-*O*-acetyl-1-*O*-phenoxycarbonyl-β-D-glucopyranose at 170°C afforded phenyl 2,3,4,6-tetra-*O*-acetyl-β-D-glycopyranoside in moderate yield (● *Scheme 71*) [82,83]. Ikegami [84] and Schmidt [85] et al. developed a novel intramolecular decarboxylative glycosylation via mixed carbonate as a two-step glycosylation procedure, which involves linking two sugars by using the carbonate as a connector and subsequent extrusion of carbon dioxide to form a glycosyl bond by the aid of Lewis acid, TMSOTf or TBSOTf (● *Scheme 72*). Ikegami et al. also reported that



Scheme 72
(Ref. [84,85])



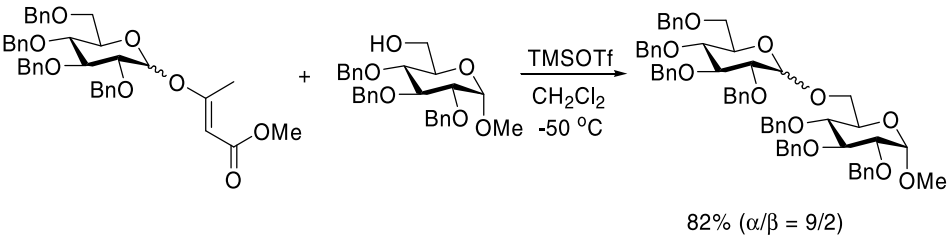
Scheme 73
(Ref. [86,87,88])



Scheme 74
(Ref. [89])

a catalytic amount of $SnCl_4$-$AgClO_4$ or Cp_2HfCl_2-$2AgClO_4$ promoted decarboxylation of an 1-*O*-carbonate sugar to afford the α-glycoside stereoselectively (*Scheme 73*) [86]. At the same time, they reported the synthesis of mixed β-carbonates of acyl-protected sugar and their decarboxylative glycosylations promoted by TMSOTf [87] or $Hf(OTf)_4$ [88] and obtained the β-glycosides stereoselectively.

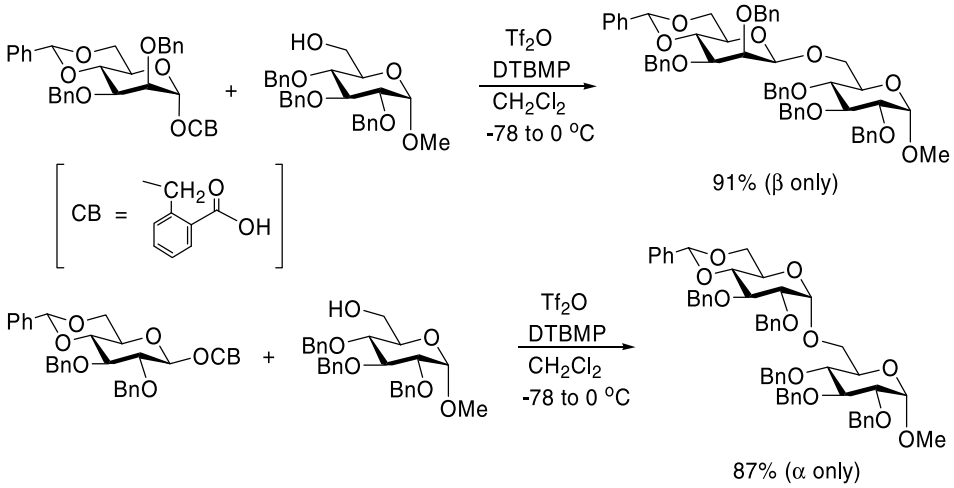
On the other hand, Mukai et al. reported a glycosylation reaction based on the alkyne-$Co_2(CO)_6$ complex, thus *O*-protected glycopyranose possessing 4-(*O*-protected-glycosyl)-6-phenyl-5-hexynoate residue at an anomeric position was converted to the corresponding cobalt complex, which on exposure to TMSOTf afforded the product disaccharide in good yield via the internal delivery of the glycosyl acceptor (*Scheme 74*) [89].



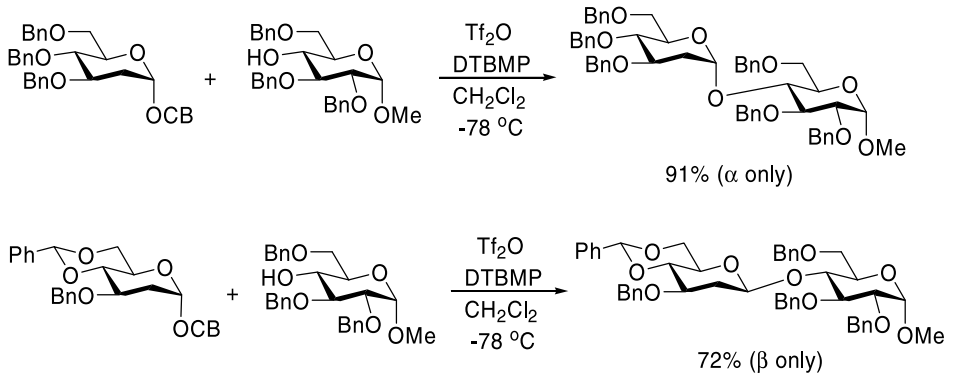
Scheme 75
(Ref. [90])

6 Glycosylation with Donors Having Remote Acyl Groups

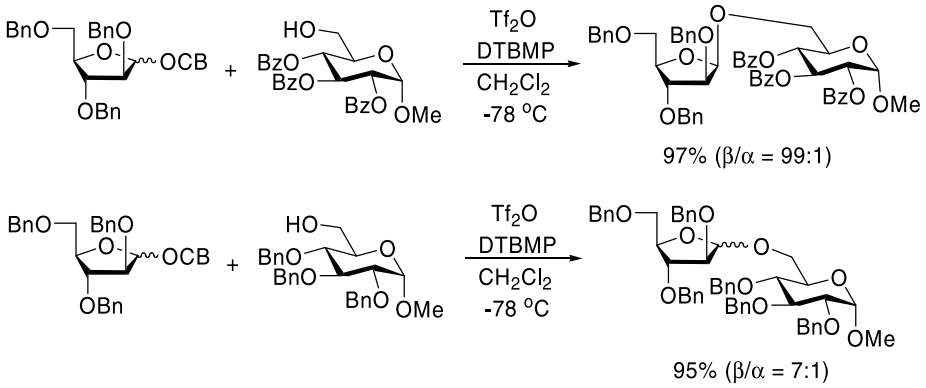
Takeda et al. reported a glycosylation method using glycosyl donors having the enol ether conjugated with the ester functionality as a leaving group and TMSOTf as the promoter (Scheme 75) [90]. Kim et al. introduced a stereocontrolled glycosylation method employing 2′-carboxybenzyl (CB) glycosides as glycosyl donors, which was found to be a useful tool for the stereoselective β-mannopyranosylation using CB 4,6-*O*-benzylidene-2,3-di-*O*-benzyl-D-mannoside and triflic anhydride (Tf_2O) as the promoter, and for the stereoselective α-glucopyranosylation using CB 4,6-*O*-benzylidene-2,3-di-*O*-benzyl-D-glucoside (Scheme 76) [91]. Next, they showed a highly α- and β-stereoselective (dual stereoselective)



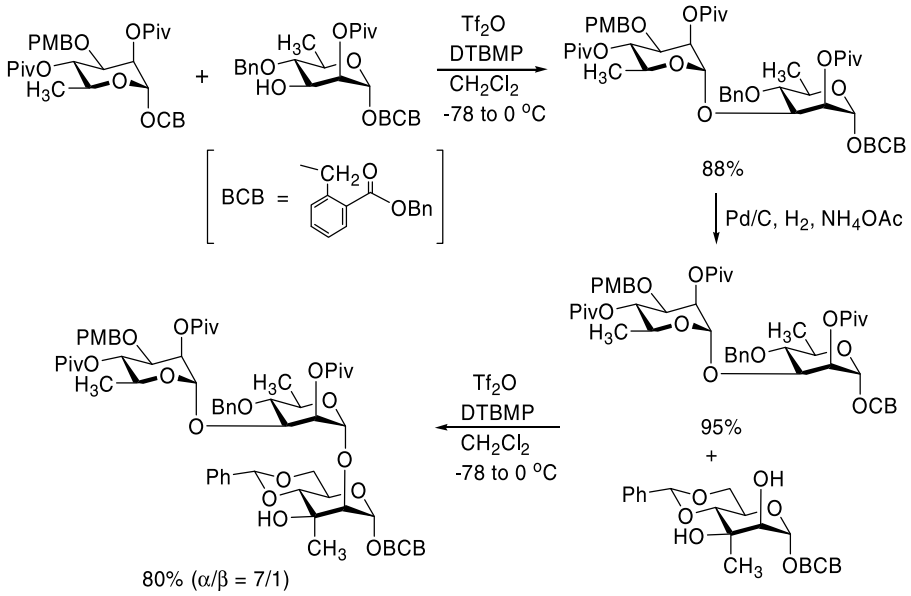
Scheme 76
(Ref. [91])



Scheme 77
(Ref. [92])

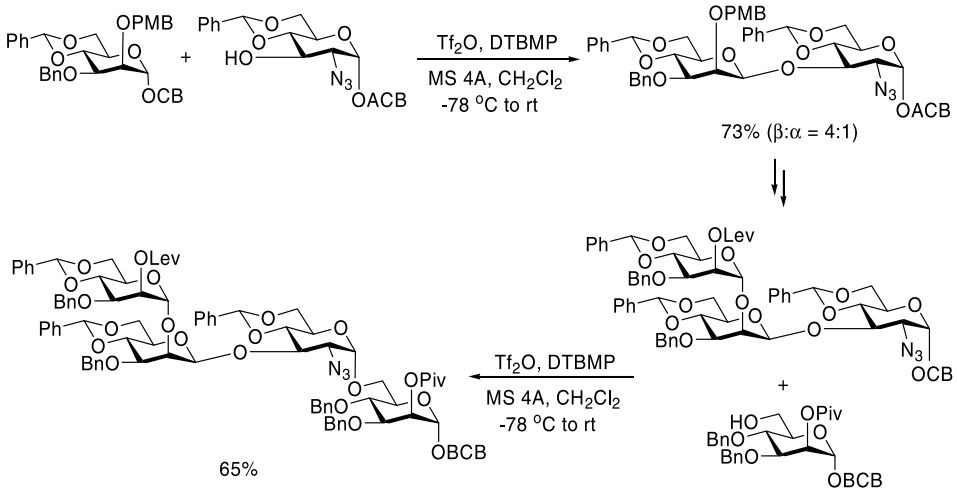


Scheme 78
(Ref. [93])

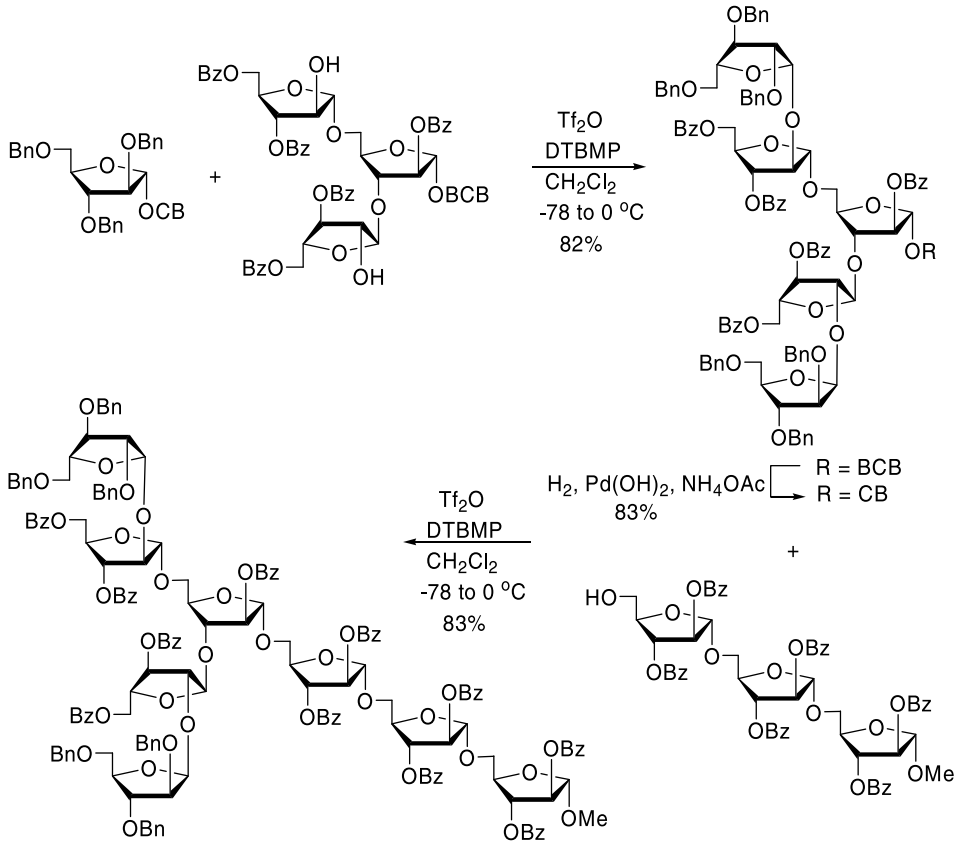


Scheme 79
(Ref. [94])

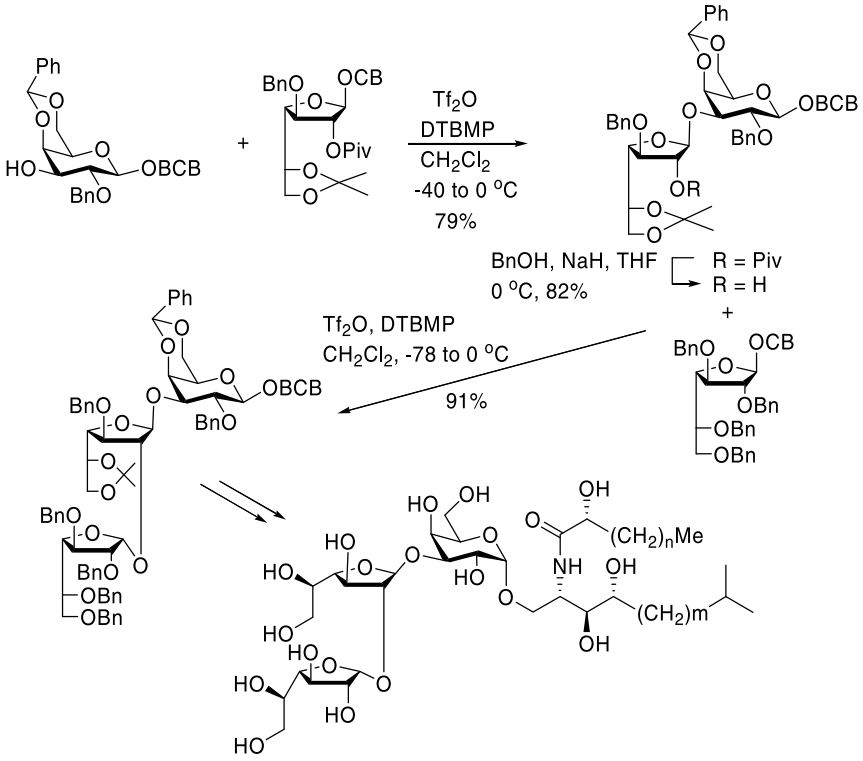
method for the synthesis of 2-deoxyglycosides by employing CB 2-deoxyglycosides as glycosyl donors, in which glycosylation of the 4,6-*O*-benzylidene-protected glycosyl donor with secondary alcohols afforded predominantly β-glycosides whereas glycosylation of the benzyl-protected glycosyl donor with secondary alcohols afforded α-glycosides (*Scheme 77*) [92]. They also established a reliable and generally applicable direct method for the stereoselective β-arabinofuranosylation employing CB tri-*O*-benzylarabinoside as the glycosyl donor, in



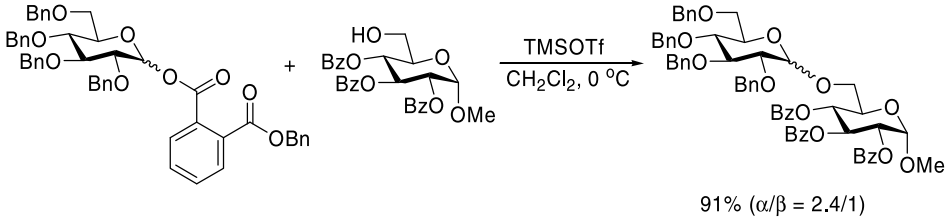
Scheme 80
(Ref. [95])



Scheme 81
(Ref. [93])



Scheme 82
(Ref. [97])



Scheme 83
(Ref. [98])

which the acyl-protective group on glycosyl acceptors was essential for the β-stereoselectivity (*Scheme 78*) [93]. Furthermore, they applied this CB glycoside methodology to the synthesis of the repeat unit of the *O*-antigen polysaccharide of the lipopolysaccharide (LPS) from Gram-negative bacteria; a protected form of the trisaccharide repeat unit of the atypical *O*-polysaccharide from Danish *Helicobacter pylori* strains (*Scheme 79*) [94], and a protected form of the tetrasaccharide repeat unit of the *O*-antigen polysaccharide from the *E. coli* lipopolysaccharide (*Scheme 80*) [95]. An octaarabinofuranoside in arabinogalactan and

lipoarabinomannan, found in the mycobacterial cell wall, was also synthesized by employing the latent (BCB)-active (CB) glycosylation method (*Scheme 81*) [93]. Very recently, they reported the total synthesis of agelagalastatin, which is a trisaccharide sphingolipid and displays significant in vitro inhibitory activities against human cancer cell growth [96]. Among three glycosyl linkages in agelagalastatin, α-galactofuranosyl and β-galactofuranosyl linkages were stereoselectively constructed employing the CB glycoside method (*Scheme 82*) [97]. On the other hand, they reported the glycosylation of various acceptors with glycosyl benzyl phthalates as glycosyl donors using TMSOTf as the promoter (*Scheme 83*) [98].

References

1. Helferich B, Shimitz-Hillebrecht E (1933) Chem Ber 66:378
2. Boivin J, Monneret C, Pais M (1978) Tetrahedron Lett 19:1111
3. Lemieux RU, Shyluk WP (1953) Can J Chem 31:528
4. Hanessian S, Banoub J (1977) Carbohydr Res 59:261
5. Honma K, Nakazima K, Uematsu T, Hamada A (1976) Chem Pharm Bull 24:394
6. Dayal B, Salen G, Padia J, Shefer S, Tint GS, Sasso G, Williams TH (1993) Carbohydr Res 240:133
7. Benjahad A, Benhaddou R, Granet R, Kaouadji M, Krausz P, Piekarski S, Thomasson F, Bosgiraud C, Delebassée S (1994) Tetrahedron Lett 35:9545
8. Gaud O, Granet R, Kaouadji M, Krausz P, Blais JC, Bolbach G (1996) Can J Chem 74:481
9. Kiso M, Anderson L (1979) Carbohydr Res 72:C15
10. Lerner LM (1990) Carbohydr Res 207:138
11. Katsuraya K, Ikushima N, Takahashi N, Shoji T, Nakashima H, Yamamoto N, Yoshida T, Uryu T (1994) Carbohydr Res 260:51
12. Chatterjee SK, Nuhn P (1998) Chem Commun 1729
13. Dahmén J, Frejd T, Magnusson G, Noori G (1983) Carbohydr Res 114:328
14. Gurjar MK, Viswanadham G (1991) Tetrahedron Lett 32:6191
15. Elofsson M, Walse B, Kihlberg J (1991) Tetrahedron Lett 32:7613
16. Steffan W, Schutkowski M, Fischer G (1996) J Chem Soc Chem Commun 313
17. Gangadhar BP, Jois SDS, Balasubramaniam A (2004) Tetrahedron Lett 45:355
18. Wang TC, Furukawa H, Nihro Y, Kakegawa H, Matsumoto H, Satoh T (1994) Chem Pharm Bull 42:570
19. (a) Ludin C, Weller T, Seitz B, Meier W, Erbeck S, Hoenke C, Krieger R, Keller M, Knothe L, Pelz K, Wittmer A, Prinzbach H (1995) Liebigs Ann 291;
(b) Erbeck S, Prinzbach H (1997) Tetrahedron Lett 38:2653;
(c) Erbeck S, Liang X, Hunkler D, Krieger R, Prinzbach H (1998) Eur J Org Chem 1935
20. Lam SN, Gervay-Hague J (2005) J Org Chem 70:8772
21. Kelly NM, Jensen KJ (2001) J Carbohydr Chem 20:537
22. Ikeda K, Torisawa Y, Nishi T, Minamikawa J, Tanaka, K, Sato M (2003) Bioorg Med Chem 11:3073
23. Ogawa T, Beppu K, Nakabayashi S (1981) Carbohydr Res 93:C6
24. Ogawa T, Beppu K (1982) Carbohydr Res 101:271
25. (a) Kovac P (1986) Carbohydr Res 153:237;
(b) PozsgayV, Brisson JR, Jennings HJ (1987) Can J Chem 65:2764;
(c) Pozsgay V, Brisson JR, Jennings HJ (1990) Carbohydr Res 205:133;
(d) Nifant'ev NE, Lipkind GM, Shashkov AS, Kochetkov NK (1992) Carbohydr Res 223:109
26. Nifant'ev NE, Backinowsky LV, Kochetkov NK (1988) Carbohydr Res 174:61
27. Nifant'ev NE, Khatuntseva EA, Shashkov AS, Bock K (1996) Carbohydr Lett 1:399
28. Trumtel M, Tavecchia P, Veyriéres A, Sinaÿ P (1989) Carbohydr Res 191:29
29. Paulsen H, Paal M (1984) Carbohydr Res 135:53
30. Paulsen H, Tietz H (1985) Angew Chem Int Ed Engl 24:128
31. Pozsgay V, Brisson JR, Jennings HJ (1991) J Org Chem 56:3377
32. Rainer H, Scharf HD, Runsink J (1992) Liebigs Ann Chem 103

33. Lafont D, Boullanger P, Banoub J, Descotes G (1990) Can J Chem 68:828
34. (a) Sebesta DP, Roush WR (1992) J Org Chem 57:4799;
(b) Roush WR, Briner K, Sebesta DP (1993) Synlett 264
35. Roush WR, Narayan S (1999) Org Lett 1:899
36. Durham TB, Roush WR (2003) Org Lett 5:1871
37. Arai M, Kaneko S, Konosu T (2002) Tetrahedron Lett 43:6705
38. Roush WR, Bennett CE, Roberts SE (2001) J Org Chem 66:6389
39. Mergott DJ, Frank SA, Roush WR (2004) Proc Natl Acad Sci USA 101:11955
40. Gargiulo D, Blizzard TA, Nakanishi K (1989) Tetrahedron 45:5423
41. Satto S, Kuroda K, Hayashi Y, Sasaki Y, Nagamura Y, Nishida K, Ishiguro I (1991) Chem Pharm Bull 39:2333
42. Dondoni A, Marra A, Rojo I, Scherrmann MC (1996) Tetrahedron 52:3057
43. Mukaiyama T, Kobayashi S, Shoda S (1984) Chem Lett 907
44. (a) Evans DA, Kaldor SW, Jones TK, Clardy J, Stout TJ (1990) J Am Chem Soc 112:7001;
(b) Evans DA, Black WC (1993) J Am Chem Soc 115:4497
45. Mukaiyama T, Shimpuku T, Takashima T, Kobayashi S (1989) Chem Lett 145
46. Mukaiyama T, Takashima T, Katsurada M, Aizawa H (1991) Chem Lett 533
47. Mukaiyama T, Katsurada M, Takashima T (1991) Chem Lett 985
48. Mukaiyama T, Sasaki T, Iwashita E, Matsubara K (1995) Chem Lett 455
49. Matsushima Y, Itoh H, Nakayama T, Horiuchi S, Eguchi T, Kakinuma K (2002) J Chem Soc Perkin Trans 1 949
50. Florent JC, Monneret C (1987) J Chem Soc Chem Commun 1171
51. Yamanoi T, Nagayama S, Ishida H, Nishikido J, Inazu T (2001) Synth Commun 31:899
52. Higashi K, Susaki H (1992) Chem Pharm Bull 40:2019
53. Susaki H, Higashi K (1993) Chem Pharm Bull 41:201
54. Matsubara K, Sasaki T, Mukaiyama T (1993) Chem Lett 1373
55. Yu CF, Li ZJ, Cai MS (1990) Synth Commun 20:943
56. Li ZJ, Huang HQ, Cai MS (1994) Carbohydr Res 265:227
57. Mao J, Chen H, Zhang J, Cai M (1995) Synth Commun 25:1563
58. Li ZJ, Huang HQ, Cai M (1996) J Carbohydr Chem 15:501
59. Charette AB, Marcoux JF, Côté B (1991) Tetrahedron Lett 32:7215
60. Kimura Y, Suzuki M, Matsumoto T, Abe R, Terashima S (1984) Chem Lett 501
61. Jütten P, Scharf HD, Raabe G (1991) J Org Chem 56:7144
62. Nicotra F, Panza L, Romanò A, Russo G (1992) J Carbohydr Chem 11:397
63. Koide K, Ohno M, Kobayashi S (1991) Tetrahedron Lett 32:7065
64. Carcía-Barrientos A, García-López JJ, Isac-García J, Ortega-Caballero F, Uriel C, Vargas-Berenguel A, Santoyo-González F (2001) Synlett 1057
65. Oßwald M, Lang U, Friedrich-Bochnitschek S, Pfrengle W, Kunz H (2003) Z Naturforsch 58b:764
66. Mann MC, Islam T, Dyason JC, Florio P, Trower CJ, Thomson RJ, von Itzstein M (2006) Glycoconj J 23:127
67. Kunz H, Wernig P, Schultz M (1990) Synlett 631
68. Lopez JC, Fraser-Reid B (1991) J Chem Soc Chem Commun 159
69. Choi TJ, Baek JY, Jeon HB, Kim KS (2006) Tetrahedron Lett 47:9191
70. Baek JY, Choi TJ, Jeon HB, Kim KS (2006) Angew Chem Int Ed 45:7436
71. Ikeda K, Fukuyo J, Sato K, Sato M (2005) Chem Pharm Bull 53:1490
72. Imagawa H, Kinoshita A, Fukuyama T, Yamamoto H, Nishizawa M (2006) Tetrahedron Lett 47:4729
73. Mukaiyama T, Miyazaki K, Uchiro H (1998) Chem Lett 635
74. Mukaiyama T, Wakiyama Y, Miyazaki K, Takeuchi K (1999) Chem Lett 933
75. Mukaiyama T, Ikegai K, Jona H, Hashihayata T, Takeuchi K (2001) Chem Lett 840
76. (a) Boursier M, Descotes G (1989) C R Acad Sci Ser II 308:919;
(b) Sinaÿ P (1991) Pure Appl Chem 63:519;
(c) Marra A, Esnault J, Veyrières A, Sinaÿ P (1992) J Am Chem Soc 114:6354
77. Ford MJ, Ley SV (1990) Synlett 255
78. Kunz H, Zimmer J (1993) Tetrahedron Lett 34:2907
79. Hinklin RJ, Kiessling LL (2001) J Am Chem Soc 123:3379

80. Hanessian S, Huynh HK, Reddy GV, Duthaler RO, Katopodis A, Streiff MB, Kinzy W, Oehrlein R (2001) Tetrahedron 57:3281
81. Hanessian S, Mascitti V, Rogel O (2002) J Org Chem 67:3346
82. Inaba S, Yamada M, Yoshino T, Ishido Y (1973) J Am Chem Soc 95:2062
83. Ishido Y, Inaba S, Matsuno A, Yoshino T, Umezawa H (1977) J Chem Soc Perkin Trans 1 1382
84. Iimori T, Shibazaki T, Ikegami S (1996) Tetrahedron Lett 37:2267
85. Scheffler G, Schmidt RR (1997) Tetrahedron Lett 38:2943
86. Iimori T, Azumaya I, Shibazaki T, Ikegami S (1997) Heterocycles 46:221
87. Azumaya I, Niwa T, Kotani M, Iimori T, Ikegami S (1999) Tetrahedron Lett 40:4683
88. Azumaya I, Kotani M, Ikegami S (2004) Synlett 959
89. Mukai C, Itoh T, Hanaoka M (1997) Tetrahedron Lett 38:4595
90. Osa Y, Takeda K, Sato, T, Kaji E, Mizuno Y, Takayanagi H (1999) Tetrahedron Lett 40:1531
91. Kim KS, Kim JH, Lee YJ, Lee YJ, Park J (2001) J Am Chem Soc 123:8477
92. Kim KS, Park J, Lee YJ, Seo YS (2003) Angew Chem Int Ed 42:459
93. Lee YJ, Lee K, Jung EH, Jeon HB, Kim KS (2005) Org Lett 7:3263
94. Kwon YT, Lee YJ, Lee K, Kim KS (2004) Org Lett 6:3901
95. Lee BR, Jeon JM, Jung JH, Jeon HB, Kim KS (2006) Can J Chem 84:506
96. Pettit GR, Xu JP, Gingrich DE, Williams MD, Doubek DL, Chapius JC, Schmidt JM (1999) Chem Commun 915
97. Lee YJ, Lee BY, Jeon HB, Kim KS (2006) Org Lett 8:3971
98. (a) Kim KS, Lee YJ, Kim HY, Kang SS, Kwon SY (2004) Org Biomol Chem 2:2408;
(b) Kwon SY, Lee BY, Jeon HB, Kim KS (2005) Bull Korean Chem Soc 26:815

3.4 *O*-Glycosyl Donors

J. Cristóbal López
Instituto de Química Orgánica General, CSIC,
Juan de la Cierva 3, 28006 Madrid, Spain
clopez@iqog.csic.es

In: *Glycoscience*. Fraser-Reid B, Tatsuta K, Thiem J (eds)
Chapter-DOI 10-1007/978-3-540-30429-6_13:

Abstract

O-Glycosyl donors, despite being one of the last successful donors to appear, have developed themselves into a burgeoning class of glycosyl donors. They can be classified in two main types: *O*-alkyl and *O*-aryl (or hetaryl) glycosyl donors. They share, however, many characteristics, they can be (1) synthesized from aldoses, either by modified Fisher glycosidation (*O*-alkyl) or by nucleophilic aromatic substitution (*O*-aryl or *O*-hetaryl), (2) stable to diverse chemical manipulations, (3) directly used for saccharide coupling, and (4) chemoselectively activated. Among these, *n*-pentenyl glycosides stand apart. They were the first *O*-alkyl glycosyl donors to be described and have paved the way to many conceptual developments in oligosaccharide synthesis. The development of the chemoselectivity-based "armed-disarmed" approach for saccharide coupling, including its stereoelectronic or torsional variants, now extended to other kinds of glycosyl donors, was first recognized in *n*-pentenyl glycosides. The chemical manipulation of the anomeric substituent in the glycosyl donor to induce reactivity differences between related species (sidetracking) was also introduced in *n*-pentenyl glycosides. An evolution of this concept, the "latent-active" strategy for glycosyl couplings, first described in thioglycosyl donors (vide infra), has been elegantly applied to *O*-glycosyl donors. Thus, allyl and vinyl glycosides, 2-(benzyloxycarbonyl)benzyl (BCB) glycosides and 2′-carboxybenzyl (CB) glycosides are useful "latent-active" glycosyl pairs. Finally, unprotect-

ed 3-methoxy-2-pyridyl (MOP) glycosides have been used in glycosylation processes with moderate success.

Keywords

2′-carboxybenzyl (CB) glycosides; 3-methoxy-2-pyridyl glycosides; Armed–disarmed; DISAL glycosyl donors; Halonium ion transfer; Latent-active glycosylation; *n*-Pentenyl glycosides; *O*-heteroaryl glycosyl donors; Oligosaccharide synthesis; Vinyl glycosides

Abbreviations

BCB	2-(benzyloxycarbonyl)benzyl
CAN	cerium ammonium nitrate
CB	2′-carboxybenzyl
DAST	(diethylamino)-sulfur trifluoride
DDQ	2,3-dichloro-5,6-dicyano-*p*-benzoquinone
DISAL	a dinitrosalicylic acid glycoside derivative
DMAP	4-(*N*,*N*-dimethylamino)pyridine
DTBMP	di-*tert*-butylmethylpyridine
IDCP	iodonium di-*sym*-collidine perchlorate
MOP	3-methoxy-2-pyridyl
NBS	*N*-bromosuccinimide
NIS	*N*-iodosuccinimide
NMP	1-methylpyrrolidin-2-one
NPG	*n*-pentenyl glycosides
NPhth	*N*-phthaloyl
NPOE	*n*-pentenyl orthoester
PMB	*p*-methoxybenzyl
PMP	*p*-methoxyphenyl
TBAF	tetra-*n*-butylammonium bromide
TBAI	tetra-*n*-butylammonium iodide
TBSOTf	*tert*-butyldimethylsilyl trifluoromethanesulfonate
TESOTf	triethylsilyl trifluoromethanesulfonate
Tf_2O	trifluoromethanesulfonic anhydride
TfOH	trifluoromethanesulfonic acid
THF	tetrahydrofuran
TMSOTf	trimethylsilyl trifluoromethanesulfonate
TPSOTf	*tert*-butyldiphenylsilyl trifluoromethane sulfonate
Troc	*N*-trichloroethoxycarbonyl
TTCP	*N*-tetrachlorophthaloyl

1 Introduction

From the early days, chemists involved in chemical glycosylation have been trying to develop successful glycosyl donors. In general, the characteristics of a successful glycosyl donor might

include: (*a*) preparation under mild reaction conditions, (*b*) selective activation by reagents that would not interfere with the protecting and functional groups present in the donor and the glycosyl acceptor, and (*c*) good reactivity [1,2,3,4]. More recently, the advent of convergent block synthesis to tackle complex oligosaccharide preparations have also demanded that glycosyl donor building blocks (*d*) are sufficiently stable to be purified and stored for considerable periods of time, and (*e*) are resistant towards a wide range of reaction conditions [5,6,7,8,9]. According to this, *O*-glycosyl donors (the topic of this chapter), because of their remarkable "shelf-life" and stability (conditions *d*, *e*) will be attractive candidates for oligosaccharide block synthesis, provided that conditions *a*, *b*, and *c* are also met.

A chronology, displayed in ● *Fig. 1*, highlights the relatively recent arrival of *O*-glycosyl donors to the assortment of relevant glycosyl donors. In fact, the first *O*-alkyl glycosyl donor (*n*-pentenyl glycoside) [10], was introduced more than a century after the first glycosylation was described (synthesis of aryl glycosides from glycosyl chlorides) [11].

This late arrival is understandable on the basis of the outline in ● *Scheme 1a*, which makes it obvious that in situ transformation of one alkyl glycoside donor into a disaccharide (or into another alkyl glycoside) could be problematic. The acidic conditions normally used to cleave alkyl glycosides (**1**), generating oxocarbenium ion **2**, could tamper with the newly formed intersaccharidic linkage in **3**, notwithstanding the liberation of alkanol that might compete for glycosylation with the sought glycosyl acceptor, thus regenerating **1** (● *Scheme 1b*). The successful implementation of the strategy represented in ● *Scheme 1a* would imply that: (*a*) the alkanols have to be released under a non-nucleophilic form, and (*b*) the newly formed glycosidic linkage must be compatible with the promoter employed.

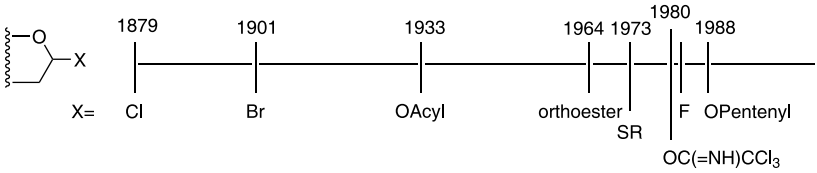
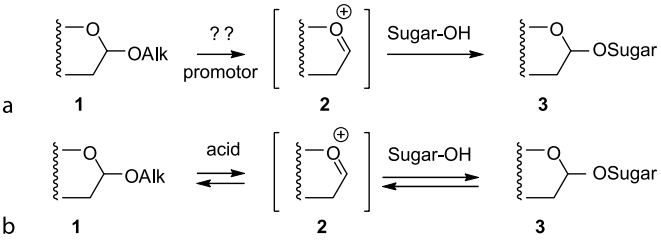


Figure 1
Chronology of selected glycosyl donors



Scheme 1
***O*-Alkyl-glycosyl donors in glycosylation**

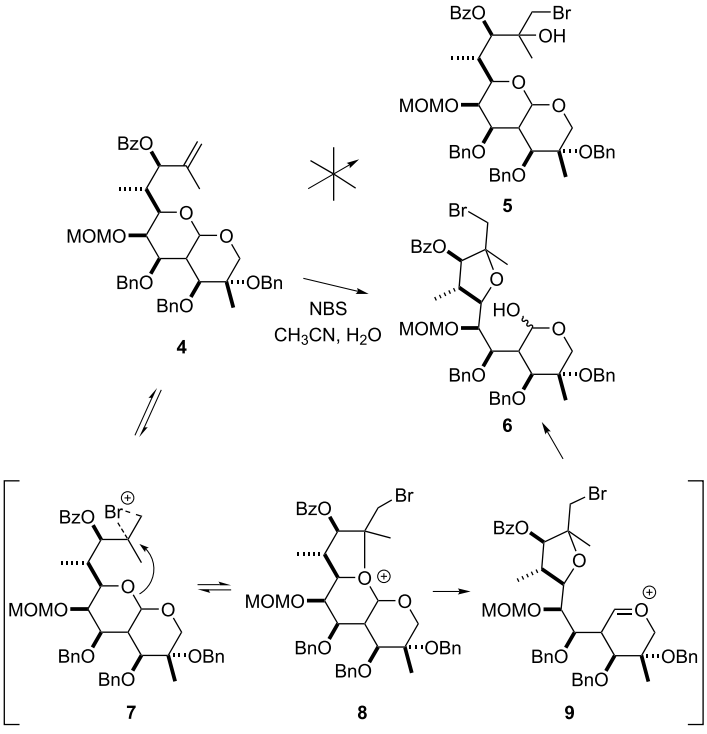
2 *n*-Pentenyl Glycosides

2.1 Introduction

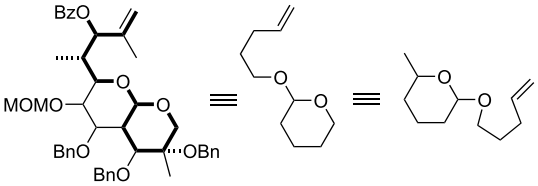
2.1.1 The Origin of *n*-Pentenyl Glycosides (NPGs)

The discovery of *n*-pentenyl glycosides (NPGs) [12], was derived from an observation made by Mootoo and Fraser-Reid in a completely unrelated project [13]. Attempted formation of bromohydrin **5** by reaction of **4** with NBS in 1% aqueous acetonitrile led, instead to bromomethyl tetrahydrofuran **6** (● *Scheme 2*) [14]. To rationalize this transformation (**4**→**6**), the authors invoked a 5-*exo*-cyclization [15] of the pyranosidic oxygen in **7** leading to cationic intermediate **8**, and thence to oxocarbenium ion **9**, that upon capture of H_2O led to hemiacetal **6**. The overall result of the process had been a, nonhydrolytic, electrophilic unravelling of the glycosidic-type bond in **4**.

The overlap between structures **4** and **10** permitted the authors to design structure **11**, as a candidate for testing electrophilic deprotection at the anomeric center of a pyranose (● *Scheme 3*). It has now become clear, 20 years after this observation, that the correlation shown in ● *Scheme 3* led to a breakthrough in glycoside synthesis.



Scheme 2
The origin of *n*-pentenyl glycosides



4 **10** **11**

Scheme 3
The design of *n*-pentenyl glycosides

2.1.2 Chemoselective Liberation of the Anomeric Group in NPGs

To test the validity of their assumption, Mootoo and Fraser-Reid prepared NPGs **12**–**18** and treated them with NBS in 1% aqueous acetonitrile [16]. Their results, summarized in ❯ *Table 1*, showed that differently substituted NPGs could be chemoselectively liberated at the anomeric center to yield hemiacetals **19**–**24**. Furthermore, benzylidene, silyl, p-methoxybenzyl (PMB), ethoxyethyl, and allyl protecting groups proved to be compatible with the conditions employed in the deprotection of the anomeric pent-4-enyl group. Diol **18**, however, furnished a complex reaction mixture probably related to competing glycosylation processes, vide infra.

2.2 NPGs as Glycosyl Donors

To test the potential of NPGs as glycosyl donors, Fraser-Reid et al. first examined the reaction of compound **12** with NBS in MeCN-MeOH [10]. The reaction took place in 3 h, at room temperature yielding methyl glucoside **25** in 85% yield as a 1:3 (α:β) anomeric mixture (❯ *Table 2*, entry i). The utilization of iodonium di-*sym*-collidine perchlorate (IDCP) [17] as a promoter resulted in a faster reaction (0.5 h), which maintained the previous anomeric mixture (❯ *Table 2*, entry ii). The use of a 4:1 mixture of Et_2O-CH_2Cl_2 as solvent, to favor α-glucoside formation while solubilizing IDCP, resulted in a 3:1 (α:β) anomeric mixture of **25** (❯ *Table 2*, entry iii). When CH_2Cl_2 was used as a solvent a 1.2:1 (α:β) anomeric mixture was obtained.

NPGs were next tested in the elaboration of disaccharides by glycosylation of monosaccharide acceptors. Gluco- (**12**), *manno*- (**26**), and 2-deoxy- (**27**) NPGs reacted with sterically demanding methyl glucoside **28**, to give disaccharides **29**–**31** (❯ *Table 3*). Gluco-derivatives gave the best α versus β solvent dependence, MeCN favoring β, and Et_2O favoring α (the same trend as noted for MeOH, ❯ *Table 2*). For the *manno*- and 2-deoxy donors **26** and **27**, no consistent pattern of solvent dependence was noticed. 2-Deoxy-donor **27**, gave appreciable α-selectivities with secondary acceptors (❯ *Table 3*, entries vii–ix). MeCN gave the lowest overall yield of disaccharide products (❯ *Table 3*, entries i, iv, vii). Reactions of 2-deoxy NPG **27** (❯ *Table 3*, entries vii–ix) were generally much faster than those of either **12** or **26**, an observation that parallels the observed trends in acid lability of the three donors (❯ *Table 3*, entries i–vi). With Et_2O as solvent, reactions with primary alcohol acceptors displayed more stereoselectivity than reactions with secondary hydroxyl acceptors.

Table 1
Oxidative hydrolysis of some NPGs with NBS in 1% aqueous acetonitrile

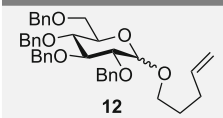
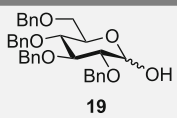
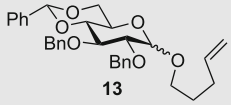
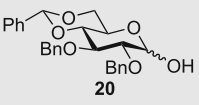
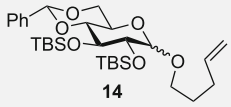
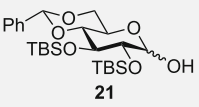
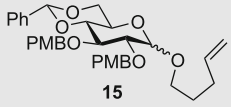
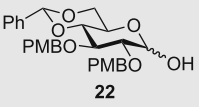
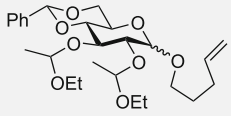
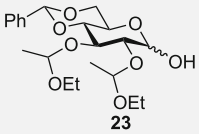
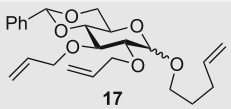
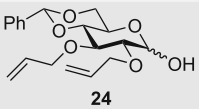
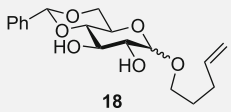
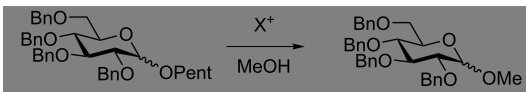
Substrate	Product	Yield (%)
 **12**	 **19**	85
 **13**	 **20**	70
 **14**	 **21**	90
 **15**	 **22**	68
 **16**	 **23**	63
 **17**	 **24**	72
 **18**	complex mixture	

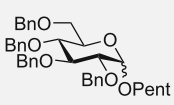
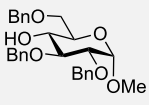
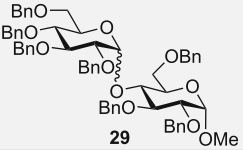
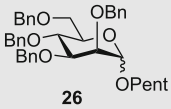
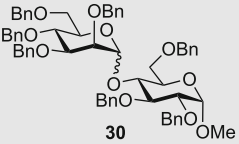
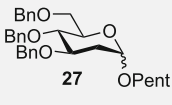
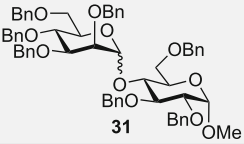
Table 2
Reaction of NPG 12 with methanol in the presence of halonium ions



12 **25**

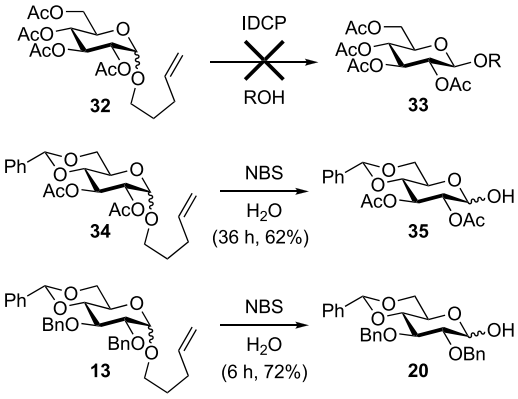
Entry	Promoter	Time (h)	Solvent	α:β	Yield (%)
I	NBS	3	MeCN	1:3	85
ii	$I(collidine)_2ClO_4$	0,5	MeCN	1:3	75
iii	$I(collidine)_2ClO_4$	0,5	CH_2Cl_2	1.2:1	85
iv	$I(collidine)_2ClO_4$	24	CH_2Cl_2-Et_2O	3:1	75

Table 3
Direct elaboration of NPGs into disaccharides on treatment with IDCP

Entry	NPG	Acceptor	Time (h)	Solvent α:β	Product	Yield (%)
i	 **12**	 **28**	1–2	MeCN 1:2	 **29**	20
ii	12	28	1–2	CH_2Cl_2 1.2:1	29	75
iii	12	28	16–24	Et_2O/CH_2Cl_2 3:1	29	95
iv	 **26**	28	2–5	MeCN 6:1	 **30**	36
v	26	28	2–5	CH_2Cl_2 9:1	30	76
vi	26	28	24	Et_2O/CH_2Cl_2 α only	30	92
vii	 **27**	28	0,5	MeCN 7:3	 **31**	51
viii	27	28	0,5	CH_2Cl_2 4:1	31	57
ix	27	28	4–6	Et_2O/CH_2Cl_2 4:1	31	61

2.2.1 Acyl-Substituted NPG Donors

The results in *Table 3* made it clear that the use of different solvents to induce (α versus β) stereoselectivity in glycosyl couplings of NPGs was only moderately successful, generally leading to anomeric mixtures [10]. As it has been established, good stereocontrol in the formation of 1,2-*trans* glycosidic linkages can be conveniently obtained with the assistance of a neighboring participating group, generally an acyl moiety [18]. In this context, Fraser-Reid and co-workers examined the glycosylation of NPG **32** for the preparation of 1,2-*trans* glycoside **33** (*Scheme 4*). Unfortunately, the reaction did not lead to glycoside **33**, but to compounds resulting from addition across the terminal double bond of the pent-4-enyl moiety. Along this line, the authors had previously noticed that hydrolysis of acyl derivative **34** was considerably slower than that of **13** [16]. These results paralleled previous observations by Paulsen [1] on the deactivating effect of esters versus ether protecting groups upon differently substituted glycosyl halides. These observations, however, according to the state-of-the-art in glycosylation in 1988, only meant that acyl-NPGs would not be useful as glycosyl donors.



[a] The reaction was stopped at 50% conversion, the yield is based on recovered **34**.

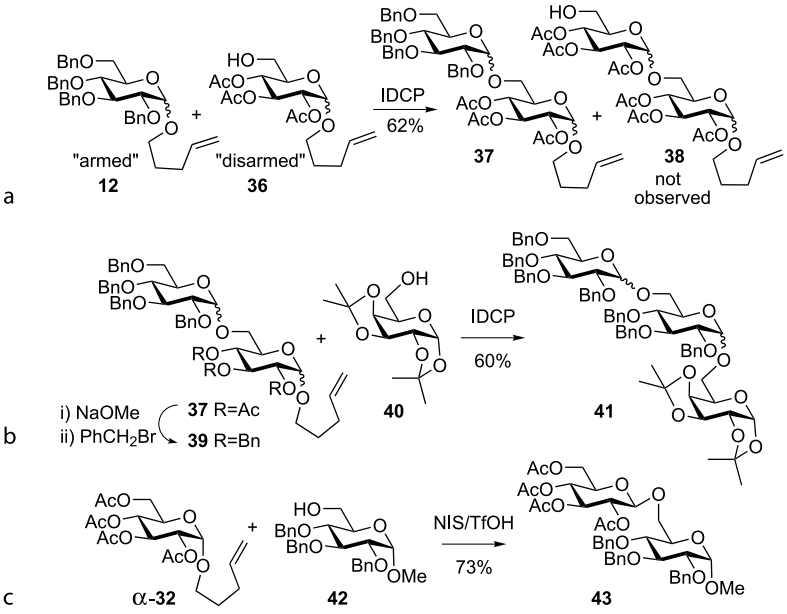
Scheme 4
Reaction of acylated NPGs with halonium ions

2.3 *Armed–Disarmed* Strategy for Glycosyl Coupling

Fraser-Reid and co-workers, however, anticipated that the difference in reactivity found in differently substituted NPGs could be applied in a chemoselective protocol for glycosyl coupling [19]. The activated and the deactivated NPGs were termed "armed sugar" and "disarmed sugar", respectively. Thus, as illustrated in *Scheme 5a*, coupling of **12** and **36**, mediated by IDCP afforded a 62% yield of disaccharide **37** [19]. Therefore, the acyl groups of **36** indeed "disarmed" the NPG, thereby ensuring that **12** served as the only glycosyl donor. No evidence for a hexaacetyl disaccharide **38**, arising from self-condensation of **12** was found, nor of further reaction of (disarmed) disaccharide **37** with the acceptor **36**.

The chemoselective coupling, however, is not the only quality of the armed-disarmed strategy for glycosyl assembly. An additional aspect of this strategy is the ability to "rearm" disarmed glycosyl donors for further glycosyl coupling. Thus, "disarmed" **37** was converted to "armed" disaccharide **39**, (by replacing the acetyl groups with benzyl substituents) which could then glycosylate galacto- derivative **40**, to yield trisaccharide **41**, in 60% yield (*Scheme 5b*). An alternative way of "rearming" NPGs by increasing the potency of the promoter used for glycosylation was also introduced by the same authors [20]. According to that, an iodonium ion generated in situ from NIS and TfOH was able to promote the coupling of "disarmed" pent-4-enyl glycosides (e. g. α-**32**, *Scheme 5c*) with acceptors to give 1,2-*trans* disaccharides, e. g. **43**, via neighboring group participation [18].

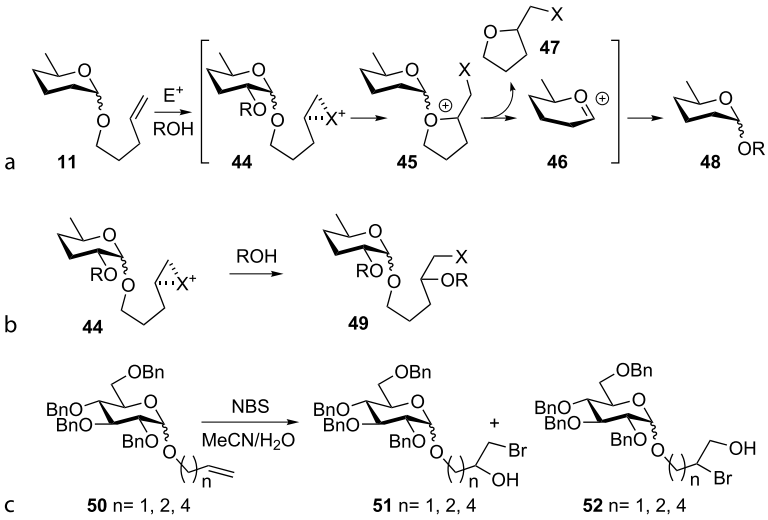
The armed-disarmed concept takes advantage of reactivity differences induced by the ring substituents on the anomeric leaving group and, although originally described for NPG donors, it has been extended to various types of glycosyl donors. These include thioglycosides [21,22], glycals [23], glycosyl fluorides [24], selenoglycosides [25], glycosyl phosphoroamidates [26,27], glycosyl thioformimidates [28,29,30], and *S*-benzoxazolyl glycosides [31,32,33].



Scheme 5
Armed-disarmed strategy for chemoselective glycosyl couplings

2.3.1 Mechanistic Aspects of the Oxidative Hydrolysis of NPGs

The currently accepted mechanism for the reaction of NPGs (e. g. **11**) with halonium ions is outlined in *Scheme 6a*. The oxygen in the acetal function participates in a favored 5-*exo-tet* ring opening of an intermediate cyclic halonium ion, **44** [15]. The ensuing furanilium ion **45**, evolves by splitting off a non-nucleophilic halotetrahydrofuran **47** [34], thus leading to oxocarbenium ion **46**, that can trap the nucleophile (ROH). The overall result is the cleavage of the acetal moiety with the formation of a new glycosyl derivative, **48**. Madsen and Fraser-Reid have demonstrated that even when the system NIS/TESOTf [20] is used to promote the cleavage of NPGs, the reaction is not acid catalyzed but still halonium ion catalyzed [35].
In this connection, the question of why the reaction of an NPG in the presence of water leads to an aldose **48** (R=H), rather than to a halohydrin **49** (R=H) was raised (*Scheme 6b*). In fact, the successful cleavage observed for NPGs rests on two issues: (*a*) the concentration of nucleophile (water in the case of hydrolysis), and (*b*) the rate of the 5-*exo-tet* cyclization, **44**→**45**. Pertinent to question *a* the intramolecular reaction **44**→**45**, is preferred to the bimolecular reaction with water leading to a halohydrin, **44**→**49**, under the conditions used by the authors. An increase in the concentration of water would enhance the rate of the bimolecular reaction, without affecting the intramolecular process. Indeed added water led to the formation of bromohydrin **49** (R=H) [36]. Related to the second issue, Rodebaugh and Fraser-Reid examined the same reaction with allyl, butenyl, and hexenyl glycosides **50** ($n = 1, 2, 4$), differing on the rate of cyclization compared to NPGs [37,38]. They found that, unlike NPGs, they all gave rise to isomeric halohydrins **51** and **52**, upon treatment with NBS in aqueous MeCN (*Scheme 6c*).



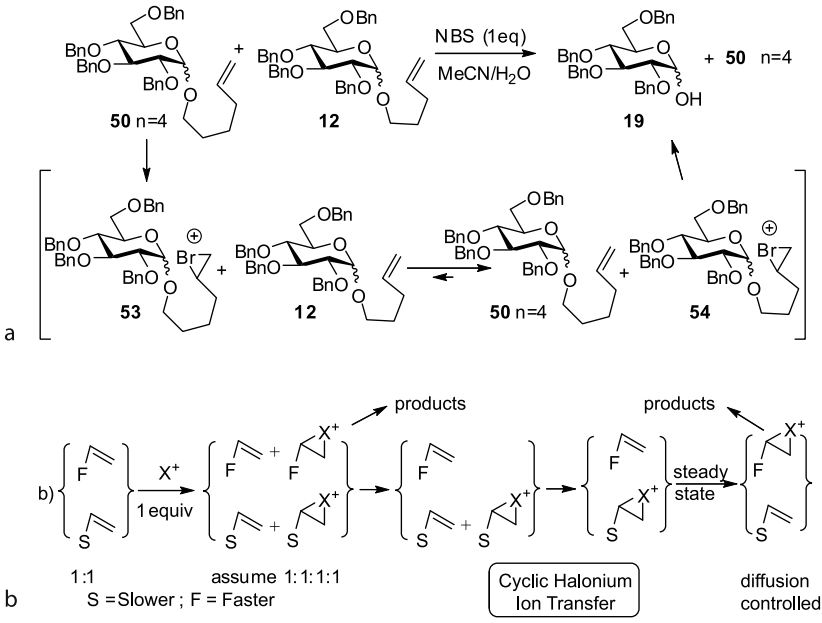
Scheme 6
Oxidative cleavage of NPGs

2.3.2 Evidence for Intermolecular Halonium-Ion Transfer

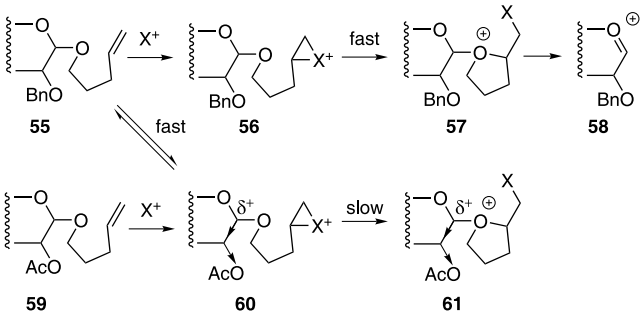
In a related experiment, hexenyl glycoside **50** ($n = 4$), that has been found to react 2.3-times slower than NPG **12** [37,38], was made to compete with **12** for 1 equiv. of NBS. The hexenyl glycoside **50** ($n = 4$) was recovered unchanged together with hemiacetal **19**, arising from the hydrolysis of **12** (*Scheme 7a*). Rodebaugh and Fraser-Reid proved that this phenomenon was due to a diffusion-controlled intermolecular halonium-ion transfer (a similar process has been previously noted by Brown and co-workers for "sterically encumbered olefins" [39]). Accordingly, bromonium species **53**, obtained by irreversible reaction of **50** with NBS [37,38], would undergo a fast bromonium ion transfer to NPG **12** leading to halonium **54** (i. e. **44**, *Scheme 6a*) which will undergo a fast transformation to **19**. The general process, a classic example of Le Chatelier's principle, is represented in *Scheme 7b*. When two alkenes are made to compete for one equivalent of halonium ion X^+, a steady-state regime can be envisaged whereby the faster (F) (e. g. NPG **12**) reacts completely, and the slower (S) (e. g. **50**, $n = 4$) is recovered completely.

2.3.3 Intermolecular Halonium-Ion Transfer: A Key Factor in the Implementation of the Armed-Disarmed Protocol

This intermolecular halonium ion transfer had indeed been postulated earlier as the key factor to account for the absence of self-coupling product **38**, when armed and disarmed NPGs, **12** and **36** respectively, were made to compete for one equivalent of NBS (*Scheme 5a*). The observed 6-fold difference in the hydrolysis rates of **36** and **12** should have resulted in the presence of **38**, in at least 10% [40]. Irreversible reaction of NPGs **55** and **59** with NBS leads to halonium ions **56** and **60**, respectively (*Scheme 8*). The transfer of halonium (e. g. **60**→**55**) is reversible and rapid compared with the subsequent steps leading to glycoside formation



Scheme 7
Intermolecular halonium-ion transfer



Scheme 8
Halonium ion transfer: a key factor in armed-disarmed couplings

(**56**→**57**→**58**). By corollary, the inherent reactivity of the glycosyl donors is thus revealed in the final product distribution. If the acceptor functionality is located in the less reactive component, selective glycosylations take place leading to a specific disaccharide.

2.3.4 Torsional Disarming of NPGs

Fraser-Reid and co-workers found that widely used cyclic acetals also affected anomeric reactivity [41]. They showed that these reactivity differences could be applied to an armed-

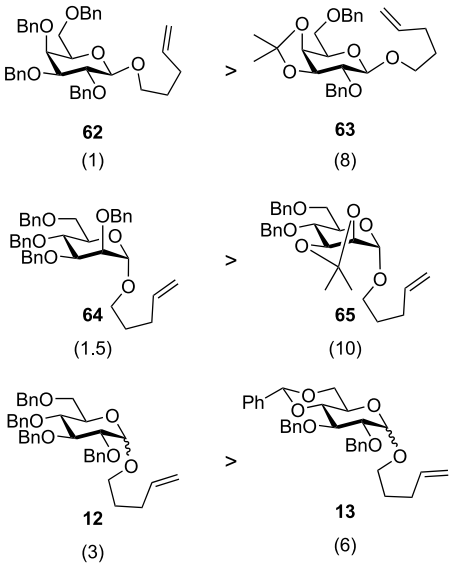
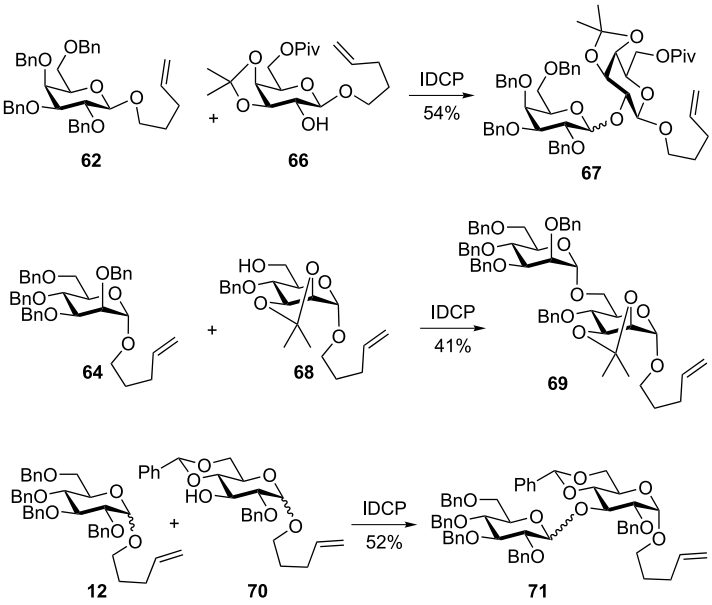


Figure 2
Relative rates of oxidative hydrolysis for acetal-protected NPGs

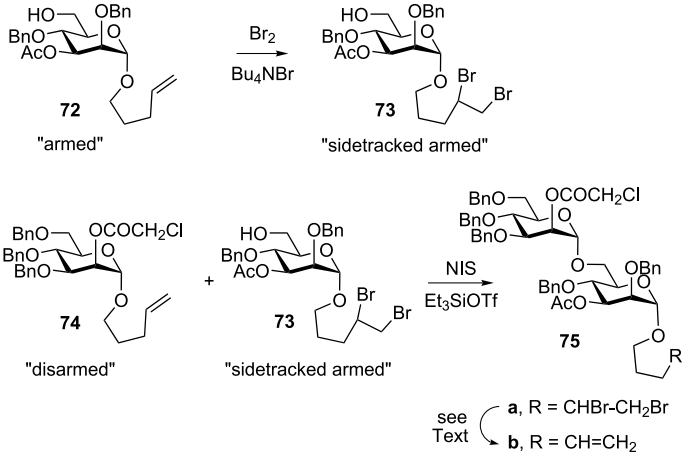
disarmed protocol based on torsional, rather that electronic, effects. The measured experimental relative rates of oxidative hydrolysis for some pairs of *galacto*- (**62**, **63**), *manno*- (**64**, **65**), and *gluco*- (**12**, **13**), acetalated and nonacetalated NPGs are displayed in *Fig. 2*. From these data, the authors were able to design the successful armed-disarmed couplings shown in *Scheme 9*. These reactivity differences were ascribed to the fact that *trans*-fused protection restricts the molecule from ring flexibility, thereby making it increasingly difficult to reach a half-chair transition state from a chair ground state.

2.3.5 Sidetracking of NPGs: A Reversal for the Armed-Disarmed Strategy

In the armed-disarmed protocol, the reactivity differences induced by the ring substituents upon the anomeric center were exploited for chemoselective couplings. In these protocols, the more reactive "armed" NPG always glycosylates the "disarmed" NPG. On the basis of chemical manipulation of the pent-4-enyl moiety, rather than the ring substituents, Fraser-Reid and co-workers showed that "disarmed" NPGs could be used to glycosylate "armed" NPGs [40]. Treatment of NPG **72** with bromine and tetra-*n*-butylammonium bromide (TBAF), in a bimolecular reaction (e. g. **44**→**48**, *Scheme 6b*) yielded dibromoderivative (*Scheme 10*) **73**. Glycosylation of the latter with "disarmed" **74**, under the agency of NIS/TESOTf, yielded disaccharide **75a**, which could be transformed to the pentenyl disaccharide **75b**. Several methods proved to be successful for the restoration of the double bond from the dibromoderivative including, (a) Zn/TBAI in sonicating EtOH, (b) NaI in methyl ethyl ketone, and (c) SmI_2 in THF [42]. The choice of the reagent will vary with the



Scheme 9
Armed-disarmed couplings based on torsional effects



Scheme 10
Sidetracking of NPGs in saccharide synthesis

reactivity of the substrate, as well as the protecting groups thereon. More recently, a milder brominating system, the combination of $CuBr_2$ and LiBr in MeCN:THF (3:1), has been used to brominate *n*-pentenyl glycosides containing *O*-benzyl, *O*-*p*-methoxybenzyl, *N*-phthaloyl, and *N*-tetrachlorophthaloyl protecting groups [43].

2.4 Conversion of NPGs to Other Glycosyl Donors

NPGs have been converted into different glycosyl donors.

2.4.1 Conversion to Glycosyl Bromides

Konradsson and Fraser-Reid [44] reported the conversion of NPGs into glycosyl bromides, e. g. **76**, by treatment of the corresponding pentenyl glycoside with a dilute dichloromethane solution of bromine, conditions that favor unimolecular reaction. The reaction was shown to be compatible with acetals, benzyl, silyl, and allyl protecting groups in the NPG (● *Scheme 11a*).

2.4.2 Conversion to Glycosyl Phosphates

Pale and Whitesides [45] described the synthesis of glycosyl phosphates **77** [46,47], by reaction of dibenzyl phosphate with NPG **12**, with the use of either IDCP or NBS as promoters. The authors noted the influence of the solvent (MeCN, Et_2O, CH_2Cl_2) and the promoter in the α/β selectivity of the glycosyl phosphates formed (● *Scheme 11b*).

2.4.3 Conversion to Glycosyl Fluorides

Clausen and Madsen [48] reported the transformation of NPG **78** into glycosyl fluoride **79** by treatment with NBS and (diethylamino)-sulfur trifluoride (DAST) (● *Scheme 11c*).
López et al. [49] described the preparation of glycosyl fluorides **80**, by reaction of NPGs with bis(pyridinium) iodonium (I) tetrafluoroborate (IPy_2BF_4) in the presence of tetrafluoroboric acid (● *Scheme 11d*). The process was compatible with the presence of silyl and benzyl groups in the NPG.

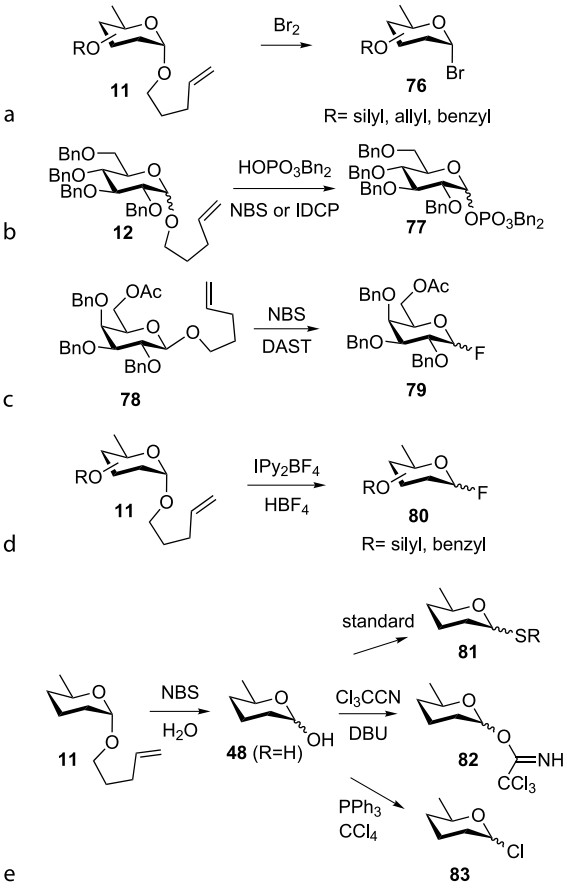
2.4.4 Chemoselective Liberation Followed by Anomeric Activation

The ability to chemoselectively deprotect pent-4-enyl glycosides opens an avenue for a two-step transformation of NPGs into different glycosyl donors. In this context, NPGs can be transformed [50] into thioglycosides **81** [51,52,53], glycosyl trichloroacetimidates **82** [54], and glycosyl chlorides **83** [55] (● *Scheme 11e*).

2.5 NPGs in the Stereocontrolled Assembly of α- and β- Glycoproteins

2.5.1 Pyranosylacetonitrilium Ions from NPGs

Ratcliffe and Fraser-Reid found that acetonitrile was able to trap glycosyl oxocarbenium ions (e. g. **46**), arising from NPGs, to give acetonitrilium ions, e. g. **84** (● *Scheme 12a*) [56]. The latter reacted with water to produce intermediate **85** that evolves to α-amide **86**, in a Ritter-type reaction [57] (● *Scheme 12a*).
This transformation was significant from a mechanistic standpoint. The formation of the α-acetonitrilium ion was not expected on the basis of the reverse anomeric effect (originally defined as the tendency of positively charged substituents at C-1 of a pyranose ring to adopt the equatorial orientation [58]). The authors, however, unambiguously established the α-orientation of

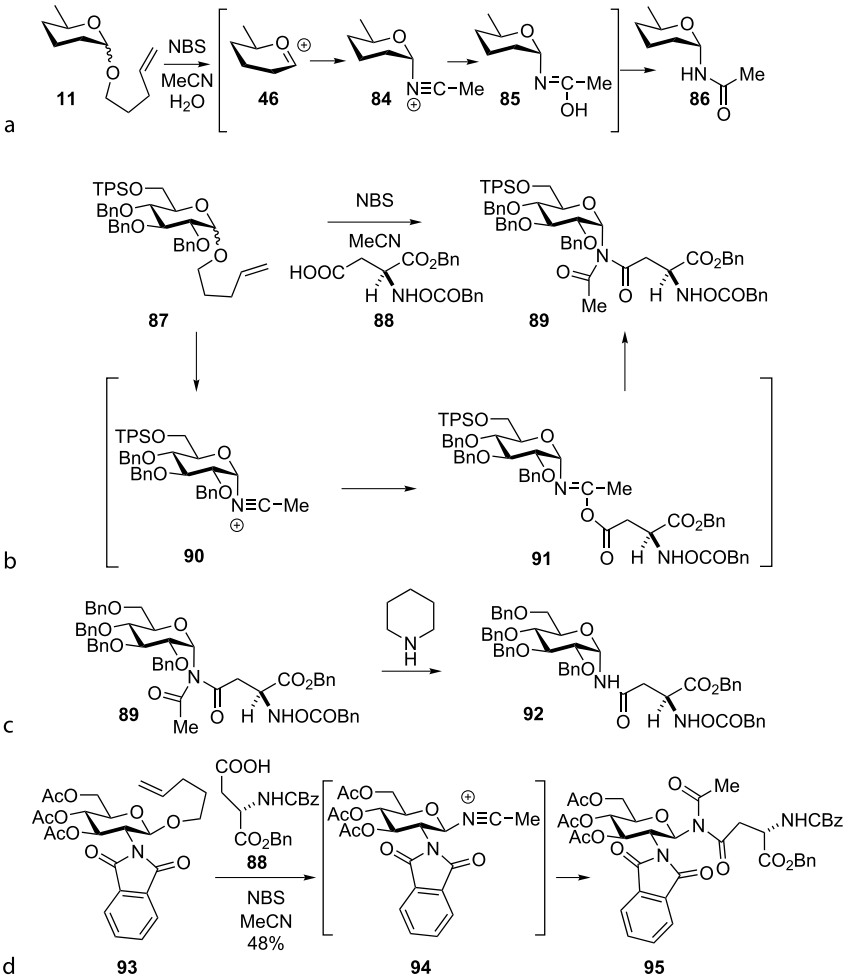


Scheme 11
Conversion of NPGs into different glycosyl donors

the amide and explained this result assuming the formation of the kinetically favored α-D-glucopyranosylacetonitrilium ion, **84** [59].

2.5.2 Synthesis of *N*-α-Linked Glycoproteins from Pyranosylacetonitrilium Ions

The synthetic value of the above-mentioned transformation was considerably enhanced when a carboxylic acid, rather than water, was used to trap the pyranosylacetonitrilium ion (*Scheme 12b*) [59]. Reaction of **87** with aspartic acid derivative **88**, in dry acetonitrile containing NBS, led to α-imide **89** in 61% yield [60,61]. The acetonitrilium ion **90** was trapped by carboxylic acid **88**, to give an imidic anhydride **91**, which rearranged in situ to give the *N*,*N*-diacyl derivative **89**. The route to 2-acetamido-1-*N*-(L-aspart-4-oyl)-2-deoxy-β-D-glycopyranosyl-amine **92**, was completed by selective *N*-deacetylation of **89** with piperidine (*Scheme 12c*) [59].



Scheme 12
Reactions of pyranosylacetonitrilium ions arising from NPGs

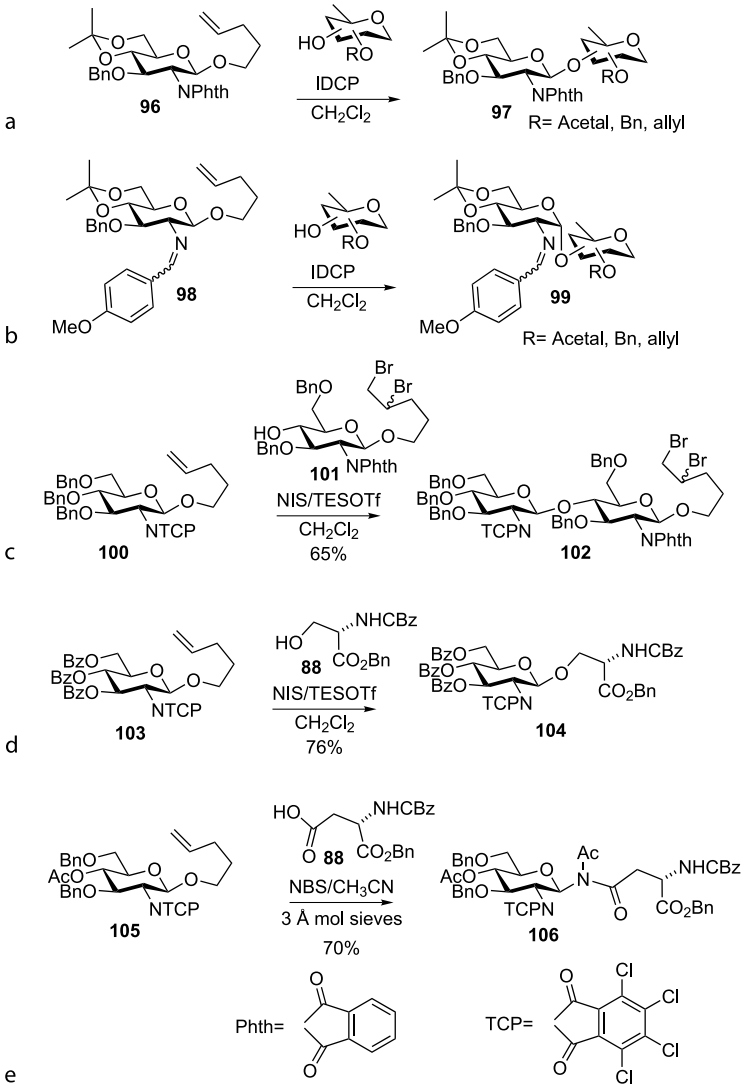
2.5.3 Synthesis of *N*-β-Linked Glycoproteins from Pyranosylacetonitrilium Ions

More interestingly, the presence of a neighboring participating group at C2 induces the formation of β-nitrilium ion intermediates, e. g. **94** (*Scheme 12d*), thus paving the way to β-linked glycoproteins [62]. Accordingly, phthalimido NPG **93**, reacted with aspartic acid derivative **88**, in acetonitrile using NBS as promoter, via the β-nitrilium intermediate **94**, to give the β-asparagine-linked product **95** in 48% yield.

More recently, this method has been elaborated in a three-component-reaction (NPG, acetonitrile, carboxylic acid) route to *N*-glycosylamines [63].

2.6 *n*-Pentenyl 2-Amino-2-Deoxy Glycoside Derivatives as Glycosyl Donors

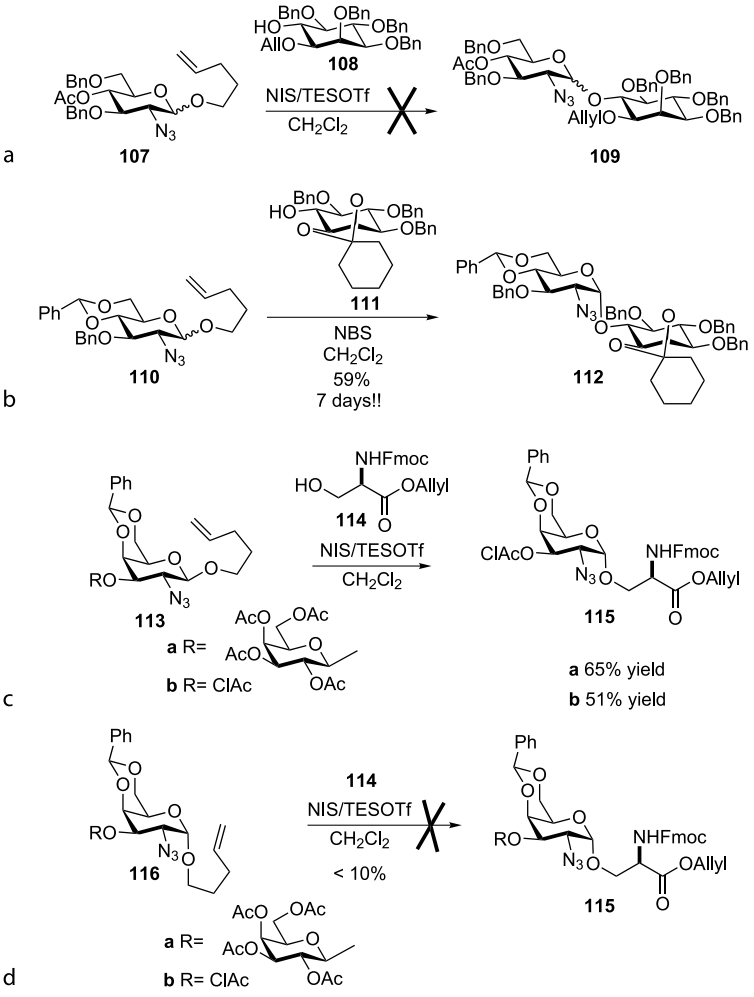
Several pent-4-enyl 2-amino-2-deoxy glycoside derivatives were evaluated as glycosyl donors for the synthesis of 2-amino-2-deoxy oligosaccharides [64]. 2-Deoxy-2-phthalimido **96**, and 2-anisylimino-2-deoxy-D-glucopyranosides **98**, underwent IDCP-induced coupling with



Scheme 13
Glycoside formation from pent-4-enyl 2-amino-2-deoxy glycosides

a variety of sugar alcohols to give β and α disaccharides **97** and **99**, respectively, in moderate to good yields (● *Scheme 13a,b*) [65,66]. 2-Deoxy-2-*N*-tetrachlorophthaloyl NPGs, e. g. **100**, **103**, [67,68,69] are useful donors for the stereocontrolled access to 1,2-*trans* glycosides as exemplified in ● *Scheme 13c,d*. Good yields of disaccharide **102**, and aminoacid **104** were obtained by the use of NIS/TESOTf as promoter. β-*N*-Linked glycopeptide **106** was prepared by treatment of **105** with acid **88** in dry MeCN containing NBS (● *Scheme 13e*).

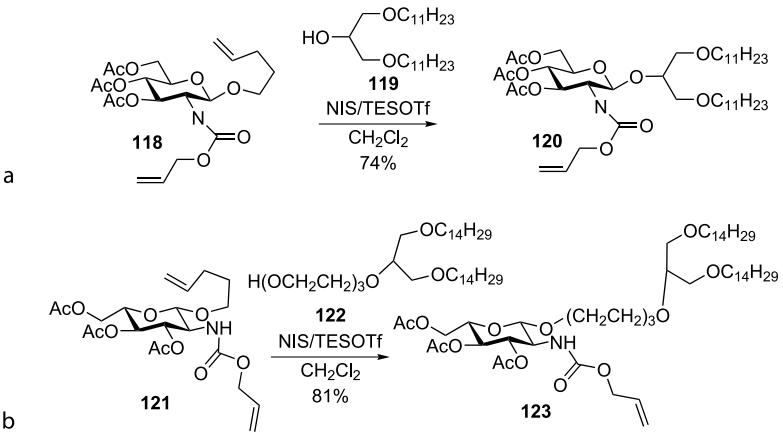
Controversial results have been reported when 2-deoxy-2-azido NPGs were used as glycosyl donors (● *Scheme 14*). Fraser-Reid and co-workers reported that **107** failed to give pseudo-disaccharide **109** upon reaction with acceptor **108** under the agency of NIS/TESOTf



Scheme 14
Glycoside formation from pent-4-enyl 2-deoxy-2-azido NPGs

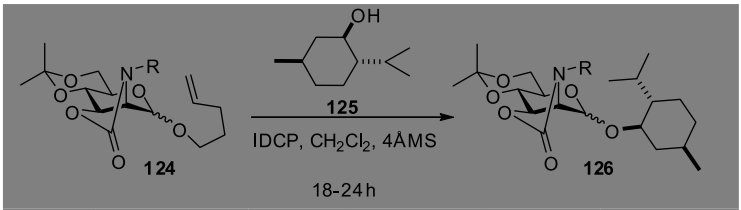
(Scheme 14*a*) [70]. By contrast, good results were obtained with the benzylidenated derivative **110** (Scheme 14*b*). The authors ascribed this result to the conformational constraint imposed by the benzylidene ring, in keeping with their precedents [41].

Svarovsky and Barchi [71] observed a striking reactivity difference between pent-4-enyl β- and α-2-azido-2-deoxy galactosides **113** and **116**, respectively (Scheme 14*c*-X). Thus, where-



Scheme 15
Pent-4-enyl 2-allyloxycarbonyl-2-deoxy-D- and L-glucopyranosides

Table 4
Glycosidation of (−)-menthol (125) by oxazolidinone protected NPGs, 124



Entry	Oxazolidinone	R[a]	IDCP (equiv)	α:β	Yield (%)
i	α	H	2	4:1	53
ii	β	H	2	5:1	24
iii	α	CBz	4	α only	71
iv	β	CBz	4	α only	4
v	α	TCBoc	4	α only	41
vi	α	Troc	4	α only	50
vii	α	Boc	4	α only	63

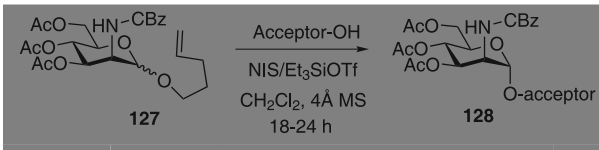
[a] CBz = (benzyloxy)carbonyl, TCBoc = (2,2,2-trichloro-1,1-dimethylethoxy)carbonyl, Troc = (2,2,2-trichlroethoxy)carbonyl, Boc = *tert*-butoxycarbonyl

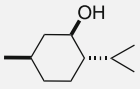
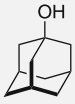
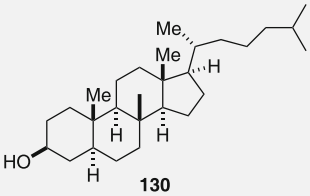
as β-NPGs **113a** and **113b** reacted with serine derivative **114** to give the sought α-glycosyl aminoacids **115a,b**, with complete stereocontrol, the corresponding α-anomers **116a,b** gave very poor yields of **115a,b** (< 10%) and much slower reaction rates.

Fraser-Reid's group advanced the synperiplanar lone-pair hypothesis (SLPH), to account for the fact that β-D-glycopyranosides hydrolyze ≈2–3 times faster [72] than the corresponding α-anomers [73]. This theory advocates that as the reaction progresses synperiplanar lone-pair interactions in the energetically accessible half-chair conformation of the β-anomer are equivalent to the antiperiplanar interactions in the half-chair of the α-anomer (antiperiplanar lone pair hypothesis, ALPH) [74]. On the other hand, the torsional effects associated with the conformational restraint imposed by the presence of the benzylidene ring might enhance this β/α reactivity difference to the point that the α-anomer hardly reacts [75].

2-Allyloxycarbonylamino-2-deoxy D- and L-glucopyranosides **118** and **121**, respectively, have been reported by Lafont and Boullanger [76], to successfully glycosylate 10-tetradecyloxymethyl-3,6,9,12-tetraoxahexacosanol (**119**) and 1,3-bis(undecyloxy)propan-2-ol (**122**), in the course on their studies on neoglycolipids for monolayers (*Scheme 15*). In this case, the chemoselectivity in the reaction of the anomeric pent-4-enyl moiety in the presence of the

Table 5
Stereocontrolled glycosylation using *N*-CBz NPG donor 127



Entry	Donor	Acceptor-OH	α:β	Yield (%)
				
i	α	**125**	α only	63
ii	β	125	α only	60
				
iii	α	**129**	α only	66
iv	β	129	α only	60
				
v	α	**130**	α only	68
vi	β	130	α only	63

[a] CBz = (benzyloxy)carbonyl

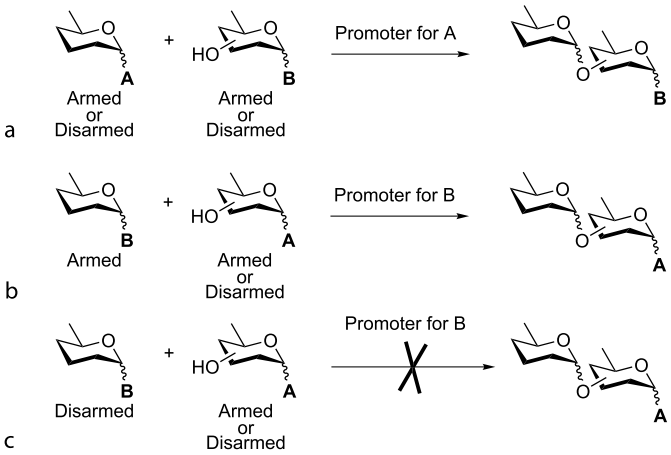
allyloxycarbonyl group is noteworthy. Related *n*-pentenoyl derivatives have been reported as efficient protecting groups for amines by its mild deprotection with iodine in THF-water [77]. Rojas and co-workers [78] described a novel synthetic route to α-linked 2-deoxy-2-mannosamine derivatives, which involved a stereocontrolled glycosidation step of NPG oxazolidinones (e. g. **124**, ❯ *Table 4*) and *N*-CBz NPGs (e. g. **127**, ❯ *Table 5*). The authors found a striking difference in reactivity between α- and β-anomers of oxazolidinones **124**. α-NPG oxazolidinones served as highly stereoselective donors (❯ *Table 4*, entries iii, v–vii), whereas the β-anomer was nearly inert (❯ *Table 4*, entries ii, iv). However, regioselective *N*-CBz oxazolidinone ring opening to **127**, prior to glycosylation permitted elaboration of either NPG anomer to the desired α-Man-NCBz products **128** (❯ *Table 5*).

2.7 Semi-Orthogonal Couplings of NPGs

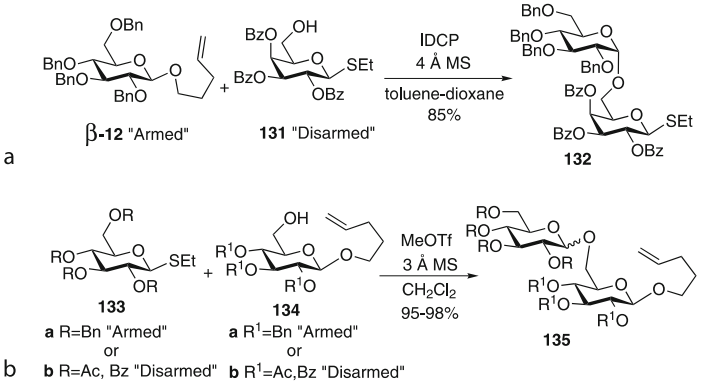
The term "semi-orthogonality" between glycosyl donors (e. g. **A** and **B**, ❯ *Scheme 16*) was introduced by Demchenko [79]. It indicates that whereas selective activation of armed and disarmed glycosyl donor **A** can be effected in the presence of either armed or disarmed donor **B** (❯ *Scheme 16a*), the opposite is not feasible. Thus, in semi-orthogonal donors, the selective activation of disarmed glycosyl donor **B** in the presence of glycosyl donor **A** can not be accomplished (❯ *Scheme 16c*).

2.7.1 Semi-Orthogonality of *O*-Pentenyl and *S*-Ethyl Glycosides

Demchenko and De Meo found conditions for the selective activation of NPGs and ethyl 1-thioglycosides (❯ *Scheme 17*) [79]. They demonstrated that armed NPGs (e. g. β-**12**)



Scheme 16
Semi-orthogonality of glycosyl donors

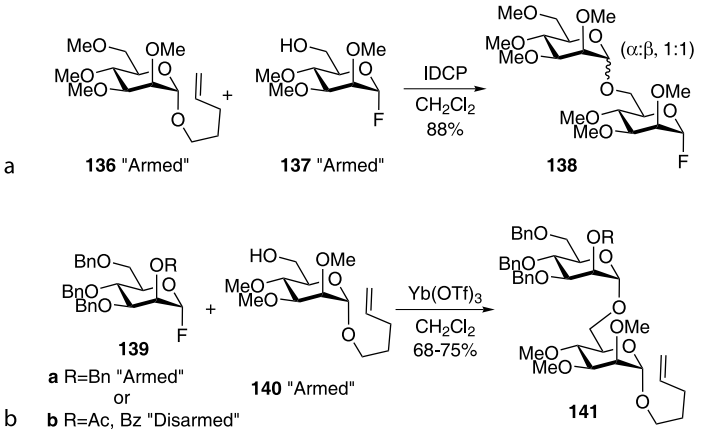


Scheme 17
Semi-orthogonality of *O*-pentenyl and *S*-ethyl-glycosides

could be activated in the presence of thioglycosides (e. g. **131**) with IDCP as the promoter (*Scheme 17a*). On the other hand, the use of methyl triflate (MeOTf) permitted the activation of disarmed thioglycosides (e. g. **133b**) in the presence of armed or disarmed NPGs, **134** (*Scheme 17b*).

2.7.2 Semi-Orthogonality of NPGs and Glycosyl Fluorides

López et al. reported the selective activation of armed NPGs (e. g. **136**) in the presence of armed glycosyl fluorides (e. g. **137**) on treatment with IDCP (*Scheme 18a*) [49]. On the other hand, armed and disarmed glycosyl fluorides **139**, could be activated in the presence of armed NPGs (e. g. **140**) on treatment with ytterbium triflate ($Yb(OTf)_3$) (*Scheme 18b*).

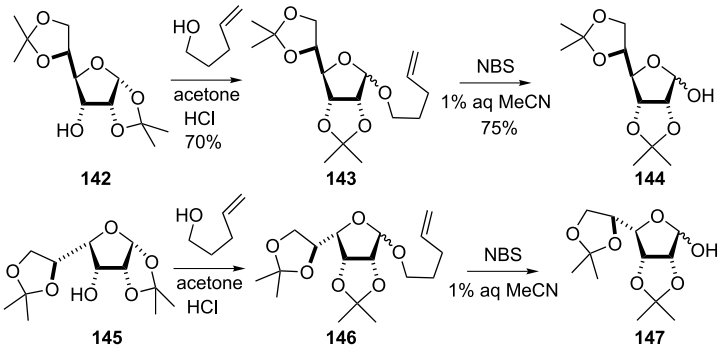


Scheme 18
Semi-orthogonality of NPGs and glycosyl fluorides

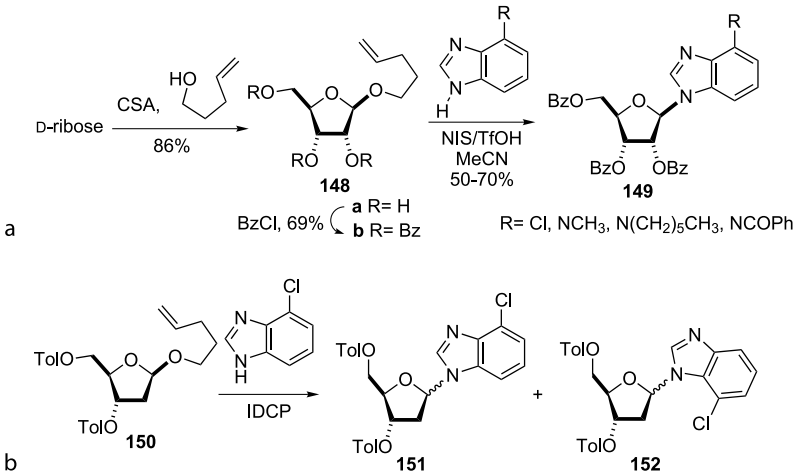
2.8 *n*-Pentenyl Furanoside Donors

2.8.1 Chemoselective Deprotection of the Anomeric Center

Unlike *n*-pentenyl pyranosides, the corresponding furanosides have attracted comparatively little attention. Sharma and Rao reported the preparation of *n*-pentenyl D-allo-, and D-gulo-furanosides **143** and **146**, respectively (*Scheme 19*) [55]. They made use of an efficient acid-induced rearrangement of diacetonides **142** and **145**, in the presence of *n*-pentenyl alcohol. The ensuing pent-4-enyl diacetonides **143** and **146**, were chemoselectively cleaved to hemiacetals **144** and **147**.



Scheme 19
Pent-4-enyl furanosides from D-allose and D-gulose diacetonides



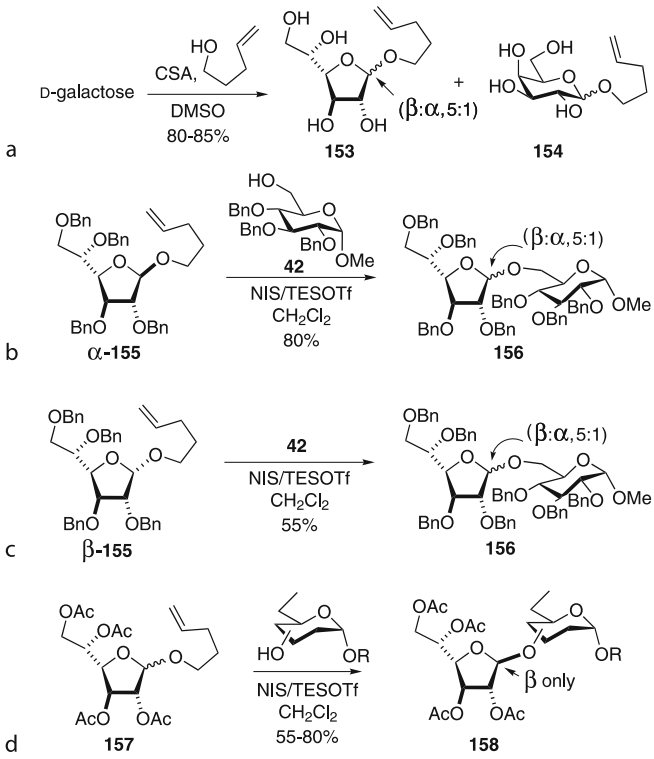
Scheme 20
Pentenyl ribofuranosides in the synthesis of purine nucleosides

2.8.2 Application to the Synthesis of Nucleosides

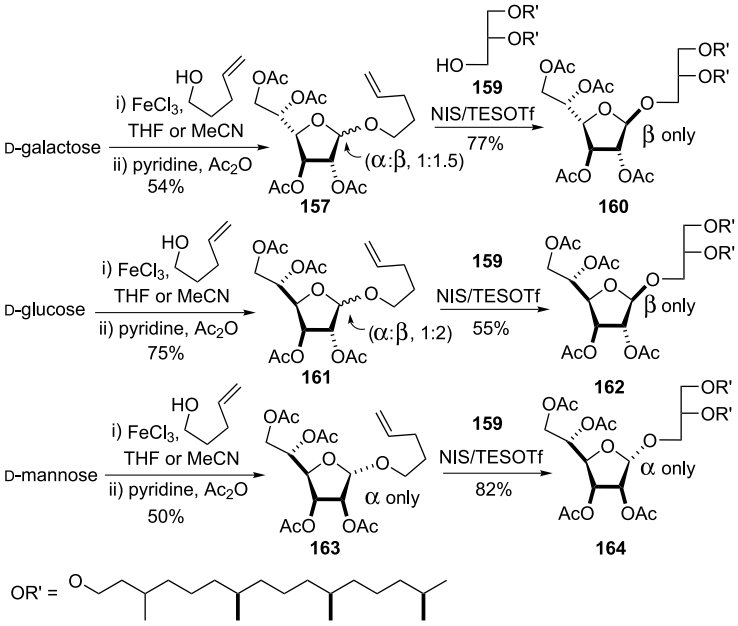
Chapeau and Marnett developed a synthetic route to purine nucleosides from *n*-pentenyl ribosides (Scheme 20) [80]. The authors used a Fischer glycosylation of D-ribose with 4-pentenol to produce pent-4-enyl β-D-erythro-pentofuranoside **148a**, in 86% yield. Reaction of the latter with benzoyl chloride produced their glycosyl donor **148b**, in 69% yield. Addition of TfOH to acetonitrile solutions containing **148b**, the selected purine, and NIS, resulted in a rapid coupling to form the desired nucleosides **149**, in a stereocontrolled manner with yields ranging from 50 to 70% (Scheme 20a). The absence of an acyl group at *O*2 in 2-deoxy NPG **150**, enhanced its reactivity to iodonium sources so IDCP could be used as the promoter. Thus, reaction of **150** with 6-chloropurine in acetonitrile was neither regio- nor stereo-selective, yielding four coupling products **151**α,β and **152**α,β, in similar amounts (Scheme 20b).

2.8.3 *n*-Pentenyl Furanosides as Glycosyl Donors

Arasappan and Fraser-Reid described the preparation of *n*-pentenyl galactofuranosides and evaluated their prospects as glycosyl donors (Scheme 21) [81]. Fischer glycosidation of D-galactose under kinetic conditions using *n*-pentenyl alcohol and DMSO as co-solvent [82]



Scheme 21
Pentenyl galactofuranosides



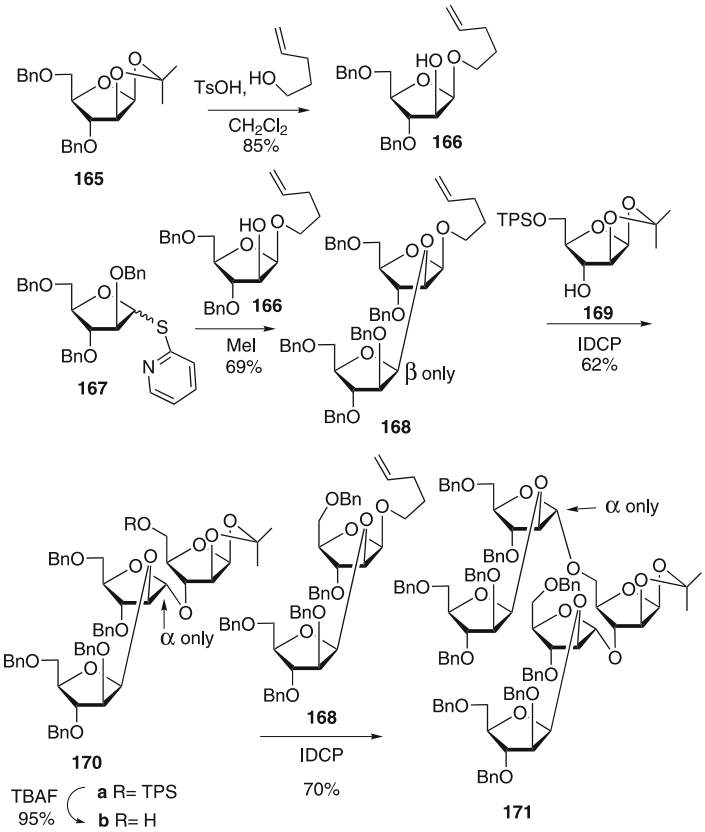
Scheme 22
Synthesis of archaeol glycolipid analogues from pent-4-enyl furanosides

afforded an anomeric α/β (1:3) mixture of *n*-pentenyl galactofuranosides **153**, (≈80–85% yield), contaminated with small amounts of the corresponding *n*-pentenyl galactopyranosides, **154** (*Scheme 21a*). Glycosylation with α- or β- pentenyl glycosides **155** was irrelevant to the product β/α ratio, both favoring the β-furanoside β-**156** (*Scheme 21b,c*). Reactions with donors α- or β-**157** resulted in the β-linkage product **158**, exclusively (saccharide acceptors with free hydroxyl groups at C-2, C-4, and C-6 were assayed), presumably due to the neighboring group participation of the C-2 ester functionality (*Scheme 21d*).

Plusquellec and co-workers reported an improved method for the preparation of *n*-pentenyl furanosides [83] based on their previously described use of $FeCl_3$ as a catalyst in Fischer-type glycosylations [84]. Accordingly, D-glucose, D-galactose, and D-mannose upon treatment with $FeCl_3$ and *n*-pentenyl alcohol followed by in situ acetylation, yielded pent-4-enyl D-gluco-, D-galacto-, and D-mannofuranoside derivatives **157**, **161**, and **163**, respectively in yields ranging from 50 to 75%. Glycosylation of glycerol diether **159** with these donors, promoted by NIS/TESOTf yielded glycolipids **160**, **162**, and **164** in high yields and with excellent 1,2-*trans* stereoselectivity (*Scheme 22*).

2.8.4 *n*-Pentenyl Arabinofuranosides in the Assembly of Oligoarabinans of *Mycobacterium tuberculosis*

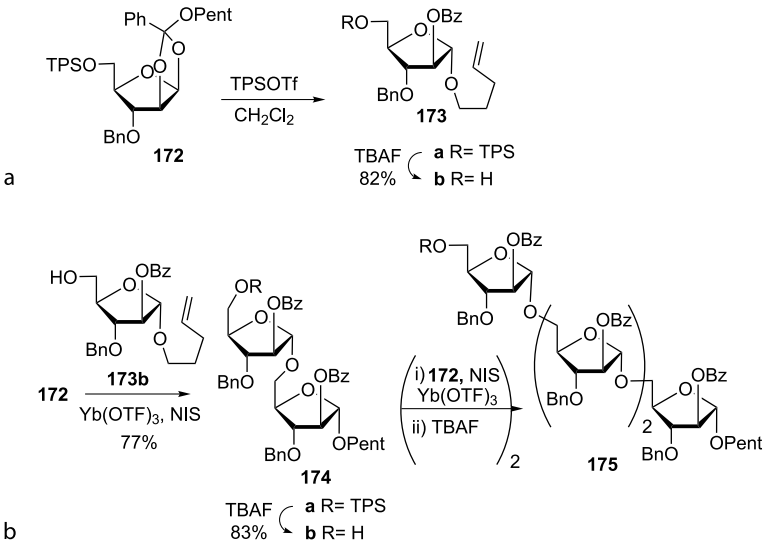
Recent interest in oligoarabinans, have been triggered by their presence in the lipoarabinomannan polysaccharide component of the cell wall complex of mycobacteria [85]. Several research groups have employed *n*-pentenyl arabinofuranosides in their approaches to oligoarabinans.



Scheme 23
Synthesis of a pentaarabinofuranosyl structure motif of *Mycobacterium tuberculosis*

n-Pentenyl β-D-arabinofuranoside **166**, readily prepared from **165**, was employed as acceptor/donor in Gurjar's approach to arabinosyl pentasaccharide **171** (*Scheme 23*) [86]. Accordingly, **166** was glycosylated with *S*-(2-pyridyl)-1-thiofuranose **167** to yield, in a stereoselective manner, β-disaccharide **168**. The latter, itself a pentenyl donor, was then used as a glycosyl donor in two glycosylation events. First, IDCP-promoted glycosylation of silyl derivative **169** yielded trisaccharide **170a** in 62% yield, and in a stereocontrolled manner. Desilylation of **170a** furnished **170b**, which then functioned as the acceptor in the second IDCP-induced glycosylation with **168** to produce pentasaccharide **171**.

More recent studies by Fraser-Reid's group, have focused on the use of NPOEs both as arabinofuranosyl donors, and as convenient starting materials for the preparation of *n*-pentenyl arabinofuranosyl acceptors [87]. TPSOTf-induced rearrangement of NPOE **172**, followed by desilylation afforded pentenyl glycoside **173b** (*Scheme 24*). Glycosylation of pentenyl glycoside **173b** with NPOE **172** was carried out using NIS/Yb(OTf)$_3$, a chemospecific promoter for NPOEs [88,89]. Iteration of the sequence permitted the preparation of the α-1,5-linked arabinan segment of the complex lipoarabinomanan cell wall array of *Mycobacterium tuberculosis*, **175**.



Scheme 24
NPOEs in the synthesis of arabinofuranosyl donors

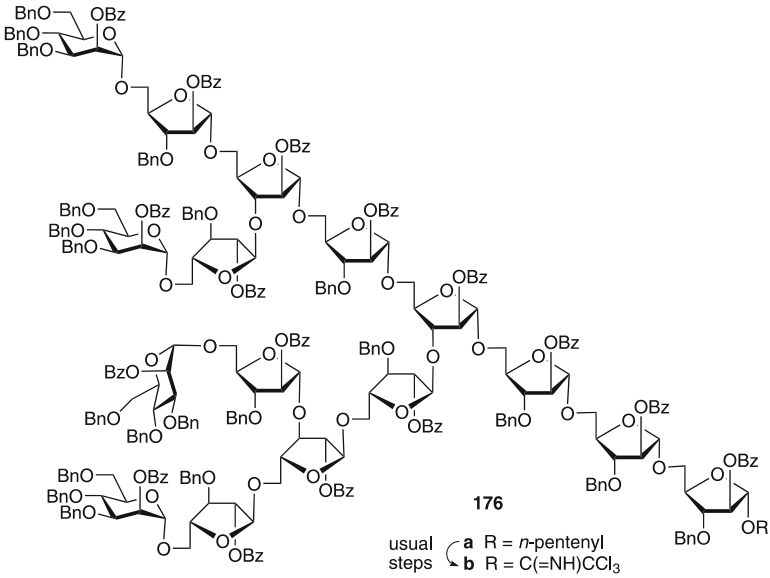


Figure 3
Mannose-capped multibranched dodecafuranoarabinan of *Mycobacterium* species

The potency of this strategy relies ultimately in the sturdiness, and yet the possibility for chemoselective cleavage, of pentenyl arabinofuranosides (e. g. **175**) [90]. Its value has been demonstrated recently with the synthesis of the pentenyl glycoside of mannose-capped dode-

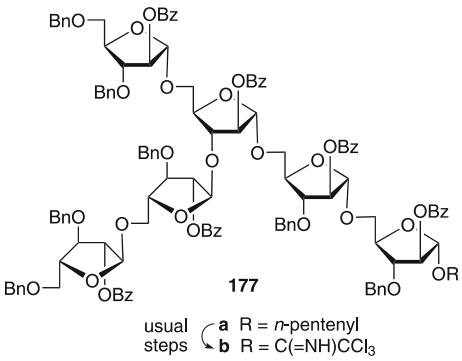
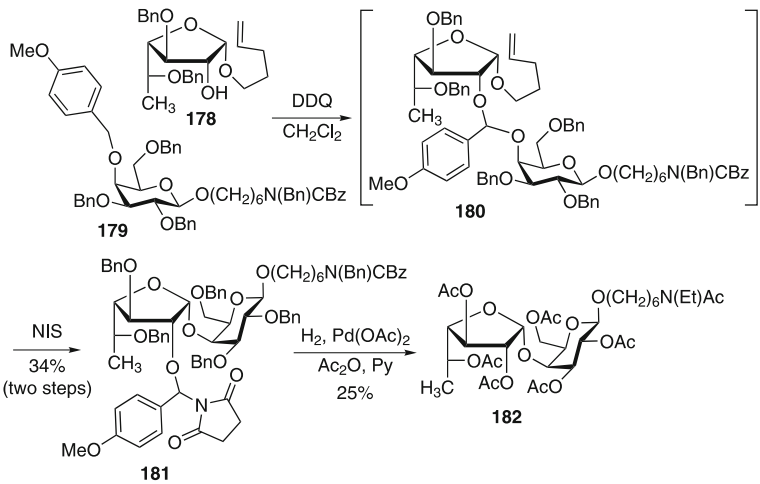


Figure 4
Multibranched hexafuranoarabinan of *Mycobacterium* species

cafuranoarabinan of *Mycobacterium* species, **176a** (*Fig. 3*). The final NPG→trichloroacetimidate transformation (**176a**→**176b**) made possible the coupling of this arabinan segment to an oligomannan acceptor, thus resulting in the synthesis of the largest heterooligosaccharide to date, a 28-mer arabinomannan [91].

n-Pentenyl arabinofuranosides have also been used by Seeberger and co-workers in the final stages of their synthesis of a 12-mer component of *Mycobacterium tuberculosis* [92]. *n*-Pentenyl arabinan hexasaccharide **177a**, was transformed to the corresponding trichloroacetimidate **177b** and coupled with a mannan hexasaccharide acceptor to yield the sought arabinomannan dodecasaccharide (*Fig. 4*).



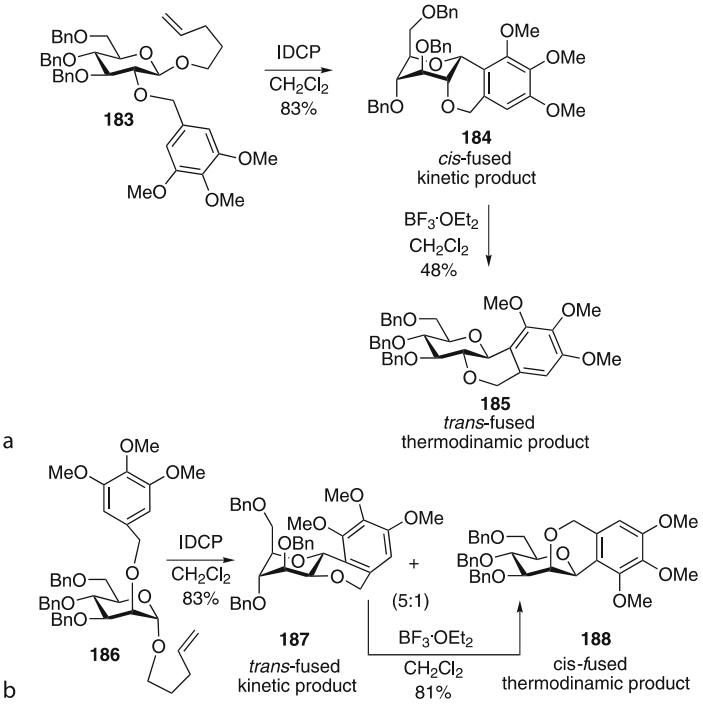
Scheme 25
NPGs as glycosyl donors in intramolecular aglycon delivery

2.8.5 Intramolecular Aglycon Delivery from *n*-Pentenyl Glycofuranosides

An approach to α-D-fucofuranosyl glycosides that makes use of the intramolecular aglycon delivery (IMAD) [93,94,95,96] starting from an *n*-pentenyl fucofuranoside has been described (*Scheme 25*) [97]. *n*-Pentenyl fucofuranoside **178** bearing a free 2-OH group was attached to a 4-*O*-PMB-protected galactopyranoside **179**, upon treatment of the mixture with DDQ. The unstable tethered compound **180**, could be activated with NIS, in the absence of even catalytic amounts of acid, to undergo an efficient *p*-methoxybenzyl-assisted aglycon delivery [95] leading to the desired glycoside **181**. The unusual structure **181**, resulting from quenching of the benzylic cation with *N*-succinimide, was then processed to α-D-fucofuranoside **182**.

2.8.6 Intramolecular *C*-Glycosylation of NPGs

The intramolecular *C*-glycosylation of NPGs has been studied by Martin's group in the course of their approaches to bergenin [98] and related natural products [99,100]. The treatment of pentenyl β-D-glucopyranose **183** with IDCP promoted an internal, Friedel–Crafts type, *C*-arylation reaction in excellent yield (*Scheme 26a*). The resulting product was exclusively the kinetically favored, *cis*-fused tricyclic system **184**. Treatment of the latter with an oxophilic Lewis acid ($BF_3{\cdot}OEt_2$) led to the *trans*-fused (β-linked) **185**. On the contrary, analogous reac-



Scheme 26
Intramolecular *C*-glycosylation of NPGs

tion of α-D-mannopyranoside **186** led to a mixture of *trans*- and *cis*-fused compounds **187**, and **188** (5:1), where the major *trans*-fused (α-linked) product **187**, was this time the kinetic product (*Scheme 26b*). Treatment of the latter with $BF_3{\cdot}OEt_2$ promoted the epimerization to the, more stable, 1,2-*cis* epimer **188**, in 81% yield.

2.8.7 NPGs of *N*-Acetylneuraminic Acids (Neu5Ac)

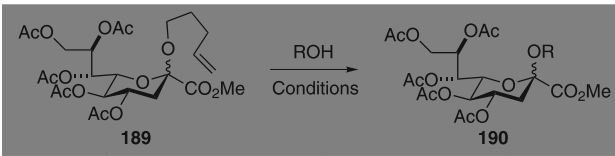
One report describing *O*-sialylation of 4-pentenyl glycosides of Neu5OAc, e. g. **189**, has appeared (*Table 6*) [101]. Good α/β selectivity (11:1) was attained in the glycosylation of primary acceptor **40** in MeCN using NIS/TfOH as the promoter (*Table 6*, entry i), however, with the secondary acceptor **191** the α/β selectivity dropped to 4:1 (entry iii). The use of Et_2O as solvent produced a 1:1 mixture of anomers (entry ii).

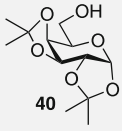
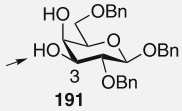
2.8.8 NPGs of L-Iduronic Acid as Glycosyl Donors

In their studies on heparin/heparin sulfate, and dermatan sulfate, Petitou, Sinaÿ and co-workers found that *n*-pentenyl glycosides of L-Iduronic acid, e. g. **192**, were efficient glycosyl donors [102]. In contrast, the corresponding thioglycosides, and glycosyl fluorides did not give the expected disaccharides. Reaction of *n*-pentenyl glycosyl donors **192** (α or β) with acceptors **194**–**197** (*Table 7*) was carried out in CH_2Cl_2 with NIS/TfOH to furnish the corresponding α-disaccharides **193**, in good yields.

More recently, Reichardt and Martín-Lomas have evaluated *n*-pentenyl glycosides of glucosamine $\alpha 1 \rightarrow 4$L-iduronic acid disaccharide, as substrates for autocondensation in their approach to heparin oligosaccharide fragments. However, the NIS used as promoter, being

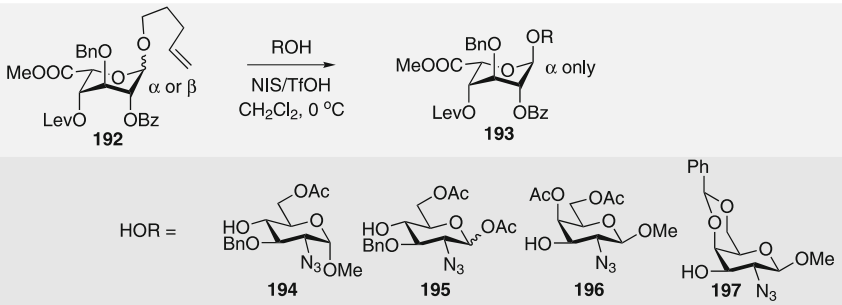
Table 6
***n*-Pentenyl glycosides of Neu5Ac in glycosylation**



Entry	ROH	Conditions	Yield (%) (α:β)
i	 40	NIS/TfOH MeCN, −40 °C	60 (11:1)
ii	40	NIS/TfOH Et_2O, −40 °C	33 (1:1)
iii	 191	NIS/TfOH MeCN, −40 °C	37[a] (4:1)

[a] The glycosylation is regioselective at *O*-3

Table 7
***n*-Pentenyl glycosides of L-iduronic acids as glycosyl donors**



Yields of disaccharide	80% from α-192 83% from β-192	75% from α-192 77% from β-192	72% from α-192	90% from α-192

itself a nucleophile, competes with the acceptor disaccharide in the polycondensation process, which results in fast chain-reaction termination and a low yield and degree of polymerization [103].

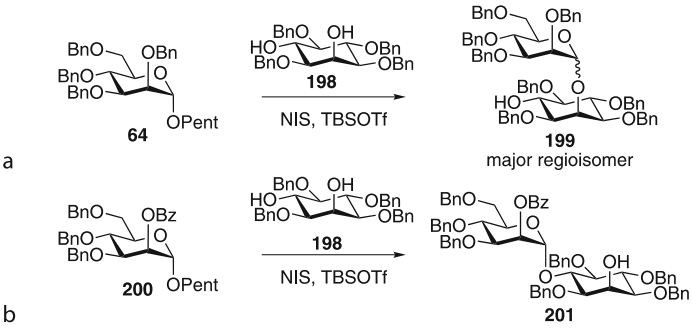
2.9 NPGs in Regioselective Couplings

The synthesis of branched saccharides by multiple glycosylations onto a central monosaccharide normally requires the use of orthogonal protecting groups in the acceptor. In this context, regioselective glycosylation of diols or polyols would ease the number of protection-deprotection steps in these synthetic protocols.

2.9.1 The Role of the *O*-2 Substituent in Regioselective Couplings

In their studies on *myo*-inositol glycosylation, Fraser-Reid and co-workers made the observations summarized in ❯ *Scheme 27* [104]. In the hope of achieving selective glycosidation of the equatorial-OH, they treated diol **198** with the armed *n*-pentenyl donor **64** (❯ *Scheme 27a*). However, the major product was the mixture of α/β glycosides **199** from glycosidation at the axial-OH (❯ *Scheme 27a*). In order to improve α anomeric stereoselectivity they selected the corresponding disarmed NPG **200**, as the donor (❯ *Scheme 27b*). Surprisingly, the only product obtained was the disaccharide **201** from glycosidation at the equatorial-OH.

In a series of subsequent papers Fraser-Reid and co-workers confirmed these discrepancies, and showed that the *O*-2 substituent in glycosyl donors, besides its recognized role for stereocontrol, exerts a profound influence in eliciting regioselective glycosyl couplings [105,106,107]. In most cases, 2-*O*-acyl NPGs and NPOEs shared the same regiopreferences, which were usually different from the ones displayed by 2-*O*-alkyl NPGs. The regiopreferences of the former were generally more pronounced or even exclusive.

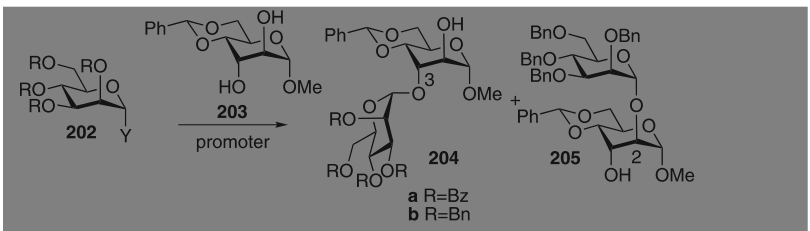


Scheme 27
Influence of the *O*-2 protecting group in regioselective glycosylations

2.9.2 Reciprocal Donor Acceptor Selectivity (RDAS)

The influence of the *O*-2 substituent in regioselective couplings is not limited to pentenyl glycoside donors. Thioglycoside and trichloroacetimidate donors have shown the same tendency [108]. The glycosylation of allose diol **203** with donors **64** and **202a–f** (*Table 8*) illustrates this point. NPOE **202a** (that shows the same regiopreferences as disarmed NPGS), disarmed thiomannoside **202c**, and disarmed trichloroacetimidate **202e**, exhibited the same preference for the *O*3 of allose acceptor **203** (*Table 8*, entries i, iii, v). On the contrary, armed donors **64**, **202c**, and **202e** furnished a 2:1 mixture of disaccharides **204b** and **205**. The above-mentioned examples indicate that each donor expresses preference for one of the diol–OHs in the acceptor and vice versa. The authors coined the term *Reciprocal Donor Acceptor Selectivity* (RDAS) [109] to account for these findings.

Table 8
Influence of the *O*-2 substituent in the regioselective coupling of various glycosyl donors with allose diol 203



Entry	Donor (202)	Promoter (Temp °C)	Products (ratio 03:02)	Yield %
i	a Y=orthoester; R=Bz	NIS/BF_3·Et_2O −30	204a only	92
ii	64 Y=OPent; R=Bn	NIS/BF_3·Et_2O −30	204b + 205 (2:1)	37
iii	b Y=SPh; R=Bz	NIS/BF_3·Et_2O −30	204a only	58
iv	c Y=SPh; R=Bn	NIS/BF_3·Et_2O −30	204b + 205 (2:1)	66
v	d Y=OC(NH)CCl_3; R=Bz	BF_3·Et_2O −78	204a only	65
vi	e Y=OC(NH)CCl_3; R=Bn	BF_3·Et_2O −78	204b + 205 (2:1)	54

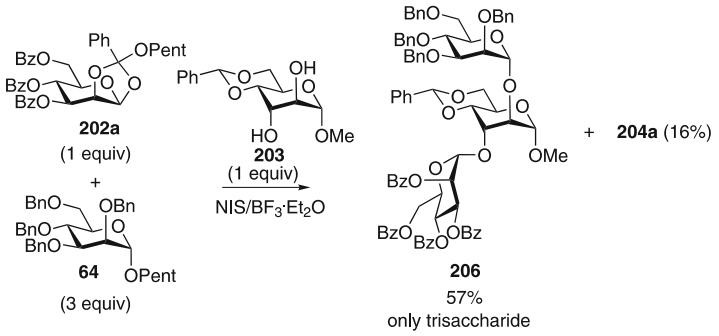
2.9.3 In Situ Double Differential Glycosylations of Two Donors with One Acceptor

The practical utility of this concept was further demonstrated when diol acceptor **203**, NPOE **202a**, and armed-NPG **64** were treated with NIS/BF_3·Et_2O to give one single trisaccharide **206**, in 57% yield (● *Scheme 28*) [110]. It seemed that in the formation of the trisaccharide the regiopreferences of the NPOE **202a** and the armed NPG **64**, displayed in ● *Table 8*, have been followed.

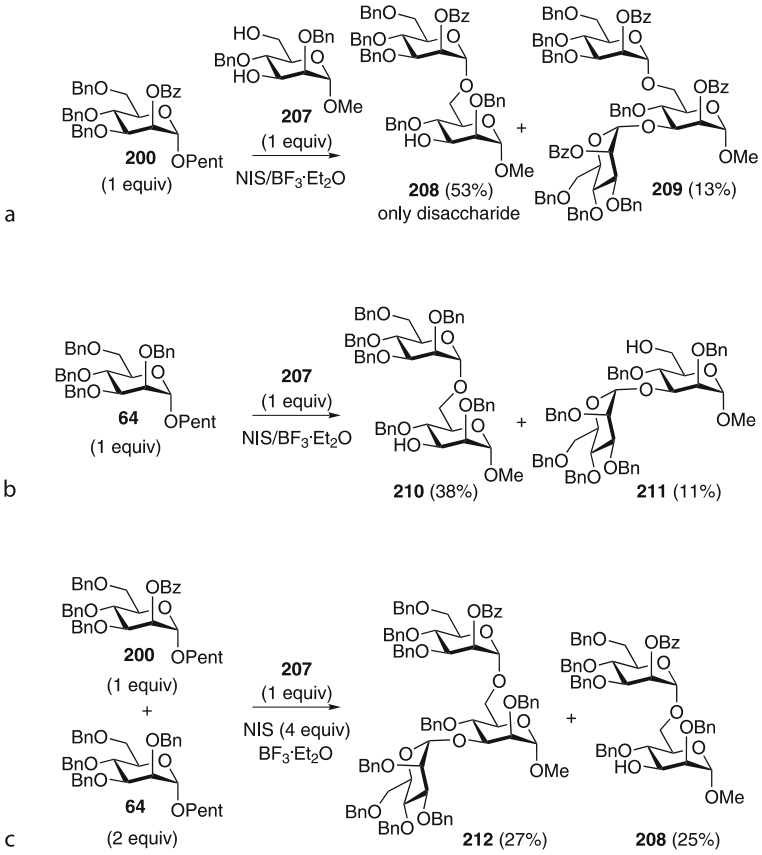
Analogously, the regiopreferences (RDAS) of disarmed NPG **200**, and armed NPG **64** vis a vis mannose diol **207**, were evaluated (● *Scheme 29a,b*). With the disarmed donor **200** mannosylation occurred at the (C6)-OH only to give **208** in 53% yield, and also the symmetrical trisaccharide **209** in 13% yield, but with no evidence for the dimannan resulting from glycosylation of the (C3)-OH (● *Scheme 29a*). By contrast, the armed donor **64** gave a 38% yield of the *O*-6 product, **210**, but also 11% of the *O*3 regioisomer **211** (● *Scheme 29b*). Analysis of these results according to conventional wisdom, dictates that the preference of both donors, **200** and **64**, for the primary −OH was to be expected [1] but raised the question of the possible outcome of a three-components double glycosylation when **200** and **64** compete for diol **207**. Previous calculations had shown that the relative reactivity of these donors (k_{64}/k_{200}) is 3.2 [111]. Hence, it was expected that *O*6 mannosylation by the armed donor, **64**, would predominate in any trimannan produced. Surprisingly, a single trimannan **212**, in which the *less* reactive donor **200** ended up at *O*6 was obtained, even in the presence of 2 equiv. of the "*more reactive*" **64**.

2.10 The Origin of Regioselectivity in Three-Component Couplings

In searching for the origin of the regioselectivity observed in the formation of trisaccharides **206** and **212** (● *Scheme 28* and ● *Scheme 29*) several factors were considered. The reactions in ● *Scheme 28* and ● *Scheme 29c* were carried out with excess NIS promoter, conditions under which the intermolecular halonium ion transfer (responsible for the armed-disarmed effect) is not operative. A study of the three types of *n*-pentenyl donors indicated that their relative reactivities were in the order NPOE > armed > disarmed (e. g. **202a** > **64** > **200**) [111]. Therefore, the most and the least reactive donors have "chosen" their preferred −OH in the final trisac-



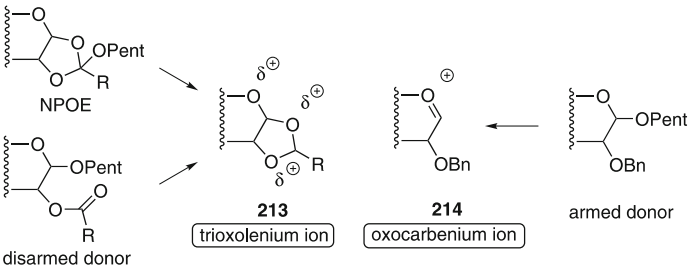
Scheme 28
In situ three-component double differential glycosylation of two donors and one diol acceptor



Scheme 29
In situ three-component double differential glycosylation of two donors and one diol acceptor

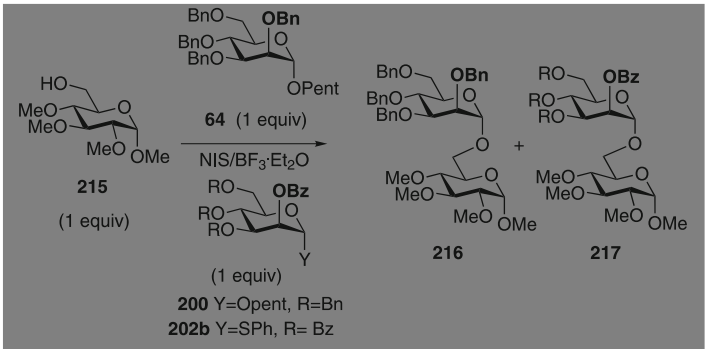
charide. On the other hand, the most and the least reactive donors give rise to the highly delocalized, more stable intermediate **213**, while the armed donor gives the less stable oxocarbenium ion **214** (*Scheme 30*) [112]. The conclusion was that in *competitive* glycosylations the more *stable* donor/intermediate (not the most *reactive* donor) controls regioselectivity, resulting in the formation of the single trisaccharides **206** and **212** and the single disaccharides **204a** and **208**.

In order to confirm this assumption the authors performed the experiments in *Table 9* [113]. Equimolar amounts of armed and disarmed donors **64**, and **200** or **202b** were allowed to compete for one equivalent of acceptor **215** under the agency of NIS. When one equivalent of NIS was used the major product obtained was that of glycosylation of armed NPG **216**, thus in agreement with a process of intermolecular halonium transfer and preferred reaction of the more reactive donor (*Table 9*, entries i, iii). When the amount of NIS was increased to three equivalents, the observed ratio of compounds **216** and **217** indicated enhanced coupling of the disarmed donor (*Table 9*, entries ii, iv), thus in agreement with the proposed rationalization for the regiopreferences observed in the three-component reactions.



Scheme 30
Reactive intermediates from different glycosyl donors

Table 9
Competition studies on the glycosylation of acceptor 215 in the presence of glycosyl donors (64 and 200 or 202b) with variable amounts of NIS



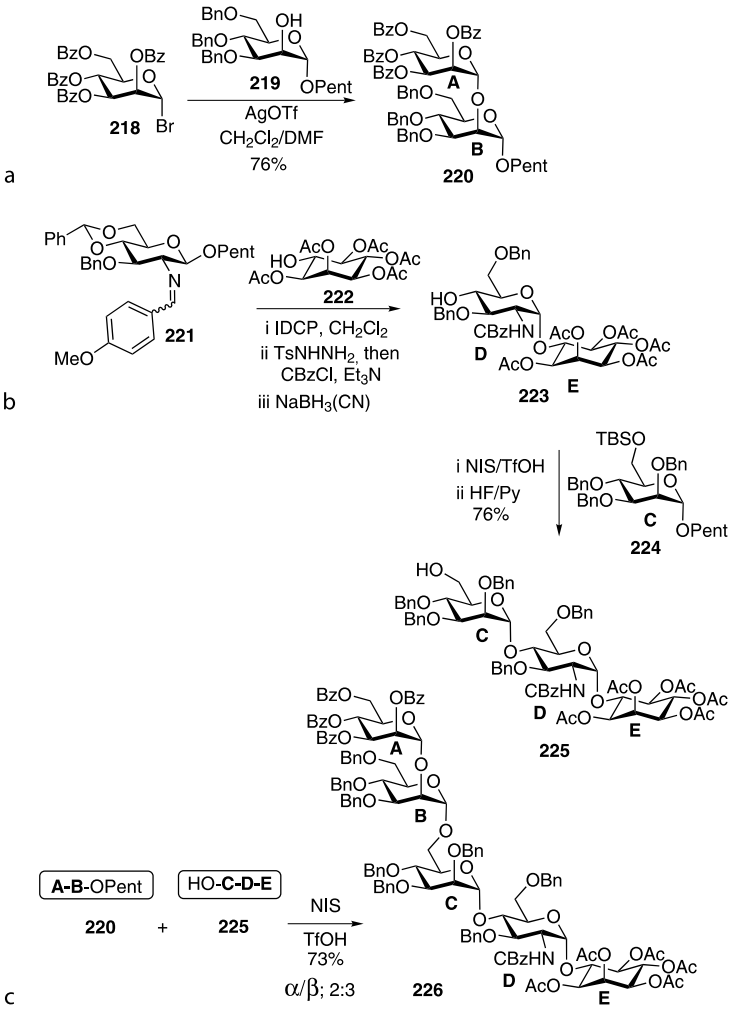
Entry	Donors		NIS	Product ratio		Yield
	armed/disarmed		(equiv)	216	217	(%)
i	64	200	1	2	1.1	57
ii	64	200	3	1.1	1	69
iii	64	202b	1	1	0	52
iv	64	202b	3	1	2.2	83

2.11 NPGs in Oligosaccharide Synthesis

Since their discovery, the unique properties of NPGs have allowed the preparation of several oligosaccharides. The pentenyl moiety may be installed early in the synthetic sequence and can survive many types of protecting group manipulations. Some selected syntheses of oligosaccharides are briefly discussed below.

2.11.1 The Pentasaccharide Core of the Protein Membrane Anchor Found in *Trypanosoma brucei*

Fraser-Reid and co-workers described a block (i. e., convergent), and a linear approach to the title compound, **226** (*Scheme 31*) [114]. The convergent approach, outlined in *Scheme 31*,



Scheme 31
The pentasaccharide core of the protein membrane anchor of *Trypanosoma brucei*

makes use of the stereocontrolled glycosylation of inositol derivative **222** with 2-deoxy-2-imino NPG **221** (*Scheme 31b*). Protecting group manipulations led to acceptor **223** that was glycosylated with NPG **224**, to furnish, after desilylation, the acceptor **CDE** block, **225**. The donor counterpart **220**, had been readily prepared by Koenigs–Knorr [115] coupling of NPG **219** with glycosyl bromide **218** (*Scheme 31a*). Finally, coupling of fragments AB (**220**) and CDE (**225**) promoted by NIS/TfOH led to pentasaccharide **226**, in 73% as an α/β (2:3) mixture (*Scheme 31c*).

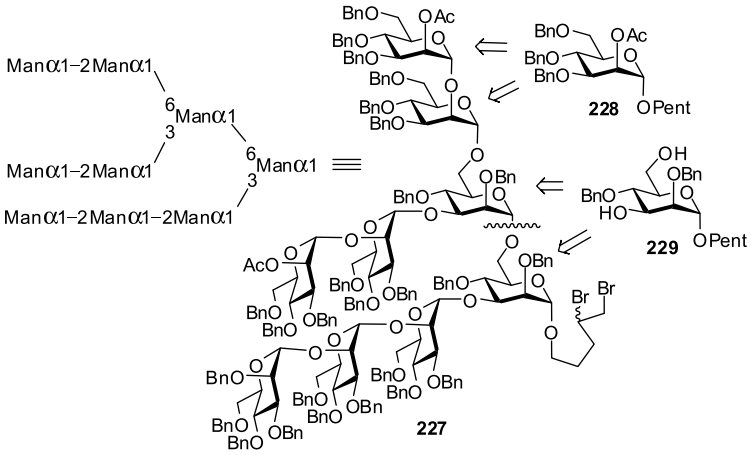
2.11.2 The Nonamannan Component of High Mannose Glycoproteins

The concise approach to nonamannan **227** (› *Scheme 32*), was greatly simplified with the sidetracking of NPGs [116,117] that allows the same NPG synthon to function as glycosyl donor or as glycosyl acceptor.

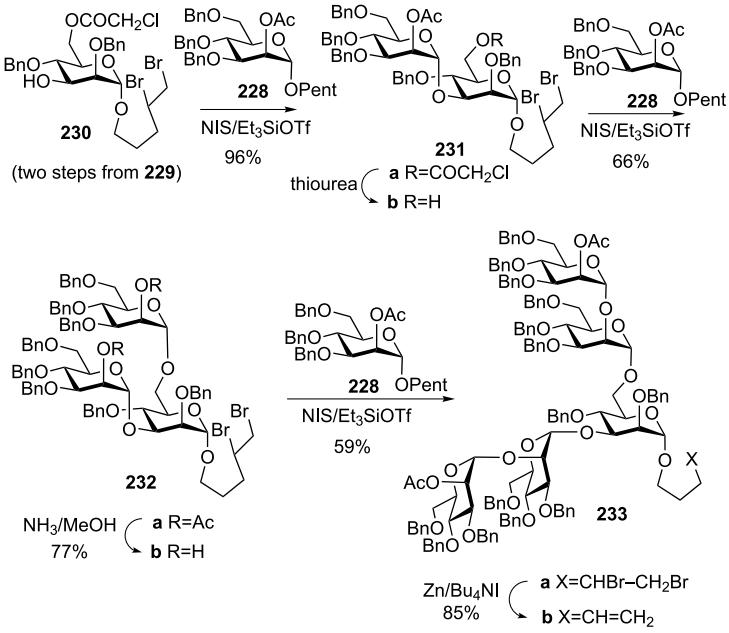
In the retrosynthesis of **227**, the authors identified three types of elements depending on the number of sugar units attached to them. Two components carried sugars at *O*3 and *O*6, four held substituents at *O*2, and the last three had no monosaccharides attached. According to that, the nonasaccharide target could be correlated with only two mannopyranose precursors **228** and **229**, since synthon **228** could be used to access the last two kinds of sugars. The approach featured the final link of a pentasaccharide donor with a tetrasaccharide acceptor, as outlined in › *Scheme 32*.

The synthesis of pentasaccharide donor **233** started with the mannosylation of sidetracked NPG **230** with disarmed NPG donor **228** (› *Scheme 33*). The ensuing, stereoselectively formed disaccharide **231a**, after unveiling of its C6-OH, underwent a second mannosylation with **228**. Removal of the acetates in **232a** with NH_3/MeOH led to diol trisaccharide **232b**, which was bis-mannosylated with **228** to give pentasaccharide **233a** in 59% yield. Regeneration of the pentenyl moiety in sidetracked **233a**, with Zn/nBu_4NI, granted access to pentasaccharide NPG donor, **233b**.

The lowest antenna of **227** was built from **234** (› *Scheme 34*). By taking advantage of the sidetracking concept, compound **228** could be used as a glycosyl donor or, after dibromination and deacetylation as the glycosyl acceptor **234**, thereby facilitating the rapid assembly of trisaccharide fragment **235**. Thus, coupling of **234** and **228** afforded the expected disaccharide in 73% yield, deacetylation and additional coupling with **228** led to trisaccharide **235a** in 62% yield. The latter was transformed in glycosyl donor **235b** by reductive elimination, and coupled with **230** to give sidetracked tetrasaccharide **236a**. Dechloroacetylation of the latter led to **236b** that was glycosylated by pentasaccharide donor **233b** to give nonasaccharide **227** in 57% yield.



Scheme 32
The nonamannan component of high mannose glycoproteins

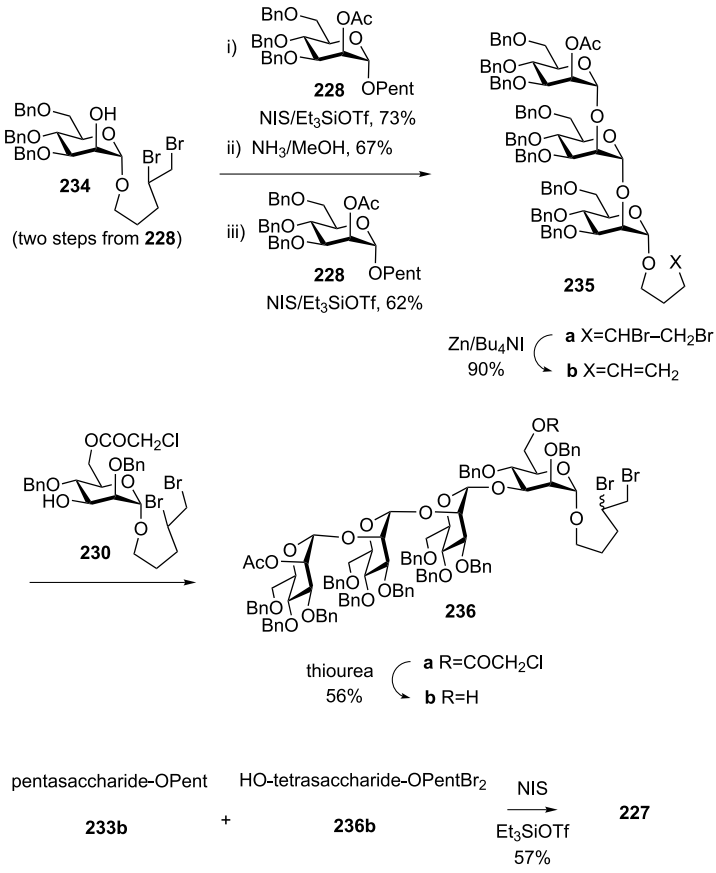


Scheme 33
Convergent synthesis of nonamannan 227. Synthesis of the pentasaccharide donor 233

2.11.3 Synthesis of NodRf-III (C18:1) (MeFuc)

Nodulation factors comprise a family of unique oligosaccharides composed substantially of glucosamine (2-amino-2-deoxy-D-glucose) units that are *N*-acylated with acetic acid and fatty acids residues, the latter residing at the nonreducing terminus [118]. The block synthesis of NodRf-III (C18:1) (MeFuc) **237**, a nod factor produced by *Rhizobium fredii*, is an illustrative example of the chemistry developed around 2-amino-2-deoxy-NPGs (*Scheme 35*) [119]. The key elements in this stereocontrolled synthesis are: (a) the use of the TCP protecting group, which provides a facile method for *N*-differentiation in the glucosamine oligomer, (b) the assistance of the sidetracking methodology, (c) a solvent-assisted stereoselective α-fucosylation, (d) a β-selective, neighboring group assisted, glycosidation, and (e) the use of $FeCl_3$ for late-stage debenzylation of the oligosaccharide moiety [120,121]. In the retrosynthesis (*Scheme 35*), the authors selected a TCP as protecting group for the nitrogen atom that would bear the unique fatty acid, while the repeating unit would be a 2-deoxy-2-*N*-phthaloyl NPG capable of acting as a glycosyl donor (e. g. **238**). The reducing end retron was identified with benzyl glycoside **239**.

The disaccharide acceptor **239** was prepared in 85% yield by coupling (NIS/TESOTf) of acceptor **241** with *n*-pentenyl fucoside **240** in Et_2O:CH_2Cl_2, (5:1) (*Scheme 36*). The disaccharide donor **238** was assembled in 71% by coupling of NPG **242** with sidetracked acceptor **243** (NIS/TESOTf), followed by reinstating of the pent-4-enyl moiety from the dibromo pentenyl residue in **244**. Final coupling (NIS/TESOTf) of donor **238** with acceptor **239** yielded

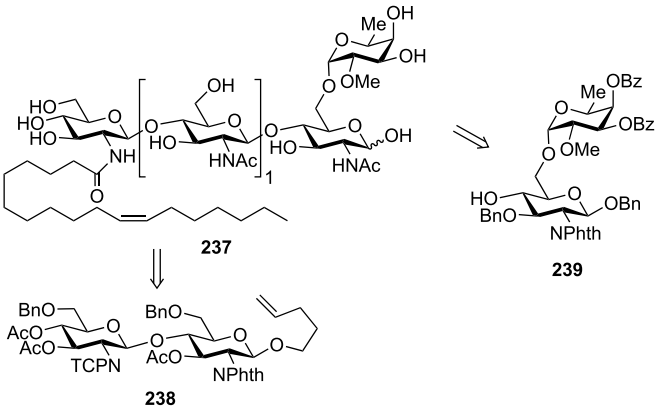


Scheme 34
Convergent synthesis of nonamannan 227. Synthesis of the tetrasaccharide acceptor 236b and final assembly

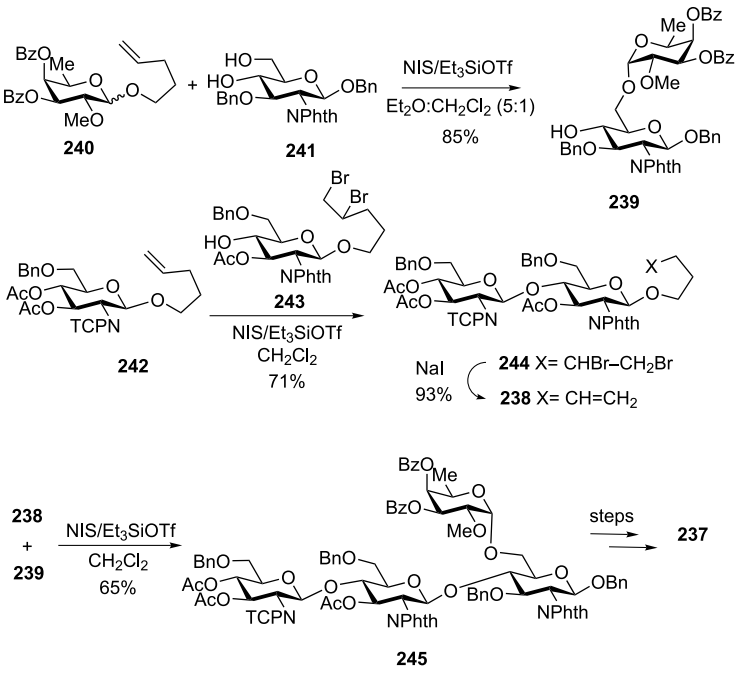
tetrasaccharide **245** in 65% yield. The final stages in the preparation of **237** involved: (i) $FeCl_3$ debenzylation, (ii) silylation of the resulting free –OHs, (iii) deprotection of the TCP and condensation with an activated fatty acid, (iv) removal of the phthalimido protecting groups, and (v) acylation, saponification, and desilylation.

2.11.4 Synthesis of Phosphorylated Rat Brain Thy-1 Glycosylphosphatidylinositol Anchor

Glycosylphosphatidylinositol (GPI) membrane anchors constitute a class of glycolipids that covalently link certain proteins to cell and virion surfaces [122,123]. A boost in their chemistry occurred in 1988 when Ferguson et al. reported the first covalent structure of a member of this family [124,125]. The first synthesis of a fully phosphorylated GPI, compound **246** (*Fig. 5*), was accomplished by Fraser-Reid's group based entirely on NPG chemistry [126,127,128,129,130].



Scheme 35
Retrosynthesis of nodulation factor NodRf-III (C18:1) (MeFuc)



Scheme 36
Synthesis of nodulation factor NodRf-III (C18:1) (MeFuc)

The retrosynthetic analysis dictated a heptasaccharide **251**, with all free hydroxyls in the final product benzylated, and the three sites of phosphorylation all differentially protected, so that all three can be manipulated separately for maximal flexibility. The free amine of glucosamine is protected as an azide, which can be taken through multiple transformations and will only be unmasked at the end of the synthesis (*Scheme 37*). The heptasaccharide is in turn put

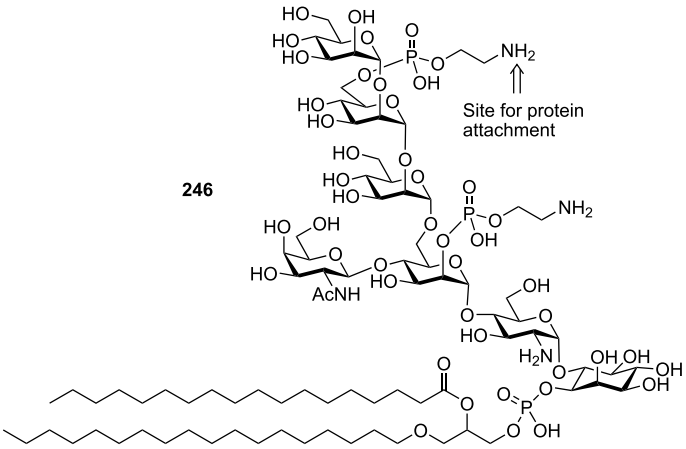
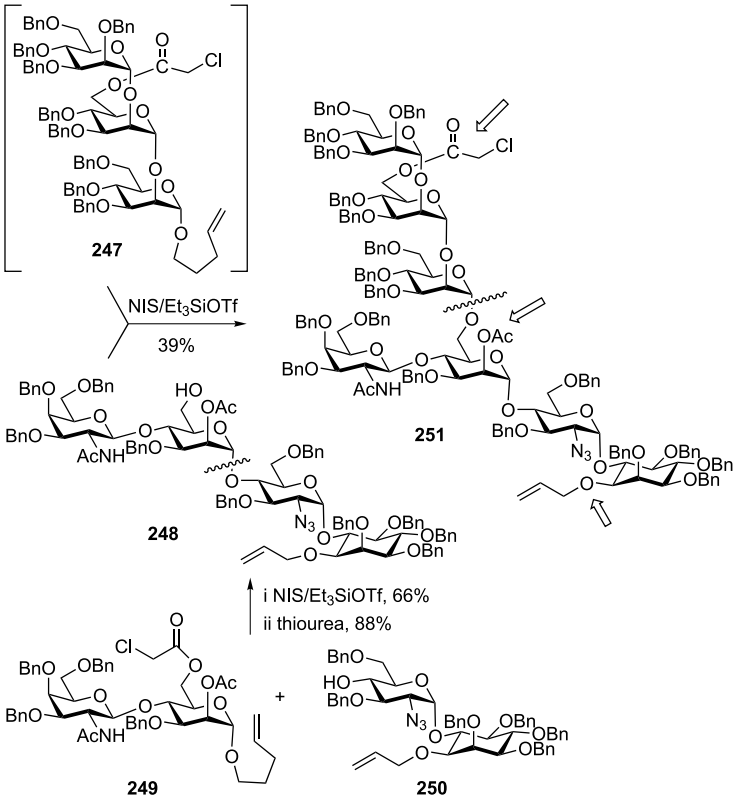
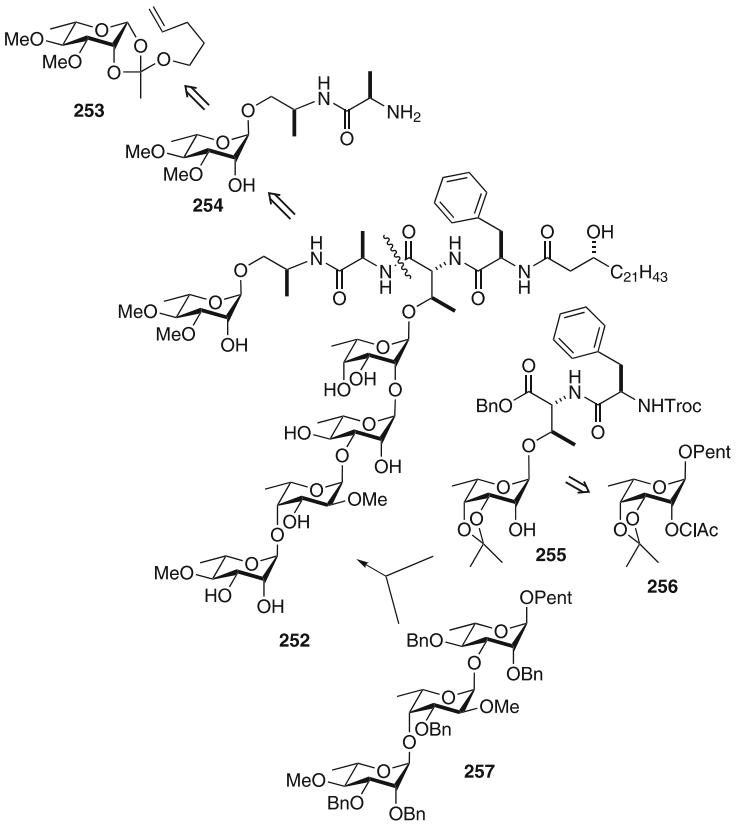


Figure 5
Rat brain Thy-1 GPI anchor 1, 246



Scheme 37
Synthesis of rat brain Thy-1 GPI anchor 1



Scheme 38
Synthesis of the glycopeptidolipid of *Micobacterium avium* Serovar 4, 252

together in three portions, a galatosaminylmannose **249** being coupled to azidoglucosylinositol **250** and then a trimannose, **247**, being coupled to that moiety. Coupling of glycosyl donor **249** with the disaccharide acceptor **250** was carried out with NIS/TESOTf to give α-linked tetrasaccharide **248** in 66% yield. Notably, the allyl protecting group survived the treatment with NIS. The *O*6 of the mannose residue was deprotected by removal of the chloroacetate moiety with thiourea and glycosylated with pentenyl trimannoside **247** to give the fully protected heptasaccharide **251** in 39% yield. The three positions (marked with arrows) were then deprotected and phosphorylated according to the following sequence: dechloroacetylation with thiourea, saponification of the acetate with methoxide, and deallylation with $PdCl_2$. Complete debenzylation, then culminated the synthesis of **246**.

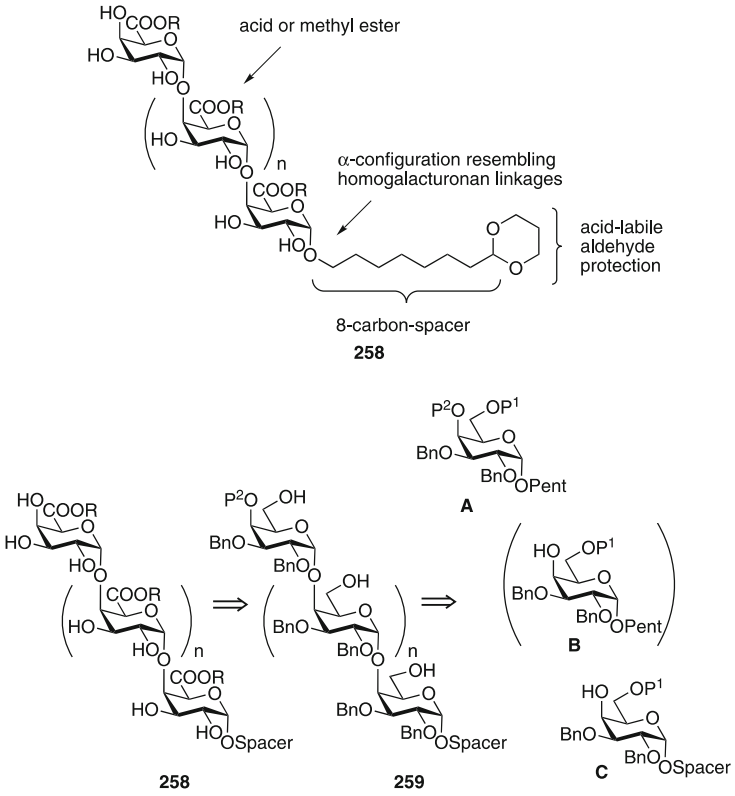
2.11.5 Synthesis of the Glycopeptidolipid of *Micobacterium avium* Serovar 4

Heidelberg and Martin described the first synthesis of the "polar mycoside C" **252** (*Scheme 38*) [131]. The synthesis was based on the disconnection of the final structure into three saccharidic building blocks, an L-rhamnosyl pseudodipeptide **254**, a 6-deoxy-L-talosyl

dipeptide **255**, and a pentenyl trisaccharide donor **257**. The key steps were the creation of the glycosidic linkage between the trisaccharide donor **257**, and the 6-deoxy-L-talose unit **255** (IDCP, 60% yield), and the final coupling of the two glycopeptide fragments. Other pentenyl mediated couplings were the glycosylation of orthoester **253** leading to **254** (NIS/TMSOTf, 81%) and the stereoselective α-coupling of disarmed NPG **256** to give glycodipeptide **255** (NIS/TMSOTf, 70%).

2.11.6 Synthesis of Oligogalacturonates Based on NPGs

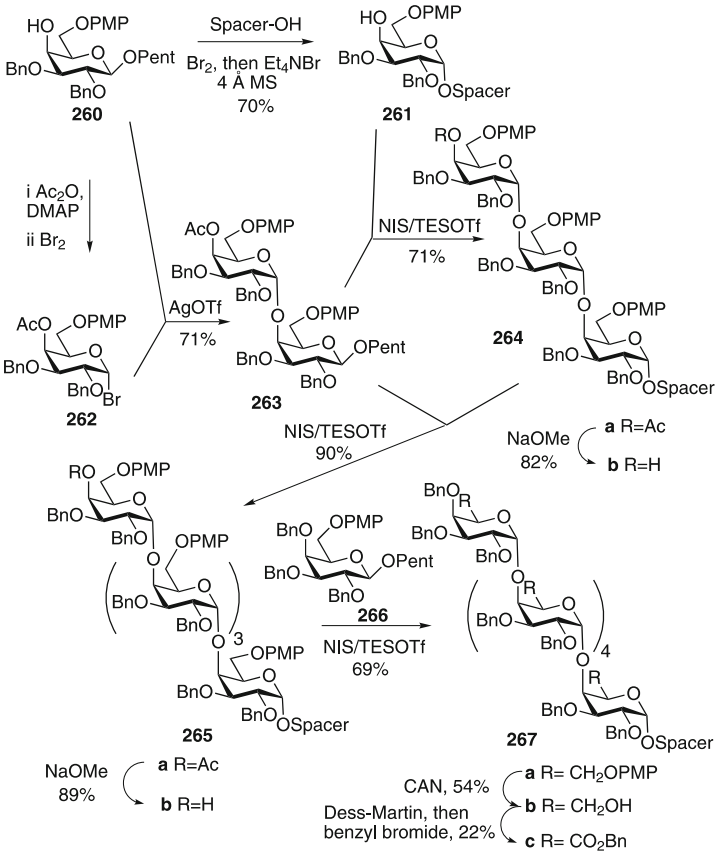
Madsen and co-workers described a concise approach to oligogalacturonates (e. g. **258**, ● *Scheme 39*) conjugated to bovine serum albumin (BSA) based on NPGs. They synthesized several oligogalacturonates, which were linked to the BSA by reductive amination via an aldehyde spacer at the reducing end (● *Scheme 39*) [48,132,133]. Their strategy called for two orthogonal protecting groups (P^1 and P^2), and three different monomeric building blocks: a spacer galactoside **C** to serve as glycosyl acceptor for the reducing end, and two glycosyl donors **A** and **B**, the former for the nonreducing end and the latter for the galacturonic



■ Scheme 39
Retrosynthesis of oligogalacturonates for conjugation to bovine serum albumin

repeating unit. *p*-Methoxyphenyl (PMP) and acetyl groups were used as protecting groups. The methodology was then based on the repeated coupling of galactose donors onto galactose acceptors followed by deprotection at *O*6, as in **259**, which permitted the oxidation of these primary positions to either the carboxylic acids or methyl esters.

The attaching of the spacer to galactose (i. e. **261**, building block **C**) was carried out by glycosylation, under Lemieux conditions [134], of a glycosyl bromide readily obtained from NPG, **260** (*Scheme 40*). Coupling of NPG **260** with galactosyl bromide **262** (AgOTf, 71%) led to pentenyl disaccharide **263** that glycosylated acceptor **261** (NIS/TESOTf, 71%) to give trisaccharide **264a**. Deprotection of the latter to **264b**, and glycosylation with glycosyl donor **263** (NIS/TESOTf, 90%) led to pentasaccharide **265a**. Glycosyl assembly to hexasaccharide **267**, included the glycosylation of pentenyl donor **266** with acceptor **265b** (NIS/TESOTf, 69%), obtained by NaOMe treatment of **265a**. The final stages of the synthesis included CAN-mediated deprotection of the *p*-methoxyphenyl groups, Dess–Martin oxidation and esterification.



Scheme 40
Synthesis of oligogalacturonates

2.11.7 Miscellaneous

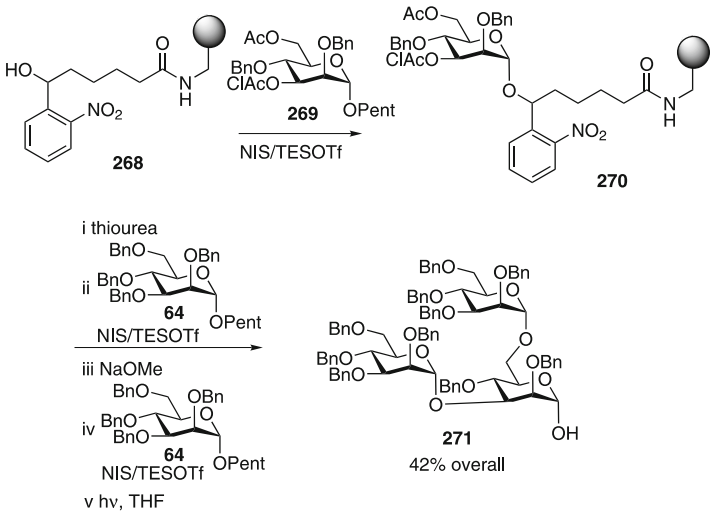
Arasappan and Fraser-Reid reported an NPG-based methodology for the stereoselective construction of the tetrasaccharyl cap portion of Leishmania Lipophosphoglycan [135].
Kuzuhara and co-workers have reported the use of a NPG disaccharidic synthon as the chain elongating unit in the synthesis of amphiphilic chitopentaose and chitoheptaose derivatives [136].
Toshida and co-workers have described the synthesis of a set of di- and tri-sulfated galabioses by using an *n*-pentenyl galactoside donor and IDCP as the catalyst [137].

2.12 NPGs in Solid-Phase Oligosaccharide Synthesis

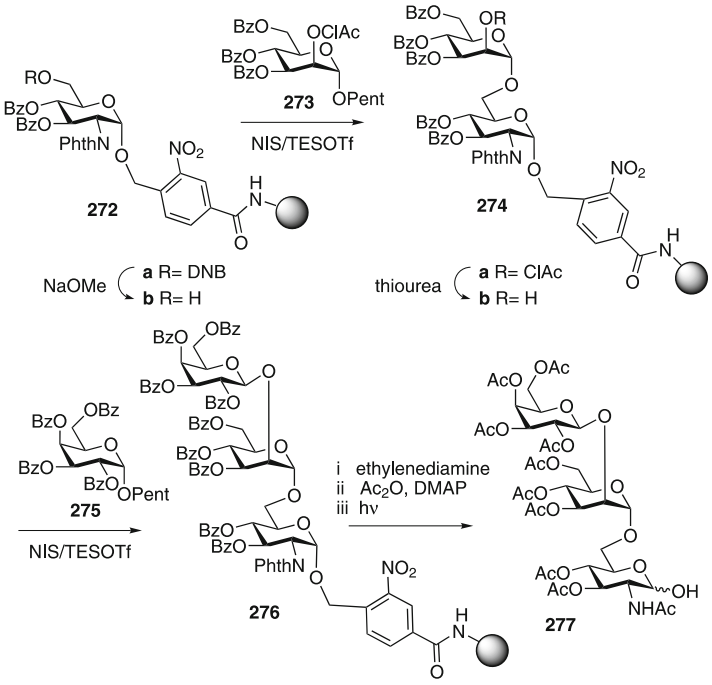
Solid-phase oligosaccharide synthesis has received considerable attention in the last years [138]. Some of the approaches that involve NPGs are discussed below.

2.12.1 Glycosylation of Supported Alcohol Acceptors with NPG Donors

Fraser-Reid and co-workers designed a photolabile *o*-nitrobenzylic linker **268**, which was used in the synthesis of a branched trimannan oligosaccharide **271** (*Scheme 41*) [139]. Differentially protected NPG **269** was coupled to the resin via linker **268**. Selective removal of the C6 chloroacetyl and C3 acetyl groups, followed each by mannosylation (NIS/TESOTf) with NPG **64**, afforded trimannan **271**, in 42% overall yield after photolytic cleavage.
In a related approach, Fraser-Reid and co-workers used Chiron's polystyrene-grafted "crowns" with Rich's photocleavable o-nitrobenzyl linker [140] and NPG donors in the synthesis of trisaccharide **277** (*Scheme 42*) [141]. After attachment of the first aminoglucosyl moiety to



Scheme 41
NPGs in solid-phase oligosaccharide synthesis

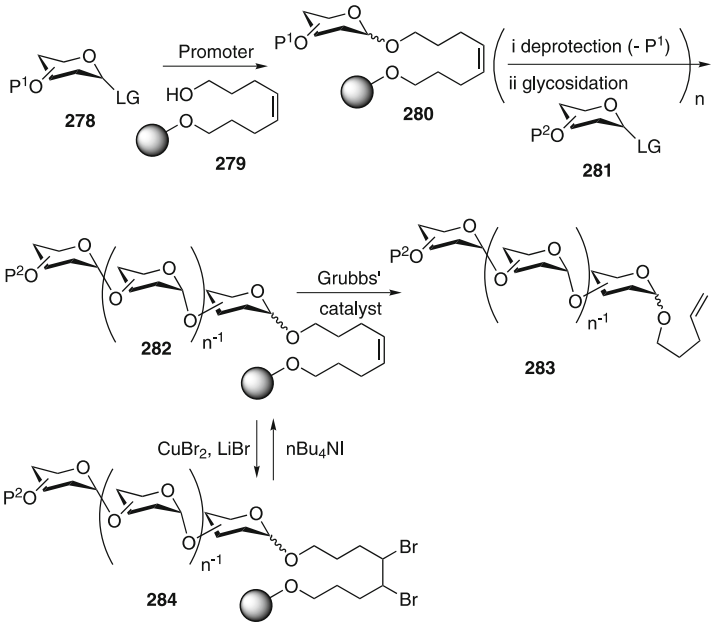


Scheme 42
NPGs in solid-phase oligosaccharide synthesis

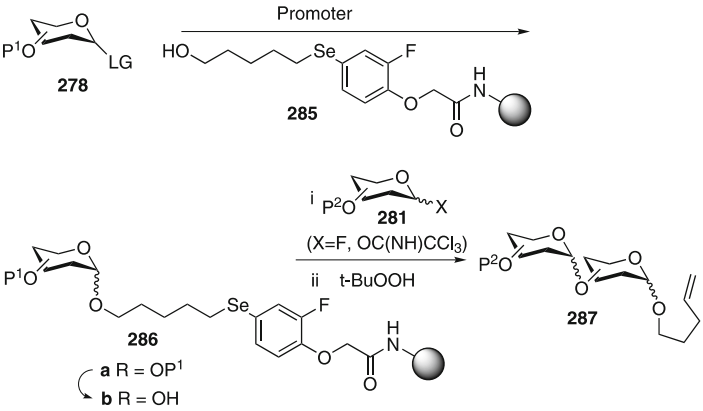
the linker, via its corresponding NPG, the C6 dinitrobenzoyl (DNB) group was removed to give **272b**. Coupling with mannose donor **273**, deprotection of the *O*2 chloroacetyl group, and galactosylation with NPG **275** furnished trisaccharide **276**. Global deprotection followed by peracetylation and photolytic cleavage from the solid support provided trisaccharide **277**.

2.12.2 Pentenyl Glycoside-Based Linkers

Seeberger and co-workers developed a new linker concept in solid-phase oligosaccharide synthesis. They designed a new NPG-based linker that upon deprotection rendered an oligosaccharide NPG suitable for further glycosylations in fragment couplings (*Scheme 43*) [142]. The first carbohydrate moiety (e. g. **278**) was connected via a glycosidic bond to octenediol-functionalized Merrifield's resin, **279**. Resins with loadings of up to 0.65 mmol/g were obtained and employed in oligosaccharide synthesis. Glycosylation events can now take place in deprotected saccharide **280** to yield oligosaccharide **282**. The octenediol linker was then cleaved by olefin cross metathesis using Grubbs' catalyst under an atmosphere of ethylene to afford fully protected oligosaccharides in the form of NPGs, e. g. **283**. A further refinement of this strategy is that it can be made compatible with glycosyl donors that require electrophiles as activators by sidetracking the linker to the corresponding dibromooctane derivative (e. g. **284**) [143]. Seeberger and co-workers have illustrated the potency of this strategy with several oligosaccharide syntheses [144,145,146,147].



Scheme 43
Seeberger's NPG-based linkers for oligosaccharide synthesis



Scheme 44
A fluorinated selenide linker for solid-phase synthesis of NPGs

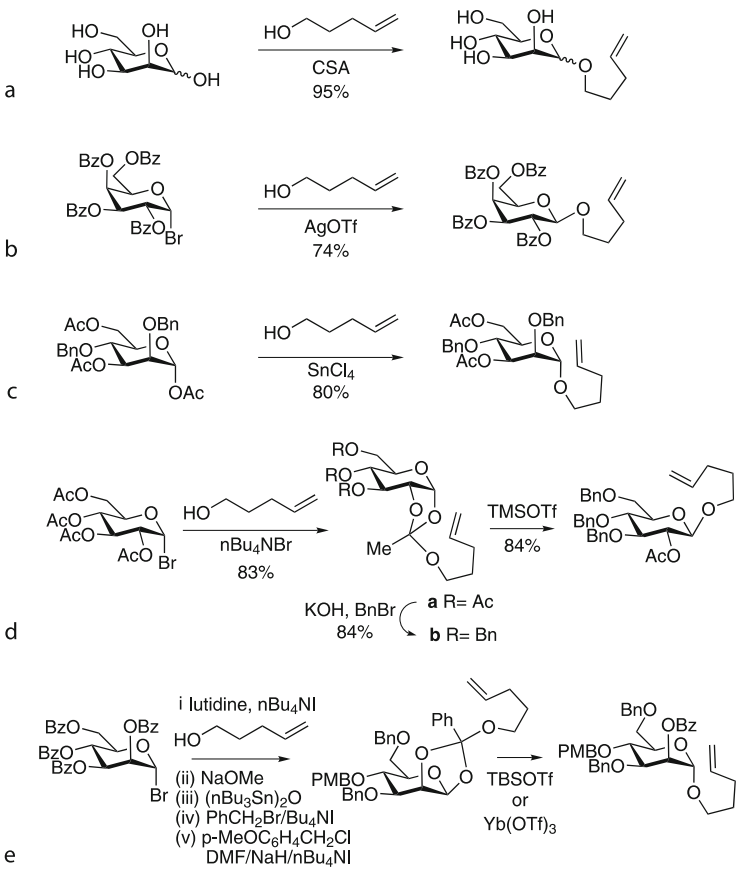
Mogemark et al. described a fluorinated selenide linker **285**, for solid-phase synthesis of NPGs (*Scheme 44*) [148]. The resin-bound linker could be glycosylated both with trichloroacetimidates and glycosyl fluorides to give anchored saccharides, e. g. **286**, that can be submitted to glycosylation once deprotected. After oxidation to a selenoxide with t-BuOOH the linker undergoes β-elimination upon heating, and releases the NPG **287**, in excellent yield.

2.13 Miscellaneous Uses of NPGs

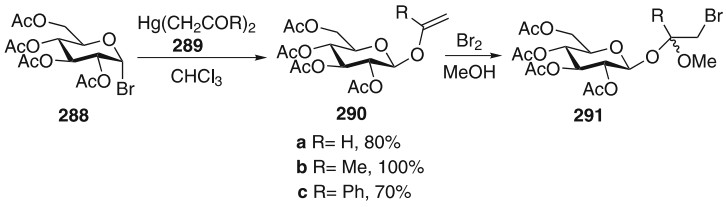
The versatility of NPGs has been further enhanced by chemical modifications of the pent-4-enyl moiety itself. In this context, the pentenyl moiety has been transformed in many spacer functionalities [149,150], used as a handle to incorporate amino-acid moieties [151,152,153,154], used as a monomer in copolymerization strategies [155,156,157], used in the formation of dendrimers [158], and converted to dimeric and trimeric structures for multivalent presentations [159,160]. These applications fall beyond the scope of this chapter.

2.14 Preparation of NPGs

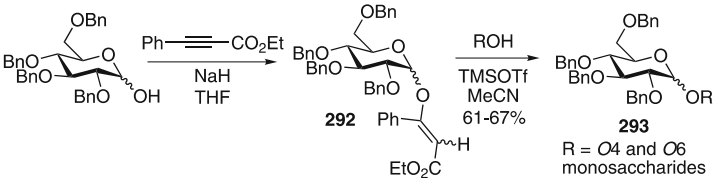
NPGs, being normal *O*-glycosides, can be readily obtained by application of the standard procedures for preparing such derivatives [12]. They can be obtained by Fischer glycosida-



Scheme 45
Methods for the preparation of NPGs



Scheme 46
Synthesis of 1′-substituted vinyl glucosides



Scheme 47
Synthesis of 1′-substituted vinyl glycosides

tion (*Scheme 45a*) [82,161,162]. An obvious advantage of this procedure is that the *n*-pentenyl group can be installed right at the outset of the synthesis; however, the formation of α/β anomers might sometimes be a drawback. Use of the Koenigs–Knorr coupling [115] permits the stereocontrolled preparation of NPGs (*Scheme 45b*). $SnCl_4$ facilitates the formation of NPGs from acetyl mannosides (*Scheme 45c*) [117]. The most useful method for the preparation of NPGs is, arguably, the acid-catalyzed rearrangement of NPOEs, prepared under Lemieux–Morgan conditions (*Scheme 45d,e*) [163]. This method permits the stereoselective synthesis of NPGs with different protecting groups [50,164]. Rousseau and Martin described the rearrangement of acetyl NPOEs with TMSOTf (*Scheme 45d*) [99], and Fraser-Reid and co-workers have used TBSOTf [165] or ytterbium triflate [166] to rearrange benzoyl-substituted NPOEs (*Scheme 45e*).

3 Enol Ether-Type Glycosides

3.1 Early Contributions

De Raadt and Ferrier were the first to report the preparation and attempted glycosylation of 1′-substituted-vinyl glycosides [167]. Reaction of tetra-*O*-acetyl-α-D-glucopyranosyl bromide (**288**) with bis(acetonyl)mercury derivatives **289a–c** in refluxing chloroform afforded vinyl-, isopropenyl-, and styryl-β-D-glucosides **290a–c** in excellent yields (*Scheme 46*). However, when **290a–c** were each treated with either NBS or bromine/$AgClO_4$ in the presence of methanol no glycosides were formed, the products in each case being mixed stereoisomers of the glycosyl acetals **291a–c**.

Schmidt and co-workers described the preparation of vinyl glucosides **292** from the reaction of tetra-*O*-benzyl glucose with ethyl phenyl propiolate under the agency of sodium hydride

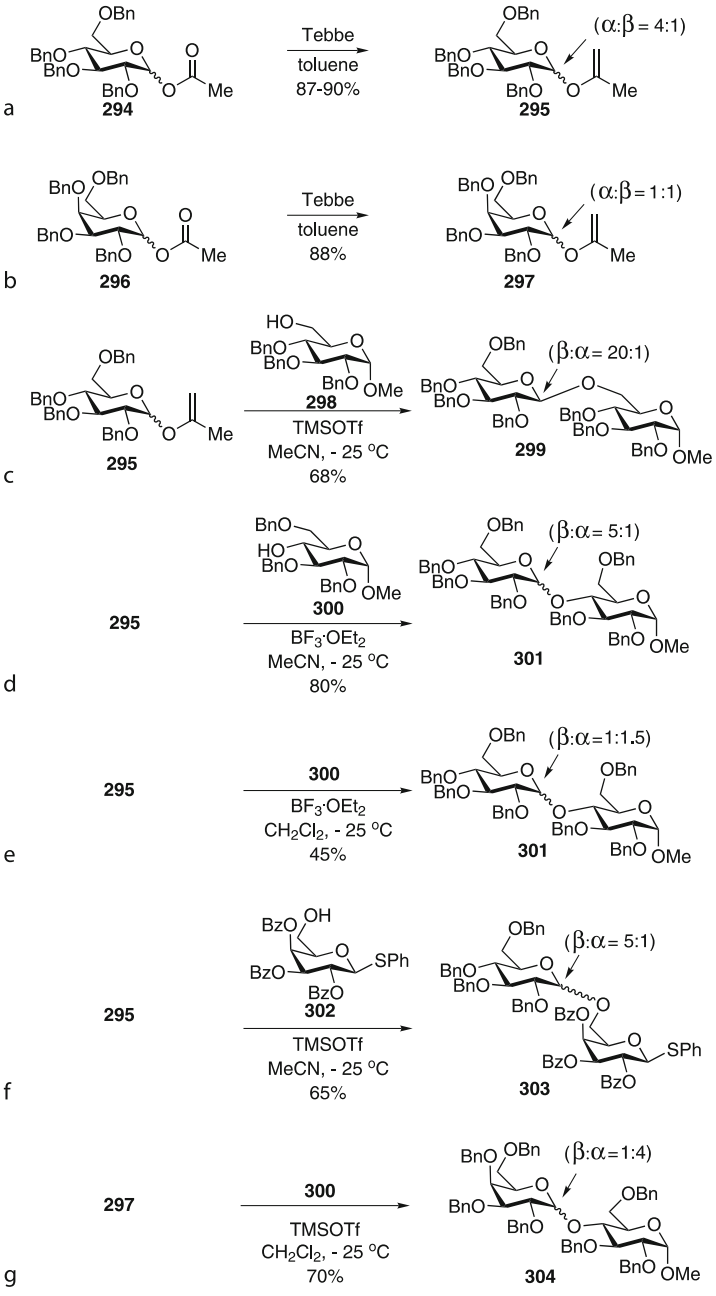
(Scheme 47) [168]. The reaction of **292**, as an anomeric mixture, with various acceptors was examined in acetonitrile at −40 °C in the presence of TMSOTf as catalyst. Reaction of **2** with 6-OH and 4-OH methyl glucosides as acceptors gave the corresponding disaccharides in 61 and 67% yield and as 85:15 and 75:25 β/α mixtures, respectively. Similar results were obtained for tetra-*O*-benzyl galactose.

3.2 Isopropenyl Glycosides

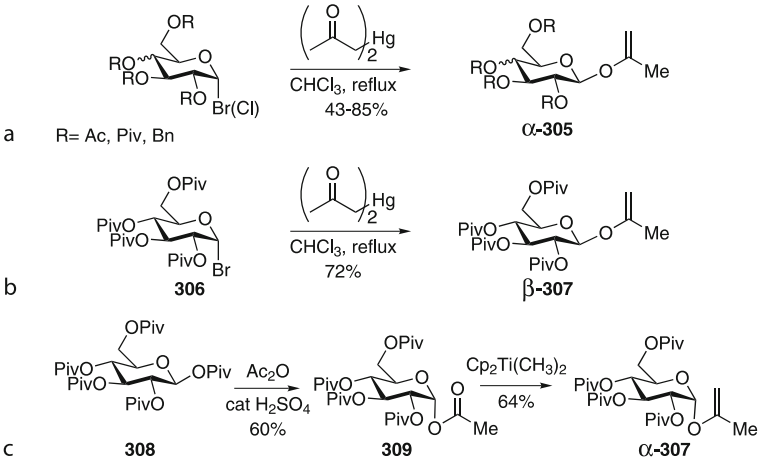
Sinaÿ and co-workers described the synthesis of isopropenyl glycosides [169] by reaction of the corresponding anomeric acetates with the Tebbe reagent [170]. Reaction of 1-*O*-acetyl-2,3,4,6-tetra-*O*-benzyl-D-glucopyranose (**294**) with a solution of Tebbe reagent in toluene gave the isopropenyl glycosides **295**, in 87–90% yields (Scheme 48*a*). Likewise, isopropenyl galactoside **297** ($\alpha:\beta \approx 1:1$) was prepared from the corresponding acetate **296** by Tebbe methylenation in 88% yield (Scheme 48*b*). Treatment of **295** ($\alpha:\beta = 4:1$) in MeCN at −25 °C with the primary hydroxyl acceptor **298**, in the presence of TMSOTf gave the disaccharides **299** (68%) with an excellent β-selectivity (20:1) (Scheme 48*c*). The condensation of **295** with the secondary alcohol **300** in MeCN at −25 °C in the presence of $BF_3{\cdot}Et_2O$ afforded the disaccharide **301** in good yield, albeit with reduced stereoselectivity ($\beta:\alpha = 5:1$) (Scheme 48*d*). When the same glycosylation was carried out in CH_2Cl_2 instead of MeCN the disaccharide **301** was obtained in limited yield (Scheme 48*e*). The successful glycosylation of phenyl 1-thio-glycoside **302** with **295** in the presence of TMSOTf illustrates the usefulness of isopropenyl glycosides in the synthesis of thiophenyl disaccharides (e. g. **303**, Scheme 48*f*). The authors found no significant variations on yield or stereoselectivity by the use of either mainly α or mainly β isopropenyl derivatives. The best results for galactosylation were achieved in CH_2Cl_2 with TMSOTf as promoter (Scheme 48*g*).

Chenault and co-workers reported the use of *O*-isopropenyl glycosides bearing ester protecting groups [171,172]. These compounds are stable at room temperature and can be readily purified by column chromatography on silicagel, moreover their glycosylation would proceed to give β-glycosides via neighboring group participation. The reaction of bis(acetonyl)mercury [173] with glycopyranosyl halides proved to be a good method for the preparation of isopropenyl β-glycopyranosides (e. g. **305**, Scheme 49*a*). The authors described routes to *O*-isopropenyl 2,3,4,6-tetra-*O*-pivaloyl-α, and β-D-glucopyranosides **α-307** and **β-307**, respectively. Reaction of 2,3,4,6-tetra-*O*-pivaloyl-α-D-glucopyranosyl bromide (**306**) with diacetonyl mercury led to **β-307** (Scheme 49*b*), whereas regioselective methylidenation [174] of **309** (prepared stereoselectively by acid-catalyzed exchange of the anomeric pivaloyloxy group of penta-*O*-pivaloyl-β-D-glucopyranose, **308**) generated **α-307** as the only product (Scheme 49*c*). The β-isomer, however, exhibited greater shelf life than the latter.

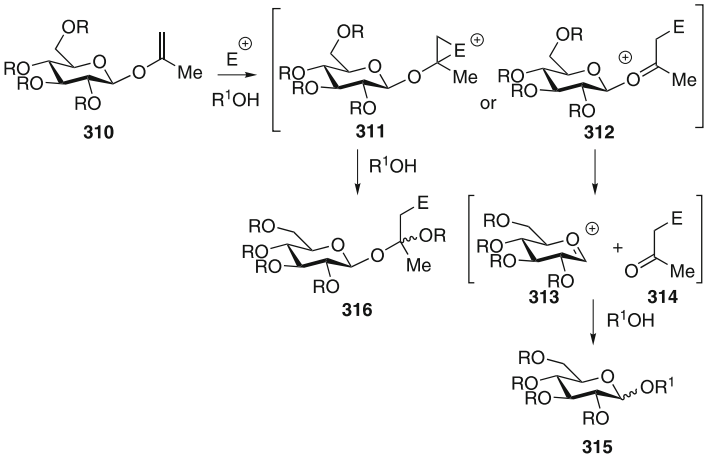
On the basis of the reaction of NPGs with electrophiles, Chenault et al. considered the possible activation of isopropenyl glycosides with electrophiles. The mechanism of activation was expected to involve initial capture of the electrophile (E^+) by the vinyl ether double bond of **310** leading to the formation of cation **311** or **312** (Scheme 50). Collapse of **311** or **312** to form glycosyl oxocarbenium ion **313** and acetone derivative **314** would be followed by nucleophilic attack on **313** to generate glycoside **315**. An alternative reaction would involve direct nucleophilic attack on **311** or **312** to generate the addition product **316**.



Scheme 48
Synthesis and glycosylation reactions of isopropenyl glycosides

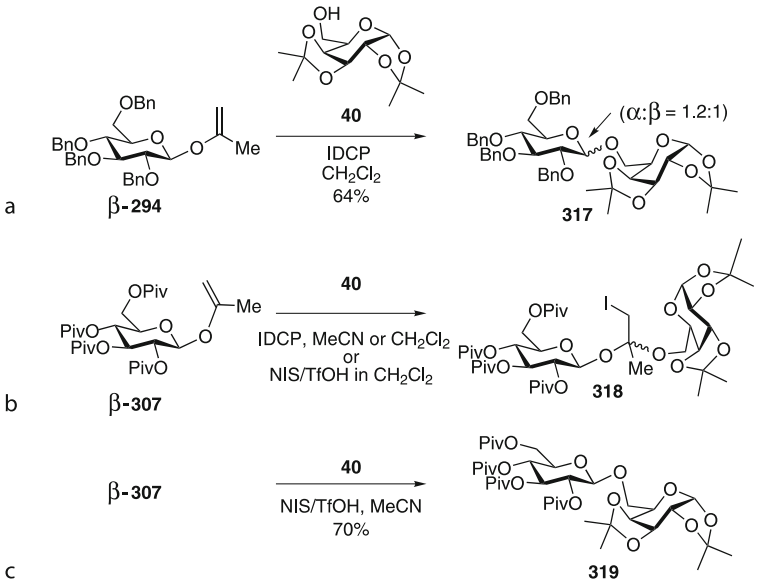


Scheme 49
Synthesis of pivaloyl isopropenyl glycosides



Scheme 50
Activation of isopropenyl glycopyranosides

The authors found that "armed" and "disarmed" isopropenyl glycosides displayed different behavior towards electrophiles (*Scheme 51*). Armed isopropenyl glycoside ***β*-294**, glycosylated acceptor **40** to give disaccharide **317** under the agency of IDCP, a relatively weak electrophile, in a nonpolar solvent (CH_2Cl_2) (*Scheme 51a*). On the other hand, disarmed glycoside ***β*-307**, led under the same conditions to the electrophilic addition product **318**. Use of a more potent electrophile (NIS/TfOH) in CH_2Cl_2 also resulted in the formation of the addition product **318** (*Scheme 51b*). However, NIS/TfOH in more polar MeCN successfully promoted the glycosidic coupling (*Scheme 51c*). Apparently, the relatively electron-releasing ethereal protecting groups lower the energy barrier to oxocarbenium ion formation from



Scheme 51
Reaction of armed and disarmed isopropenyl glycopyranosides

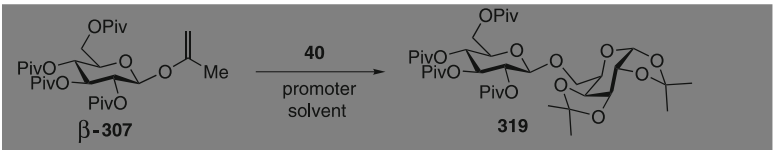
armed ***β*-294** relative to that from disarmed glycoside ***β*-307**. In general, factors which favor the formation of the glycosyl oxocarbenium ion (strong electrophile, polar solvent, electron-releasing protecting groups on the glycosyl donor) lead to transglycosylation. Factors which retard the formation of the glycosyl cation (weak electrophile, nonpolar solvent, electron-withdrawing protecting groups on the glycosyl donor) lead to addition across the isopropenyl ether double bond.

The ability of various electrophiles to promote transglycosylation of disarmed isopropenyl glycosides is outlined in *Table 10*. NIS/TfOH, TMSOTf, and Tf_2O in MeCN, all led to the formation of disaccharide **319** in good yield (*Table 10*, entries i–iii). Reactions were carried out at 0 °C and were complete within 2–5 min. With silver triflate (AgOTf) the reaction was slower and gave a lower yield of disaccharide **319** (*Table 10*, entry iv). When TfOH, NIS, or NBS were used alone ***β*-307** failed to react and the glycosyl donor was recovered unchanged (*Table 10*, entries vi, viii). Thus, neither NIS/TfOH, TMSOTf, Tf_2O, nor AgOTf seem to activate isopropenyl glycosides by acting as a source of TfOH. Dimethyl(methylthio)-sulfonium triflate (DMTST) was the only promoter that led exclusively to the formation of disaccharide **319** from ***β*-307** when CH_2Cl_2 was used as the solvent.

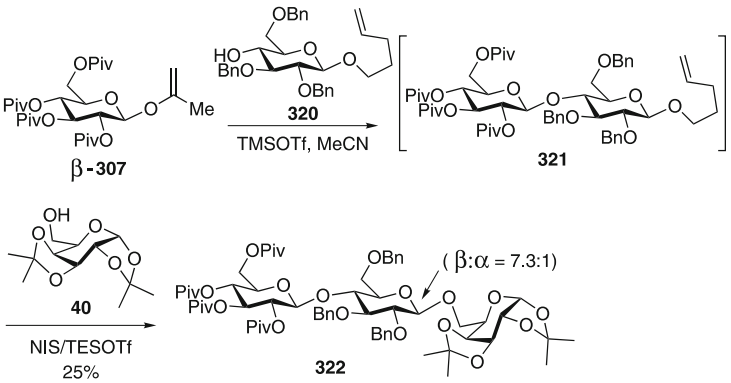
In terms of glycosyl donors, either ***α*-307** or ***β*-307** gave the same results in terms of yields. Likewise, isopropenyl galactopyranosides reacted in a similar manner to glucopyranosides. Acylated isopropenyl donors gave lower yields than pivaloyl analogs, presumably because of complications due to orthoester formation [175].

Isopropenyl glycosides could be activated selectively in the presence of armed NPGs, and that allowed a one-pot synthesis of trisaccharide **322** involving the successive glycosyl coupling of a vinyl glycoside ***β*-307**, and an NPG, **321** (*Scheme 52*).

Table 10
Evaluation of promoters for transglycosylation of β-307



Entry	Promoter	Solvent	Yield
i	NIS/TfOH	MeCN	70%
ii	TMSOTf	MeCN	69%
iii	Tf_2O	MeCN	65%
iv	silver triflate	MeCN	24% (24h)
v	DMTST	CH_2Cl_2	48%
vi	TfOH	MeCN	no reaction
vii	trimethylsilyl iodide	MeCN	no reaction
viii	NIS or NBS	MeCN	no reaction



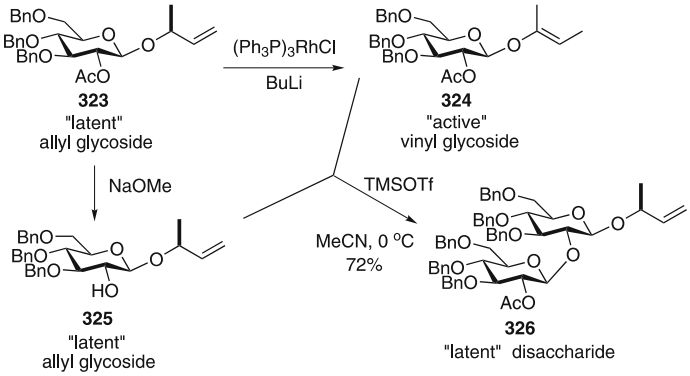
Scheme 52
One-pot synthesis of trisaccharide 322

3.3 3-Butene-2-yl Glycosides as Precursors for Vinyl Glycosides

Boons and co-workers introduced stable allyl glycosides (e. g. **323**, *Scheme 53*), which are converted to the enol ether-type glycosides **324**, prior to glycosylation [176].

3.3.1 Latent-Active Glycosylation Strategy

The allyl glycoside **323**, can be considered a "latent" [177] form of a glycosyl donor which can be efficiently isomerized to the "active" vinyl glycoside, **324**. The isomerization reaction was performed by a rhodium catalyst obtained by treating the Wilkinson's catalyst, $(Ph_3)P_3RhCl$,

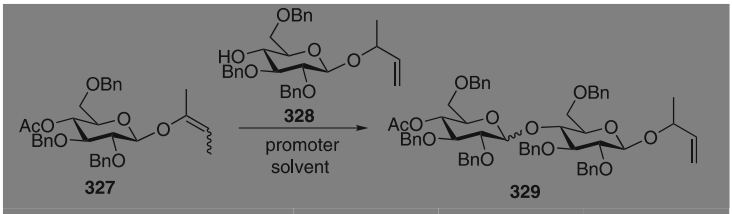


Scheme 53
Vinyl glycoside-based latent-active strategy for glycosyl coupling

with BuLi [178]. Base labile functionalities in the molecule are compatible with these isomerization conditions [179]. The "active" vinyl glycoside **324**, undergoes Lewis acid-catalyzed glycosylation reactions with "latent" allyl glycoside **325**, to give "latent" disaccharide **326** (*Scheme 53*). Unlike isopropenyl glycosides, which require stoichiometric amounts of Lewis acids for activation [169], the reaction of Boons' vinyl glycosides only demands catalytic amounts of TMSOTf. The higher reactivity of the substituted vinyl glycoside was ascribed to the additional methyl substituent of the vinyl moiety that makes the double bond more electron rich. Although racemic 3-buten-2-ol could be used for the preparation of **323** without affecting its reactivity, the use of diastereomeric allyl glycosides can be avoided with the use of optically pure 3-buten-2-ol, easily obtainable in multigram amounts.

The use of neighboring participating groups permits the formation of 1,2-*trans* glycosides (e. g. **326**, *Scheme 53*). The choice of solvent and, to some extent, the choice of activator, was used to control the α/β ratio in glycosyl donors without participating groups at *O*2. TMSOTf-

Table 11
Effect of the solvent and promotor in α/β selectivity

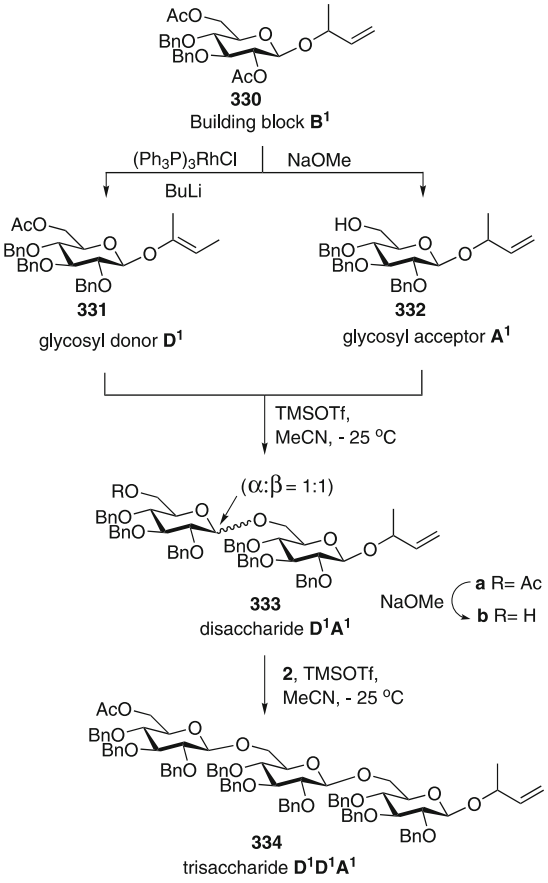


Entry	Promoter	Solvent	α/β ratio	Yield
i	TMSOTf	MeCN	1:8	78%
ii	TMSOTf	CH_2Cl_2	1.5:1	69%
iii	$BF_3{\cdot}OEt_2$	Et_2O	1.5:1	69%
iv	NIS/TfOH	Et_2O/dichloroeth	3:1	73%

promoted condensation of **327** with **328** in MeCN gave disaccharide **329** as the β-anomer mainly (α/β = 1:8) (*Table 11*, entry i). An improved α-selectivity was obtained (73%, α/β = 3:1) when the coupling was performed in ether/dichloroethane (*Table 11*, entry iv).

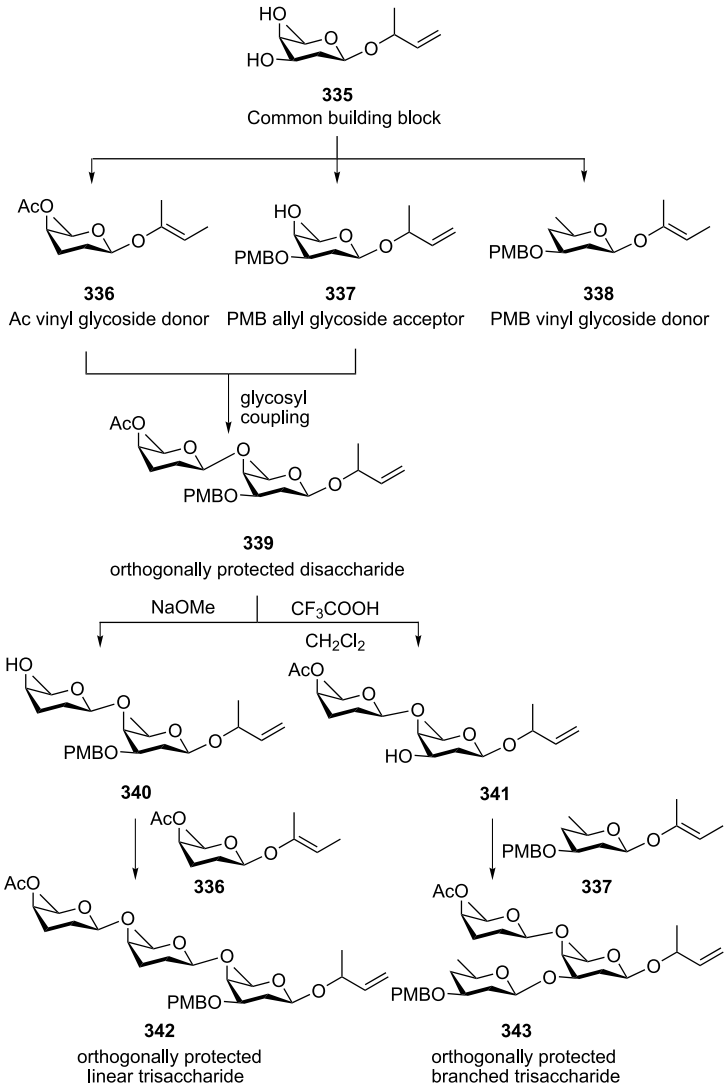
3.3.2 Preparation of Trisaccharide Libraries

Linear Trisaccharide Libraries Boons et al. described an approach to combinatorial synthesis of trisaccharide libraries based on their latent-active glycosylation strategy [180]. One major building block, **330** (i. e. $\mathbf{B}^1$, *Scheme 54*) can be converted into a glycosyl donor **331** (i. e. $\mathbf{D}^1$) and a glycosyl acceptor **332** (i. e. $\mathbf{A}^1$). Coupling of compounds **331** and **332** gives disaccharide **333a** in excellent yield (the anomeric ratio can be greatly influenced by changes in the temperature: α/β = 1:20 at low temperature; α/β = 1:1 at ambient temperature). The latter can be converted into a glycosyl acceptor **333b** (i. e. $\mathbf{D}^1\mathbf{A}^1$) by removing the acetyl protecting



Scheme 54
Preparation of linear trisaccharide libraries

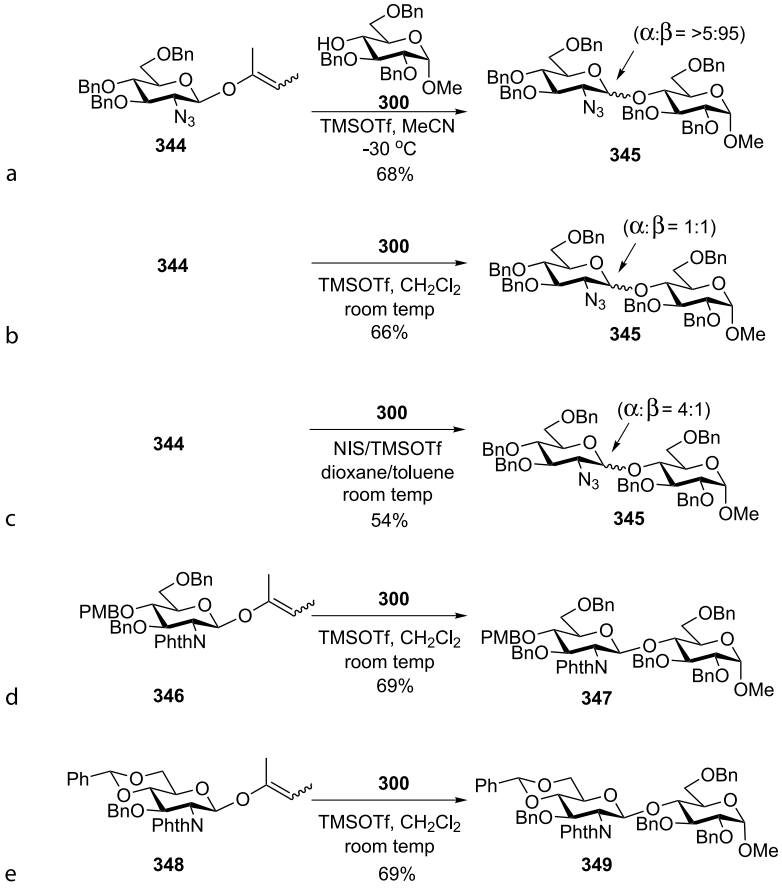
group and into a glycosyl donor by isomerizing the allyl moiety. These compounds can be used in oligosaccharide synthesis, as outlined in *Scheme 54*, for example by coupling **333b** with **332** to give trisaccharide **334** (i. e. $\mathbf{D}^1\mathbf{D}^1\mathbf{A}^1$). Application of this strategy to four allyl building blocks ($\mathbf{B}^{1\rightarrow4}$) would lead to four vinyl glycosyl donors ($\mathbf{D}^{1\rightarrow4}$) and four allyl glycosyl acceptors ($\mathbf{A}^{1\rightarrow4}$). Individual glycosylations of each donor with each acceptor will furnish 16 disaccharides ($\mathbf{D}^{1\rightarrow4}\mathbf{A}^{1\rightarrow4}$) (if glycosylations are stereoselective, or 32 disaccharides if conditions are met for 1:1 anomeric selectivity). Next, the disaccharides can be mixed, and removal



Scheme 55
Preparation of branched trisaccharide libraries

of the acetyl groups will give an assortment of acceptors. The pool of compounds can be split, and in combinatorial steps each pool of glycosyl acceptors can be coupled with a particular glycosyl donor ($\mathbf{D}^{1\rightarrow4}$) resulting in four libraries of 32 (or 64, as above) trisaccharides each.

Branched Trisaccharide Libraries Biologically important oligosaccharides often contain more complex features such as branching points and further functional groups. In this context, Boons and co-workers, using the latent-active strategy, designed a synthetic method to create orthogonally protected saccharides (acetyl and p-methoxybenzyl groups were used as orthogonal protecting groups) that could be easily further derivatized [181]. Thus, a common allyl glycoside building block (e. g. **335**, ❯ *Scheme 55*) can be converted to two vinyl glycoside donors bearing orthogonal protecting groups (e. g. **336** and **338**), and to an allyl glycosyl acceptor **337**, bearing one free hydroxyl and one selectively removable PMB ether. The latter will be coupled with donor **336** bearing an acetyl protecting group to give an orthogonally



Scheme 56
3-Buten-2-yl derivatives for the synthesis of amino sugar containing disaccharides

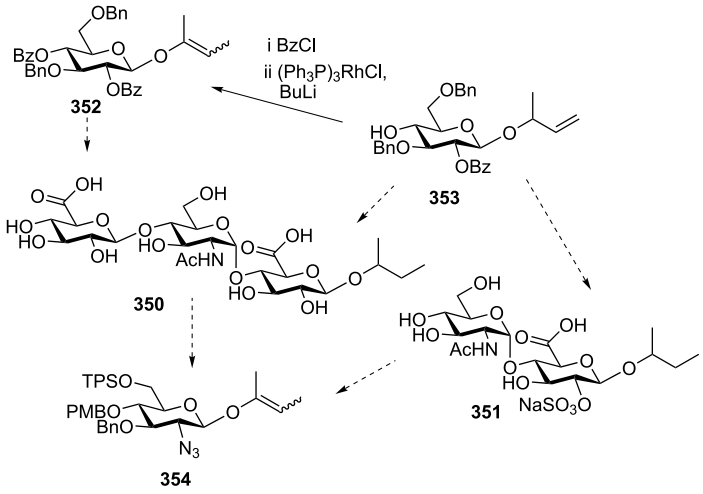
protected disaccharide, **339**. Compound **339** can be elaborated into linear or branched trisaccharides **342** and **343**. Thus, deprotection of the acetyl group in **339**, and glycosylation with vinyl donor **336** will yield linear trisaccharide **342**, whereas removal of the PMB group and coupling with **337** will produce orthogonally protected branched trisaccharide **343**.

3.3.3 3-Buten-2-yl 2-amino-2-deoxy Glycosides as Glycosyl Donors

Boons and co-workers studied the use of 3-buten-2-yl 2-azido-2-deoxy, and 2-deoxy-2-phthalimido glycosides, as building blocks for the preparation of sugar containing oligosaccharides [182]. Vinyl glycoside donors **344**, **346**, and **348**, were uneventfully prepared by isomerization of the corresponding 3-buten-2-yl glycosides with $(Ph_3)P_3RhCl/BuLi$ in yields exceeding 90%. Several glycosyl acceptors were used in the study, although representative data in ❯ *Scheme 56* refer solely to acceptor **300**. The glycosylation with azido donor **344**, in MeCN using TMSOTf as the promoter at −30 °C, proceeded with high β-selectivity (❯ *Scheme 56a*), whereas NIS/TMSOTf in a dioxane/toluene mixture gave good α-selectivities (❯ *Scheme 56c*). 2-Buten-2-yl 2-deoxy-2-phthalimido glycosides **346** and **348** reacted in CH_2Cl_2 in the presence of a catalytic amount of TMSOTf to give only the β-linked disaccharides **347** and **349**, respectively.

3.3.4 An Approach for Heparin Synthesis Based on 3-Buten-2-yl Glycosides

Haller and Boons described an approach fully based on 3-buten-2-yl glycosides for the synthesis of trisaccharide **350** and sulfated disaccharide **351** (❯ *Scheme 57*) [183]. In their strategy the glucuronic acid moieties were introduced at a late stage of the synthetic sequence by selective oxidation of primary hydroxyl groups with TEMPO and NaOCl. “Latent” allyl glycoside **353** functioned as an acceptor for the reducing end in compounds **350** and **351**, and was also transformed to “active” vinyl glycoside **352**, for the nonreducing unit of **350**. The 2-aceta-



Scheme 57
3-Buten-2-yl derivatives for the synthesis of amino sugar containing oligosaccharides

mido-2-deoxy unit in **350**, was retrosynthetically correlated with 2-azido-2-deoxy glycosyl donor **354**.

3.3.5 Conversion of 2-Buten-2-yl Glycosides to Other Glycosyl Donors

Treatment of "active" vinyl glycosides with NIS/TMSOTf in CH_2Cl_2 in the presence of dibenzyl phosphate gives good yield of glycosyl phosphates [184].
2-Buten-2-yl glycosides can also be transformed to glycosyl fluorides and trichloroacetimidates by hydrolysis to the corresponding hemiacetal (HgO, $HgBr_2$, aq. acetone) followed by standard treatment (CCl_3CN, DBU, CH_2Cl_2 or DAST, THF, respectively) [183].

3.3.6 Synthesis of 3-Buten-2-yl Glycosides

Being normal alkyl glycopyranosides, 3-buten-2-yl glycosides can be prepared by standard glycosylation methods, as previously mentioned for NPGs.

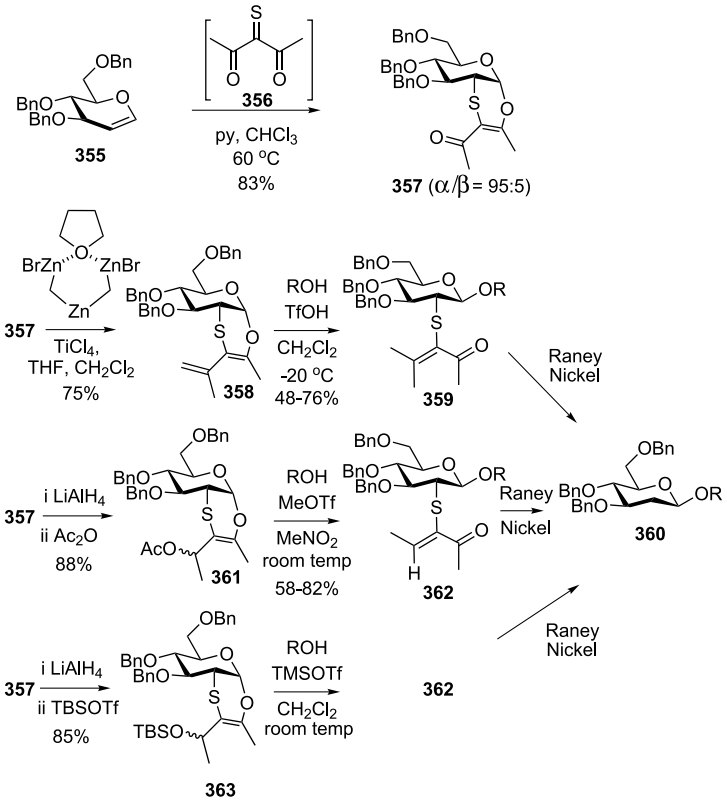
3.4 Oxathiines: Vinyl Glycosyl Donors for the Synthesis of 2-Deoxy Glycosides

Cycloadduct **357**, readily available by cycloaddition of tri-*O*-benyzl glucal (**355**) with the electron-poor 3-thioxopentane-2,4-dione (**356**) [185] has been used by two research groups as precursor glycosyl for vinyl glycosyl donors **358**, **361**, and **363** (◉ *Scheme 58*). Franck and co-workers showed that glycoside **358**, prepared by methylenation of **357**, underwent β-selective glycosylation with a variety of glycosyl acceptors in the presence of TfOH to give glycosides **359**, in good yields [186]. Moreover, Raney nickel desulfurization of **359** granted access to 2-deoxy-glycosides **360** [187]. Capozzi and co-workers reported that acetyl [188], and silyl [189] derivatives **361** and **363**, also functioned as glycosyl donors in reactions catalyzed by MeOTf in nitromethane and TMSOTf in CH_2Cl_2, respectively. The timing in the quenching of the reactions is crucial for obtaining completely selective β-glycosylations, and prolonged reaction times led to α/β anomeric mixtures. The total β-stereoselectivity of the coupling was ascribed by Capozzi and co-workers to an S_N2 type reaction (**361**→**362**, ◉ *Scheme 59*) that induces β-stereospecific glycosylation [188]. The observed subsequent α/β–equilibration presumably proceeds through an oxonium intermediate **364** (◉ *Scheme 59*).

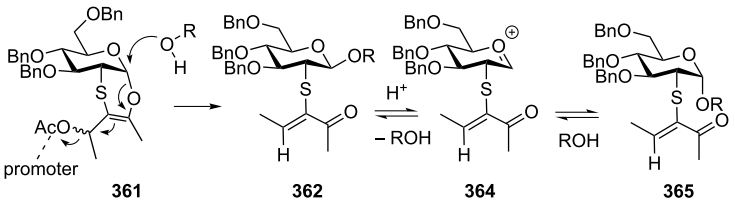
4 DISAL Glycosyl Donors

4.1 Synthesis and Glycosylation Reactions

Petersen and Jensen reasoned that glycosides of phenols (e. g. **366**) carrying sufficiently electron-withdrawing substituents could possibly serve as *O*-glycosyl donors under neutral or mildly basic conditions (◉ *Scheme 60*) [190,191]. Carbohydrate hemiacetals have been used as nucleophiles in aromatic substitutions using activated fluoroarenes [192,193]. Accordingly, glycosides of methyl 2-hydroxy-3,5-dinitrobenzoate (DISAL, a *DI*nitro*SAL*icylic acid derivative), e. g. **367** (◉ *Scheme 61*), and methyl 4-hydroxy-3,5-dinitrobenzoate (*para*-isomer)



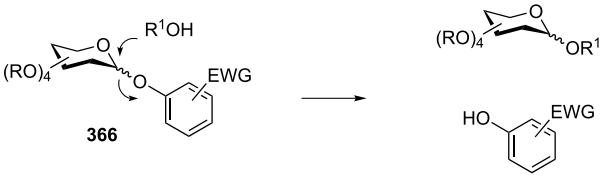
Scheme 58
Oxathiines: vinyl glycoside donors for the synthesis of 2-deoxy glycosides



Scheme 59
Stereocontrol in the formation of 2-deoxy glycosides from oxathiines

were prepared by reaction of carbohydrate hemiacetals with the corresponding activated fluoroarenes in the presence of a base. The use of 4-(*N*,*N*-dimethylamino)pyridine (DMAP) gave an α/β ratio similar to the starting 1-OH, i. e. predominantly α. In contrast, the formation of β-DISAL donors was favored using 1,4-dimethylpiperazine as base. The fluoroarenes were prepared by nitration of 2-fluoro- or 4-fluoro-benzoic acid.

The preparation of disaccharides, from benzyl-protected DISAL donors (e. g. **367**, ◉ *Scheme 61*), was best carried out in 1-methylpyrrolidin-2-one (NMP), a high polar, aprotic solvent,



Scheme 60
Glycosides of phenols with electron-withdrawing substituents as glycosyl donors

at 40 °C, in the absence of Lewis acids (*Scheme 61a,b*). The fact that glycosylations also occurred in the presence of base (e. g. Et_3N, 2,6-lutidine) indicated that the glycosylations were not auto-catalytically promoted by the released phenol. Under these conditions, galactose derivative **40** was glycosylated with DISAL donor **367** (1.5 equiv.) to give disaccharide **368** in 90% yield (α/β = 2.4:1). Glycosylation of a secondary hydroxyl group with DISAL donor **367** required increasing the temperature to 60 °C, and resulted in the formation of disaccharide **370** as the α-glycoside in 74% yield (*Scheme 61b*). The *para*-glycosyl donor, (vide supra) also proved effective in analogous glycosylations.

Unlike benzyl-protected DISAL donors, benzoyl-protected donors, e. g. **371**, did not give the expected glycosides under these neutral conditions, in part due to trapping of intermediates as the orthoesters (*Scheme 61c*). Lewis acids, such as $BF_3{\cdot}Et_2O$ or TMSOTf, activated the acylated DISAL donor **367**, albeit diisopropylidene acceptors **40** and **369** were not stable in the reaction media [194]. More robust benzyl-protected acceptors were glycosylated with alkylated and acylated DISAL donors in the presence of $BF_3{\cdot}Et_2O$ to give disaccharides **372** and **373** in 82 and 46% yield, respectively (*Scheme 61d,e*). Interestingly, $LiClO_4$ was found to be an efficient additive for activation of DISAL donors in nonpolar solvents, giving significantly higher yields of disaccharides than $BF_3{\cdot}Et_2O$ (*Scheme 61f*). Acylated DISAL donor **371** did not give good yield of disaccharides when reacting with secondary hydroxyl acceptors (*Scheme 61g*). More recently, Jensen and co-workers have shown that high-temperature glycosylation of DISAL donors using precise microwave heating results in improved yield of disaccharides (*Scheme 61h*) [195].

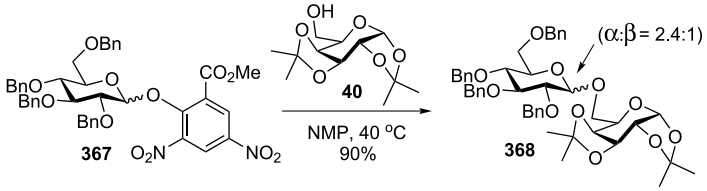
4.2 DISAL Donors in Solid-Phase Synthesis

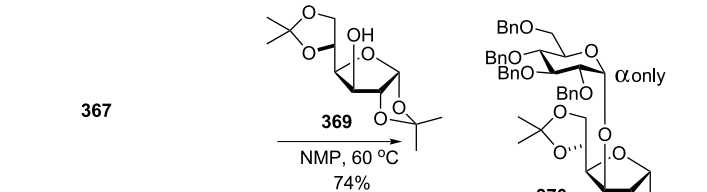
This approach was extended to solid-phase glycosylation of D-glucosamine derivatives anchored by the 2-amino group through a Backbone Amide Linker to a solid support [196].

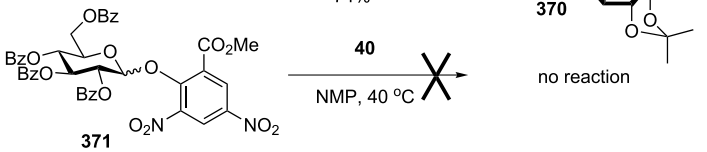
4.3 Intramolecular Glycosylation Approach to the Synthesis of 1,4-Linked Disaccharides

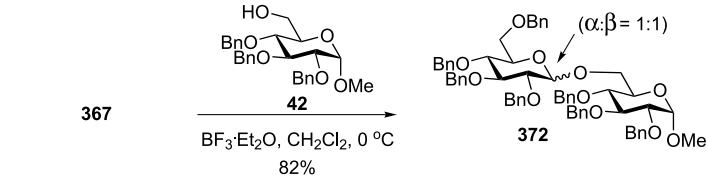
The DISAL donor concept was developed further to allow intramolecular glycosylations [197]. The glycosyl donor and acceptor were linked through the DISAL leaving group positioned to facilitate intramolecular glycosyl transfer to 4-OH by a 1,9-glycosyl shift (*Scheme 62a*).

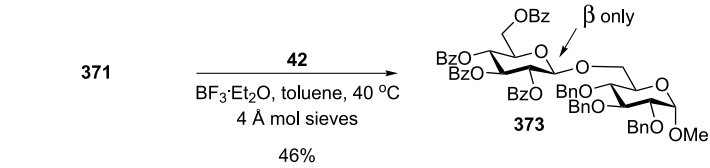
Scheme 61
Glycosylation with DISAL glycosyl donors

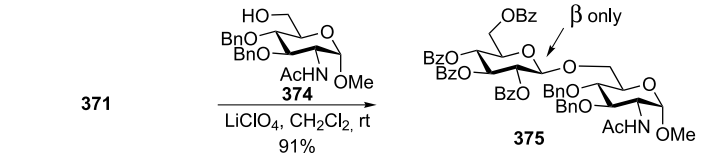
a

OBn
BnO
BnO
BnO
O
CO2Me
O2N
NO2
367
OH
40
NMP, 40 °C
90%
OBn
(α:β = 2.4:1)
368

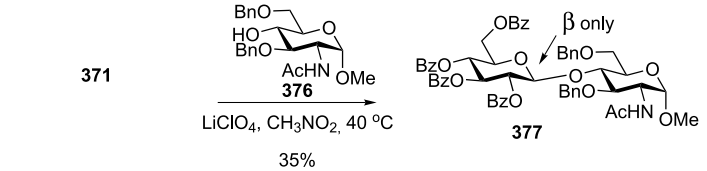
b

367
OH
369
NMP, 60 °C
74%
BnO
BnO
BnO
BnO
α only
370

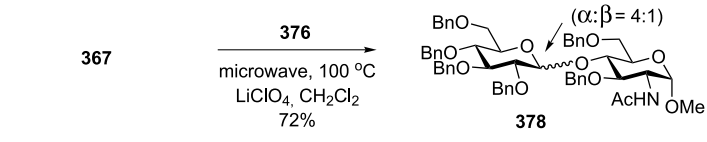
c

OBz
BzO
BzO
BzO
CO2Me
O2N
NO2
371
40
NMP, 40 °C
no reaction

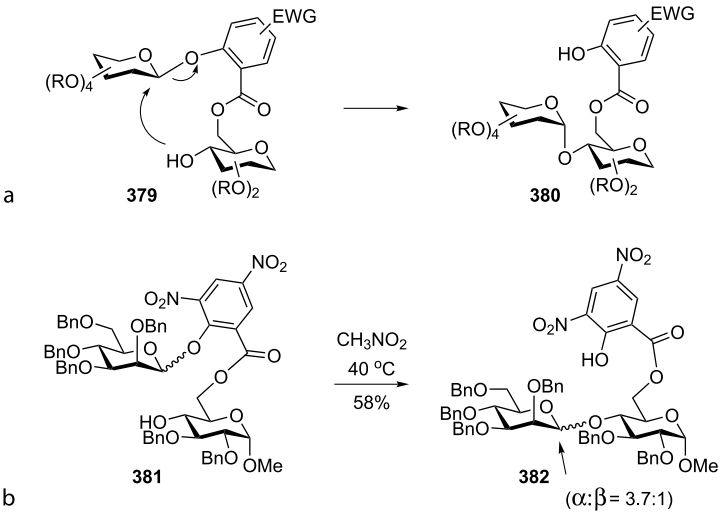
d

367
HO
BnO
BnO
BnO
OMe
42
BF3·Et2O, CH2Cl2, 0 °C
82%
OBn
(α:β = 1:1)
372

e

371
42
BF3·Et2O, toluene, 40 °C
4 Å mol sieves
46%
OBz
β only
373

f

371
HO
BnO
BnO
AcHN
OMe
374
LiClO4, CH2Cl2, rt
91%
OBz
β only
375

g

371
BnO
HO
BnO
AcHN
OMe
376
LiClO4, CH3NO2, 40 °C
35%
OBz
β only
377

h

367
376
microwave, 100 °C
LiClO4, CH2Cl2
72%
BnO
(α:β = 4:1)
AcHN
OMe
378



Scheme 62
DISAL-Based intramolecular glycosylation approach to 1,4-linked disaccharides

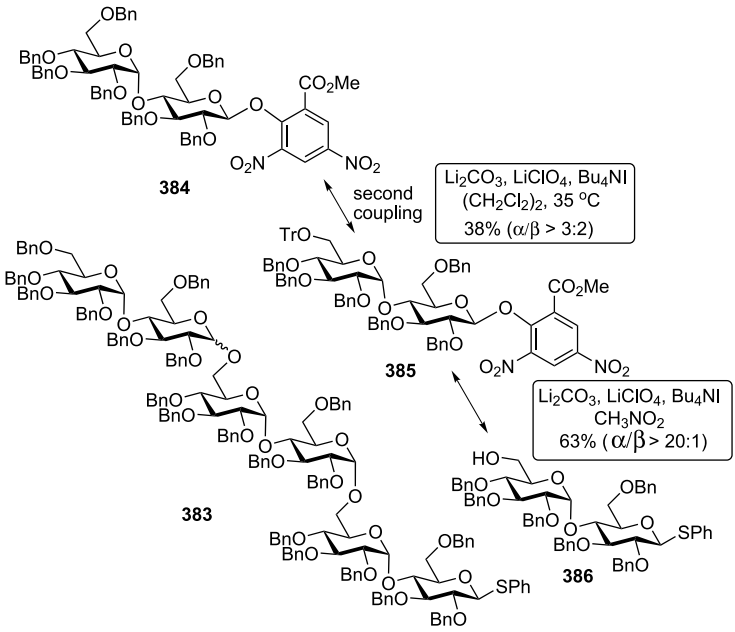
The tethered glycoside **381** underwent intramolecular transglycosylation to form the 1,4-linked mannoside **382** as an anomeric mixture (α/β = 3.7:1) in moderate yield (*Scheme 62b*).

4.4 Application of DISAL Donors to Oligosaccharide Synthesis

Jensen and co-workers reported the synthesis of hexasaccharide **383**, a starch-related hexasaccharide (*Scheme 63*) [198]. Their approach was based on the use of DISAL disaccharides **384** and **385**, readily obtained from the corresponding disaccharide hemiacetals, for sequential glycosylations. Glycosylation of phenyl 1-thio disaccharide **386** with DISAL donor **385** took place with good yield and excellent α-selectivity in CH_3NO_2 in the presence $LiClO_4$ and Li_2CO_3. The trityl group that have survived the coupling, was next removed and the ensuing tetrasaccharide glycosylated with DISAL donor **384** ($LiClO_4$, Li_2CO_3, $(CH_2Cl_2)_2$, 35 °C, 38% yield, α/β = 3:2).
DISAL donors have also been used in the preparation of phenazine natural products and analogs [199].

4.5 2-Deoxy-2-amino Derivatives as DISAL Donors

Jensen and co-workers evaluated the behavior of different glucosamine-derived DISAL donors in glycosylation reactions [200]. *N*-tetrachlorophthaloyl (TCP), *N*-trifluoroacetyl (TFAc), and *N*-trichloroethoxycarbonyl (Troc) DISAL donors **387**, **388**, and **389** and **390**, respectively, were prepared from the corresponding hemiacetals (*Fig. 6*). Glycosylation of cyclohexanol, in NMP at 60 °C, with these donors took place with yields ranging from 35 to 76%. The *N*-TCP protected donor **387**, was the least reactive. *N*-Troc protected donors **389** and **390**, gave the highest glycosylation yields with monosaccharides (63–71% yield), although they displayed



Scheme 63
Synthesis of hexasaccharide 383 based on DISAL donors

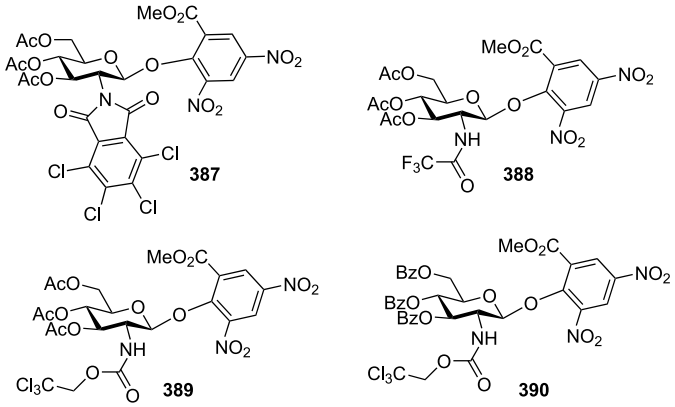
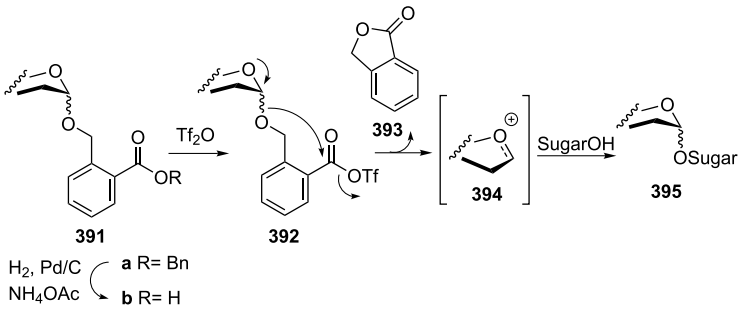


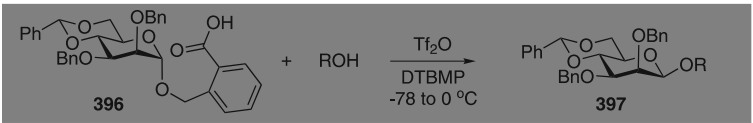
Figure 6
Glucosamine-derived DISAL donors

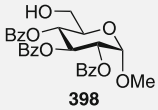
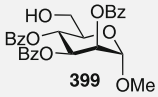
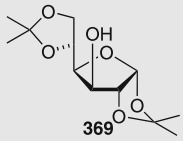
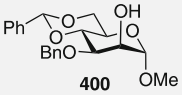
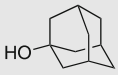
lower selectivities with primary hydroxyl acceptors (α/β ratio, from 1:1 to 1:7). A secondary hydroxyl acceptor was glycosylated with *N*-Troc DISAL donor **389** under microwave heating (130 °C, CH_3NO_2, $LiClO_4$) to give the corresponding disaccharide in 38% yield (β-anomer only). *N*-TFAc DISAL donor **388** gave even lower yields on coupling reactions with primary hydroxyl acceptors (35–45%) although β-disaccharides were obtained exclusively.



Scheme 64
2′-Carboxybenzyl (CB) glycosides

Table 12
***β*-Mannopyranosylation with 2′-carboxybenzyl glycosides**



Entry	Glycosyl acceptor (ROH)	Solvent	β/α ratio	Yield (%)
i	 **398**	CH_2Cl_2	β only	91
ii	389	toluene	20:1	95
iii	 **399**	CH_2Cl_2	β only	91
iv	 **369**	CH_2Cl_2	> 20:1	91
v	 **400**	CH_2Cl_2	17:1	96
vi	 **129**	CH_2Cl_2	> 20:1	90

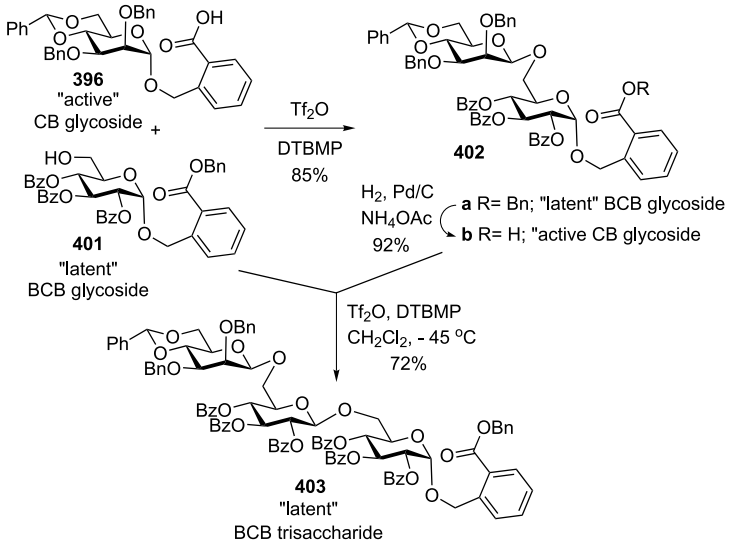
5 2′-Carboxybenzyl (CB) Glycosides

Kim and co-workers introduced a novel type of *O*-glycosyl donor, the 2′-carboxybenzyl (CB) glycoside **391b**, readily available by selective hydrogenolysis of the benzyl ester functionality

of 2-(benzyloxycarbonyl)benzyl (BCB) glycosides, **391a** [201,202,203]. Lactonization of the glycosyl triflate **392**, which was derived from the CB glycoside **391b**, is the driving force for the facile generation of the oxocarbenium ion **394** (◉ *Scheme 64*). Reaction of **394** with the glycosyl acceptor (Sugar–OH) would give the desired saccharide **395**. In the course of the transformation, a non-nucleophilic phthalide **393** is extruded. Treatment of CB glycosides with Tf_2O in the presence of di-tert-butylmethylpyridine (DTBMP) at −78 °C and subsequent addition of the glycosyl acceptor afforded the expected disaccharides in excellent yields.

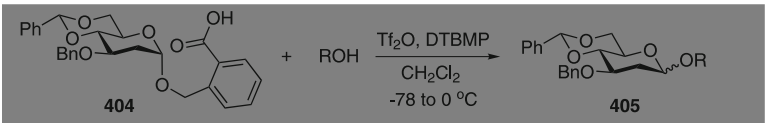
5.1 β-D-Mannosylation Employing 2′-Carboxybenzyl Glycosyl Donors

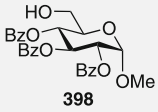
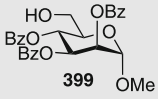
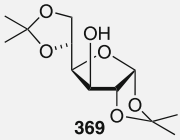
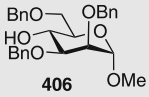
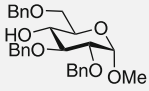
The stereospecific formation of β-mannopyranosyl linkages is a challenging task in oligosaccharide synthesis [204]. Crich and co-workers found that 4,6-*O*-benzylidene-protected glycosyl sulfoxides or thioglycosides are useful donors in the construction of β-mannopyranosyl linkages [205,206,207,208,209]. Kim and co-workers have shown that CB glycosides with a 4,6-benzylidene group can also be applied for stereoselective β-mannopyranosylation. Glycosylations of primary alcohol acceptors, **398** and **399**, in CH_2Cl_2 were completed in 1 h at −78 °C to afford only β-mannosides in high yields (◉ *Table 12* entries i, iii). Toluene was also found to be a good solvent (◉ *Table 12*, entry ii). This high β-selective mannosylation was also achieved with secondary alcohols, e. g. **369**, **400**, and with hindered tertiary alcohol **129** (◉ *Table 12*, entries iv–vi). Glucosyl CB donors possessing the 4,6-benzylidene group gave high yields of α-glucosides.



Scheme 65
2′-Carboxybenzyl glycoside-based "latent-active" strategy for glycosyl coupling

Table 13
***β*-Selective glycosylation of secondary hydroxyl acceptors**



Entry	Glycosyl acceptor (ROH)	α/β ratio	Yield (%)
i	 **398**	1:1	92
ii	 **399**	1:1.2	80
iii	 **369**	1:10	76
iv	 **406**	1:10	78
v	 **300**	*β* only	72

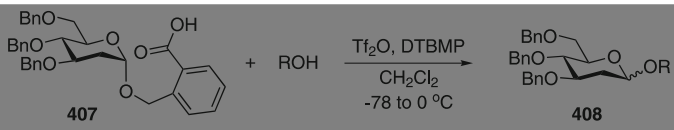
5.2 Latent-Active Glycosylation Strategy

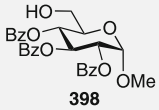
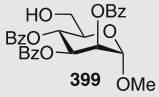
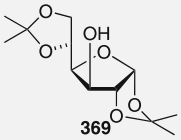
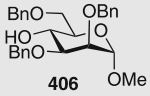
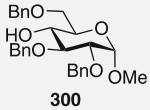
A remarkable feature of 2′-carboxybenzyl glycosides (e. g. **391b**, ➤ *Scheme 64*) is that they can be used as a latent-active pair, together with their synthetic precursors 2-(benzyloxycarbonyl)benzyl (BCB) glycosides (e. g. **391a**, ➤ *Scheme 64*). The successful mannosylation of "latent" BCB-glycoside **401** with "active" CB glycoside **396**, to give disaccharide **402a** indicated that a sequential glycosylation strategy for oligosaccharide synthesis would be possible (➤ *Scheme 65*). Thus, BCB disaccharide **402a** was readily converted into the active CB disaccharide **402b** by selective hydrogenolysis (92%, in the presence of benzyl and benzylidene groups), which upon treatment with Tf_2O/DTBMP glycosylated the latent BCB glycoside **401** to yield trisaccharide **403** in 72% yield.

5.3 Stereoselective Construction of 2-Deoxyglycosyl Linkages

Kim and co-workers have developed a highly *α*- and *β*- stereoselective (dual stereoselective) [210] method for the synthesis of 2-deoxyglycosides by employing CB 2-deoxyglycosides as glycosyl donors. Glycosylation of the 4,6-*O*-benzylidene-protected glycosyl donor

Table 14
α-Selective glycosylation of secondary hydroxyl acceptors



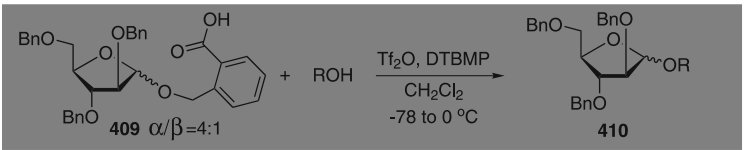
Entry	Glycosyl acceptor (ROH)	β/α ratio	Yield (%)
i	 **398**	1:1	98
ii	 **399**	1:1.2	93
iii	 **369**	α only	91
iv	 **406**	α only	91
v	 **300**	α only	88

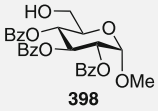
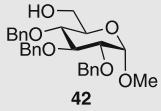
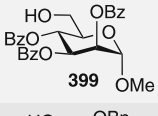
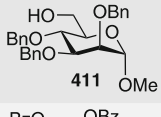
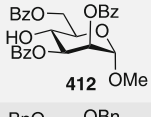
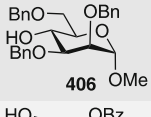
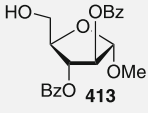
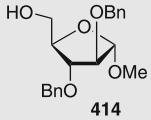
404 with secondary alcohols afforded predominantly β-glycosides (*Table 13*, entries iii–v). Complete reversal of the stereoselectivity, from β to α, was observed in the glycosylation of secondary alcohols with benzyl-protected glycosyl donor **406** (*Table 14*, entries iii–v). On the other hand, glycosylation of primary hydroxyl acceptors with both donors did not show appreciable stereoselectivity (*Table 13* and *Table 14*, entries i, ii). The authors suggested that the secondary hydroxyl acceptors formed β-disaccharides by S_N2-like displacement of an α-triflate favored in 4,6-*O*-benzylidene derivatives, as previously mentioned in the formation of β-mannosides of 4,6-*O*-benzylidene derivatives. No or poor β-selectivity in the reaction of **404** with primary alcohols was interpreted assuming that the more reactive primary alcohols reacted both with the α-triflate and an oxocarbenium ion.

5.4 2′-Carboxybenzyl Furanosyl Donors. Acceptor-Dependent Stereoselective β-D-Arabinofuranosylation

Kim and co-workers reported recently that CB tribenzyl-D-arabino furanoside **409** (easily available from methyl tribenzyl-D-arabinofuranoside) could be efficiently applied in stereos-

Table 15
β-Selective arabinofuranosylation of acyl-protected acceptors



Entry	Glycosyl acceptor (ROH)	β/α ratio	Yield β/α ratio (%)
i	 **398**	99:1	97
ii	 **42**	7:1	95
iii	 **399**	β only	95
iv	 **411**	4:1	95
v	 **412**	20:1	86
vi	 **406**	4:1	95
vii	 **413**	β only	92
viii	 **414**	2.2:1	95

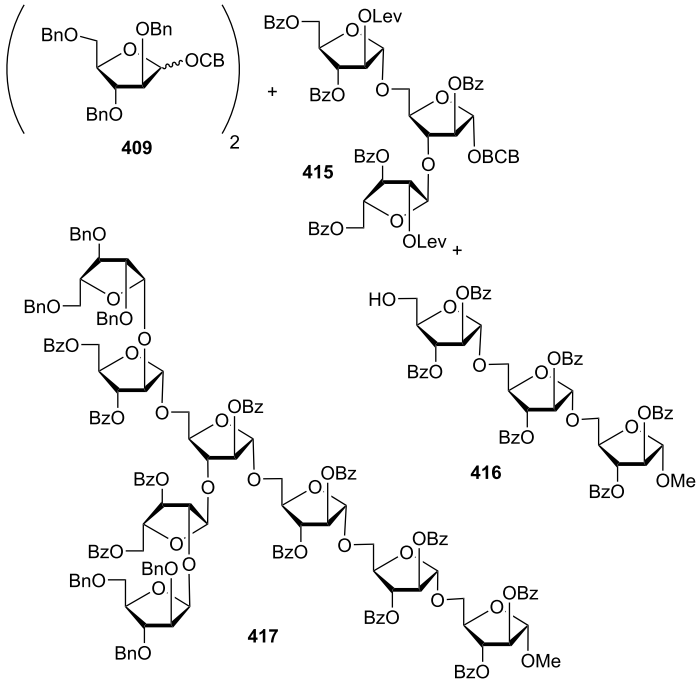
elective β-arabinofuranosylation processes [211]. They found that the presence of acyl-protective groups on the glycosyl acceptors was essential for attaining β-stereoselective glycosyl couplings. Thus, reaction of donor **409** with acceptor **398** having benzoyl-protective groups afforded a β-disaccharide almost exclusively (β/α = 99:1) in 97% yield (*Table 15*, entry i), while the same reaction with acceptor **42** having benzyl-protective groups gave a mixture of α- and β-disaccharides (β/α = 7:1) (*Table 15*, entry ii). Further examples in *Table 15*

clearly showed that the protective groups in the acceptors, regardless of pyranoses or furanoses and of primary alcohols or secondary alcohols, were the crucial factor for the outcome of the stereochemistry in glycosylations with **409**. This observed stereoselectivity was also donor dependent, since glycosylation with 2-benzyl-3,5-dibenzoyl CB arabinofuranoside was not as stereoselective [211].

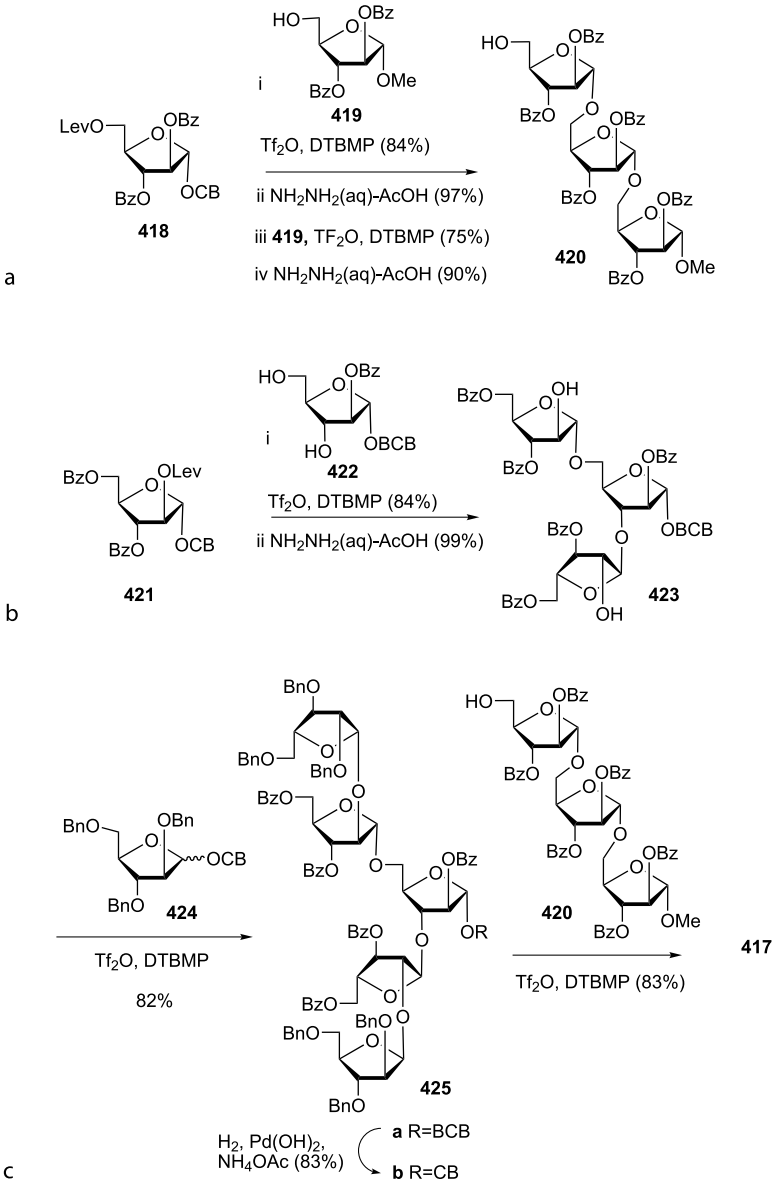
5.4.1 Synthesis of an Octaarabinofuranoside Based on Stereoselective β-D-Arabinofuranosylation

The authors applied this acceptor-dependent β-arabinofuranosylation method to the synthesis of octaarabinofuranoside **417**. Their retrosynthesis of compound **417** led to three components, a linear methyl trisaccharide **416**, a branched BCB trisaccharide **415**, and to CB furanosyl donor **409**. Levulinyl protective groups were chosen in fragments **415** and **416** for selective deprotection prior to furanosyl coupling (*Scheme 66*). Three arabinose building blocks were used in the assembly.

Arabinofuranosyl donor **418** glycosylated acceptor **419**, to yield after levulinyl-deprotection and repetitive glycosylation with **419**, the linear trisaccharide **420** (*Scheme 67a*). Coupling of latent BCB donor **422**, with active CB donor **421**, led after deprotection of the levulinyl groups to diol **423** (*Scheme 67b*). The crucial double β-arabinofuranosylation of diol **423**



Scheme 66
Retrosynthesis of octaarabinose 417. A component of lipoarabinomannan in mycobacterial cell wall

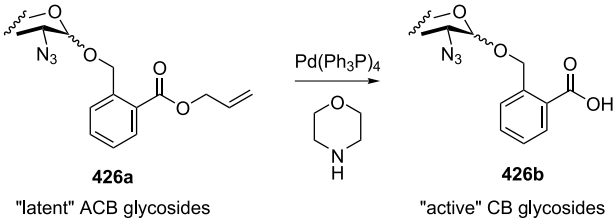


Scheme 67
Synthesis of octaarabinose 417

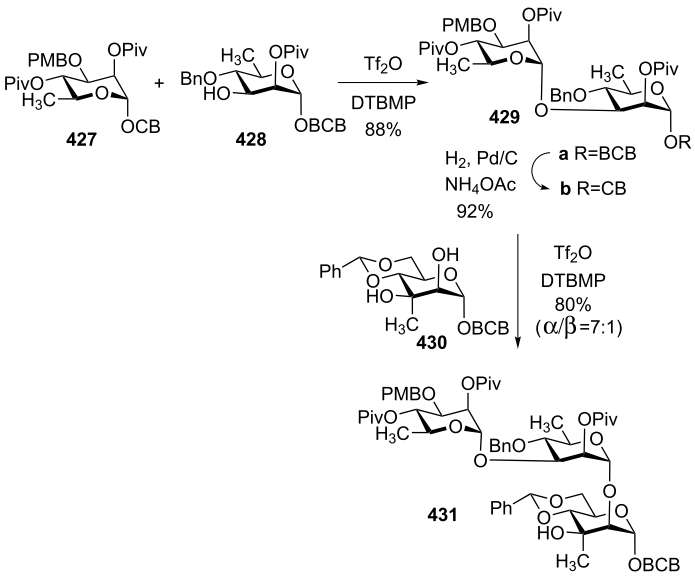
with 3.7 equiv. of the arabinofuranosyl donor **424**, paved the way to pentaarabinofuranoside **425a** (82% yield) with complete β-selectivity. The latent BCB arabinofuranoside **425a** was converted into the active CB arabinoside **425b**. Finally, coupling of the latter with triarabinofuranosyl acceptor **420**, afforded octaarabinofuranoside **417**, in 83% yield.

5.5 2′-(Allyloxycarbonyl)benzyl (ACB) Glycosides: New "Latent" Donor for the Preparation of "Active" 2-Azido-2-deoxy BC Glycosyl Donors

Kim and co-workers introduced 2′-(allyloxycarbonyl)benzyl (ACB) glycosides, e. g. **426a**, as new "latent" glycosyl donors for 2-azido-2-deoxy-glucosides [212]. Introduction of the new ACB group in the place of the previously used BCB group was necessary because the azide functionality at C-2 was also reduced during the conversion of the BCB group into the CB group under the normally used hydrogenolysis conditions (Pd/C, H_2, NH_4OAc, MeOH). 2-Azido-2-deoxy ACB glycosides could be converted into active CB glycosyl donors (e. g. **426b**, ➤ *Scheme 68*) without affecting the azide functionality on treatment with a catalytic amount of $Pd(Ph_3P)_4$ in the presence of morpholine [213].



Scheme 68
2′(Allyloxycarbonyl)benzyl (ACB) glycosides



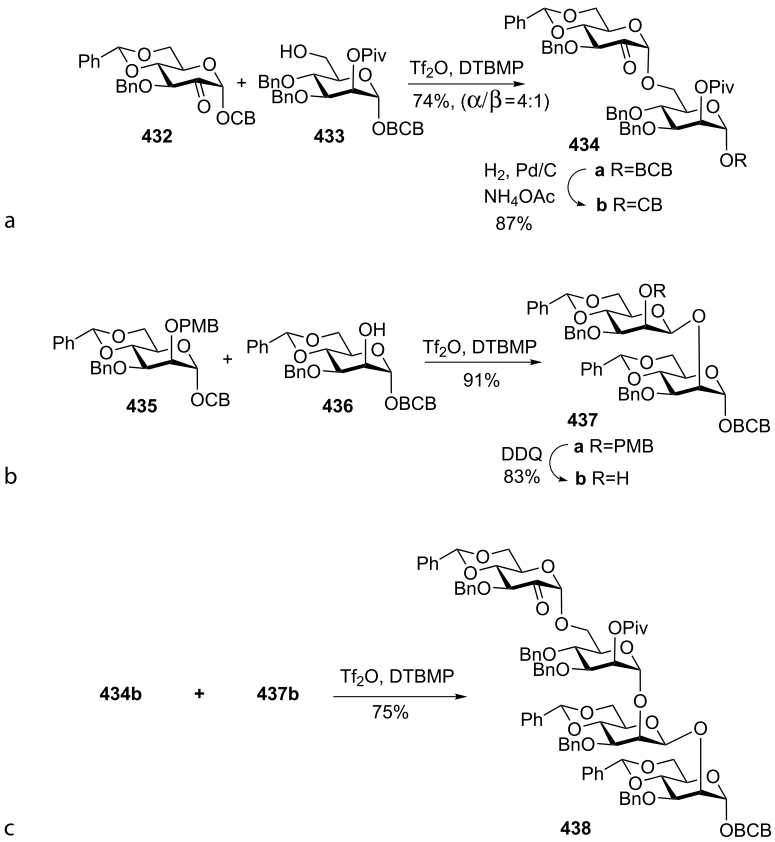
Scheme 69
Synthesis of trisaccharide 431

5.6 Synthesis of Oligosaccharides Based on BC Glycosyl Donors

The CB glycoside methodology by means of the "latent" BCB (or ACB) glycoside and the "active" CB glycoside has proved itself as a reliable method for the synthesis of complex oligosaccharides.

5.6.1 Synthesis of Trisaccharide 431, the Repeat Unit of the *O*-Antigen Polysaccharide from Danish *Helicobacter pylori* Strains

Kim and co-workers synthesized the repeat unit of the *O*-antigen polysaccharide from Danish *Helicobacter pylori* strains, **431** (Scheme 69) [214]. Coupling of donor CB L-rhamnoside **427** and acceptor BCB D-rhamnoside **428** gave α-disaccharide **429a** in 88% yield. Selective hydrogenolysis of "latent" BCB disaccharide afforded "active" CB disaccharide **429b** in



432 **433** **434** **435** **436** **437** **438**

Scheme 70
Synthesis of tetrasaccharide 438

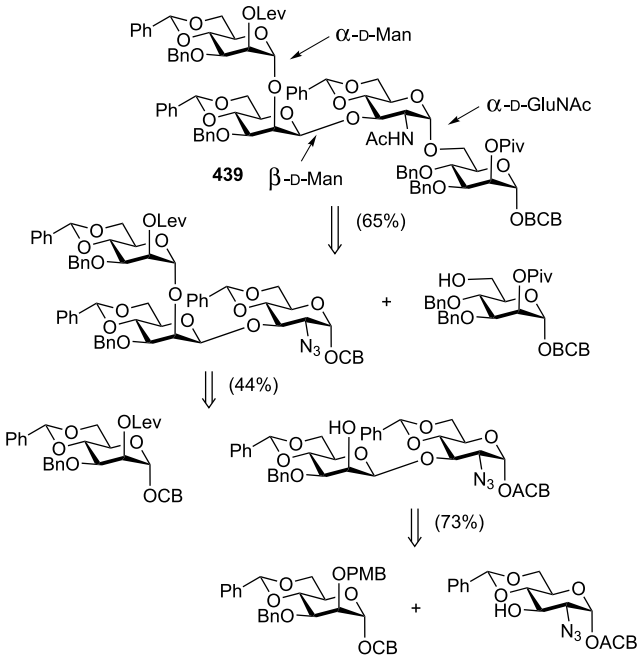
92% yield. Finally, glycosylation of 3-*C*-methyl mannoside, **430**, with **429b** yielded the target α-trisaccharide **431**, along with its β-anomer in 7:1 ratio in 80% yield. A result that indicated that neighboring group participation is operative in CB glycosides.

5.6.2 Synthesis of Tetrasaccharide 438

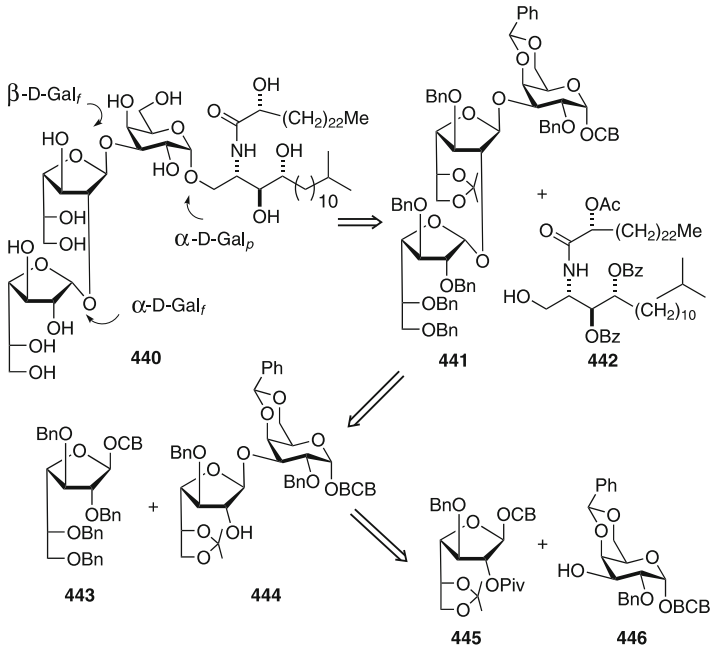
The CB methodology was also applied to the synthesis of protected tetrasaccharide **438**, an analogue of the tetrasaccharide repeat unit of the *O*-antigen polysaccharide from the *E. coli* lipopolysaccharide (*Scheme 70*) [215]. Coupling of "latent" BCB acceptor **433** with "active" CB glycosyl donor **432** gave a mixture of α-disaccharide **434a** along with its β-isomer (4:1) in 74% yield (*Scheme 70a*). Glycosylation of acceptor **436** with donor **435** gave β-mannoside **437a**, that after removal of the PMB protecting group led to **437b** (*Scheme 70b*). Finally, coupling of the latter with active donor **434b**, prepared from latent **434a**, yielded tetrasaccharide **438**, in 75% yield (*Scheme 70c*).

5.6.3 Synthesis of Tetrasaccharide Repeat Unit from *E. coli* O77

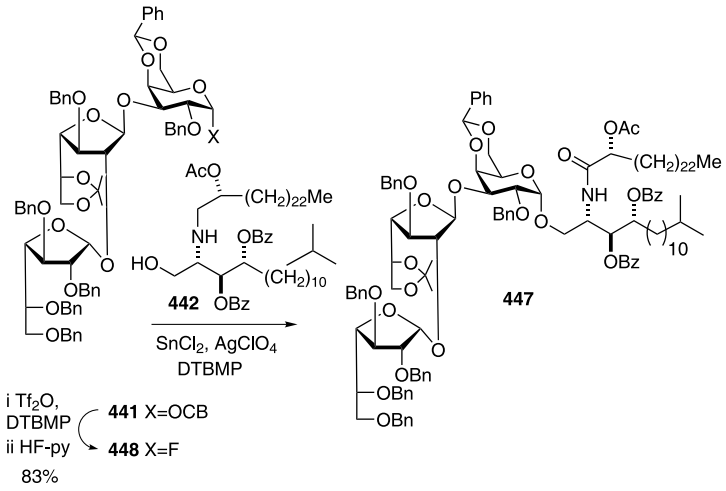
A route to a tetrasaccharide **439** was reported, which made use of the previously mentioned "latent" 2′-(allyloxycarbonyl)benzyl (ACB) glycosides in combination with "latent" and "active" BCB and CB glycosides, respectively [212]. The retrosynthesis is outlined in



Scheme 71
Retrosynthesis of tetrasaccharide 439



Scheme 72
Retrosynthesis of Agelagalastatin 440



Scheme 73
Transformation of CB glycoside 441 into glycosyl fluoride 448 and final coupling of Agelagalastatin

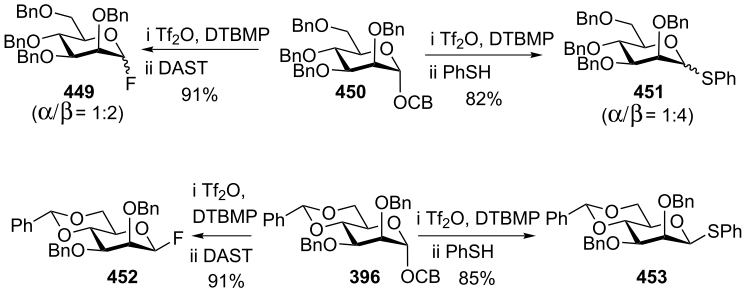
Scheme 71. All glycosyl couplings were based on the "latent-active" methodology, and all were stereoselective. A slightly modified synthesis of **439** has been reported including one CB mediated coupling [216].

5.6.4 Total Synthesis of Agelagalastatin

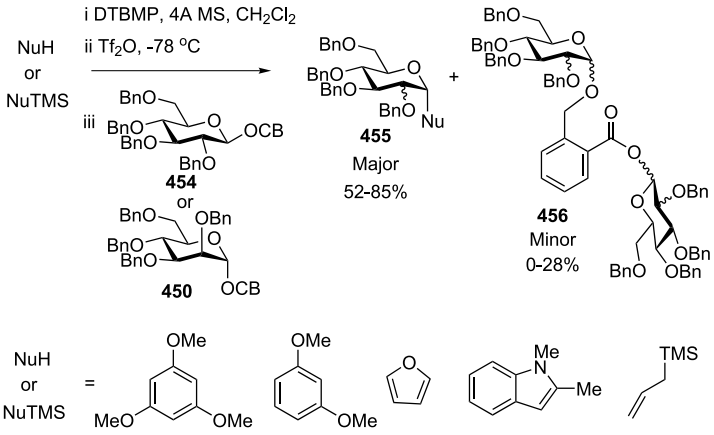
The total synthesis of agelagalastatin, an antineoplastic glycosphingolipid, has been described by Kim and co-workers [217]. The retrosynthesis, outlined in *Scheme 72*, involved a β-D-galactofuranosylation, an α-D-galactofuranosylation, and a final α-D-galactopyranosylation. The β-D-galactofuranosylation was achieved in 79% yield via neighboring group participation of the pivaloyl group at *O*2 in compound **445**. The α-D-galactofuranosylation to **441**, took place with 91% yield with a nonparticipant benzyl group at *O*2 in donor **443**. The final α-D-galactopyranosylation (*Scheme 73*) was carried out with CB trisaccharide donor **441** furnishing compound **447** in 77% yield as a 1.4:1 (α/β) mixture of saccharides. The efficiency of this coupling was improved by conversion of the CB trisaccharide donor to glycosyl fluoride **448**. Treatment of **441** with TF_2O/DTBMP followed by HF-pyridine, as a source of fluoride, yielded glycosyl fluoride **448**. The glycosylation of acceptor **442** with glycosyl fluoride **448**, then gave the target protected agelagalastatin **447**, in 72% yield as the pure α-isomer.

5.7 Conversion of 2′-Carboxybenzyl Glycosides into Other Glycosyl Donors

CB glycosides have been converted to phenyl 1-thio glycosides and glycosyl fluorides in one-pot operations [218]. Thus, treatment of CB glycosyl donors with TF_2O/DTBMP in CH_2Cl_2 at −78 °C for 10 min followed by addition of PhSH furnished thioglycosides, e. g. **451**, **453** (*Scheme 74*), whereas treatment with DAST or HF-pyridine (see *Scheme 73*) yielded the corresponding glycosyl fluorides, e. g. **449**, **452** (*Scheme 74*). The high β-selectivity observed in the formation of glycosyl fluoride **452** and thioglycoside **453** from 4,6-*O*-benzylidenemannopyranoside **396** was ascribed to the presence of a highly reactive 4,6-*O*-benzylidenemannopyranosyl α-triflate, in keeping with previously mentioned findings.



Scheme 74
Conversion of CB glycosides to glycosyl fluorides and thiophenyl glycosides



Scheme 75
CB glycosides in the synthesis of α-*C*-glycosides

5.8 2′-Carboxybenzyl Glycosides as Glycosyl Donors for *C*-Glycosylation

Glycosylation of various glycosyl acceptors (NuH or NuTMS, *Scheme 75*) with *manno*- and *gluco*- CB glycosyl donors **450** and **454**, respectively afforded α-*C*-glycosides **455**, exclusively or predominantly in good yields [218]. Experimentally these reactions were carried out by addition of the donor to a solution of the acceptor, DTBMP, and Tf_2O in CH_2Cl_2 at −78 °C. These modified conditions led to increased yields of *C*-glycosides and minimized the amount of self-condensed esters **456**.

6 *O*-Heteroaryl Glycosyl Donors

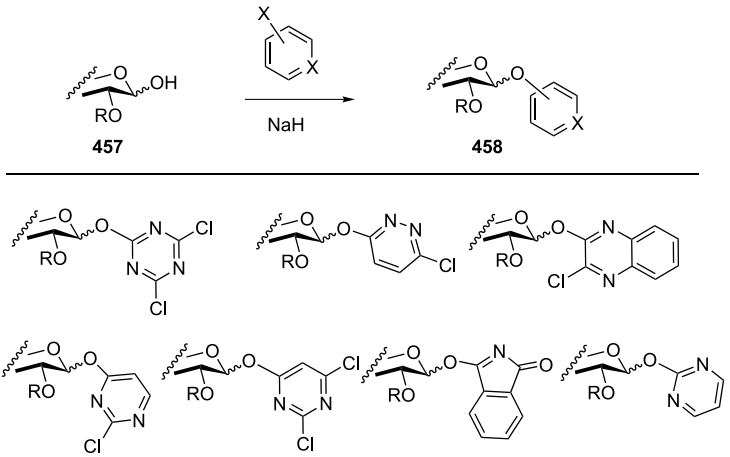
Glycosides of some heterocycles have also been investigated as glycosyl donors.

6.1 2-Pyridyl 2,3,4,6-tetra-*O*-benzyl-D-glucosides

The first example, reported by Nikolaev and Kochetkov [219], dealt with the use of 2-pyridyl 2,3,4,6-tetra-*O*-benzyl-β-D-glucoside in glycosylation. This heteroaryl glycoside was prepared by glycosylation of 2(1H)-pyridinone by the corresponding sugar chloride, and was activated by electrophiles, such as MeOTf and $Et_3O{\cdot}BF_4$, to give mixtures of *cis*- and *trans*-glycosides.

6.2 *O*-Hetaryl Glycosides by Schmidt's Group

Schmidt and co-workers [168,220,221] reported the preparation, and use in glycosylation reactions of several *O*-hetaryl glycosides, e. g. **458**, conveniently prepared by anomeric



Scheme 76
***O*-Hetaryl glycosides synthesized by Schmidt's group**

O-hetarylation of hexoses, e. g. **457**, with the corresponding electron-deficient heteroaromatic/heterocyclic systems (*Scheme 76*). The best results in terms of glycosylation were obtained with tetrafluoropyridyl glycosides **460** and **462**, obtained by reaction of hexoses **459** and **461** with 2,3,4,5,6-pentafluoro pyridine (*Scheme 77a,b*). Under TMSOTf catalysis, in CH_2Cl_2 at room temperature, they furnished the corresponding α- and β-disaccharides **463** and **464** in 98 and 74% yield, respectively (*Scheme 77c,d*).

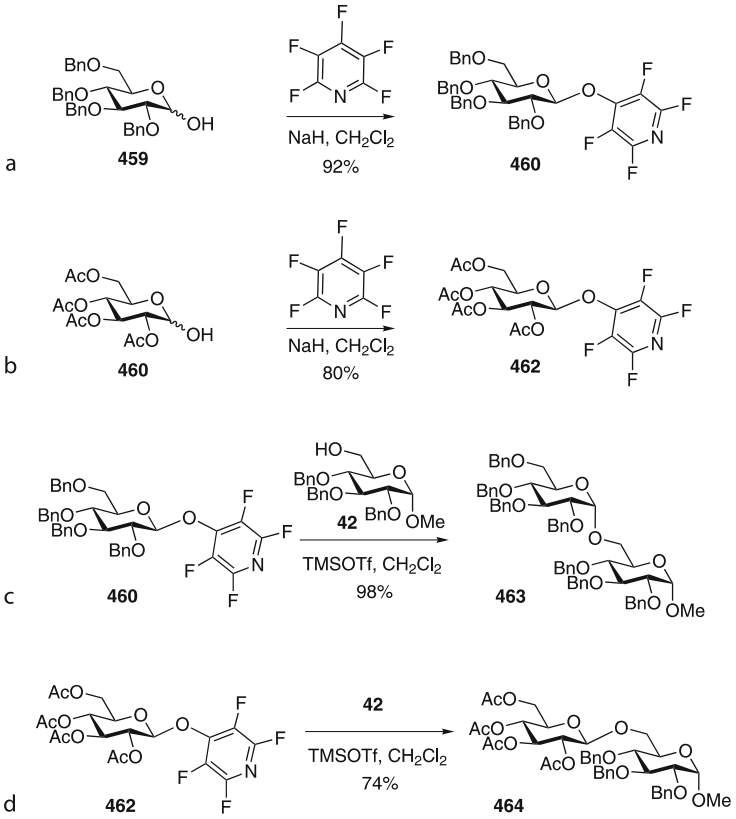
6.3 3-Methoxy-2-pyridyl (MOP) Glycosides

On the basis of the concept of remote activation [222], first applied to pyridine thioglycosides [223], Hanessian and co-workers introduced 3-methoxy-2-pyridyl (MOP) glycosides [224,225]. They first reported the usefulness of ribofuranosyl MOP donor **465** in the coupling with silylated pyrimidine bases, by activation with TMSOTf, to give 1,2-*cis* nucleosides, **466**, with high selectivity (*Scheme 78*) [226]. These glycosides react in MeOTf-, $Cu(OTf)_2$-, TfOH-, or $Yb(OTf)_3$-promoted reactions to give disaccharides [227].

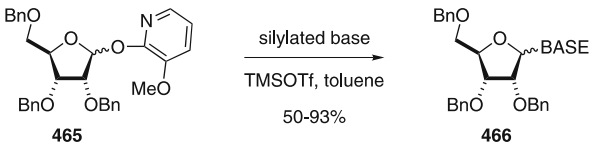
6.3.1 Coupling of Unprotected MOP Glycosyl Donors

Interestingly, unprotected MOP glycosides could also be used as donors. In fact, when using an excess of glycosyl acceptor (≈10 equiv.), disaccharides are obtained in reasonable yields, as illustrated in *Scheme 79*.

It was also found that introduction of any protecting group on the unprotected MOP glycosyl donors resulted in a significant decrease of the reactivity. This deactivation was considerable when *p*-fluorobenzoates (FBz) were used as protecting groups, and it was applied to the synthesis of disaccharides, and to iterative oligosaccharide synthesis (*Scheme 80*).



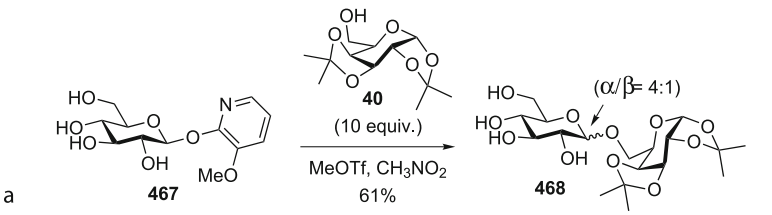
Scheme 77
2,3,5,6-Tetrafluoro-4-(2,3,4,6-tetra-*O*-acetyl or *O*-benzyl-β-D-glucopyranosyloxy) pyridine as glycosyl donors

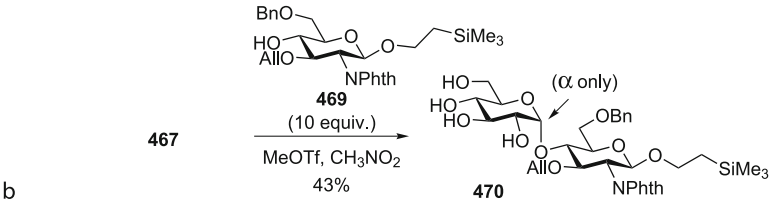


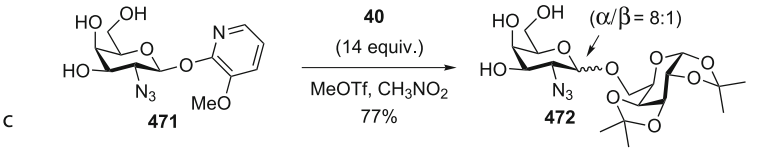
Scheme 78
MOP Ribofuranosyl donors in the synthesis of 1,2-*cis* furanosyl nucleosides

6.3.2 Esterification and Phosphorylation of Unprotected MOP Glycosides

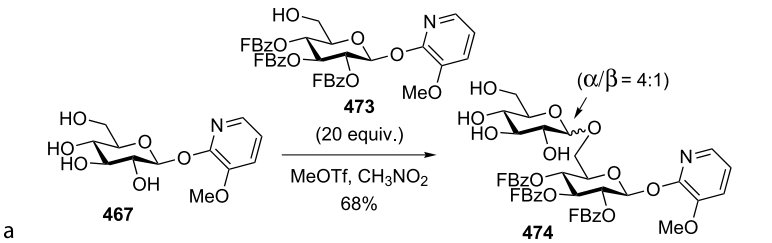
MOP glycosyl donors have been used in stereocontrolled esterification and phosphorylation, leading to glycosyl 1,2-*cis*-1-carboxylates or glycosyl 1,2-*cis*-glycosyl-1-phosphates in one step. Treatment of MOP donor **475** in acetonitrile or DMP with an excess (20–200 equiv.) of a carboxylic acid under anhydrous conditions led to the corresponding D-glycosyl carboxylate **477**, in excellent yields [228]. Moreover, treatment of 6-*O*-tert-butyldiphenylsilyl MOP donor **478** with only 1.5 equiv. of the corresponding carboxylic acid in CH_2Cl_2 resulted in the formation of 1,2-*cis*-glycosyl carboxylate **479** (*Scheme 81a*). Treatment of β-D-galactopyranosyl, and 2-azido-2-deoxy-α-D-galactopyranosyl MOP donors **475** and **471**, respectively with

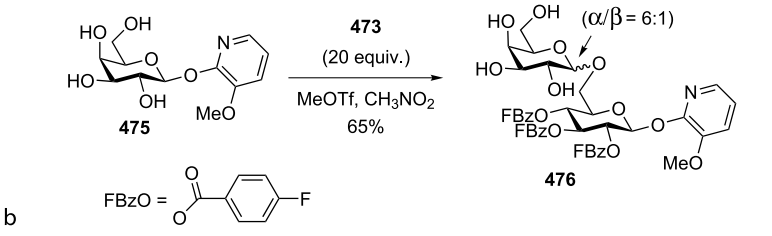




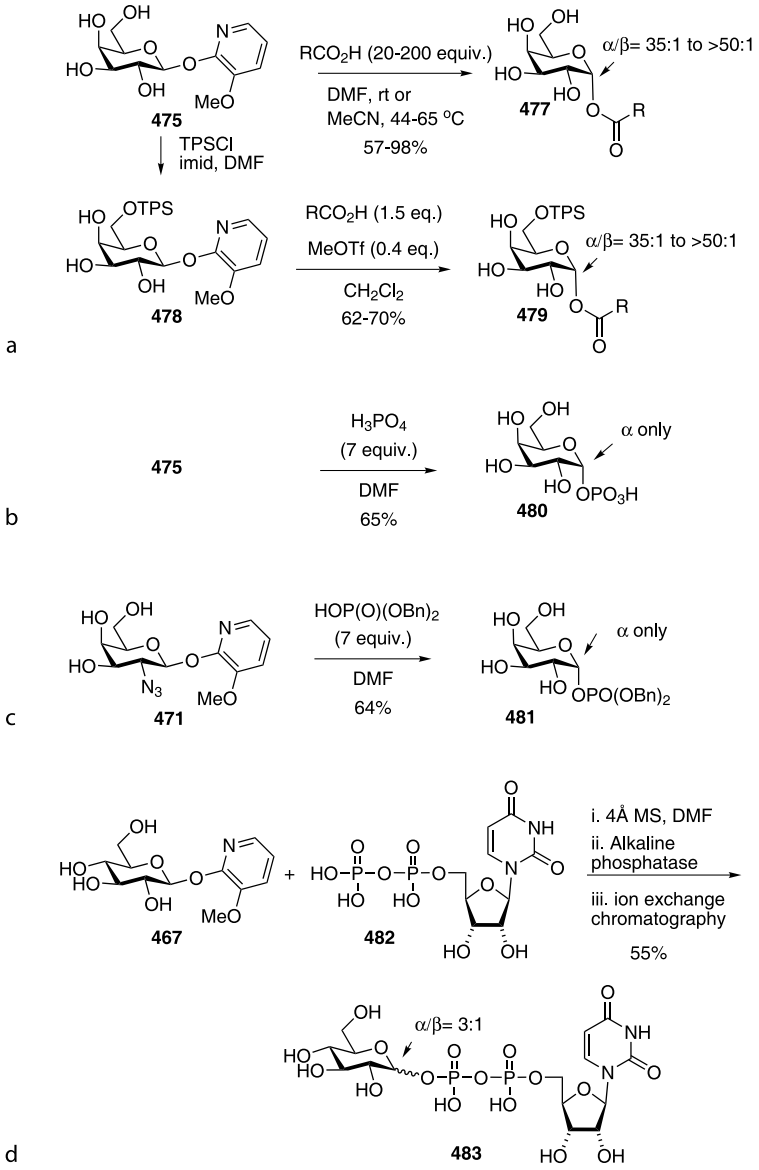


Scheme 79
Disaccharide synthesis with unprotected MOP donors





Scheme 80
Selective activation of unprotected- versus protected-MOP glycosides



Scheme 81
Stereocontrolled synthesis of glycosyl 1-carboxylates and 1-phosphates

7 equiv. of phosphoric acid or dibenzyl phosphate in DMF led to the corresponding α-glycosyl phosphates **480** and **481**, respectively. The same donors are also capable of transferring glucopyranosyl (e. g. **467**, *Scheme 81d*) and galactopyranosyl units to UDP-free acid **482**, to afford the corresponding uridine 5′ diphosphosugars (e. g. **483**) in one step [229].

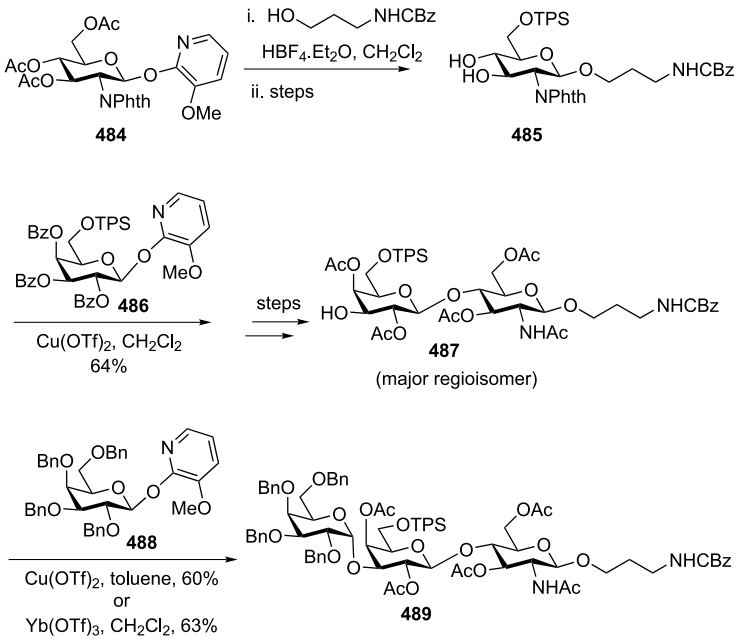
6.3.3 MOP Glycosides in Oligosaccharide Synthesis

A solid-phase oligosaccharide synthesis based on the MOP donor/acceptor methodology was developed by Hanessian and co-workers [224]. Thus, an *O*-unprotected polymer-phase bound MOP donor is coupled with an excess of a partially esterified MOP acceptor. Selective removal of the ester (or related protecting groups) from the new saccharides generates a new *O*-unprotected MOP donor to engage in a subsequent iteration.

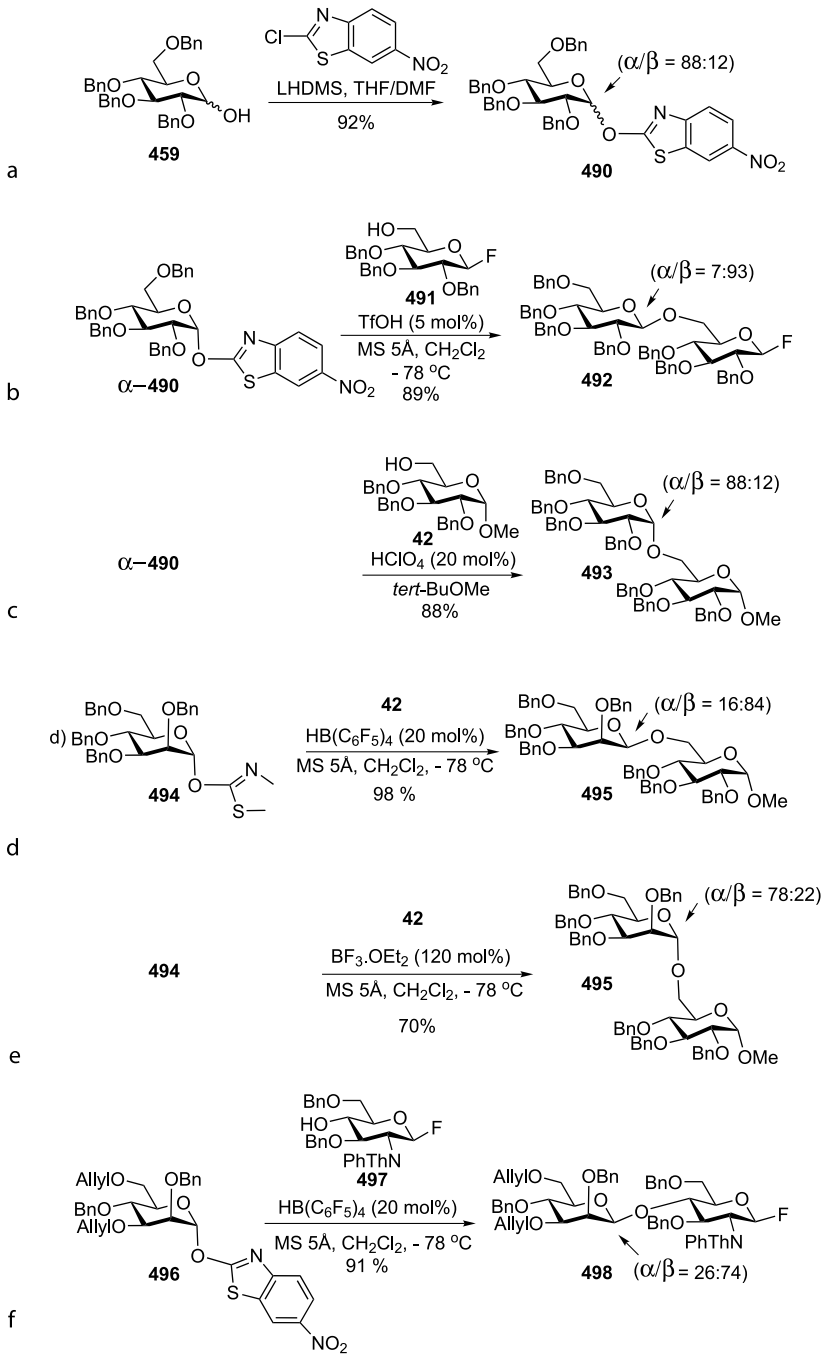
Hanessian and co-workers also illustrated the usefulness of the MOP methodology with some syntheses of oligosaccharides [230,231]. They reported a concise synthesis of a Galα1→3Galβ1→4GlcNAcOR trisaccharide **489**, outlined in ❯ *Scheme 82* [232]. Treatment of MOP donor **484** with 3-benzyloxycarbonylamino 1-propanol in the presence of $HBF_4{\cdot}Et_2O$ in CH_2Cl_2 led to the expected β-glycoside **485**. Protecting group manipulation and glycosylation with MOP galactopyranosyl **486** in the presence of $Cu(TfO)_2$ as activator gave the intended β-disaccharide **487**. Final glycosylation of **487** with β-galacto MOP donor **488** ($Cu(OTf)_2$) or $Yb(OTf)_3$, as promoters) led to protected trisaccharide **489**.

6.4 6-Nitro-2-benzothiazolyl Glycosides

Mukaiyama et al. described glycosyl 6-nitro-2-benzothiazoates (e. g. **490**) as useful glycosyl donors [233]. They are prepared by reaction of glycose derived hemiacetals (e. g. **459**) with 2-chloro-6-nitro-2-benzothiazoate (❯ *Scheme 83a*). The purified α-isomer **α-490**, react-



❯ Scheme 82
Synthesis of a trisaccharide using MOP glycosyl donors



Scheme 83
Glycosyl 6-nitro-2-benzothiazoate as a glycosyl donor

ed with primary hydroxyl acceptors in the presence of catalytic TfOH at −78 °C to give mainly β-glucosides, e. g. **492** (Scheme 83*b*). Although, a highly stereoselective α-glucosylation (α/β = 88:12) was carried out in high yield using 20 mol% of $HClO_4$ in *tert*-BuOMe (Scheme 83*c*). 6-Nitro-2-benzothiazolyl α-mannosides (e. g. **494**) effected stereoselective β-mannosylation with several glycosyl acceptors [234,235]. The highest β-stereoselectivity was achieved when tetrakis(pentafluorophenyl)boric acid [$HB(C_6F_5)_4$] [236] was employed as catalyst (Scheme 83*d*). $BF_3{\cdot}Et_2O$, a weaker Lewis acid, showed a reversed stereoselectivity [237] (Scheme 83*e*). The β-selective coupling was employed by Mukaiyama and co-workers in the formation of the β-Man(1→4)GlcN linkage, e. g. **498**, that exists in *N*-linked glycans (Scheme 83*f*) [238,239].

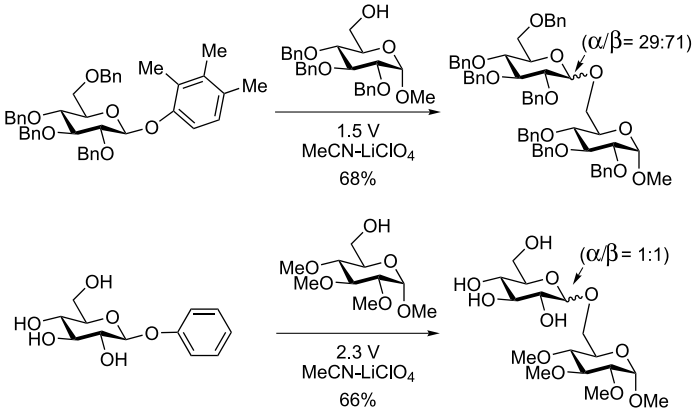
7 Miscellaneous *O*-Glycosyl Donors

Noyori and Kurimoto [240] described that hydroxyl-protected and -unprotected glycosyl aryloxides reacted with alcohols under mild electrolytic conditions to give the corresponding glycosides. They hypothesized that the glycosylation reaction proceeded via oxocarbenium ion intermediates generated from the radical cation of the easily oxidizable aryloxy substrate (Scheme 84).

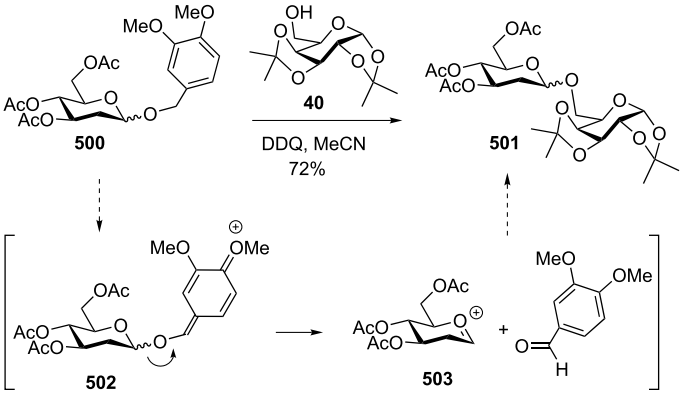
A combination of trimethylsilyl bromide and zinc triflate promoted the glycosylation of benzyl-, isopropyl-, and methyl glycosides with several glycosyl acceptors in moderate to good yields [241,242].

2-Deoxyglycosides, e. g. **501**, were obtained by DDQ oxidation of 3,4-dimethoxybenzyl glycosides **500**, in MeCN in the presence of primary, secondary, and tertiary alcohols (Scheme 85) [243].

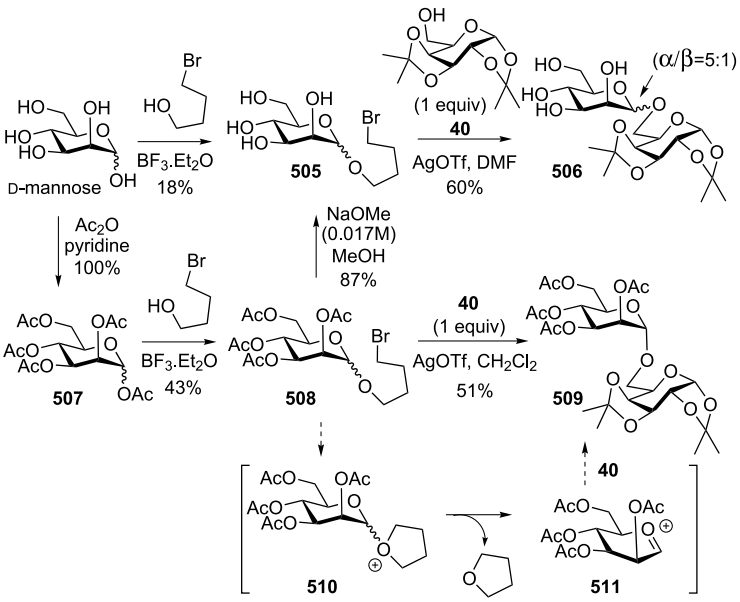
Davis and co-workers [244] examined the self-activating properties of unprotected and acetylated bromobutyl glycosides **505** and **508**, respectively (Scheme 86). These readily available compounds reacted with galactose acceptor **40** (1 equiv.) in the presence of a halophilic Lewis



Scheme 84
Electrochemical glycosylation of glycosyl aryloxides



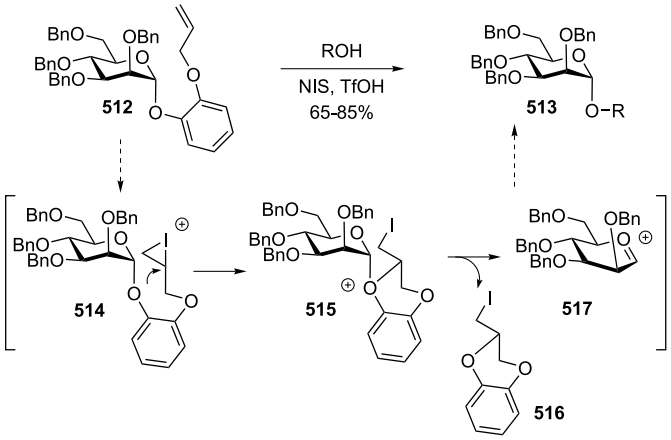
Scheme 85
3,4-Dimethoxybenzyl 2-deoxy glycosides as glycosyl donors



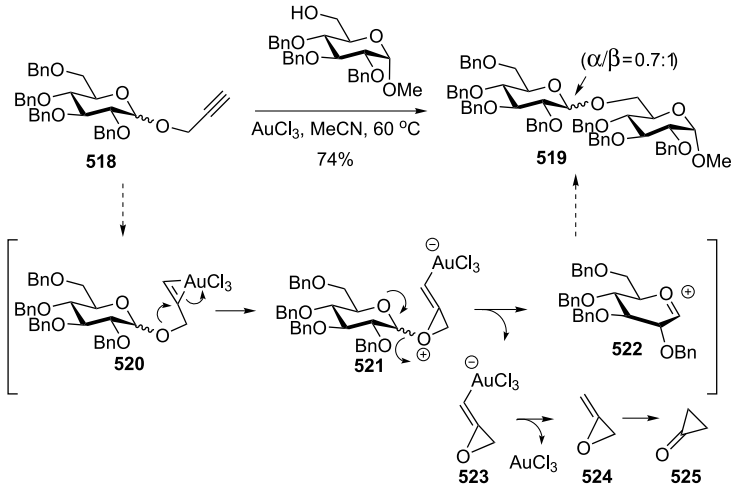
Scheme 86
Bromobutyl glycosides as glycosyl donors

acid promoter (AgOTg) to give disaccharides **506** and **509** in moderate yields. The suggested reaction pathway involved a spontaneous, or acid, triggered, 5-*exo-tet* cyclization of the bromobutyl glycoside, e. g. **508→510**, to form an anomeric furanosyl cation **510**, which would evolve to give non-nucleophilic volatile tetrahydrofuran, and oxocarbenium ion **511**. The latter will then react with acceptor **40** to furnish the disaccharide.

Hung and co-workers have reported on the use of 2-allyloxyphenyl mannoside **512** as a useful glycosyl donor [245]. Mannoside **512** reacted in the presence of NIS/TfOH in CH_2Cl_2 at room



Scheme 87
2-Allyloxyphenyl mannosides as glycosyl donors



Scheme 88
Propargyl glycosides as glycosyl donors

temperature, with a series of primary and secondary hydroxyl acceptors to give α-mannosides, e. g. **513**, in good yields (*Scheme 87*). The proposed mechanism for the formation of the oxocarbenium ion **517**, outlined in *Scheme 87*, implies a 6-*exo-tet* cyclization on halonium ion **514**, and the ejection of the non-nucleophilic species **516**.

Hotha and Kashyap have identified propargyl glycosides **518**, as new glycosyl donors (*Scheme 88*) [246,247]. Various aglycones, including primary and secondary alcohols, reacted with propargyl glycosides in the presence of 3 mol% of $AuCl_3$ in MeCN at 60 °C, to give α/β-mixtures of glycosides and disaccharides in good yields. The α,β-ratio of the

transglycosylation products was found to be independent of the anomeric ratio of the donor. per-*O*-Acylated propargyl glycosides did not give transglycosylation products. A possible reaction pathway to the generation of an intermediate oxocarbenium ion, based on the alkynophilicity of gold catalysts, was advanced by the authors (› *Scheme 88*). Coordination of $AuCl_3$ to the glycosyl donor **518** would be followed by formation of the cyclopropyl gold carbene intermediate (**520**) that could evolve to intermediate **521**, which would lead to oxocarbenium ion **522**, and alkenyl gold complex **523**. The latter upon protodemetalation will generate $AuCl_3$ and cyclopropanone **525** via intermediate **524**.

Acknowledgements

The author thanks the *Dirección General de Enseñanza Superior* (Grant CTQ2006-15279-C03-02) for financial support.

References

1. Paulsen H (1982) Angew Chem Int Ed Engl 21:155
2. Schmidt RR (1986) Angew Chem Int Ed Engl 25:212
3. Paulsen H (1990) Angew Chem Int Ed Engl 29:823
4. Toshima K, Tatsuta K (1993) Chem Rev 93:1503
5. Schmidt RR (1986) Angew Chem Int Ed Engl 25:212
6. Boons GJ (1996) Tetrahedron 52:1095
7. Boons GJ (1996) Contemp Org Synth 3:173
8. Davis BG (2000) J Chem Soc Perkin Trans 1 2137
9. Demchenko AV (2005) Lett Org Chem 2:580
10. Fraser-Reid B, Konradsson P, Mootoo DR, Udodong U (1988) J Chem Soc Chem Commun 823
11. Michael A (1879) Am Chem J 1:305
12. Fraser-Reid B, Udodong UE, Wu Z, Ottoson H, Merrit JR, Rao CS, Roberts C, Madsen R (1992) Synlett 927
13. Mootoo DR, Date V, Fraser-Reid B (1988) J Chem Soc Chem Commun 1462
14. Mootoo DR, Fraser-Reid B (1986) J Chem Soc Chem Commun 1570
15. Baldwin JE (1976) J Chem Soc Chem Commun 734
16. Mootoo DR, Date V, Fraser-Reid B (1988) J Am Chem Soc 110:2662
17. Lemieux RU, Morgan AR (1965) Can J Chem 43:2190
18. Goodman L (1967) Adv Carbohydr Chem Biochem 22:109
19. Mootoo DR, Konradsson P, Udodong U, Fraser-Reid B (1988) J Am Chem Soc 110: 5583
20. Konradsson P, Mootoo DR, McDevitt RE, Fraser-Reid B (1990) J Chem Soc Chem Commun 270
21. Veeneman GH, van Boom JH (1990) Tetrahedron Lett 31:275
22. Veeneman GH, van Leeuwen SH, van Boom JH (1990) Tetrahedron Lett 31:1331
23. Friesen RW, Danishefsky SJ (1989) J Am Chem Soc 111:6656
24. Barrena MI, Echarri R, Castillón S (1996) Synlett 675
25. Cheung MK, Douglas NL, Hinzen B, Ley SV, Pannecoucke X (1997) Synlett 257
26. Hashimoto S, Sakamoto H, Honda T, Ikegami S (1997) Tetrahedron Lett 38:5181
27. Hashimoto S, Sakamoto H, Honda T, Abe H, Nakamura S, Ikegami S (1997) Tetrahedron Lett 38:8969
28. Mukaiyama T, Chiba H, Funasaka S (2002) Chem Lett 392
29. Chiba H, Funasaka S, Kiyota K, Mukaiyama T (2002) Chem Lett 746
30. Chiba H, Funasaka S, Mukaiyama T (2003) Bull Chem Soc Jpn 76:1629
31. Demchenko AV, Kamat MN, De Meo C (2003) Synlett 1287
32. Kamat MN, Demchenko AV (2005) Org Lett 7:3215
33. Bongat AFG, Kamat MN, Demchenko AV (2007) J Org Chem 72:1480

34. Llera JM, López JC, Fraser-Reid B (1990) J Org Chem 55:2997
35. Madsen R, Fraser-Reid B (1995) J Org Chem 60:772
36. Fraser-Reid B, Merrit RJ, Handlon AJ, Andrews CW (1993) Pure Appl Chem 65:779
37. Rodebaugh R, Fraser-Reid B (1994) J Am Chem Soc 116:3155
38. Rodebaugh R, Fraser-Reid B (1996) Tetrahedron 52:7663
39. Brown RS (1997) Acc Chem Res 30:131
40. Fraser-Reid B, Wu Z, Udodong UE, Ottosson, H (1990) J Org Chem 55:6068
41. Fraser-Reid B, Wu Z, Andrews CW, Skowronski E, Bowen JP (1991) J Am Chem Soc 113: 1434
42. Merrit JR, Debenham JS, Fraser-Reid B (1996) J Carbohydr Chem 15:65
43. Rodebaugh R, Debenham JS, Fraser-Reid B, Snyder JP (1999) J Org Chem 64:1758
44. Konradsson P, Fraser-Reid B (1988) J Chem Soc Chem Commun 1124
45. Pale P, Whitesides GM (1991) J Org Chem 56:4547
46. Plante OJ, Andrade RB, Seeberger PH (1999) Org Lett 1:211
47. Hashimoto S, Honda T, Ikegami S (1999) J Chem Soc Chem Commun 685
48. Clausen MH, Madsen R (2003) Chem Eur J 9:3281
49. López JC, Uriel C, Guillamón-Martín A, Valverde S, Gómez AM (2007) Org Lett 9:2759
50. Fraser-Reid B, Lu J, Jayaprakash KN, López JC (2006) Tetrahedron Asymmetry 17:2449
51. Ferrier RJ, Hay RW, Vethaviyasar N (1973) Carbohydr Res 27:55
52. Fügedi P, Garegg PJ, Lönn H, Norberg T (1987) Glycoconjugate J 4:97
53. Codée JDC, Litjens REJN, van den Bos LJ, Overkleeft HS, van der Marel GA (2005) Chem Soc Rev 34:769
54. Schmidt RR, Michel J (1980) Angew Chem Int Ed Engl 19:731
55. Sharma GVM, Rao SM (1992) Tetrahedron Lett 33:2365
56. Ratcliffe AJ, Fraser-Reid B (1989) J Chem Soc Perkin Trans 1 1805
57. Ritter JJ, Minieri PP (1948) J Am Chem Soc 70:4045
58. Lemieux RU, Morgan AR (1965) Can J Chem 43:2205
59. Ratcliffe AJ, Fraser-Reid B (1990) J Chem Soc Perkin Trans 1 747
60. Ratcliffe AJ, Konradsson P, Fraser-Reid B (1990) J Am Chem Soc 112:5665
61. Ratcliffe AJ, Konradsson P, Fraser-Reid B (1991) Carbohydr Res 216:323
62. Handlon AL, Fraser-Reid B (1993) J Am Chem Soc 115:3796
63. Nair LG, Fraser-Reid B, Szardenings AK (2001) Org Lett 3:317
64. Banoub J, Boullanger P, Lafont D (1992) Chem Rev 92:1167
65. Mootoo DR, Fraser-Reid B (1989) Tetrahedron Lett 30:2363
66. Klaffke W, Warren CD, Jeanloz RW (1992) Carbohydr Res 244:171
67. Debenham JS, Fraser-Reid B (1994) XVIIth International Carbohydrate Symposium, Ottawa, Canada
68. Debenham JS, Madsen R, Roberts C, Fraser-Reid B (1995) J Am Chem Soc 117:3302
69. Debenham JS, Fraser-Reid B (1996) J Org Chem 61:432
70. Fraser-Reid B, Anilkumar G, Nair LG, Olsson L, García Martín M, Daniels JK (2000) Isr J Chem 40:255
71. Svarovsky SA, Barchi JJ (2003) Carbohydr Res 338:1925
72. Feather MS, Harris JF (1965) J Org Chem 30:153
73. Andrews CW, Fraser-Reid B, Bowen JP (1993) Involvement of $n\sigma^*$ interactions in glycoside cleavage. In: Thatcher GRJ (ed) The anomeric effect and associated stereoelectronic effects. ACS Symposium Series 539, American Chemical Society, Washington, DC, p 114
74. Deslongchamps P (1983) Stereoelectronic effects in organic chemistry. Pergamon Press, New York, pp 30–35
75. Ratcliffe AJ, Mootoo DR, Andrews CW, Fraser-Reid B (1989) J Am Chem Soc 111:7661
76. Lafont D, Boullanger P (2006) Tetrahedron Assymmetry 17:3368
77. Madsen R, Roberts C, Fraser-Reid B (1995) J Org Chem 60:7920
78. Bodner R, Marcellino BK, Severino A, Smenton AL, Rojas CM (2005) J Org Chem 70:3988
79. Demchenko, AV, De Meo C (2002) Tetrahedron Lett 43:8819
80. Chapeau M-C, Marnett LJ (1993) J Org Chem 58:7258
81. Arasappan A, Fraser-Reid B (1995) Tetrahedron Lett 36:7967
82. Konradsson P, Roberts C, Fraser-Reid B (1991) Recl Trav Chim Pays-Bas 110:23

83. Velty R, Benvegnu T, Plusquellec D (1996) Synlett 817
84. Ferrières V, Bertho J-N, Plusquellec D (1995) Tetrahedron Lett 36:2749
85. Brennan PJ, Nikaido H (1995) Annu Rev Biochem 64:29
86. Mereyala HB, Hotha S, Gurjar MK (1998) 685
87. Lu J, Fraser-Reid B (2004) Org Lett 6: 3051
88. Jayaprakash KN, Radhakrishnan KV, Fraser-Reid B (2002) Tetrahedron Lett 43:6955
89. Jayaprakash KN, Fraser-Reid B (2004) Synlett 301
90. Lu J, Fraser-Reid B (2005) Chem Commun 862
91. Fraser-Reid B, Lu J, Jayaprakash KN, López JC (2006) Tetrahedron Asymmetry 17:2449
92. Hölemann A, Stocket BL, Seeberger PH (2006) J Org Chem 71:8071
93. Barresi F, Hindsgaul O (1991) J Am Chem Soc 113:9337
94. Stork G, Kim G (1992) J Am Chem Soc 114:1087
95. Ito Y, Ogawa T (1994) Angew Chem Int Ed Engl 33:1765
96. Krog-Jensen C, Oscarson S (1996) J Org Chem 61:4512
97. Gelin M, Ferrières V, Lefeuvre M, Plusquellec D (2003) Eur J Org Chem 1285
98. Hay JE, Haynes LJ (1958) J Chem Soc 2231
99. Rousseau C, Martin OR (2000) Tetrahedron Asymmetry 11:409
100. Girard N, Rousseau C, Martin OR (2003) Tetrahedron Lett 44:8971
101. Ikeda K, Fukuyo J, Sato K, Sato M (2005) Chem Pharm Bull 53:1490
102. Tabeur C, Machetto F, Mallet J-M, Duchaussoy P, Petitou M, Sinaÿ P (1996) Carbohydr Res 281:253
103. Reichardt N-C, Martín-Lomas M (2005) Arkivoc ix:133
104. Anilkumar G, Nair LG, Fraser-Reid B (2000) Org Lett 2587–2589
105. Fraser-Reid B, López JC, Radhakrishnan KV, Mach M, Schlueter U, Gómez AM, Uriel C (2002) J Am Chem Soc 124:3198
106. Fraser-Reid B, Anilkumar GN, Nair LG, Radhakrishnan KV, López JC, Gómez AM, Uriel C (2002) Aust J Chem 55:123
107. Fraser-Reid B, López JC, Radhakrishnan KV, Mach M, Schlueter U, Gómez AM, Uriel C (2002) Can J Chem 80:1075
108. López JC, Gómez AM, Uriel C, Fraser-Reid B (2003) Tetrahedron Lett 44:1417
109. Fraser-Reid B, López JC, Gómez AM, Uriel C (2004) Eur J Org Chem 1387
110. Fraser-Reid B, López JC, Radhakrishnan KV, Nandakumar MV, Gómez AM, Uriel C (2002) Chem Común 2104
111. Wilson BG, Fraser-Reid B (1995) J Org Chem 60:317
112. Fraser-Reid B, Grimme S, Piacenza M, Mach M, Schlueter U (2003) Chem Eur J 9:4687
113. Uriel C, Gómez AM, López JC, Fraser-Reid B (2003) Synlett 2203
114. Mootoo DR, Konradsson P, Fraser-Reid B (1989) J Am Chem Soc 111:8540
115. Koenigs W, Knorr E (1901) Ber 34:957
116. Merritt JR, Fraser-Reid B (1992) J Am Chem Soc 114:8334
117. Merritt JR, Naisang E, Fraser-Reid B (1994) J Org Chem 59:4443
118. Lerouge P (1994) Glycobiology 4:127
119. Debenham JS, Rodebaugh R, Fraser-Reid B (1996) J Org Chem 61:6478
120. Rodebaugh R, Debenham JS, Fraser-Reid B (1996) Tetrahedron Lett 37:5477
121. Park MH, Takeda R, Nakanishi K (1987) Tetrahedron Lett 28:3823
122. Ferguson MAJ (1991) Curr Opin Struct Biol 1:52
123. Ferguson MAJ, Low MG, Cross GAM, (1985) J Biol Chem 260:14547
124. Ferguson MAJ, Homans SW, Dwek RA, Rademacher TW (1988) Science 239:753
125. Homans SW, Ferguson MAJ, Dwek RA, Rademacher TW, Anand R, Williams AF (1988) Nature 333:269
126. Campbell AS, Fraser-Reid B (1995) J Am Chem Soc 117:10387
127. Roberts C, Madsen R, Fraser-Reid B (1995) J Am Chem Soc 117:1546
128. Madsen R, Udodong UE, Roberts C, Mootoo DR, Konradsson P, Fraser-Reid B (1995) J Am Chem Soc 117:1554
129. Udodong UE, Madsen R, Roberts C, Fraser-Reid B (1993) J Am Chem Soc 115:7886
130. Bridges AJ (1996) Chemtracts Org Chem 9:215
131. Heidelberg T, Martin OR (2004) J Org Chem 69:2290
132. Clausen MH, Madsen R (2004) Carbohydr Res 339:2159
133. Clausen MH, Jørgensen MR, Thorsen J, Madsen R (2001) 543
134. Lemieux RU, Hendricks KB, Stick RV, James K (1975) J Am Chem Soc 97:4056

135. Arasappan A, Fraser-Reid B 1996 J Org Chem 61:2401
136. Hinou H, Umino A, Matsuoka K, Terunuma D, Takahashi S, Esumi Y, Kuzuhara H (2000) Bull Chem Soc Jpn 73:163
137. Yoshida T, Chiba T, Yokochi T, Onozaki K, Sugiyama T, Nakashima I (2001) Carbohydr Res 335:167
138. Seeberger PH, Haase WH (2000) Chem Rev 100:4349
139. Rodebaugh R, Fraser-Reid B, Geysen HM (1997) Tetrahedron Lett 44:7653
140. Rich DH, Gurwara SK (1975) J Am Chem Soc 97:1575
141. Rodebaugh R, Joshi S, Fraser-Reid B, Geysen HM (1997) J Org Chem 62:5660
142. Andrade RB, Plante OJ, Melean LG, Seeberger PH (1999) Org Lett 1:1811
143. Melean LG, Haase WC, Seeberger PH (2000) Tetrahedron Lett 41:4329
144. Hewitt MC, Seeberger PH (2001) J Org Chem 66:4233
145. Ratner DM, Plante OJ, Seeberger PH (2002) Eur J Org Chem 826
146. Hewitt MC, Zinder DA, Seeberger PH (2002) J Am Chem Soc 124:13434
147. Palmacci ER, Plante OJ, Hewitt MC, Seeberger PH (2003) Helv Chim Acta 86:3975
148. Mogemark M, Gustafsson L, Bengtsson C, Elofsson M, Kihlberg J (2004) Org Lett 6:4885
149. Buskas T, Söderberg E, Konradsson P, Fraser-Reid B (2000) J Org Chem 65:958
150. Bindschädler P, Noti C, Castagnetti E, Seeberger PH (2006) Helv Chim Acta 89:2591
151. Allen JR, Allen JG, Zhang X-F, Williams LJ, Zatorski A, Ragupathi G, Livingston PO, Danishefsky SJ (2000) Chem Eur J 6:1366
152. Allen JR, Danishefsky SJ (2000) J Prakt Chem 342:736
153. Allen JR, Ragupathi G, Livingston PO, Danishefsky SJ (1999) J Am Chem Soc 121:10875
154. Allen JR, Harris CR, Danishefsky SJ (2001) J Am Chem Soc 123:1890
155. Nishimura SI, Matsuoka K, Furuike T, Ishii S, Kurita K, Nishimura KM (1991) Macromolecules 24:4236
156. Nishimura SI, Furuike T, Matsuoka K, Maruyama K, Nagata K, Kurita K, Nishi N, Tokura S (1994) Macromolecules 27:4876
157. Miyagawa A, Kurosawa H, Watanabe T, Koyama T, Terunuma D, Matsuoka K (2004) Carbohydr Polymers 57:441
158. Yamada A, Hatano K, Koyama T, Matsuoka K, Esumi Y, Terunuma D (2006) Carbohydr Res 341:467
159. Rele SM, Iyer SS, Baskaran S, Chaikof EL (2004) J Org Chem 69:9159
160. Lu J, Fraser-Reid B, Chowda G (2005) Org Lett 7:3841
161. Fischer E (1893) Ber 26:2400
162. Fischer E (1895) Ber 28:1145
163. Lemieux RU, Morgan AR (1965) Can J Chem 43:2199
164. Lu J, Jayaprakash KN, Schlueter U, Fraser-Reid B (2004) J Am Chem Soc 126:7540
165. Mach M, Schlueter U, Mathew F, Fraser-Reid B, Hazen KC (2002) Tetrahedron 58:7345
166. Jayaprakash KN, Fraser-Reid B (2004) Org Lett 6:4211
167. de Raadt A, Ferrier RJ (1991) Carbohydr Res 216:93
168. Vankar YD, Vankar PS, Behrendt M, Schmidt RR (1991) Tetrahedron 47:9985
169. Marra A, Esnault J, Veyrières A, Sinaÿ P (1992) J Am Chem Soc 114:6354
170. Tebbe FN, Parshall GW, Reddy GS (1978) J Am Chem Soc 100:3611
171. Chenault HK, Castro A (1994) Tetrahedron Lett 35:9145
172. Chenault HK, Castro A, Chafin LF, Yang Y (1996) J Org Chem 61:5024
173. Lutsenko IF, Khomutov RV (1955) Doklady Akad Nauk SSSR 102:97
174. Petasis NA, Bzowej EI (1990) J Am Chem Soc 112:6392
175. Kunz H, Harreus A (1982) Liebigs Ann Chem 41
176. Boons GJ, Isles S (1994) Tetrahedron Lett 35:3593
177. Roy R, Andersson FO, Letellier M (2002) Tetrahedron Lett 33:6053
178. Boons GJ, Burton A, Isles S (1996) J Chem Soc Chem Commun 141
179. Boons GJ, Isles S (1996) J Org Chem 61:4262
180. Boons GJ, Heskamp B, Hout F (1996) Angew Chem Int Ed Engl 35:2845
181. Johnson M, Aries C, Boons GJ (1998) Tetrahedron Lett 39:9801
182. Bai Y, Boons GJ, Burton A, Johnson M, Haller M (2000) J Carbohydr Chem 19:939
183. Haller M, Boons GJ (2001) J Chem Soc Perkin Trans 1 814
184. Boons GJ, Burton A, Wyatt P (1996) Synlett 310

185. Capozzi G, Dios A, Franck RW, Geer A, Marzabadi C, Menichetti S, Nativi C, Tamarez M (1996) Angew Chem Int Ed Engl 35:777
186. Marzabadi CH, Franck RW (1996) Chem Commun 2651
187. Marzabadi CH, Franck RW (2000) Tetrahedron 56:8385
188. Capozzi G, Mannocci F, Menichetti S, Nativi C, Paoletti S (1997) Chem Commun 2291
189. Bartolozzi A, Capozzi G, Menichetti S, Nativi C (2001) Eur J Org Chem 2083
190. Petersen L, Jensen KJ (2001) J Org Chem 66:6268
191. Jensen KJ (2002) J Chem Soc Perkin Trans 1 2219
192. Koeners HJ, de Kok AJ, Romers C, van Boom JH (1980) Recl Trav Chim Pays-Bas 99:355
193. Sharma SK, Corrales G, Penadés S (1995) Tetrahedron Lett 36:5627
194. Petersen L, Jensen KJ (2001) J Chem Soc Perkin Trans 1 2175
195. Larsen K, Worm-Leonhard K, Olsen P, Hoel A, Jensen KJ (2005) Org Biomol Chem 3:3966
196. Tolborg JF, Jensen KJ (2000) Chem Commun 147
197. Laursen JB, Petersen L, Jensen KJ (2001) Org Lett 3:687
198. Petersen L, Laursen JB, Larsen K, Motawia MS, Jensen KJ (2003) Org Lett 5:1309
199. Laursen JB, Petersen L, Jensen KJ, Nielsen J (2003) Org Biomol Chem 1:3147
200. Grathe S, Thygessen MB, Larsen K, Petersen L, Jensen KJ (2005) Tetrahedron Asymmetry 16:1439
201. Kim KS, Lee YJ, Kim HY, Kang SS, Kwon SY (2001) J Am Chem Soc 123:8477
202. Kim KS, Lee ME, Cho JW (2004) Bull Korean Chem Soc 25:139
203. Kim KS, Jeon HB (2007) In: Demchenko AV (ed) Frontiers in modern carbohydrate chemistry. ACS Symposium Series, vol 960. ACS, Washington, DC, Ch 9
204. Demchenko AV (2003) Curr. Org. Chem. 7:35
205. Crich D, Sun S (1997) J Am Chem Soc 119:11217
206. Crich D, Sun S (1996) J Org Chem 61:4506
207. Crich D, Sun S (1997) J Org Chem 62:1198
208. Crich D, Sun S (1998) Tetrahedron 54:8321
209. Crich D, Smith M (2000) Org Lett 2:4067
210. Kim KS, Park J, Lee YJ, Seo YS (2003) Angew Chem Int Ed 42:459
211. Lee YJ, Lee K, Jung EH, Jeon HB, Kim KS (2005) Org Lett 7:3263
212. Lee R, Jeon JM, Jung JH, Jeon HB, Kim KS (2006) Can J Chem 84:506
213. Kunz H, Waldmann H (1984) Angew Chem Int Ed Engl 23:71
214. Kwon YT, Lee YJ, Lee K, Kim KS (2004) Org Lett 6:3901
215. Kim KS, Kang SS, Seo YS, Kim HJ, Lee YJ, Jeong KS (2003) Synlett 1311
216. Lee BY, Baek JY, Jeon HB, Kim KS (2007) Bull Korean Chem Soc 28:257
217. Lee YJ, Lee BY, Jeon HB, Kim KS (2006) Org Lett 8:3971
218. Lee YJ, Baek JY, Lee BY, Kang SS, Park HS, Jeon HB, Kim KS (2006) Carbohydr Res 341:1708
219. Nikolaev AV, Kochetkov NK (1986) Izv Akad Nauk SSSR Ser Khim 2556
220. Huchel U, Schmidt C, Schmidt RR (1995) Tetrahedron Lett 36:9457
221. Huchel U, Schmidt C, Schmidt RR (1998) Eur J Org Chem 1353
222. Hanessian S (1997) In: Hanessian S (ed) Preparative carbohydrate chemistry. Marcel Dekker Inc., New York, p 381
223. Hanessian S, Bacquet C, Lehong N (1980) Carbohydr Res 80:C17
224. Hanessian S, Lou B (2000) Chem Rev 100:4443
225. Lou B, Reddy GV, Wang H, Hanessian S (1997) In: Hanessian S (ed) Preparative carbohydrate chemistry. Marcel Dekker Inc., New York, p 389
226. Hanessian S, Condé JJ, Lou B (1995) Tetrahedron Lett 36:5865
227. Lou B, Huynh HK, Hanessian S (1997) In: Hanessian S (ed) Preparative carbohydrate chemistry. Marcel Dekker Inc., New York, p 413
228. Hanessian S, Mascitti V, Lu PP, Ishida H (2002) Synthesis 1959
229. Hanessian S, Lu PP, Ishida H (1998) J Am Chem Soc 120:13296
230. Lou B, Eckhardt E, Hanessian S (1997) In: Hanessian S (ed) Preparative carbohydrate chemistry, Marcel Dekker Inc., New York, p 449
231. Hanessian S, Huynh HK, Reddy GV, Duthaler RO, Katopodis A, Streiff MB, Kinzy W, Oehrlein R (2001) Tetrahedron 57:3281
232. Hanessian S, Saavedra OM, Mascitti V, Marterer W, Oehrlein R, Mak CP (2001) Tetrahedron 57:3267
233. Mukaiyama T, Hashihayata T, Mandai H (2003) Chem Lett 32:340

234. Hashihayata T, Mandai H, Mukaiyama T (2003) Chem Lett 32:442
235. Hashihayata T, Mukaiyama T (2003) Heterocycles 61:51
236. Jona H, Mandai H, Chavasri W, Takeuchi K, Mukaiyama T (2002) Bull Chem Soc Jpn 75:291
237. Hashihayata T, Mandai H, Mukaiyama T (2004) Bull Chem Soc Jpn 77:169
238. Mandai H, Mukaiyama T (2005) Chem Lett 34:702
239. Mandai H, Mukaiyama T (2006) Bull Chem Soc Jpn 79:479
240. Noyori R, Kurimoto I (1986) J Org Chem 51:4320
241. Higashi K, Nakayama K, Shioya E, Kusama T (1992) Chem Pharm Bull 40:1042
242. Higashi, K, Susaki H (1992) Chem Pharm Bull 40:2019
243. Inanaga J, Yokoyama Y, Hanamoto T (1993) Chem Lett 85
244. Davis BG, Word, SD, Maughan MAT (2002) Can J Chem 80:555
245. Lee JC, Pan GR, Kulkarni SS, Luo SY, Liao CC, Hung SC (2006) Tetrahedron Lett 47:1621
246. Hotha S, Kashyap S (2006) J Am Chem Soc 128:9620
247. Hotha S, Kashyap S (2006) J Am Chem Soc 128:17153

3.5 *S*-Glycosylation

Stefan Oscarson
Department of Organic Chemistry, Arrhenius Laboratory,
Stockholm University, Stockholm 106 91, Sweden
stefan.oscarson@ucd.ie

In: *Glycoscience*. Fraser-Reid B, Tatsuta K, Thiem J (eds)
Chapter-DOI 10-1007/978-3-540-30429-6_14:

Abstract

Published methods to synthesize thioglycosides have been reviewed. Two major pathways are recognized, either a displacement of an anomeric leaving group with a thiol acceptor to give the thioglycoside directly, or an initial formation of an anomeric thiol, which then is alkylated to give the thioglycoside. For the direct approach a listing is made based on the diverse anomeric leaving groups that have been utilized, acetates and halides being the most common. For the second pathway, various ways to form an anomeric thiol in a stereoselective manner are summarized, hydrolysis of anomeric thioacetates or pseudothiouronium salts being the most frequent, as well as different methodologies to perform the subsequent alkylation or arylation. Some miscellaneous pathways to thioglycosides, mainly through rearrangement of special *O*-glycosyl derivatives, are also discussed. In the last part an overview of the use of the discussed methodologies in the synthesis of thiooligosaccharides, i. e. oligosaccharides with interglycosidic thiolinkages, is presented.

Keywords

Thioglycosylation; Thioglycosides; Thiosaccharides; Synthesis

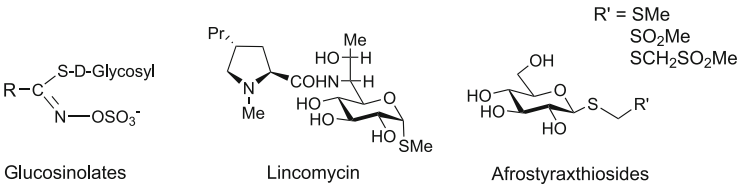
Abbreviations

AIBN azobisisobutyronitrile
CAN ceric ammonium nitrate
DMF dimethylformamide
DTT dithiothreitol
HMPA hexamethylphosphoramide
IDCP iodonium dicollidine perchlorate
Kdo 3-deoxy-D-*manno*-2-octulosonic acid
mCPBA *meta*-chloroperoxybenzoic acid
NIS *N*-iodosuccinimide
OTMS octadecyltrimethoxysilane
TFA trifluoroacetic acid
TMSO *N*-trimethylsilyl-2-oxazolidinone
UDP uridine diphosphate

1 Introduction

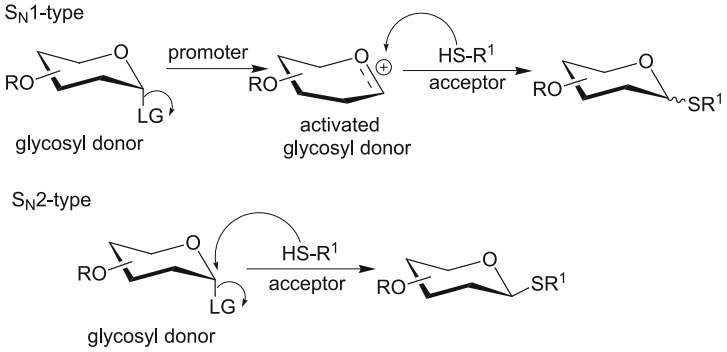
Carbohydrate derivatives with anomeric sulfur are not very common in nature, the only examples found are the various glucosinolates (“mustard oil glycosides”) mainly from the Brassicaceae family [1,2,3], the simple alkyl thioglycosides of lincomycin and structurally relat-

ed antibiotics found in *Streptomyces* species [4,5,6] and quite recently the methylthiomethyl and methylsulfonylmethyl thioglucosides from the Huaceae family [7] (● *Fig. 1*). However, there is an increasing interest in the use of thioglycosides, and other derivatives with anomeric sulfur, as efficient glycosyl donors, as ligands in chiral catalysts [8,9] and as stable *O*-glycoside analogues to be used in various biological context, e. g., as glycosidase inhibitors or vaccines [10], and, thus, their efficient syntheses are of great importance. The chemistry of thiosugars [1], their use as glycosyl donors [11,12,13,14], and the synthesis and biological application of thiooligosaccharides [15,16,17,18] have been reviewed extensively.



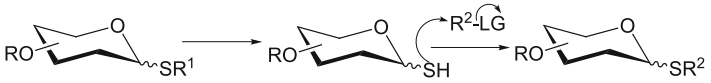
● Figure 1
Naturally occurring thioglycosides

Two general approaches to thioglycosides are used: either a direct introduction of a mercaptan through a displacement reaction of an anomeric leaving group, sometimes aided by a promoter, in a way similar to an *O*-glycosylation reaction (● *Scheme 1*), or a two (or more)-step procedure in which a non-thiol anomeric thio group is first introduced as above and then cleaved or rearranged in different fashions to give either directly a thioglycoside or an anomeric thiol or thiolate, which is then reacted with an alkyl electrophile to give the target thioglycoside (● *Scheme 2*). In this chapter the synthesis of thiomono- and thiooligosaccharides using these general pathways and also some other miscellaneous methods are discussed.



● Scheme 1

Two other types of glycosides worth mentioning in this context are glycosyl sulfoxides and selenoglycosides. Glycosyl sulfoxides were introduced as glycosyl donors by Kahne [19] and are often reviewed together with thioglycosides [11,12,14,20]. They are prepared by oxidation of the corresponding thioglycosides. Various reagents have been used, the most common



R^1 = e.g., Ac, CN, C(S)OMe

Scheme 2

being *m*CPBA [19,21], also chiral glycosyl sulfoxides have been synthesized [22]. The interest in selenoglycosides, i. e., the equivalent of thioglycosides but with a selenium atom instead of a sulfur atom, has also increased recently, partly because they as well have been found to be good glycosyl donors, but with a different and complementary reactivity as compared to thioglycosides [23,24,25]. Thus, selenoglycosides can be activated in the presence of thioglycosides giving the possibility to build up thioglycoside building block donors. Also, selenoglycosides have proven to be very effective in helping and simplifying the phasing of protein-carbohydrate complex crystals in X-ray crystallography [26,27]. The syntheses of selenoglycosides follow the same general routes as are discussed for thioglycosides in this chapter [24,28]. Other thiosugars apart from thioglycosides (1-thiosugars), i. e., with the sulfur atom substituting another oxygen atom than the anomeric one and especially the ring oxygen, are also of interest as carbohydrate mimics [29]. 2-, 3-, 4- and 6-thiosugars are often synthesized as precursors (acceptors) for the synthesis of oligosaccharides containing a sulfur atom in the interglycosidic linkage (see *Sect. 3.1*).

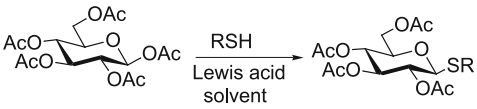
2 Synthesis of Monothioglycosides

In the synthesis of *S*-monosaccharides both the above-mentioned pathways (*Scheme 1*, *Scheme 2*) have been used extensively. The first step is the same in both routes: the displacement of an anomeric leaving group by a thionucleophile. Various leaving groups as well as thionucleophiles have been used. In *Sect. 2.1* is a listing organized by the nature of the leaving group. The emphasis will be on reactions with thiols and derivatives thereof as acceptors to give thioglycosides directly, but examples of other thio compounds, e. g., thioacetic acid, as acceptors are also included. The formation of the latter compounds and their transformation into proper thioglycosides, i. e., with a thiol aglycone part, will be predominantly discussed in *Sect. 2.2*.

2.1 Direct Synthesis

2.1.1 From Anomeric Acetates

An anomeric acetate is efficiently displaced by a thiol under the influence of an acidic catalyst, and this is probably presently the most used and efficient way to produce thioglycosides of simple mercaptans, especially on a large scale. The standard procedure is to react a peracetylated aldose dissolved in dichloromethane with a slight excess of thiol using a hard Lewis acid as promoter, which generally gives a high yield of mainly the 1,2-*trans* product (*Scheme 3*) [11].



Scheme 3

Early studies employing hydrogen chloride as catalyst on perbenzoylated sugars showed complicated product mixtures [1,30], while Lewis acid-catalyzed mercaptolysis of peracetylated monosaccharides afforded the corresponding thioglycosides in good yields [31,32,33]. However, from a preparative point of view (smell, work-up), the present use of a solvent and only a small excess of thiol is definitely an advantage.

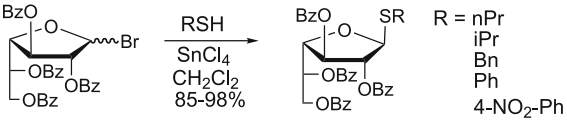
Various Lewis acids have been employed, most commonly BF_3-etherate [34,35], but also zinc chloride, stannic chloride, ferric chloride [36], zirconium chloride [37] and titanium chloride/indium chloride [38], with similar results. Recently 18 different oxometallic species were investigated as catalysts, and of these MoO_2Cl_2 (3 mol%) was found to be the most efficient affording thioglycosides in 75–94% yield [39].

β-Acetates react faster than α-acetates, which sometimes are difficult to induce to react at all [31,40]. With simple mercaptans the reaction can be performed almost stereospecifically to give the 1,2-*trans* product due to neighboring group participation, minor 1,2-*cis* impurities are often possible to remove by crystallization. However, if less reactive mercaptans [41,42,43,44] or sugar precursors (e. g., α-acetates or glucuronic acids [45,46]) are used, and, thus, prolonged reaction times have to be employed, more of the 1,2-*cis*-α product is generally formed, because of an anomerization of the kinetically favored β-thioglycoside to the more thermodynamically stable α-glycoside under the reaction conditions [47]. High stereoselectivity is also observed in the synthesis of thioglycosides of Kdo [48,49] and neuraminic acid [50] according to this approach, in spite of their 3-deoxy functions and thus lack of neighboring group participation. Both sugars yield preferentially the β-thioglycoside.

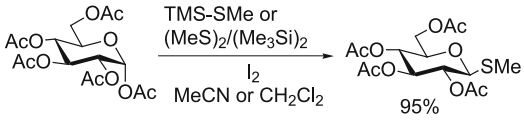
The approach works equally well with thio reagents other than mercaptans, e. g., thioacetic acid [37], thiourea [51] or *O*,*O*-dialkylphosphorodithioic acid derivatives [52], to yield the corresponding derivatives with an anomeric sulfur in high yields and stereoselectivity.

With the more reactive furanosides the per-benzoylated derivatives can also effectively be used as precursors (*Scheme 4*) [53].

With more elaborate protecting group patterns, introduced to allow consecutive oligosaccharide synthesis, optimal conditions for the specific transformation often have to be worked out. Obviously, the conditions used are not compatible with acid-labile groups. Problems encountered might be cleavage of ether protecting groups, e. g. benzyl groups [54]. With the levulinoyl



Scheme 4



Scheme 5

protecting group, the keto function is transformed into a dithioacetal as expected, but this can be chemoselectively hydrolyzed in the presence of the thioglycoside by following treatment with $AgNO_3/Ag_2O$ [55]. However, as mentioned above, in most oligosaccharide synthetic pathways the thioglycoside is formed first using the per-acetylated derivative and subsequently the desired elaborate protecting group pattern is introduced.

One-pot syntheses, via per-acetylated intermediates, of acetylated thioglycosides from unprotected reducing sugars have been reported. Either the same (BF_3-Et_2O [56]) or two different Lewis acids ($Cu(OTf)_2$ or $LiClO_4$ followed by BF_3-Et_2O [57,58]) were used for the initial acetylation and the succeeding phenylthioglycosylation step. No solvent other than Ac_2O was necessary and the yields were in the 70–75% range. One-pot synthesis via acetylated halide sugar intermediates has also been described (see ❯ *Sect. 2.1.2*).

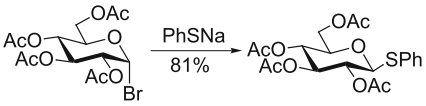
Variants of the method are the use of tributylstannyl [59] or trimethylsilyl [60] derivatives of the thiol, which give less odor and sometimes higher yields, due to less formation of 1,2-*cis*-products and dithioacetals, but these are also more expensive reagents. With the latter type of reagents α-acetates have also successfully been used as precursors using either zinc iodide [61] or iodine [62] as promoter. In the latter publication the authors could not detect any difference in reactivity between the α- and the β-acetate precursor under these reaction conditions. Furthermore, a cheaper version utilizing dimethyl sulfide/hexamethyldisilane as reagent was found to be equally efficient (❯ *Scheme 5*).

Also perbenzylated anomeric acetates have been employed as precursors using the stannylated thiol approach to give good yields of products, but with, as expected, low stereoselectivity [54,63]. A more preferred way to perbenzylated thioglycosides is to first construct the peracetylated thioglycoside stereospecifically and then change the acetyl groups for benzyl groups.

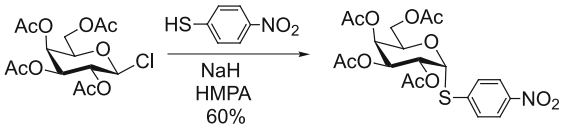
Modern laboratory techniques have been used to further improve the efficacy of these reactions. Thus, the use of standard conditions (peracetylated monosaccharides, thiophenol, BF_3-Et_2O, CH_2Cl_2) in combination with sonication gave the target 1,2-*trans*-thioglycosides in almost quantitative yield after less than 15 min, even when starting from the α-acetate [64]. With 2-acetamido-2-deoxy-sugars the reaction time was 2 h, but still only 1,2-*trans*-products were obtain in high yields (83–87%). Also microwave heating has been utilized but then with thiouronium salt precursors (see ❯ *Sect. 2.2.2*).

2.1.2 From Halogen Sugars

A classical route to thioglycosides is the reaction between an acetobromo sugar and a thiolate anion (❯ *Scheme 6*). The first thioglycoside ever synthesized was prepared by this method [65]. However, it is also a general method, and a good alternative to the use of peracetylated donors, and numerous thioglycosides, both alkyl and aryl have been prepared



Scheme 6



Scheme 7

accordingly [66,67,68]. Thioglycosides containing a heterocyclic aglycon, e. g. various nucleoside analogues, are preferentially synthesized from halogen sugars, although other precursors are also utilized [69]. The high nucleophilicity of the sulfur towards the anomeric position combined with its rather low basicity makes it possible to perform the reaction in acetone or methanol or even acetone-water. Since, especially with alkyl thiolates, de-*O*-acetylation is a frequent concomitant side-reaction, reacetylation prior to work-up is often necessary.

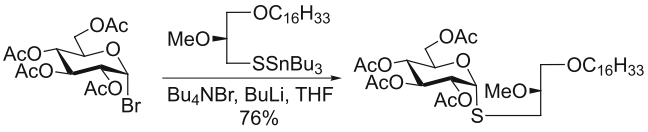
Usually a 1,2-*trans* product is obtained, possibly through participation of the 2-*O*-acyl group, but if the conditions are carefully selected a direct S_N2 displacement reaction can take place, to give, e. g., α-thioglucosides from β-halogen precursors or β-thiomannosides from α-halogen precursors, in spite of the presence of 2-*O*-participating groups. Thus, using an aprotic polar solvent (HMPA), the 1,2-*cis* *p*-nitrophenyl thioglycosides of mannose, galactose, and glucose were prepared (*Scheme 7*) [70,71].

Instead of thiolate anions, stannyl sulfides have been utilized, which necessitate elevated temperature during the formation of the thioglycoside, due to the lower reactivity of the stannyl derivatives. However, if a promoter ($SnCl_4$) is used, reactions at room temperature or below can be performed, this, though, also causes anomerization and α/β product mixtures [59].

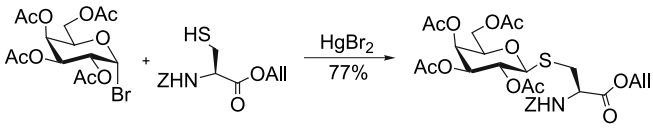
A variant is to use phase-transfer conditions, which allows the use of free thiols and non-polar solvents [72,73,74]. These conditions gave, starting from acetylated bromosugars, high yields of the corresponding 1,2-*trans*-thioglycoside both with various alkylthiols and phenylthiols, although ethyl mercaptan gave considerable deacetylation, which could be circumvented by using the perbenzoylated derivative instead. Interestingly, the β-chloroglucosyl donor gave exclusively the 1,2-thioorthoester under these conditions, which might indicate that an acyloxonium ion is not an intermediate in the reaction with α-bromo precursors [72]. If chlorinated solvents were used, competing displacement reactions with the solvents occurred, making toluene and ethyl acetate better choices as solvents [73].

In addition, thioglycosylations using halide-assisted conditions (*Scheme 8*) [75] and Koenigs–Knorr conditions (*Scheme 9*) [76], the latter both with stannyl sulfides and free thiols, have been described.

One-pot syntheses from unprotected reducing sugars via acetylated halide sugars have been described, too. Using iodine and iodine/HMDS as catalyst the iodosugar was prepared in situ and subsequently treated with dimethyldisulfide (compare *Scheme 5*) to give the acetylated



Scheme 8



Scheme 9

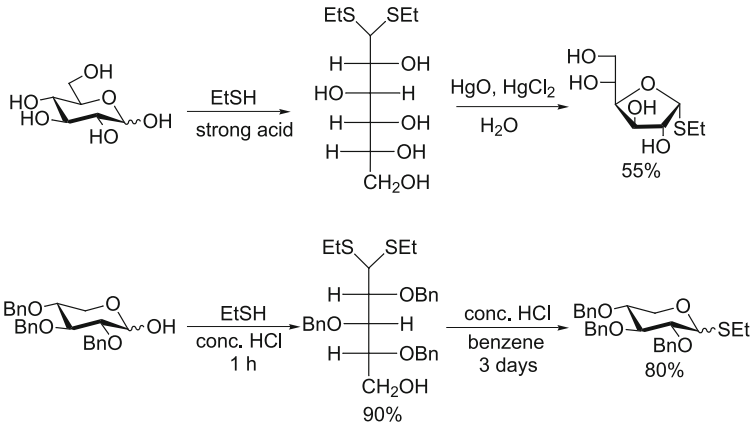
methyl thioglycoside in high yield (~80%) [77]. When HBr/HOAc were used to affect both acetylation and glycosyl bromide formation, the following thioglycosylation was performed under phase transfer conditions with various mercaptans (ethyl, phenyl, toluyl) to give the target thioglycosides in comparable high yields [78].

2.1.3 From Hemiacetals

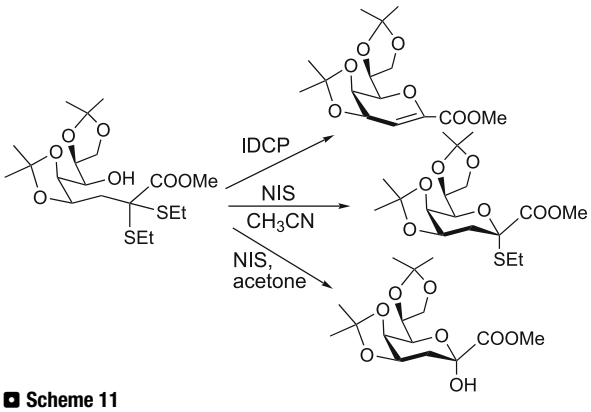
The problem when making thioglycosides from hemiacetals, either a free sugar or a protected derivative with a free anomeric position, is that the higher nucleophilicity of sulfur, as compared to oxygen, facilitates the consecutive reaction with an additional thiol to give the dithiomercaptal. Thus, under mercaptolytic conditions (thiol, strong acid), most free sugars directly form the open type dithioacetals, some of these (especially mannose) after standing and at elevated temperature then producing thioglycosides in low yields. The isolated dithioacetal can sometimes be transformed in fair yield to the thioglycoside, but usually an equilibrium is established containing both the dithioacetal and the furanosidic and pyranosidic thioglycosides (*Scheme 10*) [1]. However, with partially protected derivatives these ring-closures, promoted by standard thiophilic promoters (usually NIS), are often quite high-yielding and is a reaction now and then used as final step in de novo synthesis of thioglycoside monosaccharides (*Scheme 11* and *Scheme 12*) [79,80,81].

Another route to thioglycosides from hemiacetals involving a cyclization of an open-chain sulfide has been described recently. A Wittig–Horner reaction afforded the α,β-unsaturated thiophenyl derivative, which then was treated with iodonium reagents to induce cyclization to afford 2-iodo-1-thioglycosides, which are good precursors for 2-deoxy-sugars (*Scheme 13*) [167].

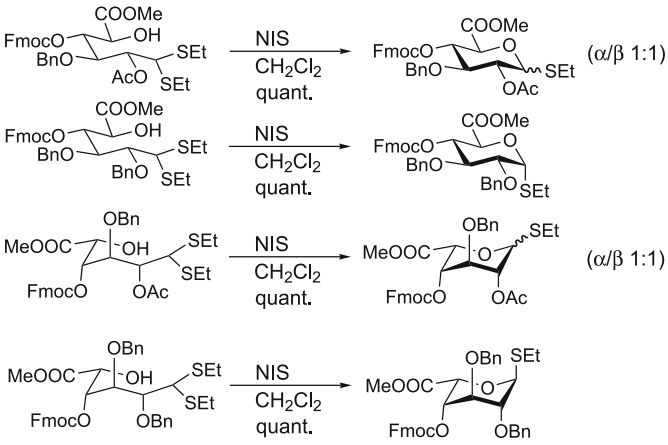
In addition, a number of methods to synthesize thioglycosides direct from hemiacetals (especially perbenzylated) have been reported. Hence, treatment of fully protected hemiacetal monosaccharides (both furanoses and pyranoses) with a trialkylphosphine and a diaryl disulfide gave high yields of the corresponding aryl thioglycosides, predominantly with converted stereochemistry at the anomeric center, indicating a S_N2 type of reaction (*Scheme 14*) [82,83,84].



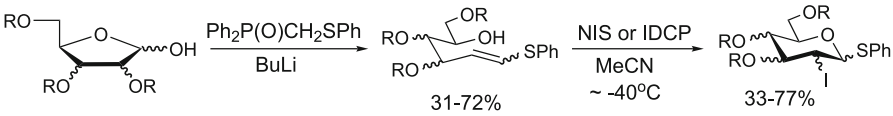
Scheme 10



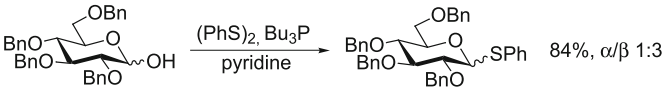
Scheme 11



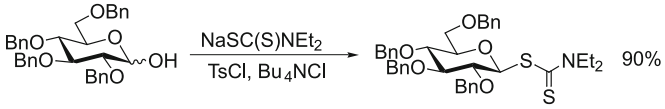
Scheme 12



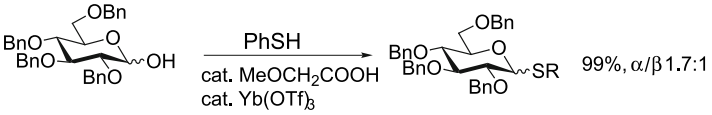
Scheme 13



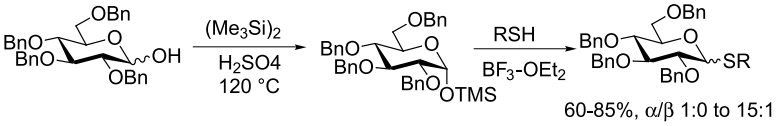
Scheme 14



Scheme 15



Scheme 16

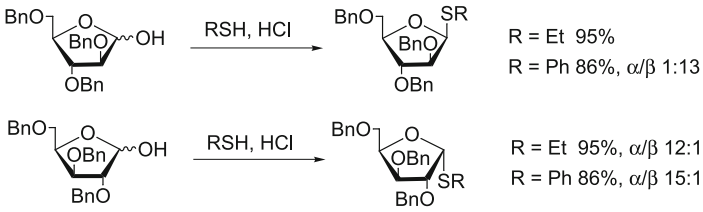


Scheme 17

The use of diphenyl phosphorochloridate or tosyl chloride and sodium *N*,*N*-diethyldithiocarbamates under phase-transfer conditions yields the *S*-glycosyl diethyldithiocarbamates through intermediate 1-*O*-diphenylphosphates or tosylates, respectively [85] (*Scheme 15*).

Similarly, perbenzylated glucopyranose and ribofuranose almost quantitatively yield the phenyl thioglycoside when reacted with thiophenol and catalytical amounts of methoxyacetic acid and ytterbium triflate (*Scheme 16*) [86].

An alternative is, of course, to convert the hemiacetals to isolable derivatives with better anomeric leaving groups, e. g., acetates (*Sect. 2.1.1*), halogens (*Sect. 2.1.2*), or trichloroacetimidates (*Sect. 2.1.4*). 1-*O*-Trimethylsilyl derivatives can also be smoothly converted to the corresponding thioglycoside through the reaction with a thiol and BF_3-etherate (*Scheme 17*) [83]. Since no participating groups were involved predominantly α-thioglycosides were produced.



Scheme 18

In special cases also the use of a strong proton acid as catalyst can result in an efficient synthesis of the thioglycoside. Thus, perbenzylated pentofuranoses, when treated with ethyl or phenyl mercaptan and a catalytic amount of HCl, gave predominantly β-linked thioglycosides in high yields (*Scheme 18*), whereas perbenzylated arabinopyranose (as did unprotected arabinose) with ethyl mercaptan gave exclusively the dithioacetal, which was slowly converted to the ethyl thioarabinopyranoside upon standing (see *Scheme 10*) [87]. Perbenzylated hexopyranoses (glucose and galactose) were inert under these conditions.

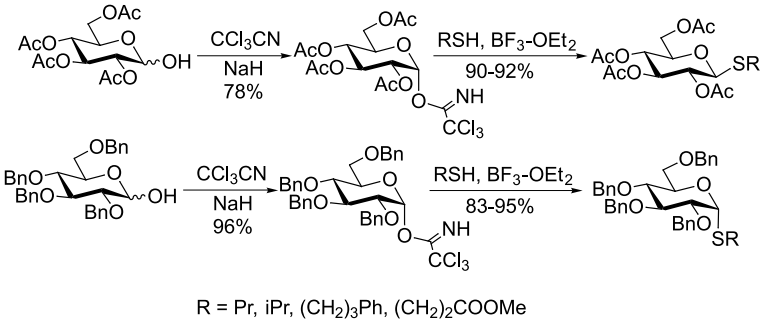
2.1.4 From Trichloroacetimidates

Since the activation of trichloroacetimidates involves hard Lewis acids like BF_3-etherate or TMS-triflate, promoters which usually do not activate thioglycosides, it is possible to effectively prepare thioglycosides from this type of donor [88].

Both acetylated and benzylated precursors have been employed. In the *gluco*-series using the α-imidates, an alkylthiol, and BF_3-activation, the acetylated derivative gave, as expected, the β-thioglycoside, whereas the benzylated donor gave, surprisingly, exclusively the α-linked product (*Scheme 19*). Not even at low temperatures (−42 °C) was any β-linked product observed, giving evidence that this reaction does not proceed through an S_N2-mechanism, which is the case using alcohol nucleophiles, and probably not through an S_N1-mechanism either, since this would give an α/β-mixture. The authors suggest a mechanism involving an intramolecular reaction of tight ion-pairs to give the retention of configuration. However, the use of a L-cysteine derivative as thionucleophile under identical coupling conditions gave a 5:1 α/β-mixture (*Scheme 20*) [89] (see also *Sect. 3.1.2*).

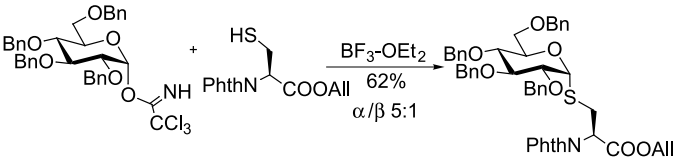
2.1.5 From Glycosides

Mercaptolysis (excess thiol (often ethyl mercaptan) in the presence of a strong proton acid) of a glycosidic linkage generally yields the open dithioacetal as the main product [1,90]. *O*-Glycosides can, however, be converted to *S*-glycosides in quite acceptable yields by treatment with a thiol or a thiotrimethylsilane and a Lewis acid. The conditions needed are still rather harsh and acid-sensitive protecting groups, including methyl (primary) and benzyl ethers, are often concomitantly cleaved [91,92,93]. Treatment of permethylated or persilylated methyl α-D-glucopyranoside with phenylthio(trimethyl) silane and zinc iodide gave the permethylated and the unprotected phenyl thioglycoside, respectively, as α/β-mixtures with the α-anomer predominating (*Scheme 21*) [91]. TM*S*-triflate has also been used as Lewis acid [92], as well as zinc bromide in the thiolysis of permethylated cyclodextrins [94].

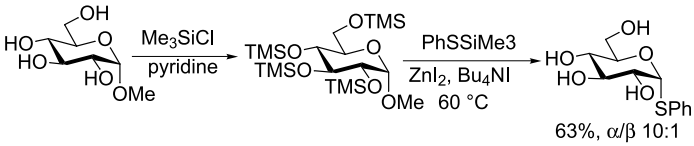


R = Pr, iPr, $(CH_2)_3Ph$, $(CH_2)_2COOMe$

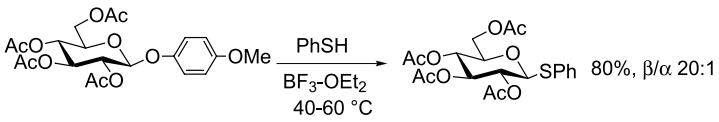
Scheme 19



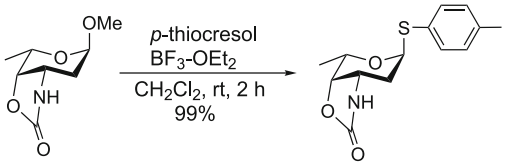
Scheme 20



Scheme 21



Scheme 22



Scheme 23

With a more acid-labile *O*-glycoside, the transformation into a thioglycoside can be performed more efficiently. Accordingly, variously protected *p*-methoxyphenyl mono- and disaccharides have been smoothly converted into the corresponding phenyl thioglycosides using thiophenol and BF_3-etherate (➤ *Scheme 22*) [95]. With a 2-participating group the β-selectivity of the reaction is high, but, as above, with a non-participating group in the 2-position the α-anomer heavily prevails in the product thioglycoside.

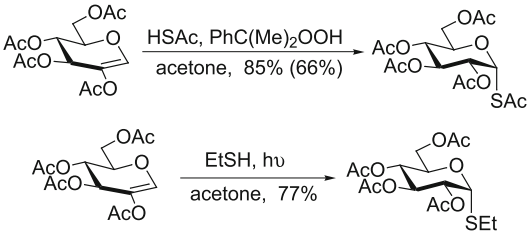
Deoxy functions in the sugar ring also increase the acid lability of the glycoside. Thus, was the methyl glycoside of daunosamine (2,3,6-trideoxy-3-amino-L-*lyxo*-hexopyranose) efficiently transformed into the corresponding aryl thioglycoside by treatment with a thiol and BF_3-etherate (➤ *Scheme 23*) [96]. The reaction was highly α-selective.

2.1.6 From Glycals

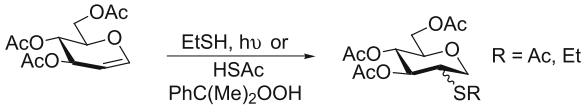
Radical addition to peracetylated 2-hydroxyglucals proceeds efficiently using either photochemical initiation and thiols or *t*-butyl or cumene peroxides and thioacetic acid to give the 1-thio-α-D-*gluco* derivative (➤ *Scheme 24*) [97,98].

Radical addition to peracetylated glycals, however, yields the opposite regioselectivity and poorer stereoselectivity and 2-thio-D-glucitols and mannitols are here the normal products (➤ *Scheme 25*) [97,99,100].

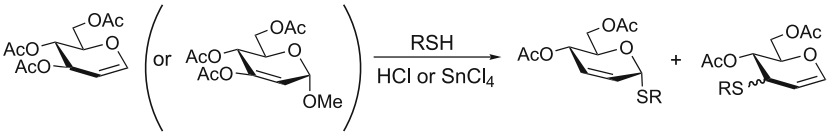
Acid catalyzed reactions of glycals with thiols on the other hand, involve an allylic displacement reaction and a product mixture of the 2,3-dideoxy-1-thioglycoside (Ferrier rearrangement product) and the 3-*S*-alkyl-3-thioglycal is obtained (➤ *Scheme 26*) [101,102]. The ratio between these products is dependent both on the substrate, the thio compound, and the acid used. The thioglycoside is the kinetic product and prolonged treatment with $SnCl_4$ gives a thermodynamic equilibrium where the 3-thioglycal is dominant (about 95%). Shorter reaction time and the use of, e. g., trimethylsilylated thiols, BF_3-etherate, $Sc(OTf)_3$ [103] or $ZrCl_4$ [104] as Lewis acid gave almost exclusively the 1-thio compound (➤ *Scheme 27*). With thiols (also trimethylsilylated) preferentially α-thioglycosides were formed, but (thionoace-



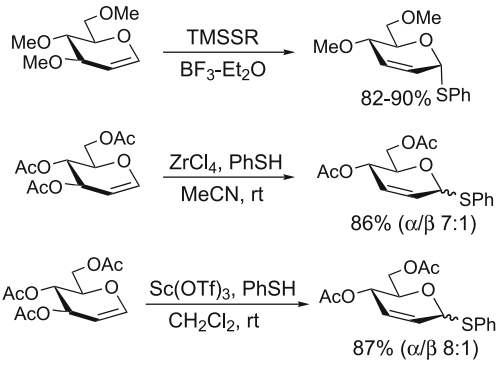
Scheme 24



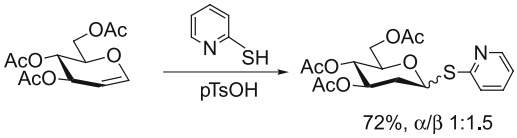
Scheme 25



Scheme 26



Scheme 27

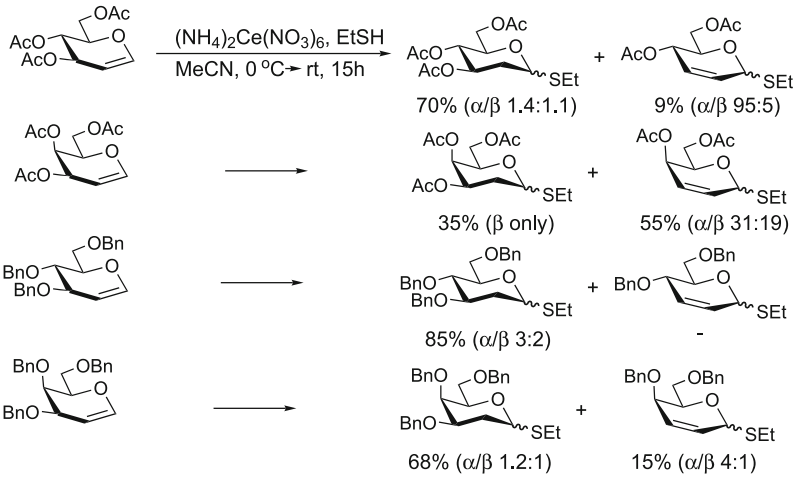


Scheme 28

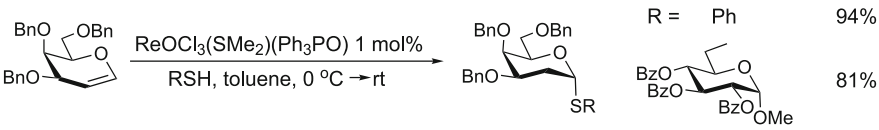
toxy)trimethylsilane ($Me_3SiOC(S)Me$) gave good β-selectivity, especially with the poorer OMe leaving group (as compared to OAc).

Under some conditions the formation of the intermediate allylic cation can be prevented and direct addition to the double bond is the dominant reaction to give 2-deoxy thioglycosides. The treatment of tri-*O*-acetylglucal (and -galactal and -rhamnal) with 2-mercaptopyridine and *p*-toluenesulfonic acid resulted in an addition reaction and gave the 2-deoxythioglycoside in good yield (*Scheme 28*) [105]. Use of BF_3-etherate as catalyst, however, gave the 2,3-dideoxy-1-thio-α-D-glycoside. Recently, use of CAN, a Re(V)-oxo complex, or $GaCl_3$ were reported to give good selectivity for the 2-deoxythioglycoside (*Scheme 29*, *Scheme 30* and *Scheme 31*) [106,107,108], although, as expected, the outcome is also dependent on the glycal and the thiol used.

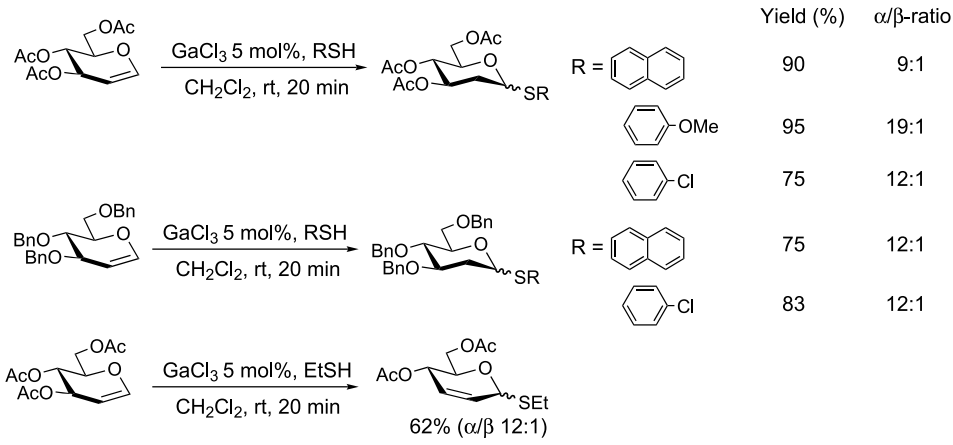
Thioglycosides from glycals can also be obtained in a two-step procedure via a 1,2-anhydro intermediate (*Sect. 2.1.7*).



Scheme 29



Scheme 30

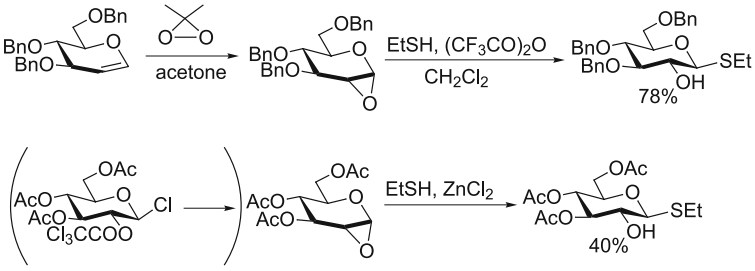


Scheme 31

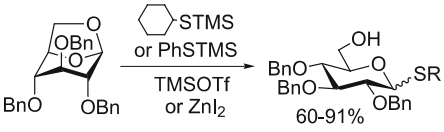
2.1.7 From Anhydro Sugars

Similar to glycosides, anhydro derivatives involving the anomeric oxygen, which can be considered as internal glycosides, can also be cleaved by a Lewis acid in the presence of a thiol, or a TMS-derivative thereof, to give thioglycosides. Examples with 1,2- and 1,6-anhydro

sugars have been published (Scheme 32 and Scheme 33). The 1,2-anhydro precursors, obtained, e. g., from glycals, gave exclusively the 1,2-*trans* products [109,110], whereas the 1,6-anhydro precursors gave 1,2-*trans* products with a 2-participating group, but α/β-mixtures otherwise [111,112,113]. The ratio obtained was dependent both on the anhydro derivative, the thiol, and the promoter used and the proportion between these. Starting from a *gluco*-derivative the α-anomer was always predominant, but the use of TMSOTf as promoter gave much higher α/β-ratios (about 20:1) than ZnI_2 (2–5/1) [112].

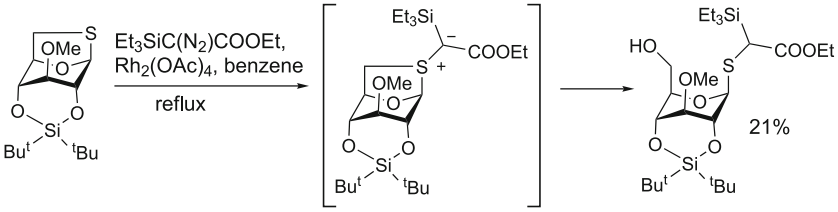


Scheme 32



Scheme 33

The corresponding thio derivative, 1,6-dideoxy-1,6-epithio sugars, are not good precursors for thioglycosides, since the sulfur normally ends up in the 6-position when the sulfur bridge is opened [114]. However, with soft nucleophiles, like thiols, attack on the 6-position is observed to afford 1,6-dithio-glycosides (compare Scheme 60, 3.2.1) [115]. Also, using a carbene-mediated reaction a thioglycoside was produced in low yield via hydrolysis of an intermediate sulfur ylid (Scheme 34) [116].



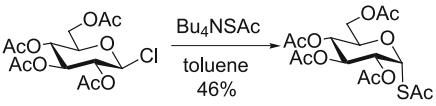
Scheme 34

2.2 Indirect Synthesis

The second main pathway to thioglycosides involves, as mentioned in the introduction, at least two steps: first the introduction of the anomeric sulfur using one of the ways described above, most often displacement of an anomeric halogen or acetate with a thionucleophile to give a 1-thio derivative, which then is directly transformed into a thioglycoside or cleaved to give a thioaldose or its thiolate salt, which finally is alkylated (*Scheme 2*). Recently it was also shown that anomeric thiols can be produced directly from protected hemiacetal saccharides by treatment with Lawesson's reagent [197]. Sometimes the uncovered thiol is protected, e. g., by a xanthenyl [117] or a trityl [118] group, to allow various manipulations prior to re-exposure of the thiol and eventual formation of a thioglycoside. The 1-thio precursors mainly employed are glycosyl thioacetates, thiouronium salts, thiocyanates, and xanthates. Their syntheses, as well as their transformations into thioglycosides, are discussed below.

2.2.1 Via Glycosyl Thioacetates

Glycosyl thioacetates are prepared from glycopyranosyl halides or acetates employing either thioacetic acid or potassium thioacetate as reagent under various conditions as discussed above [1,37,119]. The higher nucleophilicity of sulfur, as compared to oxygen, is evident since only 1-thioacetates are obtained. As with mercaptans under controlled conditions, a S_N2-displacement reaction can be performed, e. g., to give the α-product from a β-chloro sugar (*Scheme 35*) [70,120].

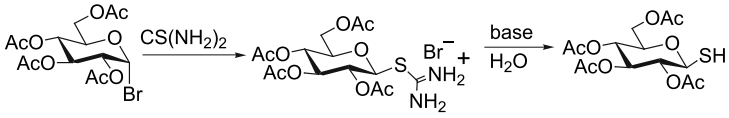


Scheme 35

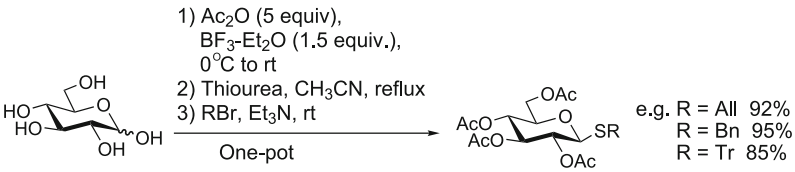
Chemoselective cleavage of anomeric thioacetates to the thioaldose in the presence of *O*-acetyl groups is possible under various conditions, e. g., methanolic sodium methoxide at low temperatures [121], phenylmercury acetate followed by demercuration [122], 2-aminoethanethiol [123], or hydrazinium acetate [124], and these methods have therefore often been used in the synthesis of thiooligosaccharides (*Sect. 3.2.1*). As mentioned, glycosyl thioacetates can also be prepared from the 2-acetoxyglycal derivative (see *Scheme 24*).

2.2.2 Via Glycosyl Pseudothiouronium Salts

These derivatives are also obtained by a displacement reaction of a sugar halide, this time with thiourea as nucleophile [125,126]. The higher nucleophilicity of sulfur as compared to nitrogen towards the (soft) electrophilic anomeric center ensures the sole formation of the 2-glycosyl-2-pseudothiourea salt (i. e., with the sulfur attached to the glycosyl moiety) which often can be isolated by direct crystallization. An advantage with these derivatives is their very mild conversion into thioaldoses (*Scheme 36*). Thus, the amidino group can be cleaved with potassium carbonate or pyrosulfite or by simply heating the bicarbonate salt avoiding de-*O*-acetylation, which is often a concomitant reaction in the fission of the other anomeric



Scheme 36



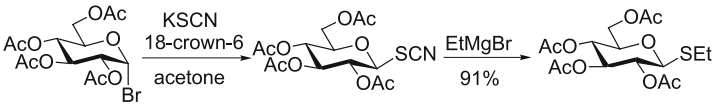
Scheme 37

thio groups discussed. This allows the easy and non-smelling formation of various acetylated alkyl thioglycosides on a large scale [125], a drawback is the use of the probably carcinogenic substances thiourea and alkyl iodide.

Recently, Et_3N/CH_3CN has been introduced as an efficient reagent for the conversion into the thiol. This procedure, furthermore, allows the in situ formation of an array of thioglycosides [127]. A one-pot protocol for the transformation of unprotected reducing sugars into thioglycosides via pseudothiouronium salts has been described [128]. In this protocol the thiouronium salt is formed directly from the in situ formed (BF_3-etherate/Ac_2O) peracetate by treatment with thiourea in refluxing acetonitrile (*Scheme 37*). Also, microwave heating has been used to accelerate the formation of the thiouronium salt as well as the subsequent transformations into free anomeric thiol and thioglycoside [129,130].

2.2.3 Via Glycosyl Thiocyanates

Glycosyl thiocyanates are prepared from the corresponding bromo sugar through a displacement reaction using alkali thiocyanates as nucleophiles [131,132] (*Scheme 38*). Peracylated bromo sugars give the 1,2-*trans* linked product, whereas perbenzylated derivatives in the *gluco*- and *galacto*-series give predominantly the 1,2-*cis* derivatives. However, if the reaction is carefully monitored the 1,2-*trans* product can be obtained in good yield even without a participating group [134]. A competing reaction is always the formation of the 1-isothiocyanates. If the silver [133,134] or lead salt [135], or potassium salt under phase-transfer [136] or solvent-free conditions [137] are used, the product will be the isothiocyanate, which also can be obtained by rearrangement of the thiocyanate by heating.



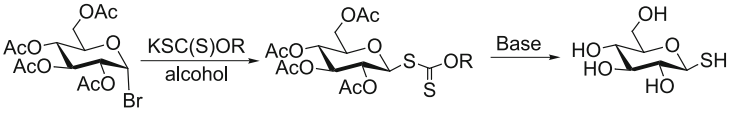
Scheme 38

The glycosyl thiocyanates can be transformed into thioglycosides through cleavage to the thioaldose by treatment with sodium methoxide under mild conditions followed by alkylation, or by a Grignard reaction which directly gives the thioglycoside (mannose-α-D-thiocyanate was inert) (● *Scheme 38*) [138].

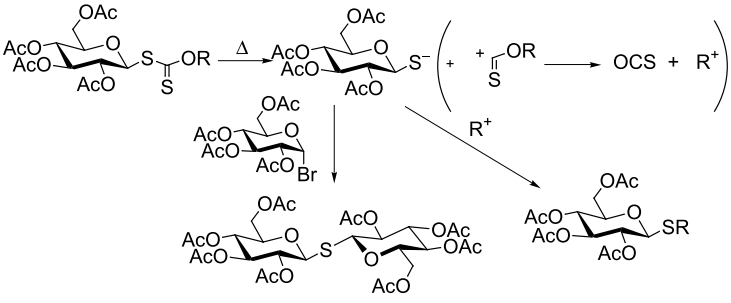
2.2.4 Via Glycosyl Xanthates

Glycosyl xanthates (dithiocarbonates) are synthesized by the reaction of an acylated glycopyranosyl halide with a potassium alkyl (or benzyl) xanthate in an alcohol [139] or using phase-transfer conditions [140,141] (● *Scheme 39*). Also, glycosyl alkyl *N*-xanthates (dithiocarbamates) have been produced utilizing the same procedure with a bromo sugar and a sodium alkyl *N*-xanthate in DMF [142]. The glycosyl xanthates are converted to 1-thioaldoses by treatment with a base, generally ammonia or sodium methoxide, to allow later alkylation to afford the thioglycosides.

However, if the primary formation of the xanthate is performed with acetone instead of alcohol as solvent and at elevated temperature, the initially formed glycosyl xanthate will react further to give the bis(glycoside) sulfide and the alkyl thioglycoside as main products [143]. Thus, under these conditions the xanthate is cleaved to give an alkoxythiocarbonyl cation and the glycosyl thiolate, which then can react with the bromo sugar, to give the bis(glycosyl) sulfide, or an alkyl cation, obtained through decomposition of the alkoxythiocarbonyl cation, which then produces the thioglycoside (● *Scheme 40*, compare also ● *Scheme 51*) [144]. This rearrangement is most effective in polar solvents and with an added salt, e. g., NaI, also thermal decomposition is effective but slightly slower. To avoid the formation of bis(glycosyl) sulfides, the xanthate can first be isolated and then rearranged to give high yields of thioglycosides [45,144].



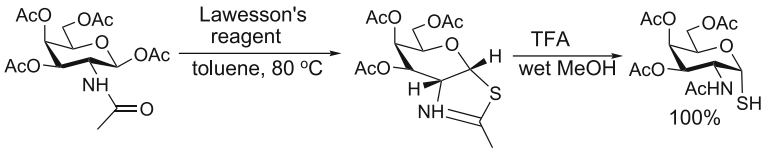
■ **Scheme 39**



■ **Scheme 40**

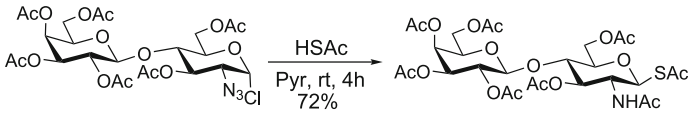
2.2.5 Via 1,2-Thiazoline Derivatives *(2-Acetamido Donors)*

Treatment of peracetylated 2-acetamido-β-D-gluco- or galacto-pyranose with the Lawesson's reagent gave the corresponding 1,2-thiazoline derivatives in quantitative yield (*Scheme 41*). Consecutive hydrolysis using TFA in wet methanol yielded the corresponding anomeric α-thiols, which were then acetylated to the α-thioacetate or alkylated in various ways (compare *Sect. 2.3* below) to give a number of different α-thioglycosides [145,146,147].



Scheme 41

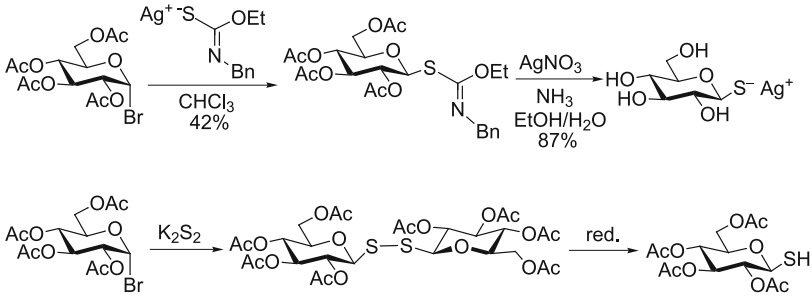
Another efficient way to 2-acetamido thiosugars is described from a 2-azido-1-chloro precursor. The β-thioacetate of 2-acetamido-2-deoxy-lactose was obtained by treatment of the 2-azido-2-deoxy α-chloride with a large excess of thioacetic acid, which both displaced the chloride and concomitantly reduced and acetylated the azido group (*Scheme 42*) [148].



Scheme 42

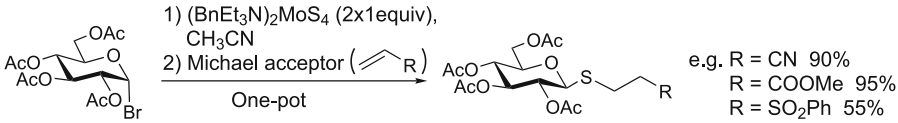
2.2.6 Via Glycosyl Thioimidocarbonates and Diglycosyl Disulfides

Other miscellaneous 1-thio derivatives convertible to thioaldoses are glycosyl thioimidocarbonates and diglycosyl disulfides [1]. Their syntheses are similar to the ones described above, i. e., displacement reactions of an acylated glycopyranosyl halide with a thionucleophile, an



Scheme 43

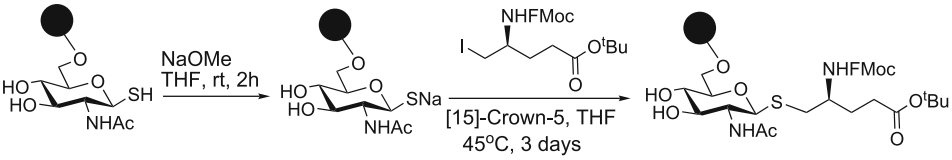
O-ethylbenzyl thioimidocarbonate, or potassium disulfide, respectively (● *Scheme 43*). Diglycosyl disulfides have also been prepared using tetrathiomolybdate as the sulfur transferring reagent [149]. The former product is saponified, whereas in the latter product the disulfide bridge is reduced to give the thioaldose. With the tetrathiomolybdate reagent this could be performed in a tandem sulfur transfer/reduction/Michael reaction to afford thioglycosides directly in a one-pot reaction (● *Scheme 44*) [150].



● **Scheme 44**

2.3 Alkylation of Thioaldoses

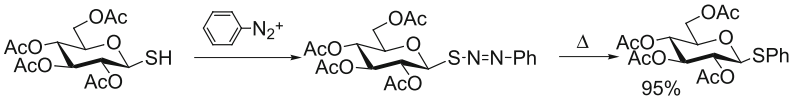
Thialdoses are easily *S*-alkylated to give the corresponding alkyl thioglycoside using alkyl iodides or bromides under basic conditions [1]. The pseudothiouronium salt derivative (● *Scheme 36*) can be used as precursor in a one-pot procedure, since the thioaldose is obtained in situ under these conditions [1,127]. Other leaving groups, like tosylates [151], have been used as well as a solid phase approach with the thioaldose attached to the resin (● *Scheme 45*, compare ● *Scheme 61*) [152]. Furthermore, thioalkylation using an alcohol and Mitsunobu conditions has been reported [153]. Arylations have been performed with aryl halides susceptible to nucleophilic displacement, e. g., 4-nitro- and 2,4-dinitrophenyl fluoride, in the presence of, e. g., potassium carbonate [127,154,155]. Also reactions with triazene substituted aryl iodides have been carried out using copper mediated arylation [156]. Unsubstituted iodobenzene did not react properly under the former conditions, but the thiophenyl glucoside can be obtained by refluxing the peracetylated thioaldose in DMF in the presence of phenylboronic acid and $Cu(OAc)_2$ [157].



● **Scheme 45**

2.3.1 Via Aryldiazonium Derivatives

Apart from direct arylation using aryl halides, aryl thioglycosides can be prepared most efficiently by treatment of the thioaldose with an aryldiazonium salt followed by thermal decomposition of the produced diazo derivative (● *Scheme 46*) [158]. The procedure is usually very

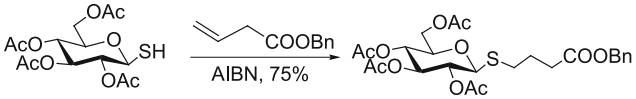


Scheme 46

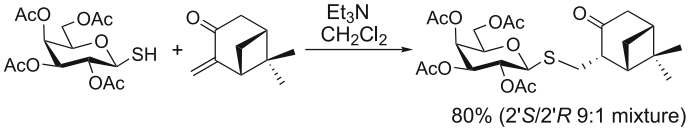
high-yielding and is for many aryl groups (e. g., not containing electron-withdrawing groups) superior to the direct arylation pathway.

2.3.2 By Addition to Alkenes

The well-known reaction of radical thiol addition to unsaturated compounds has been utilized to produce thioglycosides from thioaldoses. A number of alkenes, including bifunctional ones, were coupled to various acetylated thioaldoses using AIBN as radical initiator to give alkyl and spacer thioglycosides in good to high yields (50–93%) (*Scheme 47*) [159].
Addition of thiols to Michael acceptors is another classical reaction that has been used extensively to produce thioglycosides from the corresponding thioaldose nucleophile (*Scheme 44* and *Scheme 48*) [150,160]. Both monosaccharide and oligosaccharide thioglycosides (compare *Scheme 59*) have been prepared.



Scheme 47



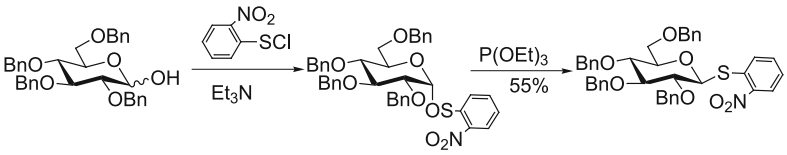
Scheme 48

2.4 Miscellaneous Methods

A few syntheses of thioglycosides by rearrangement of non-1-thio compounds have been reported.

2.4.1 Rearrangement of 1-Sulfenates

1-Sulfenates, prepared from the hemiacetal by treatment with sulfenyl chlorides and triethylamine, can be rearranged in the presence of a phosphorus(III) reagent, trialkyl phosphites or trialkylphosphines, to give the corresponding thioglycoside with inversed configuration (*Scheme 49*) [161,162].

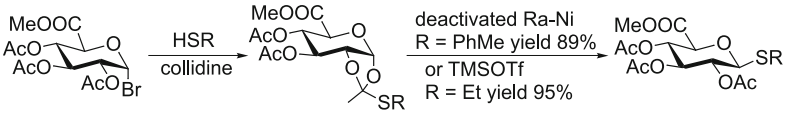


Scheme 49

The suggested mechanism involves an attack of the nucleophilic phosphorus reagent on the anomeric oxygen to afford an anomeric phosphonium ion and a thiolate counterion, which then attacks in a S_N2 reaction the anomeric position to give the thioglycoside and trialkyl phosphate or trialkylphosphine oxide.

2.4.2 Rearrangement of Thioorthoesters

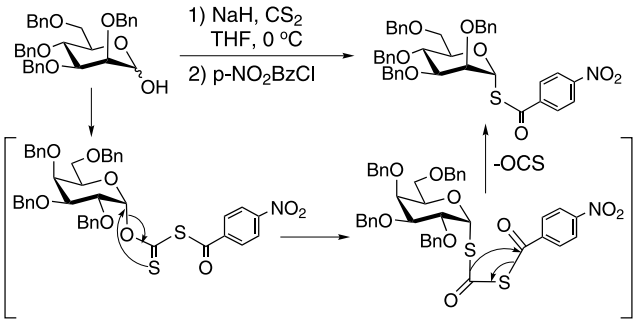
Thioorthoesters, easily prepared from the corresponding acetobromo sugar by treatment with a thiol and collidine, can be rearranged to the 1,2-*trans*-thioglycoside (*Scheme 50*). Deactivated Raney-nickel or TMSOTf have been used as promoters to give excellent yields of the thioglycosides of glucose, galactose, xylose, lactose, and glucuronic acid [163,164].



Scheme 50

2.4.3 Rearrangement of *p*-Nitrobenzoylxanthates

In a one-pot, two-step procedure anomeric thio *p*-nitrobenzoyl esters were formed from the corresponding free hemiacetal precursor via the rearrangement of in situ formed 1-*O*-(*S*-*p*-nitrobenzoyl) dithiocarbonate glycopyranose intermediates (*Scheme 51*) [165,166].



Scheme 51

The mechanism involves the rearrangement of the intermediate into the 1-*S*-(*S*-thio-*p*-nitrobenzoyl)dithiocarbonyl glycopyranose followed by a subsequent rearrangement with loss of carbon oxide sulphide to give the product (compare ❯ *Scheme 40*).

3 Synthesis of Thiooligosaccharides

As mentioned in the introduction, thiooligosaccharides (i. e., oligosaccharides in which at least one interglycosidic oxygen atom is substituted by a sulfur atom) are of great value as glycosylhydrolase stable analogues of *O*-glycosides and an increasing number of syntheses of complex thiooligosaccharides is continuously being reported [15,16,17,18,168].

The two general approaches to thiooligosaccharides are identical to the ones discussed above in the synthesis of thiomonosaccharides. Either a "normal" glycosylation is performed, but with a saccharide thiol acceptor, or the anomeric thio function is introduced first to yield an anomeric thiol or thiolate which then is reacted with a saccharide electrophile to form the inter-*S*-glycosidic linkage through a S_N2 displacement reaction. Both approaches are commonly used, but perhaps the second approach, where the anomeric conformation is more easily controlled, since it is constructed during the formation of the anomeric thiol function and not during the coupling reaction, has been the most abundant.

General problems, as compared to the formation of *O*-glycosides, are the incompatibility between catalytic hydrogenolysis and sulfur functions, which complicates the use of benzyl ethers as protecting groups, although Birch reduction might be an alternative, and the easy formation of disulfides from thiols, irrespective of if they are used as donors or acceptors.

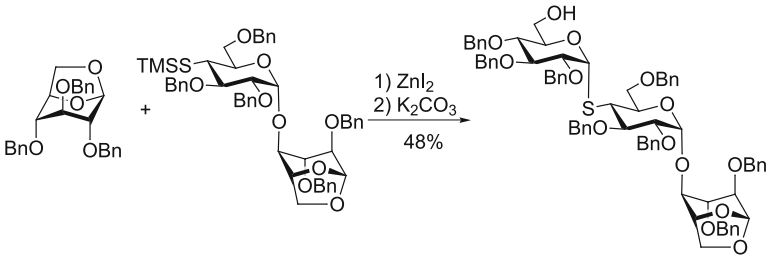
An interesting new possibility is to use enzymes in the synthesis of thiooligosaccharides, a methodology that has been made possible by the development of "thioglycoligases" by point mutation of glycosidases (See ❯ *Sect. 3.3*).

3.1 Glycosylations with Thiol Acceptors

Under this heading methods according to ❯ *Scheme 1* will be discussed, i. e., with thiol saccharide acceptors. As compared to the synthesis of monosaccharides along this pathway, the formation of thiooligosaccharides is complicated by the fact that now the acceptor moiety is often the most laborious to construct and thus cannot, if necessary, be used in large excess to effect the glycosidic bond formation. Only glycosylation methods using promoters that are not thiophilic can obviously be used, and this disqualifies many *O*-glycosylation methods. More or less only the trichloroacetimidate method has been generally applied together with a few examples utilizing anhydro sugars and glycals as donors. Also thioglycosylation reactions under basic conditions have been employed quite frequently.

3.1.1 Anhydro Sugars as Donors

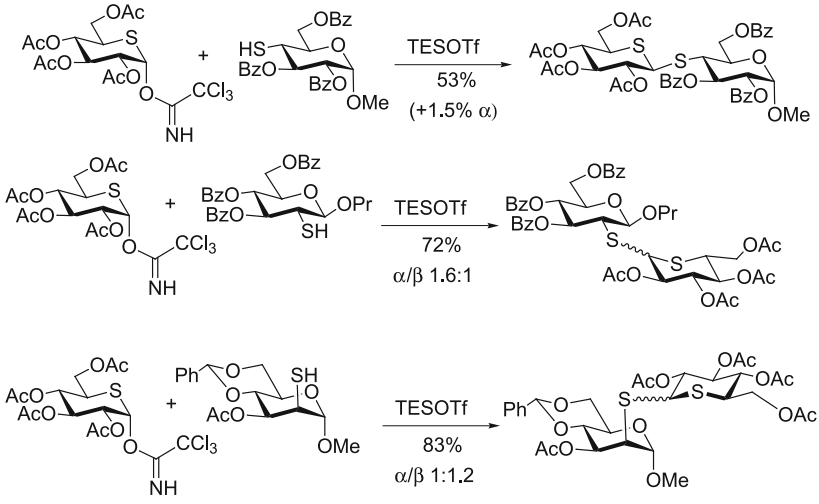
Benzylated 1,6-anhydroglucose has been used as donor in the construction of maltose and maltotriose structures with high α-selectivity (❯ *Scheme 52*) [112]. Interestingly, 1,6-anhydrodisaccharides were found to be unreactive as donors with 4-thiosaccharide acceptors, and could consequently be used as alternative acceptors.



Scheme 52

3.1.2 Trichloroacetimidates as Donors

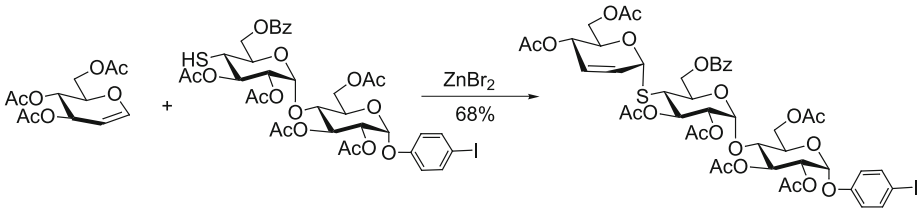
Trichloroacetimidates function well as donors also with saccharide thiol acceptors. A number of thiodisaccharides have accordingly been prepared using this method [169,170,171,172,173]. Possible by-products are various elimination products and orthoester formation. The stereochemical outcome is not easy to predict or control and, as so often in carbohydrate chemistry, an optimization of conditions is usually necessary for each glycosylation. For example, the use of an acetylated α-trichloroacetimidate glucose donor can yield not only an almost exclusively β-linked disaccharide product but also preferentially an α-linked product, depending on the acceptor and conditions used (*Scheme 53*) (see also *Sect. 3.3*).



Scheme 53

3.1.3 Glycals as Donors

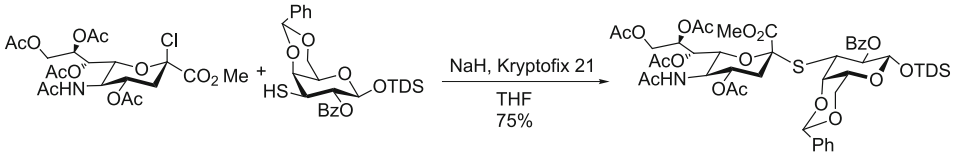
Addition reactions of saccharide thiols to glucals have been performed (*Scheme 54*, compare with *Scheme 26*). The α-linked thiooligosaccharide was the main product but the β-anomer and the product resulting from a displacement of the 3-*O*-acetate were also isolated (ratio 5 : 1 : 1) [174].



Scheme 54

3.1.4 Base-Promoted Glycosylations

Also the S_N2-displacement of an anomeric halogen by saccharide thiolates, even secondary ones, is quite feasible. Using this approach and α-acetobromo sugars for construction of β-linkages and β-acetochloro sugars for the construction of α-linkages, the thiodisaccharide analogues of gentiobiose, maltose, lactose, cellobiose, and galabios, have been synthesized in yields of around 50% [70,175,176,177]. Treatment of the β-chloro sugar of peracetylated neuraminic acid with a 3-thiogalactopyranoside derivative in the presence of NaH and a crown ether yielded the α-thio-linked disaccharide in high yield and stereospecificity (*Scheme 55*) [178].



Scheme 55

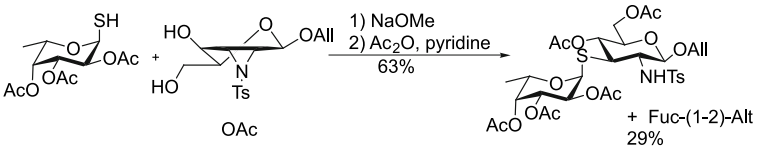
3.2 Displacement Reactions with 1-Thioglycoses

The second approach to thiooligosaccharides follows that outlined in *Scheme 2*, i. e., an anomeric thiol (or thiolate) is first constructed and then reacted with a saccharide electrophile to give the thiooligosaccharide through an S_N2 displacement reaction. Problems to consider are the stereospecific introduction of the anomeric sulfur group, its (preferably selective) conversion into a thiol, and the nature of the electrophile. Side-reactions encountered in the couplings are elimination of the electrophile and anomerization of the thiol. In a model study of the synthesis of methyl 4-thiocellobioside, this approach was found to be superior to the one using thiol acceptors [179].

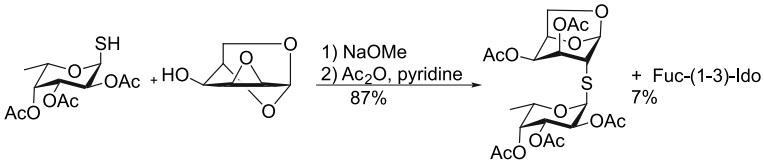
3.2.1 Formation of Anomeric Thiols

Of all the possible routes to anomeric thiols discussed in Sects. 1 and 2, almost exclusively thioacetates and thiourea salts have been used as intermediates in the synthesis of thiooligosaccharides, above all, since they can be selectively cleaved into the thiol in the presence of acyl protecting groups (*Sects. 2.2.1* and *2.2.2*). The synthesis of the 1,2-*trans*-1-thio deriva-

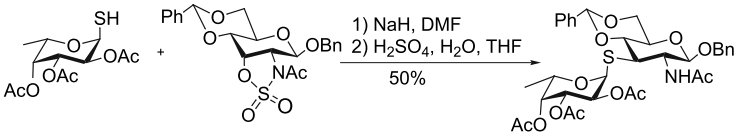
tives is usually straight-forward using participating protecting groups. However, also the construction of 1,2-*cis* compounds is usually performed using acetyl protecting groups, but then starting from 1,2-*trans* halogeno sugars employing conditions favoring S_N2-type reactions to give the 1,2-*cis* products. The anomeric thiol can either be isolated (with the concomitant problem of disulfide formation) and is then often activated as the thiolate during the coupling with the electrophile using, e. g., sodium hydride, or the cleavage and activation can be accomplished in situ. The displacement reaction is performed using an aprotic polar solvent or a phase-transfer reagent and non-polar solvents. Activation through complete deacetylation



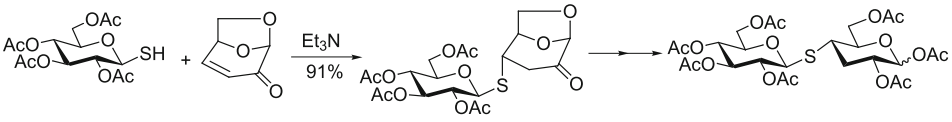
Scheme 56



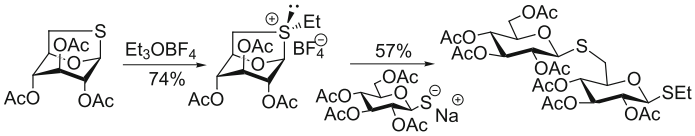
Scheme 57



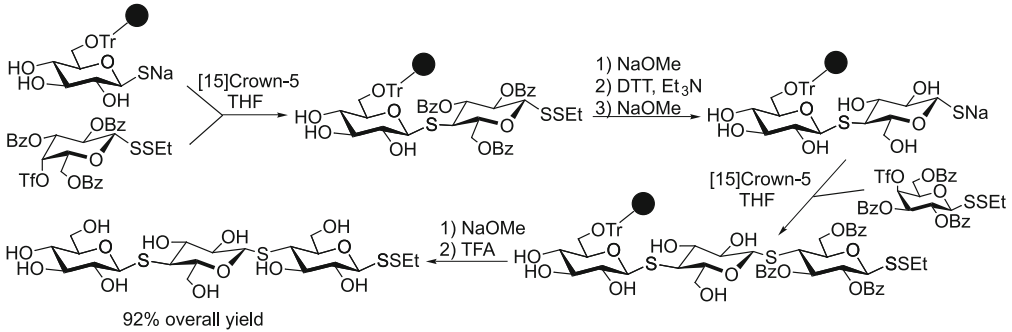
Scheme 58



Scheme 59



Scheme 60


[15]Crown-5
THF
1) NaOMe
2) DTT, Et3N
3) NaOMe
[15]Crown-5
THF
1) NaOMe
2) TFA
92% overall yield

Scheme 61

and formation of the unprotected anomeric thiolate, utilizing the higher nucleophilicity (and lower basicity) of sulfur in the consecutive substitution reaction is also an often used alternative.

3.2.2 The Nature of the Electrophile (Leaving Group)

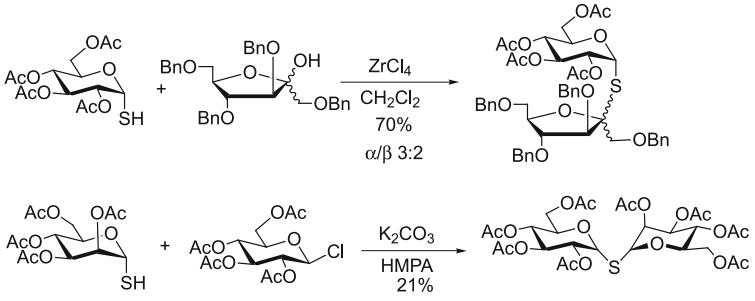
In the synthesis of (1→6)-linked structures, with a primary leaving group, even less effective leaving groups can be used. Bromides, iodides, tosylates, and triflates have been employed, bromides and tosylates requiring harsher conditions (sodium salt, high temperature) for reaction. Obviously there is also no problem with stereochemistry, neither in the construction of the electrophile nor in the displacement reaction. However, at secondary positions only triflates are good enough leaving groups to allow the displacement by anomeric thiolates and the construction of thiooligosaccharides. The substitution reactions yield the inverted configuration, which might pose a problem if the required electrophile has an unusual configuration. Accordingly, (1→4)-linked galactose and glucose derivatives are conveniently converted into each other, as is the case with (1→2)-linked glucose and mannose derivatives. For other oligosaccharides, for instance (1→3)-linked, a more laborious construction of the electrophile is often necessary, for example through a double inversion pathway.

Various other electrophiles have also been utilized, e. g., when their construction is more convenient than that of the corresponding triflate or when the triflate has failed to give the right product. Ring opening of a 2,3-*allo*-tosylaziridine, easily available from a GlcNAc precursor, with a 1-thiofucopyranosyl derivative gave with high regioselectivity the (1→3)-linked disaccharide (● *Scheme 56*) [180,181].

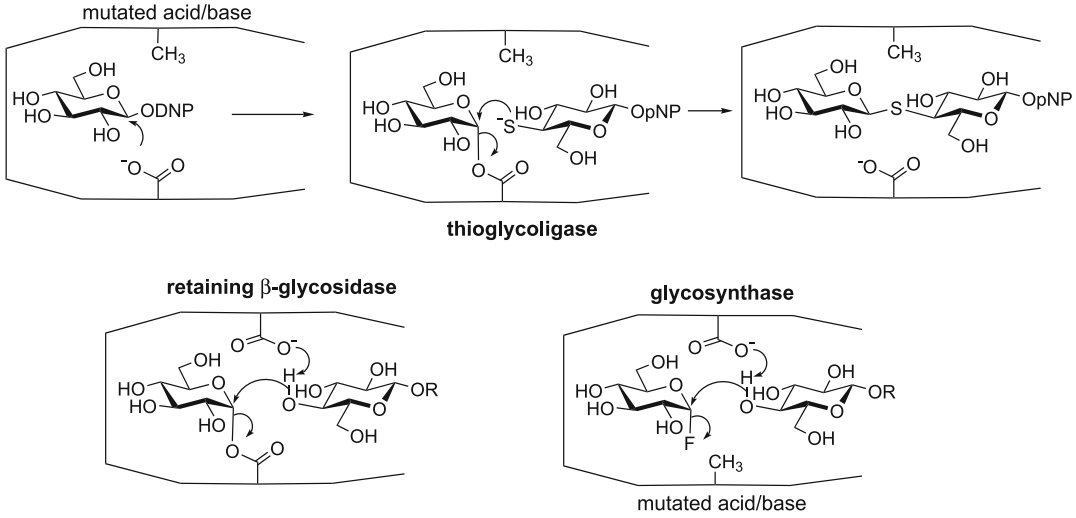
Similarly, ring opening of a 2,3-*talo*-oxirane with the same thionucleophile gave mainly the (1→2)-linked product (● *Scheme 57*) [180,181].

After failures in the synthesis and in the use of the 3-*O*-*allo*-triflate, the corresponding cyclic sulfamidate was constructed and reacted with the same 1-thiofucopyranosyl derivative to yield the desired (1→3)-linked thiodisaccharide (● *Scheme 58*) [182].

Other saccharide electrophiles utilized have been an α,β-unsaturated keto derivative and ethylated thiolevoglucosan. The former yielded in a Michael addition with tetra-*O*-acetyl-1-thio-β-D-glucopyranose the (1→4)-linked disaccharide in an excellent yield (● *Scheme 59*) [183,184]. The latter, when reacted with the sodium salt of the above glucose thiol, gave the (1→6)-linked disaccharide as its thioethyl glycoside (● *Scheme 60*) [185].



Scheme 62


mutated acid/base
thioglycoligase
retaining β-glycosidase
glycosynthase
mutated acid/base

Scheme 63

3.2.3 Solid-Phase Synthesis of Thiooligosaccharides

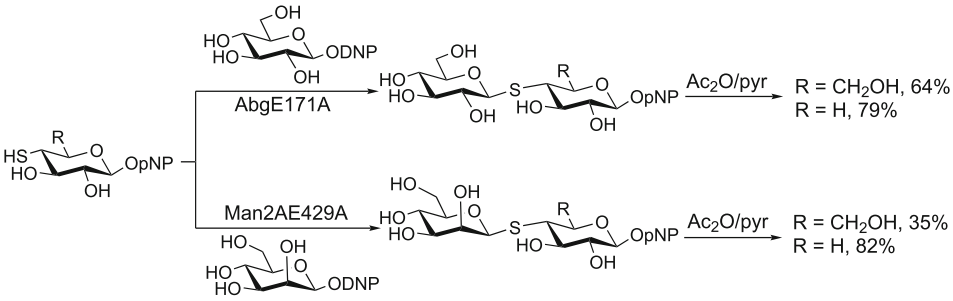
Solid-phase syntheses of thiooligosaccharides in excellent yields have been described [186]. Unprotected monosaccharides with an anomeric ethyl disulfide group were attached to a solid phase via their primary hydroxy group. Activation through the formation of the sodium thiolate and subsequent reaction with a saccharide triflate gave thiodisaccharides, which can be cleaved from the resin by treatment with TFA. If the electrophile contains an anomeric ethyl disulfide group, the reaction sequence can be repeated to give thiooligosaccharides (➤ *Scheme 61*).

3.2.4 Synthesis of Non-Reducing Sugars

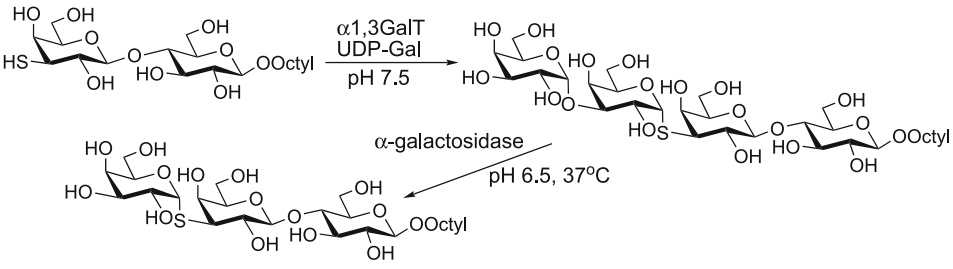
The syntheses of thio analogues of non-reducing disaccharides could be categorized under any of the above two headings since they involve both an anomeric thiol and an anomeric electrophile. Both halo sugars and hemiacetals have been used as electrophiles (➤ *Scheme 62*) [70,187].

3.2.5 Enzymatic Synthesis

Attempts to use glucosidases to form interglycosidic thioglycoside linkages from thiol precursors have not been successful. Neither was the use of glycosynthases, mutant glycosidases optimized to catalyze the formation and not the hydrolysis of *O*-glycosidic bonds where one (the nucleophilic) of the two carboxylate residues in the active site of the glycosidases has been changed to a non-nucleophilic (methyl) group. However, when the other carboxylic acid



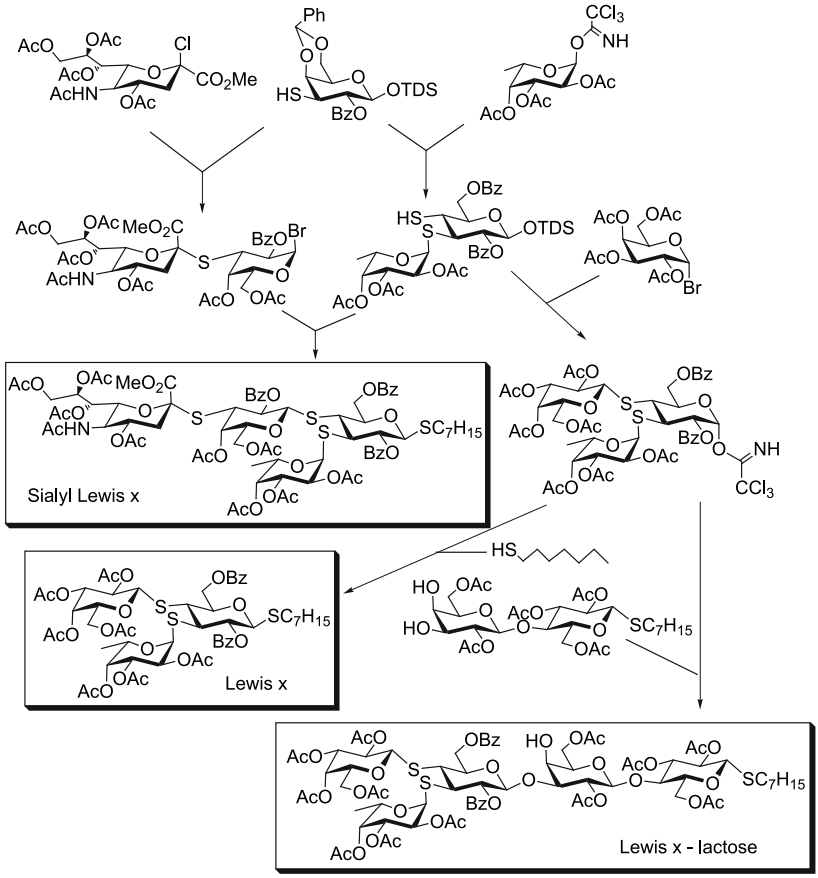
Scheme 64



Scheme 65

residue (the general acid/base) was mutated, a new type of mutant glycosidases was created, named thioglycoligases by the developers, which were able to catalyze the reaction between a glycosyl donor and a thiol acceptor to efficiently give thioglycosidically linked oligosaccharides (*Scheme 63*) [188]. Further studies showed that also the "double mutant" (called thioglycosynthases), where both the carboxylates were removed could catalyze this type of reaction [189]. So far only β-(1→4)-linkages have been produced using glucose and xylose acceptors and glucose and mannose donors (*Scheme 64*) [190,191,192,193].

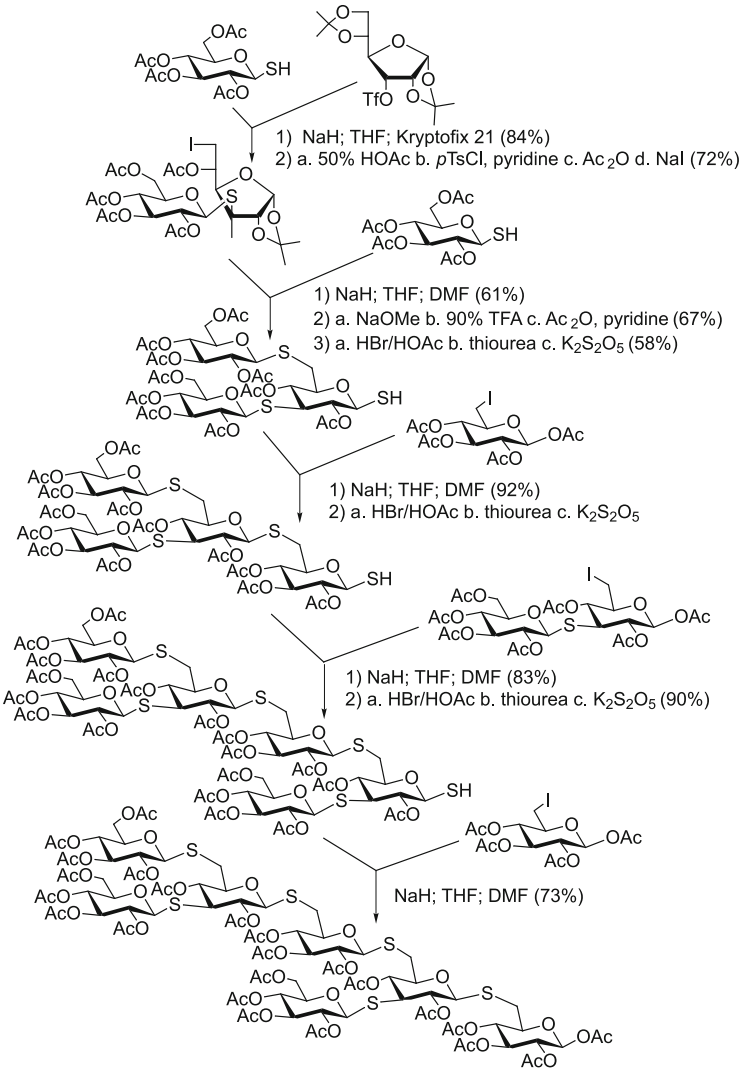
Interestingly, glycosyl transferases seem to tolerate thiol acceptors. Thus, using an α1,3-Gal-transferase and an octyl 3′-thio-lactoside acceptor a thio-interglycosidic linkage was formed in a very high yield. Surprisingly, the product was not the expected trisaccharide, but a tetrasaccharide with an extra (1→3)-*O*-linked galactose residue added. Here, the thiolinkages stability towards glycosidases could be utilized to selectively cleave the Gal-(1→3)-Gal *O*-glycosidic bond to obtain the desired trisaccharide product (*Scheme 65*). In addition, MS-data were reported to show that the 3′-thiolactoside was also accepted by a β1,3GlcNac-transferase [194].



Scheme 66

3.3 Examples

With the continuous development of new methods and improvement of old methods, the construction of more and more complex thiooligosaccharides has become possible. Two recently published examples of this are shown in Scheme 66 and Scheme 67. In Scheme 66 the synthesis of thio analogues of Lewis x and sialyl-Lewis x, using exclusively non-anomeric thiols in both acid-and base-catalyzed glycosylation reactions, is summarized [195].



Scheme 67

In *Scheme 67* is shown the assembly of the thio analogue of a phytoelicitor active heptasaccharide. In contrast to the above example, here all the thio-interglycosidic linkages are formed through the reaction of anomeric thiols with non-anomeric electrophiles [196].
In spite of the difference in methodology between these syntheses, both are impressive and excellent representatives of the state of the art of thiooligosaccharide synthesis.

References

1. Horton D, Hutson DH (1963) Adv Carbohydr Chem 18:123
2. Gil V, MacLeod AJ (1980) Phytochemistry 19:2071
3. Fahey JW, Zalcmann AT, Talalay P (2001) Phytochemistry 56:5
4. Herr RR (1967) J Am Chem Soc 89:2444
5. Hoeksema H, Bannister B, Birkenmeyer RD, Kagan F, Magerlein BJ, MacKellar FA, Schroeder W, Slomp G, Herr RR (1964) J Am Chem Soc 86:4223
6. Hoeksema H (1964) J Am Chem Soc 86:4224
7. Ngane AN, Lavault M, Seraphin D, Landreau A, Richomme P (2006) Carbohydr Res 341:2799
8. Khiar N, Suarez B, Valdivia V, Fernandez I (2005) Synlett 19:2963
9. Khiar N, Araujo CS, Suarez B, Fernandez I (2006) Eur J Org Chem 7:1685
10. Bundle DR, Rich JR, Jacques S, Yu HN, Nitz M, Ling C-C (2005) Angew Chem Int Ed 44:7725
11. Norberg T (1996) In: Khan SH, O'Neill RA (eds) Modern Methods in Carbohydrate Synthesis. Harwood, London, p 82
12. Garegg PJ (1997) Adv Carbohydr Chem Biochem 52:179
13. Oscarson S (2000) In: Ernst B, Hart G, Sinaÿ P (eds) Oligosaccharides in Chemistry and Biology: A Comprehensive Handbook. Wiley-VCH, Weinheim
14. Kartha KPR, Field RA (2003) In: Osborn HMI (ed) Carbohydrates. Elsevier Science Ltd, Oxford, p 121
15. Defaye J, Gelas J (1991) In: Atta-ur-Rahman (ed) Studies in Natural Products Chemistry, vol 8 E. Elsevier, Amsterdam, p 315
16. Defaye J (1995) In: Petersen SB, Svensson B, Petersen S (eds) Progress in Biotechnology, vol 10. Elsevier, Amsterdam, p 113
17. Driguez H (1997) Topics Curr Chem 187:85
18. Fairweather JK, Driguez H (2000) In: Ernst B, Hart G, Sinaÿ P (eds) Oligosaccharides in Chemistry and Biology: A Comprehensive Handbook. Wiley-VCH, Weinheim, p 531
19. Kahne D, Walker S, Cheng Y, van Engen D (1989) J Am Chem Soc 111:6881
20. Crich D, Lim LBL (2004) Org React 64:115
21. Agnihotri G, Misra AK (2005) Tetrahedron Lett 46:8113
22. Fernandez I, Khiar N (2003) Chem Rev 103:3651
23. Mehta S, Pinto BM (1991) Tetrahedron Lett 32:4435
24. Mehta S, Pinto BM (1996) In: Khan SH, O'Neill RA (eds) Modern Methods in Carbohydrate Synthesis. Harwood, London, p 107
25. van Well RM, Kaerkkaeinen TS, Kartha KPR, Field RA (2006) Carbohydr Res 341:1391
26. Buts L, Bouckaert J, De Genst E, Loris R, Oscarson S, Lahmann M, Messens J, Brosens E, Wyns L, De Greve H (2003) Acta Cryst D 59:1012
27. Cioci G, Mitchell EP, Chazalet V, Debray H, Oscarson S, Lahmann M, Breton C, Gautier C, Perez S, Imberty A (2006) J Mol Biol 357:1575
28. Witczak ZJ, Czernecki S (1998) Adv Carbohydr Chem Biochem 53:143
29. Robina I, Vogel P, Witczak ZJ (2001) Curr Org Chem 5:1177
30. Brigl P, Schinle R (1932) Ber dtsch chem Ges 65:1890
31. Lemieux RU (1951) Can J Chem 29:1079
32. Lemieux RU, Brice C (1955) Can J Chem 33:109
33. Hough L, Taha MI (1956) J Chem Soc 2042
34. Ferrier RJ, Furneaux RH (1976) Carbohydr Res 52:63
35. Das SK, Roy N (1996) Carbohydr Res 296:275
36. Dasgupta F, Garegg PJ (1989) Acta Chem Scand 43:471
37. Contour M-O, Defaye J, Little M, Wong E (1989) Carbohydr Res 193:283
38. Das SK, Roy J, Reddy KA, Abbineni C (2003) Carbohydr Res 338:2237
39. Weng S-S, Lin Y-D, Chen C-T (2006) Org Lett 8:5633
40. Vesely J, Ledvina M, Jindrich J, Saman D, Trnka T (2003) Collect Czech Chem Comm 68: 1264

41. Dahlhoff WV (1990) Liebigs Ann Chem 1025
42. Doren HAV, Geest RVD, Kellogg RM, Wynberg H (1989) Carbohydr Res 194:71
43. Galema SA, Engberts JBFN, van Doren HA (1997) Carbohydr Res 303:423
44. Boons GJ, Geurtsen R, Holmes D (1995) Tetrahedron Lett 36:6325
45. Garegg PJ, Olsson L, Oscarson S (1995) J Org Chem 60:2200
46. Nakano T, Ito Y, Ogawa T (1990) Tetrahedron Lett 31:1597
47. Lindberg B, Erbing B (1976) Acta Chem Scand Ser B 30:611
48. Boons GJPH, Delft FL, vanKlein PAM, van der Marel GA, van Boom JH (1992) Tetrahedron 48:885
49. Mannerstedt K, Ekelöf K, Oscarson S (2006) Carbohydr Res 342:631
50. Marra A (1989) Carbohydr Res 187:35
51. Ibatullin FM, Shabalin KA, Jänis JV, Shavva AG (2003) Tetrahedron Lett 44:7961
52. Kudelska W, Michalska M (1995) Synthesis 1539
53. Marino C, Marino K, Miletti L, Alves MJM, Colli W, de Lederkremer RM (1998) Glycobiology 8:901
54. Helander A, Kenne L, Oscarson S, Peters T, Brisson JR (1992) Carbohydr Res 230:299.
55. Adamo R, Kovac P (2006) Eur J Org Chem 2803
56. Agnihotri G, Tiwari P, Misra AK (2005) Carbohydr Res 340:1393
57. Tai CA, Kulkarni SS, Hung SC (2003) J Org Chem 68:8719
58. Lin CC, Huang LC, Liang PH, Liu CY (2006) J Carbohydr Chem 25:303
59. Ogawa T, Matsui M (1977) Carbohydr Res 54:C17
60. Pozsgay V, Jennings HJ (1987) Tetrahedron Lett 28:1375
61. Takahashi S, Terayama H, Kuzuhara H (1992) Tetrahedron Lett 33:7565
62. Kartha KPR, Field RA (1998) J Carbohydr Chem 17:693
63. Sato T, Fujita Y, Otera J, Nozaki H (1992) Tetrahedron Lett 33:239
64. Deng S, Gangadharmath U, Chang CWT (2006) J Org Chem 71:5179
65. Fischer E, Delbrück K (1909) Ber dtsch chem Ges 42:1476
66. Schneider W, Sepp J, Stiehler O (1918) Ber dtsch chem Ges 51:220
67. Helferich B, Grünewald H, Langenhoff F (1953) Chem Ber 86:873
68. Pedretti V, Veyrières A, Sinaÿ P (1990) Tetrahedron 46:77
69. El Ashry ESH, Awad LF, Atta AI (2006) Tetrahedron 62:2943
70. Blanc-Muesser M, Defaye J, Driguez H (1978) Carbohydr Res 67:305
71. Apparu M, Blanc-Muesser M, Defaye J, Driguez H (1981) Can J Chem 59:314
72. Bogusiak J, Szeja W (1985) Pol J Chem 59:293
73. Tropper FD, Andersson FO, Grand-Maître C, Roy R (1991) Synthesis 734
74. Meunier SJ, Andersson FO, Letellier M, Roy R (1994) Tetrahedron Asymm 5:2303
75. Gerz M, Matter H, Kessler H (1993) Angew Chem Int Ed Engl 32:269
76. Byun H-S, Bittman R (1995) Tetrahedron Lett 36:5143
77. Mukhopadhyay B, Kartha KPK, Russell DA, Field RA (2004) J Org Chem 69:7758
78. Kumar R, Tiwari P, Maulik PR, Misra AK (2006) Eur J Org Chem 74
79. van der Klein PAM, Boons GJPH, Veeneman GH, van der Marel GA, van Boom JH (1990) Synlett 6:311
80. Reimer M, Schmidt RR (2000) Tetrahedron Asymm 11:319
81. Timmer MSM, Adibekian A, Seeberger PH (2005) Angew Chem Int Ed 44:7605
82. Stewart AO, Williams RM (1985) J Am Chem Soc 107:4289
83. Li ZJ, Liu PL, Qui DX (1990) Synth Commun 20:2169
84. Fürstner A (1993) Liebigs Ann Chem 1211
85. Szeja W, Bogusiak J (1988) Synthesis 224
86. Inanaga J, Yokoyama Y, Hanamoto T (1993) J Chem Soc Chem Commun 1090
87. Hiranuma S, Kajimoto T, Wong C-H (1994) Tetrahedron Lett 35:5257
88. Schmidt RR, Stumpp M (1983) Liebigs Ann Chem 1249
89. Käsbeck L, Kessler H (1996) Liebigs Ann/Recueil 165
90. Wolfrom ML, Thompson A (1963) Methods Carbohydr Chem 3:150
91. Hannesian S, Guindon (1980) Carbohydr Res 86:C3
92. Nicolau KC, Seitz SP, Papahatjis DP (1983) J Am Chem Soc 105:2430
93. Liu D, Chen R, Hong L, Sofia MJ (1998) Tetrahedron Lett 39:4951
94. Sakairi N, Kuzuhara H (1996) Carbohydr Res 280:139

95. Zhang Z, Magnusson G (1996) Carbohydr Res 295:41
96. Mendlik TM, Tao P, Hadad CM, Coleman RS, Lowary TL (2006) J Org Chem 71: 8059
97. Araki Y, Matsuura K, Ishido Y, Kushida K (1973) Chem Lett 383
98. Gadelle A, Defaye J, Pedersen C (1990) Carbohydr Res 200:497
99. Lehmann J (1966) Carbohydr Res 2:486
100. Igarashi K, Honma T(1970) J Org Chem 35:606
101. Priebe W, Zamojski A (1980) Tetrahedron 36:274
102. Dunkerton LV, Adair NA, Euske JM, Brady KT, Robinson PD (1988) J Org Chem 53:845
103. Yadav JS, Reddy BVS, Geetha V (2003) Synth Comm 33:717
104. Smitha G, Reddy S (2004) Synthesis 834
105. Mereyala HB (1987) Carbohydr Res 168:136
106. Paul S, Jayaraman N (2004) Carbohydr Res 339:2197
107. Sherry BD, Loy RN, Toste FD (2004) J Am Chem Soc 126:4510
108. Yadav JS, Reddy BVS, Bhasker EV, Raghavendra S, Narsaiah AV (2007) Tetrahedron Lett 48:677
109. Weygand F, Ziemann H (1962) Liebigs Ann Chem 657:179
110. Seeberger PH, Eckhardt M, Gutteridge CE, Danishefsky SJ (1997) J Am Chem Soc 119:10064
111. Sakairi N, Hayashida M, Kuzuhara H (1987) Tetrahedron Lett 28:2871
112. Wang LX, Sakairi N, Kuzuhara H (1990) J Chem Soc Perkin Trans 1 1677
113. Garegg PJ, Oscarson S, Tedebark U (1998) J Carbohydr Chem 17:587
114. Skelton BW, Stick RV, Tilbrook DMG, White AH, Williams, SJ (2000) Austr J Chem 53:389
115. Lundt I, Skelbaeck-Pedersen B (1981) Acta Chem Scand Ser B 35:637
116. Plet JRH, Porter MJ (2006) Chem Comm 1197
117. Falconer RA (2002) Tetrahedron Lett 43:8503
118. Zhu X (2006) Tetrahedron Lett 47:7935
119. Gehrke M, Kohler W (1931) Ber dtsch chem Ges 64:2696
120. Blanc-Muesser M, Defaye J, Driguez H (1982) J Chem Soc Perkin Trans 1:15
121. Hasegawa A, Nakamura J, Kiso M (1986) J Carbohydr Chem 5:11
122. Ferrier RJ, Furneaux RH (1977) Carbohydr Res 57:63
123. Defaye J, Guillot JM (1994) Carbohydr Res 253:185
124. Park WKC, Meunier SJ, Zanini D, Roy R (1995) Carbohydr Lett 1:179
125. Horton D (1963) Methods Carbohydr Chem 2:433
126. Cerny M, Stanek J, Pacák J (1963) Monatsh Chem 94:290
127. Ibatullin FM, Selivanov SI, Shavva AG (2001) Synthesis 419
128. Tiwari P, Agnihotri G, Misra AK (2005) J Carbohydr Chem 24:723
129. El Ashry ESH, Awad LF, Hamid HMA, Atta AI (2005) J Carbohydr Chem 24:745
130. El Ashry ESH, Awad LF, Hamid HMA, Atta AI (2006) Synth Comm 36:2769
131. Fischer E, Helferich B, Ostman P (1920) Ber dtsch chem Ges 53:873
132. Müller A, Wilhelms A (1941) Ber dtsch chem Ges 74:698
133. Kochetkov NK, Klimov E, Malysheva NN, Demchenko AV (1991) Carbohydr Res 212:77
134. Fischer E (1914) Ber dtsch chem Ges 47:1378
135. Ogura H, Takahashi H (1982) Heterocycles 17:87
136. Camarasa MJ, Fernández-Resa P, López MTG, Heras FGdL, Méndez-Castrillón PP, Felix AS (1984) Synthesis 509
137. Lindhorst TK, Kieburg C (1995) Synthesis 1228
138. Pakulski Z, Pierozynski D, Zamojski A (1994) Tetrahedron 50:2975
139. Schneider W, Gille R, Eisfeld K (1928) Ber dtsch chem Ges 61:528, 1244
140. Tropper FD, Andersson FO, Cao S, Roy R (1992) J Carbohydr Chem 11:741
141. Bogusiak J, Wandzik I, Szeja W (1996) Carbohydr Lett 1:411
142. Fügedi P, Garegg PJ, Oscarson S, Rosén G, Silwanis BA (1991) Carbohydr Res 211:157
143. Sakata M, Haga M, Tejima S, Akagi M (1963) Chem Pharm Bull 11:1081
144. Sakata M, M MH, Tejima S (1970) Carbohydr Res 13:379
145. Knapp S, Gonzales S, Myers DS, Eckman LL, Bewley CA (2002) Org Lett 4:4337
146. Knapp S, Myers DS (2001) J Org Chem 66:3636
147. Knapp S, Myers DS (2002) J Org Chem 67:2995
148. Matsuoka K, Ohtawa T, Hinou H, Koyama T, Esumi Y, Nishimura SI, Hatano K, Terunuma D (2003) Tetrahedron Lett 44:3617
149. Bhar D, Chandrasekaran S (1997) Carbohydr Res 301:221
150. Sridhar PR, Prabhu KR, Chandrasekaran S (2004) Eur J Org Chem 4809

151. McDougall JM, Zhang XD, Polgar WE, Khroyan TV, Toll L, Cashman JR (2004) J Med Chem 47:5809
152. Jobron L, Hummel G (2000) Org Lett 2:2265
153. Falconer RA, Jablonkai I, Toth I (1999) Tetrahedron Lett 40:8663
154. Falco EA, Hitchings GH, Russell PB (1949) J Am Chem Soc 71:362
155. Driguez H, Szeja W (1994) Synthesis 1413
156. Naus P, Leseticky L, Smrcek S, Tislerova I, Sticha M (2003) Synlett 2117
157. Herradura PS, Pendola KA, Guy RK (2000) Org Lett 2:2019
158. Cerny M, Zachystalova D, Pacák J (1961) Coll Czech Chem Commun 26:2206
159. Lacombe JM, Rakotomanomana N, Pavia AA (1988) Tetrahedron Lett 29:4293
160. Bourgeois MJ, Gueyrard D, Montaudon E, Rollin P (2005) Lett Org Chem 2:665
161. Fokt I, Szeja W (1992) Carbohydr Res 232:169
162. Fokt I, Bogusiak J, Szeja W (1998) Carbohydr Lett 3:191
163. Magnusson G (1977) J Org Chem 42:913
164. Krog-Jensen C, Oscarson S (1998) Carbohydr Res 308:287
165. Josse S, Le Gal J, Pipelier M, Cleophax J, Olesker A, Pradere JP, Dubreuil D (2002) Tetrahedron Lett 43:237
166. Ane A, Josse S, Naud S, Lacone L, Vidot S, Fournial A, Kar A, Pipelier M, Dubreuil D (2006) Tetrahedron 62:4784
167. Rodriguez MA, Boutureira O, Arnes X, Matheu MI, Diaz Y, Castillon S (2005) J Org Chem 70:10297
168. Driguez H (2001) Chem Bio Chem 2:311
169. Mehta S, Andrews JS, Johnston BD, Pinto BM (1994) J Am Chem Soc 116:1569
170. Mehta S, Andrews JS, Johnston BD, Svensson B, Pinto BM (1995) J Am Chem Soc 117:9783
171. Andrews JS, Pinto BM (1995) Carbohydr Res 270:51
172. Andrews JS, Johnston BD, Pinto BM (1998) Carbohydr Res 310:27
173. Johnston BD, Pinto BM (1998) Carbohydr Res 310:17
174. Blanc-Meusser M, Driguez H (1988) J Chem Soc Perkin Trans 1 3345
175. Hutson DH (1967) J Chem Soc (C) 442
176. Reed LA, Goodman L (1981) Carbohydr Res 94:91
177. Nilsson U, Johansson R, Magnusson G (1996) Chem Eur J 2:295
178. Eisele T, Toepfer A, Kretzschmar G, Schmidt RR (1996) Tetrahedron Lett 37:1389
179. Moreau V, Norrild J C, Driguez H (1997) Carbohydr Res 300:271
180. Hashimoto H, Shimada K, Horito S (1993) Tetrahedron Lett 31:4953
181. Hashimoto H, Shimada K, Horito S (1994) Tetrahedron Asymm 5:2351
182. Aguilera B, Fernández-Mayoralas A (1996) J Chem Soc Chem Commun 127
183. Witczak ZJ, Sun J, Mielguj R (1995) Bioorg Med Chem Lett 5:2169
184. Witczak ZJ, Chhabra R, Chen H, Xie X-Q (1997) Carbohydr Res 301:167
185. Lundt I, Skelbaeck-Pedersen B (1981) Acta Chem Scand Ser B 35:637
186. Hummel G, Hindsgaul O (1999) Angew Chem Int Ed 38:1782
187. Defaye J, Driguez H, Poncet S, Chambert R, Petit-Glatron M-F (1984) Carbohydr Res 130:300
188. Jahn M, Marles J, Warren RAJ, Withers SG (2003) Angew Chem Int Ed 42:352
189. Jahn M, Chen H, Mullegger J, Marles J, Warren RAJ, Withers SG (2004) Chem Comm 274
190. Jahn M, Marles J, Warren RAJ, Withers SG (2003) Chem Comm 1327
191. Jahn M, Withers SG (2003) Biocat Biotrans 21:159
192. Mullegger J, Jahn M, Chen HM, Warren RAJ, Withers SG (2005) Protein Eng Design Select 18:33
193. Stick RV, Stubbs KA (2005) Tetrahedron Asymm 16:321
194. Rich JR, Szpacenko A, Palcic MM, Bundle DR (2004) Angew Chem Int Ed 43: 613
195. Eisele T, Schmidt RR (1996) Liebigs Ann/Recueil 1303
196. Ding Y, Contour-Galcera M-O, Ebel J, Ortiz-Mellet C, Defaye J (1999) Eur J Org Chem 1143
197. Bernardes JLG, Gamblin DP, Davis BG (2006) Angew Chem Int Ed 45:4007

3.6 Glycal Derivatives

Waldemar Priebe[*1], *Izabela Fokt*[1], *Grzegorz Grynkiewicz*[2]
[1] M. D. Anderson Cancer Center, The University of Texas, Houston, TX 77030, USA
[2] Pharmaceutical Research Institute, Rydgiera 8, 01-793 Warsaw, Poland
wpriebe@mdanderson.org, wp@wt.net, g.grynkiewicz@ifarm.waw.pl

Abstract

Glycals are unsaturated sugar derivatives in which the double bond engages the anomeric carbon atom. Such cyclic vinyl ethers are characterized by high reactivity, allowing for regio- and stereoselective transformations, directly or indirectly related to glycosylation, as well as to formation of carbon–carbon and carbon–heteroatom bonds at the anomeric center. This review provides a systematic survey of chemical synthetic methods, by which the carbon–carbon double bond is introduced next to the ring oxygen, in an endo- or exocyclic position. Some mechanistic aspects are discussed in relation to traditional methods of glycal preparation, which rely on elimination reactions. Glycal-to-glycal rearrangement and applications of organometallic chemistry and heteroatom-induced transformations for syntheses and activation of glycals are also highlighted.

Keywords

Glycal; Synthesis; Carbohydrate; Monosaccharides

In: *Glycoscience*. Fraser-Reid B, Tatsuta K, Thiem J (eds)
Chapter-DOI 10-1007/978-3-540-30429-6_15:

Abbreviations

AIBN	azobisisobutyronitrile
DBU	1,8-diazabicyclo[5.4.0]undec-7-ene
HMPA	hexamethylphosphoramide
HSAB	hard and soft acids and bases
LAH	lithium aluminum hydride
LDA	lithium diisopropylamide
m-CPBA	3-chloroperbenzoic acid
NIS	*N*-iodosuccinimide
THF	tetrahydrofuran
TMSBr	trimethylsilyl bromide

1 Introduction

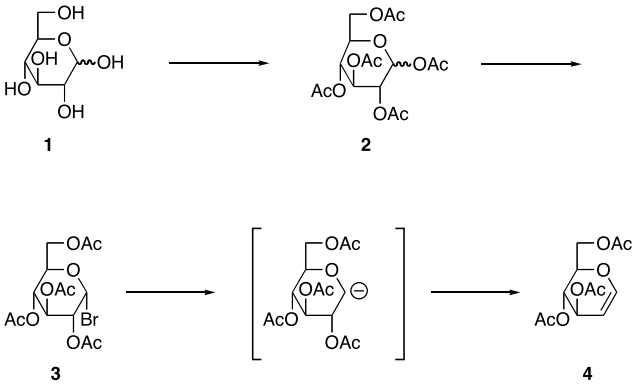
Monosaccharides possessing a double bond between C1 and C2 are known under their trivial name as glycals. The name glycals was coined after the product of the first synthesis of 1,2-unsaturated sugar from D-glucose **1** which was named D-glucal [1]. Consequently, the generic name "glycals" was adopted for this class of unsaturated sugars, which are in fact extensively substituted cyclic vinyl ethers. Such cyclic vinyl ethers are also encountered as structural motifs in total synthesis of some complex natural products [2,3].

Systematic naming [4,5] uses alditol descriptors and not the parent aldose, as is the case in commonly used trivial names. Thus, **4** should be named 1,5-anhydro-2-deoxy-D-*arabino*-hex-1-enitol. However, trivial names are being widely used because of their simplicity. Advances in the synthesis of glycals have been summarized regularly in comprehensive reviews [2,6,7,8] and new reactions were regularly summarized in annual reports [9]. Recent developments in the chemistry of glycals have expanded their importance as useful organic substrates and intermediates far beyond the limits set by the traditional placement of somewhat esoteric unsaturated sugars. Glycals are extensively used as glycosylating agents, [10,11,12] especially in the synthesis of 2-deoxy sugars [13,14,15,16] and oligosaccharides [17], and also as precursors of *C*-glycosyl compounds [18,19,20]. They are also widely recognized as versatile chiral building blocks [21,22,23]. Selectively protected glycals are being used in chemical and enzymatic synthesis of complex carbohydrates and in the synthesis of *C*-oligosaccharides and glycoproteins [24,25]. There is an increased interest in exploring carbohydrate scaffolds in drug discovery as a consequence of identifying the role of glycoconjugates in biological recognition and signal transduction processes [26]. After genomics and proteomics, glycomics appears to be the next significant frontier for the development of new medical therapeutics, where progress will require concerted multidisciplinary efforts. Specifically, chemical glycomics will require the maturation of synthetic carbohydrate methodologies that would allow the preparation of glycoconjugates in a parallel, automated synthesis [27,28]. It now appears that unsaturated sugar derivatives, like glycals, can be used as useful substrates for such a purpose [29,30].

Various aspects of essential glycal chemistry connected to Ferrier rearrangement and anomeric carbanions have been extensively reviewed [31,32,33,34].

2 Reductive Elimination of Glycosyl Halides

Because of their facile preparation and reactivity, glycosyl halides are commonly used in carbohydrate chemistry. They are primarily used as glycosyl donors [35,36] and are the most important substrates in the synthesis of glycals from monosaccharides. Especially useful are glycosyl bromides **3**; they are conveniently prepared from glycosyl esters **2** [37,38] (one-pot procedures starting from reducing sugars are also available [39,40] and are often used without purification in the next step of synthesis (*Scheme 1*)). Reductive elimination of glycosyl bromides, prepared in situ with zinc generates glycals with good to excellent yields [41]. The mechanism of reductive elimination is believed to involve a two-electron reduction process localized at the anomeric center, followed by elimination of an acetoxonium anion from the vicinal position [42].



Scheme 1

Fischer's original procedure [1] as well as subsequent improvements dating from the early 1900s, are practically unavailable; however, they are summarized and in some cases specifically described in *Methods in Carbohydrate Chemistry* [43,44,45,46]. All these procedures for reduction of glycosyl bromides into acetylated glycals used zinc powder alone or in combination with salts of other metals in aqueous acetic acid, and several generations of chemists used them as standard procedures for the preparation of glycals. Specifically, in the synthesis of 3,4-di-*O*-acetyl-D-xylal, the recommended reductive agent was zinc dust alone in 50% acetic acid (at a temperature below −10 °C) [47]. However, heterogeneous reaction conditions can sometimes create problems of reproducibility, and although the quality of the zinc powder used was not as critical as it is in typical metaloorganic transformations [48], it was often mentioned as a factor affecting yield [49]. To overcome this problem, pre-washing with diluted hydrochloric acid or the addition of platinum [50] and copper [51] salts were proposed for more effective reduction on a metal surface. Such improvements are now part of the standard protocols for preparation of L-arabinal [43] and D-glucal [44], where zinc is activated with cupric sulfate pentahydrate. On the other hand, in the protocol for preparation of D-galactal, zinc is activated with platinic chloride [46]. In general, the yields for most glycals are good to excellent, and the final product can be purified by distillation or crystallization.

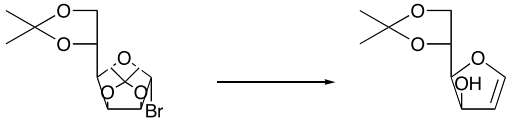
Another approach (and one more important for assessing the mechanisms of glycal formation than for practical use) involved the electro-reduction of 2,3,4,6-tetra-*O*-acetyl-D-glucopyranosyl bromide **3** on mercury electrode in dipolar aprotic solvents containing tetraethylammonium perchlorate to support a mechanistic two-electron reduction hypothesis [52]. However, 3,4,6-tri-*O*-acetyl-D-glucal **4** was obtained in only 55% yield. The main product **4** was accompanied by penta-*O*-acetyl-D-glucopyranose (35%) and 2-acetoxy-3,4,6-tri-*O*-acetyl-D-glucal (10%), indicating the extent of secondary solvolytic processes in this method.

The composition of reaction mixtures obtained under classical Fischer–Zach conditions was determined by capillary gas chromatography [53] of per-*O*-silylated and per-*O*-acetylated components of the crude reaction mixtures, and some progress was made in explaining the mechanism behind typical side-product formation. The compounds identified included products of solvolytic displacement of an anomeric bromine by acetic acid and water as well as 2,3-unsaturated derivatives. Interestingly, there was a difference in the reaction outcome of pentopyranosyl versus hexopyranosyl substrates. In the former case, glycals were accompanied by acetylated 1,5-anhydropentitols; in the latter, hex-1-enitols contained admixtures of peracetylated 2-deoxyhexopyranosides.

One failed attempt to convert peracetylated glycosyl chlorides into the corresponding glycals used zinc dust in aqueous acetic acid [54]. However, when an excess of chromium(II) diacetate dimer monohydrate was used together with 1,2-ethanediamine in dimethylformamide to generate α-pyranosyl chlorides of D-glucose, D-xylose, and D-arabinose, the corresponding glycals were obtained in yields of 82–97%.

Aluminum amalgam has been shown to be an effective reducing-eliminating reagent for glycosyl bromides when used in aqueous tetrahydrofuran [55], producing 60–85% yields of product. From the design of alternative procedures employing aprotic reaction conditions have come synthetic methodologies for the preparation of glycals from pyranose derivatives bearing acid-sensitive protecting groups and even from furanose derivatives. The first attempts in this direction, consisting of treating pentofuranosyl bromides bearing good leaving groups at C-2 with sodium iodide in dry acetone [56], generated a hitherto unknown class of furanosyl glycals. However, it also became immediately obvious that, to prevent Ferrier rearrangement, the oxygen at C-3 would require protection of the ether type [57]. Later, 2,3:5,6-di-*O*-isopropylidene-D-mannofuranosyl bromide **5** was shown to react smoothly with an excess of sodium or potassium in dry tetrahydrofuran to give corresponding glycal **6** in 59% yield [58], whereas lithium and magnesium failed to produce glycals (❯ *Scheme 2*).

Sodium naphtalide was found to be a good replacement for the sodium and to afford good yields of pyranoid and furanoid glycals from glycosyl halides having alkali-stable protecting groups [59]. The furanoid glycal **6** was subsequently prepared using modifications of this pro-



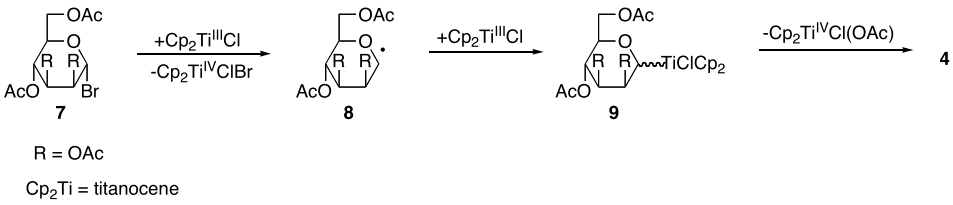
5 **6**

Scheme 2

cedure [60,61]. Meanwhile, Ireland's innovative application of lithium in liquid ammonia as a reducing agent for furanosyl halides [62,63] became a widespread procedure [64]. However, it has been demonstrated that, in special cases (i. e., when there are benzyl protective groups present and there is no heteroatom substitution at C-3), even furanoid glycals can be prepared by the zinc/acetic acid method [65].

Considerable attention has been given to reactive metal–graphite combinations, which were found to reduce both *O*-acylated pyranosyl and *O*-alkylidene furanosyl halide substrates under neutral conditions and to give excellent yields of high-purity glycals. Zinc/silver-graphite in tetrahydrofuran [66,67] and potassium graphite laminate [68] in the same solvent are very reactive, which allows the reaction work-up to be reduced to simple filtration and evaporation. The potassium graphite reagent is unique in effectively eliminating the 2-benzyloxy substituent from per-*O*-benzylated glycosyl halides, thus affording direct access to 3,4,6-tri-*O*-benzyl-D-glucal [69,70] and circumventing, in this case, the need for benzylation of preformed glucal derivatives.

Two other interesting reagents are samarium iodide and dimeric cyclopentadiene-titanium complex $(Cp_2TiCl)_2$. When used at 6 equivalents, samarium iodide can convert acetylated glucopyranosyl bromide into D-glucal **4** at yields of 90% [71]. Dimeric cyclopentadiene-titanium complex $(Cp_2TiCl)_2$, which is known to react with activated alkyl halides via halogen atom abstraction forming alkyl radicals, also reacts smoothly with glycosyl bromides in tetrahydrofuran at room temperature, affording corresponding glycals in very good yields [72]. The proposed reaction scheme, involving repeated transformation of Ti(III) into Ti(IV) in intermediate titanocene species [73], is shown on ❯ *Scheme 3*.

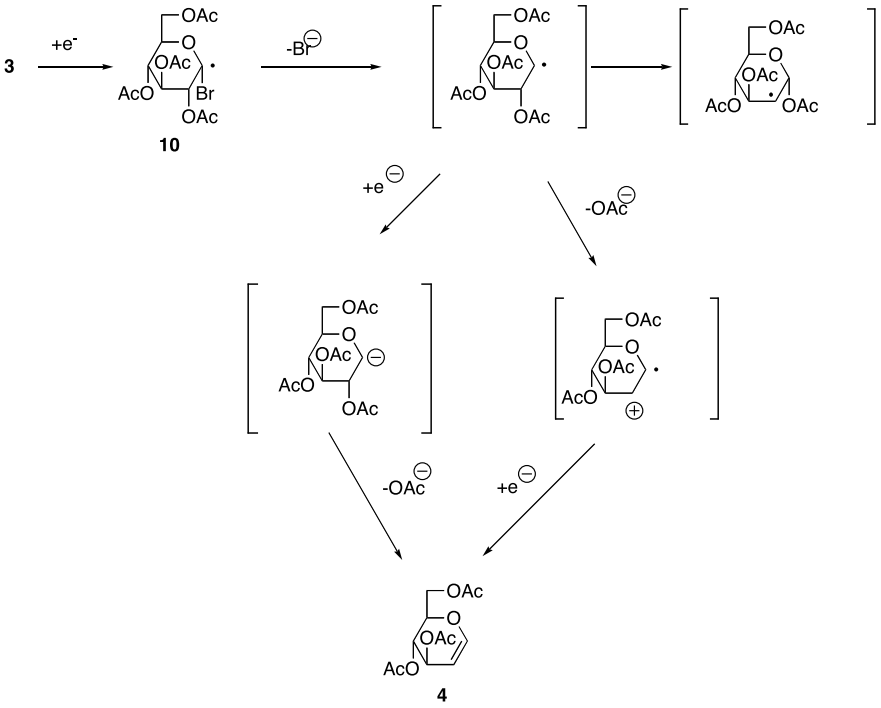


Scheme 3

The intermediate glycosylic-titanium complex is likely to be a mixture of anomers. But, when taking into account the facile formation of acetylated D-glucal from both -D-glucopyranosyl and -D-mannopyranosyl bromides under the conditions discussed above, it seems that both *syn* and *anti* elimination of $Cp_2TiCl(OAc)$ is possible. Glycosyl radical intermediates, on the other hand, show a strong preference for alpha orientation, as evidenced by the configuration of *C*-glycosyl compounds obtained by trapping with unsaturated species [74].

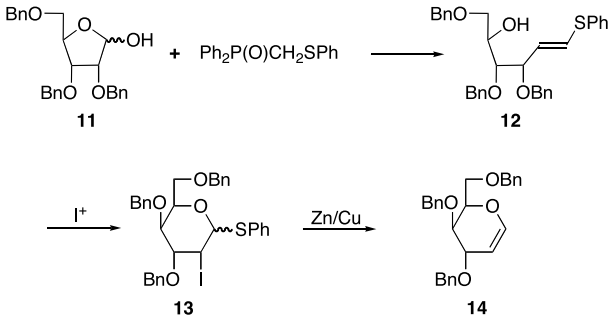
In the case of organochromium(III) complex intermediates obtained from glycosyl halides, work by Somsak and coworkers showed that their decay into glycals [75] is marked by the remarkable stability of the organometalic compounds in aqueous solutions. From the same group came the radically new concept of using zinc metal under aprotic conditions [76]. In brief, they proposed that acetylated glycosyl bromides be treated with zinc dust in the presence of *N*-heterocyclic bases (e. g., 4-methylpyridine or 1-methylimidazole) in a variety of aprotic solvents (i. e., benzene, ethyl acetate, tetrahydrofuran, acetone, dichloromethane) to afford the

corresponding glycals. The beneficial influence of triethylamine and pyridine addition on zinc-mediated reductive elimination was first noticed in the case of acylated 1-bromo-D-glycosyl cyanides [77] and was further noticed in other monosaccharide derivatives substituted with an electron-withdrawing group and bromine atom at the anomeric center [78,79]. To identify the reaction mechanism operating under aprotic solvent conditions, two model reactions were chosen and compositions of corresponding reaction mixtures were examined using capillary gas chromatography and a standard co-injection procedure [80]. The postulated reaction scheme, based on isolable product distribution as well as free radical inhibition and trapping experiments, involved glycosyl and 2-deoxyglycosyl radicals in addition to a glycosyl carbanion (1,5-anhydro-2,3,4,6-tetra-*O*-acetyl-D-glucitol-1-ide) [5]. This postulated carbanion, however, resisted attempts to trap it in the presence of a benzaldehyde serving as an electrophile (*Scheme 4*).



Scheme 4

Anomeric substituents, other than halides, are also capable of undergoing 1,2-elimination under a variety of reduction conditions. Anomeric 2-pyridylsulfonyl substituents can undergo 1,2-elimination, yielding hex-2-enitols, in reaction with samarium diiodide, particularly in the presence of HMPA [81]. Less obvious, but apparently general transformation leading to glycals has been discovered by Spanish researchers [82]. 2-Deoxy-2-iodo-1-phenylthioglycoside undergoes 1,2-elimination under various reducing conditions, preferably in the presence of metallic zinc or zinc-copper couple. Substrates for this reaction are easily accessible from



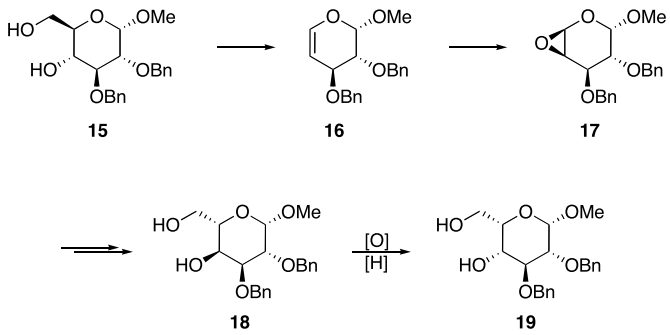
Scheme 5

1-OH pentofuranoses, as shown in *Scheme 5*. Thus, otherwise scarcely available D-allal and D-gulal derivatives can be readily prepared from D-ribose and D-arabinose, respectively [82]. It has recently been demonstrated that reductive elimination of β-hydroxycarboxylic acid can be applied for preparation of hex-1-enitols. The procedure involves refluxing appropriate substrate in toluene with *N*,*N*-dimethylformamide dineopentyl acetal.

By using such a transformation, uronic acid obtained from methyl D-glucopyranoside was converted into intermediate 4-deoxypentenoside and then via epoxidation followed by a nucleophile attack, into a variety of pyranosides of the L-configuration series, as shown in *Scheme 6* [83].

It should be noted that, under radical-generating conditions, a number of vicinal 1,2-substituents in cyclic sugar derivatives will tend to undergo elimination. This tendency has been demonstrated in a series of pyranosyl 1-phenylselenide-2-azides, which form corresponding glycals in yields close to quantitative upon treatment with tributyltin hydride [84].

Under the same conditions, glycals are also formed from 1-thioglycoside-2-xanthates. Phenyl thioglycopyranosides easily undergo reductive lithiation with lithium naphtalenide in tetrahydrofuran at low temperatures and subsequent elimination of C-2 substituents [85,86]. Such reaction conditions are compatible with labile acid- and base-protecting groups and yields



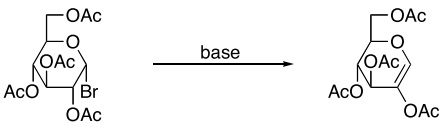
Scheme 6

of isolated glycals are high. Corresponding phenyl sulfones follow the same reaction pathway [87]. Anomeric sulfoxides form glycals upon treatment with 3 equivalents of butyllithium in THF at −78 °C. The yield of products generated from pyranosyl derivatives protected with acetal and ether groups ranges from 51 to 86% [88].

Although there is much new evidence testifying to the potential complexity of glycosyl halide reductive transformations, this evidence has not significantly influenced laboratory practices for the preparation of glycals from simple sugars. One paper [89] outlined a general method amenable to large-scale synthesis; the method involved a one-pot, three-step procedure involving peracetylation of an aldose, synthesis of a glycosyl bromide, and bromide reductive elimination. For D-glucose, the last step uses a Zn/AcOH/AcONa/$CuSO_4$ mixture that yields 98% 3,4,6-tri-*O*-acetyl-D-glucal. Another improved version of glycal synthesis, devised by Franck [90], uses vitamin B-12 as a catalyst for reductive eliminations [91]. With methanol as solvent, acetylated glycals were conveniently prepared by this method in the presence of metallic zinc and ammonium chloride in yields exceeding 90%. In the case of an L-rhamnal derivative, however, column chromatography had to be applied to isolate product in a 45% yield. Both the original and modified procedures of glycal preparation are reasonably tolerant of the inclusion of commonly applied functionalities, including glycosidic bonds. In fact, a number of glycals have already been prepared by different procedures from peracetylated disaccharides and trisaccharides [92,93,94,95,96,97,98,99,100].

3 Base-Catalyzed Eliminations of Glycosyl Halides to Yield 2-Oxy-hex-1-enitols

Peracylated glycosyl halides **3** are easily dehydrohalogenated to per-*O*-acyl-2-hydroxy-hex-1-enitols **20** (acylated 2-hydroxyglycals) under a variety of basic conditions (*Scheme 7*) [101]. The first transformation of this kind was recorded during the attempted *N*-glycosylation of secondary amines. As a result, diethylamine in benzene [102,103] or in chloroform [104,105] was used routinely in early preparations to effect such elimination. Since then, dehydrobromination has emerged as one of the competing reactions, and factors influencing their relative rates have been assessed [106]. For example, formation of 2-*O*-acyl-hex-1-enitols was often observed during attempted glycosylation of phenolates or carboxylates under basic (including phase-transfer) conditions. A reaction of 2,3,4-tri-*O*-acetyl-L-rhamnopyranosyl bromide with 4-methylbenzenethiol in the presence of 2,4,6-collidine in nitromethane resulted in a mixture containing not only the desired thioglycosides but also the elimination product (2,3,4-tri-*O*-acetyl-1,5-anhydro-6-deoxy-L-*arabino*-hex-1-enitol) and two isomeric thiophenyl

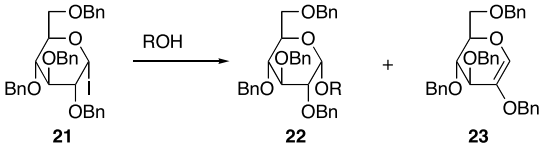


3 **20**

Scheme 7

orthoesters [107]. In the case of ethanethiol, glycosylation in the presence of triethylamine in acetonitrile resulted in elimination [108].

Other studies led to the successful yet sometimes unintentional isolation of glycals. For instance, heating acetylated glucopyranosyl bromide in nitromethane with nicotinamide gave peracetyl-2-hydroxy-D-glucal **21** along with both anomers of quaternary *N*-glycosides [109]. Studies of reactions of peracetylated -D-hexopyranosyl bromides with sodium 4-nitrophenoxide in dimethylformamide [110] revealed a strong correlation between sugar configuration and elimination product. Depending on the sugar, the yields of 2-acetoxyglycals diminished from 61 to 8% (*allo* > *gulo* > *gluco* > *galacto*), while the effectiveness of substitution with phenoxide anion followed the reverse order. A dispute over the scope of phase-transfer glycosylating reactions employing acylated glycosyl bromides and protic acid substrates led to the finding that application of 2,3,4,6-tetra-*O*-acetyl-D-glucopyranosyl bromide [111] or benzoylated glycosyl bromide [112] can give acylated glycals as common side products. For example, in phase-transfer reactions with a series of 44 phenolic substrates, acetylated α-D-galactopyranosyl bromide afforded the expected aryl β-galactopyranosides in yields ranging from 29 to 87% but no elimination product [113]; in contrast, during attempted glycosylation of a sugar substrate in the presence of silver trifluoromethanesulfonate and 2,4,6-trimethylpyridine, the benzoylated analog formed as much as 40% of the corresponding 2-*O*-benzoyl-D-galactal derivative [114]. When reacted with 1,2:5,6-di-*O*-isopropylidene-D-glucofuranose in the presence of diisopropylethylamine and tetrabutylammonium iodide in refluxing benzene, 2,3,4,6-tetra-*O*-benzyl-D-glucopyranosyl iodide **21** gave equimolar amounts of the expected product **22** (linked disaccharide) and the glycal **23** (◉ *Scheme 8*) [115].



R = 1,2:5,6-di-O-isopropylidene-α-D-glucofuranos-3-yl

◘ Scheme 8

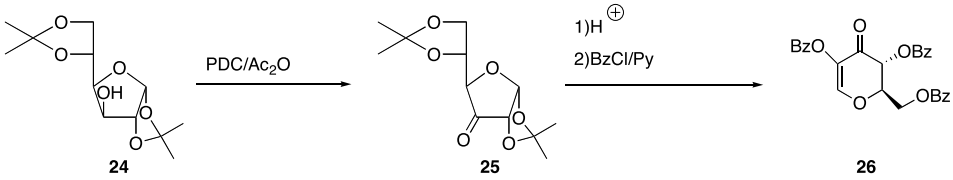
Such isolations of 2-*O*-substituted glycals were, however, unintentional. Therefore, more purposeful and effective methods and conditions for glycal preparation were sought. Excellent yields of disaccharide 2-*O*-benzoylated glycals were obtained by supplementing diethylamine with sodium iodide and carrying out the elimination reaction in acetone at ambient temperature [116]. Application of tetrabutylammonium bromide, as an addition to diethylamine, proved somewhat less effective [96,117,118]. When tetra-*O*-benzoyl-D-glucopyranosyl bromide was subjected to treatment with NaH in dimethylformamide, the corresponding glycal was isolated in yields of 52% [119]. Sodium hydride-promoted glycosylation of an enolic substrate gave a similar yield [120]. Cesium fluoride in acetonitrile was applied successfully for efficient elimination of acetylated glycosyl bromides with D- and L-*manno* configuration [121]. Introduction of 1,5-diazabicyclo bases as reagents for dehydrohalogenation [122] led to the establishment [123] of 1,8-diazabicyclo[5.4.0.]undec-7-ene (DBU) as a very effective agent for the elimination of hydrogen bromide from acetylated [124], benzoylated [125], and pival-

oylated [126] glycosyl bromides. In the case of 2-*O*-acetyl-3,4,6-tri-*O*-benzyl-D-glucopyranosyl bromide, DBU was applied, together with sodium iodide, in acetone [117].

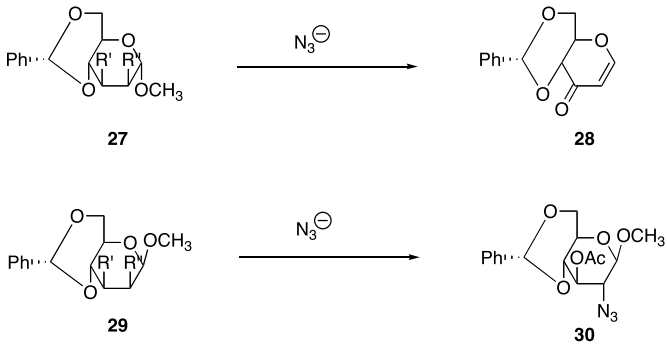
Obviously, the tendency to eliminate HBr from glycosyl bromides, or other protic acids from a sugar derivative bearing a good leaving group at an anomeric position, can be considerably enhanced by exploiting the possibility of forming a conjugated double bond system. Such an arrangement of functional groups exists in natural products (for example, in members of the neuraminic acid family [127,128,129], and it is well known that corresponding sialic acid glycals [130,131,132,133] can be prepared by treating peracetylated derivatives with substoichiometric amounts of trimethylsilyl triflate.

On the other hand, an electron-withdrawing functional group can be introduced in position C-3 of the sugar ring. A good example of this approach is the preparation of chiral pyranone synthon **26** from di-*O*-isopropylidene-D-glucofuranose **24** as shown below (*Scheme 9*) [134]. Other benzoylated 3-enones were obtained using the same approach [135].

Analogously, various 2-aminoglycals were obtained through oxidation of a C-3 hydroxyl group in 4,6-benzylidene derivatives of *N*-protected 2-amino-D-pyranosides [136]. 2-Acetamido-*O*-acetylated glycals, on the other hand, were obtained under acidic conditions, during reaction of protected 2-amino-2-deoxypyranosyl chlorides with propenyl acetate [137,138]. Attempted displacement of mannoside 2-triflate **27** with an azide anion resulted in the formation of enone **28**, while attempted displacement of the corresponding anomer gave, as expected, 2-azide **30** in the D-gluco configuration (*Scheme 10*) [139].

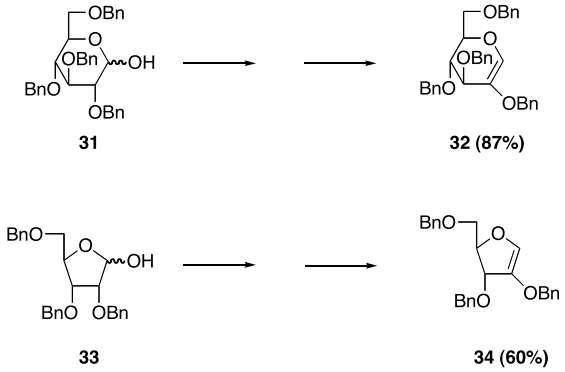


Scheme 9



Scheme 10

In one interesting case, hexopyranoses **31** and pentofuranoses **33** protected with nonparticipating groups were formally dehydrated in a two-step process involving preparation of anomeric mesylate and palladium catalysis [140]. The researchers who observed this case, however, ruled out the simple explanation of bases assisting in Bronsted acid removal. Instead, they claimed that treatment of an intermediate anomeric mesylate with 0.9–5% $Pd(Ph_3P)_4$ at 50 °C (in a solution of methylene dichloride and *s*-collidine) resulted in oxidative addition followed by hydride elimination, consequently affording *O*-protected glycals **32** and **34** (*Scheme 11*).



Scheme 11

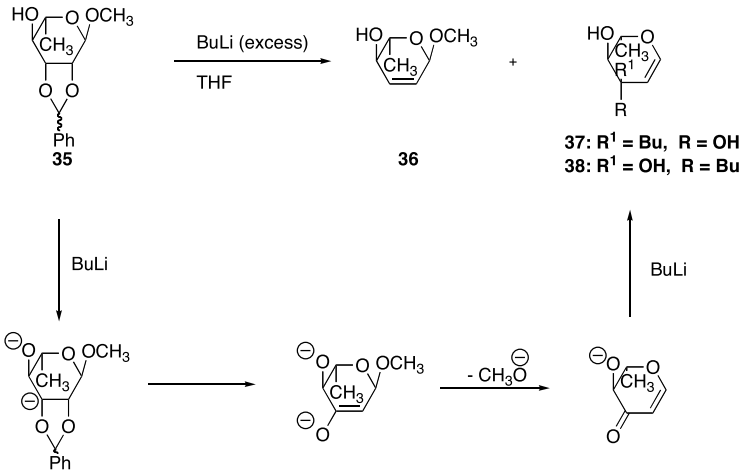
4 Eliminations Induced by Alkyllithium Reagents

Deprotonation of heterosubstituted carbon in protected sugar derivatives, as well as deprotonation of a hydroxyl group, can start anion fragmentation that, as illustrated below, can lead to unsaturated products. Subsequent deprotonation of glycals at C-1 is also possible and has already been expertly covered in a monograph [141]. This section, in addition to discussing the synthesis leading to glycals, will also discuss deprotonation at C-1 as a general approach to obtaining a great variety of C-1 heterosubstituted and C–C linked products.

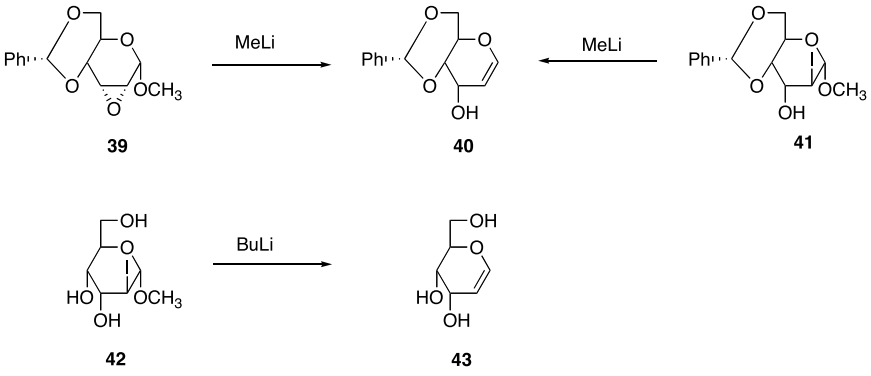
During a study devoted to fragmentation of sugar acetals in the presence of butyllithium [142], the formation of unsaturated pyranoses (**36**–**38**) from L-rhamnose derivative **35** was observed and explained in terms of anionic species rearrangement (*Scheme 12*).

In a related process, D-allal derivative **40** was first obtained [143] from the *allo* epoxide **39** by treatment with methyllithium. Subsequently, it was demonstrated that 2-iodo-D-altroside [144] **41** and its deprotected analog [145] **42** gave even higher yields of the allal (**40** and **43**, respectively) when treated with an excess of alkyllithium (*Scheme 13*). Incidentally, benzylidene derivative **40** could not be deprotected to **43** without decomposition.

2-Deoxy-2-iodopyranosides constitute a class of synthetic sugar derivatives that are useful intermediates for preparation of 2-deoxy glycosides, which in turn are natural compounds of considerable medicinal importance [15]. These iodopyranosides can be conveniently prepared by using *N*-iodosuccinimide (NIS)-assisted addition of hydroxylic substrates to glycals or using oxirane ring opening by iodine anion of 2,3-anhydro pyranosides [12]. The resulting



Scheme 12

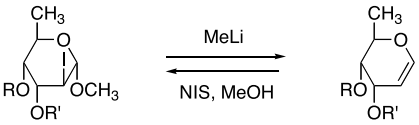


Scheme 13

sugar iodohydrines and their acylated derivatives can be transformed into unsaturated sugars using methyllithium [146,147]. The formal reversibility of NIS-promoted addition to glycals can be illustrated by the transformation of D-digitose-derived disaccharide synthons **44** and **45** (*Scheme 14*).

Only one-step is required to derive glycals from antibiotic sugars. For example, L-mycarose and L-olivose can be derived in one step from methyl 2,3-benzylidene-L-rhamnoside by applying the same principle of acetal fragmentation in reaction with a 6-fold excess of methyllithium [148]. Glycosyl sulfoxides with nonparticipating protecting groups require only 3 equivalents of butyllithium (at −78 °C in THF) for smooth conversion to the corresponding glycals [88].

As these examples indicate, introduction of a double bond via deprotonation of sugar derivatives having nonparticipating protecting groups can be synthetically useful. In practical syn-



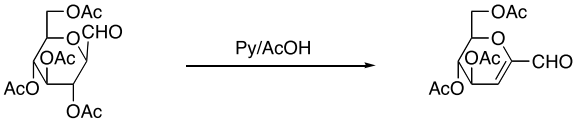
44 **45**

R - monosaccharide
R' - H, Ac

Scheme 14

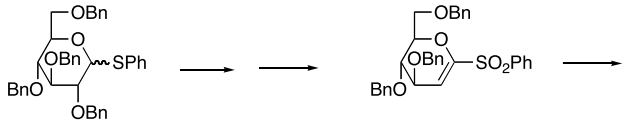
theses of glycals, deprotonation is facilitated by placing a strong electron-withdrawing group at the anomeric position. For example, *C*-glycosyl derivative **46** undergoing acetic acid elimination upon treatment with pyridine at room temperature easily affords **47** in a 75% yield (*Scheme 15*) [149].

Easily available anomeric sulfones, which have already been mentioned as glycal precursors, readily undergo base-induced eliminations resulting in formation of 1,2-unsaturated pyranose derivatives in which the sulfonyl group has been retained [150,151]. Such compounds have considerable synthetic potential, particularly for C–C bond formation. For example, compound **49** can be lithiated with LDA/THF at 80 °C to afford C-2 branched glycals **50** (*Scheme 16*) [152].

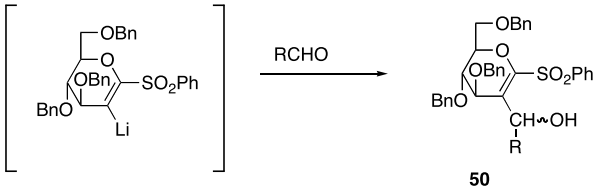


46 **47**

Scheme 15



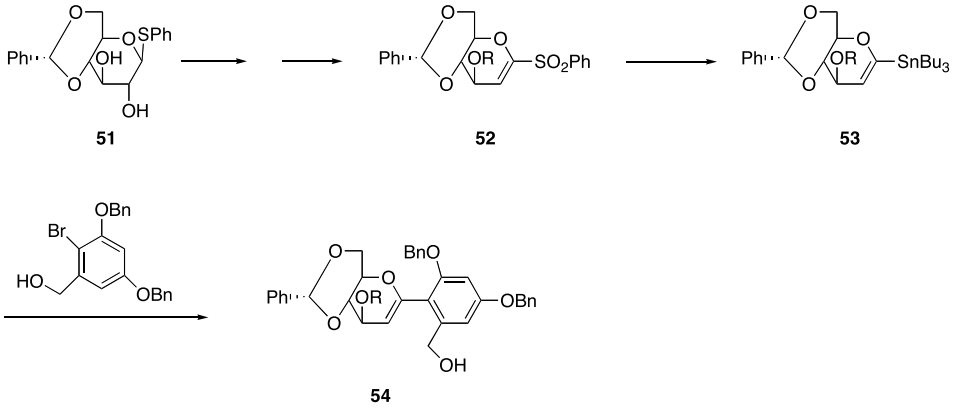
48 **49**



50

Scheme 16

Sulfonylated glycals are also useful precursors of 1-unsaturated *C*-glycosyl compounds. For example, treatment of **52** with an excess of tri-*n*-butyltin hydride in the presence of a free radical initiator (AIBN) in refluxing toluene leads to exchange of the anomeric sulfonyl group resulting in 1-stannyl glycal **53**, which in turn can be coupled with bromoaryl substrates in a palladium-catalyzed reaction to **54** (● *Scheme 17*) [153].



● **Scheme 17**

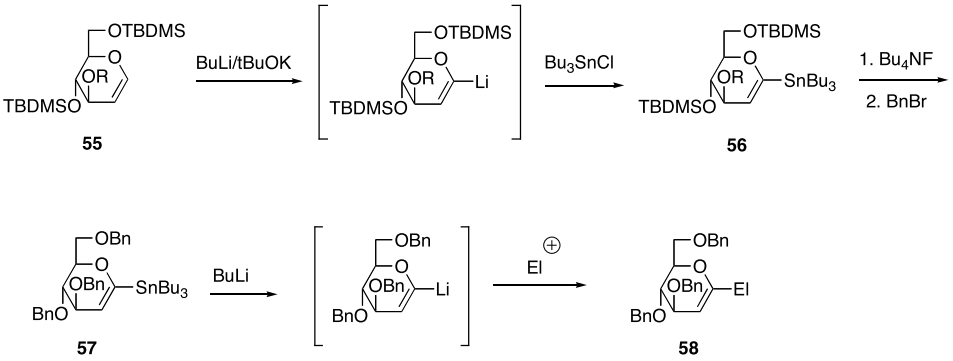
Since stannylated glycals are of particular interest as prospective components of Pd(0)-catalyzed cross-couplings (i. e. Stille reactions) [154], their preparation deserves some attention. Reactions involving the direct deprotonation of glycal derivatives, which require *tert*-butyl lithium, were first reported in 1986 [155,156]. Use of silyl ethers as protective groups is considered critical for the success of this transformation. In fact, one report recommends using triisopropylsilyl protective groups and 4 equivalents of *tert*-butyllithium for best results [157]. As shown below, lithiated glycals can be trapped by a variety of electrophiles, including tributylstannyl chloride. The stannyl derivatives **56** thus formed can be used to generate reactive intermediates that are more convenient than primary lithiation products, via exchange of protective groups for more generally compatible benzyl ethers (● *Scheme 18*) [141].

Lithiated glycals can also be transformed into more stable and more synthetically useful organozinc reagents [158]. Nevertheless, as has already been shown, even the primary lithiation products of the furanoid glycals **59** can be efficiently trapped by appropriate electrophiles to **60** (● *Scheme 19*) [159].

5 Thermal Elimination of Sulfoxides and Selenoxides to Glycals

Sulfoxides and selenoxides are known to easily undergo beta-elimination reactions. Sugar sulfoxides obtained by oxidation of respective thioglycosides were efficiently transformed to glycals upon heating in toluene (● *Scheme 20*) [160].

Analogous oxidation of selenoglycosides has been used to directly synthesize glycals due to the low temperature spontaneous *syn*-elimination of selenoxides [161,162].



R = TBDMS
El = Me, CH(OH)Ph etc.

Scheme 18

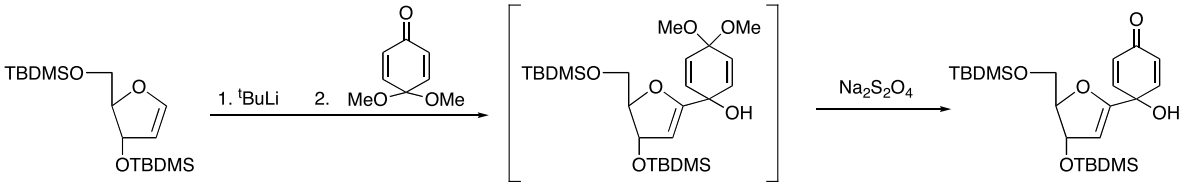
Furanoid glycals utilizing elimination of selenoxides were prepared from unusual substrates, the 4-phenylselenyltetrahydrofurans. These intermediates were obtained with good yields from D-glyceraldehyde (*Scheme 21*). Their oxidations have led to selenoxides, however, the elimination reaction required heating. The best yields were noted for reactions refluxed in 1,2-dichloroethane and yields ranged from 62 to 95%.

6 Glycals from 2-Deoxysaccharides

Methods for the preparation of glycals from 2-deoxy sugars, although very efficient, are limited by the availability of 2-deoxy sugars. Very often the best substrates for 2-deoxy sugars are glycals themselves. Therefore, such methods are not practical for the preparation of simple glycals, which can be easily obtained directly from commercially available monosaccharides. However, such approaches are very useful in the preparation of glycals of rare sugars or glycals of naturally occurring 2-deoxy sugars. Modified glycals are often needed for the preparation of biologically important glycosides, oligosaccharides, and sugars modified at C-2.

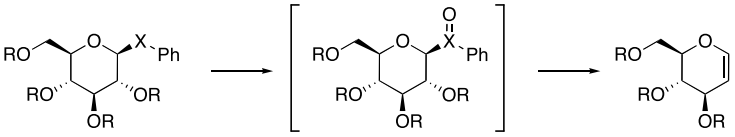
1-*O*-esters of 2-deoxyhexopyranoses or respective glycosyl halides appear to be useful substrates in the synthesis of glycals. When added to a silica gel suspended in dry xylene and upon heating, 1,4-di-O acetyl-*N*-trifluoroacetyl-α,β-daunosamine **68**, formed the respective glycal **69** (4-*O*-acetyl 1,5-anhydro-2,3,6-trideoxy-3-trifluoroacetamido-L-lyxo-hex-1-enitol) in 60% yield (*Scheme 22*) [163]. In fact, the whole spectrum of 3-aminoglycals was prepared from 2-deoxy sugars starting from either esters or glycosyl halides. These transformations along with other approaches to 3-aminoglycals are described in detail by Pelyvas et al. [164].

Biologically interesting unsaturated derivatives of *N*-acetylneuraminic acid (Neu5Ac), a potent inhibitors of sialidase, were also obtained from peracetylated substrates. Unsaturated peracetate methyl ester of Neu5Ac2en **71** was obtained in over 90% yield via the $CF_3SO_3SiMe_3$-mediated elimination of acetic acid from **70** (*Scheme 23*). Other unsaturated products were similarly prepared [133].



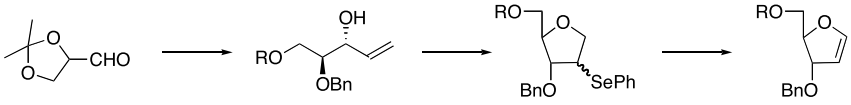
59 **60**

Scheme 19



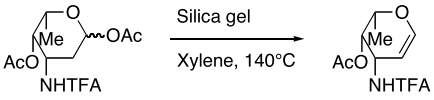
61 62 63

Scheme 20



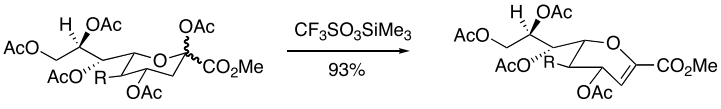
64 65 66 67

Scheme 21



68 69

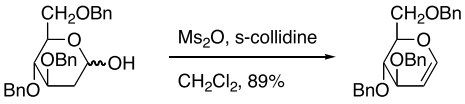
Scheme 22



70 71

R = NHAc

Scheme 23

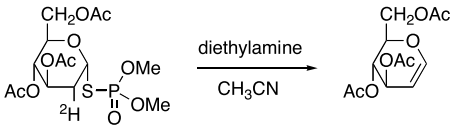


72 73

Scheme 24

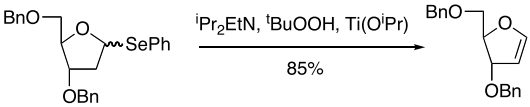
Glycals can also be obtained in a one-pot reaction from 2-deoxyhexopyranoses having a free hydroxyl at C-1. An example of such a transformation is high-yield (89%) synthesis of tri-*O*-benzyl-D-glucal **73** from 2-deoxy-3,4,6-tri-*O*-benzyl-D-glucopyranose **72** via 1-*O*-mesylate (*Scheme 24*) [165].

An interesting demonstration of base-catalyzed *cis* elimination leading to 3,4,6-tri-*O*-acetyl-glucal (**4**) used O,O-dimethyl-*S*-(3,4,6-tri-*O*-acetyl-2(R)-2-deoxy-2-deuterio-α-D-arabinohexopyranosyl)phosphorodithioate **72** (*Scheme 25*) [166].



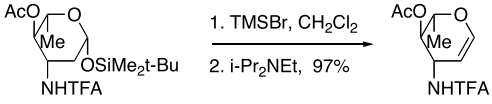
72 4

Scheme 25

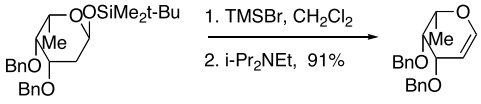


75 76

Scheme 26



77 78



79 80

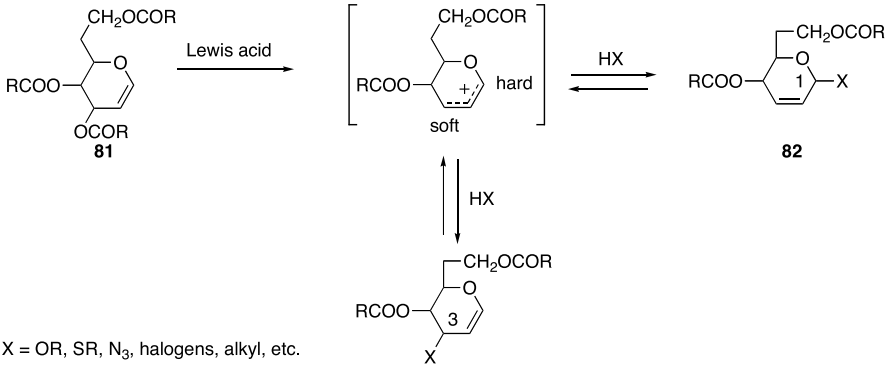
Scheme 27

An elimination reaction was also successfully used in the synthesis of furanoid glycals from 2-deoxyribose. The substrates used were 2-deoxy-1-seleno-furanosides **75**, which were transformed to their respective glycals **76** via selenoxide elimination in yields ranging from 71 to 85% (*Scheme 26*) [167].

A one-pot, two-step reaction leading to glycals from 2-deoxyhexopyranoses silylated or acetylated at the anomeric center or from 2-deoxyhexopyranosides can be considered a method of choice because of its high yield and simplicity. The two-step reaction includes addition of trimethylbromosilane (TMSBr) to **77** or **79** (generation of glycosyl bromide) and subsequent addition of *N*,*N*-diisopropylethylamine to cause elimination to glycals **78** and **80**, respectively(*Scheme 27*). In all cases, yields of isolated product exceeded 90% [163,168,169].

7 Reactions Involving 2,3-Unsaturated Saccharide-to-Glycal and Glycal-to-Glycal Rearrangements

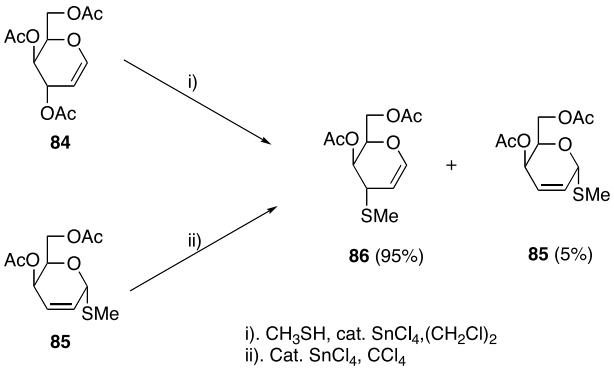
Esterified glycals **81** easily undergo reaction with nucleophiles, typically under Lewis acid catalysis to 2,3-unsaturated pyranoses **82**. This reaction is known as Ferrier rearrangement [7].



Scheme 28

The reaction initially leads to the formation of carbenium ions leaving open the possibility of attack at C-1 or C-3. Even though initial studies showed that reaction of glycals with alcohols led to 2,3-unsaturated products but no 3-*O*-alkylated glycals **83**, later studies explained the regioselectivity of the Ferrier rearrangement and clearly demonstrated the possibility of selecting attack at the C-3 carbocation to form C-3-modified glycals with good yields [170]. Use of the reverse Ferrier rearrangement of 2,3-unsaturated pyranose derivatives to produce glycals **82** to **83** is not widespread but should nevertheless be considered along with glycal-to-glycal transformation **81** to **83** (*Scheme 28*).

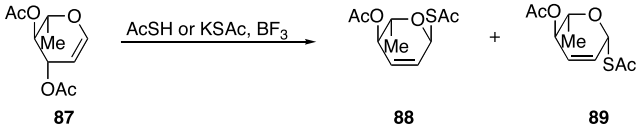
By testing nucleophilic substrates categorized according to Pearson's Theory of Hard and Soft Acids and Bases (HSAB) [171,172] Priebe and Zamojski proposed that the regioselectivity of the Ferrier rearrangement reaction could be rationalized in terms of the HSAB principle, with the carbenium ion center at C-1 being hard and C-3 being soft [170]. They demonstrated that hard bases preferentially form 2,3-unsaturated products (via attack at C-1) following typical



Scheme 29

Ferrier rearrangement prediction, that soft bases preferentially form bonds with a carbenium ion at C-3, leading to novel glycals modified at C-3 (see below the synthesis of 3-thiomethyl glycal **86**), and that borderline bases can form both C-1 and C-3 substituted products [170]. Priebe and Zamojski also clearly demonstrated that reaction of glycals with nucleophiles is reversible and that substrates, as well as kinetically controlled products, are prone to generate intermediate carbenium ions before attaining thermodynamic equilibrium [170]. Thus, alcohols and other *O*-nucleophiles end up as anomeric (a hard acidic center) substituents, while *S*-(**86**) [170], *N*- [173,174], and C- [175] anions tend to form a new bond at C-3 (soft center), which leads to substituted glycals (*Scheme 29*).

Further support for the HSAB-based rationalization comes also from later studies demonstrating that making sulfur harder (*S*-alkyl to *S*-acetyl) dramatically alters regioselectivity and leads to 2,3-unsaturated products [176,177]. In the reaction from 3,4-di-*O*-acetyl-L-rhamnal **87** with thioacetic acid or potassium thioacetate in the presence of $BF_3 \cdot Et_2O$, only products substituted at C-1 **88** and **89** were isolated (*Scheme 30*) [177].

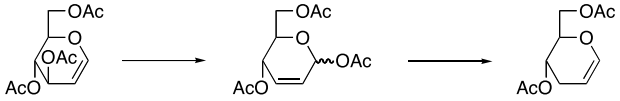


87 **88** **89**

Scheme 30

In the 1970s and 1980s, unsaturated pyranosyl azides were studied extensively because of their potential use in synthesis of aminated antibiotic sugars [175,178,179,180,181,182]. These studies showed that both regioisomers were being formed. This can be explained by the borderline nature of azides according to the principles of HSAB. It is interesting to note that, according to an HASB-based rationalization, a hydride anion will, as a soft base in reaction with 2,3-unsaturated pyranosides, produce 3-deoxyglycals [183,184,185].

Once the stereospecific character of hydride delivery was demonstrated in experiments with deuterated LAH, the reductive removal of glycosidic functions were subsequently used to prepare pyranosyl dienes [186], which are precursors of *C*-disaccharides. The usefulness of this transformation even extends to tertiary C-4 branched hex-2-enopyranosides, provided the leaving oxygen substituents of the substrate are in a relative 1,4-*cis* configuration [187]. A similar type of reductive rearrangement was observed [188] during catalytic hydrogenation of acetylated hex-2-enopyranose **90** over platinum in the presence of amine. Under carefully selected conditions, the 3-deoxy glycal **91** could be isolated in 83% yield using a substrate obtained by $ZnCl_2$-catalyzed rearrangement of **4** in acetic anhydride (*Scheme 31*).



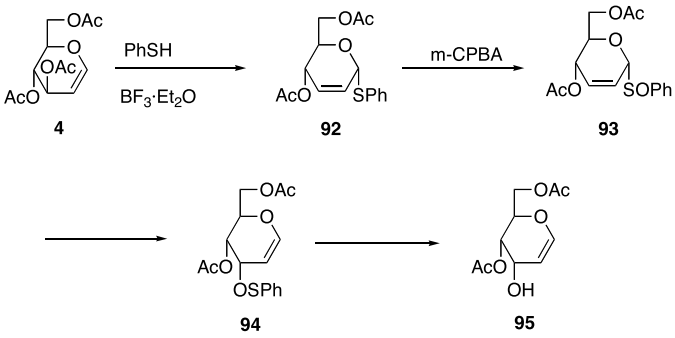
4 **90** **91**

Scheme 31

Subsequently, it has been demonstrated that, under aprotic reaction conditions, various metal chlorides can catalyze not only double bond migration but also C-3 epimerization [189].

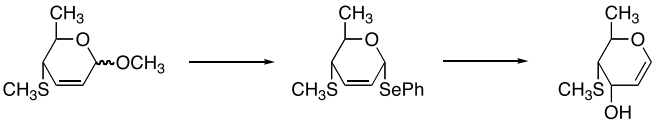
Some caution is required in predicting regioselectivity in analyzing Ferrier type rearrangements that are not catalyzed by Lewis acid or that occur under conditions that favor kinetically controlled products. For example, nucleophilic agents can sometimes react with glycals in the absence of Lewis catalysts to produce 2,3-unsaturated products [2]. Thus, as observed in one study, condensation of 3,4-bis-*O*-(4-nitrobenzoyl)-D-xylal with purines and pyrimidines without an acid catalyst led to anomeric mixtures of 2′,3′-unsaturated nucleosides [190]. Moreover, **4** can react with thiophenol to give kinetically favored phenyl 4,6-di-*O*-acetyl-2,3-dideoxy-1-thio-D-*erythro*-hex-2-enopyranoside **92** [191] and only a minor admixture of the expected 3-regioisomer. The tendency of this unsaturated thioglycoside to undergo rearrangement was exploited in order to synthesize 1-*C*-substituted glycals and related bicyclic spiroketals [192]. Also, oxidation of the phenyl thioglycoside **92** with m-CPBA to **93**, followed by treatment with piperidine, led to the allal derivative **95**, apparently via sulfenate **94** as intermediate (● *Scheme 32*) [193]. However, because m-CPBA can give rise to complex mixtures upon oxidation of unsaturated thioglycosides, use of 3,3-dimethyldioxirane, which cleanly affords the desired sulfoxides, is recommended [194].

The same synthetic approach was applied in obtaining a unique thiosugar and constituent of ene-diyne antitumor antibiotics; in contrast, unsaturated glycosides **96** were treated with [195] phenylselenol at −90 °C in the presence of $BF_3{\cdot}Et_2O$ to give mainly unsaturated glycoside **97** [196,197], which in turn was rearranged into glycal **98** after oxidation with m-CPBA (● *Scheme 33*). An analogous sequence of reactions employing a dinitrophenyl thiol protective group has been used to obtain 2,3-unsaturated thiophenyl glycosides [198].



4 **92** **93** **94** **95**

Scheme 32



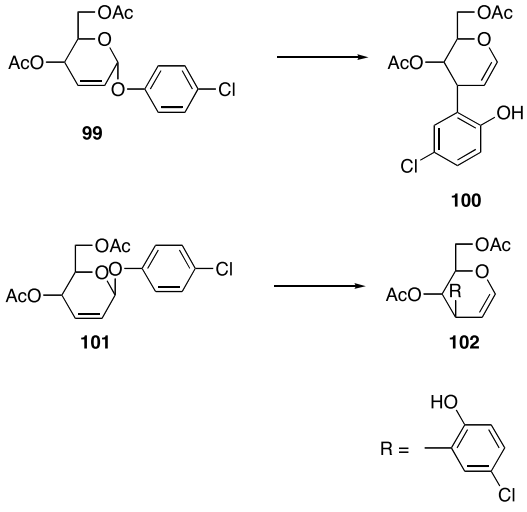
96 **97** **98**

Scheme 33

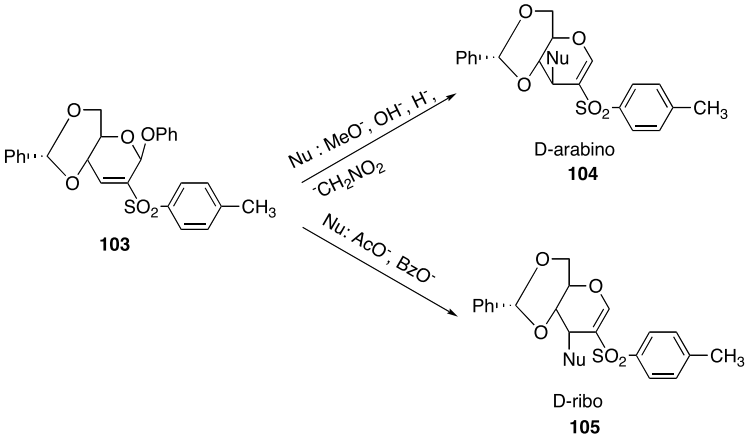
The above remarks concerning the regioselectivity of the Ferrier rearrangement apply also to reactions with *C*-nucleophiles but not to reactions with C-1-substituted kinetic products since such kinetic products are not susceptible to equilibration and therefore become final products. Anomeric allylation of glycals [199] and their reactions with enol ethers [200] are good examples of this. On the other hand, reaction of **4** with dimethylcuprate afforded mainly the S_N2 substitution product [201]. However, reaction with a number of *C*-nucleophiles allowed formation of all four possible alkylation products [202,203]. 2,3-Unsaturated vinyl glycosides were shown to exhibit regioselective *C*-alkylating properties depending on the conditions under which they were applied. Lewis acid catalysis afforded anomeric mixtures of *C*-glycosides, while thermal Claisen rearrangement [204] afforded stereospecific 3-*C*-branched glycals [205,206,207]. An analogous transformation of aryl glycosides (**99** and **101**) was achieved by heating substrates in *N*,*N*-diethylaniline [208]; as a result and as expected, the aryl substituents remained on the same face of the pyranoid ring (**100** and **102**) (*Scheme 34*). Reaction rates differed markedly between the stereoisomers. For example, 1,4-*cis* substrates took almost 100-times longer to complete their migration.

In other experiments, a range of nucleophilic agents were reacted with sulfonyl hex-2-enopyranoside **103**, and corresponding glycals **104** and **105** were obtained stereoselectively (*Scheme 35*) [209]. In this case, C-3 stereochemistry was regulated by the nature of the incoming anion.

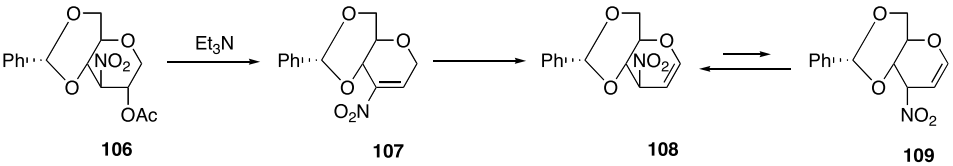
The stereochemistry of 3-*C*-nitro glycals has been studied in some detail [210]. For example, when nitroanhydroglucitol **106** was subjected to reaction with triethylamine, it gave the elimination product **107**, which was then rearranged to an equilibrated mixture of glycals **108** and **109** (*Scheme 36*) [211,212]. Some researchers have tried to explain this equilibrium shift by arguing that the quasi-equatorial anomeric proton is made more acidic by the stereoelectronic effect [213].



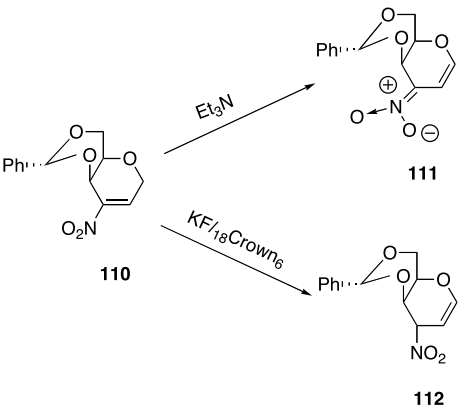
Scheme 34



Scheme 35

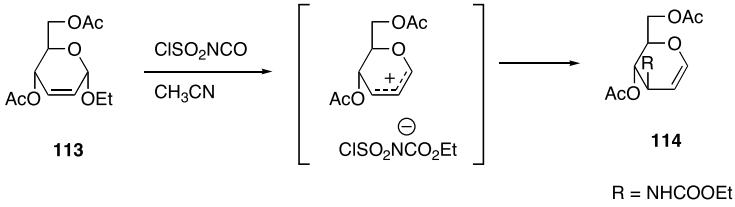


Scheme 36



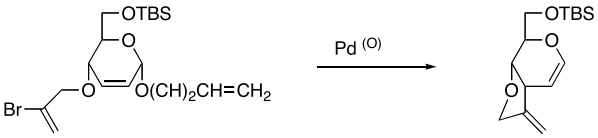
Scheme 37

In the case of the D-*threo* isomer **110**, nitronate **111** was isolated under the same conditions, while 3-nitroglycal **112** was isolated under conditions that included treatment with KF in the presence of 18-crown-6-ether (*Scheme 37*).



113 114

Scheme 38



115 116

Scheme 39

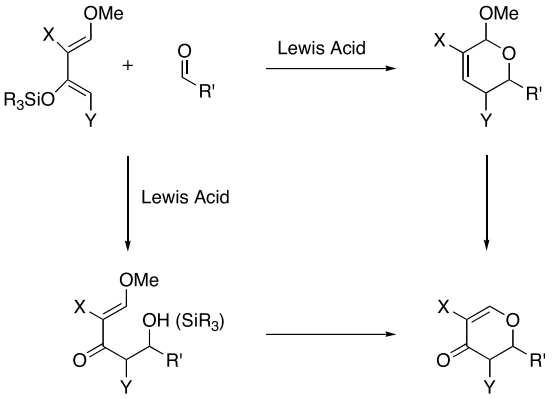
3-Aminoglycal derivatives are of special interest in connection with the synthesis of sugar constituents of antibiotics [164], and their preparation is covered above in *Sect. 6*. However, reactions of hex-2-enopyranosides with chlorosulfonyl isocyanate [214,215], which involve double bond migration, (**113** to **114**), should also be mentioned since they are one-step procedures of considerable synthetic potential (*Scheme 38*).

In recent years unsaturated sugars have evoked much interest as chiral templates for assembling polycyclic systems in a stereocontrolled way (e. g., by palladium-catalyzed transformation). An example of palladium-catalyzed transformation involving the formation of a 1,2-double bond **116** from a 2,3-unsaturated precursor **115** is given below (*Scheme 39*) [216,217]. Many more "annulated sugars" of the same general structure have been obtained by using an analogous methodology [218].

8 Cycloadditions and Cyclizations

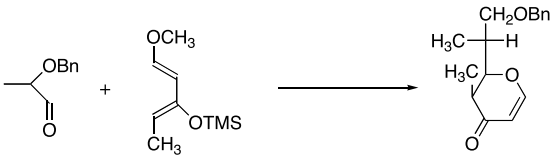
The potential of the hetero Diels–Alder reaction as a means of synthesizing sugar derivatives has long been recognized and explored [219,220]. More recently, introduction of a new generation of electron-rich dienes [221] and stereoselective catalysts of cycloaddition [222] have expanded the usefulness and practicality of this approach. The reaction of 1-alkoxy-3-siloxy-1,3-butadienes with aldehydes shown below has been studied in great detail with respect to the influence exerted by the structural features of the reagents, as well as by the catalysts (*Scheme 40*). Two general pathways—cycloaddition and aldol-like condensation followed by heterocyclization—have been identified [223,224], and both are relevant to glycal chemistry.

Indeed, pent-1-enitols [225], glucal derivatives [226], and galactal and fucal analogs have all been obtained via 1-en-3-ones by using the above-mentioned approach. Thermal cyclocondensation of aldehyde **117** with diene **118** gave racemic enone **119** (*Scheme 41*), which is used as a synthon in the preparation of the rifamycin S chain [227].



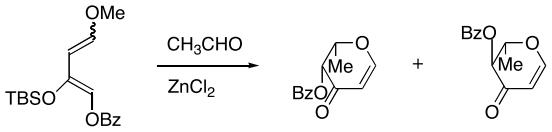
R', X, Y, = various unspecified substituents

Scheme 40



117 **118** **119**

Scheme 41

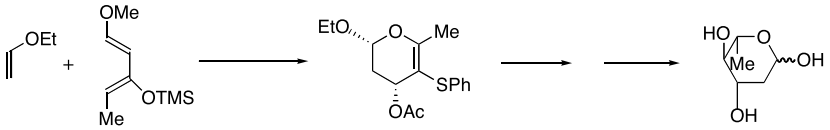


120 **121** **122**

Scheme 42

Reaction of isomeric dienes **120** with acetaldehyde in the presence of zinc chloride gave a mixture of C-4 epimeric enones **121**, **122** (*Scheme 42*) [224], which are convenient substrates for synthesis of all four racemic 3-amino-2,3,6-trideoxy pyranoses.

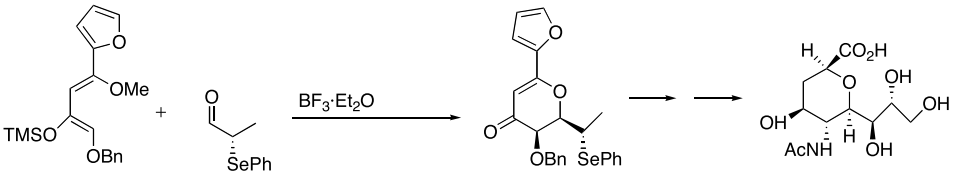
Similarly, reaction of 1,4-diacetoxy-1,3-butadiene with methyl glyoxalate afforded glucuronate glycal [228] and a 2,3-unsaturated isomer. Cycloaddition reactions performed on substrates with inverse electronic properties (e. g., enol ether as a dienofile and unsaturated carbonyl compound as a heterodiene) afford not only the expected products but also ones having high *endo* selectivity [229,230] as exemplified below by olivose **126** synthesis (*Scheme 43*).



123 **124** **125** **126**

endo : exo > 10 : 1

Scheme 43

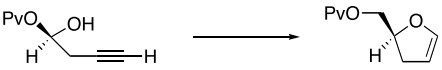


127 **128** **129** **130**

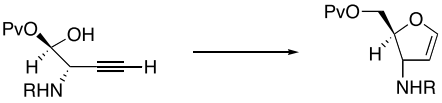
Scheme 44

Studies of hetero Diels–Alder condensation that used enantiopure substituted aldehydes found that the catalyzed additions had a high and predictable stereoselectivity [231]. The synthesis of *N*-acetylneuraminic acid **130** [232] from diene **127** and (S)-selenoaldehyde **128**, which is shown below, illustrates this point (*Scheme 44*).

A brand-new methodology for synthesizing glycals from noncarbohydrate precursors, one based on cyclization of acetylenic alcohols, has emerged from the field of metalorganics. Molybdenum pentacarbonyl-trialkylamine complexes have been found to efficiently catalyze cyclization of 1-alkyn-4-ols to substituted dihydrofurans [233,234]. This same transformation has been successfully carried out on asymmetrically substituted alcohols; the furanoid glycals **132**, **134**, and **135** (*Scheme 45*) so obtained have in turn been used as intermediates in the synthesis of nucleosides [235].



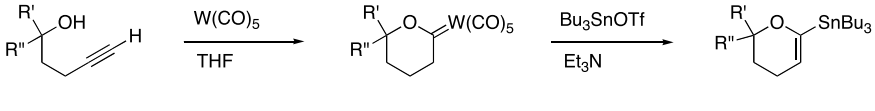
131 **132**



133

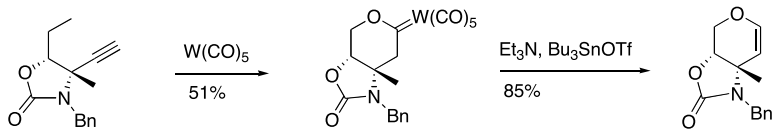
134 R = CH_3CO (89%)
135 R = CF_3CO (92%)

Scheme 45



136 137 138

Scheme 46



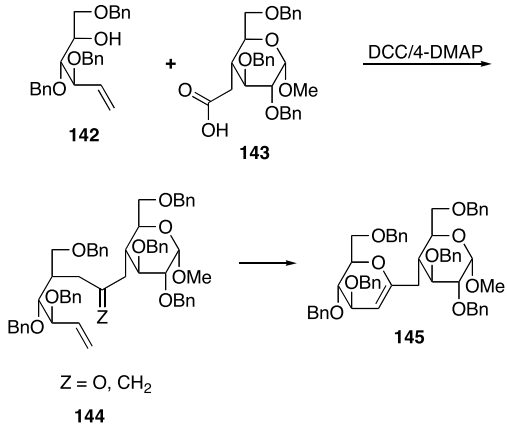
139 140 141

Scheme 47

The transformation has been extended even further to pyran derivatives by using a tungsten pentacarbonyl-tetrahydrofuran complex. Primary, secondary, and tertiary 1-alkyn-5-ols **136** undergo cyclization to dihydropyranylidene carbenes, which can then be converted into the corresponding stannyl dihydropyranes **138** upon treatment with tributyltin triflate and triethylamine (*Scheme 46*) [236].

3-Aminosugar-derived glycal **141** was obtained by this method from an appropriate chiral precursor **139** [237]. Unexpectedly, however, reaction of tungsten intermediate **140** with tri-*n*-butyltin triflate and triethylamine in ether gave the glycal **141** (in 85% yield) instead of the 1-stannyl derivative (*Scheme 47*). The molybdenum carbonyl-triethylamine complex afforded the same product, although in a much lower yield.

Glycals can also be obtained from suitable substrates, by ring-closing olefin metathesis reactions. In a general approach to variously linked *C*-disaccharides, illustrated in *Scheme 48*, nonreducing-end glycals are typical intermediates, finally subjected to hydroboration or dihydroxylation to afford the desired *C*-analogs of *O*-disaccharides [238,239].



142 143 144 145

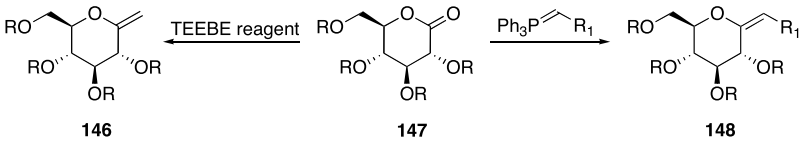
Scheme 48

9 Preparation of exo-Glycals

Exomethylene monosaccharide derivatives became of interest in connection with increased application of *C*-glycosyl synthons in the preparation of glycomimetics intended as glycosidase inhibitors and glycoconjugates. Their syntheses and uses have been comprehensively reviewed [240].

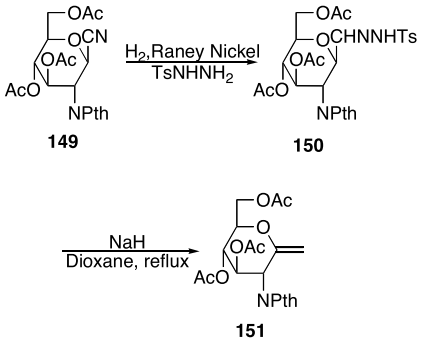
Sugar lactones are the most convenient substrates in the synthesis of exo-glycals. Their methylenation using Tebbe's reagent has been efficiently used but application of the Wittig reaction also allowed for the preparation of a series of exo-glycals (*Scheme 49*) [241,242]. Suitably protected sugar lactones can also easily undergo addition of nucleophiles, with creation of a ketosyl anomeric center from which the hydroxyl group can be easily eliminated to exo-glycals by treatment with trifluoroacetic anhydride in the presence of pyridine [243]. The same group of authors described simple and general synthesis of conjugated *exo*-glycals. In the key step β-C-1 carbonyl sugar derivative was reacted with Wittig reagent followed by phenylselenyl chloride to place the substituent which undergoes easy elimination after oxidation to sulfoxide [244]. Glycosyl cyanides (2,5-anhydroaldonitriles) are easily converted into corresponding aldehyde tosylhydrazones, which in turn undergo base-catalyzed decomposition to *C*-glycosylmethylene carbenes and eventually to *exo*-glycals (*Scheme 50*) [245,246]. These reaction conditions are compatible with the presence of acyl protecting groups.

Application of Ramberg–Bäcklund rearrangement provides a very effective approach to exo-glycals, in which α-halogenated sulfones are converted into olefins, with concomitant expulsion of sulfur dioxide. Since thioglycosides are easily available and there are several methods for their efficient oxidation to sulfones, this method seems to offer a wide scope, frequently



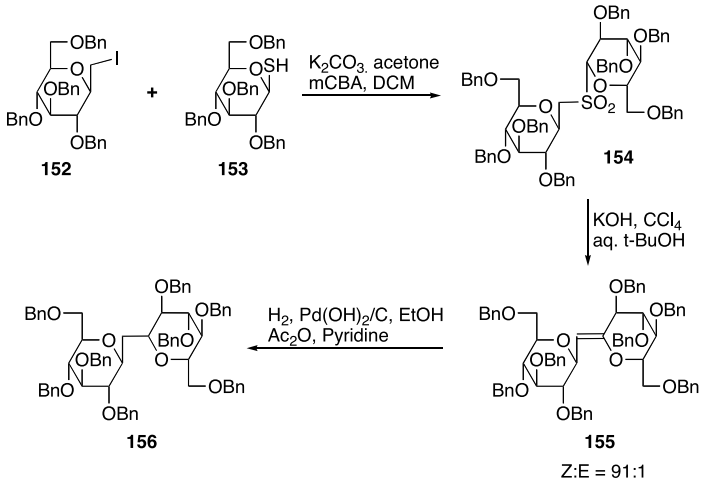
146 147 148

Scheme 49



149 150 151

Scheme 50



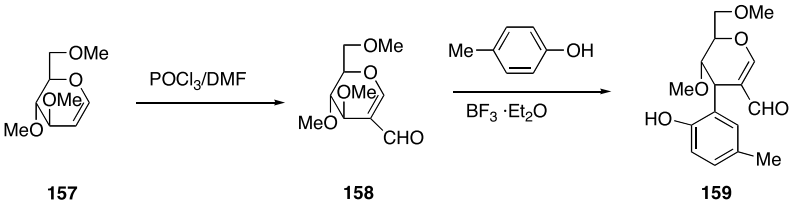
Scheme 51

lacking in the case of previously described methods. In particular, this method is suitable for the synthesis of *C*-linked disaccharides and their analogs via reaction of 1-thio-sugars with iodo-sugars as shown in *Scheme 51*.

The application of the Ramberg–Bäcklund reaction for the synthesis of *C*-glycosides and related compounds has been recently discussed in detail [247].

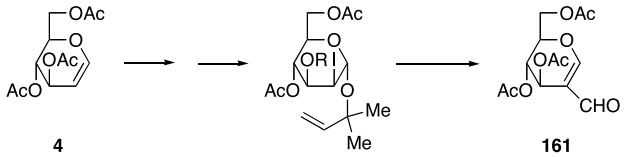
10 Miscellaneous Methods

Specifically substituted glycals, which are often used as chiral synthons or key substrates, can be prepared in many different ways. Double bond formation frequently precedes the assembly of the complete carbon framework designed for a cyclic enol ether intermediate. Of particular interest here are glycals substituted at C-1 and C-2, which are obtained by functionalization of the existing double bond. For example, C-2-acylated glycals offer a variety of ways to generate synthetically useful transformations. C-2 formyl glycals were first obtained by syntheses based on dithiane [248] and enol ether [249]. Exhaustive hydrogenation of the enone, which leads to C-2-methyl pyrane, has been applied as the final step in the total synthesis of the antibiotic restricitin [250]. In an alternative approach, direct formylation of methyl **157** or benzyl ethers of glycals with Vilsmeier–Haack reagent has been achieved [251], and it has been shown that these products **158** can undergo BF_3/Et_2O catalyzed C-3 arylation to **159** (*Scheme 52*) [252]. The C-2 formyl glycals have an obvious connection, through Wittig chemistry, with C-2 vinyl glycals. The C-2 vinyl glycals, in their turn, have been studied as the diene components of Diels–Alder cycloaddition reactions [253]. Interestingly, C-2 formylation can also be achieved by free radical-promoted cyclization-fragmentation of suitably arranged pyranose derivatives. For example, in one study, 2-deoxy-2-iodo– mannoside **160** was converted into formylated glycal **161** [254] in 65% yield (*Scheme 53*), while the corresponding benzylated glycoside afforded saturated C-2 branched products.



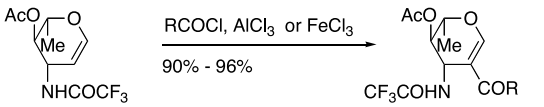
157 **158** **159**

Scheme 52



4 **160**: R = Ac **161**

Scheme 53



162 **163**

R = alkyl; C1 - C17

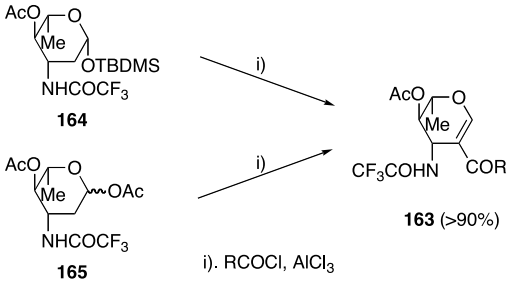
Scheme 54

Although the general scope of direct acylation of olefins under Friedel–Crafts conditions is rather limited, it has been demonstrated that such reactions can be applied to glycals like **162** to produce C-2 acylated glycals **163** in excellent yields and selectivity (*Scheme 54*). Moreover, as reagents, the long-chain fatty acid based acylating species have been shown to be as effective as acetyl chloride [255].

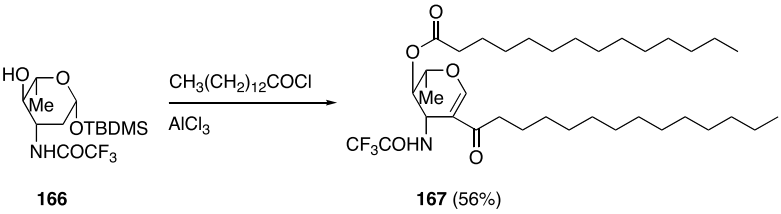
Interestingly, such C-acylated glycals **163** can also be obtained in a one-pot reaction from 2-deoxyhexopyranoses **164** and **165** in surprisingly good yield (*Scheme 55*). Friedel–Crafts acylation is only one of several sequentially occurring reactions [255].

This approach also allowed preparation, in a one-pot reaction, of the 2-C and 4-O acylated glycal **167** [255] starting from 2-deoxyhexopyranose **166** having a free hydroxyl at C-4 (*Scheme 56*).

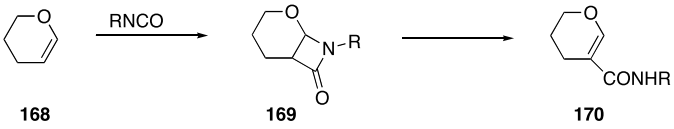
The addition of isocyanates to glycals, which has been studied thoroughly in connection with new methods of synthesizing lactams [256], usually affords a mixture of [2+2] and [4+2] cycloadducts. These primary reaction products are rather unstable and slowly rearrange to unsaturated amides, which can often be isolated directly from reaction mixtures [257,258]. In the case of simple dihydropyran derivatives **168**, unsaturated *N*-substituted C-2 formamides **170** are the only isolable products (*Scheme 57*) [259,260].



Scheme 55

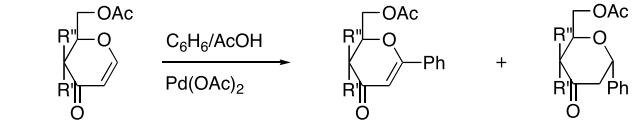


Scheme 56



R = Ts, $CCl_3CH_2OSO_2$, $CCl_3CH_2SO_2$, CF_3CO

Scheme 57



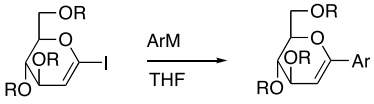
171: R' =OAc, R'' = H
172: R' = H, R'' = OAc

173 (49%)
174 (38%)

175 (21%)
176 (32%)

Scheme 58

Glycals have been used extensively as substrates in various approaches to *C*-glycosyl compounds. Among the studies devoted to the Heck-type catalytic coupling [261] between carbohydrate derivatives and aromatic compounds, there are several examples of the glycal double bond being preserved during such reaction. For example, in reaction with benzene, peracylated pyranosyl eneones **171** and **172** mainly form unsaturated 1-*C*-phenyl substituted products **173–176** (*Scheme 58*) [262].



177 **178**

R = TIPS

Scheme 59

The discovery (already mentioned) that benzylated or silylated 1-tri-n-butylstannyl derivatives of glucal undergo palladium-catalyzed cross-couplings with aryl or alkenyl halides [263,264] facilitated synthesis of naturally occurring *C*-glycosyl compounds [264]. Interestingly, silylated 1-iodo D-glucal **177**, obtained by a two-step procedure consisting of stannylation and tin-iodine exchange, can also be coupled with a variety of metalated aromatic derivatives to afford a high yield of 1-*C*-arylated compounds **178** (*Scheme 59*) (the best results were obtained using arylboronic acids and ArZnCl derivatives) [265]. 1-*C*-arylated compounds, in turn, are useful precursors of more complex *C*-glycosyl natural products. In an analogous way, iodoglucal can be obtained in unexpectedly high yield (75%) upon attempted coupling of C-1 stannylated derivative to 3-iodo-2-propyn-1-ol [263].

2-Cyanoglycals [266], which have found application in the synthesis of branched sugars [267], can be obtained in moderate yields by addition of chlorosulfonyl isocyanate to acylated glycals and treatment of intermediate *N*-chlorosulfonamides with triethylamine. Potentially useful *S*-heterocyclic analogs of glycals have been obtained predominantly by adaptation of the reductive elimination method described in the first part of this chapter [268,269].

References

1. Fischer E, Zach K (1913) Sitzber Kgl Preuss Akad Wiss 16:311
2. Helferich B (1952) Adv Carbohydr Chem Biochem 7
3. Nicolau KC, Sorensen EJ (1966) Classics in Total Synthesis. VCH, Weinheim
4. McNaught AD (1997) Adv Carbohydr Chem Biochem 52
5. McNaught AD (1997) J Carbohydr Chem 16:1191
6. Blair MG (1954) Adv Carbohydr Chem 9:97
7. Ferrier RJ (1965) Adv Carbohydr Chem 20:67
8. Ferrier RJ (1969) Adv Carbohydr Chem 24:199
9. Ferrier RJ (1967–1999) In: Ferrier RJ (ed) (1967–1999) A Specialist Periodical Report; Carbohydr Chem 1–33
10. Danishefsky SJ, Bilodeau MT (1996) Angew Chem Int Ed Engl 35:1380
11. Seeberger PH, Bilodeau MT, Danishefsky SJ (1997) Aldrichimica Acta 30:75
12. Di Busolo V, Kim YJ, Gin DY (1998) J Am Chem Soc 120:13515
13. Thiem J, Karl H, Schwentner J (1978) Synthesis: 696
14. Gervay J, Danishefsky SJ (1991) J Org Chem 56:5448
15. Thiem J, Klaffke W (1990) Top Curr Chem 154:285
16. Grewal G, Kaila N, Franck RW (1992) J Org Chem 57:2084
17. Seeberger PH, Danishefsky SJ (1998) Acc Chem Res 31:685
18. Daves JGD (1990) Acc Chem Res 23:201
19. Postema MHD (1995) C-Glycoside Synthesis. CRC Press, Boca Raton, FL
20. Levy DE, Tang C (1995) The Chemistry of C-glycosides. Pergamon Press, Oxford

21. Bols M (1996) Carbohydrate building blocks. Wiley, New York
22. Hanessian S (1993) Total Synthesis of Natural Products: The "Chiron" Approach. Pergamon Press, Oxford
23. Klingler FD, Psiorz M (1992) Chimica Oggi 3:47
24. Yuan X, Linhardt RJ (2005) Curr Top Med Chem 5:1393
25. Hang HC, Bertozzi CR (2001) Accounts Chem Res 34:727
26. Meutermans W, Le GT, Becker B (2006) Chem Med Chem 1:1164
27. Sofia MJ (1998) Mol Divers 3:75
28. Chakraborty TK, Gosh S, Jayaprakash S (2002) Curr Med Chem 102:49
29. Seeberger PH (2003) Chem Comm: 1115
30. Palmacci ER, Plante OJ, Hewitt MC, Seeberger PH (2003) Helvetica Chimica Acta 86:3975
31. Ferrier RJ (2001) Topics Curr Chem 215:153
32. Ferrier RJ, Hoberg JO (2003) Adv Carbohydr Chem Biochem 58:55
33. Somsak L (2001) Chem Rev 101:81
34. Ferrier RJ, Zubkov OA (2003) Org React 62:569
35. Wulff G, Rohle G (1973) Angew Chem Int Ed Engl 13:157
36. Paulsen H (1982) Angew Chem Int Ed Engl 21:155
37. Jeanloz RW, Stoffyn PJ (1962) In: Whistler RL (ed) Methods in Carbohydrate Chemistry, vol I. Academic Press, New York, p 221
38. Wolfrom ML, Thompson A (1963) In: Whistler RL (ed) Methods in Carbohydrate Chemistry, vol II. Academic Press, New York, p 211
39. Redemann CE, Niemann C (1942) Org Synth 22:1
40. Kartha KPR, Jennings HJ (1990) J Carbohydr Chem 9:777
41. Helferich B, Mulcahy EN, Zeigler H (1954) Chem Ber 87:233
42. Collins P, Ferrier RJ (1995) Monosaccharides: Their Chemistry and Their Roles in Natural Products. Wiley, Chichester, p 316
43. Humoller FL (1962) Methods in Carbohydrate Chemistry, vol 1. Academic Press, New York, p 83
44. Roth W, Pigman W (1963) In: Whistler RL (ed) Methods in Carbohydrate Chemistry, vol II. Academic Press, New York, p 405
45. Blair MG (1963) Methods in Carbohydrate Chemistry, vol II. Academic Press, New York, p 411
46. Shafizadeh F (1963) Methods in Carbohydrate Chemistry, vol II. Academic Press, New York, p 409
47. Weygand F (1962) Methods in Carbohydrate Chemistry, vol 1. Academic Press, New York, p 182
48. Erdik E (1987) Tetrahedron 43:2203
49. Hurd CD, H. J (1966) Carbohydr Res 2:240
50. Deriaz RE, Overend WG, Stacey M, Teece EG, Wiggins LF (1949) J Chem Soc: 1879
51. Iselin B, Reichstein T (1944) Helv Chim Acta: 1146
52. Maran FEV, Catelani G, D'Angeli F (1989) Electrochimica Acta 34:587
53. Wisniewski A, Skorupowa ERW, Glod D (1991) Pol J Chem 65:875
54. Pollon JHP, Llewellyn G, Williams JM (1989) Synthesis: 758
55. Jain S, Suryawaneshi SN, Bhakuni DS (1987) Indian J Chem 26B:866
56. Ness RK, Fletcher JHG (1963) J Org Chem 28:435
57. Haga M, Ness RK (1965) J Org Chem 30:158
58. Eitelman SJ, Jordaan A (1977) J Chem Soc Chem Commun: 552
59. Eitelman SJ, Hall RH, Jordaan A (1978) J Chem Soc Perkin Trans I: 595
60. Ireland RE, Thaisrivongs S, Vanier N, Wilcox CS (1980) J Org Chem 45:48
61. Armstrong PL, Coull IC, Hewson AT, Slater MJ (1995) Tetrahedron Lett 36:4311
62. Ireland RE, Wilcox CS (1977) Tetrahedron Lett: 2389
63. Ireland RE, Wilcox CS, Thaisrivongs S (1978) J Org Chem 43:786
64. Cheng JCY, Hacksell U, Daves Jr GD (1985) J Org Chem 50:2778
65. Holzapfel CW, Koekemoer JM, Verdoorn GH (1986) S Afr J Chem 39:151
66. Csuk R, Furstner A, Glanzer BI, Weidmann H (1986) J Chem Soc Chem Commun: 1149
67. Csuk R, Glanzer BI, Furstner A, Weidmann H, Formacek V (1986) Carbohydr Res 157:235
68. Furstner A, Weidmann H (1988) J Carbohydr Chem 7:773
69. Boullanger P, Martin JC, Descotes G (1975) J Heterocycl Chem 12:91
70. Chmielewski M, Fokt I, Grodner J, Grynkiewicz G, Szeja W (1989) J Carbohydr Chem 8:735

71. de Pouilly P, Vauzeilles B, Mallet JM, Sinay P (1991) CR Acad Sci Paris 313:1391
72. Cavallaro CL, Schwartz J (1995) J Org Chem 60:7055
73. Spencer RP, Schwartz J (1996) Tetrahedron Lett 37:4357
74. Spencer RP, Schwartz J (1997) J Org Chem 62:4204
75. Kovacs G, Gyarmati J, Somsak L, Micskei K (1996) Tetrahedron Lett 37:1293
76. Somsak L, Nemeth I (1993) J Carbohydr Chem 12:679
77. Somsak L, Bajza I, Batta G (1990) Liebids Ann Chem: 1265
78. Mahmoud SH, Somsak L, Farkas I (1994) Carbohydr Res 254:91
79. Kiss L, Somsak L (1996) Carbohydr Res 291:43
80. Somsak L, Madaj J, Wisniewski A (1997) J Carbohydr Chem 16:1075
81. Chemla F (2002) J Chem Soc Perkin Trans 1:275
82. Boutureira O, Rodriguez MA, Matheu MI, Diaz Y, Castillion S (2006) Org Lett 8:673
83. Boulineau FP, Wei A (2002) Org Lett 4:2281
84. Santoyo-Gonzales F, Calvo-Flores FG, Hernandez-Mateo F, Garcia-Mendoza P, Isac-Garcia J, Perez-Alvarez D (1994) Synlett: 454
85. Lancelin JM, Morin-Allory L, Sinay P (1984) J Chem Soc Chem Commun: 355
86. Fernandez-Mayorales A, Marra A, Trumtel M, Veyriers A, Sinay P (1989) Tetrahedron Lett 30:2537
87. Fernandez-Mayorales A, Marra A, Trumtel M, Veyriers A, Sinay P (1989) Carbohydr Res 188:81
88. Casillas M, Gomez AM, Lopez CJ, Valverde S (1996) Synlett: 628
89. Shull BK, Wu Z, Koreeda M (1996) J Carbohydr Chem 15:955
90. Forbes CL, Franck RW (1999) J Org Chem 64:1424
91. Scheffold R, Rytz G, Walder R (1983) In: Scheffold R (ed) Modern Synthetic Methods, vol 3. Springer, Berlin Heidelberg New York, p 355
92. Bergmann M (1923) Justus Liebigs Ann Chem 434:79
93. Bergmann M, Freudenberg K (1929) Ber Dtsch Chem Ges 62
94. Haworth WN (1930) J Chem Soc: 2644
95. Dauben WL, Evans X (1938) J Am Chem Soc 60:886
96. Rolland N, Vass G, Cleophax J, Sepulchre AM, Gero SD, Cier A (1982) Helv Chim Acta 65:1627
97. Wang LX, Sakairi N, Kuzuhara H (1991) Carbohydr Res 219:133
98. Spohr U, Bach M, Spiro RG (1993) Can J Chem 71:1919
99. Ziegler T (1995) Liebigs Ann Org Bioorg Chem 6:949
100. Kanemitsu T, Ogihara Y, Takeda T (1997) Chem Pharm Bull 45:643
101. Ferrier RJ (1980) In: Pigman W (ed) The Carbohydrates, Chemistry and Biochemistry, vol IB. Academic Press, New York, p 843
102. Maurer K, Mahn H (1927) Ber Dtsch Chem Ges 60:1316
103. Bredereck H, Wagner A, Faber G, Huber W, Immel G (1958) Chem Ber 91:2891
104. Zemplen G, Bruckner X (1928) Ber Dtsch Chem Ges 61:2481
105. Maurer K (1930) Ber Dtsch Chem Ges 63:25
106. Lemieux RU, Lineback DR (1965) Can J Chem 43:94
107. Backinovski LV, Tsvetkov YE, Bairamova NE, Balan NF, Kochetkov NK (1980) Izv Akad Nauk SSSR, Ser Khim 8:1905
108. Tsvetkov YE, Byramova NE, Backinovski LV (1983) Carbohydr Res 115:254
109. Honda T, Inoue M, Kato M, Shima K, Shimamoto T (1987) Chem Pharm Bull 35:3975
110. Shah RH, Bahl OP (1979) Carbohydr Res 74:105
111. Dess D, Kleine HP, Weinberg DV, Kaufman RJ, Sidhu RS (1981) Synthesis: 883
112. Loganathan D, Trivedi GK (1987) Carbohydr Res 162:117
113. Kleine HP, Weinberg DV, Kaufman RJ, Sidhu RS (1985) Carbohydr Res 142:333
114. Kovac P, Taylor RB (1987) Carbohydr Res 167:153
115. Hadd MJ, Gervay J (1999) Carbohydr Res 320:61
116. Lichtenthaler FW, Kaji E, Weprek S (1985) J Org Chem 50:3505
117. Kaji E, Osa Y, Takahashi K, Hirooka M, Zen S, Lichtenthaler FW (1994) Bull Chem Soc Jpn 67:1130
118. Konda Y, Iwasaki Y, Takahata S, Arima S, Toida T, Kaji E, Takeda K, Harigaya Y (1997) Chem Pharm Bull 45:626
119. Chretien F (1989) Synth Commun 19:1015
120. Marais C, Steenkamp JA, Ferreira D (1996) J Chem Soc Perkin Trans 1:2915

121. Jain S, Suryawanshi SN, Misra S, Bhakuni DS (1988) Indian J Chem, Sect B 27:866
122. Oedinger H, Moeller F (1967) Angew Chem Int Ed Engl 6:76
123. Rao DR, Lerner LM (1972) Carbohydr Res 22:345
124. Lain V, Coste-Sarguet A, Gadelle A, Defay J, Perly B, Djedaini-Pilard F (1995) J Chem Soc Perkin Trans 2:1479
125. Varela O, de Fina MG, de Lederkremer RM (1987) Carbohydr Res 167:187
126. Lichtenthaler FW, Klaeres U, Lergenmueller M, Schwidetzky S (1992) Synthesis: 179
127. Schauer R (1982) Adv Carbohydr Chem Biochem 40:131
128. Reutter PW, Kottgen E, Bauer C, Gerok W (1982) Sialic Acids: Chemistry, Metabolism and Function. In: Cell Biology Monographs, vol 10. Springer, Berlin Heidelberg New York, p 263
129. von Itzstein M, Thomson RJ (1986) Glycoscience: Synthesis of Glycoconjugates and Oligosaccharides, vol 186. Springer, Berlin Heidelberg New York
130. Zbiral E, Schreiner E, Christian R, Kleineidam RG, Schauer R (1989) Liebids Ann Chem: 159
131. Kumar V, Tanenbaum SW, Flashner M (1982) Carbohydr Res 101:155
132. Kumar V, Kessler J, Scott ME, Patwardhan BH, Tanenbaum SW, Flashner M (1981) Carbohydr Res 94:123
133. Kok GB, Mackey BL, von Itzstein M (1996) Carbohydr Res 289:67
134. Lichtenthaler FW (1992) Modern Synthetic Methods 1992, vol 6. Verlag Helvetica Chimica Acta, Verlag Chemie, Weinheim, p 273
135. Lichtenthaler FW, Nishiyama S, Weimer T (1989) Liebigs Ann Chem: 1163
136. Iglesias-Guerra F, Candela JI, Espareto JL, Vega-Perez JM (1994) Tetrahedron Lett 35:5031
137. Pravdic N, Fletcher JHG (1967) J Org Chem 32:1806
138. Pravdic N, Franjic-Mihalic I, Danilov B (1975) Carbohydr Res 45:302
139. Vos JN, van Boom JH, van Boeckel CAA, Beetz T (1984) J Carbohydr Chem 3:117
140. Jones GS, Scott WJ (1992) J Am Chem Soc 114:1491
141. Beau JM, Gallagher T (1997) Top Curr Chem, Glycoscience, vol 187. Springer, Berlin Heidelberg New York, 2
142. Clode DM, Horton D, Weckerle W (1976) Carbohydr Res 49:305
143. Feast AA, Overend WG, Williams NR (1965) J Chem Soc: 7378
144. Lemieux RU, Frage E, Watanabe KA (1968) Can J Chem 46:61
145. Guthrie RD, Irvine RW (1979) Carbohydr Res 72:285
146. Thiem J, Ossowski P, Schwentner J (1979) Angew Chem Int Ed Engl 18:222
147. Thiem J, Ossowski P, Schwentner J (1980) Chem Ber 113:955
148. Jung G, Klemer A (1981) Chem Ber 114:740
149. Dettinger HM, Kurz G, Lehmann J (1979) Carbohydr Res 74:301
150. Lesimple P, Beau JM, Jaurand G, Sinay P (1986) Tetrahedron Lett 27:6201
151. Qiu D, Schmidt RR (1990) Synthesis: 875
152. Qiu D, Schmidt RR (1995) Carbohydr Lett 1:291
153. Dubois E, Beau JM (1990) Tetrahedron Lett 31:5165
154. Stille JK (1986) Angew Chem Int Ed Engl 25:508
155. Nicolau KC, Hwang CK, Duggan ME (1986) J Chem Soc Chem Commun: 925
156. Hanessian S, Martin M, Desai RC (1986) J Chem Soc Chem Commun: 926
157. Friesen RW, Sturino CF, Daljeet AK, Kolaczewska AE (1991) J Org Chem 56:1944
158. Tius MA, Gu Y, Gomez-Galeno J (1990) J Am Chem Soc 112:8188
159. Parker KA, Su DS (1996) J Org Chem 61:2191
160. Liu J, Huang CH, Wong CH (2002) Tetrahedron Lett 43:3447
161. Chambers DJ, Evans GR, Fairbanks AJ (2003) Tetrahedron Lett 44:5221
162. Chambers DJ, Evans GR, Fairbanks AJ, Fairbanks AJ (2004) Tetrahedron 60:8411
163. Horton D, Priebe W, Sznaidman M (1989) Carbohydr Res 187:145
164. Pelyvas IF, Monneret C, Herczegh P (1988) Synthetic Aspects of Aminodeoxy Sugars of Antibiotics. Springer, Berlin Heidelberg New York
165. Charette AB, Cote B (1993) J Org Chem 58:933
166. Borowiecka J, Lipka P, Michalska M (1988) Tetrahedron 44:2067
167. Kassou M, Castillon S (1994) Tetrahedron Lett 30:5513
168. Priebe W, Grynkiewicz G, Krawczyk M, Fokt I (1994) Abstr Papers. Am Chem Soc 207:78-CARB
169. Priebe W, Grynkiewicz G, Krawczyk M, Fokt I (1994) Proc XVII IUPAC Int Carbohydr Symp, Ottawa, Canada, B2.55

170. Priebe W, Zamojski A (1980) Tetrahedron 36:287
171. Pearson J (1963) J Am Chem Soc 85:3533
172. Ho TL (1975) Chem Rev 75:1
173. Leutzinger EE, Meguro T, Towwnsend LB, Shuman DA, Schweizer MP, Stewart CM, Robins RK (1972) J Org Chem 37:3695
174. De las Heras FG, Stud M (1977) Tetrahedron 33:1513
175. Heyns K, Park JI (1976) Chem Ber 109:3262
176. Dunkerton LV, Adair NK, Euske JM, Brady KT, Robinson PD (1988) J Org Chem 53:845
177. Priebe W, Grynkiewicz G, Neamati N (1991) Tetrahedron Lett 32:3313
178. Heyns K, Lim MT, Park JI (1976) Tetrahedron Lett: 1477
179. Guthrie RD, Irvine RW (1980) Carbohydr Res 82:207
180. Guthrie RD, Irvine RW (1980) Carbohydr Res 82:225
181. Boivin J, Pais M, Monneret C (1980) Carbohydr Res 79:193
182. Thiem J, Springer D (1985) Carbohydr Res 136:325
183. Fraser-Reid B, Radatus B (1970) J Am Chem Soc 92:6661
184. Achmatowicz O, Szechner B (1972) Tetrahedron Lett: 1205
185. Martin A, Pais M, Monneret C (1983) Carbohydr Res 113:21
186. Fraser-Reid B, Radatus B (1975) Acc Chem Res 8:192
187. Achmatowicz O, Szechner B (1997) Tetrahedron Lett 38:4701
188. Habus I, Sunjic V (1985) Croatica Chem Acta 58:321
189. Inaba K, Matsumura S, Yoshikawa S (1991) Chem Lett: 485
190. Doboszewski B, Blaton N, Herdewijn P (1995) J Org Chem 60:7909
191. Valverde S, Garcia-Ochoa S, Martin-Lomas M (1987) J Chem Soc Chem Commun: 383
192. Gomez AM, Valverde S, Fraser-Reid B (1991) J Chem Soc Chem Commun: 1207
193. Wittman MD, Halcomb RL, Danishefsky SJ, Golik J, Vyas D (1990) J Org Chem 55:1979
194. Griffith DA, Danishefsky SJ (1996) J Am Chem Soc 118:9526
195. Halcomb RL, Boyer SH, Wittman MD, Olson SH, Denhart DJ, Liu KKC, Danishefsky SJ (1995) J Am Chem Soc 117:5720
196. Dupradeau FY, Allaire S, Prandi J, Beau JM (1993) Tetrahedron Lett 34:4513
197. Dupradeau FY, Prandi J, Beau JM (1995) Tetrahedron 51:3205
198. Danishefsky SJ, Shair MD (1996) J Org Chem 61:16
199. Dawe RD, Fraser-Reid B (1981) J Chem Soc Chem Commun: 1180
200. Grynkiewicz G, BeMiller JN (1982) J Carbohydr Chem 1:121
201. Kawauchi N, Hashimoto H (1987) Bull Chem Soc Jpn 60:1441
202. Marco-Contelles JL, Fernandez C, Gomez A, Martin-Leon N (1990) Tetrahedron Lett 31: 1467
203. Priebe W, Grynkiewicz G, Neamati N (1991) Monatsh Chem 121:419
204. Wipf P (1997) In: Trost BM (ed) Comprehensive Organic Synthesis, vol 5. Pergamon Press, Oxford, p 827
205. Heyns K, Hohlweg R (1978) Chem Ber 111:1632
206. Cottier L, Remy G, Descotes G (1979) Synthesis: 711
207. de Raadt A, Ferrier RJ (1991) Carbohydr Res 216:93
208. Balasubramanian KK, Rhamesh NG, Pramanik A, Chandrasekar J (1994) J Chem Soc Perkin Trans 2:1399
209. Takai I, Yamamoto A, Ishido Y, Sakakibara T, Yagi E (1991) Carbohydr Res 220:195
210. Seta A, Nagano C, Ito S, Tokuda K, Tamura T, Kamitani T, Sakakibara T (1998) Tetrahedron Lett 39:591
211. Sakakibara T, Nishitani A, Seta A, Nakagawa T (1991) Tetrahedron Lett 32:5809
212. Sakakibara T, Nomura Y, Sudoh R (1983) Carbohydr Res 124:53
213. Sakakibara T, Ito S, Ikegawa H, Matsuo I, Seta A (1993) Tetrahedron Lett 34:3429
214. Hall RH, Jordaan A, Lourens GJ (1973) J Chem Soc Perkin Trans 1:38
215. Hall RH, Jordaan A, Villiers OG (1975) J Chem Soc Perkin Trans 1:626
216. Nguefack JF, Bolitt V, Sinou D (1996) Tetrahedron Lett 37:59
217. Nguefack JF, Bolitt V, Sinou D (1997) J Org Chem 62:6827
218. Holzapfel CW, Engelbrecht GJ, Marais L, Toerien F (1997) Tetrahedron 53:3975
219. Zamojski A, Banaszek A, Grynkiewicz G (1982) Adv Carbohydr Chem Biochem 40:1
220. Zamojski A, Grynkiewicz G (1984) ApSimon J (ed) Total Synthesis of Natural Products. Wiley, New York 6:141

221. Danishefsky S (1980) Acc Chem Res 14:400
222. Carruthers W (1990) Cycloaddition Reactions in Organic Synhesis. Pergamon Press, Oxford
223. Larson E, Danishefsky S (1982) J Am Chem Soc 104:6458
224. Danishefsky S, Larson E, Askin D, Kato N (1985) J Am Chem Soc 107:1246
225. Danishefsky SJ, Webb RR (1984) J Org Chem 49:1955
226. Danishefsky SJ, Bednarski M (1985) Tetrahedron Lett 26:3411
227. Danishefsky SJ, Myles DC, Harvey DF (1987) J Am Chem Soc 109:862
228. Schmidt RR, Angerbauer R (1981) Carbohydr Res 89:159
229. Schmidt RR, Maier M (1985) Tetrahedron Lett 26:2065
230. Schmidt RR (1986) Acc Chem Res 19:250
231. Danishefsky SJ, Kobayashi S, Kerwin JF (1982) J Org Chem 104:358
232. Danishefsky SJ, De Ninno MP, Chen S (1988) J Am Chem Soc 110:3929
233. McDonald FE, Schultz CC, Chatterjee AK (1995) Organomtallics 14:3628
234. McDonald FE, Gleason MM (1995) Angew Chem Int Ed Engl 34:350
235. McDonald FE, Gleason MM (1996) J Am Chem Soc 118:6648
236. McDonald FE, Bowman JL (1996) Tetrahedron Lett 37:4675
237. McDonald, Zhu HYH (1997) Tetrahedron 53:11061
238. Postema MHD, Piper JL, Liu L, Shen J, Faust M, Andreana P (2003) J Org Chem 68:4748
239. Postema MHD, Piper JL, Betts RL, Valeriote FA, Pietraszkiewicz (2005) J Org Chem 70:829
240. Taillefumier C, Chapleur Y (2004) Chem Rev 104:263
241. Ali MH, Collins PM, Overend WG (1990) Carbohydr Res 205:428
242. Gascon-Lopez M, Motevalli M, Paloumbis G, Bladon P, Wyatt PB (2003) Tetrahedron 59:9349
243. Yang WB, Wu CY, Chang CC, Wang SH, Teo CF, Lin CH (2001) Tetrahedron Lett 42:6907
244. Yang WB, Yang YY, Gu YF, Wang SH, Chang CC, Lin CH (2002) J Org Chem 67:3773
245. Toth M, Somsak L (2001) J Chem Soc Perkin Trans 1:942
246. Toth M, Kover KE, Benyei A, Somsak L (2003) Org Biomol Chem 1:4039
247. Taylor RJK, Mc Allister GD, Franck RW (2006) Carbohydr Res 341:1298
248. Lopez JC, Lameignere E, Lukacs G (1988) J Chem Soc Chem Commun: 514
249. Burnouf C, Lopez JC, Laborde MA, Olesker A, Lukacs G (1988) Tetrahedron Lett 28:5533
250. Jedrzejewski S, Ermann P (1993) Tetrahedron Lett 34:615
251. Ramesh NG, Balasubramanian KK (1991) Tetrahedron Lett 32:3875
252. Booma C, Balasubramanian KK (1991) Tetrahedron Lett 32:3875
253. Lopez CJ, Lukacs G (eds) (1992) Cycloaddition Reactions in Carbohydrate Chemistry, vol 494. American Chemical Society, Washington, DC
254. Jung ME, Choe SWT (1993) Tetrahedron Lett 34:6247
255. Priebe W, Grynkiewicz G, Neamati N (1992) Tetrahedron Lett 33:7681
256. Chmielewski M, Kaluza Z, Grodner J, Urbanski R (1992) In: Guiliano RM (ed) Cycloaddition Reactions in Carbohydrate Chemistry, vol 494. American Chemical Society, Washington, DC
257. Chmielewski M, Kaluza Z, Belzecki C, Salanski P, Jurczak J, Adamowicz H (1985) Tetrahedron 41:2441
258. Chmielewski M, Kaluza Z, Mostowicz D, Belzecki C, Baranowska E, Jacobsen JP, Salanski P, Jurczak J (1987) Tetrahedron 43:4563
259. Barret AGM, Fenwick A (1983) J Chem Soc Chem Commun: 299
260. Chan JH, Hall SS (1984) J Org Chem 49:195
261. Heck RF (1991) In: Semmelhack MF (ed) Comprehensive Organic Synthesis, vol 4: Additions to and substitutions at C-C π-bonds. Pergamon Press, Oxford, p 833
262. Benhaddou R, Czernecki S, Ville G (1992) J Org Chem 57:4612
263. Dubois E, Beau JM (1990) J Chem Soc Chem Commun: 1191
264. Friesen RW, Sturino CF, Daljeet AK, Kolaczewska AE (1990) J Org Chem 55:5808
265. Friesen RW, Loo RW (1991) J Org Chem 56:4821
266. Hall RH, Jordaan A (1973) J Chem Soc Perkin Trans 1:1059
267. Bischofberger K, Hall RH, Jordaan A, Woolard GR (1980) S Afr J Chem 33:92
268. Korytnyk W, Angelino N, Dodson-Simmons O, Hanchak M, Madson M, Valentkovic-Horvath S (1983) Carbohydr Res 113:166
269. Bozo E, Boros S, Kuszmann J (1997) Carbohydr Res 299:59

3.7 Anomeric Anhydro Sugars

*Nathan W. McGill, Spencer J. Williams**
School of Chemistry, The University of Melbourne,
Parkville, VIC 3052, Australia
n.mcgill@pgrad.unimelb.edu.au, sjwill@unimelb.edu.au

Abstract

Anomeric anhydro sugars are sugar derivatives where the anomeric carbon participates in an acetal linkage with two of the hydroxyl groups of the sugar. They are essentially intramolecular glycosides, and their bicyclic nature provides a powerful conformational constraint that greatly influences their reactivity. This chapter reviews the occurrence, properties, formation, and reactions of anomeric anhydro sugars. Particular emphasis is placed on 1,2- and 1,6-anhydropyranoses, including conformational aspects and ring-opening reactions. Epoxide-containing 1,6-anhydro sugars (Černý epoxides) are briefly reviewed, and the formation and some reactions of the 1,6-anhydro sugar enone, levoglucosenone, is covered. An overview is given of the use of 1,2-anhydro sugar as glycosyl donors. Also discussed are the formation and reactions of anomeric anhydro sugars containing nitrogen, sulfur, or selenium.

Keywords

Anhydro sugars; Dianhydro sugars; Epoxide; Glucosan; Glycosan; Levoglucosan; Levoglucosenone; Seleno sugar; Selenolevoglucosan; Thiolevoglucosan

Abbreviations

CSA camphorsulfonic acid
DMDO 3,3-dimethyldioxirane

In: *Glycoscience*. Fraser-Reid B, Tatsuta K, Thiem J (eds)
Chapter-DOI 10-1007/978-3-540-30429-6_16:

mw microwave
NIS *N*-iodosuccinimide
TESOTf triethylsilyl triflate

1 Introduction

Anomeric anhydro sugars are sugar derivatives in which the anomeric carbon participates in an acetal linkage with two of the hydroxyl groups of the sugar. They are essentially intramolecular glycosides, which may be formally derived from the parent sugar by the loss of a molecule of water. Importantly, the intramolecular acetalation of the anomeric center acts to protect this position and two other hydroxyl groups. The bicyclic structure of anomeric anhydro sugars ensures that they are restricted to a limited conformational coordinate that strongly influences the reactivity of the remaining hydroxy groups. Many anhydro sugars are highly crystalline and this feature makes them easy compounds to work with and to purify. This chapter focuses on anomeric anhydro sugars of synthetic utility and on the more common reactions known for this class of sugars.

1.1 General Remarks

Solutions of aldohexoses in aqueous acid at equilibrium contain varying amounts of anhydro sugars, particularly the 1,6-forms. The composition of the equilibrium mixture has been determined for a number of aldohexoses in the presence of acidic ion-exchange resin or in 0.25-M H_2SO_4 at 100 °C [1]. For D-glucose and D-galactose, the proportion of anhydro sugars present at equilibrium is small; for D-altrose, D-gulose, and D-idose the proportion of the 1,6-anhydropyranose present is greater than 50%. The formation of 1,6-anhydrofuranoses under these conditions is much less favorable. Of the eight aldohexoses, D-talose forms an anhydrofuranose to the greatest degree and, in this case, only to the extent of 2.5%. As a rule, as the number of hydroxy groups oriented in the axial direction in the 'usual' $^4C_1$ conformation of an aldose increases, so does its tendency to form 1,6-anhydro derivatives. The acetal of anomeric anhydro sugars may be cleaved by solvolysis under acidic conditions. In the absence of a nucleophilic solvent, Lewis acid-catalyzed polycondensation of anomeric anhydro sugar monomers affords glycans. This approach provides an efficient method for the preparation of certain stereoregular polysaccharides [2].

2 1,6-Anhydrohexopyranoses

2.1 Occurrence and Formation

Of all the anhydro sugars, 1,6-anhydrohexopyranoses are the most studied and are of the greatest synthetic utility. 1,6-Anhydrohexopyranoses have been the subject of several comprehensive reviews, which the interested reader should consult for more details [3,4,5,6]. 1,6-Anhydro-β-D-glucopyranose (**1**; ❯ *Fig. 1*) (also known as 'glucosan' or 'levoglucosan' on account

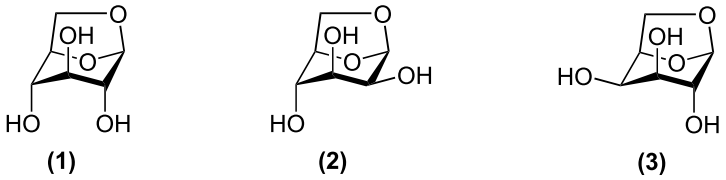
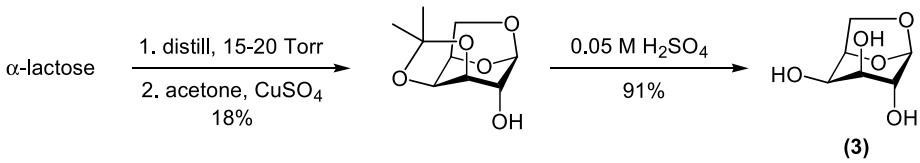


Figure 1
Common 1,6-anhydro sugars: D-glucosan (1), D-mannosan (2) and D-galactosan (3)

of its significantly negative optical rotation) was first isolated in 1894 by Tanret upon treatment of the naturally occurring phenolic glycosides, picein, salicin, and coniferin, with barium hydroxide [7]. On the basis of the molecular formula of $C_6H_{10}O_5$, Tanret correctly described this crystalline material as 'glucose anhydride,' which is often shortened to 'glucosan.' The structure of this anhydride remained unknown until 1920, when Pictet proposed the existence of a 1,6-anhydro bridge. As a result of the 1,6-anhydro bridge, levoglucosan adopts a $^1C_4$ conformation in which the hydroxyl groups are axially disposed, and opposite to the normal $^4C_1$ conformation observed for D-glucopyranose. The crystal structure [8] and ^{1}H-NMR spectral data [9] are in agreement with a $^1C_4$ conformation, which is common to all eight 1,6-anhydropyranoses. Despite the obvious structural differences, many of the chemical properties of levoglucosan are similar to methyl β-D-glucopyranoside: it is stable to alkali, but is hydrolyzed by acid.

In 1918, Pictet and Sarasin isolated 1,6-anhydro-β-D-glycopyranose from the distillate obtained from the pyrolysis of cellulose [10]. Indeed, the pyrolysis of oligo- and polysaccharide-rich biomass remains an effective method for the preparation of simple 1,6-anhydro sugars. Levoglucosan has been produced on a large scale by the uncatalyzed pyrolysis of powdered corn starch [11] or cellulose [12]. The pyrolysis of cellulosic-biomass during bushfires or residential wood combustion produces large amounts of smoke aerosol, of which levoglucosan is a significant component. Consequently, levoglucosan has been used as a molecular marker for quantifying the contribution of woodsmoke to atmospheric pollution; levoglucosan can be quantified by GC-MS or LC-ESI-MS analysis [13]. Mannosan (**2**) and galactosan (**3**) (*Fig. 1*) may be produced by pyrolysis of the seeds of the ivory nut palm, *Phytelephas macrocarpa* [14], and of α-lactose monohydrate [15], respectively (*Scheme 1*). In both cases, purification of the anhydro sugars from other pyrolyzates is facilitated by conversion of each into the corresponding isopropylidene acetal. Extraction of this product into an organic solvent and subsequent removal of the acetonide group yields the parent glycosans.



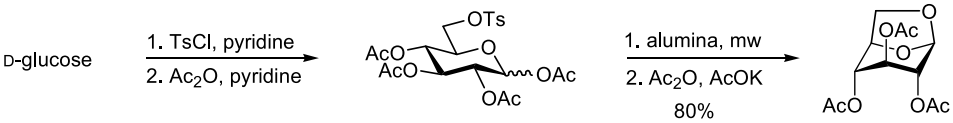
Scheme 1

Pyrolysis of a variety of monosaccharides has been shown to afford the corresponding 1,6-anhydro sugars; however, the yields are lower than that seen when using the glycan precursors. It has been proposed that the pyrolysis of simple sugars proceeds by way of an initial condensation to intermediate polysaccharides, which then thermally depolymerize, affording the 1,6-anhydro sugars [3].

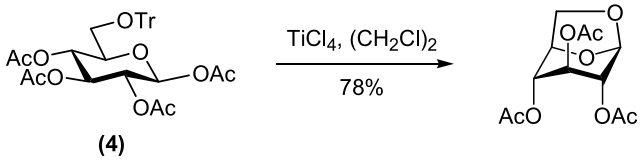
In addition to the pyrolytic procedures described above, there are three main synthetic routes to 1,6-anhydro sugars. The first method relies on selective activation of C-1 or C-6 with a suitable leaving group, such as a halide or sulfonate, and displacement of this group by an oxyanion, generated with base, at the alternate position. Large scale, one-pot syntheses of tri-*O*-acetyl-1,6-anhydro-β-D-glucopyranose and -β-D-mannopyranose, involving the selective tosylation of D-glucose and D-mannose at *O*-6, and subsequent treatment of the intermediate sulfonates with sodium hydroxide, have been described by Fraser-Reid [16]. More recently, Cleophax reported a high yielding, solvent-free synthesis of tri-*O*-acetyl-1,6-anhydro-β-D-glycopyranose. In an analogous approach to that described by Fraser-Reid, selective tosylation of D-galactose, D-glucose, and D-mannose at *O*-6, followed by microwave (mw) irradiation of the sulfonate intermediates in the presence of basic alumina, afforded the desired 1,6-anhydrohexopyranoses (*Scheme 2*) [17].

In analogy to Tanret's original observation, treatment of a number of 1,2-*trans*-glycosides with base affords 1,6-anhydropyranoses, by way of the 1,2-anhydropyranose intermediate. In this way, a large-scale synthesis of tri-*O*-acetyl-1,6-anhydro-β-D-glucopyranose, utilizing the basic hydrolysis of phenyl tetra-*O*-acetyl-β-D-glucopyranoside as the key step, has been described [18]. More recently, Boons and co-workers have shown that treatment of pentabromophenyl β-D-glycosides with Amberlite IRA-400 (OH^-) resin affords 1,6-anhydro derivatives in high yields [19].

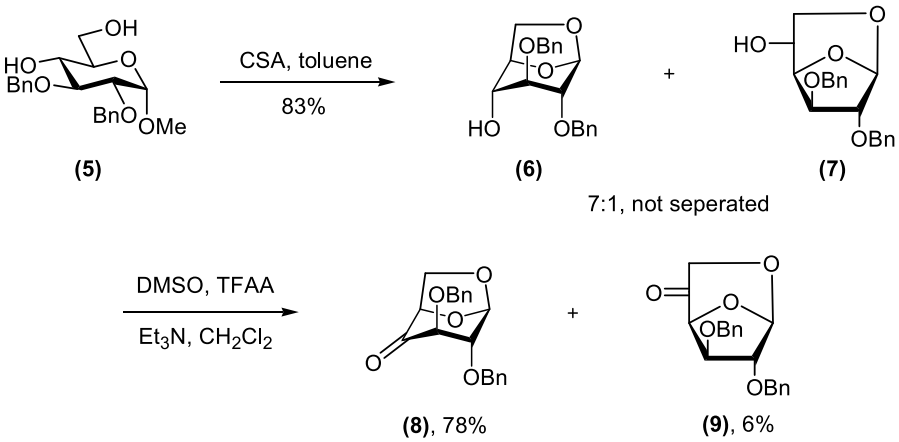
In the second approach to 1,6-anhydro sugars, anomeric esters, particularly glycosyl acetates, are activated by Lewis acids and displaced by 'labile' ethers at *O*-6. Rao has reported an efficient synthesis of tri-*O*-acetyl-1,6-anhydro-β-D-glucopyranose by treatment of the trityl ether (**4**) with titanium chloride (*Scheme 3*) [20].



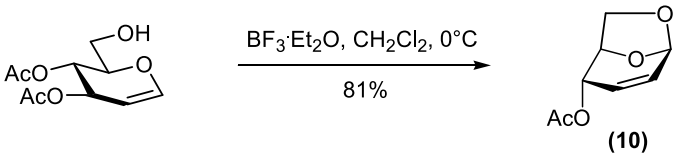
Scheme 2



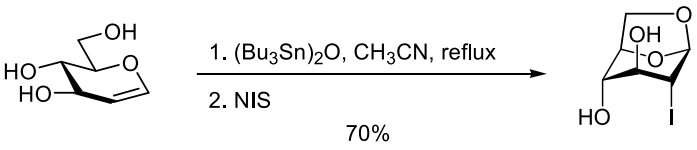
Scheme 3



Scheme 4



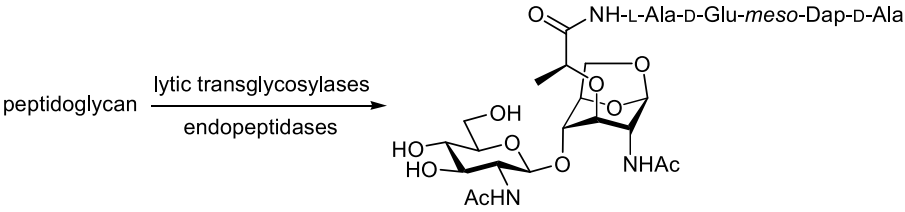
Scheme 5



Scheme 6

In a similar approach, treatment of methyl glycosides with camphorsulfonic acid (CSA) in toluene under Dean–Stark conditions provides a good route to functionalized 1,6-anhydro sugars. Heathcock and co-workers showed that oxidation of the difficult-to-separate mixture of the 1,6-anhydro-β-D-glucopyranose (**6**) and 1,6-anhydro-β-D-glucofuranose (**7**) allowed for easy isolation of the desired ketone (**8**), by either chromatography or direct crystallization, depending on the reaction scale (*Scheme 4*) [21].

In the final method, glycals are transformed into various 1,6-anhydro sugar derivatives. Glycals may be efficiently cyclized to yield 1,6-anhydro-2,3-dehydro sugars, such as (**10**), by the action of Lewis or Brønsted acids; this reaction is an intramolecular Ferrier glycosylation (*Scheme 5*) [22,23]. 1,6-Anhydro-2-deoxy-2-halo sugars are accessed through 'halocyclization' reactions of *O*-stannylated glycals with a source of halogen such as iodine, bromine, or NIS (*Scheme 6*) [22,24,25,26].



Scheme 7

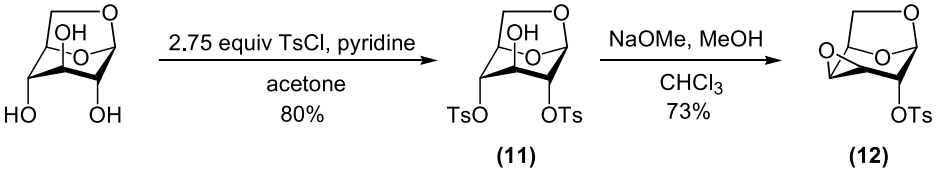
While the aforementioned methodologies constitute the most useful and general synthetic procedures for the preparation of 1,6-anhydropyranoses, a number of alternative routes have been described [27]. As mentioned previously, solutions of a number of sugars in aqueous acid at equilibrium contain considerable amounts of 1,6-anhydro sugars. A simple synthesis of tri-*O*-acetyl-1,6-anhydro-β-L-idopyranose has been reported by Stoffyn and Jeanloz that utilizes the acid-catalyzed dehydration of L-idose to form the anhydro ring [28].

A 1,6-anhydromuropeptide is formed during the recycling of cell wall peptidoglycan in *E. coli* (Scheme 7) [29,30]. Upon cleavage of the GlcNAc and peptide moieties, the liberated 1,6-anhydro-*N*-acetylmuramic acid residue is converted in two steps to *N*-acetyl-D-glucosamine-6-phosphate which enters an established pathway for amino sugar metabolism [31].

Interestingly, a 1,6-anhydromuropeptide is formed during the recycling of cell wall peptidoglycan in *E. coli* (Scheme 7) [29,30]. Upon cleavage of the GlcNAc and peptide moieties, the liberated 1,6-anhydro-*N*-acetylmuramic acid residue is converted in two steps to *N*-acetyl-D-glucosamine-6-phosphate [31].

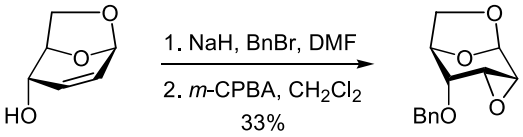
2.2 Reactions

Aside from the protection afforded to the anomeric and 6 positions of 1,6-anhydro sugars, considerable differences in the reactivities at the remaining ring-positions are observed. The reactivities of the hydroxy groups of a number of 1,6-anhydrohexopyranoses towards sulfonylation with tosyl chloride have been determined and were shown to depend upon both the configuration, i. e. axial or equatorial, and the position around the ring [3]. In one case, Pulido has reported a regioselective 4-*O*-acylation of levoglucosan by action of a lipase isolated from *Candida antarctica* [32]; under similar conditions D-glucose is acylated at the primary hydroxyl group [33]. Epoxides of 1,6-anhydrohexopyranoses, commonly called 'Černý' epoxides [25], are useful compounds for the selective introduction of substituents onto the sugar ring. Owing to the conformational restraint in a $^1C_4$ conformation imposed on the sugar ring by the 1,6-anhydro bridge, these reactions generally proceed with reversed regioselectivity when compared to the more usual $^4C_1$ conformation of most hexopyranoses. This is in accordance with the Fürst–Plattner rule that predicts diaxial opening of constrained epoxides. The preparation of such dianhydro sugars from a number of 1,6-anhydropyranoses is relatively straightforward. For example, treatment of 1,6-anhydro-β-D-glucopyranose with tosyl chloride affords the ditosylate (**11**) that, upon treatment with base, is converted in high yield into the 1,6:3,4-dianhydro-β-D-galactoside (**12**) (Scheme 8). In another approach, the 1,6-anhydro-2,3-didehydro sugar (**13**) upon treatment with peracid under Prilezhaev conditions fur-



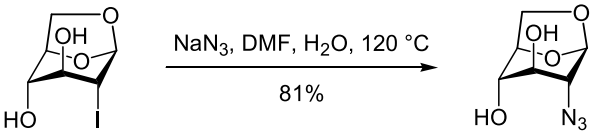
(11) (12)

Scheme 8



(13) (14)

Scheme 9



(15)

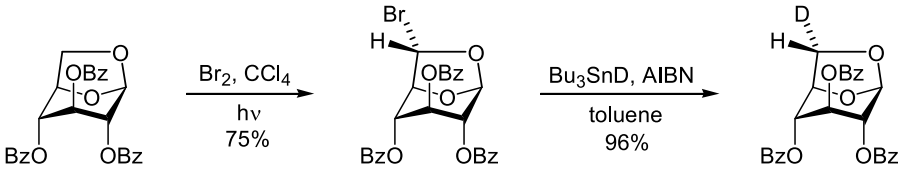
Scheme 10

nishes the 1,6:2,3-dianhydro-2,3-β-D-allopyranose (**14**) (*Scheme 9*) [24]. The presence of the 1,6-anhydro bridge directs formation of the oxirane ring to the *exo*-face.

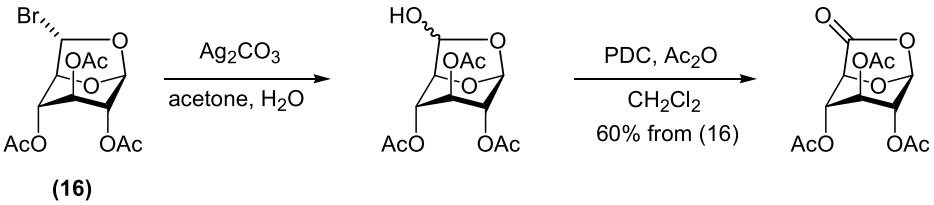
Treatment of the 2-iodo sugar (**15**) (derived by iodocyclization of D-glucal) with sodium azide affords the 2-azido-2-deoxy sugar with retention of stereochemistry, presumably by way of the 1,6:2,3-dianhydro sugar (*Scheme 10*) [26].

Acylated 1,6-anhydro sugars readily undergo photobromination with high regio- and stereoselectivity, affording *exo*-6-bromo sugars (*Scheme 11*) [34,35], and prolonged reaction times yield small amounts of reactive 6,6-dibromo anhydro sugars [36]. The 6-bromo anhydro sugars can be reduced with tributyltin deuteride to the 6-deutero compounds with retention of stereochemistry (*Scheme 11*) [35]. After cleavage of the anhydro ring, this method provides a good route to D-glucose stereoselectively labeled with deuterium at C-6.

6-Bromo anhydro sugars react with nucleophiles, generally with inversion, affording the corresponding *endo*-products [35]; however, many carbon nucleophiles react with high *exo*-selectivity [37]. The reaction of 6-bromo sugars with carbon nucleophiles allows for stereoselective carbon-chain elongation at C-6. This is a useful alternative to the reaction of hexopyranose 6-aldehydes with Grignard reagents. A procedure has been devised for the hydrolysis of the 6-bromo anhydro sugar (**16**); oxidation of the intermediate hemiacetals affords a urono-6,1-lactone (*Scheme 12*) [38].



Scheme 11



(16)

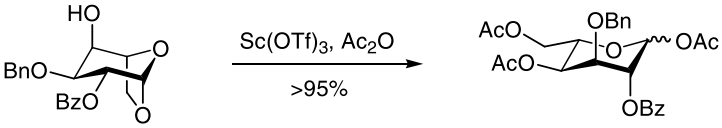
Scheme 12

2.3 Cleavage

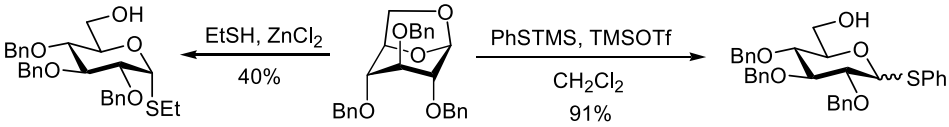
Cleavage of the 1,6-anhydro bridge can be effected by aqueous mineral acids, such as hydrochloric or sulfuric acid. Hydrolysis of 1,6-anhydro-β-D-glucopyranose is achieved in 5 h by 0.5-M sulfuric acid at 100 °C [7]. Glycosans are also susceptible to acetolysis by the action of acetic anhydride and catalytic Lewis acid. The treatment of a number of glycosans with triethylsilyl triflate (TESOTf) in acetic acid provides a rapid cleavage of the 1,6-anhydro bridge in good to excellent yields [39]. Treatment of various protected 1,6-anhydropyranoses with catalytic scandium triflate and a minimum of acetic anhydride leads to the expected products in crude yields exceeding 95% (*Scheme 13*) [40]. This method is particularly attractive as it circumvents the use of moisture sensitive Lewis acids.

The rate of acetolysis of 1,6-anhydro sugars is dependent upon the electron-withdrawing ability of the protecting groups and, in particular, of the group at C-2. Thus, ester protecting groups lower the rate of acetolysis compared to their benzylated counterparts [41].

Methods for the direct conversion of 1,6-anhydropyranoses to thioglycosides have been described using either substituted thio(trimethyl)silanes or ethanethiol in the presence of Lewis acid [42,43]; the resulting thioglycosides are furnished with a free primary hydroxyl group that is available for further reaction (*Scheme 14*).



Scheme 13



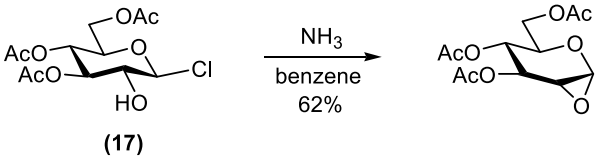
Scheme 14

Ring-opening polymerization of 1,6-anhydro sugars provides an efficient route to high molecular weight, stereoregular glycans. For example, the first reported synthesis of stereoregular (1→6)-α-D-glucopyranan (dextran) was achieved through the ring-opening polymerization of benzylated levoglucosan [44]. More recently, this approach has been employed in the synthesis of dextran sulfonates that possess anti-coagulant activity and inhibitory effects on HIV infection in vitro [45].

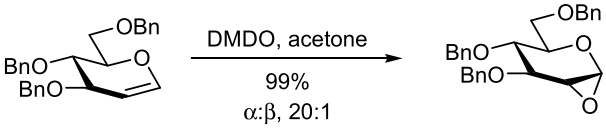
3 1,2-Anhydro Sugars

1,2-Anhydro sugars have proven to be versatile glycosyl donors for the preparation of many glycoconjugates. The 1,2-anhydro bridge is an epoxide that is also part of an acetal and is responsible for the unique reactivity of 1,2-anhydrohexopyranoses. The first well-defined member of this class was prepared by treatment of the partially functionalized pyranose (**17**) with ammonia in benzene, and is commonly referred to as Brigl's anhydride (*Scheme 15*) [46].

Brigl's anhydride was utilized by Lemieux and Huber in the first chemical synthesis of sucrose; a milestone achievement in synthetic organic chemistry [47]. Non-protected derivatives of Brigl's anhydride have not been isolated on account of their high reactivity; however, they have been implicated as intermediates in the alkali-mediated formation of 1,6-anhydro sugars. There exists three main routes to 1,2-anhydropyranoses. The first uses the approach of Brigl: treatment of 2-hydroxy, C-1 activated sugars with base results in an intermediate oxyanion that displaces the leaving group at C-1, thus leading to 1,2-anhydro sugars directly. This method necessitates a *trans*-arrangement of the groups at the C-1 and C-2 positions of the pyranose ring. Alternatively, treatment of 2-*O*-tosyl or 2-*O*-bromo-2-deoxy derivatives of reducing sugars with base yields 1,2-anhydro sugars through *trans*-cyclization [47]. The third method for synthesis of 1,2-anhydro sugars involves epoxidation of glycals. In pioneering work, Halcomb and Danishefsky reported that epoxidation of glycals with 3,3-dimethyldioxirane (DMDO) directly affords 1,2-anhydropyranoses (*Scheme 16*) [48]. DMDO epoxidation enjoys a sim-



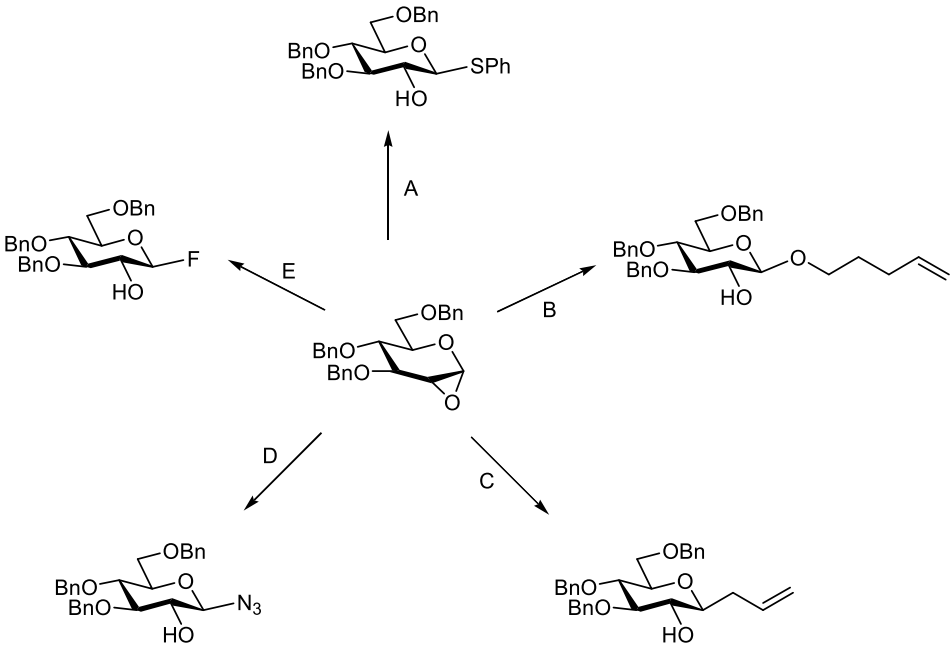
Scheme 15



Scheme 16

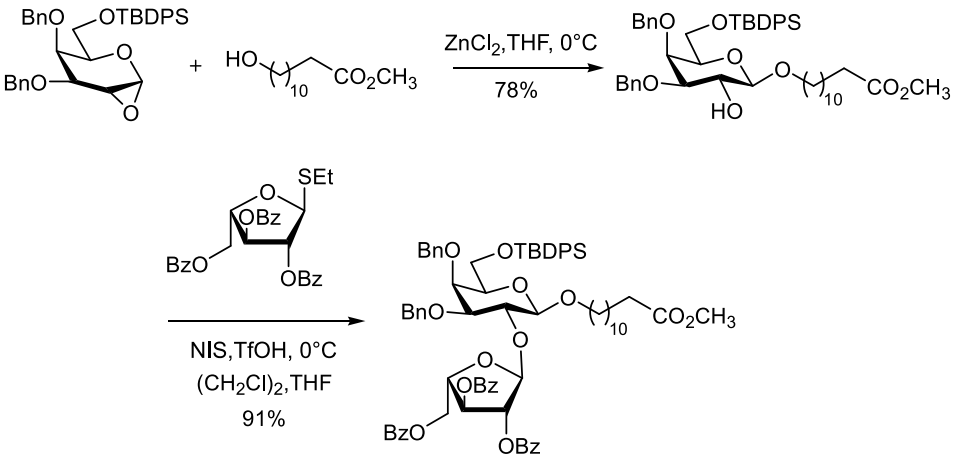
ple work-up, which requires only evaporation. The stereochemistry of the resulting epoxide is usually *trans*- to that of the substituent at C-3 and, as a result, the first two procedures may be necessary for access to certain 1,2-anhydro sugars. It is noteworthy that the use of other epoxidation reagents, such as peracids, usually fail owing to the susceptibility of the anomeric oxirane ring to nucleophilic attack by the solvent or carboxylic acid by-products of the peracid reagents [48].

Interest in 1,2-anhydro sugars stems largely from their ability to function as glycosyl donors under neutral or mildly acidic conditions. Danishefsky and Bilodeau have reported routes to several complex oligosaccharides in which glycosylation steps were carried out using 1,2-anhydro sugars as glycosyl donors [49]. In each case, the epoxide donors were activated with Lewis acid and treated with suitably protected alcohols to afford glycosides in good to excellent yields. For reactions of 1,2-anhydro sugars with alcohol acceptors, including



Scheme 17

***Reagents*: (A) Bu_4NSPh, α:β=1:6, 50%; (B) pent-4-en-1-ol, $ZnCl_2$, 63%; (C) AllMgBr, 75%; (D) DIBAL-LiN_3, 73%; (E) TBAF, 53%**



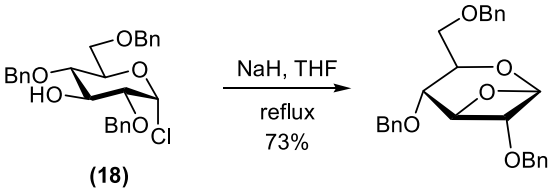
Scheme 18

those in which the hydroxyl function is part of a saccharide, the catalyst most frequently employed is anhydrous zinc chloride [49]. Importantly, glycal epoxides can be easily converted into a number of glycosides or glycosylating agents such as thioglycosides, *n*-pentenyl glycosides, glycosyl fluorides, glycosyl azides, and *C*-glycosides (*Scheme 17*) [49,50,51]. A propitious feature of these conversions is that *O*-2 is rendered free of protecting groups and this free hydroxyl group is amenable to glycosylation in a subsequent step, allowing for the ready preparation of 1,2-glycosides (*Scheme 18*) [52]. The use of glycals as precursors for 1,2-anhydro sugar glycosyl donors has been developed by the Danishefsky group and termed the 'glycal approach' or 'glycal assembly method.' This approach can be used in an iterative fashion for the stepwise elongation of saccharides; thus, a glycal is converted to a 1,2-anhydro sugar and used as a glycosyl donor to glycosylate a selectively protected glycal acceptor; the resultant glycoside is a glycal and sequential epoxidation/glycosylations can be performed. The 'glycal approach' has been utilized on a solid phase [53].

4 Miscellaneous Anhydro Sugars

4.1 1,3-Anhydro Sugars

Interest in 1,3-anhydrosugars stems predominately from their analogy to the dioxabicyclo[3.1.1]heptane ring system present in the blood-platelet aggregation factors, the thromboxanes. The first synthesis of a thromboxane-like nucleus was achieved by the action of sodium hydride on the glycosyl chloride (**18**) (*Scheme 19*) [54]. In general, 1,3-anhydrosugars are prepared in an analogous manner to 1,6-anhydro sugars through treatment of activated sugars (at C-1 or C-3) with base. Still and co-workers have described a procedure for the synthesis of 1,3-anhydro sugars using a modified Mitsunobu reaction on a sugar hemiacetal unprotected at C-3 [55]. 1,3-Anhydrosugars have been utilized in the synthesis of (1→3)-β-D-glucans by

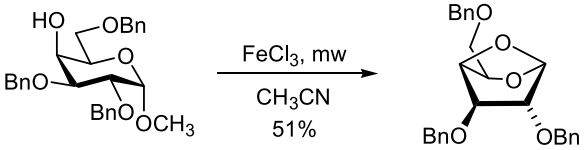


Scheme 19

stereoregular polymerization [54]; polymeric glycoforms of this type are commonly found in plants, bacteria, yeast, and fungi.

4.2 1,4-Anhydropyranoses (1,5-Anhydrofuranoses)

Relatively few studies have been undertaken on 1,4-anhydro sugars. 1,4-Anhydropyranoses (which may be considered as 1,5-anhydrofuranoses) have been synthesized by treatment of glycosyl fluorides possessing an unprotected C-4 hydroxy group with Lewis acid, or by treatment of 4-*O*-sulfonyl (usually mesyl) derivatives with bases such as sodium azide [56,57,58]. Reaction times may be shortened by mw irradiation (*Scheme 20*) [59], however, as with conventional heating, the anhydro bridge remains susceptible to ring-opening and formation of self-polymerization by-products.

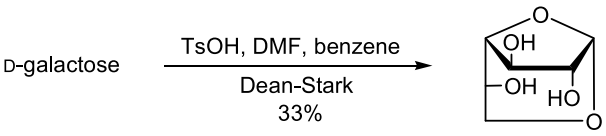


Scheme 20

4.3 1,6-Anhydrohexofuranoses

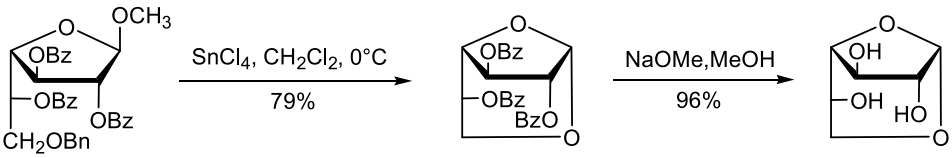
1,6-Anhydrohexofuranoses are present to a small extent in equilibrium mixtures of aldohexoses in water [1] and have been isolated from the pyrolysis distillates of sugars. 1,6-Anhydro-β-D-glucofuranose and 1,6-anhydro-α-D-galactofuranose were first isolated from pyrolysis residues after removal of the bulk of the more abundant 1,6-anhydrohexopyranoses [60,61]. These two sugars have been the subject of review [3,62]. Specific methods for the synthesis of 1,6-anhydrohexofuranoses are somewhat limited; however, Angyal and Beveridge have reported that treatment of dilute solutions of a number of aldohexoses in DMF/benzene with 4-toluenesulfonic acid under Dean–Stark conditions affords moderate yields of 1,6-anhydrohexofuranoses such as (**19**) (*Scheme 21*) [63].

As described earlier, treatment of methyl α-D-glucoside (**5**) with camphorsulfonic acid under Dean–Stark conditions yields a 7:1 mixture of 1,6-anhydro-β-D-glucopyranose (**6**)



(19)

Scheme 21



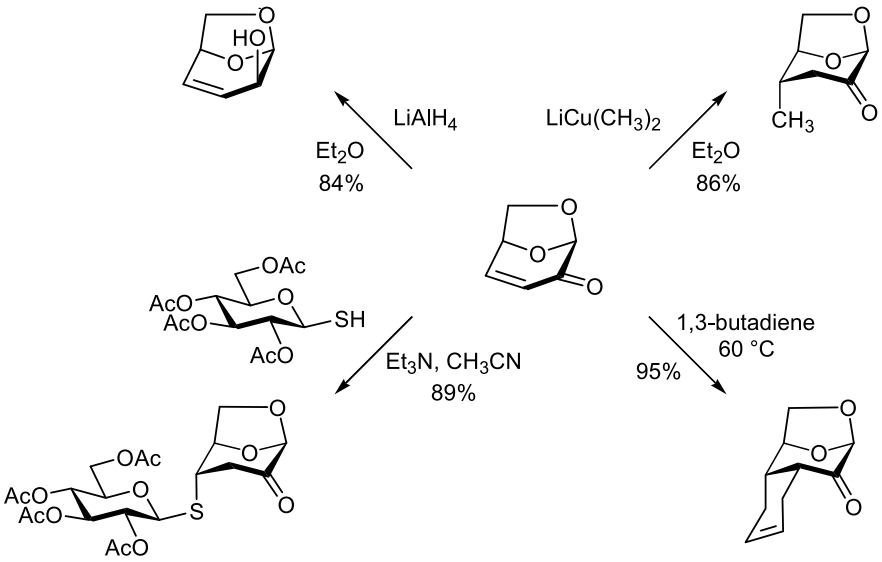
(20) **(19)**

Scheme 22

and 1,6-anhydro-β-D-glucofuranose (**7**) (*Scheme 4*). Dehydration of methyl α-D-galactopyranoside with CSA yields an equimolar mixture of the furanose and pyranose anhydro derivatives [21]. In an alternative approach, synthesis of 1,6-anhydro-α-D-galactofuranose (**19**) through stannic chloride catalyzed ring-closure of the fully protected galactofuranose (**20**) was achieved in two steps in excellent yield (*Scheme 22*) [64]. It has been observed that the *trans*-oriented 2,3-diols of these bicyclic compounds possess a remarkable resistance to both periodate and lead tetraacetate oxidation, a consequence of the rigidity imposed upon these molecules by the 1,6-anhydro bridge.

4.4 Levoglucosenone

Levoglucosenone is an enone-containing 1,6-anhydro sugar, the considerable differences of which merit a separate section to 1,6-anhydro sugars, although only a brief overview of the extensive chemistry pertaining to this compound is possible here [65]. The presence of the acetal and enone moieties of levoglucosenone provides ready access to a wide range of functionalized sugars from a single chemical feedstock. Halpern and co-workers properly characterized levoglucosenone in 1973, correcting several structures proposed by other workers [66]. Levoglucosenone may be prepared by the Brønsted-acid catalyzed pyrolysis of cellulosic materials and is easily isolated in multi-gram quantities from the pyrolyzate of waste paper [67]. The 1,6-anhydro bridge sterically hinders the *endo*-face of levoglucosenone and, as a consequence, nucleophilic addition is directed to the non-hindered *exo*-face. For example, reduction of the carbonyl group with sodium borohydride or lithium aluminum hydride leads to the expected allylic alcohol in good yield [68,69]. Carbon nucleophiles derived from Grignard or Gilman reagents add to the β-carbon yielding the corresponding Michael adducts. In this way, stereospecific addition of methylmagnesium iodide or lithium dimethylcuprate

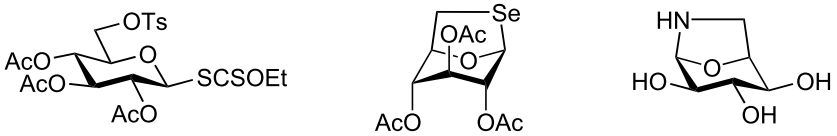


Scheme 23

to the enone of levoglucosenone leads to C-4 methyl-substituted carbohydrates [69,70]. Furthermore, conjugate addition of glycosyl nitronates [71] or thioglycosides [72] with levoglucosenone affords *C*- or *S*-linked disaccharides, respectively. The alkene of levoglucosenone acts as a dienophile in Diels–Alder reactions [73]. Some transformations of levoglucosenone are summarized in *Scheme 23*.

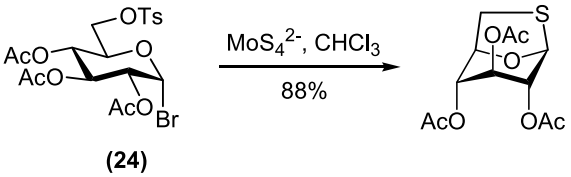
5 Anhydro Sugars Containing Nitrogen, Sulfur, or Selenium

A number of anhydro sugars are known that contain nitrogen, sulfur, or selenium. Tri-*O*-acetyl thioglucosan was first prepared in 1963 by Akagi and co-workers by treatment of *S*-glucosyl dithiocarbonate (**21**; *Fig. 2*) with sodium methoxide [74]. Since that time, numerous syntheses have been reported that rely on the introduction of a thioester at C-1 or C-6 and the activation of the alternate position with a halide or sulfonate group [75,76]. Treatment of a 'doubly activated' hexopyranose such as (**24**) with hydrogen sulfide and triethylamine



(21) **(22)** **(23)**

Figure 2



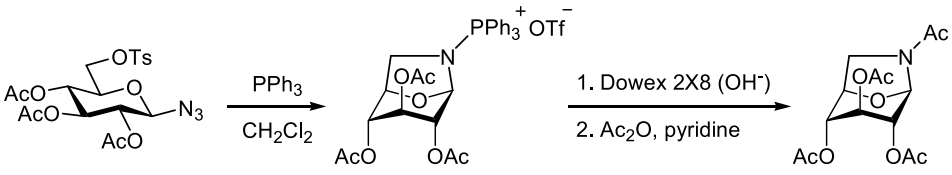
(24)

Scheme 24

leads to thioglycosans in good yields [77]. Treatment of doubly activated sugars with the sulfur-transfer reagent benzyltriethylammonium tetrathiomolybdate [78] now represents the most expeditious route to thioglycosans (*Scheme 24*) [77,79].

Thioglycosans are stable under basic conditions but are cleaved either by sulfuric acid in acetic anhydride to afford a 6-*S*-acetyl anomeric acetate [80] or by dilute hydrochloric acid in water to yield the 6-thioglycose [74]. The sulfur atom of thiolevoglucosan is susceptible to oxidation and efficient routes to the corresponding sulfoxides and sulfones have been reported [81]. The sulfoxides undergo stereospecific Pummerer rearrangement to *exo*-α-acetoxy sulfides, which can be oxidized in turn to α-acetoxy sulfoxides and sulfones [81]. Thioglycosans are intramolecular thioglycosides and this observation has led to their investigation as glycosyl donors [75]. Stick and co-workers have shown that thioglucosans are effective glycosyl donors when activated with the *N*-iodosuccinimide/TfOH couple [82]. The initial products are disulfide-linked disaccharides and desulfurization (with Raney nickel) leads to 6-deoxy glycosides. A selenium counterpart, tri-*O*-acetyl selenoglucosan (**22**; *Fig. 2*) was prepared through the agency of sodium hydrogen selenide (prepared from sodium borohydride and selenium in anhydrous ethanol) and the doubly activated sugar (**24**) [77]. Selenoglycosans can also act as glycosyl donors in the presence of NIS/TfOH, and the weak nature of the carbon–selenium bond allows a mild tributylstannane-mediated reduction of intermediate 6-seleno sugars to afford the 6-deoxy glycosides [83]. Detailed structural analyses of a number of sulfur and selenium-containing 1,6-anhydro sugars have been reported [81,84].

Several nitrogen-containing anhydro sugars are known. These compounds, while formally anhydro amino sugars, are better considered aza sugars. In alkaline media, equilibrated solutions of 6-amino-6-deoxy aldoses contain varying amounts of the corresponding aza sugars, such as the amine (**23**; *Fig. 2*) [3,85]. Treatment of C-6 activated anomeric azides with triphenylphosphine yields anomeric iminophosphoranes that rearrange in situ by elimination of the *O*-6 sulfonate affording, after ion-exchange chromatography and *N*-acylation, 6-amino-1,6-anhydro-6-deoxy-β-D-glycopyranoses (*Scheme 25*) [86].



(25) 84% from **(25)**

Scheme 25

References

1. Angyal SJ, Dawes K (1968) Aust J Chem 21:2747
2. Kochetkov NK (1987) Tetrahedron 43:2389
3. Černý M, Stanek J Jr (1977) Adv Carbohydr Chem Biochem 34:23
4. Bols M (1996) 1,6-Anhydro sugars. Carbohydrate building blocks. Wiley, New York
5. Peat S (1946) Adv Carbohydr Chem 2:37
6. Witczak ZJ (1994) Selective protection of levoglucosan derivatives. In: Witczak ZJ (ed) Frontiers in biomedicine and biotechnology. Levoglucosenone and levoglucosans, chemistry and applications. ATL Press, Mount Prospect, 2:165
7. Tanret M (1894) Bull Soc Chim Fr 211:944
8. Park YJ, Kim HS, Jeffrey GA (1971) Acta Cryst B 27:220
9. Heyns K, Weyer J (1968) Liebigs Ann Chem 718:224
10. Pictet A, Sarasin J (1918) Helv Chim Acta 1:87
11. Ward RB (1963) Methods Carbohydr Chem 2:394
12. Shafizadeh F, Furneaux RH, Stevenson TT, Cochran TG (1978) Carbohydr Res 61:519
13. Schkolnik G, Rudich Y (2006) Anal Bioanal Chem 385:26
14. Knauf AE, Hann RM, Hudson CS (1941) J Am Chem Soc 63:1447
15. Hann RM, Hudson CS (1942) J Am Chem Soc 64:925
16. Zottola MA, Alonso R, Vite GD, Fraser-Reid B (1989) J Org Chem 54:6123
17. Cleophax J (2003) Synthesis 1015
18. Coleman GH (1963) Methods Carbohydr Chem 2:397
19. Boons G-J, Isles S, Setälä P (1995) Synlett 755
20. Rao MV, Nagarajan M (1987) Carbohydr Res 162:141
21. Caron S, McDonald A, Heathcock CH (1996) Carbohydr Res 281:179
22. Haeckel R, Lauer G, Oberdorfer F (1996) Synlett 21
23. Mereyala HB, Venkataramanaiah KC, Dalvoy VS (1992) Carbohydr Res 225:151
24. Lauer G, Oberdorfer F (1993) Angew Chem Int Ed Engl 32:272
25. Černý M (1994) 1,6:2,3- and 1,6:3,4-Dianhydro-β-D-hexopyranoses. Synthesis and preparative applications. In: Witczak ZJ (ed) Frontiers in biomedicine and biotechnology. Levoglucosenone and levoglucosans, chemistry and applications. ATL Press, Mount Prospect, 2:121
26. Tailler D, Jacquinet JC, Noirot AM, Beau JM (1992) J Chem Soc Perkin 1 3163
27. Černý M (2003) Adv Carbohydr Chem Biochem 58:122
28. Stoffyn PJ, Jeanloz RW (1960) J Biol Chem 235:2507
29. Hölte JV (1998) Microbiol Mol Biol R 62:181
30. Hölte JV, Mirelman D, Sharon N, Schwarz U (1975) J Bacteriol 124:1067
31. Uehara T, Suefuji K, Valbuena N, Meehan B, Donegan M, Park J (2005) J Bacteriol 187:3643
32. Pulido R, Gotor V (1994) Carbohydr Res 252:55
33. Ljunger G, Adlercreutz P, Mattiasson B (1994) Biotechnol Lett 16:1167
34. Somsák L, Ferrier RJ (1991) Adv Carbohydr Chem Biochem 49:37
35. Ohrui H, Horiki H, Kishi H, Meguro H (1983) Agric Biol Chem 47:1101
36. Ferrier RJ, Furneaux RH (1980) Aust J Chem 33:1025
37. Nishikawa T, Mishima Y, Ohyabu N, Isobe M (2004) Tetrahedron Lett 45:175
38. Vogel C, Liebelt B, Steffan W, Kristen H (1992) J Carbohydr Chem 11:287
39. Zottola M, Rao BV, Fraser-Reid B (1991) J Chem Soc Chem Commun 969
40. Lee JC, Tai CA, Hung SC (2002) Tetrahedron Lett 43:851
41. Burgey CS, Vollerthun R, Fraser-Reid B (1994) Tetrahedron Lett 35:2637
42. Wang L-X, Sakairi N, Kuzuhara H (1990) J Chem Soc Perkin 1 1677
43. Koto S, Uchida T, Zen S (1972) Chem Lett 1049
44. Ruckel ER, Schuerch C (1966) J Org Chem 31:2233
45. Yoshida T, Nakashima H, Yamamoto N, Uryu T (1993) Polymer J 25:1069
46. Brigl P (1922) Hoppe-Seyler's Z Physiol Chem 122:245
47. Lemieux RU, Huber G (1956) J Am Chem Soc 78:4117
48. Halcomb RL, Danishefsky SJ (1989) J Am Chem Soc 111:6661
49. Danishefsky SJ, Bilodeau MT (1996) Angew Chem Int Ed Engl 35:1380
50. Collins P, Ferrier R (1995) Monosaccharides: Their Chemistry and Their Roles in Natural Products. Wiley, Chichester

51. Leeuwenburgh MA, van der Marel GA, Overkleeft HS, van Boom JH (2003) J Carbohydr Chem 22:549
52. Timmers CM, Wigchert SCM, Leeuwenburgh MA, van der Marel GA, van Boom JH (1998) Eur J Org Chem 1998:91
53. Seeberger PH, Danishefsky SJ (1998) Acc Chem Res 31:685
54. Ito H, Eby R, Kramer S, Schuerch C (1980) Carbohydr Res 86:193
55. Bhagwat SS, Hamam PR, Still WC (1985) J Am Chem Soc 107:6372
56. Kops J, Schuerch C (1965) J Org Chem 30:3951
57. Bullock C, Hough L, Richardson AC (1990) Carbohydr Res 197:131
58. Thiem J, Wiesner M (1993) Carbohydr Res 249:197
59. Nokami T, Werz DB, Seeberger PH (2005) Helv Chim Acta 88:2823
60. Dimler RJ, Davis HA, Hilbert GE (1946) J Am Chem Soc 68:1377
61. Alexander BH, Dimler RJ, Mehltretter CL (1951) J Am Chem Soc 73:4658
62. Dimler RJ (1952) Adv Carbohydr Chem 7:37
63. Angyal SJ, Beveridge RJ (1978) Aust J Chem 31:1151
64. Sarkar SK, Choudhury AK, Mukhopadhyay B, Roy N (1999) J Carbohydr Chem 18:1121
65. Witczak ZJ (1994) Levoglucosenone; past, present and further applications In: Witczak ZJ (ed) Frontiers in biomedicine and biotechnology. Levoglucosenone and levoglucosans, chemistry and applications. ATL Press, Mount Prospect, 2:3
66. Halpern Y, Riffer R, Broido A (1973) J Org Chem 38:204
67. Shafizadeh F, Furneaux RH, Stevenson TT (1979) Carbohydr Res 71:169
68. Brimacombe JS, Hunedy F, Tucker LCN (1978) Carbohydr Res 60:C11
69. Shafizadeh F, Chin PPS (1977) Carbohydr Res 58:79
70. Mori M, Chuman T, Kato K (1984) Carbohydr Res 129:73
71. Witczak ZJ (1994) Pure Appl Chem 66:2189
72. Witczak ZJ, Chhabra R, Chen H, Xie X-Q (1997) Carbohydr Res 301:167
73. Ward DD, Shafizadeh F (1981) Carbohydr Res 95:155
74. Akagi M, Tejima S, Haga M (1963) Chem Pharm Bull 11:58
75. Lundt I, Skelbæk-Pedersen B (1981) Acta Chem Scand B35:637
76. Whistler RL, Seib PA (1966) Carbohydr Res 2:93
77. Driguez H, McAuliffe JC, Stick RV, Tilbrook DMG, Williams SJ (1996) Aust J Chem 49:343
78. Ramesha AR, Chandrasekaran S (1992) Synth Commun 22:3227
79. Sridhar PR, Saravanan V, Chandrasekaran S (2005) Pure Appl Chem 77:145
80. Yamamoto K, Haga M, Tejima S (1975) Chem Pharm Bull 23:233
81. Skelton BW, Stick RV, Tilbrook DMG, White AH, Williams SJ (2000) Aust J Chem 53:389
82. Stick RV, Tilbrook DMG, Williams SJ (1997) Tetrahedron Lett 38:2741
83. Stick RV, Tilbrook DMG, Williams SJ (1999) Aust J Chem 52:685
84. Buděšínský M, Poláková J, Hamerníková M, Císařová I, Trnka TS, Černý M (2006) Coll Czech Chem Commun 71:311
85. Paulsen H, Todt K (1967) Chem Ber 100:512
86. Lafont D, Wollny A, Boullanger P (1998) Carbohydr Res 310:9

3.8 *C*-Glycosylation

Toshio Nishikawa, *Masaatsu Adachi*, *Minoru Isobe*
Graduate School of Bioagricultural Sciences, Nagoya University,
464-8601 Nagoya, Japan
nisikawa@agr.nagoya-u.ac.jp, madachi@agr.nagoya-u.ac.jp,
isobem@agr.nagoya-u.ac.jp

Abstract

This chapter deals with *C*-glycosylation, a carbon–carbon bond-forming reaction at an anomeric carbon of carbohydrate and its derivatives. Since synthesis of natural products containing the *C*-glycosidic linkage was one of the major issues in organic synthesis in the 1970s, extensive efforts have been devoted to developing stereoselective synthesis of *C*-glycoside, which includes a variety of C–C bond formations by using anomeric cations (oxocarbenium cations), anomeric radicals, and anomeric anions, as well as sigmatropic

In: *Glycoscience*. Fraser-Reid B, Tatsuta K, Thiem J (eds)
Chapter-DOI 10-1007/978-3-540-30429-6_17:

rearrangement and transition metal-catalyzed reactions. Among these reactions, useful *C*-glycosylations with high stereoselectivities are reviewed in this chapter with an emphasis on the reaction mechanisms.

Keywords

C-Glycosylation; Oxocarbenium cation; Anomeric radical; Anomeric anion; Stereoelectronic effect

Abbreviations

HMPA hexamethylphosphoramide
MCPBA *m*-chloroperoxybenzoic acid
TIPS triisopropylsilyl
TMSCN trimethylsilyl cyanide

1 Introduction

C-Glycosylation, a carbon–carbon bond-forming reaction at an anomeric center of carbohydrates and related compounds, has been the subject of great attention due to the importance of *C*-glycoside-containing compounds such as *C*-glycoside mimics of glycolipids, oligosaccharides, and glycoproteins, and natural products in chemistry and biology.

In 1993, researchers of the Kirin Brewery Corporation, Japan, reported the isolation and structure of agelasphines, novel anticancer sphingolipids isolated from an Okinawan marine sponge, *Agelas mauritianus* (◉ *Fig. 1*) [1]. Since these compounds are the first examples of α-galactosyl ceramide found in nature, and a novel mode of action for their anticancer activity was expected, development of an antitumor drug candidate based on agelasphines led to KRN7000, a synthetic sphingolipid. Extensive studies regarding the mode of action revealed that KRN7000 activated natural killer T-cells by binding CD1d, an antigen presentation protein, thus releasing cytokines including Interferon-γ (IFN-γ) causing tumor rejection and Interleukin-4 (IL-4) suppressing autoimmune diseases. Interestingly, OCH, a truncated synthetic analog of KRN7000 induced a predominant production of IL-4 over IFN-γ, while KRN7000 induced production of both cytokines. In 2003, Franck and co-workers reported the synthesis of a *C*-glycoside analog of KRN7000, surprisingly, this compound exhibited 1000-fold more activity than KRN7000 in a mouse malaria assay as well as a 100-fold activity increase in a mouse melanoma model. Significant efforts have been devoted towards the synthesis of *C*-glycoside that specifically releases one of these cytokines in order to yield compounds for chemotherapeutic application. This recent example demonstrates the importance of *C*-glycoside analogs of biologically active glycoconjugates in biology and pharmacy. These results are stimulating development of stereoselective and efficient methods for synthesizing *C*-glycoside analogues of KRN7000 (◉ *Fig. 1*).

The *C*-glycosidic linkage is also found in many biologically important natural products such as marine natural toxins and antibiotics as shown in ◉ *Fig. 2*. Since synthesis of these structurally complex natural products has been a major subject in modern organic synthesis, stere-

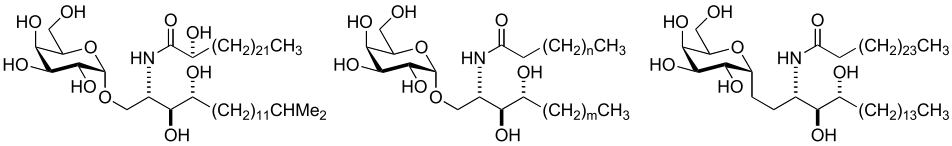


Figure 1
Agelasphines, novel anticancer sphingolipids and its related synthetic compounds

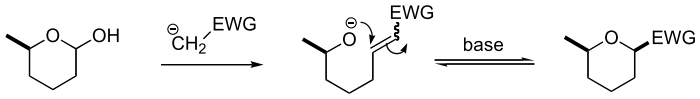
oselective *C*-glycosylation has been extensively studied. In the synthesis of these molecules, carbohydrates have been employed as chiral starting materials (chiral pool), because a variety of carbohydrates with multi-functionality are readily available from nature [2]. In *Fig. 2*, the arrows indicate the C–C bonds that were synthesized by "*C*-glycosylation". It is worth noting that *C*-glycosylation has been also employed for the synthesis of hidden *C*-glycosidic bonds (see the structure of brevetoxin A).
This chapter will survey *C*-glycosylation by classifying the reactions mainly according to the type of reaction and in some cases the structure of the substrate with focus on the selectivity and its applicability [3].

2 Nucleophilic Addition to Electrophilic Carbohydrate Derivatives

Addition reaction of carbon nucleophile to electrophilic carbohydrate derivative is one of the most widely used methods for *C*-glycosylation. Since there are many carbohydrate derivatives that have been reported for *O*-glycosylation, these materials have also been employed for *C*-glycosylations. The difference between *O* and *C*-glycosylation is lack of the thermodynamic anomeric effect to control stereochemistry, and neighboring group participation has been rarely operated in *C*-glycosylation. In this section, the reactions are classified by the structure of the electrophilic carbohydrate such as (a) lactol, (b) lactone, (c) glycoside, (d) glycal, and (e) anhydrosugar.

2.1 Lactol as an Electrophilic Carbohydrate

Reaction of lactol with stabilized nucleophilic reagents such as 1,3-diketone and the Wittig reagent is a classic method and is still a useful method for the synthesis of *C*-glycosides. The reaction undergoes formation of olefin with an electron-withdrawing group followed by intramolecular conjugate addition with the resulting alkoxide anion (*Eq. 1*). As the process is reversible, thermodynamically stable *C*-glycoside would be obtained as the major product. Among many of these reactions reported to date, several recent examples with special characters are shown below.



(1)

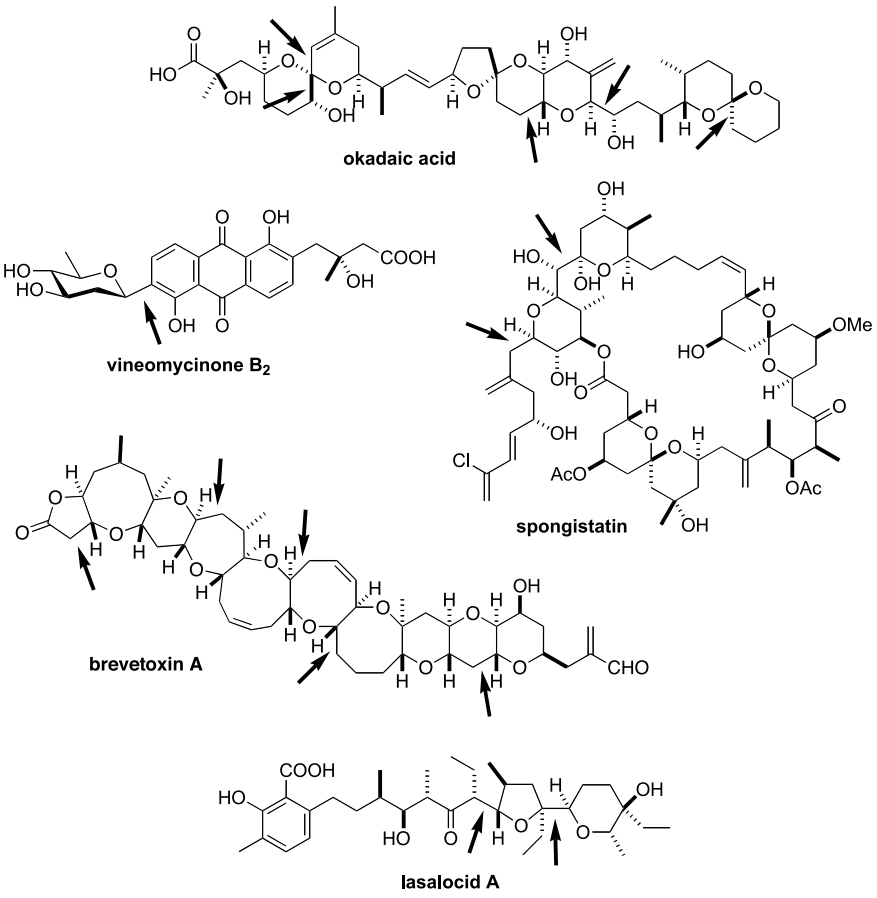
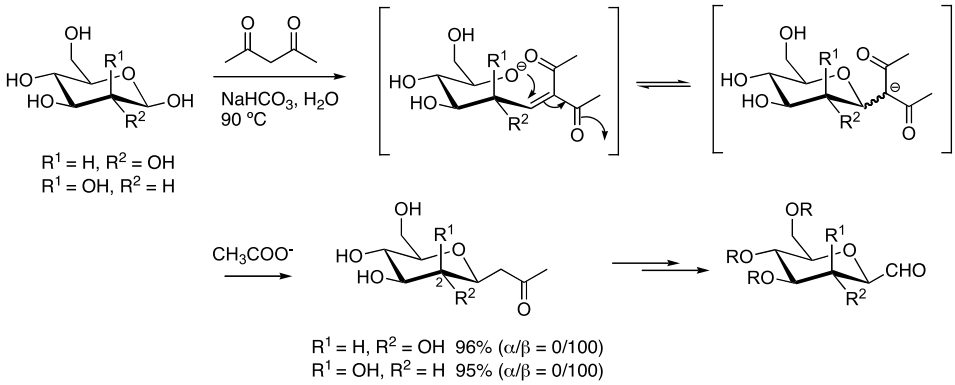


Figure 2
Naturally occurring compounds containing the *C*-glycosidic linkage

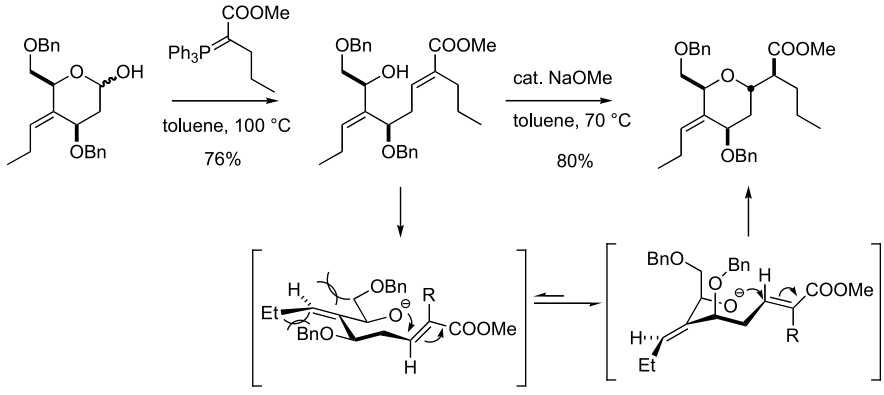
One-step synthesis of β-glycosyl ketone from unprotected carbohydrate in aqueous media was reported by Rubineau and co-workers (*Scheme 1*) [4]. Unprotected D-glucose, mannose, and galactose were treated with pentane-2,4-dione at 90 °C in aqueous sodium bicarbonate to afford β-glycosyl ketones in high yields. The reaction began with Knoevenagel condensation to give the unsaturated ketone, which underwent an intramolecular conjugate addition of the alkoxide. The reaction conducted at room temperature gave a mixture of the four possible furanosides and pyranosides. It is worth noting that epimerization at the C-2 position did not occur in the reaction with glucose and mannose, while the same reactions with *N*-acetyl-gluco-, *manno*-, and galactosamine afforded a mixture of the C-2 epimeric products. The β-glycosyl ketone was transformed into β-1-formyl sugar, a versatile synthetic intermediate for a variety of *C*-glycosidic compounds.



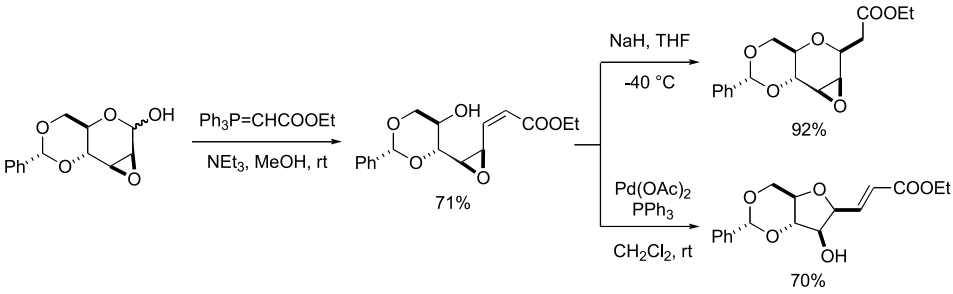
Scheme 1

This reaction generally gives β-*C*-hexopyranose, however, α-*C*-glycoside could be obtained in a special case as shown in *Scheme 2* [5]. Reaction of 2-deoxylactol with a Wittig reagent yielded an unsaturated ester, which was treated with a catalytic amount of NaOMe to afford α-*C*-glycoside as a single product in good yield. Preferential formation of the *C*-glycoside with α-configuration is rationalized by the transition state conformers shown in *Scheme 2*. The right conformer with the two axial substituents is involved in the recyclization step, affording α-*C*-glycoside, because the left conformer is much destabilized by A-strain between the exo-olefin and the two equatorial substituents. The product was further transformed into luminacins, angiogenesis inhibitors isolated from a *Streptomyces*.

Epoxy lactol was employed for the *C*-glycosylation depicted in *Scheme 3* [6]. The lactol was treated with a stabilized Wittig reagent to preferentially give the *Z*-unsaturated ester, which upon exposure to NaH gave the thermodynamically stable β-*C*-pyranoside through conjugate addition of the resulting alkoxide. On the other hand, the same unsaturated epoxide was treated with a palladium catalyst to give β-*C*-furanoside as a single product via a π-allyl palladium complex with net retention of configuration.



Scheme 2



Scheme 3

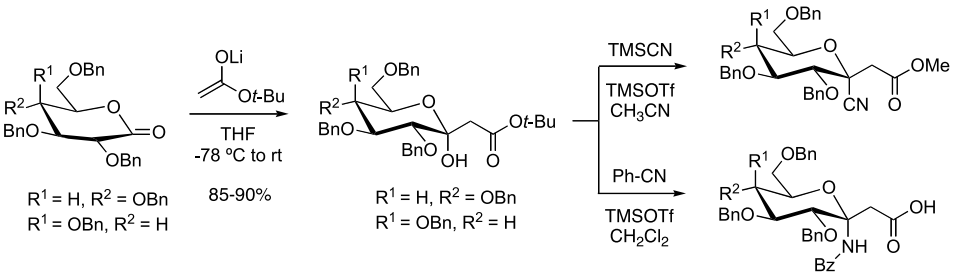
2.2 Lactone as an Electrophilic Carbohydrate

Glyconolactones (Sugar lactones) have been widely employed as electrophilic donors for *C*-glycosylation, because of their easy availability from oxidation of the corresponding hemiacetal. Reaction of the lactone with a carbon nucleophile gives lactol, which can be further manipulated into *C*-glycoside or olefin at the anomeric position. This section deals with *C*-glycosylations starting from sugar lactone classified by the nucleophiles; (1) enolate and its related anion stabilized by carbonyl, sulfonyl group, (2) Wittig and its related reagents [7], and (3) organometallic reagents such as Grignard and organolithium reagents followed by reduction.

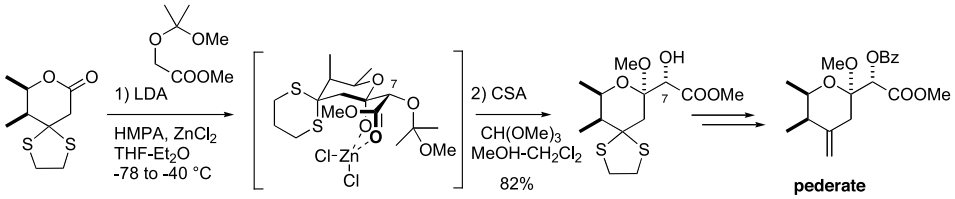
2.2.1 Reaction with Enolate and Related Anions

Reaction of sugar lactone with ester enolate (the Claisen condensation) gives β-keto ester, the resulting 1,3-dicarbonyl compound immediately reacts with the hydroxyl group present in the same molecule to afford a lactol with an axial hydroxyl group. In a recent example shown in *Scheme 4* [8], the product of the reaction was further transformed to the corresponding nitrile or amide by addition of cyanide or a Ritter reaction with benzonitrile in the presence of TMSOTf as a Lewis acid.

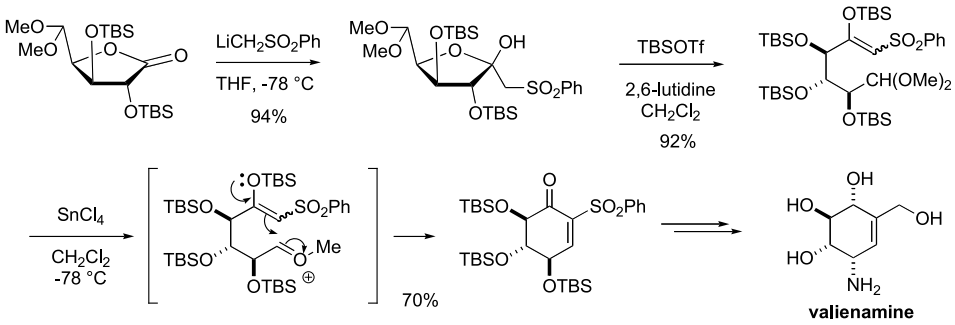
In the reaction with an enolate of α-substituted ester, stereocontrol of the α-position is generally difficult because of the easy epimerization. However, Nakata reported a highly stereose-



Scheme 4



Scheme 5



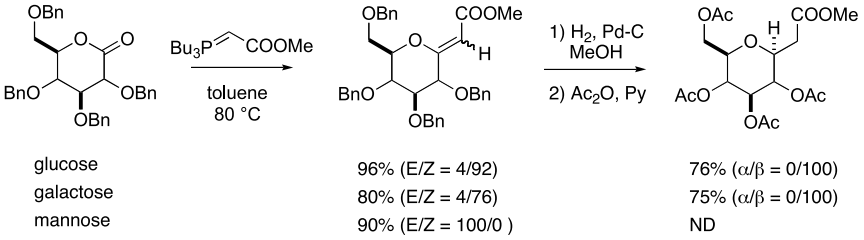
Scheme 6

lective reaction of lactone with enolate derived from glycolate in the presence of $ZnCl_2$ and HMPA (*Scheme 5*) [9]. The high stereoselectivity was explained by formation of the thermodynamically stable chelate intermediate, which has an equatorial hydroxyl group at the C-7 position. The product was further transformed to pederate, the left part of pederine, an insect toxin.

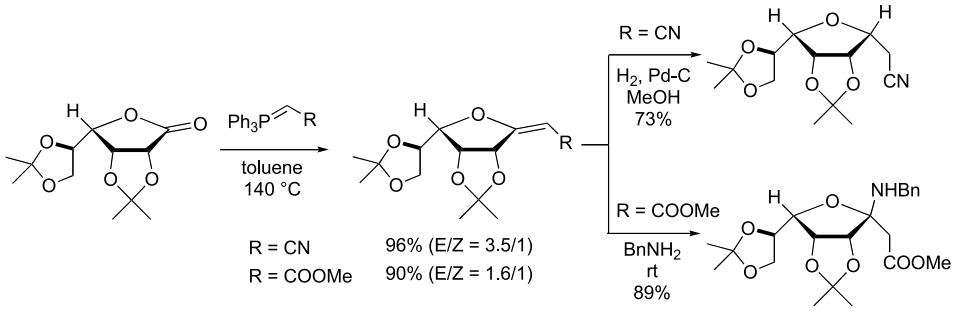
Many carbanions stabilized by other electron-withdrawing groups have also been employed as nucleophiles. For example, the lithium anion of methyl phenyl sulfone reacted with furanolactone to give a hemiacetal, which was transformed to silyl enol ether (*Scheme 6*) [10]. Upon treatment with $SnCl_4$ as a Lewis acid, an intramolecular reaction of the silyl enol ether (the so-called Mukaiyama reaction) takes place followed by elimination of methanol, giving the highly substituted cyclohexenone, an important intermediate for valienamine, a glycosidase inhibitor.

2.2.2 Reaction with Wittig Reagents and Related Reagents

Reaction of ester with Wittig reagents has rarely been employed, because of the low reactivity compared with aldehyde and ketone. However, sugar lactone with suitable protecting groups reacts with (carbomethoxymethylene)tributylphosphorane at elevated temperature to give *exo*-glycal (also known as glycosylidene) in good yield. Examples of Wittig reactions with glucono-, galactono-, and mannolactone are shown in *Scheme 7* [11]. Lactones derived from amino sugars such as GalNAc and GlcNAc also underwent the Wittig reaction giving the corresponding *exo*-glycals in good yields [12]. The carbon–carbon double bond can be stereoselectively reduced by hydrogenation to give β-*C*-glycoside.



Scheme 7

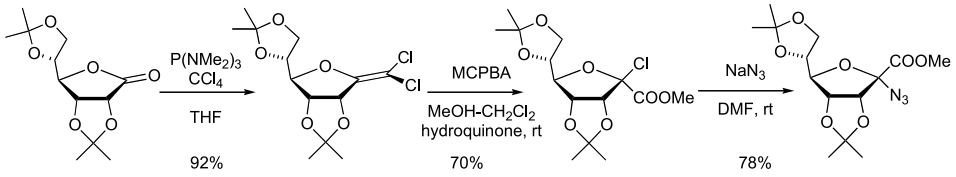


Scheme 8

Wittig reactions of furanolactone have also been reported (Scheme 8) [13]. Use of microwave irradiation reduces the reaction time due to the effect of microwave flash heating. The resulting *exo*-methylene moiety can be utilized as a stepping stone for further modifications. Catalytic hydrogenation gave α-*C*-glycoside, while conjugate addition of benzylamine afforded β-amino acid derivatives in a stereoselective manner [14].

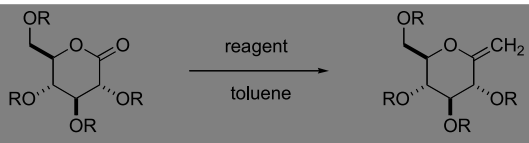
Sugar lactone was treated with tris(dimethylamino)phosphine-carbon tetrachloride or triphenylphosphine-carbon tetrachloride to give *exo*-dichloromethylene as shown in Scheme 9 [15]. A wide range of sugar lactones including furanolactone and pyranolactone with a variety of protecting groups have been employed as a substrate for this reaction. The *exo*-methylene moiety was transformed to an amino acid precursor by oxidation with MCPBA followed by addition of azide to the resulting chloride [16].

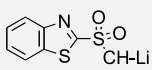
Simple methylenation of sugar lactone was carried out by a titanium-based reagent such as Tebbe and Petasis reagent (Me_2TiCp_2) (Table 1) [17]. Since the carbonyl group of ester also reacts with these reagents, ester should be avoided as a protecting group. The Tebbe reagent reacted with the lactone at a low temperature (entry 1), while reaction with the Petasis



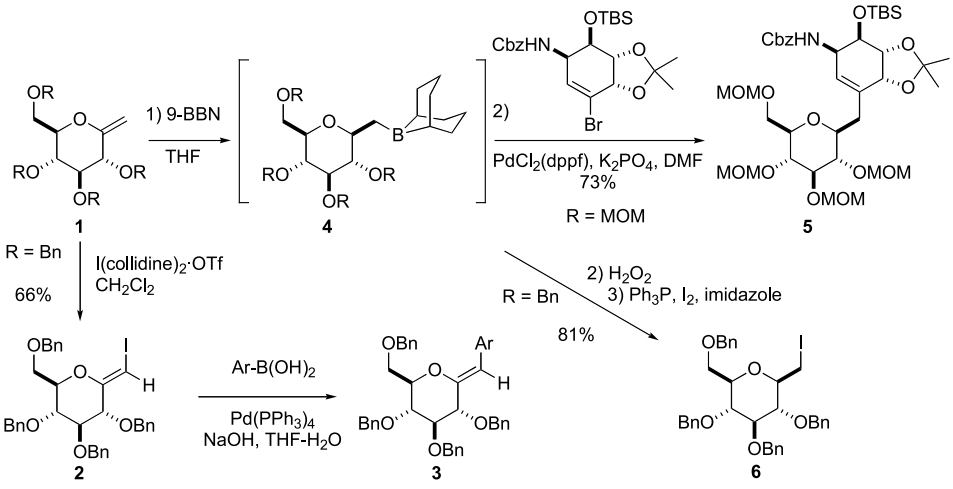
Scheme 9

Table 1
[17]



Entry	R	Reagent	Temp	Yield (%)	Ref.
1	Bn	Cp, Cl, Me, Ti, Al, Cp, Me (Tebbe)	−40 to 0 °C	82	[17a]
2	Bn	Me_2TiCp_2 (Petasis)	70 °C	94	[17b]
3	TMS	Cp, Cl, Me, Ti, Al, Cp, Me (Tebbe)	−40 to 0 °C	54	[17a]
4	MOM	Me_2TiCp_2 (Petasis)	70 °C	79	[17c]
5	TES	i)  ii) DBU, rt	−78 °C	74	[17d]

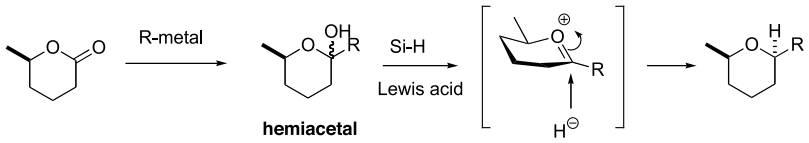
reagent required heating at 60–70 °C. Recently, methylenation of the lactone by Julia coupling was also reported (entry 5). The product, *exo*-glycal is a versatile synthetic intermediate for a variety of *C*-glycosides by means of Suzuki–Miyaura coupling (*Scheme 10*). For example, alkenyl iodide **2** prepared from **1** was coupled with an arylboron reagent to **3** in the presence of a palladium catalyst [17b]. On the other hand, hydroboration of *exo*-glycal **1** with 9-BBN gave alkylborane intermediate **4** [18], which was coupled with alkenyl bromide in the presence of a palladium catalyst to give *β*-*C*-glycoside **5** [17c]. This synthetic methodology has been extensively used for the synthesis of marine polyether toxins. The related reactions will be discussed in *Sect. 6*.



Scheme 10

2.2.3 Reaction with Organometallic Reagents Followed by Reduction of the Newly Formed Hemiacetal

Reaction of glyconolactone with organometallic reagents such as Grignard and organolithium reagents gives hemiacetal (ketose). The resulting hemiacetal is reduced with Et_3SiH in the presence of a Lewis acid such as TMSOTf or $BF_3{\cdot}OEt_2$ to afford *C*-glycoside (*Eq. 2*). In the reduction of a hexopyranose intermediate, the hydride attacks the oxocarbenium cation generated from the reaction between the hemiacetal and the Lewis acid from the axial position, because of the stereoelectronic effect, affording β-*C*-glycoside. This two-step procedure established by Kishi is one of the most widely used reactions for β-*C*-hexopyranoside (*Table 2*) [19].

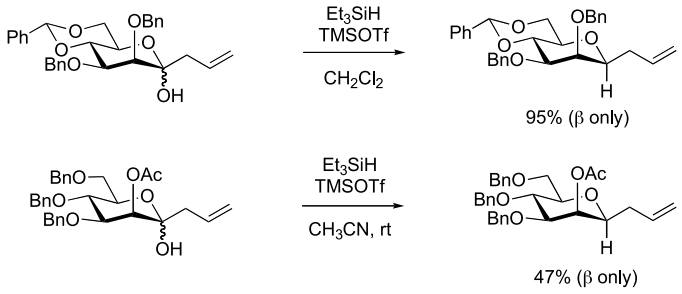


(2)

As shown in *Table 2*, this procedure enables the synthesis of a variety of β-*C*-glucosides and β-*C*-galactosides with substituents such as alkyl, allyl, vinyl, alkynyl, propargyl, aromatic, heteroaromatic [20], and acetyl groups.

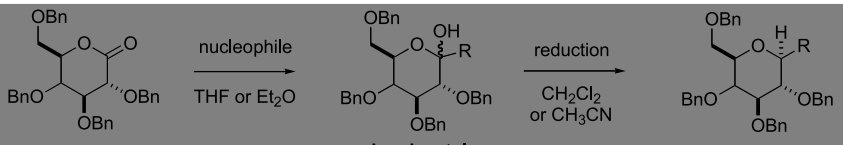
In contrast, the reduction of mannose-derived hemiacetals under similar conditions gave lower selectivity (*Table 2*) [19]. In order to improve the stereoselectivity, several methods have been proposed for special substrates; use of sterically hindered reducing agents such as TIPS-H (i-Pr_3SiH) or $(TMS)_3SiH$ (entry 21 in *Table 2*), restricted conformation of pyranose with 4,6-benzylidene acetal [21], and neighboring group participation of an acetyl group at the C-2 position (*Scheme 11*) [22].

The same procedure was applied to *C*-glycosylations of furanose-derived lactones (*Table 3*) [23]. The lactone reacted with organolithium reagents such as alkyl, alkynyl and heteroaromatic lithium to give the corresponding hemiacetals, which were reduced with triethylsilane in the presence of a Lewis acid. Since the reaction proceeds through an oxocarbenium cation intermediate, the hydride attack determines the stereochemistry of the product. However, rational explanation of these selectivities is difficult at present, because the selectivity is influenced by many factors such as steric and stereoelectronic effects together with the kind of sugar used in the reaction, the substituent, and the protective group.



Scheme 11

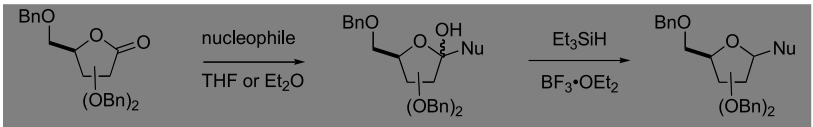
Table 2
[19]

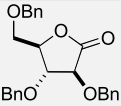
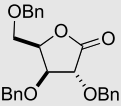
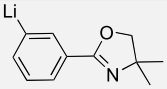
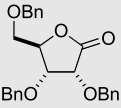
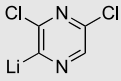


Entry	Sugar	Nucleophile	Hemiacetal	Reduction	Product	Ref.
				Silane, Lewis acid	Yield (α/β)	
1	Glucose	MeMgBr	90%	Et_3SiH, TMSOTf	81% (0/100)	[19a,b]
2			ND	Et_3SiH, $BF_3{\cdot}OEt_2$	85% (10/90)[a]	[19c]
3			82%	Et_3SiH, TMSOTf	79% (10/90)	[19a,b]
4			ND	Et_3SiH, $BF_3{\cdot}OEt_2$	49% (0/100)[a]	[19d]
5			ND	Et_3SiH, $BF_3{\cdot}OEt_2$	74% (0/100)	[19e]
6			Quant	Et_3SiH, $BF_3{\cdot}OEt_2$	72% (0/100)	[19f]
7		PhLi	85%	Et_3SiH, $BF_3{\cdot}OEt_2$	80% (1/4)	[19g,h]
8			ND	Et_3SiH, $BF_3{\cdot}OEt_2$	77% (0/100)[a]	[19g]
9			ND	Et_3SiH, $BF_3{\cdot}OEt_2$	72% (0/100)[a]	[19c]
10			74%	Et_3SiH, $BF_3{\cdot}OEt_2$	99% (0/100)	[19i]
11	Galactose		ND	Et_3SiH, $BF_3{\cdot}OEt_2$	81% (0/100)[a]	[19j]
12			ND	Et_3SiH, $BF_3{\cdot}OEt_2$	76% (10/90)[a]	[19c]
13			ND	Et_3SiH, $BF_3{\cdot}OEt_2$	50% (0/100)[a]	[19d]
14			91%	Et_3SiH, $BF_3{\cdot}OEt_2$	71% (0/100)[b]	[19k]
15			75%[c]	Et_3SiH, TMSOTf	96% (0/100)	[19l]
16			ND	Et_3SiH, $BF_3{\cdot}OEt_2$	79% (0/100)[a]	[19c]
17	Mannose	MeMgBr	86%	Et_3SiH, TMSOTf	68% (22/78)	[19a,b]
18			ND	Et_3SiH, $BF_3{\cdot}OEt_2$	67% (1/1)[a]	[19c]
19			ND	Et_3SiH, TMSOTf	61% (33/67)[a]	[19a,b]
20			93%	Et_3SiH, $BF_3{\cdot}OEt_2$	77% (1/2.5)[b]	[19k]
21			–	$(TMS)_3SiH$, $BF_3{\cdot}OEt_2$	76% (0/100)[b]	[19m]
22		PhLi	85%	Et_3SiH, TMSOTf	53% (0/100)	[19b]
23			78%[c]	Et_3SiH, TMSOTf	97% (0/100)	[19l]
24			ND	Et_3SiH, $BF_3{\cdot}OEt_2$	66% (3/1)[a]	[19c]

[a]: 2-step yield; [b]: after removal of TMS with TBAF; [c]: product was the corresponding acetate

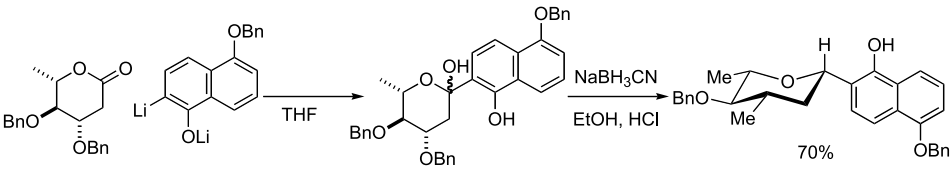
Table 3
[23]



Entry	Furanolactone	Nucleophiles	Hemiacetal	Product Yield (α/β)	Ref.
1		MeLi	ND	53% (100/0)*	[23a]
2			ND	81% (5/1)*	[23a]
3		$LiCH_2PO(OMe)_2$	93%	95% (100/0)	[23b]
4		MeLi	81%	76% (1/5)	[23a]
5			ND	87% (0/100)*	[23c]
6			ND	72% (0/100)*	[23c]

*: yield over 2 steps

Synthesis of aryl-*C*-glycoside has been reported by the same methodologies. One recent example is shown in ● *Scheme 12*. 2-Deoxysugar reacts with napthtyl lithium to give the corresponding hemiacetal, which was reduced with $NaBH_3CN$ in an acidic aqueous medium to afford α-*C*-glycoside exclusively [24].



Scheme 12

2.3 Glycosides (Halides, Oxygen, and others) as Electrophilic Carbohydrates

C-Glycosylation reactions of glycosyl-halide, -ester and -ether with carbon nucleophiles are divided into two types: the S_N2 type of substitution with carbon nucleophiles such as Grignard reagents, and addition of carbon nucleophiles to oxocarbenium cations generated from

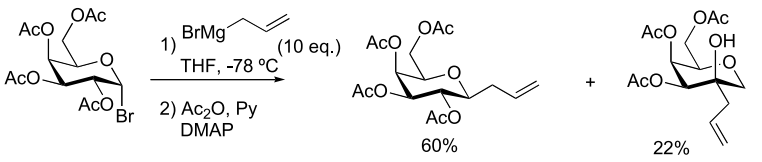
the glycoside upon reaction with Lewis acid. In the latter case, the stereochemistry of the addition is controlled by steric hindrance and/or stereoelectronic effects. This section deals with these types of reactions classified by nucleophiles such as (i) Grignard reagents and organoaluminum reagents, (ii) organosilane and tin reagents, (iii) aromatic nucleophiles (for Friedel–Crafts reaction).

2.3.1 Reaction with Grignard Reagents and Organoaluminum Reagents

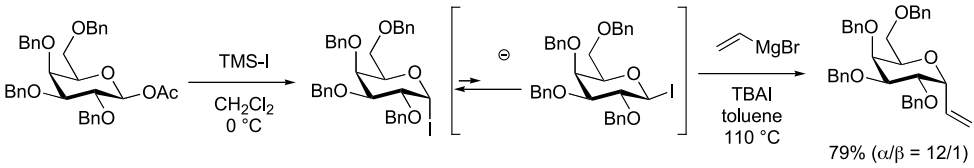
Addition of a Grignard reagent to glycosyl halide has been employed as one of the classic methods for *C*-glycoside synthesis. However, the use of this method is limited to the reaction in which the desired product is easily separable. ❯ *Scheme 13* shows one such example [25]; reaction of a commercially available peracetyl α-galactosyl bromide (or chloride) with a large excess of allyl magnesium bromide (ca. 10 equiv.) gave a mixture of the products depicted in ❯ *Scheme 13*. Since the acetyl groups were deprotected under the reaction conditions, the crude products were reacetylated and then separated by column chromatography. The major *C*-glycoside product was found to have β-configuration indicating preferential inversion of configuration at the anomeric position.

Despite instability of glycosyl halides having benzyl protective groups, the preparation of α-galactosyl iodide was recently reported from the corresponding acetate with TMS-I at a low temperature (❯ *Scheme 14*) [26]. Reaction of galactosyl iodide with vinyl Grignard reagent in the presence of TBAI (n-Bu_4NI) in toluene at 110 °C gave α-vinyl-*C*-galactoside, stereoselectively. Under these conditions, vinyl magnesium bromide preferentially reacts with the more reactive β-iodo galactoside equilibrated with the corresponding α-iodide through in situ anomerization, to afford α-*C*-galactoside in an S_N2 manner. It is noted that this isomerization under thermal conditions would be much faster than *C*-glycosylation. The vinyl product was further transformed into a *C*-analogue of the antitumor glycolipid, KRN7000 through olefin cross metathesis.

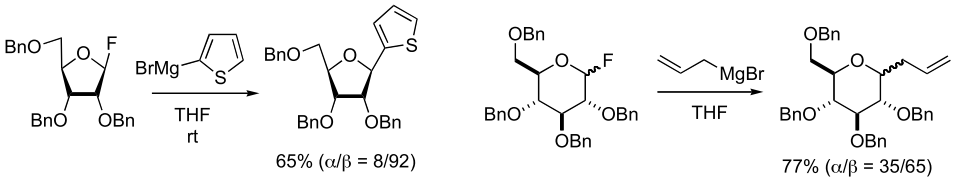
Glycosyl fluoride has been widely used as a stable donor compared with the corresponding glycosyl bromide and chloride in *O*-glycosylation. Two examples of reactions between glycosyl fluoride and Grignard reagents are shown in ❯ *Scheme 15*, indicating that the reactions take place through an oxocarbenium cation intermediate [27].



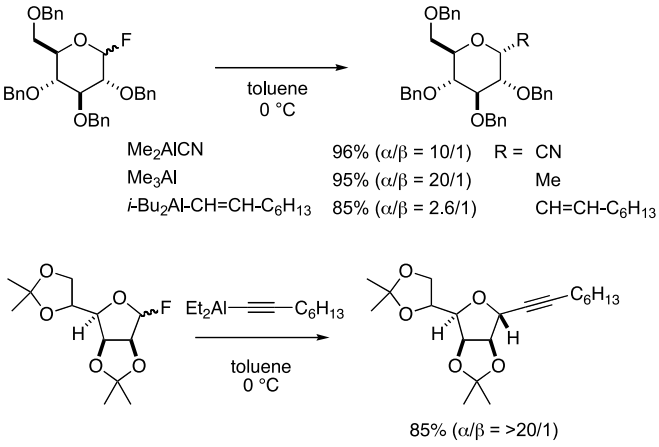
Scheme 13



Scheme 14



Scheme 15

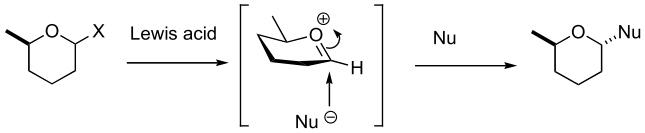


Scheme 16

Glycosyl fluoride reacts with a variety of organoaluminum reagents having alkyl, alkenyl, alkynyl, and aryl groups without the use of a promoter, this is due to high affinity between aluminum and fluoride. The reactions shown in *Scheme 16* indicate that the reaction proceeded via an oxocarbenium cation, thereby giving α-*C*-glycosides as the major product from hexopyranose, and α-*C*-glycoside was obtained as the major product from the corresponding furanose [28].

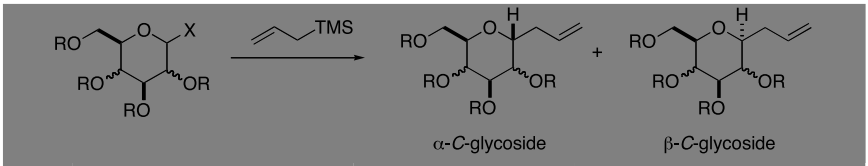
2.3.2 Reaction with Organosilane and Organotin Reagents

Reaction of glycosyl donors with organosilane or organotin reagents as nucleophiles in the presence of a Lewis acid has been widely employed as one of the most important *C*-glycosidation methods. The reaction takes place through addition of the nucleophiles to an oxocarbenium cation generated from the glycoside by using a Lewis acid (*Eq. 3*). In the case of the hexopyranose derivative, attack of the nucleophile occurs preferentially from the axial direction, due to stereoelectronic effects, leading to α-*C*-glycoside in a highly stereoselective manner.



(3)

Table 4
[29]

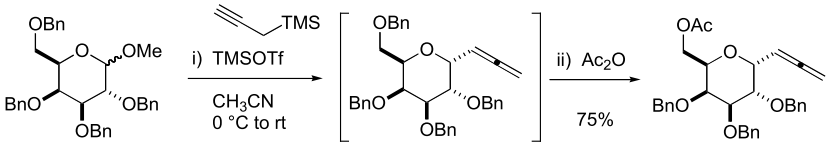


Entry	Carbohydrate			Conditions	Products		Ref.
		R	X (α/β)	Solvent, Lewis acid	Yield (%), α/β		
1	Glc	Bn	*p*-NBz (α)	CH_3CN, $BF_3{\cdot}OEt_2$	55	10/1	[19c]
2		Bn	OMe (α)	CH_3CN, TMSOTf	86	91/9	[29a]
3		Bn	F (1/1)	CH_2Cl_2, $BF_3{\cdot}OEt_2$	95	>20/1	[29b]
4		Bn	Ac (mix)	CH_3CN, $BF_3{\cdot}OEt_2$	89	α	[29e]
5		Ac	Ac (β)	CH_3CN, $BF_3{\cdot}OEt_2$	81	95/5	[29c]
6	Gal	Bn	*p*-NBz (α)	CH_3CN, $BF_3{\cdot}OEt_2$	79	10/1	[19c]
7		Bn	OMe (α)	CH_3CN, TMSOTf	91	α	[29f]
8		Ac	Ac (β)	CH_3CN, $BF_3{\cdot}OEt_2$	80	95/5	[29c]
9	Man	Bn	*p*-NBz (α)	CH_3CN, $BF_3{\cdot}OEt_2$	79	>10/1	[19c]
10		Bn	OMe (α)	CH_3CN, TMSOTf	87	α	[29a]
11		Bn	$PO(OBn)_2$ (α)	CH_2Cl_2, TMSOTf	93	α	[29d]
12		Ac	Ac (β)	CH_3CN, $BF_3{\cdot}OEt_2$	68	4/1	[29c]

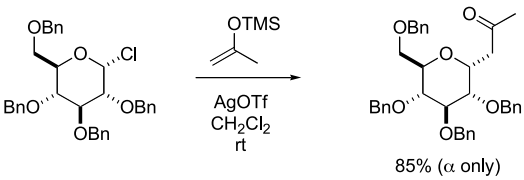
Since Kishi reported allylation of glycosyl *p*-nitrobenzoate with allyltrimethylsilane in the presence of a Lewis acid (entry 1 in *Table 4*) [19c], a variety of leaving groups such as halogen, ester, ether, imidate, and phosphate have been employed. Typical examples are shown in *Table 4*, indicating the generality of this method for the synthesis of α-allyl-*C*-glycosides [29].

Under similar conditions, allenyl and ketone could be installed by reaction with propargyl silane and silyl enol ether. Methyl galactoside was treated with propargylsilane in the presence of TMSOTf to give α-allenyl-*C*-glycoside, which was directly treated with acetic anhydride to afford the corresponding acetate at the 6-position (*Scheme 17*) [30]. Glucosyl chloride reacted with silyl enol ether in the presence of silver salt to yield α-glycosyl ketone in a highly stereoselective manner (*Scheme 18*) [31].

The acetylene unit is also introduced by the reaction with silyl or tin acetylene in the presence of a Lewis acid. Tin acetylene was used in order to introduce acetylene into glucose, galactose, and mannose (*Table 5*) [32], while *C*-glycosylation of deoxyhexopyranose was carried out with silylacetylene (*Table 6*) [33]. These examples indicate an easy and reliable route for the preparation of α-glycosylacetylene (α-sugar acetylene).

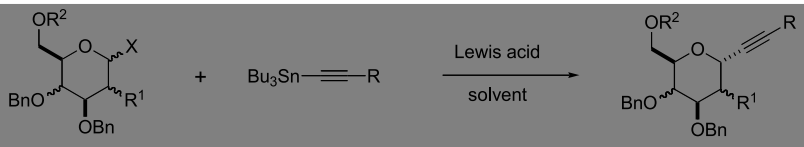


Scheme 17



Scheme 18

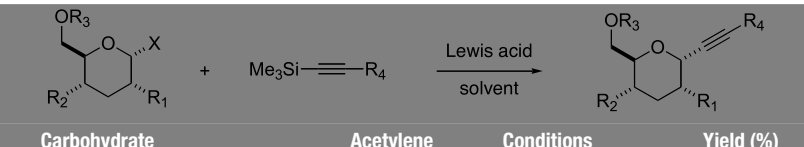
Table 5
[32]



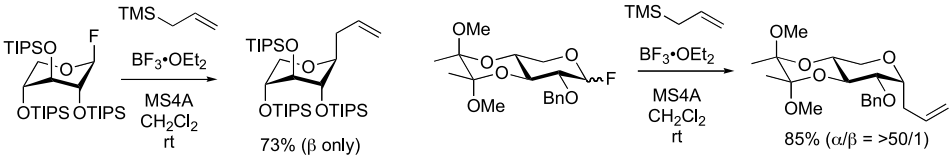
	Carbohydrate			Acetylene	Conditions	Products		Ref.
	R^1	R^2	X	R	Lewis acid, solvent	R	Yield (%)	
Glc	OBn	Bn	Br	Ph	$ZnCl_2$, CCl_4	Ph	61	[32a]
	OBn	Ac	Cl	Ph	$AgBF_4$, $(CH_2Cl)_2$	Ph	73	[32b]
	OBn	Bn	OAc	TMS	TMSOTf, CH_2Cl_2	H	71[a]	[32c]
	N_3	Bn	OAc	TMS	TMSOTf, CH_2Cl_2	TMS	43	[32d]
	N_3	Ac	Br	C_6H_{13}	$AgBF_4$, $(CH_2Cl)_2$	C_6H_{13}	76	[32b]
Man	OBn	Ac	OAc	TMS	TMSOTf, CH_2Cl_2	H	83[a]	[32c]
	OBn	Bn	OAc	TMS	TMSOTf, CH_2Cl_2	H	65[a]	[32c]
	N_3	Bn	OAc	TMS	TMSOTf, CH_2Cl_2	TMS	35	[32d]
Gal	OBn	Bn	OAc	TMS	TMSOTf, CH_2Cl_2	H	54[a]	[32c]
	N_3	Bn	OAc	TMS	TMSOTf, CH_2Cl_2	TMS	63	[32d]

[a]: The yields after desilylation with TBAF in THF or K_2CO_3 in MeOH

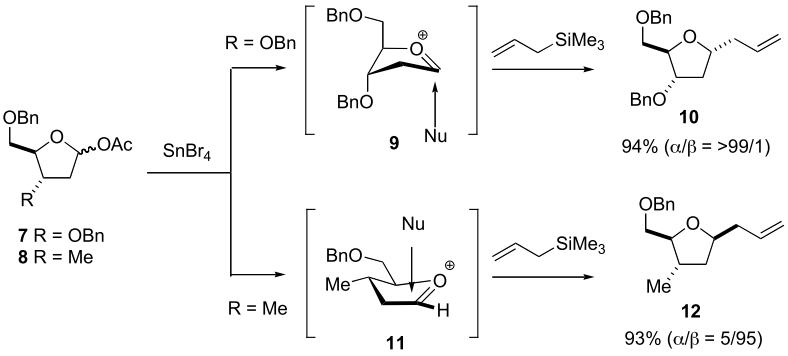
Table 6
[33]



Entry	Carbohydrate				Acetylene	Conditions	Yield (%)	Ref.
	R_1	R_2	R_3	X	R_4	Lewis acid, solvent		
1	OAc	OAc	Ac	OAc	TMS	$SnCl_4$, CH_2Cl_2	73	[33a]
2	H	OAc	Ac	OAc	SPh	$BF_3{\cdot}OEt_2$, CH_3CN	72	[33b]
3	OAc	OAc	Ac	OAc	≡—TMS	$SnCl_4$, CH_2Cl_2	71	[33a]
4	OAc	OAc	Ac	OAc		$SnCl_4$, CH_2Cl_2	83	[33a]
5	Me	H	TBDPS	O*i*-Pr	SPh	$BF_3{\cdot}OEt_2$, CH_3CN	61	[33c]
6	H	Me	H	OEt	SPh	$BF_3{\cdot}OEt_2$, CH_3CN	62	[33d]



Scheme 19



Scheme 20

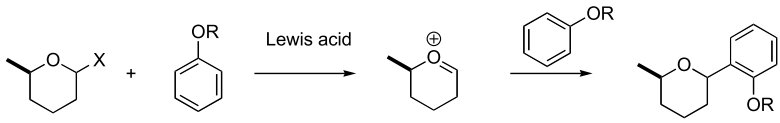
Since the stereoelectronic effect determines the stereochemical outcomes in the above-mentioned *C*-glycosidation of pyranose derivatives, application of the same method for the synthesis of β-*C*-glycosylation seems to be difficult. However, when the conformation of the substrate can be inverted, axial attack of a nucleophile from the β face would be predominant for stereoelectronic reasons, yielding β-*C*-glycoside. On the basis of the above outlined hypothesis, Shuto and co-workers demonstrated β-*C*-allylation of pentoxylose protected with TIPS (tri-isopropylsilyl) (*Scheme 19*) [34]. Installation of the sterically hindered protecting group fixed the inverted conformation ($^1C_4$) to avoid the gauche repulsion, thereby axial attack of allylsilane from the β face resulted in the β-*C*-glycoside, while allylation of 1-fluoroxylose protected with butane 2,3-diacetal (BDA) gave α-*C*-glycoside exclusively.

The stereochemical outcome of the similar *C*-allylation of pentofuranose does not depend on the substituent at the C-4 position of the carbohydrate, but rather depends on the type of substituent at the C-3 position as shown in *Scheme 20*. Woerpel explained the stereochemical outcome by stereoelectronically controlled attack of allylsilane to the lower energy envelope conformer as shown below [35]. The reaction of **7** gave α-*C*-products through the transition state shown for intermediate **9**, while the reaction of **8** afforded β-allyl product **12** through a different transition state **11**, in which the methyl substituent occupied the equatorial position.

2.3.3 Aryl-*C*-Glycosylation with Aromatic Nucleophiles

Stereoselective formation of the aryl-*C*-glycosidic linkage has been developed for synthesis of biologically important antibiotics and plant constituents. The most direct and widely used synthetic method is the Friedel–Crafts type reaction between electrophilic carbohydrate

and electron-rich aromatic compounds. Carbohydrates with an appropriate leaving group react with an aromatic derivative in the presence of various Lewis acids to afford aryl-*C*-glycoside through addition of aromatic compounds to an oxocarbenium ion generated by a Lewis acid (● *Eq. 4*). Generally, the thermodynamically stable product is preferentially obtained with an equatorial *C*-glycosidic bond.

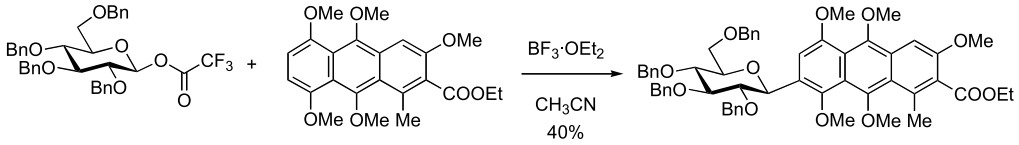


(4)

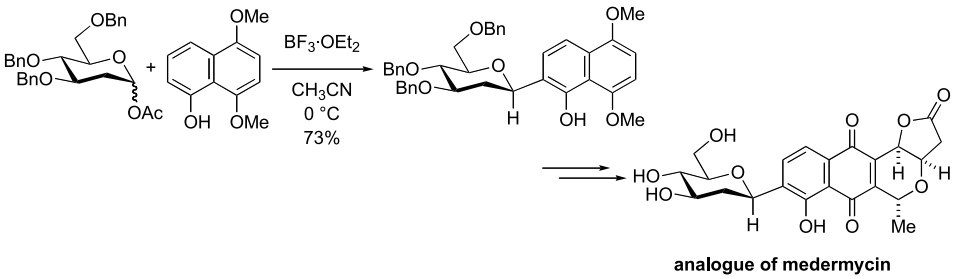
Typical examples are shown below. Combination of the leaving group with Lewis acid is critical in some special cases. Acetate, trifluoroacetate, and imidate were employed as leaving groups of the glycosyl donor, which were activated with a Lewis acid such as $BF_3{\cdot}OEt_2$, TMSOTf, and $SnCl_4$, etc. In all cases, products possessing an equatorial *C*-glycosidic linkage were exclusively obtained and transformed to naturally occurring aryl-*C*-glycosides such as carminic acid (● *Scheme 21*) [36], an analogue of medermycin (● *Scheme 22*) [37], and a flavone from oolong tea (● *Scheme 23*) [38]. Combination of AgOTFA ($AgOCOCF_3$) and $SnCl_4$ exhibited a superior reactivity for these types of reaction, and by such means it was possible to form α-*C*-arylation of *N*-acetyl nueraminic acid, which had not been reported previously (● *Scheme 24*) [39]. *C*-Glycosylation of a complete unprotected 2,6-deoxysugar with naphthalenediol has also been reported for the synthesis of urdamycinone B (● *Scheme 25*) [40]. Among the examples shown above, reaction with a phenol derivative may take place through *O*- to *C*-glycoside rearrangement (vide infra).

The general problems in the above-mentioned Friedel–Crafts arylations are the regioselectivity of the aromatic substitution and stereoselectivity of the aryl-*C*-glycoside. Since many naturally occurring aryl-*C*-glycoside compounds contain the *C*-glycosidic linkage at the phenol *ortho*-position, an *O*- to *C*-glycoside rearrangement strategy has been developed by Suzuki and co-workers in order to solve the problems (● *Scheme 26*) [41]. Lewis acid-promoted *O*-glycosylation of a phenol derivative is followed by rearrangement to *C*-glycoside affording aryl-*C*-glycoside. The mechanism of the rearrangement is not as simple as it seems. The reaction proceeds through an ion pair generated from *O*-glycoside, followed by re-combination, resulting in *ortho*-substituted *C*-glycoside. Under the Lewis acid conditions, anomerization via a quinone methide occurs to afford the thermodynamically stable product.

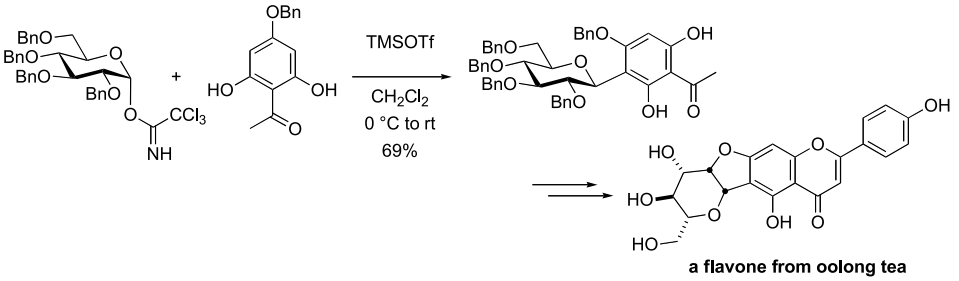
$Sc(OTf)_3$ was reported as an efficient catalyst for the *O*- to *C*-glycoside rearrangement (● *Scheme 27*) [42]. The conditions enabled an efficient coupling between glycosyl acetate and an electron-deficient phenol derivative that reduced the nucleophilicity of the donor [● *Scheme 27* (a)], while the same reaction with $BF_3{\cdot}OEt_2$ gave a 38% combined yield of the products with a poor selectivity (α/β = 2/3). Total synthesis of ravidomycin demonstrated the power of the *O*- to *C*-glycoside rearrangement methodology for the synthesis of structurally complex aryl-*C*-glycoside antibiotics [● *Scheme 27* (b)] [43]. *O*-Glycosylation of glycosyl fluoride with phenol was carried out upon treatment with Cp_2HfCl_2 and $AgClO_4$, which promoted the subsequent rearrangement to *C*-glycoside.


OBn
BnO
BnO
BnO
O
O
CF3
O
+
OMe OMe
OMe
COOEt
OMe OMe Me
BF3·OEt2
CH3CN
40%
OBn
BnO
BnO
BnO
O
OMe OMe
OMe
COOEt
OMe OMe Me

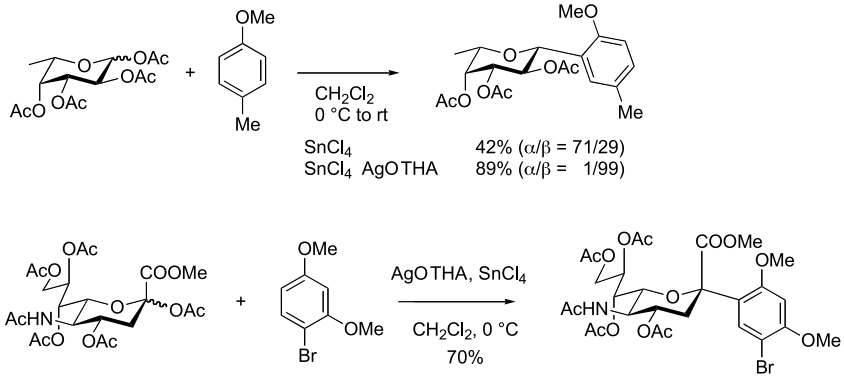
Scheme 21



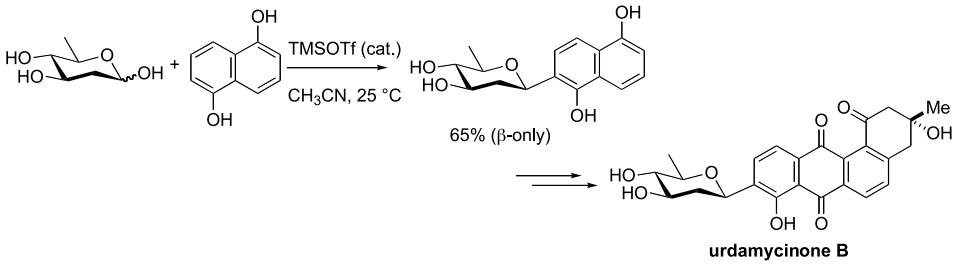
Scheme 22



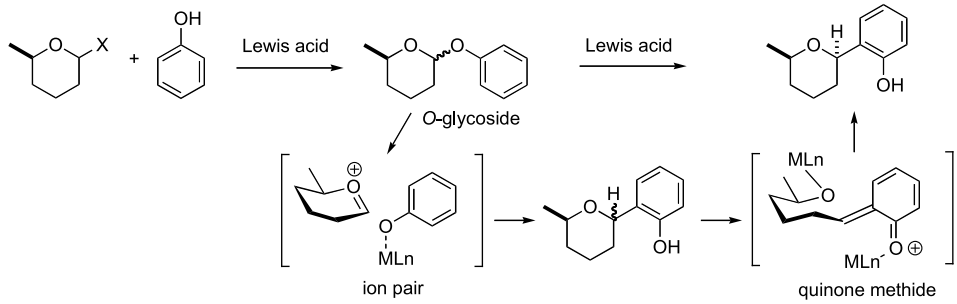
Scheme 23



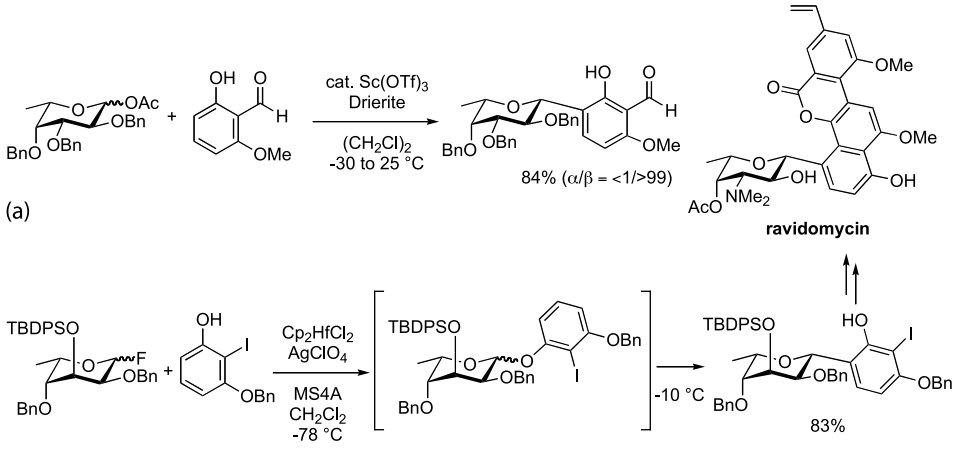
Scheme 24



Scheme 25



Scheme 26



Scheme 27

2.4 Glycal as an Electrophilic Carbohydrate

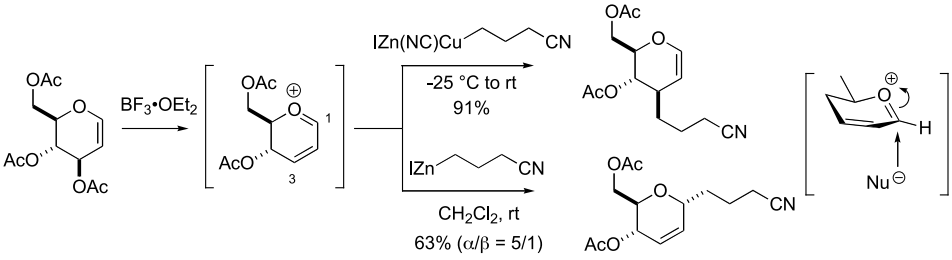
Glycals, easily available electrophilic carbohydrates with great diversity of functionality, react with a variety of carbon nucleophiles in the presence of Lewis acid to give unsaturated *C*-glycosides. The reaction seemingly proceeds in S_N2' manner with migration of the double bond (so-called Ferrier rearrangement). However, in the reaction mechanism the Lewis acid activates the acyl group at the C-3 position to generate an oxocarbenium cation, which is attacked by a carbon nucleophile (*Scheme 28*). The cation is an ambident electrophile that will react at the C-1 or C-3 position with nucleophiles. Most nucleophiles such as organozinc reagent, silane, and stannane react at the C-1 position to give *C*-glycoside [44], while organocopper reagents react at the C-3 position to afford a carbon-branched carbohydrate [45]. This section deals with the reactions with organozinc reagents, silanes, and stannanes in the mentioned order.

The stereochemical outcome is determined by the nucleophilic attack of a carbon nucleophile to an oxocarbenium cation; in the case of hexopyranose as tri-*O*-acetyl-D-glucal, axial attack of the nucleophile takes place at the anomeric position due to stereoelectronic effects, giving α-*C*-glycoside as the major product.

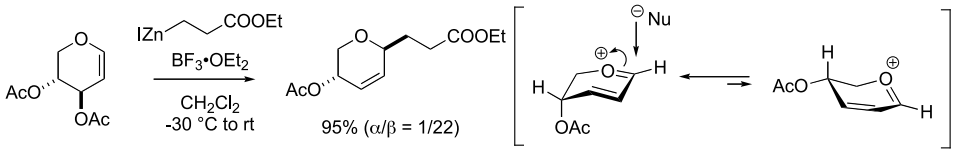
On the other hand, in the case of pentopyranose, the 1,4-*anti*-adduct was preferentially obtained (*Scheme 29*), due to the preferred conformation of the intermediate oxonium cation derived from the glycal [46,51a].

Aryl- and alkylzinc reagents having functional groups reacted with glycals without the addition of Lewis acid to give the corresponding *C*-glycosides in high α-stereoselectivity (*Scheme 30*) [47].

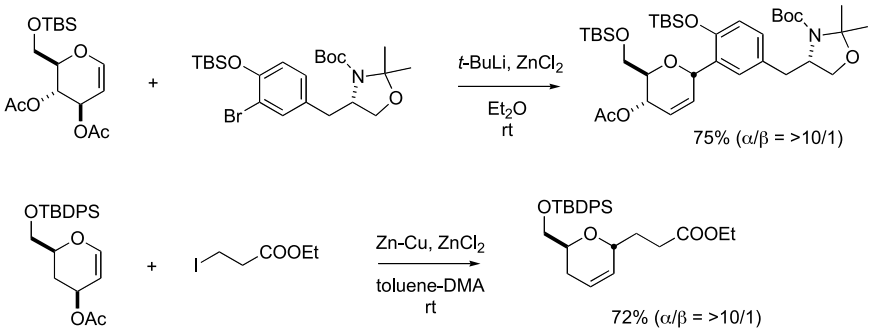
All of the above-mentioned reactions using glycals derived from hexopyranoses gave stereoelectronically controlled α-*C*-glycoside products. On the other hand, it is difficult to synthesize the corresponding β-*C*-glycoside by using the same method. Recently, reaction of



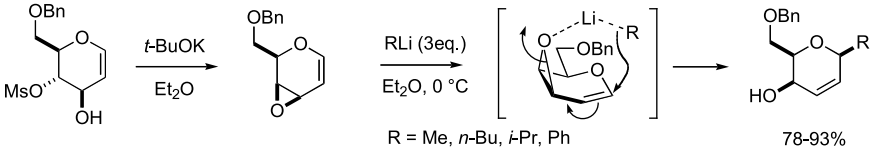
Scheme 28



Scheme 29



Scheme 30



Scheme 31

3,4-β-epoxyglycal with alkyl lithium was reported to give β-*C*-glycoside (*Scheme 31*) [48]. This reaction proceeded intramolecularly through chelation between the epoxide oxygen and the lithium reagent. The same mechanism also worked in the reaction of the corresponding 3,4-α-epoxyglycal with alkyl lithium, giving α-*C*-glycoside, exclusively.

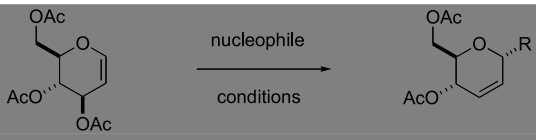
The most widely employed nucleophiles in the *C*-glycosylation of glycals are silicon-based reagents, such as allylsilane, silylacetylene [49], silylenol ether and its related reagents. *Table 7* lists typical examples of such nucleophiles, indicating the versatility of the reaction [50]. Reactions between glycal and these nucleophiles take place in the presence of conventional Lewis acids to give the corresponding *C*-glycosidic products generally with very high stereoselectivities. Recently, many other Lewis acids including $Sc(OTf)_3$, $Yb(OTf)_3$, InX_3 (X = Cl, Br), and I_2 were reported as catalysts for the reaction. The observed stereochemistries of the products are rational in light of the stereoelectronic effect. The reactions with TMSCN and silyl enol ether exhibited lower selectivities.

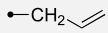
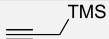
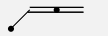
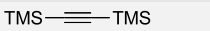
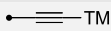
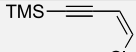
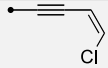
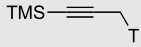
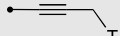
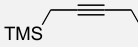
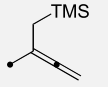
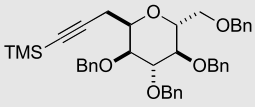
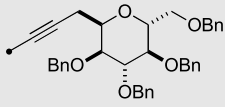
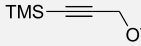
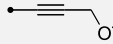
The methodology has been extended to glycal derivatives of pentopyranose. *Table 8* [51] and *Table 9* [51a] show typical examples of *C*-glycosylation of arabinal and xylal, indicating the reliability of the reaction.

These reactions have been extensively employed for the synthesis of complex natural products due to easy installation of various substituents with high stereoselectivities. In particular, the resulting sugar acetylenes are attractive for further transformation by means of Co-mediated reactions developed by Isobe and co-workers [33a].

The configuration of the acetylene group was inverted through an acetylene-cobalt complex prepared by treatment of the acetylene with $Co_2(CO)_8$ (*Scheme 32*) [52]. Upon treatment of the complex with acid, epimerization occurred via a propargylic cation intermediate stabilized by the cobalt complex to afford the thermodynamically more stable β-*C*-glycoside as the

Table 7
[50]



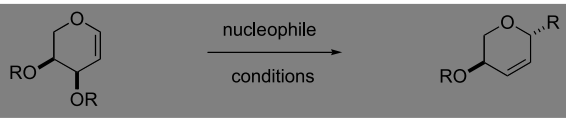
Entry	Nucleophile	Conditions	Products		Ref.
		Lewis acid, solvent	R	Yield (α/β)	
1		$TiCl_4$, CH_2Cl_2		85% (16/1)	[50a]
2		$Yb(OTf)_3$ (10 mol%), CH_2Cl_2		94% (1/0)	[50b]
3		$TiCl_4$, CH_2Cl_2		89%	[50c]
4	TMS-CN	I_2 (5 mol%), CH_2Cl_2		80% (6/4)	[50d]
5		$SnCl_4$, CH_2Cl_2		99%	[50e]
6		$SnCl_4$, CH_2Cl_2		96%	[50e]
7		$SnCl_4$, CH_2Cl_2		81%	[50e]
8		I_2, CH_2Cl_2		53%	[50f]
9		$BF_3{\cdot}OEt_2$, CH_2Cl_2		82%	[50g]
10		I_2, CH_2Cl_2		88%	[50f]
11		$SnCl_4$, CH_2Cl_2		84%	[50g]

major product. Decomplexation with iodine afforded β-sugar acetylene. The acetylene cobalt complex was transformed into a variety of structures such as *cis*-olefin and vinylsilane by treatment with Bu_3SnH and Et_3SiH, respectively [53].

An acetylene cobalt complex stabilizes the propargyl cation, which can be captured by a variety of nucleophiles (Nicholas reaction). By using the unique properties of these complexes, an efficient synthetic methodology for the preparation of medium-sized cyclic ether has been developed (*Scheme 33*) [54]. An acetylene-cobalt complex prepared from sugar acetylene **13** was transformed to the acyclic cobalt complex **14**, which was exposed to pivalic anhydride under acidic conditions to give **16**. Upon treatment with a Lewis acid, the seven-membered cyclic ether was efficiently obtained through a stabilized propargyl cation **15** by the Co-complex. Hydrogenation with Wilkinson catalyst gave the ABC ring model compound **17** of ciguatoxin, a highly toxic marine polyether natural product.

Sugar acetylene reacted with a Fisher chromium carbene complex to give a phenol derivative, a possible intermediate for naturally occurring aryl-*C*-glycoside antibiotics (*Scheme 34*) [55].

Table 8
[51]



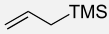
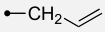
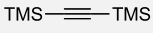
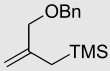
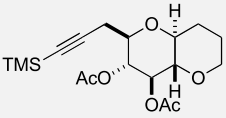
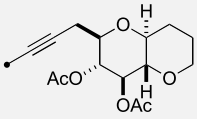
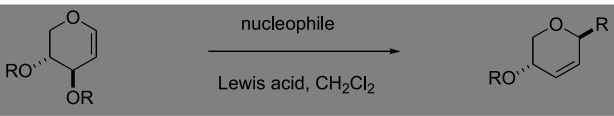
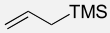
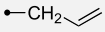
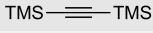
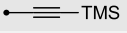
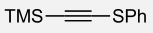
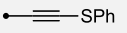
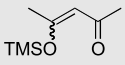
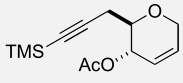
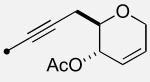
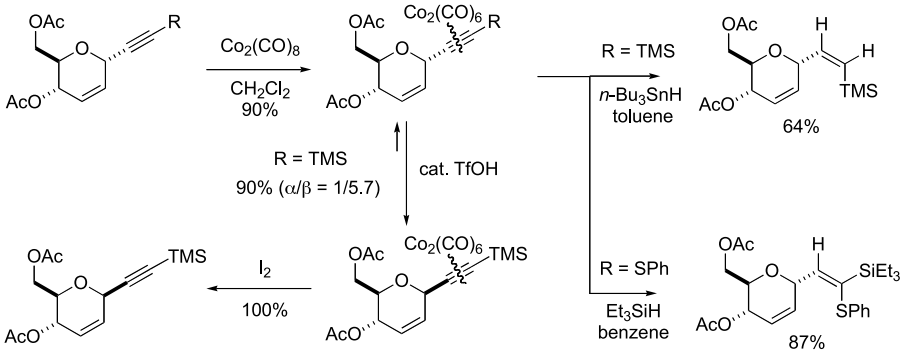
Entry	R	Nucleophile	Conditions	Products		Ref.
			Lewis acid, solvent	R	Yield (α/β)	
1	Ac		$BF_3{\cdot}OEt_2$, CH_2Cl_2		95% (95/5)	[51a]
2	Ac	TMS-CN	$InCl_3$(20mol%), CH_3CN		72% (10/1)	[51b]
3	Ac		$TiCl_4$, CH_2Cl_2		97% (95/5)	[51a]
4	Ac		$Yb(OTf)_3$(10mol%), CH_2Cl_2		85%	[51c]
5	Piv		$SnCl_4$, CH_2Cl_2		83% (95/5)	[51a]

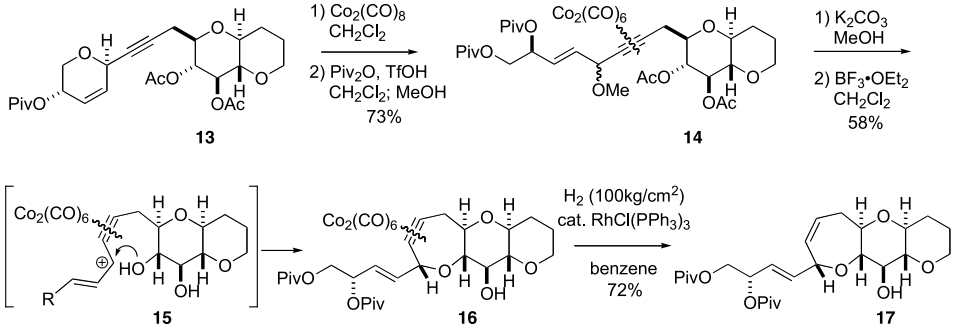
Table 9
[51a]



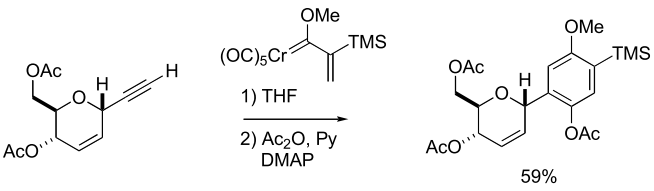
Entry	R	Nucleophile	Conditions	Products		Ref.
			Lewis acid	R	Yield (α/β)	
1	Ac		$BF_3{\cdot}OEt_2$		99% (5/95)	[51a]
2	Ac		$TiCl_4$		73% (5/95)	[51a]
3	Ac		$BF_3{\cdot}OEt_2$		88% (5/95)	[51a]
4	Ac		$BF_3{\cdot}OEt_2$		65% (15/85)	[51a]
5	Piv		$TiCl_4$		87% (5/95)	[51a]



Scheme 32



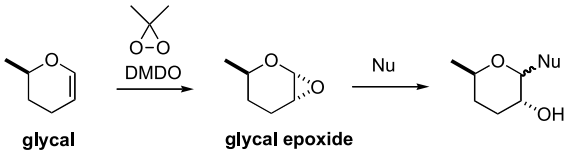
Scheme 33



Scheme 34

2.5 Anhydrosugar as an Electrophilic Carbohydrate

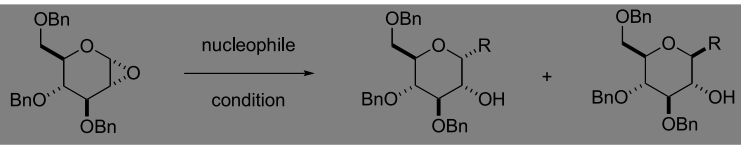
Among anhydrosugar derivatives, 1,2-anhydrosugar (so-called glycal epoxide) has been employed as the most useful donor for *C*-glycosylation. Since these unstable materials are easily prepared by epoxidation of glycal with DMDO (dimethyldioxorane) under very mild conditions, *C*-glycosylations of the glycal epoxide with various organometallic reagents have been extensively explored (*Eq. 5*).

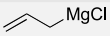
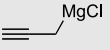
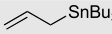
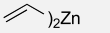
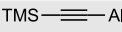
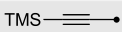
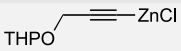
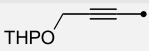
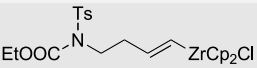
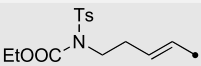


(5)

The stereoselectivity of the reaction with organometallic reagents is strongly dependent on the Lewis acidity as well as the nucleophilicity of the reagents. ❯ *Table 10* shows reactions of glucal epoxide with a variety of reagents [56]. When organocopper reagents such as dimethyl and diphenyl cuprates were employed as nucleophiles, opening of the epoxide occurred with inversion of configuration to give β-*C*-glycosides (entries 1 and 2). Reactions with allyl and propargyl Grignard reagents also gave the corresponding β-*C*-glycosides in high stereoselectivities (entries 3 and 4). In the case of vinyl Grignard reagents, reaction temperature affected the stereoselectivity; the reaction with vinyl Grignard reagents at 0 °C showed no selectivity, while the same reaction at −40 °C gave β-*C*-glycoside exclusively (entry 5). Allylation was also carried out by allylstannane with a catalytic amount of tributyltin triflate (entry 6). Sodi-

Table 10
[56]



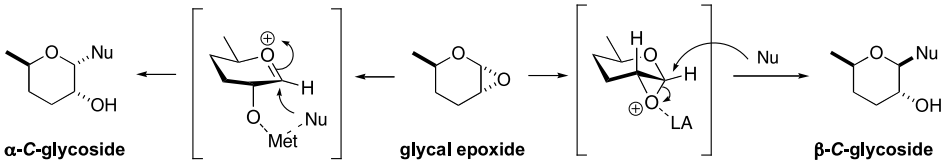
Entry	Nucleophile	Conditions		Products		Ref.
		Solvent	Temp.	R	Yield (α/β)	
1	Me_2CuLi	THF	0 °C	Me	82% (0/100)	[56a]
2	Ph_2CuLi	Et_2O	0 °C	Ph	84% (0/100)	[56a]
3		THF	0 °C		82% (0/100)	[56a]
4		CH_2Cl_2	0 °C		78% (0/100)	[56a]
5		CH_2Cl_2	−40 °C		57% (0/100)	[56a]
6		Bu_3SnOTf, CH_2Cl_2	−78 °C		57% (5/95)	[56c]
7	$NaCH(COOEt)_2$	$ZnCl_2$, THF	ND	$CH(COOEt)_2$	76% (0/100)	[56d]
8	$AlMe_3$	CH_2Cl_2	−95 °C	Me	82% (1/0)	[56a]
9	Ph_3Al	CH_2Cl_2	−65 °C to rt	Ph	79%	[56a]
10		CH_2Cl_2	−60 °C		70% (13/1)	[56b]
11		THF	0 °C		50% (1/0)	[56e]
12		THF	−95 °C		80% (1/0)	[56a]
13		THF	−50 °C to rt		72% (1/0)	[56e]
14		$AgClO_4$, CH_2Cl_2	rt		70% (1/0)	[56f]

um malonate ester as the nucleophile reacted with the glucal epoxide in the presence of $ZnCl_2$ to afford β-*C*-glycoside (entry 7). In this case, the thermodynamically stable β-*C*-glycoside was preferentially obtained after equilibrium through a retro-Michael reaction followed by conjugate addition.

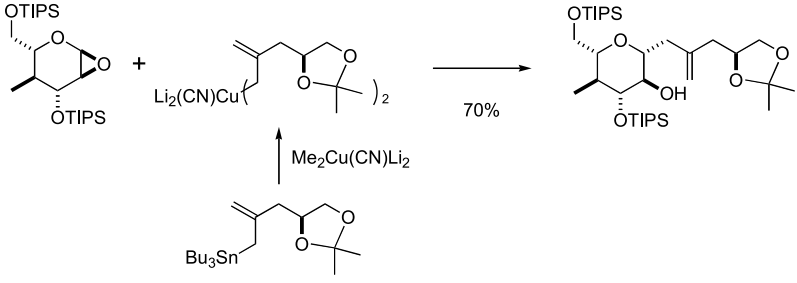
On the other hand, organoaluminum reagents such as trialkyl, triaryl, and trivinyl aluminum except allyl aluminum gave high α-stereoselectivities (entries 8 and 9). For the synthesis of α-allyl-*C*-glycoside, allyl borane has been shown to be the best reagent (entry 10). Vinyl zinc and vinyl-zirconium reagents prepared by hydrozirconation of alkyne served as good nucleophiles for the synthesis of alkenyl-α-*C*-glycoside (entries 11 and 14). Reaction with aluminum and zinc acetylides afforded α-*C*-glycosides in high stereoselectivities (entries 12 and 13).

The stereochemical outcomes were explained by the following mechanisms (*Scheme 35*); in the reactions with copper reagents, Grignard reagents, or a combination of allyl stannanes and Bu_3SnOTf, a mild Lewis acid, a S_N2-like reaction with the nucleophile took place at the anomeric position to afford the β-*C*-glycoside. On the other hand, aluminum reagents promoted formation of the oxocarbenium ion due to the strong Lewis acidity, followed by intramolecular transfer of the nucleophile to the cation through chelation giving the α-*C*-glycoside.

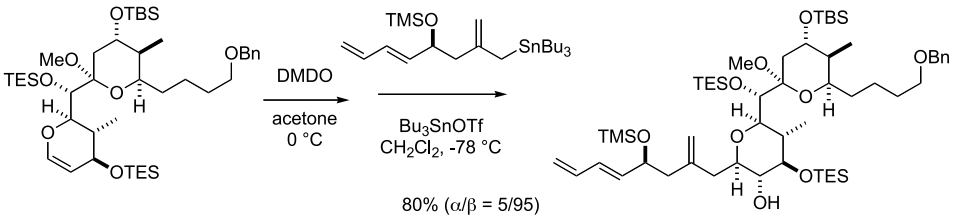
Since some of the above-mentioned *C*-glycosylations are mild enough and compatible with some functionalities, the reaction has been widely utilized in the synthesis of complex natural products. In the total synthesis of a complex marine macrolide, altohyrtin (spongistatin), Kishi and Evans independently employed the reaction between glycal epoxide and an organometallic reagent for the synthesis of the substituted tetrahydropyrane structure. An allyl cuprate prepared from allylstannane and cyanocuprate, reacted with the glycal epoxide to give 1,5-*syn*-pyranose (*Scheme 36*) [57]. Reaction of glycal epoxide with allylstannane in the presence of tributyltin triflate gave the 1,5-*syn*-pyranose product in good yield with very high stereoselectivity (*Scheme 37*) [58].



Scheme 35

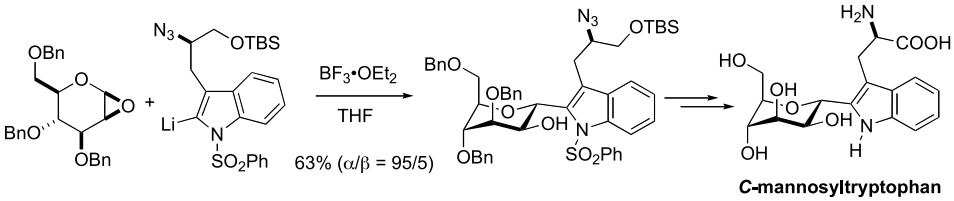


Scheme 36



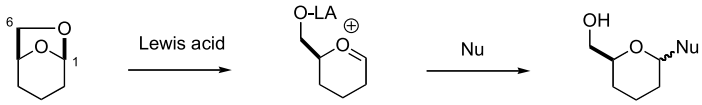
Scheme 37

Reaction of 1,2-anhydro-mannose with lithiated indole in the presence of $BF_3{\cdot}OEt_2$ as Lewis acid yielded α-*C*-mannosylindole, which was further transformed into *C*-mannosyltryptophan, a naturally occurring *C*-glycosyl amino acid found in some proteins (Scheme 38) [59].



Scheme 38

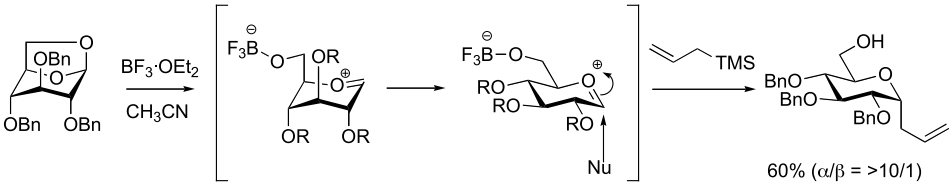
In contrast to the reaction of 1,2-anhydrosugar, *C*-glycosylation of 1,6-anhydropyranose requires a strong Lewis acid in order to open the anhydro bridge, generating the corresponding oxocarbenium cation (Eq. 6). Addition of nucleophiles to the cation could be controlled by the stereoelectronic effect, chelation between the nucleophile and the Lewis acid coordinated to the hydroxyl group, or intramolecular delivery of nucleophiles as shown below.



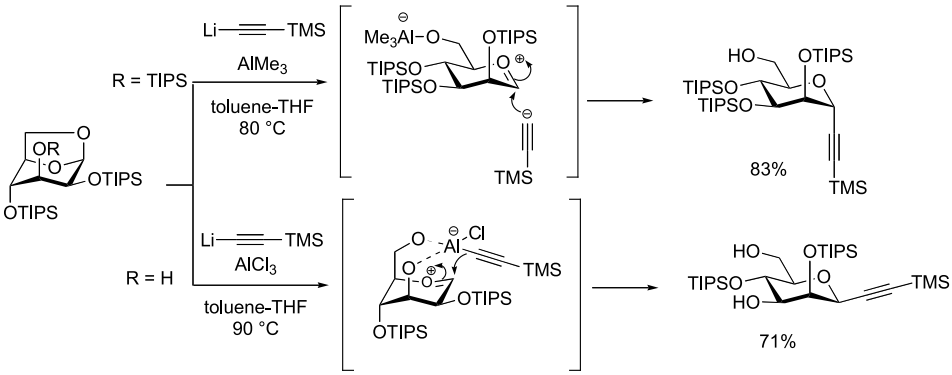
(6)

For example, reaction of 1,6-anhydroglucopyranose with allylsilane in the presence of $BF_3{\cdot}OEt_2$ gave α-*C*-glycoside in a high stereoselective manner (Scheme 39) [19c]. In this reaction, an initially formed oxocarbenium cation changed the conformation of the substrate to a more stable one ($^4C_1$), followed by attack of the allylsilane to the cation from the axial direction, affording α-*C*-glycoside.

1,6-Anhydromannopyranose (R = TIPS) reacted with lithium acetylide in the presence of Me_3Al to afford α-*C*-glycoside (Scheme 40) [60]. On the other hand, when a similar substrate (R = H) was treated with lithium acetylide in the presence of $AlCl_3$ instead of Me_3Al, β-*C*-glycoside was exclusively obtained [61]. The mechanism for the reaction was proposed as follows; the acetylide reacted with $AlCl_3$ to form the alkynylaluminum species, which opened the anhydro bridge by participating with the hydroxyl group at the C-3 position. Intramolecular transfer of the acetylene to the resulting oxocarbenium cation afforded β-*C*-glycoside stereoselectively.

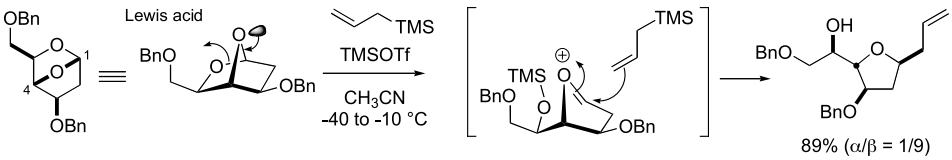


Scheme 39



Scheme 40

There are a few reports of *C*-glycosylation by using 1,4-anhydrosugar as a glycosyl donor. The reaction proceeds through an oxocarbenium cation as in the case for 1,6-anhydrosugar. Under similar conditions, reaction of 1,4-anhydrosuger with allylsilane did not give the allyl-*C*-pyranoside, instead allyl-*C*-furanoside was exclusively obtained (*Scheme 41*) [62]. The stereochemistry was explained by attack of allylsilane to the oxocarbenium cation from the less hindered face.



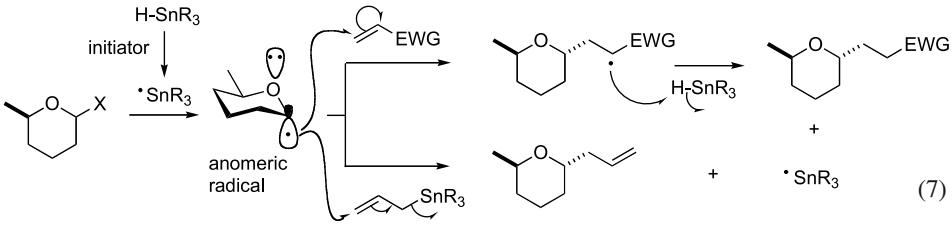
Scheme 41

3 Anomeric Radical Intermediate

Free radical reactions proceed under essentially neutral conditions that are compatible with various functional and protective groups used for saccharides. *C*-Glycosylation by means of radical reaction has been extensively developed [63]. This section will discuss intermolecular and intramolecular reactions for control of the stereochemistry of the *C*-glycoside linkage.

3.1 Intermolecular Radical Reaction

C-Glycosylation by means of anomeric radical species was first reported by Giese and coworkers in 1983 [64]. In general, an anomeric radical is readily generated from glycosyl halides (Cl or Br) or phenylseleno-glycoside by using tributyltin hydride and AIBN or irradiation as initiator. The resulting anomeric radical possesses a nucleophilic character and reacts with activated olefins such as acrylonitrile, acrylate ester, and allyltin reagents to afford chain extension products (*C*-glycosides). The general mechanism is outlined in *Eq. 7*. The tin radical generated by an initiator abstracts the X group to give an anomeric radical, which reacts with unsaturated olefin or allyltin reagent to form the *C*-glycosidic bond. Radical-mediated *C*-glycosylation of hexopyranose predominantly affords the α-*C*-glycosidic bond, independent of configuration at the anomeric position of the starting material. In the transition state, the anomeric radical is pyramidal (sp^3-like) and the axial radical preferentially reacts with olefine to afford α-*C*-glycoside [65].

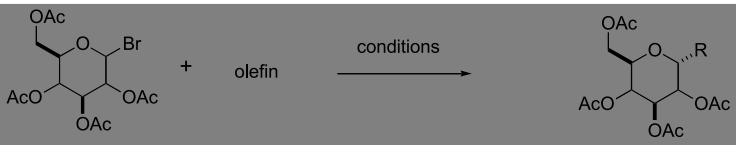


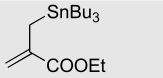
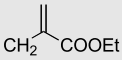
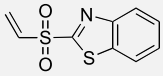
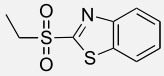
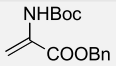
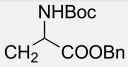
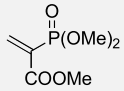
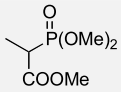
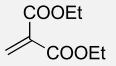
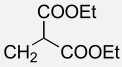
(7)

α-Glycosyl bromides have been widely studied as glycosyl donors for the radical *C*-glycosylation (*Table 11*) [66]. A variety of terminal olefins with electron-withdrawing groups can be installed into glucose, galactose, and mannose derivatives in high α-selectivities by using this methodology. It is also possible to conduct this chemistry without the use of a tin reagent (entries 2 and 9).

Radical reactions have also been employed for the synthesis of *C*-glycoside of 2-amino sugars as shown in *Table 12* [67] in which most of the reaction gave α-*C*-glycosides. However, when a sterically hindered protective group such as phtaloyl or tetrachlorophtaloyl was installed in order to protect the amino functionality at the C-2 position, the reverse stereoselectivity was observed (entry 2). In the reaction with an activated olefin such as acrylate ester or styrene, a milder reaction by using phenylseleno glycoside with Et_3B as the radical initiator was required (entry 5). The conventional conditions resulted in simple reduction of the glycosyl halide as a serious side reaction. Et_3B in the presence of traces of oxygen generates a radical species at room temperature. This method could be applied for *C*-glycosylation of

Table 11
[66]



Entry	Sugar	Olefin	Conditions	Products		Ref.
				R	Yield (α/β)	
1	Glc	CH_2=CH-CN	Bu_3SnH, benzene, *hv*	CH_2CH_2CN	75% (93/7)	[66a]
2		CH_2=CH-CN	vitamine B_{12}, Zn, NH_4Cl/DMF	CH_2CH_2CN	68% (13/1)	[66b]
3		CH_2=CH_2-CH_2-$SnBu_3$	AIBN, benzene, 80 °C	CH_2-CH=CH_2	64% (ND)	[66c]
4			AIBN, benzene, 80 °C		84% (ND)	[66a]
5			*n*-Bu_3SnH, AIBN Et_2O, reflux		66%	[66d]
6	Gal	CH_2=CH-CN	Bu_3SnH, Et_2O, *hv*	CH_2CH_2CN	70%	[66a]
7			Bu_3SnH, AIBN, toluene		61%	[66e]
8		CH_2=CH_2-CH_2-SPh	$(Bu_3Sn)_2Sn$, benzene, *hv*	CH_2CH=CH_2	73%	[66f]
9			$(TMS)_3SiH$ Et_2O, *hv*		80% (98/2)	[66g]
10	Man	CH_2=CH-CN	Bu_3SnH, Et_2O, *hv*	CH_2CH_2CN	65%	[66a]
11			Bu_3SnH, AIBN toluene, 67 °C		99%	[66a]

other 2-acetamide sugars such as GalNAc and ManNAc. The use of tristrimethylsilylsilane as an alternative to tin hydride has also been reported (entry 4).

The above-mentioned radical *C*-glycosylation has been applied for the synthesis of complex compounds. *Scheme 42* shows two such examples of the radical *C*-glycosylation of mannosyl bromide with complex olefins such as unsaturated ketone and unsaturated lactone. This highlights the mildness of the reaction and these conditions have compatibility with many functional groups [68]. It is noteworthy that an acceptor bearing seleno and chloro groups can be employed without interference under the radical conditions. In both reactions, the α-*C*-glycosidic products were exclusively obtained, although in moderate yields.

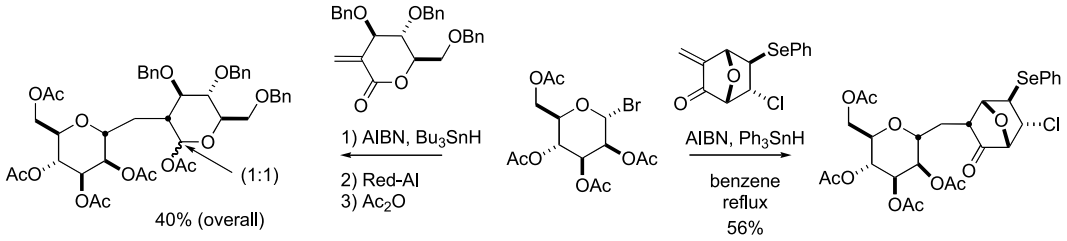
Since the conformation of hexopyranose is fixed by the presence of the equatorial hydroxymethyl group at the C-5 position, the corresponding radical intermediate possesses the axial radical in the α-direction, which reacted with the olefine to form α-*C*-glycoside. On the other hand, the conformation of pentopyranose is rather flexible, and easily inverted by several modifications. The anomeric radical adopting the $^1C_4$ conformation reacts with the olefin to give β-*C*-glycoside. Shuto and co-workers reported formation of β-*C*-glycosides by using

Table 12
[67]

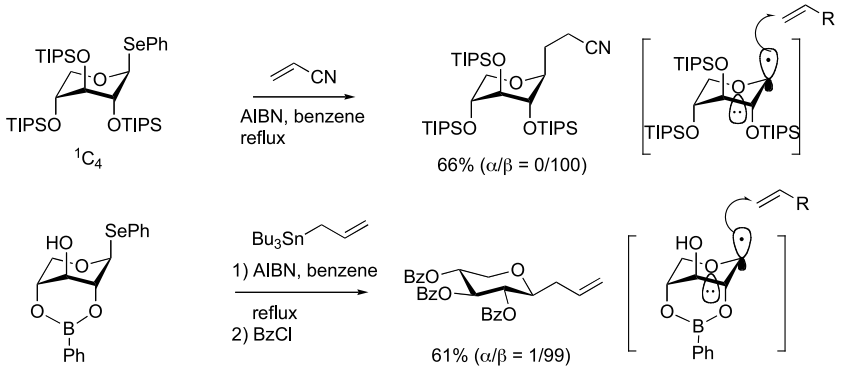
Entry	Carbohydrate			Olefine	Condition	Products		Ref.
		R_1	X			R	yield (α/β)	
1	GlcNAc	Ac	Cl	$SnBu_3$	AIBN, toluene	CH_2-CH=CH_2	73% (12/1)	[67a]
2		TCP	Br		AIBN, benzene[a]		77% (1/20)	[67b]
3		Ac	SePh	COO-*t*-Bu	*n*-Bu_3SnH, AIBN, benzene[a]	CH_2CH_2COO-*t*-Bu	76% (α)	[67c]
4		Ac	SePh		$(TMS)_3SiH$, AIBN, benzene[a]		93% (α)	[67c]
5		Ac	SePh		Et_3B, AIBN, benzene, rt		71% (α)	[67c]
6		Ac	SePh	Ph	*n*-Bu_3SnH, AIBN, benzene[a]	CH_2CH_2Ph	30% (α)	[67c]
7	GalNAc	Ac	Br	SO_2Ph	$(Bu_3Sn)_2$, hv, benzene	CH_2-CH=CH_2	42% (α)	[67a]
8		Ac	SePh	COO-*t*-Bu	*n*-Bu_3SnH, AIBN, benzene[a]	CH_2CH_2COO-*t*-Bu	68% (α)	[67c]
9		Ac	SePh		Et_3B, *n*-Bu_3SnH, toluene, *hv*		90% (α)	[67d]
10		Ac	SePh	Ph	*n*-Bu_3SnH, AIBN, benzene[a]	CH_2CH_2Ph	41% (α)	[67a]
11	ManNAc	Ac	SePh	COO-*t*-Bu	Et_3B, *n*-Bu_3SnH, toluene, *hv*	CH_2CH_2COO-*t*-Bu	64% (α)	[67c]
12		Ac	SePh	Ph	Et_3B, *n*-Bu_3SnH, toluene, *hv*	CH_2CH_2Ph	28% (α)	[67c]
13		Ac	SePh	$SnBu_3$, COOMe	Ph_3SnH, toluene	CH_2, COOMe	80% (α)[b]	[67e]

[a]: The reaction was carried out at a reflux. [b]: The yield of ketoester after ozonolysis of the product

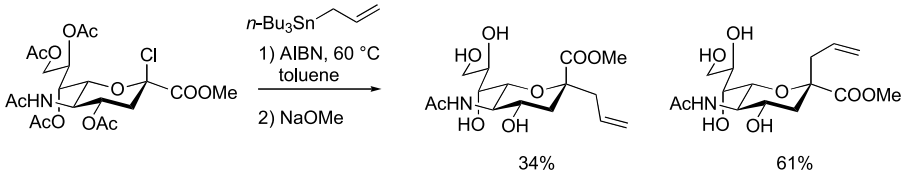
substrates with an inverting conformer (*Scheme 43*) [65]. When xylose was protected with a sterically hindered TIPS group, the conformation was fixed to be $^1C_4$, due to gauche repulsion between the TIPS groups. By utilizing a cyclic boronate, it is also possible to fix the $^1C_4$ conformation. Radical *C*-glycosylation of these substrates afforded β-*C*-glycoside in high stereoselectivities through the transition state as shown in the parentheses in *Scheme 43*.
The radical *C*-glycosylation of *N*-acetylneuraminic acid was independently reported by Paulsen and Bednarski (*Scheme 44*) [69]. The products were obtained as an anomeric mixture, which were easily separated after deprotection. The radical allylation is one of a few synthetic methods for formation of *C*-glycoside with neuraminc acid derivatives.


BnO
OBn
OBn
OAc
O
O
AcO
OAc
OAc
OAc
(1:1)
40% (overall)
BnO
OBn
OBn
O
O
1) AIBN, Bu3SnH
2) Red-Al
3) Ac2O
OAc
O
Br
AcO
OAc
OAc
SePh
O
O
Cl
AIBN, Ph3SnH
benzene
reflux
56%
OAc
SePh
O
Cl
AcO
OAc
OAc
O

Scheme 42



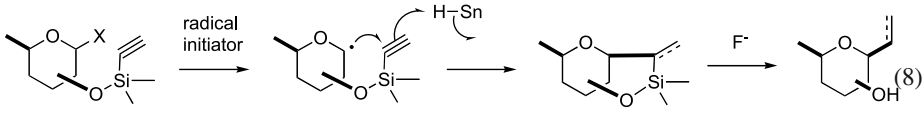
Scheme 43



Scheme 44

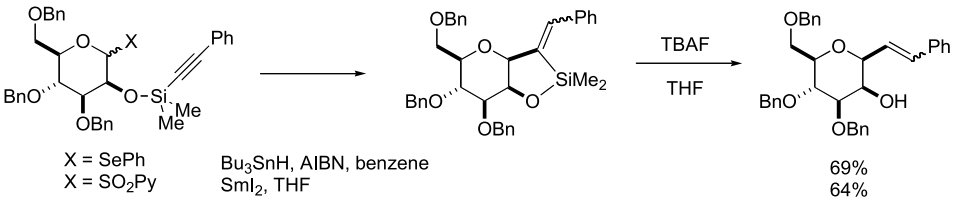
3.2 Intramolecular Radical Reaction

Intermolecular *C*-Glycosylation of hexopyranoses inherently provides α-*C*-glycoside in high stereoselectivity, while intramolecular radical *C*-glycosylation controls the stereochemistry at the anomeric position by the configuration of the hydroxyl group that is connected to the radical acceptor through a tether. Therefore, the intramolecular reaction is of significance despite the fact that several additional steps are necessary in order to synthesize the precursor and for removal of the tether after conducting the radical reactions. Phenylselenoglycoside and silicon tethers have been widely employed for this purpose [70]. After *C*-glycoside bond formation, the silicon tether was removed to give the product with the desired configuration (*Eq. 8*).

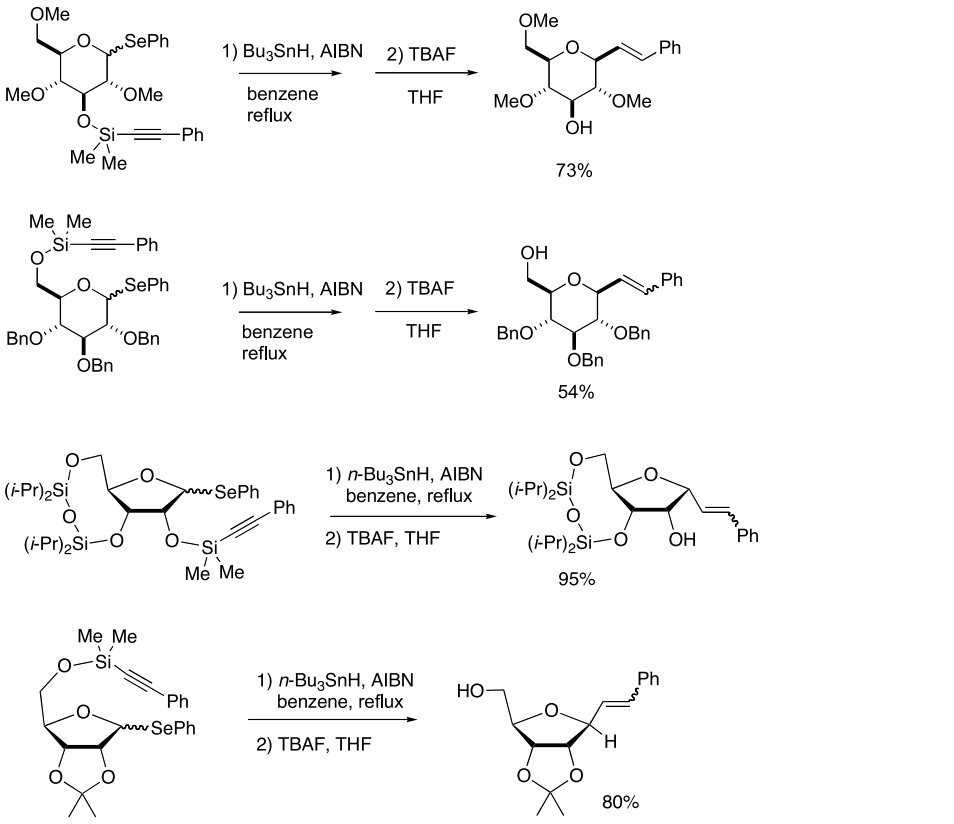


β-*C*-Mannoside is one of the most difficult to synthesize among *C*-glycosylations, however, intramolecular radical *C*-glycosylation of mannose by using a silicon tether at the C-2 position, afforded β-*C*-mannoside, exclusively (*Scheme 45*) [71]. Phenylsulfonyl- or pyridylsulfonyl groups can be employed as glycosyl donors for the same purpose by using SmI_2. This method avoids the use of toxic tin hydride and the phenylseleno group [72].

Hydroxyl groups at the C-3 and -6 position of glucoside were employed for the introduction of the silicon tether, leading to β-*C*-glycoside (*Scheme 46*). The resulting styryl group can be transferred to a variety of functional groups. The strategy was also applied to the synthesis of *C*-furanosides.

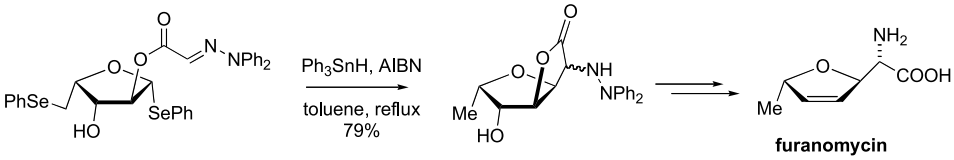


Scheme 45



Scheme 46

Carbohydrates with a hydrazono ester at the C-2 position underwent intramolecular radical *C*-glycosylation upon reaction with tin hydride and AIBN, to afford the corresponding lactone in good yield (*Scheme 47*). Although the stereoselectivities of the amino acid moiety were generally low, the methods have been used for the synthesis of a variety of *C*-glycosyl amino acids [73]. The following reaction was a key step in the total synthesis of furanomycin, an antibiotic.



Scheme 47

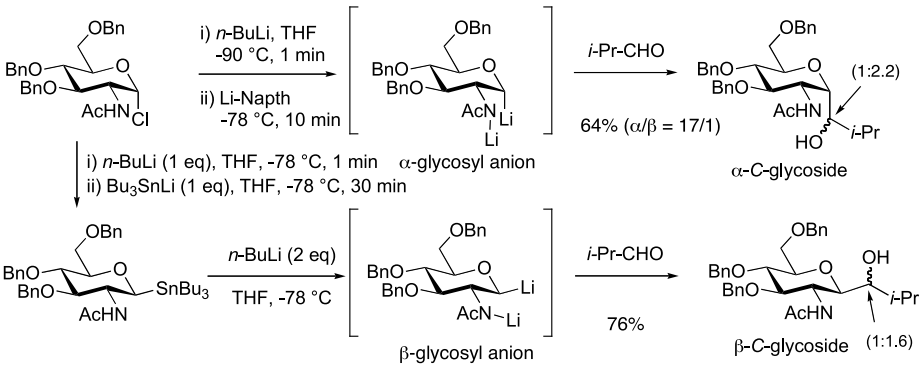
4 Anomeric Anion

In contrast to *C*-glycosylation by nucleophilic addition to an electrophilic carbohydrate, the use of an anomeric anion has found limited use for such a reaction due to instability. However, *C*-glycosylation by means of anomeric anion species is of significance in some special cases [74]. The anomeric anion intermediate can be stereoselectively prepared by (i) reductive metalation of anomeric halides or sulfones, (ii) transmetalation of glycosyl stannanes, and (iii) direct deprotonation of the anomeric proton. This section deals with three types of the reaction using different anomeric anions.

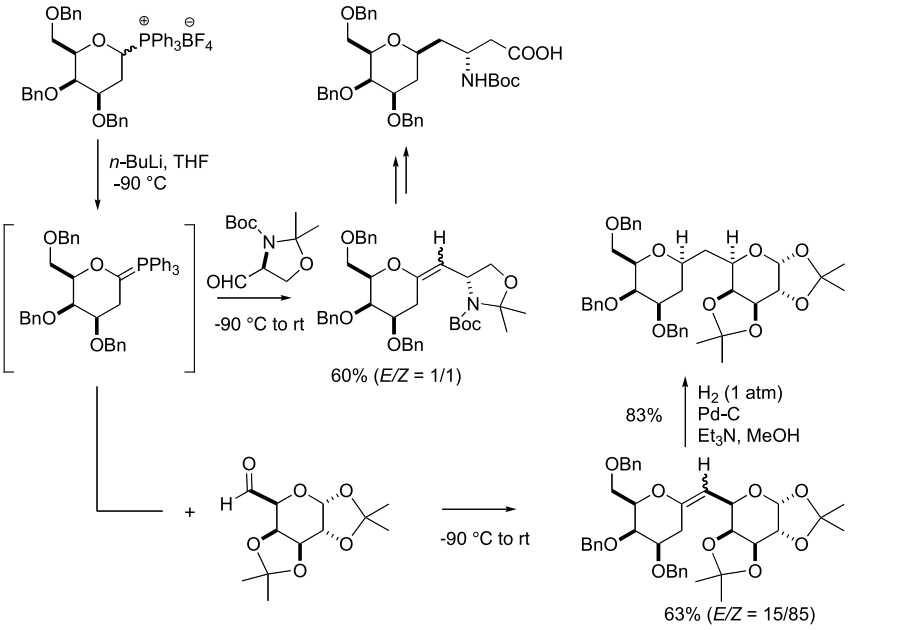
4.1 Lithium Anomeric Anion

Anomeric anion (sp^3) species without a stabilizing group can be employed in stereoselective *C*-glycosylation, because the anions are configurationally stable at low temperature, and react with a variety of electrophiles such as alkyl halides and compounds containing carbonyl groups. However, the method has not been well exploited probably due to strict anhydrous conditions that are required in order to conduct these reactions. A typical example of *N*-acetylglucosamine is shown in *Scheme 48* [75]. In this case, dianion chemistry was employed to prevent elimination of the substituent at the C-2 position; *N*-acetyl-glucosyl chloride was treated with *n*-butyl lithium to abstract the amide proton, followed by reduction with lithium naphthalenide at −78 °C to generate the α-anomeric anion with retention of configuration. On the other hand, the same glycosyl chloride was treated with 1 equivalent of *n*-butyl lithium followed by 1 equivalent of tributyltin lithium to give β-stannane with inversion of configuration. Upon treatment with 2 equivalents of *n*-butyl lithium, deprotonation and transmetalation took place to generate the β-anomeric anion with retention of configuration. These stereoselectively prepared anomeric anions reacted with aldehydes, CO_2, and alkyl iodide to afford the corresponding α- and β-*C*-glycosides, respectively. This method is applied to *C*-glycosylation of other carbohydrates such as 2-hydroxy and 2-deoxy sugars. It is worth noting that a carbohydrate with a protected hydroxyl group at the C-2 position undergoes β-elimination to give the corresponding glycal under these conditions.

C-Glycosylation by means of reaction with an anomeric anion containing stabilizing groups such as nitro, cyano, and sulfonyl groups has been extensively studied. The anomeric anion is usually prepared by proton abstraction with a base, and reacts with electrophiles to afford the *C*-glycoside product. However, due to the potential β-elimination to the corresponding glycal, the use of the stabilizing anomeric anion for *C*-glycosylation is limited. Since comprehensive



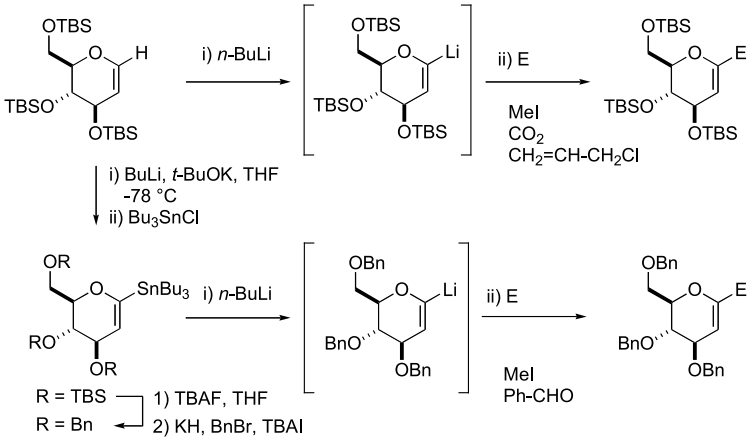
Scheme 48



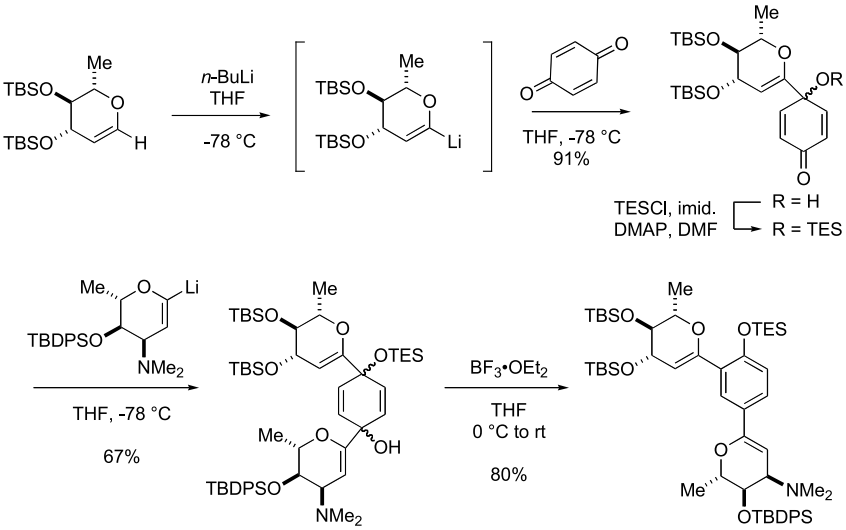
Scheme 49

reviews for this type of *C*-glycosylation have been published [74], this section will only deal with a recent example utilizing glycosyl ylide. The ylide prepared from a galactose derivative reacted with Garner aldehyde and sugar aldehyde to give the corresponding *exo*-olefinic products as a mixture of the geometric isomers (*Scheme 49*) [76]. Catalytic hydrogenation gave a single isomer of *β*-*C*-glycosyl amino acid and *C*-disaccharide.

Lithio glycal is prepared by direct abstraction of a vinylic proton with butyl lithium or Schlosser's base (*n*-BuLi and *t*-BuOK). Alternatively, transmetalation of 1-stannyl-glycal with butyl



Scheme 50



Scheme 51

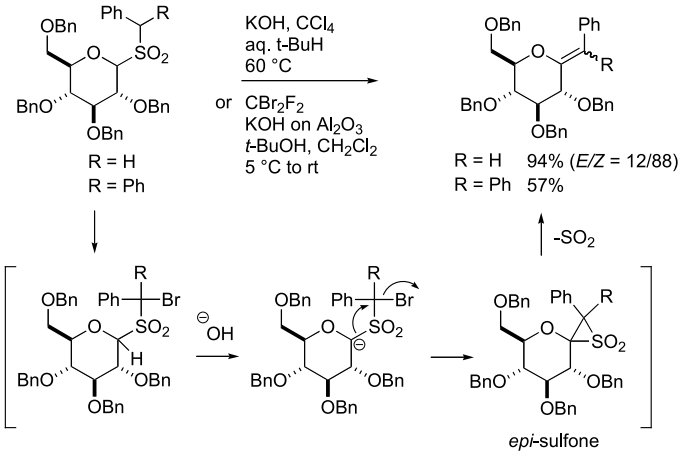
lithium has been frequently employed as exemplified in *Scheme 50* [77] due to the inertness of the resulting tetrabutyltin, the side product. The resulting alkenyl anion reacts with a wide variety of electrophiles such as alkyl iodide, carbon dioxide, allyl halide, and quinone. A recent application of this methodology is shown in *Scheme 51* [78]. Iterative addition of 1-lithioglycal to *p*-quinone was followed by 1,2-migration of one of the glycals (dienol phenol-type rearrangement) by treatment with a Lewis acid to afford a model compound of pluramycin, an aryl-*C*-glycoside antibiotic.

4.2 Ramberg–Bäcklund Reaction

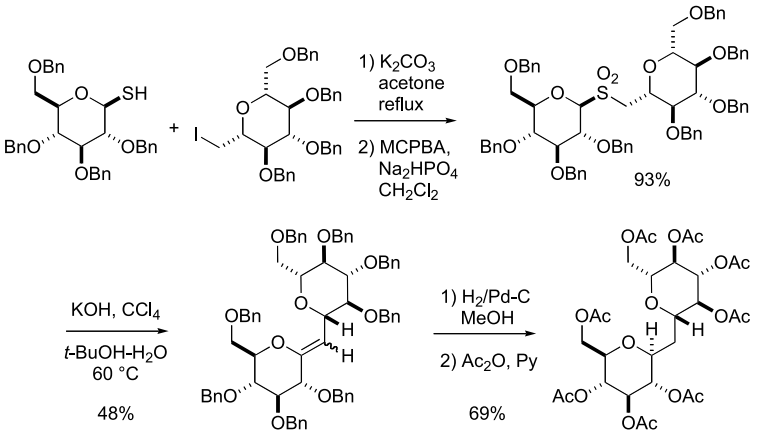
Another valuable type of *C*-glycosylation by using an anomeric anion is known as the Ramberg–Bäcklund reaction, which is classified as an intramolecular glycosylation. The reaction is, in general, a transformation of α-halosulfone to olefin under alkaline conditions. A typical example of *C*-glycosylation by the use of this reaction is shown in *Scheme 52* [79]. When glucosyl sulfone was treated with pulverized KOH suspended in a mixture of CCl_4 and *t*-BuOH or alumina-supported KOH in CBr_2F_2/*t*-BuOH, halogenation of the sulfone moiety occurred, which resulted in the formation of *epi*-sulfone. Elimination of SO_2 gave the *exo*-glycal in good to moderate yield. The product was, in most cases, obtained as a mixture of regioisomers. The selectivity is influenced by the stereochemistry of the *S*-glycoside. It is worthwhile to note that β-alkoxy elimination dose not occur even though an anomeric anion is generated as the intermediate. Since preparation of glycosyl sulfone, a precursor for the Ramberg–Bäcklund reaction, is easy, the method has been employed for a wide range of applications, giving a variety of *C*-glycosides. *Scheme 53* shows an example of the synthesis of *C*-disaccharide commencing with the synthesis of the corresponding sulfide [80].

4.3 Samarium Anion Mediated Reactions

Samarium diiodide (SmI_2)-reduction of glycosyl pyridyl sulfone with aldehydes or ketones under Barbier conditions provides an easily accessible method for the formation of *C*-glycosides. Since the conditions are very mild compared to the conditions used for lithium anomeric anions, a variety of substrates could be used [81]. Stereochemistry and yield of the *C*-glycosylation highly depends on the kind of carbohydrate and the protective groups used. Typical examples are listed in *Table 13*. Samarium-mediated *C*-glycosylation of mannosyl pyridyl sulfone with cyclohexanone gave α-*C*-glycoside in good yield without giving glycal as a by-

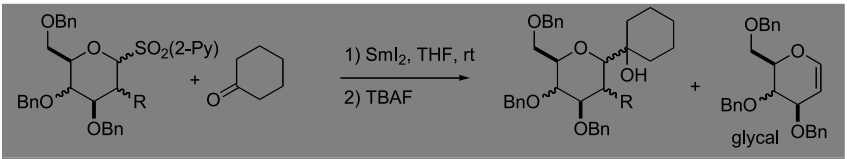


Scheme 52



Scheme 53

Table 13
[81]

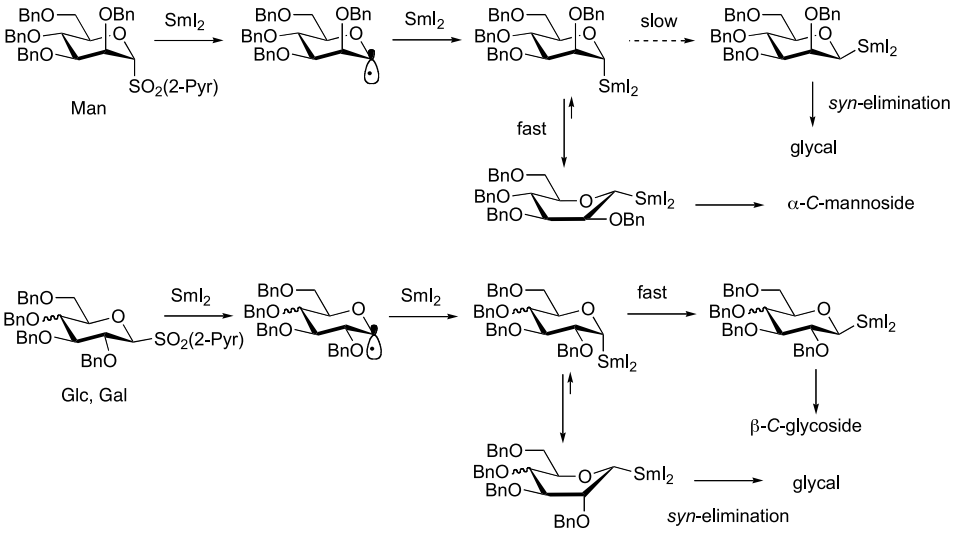


Entry	Glycosyl pyridyl sulfone			Product		
		Anomer	R	C-Glycoside	α/β	Glycal
1	Man	α	OTBS	80%[a]	1/0	0
2	Glc	β	OTBS	57%	0/1	21%
3	Gal	β	OTBS	22%	0/1	32%
4	2-Deoxy Gal	Mixture	H	86%	1/1	0

[a]: The yield was before desilyation

product (entry 1). On the other hand, *C*-glycosylation of glucosyl and galactosyl pyridyl sulfone under the same reaction conditions gave β-*C*-glycosylated products in lower yields with a considerable amount of glycal formation (entries 2 and 3). The configuration of the glycosyl pyridyl sulfone substrate does not influence the stereoselectivity of the reaction.

The stereochemical outcome of this *C*-glycosylation is explained by the following mechanism. Single electron transfer from SmI_2 to glycosyl sulfone leads to the formation of the thermodynamically more stable α-anomeric radical, which is further reduced by SmI_2 to give the α-samarium anion (*Scheme 54*). In order to avoid a repulsive interaction between the Sm-anion and the lone-pair electron of the ring oxygen in the case of mannose, conformation of the samarium anion flips to a boat-like conformer, which reacts with aldehyde to afford the α-*C*-mannoside. In the case of glucose or galactose derivatives, the conformation of the anion is equilibrated with a boat-like conformation, which leads to formation of glycal via a facile

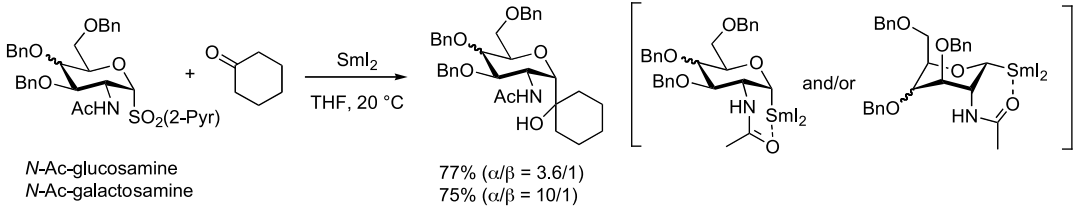


Scheme 54

syn-elimination. On the other hand, the configurational change (inversion of the anomeric configuration) gives a more stable β-samarium anion, which reacts with carbonyl compounds to form β-*C*-glycosides. It is worth noting that *C*-glycosylation of 2-deoxysugar gave a good yield of the product, however no stereoselectivity was observed (entry 4). Glycosyl phosphates instead of glycosyl pyridyl sulfone can also be employed as the substrate for samarium-mediated *C*-glycosylation [82].

Samarium-mediated *C*-glycosylation in order to form *N*-acetyl-sugars such as *N*-acetyl-glucosamine and *N*-acetyl-galactosamine has also been reported (*Scheme 55*) [83]. These *C*-glycosylations exhibited α-selectivity, presumably because the strong coordination of the *N*-acetyl group to samarium (III) in an intermediate fixed the α-configuration of the samarium anomeric anion. This chemistry has also been employed for the synthesis of *C*-glycosyl amino acid such as the mimic of the tumor-associated antigen depicted in *Scheme 56* [84].

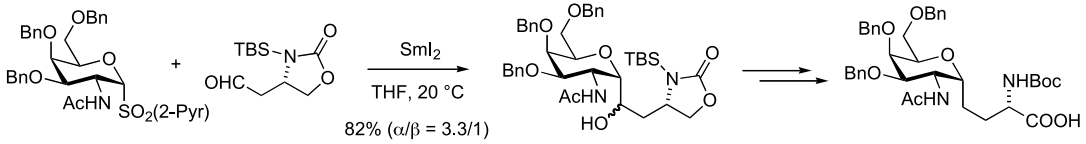
In addition to radical allylation of sialic acid, Sm-mediated *C*-glycosylation of sialic acids has been recognized as a reliable method. Sialic acid derivatives with a suitable leaving group at the C-2 position are reduced with SmI_2 in the presence of a ketone or an aldehyde, giving the α-*C*-glycoside in high yields [85]. A special characteristic of this reaction is the exclusive production of α-*C*-glycoside, because the samarium enolate intermediate chelates with ring oxygen resulting in the reaction of the carbonyl from the less hindered face. As leaving groups, phenylsulfone, 2-pyridyl sulfone, phenyl sulfide, acetate, and chloride can be employed (*Scheme 57*). The method can be applied to the synthesis of *C*-glycoside of other ulosonic acids such as KDN and KDO. Many applications of this methodology for the synthesis of *C*-disaccharide have been reported, some of which are shown in *Scheme 58* [86].



N-Ac-glucosamine
N-Ac-galactosamine

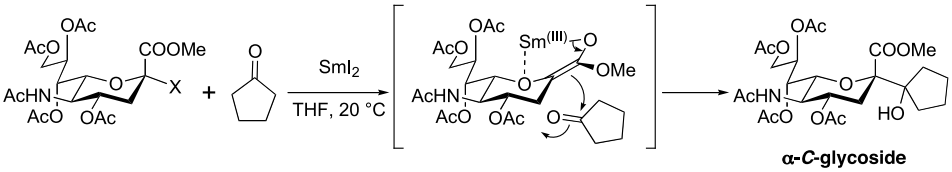
77% (α/β = 3.6/1)
75% (α/β = 10/1)

Scheme 55



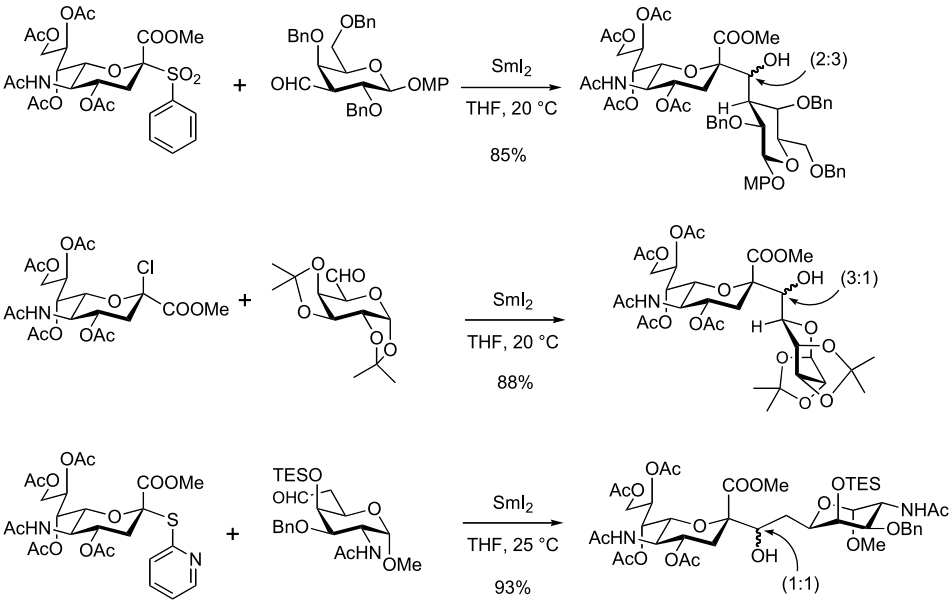
82% (α/β = 3.3/1)

Scheme 56



X = SO_2Ph (95%), SPy (76%), OAc (97%), Cl (95% in the reaction with acetone)

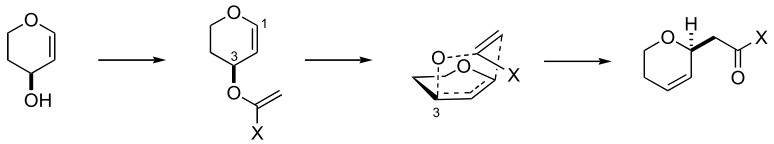
Scheme 57



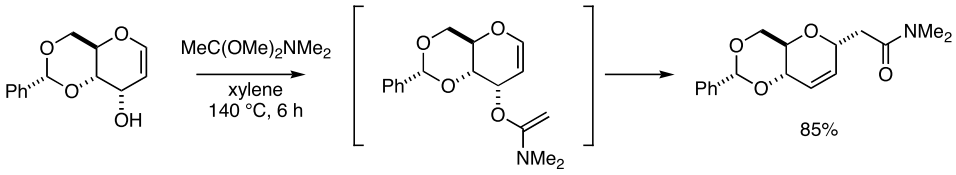
Scheme 58

5 Sigmatropic Rearrangement

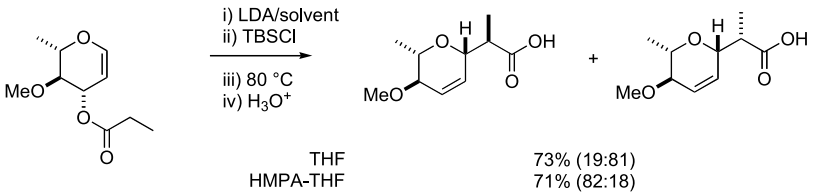
The Claisen rearrangement is the most important reaction among sigmatropic rearrangements for *C*-glycosylation. In the late 1970s, R. Ireland and B. Fraser-Reid reported Claisen rearrangement of glycal ester and its derivative [87] [88]. The Claisen rearrangement generally proceeds through a 6-membered chair-like transition-state, however, in the case of reaction with a glycal ester, a boat-like transition state was proposed as shown in *Eq. 9*. Since the reaction proceeds in a concerted manner, chirality at the C-3 position is transferred to the newly formed *C*-glycoside.



(9)



Scheme 59



Scheme 60

The reaction outlined in *Scheme 59* is an example of a variant of the Claisen rearrangement of allyl ketene aminal (so-called Eschenmoser–Claisen rearrangement) [87]. The reaction dose not require an acid catalyst; glycal was just heated with dimethylacetamide dimethyl acetal to form ketene aminal, which underwent the sigmatropic rearrangement to form the corresponding γ,δ-unsaturated amide.

Claisen rearrangement of ketene silyl acetal derived from glycal ester, so-called Ireland–Claisen rearrangement, has been extensively developed for synthesis of polyether antibiotics. Glycals of pyranose or furanose with an enolizable ester can be employed as substrate for the reaction. However, it has been generally difficult to control the configuration at the α-position of the resulting acetic acid, although some improvement has been reported as shown in *Scheme 60* [88]. In some reactions, epimerization of the products was observed under the conditions.

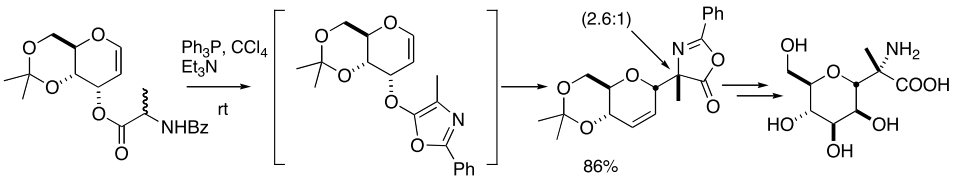
Ireland demonstrated the usefulness of this reaction in the total synthesis of the complex polyether antibiotic, lasalocid A (X537A) [89]. Two Claisen rearrangements were used as key steps for the introduction of the carbon chain to the tetrahydrofuran and tetrahydropyrane rings (arrows indicate the bonds generated by the Claisen rearrangement in *Scheme 61*).

Claisen rearrangement of an oxazole derived from ester of alanine has also been reported (*Scheme 62*) [90]. The product was further transformed into *C*-mannosyl-alanine.

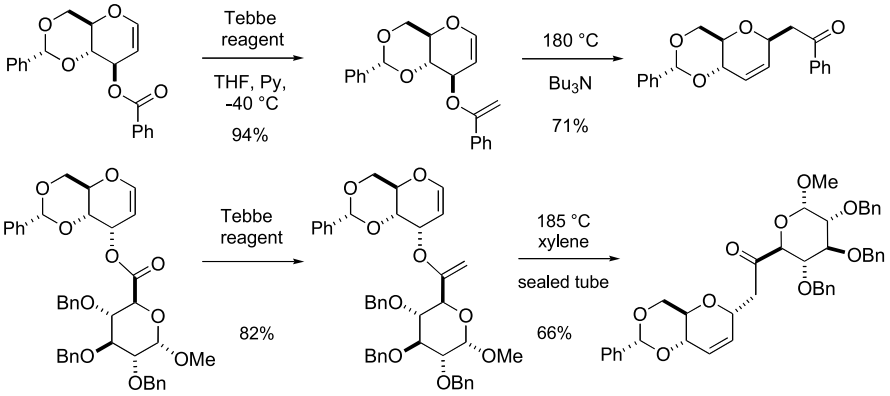
Recently, Claisen rearrangement of allyl-vinyl ether prepared from glycal ester with Tebbe reagent was reported [91]. In contrast to the Ireland–Claisen rearrangement, in principle, a non-enolizable ester can be employed (*Scheme 63*). This method was applied for the synthesis of *C*-disaccharide.

A few examples of [2,3] sigmatropic rearrangements to form *C*-glycoside have been reported [92]. On the other hand, [1,2] Wittig rearrangement has been extensively developed for the synthesis of *C*-glycoside. *O*-Benzyl, allyl, and propargyl glycoside is treated with base (BuLi) to generate a carbanion, which undergoes the [1,2] rearrangement to form *C*-glycoside with

Scheme 61

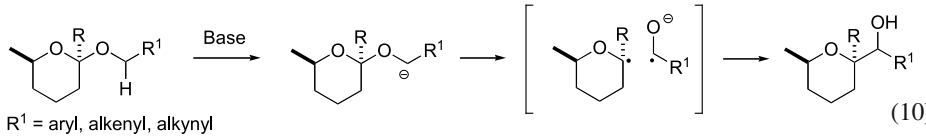


Scheme 62



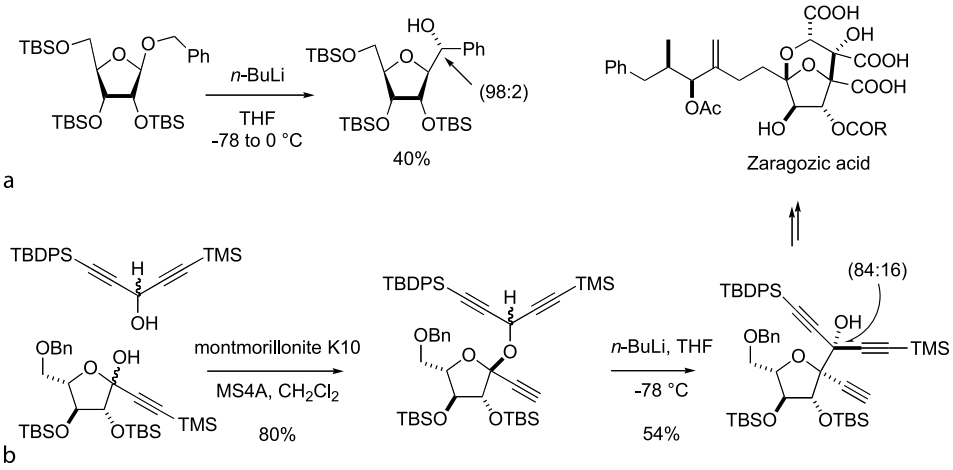
Scheme 63

retention of configuration. The reaction mechanism was reported to be non-concerted but radical-dissociation and recombination in a cage (*Eq. 10*). However, this section will deal with the [1,2] Wittig reaction.



(10)

For example, upon lithiation at the benzylic position of β-*O*-benzyl ribose with butyl lithium, the rearrangement takes place to give the β-*C*-glycosidic product [*Scheme 64*(a)] [93]. The stereochemistry of the secondary alcohol on the side chain was controlled by chelation between lithium alkoxide and ring oxygen. In the total synthesis of zaragozic acid, a potent inhibitor of cholesterol biosynthesis, [1,2] Wittig rearrangement of bis-ethynylalcohol-α-*O*-glycoside was efficiently employed as one of the key steps for constructing the highly congested stereochemistries of zaragozic acid [*Scheme 64*(b)].



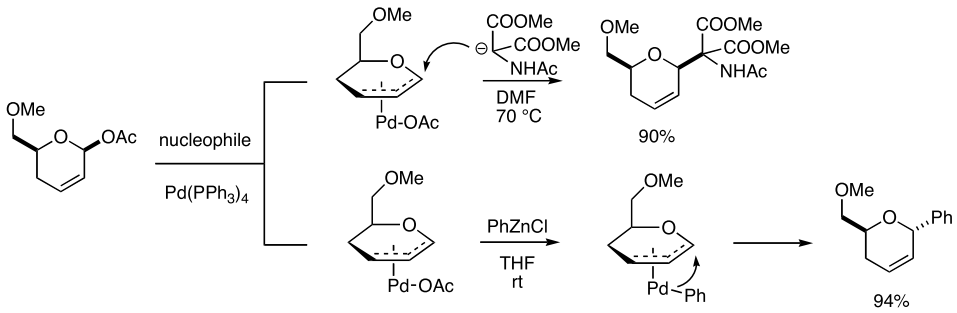
Scheme 64

6 Transition Metal-Catalyzed Reactions

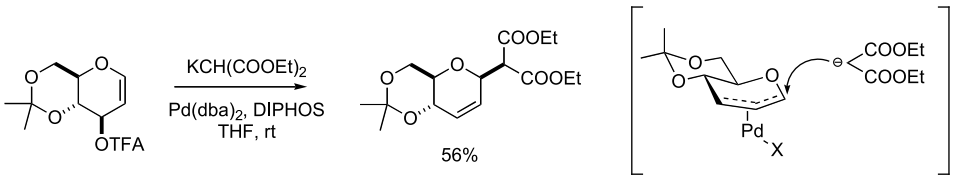
Since transition metal-catalyzed C–C bond formations have been extensively studied, a number of reliable reactions with palladium or nickel catalysts have been reported for the synthesis of *C*-glycosidic linkages due to the high functional group compatibility.

Reaction of the π-allyl-palladium complex with a stabilized anion (this reaction is known as the Tsuji–Trost reaction) has been applied for the stereoselective *C*-glycosylation of unsaturated sugars. 2,3-Unsaturated sugars are treated with a carbon nucleophile in the presence of a palladium(0) catalyst to give *C*-glycosides with migration of the double bond. Stereochemistry of the products mainly depends on the nature of the nucleophiles. The reaction mechanism involves formation of an electrophilic π-allyl-palladium complex by nucleophilic attack of palladium(0) catalyst to allyl ester with inversion of configuration (*anti* attack). When a stable anion such as malonate ester (i. e. soft nucleophile) was employed, the nucleophile attacked the anomeric carbon from the opposite side of the palladium complex to afford the C-1 substituted product with net retention of configuration (*Scheme 65*) [94]. On the other hand, when an unstable anion nucleophile such as organo zinc reagent (i. e. hard nucleophile) was used, the anion attacked the palladium and then the nucleophile was transferred to the C-1 position to give the product with net inversion of configuration. Epimerization might occur when an unsubstituted malonate is employed as the nucleophile due to retro-Michael reaction followed by ring closing.

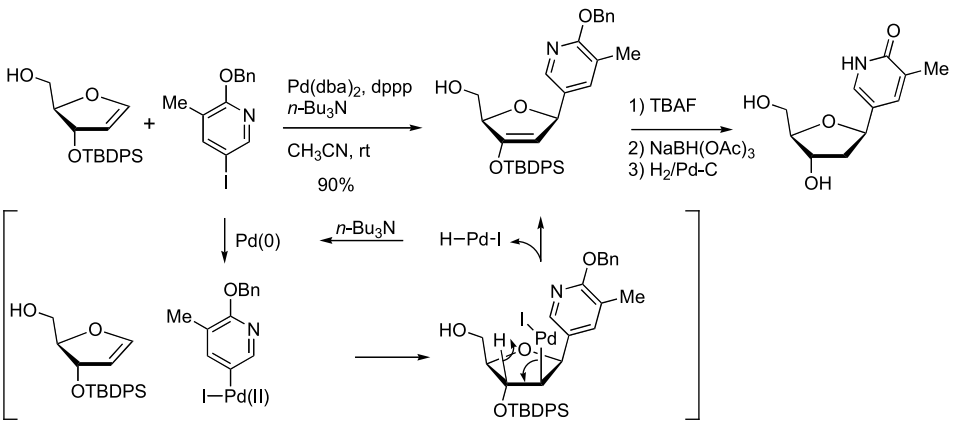
Acetyl glycal has not been employed for the Tsuji–Trost reaction due to the fact that the vinyl ether is inert under the conditions. However, trifluoroacetyl glycal is reactive enough for the reaction with the potassium salt of malonate ester in the presence of Pd(0) catalyst to form the product with migration of the double bond (*Scheme 66*) [95]. According to the mechanisms mentioned above, the reaction proceeded with double inversion of configuration to give β-*C*-glycoside.



Scheme 65

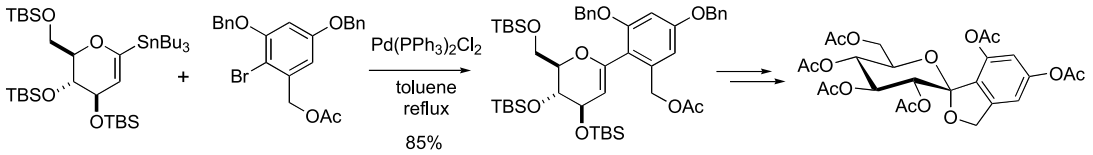


Scheme 66



Scheme 67

Palladium-catalyzed arylation and vinylation of alkene is referred to as the Mizoroki–Heck reaction and is one of the most widely used Pd(0)-catalyzed C–C bond formations in organic synthesis. However, the reaction has not been extensively employed for *C*-glycosylation [96]. The example shown in *Scheme 67* outlines the reaction of iodopyridine and furanose glycal for the synthesis of *C*-nucleoside [97]. The mechanism began with the oxidative addition of iodopyridine to Pd(0) catalyst, and the resulting organo-palladium species was inserted by


TBSO
SnBu3
TBSO
OTBS
+
BnO
OBn
Br
OAc
Pd(PPh3)2Cl2
toluene
reflux
85%
TBSO
BnO
OBn
TBSO
OTBS
OAc
AcO
AcO
AcO
AcO
O
OAc
OAc
O

Scheme 68

glycal double bond with *syn* fashion, followed by elimination of palladium hydride in stereospecific *syn* manner to give β-*C*-glycoside with migration of the double bond. The stereochemistry was explained by the preferential addition of the organo-palladium intermediate to the less hindered β-face of the double bond due to the siloxy group at the C-3 position. In the reaction, dppp (1,3-bis(diphenylphosphino)propane), a bidentate ligand, was critical in order to obtain the high yield of the product. The resulting vinylsilyl ether was desilylated upon treatment with fluoride and the ketone was stereoselectively reduced with triacetoxyborohydride to form the aryl-*C*-nucleoside.

Pd-catalyzed coupling between aryl or alkenyl halides and an organostannane is known as Migita–Stille coupling. Since organostannanes are stable to air and moisture, and the reaction proceeds under essentially neutral conditions, the Migita–Stille coupling has been employed in the synthesis of complex molecules, despite stoichiometric use of toxic organostannane as a coupling partner [98]. ❯ *Scheme 68* shows the coupling between the sterically hindered aryl group and stannylglucal [99]. The product was further transformed into the spiroketal core structure of papulacandin, an antifungal antibiotic.

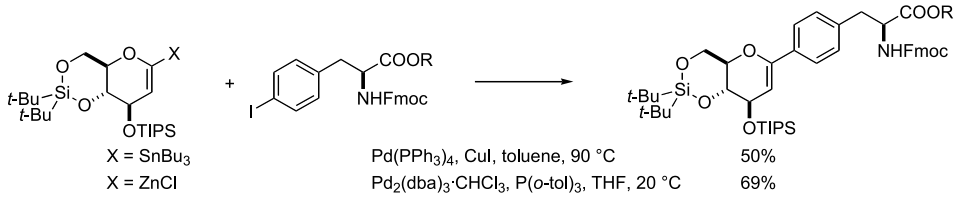
On the other hand, similar coupling by using organozinc reagents as the nucleophilic components is called Negishi coupling. Since organozinc reagents with functional groups such as ester, amide, and carbamate can be prepared, a highly convergent synthesis of multi-functional compounds can be carried out. When both organotin and organozinc reagents with the same substituents are available, the Negishi coupling with the organozinc reagent generally gives better results than that of the Stille coupling (❯ *Scheme 69*) [100]. ❯ *Scheme 70* demonstrates the power of this reaction in the total synthesis of vineomycinone, a naturally occurring aryl-*C*-glycoside antibiotic [101].

In place of the aryl halide or alkenyl halide, the corresponding phosphate or triflate have been employed for the above cross-coupling reaction (❯ *Scheme 71*) [102]. These substrates were easily prepared from the corresponding lactone. The reaction of alkenyl phosphate or triflate has frequently been employed in the synthesis of marine polycyclic ethers such as brevetoxin etc (❯ *Scheme 72*) [103].

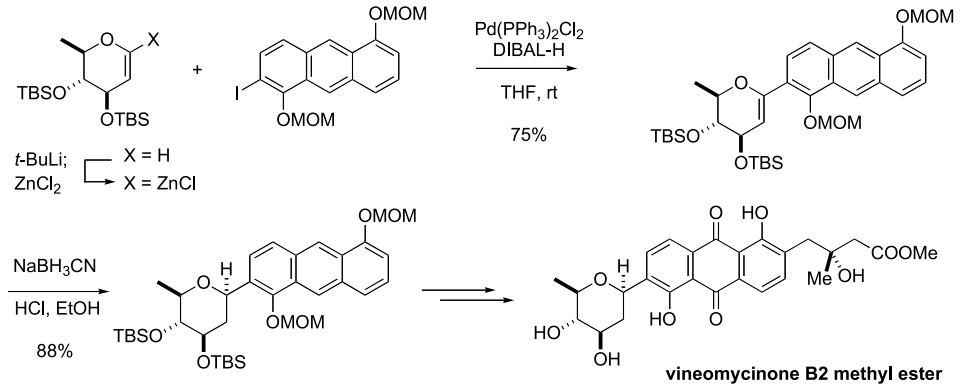
Recently, the reaction of glycosyl halides with organozinc reagent in the presence of a nickel catalyst was reported (❯ *Scheme 73*) [104]. This is a rare example of cross-coupling between sp^3-sp^3 carbons. Stereoselectivities of the reaction between glucosyl or galactosyl halide and the organozinc reagents are very low, while *C*-glycosylation of mannose under the same conditions gave α-*C*-glycosides in high selectivity, due to the steric hindrance of the protected hydroxyl group at the C-2 position.

Palladium-catalyzed cross-coupling reactions between organo borane compound and organic halides (or triflate, phosphate) provide a powerful and reliable method for C–C bond formation known as the Suzuki–Miyaura coupling. sp^3-Alkyl boranes prepared by hydroboration of alkane can be employed as the coupling partner for 1-iodoglucal as shown in ❯ *Scheme 74* [105]. When the reaction was conducted under a carbon monoxide atmosphere, carbonylative coupling occurred. This coupling has been frequently used as a key reaction in the synthesis of marine polyether toxins (❯ *Scheme 75*) [106]. For more examples, see ❯ *Sect. 2.2.2*.

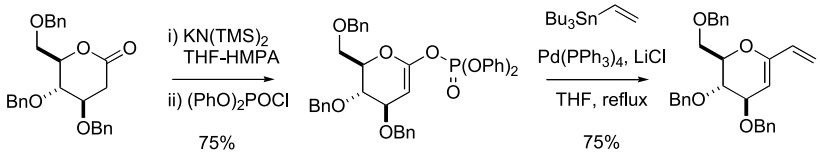
A useful *C*-glycosylation reaction employing a cobalt catalyst is shown in ❯ *Scheme 76* [107]. Glycosyl acetate was treated with diethylmethylsilane and catalytic amounts of $Co_2(CO)_8$ under a carbon monoxide atmosphere, stereospecific hydroxymethylation occurred in 1,2-*trans* manner through neighboring group participation with the acetyl group.



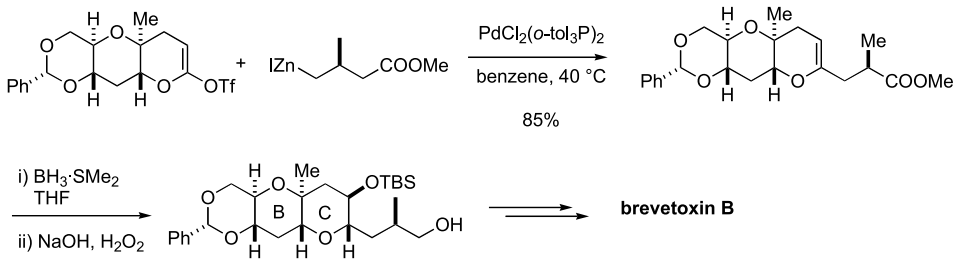
Scheme 69



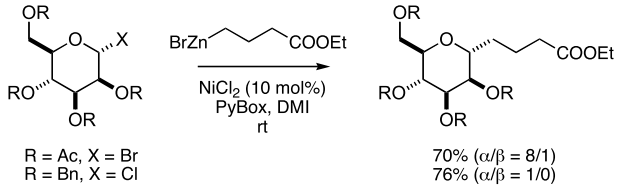
Scheme 70



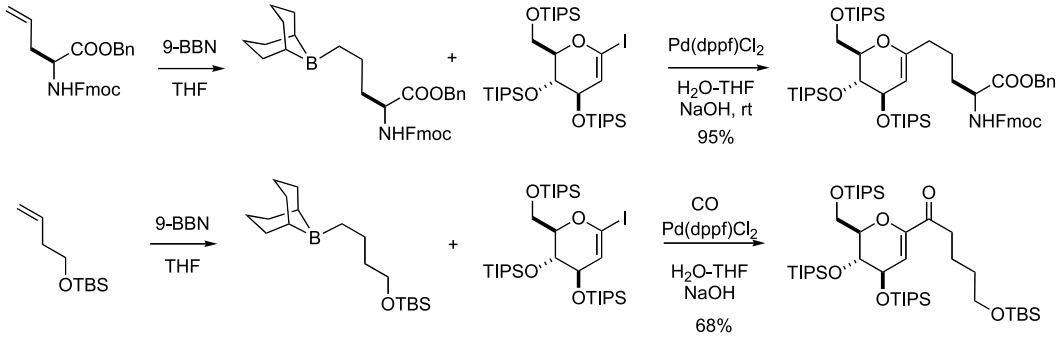
Scheme 71



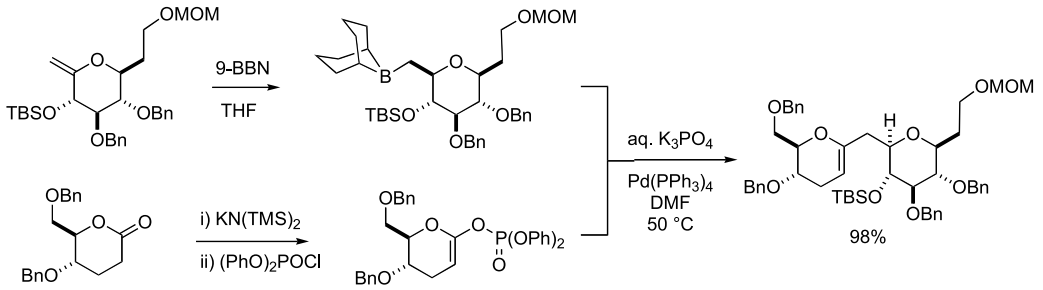
Scheme 72



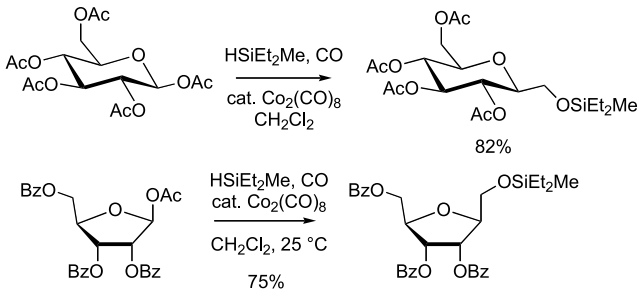
Scheme 73



Scheme 74



Scheme 75



Scheme 76

References

1. Frank RW, Tsuji M (2006) Acc Chem Res 39:692
2. Okadaic acid: (a) Isobe M, Ichikawa Y, Bai DL, Masaki H, Goto T (1986) Tetrahedron 27:4767 and references cited therein; (b) Ley SV, Humphries AC, Eick H, Downham R, Ross AR, Boyce RJ, Pavey JBJ, Pietruszka J (1998) J Chem Soc Perkin Trans 1 3907; (c) Forsyth CJ, Sabes SF, Urbanek RA (1997) J Am Chem Soc 119:8381; Brevetoxin A: Nicolaou KC, Yang Z, Shi GQ, Gunzner JL, Agrios KA, Gärtner P (1998) Nature 392:264; Review of aryl-*C*-glycosides: Jaramillo C, Knapp S (1994) Synthesis 1
3. Previous reviews on synthesis of *C*-glycosides: (a) Postema MHD (1995) *C*-Glycoside Synthesis. CRC Press, London; (b) Levy DE, Tang C (1995) The Chemistry of *C*-Glycosides. Pergamon, Oxford; (c) Postema MHD (1992) Tetrahedron 48:8545; (d) Du Y, Linhardt RJ, Vlahov IR (1998) Tetrahedron 54:9913; (e) Meo P, Osborn HMI (2003) In: Osborn HMI (ed) Carbohydrate. Academic Press, Amsterdam p 337
4. (a) Rodrigues F, Canac Y, Lubineau A (2000) J Chem Soc Chem Commun 2049; (b) Bragnier N, Scherrmann MC (2005) Synthesis 814
5. Tatsuta K, Nakano S, Narazaki F, Nakamura Y (2001) Tetrahedron Lett 42:7625
6. Harvey JE, Raw SA, Taylor RJK (2004) Org Lett 6:2611
7. Review of exo-glycal: Taillefumier C, Chapleur Y (2004) Chem Rev 104:263
8. (a) Schweizer F, Otter A, Hindsgaul O (2001) Synlett 1743; (b) Penner M, Taylor D, Desautels D, Marat K, Schweizer F (2005) Synlett 212
9. Trotter NS, Takahashi S, Nakata T (1999) Org Lett 1:957
10. Tatsuta K, Mukai H, Takahashi M (2000) J Antibiotics 53:430
11. (a) Gascón-López M, Motevalli M, Paloumbis G, Bladon P, Wyatt PB (2003) Tetrahedron 59:0349; (b) Coumbarides GS, Motevalli M, Muse WA, Wyatt PB (2006) J Org Chem 71:7888
12. Molina A, Czernecki S, Xie J (1998) Tetrahedron Lett 39:7507
13. (a) Lakhrissi M, Chapleur Y (1996) Angew Chem Int Ed 35:750; (b) Lakhrissi M, Taillefumier C, Lakhrissi M, Chapleur Y (2000) Tetrahedron Asymmetry 11:417
14. Taillefumier C, Lakhrissi Y, Lakhrissi M, Chapleur Y (2002) Tetrahedron Asymmetry 13:1707
15. Lakhrissi M, Chapleur Y (1994) J Org Chem 59:5752
16. Lakhrissi M, Chapleur Y (1998) Tetrahedron Lett 39:4659
17. (a) RajanBabu TV, Reddy GS (1986) J Org Chem 51:5458; (b) Gómez AM, Pedregosa A, Valverde S, López JC (2003) Tetrahedron Lett 44:6111; (c) John BA, Pan YT, Elbein AD, Johnson CR (1997) J Am Chem Soc 119:4856; (d) Gueyrard D, Haddoub R, Salem A, Bacar NS, Goekjian PG (2005) Synlett 520
18. Alcaraz ML, Griffin FK, Paterson DE, Taylor RJK (1998) Tetrahedron Lett 39:8183
19. (a) Yang WB, Yang YY, Gu YF, Wang SH, Chang CC, Lin CH (2002) J Org Chem 67:3773; (b) Terauchi M, Abe H, Matsuda A, Shuto S (2004) Org Lett 6:3751; (c) Lewis MD, Cha JK, Kishi Y (1982) J Am Chem Soc 104:4976; (d) Xie J, Durrat F, Valéry JM (2003) J Org Chem 68:7896; (e) Pulley SR, Carey JP (1998) J Org Chem 63:5275; (f) Lancelin JM, Zollo

PHA, Sinaÿ P (1983) Tetrahedron Lett 24:4833; (g) Czernecki S, Ville G (1989) J Org Chem 54:610; (h) Ellsworth BA, Doyle AG, Patel M, Caceres-Cortes J, Meng W, Deshpande PP, Pullockaran A, Washburn WN (2003) Tetrahedron Asymmetry 14:3243; (i) Labéguère F, Lavergne JP, Martinez J (2002) Tetrahedron Lett 43:7271; (j) Wellner E, Gustafsson T, Bäcklund J, Holmdahl R, Kihlberg J (2000) ChemBioChem 1:272; (k) Lowary T, Meldal M, Helmboldt A, Vasella A, Bock K (1998) J Org Chem 63:9657; (l) Dondoni A, Scherrmann MC (1994) J Org Chem 59:6404; (m) Nishikawa T, Koide Y, Kanakubo A, Yoshimura H, Isobe M (2006) Org Biomol Chem 4:1268

20. Review of thiazole as a nucleophile: Dondoni A, Marra A (1999) J Chem Soc Chem Commun 2133
21. Beignet J, Tiernan J, Woo CH, Kariuki BM, Cox LR (2004) J Org Chem 69:6341
22. Terauchi M, Abe H, Matsuda A, Shuto S (2004) Org Lett 6:3751
23. (a) Calzada E, Clarke CA, Roussin-Bouchard C, Wightman RH (1995) J Chem Soc Perkin Trans 1 517; (b) Centrone CA, Lowary TL (2002) J Org Chem 67:8862; (c) Krohn K, Heins H, Wielckens K (1992) J Med Chem 35:511
24. (a) Boyd VA, Drake BE, Sulikowski GA (1993) J Org Chem 58:3191; (b) Kaelin DE Jr, Lopez OD, Martin SF (2001) J Am Chem Soc 123:6937
25. Uchiyama T, Woltering TJ, Wong W, Lin CC, Kajimoto T, Takebayashi M, Weitz-Schmidt G, Asakura T, Noda M, Wong CH (1996) Bioorg Med Chem 4:1149
26. Kulkarni SS, Gervay-Hague J (2006) Org Lett 8:5765
27. (a) Toshima K (2000) Carbohydr Res 327:15; (b) Yokoyama M, Toyoshima H, Shimizu M, Mito J, Togo H (1998) Synthesis 409
28. (a) Nicolaou KC, Chucholowski A, Dolle RE, Randall JL (1984) J Chem Soc Chem Commun 1155; (b) Posner GH, Haines SR (1985) Tetrahedron Lett 26:1823
29. (a) Hosomi A, Sakata Y, Sakurai H (1984) Tetrahedron Lett 25:2383; Hosomi A, Sakata Y, Sakurai H (1987) Carbohydr Res 171:223; (b) Nicolaou KC, Dolle RE, Chucholowski A, Randall JL J Chem Soc Chem Commun (1984) 1153; (c) Giannis A, Sandhoff K (1985) 26:1479; (d) Palmacci ER, Seeberger PH (2001) Org Lett 3:1547; Plante OJ, Palmacci ER, Andrade RB, Seeberger PH (2001) J Am Chem Soc 123:9545; (e) Brenna E, Fuganti C, Grasselli P, Serra S, Zambotti S (2002) Chem Eur J 8:1872; (f) Fletcher S, Jorgensen MR, Miller AD (2004) Org Lett 6:4245
30. Hung SC, Lin CC, Wong CH (1997) Tetrahedron Lett 38:5419
31. Allevi P, Anastasia M, Ciuffreda P, Fiecchi A, Scala A (1987) J Chem Soc Chem Commun 101
32. (a) Zhai D, Zhai W, Williams RM (1998) J Am Chem Soc 110:2501; (b) Jobron L, Leteux C, Veyrières A, Beau JM (1994) J Carbohydr Chem 13:507; Désiré J, Veyrières (1995) Carbohydr Res 268:177; (c) Nishikawa T, Ishikawa M, Isobe M (1999) Synlett 123; Nishikawa T, Koide Y, Kajii S, Wada K, Ishikawa M, Isobe M (2005) Org Biomol Chem 3:687; (d) Dondoni A, Mariotti G, Marra A (2002) J Org Chem 67:4475
33. (a) Isobe M, Nishizawa R, Hosokawa S, Nishikawa T (1998) J Chem Soc Chem Commun 2665; (b) Jiang Y, Isobe M (1996) Tetrahedron 52:2877; (c) Tsuboi K, Ichikawa Y, Isobe M (1997) Synlett 713; (d) Jiang Y, Ichikawa Y, Isobe M (1997) Tetrahedron 53:5103
34. Tamura S, Abe H, Matsuda A, Shuto S (2003) Angew Chem Int Ed 42:1021
35. (a) Smith DM, Woerpel KA (2006) Org Biomol Chem 4:1195; (b) Larsen CH, Ridgway BH, Shaw JT, Smith DM, Woerpel KA (2005) J Am Chem Soc 127:10879
36. Brimble MA, Brenstrum TJ (2000) Tetrahedron Lett 41:2991
37. Allevi P, Anastasia M, Ciuffreda P, Fiecchi A, Scala A, Bingham S, Muir M, Tyman J (1991) J Chem Soc Chem Commun 1319
38. Furuta T, Kimura T, Kondo S, Mihara H, Wakimoto T, Nukaya H, Tsuji K, Tanaka K (2004) Tetrahedron 60:9375
39. (a) Kuribayashi T, Ohkawa N, Satoh S (1998) Tetrahedron Lett 39:4537; (b) Kuribayashi T, Mizuno Y, Gohya S, Satoh S (1999) J Carbohydr Chem 18:371
40. Matsuo G, Miki Y, Nakata M, Matsumura S, Toshima K (1999) J Org Chem 64:7101
41. Matsumoto T, Katsuki M, Suzuki K (1988) Tetrahedron Lett 29:6935
42. Ben A, Yamauchi T, Matsumoto T, Suzuki K (2004) Synlett 225
43. Futagami S, Ohashi Y, Imura K, Hosoya T, Ohmori K, Matsumoto T, Suzuki K (2000) Tetrahedron Lett 41:1063
44. Thorn SN, Gallagher T (1996) Synlett 185
45. Thorn SN, Gallagher T (1996) Synlett 856

46. Cook MJ, Fletcher MJE, Gray D, Lovell PJ, Gallagher T (2004) Tetrahedron 60:5085
47. Steinhuebel DP, Fleming JJ, Du Bois J (2002) Org Lett 4:293
48. (a) Bussolo VD, Caselli M, Romano MR, Pineschi M, Crotti P (2004) J Org Chem 69:7383; (b) Bussolo VD, Caselli M, Romano R, Pineschi M, Crotti P (2004) J Org Chem 69:8702
49. Saeeng R, Isobe M (2006) Chem Lett Highlight Review 35:552
50. (a) Danishefsky SJ, Kerwin JF Jr (1982) J Org Chem 47:3805; (b) Takhi M, Rahman AA-HA, Schmidt RR (2001) Tetrahedron Lett 42:4053; (c) Zhu YH, Vogel P (2001) Synlett 82; (d) Yadav JS, Reddy BVS, Rao CV, Chand PK, Prasad AR (2001) Synlett 1638; (e) Tsukiyama T, Isobe M (1992) Tetrahedron Lett 33:7911; (f) Saeeng R, Sirion U, Sahakitpichan P, Isobe M (2003) Tetrahedron Lett 44:6211; (g) Saeeng R, Isobe M (2005) Org Lett 7:1585; (h) Isobe M, Saeeng R, Nishizawa R, Konobe M, Nishikawa T (1999) Chem Lett 467
51. (a) Hosokawa S, Kirschbaum B, Isobe M (1998) Tetrahedron Lett 39:1917; (b) Das SK, Reddy KA, Abbineni C, Roy J, Rao KVLN, Sachwani RH, Iqbal J (2003) Tetrahedron Lett 44:4507; (c) Shoji M, Akiyama N, Tsubone K, Lash LL, Sanders JM, Swanson GT, Sakai R, Shimamoto K, Oikawa M, Sasaki M (2006) J Org Chem 71:5201
52. Tanaka S, Tsukiyama T, Isobe M (1993) Tetrahedron Lett 34:5757
53. Hosokawa S, Isobe M (1998) Tetrahedron Lett 39:2609
54. Hosokawa S, Isobe M (1999) J Org Chem 64:37
55. Pulley SR, Carey JP (1998) J Org Chem 63:5257
56. (a) Allwein SP, Cox JM, Howard BE, Johnson HWB, Rainier JD (2002) Tetrahedron 58:1997; (b) Rainier JD, Cox JM (2000) Org Lett 2:2707; (c) Evans DA, Trotter BW, Côté B (1998) Tetrahedron Lett 39:1709; (d) Timmers CM, Dekker M, Buijsman RC, van der Marel GA, Ethell B, Burchell AB, Mulder GJ, van Boom JH (1997) Bioorg Med Chem Lett 7:1501; (e) Leeuwenburgh MA, van der Marel GA, Overkleeft HS, van Boom JH (2003) J Carbohydr Chem 22:549; (f) Wipf P, Pierce JG, Zhuang N (2005) Org Lett 7:483
57. Hayward MM, Roth RM, Duffy KJ, Dalko PI, Stevens KL, Guo J, Kishi Y (1998) Angew Chem Int Ed 37:192
58. Evans DA, Trotter BW, Coleman PJ, Côté B, Dias LC, Rajapakse HA, Taylor AN (1999) Tetrahedron 55:8671
59. Manabe S, Marui Y, Ito Y (2003) Chem Eur J 9:1435
60. Stichler-Bonaparte J, Vasella A (2001) Helv Chim Acta 84:2355
61. Stichler-Bonaparte J, Bernet B, Vasella A (2002) Helv Chim Acta 85:2235
62. Jaouen V, Jégou A, Veyrières A (1996) Synlett 1218
63. Review: Togo h, He W, Waki Y, Yokoyama M (1998) Synlett 700
64. Giese B (1989) Angew Chem Int Ed 28:969
65. Abe H, Shuto S, Matsuda A (2001) J Am Chem Soc 123:11870
66. (a) Giese B, Linker T, Muhn R (1989) Tetrahedron 45:935; Giese B, Dupuis J (1983) Angew Chem Int Ed 22:622; Giese B, Dupuis J, Nix M (1987) Org Synth 65:236; Giese B, Dupuis J, Leising M, Nix M, Lindner HJ (1987) Carbohydr Res 171:329; (b) Harenbrock M, Matzeit A, Schäfer HJ (1996) Liebigs Ann Chem 55; (c) Keck GE, Enholm EJ, Yates JB, Wiley MR (1985) Tetrahedron 41:4079; (d) Krishna PR, Lavanya B, Jyothi Y, Sharma GVM (2003) J Carbohydr Chem 22:423; (e) Kessler H, Wittmann V, Köck M, Kottenhahn M (1992) Angew Chem Int Ed 31:902; (f) Pontén F, Magnusson G (1996) J Org Chem 61:7463; (g) Junker HD, Phung N, Fessner WD (1999) Tetrahedron Lett 40:7063
67. (a) Bouvet VR, Ben RN (2006) J Org Chem 71:3619; (b) Roe BA, Boojamra CG, Griggs JL, Bertozzi CR (1996) J Org Chem 61:6442; (c) Grant L, Liu Y, Walsh KE, Walter DS, Gallagher T (2002) Org Lett 4:4623; (d) SanMartin R, Tavassoli B, Walsh KE, Walter DS, Gallagher T (2000) Org Lett 2:4051; (e) Abel S, Linker T, Giese B (1991) Synlett 171
68. (a) Giese B, Hoch M, Lamberth C, Schmidt RR (1988) Tetrahedron Lett 29:1375; (b) Pasquarello C, Demange R, Vogel P (1999) Bioorg Med Chem Lett 9:793
69. (a) Paulsen H, Matschulat P (1991) Liebigs Ann Chem 487; (b) Nagy JO, Bednarski MD (1991) Tetrahedron Lett 32:3953
70. Bois M, Skrydstrup T (1995) Chem Rev 95:1253
71. Stork G, Suh HS, Kim G (1991) J Am Chem Soc 113:7054
72. Mazéas D, Skrydstrup T, Doumeix O, Beau JM (1994) Angew Chem Int Ed 33:1383

73. (a) Zhang J, Clive DLJ (1999) J Org Chem 64:770; (b) Zhang J, Clive DLJ (1999) J Org Chem 64:1754
74. Review: (a) Somsák L (2001) Chem Rev 101:81; (b) Beau JM, Gallagher T (1997) In Topics in Current Chemistry 187:1. Springer, Berlin Heidelberg New York
75. (a) Hoffmann M, Kessler H (1994) Tetrahedron Lett 35:6067; (b) Burkhart F, Kessler H (1998) Tetrahedron Lett 39:255
76. (a) Lieberknecht A, Griesser H, Krämer B, Bravo RD, Colinas PA, Grigera RJ (1999) Tetrahedron 55:6475; (b) Colinas PA, Ponzinibbio A, Lieberknecht A, Bravo RD (2003) Tetrahedron Lett 44:7985
77. (a) Lesimple P, Beau JM, Jaurand G, Sinaÿ P (1986) Tetrahedron Lett 27:6201; (b) Hanessian S, Martin M, Desai RC (1986) J Chem Soc Chem Commun 926
78. Parker KA, Su DS (2005) J Carbohydr Chem 24:199
79. (a) Griffin FK, Murphy PV, Paterson DE, Taylor RJK (1998) Tetrahedron Lett 39:8179; (b) Belica PS, Franck RW (1998) Tetrahedron Lett 39:8225
80. Griffin FK, Paterson DE, Murphy PV, Taylor RJK (2002) Eur J Org Chem 1305
81. Skrydstrup T, Jarreton O, Mazéas D, Urban D, Beau JM (1998) Chem Eur J 4:655
82. Hung SH, Wong CH (1996) Angew Chem Int Ed 35:2671
83. (a) Andersen L, Mikkelsen LM, Beau JM, Skrydstrup T (1998) Synlett 139; (b) Urban D, Skrydstrup T, Beau JM (1998) J Org Chem 63:2507
84. Urban D, Skrydstrup T, Beau JM (1998) J Chem Soc Chem Commun 955
85. (a) Du Y, Linhardt RJ (1998) Carbohydr Res 308:161; (b) Polat T, Du Y, Linhardt RJ (1998) Synlett 1195; (c) Malapelle A, Abdallah Z, Doisneau G, Beau JM (2006) Angew Chem Int Ed 45:6016
86. (a) Bazin HG, Du Y, Polat T, Linhardt RJ (1999) J Org Chem 64:7254; (b) Abdallah Z, Doisneau G, Beau JM (2003) Angew Chem Int Ed 42:5209
87. Fraser-Reid B, Dawe RD, Tulshian DB (1979) Can J Chem 57:1746
88. Ireland RE, Wilcox CS, Thaisrivongs S, Vanier NR (1979) Can J Chem 57:1743
89. Ireland RE, Anderson RC, Badoud R, Fitzsimmons BJ, McGarvey GJ, Thaisrivongs S, Wilcox CS (1983) J Am Chem Soc 105:1988
90. Colombo L, Di Giacomo M, Ciceri P (2002) Tetrahedron 58:9381
91. (a) Godage HY, Fairbanks AJ (2000) Tetrahedron Lett 41:7589; (b) Chambers DJ, Evans GR, Fairbanks AJ (2005) Tetrahedron Asymmetry 16:45
92. Tulshian DB, Fraiser-Reid B (1984) J Org Chem 49:518
93. (a) Tomooka K, Kikuchi M, Igawa K, Keong PH, Nakai T (1999) Tetrahedron Lett 40:1917; (b) Tomooka K, Kikuchi M, Igawa K, Suzuki M, Keong PH, Nakai T (2000) Angew Chem Int Ed 39:4502
94. (a) Dunkerton LV, Serino AJ (1982) J Org Chem 47:2814; (b) Moineau C, Bolitt V, Sinou D (1998) J Org Chem 63:582
95. RajanBabu TV (1985) J Org Chem 50:3642
96. Daves GD Jr (1990) Acc Chem Res 23:201
97. Hsieh HP, McLaughlin LW (1995) J Org Chem 60:5356
98. Friesen RW (2001) J Chem Soc Perkin Trans 1 1969
99. Friesen RW, Sturino CF (1990) J Org Chem 55:5808
100. Boucard V, Larrieu K, Lubin-Germain N, Uziel J, Auge J (2003) Synlett 1834
101. Tius MA, Gu XQ, Gomez-Galeno J (1990) J Am Chem Soc 112:8188
102. Nicolaou KC, Shi GQ, Gunzner L, Gärtner P, Yang Z (1997) J Am Chem Soc 119:5467
103. Kadota I, Takamura H, Nishii H, Yamamoto Y (2005) J Am Chem Soc 127:9246
104. Gong H, Sinisi R, Gagné MR (2007) J Am Chem Soc 129:1908
105. Potuzak JS, Tan DS (2004) Tetrahedron Lett 45:1797
106. Sasaki M, Fuwa H, Ishikawa M, Tachibana K (1999) Org Lett 1:1075
107. Chatani N, Ikeda T, Sano T, Sonoda N, Kurosawa H, Kawasaki Y, Murai S (1988) J Org Chem 53:3387

Part 4
Monosaccharides

4.1 Monosaccharides: Occurrence, Significance, and Properties

Zbigniew J. Witczak
Department of Pharmaceutical Sciences, Nesbitt School of Pharmacy, Wilkes University, 84 W. South Street, Wilkes-Barre, PA 18766, USA
zbigniew.witczak@wilkes.edu

Abstract

This chapter is exclusively devoted to the occurrence, significance, and physical and randomly selected chemical properties of various groups of sugars and their implications in synthetic applications of these specified classes of monosaccharides.

Keywords

Monosaccharides; Anhydro-sugars; Amino-sugars; Deoxy-sugars; Fluoro-sugars; Nitro-sugars; Thio-sugars; Carba-sugars; Branched-chain sugars; Mass spectrometry; Infrared spectroscopy

In: *Glycoscience*. Fraser-Reid B, Tatsuta K, Thiem J (eds)
Chapter-DOI 10-1007/978-3-540-30429-6_18:

Abbreviations

CD	circular dichroism
CI	chemical ionization
DMSTS	dimethyl(methylthio)sulfonium triflate
EI	electron impact
ESI	electrospray ionization
FAB	fast atom bombardment
FD	field desorption
IDCP	iodonium dicollidine perchlorate
IR	infrared
MALDI	matrix-assisted laser desorption ionization
NIS	*N*-iodosuccinimide

1 Introduction

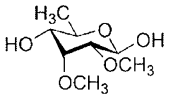
Over a century has passed since Emil Fischer introduced his pioneering concepts of the tetrahedral carbon and its stereochemical consequences: diastereisomerism and selective functional group transformations. Those concepts then became important principles of modern mechanistic and synthetic carbohydrate chemistry. Since then numerous examples of carbohydrate functionalization in almost all classes of monosaccharides have confirmed his hypothetical concepts. Without these classical requirements, today's systematic, mechanistic and synthetic developments in monosaccharide chemistry would be impossible to achieve. Advances in protecting functional groups and their applications in modern synthetic carbohydrate chemistry are nothing else but the legacy of Fischer's rules. Monosaccharides, which are regarded as typical natural polyfunctional carbohydrates possess a unique polyhydroxylated enol character. Of these, the most widely distributed and most intensively investigated are those with an aldehyde function. Their strongly hydrophilic character is attributed to the presence of primary and secondary hydroxyl groups. Additionally, their solubility in water and aprotic polar organic solvents is critical for the physical and chemical properties that Fischer predicted.

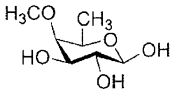
During the past decade, many excellent reviews covering the occurrence, properties, and preparation of naturally occurring monosaccharides, and properties of 1,6-anhydrosugars [1], aminosugars [2], deoxy-sugars [3], fluoro-sugars [4], nitro-sugars [5], thio-sugars [6,7], carbasugars [8], and branch-chain sugars [9] have been published. These groups of sugars are perhaps the most common monosaccharides and are considered the core of carbohydrate chemistry. Inorganic complexes of sugars have also been discussed in detail [10].

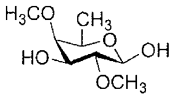
This chapter is exclusively devoted to the physical and randomly selected chemical properties of various groups of sugars and their implications in synthetic applications of these specified classes of monosaccharides.

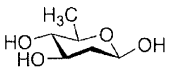
2 Occurrence

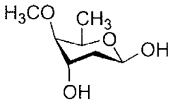
Carbohydrates occur in almost all living matter in the biosphere, but not all organisms make this class of compounds in equal amounts.

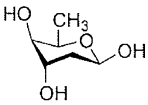

Mycinose

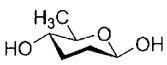

Curacose

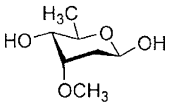

Labilose


Olivose

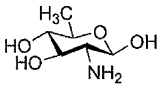

Olivomose

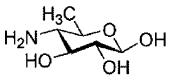

Boivinose

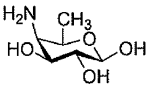

Amicetose

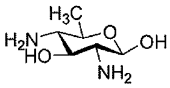

Cymarose

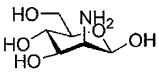
Scheme 1

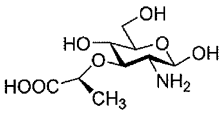

Quinovosamine

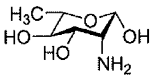

Viosamine

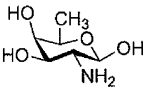

Thomasamine


Bacilosamine

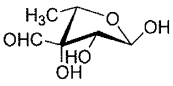

Mannosamine

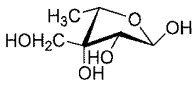

Muramic acid

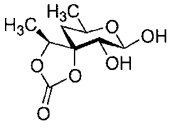

Rhamnosamine

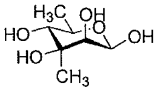

D-Fucosamine

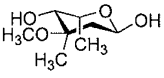
Scheme 2

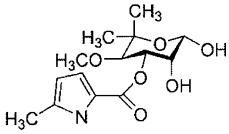

Streptose

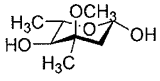

Dihydrostreptose

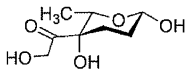

Aldgraose

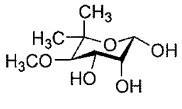

Evalose

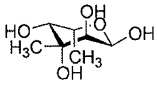

Arcanose

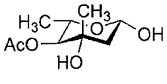

Coumerose

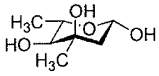

Cladinose

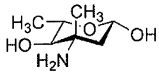

L-Pilarose

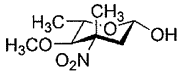

Noviose


Vinelose


Olivomycose


L-Mycarose


Vancosamine


Evernitrose

Scheme 3

The de novo biosynthesis of organic material in green plants found on Earth produces approximately 2×10^{11} metric tons of carbohydrates per year. The dominant part of this (ca. 90%) is found in trees [1]. Bacteria and lower organisms, such as fungi and algae, show much wider variation. The fungi (including yeast and lichen) as eukaryotes can be placed in a middle position between the plant and animal kingdoms. The death of plants, bacteria, and fungi are the means by which the enormous amount of decomposed biomass returns to the cycle of organic compounds including monosaccharides, oligosaccharides, and polysaccharides.

In total about 250 different, naturally occurring monosaccharides have been isolated and identified. Interestingly, the monosaccharides produced by nature are almost exclusively in the form of pentoses and hexoses, with the higher carbon sugars (more than nine) rarely occurring in nature.

Among eight diastereomeric aldohexose pairs, only three, namely glucose, mannose, and galactose are found in nature in both D- and L-configuration.

The most important monosaccharides which exist in pyranose rings found in higher mammalian organisms are: D-glucose, D-galactose, D-mannose, D-glucosamine, D-galactosamine, L-fucose, and neuraminic acid. With these seven monosaccharides, nearly all oligosaccharides and polysaccharides occurring in the mammalian system can be assembled.

In *Scheme 1*, *Scheme 2* and *Scheme 3*, some of the important representatives of natural monosaccharides are presented and compiled by the three different classes: deoxy-sugars, amino-sugars, and branched-chain sugars. *Table 1* lists their names and natural sources of occurrence.

3 Significance

Sucrose is probably the most abundantly produced bulk chemical. It can be degraded to its hexose component sugars D-glucose and D-fructose.

Among the pentoses, D-xylose is the cheapest and is readily accessible from wood- or straw-derived xylans. L-Sorbose is the most readily available L-sugar on a large scale due to its technological accessibility from D-sorbitol in the vitamin C industrial production process.

The sugar alcohols D-xylitol and D-sorbitol as industrial bulk chemicals are mostly used as food ingredients due to their excellent sweetening properties and clear taste profile.

D-Gluconic acid as an important industrial bulk chemical is widely used in textile industries as a chelating agent and as a catalyst for textile printing.

4 Physical Properties

The physical properties and many chemical properties of monosaccharides depend on the molecular shape. Equilibria of pyranoid compounds depend largely upon the *axial-equatorial* relationship between substituents on the rings. Thus, the α/β ratio of pyranoses is governed to some extent by the tendency of the anomeric hydroxyl group to occupy the less hindered equatorial substituent orientation.

Equilibria between different sugars (or their derivatives) are controlled in a similar fashion. One of the classical examples of this phenomenon is a rapid configurational change or interconversion in aqueous solution often reaching the characteristic population of conformers as illustrated in *Table 2*.

Table 1
Occurrence of monosaccharides in nature*

Name	Occurrence
Pentoses	
L-Arabinose	Plant polysaccharides
D-Arabinose	Glycosides of Aloe, bacterial polysaccharides
D-Lyxose	Free form in human heart muscle
L-Lyxose	Constituent of the antibiotic *Curamycin*
D-Ribose	Constituent of nucleic acids (RNA)
D-Xylose	Constituent of hemicellulose
Pentuloses	
D-Ribulose	Free compound in seaweed, leaves of sugar beet, and leaves of barley germ
Hexoses	
D-Glucose	Widespread in plants, animals, and microorganisms; constituent of starch and cellulose
D-Galactose	As the free sugar only in small amount in wine and fermented milk products (cheese and yogurt). Constituent of oligosaccharides and polysaccharides
D-Mannose	As the free sugar only in traces in fruits such as apples, peaches, constituent of oligosaccharides and polysaccharides
D-Talose	Constituent of the antibiotic *Hygromycin B*
Hexuloses	
D-Fructose	As the free sugar in many fruits, plants, and in honey, unit of sucrose and inulin
D-Psicose	As the free sugar in residue of fermented molasses
L-Sorbose	Component of pectins from the skin of the fruit of *Passiflora edulis* (passion fruit)
D-Tagatose	Component of the gum exudates of *Sterculia setigera*
L-galactose	Component of algal polysaccharides such as agar
Heptuloses	
Sedoheptulose (D-Altroheptulose)	As the free sugar in all green plants, intermediate of photosynthesis, isolated from Sedum leaves
Volemose (D-Mannoheptulose)	Free compound in avocado
Octuloses	
D-Glycero-D-mannooctulose	Free compound in California avocado
Nonuloses	
D-Erythro-L-glucononulose	Free compound in avocado
Deoxysugars	
D-Allomethylose (6-Deoxy-D-allose)	Cardiac glycosides, of *Digitalis*
Antiarose (6-Deoxy-D-gulose	Glycosides of *Digitalis canariensis*
Boivinose (2,6-Dideoxy-D-xylohexose)	Glycosides of *Strophantus*

Table 1 (continued)

Name	Occurrence
D-Chinovose (6-Deoxy-D-glucose)	Glycosides of the bark of *Cinchona*
Cymarose (2,6-Dideoxy-3-O-methyl-D-allose)	Glycosides of *Strophantus*
Digitoxose (2,6-Dideoxy-D-allose)	Glycosides of *Digitalis*
Epifucose (6-Deoxy-L-talose)	Constituent of the cell wall polysaccharide of a strain of *Actinomyces bovis* and the K-antigen
L-Fucose (6-Deoxy-L-galactose)	Constituent of oligosaccharide of human milk, glycosides of some plants, some seaweeds (fucoidan) and some glycoproteins
2-Deoxy-D-ribose	Constituent of DNA
D-Rhamnose (6-Deoxy-D-mannose)	Gram (–) bacteria strain GS
L-Rhamnose (6-Deoxy-L-mannose)	As the free sugar in small amounts in wine and in the extract of the leaves of poison ivy (*Rhus toxidendron*), and in some microbial polysaccharides
Thalomethylose (6-Deoxy-D-talose)	Constituent of capsular polysaccharide of Gram (–) bacteria of strain GC
Branch-chain sugars	
D-Apiose	Glycosides of parsley and celery, bark of *Hevea brasiliensis*
Hammamelose	As the diphosphate in spinach, constituent of hamamellitannins of bark of witch hazel
Streptose (5-deoxy-3-C-formyl-L-lyxose)	Constituent of antibiotic Streptomycin
Uronic acids	
D-Iduronic acid	Unit of the mammalian polysaccharides, chondroitin sulfuric acid, and heparin
D-Glucuronic acid	Unit of polysaccharides of connective tissue, plant polysaccharides gum arabic and wood hemicelluloses
D-Galacturonic acid	Unit of polysaccharides, pectin, plant gums, tragacanth or karya gum
D-Guluronic acid	Unit of polysaccharide algin
D-Mannuronic acid	Unit of polysaccharide algin
Amino sugars carboxylic acids	
D-Glucosaminic acid (2-Amino-2-deoxy-D-gluconic acid)	Metabolism product of D-glucosamine with *Pseudomonas fluorescente*
D-Galactosamineuronic acid (2-amino-2-deoxy-D-galacturonic acid)	Component of the Vi-Antigen of *Escherichia coli*
Muramic acid (2-amino-2-deoxy-3-*O*-(α-D-carboxyethyl-D-glucose)	Component of cell walls and spores of many bacteria

Table 1 (continued)

Name	Occurrence
Sialic acids: Mono-and diacetylneuraminic Neuraminic acid (D-glycero-D-galacto-5-amino-2,5-dideoxy-2-ketononic acid)	Widespread component of gangliosides, acids, lipopolysaccharides, oligosaccharides and glycoprotein in human and animal organisms
Amino cyclitols	
Streptamine (1,3-Diamino-1,3-dideoxy-scyllo-inositol)	Degradation product of *Streptidin* component of the antibiotic Streptomycin
2-Deoxystreptamine (1,3-Diamino-1,2,3-trideoxy-scyllo-inositol)	Component of the antibiotics *Kanamycin* and *Neomycin*
Amino-sugars	
D-Glucosamine (Chitosamine) (2-Amino-2-deoxy-D-glucose) and *N*-acetyl-D-glucosamine (2-Acetamido-2-deoxy-D-glucose)	As the *N*-acetyl derivative unit of several oligo- and polysaccharides, glycoprotein, chitin and several glycosaminoglycans
D-Galactosamine (2-Amino-2-deoxy-D-galactose) and *N*-acetyl-D-galactosamine (2-Acetamido-2-deoxy-D-galactose)	The acetylated compound is a widespread unit of animal glycosaminoglycans in connective tissue (Chondroitin sulfate, dermatan sulfate) and component of many bacterial immunopolysaccharides and glycoproteins
D-Mannosamine (2-Amino-2-deoxy-D-mannose)	Degradation product of sialic acids
D-Gulosamine (2-Amino-2-deoxy-D-gulose)	Degradation product of the *Streptomycin* antibiotics, *Strepthothricin* and *Streptolin B*
D-Fucosamine (2-Amino-2-deoxy-D-fucose)	Unit of lipopolysaccharides of *Chromobacterium violaceum*
L-Fucosamine (2-Amino-2-deoxy-L-fucose)	Unit of capsular polysaccharides of *Pneumoccocus Type V*
3-Amino-3-deoxy-β-D-ribose	Component of antibiotic *Puromycin*
Kanosamine (3-Amino-3-deoxy-D-glucose)	Component of antibiotic *Kanamycin*
Mycosamine (3-Amino-3,6-dideoxy-D-mannose)	Component of antibiotics *Natamycin*, *Amphothericin*, and *Nystatin*
Mycaminose (3,6-Dideoxy-3 dimethylamino-D-glucose)	Component of antibiotic *Carbomycin*
D-Fructosamine (1-Deoxy-1-amino-D-fructose)	As the free sugar in human and mammalian blood serum
Cyclitols	
Conduritol	Free compound in the bark of the condurango tree

Table 1 (continued)

Name	Occurrence
D-Inositol	Free compound in the wood of California sugar pine
L-Inositol	Free compound in the latex of *Hevea brasiliensis*
myo-Inositol	Free compound in many plants, animals and microorganisms, as phosphate (phytic acid) in many plants
Scyllitol, Cocositol	Free compound in a few plants (acorns), in fish (shark), also in ray
Mytillitol, *C*-Methyl-*scyllo*-inositol	Free compound in sphincter of mussels
Quebrachitol, (2-Methyl-L-Inositol)	Free compound in *Quebracho* barks and in the latex of *Hevea brasiliensis*
D-Quercitol (1-Deoxy-*muco*-inositol)	Free compound in the bark of oak tree
L-Viburnitol (1-Deoxy-myo-inositol)	Free compound in the leaves of *Viburnum sinus*
Sugar alcohols	
Pentitols	
Adonitol, Ribitol	Free compound in *Adonis vernalis*; in the roots of *Bapleurum* sp., unit of Vitamin B_2
D-Arabitol, D-Lyxitol	Free compound in *Boletus bovines* mushrooms, also in some lichens (*Ramalina scopulorum*)
Hexitols	
D-Glucitol, D-Sorbitol	Free compounds in fruits, plums, apples, pears, and some berries
D-Mannitol	Free compound in *Chanterelle* mushrooms also in marine brown algae and bacterial fermentations
Galactitol, Dulcitol	Free compound in plants, in substantial amount in Madagascar manna
Heptitols	
Perseitol (L-glycero-D-mannoheptitol)	Free compound in certain seeds of avocado
Volemito (D-glycero-D-mannoheptitol)	Free compound in *Lactarius volemus* mushroom

*Compiled from Carbohydrates, Collins PM (ed), Chapman and Hall Chemistry 1987

Functional group reactivity also depends largely on conformation. This difference is directly related to the total number of equatorial groups. The greater the numbers of axial substituents present in the molecule, the greater the ring strain, the higher the energy of the molecule. Equatorial anomeric hydroxyl groups of pyranoses are oxidized with aqueous bromine solution more rapidly than are axial hydroxyls. Thus, a lower conformational instability of the molecule corresponds to the larger differences in the oxidation rates.

Other physical parameters that correlate with conformation include melting point, (*Table 3*) solubility, refractive index, chromatographic mobilities, infrared (IR), and ^{1}H and ^{13}C NMR spectra, as well as circular dichroism (CD).

Table 2
The percentage composition of free sugars in aqueous solution at equilibrium[a]

Sugar	Cyclic form				
	α-Pyranose [%]	β-Pyranose [%]	α-Furanose [%]	β-Furanose [%]	Temp. [°C]
Allose	14.0	77.5	3.5	5.0	31
Altrose	27.0	43.0	17.0	13.0	22
Arabinose	60.0	35.5	2.5	2.0	31
Fructose	2.0	70.0	5.0	23.0	30
Fucose	28.0	67.0		5.0	31
Galactose	30.0	64.0	2.5	3.5	31
Glucose	37.3	62.6		0.1	22
Gulose	16.0	81.0		3.0	22
Idose	38.5	36.0	11.5	14.0	31
Lyxose	70.0	28.0	1.5	0.5	31
Mannose	63.7	35.5	0.6	0.2	36
Psicose	22.0	24.0	39.0	15.0	27
Ribose	21.5	58.5	6.5	13.5	31
Sorbose	93.0	2.0	4.0	1.0	31
Tagatose	79.0	16.0	1.0	4.0	27
Talose	42.0	29.0	16.0	13.0	22
Xylose	36.5	63.0		< 1.0	31

[a]Compiled from: Carbohydrates, Collins PM (ed), Chapman and Hall Chemistry, 1987

Monosaccharides, disaccharides, and the lower oligosaccharides have distinct melting points. Some boast two melting points, the first as a hydrate and the second for the anhydrous form. A classical example of this phenomenon is D-galactose monohydrate (m.p. 118–120 °C) and an anhydrous anomeric mixture of D-galactose (m.p. 167 °C).

Many properties depend on the symmetry of the molecule, including the melting point, which correlates with predicted energy levels on the basis of a stable configuration. The melting point is the lowest temperature at which the crystal lattice is disrupted and overcomes the energetically favorable crystalline array. Usually, a sharp transition could be attributed to the rigorous geometrical criteria that have to be fulfilled for lattice formation.

Similar factors of symmetry should be also considered during the physical process of solubility in water as well as in aprotic solvents. A solvent dissolves a crystalline structure when the forces interacting between the solvent and the molecule exceed those between soluble molecules. Symmetrical compounds have lower solubilities than their "less ordered" counterparts. Through identical logic one can surmise that conformationally unstable forms will be highly soluble, and most importantly, extremely difficult to crystallize. *Table 4* lists selected solubility data.

Highly specific parameters that depend largely on conformer stability, such as refractive index and chromatographic behavior (mobility), will be more directly related to the magnitude of instability present in any single form.

Table 3
Melting points of selected free monosaccharides*

Sugar	Form	m.p. [°C]	Salt
2-acetamido-2-deoxyglucose	α-D-pyranose	202–204	
(acetylglucosamine)	β-D-pyranose	182–184	
allose	β-D-pyranose	128–129	
altrose	β-D-pyranose	103–105	
	β-L-pyranose	107–109.5	
2-acetamido-2-deoxygalactose	α-D-pyranose	159–160	
(acetylgalactosamine)		172–173	
2-amino-2-deoxyglucose	α-D-pyranose	88	
(glucosamine)	α-D-pyranose	190–210	HCl
	β-D-pyranose	110–111	
3-amino-3-deoxyglucose (kanosamine)	α-D-pyranose	132–137	HCl
2-amino-2-deoxyribose (ribosamine)	D-pyranose	147–148	HCl
3-amino-3-deoxyribose	α-D-furanose	164–165	HCl
1,6-anhydrogalactopyranose		222–233	
1,6-anhydroglucopyranose		179–180	
1,6-anhydrogalactofuranose		183–184	
arabinose	α-D-pyranose	159–160	
ascorbic acid	L-furanose	190–192	
2-deoxygalactose	β-D-pyranose	120–121	
2-deoxyglucose	α-D-pyranose	128–129	
	β-D-pyranose	146	
3-deoxyglucose	α-D-pyranose	105–107	
4-deoxyglucose	α-D-pyranose	136–140	
6-deoxyidose	α-L-pyranose	98–100	
fructose	β-D-pyranose	102–104	
fucose	β-L-pyranose	140–141	
galactose	α-D-pyranose	167 (162–164)	
glucose	α-D-pyranose	83	(monohydrate)
		146	(anhydrous)
	β-D-pyranose	148–150	(anhydrous)
gulose	α-L-pyranose	200–203	$CaCl_2$ complex
lyxose	α-D-pyranose	117–188 (106–107)	
mannose	α-D-pyranose	132–133	
psicose			
ribose	β-D-pyranose	95 (87)	
sorbose	α-D-pyranose	165	
tagatose	α-D-pyranose	134–135	
talose	α-D-pyranose	133–134	
xylose	β-D-pyranose	144–145 (153)	

*Compiled from Carbohydrates, (1987) Collins PM (ed) Chapman and Hall Chemistry

Table 4
Solubility of selected free monosaccharides

Sugar	Solubility	
	H_2O (% by wt. at 25 °C)	Aq. 80% EtOH (g/ml at 20 °C)
α-D-galactose		0.65
α-D-glucose	62	4.5
α-D-glucose hydrate	30.2	
β-D-glucose	72	
β-D-fructose		27.4
β-D-mannose		13.0
β-D-arabinose	52	12.0

*Compiled from Merck Index

4.1 Infrared Spectroscopy

Infrared spectroscopy is an extremely useful tool for detecting the presence and type of functional group. Strong absorption bands are characteristic of various classes of monosaccharides containing a carbonyl group.

Functionalized monosaccharides, peracetylated aryl glycosides for example, exhibit bands in the 800–1000 cm^{-1} region and are characteristic for differentiating α- and β- anomers. The relative intensities of C-O-C stretching vibrations in the 1000–1100 cm^{-1} region and a band near 300 cm^{-1} for β-anomers are criteria for differentiating anomeric peracetylated alkyl and aryl glycosides [11].

Infrared spectroscopy has also proved to be a very useful tool for stereochemical analyses of pyranoid branched-chain compounds. The method is highly capable of detecting structural features that have the hydroxyl group as well as a substituted hydroxyl group in a 1,3-diaxial relationship to each other. In such systems the free hydroxyl groups can form hydrogen bonds to the substituted oxygen atoms, which can be easily detected by precise measurement of stretching frequency of the O–H bonds in a dilute carbon tetrachloride solution. Compounds with such bonding have stretching frequencies near 3500 cm^{-1}, which is more than 100 cm^{-1} lower then the value for the stereoisomers without bonded hydroxyl groups.

4.2 Circular Dichroism Spectrometry

The analytical methodology, circular dichroism spectroscopy is useful in determining the enantiomeric purity of molecules and the configuration of naturally occurring sugars. Chiral derivatives absorb left- and right-circularly polarized light differently, thereby creating the Cotton effect, which is positive for α-D-glycoside and negative for β-D-glycoside, respectively [12].

In particular, the relation of the sign to the Cotton effect in the $\pi-\pi^*$ intramolecular charge transfer transition near 225 nm in dibenzoates of α-diols is the basis of the "exciton chirality method" developed by Nakanishi [13]. This simple analytical technique determines the chirality of the diols and their configuration and conformation.

Table 5
Chemical shifts (δ) of methyl protons in ^{1}H-NMR of peracetylated monosaccharides*

Axial		Equatorial	
		Primary (C-6)	Secondary
C-2	2.19–2.16		2.09–2.04
C-3	2.16–2.15		2.09–2.02
C-4	2.19–2.15	2.02–1.98	2.11–2.13

*Compiled from Richardson AC, McLauchlan KA (1962) J Chem Soc 2499

The CD and conformational properties of acyclic derivatives of fructose, sorbose, tagatose, and four deoxy derivatives of fructose have been reviewed [14].
Over 15 years ago a new technique [15] using new instrumentation and known as FT-IR vibrational CD was developed and has been continuously used since then.

4.3 Mass Spectrometry

Mass spectrometry is an important method in determining carbohydrate structures, and there are excellent reviews of the most recent mass spectrometric methods for complex carbohydrates [16,17]. Among many useful techniques, classical ionization of volatile molecules through electron impact (EI) or chemical ionization (CI) [18,19], electrospray ionization (ESI) [20], and field desorption techniques (FD) [21] are frequently employed in structural analysis.
Structural information obtained from these modern methods is complemented by fast atom bombardment (FAB) mass spectra [22] of both free and functionalized monosaccharides with a variety of functional groups such as thiols, nitro, keto, benzyl, azido, epoxy, esters, etc. FAB mass spectrometry is one of the most reliable and sophisticated techniques developed thus far. Moreover, one can predict the specific fragmentation pattern for the selected class of functionalized monosaccharides through the FAB method as well as the matrix assisted laser desorption ionization (MALDI) mass spectrometry method.

4.4 ^{1}H and ^{13}C-NMR Spectroscopy

A very useful source on various NMR techniques and chemical shifts of functionalized carbon atoms that contains tabulated data on monosaccharides is Reference [23]. Specifically correlating functional groups with chemical shifts of protons and the associated carbon is an important tool for elucidating the structure of the monosaccharide moiety of many complex oligosaccharides. Since monosaccharides contain one or more chiral centers, absolute configuration must be known in order to predict the conformational shape of the molecules. The outcome of many stereoselective functionalizations of free and partially functionalized monosaccharides depends on conformation.
These aspects are discussed in an excellent recent review [24]. It must also be noted that NMR experiments are indispensable tools in structure elucidation and furnish specific information regarding chemical shift, (δ), spin-spin coupling constants [25] (J), spin-lattice relaxation times (T_1), spin-spin relaxation times (T_2) as well as nuclear Overhauser enhancement (nOE). These

data specific to conformational properties of various carbohydrate molecules also provide traditional information on chemical shifts of peracetylated pentoses and hexoses and show the methyl protons of an axial oriented acetoxy group to have lower chemical shifts than those of equatorially oriented groups [26].
Additionally, ^{1}H and ^{13}C NMR spectroscopy provide valuable means for ascertaining the ring size and a preferred conformation of functionalized protected monosaccharides such as benzylidene acetals [27,28].

5 Chemical Properties and Reactivity

The chemical properties of typical monosaccharides are a consequence of different levels of reactivity of primary and secondary hydroxyls, during specific functionalization of unprotected as well partially functionalized sugar derivatives. The hydroxyl group is nucleophilic, acidic (pKa 10–18), and is also easily oxidized by a wide range of reagents such as pyridinium dichromate (PDC), chromium trioxide (CrO_3), manganese oxide (MnO_2), and various combinations. Because the hydroxyl group can participate in numerous transformations under relatively mild conditions, a main obstacle in organic synthesis is to ensure that a specific hydroxyl function in a multifunctional molecule is protected from unwanted reactions altogether or until such time that intrinsic reactivity is specifically required. In specific cases, that can be easily done by a number of protecting agents.
The presence of an electron-withdrawing group at the primary position as well as the secondary (for example $–NO_2$, $–COCH_3$), makes the sugar molecule extremely reactive towards certain reducing agents such as sodium borohydride and other boranes.
Moreover, the concept of reactivity of sugar molecules (free or functionalized) is heavily governed by the geometry of the molecule. The specific conformation $^{1}C_4$ or $^{4}C_1$ of selected sugar molecules (aldoses or ketoses) causes different reactivities of the secondary hydroxyl groups. This particular natural characteristic of monosaccharide reactivity is an important and critical factor in selective functionalization by a variety of protecting groups.
The acyl derivatives of thiols, alcohols, carboxyls, and amino groups are amongst the oldest protecting groups still in use. The ester groups formed are not only good protecting groups during oxidation or glycosidation but during peptide coupling reactions as well. They can be attacked by stronger nucleophiles.
The electrophilicity of the carbonyl group of the ester functionality is a dominant feature of its chemistry and shields carbonyls from nucleophilic attack. However, its electrophilic properties must be considered as a major obstacle in any synthetic approach of stereoselective functionalization. This difficult task of protecting the carbonyl group can be achieved by highly stereoselective functionalization with very strong nucleophiles. This selective protection by functionalization with various blocking groups such as benzylidene, isopropylidene, and cyclohexylidene acetals is reported [27].

5.1 Anhydro-Sugars

Among the anhydro sugar family of monosaccharides, 1,6-anhydro-sugars are the most common and are particularly suitable for highly stereoselective protection of the remaining three

free secondary hydroxyl groups at the C-2, -3, and -4 positions. The 1,6-anhydro bridge eliminates the need for protecting groups at the anomeric carbon and at the 6-OH. Moreover, it fixes the conformation of the system of this bicyclic molecule and most importantly sterically hinders the β-D-face of the molecule. Another very important factor is that the pyranose ring in 1,6-anhydrohexopyranoses is locked in the $^1C_4$ conformation (and it is thus relatively energetically disfavored). The $^1C_4$ conformation also generates sterically hindered stereocenters, contrary to those encountered with the conventional $^4C_1$ conformer.

These important factors constitute the proper strategy for the regioselective protection of the remaining active centers of the molecules as their reactivity varies significantly. Thus, the relative reactivity of the various –OH groups upon stereoselective tosylation of 1,6-anhydro sugars decreases according to the determined order as depicted in *Fig. 1* [27].

Because of the relatively high reactivity of the tosyl and mesyl functional groups, tosylates and mesylates are convenient intermediates to the synthetically useful epoxides. This particular route exploits intramolecular displacement reactions involving participation from β-hydroxyl groups. The specific requirement of the *trans*-diaxal orientation of the leaving group and the hydroxyl group for the formation of epoxide must be observed in order to achieve the formation of these bicyclic epoxides. A classical example of epoxide formation's high efficiency is the base catalyzed conversion of 1,6-anhydro-2-mesyl-β-D-galactopyranose with a *trans*-diaxial arrangement at C-2 and C-3 to the corresponding 1,6:2,3-dianhydro-β-D-talopyranose [28]. In contrast, 1,6-anhydro-2-tosyl-β-D-altropyranose with a *trans*-diequatorial disposition of the vicinal hydroxyl and sulfonate group is unfavorable for the ring closure, as no epoxide formation is observed [29].

As a consequence of the target epoxide's specific stereochemical formation requirements, the aforementioned anhydrosugar conformation forces the hexose out of the stable $^1C_4$ conformation into the $^5H_0$ conformation, which is higher in energy and easily attacked by nucleophiles, thus enabling a variety of applications. This particular property of high chemical reactivity

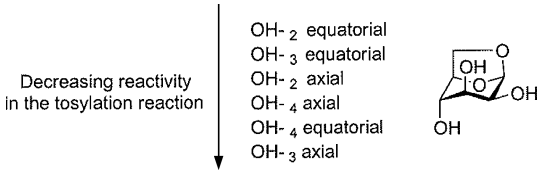


Figure 1
Decreasing reactivity of 1,6-anhydrosugars

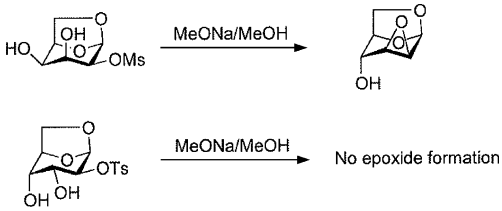


Figure 2
Epoxide formation [29]

makes the compound an ideal starting material for the selective formation of pyranose derivatives functionalized at C-2, C-3, or C-4.

An interesting stereoselective route to this conformationally fixed derivative called "Cerny epoxide" has been developed [30]. This approach proceeds directly from unprotected D-glucal through intramolecular condensation with allylic rearrangement with the formation of the pseudoglucal derivative. The 1,6-anhydro ring formation proceeds in this case via intramolecular cyclization promoted by copper sulfate under strictly anhydrous conditions (*Fig. 3*).

One of the most effective methods of regioselective deactivation of the dianion of carbohydrate diols by complexing 1,6-anhydro-4-*O*-benzyl-β-D-mannopyranose with copper salts [31,32,33,34] has been employed for their effective monobenzylation. Interestingly, when benzyl iodide was used, the yield of the monosubstituted products was higher than 85%. This very effective approach was used even for the unfavorable axial position at C-3; the yield of the expected ether was very good (65%) (*Fig. 4*).

These two referred, regioselective functionalizations of 1,6-anhydrosugars are based on a high degree of selectivity during the nucleophilic reactions and constitute convenient approaches to other classes of functionalized monosaccharides including inositol derivatives. The latter example of such functionalization through intramolecular rearrangement was serendipitously reported by Klemer and Kohla [35] during strong base (BuLi/THF) catalyzed treatment of 1,6-anhydro-3,4-isopropylidene-β-D-galactopyranose with the formation of the *C*-butyl inositol derivative in 58% yield. This intramolecular rearrangement proceeds via abstraction of H-5 and the formation of an intermediate carbanion. Rearrangement with ring opening affords the enolate and its ring closure to the cyclitol following butyllithium addition to the carbonyl group as shown in *Fig. 5*. Interestingly, the reaction can be applied with other 1,6-anhydrohexopyranose derivatives [36] but clearly it is restricted by the strongly basic conditions used and by the specific natures of substrates.

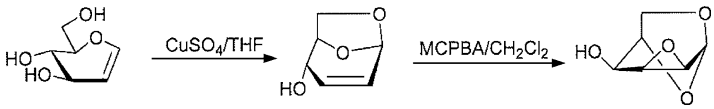


Figure 3
Cerny's epoxide formation

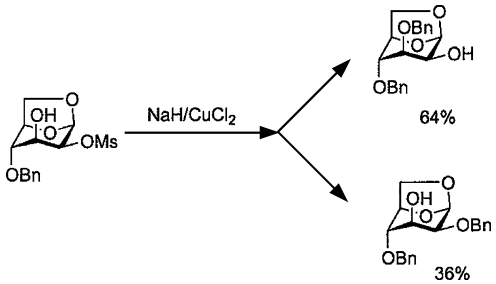


Figure 4
Monobenzylation of 1,6-anhydro-4-*O*-benzyl-β-D-mannopyranose

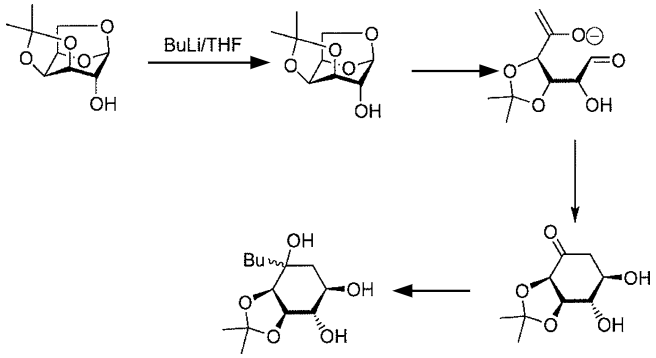


Figure 5
Formation of *C*-butylinositol

Table 6
The physical properties of selected 1,6-anhydro and 1,6:2,3 dianhydro-sugars*

Sugar		$[\alpha]_D$	m.p [°C]
1,6-anhydrogalactopyranose		–22° (H_2O)	226
	2,3,4-triacetate	–6.2° ($CHCl_3$)	73–74
	3,4-isopropylidene	–73.3° ($CHCl_3$)	151–152
1,6-anhydrogalactofuranose		+56° (c 1.0 H_2O)	183–184
	2,3,4-triacetate	+162° (c 1.0 $CHCl_3$)	79–80
		+144.9° (c 1.0 $CHCl_3$)	
1,6-anhydroglucopyranose		–66.2° (H_2O)	179–180
	2,3,5-triacetate	–50.8° ($CHCl_3$)	109–110
1,6-anhydrogulopyranose		+50.4° (c 2.8 H_2O)	154–155
	2,3,4-triacetate	+22.1° (c 1.5 $CHCl_3$)	114–115
	2,3-isopropylidene	+49.8° (c 1.0 $CHCl_3$)	93–94
1,6-anhydromannopyranose		–128° (H_2O)	211
1,6:2,3-diaanhydromannopyranose		–35° ($CHCl_3$)	68–70
	2,3,4-triacetate	–50.8° ($CHCl_3$)	109–110

*Compiled from Carbohydrates, (1987) Collins PM (ed) Chapman and Hall Chemistry

Whereas these are the only selected examples of unusual chemical reactivity of 1,6-anhydrosugars, the other anhydrosugars such as 3,6-anhydro [37] are worthy of mention on the basis of their special chemical character of increased reactivity of the epoxide toward reactive nucleophiles. Some of the physico-chemical properties of 1,6-anhydro-sugars are listed in *Table 6*.

5.2 Amino-Sugars

Nitrogen is the element that neatly divides organic chemists including carbohydrate chemists into two camps or factions. The primary disagreement focuses on the amino function whose nucleophilicity and basicity is a concern for some researchers, whereas others have fully

Table 7
The physical properties of selected amino-sugars*

Sugar	[α]D	m.p °C α/β-ratio	Solution
2-Amino-2-deoxyaltrose/HCl	+114.3° (*c* 2.2, H_2O)		Syrup
2-Amino-2-deoxygalactose/HCl	+47.3° 91.5° (c 0.95, H_2O)	178–190 dec	
2-Amino-2-deoxygulose/HCl	+5.6°–1.8° (c 2.9, H_2O)	165–170 dec	
2-Amino-2-deoxyglucose/HCl	+100° 72.5 (H_2O)	190–210 dec	
2-Amino-2-deoxyidose/HCl	+1° (*c* 3.0, H_2O)		
2-Amino-2-deoxyidose/Nac	−38° (*c* 2.0, H_2O)		Syrup
2-Amino-2-deoxymannose/HCl	−4.6° (*c* 10.0, HCl)	178–180	
2-Amino-2-deoxyribose/HCl +14°	−3 (*c* 1.0, H_2O)	147–148	
2-Amino-2-deoxytalose/HCl	−10° (*c* 0.5, H_2O)	164–165	
2-Amino-2-deoxyxylose/HCl	+80° +40° (*c* 0.8, H_2O)	165–167	
2-Amino-2,6-dideoxyglucose/HCl	+88° 53 (*c* 1.2, H_2O)	172–173	
3-Amino-3-deoxyglucose/HCl	+188° (*c* 1.56, H_2O)		Syrup
3-Amino-3,6-dideoxygalactose	+114° (*c* 0.8, H_2O)	174–176 85/15	
3-Amino-3,6-dideoxyglucose	+188° (*c* 1.56, H_2O)		Syrup
3-Amino-3,6-dideoxymannose/HCl	−11.5° = 13° (*c* 1.0, H_2O)	162	
3-Amino-3,6-dideoxymannose/Nac	−34° (*c* 1.0, EtOH)	195–196	
4-Amino-4-deoxygalactose/HCl	+48.2 (*c* 1.0, H_2O)	120	
4-Amino-4-deoxyglucose/HCl	+25.2° (*c* 1.0, H_2O)	110	
4-Amino-4,6-dideoxyglucose/HCl	+213° (*c* 1.56, $CHCl_3$)	103	
4-Amino-4,6-dideoxymannose/HCl	−20°>-13° (*c* 1.3, H_2O)	162	
4-Amino-4,6-didexymannose/NAc	+34° (*c* 2.0, H_2O)	178–180	
5-Amino-5-deoxymannose/HCl	+41° (*c* 2.0, $CHCl_3$)	90–92	
6-Amino-6-deoxyglucose/HCl		229	

*Compiled from Carbohydrates, (1987) Collins PM (ed), Chapman and Hall Chemistry

exploited its physical properties in developing new reagents for the protection of the amino group, functionalization, or simple retrogradation.

There are various levels of the amino group's basicity of this class of sugars and their imino- and aza- analogs [38,39] are an extra functionality and should be considered as a positive factor for introducing additional functionalities such as esters. However, the introduction of other groups such as a nitro group, create problems of sufficient reactivity towards specific protection reactions of the existing amino group.

Pelyvas, Moneret, and Herczegh present one of the best reviews of specific amino sugars, a leading reference of the field [40]. The characteristic properties of selected amino sugars including optical rotation are listed in *Table 7*.

5.3 Deoxy-Sugars

Deoxy-sugars are very common monosaccharides and are frequently found as components of oligosaccharides in antibiotics. A large number of the so-called bioactive lipopolysaccharides as well as several antibiotics are deoxy-sugars [41].

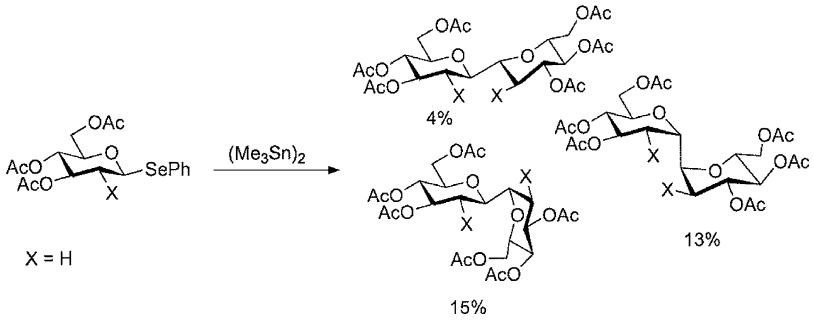


Figure 6
Dimerization of phenylselenyl 2-deoxy-D-glucopyranose [42]

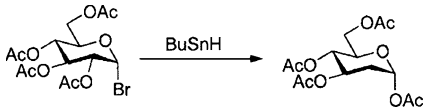


Figure 7
Rearrangement approach to 2-deoxysugars [42]

Table 8
The physical properties of selected free deoxy-sugars*

Sugar	$[\alpha]_D$	m.p. [°C]	α/β- pyranose ratio in solution %
6-deoxy-D-allose	+1.2° (H_2O)	151–152	
6-deoxy-D-altrose	+16.2° (H_2O)	Syrup	
6-deoxy-D-glucose	+75°–+30° (c 1.5 H_2O)	146	36/64
6-deoxy-D-gulose	−49.2°–−42.4° (c 0.5 H_2O)	123–126	47.5/52.5
6-deoxy-D-idose	+14.7° (c 0.7 H_2O)	Syrup	
6-deoxy-L-idose	+2.2°–−26° (H_2O)	98–100	
6-deoxy-D-talose	+20.6 (H_2O)	129–131	44/28
			16/11 fur
4-deoxy-D-galactose	+44°–+60.3° (c 2.4 H_2O)	136–140	
4-deoxy-D-glucose	+44°–+60.3° (c 2.4 H_2O)	136–140	
4-deoxy-L-ribose	+28.5° (c 0.2 H_2O)	Syrup	30/70
3-deoxy-α-D-glucose	+102°–+32.2° (c 1.0 H_2O)	105–107	24.5/55
3-deoxy-D-galactose	+6.6° (c 2.4 H_2O)	Syrup	27/53
3-deoxy-D-xylose	−6.3° (c 1.3 H_2O)		
2-deoxy-α-Dglucose	+57.6°–+46° (c H_2O)	128–129	47.5/52.5
2-deoxy-β-D-galactose	+24°–+56° (c 2.0 H_2O)	120–121	44/44
2-deoxy-D-gulose	+12° (H_2O)	Syrup	
2-deoxy-D-ribose	−56.2° (H_2O)	96–98	40/35

*Compiled from Carbohydrates, (1987) Collins PM (ed) Chapman and Hall Chemistry

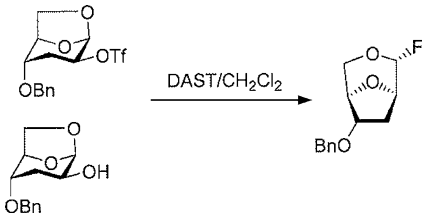


Figure 8
Ring contraction with DAST [43]

Many deoxy-sugars are synthesized by conventional methods of dehalogenation or reduction of thiocarbonates tosylates, mesylates etc. One notable exception, involving dimerization of phenylselenyl 2-deoxy-D-glucopyranose in a concentrated solution and in the absence of trapping agents is worthy of mention as it is related to the chemical character of the free radical intermediate. The approach proceeds by the formation of the axial-axial coupling product as the major one and axial-equatorial as the minor product.
Again, the shape and conformation of the anomerically functionalized 2-deoxysugar molecules are the primary factors for a particular compound's behavior.
Interestingly and presumably, this specific observation on the concentration and the presence/absence of a radical initiator led to the serendipitous discovery of a new synthetic route to 2-deoxysugars during radical debromination [42] of peracetylated hexopyranoses and furanoses. If the acceptor is absent when the radical is generated, then the radical attacks the ester carbonyl oxygen transferring the *O*-acetyl group to the anomeric carbon, and the radical formed at C-2 is then reduced by tributyltin hydride to give the 2-deoxy sugar.
Specifically, 2-deoxy-sugars exhibit specific chemical inertness to β-elimination because of the absence of the participating neighboring leaving group to form new functionalized derivatives at C-2. Some of the physico-chemical properties of selected deoxy-sugars are listed in *Table 8*.

5.4 Fluoro-Sugars

As previously reported, fluoro-sugars play an important role in the development of potential biological targets such as enzyme inhibitors and carbohydrate drug candidates.
The fluorine ion is a hard base with low nucleophilicity and displacement reactions using the fluoride ion do not proceed readily. Thus, preparing the fluoro functionalized monosaccharides is a challenging process and in many instances proceeds with unpredictable consequences. For example, treatment of 1,6-anhydro-3-deoxy-4-*O*-benzyl-β-D-arabino-hexopyranose with (DAST dialkylamino sulfur trifluoride) resulted in ring contraction resulting in a 3,8-bicyclo [3.2.1] octane derivative [43]. Analogous ring contractions, where a leaving group and the ring oxygen atom are in *anti*-periplanar disposition have previously been observed in a monocyclic pyranose system [44,45] and quite similar rearrangements are known to occur in the 1,6-anhydro system [46,47].

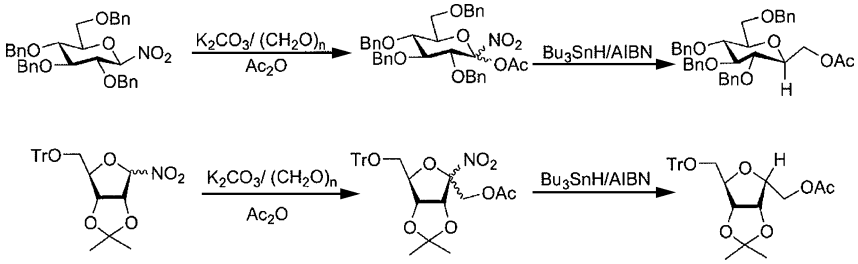


Figure 9
Reduction of anomeric nitrates with tributyltin hydride [55,56]

Csuk and Glanzer [48] published a detailed compilation of NMR data of fluorinated carbohydrates. Two other reviews on fluoro-sugars also mention chemical characteristics and biological applications [49,50].
One particular class of fluoro-sugars (i. e. glycosyl fluorides) is eminent for their role as glycosyl donors in oligosaccharide synthesis. Few excellent reports for applications of glycosyl fluorides in the synthesis of complex oligosaccharides and natural products have been published [51,52].

5.5 Nitro-Sugars

The nitro group, as one of the many functional groups with a strong electron-withdrawing effect, plays an extremely important role in the reactivity of the sugar molecule no matter where it is located. The anomeric nitrates introduced by Vasella [53] are particularly useful reactive synthons for introducing additional functionalities via formation of the C–C bond by Henry aldol condensation [54] followed by a radical denitration.
The reactivity of the anomeric nitro group is a primary factor in applying nitro-sugars to the synthesis of other functionalized monosaccharides, including branched-chain sugars, amino-sugars, and glycopeptides and as intermediates to glycoconjugates.
Notable differences in the susceptibility of the anomeric nitro group, as well as reactivity to various reducing agents including tributyltin hydride were also reported [55,56].
The chemical character of the nitro group at various positions in the sugar moiety makes this particular class of functionalized monosaccharides extremely attractive as intermediates for amino, unsaturated, oximino, and deoxy-sugars.

5.6 Thio-Sugars

Only a few reviews appeared during the last decade that dealt with the chemical character, biological activity, and functionality of thio-sugars [6,7] including thio-oligosacccharides [57]. Thio-sugars are becoming more popular as convenient probes for enzyme-inhibition studies [58,59]. Thio-sugars with sulfur in the ring possess specific biological activities, including interference with enzymes or proteins involved in the recognition of their natural coun-

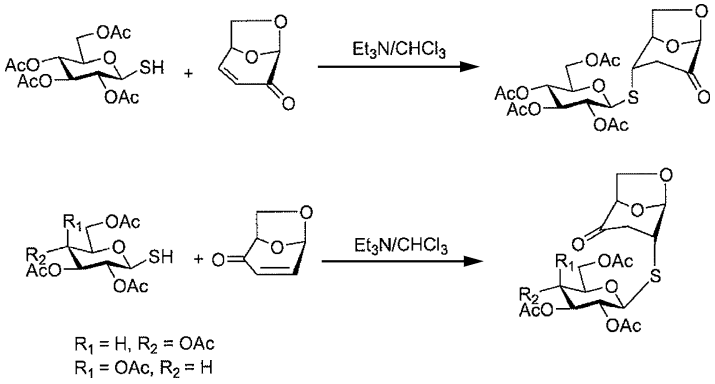


Figure 10
Michael addition of thiols to conjugated system of levoglucosenone [64,65,66,67]

terparts. These biological and physico-chemical properties are attributed to the different electronic properties of the sulfur (electronegativity) and oxygen atoms. The size of the sulfur atom and its electron density, which is more dispersed, creates significant differences. Additionally, the C–S bond is longer (ca. 1.8 Å) and the C–S–C angle (ca. 95–100°) is smaller than the corresponding oxygen-containing structure. One can thus apply rational design rather than empirical experimentation in the optimization of synthetic targets, i. e. thio-sugars.

When comparing the chemical and biological properties of oxygen and sulfur analogs, two other factors must be considered. First, thio-sugars easily form sulfoxides. Second, sulfur derivatives such as thiols can be oxidized easily to form disulfides as useful intermediates in synthetic organic chemistry approaches. These two chemical characteristics of sulfur derivatives are primary factors in considering any synthetic strategies and are the potential biochemical mechanisms for forming natural products containing sulfur.

Thio-sugars [60,61] are excellent glycosyl donors for oligosaccharide synthesis. Excellent reactivity of activated thioglycosides [62,63] by common promoters such as methyltriflate, dimethyl(methylthio)sulfonium triflate (DMSTS), iodonium dicollidine perchlorate (IDCP), and *N*-iodosuccinimide (NIS) alone or with a combination of triflic acid or silver triflate makes them versatile glycosylation donors.

A second group of thio-sugars (i. e. functionalized thiols) is attracting interest as convenient and highly reactive nucleophiles for the stereoselective functionalization of sugar enones in the formation of *S*-linked thiodisaccharides [64,65,66,67].

The reactivity of thiols [68] is sufficient to form a sulfur bridge at various positions in the sugar moieties during stereoselective coupling or conjugation with unsaturated sugars. A special category of thio-sugars are sugars that have sulfur as a heteroatom in the ring, including 5-thio-hexoses and 4-thio-pentoses. Various physicochemical properties including distortion analysis [69] and ^{13}C NMR analysis of 1-thio-sugars [70] have been published. Some of the physico-chemical properties of the selected derivatives of the above class of thio-sugars [71,72,73,74,75,76,77,78,79,80,81,82,83,84,85,86,87,88,89,90,91,92,93,94,95,96] are listed in *Table 9*.

Table 9
The physical properties of selected 6-thio-, 5-thio-, 4-thio-, and 1-thio-sugars*

Sugar	$[\alpha]_D$	m.p. [°C]	a/β-ratio in solution	Lit.
6-Thio-D-galactose/Penta-Ac	+130° (c $CHCl_3$)	113		[71]
6-Thio-D-glucose/Penta-Ac	−15° (c $CHCl_3$)	127–128		[72]
5-Thio-L-allopyranose	−73° (c 1.00 H_2O)	146–148		[73]
5-Thio-D-arabinose				[74]
5-Thio-L-fucose				[75,76]
Methyl 5-Thio-D-galactose	−347.7° (c 0.92 H_2O)			[77]
5-Thio-D-glucose	+188° (c 1.56 H_2O)	135–136	85/15	[78,79]
5-Thio-D-glucose/Penta-Ac	+213° (c 1.56 $CHCl_3$)	103		[79]
5-Thio-D-glucal				[80]
5-Thio-L-idose /Penta-Ac	+41° (c 2.0 $CHCl_3$)	90–92		[81]
5-Thio-D-ribose	+127° (c 0.6 H_2O)	160		[82]
5-Thio-D-ribose/Tetra-Ac	−61° (c 1.0.MeOH)	122–123		[83]
5-Thio-D-mannose				[84]
5-Thio-D-xylose	+202° (c 2.0 H_2O)	122–123		[85]
4-Thio-D-ribopyranose	+188° (c 1.56 H_2O)			[86]
4-Thio-D-ribose/Tetra-Ac-α	+123° (c 2.0 $CHCl_3$)			[87]
4-Thio-D-ribose/Tetra-Ac-β	−102° (c 2.0 $CHCl_3$)	64–66		[87]
1-Thio-L-arabinose/Na salt/Hyd.	+69° (c 1.56 H_2O)	155		[88]
1-Thio-L-arabinose/Tetra-Ac-β	+42° (c $CHCl_3$)			[89]
1-Thio-D-glucose/Tetra-Ac-α	+5.8° (c 1.56 H_2O)	115–17		[90]
1-Thio-D-glucose/Tetra-Ac-β	−2.1° (c 2.8 $C_2H_2Cl_4$)	118		[91]
1-Thio-D-galactose/Tetra-Ac	+3° (c $CHCl_3$)	108		[92]
1-Thio-D-mannose/Tetra-Ac-β	−29° (c $CHCl_3$)	130–131		[93]
1-Thio-D-ribose	+11° (c $CHCl_3$)			[94]
1-Thio-D-xylose	−20° (c 1.56 H_2O)			[95]
1-Thio-D-xylose/Tri-Ac	−17° (c 2.0 $CHCl_3$)	123–129		[96]

*Compiled from Carbohydrates (1987) Collins PM (ed) Chapman and Hall Chemistry and original cited references

5.7 Carba-Sugars

The important class of carba-sugars, known also as "pseudo-sugars", is attracting immense attention as newly rediscovered components of many natural products bearing traditional sugar moieties and carbocyclic units as well. Classical examples of such derivatives are aminocyclitol (aminoglycoside) antibiotics, such as streptomycins, neomycins, kanamycins, spectinomycins, gentamycins, and fortimicins. Moreover, the family of validamycins [97] consists of eight validamycins (A–H) and three validoxyamines have so far been isolated from the fermentation broth and fully characterized [98].

As an example in the 5a-carba-aldohexopyranose family, there are 32 stereoisomers theoretically possible, and all 16 of the predicted racemic forms have been prepared, as well as 15 of the enantiomers. The chemical character of this class of derivatives relies completely on the presence of functional groups, as well as the ring system. One of the specific bicyclic

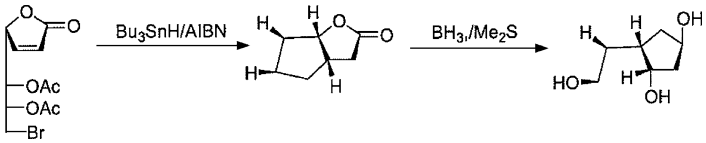
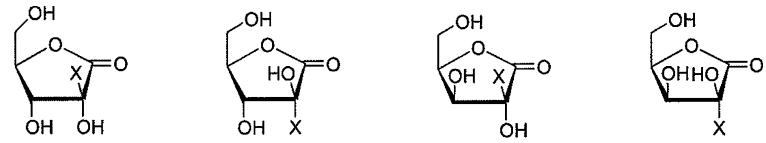


Figure 11
Bicyclic ring system of carba-sugars [100]



X = -CH_2OH

Figure 12
Branched pentenolactones [108]

ring systems of carba-sugars is easily accessible from unsaturated γ-lactones [99] constituting a new entry to carba-furanoses [100]. Ogawa and coworkers [101,102,103,104,105,106] have published numerous accounts on the chemical reactivity of the above fascinating class of compounds.

5.8 Branched-Chain Sugars

Within this extremely important group of monosaccharides [9,107] the typical representative bears a branched functional group at any position except the anomeric carbon, but most commonly at the C-2 and C-3, as in natural L-streptose, D-hamamelose, or L-apiose.

An interesting chemical characteristic of the hydroxymethylene group at C-2 of various branched penteno-γ-lactones [108] can be observed during the oxidation of vicinal diols with periodate which proceeds with extremely high selectivity with the relative rates 100:1:3:100. Thus, the isomers with *cis*- and *trans*-2,3-diols can be readily identified.

This is based on the specific chemical character of vic-*cis* diols, which reduced periodate faster than those isomers with *trans*-diol stereogeometry.

Several signals in the ^{13}C NMR spectra of dehydro branched-chain sugars can be of value in stereochemical assignments [109] and indeed they are real diagnostic factors in determining the position of the branched hydroxymethyl group. For example, the α-carbon atom of a branch group which is axially oriented on a pyranoid ring resonates upfield relative to the same carbon atom in the isomer in which the branch chain is equatorially disposed.

6 Conclusion

This short compilation of randomly selected information on occurrence, significance and properties of various classes of monosaccharides will provide additional new data essential to glycoscience field.

References

1. Lederkremer RM, Gallo-Rodriguez C (2004) Adv Carb Chem Biochem 59:9; Èerný M (1994) 1,6:2,3- and 1,6:3,4-Dianhydro-β-D-hexopyranoses. Synthesis and Preparative Applications. In: Witczak ZJ (ed) Levoglucosenone and Levoglucosans: Chemistry and Application. ATL Press Science Publishers, Mt. Prospect, IL
2. Banoub J, Bullangner P, Lafont D (1992) Chem Rev 92:1167
3. Thiem J, Klafke W (1990) Top Curr Chem 154:285
4. Yokoyama M, Nomura M, Togo H, Seki H (1996) J Chem Soc PerkinTrans I 2145
5. Baer H, Urbas L (1970) Activating and directing effect of the nitro group in aliphatic and acyclic systems. In: Feuer H (ed) The Chemistry of the Nitro and Nitroso Group. Wiley Interscience, New York, p 73
6. Driguez H (1997) Top Curr Chem 187:85
7. Witczak ZJ (1999) Curr Med Chem 5:125; Robina I, Vogel P, Witczak ZJ (2002) Curr Org Chem 5:1177; Witczak ZJ, Culhane JM (2005) Appl Microbiol Biotechnol 69:237
8. Tsuchiya T (1990) Adv Carb Chem Biochem 48:91
9. Yoshimura J (1984) Adv Carb Chem Biochem 42:69
10. Braun M, Mahler H (1989) Angew Chem Int Ed Engl 28:896
11. Tajmir-Riahi H-A (1984) Carbohydr Res 125:13
12. Johnson WC (1987) Adv Carb Chem Biochem 45:73
13. Nakanishi K (1972) Acc Chem Res 5:257
14. Lonngreen J, Svensson S (1974) Adv Carb Chem Biochem 24:41
15. Stevens ES, Sathyanaryana BK (1987) Carbohydr Res 166:181
16. Laine RA (1992) Mass spectrometry of carbohydrates. In: Allen HJ, Kisailus EC (eds) Glycononjugate Composition, Structure and Function. Marcel Dekker, New York, p 103
17. Laine RA (1990) Glycoconjugates, overview and strategy. In: McCloskey JA (ed) Methods in Enzymology. Academic Press, New York, vol 193, p 539
18. Kochetkov NK, Chizov OS (1966) Adv Carb Chem Biochem 21:39
19. Whitehouse CM, Dryer RN, Yamashita M, Fenn JB (1985) Anal Chem 57:675
20. Schulten HR (1979) Int J Mass Spectrom Ion Phys 32:97
21. Dell A (1987) Adv Carb Chem Biochem 45:19
22. Banboub J, Michon F, Shaw DH, Roy R (1984) Carbohydr Res 128:203
23. Bock K, Pedersen C, Pedersen H (1984) Adv Carbohydr Chem Biochem 42:193
24. Serianni AS (1999) Carbohydrate structure, conformation, and reactivity: NMR studies with stable isotopes. In: Hecht SM (ed) Bioorganic Chemistry Carbohydrates. Oxford University Press, New York, p 244
25. Tvaroska I, Travel FR (1996) Adv Carbohydr Chem Biochem 51:15
26. Richardson AC, McLauchlan (1962) J Chem Soc 2499; Baker SA, Homer J, Keith NC, Thomas LF (1963) J Chem Soc 1538
27. Neszmelyi A, Liptak A, Nanasi P (1977) Carbohydr Res 58:C7
28. Grindley TB, Glulasekharem V (1979) Carbohydr Res 74:7
29. Lichtenthaler FW (1992) Modern Synthetic Methods, vol 6. VCHA, Basel, p 273
30. Èerný M, Stanek J (1977) Adv Carbohydr Chem Biochem 34:23
31. Williams NR (1970) Adv Carbohydr Chem Biochem 25:109
32. Lauer G, Oberdorfer F (1993) Angew Chem Int Ed Engl 32:272
33. Eby R, Schuerch C (1982) Carbohydr Res 100:C41
34. Eby R, Webster KT, Schuerch C (1984) Carbohydr Res 129:111
35. Klemer A, Kohla M (1984) Liebigs Ann Chem 1662
36. Klemer A, Kohla M (1986) Liebigs Ann Chem 967
37. Prystas M, Gustafsson H, Sorm F (1971) Collect Czech Chem Commun 36:1487
38. Van den Broek LAGM, Vermaas DJ, Heskamp BM, van Boeckel CAA, Tan CAA, Bolscher JGGM, Ploegh HL, van Kemenade FJ, de Goede REY, Miedema F (1993) Recl Trav Chim Pays-Bas 112:82
39. Van den Broek LAGM (1997) Aza-sugars: chemistry and their biological activity as potential anti-HIV drugs. In: Witczak ZJ (ed) Carbohydrate in Drug Design. Marcel Dekker, New York, p 471

40. Pelyvas IF, Monneret C, Herczegh P (1988) Synthetic Aspects of Aminodeoxy Sugar of Antibiotics. Springer, Berlin Heidelberg New York
41. Kirshning A, Bechthold AFW, Rohr J (1977) Top Curr Chem 188:1
42. Giese B, Ruckert B, Groninger KS, Muhn R, Lindner HJ (1988) Liebigs Ann Chem 997
43. Blattner R, Furneaux RH, Mason JM, Tyler PC (1994) Synthesis of the herbicide 1,6-anhydro-4-*O*-benzyl-3-deoxy-2-*O*-methyl-β-D-ribohexopyranose and analogues from Levoglucosenone and Levoglucosan. In: Witczak ZJ (ed) Levoglucosenone and Levoglucosans: Chemistry and Application. ATL Press Science Publishers, Mt. Prospect, IL, p 43
44. Shafizadeh F (1958) Adv Carb Chem Biochem 13:9
45. Ball DH, Parish FW (1969) Adv Carb Chem Biochem 26:139
46. Kochetkow NK, Sviridov AF, Ermolenko MS, Yashunsky DV (1984) Tetrahedron Lett 25:1605
47. Baillargeon DJ, Reddy GS (1986) Carbohydr Res 154:275
48. Csuk R, Glanzer BC (1988) Adv Carb Chem Biochem 46:73
49. Penglis AAE (1981) Adv Carb Chem Biochem 38:195
50. Dax K, Albert M, Ortner J, Paul BJ (1999) Curr Org Chem 3:287
51. Nicolaou KC, Bockovich NJ (1999) Chemical synthesis of complex carbohydrates. In: Hecht SM (ed) Bioorganic Chemistry Carbohydrates. Oxford University Press, New York, p 134
52. Toshima K, Tatsuta K (1993) Chem Rev 93:1503
53. Baumberger F, Vasella A (1983) Helv Chim Acta 66:2210
54. Aebisher B, Meuvly R, Vasella A (1984) Helv Chim Acta 67:2236
55. Pham-Huu D-P, Petrusova M, BeMiller J, Petrus L (1998) Synlett 1319
56. Pham-Huu D-P, Petrusova M, BeMiller J, Petrus L (1999) Tetrahedron Lett 40:3053
57. Defaye J, Gelas J (1995) In: Atta-ur-Rahman (ed) Studies in Natural Products Chemistry, vol 8E. Elsevier, Amsterdam, pp 315–357
58. Geyer A, Hummel G, Eisele T, Reinehardt S, Schmidt RR (1996) Chem Eur J 2:981
59. Espinosa JF, Cañada FJ, Asensio JL, Dietrich H, Martin-Lomas M, Schmidt RR, Jiménez-Barbero J (1996) Angew Chem Int Ed Engl 35:303
60. Izumi M, Tsuruta O, Hashimoto H, Yazawa S (1996) Tetrahedron Lett 37:1809
61. Mehta S, Pinto BM (1992) Tetrahedron Lett 33:7675
62. Garreg P (1996) Adv Carb Chem Biochem 52:179
63. Norberg T (1996) Glycosylation properties and reactivity of thioglycosides, sulfoxides and other *S*-glycosides: Current scope and future prospects. In: Khan SH, O'Neill RA (eds) Modern Methods in Carbohydrate Synthesis. Harwood Academic Publishers, Amsterdam, p 82
64. Riedner J, Robina I, Fernandez-Bolanos JG, Gomez-Bujedo S, Fuentes J (1999) Tetrahedron: Asymmetry 10:3391
65. Witczak ZJ, Sun J, Mielguj R (1995) Bioorg Med Chem Lett 5:2169
66. Witczak ZJ, Chhabra R, Chen H, Xie X-Q (1997) Carbohydr Res 301:167
67. Becker T, Thimm B, Thiem J (1996) J Carbohydr Chem 15:1179
68. Witczak ZJ, Chen H, Kaplon P (2000) Tetrahedron: Asymmetry 11:519
69. Lambert JB, Wharry SM (1983) Carbohydr Res 115:33
70. Berman E (1983) Carbohydr Res 116:144
71. Whistler Rl, Rowell RM (1966) J Org Chem 31:816
72. Hall LD (1961) J Chem Soc 1537
73. Emery F, Vogel P (1994) J Carb Chem 13:555
74. Izumi M, Tsuruta O, Hashimoto H (1996) Carbohydr Res 280:287
75. Takahashi S, Kuzuara H (1992) Chem Lett 21; Hashimoto H, Izumi H (1992) Chem Lett 25
76. Hashimoto H, Fujimori NM, Yuasa H (1990) J Carb Chem 9:683
77. Shin JE, Perlin AS (1979) Carbohydr Res 76:165; (1980) ibid 84:315
78. Rowell RM, Whistler RL (1966) J Org Chem 31:1514
79. Driguez H, Henrisatt (1981) Tetrahedron Lett 22:5061
80. Korytnik W, Angelino N, Dodson-Simmon O, Hanchak M, Madson M, Valentekovic-Horvath M (1983) Carbohydr Res 113:166
81. Adley TJ, Owen LN (1961) Proc Chem Soc 418; Schwartz RL, Feather MS, Ingles DL (1962) J Am Chem Soc 84:122
82. Clayton CJ, Hughes NA Carbohydr Res 4:32
83. Hughes NA, Clayton CJ (1975) J Chem Soc Chem Commun 294
84. Yausa H, Izukawa Y, Hashimoto H (1989) J Carb Chem 8:753
85. Anisuzzaman AKM, Whistler RL (1977) Carbohydr Res 55:205

86. Bobek M, Whistler RL (1972) Methods Carbohydr Chem 6:292
87. Uenishi J, Motoyama M, Takahashi K (1994) Tetrahedron Asymmetry 5:101
88. Gehrke M (1931) Ber 64:2696
89. Zissis E (1966) Carbohydr Res 2:461
90. Fried J (1949) J Am Chem Soc 71:140
91. Horton D (1963) Methods Carbohydr Chem 2:433
92. Černý M (1963) Monatshefte für Chem 94: 290
93. Fragala J, Černý M, Stanek J (1975) Collect Czech Chem Commun 40:1411
94. Tejima S (1964) Chem Pharm Bull 12:528
95. Horton D (1963) Methods Carbohydr Chem 2:368
96. Stanek, J, Černý M (1975) Collect Czech Chem Commun 30:297
97. Suami T (1990) Adv Carbohydr Chem Biochem 48:21
98. Kameda Y, Asano N, Yoshikawa M, Takeuchi M, Yamaguchi T, Matsui K, Horii S, Fukase H (1984) J Antibiotic 37:1301
99. Horneman AM, Lundt I, Sotofte I (1995) Synlett 918
100. Lundt I (1997) Top Curr Chem 187:144
101. Ogawa S, Sasaki S, Furuya T (1998) Eur J Org Chem 1099
102. Ogawa S (1997) Carba-Sugars as indispensable units of aminocyclitol antibiotics and related biologically active compounds. In: Witczak ZJ (ed) Carbohydrates in Drug Design. Marcel Dekker, New York, p 433
103. Tsunoda H, Sasaki S, Furuya T, Ogawa S (1996) Liebigs Ann 159
104. Tsunoda H, Inokuchi K, Yamagishi S, Ogawa S (1995) Liebigs Ann 279
105. Ogawa S, Sasaki S, Tsunoda H (1993) Chem Lett 1587
106. Ogawa S, Aso D (1993) Carbohydr Res 250:177
107. Shafizadeh F (1956) Adv Carbohydr Chem 48(11):263
108. Roberts SM, Shoberu KA (1992) J Chem Soc Perkin Trans 1:2625
109. Aylett BJ, Harris MM (eds) (1969) Progress in Stereochemistry 4. Butterworth, London, p 43

4.2 Monosaccharides and Polyols in Foods

Robert B. Friedman
Friedmann Associates, 6654 North Mozart Street, Chicago, IL 60645, USA
bobbf@juno.com

Abstract

Monosaccharides and polyol monosaccharides are, generally, naturally occurring materials. They appear both in simple food products and in complex manufactured food systems. Their structures and functions are reviewed here, and their roles in food systems are discussed.

Keywords

Monosaccharides; Reducing sugar; Polyol; Food products

Abbreviations

FDA Food and Drug Administration
GI glycemic index
GRAS generally recognized as safe
USDA United States Department of Agriculture

1 Introduction: Uses of Monosaccharides and Polyols in Food Products

Monosaccharide sugars and polyols are carbohydrate materials of considerable commercial value and interest to the commercial food sector. They all occur in nature, and function effectively as components of substances that are used as food nutrients. They function, however,

In: *Glycoscience*. Fraser-Reid B, Tatsuta K, Thiem J (eds)
Chapter-DOI 10-1007/978-3-540-30429-6_19:

equally as well as useful ingredients providing specific functional properties in manufactured food products. This chapter will examine the roles of monosaccharides and polyols from both viewpoints: nutrients and functional ingredients. Important features of these substances will be explored as they occur in nature as nutrients, and also as they are applied in the contemporary food manufacturing scene, in the role of effective food ingredients.

The main focus in this chapter will be on monosaccharide sugars and polyols that are used commercially, or, alternatively, that have commercial opportunity or potential. Furthermore, the focus will be exclusively on monosaccharide materials, even though other low molecular weight carbohydrates, such as disaccharides, might have similar or even superior functionalities.

In simpler societies, individuals produced, gathered, or caught the raw ingredients or materials for their own food production and consumption. Preparation and preservation of the foods produced were all performed by the producers themselves. The consumed foods were physically and technically not too distant from the raw ingredients. Sweetening was performed using available materials, such as honey. Native Americans, for example, were known to have chewed some of the corn meal used in corn bread. This provided sweetness in the finished product through the action of salivary amylase activity on the starch in the corn meal, when the chewed materials were added to the dough. This indeed was an early application of biotechnology to food systems! While the role of subsistence farming in food production was significant, the importance of individual ingredients was not so great. Food was food; little regard was placed on nutritive values. As human societies became more complex, however, it became necessary to ship ingredients longer distances before they could be prepared as foods. Urbanization made this process both important and necessary. Generally, the consumed foods were still produced by the end-users on a daily basis. Quality still remained a major concern of the end users.

Our contemporary situation finds us in a state of even greater complexity. Obesity is rampant on a global basis, and considerable lip service is paid to healthy eating. The preparation of foods, however, has been given over to food producers and preparers who are miles, even continents, away. Convenience now reigns! People are so involved in their everyday tasks and duties that there is little time remaining for food preparation, or even the reading of food package ingredient labels, for that matter. Devices such as the United States Department of Agriculture (USDA) food pyramid [1] are available and useful, but few consumers have the time to learn and understand its message.

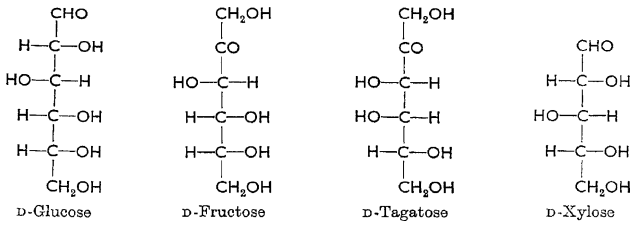
Responsibility to develop appropriate food products with balanced functionality and nutritional value has shifted, then, from domestic kitchens to the development laboratories of food manufacturing companies. A purpose of this chapter, as such, is to provide food product developers with a bit more depth of information to enhance their perception of the appropriate use and application of monosaccharide sugars and polyols in the development of new food products. Applications should reflect the chemistry of these materials, as well as the chemical composition of the other ingredients in the product and their interactions [2]. Included in this chapter is information on applications and nutrition, as well as background on regulatory requirements and use. Illustrations of structures have been minimized and simplified in order to focus on more practical aspects of the discussion.

2 Review of Monosaccharides and Polyols that Are Currently in Use by the Food Industry

A variety of monosaccharide sugars and polyols is employed in the food industry. The use of these materials reflects availability, costs, and properties. A brief review of monosaccharide structure and nomenclature will follow, though the reader is directed to other sources for a greater depth of information [e. g., 3–5].

The name "carbohydrate" arose originally from the belief that sugars were essentially hydrates of carbon. Indeed, structurally, carbohydrates can be viewed as being derivatives of glyceraldehyde. Adding one –CHOH group to the basic three carbon backbone of glyceraldehyde will yield the next higher analogue of the carbohydrate. Simple carbohydrates are classified according to the number of carbons in the chain. Simple sugars containing five carbons are called pentoses while those containing six carbons are called hexoses. Single simple sugar molecules are called monosaccharides. When two monosaccharides are joined chemically, the resulting material is called a disaccharide. When a large number of monosaccharides are joined, the end result is a polysaccharide. All these categories are found in nature. Glucose, also called dextrose, is a critically important monosaccharide in nature; sucrose, is a disaccharide comprised of glucose and fructose (*Scheme 1*). Starch is an example of an important polysaccharide comprised of long chains of glucose molecules linked end-to-end. It is a biopolymer whose sole monomeric component is the monosaccharide glucose. Complete depolymerization of starch results in a 100% glucose syrup. Partial depolymerization results in a more complex mixture of product materials [6].

The characteristic chemical functional group of monosaccharides is the hydroxyl group. Simple reducing sugars also contain a carbonyl group. When the carbonyl is on the terminal carbon of the monosaccharide chain, the carbonyl group is an aldehyde, and the saccharide is classified as an aldose. Glucose is an example of an aldose. When the carbonyl group is on one of the carbons in the chain, the carbonyl is a ketone, and the simple sugar is classified as a ketose. Fructose is an example of a ketose. Carbonyls are more reactive than hydroxyls, and many of the reactions that simple sugars undergo reflect the reactivity of these carbonyls. For example, the most common bond joining the disaccharides is an acetal, linking the aldehyde of one sugar with a hydroxyl of an adjacent sugar. Such bonds are called "glycosides".



Scheme 1
Monosaccharide reducing sugars (from *left* to *right*): Glucose (an aldohexose), Fructose (a ketohexose), Tagatose (a ketohexose), Xylose (an aldopentose). Modified from [84]

Carbonyls can be oxidized to the next level of oxidation, as well, to yield carboxylic acids. Carboxylic acids derived from carbohydrates, gluconic acid for example, are used to some extent as acidulants by the food industry. Finally, the carbonyls can be reduced by adding the elements of hydrogen across the carbonyl to yield a hydroxyl group. This chemical reduction process is employed both in nature and in industry to yield the class of materials called "polyols". These materials have no carbonyl functional groups remaining. An example of such a compound is sorbitol; it is the product of chemical reduction of glucose. Polyols are metabolized more slowly by mammalian systems. This metabolic rate is used to advantage in special applications and its utility is discussed later in this chapter. Furthermore, as a result of the current commercial importance of polyol monosaccharides, they will be discussed separately from reducing sugars. Other chemical functional groups on carbohydrate molecules are found in nature, such as amine groups that replace hydroxyls in the carbohydrate category called amino sugars. These compounds are important components of structural carbohydrates, such as connective tissue polysaccharides in mammals, or exoskeleton polysaccharides of insects or shellfish. Amino sugars will be discussed briefly below, as well.

2.1 Reducing Sugar Monosaccharides of Commercial Importance

Nutritive sweeteners have been an important component of foods since early times. They provide nutrition, bulk, sweetness, and moisture control. Furthermore, they are important elements in the flavor and color profiles of different foods. Maillard reaction products [7,8] are excellent examples of such functional properties. One of the oldest examples of a monosaccharide-based sweetener system is honey. Honey is mainly comprised of the monosaccharide reducing sugars glucose and fructose [9]. In fact, a modern challenge presented to the honey industry is the adulteration of honey using high fructose corn syrup. Sophisticated analytical methods are required to verify the purity of honey [10]. It is interesting to note that honey does have a reduced glycemic index (GI) compared to sucrose. Little difference in glycemic index, however, has been found between different sources of honey [11]. The glycemic index will also be discussed later in this chapter, in the section on polyols. Invert syrup is another example of monosaccharide-based sweeteners heavily used in the food industry [12]. Invert sugar is produced by the hydrolysis of the disaccharide sucrose to its component monosaccharides, glucose, and fructose. In the process, the optical rotation of the syrup is "inverted", hence the name "invert syrup". The monosaccharide reducing sugars have a preferred sweetness level and profile. The high level of digestibility of these materials also makes them useful in a very broad range of applications, from infant formulas to nutritional supplements for the elderly. Laxation effects are not a concern for most of the common reducing sugars.

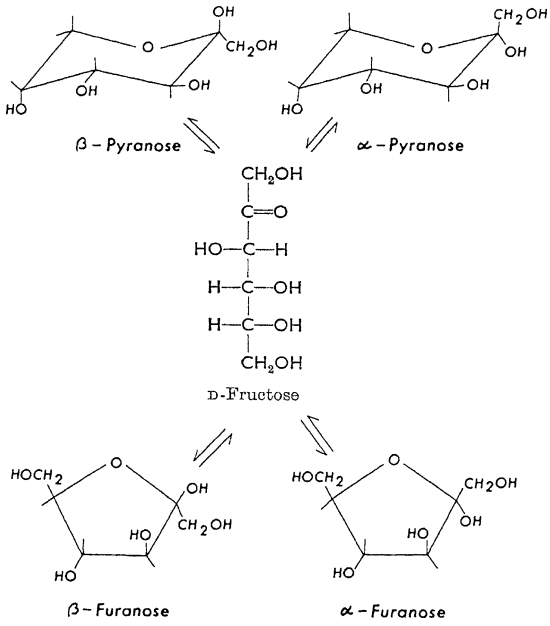
The great utility of nutritive sweeteners, especially reducing sugars, makes it difficult to find suitable replacements. The drive to replace these ingredients comes from the demand to reduce the caloric content of foods and also to avoid the cariogenicity of some carbohydrate-containing products. Glucose, and materials that release glucose on contact through the action of glucan sucrase enzymes of oral bacteria, such as sucrose, support the formation of dental plaque and dental caries [13]. These concerns have brought into focus the commercial use of polyols, or sugar alcohols, and also provided the opportunity to explore the use of less common

reducing sugars that might be metabolized in man using less calorigenic pathways. Examples of such materials and their applications will be presented later.

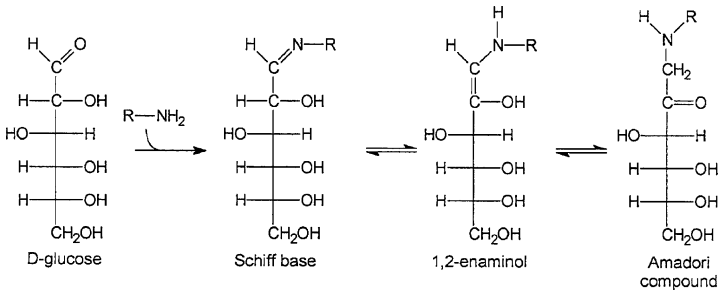
Dextrose (or D-glucose) is probably one of the most important of the monosaccharides. It not only serves as one of the principal monosaccharide building blocks of structural and energy storage polymeric carbohydrates in nature, but it is also one of the most important means for providing and transporting energy in living systems. It is not surprising; therefore, that glucose is also one of the most commonly used monosaccharides in the food industry.

D-Glucose is an aldohexose (*Scheme 1a*) containing six carbons, with a molecular weight of 180.2. Its solubility in water at 25 °C is 100 grams per 100 ml of water. The sweetness level of glucose is about 64% that of sucrose, and it is totally digested. The name glucose was coined from the Greek word "glycos" meaning sweet. In solution, reducing sugars, including glucose and fructose, exist in equilibrium between several isomeric forms (*Scheme 2*). Each form rotates polarized light differently, and the change in the optical rotation of a sugar solution until equilibrium is achieved is called mutarotation. The name "dextrose" arose due to its ability to rotate polarized light to the right (it is "dextrorotatory").

Glucose is generally produced by the complete hydrolysis (i. e., depolymerization of the biopolymer) of starch to its monomeric component, D-glucose. It is usually sold commercially as a crystalline product, or as a glassy spray-dried solid. In crystalline form the product is commonly called dextrose. It is commercially used as a sweetener and bulking agent, as well as a crystalline component of such products as coated candy. Glucose crystallizes most readily as the *alpha* D-glucopyranose monohydrate form. Using unusual conditions, processes have



Scheme 2
Tautomeric forms of fructose (modified from [85])



Scheme 3
Schiff base formation (modified from [8])

been described to form crystalline products high in anhydrous *beta*-D-glucose that is claimed to dissolve more rapidly in water [14] than the *alpha*-D-glucose monohydrate. The melting point of *alpha*-D-glucose monohydrate is 90–95 °C. The monohydrate form of glucose is less expensive than the anhydrous form, and this form of glucose, generally, is also more water-soluble than the anhydrous form. Glucose is, of course, a naturally occurring substance. It is generally recognized as a safe (GRAS), and its caloric value is about 3.5–4 calories per gram. Application areas include beverages, confections, and chewing gum. Its physical bulk properties are similar to those of sucrose. When sucrose was not readily available commercially, glucose was used to make crystallized candies. The benefit that crystalline glucose has over sucrose is that sucrose can tend to scorch and carmelize before melting, and the presence of acid causes inversion. This does not occur with crystalline glucose [15]. Glucose also has been claimed to be a superior panning sugar for color, smoothness, and cost [16,17].

As mentioned earlier, glucose is an aldose. As such, it (as well as other reducing sugars) can undergo reactions typical of aldehydes when held under favorable conditions for reaction, for example, in aqueous solution. Aldehydes can react with amines, for example, to form imines, or Schiff bases (*Scheme 3*). These compounds are important components of the Maillard reaction in that they provide browning and flavor components [7,8]. Sometimes, however, they also provide unwanted browning, as in some fruit products, and this reactivity must be arrested to extend shelf life. Similarly, when dipeptide high intensity sweeteners, such as aspartame, are employed in the presence of glucose, the potential for Maillard reaction of glucose with the aspartame is high, resulting in the loss of available sweetness from the aspartame [18,19] accompanied by the generation of undesired color in the product.

Glucose is an essential metabolic compound in living things. It is hard to visualize a more important chemical substance than glucose, except, perhaps, for water. Glucose is critical for the transmission of energy and precursors of structural components manufactured at the cellular level throughout the body via the blood stream. The body has biochemical mechanisms for the maintenance of homeostasis, though these mechanisms can be damaged by abnormally high levels of glucose [20]. Children, in particular, who are challenged by high levels of sweets, can be compromised metabolically as a result, during maturation.

d-***Fructose*** is a hexose of increasing importance. It is produced commercially by enzyme-catalyzed epimerization of dextrose (glucose). For beverage applications it is not separated from

glucose, but is sold as a syrup that approximates the composition of invert sugar syrups. These syrups are called "high fructose syrups" in the US. Caution must be used to insure that one is referring to the correct high fructose syrup when specifying syrups, since fructose syrups are marketed as 42, 55, and even 90% fructose. The remaining percentage, of course, is glucose. Fructose can also be crystallized from essentially "pure" fructose syrups. Raw materials for the epimerization reaction include depolymerized starches from various sources [21], and sucrose [22]. Fructose has also been reported to be produced by hydrolysis of naturally occurring fructose polymers, fructosans, such as inulin syrups [23]. In the older chemical literature, fructose is also called "levulose" because it rotates polarized light to the left (it is "levorotatory"). Fructose is reactive, and browns more rapidly in baked goods.

Fructose is a ketohexose (*Scheme 1b*) containing six carbons, with a molecular weight of 180.2. Its solubility in water at 25 °C is 400 grams per 100 ml of water. Fructose is hygroscopic in nature, and even the crystalline material requires special care to minimize moisture uptake. The sweetness level of fructose is about 117% that of sucrose, and it is believed to be totally digested. Fructose is, of course, a naturally occurring substance. It is GRAS, and its caloric value is about 4 calories per gram. One of the remarkable features of pure fructose is that it is not metabolized as rapidly as is glucose. Fructose has been recognized as having a relatively low glycemic index [24]. This means that crystalline fructose could be employed to provide sweetness without causing the spike in blood glucose with its resulting elevated insulin level. The sweetness quality of fructose is considered by many to be superior to that of most high intensity sweeteners, and its elevated sweetness level, compared to sucrose, also means that it can be used at a lower concentration. Furthermore, it has been reported in the literature that fructose can develop synergistic sweetness with high intensity sweeteners, such as saccharin [25,26]. Developers, however, must be sensitive to the more rapid browning potential of fructose in their formulas, as well as its higher hygroscopicity. This water holding ability actually can be put to good use, however. For example, frozen desserts that can be brought directly to the table from the freezer have been described, based on the water holding properties of fructose at low temperature [27,28]. In a similar vein, fructose has been used to hold and distribute water in an emulsified maltodextrin for improving the quality of baked goods [29]. Application areas for fructose include beverages, frozen desserts, chewing gum, and confections. Its physical bulk properties are similar to those of sucrose.

Fructose is not without health concerns. Over the years, questions have been raised over the effects that high doses of fructose might have on various metabolic processes in humans [30]. These include copper and triglyceride metabolism [31,32], as well as its intestinal absorption [33]. These problems have not been manifest in the general population in the last few decades of use, however [34], though on occasion negative articles do appear in the popular press. The reader is directed to the website of such organizations as the Corn Refiners Association for further discussion of the effective utilization of fructose-based sweeteners [35].

d-***Tagatose*** is an example of a sugar that has been proposed recently to be an alternative sugar to glucose due to its different metabolic properties [36]. It is also a hexose reducing sugar. It is slowly metabolized, however, similar to polyols, and has a low caloric content and a slow glycemic response. This could be beneficial to diabetics [37] or individuals concerned about glycemic response. However, its low laxation threshold is a matter of concern, as well as its commercial availability. D-Tagatose is currently manufactured commercially by chemical isomerization of D-galactose. Galactose, in turn, is presently produced by hydrolysis of lactose.

Lactose is a disaccharide comprised of glucose and galactose, and is a by-product of the dairy industry. This chemical conversion-based process has placed constraints on the availability of D-tagatose. There has been interest over the years to identify a convenient commercial source of tagatose through fermentation [38]. D-Tagatose is a ketohexose (*Scheme 1c*) containing six carbons, with a molecular weight of 180.2. The sweetness level of tagatose is about 90% that of sucrose, but its laxation threshold is comparable to that of the polyol mannitol, about 20–30 grams per day [39]. D-Tagatose is a naturally occurring substance. It is GRAS, and its caloric value is 1.5 calories per gram. Application areas include confections, chewing gum [40], and beverages. Its use has appeared in the latter area in connection with erythritol in icy beverages [41,42]. Its physical bulk properties are similar to those of sucrose.

Over the years other carbohydrates have been suggested to replace nutritive monosaccharides in food systems. For example, L-sugars have been proposed [43,44] as useful low calorie and slowly metabolized saccharides. Regulatory hurdles, including demonstrations of safety, generally prove to be formidable barriers to the commercialization of such new materials. L-sugars have not yet been commercialized.

2.2 Non-Reducing Sugar (Polyol) Monosaccharides of Commercial Importance

As mentioned earlier, polyols are polyhydric alcohols derived through chemical reduction (e. g., hydrogenation) of reducing sugars. Most of the commercially important polyols are produced by chemical hydrogenation of their corresponding reducing sugars. For example, sorbitol is produced by hydrogenation of dextrose (glucose), mannitol is derived through fructose hydrogenation, while xylitol is produced by hydrogenation of xylose. Erythritol, however, has always been produced commercially by fermentation [45]. Recent publications [46] suggest a breakthrough in the fermentative production of xylitol, starting with glucose, as well. Fermentative methods to produce mannitol from fructose have been published [47,48], but a commercial process has only recently been reported [49]. The potential for fermentative production generates anticipation that these products will be commercially available at reduced costs, but with greater quantity. Enzyme production is especially useful since enzymes can be more specific in their production, resulting in cleaner and more efficient reaction products. An important concern in this regard, however, is that the source of enzymes used to carry out the fermentative transformation must be from organisms that are acceptable in any country that wishes to use that product.

As pointed out earlier, polyols are metabolized much more slowly than reducing sugars. Furthermore, the caloric content of the individual polyols can be less, significantly, than those of corresponding reducing sugars. Consequently, polyols are useful replacements for more rapidly metabolized carbohydrates. They are most often used in confectionery, food, and pharmaceutical applications, as well as industrial uses [50,51].

The benefits of polyols in food and confectionery products include reduced caloric content, reduced glycemic response, and reduced cariogenicity. The dental literature is replete with reports that demonstrate unique dental benefits for specific polyols, such as xylitol [52,53], while others [54,55] show that the genuine dental benefits of some polyols are similar. Polyols generally are not fermented by oral bacteria, so that acid production is minimized while

plaque production is not supported. Xylitol has been reported actually to reduce dental caries, and, consequently, is in great demand for chewing gum and confectionery applications. In any event, polyols are useful substances for advancing dental health, particularly in confectionery [56].

Since polyols do not impact blood glucose levels to any significant degree, they are generally useful to people with diabetes, or individuals seeking food ingredients with low glycemic index values. The concept of the glycemic index was originally introduced by Jenkins [24]. It is based on actual blood glucose levels compared to equivalent carbohydrate intake. Since publication of the original article, the glycemic index has become widely accepted, even to the extent that it is now widely reported in the medical literature [57]. Tables of GI values for different food ingredients have been published [58]; polyols are featured in those tables. It is interesting to note that early on, fructose was recognized as having a low glycemic index [24], as well. The reader is directed to the website of such organizations as the Calorie Control Council for additional information concerning polyols and their functionalities [59].

Polyols can provide sweetness, texture, and bulk in many sugar-free applications. Furthermore, some crystalline forms of certain polyols can cause a cooling effect through their negative energies of solution. Polyols blend well with other sweeteners. This is particularly important, since their sweetness potential might be less than that of the reducing sugars that they could be replacing. The lack of a free aldehyde or ketone group means that polyols will also not participate in Maillard-type browning reactions. While dextrose reacts with aspartame, for example, through reaction between its aldehyde group and the amine of aspartame [18,19], sorbitol is quite stable with that high intensity sweetener. A browning agent might be needed, however, in baked goods to provide an even brown color in the finished product. Polyols can serve as humectants, cryoprotectants and aids to control freezing point depression.

One major challenge with polyols is brought about by the very beneficial properties that make their use attractive. Their slow metabolism results, generally, in a larger concentration of these materials passing into the large intestine. Here two effects are manifest. One is the attack by bacteria of the intestinal flora on the larger concentration of carbohydrate materials for which the bacteria possess the metabolic machinery to digest. The result is a dramatic increase in production of intestinal carbon dioxide and other waste gases by the intestinal flora. The other effect is brought about by the influx of undigested material. The normal response of the large intestine is to increase the flow of liquid into the intestine to control its osmotic balance [60]. The result of both phenomena is a laxation effect that is clearly not desirable. Governments have set limits on the amounts of polyols that can be used in terms of maximal limits. Furthermore, labels warning of the presence of polyols are required in some countries. A comparison of laxation thresholds of various polyols is presented in ❯ *Table 1*. ❯ *Table 2* shows the caloric content of selected polyols, while ❯ *Table 3* lists relative sweetness levels of some polyols.

Sorbitol is produced commercially by catalytic hydrogenation of D-glucose (dextrose). It is also known, particularly in older literature, as glucitol in recognition of its source. The starting material is a highly converted corn or glucose syrup that is catalytically hydrogenated under pressure at elevated temperatures. Sorbitol is a linear polyhydric alcohol (❯ *Scheme 4a*) containing six carbons, a hexitol, with a molecular weight of 182. Its solubility in water at 25 °C is 235 grams per 100 ml of water. This permits the formation of high-solids pastes that are stable under a variety of conditions. For example, sorbitol is used heavily in toothpaste to help main-

Table 1
Relative laxation effect of selected monosaccharide polyols

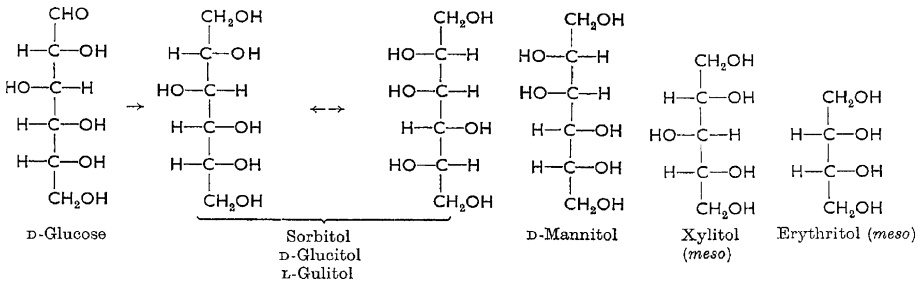
Polyol	Daily threshold (grams)
Sorbitol	30–50
Xylitol	30–50
Mannitol	20–25
Erythritol	>80

Table 2
Relative caloric content of selected monosaccharide sweeteners (based on [59,83])

Monosaccharide	Calories per gram
Glucose	4.0
Fructose	4.0
Sorbitol	2.6
Xylitol	2.4
Mannitol	1.6
Erythritol	0.2

Table 3
Relative sweetness of selected monosaccharide sweeteners (based on [59,83])

Sugar	Relative sweetness
Fructose	1.1–1.7
Sucrose	1.00
Glucose	0.5–0.8
Galactose	0.3–0.6
Sorbitol	0.6
Xylitol	0.9
Mannitol	0.5
Erythritol	0.6



Scheme 4
Monosaccharide Polyols (from *left* to *right*): Sorbitol (a hexitol), Mannitol (a hexitol), Xylitol (a pentitol), Erythritol (a tetritol). Modified from [84]

tain its creamy texture. Sorbitol has roughly 60% the sweetness of sucrose, so products must be formulated accordingly to meet the desired sweetness profile of the final product. Fortunately, in this regard, sorbitol is compatible with high intensity sweeteners. Sorbitol is highly hygroscopic. It is even more hygroscopic than fructose [61]. Crystalline sorbitol is available in different crystal forms, depending on the method of manufacture. The crystal form, however, has little impact on the quality of most products, once the final products have been manufactured. Crystalline sorbitol is used as a sweetener in sugarless chewing gum, compressed tablets, and soft or gelled confections. Sorbitol-based syrups are used in chewing gum and hard and soft confections. Sorbitol is GRAS in the US, and its caloric value is 2.6 kcal per gram. The laxation threshold for sorbitol is about 50 grams per day.

Mannitol is produced commercially by catalytic hydrogenation of D-fructose (levulose). Hydrogenation of high fructose syrup, containing fructose and glucose, results in a mixture of sorbitol and mannitol. The mannitol can be separated by differential crystallization due to mannitol's low water solubility. Its solubility in water at 25 °C is only 22 grams per 100 ml of water. Over the years, however, there has been progress in identifying a direct and more economic fermentative process to produce mannitol [62]. Mannitol is a linear polyhydric alcohol (*Scheme 4b*) containing six carbons, a hexitol, with a molecular weight of 182. It is much less soluble in water than sorbitol. One of its important useful features is that the crystalline mannitol is non-hygroscopic, making it applicable as a coating powder. It is used in sugarless chewing gum, diet chocolate, and compressed sugar-free tablets. Mannitol occurs in nature, and is considered GRAS in the US. Its caloric value is 1.6 kcal per gram. The laxation threshold for mannitol is about 20 grams per day.

Xylitol is currently manufactured commercially by catalytic hydrogenation of D-xylose. Commercial xylose is presently produced by hydrolysis of hemicellulose under alkaline conditions. This process has placed constraints on the availability of xylitol. Enzyme-catalyzed hemicellulose depolymerization is clearly desirable [63]. There has been considerable interest over the years to identify a convenient source of xylitol through a complete fermentative process [64,65]. One of the key challenges has been to find a microbial source of enzyme that will be approved by the Food and Drug Administration (FDA) as not providing any toxic co-products. Xylitol is a linear polyhydric alcohol (*Scheme 4c*) containing five carbons, a pentitol, with a molecular weight of 152. Its solubility in water at 25 °C is 200 grams per 100 ml of water. Its heat of solution at 25 °C is −36.6 calories per gram. This negative heat of solution provides a cooling effect when the crystalline polyol is incorporated in a food product. The sweetness level of xylitol is about the same as sucrose, and its laxation threshold is comparable to that of sorbitol, about 50 grams per day. Since xylitol was reported [56] actually to reduce dental caries, it has been in great demand in chewing gum and confectionery applications, as well as in beverages, and frozen desserts. Xylitol is a naturally occurring substance. It is GRAS, and its caloric value is 2.4 calories per gram.

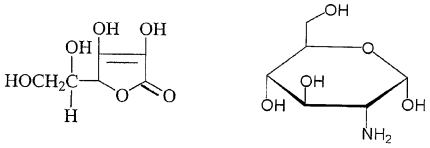
Erythritol is one of the newer polyol ingredients available in the global market place. It has always been produced commercially by fermentation [66,67]. Erythritol is a linear polyhydric alcohol (*Scheme 4d*) containing four carbons, a tetritol, with a molecular weight of 122. Its solubility in water at 25 °C is 50 grams per 100 ml of water. Its heat of solution at 25 °C is −23.7 calories per gram. This also provides a cooling effect when the crystalline polyol is incorporated in a food product. The sweetness level of erythritol is about 70% that of sucrose, but its laxation threshold is about zero per day, since it is almost completely excreted in the

urine. Erythritol is also reported [68] to reduce dental caries, and, as a result, is used in chewing gum, confectionery applications, beverages, bakery, and in frozen desserts. Erythritol is a naturally occurring substance. It is GRAS, and its caloric value (according to the FDA) is 0.2 calories per gram (glycemic index of 0).
Unique features of erythritol include low hygroscopicity and rapid crystallization. This latter feature presents a challenge in high solids applications since rapid crystallization can cause textural problems.

3 Nutritional Applications for Monosaccharides and Polyols

Users of commercial food ingredients can categorize sweeteners as being either nutritive or non-nutritive. The topic of non-nutritive sweeteners is not discussed in this chapter. Among nutritive sweeteners, reducing sugars, generally, are characterized as having a food energy value of about 4 calories per gram. Polyols have been characterized as having a caloric value of about 2.5 calories per gram. Erythritol is characterized in the US as having less than 0.5 calories per gram, though in Europe it is classified, currently, with other polyols. These caloric values, however, do not indicate the rapidity at which these materials are metabolized. From a nutritive standpoint, the rate of metabolism will have a significant impact on such biological features as the insulin response. This is characterized by the glycemic factor, discussed earlier, and is quantified by some through the glycemic index [24]. Furthermore, the body's ability to metabolize a food ingredient will impact on a person's intestinal flora, as was also mentioned earlier.
In recent years there has been increased concern over the effect that glucose-based food products have on dental decay. Oral bacteria, once attached to surfaces [69], can transfer glucose molecules from sucrose, through the action of the enzyme glucan sucrase, into long chains of insoluble polysaccharides that form plaque. Once plaque is formed, bacteria can then metabolize carbohydrate materials to form acidic compounds that degrade dental enamel. This concern about supporting dental caries is driving the increase of sugar free foods and the reduction of sugar-based products. For example, sales of sugar-free chewing gum have increased 6.5% in 2005, while sugar-based gum has decreased in the same period by 1.3% [70].
Some monosaccharide materials are used in specialty applications. One of the more important examples of such use is the monosaccharide L-ascorbic acid (*Scheme 5a*). Its function as an antiscorbutic factor from citrus fruits was recognized years ago, and it was first isolated by Albert Szent-Györgyi in 1928 [71]. It is synthesized starting with sorbitol and progresses through sorbose as an intermediate. Ultimately it yields the final product in a reasonably good yield. There is considerable interest in developing a complete fermentation-based process for this substance, as well [72]. In addition to its use today as a vitamin additive, it is used as an antioxidant ingredient in food products. It is used to stabilize color and aroma, and has been used to reduce the application of nitrates in meat. Ascorbic acid is also known, however, to participate in generating color in the Maillard-type reactions when combined with food amines [73], such as proteins and amino acids. Ascorbic acid is an unsaturated hexonic acid with a molecular weight of 176.12. It is water-soluble: 1 gram will dissolve in about 3 ml of water.

The use of vitamins as food additives has been practiced for some time. In recent years, there has been a trend to add supplementary materials that have claimed "nutraceutical" benefits. Recently, some interest has been generated concerning the use of the amino sugar glucosamine to improve the condition of joints as they age. Glucosamine is a hexose in which the hydroxyl group on the second carbon has been replaced by an amine group (◉ *Scheme 5b*). Glucosamine is a hexose whose molecular weight is 179.2. It is very soluble in water. Glucosamine has been administered by itself or in combination with chondroitin, a component of connective tissue [74] to help improve the condition of connective tissue, particularly in joints. Most of the glucosamine currently commercially available is derived by extracting shellfish exoskeletons. Commercial product that is derived from glucose fermentation, is beginning to appear on the market, though.



a Ascorbic Acid b Glucosamine

◘ Scheme 5
***a* Ascorbic acid, *b* Glucosamine**

4 Functional Applications for Monosaccharides and Polyols

In the contemporary food industry, food products are manufactured rather than prepared. Practices that enhance the efficiency and economics of product development are of prime importance. Furthermore, additional ingredients that maintain continuity of product properties and extend shelf life have considerable value. Most of the properties of the final food products result from the balance of ingredient interactions. Furthermore, much of these properties are developed as a result of the interaction of polymeric materials. Indeed, a significant amount of the food scientific literature is devoted to polymeric ingredient interactions. Some of the more innovative food science research has been devoted to the field called "food polymer science" [75]. Frequently, the monosaccharides are shunted to a secondary (but still important) role, such as a plasticizer or a water-holding agent. Monosaccharides have been used, for example, to control starch gelatinization and also control starch retrogradation [76]. These two processes are of considerable significance in food quality and shelf life stability. Starch retrogradation can be described as the processes that occur in gelatinized starch as it changes from an initially amorphous state to an ordered, crystallized state. Water content controls the rate of retrogradation, since water plasticizes the starch; monosaccharides influence the plasticizing ability of the water [75].

Control of moisture is one area of great value to food systems. Low molecular weight carbohydrate materials interact with water in ways that can be measured physically and used to predict physical properties. For example, the influence of molecular weights of saccharides on glass transition temperature has been studied [77] to understand their impact on food system proper-

ties. Water is a ubiquitous plasticizer in food systems, and moisture then influences structural properties and the migration of low molecular weight materials, such as flavors and sweeteners. Monosaccharides help control the availability of water and, hence, participate in control of properties of the final products. Monosaccharides are used to sweeten food products, in addition to their ability to provide bulk and hold water. Their sweetness intensity varies from monosaccharide to monosaccharide, and the reader has been presented with some of these properties throughout the discussion. Sweetness levels are summarized in *Table 3*. In addition to intensity, however, the question of sweetness profile must also be raised. One difference between sucrose and glucose-based coatings is that the sweetness of glucose is released much more rapidly than that of sucrose [17]. The consumer can then be presented with a different sweetness pattern, depending on the intent of the developer. Frequently the reduced sweetness intensity of a polyol can be augmented by the use of a high intensity sweetener. This extends the opportunity to the developer to expand or change the sweetness presented to the consumer. Acidulants from carbohydrates, such as gluconic acid and gluconatesalts, are known. It is not of significant commercial value, when compared to popular food acids, such as citric, malic, and fumaric, but it has some specialty applications. For example, gluconate is used as a delivery system for calcium ions. Antioxidants, such as ascorbic acid, have been discussed briefly above. Vitamin C can be added without mention on the label that it is functioning to prevent oxidation of ingredients, as well.

Finally, the importance of monosaccharides as raw materials for ingredients in food systems must also be stressed. The use of glucose and fructose to form sorbitol and mannitol has already been addressed. Chemical degradation of reducing sugars can yield colored and coloring materials, such as hydroxymethyl furfural [78], that can than either polymerize or form Maillard reaction products with amines [79]. Lichtenthaler has recently reviewed some potential opportunities to use monosaccharide ketoses as organic raw materials for industrial applications [80]. Recent instability of petroleum prices has given additional urgency to such carbohydrate chemical exploration. Finally, advances in biochemistry have expanded researchers' understanding of the role of carbohydrates in control of physiological function. This enhanced understanding of glycoscience has opened the door to the development of carbohydrate-based pharmaceuticals [81]. This exciting area, however, is outside the scope of this chapter.

Sugars are used as fermentation substrates to yield products such as lactic acid, citric acid, and ethanol. Ethanol can be used as a food product, as a food solvent, or as a source for energy. In very recent times, with rapidly escalating demand for power alcohol, concern over the economic impact of this latter usage has raised serious challenges to the availability of sufficient quantities of some cereal grains for food use [82].

References

1. http://www.MyPyramid.go. Accessed 2005
2. Gaonkar AG (1995) (ed) Ingredient Interactions. Marcel Dekker, New York
3. Cui SW (2005) (ed) Food Carbohydrates. Taylor & Francis, Boca Raton
4. Eggleston G, Côté GL (2003) (eds) Oligosaccharides in Food and Agriculture. ACS Symposium Series 849, American Chemical Society, Washington, DC
5. Eliasson AC (1996) (ed) Carbohydrates in Food. Marcel Dekker, New York
6. Friedman R (1991) In: Goldberg I, Williams R (eds) Biotechnology and Food Ingredients. Van Nostrand Reinhold, New York, pp 327–347

7. Waller GR, Feather MS (1983) The Maillard Reaction in Foods and Nutrition. American Chemical Society, Washington, DC
8. Baynes JW, Monnier VM, Ames JM, Thorpe SR (2005) (eds) The Maillard Reaction: Chemistry at the Interface of Nutrition, Aging, and Disease. NewYork Academy of Science, NY
9. Doner LW (1977) J Sci Food Agric 28:443–56
10. Doner LW, White JW Jr, Phillips JG (1979) J Assoc Off Anal Chem 62:186–9
11. Ischayek JI, Kern M (2006) J Am Diet Assoc 106:1260–1262
12. Pancoast HM, Junk WR (1980) Handbook of Sugars, 2nd edn. AVI, Westport, CT
13. Zero DT (2004) Caries Res 38:277–285
14. Daniels MJ (1982) US Patent 4,342,603
15. Schwenk M (2003) Manuf Confectioner 83:49–56
16. Horn HE (1977) Manuf Confectioner 1977:79–86
17. Nonaka HH (1991) Manuf Confectioner 1991:67–70
18. Ho CT, Huang CT, Cha AS, Sotirhos N (1988) In: Charalambous G (ed) Frontiers of Flavor. Elsevier, Amsterdam, pp 233–240
19. Huang CT, Soliman A, Rosen, R, Ho CT (1987) Food Chem 24:187–196
20. Levine R (1986) Ann Rev Nutr 6:211–224
21. van Tilburg R (1985) In: van Beynum GMA, Roels JA (eds) Starch Conversion Technology. Marcel Dekker, New York, pp 175–231
22. De Mendonca Ferreira J, Mattos Teixeira C, Stamile Soares S (2006) US Patent 7,150,794
23. Kierstan M (1980) Proc Bioch 15:2–4,32
24. Jenkins DJ, Wolever TM, Taylor RH, Barker H, Fielden H, Baldwin JM, Bowling AC, Newman HC, Jenkins AL, Goff DV (1981) Am J Clin Nutr 34:362–366
25. Hyvonen L, Kurkela R, Koivistoinen P, Ratilainen A (1978) J Food Sci 43:251–254
26. Hyvonen L, Kurkela R, Koivistoinen P, Ratilainen A (1978) J Food Sci 43:1577–1584
27. Kahn ML, Eapen KE (1979) US Patent 4,154,863
28. Kahn ML, Eapen KE (1981) US Patent 4,262,032
29. Friedman RB, West IR, Furcsik SL (1986) US Patent 4,575,461
30. Reiser S (1992) In: Bills DD, Kung S (eds) Biotechnology and Nutrition. Butterwoth-Heinemann, Boston, pp 167–189
31. Fields M, Holbrook J, Scholfield D, Smith JC Jr, Reiser S (1986) J Nutr 116:625–632
32. Vrána A, Fábry P (1983) Wld Rev Nutr Diet 42:56–101
33. Fujisawa T, Mulligan K, Wada L, Schumacher L, Riby J, Kretchmer N (1993) Am J Clin Nutr 58:75–79
34. Glinsmann WH, Bowman BA (1993) Am J Clin Nutr 58:820–823
35. http://www.corn.org
36. Zehner LR (1988) US Patent 4,786,722
37. Donner TW, Wilber JF, Ostrowski D (1999) Diabetes Obes Metab 1:285–291
38. Osama I (1993) EPO Patent 0,552,894
39. Lee A, Storey DM (1999) Regul Toxicol Pharmacol 29:78–82
40. Wolf FR (1997) PCT 97–022263A
41. Lee T (1994) US Patent 5,356,879
42. Lee T (1995) US Patent 5,447,917
43. Levin GV (1981) US Patent 4,262,032
44. Levin GV (1986) In: Nabors LO, Gelardi RC (eds) Alternative Sweeteners. Marcel Dekker, NY, pp 155–164
45. Goosens J, Roeper H (1994) Food Sci Today 38:144–149
46. Saha BC, Bothast RJ (1997) In: Saha BC, Woodward J (eds) Fuels and Chemicals from Biomass. ACS Symposium Series 666, American Chemical Society, Washington, DC, pp 307–319
47. Saha BC, Nakamura LK (2003) Biotechnol Bioeng 82:864–71
48. Saha BC (2006) Appl Micobiol Biotechnol 72:676–680
49. Saha BC (2005) US Patent 6,855,526
50. Zumbé A, Lee A, Storey D (2001) Br J Nutr 85:31–45; 51
51. Livesey G (2003) Nutr Res Revs 16:163–191
52. Mäkinen KK, Söderling E, Isokangas P, Tenovou J, Tiekso J (1989) Caries Res 23:261–267
53. Söderling E, Mäkinen KK, Chen CY, Pape HR Jr, Loesche W, Mäkinen PL (1989) Caries Res 23:378–384
54. Kawanabe J, Hirasawa M, Takeuchi T, Oda T, Ikeda T (1992) 26:358–362
55. Mäkinen KK, Saag M, Isotupa KP, Olak J, Nommela R, Söderling E, Mäkinen PL (2005) Caries Res 39:207–215
56. van Loveren C (2004) Caries Res 38:286–293
57. Ludwig DS (2002) JAMA 287:2414–2423
58. Foster-Powell K, Holt SHA, Brand-Miller JC (2002) Am J Clin Nutr 76:5–56
59. http://www.caloriecontrol.org
60. Lorenz S, Grossklaus R (1984) Nutr Res 4:447–458
61. Makkee M, Kieboom APG, van Bekkum H (1985) Starch 37:136–141

62. Soetaert W, Vanhooren PT, Vandamme EJ (1999) In: Bucke C (ed) Carbohydrate Biotechnology Protocols. Humana Press, Totowa, NJ, pp 261–275
63. Saha BC (2003) J Ind Microbiol Biotechnol 30:279–291
64. Saha BC, Bothast RJ (1999) J Ind Microb Biotech 22:633–636
65. Leathers TD (2003) FEMS Yeast Res 3:133–140
66. de Cock P (1999) In: Corti A (ed) Low-Calorie Sweeteners: Present and Future. Karger, Basel, pp 110–116
67. Roper H, Goossens J (1993) Starch//Staerke 45:400–405
68. Troostembergh G (2001) US Patent 6,177,064
69. Mobley C, Dodds MW (2003) In: Palmer CA (ed) Diet and Nutrition in Oral Health. Prentice Hall, Upper Saddle River, NJ, pp 182–201
70. Anon (2006) Manuf Confectioner 86:37–42
71. Szent-Györgyi A (1928) Biochem J 22:1387
72. Reilly CE (1991) In: Goldberg I, Williams R (eds) Biotechnology and Food Ingredients. Van Nostrand Reinhold, New York, pp 415–432
73. Wedzicha B (1995) In: Gaonkar AG (ed) Ingredient Interactions. Marcel Dekker, New York, pp 529–559
74. Clegg DO, Reda DJ, Harris CL, Klein MA, O'Dell JR, Hooper MM, Bradley JD, Bingam CO 3rd, Weisman MH, Jackson CG, Lane NE, Cush JJ, Moreland LW, Schumacher HR Jr, Oddis CV, Wolfe F, Molitor JA, Yocum DE, Schnitzer TJ, Furst DE, Sawitzke AD, Shi H, Brandt KD, Moskowitz RW, Williams HJ (2006) N Engl J Med 354:795–808
75. Slade L, Levine H (1987) In: Stivala SS, Crescenzi V, Dea ICM (eds) Industrial Polysaccharides: The Impact of Biotechnology and Advanced Methodologies. Gordon & Breach, New York, p 387
76. Eliasson AC, Gudmundsson M (1996) In: Eliasson AC (ed) Carbohydrates in Food. Marcel Dekker, New York, pp 431–503
77. Slade L, Levine H (1996) In: Eliasson AC (ed) Carbohydrates in Food. Marcel Dekker, New York, pp 41–157
78. van Dam HE, Kieboom APG, van Bekkum H (1986) Starch/Stärke 38:95–101
79. Chundury D, Szmant HH (1981) Ind Eng Chem Prod Res Dev 20:158–163
80. Lichtenthaler FW (1998) Carb Res 313:69–89
81. Turnbull JE, Linhardt RJ (2006) Nature Chem Biol 2:449–450
82. Sterk R (2007) Food Bus News Art 83055
83. Alexander RJ (1998) Sweeteners: Nutritive. Eagen Press, St. Paul
84. Stanek J, Cerny M, Kocourek J, Pacak J (1963) The Monosaccharides. Academic Press, NY
85. Shallenberer RS (1978) Pure & Appl Chem 50:1409–1420

4.3 De novo Synthesis of Monosaccharides

Pierre Vogel[1], *Inmaculada Robina*[2]
[1] Laboratoire de glycochimie et de synthèse asymétrique, Ecole Polytechnique Fédérale de Lausanne (EPFL), BCH, 1015 Lausanne-Dorigny, Switzerland
[2] Departamento de Química Orgánica, Universidad de Sevilla, 41071 Sevilla, Spain
pierre.vogel@epfl.ch, robina@us.es

In: *Glycoscience*. Fraser-Reid B, Tatsuta K, Thiem J (eds)
Chapter-DOI 10-1007/978-3-540-30429-6_20:

Abstract

Recent approaches to asymmetric total synthesis of monosaccharides and derivatives are reviewed. They imply biochemical methods or chemical methods, or combinations of both. Aldoses, alditols, deoxysugars, aminodeoxymonosaccharides, and their analogues are considered, including aza- and thiosugars (with nitrogen and sulfur in the pyranose or furanose ring). Most common monosaccharides have three to six carbon-chains, but examples of long-chain and branched-chain carbohydrates will also be described. Organization of the review is based on the synthetic methodology rather than on the types of targeted sugar, starting with aldolase-catalyzed aldol reactions of achiral precursors, and the now popular amino-acid-catalyzed

aldol and Mannich reactions. Chain elongations of small aldose derivatives with one, two, three, or four carbons remain very useful approaches as they allow one to reach a large structural and stereochemical diversity. Other methods apply Diels–Alder or hetero-Diels–Alder additions. Enantiomerically pure monosaccharides have been derived from achiral alkene, diene, and triene derivatives. Almost any targeted natural or non-natural sugar and derivative can be obtained in a few synthetic steps from inexpensive starting materials.

Keywords

Aldol reaction; Aza-sugars; Branched-chain sugars; Chain elongation; Diels–Alder; Enantioselective sharpless dihydroxylation; Furan addition; Hetero-Diels–Alder addition; Long-chain sugars

Abbreviations

9-BBN	9-borabicyclo[3.3.1]nonane
Ac	MeCO
AIBN	azoisobutyronitrile
Bn	benzyl
Boc	(*t*-Bu)OCO
BOM	$PhCH_2OCH_2$
Bu	*n*-butyl
Bz	PhCO
Cbz	BnOCO
CDI	carbonyldiimidazole
Co(II)(salen)$_2$	salcomine: *N*,*N*′-bis(salicylidene)ethylenediaminocobalt (II)
RAMP	(*R*)-(+)-amino-2-(methoxymethyl)pyrrolidine
dba	dibenzylideneacetone
DBN	1,5-diazabicyclo[4.3.0]non-5-ene
DBU	1,8-diazabicyclo[5.4.0]undec-7-ene
DDQ	dichlorodicyanobenzoquinone
DERA	2-deoxyribose-5-phosphate aldolase
DFT	density functional theory
DHA	dihydroxyacetone
DHAP	dihydroxyacetone phosphate
DIBAL-H	$(i\text{-Bu})_2AlH$
DMAP	4-(dimethylamino)pyridine
DMF	dimethylformamide
DMSO	dimethylsulfoxide
DMTC	5,5-dimethyl thiazolidinium-4-carboxylate
dr	diastereomeric ratio
ee	enantiomeric excess
FDP	fructose-1,6-diphosphate
fod	6,6,7,7,8,8,8-heptafluoro-2,2-dimethyl-3,5-octanedienoate
Grubbs I	$[(cyclohexyl)_3P]_2Ru(Cl)_2CHPh$

Grubbs II	[(cyclohexyl)$_3$P][1,3-(2,4,6-trimethylphenyl)-1,3-diazolidinylidene]Ru(Cl)$_2$CHPh
hfc	3-(heptafluoropropylhydroxymethylene)-(+)-camphorate
HMPA	hexamethyl phosphortriamide
i-Pr	isopropyl
LDA	lithium diisopropylamide
K-Selectride	KB[CH(Me)Et]$_3$H
L-(+)-DIPT	L-(+)-diisopropyl tartrate
L-KDO	3-deoxy-L-*manno*-oct-2-ulosonic acid
L-Selectride	LiB[CH(Me)Et]$_3$H
LTMP	lithium 2,2,6,6-tetramethylpiperidine
mCPBA	*meta*-chloroperbenzoic acid
MEM	MeOCH$_2$CH$_2$OCH$_2$
Ms	methanesulfonyl (MeSO$_2$)
"naked aza-sugar"	enantiomerically pure 7-oxa-bicyclo[2.2.1]hept-2-ene derivative
"naked sugar"	enantiomerically pure 7-aza-bicyclo[2.2.2]hept-2-ene derivative
NCS	*N*-chlorosuccinimide
NMO	*N*-methylmorpholine *N*-oxide
NMP	*N*-methylpyrrolidone
PEP	phosphoreol pyruvate
Piv	pivaloyl (*t*-BuCO)
PMBz	4-MeOC$_6$H$_4$CO
py	pyridine
RADO(Et)OH	(1*R*,7*R*)-3-ethyl-2-oxo-3-aza-6,8-dioxabicyclo[3.2.1]octane-7-*exo*-carboxylic acid
RAMA	rabbit muscle aldolase
Red-Al	(MeOCH$_2$CH$_2$O)$_2$AlH$_2$Na
SADO(Et)OH	(1*S*,7*S*)-3-ethyl-2-oxo-3-aza-6,8-dioxabicyclo[3.2.1]octane-7-*exo*-carboxylic acid
SAMP	(*S*)-(−)-amino-2-(methoxymethyl)pyrrolidine
Sia$_2$BH	(Me$_2$CHCHMe)$_2$BH
***S*-Pro**	L-proline
TBS	(*t*-Bu)Me$_2$Si
***t*-Bu**	*tert*-butyl
TES	Et$_3$Si
Tf	triflyl (CF$_3$SO$_2$)
THF	tetrahydrofurane
THP	2-tetrahydropyranyl
TIPS	(*i*-Pr)$_3$Si
TMS	Me$_3$Si
Ts	*para*-toluenesulfonyl (4-MeC$_6$H$_4$SO$_2$)

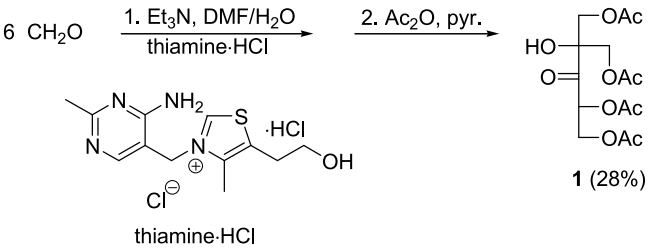
1 Introduction

Total synthesis of carbohydrates and analogs has kept chemists busy since 1861 when Butlerow [1,2,3,4,5] discovered the "formose reaction" which generates mixtures of racemic aldoses and ketoses by oligomerization of formaldehyde in the presence of $Ca(OH)_2$. Today, with the advent of highly stereoselective and enantioselective methods almost any natural or non-natural carbohydrate can be obtained readily from inexpensive starting materials in an enantiomerically pure form. D-Glucose, D-mannose, D-glucosamine, D- and L-arabinose from natural sources are certainly cheaper than from total synthesis. But when we are dealing with unnatural enantiomers of common carbohydrates, or with semi-protected derivatives or unusual derivatives in which hydroxy groups are replaced by amino moieties, by alcoxy groups, thio, halogeno, carbon-substituents, etc… total synthesis from noncarbohydrate precursors may be easy and advantageous. By total synthesis the carbohydrates are delivered in suitably protected forms. In contrast, by starting from natural sugars this sometimes requires several delicate chemical operations.

This review describes the synthetic approaches that have been developed since 2000 and summarizes the most important ones developed prior to 2000 [6,7,8,9,10]. It will concentrate on techniques generating enantiomerically enriched or pure carbohydrates and analogs. Aldoses, alditols, and their derivatives will be considered, including aza and thiosugars (with nitrogen and sulfur in the pyranose or furanose rings).

2 The Formose Reaction

The formose reaction was developed by Loew [11,12] and Fischer [13,14] who isolated *rac*-fructose osazone from the formose reaction mixture. The yield of formose sugars reaches a maximum at the so-called yellowing point [15]. On further reaction branched sugars are formed involving aldol condensations followed by cross-Cannizarro reactions [16]. Depending on the nature of the base and additives used to induce the formaldehyde oligomerization, various proportions of trioses, tetroses, pentoses, hexoses, and long-chain aldoses and ketoses are obtained [17,18,19]. When carried out in dimethylformamide (DMF) considerable control in the product distribution of the formose reaction is possible by adjustment of the water content (❯ *Scheme 1*) [20,21,22,23,24]. When, for instance, formaldehyde is heated to 75 °C for



Scheme 1
Examples of selective formose reaction

1 h with Et_3N and thiamine hydrochloride in 8:1 DMF/H_2O, DL-2-C-hydroxymethyl-3-pentulose, characterized as its tetraacetate **1** is produced in 28% yield [25].

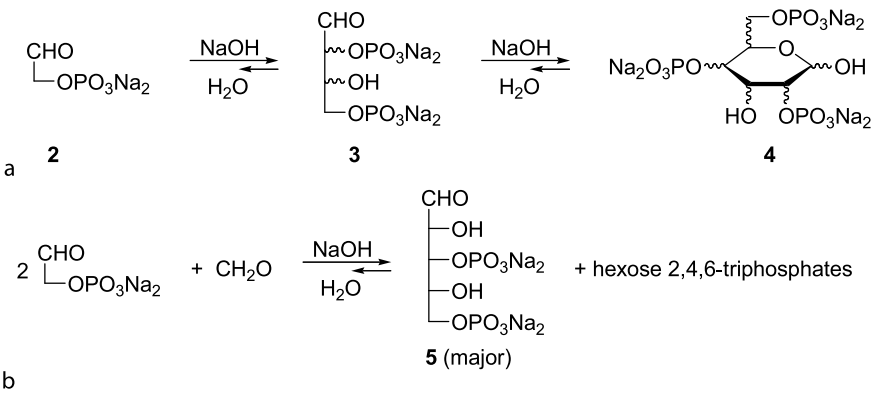
The formose reaction has been investigated using immobilized thiazolium catalyst [26]. Under these conditions the main products are dihydroxyacetone (DHA), erythrulose, and 4-hydroxymethyl-2-pentulose. The relative importance of these products depends on the amount of thiazolium salts and concentration in 1,4-dioxane [27,28,29]. A possible mechanism implies the Stetter reaction [30,31,32,33,34].

Eschenmoser and co-workers [35] studied the aldomerization of glycoaldehyde phosphate which led to mixtures containing mostly racemates of the two diastereomeric tetrose 2,4-diphosphate and eight hexose 2,4,6-triphosphates (*Scheme 2*, route A). At 20 °C in the absence of air, a 0.08-molar solution of glycolaldehyde phosphate **2** in 2-M NaOH gave 80% yield of a 1:10 mixture of tetrose **3** and hexose **4** derivatives with DL-allose 2,4,6-triphosphate comprising up to 50% of the mixture of sugar phosphate [36].

In the presence of formaldehyde (0.5 mol-equiv.) sugar phosphates were formed in up to 45% yield, with pentose 2,4-diphosphates dominating over hexose triphosphates by a ratio of 3:1 (*Scheme 2*, route B). The preference for ribose 2,4-diphosphate **5** and allose 2,4,6-triphosphate formation might have significance for the discussion concerning the origin of ribonucleic acids.

The "classical" formose reaction gives a very large number of carbohydrates including branched-chain isomers [17,18,19]. Straight-chain carbohydrates such as trioses, tetroses, pentoses, and hexoses are readily obtained in good yield by a reaction of formaldehyde with syngas in the presence of $RhCl(CO)(PPh_3)_2$ and tertiary amines [37].

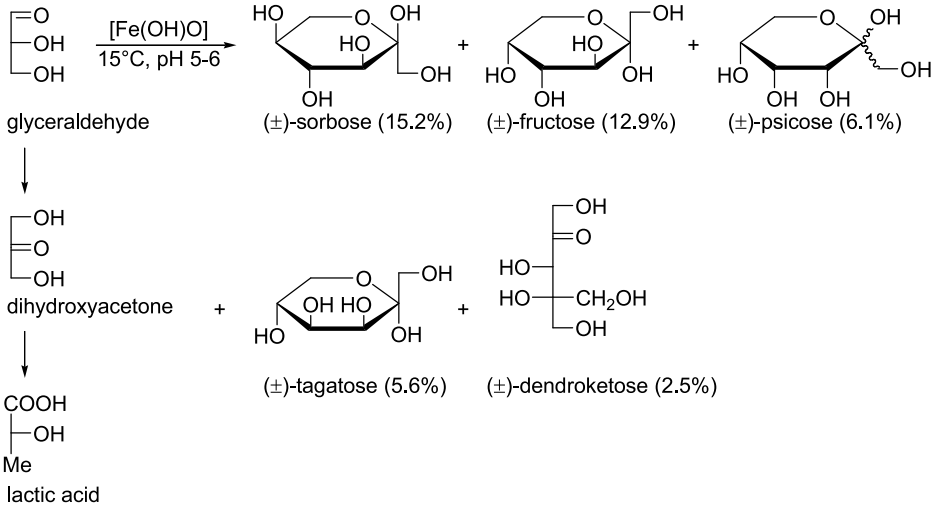
Shevlin and co-workers [38] have reported that co-condensation of carbon with H_2O and NH_3 at 77 K generates amino-acids. They showed also that atomic carbon generated by vaporizing in an arc under high pressure reacts with water at 77 K to form low yields of straight-chain aldoses with up to five carbon centers. A mechanism involving hydroxymethylene species has been supported by deuterium labeling studies [39]. Under UV irradiation neutral aqueous solutions of formaldehyde form CO, CO_2, CH_4, CH_3CH_3, and ethylene gas. At the same time,



Scheme 2
(A) Selective condensation of glycolaldehyde phosphate alone; (B) in the presence of formaldehyde

formaldehyde condenses into glycoaldehyde and glyceraldehyde, two active precursors in the formose reaction. This might correspond to reactions that occurred on prebiotic Earth which led to the first carbohydrates via the formose reaction [40].

Iron(III)hydroxide oxide [Fe(OH)O] has been shown to catalyze the condensation of 25-mM DL-glyceraldehyde to ketohexoses at 15 °C (pH 5–6). After 16 days 15.2% of sorbose, 12.9% of fructose, 6.1% of psicose, 5.6% of tagatose, and 2.5% of dendroketose are obtained. After 96 days at 15 °C this mixture was not decomposed. [Fe(OH)O] also catalyzes the isomerization of glyceraldehyde into dihydroxyacetone and of dihydroxyacetone into lactic acid [41].



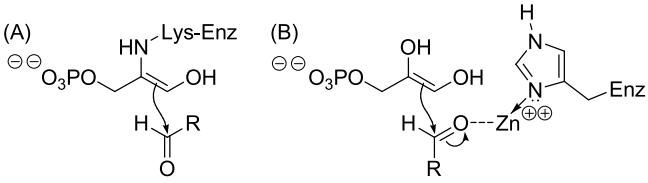
Scheme 3
[Fe(OH)O]-catalyzed reactions of D,L-glyceraldehyde

The "classical formose" conditions are not capable of producing large amounts of ribose (for RNA synthesis), nor of any other individual sugar. In contrast, the reduced sugar pentaerythritol is formed with great selectivity by the ultra-violet irradiation of 0.1-M formaldehyde. This compound may have played an important role in prebiotic chemistry [42]. The seminal work of Eschenmoser and co-workers [35] (*Scheme 2*) suggests that the "initial RNA world" might have involved glycoaldehyde phosphate [43]. In order to explain the concentration process required one can envisage that double-layer hydroxide minerals might have played a decisive role, in particular those incorporating sodium sulfite, which can absorb formaldehyde, glycoaldehyde, and glyceraldehyde by adduct formation with the immobilized sulfite anions. This translates into observable uptake at concentrations ≥ 50 mM [44]. Sugars have been proposed to be the optimal biosynthetic carbon substrate of aqueous life throughout the Universe [45].

Benner and co-workers [46] have followed the formation of pentoses under alkaline conditions from simple precursors such as formaldehyde and glycolaldehyde in the presence of borate minerals. The latter stabilize the pentose selectively by forming complexes.

3 Aldolase-Catalyzed Asymmetric Aldol Condensations

The enzymatic aldol addition represents a useful method for the synthesis of various sugars and sugar-like structures [47,48,49,50,51,52,53,54,55,56,57,58,59,60,61,62,63,64,65,66,67]. More than 20 different aldolases have been isolated (see *Table 1* for examples) and several of these have been cloned and overexpressed [68]. They catalyze the stereospecific aldol condensation of an aldehyde with a ketone donor. Two types of aldolases are known. Type I aldolases found primarily in animals and higher plants do not require any cofactor. The X-ray structure of the aldolase from rabbit muscle (RAMA = rabbit muscle aldolase) indicate that Lys-229 is responsible for Schiff-base formation with dihydroxyacetone phosphate (DHAP) (*Scheme 4*A). Type II aldolases found primarily in microorganisms use Zn^{++} as the cofactor which acts as a Lewis acid enhancing the electrophilicity of the ketone (*Scheme 4*B). In both cases the aldolases accept a variety of natural (*Table 1*) and of non-natural acceptor substrates (*Scheme 5*). *N*-Acetylneuraminic acid aldolase (Neu5Ac aldolase) from *E. coli* catalyzes the reversible aldol reaction of *N*-acetyl-D-mannosamine and pyruvate to give *N*-acetyl-D-neuraminic acid (D-sialic acid). This enzyme is quite specific for pyruvate as the donor, but flexible to a variety of D- and, to some extent, L-hexoses and L-pentoses as acceptor substrates [69,70]. Using error-prone PCR (polymerase chain reaction) for in vitro directed evolution, the Neu5Ac aldolase has been altered to improve its catalytic activity toward enantiomeric substrates such as *N*-acetyl-L-mannosamine and L-arabinose to produce L-sialic acid (a potent neuraminidase inhibitor for the treatment of flu is derived from D-sialic acid) [71], and L-KDO (3-deoxy-L-*manno*-oct-2-ulosonic acid), the enantiomers of naturally occurring D-sialic acid and D-KDO, respectively [68,72] (*Table 1*).

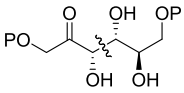
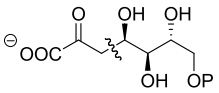
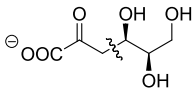
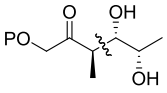
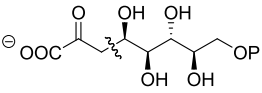
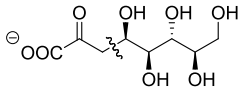
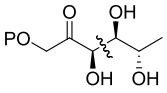
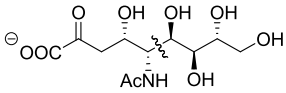
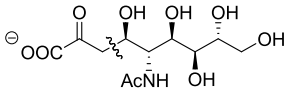
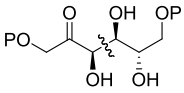
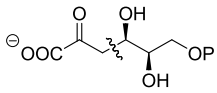


Scheme 4
(A) Type I aldolases form enamine nucleophiles (donor); (B) Type II aldolases use Zn^{++} as the cofactor activating the aldehyde (acceptor)

3.1 Resolution of Racemic Aldehydes

D-Fructose-1,6-diphosphate aldolase (FDP) catalyzes the reversible aldol addition of DHAP and D-glyceraldehyde-3-phosphate (G3P) to form D-fructose-1,6-diphosphate (FDP), for which $K_{eq} \approx 10^4\,M^{-1}$ in favor of FDP formation (*Scheme 6*). Rabbit muscle aldolase accepts a wide range of aldehyde acceptor substrates with DHAP as the donor to generate 3*S*,4*S* vicinal diols, stereospecifically (*Scheme 5*). A racemic mixture of non-natural aldehyde acceptors can be partially resolved only under conditions of kinetic control. When

Table 1 (see also on the following page)
Examples of enzymes catalyzing the equilibria of natural products with various aldol donors and various aldehydes (the *wavy line* indicating the C–C bond involved in the reversible aldol reaction). FDP = Fructose-1,6-diphosphate; DHAP = dihydroxyacetone phosphate, KDO = 3-deoxy-D-*manno*-octulosonate, P = $^{-}O_3P$

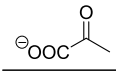
Aldol donor (nucleophiles)		
DHAP		
		
FDP aldolase	DAHP synthetase	3-deoxy-2-oxo-6-P-gluconate aldolase
		
fuculose-1-P-aldolase	KDO synthetase	KDO aldolase
		
	sialic acid synthetase	sialic acid aldolase
		
tagatose-1,6-P_2 aldolase		3-deoxy-2-oxo-6-P-galactonate aldolase

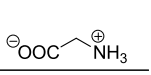
six-membered hemiacetals can be formed, racemic mixtures of aldehydes can be resolved under conditions of thermodynamic control (● *Scheme 7*).

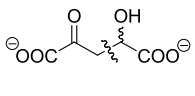
DL-Glyceraldehyde and 1,3-dihydroxyacetone are obtained from glycerol mild oxidation, for instance with hydrogen peroxide in the presence of ferrous salts as catalysts [73]. Selective formation of trioses has been observed in the formose reaction when α-ketols bearing electron-withdrawing substituents were added to the reaction mixture [74,75]. In the presence of thiazolium salts, selective conversion of formaldehyde into 1,3-dihydroxyacetone has been reported [76,77,78]. Hydration of halopropargyl alcohol followed by hydrolysis gives 1,3-dihydroxyacetone [79,80]. DHAP can be generated by three different procedures: (1) in situ from fructose 1,6-diphosphate with the enzyme triosephosphate isomerase; (2) from the dimer of dihydroxyacetone by chemical phosphorylation with $POCl_3$ (● *Scheme 8*); or (3) from dihydroxyacetone by enzymatic phosphorylation using ATP and glycerol kinase, with in-situ generation of the ATP using phosphoenol pyruvate (PEP) or acetyl phosphate as the phosphate donor (● *Scheme 9*) [73].

Table 1
(continued)

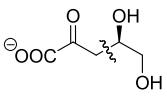
Aldol donor:



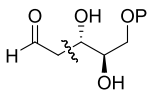




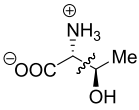
4-hydroxy-2-oxo-glutarate aldolase



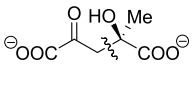
3-deoxy-2-oxo-L-arabinoate aldolase



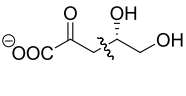
2-deoxyribose-5-P-aldolase



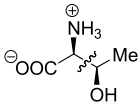
D-Thr aldolase



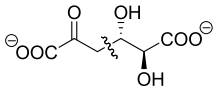
4-hydroxy-4-methyl-2-oxoglutarate aldolase



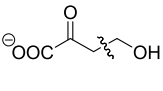
3-deoxy-2-oxo-D-pentanoate aldolase



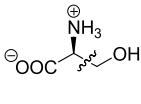
L-Thr aldolase



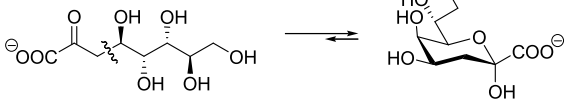
3-deoxy-2-oxo-D-glucarate aldolase



hydroxybutyrate aldolase

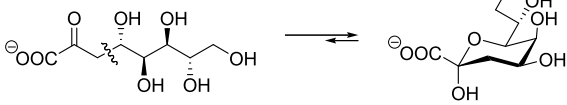


Ser-hydroxymethyl transferase



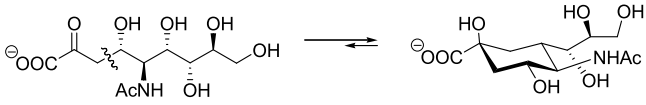
N-acetylneuraminic acid aldolase from *E. coli*

(D-KDO)

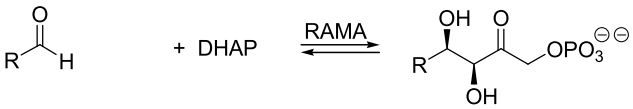


N-acetylneuraminic acid aldolase mutant

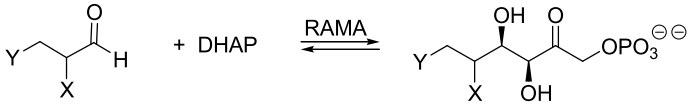
(L-KDO)



(L-sialic acid)
N-acetylneuraminic acid aldolase mutant[33]



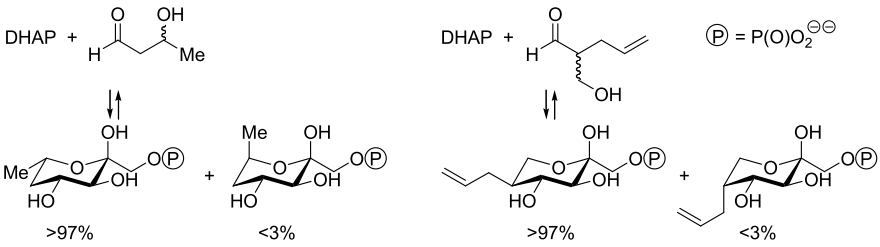
R = H, Me, $ClCH_2$, CHO, COOH, N_3CH_2CHO, $^{\ominus\ominus}O_3POCH_2$, $THPOCH_2$, $PhCOOCH_2$



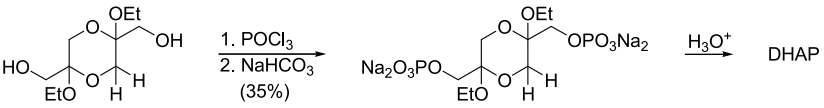
X = H, Me, OH, OMe, OAc, NHAc
Y = H, OH, $OPO_3^{\ominus\ominus}$, F, N_3

Scheme 5
Examples of RAMA-catalyzed aldol condensations

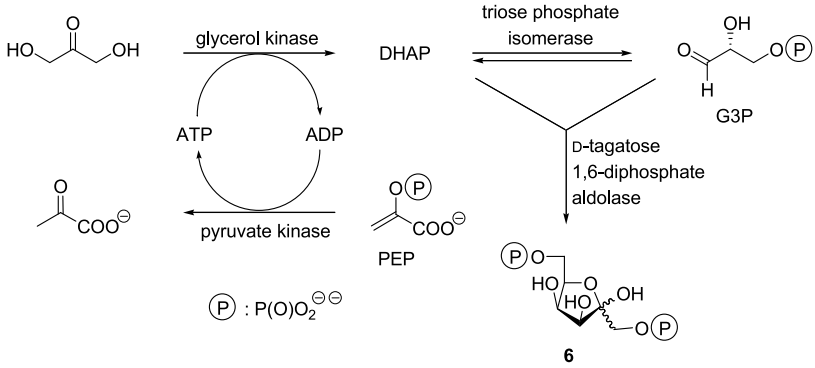
Scheme 6
Stereospecific FDP aldolase-catalyzed aldol reaction of DHAP + G3P ⇄ FDP



Scheme 7
Thermodynamically controlled resolution of racemic aldehydes with FDP aldolase from RAMA



Scheme 8
Chemical synthesis of dihydroxyacetone phosphate (DHAP)



Scheme 9
One-pot synthesis of D-tagatose 1,6-diphosphate

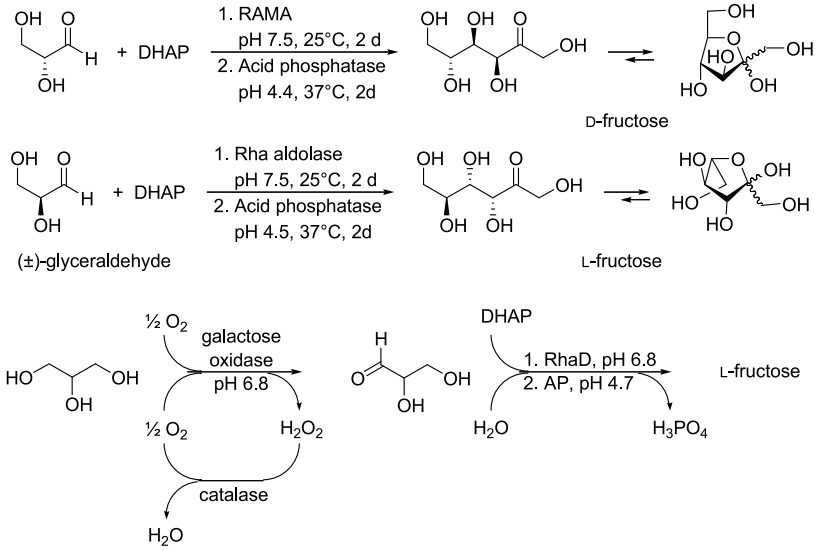
3.2 One-Pot Total Syntheses of Carbohydrates

A one-pot procedure has been proposed to convert dihydroxyacetone and phosphorenol pyruvate (PEP) into D-tagatose 1,6-diphosphate **6** (*Scheme 9*). The reaction mixture contains glycerolkinase, pyruvate kinase, triose phosphate isomerase, and a D-tagatose 1,6-diphosphate aldolase [81].

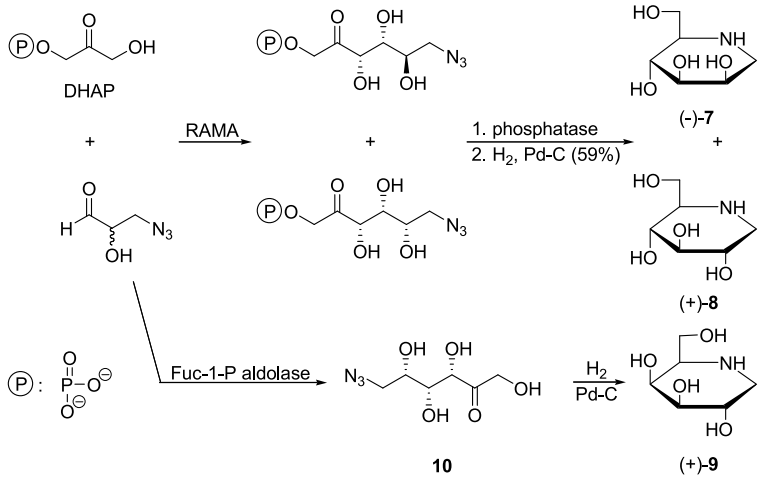
An efficient asymmetric total synthesis of L-fructose combines the Sharpless asymmetric dihydroxylation with an enzyme-catalyzed aldol reaction. L-Glyceraldehyde prepared from acrolein is condensed to DHAP in a buffered water suspension of lyzed cells of K12 *Escherichia coli* containing excess of L-rhamnulose 1-phosphate (Rha) aldolase (*E. coli* raised on L-rhamnose as the sole carbon source). The L-fructose phosphate obtained is hydrolyzed to L-fructose with acid phosphatase. Similarly, the RAMA-catalyzed condensation of D-glyceraldehyde with DHAP, followed by acid phosphatase-catalyzed hydrolysis furnishes D-fructose. A one-pot preparation of L-fructose (55% yield) starting from (±)-glyceraldehyde and DHAP has also been developed [82]. An alternative method starting from glycerol and DHAP using a coupled enzymatic system including galactose oxidase, catalase, rhamnulose-1-phosphate aldolase (RhaD), and acid phosphatase (AP) has also been presented by Wong's group [82] (*Scheme 10*). The method works also to generate 6-deoxy-D and L-galacto-2-heptulose from (*E*)-crotonaldehyde, and 6-phenyl-D and L-galacto-2-hexulose from (*E*)-cinnamaldehyde [83]. Isomerization of L-fructose catalyzed by fucose isomerase (available from commercial recombinant *E. coli* strains) furnishes L-glucose [84].

3.3 Synthesis of 1,5-Dideoxy-1,5-Iminoalditols

Two potent glycosidase inhibitors, (−)-1-deoxymannonojirimycin (−)-**7** and (+)-1-deoxynojirimycin (+)-**8** are readily obtained in three steps in which RAMA is used as the catalyst in the key C–C bond-forming step [55,56,57,58,59,60,61,62,63,64,85,86,87,88,89]. From racemic 3-azido-2-hydroxypropanal and dihydroxyacetone monophosphate, diastereomeric 6-azidoketones are formed. Following the acid phosphatase-catalyzed removal of



Scheme 10
Isomerization of L-fructose into L-glucose



Scheme 11
Chemoenzymatic synthesis of 1,5-dideoxy-1,5-imino-alditols

phosphate and subsequent reductive amination (*Scheme 11*), the products are isolated in a 4:1 ratio favoring the *manno* derivative. A similar result is obtained with fructose-1,6-diphosphate aldolase from *E. coli* [90]. Exclusive formation of (−)-**7** and (+)-**8** is observed if the respective enantiomerically pure azidoaldehydes are used as starting mate-

rials. An analogous RAMA-catalyzed aldol reaction/reductive amination procedure has been used in the total synthesis of 2-acetylamino-1,2,5-trideoxy-1,5-imino-D-glucitol and 2-acetylamino-1,2,5-trideoxy-1,5-imino-D-mannitol from (*S*)- and (*R*)-3-azido-2-acetamidopropanal, respectively [91]. The 6-deoxy analogs of the 1,5-dideoxy-1,5-iminohexitols can be obtained by direct reductive amination of the aldol products prior to removal of the phosphate group [55,56,57,58,59,60,61,62,63,64]. Fuculose-1-phosphate (Fuc-1-P) aldolase catalyzes the aldolization between DHAP and (±)-3-azido-2-hydroxypropanal leading to a ketose-1-phosphate **10** which has used the L-enantiomer of the 2-hydroxypropanal derivative (*Scheme 16*). Reduction of the azide generates an amine which cyclizes to an imine that is hydrogenated with high diastereoselectivity providing (+)-1-deoxygalactostatine (+)-**9** [55,56,57,58,59,60,61,62,63,64].

3.4 Synthesis of 2,5-Dideoxy-2,5-Iminoalditols

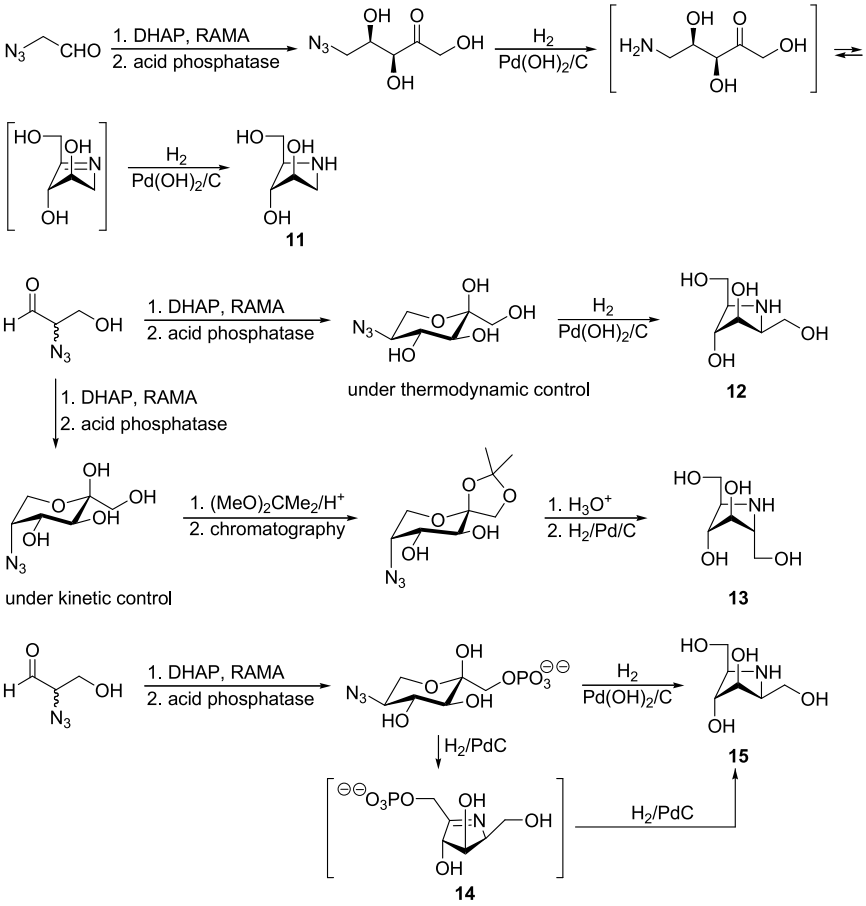
When 2-azidoaldehydes are used as substrates in the RAMA-catalyzed aldol reaction with dihydroxyacetone phosphate, the azidoketones so-obtained can be reduced into the corresponding primary amines that equilibrate with imine intermediates, the reduction of which generate the corresponding pyrrolidines (*Scheme 12*) [55,56,57,58,59,60,61,62,63,64,92,93,94]. 1,4-Dideoxy-1,4-imino-D-arabinitol **11** was prepared from azidoacetaldehyde. Both (2*R*,5*R*)- and (2*S*,5*R*)-bis(hydroxymethyl)-(3*R*,4*R*)-dihydroxypyrrolidine **12** and **13** were derived from racemic 2-azido-3-hydroxypropanol. The aldol resulting from a kinetic control was converted into the (2*R*,2*R*) derivative **12**, whereas the product resulting from a thermodynamic control gave the (2*S*,5*R*)-stereomer **13** [92,93,94]. Similar transformations with 3-acetamido-2-azidopropanal gave aza sugars structurally related to *N*-acetylglucosamine [95]. The Pd-catalyzed inductive aminations of the azidoketones are stereoselective. 6-Deoxyaza sugars and their analogs can also be prepared by direct reductive amination of the aldol products prior to removal of the phosphate group. The reaction is thought to involve an imine 6-phosphate intermediate **14** as exemplified by the synthesis of **15** (*Scheme 12*).

One of the most efficient methods to generate 2,5-dideoxy-2,5-iminogalactitol **16** relies on the fuculose-1-phosphate aldolase-catalyzed aldol condensation of 2-azido-3-hydroxypropanal with dihydroxyacetone monophosphate (*Scheme 13*). The same method applied to (2*R*)-2-azidopropanal (*R*)-**17** and to (2*S*)-2-azido-propanal (*S*)-**17** allows one to prepare 2,5,6-trideoxy-2,5-imino-D-allitol **18** and 2,5,6-trideoxy-2,5-imino-L-talitol **19**, respectively [55,56,57,58,59,60,61,62,63,64].

A facile synthesis of (3*R*,5*R*)-dihydroxy-L-homoproline, an idulonic acid mimic, was realized using L-threonine aldolase-catalyzed reaction of glycine with an aldehyde derived from L-malic acid [96].

3.5 Synthesis of Deoxythiohexoses

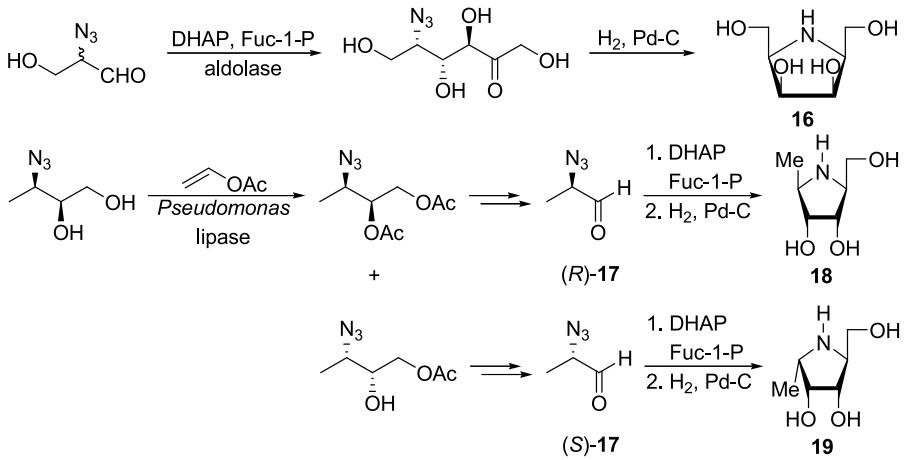
Very successful has been the aldolase-catalyzed aldol reaction as exemplified in *Scheme 14* [97]. The required (*R*)-3-thioglyceraldehyde (**20**) is obtained from regioselective epoxide ring opening of (*S*)-glycidaldehyde diethyl acetal with thioacetic acid and its potassium salt. Con-



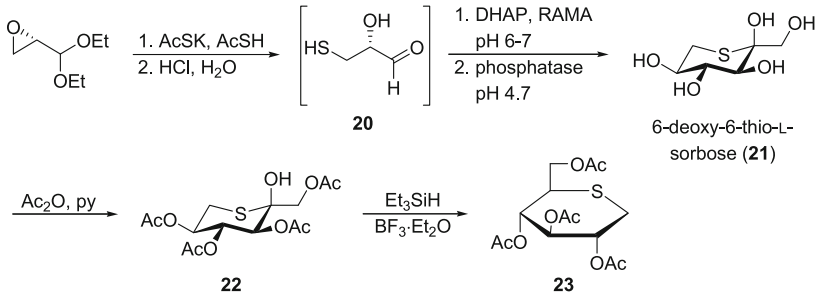
Scheme 12
Examples of chemoenzymatic syntheses of 2,5-dideoxy-2,5-iminoalditols based on RAMA-catalyzed aldol reactions

densation of the thioaldehyde **20** with DHAP-catalyzed by fructose 1,6-diphosphate aldolase from rabbit muscle, followed by removal of the phosphate group using acid phosphatase yields thio-L-sorbose **21**. Acetylation of **21** generates the tetraacetate **22** which is subsequently reduced under ionic conditions to the peracetate of 1-deoxy-5-thio-D-glucopyranose **23**. Applying similar techniques, 1-deoxy-5-thio-D-galactose, 1-deoxy-5-thio-L-altrose, 1-deoxy-5-thio-D-mannose, 1-deoxy-5-thio-L-mannose and 2-deoxy-5-thio-D-ribose have been prepared [97].

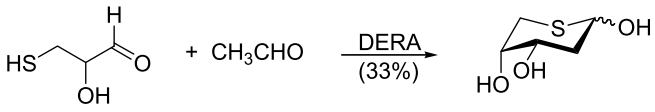
A procedure for large-scale production of 2-deoxy-5-thio-D-*erythro*-pentose (*Scheme 15*) has been developed. It uses a recombinant 2-deoxyribose-5-phosphate aldolase (DERA) from *E. coli* strain DH5α as the catalyst that combines acetaldehyde with racemic 3-thioglyceraldehyde [98].



Scheme 13
Examples of chemoenzymatic synthesis of 2,5-dideoxy-2,5-iminoalditols based on fuculose-1-phosphate aldolase-catalyzed aldol reactions.



Scheme 14
Synthesis of deoxythiosugars based on a RAMA-catalyzed aldol reaction

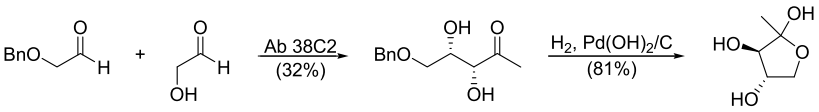


Scheme 15
Synthesis of 2,5-dideoxy-5-thio-D-*erythro*-pentose

3.6 Use of Aldolase Antibodies

Aldolase antibodies 38C2 and 33F12 are able to catalyze both the aldol addition and the retro-aldol reaction [99]. These catalysts have been employed to carry out the kinetic resolution of β-hydroxyketones [100] and have been found to catalyze the asymmetric aldol reactions of 23 donors (ketones) and 16 acceptors (aldehydes) [101]. A highly efficient enantioselective

synthesis of 1-deoxy-L-xylulose utilizing the commercially available aldolase antibody 38C2 has been proposed (Scheme 16) [102].

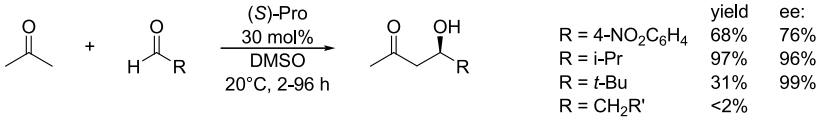


Scheme 16
Synthesis of 1-deoxy-L-xylulose by antibody catalysis

1-Deoxy-D-xylulose has been found as an intermediate in the biosynthesis of thiamine (vit. B_1) and pyridoxal (vit. B_6) [103]. It has been found also to be an alternate nonmevalonate biosynthetic precursor to terpenoid building blocks [104,105].

4 Asymmetric Synthesis of Carbohydrates Applying Organocatalysis

The asymmetric proline-catalyzed intramolecular aldol cyclization, known as the Hajos–Parrish–Eder–Sauer–Wiechert reaction [106,107], was discovered in the 1970s [108,109,110,111]. This reaction, together with the discovery of nonproteinogenic metal complex-catalyzed direct asymmetric aldol reactions (see also Sect. 5.5.1) [112,113,114], led to the development by List and co-workers [115,116] of the first proline-catalyzed intermolecular aldol reaction. Under these conditions, the reaction between a ketone and an aldehyde is possible if a large excess of the ketone donor is used. For example, acetone reacts with several aldehydes in dimethylsulfoxide (DMSO) to give the corresponding aldol in good yields and enantiomeric excesses (ee) (Scheme 17) [117].



Scheme 17
L-Proline-catalyzed asymmetric aldol reactions

In the proline-catalyzed aldol reactions, enolizable achiral aldehydes and ketones are transformed into the corresponding enamines, which can then react with less enolizable carbonyl compounds, even in one-pot protocols. These reactions, unlike most catalytic aldol reactions, do not require preformed enolates, and constitute direct aldol reactions.
Computational studies suggest that the mechanism of the proline catalyzed aldol cyclization is best described by the nucleophilic addition of the neutral enamine to the carbonyl group together with hydrogen transfer from the proline carboxylic acid moiety to the developing alkoxide. A metal-free partial Zimmerman–Traxler-type transition state involving a chair-like arrangement of enamine and carbonyl atoms and the participation of only one proline molecule has been established [118,119]. On the basis of density functional theory (DFT) calculations Córdova and co-workers [120,121] have studied the primary amino acid intermolecular aldol reaction mechanism. They demonstrated that only one amino acid molecule is involved in the

transition state. The calculations explain the origin of stereoselectivity in those reactions and demonstrate that the proposed mechanism through the enamine intermediate can predict the stereochemistry of the reaction.

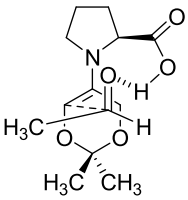


Figure 1
Postulated transition state model

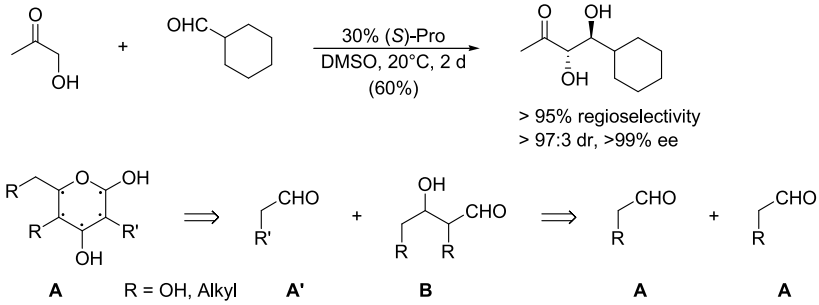
Simple L-alanine, L-valine, L-norvaline, L-isolecucine, L-serine and other linear amino acids [121] or chiral amino acids with a binaphthyl backbone [122] and peptides have also been used as asymmetric catalysts [123,124,125,126]. Solid-supported proline-terminated peptides have been used for heterogeneous catalysis of the asymmetric aldol reaction [127]. Apart from proline and derivatives, other cyclic compounds such as 5,5-dimethyl thiazolidinium-4-carboxylate (DMTC) [128], 2-*tert*-butyl-4-benzyl imidazolidinones [129], (1*R*,2*S*)-2-aminocyclopentanecarboxylic acid [130], (*S*)-5-(pyrrolidin-2-yl)tetrazole, (*S*)-1,3-thiazolidine-4-carboxylic acid, (*S*)-5,5-dimethyl-1,3-thiazolidine-4-carboxylic acid, and (*S*)-hydroxyproline are effective catalysts in asymmetric aldol reactions [126,131,132,133,134,135].

4.1 Synthesis of Ketoses

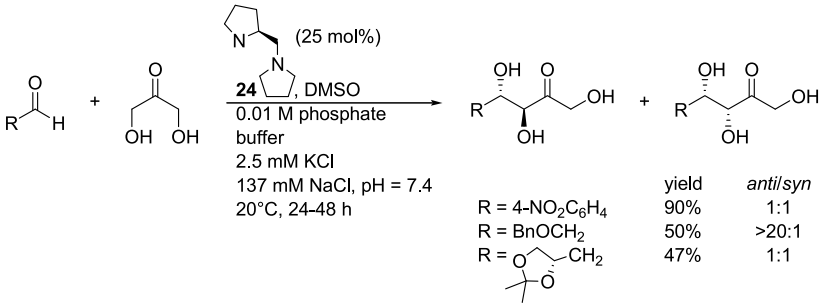
The asymmetric introduction of a hydroxy group at the α position of a carbonyl function has been carried out through organocatalytic aldol reaction and provides a new method for the de novo synthesis of carbohydrates [136] among other biologically important compounds such as antibiotics, terpenes, or alkaloids. Notz and List [137] have reported the L-proline-catalyzed aldol reaction between the hydroxyacetone and cyclohexane carboxaldehyde that furnish a pentulose framework in 60% yield with good regio- and diastereoselectivity (dr) and with high enantioselectivity (*Scheme 18*).

This procedure provides a good method for the construction of 1,2-*anti*-aldol moieties that are less accessible by the Sharpless asymmetric dihydroxylation (see *Scheme 37*, *Scheme 58*, *Scheme 92*, *Sect. 11*) [138] because the corresponding *Z*-olefins are difficult to obtain and show reduced enantioselectivity. The first demonstration of the use of the biologically significant substrate dihydroxyacetone as a donor in organocatalyzed aldol reaction was reported by Barbas III and co-workers [139]. The reactions of DHA with protected glyoxal and glyceraldehydes, in aqueous media and in the presence of enantiomerically pure diamine **24**, provide access to pentuloses and hexuloses, respectively (*Scheme 19*).

The use of protected dihydroxyacetone (eg. **25**) improves considerably the stereochemical outcome of the reaction. In this regard, Barbas III and co-workers [140] have reported the organocatalyzed aldol reaction of dihydroxyacetone variants such as 1,3-dioxan-5-one and



Scheme 18
Proline-catalyzed aldol reactions and retrosynthesis of a carbohydrate framework



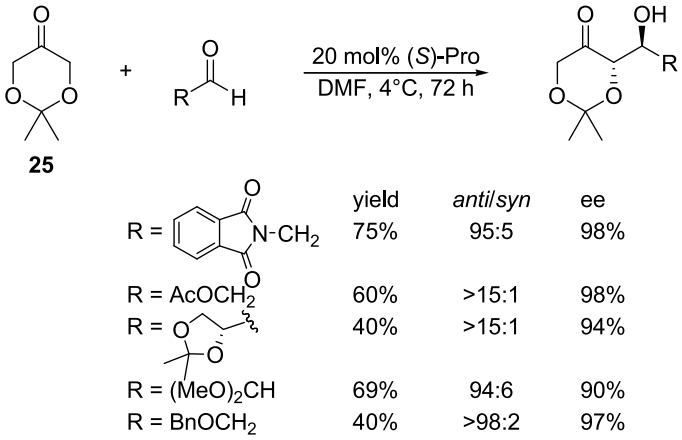
Scheme 19
Direct organocatalytic aldol reaction in buffered aqueous media (1:1 DMSO/H_2O)

2,2-dimethyl-1,3-dioxan-5-one with aldehydes in the presence of (*S*)-proline ((*S*)-Pro) and (*S*)-2-pyrrolidine-tetrazole. Reactions of 2,2-dimethyl-1,3-dioxan-5-one with appropriate aldehydes provide access to L-ribulose and D-tagatose (*Scheme 20*).

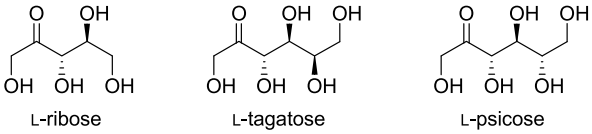
Enders and co-workers [141,142,143] also reported highly diastereo- and enantioselective direct organocatalytic aldol reactions of **25** with appropriate aldehydes in the presence of (*S*)-proline.

In this way various protected carbohydrates and aminosugars are obtained. There is a matching correspondence between α-branched (*S*) or (*R*)-configurated aldehydes and (*S*) or (*R*)-proline, respectively. Thus, the reaction of **25** with the (*R*)-configurated 2,3-di-*O*-isopropylidene-D-glyceraldehyde gives the double acetonide of D-psicose in 76% yield. Acidic deprotection with Dowex gives the parent D-psicose. A similar route has been reported by Córdova and co-workers [144].

The L-alanine catalyzed reaction of **25** and $BnOCH_2CHO$ gives 5-*O*-benzyl-1,3-di-*O*-isopropylidene-L-ribulose [121]. The direct asymmetric intermolecular aldol reactions are also catalyzed by small peptides. For instance, in the presence of 30 mol% of L-Ala-L-Ala in DMSO containing 10 equivalents of H_2O, **25** reacted with 4-cyanobenzaldehyde giving the corresponding aldols with an *anti*/*syn* ratio of 13:1 and ee of 99% for the *anti* aldol (65% yield) [145].



Scheme 20
Stereoselective L-proline-catalyzed aldol reaction



Structure 1

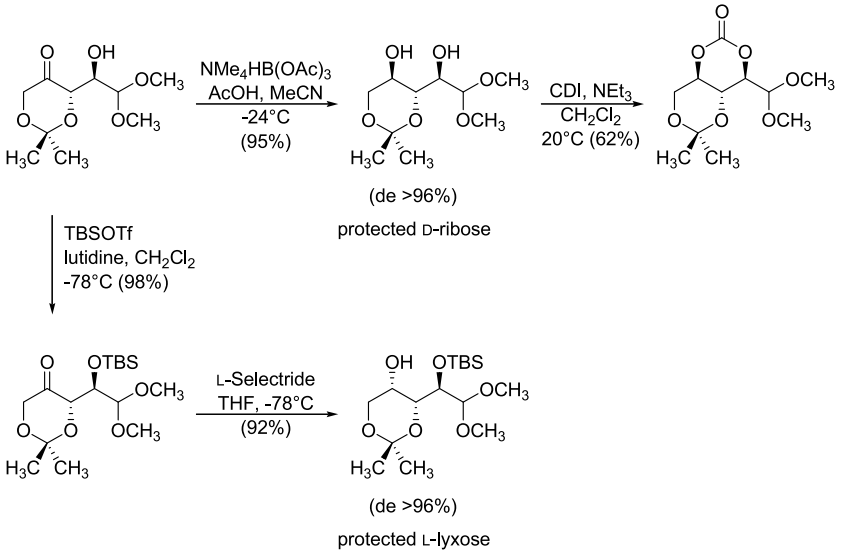
4.2 Synthesis of Aldoses

Aldopentoses such as D-ribose and L-lyxose have been prepared applying the methodology reported by Enders and co-workers [141,142,143], followed by stereoselective reduction and acetal hydrolysis (*Scheme 21*).

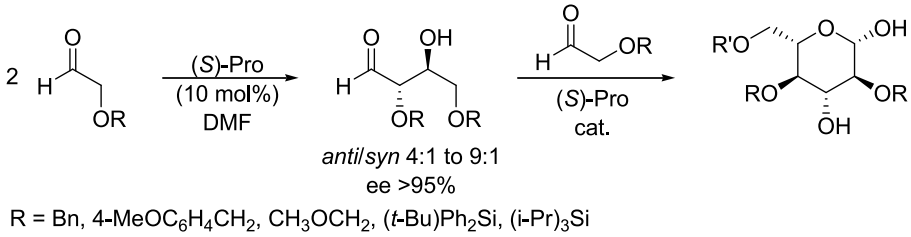
McMillan and co-workers [146] have reported the first example of direct enantioselective aldehyde–aldehyde cross-aldol reaction using small molecules as catalysts. Subsequently, they have described the enantioselective dimerization and cross-coupling of α-oxygenated aldehydes to provide erythrose architecture. A second L-proline-catalyzed aldol reaction generates hexoses (*Scheme 22*) [147].

Combining the D-erythrose derivative **26** obtained by L-proline-catalyzed dimerization of (t-Bu)Ph_2SiOCH_2CHO with enoxysilane **27** in Mukaiyama aldol reactions catalyzed by various Lewis acid, MacMillan and co-workers have realized efficient, two-step syntheses of semi-protected D-glucose (**28G**), L-mannose (**28M**) and L-allose (**28A**) (*Scheme 23*) [148]. Using D-proline to generate tetrose **29** and its condensation with **27b**, the semi-protected D-glucose derivative **30** was obtained in two steps [149].

The enamine geometry **32** is crucial for the stereocontrol in organocatalytic aldehyde–aldehyde couplings; amines of type **31** are convenient catalysts for enantioselective enamine-aldol reactions. Examples are shown in *Scheme 24* [126,131,132,133,134,135].

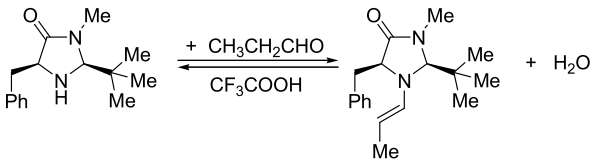


Scheme 21
Ender's synthesis of aldoses



R = Bn, 4-MeOC$_6$H$_4$CH$_2$, CH$_3$OCH$_2$, (*t*-Bu)Ph$_2$Si, (i-Pr)$_3$Si

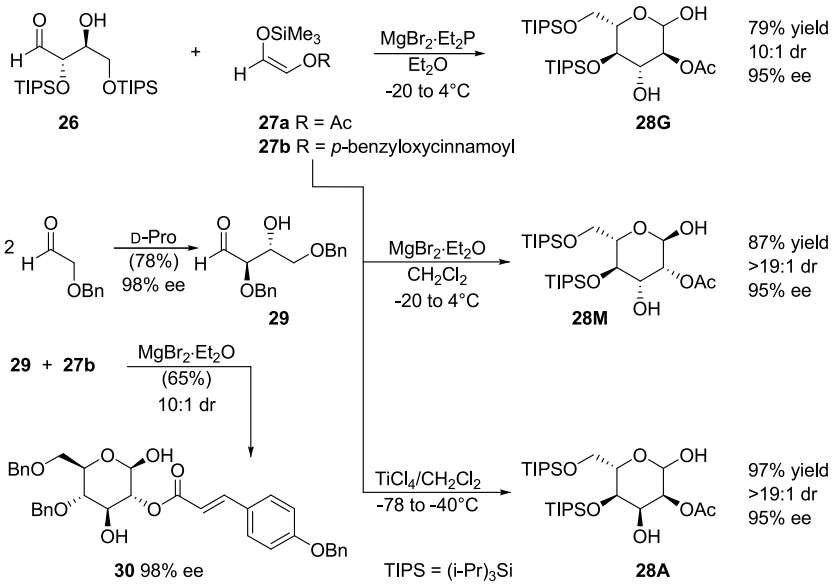
Scheme 22
MacMillan's synthesis of hexoses



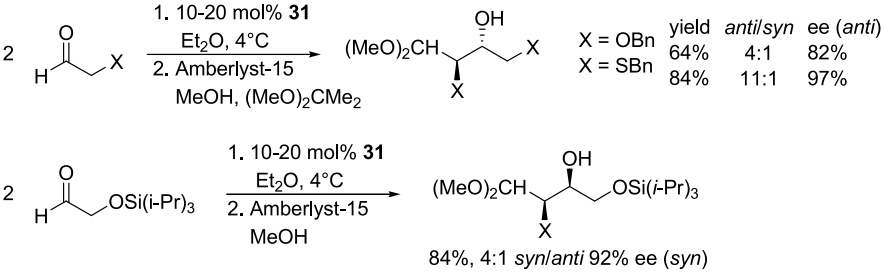
31 enamine-iminium **32**

Structure 2

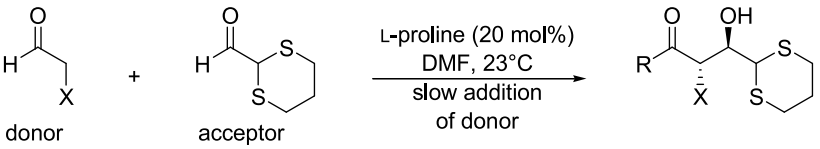
Importantly, with α-silyloxy acetaldehyde, the *syn* aldol is the major dimer (threose derivative). Thus, applying Mukaiyama condensations with **27** (see *Scheme 23*), hexoses such as idose, gulose, and galactose can be prepared. A highly stereoselective protocol for the cross coupling of aldehydes and ketones with α-thioacetal aldehydes has been developed



Scheme 23
Two-step syntheses of D-glucose, D-mannose and D-allose derivatives

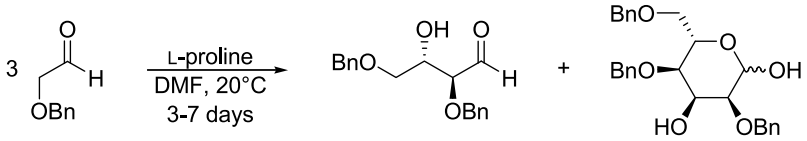


Scheme 24
diastereoselective aldol dimerizations



	yield	*anti*/*syn*	ee (*anti*)
R = H, X = Me	80%	16:1	99%
R = H, X = OSiMe$_2$(*t*-Bu)	52%	13:1	70%
R = Me, X = H	91%		98%
R = Me, X = OH	88%	>20:1	>99%

Scheme 25
Cross-aldol reactions catalyzed by L-proline

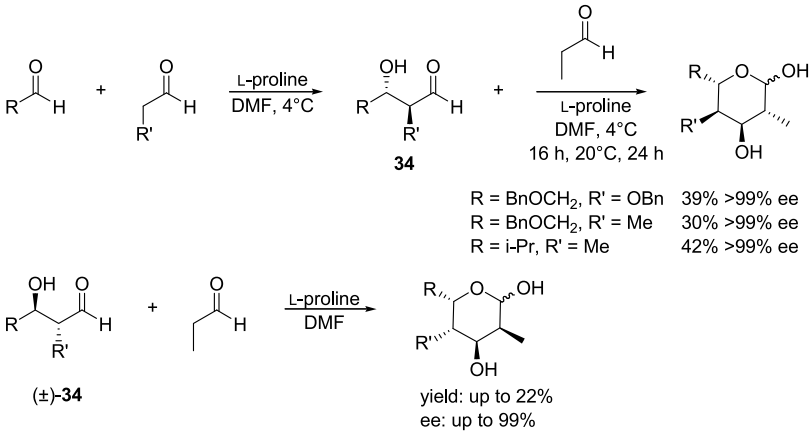


Scheme 26
Córdova's one-step synthesis of L-allose

(*Scheme 25*) [150]. The latter acts as an acceptor only because of its good electrophilic and non-nucleophilic character. The α-thioacetal functionality in this enantioselective cross-coupling allows access to highly oxidized, stereo-defined synthons of broad versatility. Moreover, the observed reactivity profile makes them pre-eminent substrates for highly selective cross-aldol reactions with ketone donors.

Córdova and co-workers have studied the double aldol reaction of benzyloxyacetaldehyde using various α-aminoacids as catalysts. With L-proline and hydroxy-L-proline a tetrose and the L-allose derivative **33** were obtained in 41 and 28% yield, respectively, and with an enantiomeric excess higher than 98% (*Scheme 26*). As expected, with D-proline as catalyst, the corresponding D-allose derivative was obtained with the same ease in a one-pot operation [151].

Out of the 16 possible stereoisomers, a single one is obtained with 99% ee. The same authors reported that the same amino acids were also efficient organocatalysts in water demonstrating the neogenesis of carbohydrates under prebiotic conditions using glycolaldehyde as the substrate. With regard to the synthesis of deoxy- and polyketide sugars, Córdova and coworkers also reported an enantioselective de novo synthesis of both enantiomers of natural or unnatural hexoses with up to 99% ee. This implied tandem two-step sugar synthesis based on direct aminoacid-catalyzed selective iterative aldol reaction with aldehydes [152]. When using racemic aldehydes (±)-**34**, propanal and L-proline catalyst, dynamic kinetic asymmetric

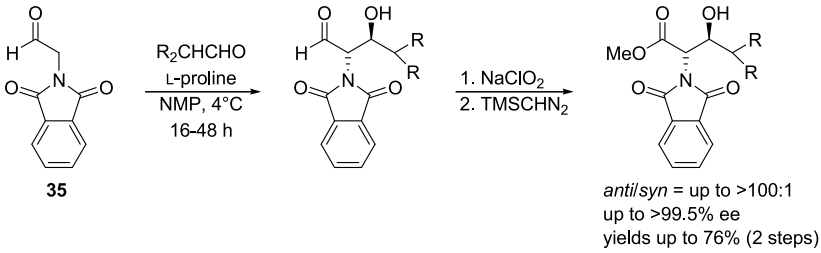


Scheme 27
Córdova's two-step syntheses of deoxyaldoses polyketides

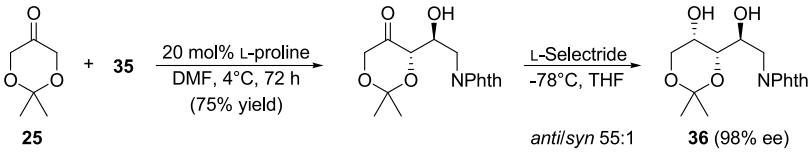
transformations occur with excellent stereoselectivities. Thus, triketide and deoxysugars can be obtained in one-step operations with up to 99% ee [153] (Scheme 27). In these reactions, the donor-aldehyde is converted into an enamine by reaction with the aminoacid catalyst, in a process analogous to the biosynthetic aldol reactions catalyzed by class I aldolases.
Silyl protected glycoaldehydes have been used also for these tandem direct aminoacid-catalytic asymmetric aldol reactions giving rise to hexoses with free hydroxyl groups at C-3 and C-1. This allows the introduction of orthogonal protecting groups in the monosaccharide. This is of importance for oligosaccharide synthesis. Further oxidation furnishes the corresponding lactones. Darbre and co-workers [154,155] have reported a Zn-proline-catalyzed aldolization of glycoaldehyde and *rac*-glyceraldehyde that give mainly tetroses and pentoses.

4.3 Synthesis of Aminosugars by Aldol and Mannich Reactions

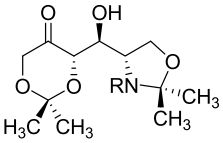
Direct organocatalytic asymmetric aldol reaction of α-aminoaldehydes **35** with other substituted aldehydes furnishes β-hydroxy-α-aminoaldehydes with high *anti*-stereoselectivity. This procedure is of importance for the synthesis of α-aminosugars and derivatives. Additionally, the oxidation of aldehydes gives rise to highly enantiomerically enriched *anti*-β-hydroxy-α-amino acids (Scheme 28) [156].
Barbas III and co-workers have used the aldol organocatalyzed condensation between **25** and achiral aldehydes **35** for the preparation of aminosugars **36** (Scheme 29) [140].
The aldol reactions of **25** with appropriate aldehydes in the presence of L-proline have been also used by Enders and co-workers [141,142,143] for the preparation of aminosugars D-*erythro*-pentos-4-ulose, 5-amino-5-deoxy-L-psicose (**37a**), and 5-amino-5-deoxy-L-tagatose (**37b**) derivatives.



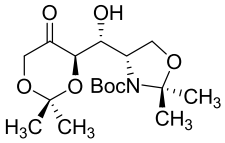
Scheme 28
Barbas' two-step synthesis of *anti*-β-hydroxy-α-aminoacids



Scheme 29
Barbas' synthesis of aminoalditols

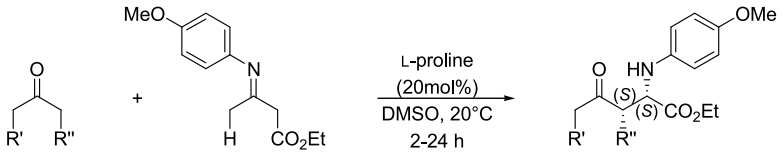


37a R = Boc, Cbz

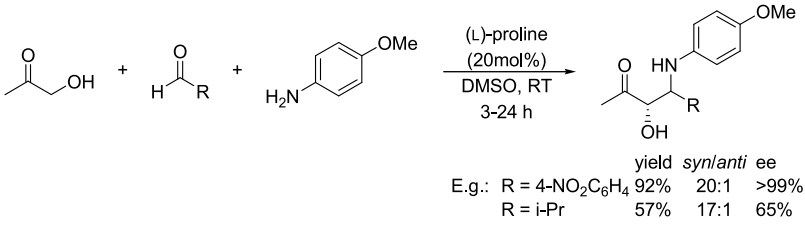


37b

Structure 3



Scheme 30
Barbas' L-proline catalyzed asymmetric Mannich reactions



Scheme 31
List's asymmetric Mannich reactions

Barbas III and co-workers reported organocatalyzed Mannich reactions between *p*-methoxyphenyl-protected α-imino ethyl glyoxolate and ketones (*Scheme 30*) [157,158,159].

The proline-catalyzed Mannich reaction has been applied also by List and co-workers (*Scheme 31*) [160,161]. In their method enolizable aldehydes, ketones, and a primary amine are mixed together with a substoichiometric quantity of L- and D-proline to give the desired β-aminocarbonyl compounds. When applied to hydroxyacetone the method furnishes 4-amino-4-deoxytetruloses.

The reaction exhibited opposite enantiofacial selectivity to the proline-catalyzed aldol reaction. The attack to the *si*-face is preferred. An explanation for this enantiofacial selectivity has been proposed by List that is based on the transition state models shown in *Fig. 2*.

Enders and co-workers [162] have reported a protocol for the synthesis of aminopentoses and aminohexoses based on the use of 2,2-dimethyl-1,3-dioxan-5-one (**25**) as the ketone donor in a three-component Mannich reaction with several aldehydes and *p*-anisidine in the presence of L-proline or (*tert*-butyl)dimethylsilyloxy-L-proline as organocatalysts.

Córdova and co-workers [163] reported simultaneously a similar approach for the synthesis of protected 4-amino-4-deoxy-threo-pentulose and 4-amino-4-deoxyfructose (*Scheme 32*). The catalyst can be L-proline, other α-aminoacids, or alanine-tetrazole [126].

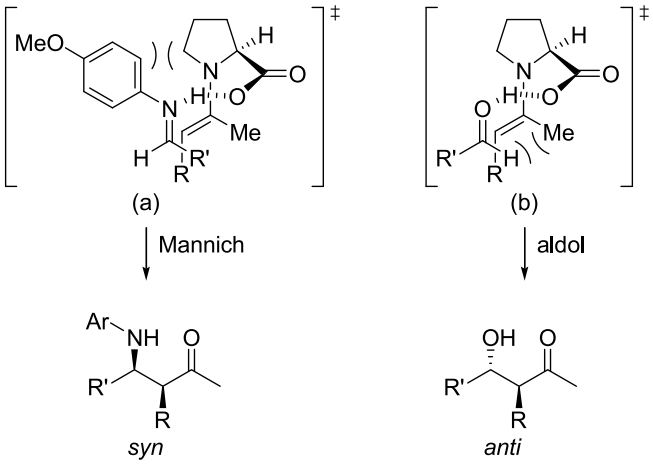
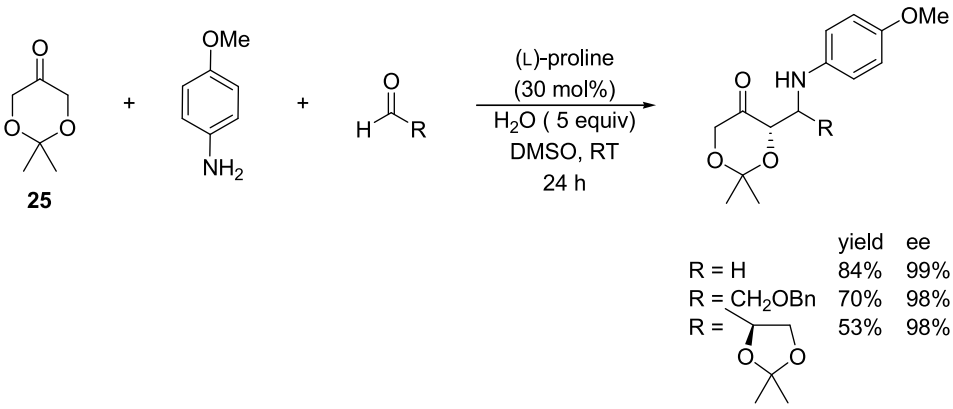


Figure 2
Proposed transition states for the L-proline-catalyzed asymmetric Mannich and aldol reactions



Scheme 32
One-step syntheses of aminosugars

The three-component Mannich reactions with various donor aldehydes have been studied also by Hayashi and co-workers [164], giving rise, after reduction, to several aminopolyols with high *syn*-diastereo and enantioselectivities.

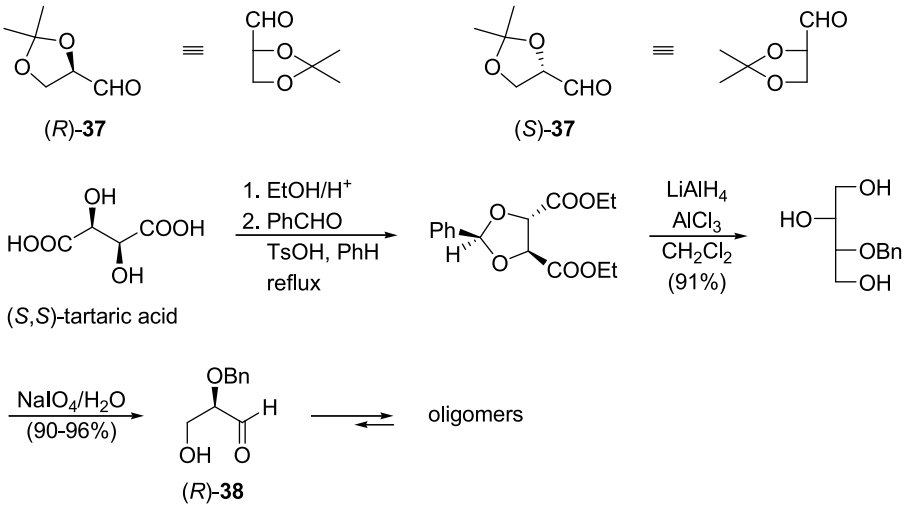
5 Chain Elongation of Aldehydes through Nucleophilic Additions

Chemical asymmetric cross-aldol condensations using enantiomerically pure Lewis acids as promoters (instead of an aldolase or α-aminoacid) have been applied to prepare monosaccharides and analogs [165,166]. If enantiomerically pure aldehydes (such as diol-protected D- or L-glyceraldehyde) are available, they can be chain-elongated by one, two, or more

carbon centers with high diastereoselectivities. The classical Kiliani–Fischer cyanohydrin synthesis [167,168,169] is a milestone in carbohydrate chemistry and has been used in numerous applications [170,171]. Nevertheless, diastereoselectivity of the nucleophile addition is often low and the harsh reaction conditions that are required to reveal the chain-elongated aldose from either their aldonic acid or directly from the cyanohydrin are serious drawbacks. Currently, there are many more flexible methods to carry out one-carbon homologations of aldehydes, including the reductive end of aldoses that will be presented below. Aldehyde allylation with allyl boronates [172,173,174,175] or with allylstannanes [176,177,178] emerge as quite useful because of their high diastereoselectivity and the diversity of modifications that can be applied to the allylic alcohols. With achiral aldehydes enantiomerically pure allylic and allenyl stannanes can be used in the asymmetric synthesis of monosaccharides and analogs [176,177,178,179].

5.1 Total Synthesis of D- and L-Glyceraldehyde and other C-3 Aldose Derivatives

The acetonide of D-glyceraldehyde (*R*-**37**: (*R*)-2,3-*O*-isopropylidene-D-glyceraldehyde) is most simply obtained from D-mannitol. D-glyceraldehyde has been derived also from D-fructose, L-glyceraldehyde from L-sorbose [180]. The acetonide of L-glyceraldehyde ((*S*)-**37**, (*S*)-2,3-*O*-isopropylidene-L-glyceraldehyde) is usually derived from ascorbic acid [181]. The aldehydes (*R*)-**37** and (*S*)-**37** are not very stable as monomers and undergo racemization on storage. Derivative (*R*)-**38** (2-*O*-benzylglyceraldehyde) has been proposed as an alternative to (*R*)-**37**. It is obtained from (*S*,*S*)-tartaric acid as shown in ● *Scheme 33* [182].

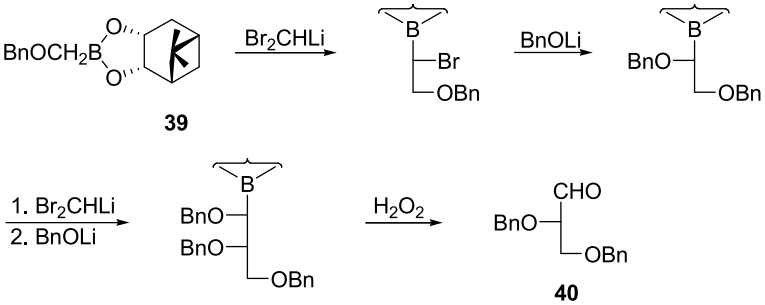


Scheme 33
Synthesis of 2-*O*-benzyl-D-glyceraldehyde

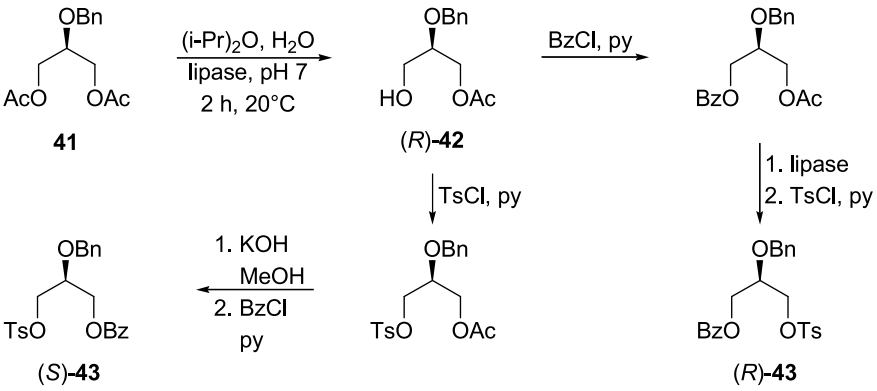
Enantiomer (*S*)-**38** can be derived from (*R*,*R*)-tartaric acid in the same way. (*R*,*R*)-Tartaric acid is obtained in large quantities from potassium hydrogen tartrate, a waste product of wineries. Racemic tartaric acid is synthesized [183] on a large scale from maleic anhydride and H_2O_2. Its resolution is carried out either by crystallization, or by enzymatic or microbiological enantiodifferentiating conversions. Thus, both (*S*,*S*)- and (*R*,*R*)-(−)-tartaric acid are supplied by the industry inexpensively [184].

Chain extension using an insertion reaction of dichloromethyllithium or dibromomethyllithium with (*S*)-pinanediol (via [(benzyloxy)methyl]boronate **39**) has been used to generate L-C_3, L-C_4, and L-C_5-aldoses [185]. In order to obtain 2,3-*O*-dibenzyl-L-glyceraldehyde **40**, the insertion reaction has to be applied twice (◉ *Scheme 34*). By repeating the process two more times, L-ribose has been prepared in this way with high enantiomeric purity [185].

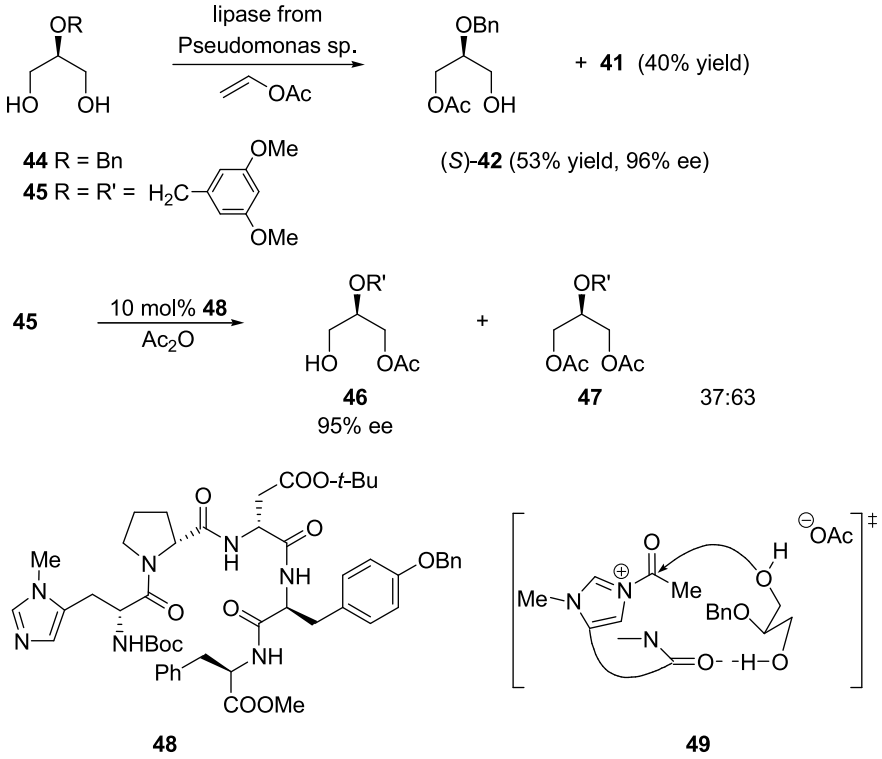
The synthesis of 3-*O*-methyl-D-glyceraldehyde starts with D-fructose [186]. The preparation of 2-*O*-methyl-D-glyceraldehyde and 2-*O*-benzyl-D-glyceraldehyde ((*R*)-**38**) starts from D-mannitol [187]. Enantiomerically pure derivatives of glycerol can be prepared on a large scale through the lipase (pig pancreas, EC 3.1.1.3)-catalyzed hydrolysis of prochiral



Scheme 34
Asymmetric chain elongation of dibromomethyllithium



Scheme 35
Desymmetrization of *meso*-diacetate by lipase-catalyzed hydrolysis synthesis of C_3-alditol derivatives



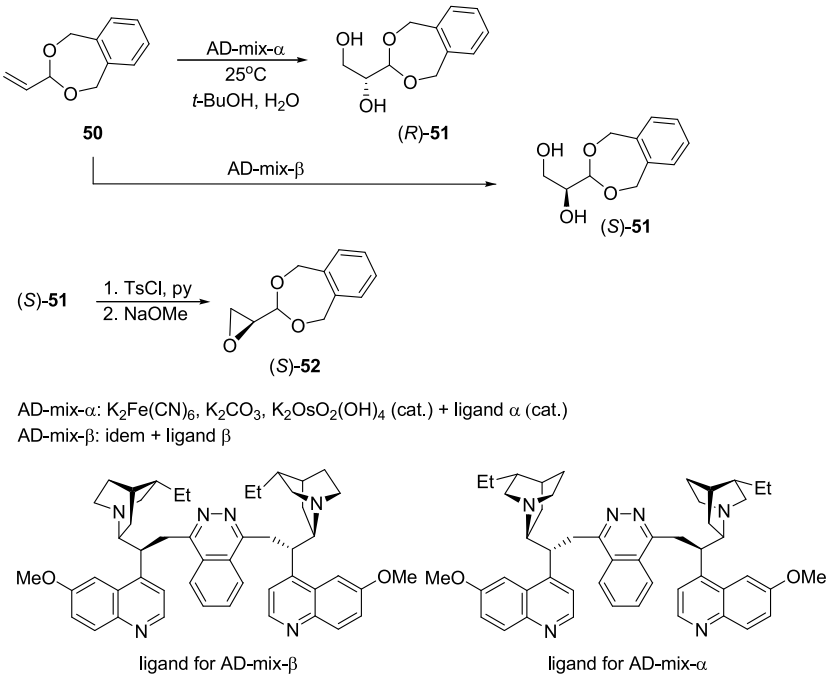
Scheme 36
Peptides as catalysts for enantioselective acetylation of alcohols

diacetate **41**. The procedure gives (*R*)-**42** (45% yield, 88% ee) which can be converted into crystalline derivative (*R*)-**43** or (*S*)-**43** (>99% ee) as shown in *Scheme 35* [188].

Instead of applying enantioselective hydrolysis of *meso* diacetates monoacetylation of *meso*-diols can generate enantiomerically enriched monoesters. The catalyst can be an esterase in vinyl acetate (e. g. **44** → (*S*)-**42**) or a short peptide derivative (e. g. **45** → **46** catalyzed by **48**) as shown in *Scheme 36*. Transition state **49** has been proposed for the asymmetric monoacetylation of diol **45** with acetic anhydride catalyzed by peptide **48** [189].

The (*R*)- and (*S*)-benzyl epoxypropyl ether have been derived from *O*-benzyl-L-serine [190]. Stable and easily-handled protected forms of L- and D-glyceraldehyde are obtained by the Sharpless asymmetric dihydroxylation of the benzene-1,2-dimethanol acetal (**50**) of acrolein (*Scheme 37*). The method produces either diol (*R*)-**51** or (*S*)-**51** with 97% ee after recrystallization from benzene. These diols can be converted into useful C-3 chiral building blocks, for instance epoxides (*R*)-**52** and (*S*)-**52**, respectively [191].

Derivatives of D- and L-glyceraldehydes such as 2-amino-2-deoxyglyceraldehyde (serinal), 3-deoxyglyceraldehyde (2-hydroxypropanal), and 2,3-dideoxy-2-aminoglyceraldehyde (2-aminopropanal) are used extensively to construct rare monosaccharides and analogs through chain-elongation applying nucleophilic addition to their carbonyl moiety. Semi-



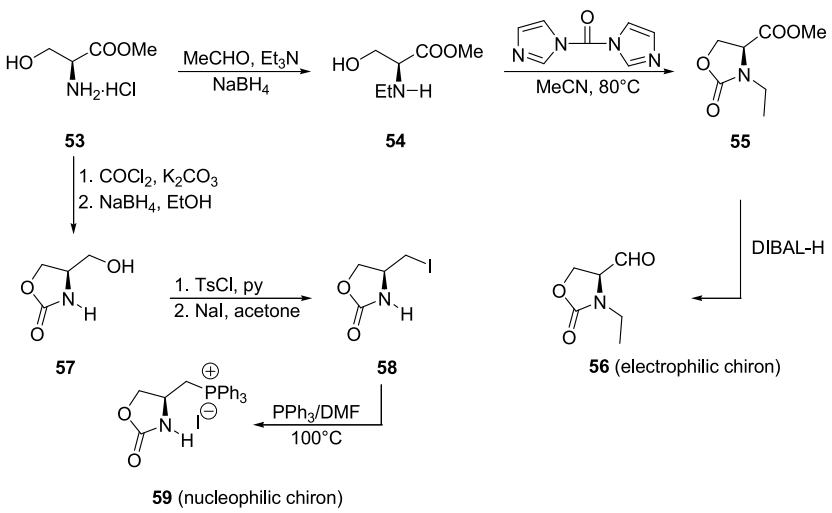
Scheme 37
Sharpless asymmetric dihydroxylation applied to the syntheses of C_3-sugar precursors

protected (*R*)- and (*S*)-2-hydroxypropanols are most simply derived from the readily available D-(−)-lactic and L-(+)-lactic acids, respectively [192]. (*S*)-2-Benzyloxypropanal can be obtained via benzylation of ethyl L-lactate, followed by reduction with $LiAlH_4$ and Swern oxidation. *N*-(*t*-Butoxycarbonyl)-L-alaninal can be obtained with high enantiomeric purity by $LiAlH_4$ reduction of the *N*-methoxy-*N*-methyl-α-(*t*-butoxycarbonylamino)carboxamide of alanine [193]. Alternatively, *N*-9-(9-phenylfluorenyl)-L-alaninal has been derived from L-alanine [194].

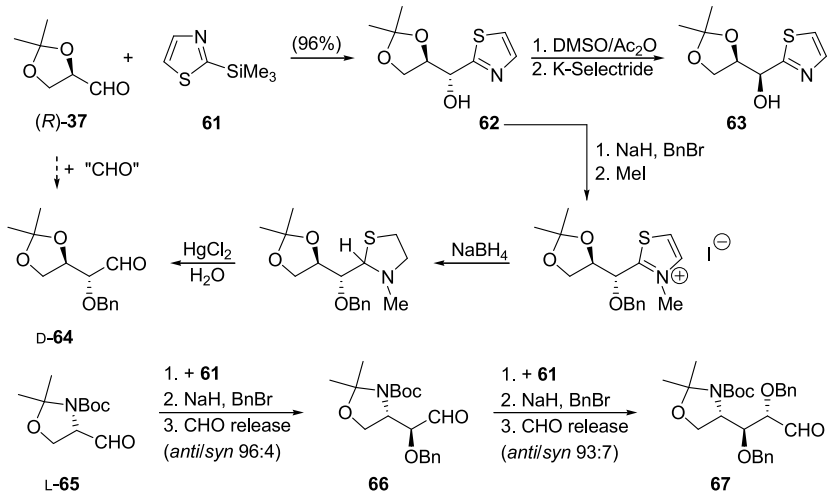
The *N*-ethyloxazolidinone **56** (◉ *Scheme 38*) is obtained from L-serine by treating (*S*)-serine methyl ester hydrochloride with Et_3N, acetaldehyde, and $NaBH_4$ to give *N*-ethylamine **54**. Oxazolidinone formation with carbonyldiimidazole leads to **55**, the reduction of which generates aldehyde **56** [195]. A nucleophilic alaninol synthon **59** has been derived from **53** by protection of the alcohol and amine moieties as a carbamate obtained by treatment with phosgene. Reduction of the ester gives the corresponding alaninol **57** which is tosylated, then displaced successively with iodide and triphenylphosphine to generate **59** (◉ *Scheme 38*) [196].

5.2 One-Carbon Homologation of Aldoses: the Thiazole-Based Method

Dondoni and co-workers [197,198,199,200,201,202,203,204,205,206] have shown that homologation of α-hydroxycarbaldehydes and α-hydroxylactones can be achieved with high



Scheme 38
Synthesis of electrophilic and nucleophilic C_3-chiron containing masked 2-amino moieties



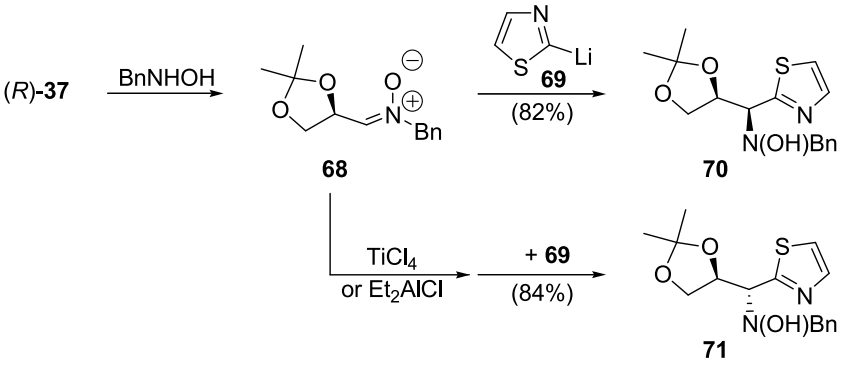
Scheme 39
Dondoni's one-carbon chain elongation

anti-selectivity by addition of 2-(trimethylsilyl)thiazole (**61**) (*Scheme 39*). For instance, D-glyceraldehyde (*R*)-**37** reacts with **61** giving **62** in 96% yields and *anti* vs. *syn* diastereoselectivity better than 95:5. Release of the carbaldehyde moiety requires protection of the alcohol as a benzyl ether, methylation of the thiazole to generate intermediate **63** that is not isolated but reduced in situ with $NaBH_4$ to give the corresponding thiazoline. Mercury(II)-

catalyzed hydrolysis liberates the semiprotected D-erythrose derivative D-**64** in 62% overall yield [207]. Methylation of the thiazole moiety can use methyl triflate instead of MeI, and copper(II) chloride can be used instead of mercury(II) chloride [208]. The iterative addition and unmasking protocols were repeated over several consecutive cycles so that the chain elongation of the triose (*R*)-**37** was brought up to a nonose derivative with all-*anti* configuration of the polyol.

For the preparation of *syn* isomers, alcohol **62** has to be oxidized into the corresponding ketone, which is reduced with K-Selectride into the *syn* isomer **64**. The α-amino aldehyde L-**65**, derived from L-serine, was converted into aminotetrose and pentose derivatives **66** and **71**, respectively. The *anti* diastereoselectivity observed for **67** of **61** to the *N*,*N*-diprotected α-amino aldehyde L-**65** can be reversed to *syn* selectivity by using a *N*-monoprotected derivative [209,210].

The reactions of *N*-benzylnitrone **68** derived from D-glyceraldehyde acetonide (*R*)-**37** with 2-lithiothiazole (**69**) gives the *syn* adduct **70** with 92% diastereoselectivity. Interestingly, the same reaction applied to **69** precomplexed with Et_2AlCl or $TiCl_4$ gives the *anti* diastereomer **71** preferentially in high yield. The method has been applied to the synthesis of all kinds of aminosugars including D-nojirimycin [204,211,212,213].

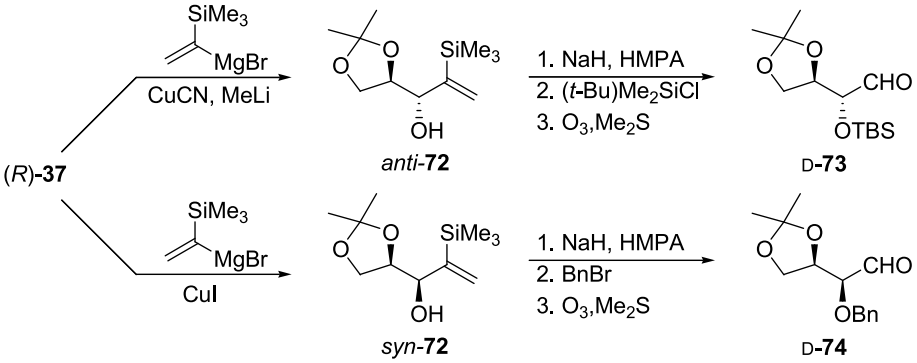


Scheme 40
Dondoni's synthesis of amino sugars

5.3 Other Methods of One-Carbon Chain Elongation of Aldoses

An alternative method (*Scheme 41*) for the homologation of D-glyceraldehyde derivative (*R*)-**37** to derivatives of D-erythrose D-**73** and D-threose D-**74** has been proposed by Kusakabe and Sato [214]. Reaction of (*R*)-**37** with appropriate 1-(trimethylsilyl)vinyl-copper reagents leads to either *anti* or *syn* stereoselective adducts *anti*-**72** *anti* (*anti*/*syn* 20:1) or *syn*-**72** *syn* (*syn*/*anti* 98:2) in 87% yield. Alcohol protection, followed by ozonolysis furnishes D-**73** and D-**74**, respectively.

The nitroaldol condensation with nitromethane (Henry's reaction), followed by Nef decomposition of the resultant nitronate under strongly acidic conditions has been used to elongate aldehydes. For instance, *N*-acetyl-D-mannosamine has been converted into *N*-acetylneu-



Scheme 41
***Anti* and *syn* diastereoselective vinylations of D-glyceraldehyde acetonide**

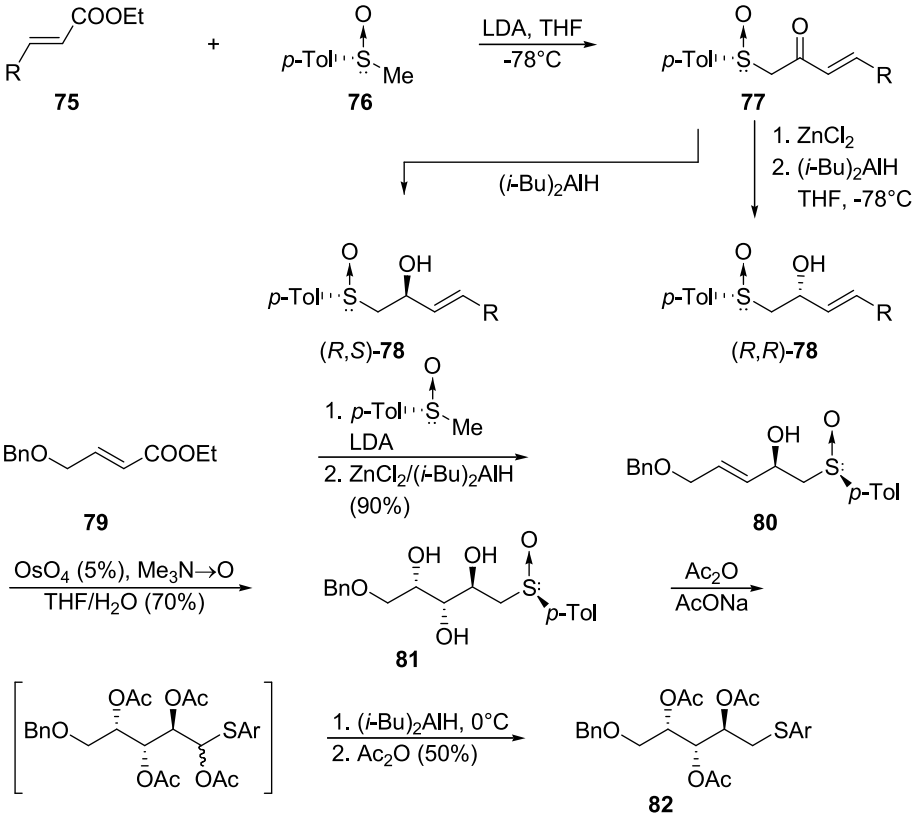
raminic acid applying this method iteratively [215]. Et_3N-catalyzed addition of CH_3NO_2 to 1,4:3,6-dianhydrofructose and subsequent Pd-C-catalyzed hydrogenation afforded 2-*C*-aminomethyl-1,4:3,6-dianhydromannitol [216]. Chikashita and co-workers [217] have reported good levels of *anti* diastereoselectivity better than 99% in an iterative *syn*-selective homologation sequence using 2-lithio-1,3-dithiane [218,219] with 2,3-*O*-cyclohexylidene-D-glyceraldehyde. Addition of MeMgI to sugar-derived aldehydes is highly diastereoselective when the aldehyde is precomplexed with $ZnCl_2$ in CH_2Cl_2/Et_2O at –78 °C. The alcohol so-obtained gives access to L-fucose, 1,5-dideoxy-1,5-imino-L-fucitol, and 5-thio-L-*fuco*pyranose [220].

The one-carbon elongation of aldoses to ketoses using iodomethyllithium adding to the corresponding aldonolactones has been used by Bessières and Morin to convert D-mannose into D-*manno*-hept-2-ulose, and L-arabinose into L-fructose [221]. The addition of Grignard reagent to lactols or aldehydes derived form D-glucose, or D-mannose, has allowed the preparation of higher-carbon sugars [222]. Syntheses of 3-deoxy-2-ulosonic acids have utilized the reactions of ketene dithio-acetals obtained via Horner–Emmons or Peterson olefination of 2-deoxy-1,5-hexono-lactones [223]. Wittig methylenation of 2,3,4-tri-*O*-benzyl-6-*O*-(4-methoxybenzyl)-D-glucopyranose gives an enitol that is converted into 7-*O*-(4-methoxybenzyl)-4,5-di-*O*-benzyl-3,6-anhydro-L-*ido*-hept-1-enitol. Further steps generate (3*S*,4*R*,5*S*)-3,4-dibenzyloxy-2-methylidene-5-vinyltetrahydrofuran. The latter allyl-vinyl ether is isomerized into a mixture of (1*R* and 1*S*,2*R*,3*S*)-1-hydroxy-2,3-dibenzyloxycyclohept-4-ene on treatment with $(i\text{-}Bu)_3Al$ in CH_2Cl_2 [224]. Wittig methylenation of 2,3-*O*-isopropylidene-D-ribofuranose and subsequent Malaprade oxidation ($NaIO_4/H_2O/CH_2Cl_2$) gives a γ,δ-unsaturated aldehyde that generates L-ribose after alkene dihydroxylation and acetonide hydrolysis [225]. L-apiose has been obtained following a similar synthetic procedure [225].

5.4 Additions of Enantiomerically Pure One-Carbon Synthon

The addition of (+)-(*R*)-methyl *p*-tolylsulfoxide (**76**) to carboxylic esters gives the corresponding β-ketosulfoxides. When applied to α,β-unsaturated esters **75**, the ketosulfoxides **77** so-obtained are reduced with high diastereoselectivity with $LiAlH_4$ or $(i\text{-}Bu)_2AlH$

(DIBAL-H) giving optically β-hydroxy sulfoxides such as (*R*,*S*)-**78** and (*R*,*R*)-**78** [226]. When applied to ester **79**, this method generates allylic alcohol **80** that undergoes highly diastereoselective osmium-catalyzed dihydroxylation giving **81**. A Pummerer rearrangement and subsequent reduction with $(i\text{-Bu})_2$AlH and acetylation furnish the L-arabinitol derivative **82** (Scheme 42) [227].

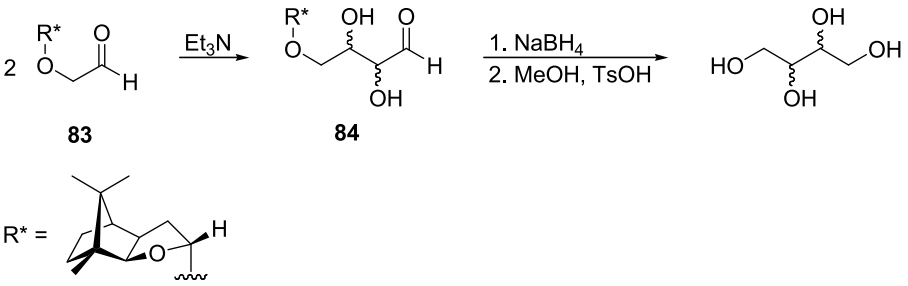


Scheme 42
Asymmetric synthesis of L-arabinitol derivatives via stereoselective dihydroxylation of an enantiomerically pure allylic alcohol

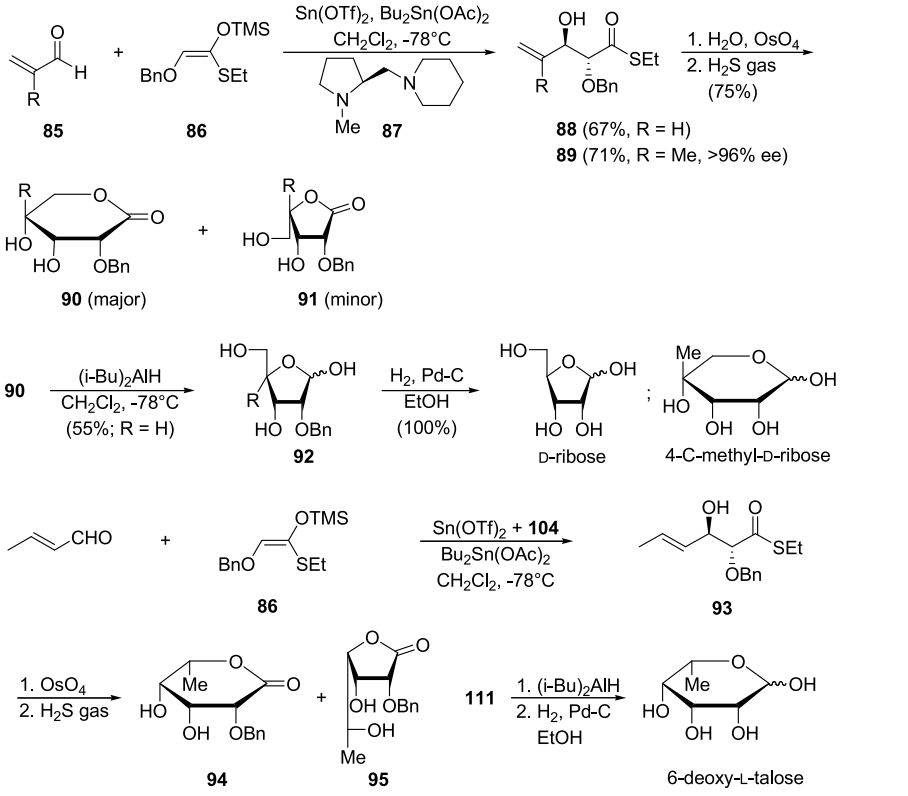
5.5 Two-Carbon Chain Elongation of Aldehydes

5.5.1 Asymmetric Aldol Reactions

Enantiomerically pure glycolaldehyde derivatives **83** undergo aldol condensations in the presence of Et_3N giving mixtures of erythrose and threose derivatives **84** (Scheme 43) for which the erythrose/threose ratio reaches 58:42 and the L/D ratio 62:38 [228,229].

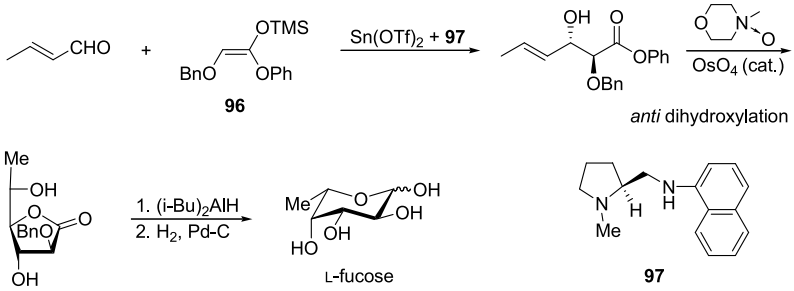


Scheme 43
Asymmetry induced by chiral auxiliary in the aldol condensation of glycoaldehyde



Scheme 44
Mukaiyama's asymmetric aldol reactions: total synthesis of D-ribose, 4-*C*-methyl-D-ribose, and 6-deoxy-L-talose

A very elegant asymmetric synthesis of D-ribose from achiral starting materials has been presented by Mukaiyama and co-workers [165]. It is based on the cross-aldolization of crotonaldehyde (**85**, R = H) and enoxysilane **86** in the presence of an enantiomerically pure diamine **87**, the chiral inducer (*Scheme 44*). High diastereoselectivity (*anti*/*syn* > 98:2) and high



Scheme 45
Kobayashi's total synthesis of L-fucose

enantioselectivity (>97% ee for *anti* aldol) are observed. The *anti*-aldol **88** is then doubly hydroxylated with moderate facial selectivity to give a 72:28 mixture of aldonolactones **90** and **91**. Reduction of the major lactone **90** provides **92**, the debenzylation of which furnishes D-ribose. The same method has been applied to prepare 4-*C*-methyl-D-ribose and 6-deoxy-L-talose starting with methacrolein (**85**, R = Me) and (*E*)-but-2-enal ((*E*)-crotonaldehyde), respectively (*Scheme 44*) [230,231].

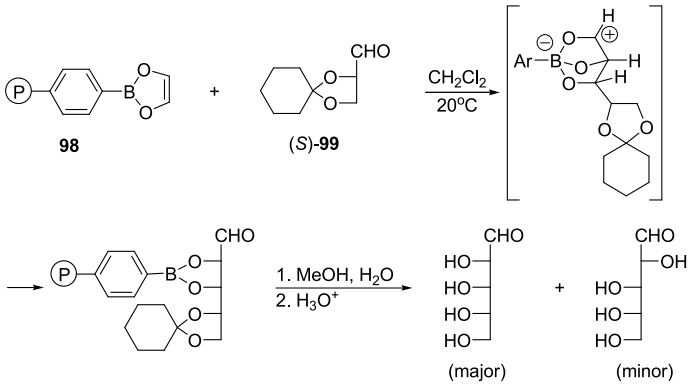
Applying an analogous method, Kobayashi and Kawasuji [166] have prepared L-fucose from (*E*)-crotonaldehyde and the benzyloxyketene acetal **96** in four steps and 49% overall yield (*Scheme 45*). The asymmetric aldol condensation is catalyzed by a complex made of $Sn(OTf)_2$ and chiral diamine **97**.

5.5.2 Nucleophilic Additions to Enantiomerically Pure Aldehydes

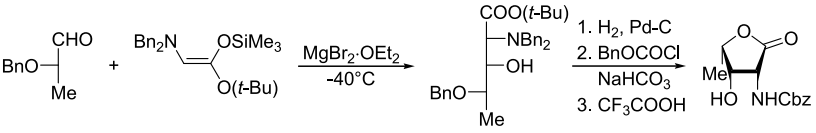
Mukaiyama and co-workers have pioneered many routes to the total syntheses of rare carbohydrates such as the 2-amino-2-deoxypentoses. In 1982 they reported that the potassium enolate derived from the magnesium salt of a (*R*)-atrolactic acid derivative adds to 2,3-*O*-isopropylidene-D-glyceraldehyde in a highly stereoselective manner giving, after alcohol protection, imine hydrolysis and amine protection, a D-arabinopentonate derivative [231]. Further elaboration leads to 2-acetamido-2-deoxy-D-arabinose. In a similar fashion, starting from (*S*)-atrolactic acid, 2-acetamido-2-deoxy-D-ribose has been prepared [232,233,234,235,236].

Several syntheses of aminodeoxypentoses have employed a similar approach in which a three-carbon starting material is condensed with a two-carbon entity [237]. A synthetic equivalent of the glycoaldehyde anion, the dioxaborole **98** (*Scheme 46*) has been used for the carbon chain elongation of aldehydes. Thus L-ribose is prepared from the addition of 2,3-*O*-cyclohexylidene-L-glyceraldehyde (*S*)-**99** to **98**. Double addition and higher addition reaction to yield polymers is alleviated by using a polymer-supported reagent [238].

Ethyl (*S*)-lactate has been the primary source of chirality in several syntheses of aminodeoxycarbohydrates. The derivative **101** of 2-amino-2-deoxy-L-lyxonic acid is the major product of condensation of 2-*O*-benzyl L-lactaldehyde with silyl ketene acetal **100**. The derived ester **101** can be converted into lactone **102** (*Scheme 47*) [239] an intermediate for the synthesis of L-daunosamine and L-vancosamine [240,241]. L-Mannose and L-altrose



Scheme 46
Synthesis of L-ribose with a polymer-supported glycoaldehyde anion equivalent



Scheme 47
Synthesis of a 2,5-dideoxy-2-aminoaldonolactone

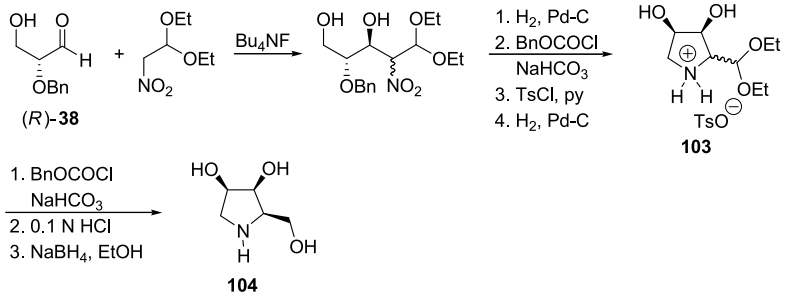
derivatives have been derived from 2,3-*O*-isopropylidene-L-glyceraldehyde via its reaction with 3-C-lithiated 5,6-dihydro-1,4-dithiine-2-yl[(4-methoxybenzyl)oxy]methane [242]. The same approach combining this organolithium reagent with (*R*)-benzyl gylcidyl ether leads to 4-deoxy-L-hexose derivatives [243].

5.5.3 Nitro-Aldol Condensations

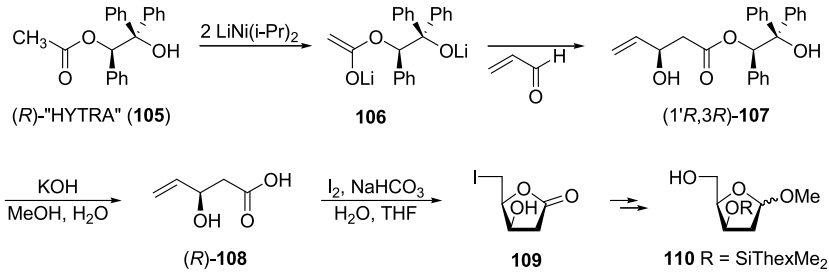
The syntheses of D- and L-2-amino-2-deoxy-arabinose and of 1,4-dideoxy-1,4-imino-D-lyxitol **104** have been achieved via the nitro-aldol condensation (Henry's reaction) of 2-*O*-benzyl-D-glyceraldehyde (*R*)-**38** and the diethyl acetal of nitroacetaldehyde (*Scheme 48*) which gives a 88:12 mixture of the *arabino* and *ribo*-adducts. Their reduction and subsequent protection of the amines so-obtained, then selective tosylation of the primary alcohol and hydrogenolysis gives **103**, which is then converted into **104** [244].

5.5.4 Nucleophilic Additions of Enantiomerically Pure Enolates

Braun's enantiomerically pure acetate **105** ((*R*)-"HYTRA") [245] can be converted to the lithium enolate **106**. Subsequent addition to acrolein predominantly gives (1′*R*,3*R*)-**107** (diastereoselectivity: 92:8). Alkaline hydrolysis of (1′*R*,3*R*)-**107** provides (*R*)-**108** with 83% ee (*Scheme 49*). On treatment of (*R*)-**108** with (*S*)-1-phenylethylamine and recrystallization,



Scheme 48
Synthesis of 1,4-dideoxy-1,4-imino-D-lyxitol



Scheme 49
Diastereoselective addition aldol reaction with (*R*)-"HYTRA"

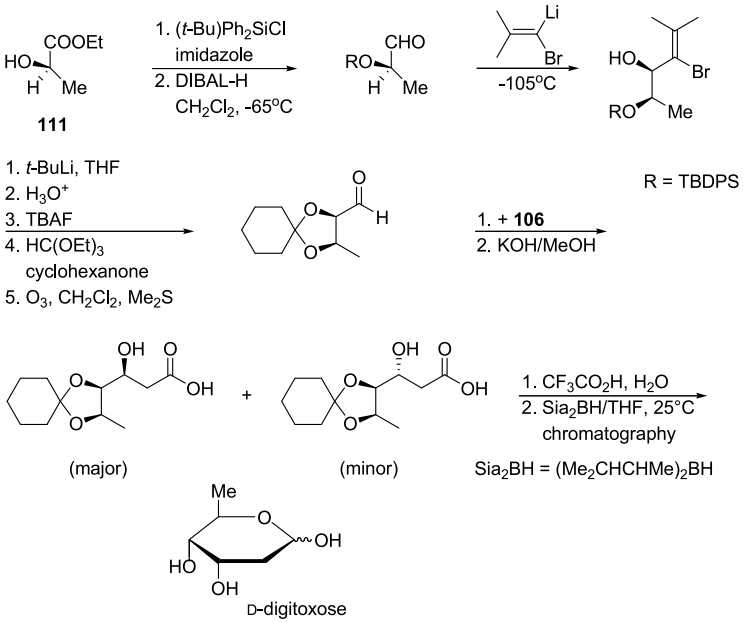
(*R*)-**108** is isolated in 42% overall yield and >99% ee. This compound has been converted into **109** by iodolactonization. It is a precursor of all kinds of 2-deoxyfuranosides [246].

D-Digitoxose, a component of cardiac glycosides found in *Digitalis purpurea* and other higher plants, can be prepared following a similar method starting from ethyl (*R*)-lactate (**111**) and (*R*)-HYTRA (*Scheme 50*) [247].

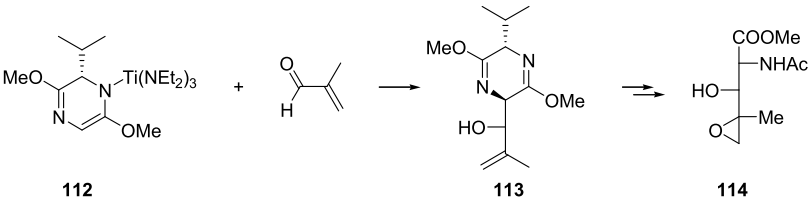
The chiral titanated bislactim ether **112** undergoes 1,2-addition to α,β-unsaturated aldehydes. With methacrolein it gives adduct **113**, the epoxidation of which and subsequent hydrolysis forms the branched-chain amino-acid derivative **114** (*Scheme 51*) [248].

The phenylalanine-derived oxazolidinone **115** undergoes a diastereoselective aldol reaction with crotonaldehyde to give the *syn* product **116**. Formation of the Weinreb amide, followed by silylation of the crude secondary alcohol provides **117**. The chiral auxiliary is recovered at this stage. Product **117** is a potential precursor of all kinds of monosaccharides and analogs, including of 1-deoxynojirimycin (*Scheme 52*) [249].

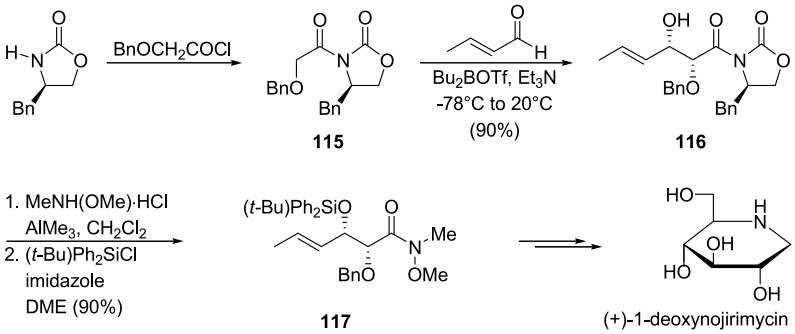
The C_{33}–C_{37}-unit of (+)-calyculin A (a marine natural product) is an amide derived from 5-*O*-methyl-4-deoxy-4-dimethylamino-D-ribonic acid that has been prepared by Evans and co-workers [250]. *N*-Protection of sarcosine as benzyl carbamate affords acid **118** which is activated and used to *N*-acylate the (*S*)-phenylalanine-derived oxazolidinone. This gives **119** that is methoxymethylated diastereoselectively (98:2) to give **120**. Reductive removal of the chiral auxiliary, followed by Swern oxidation forms aldehyde **121** with little racemization if



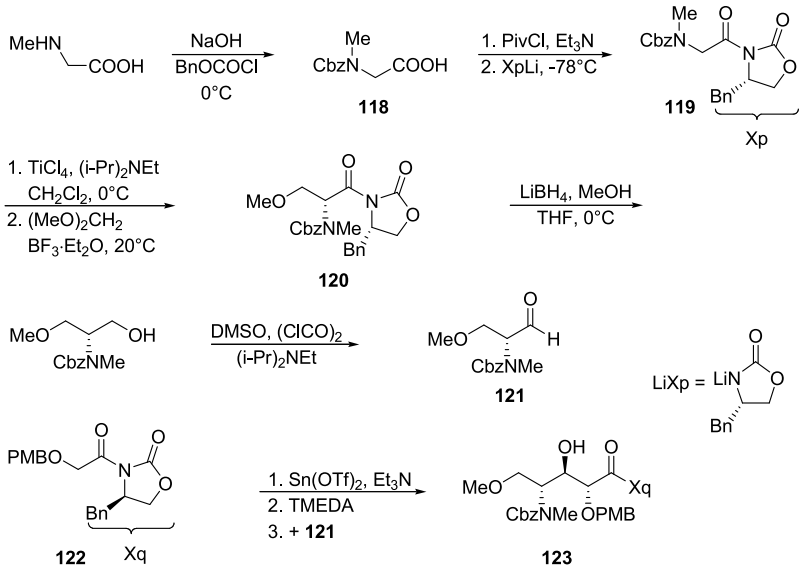
Scheme 50
Braun's synthesis of D-digitoxose



Scheme 51
Use of an enantiomerically pure bislactim ether



Scheme 52
Use of an Evans' homochiral enolate



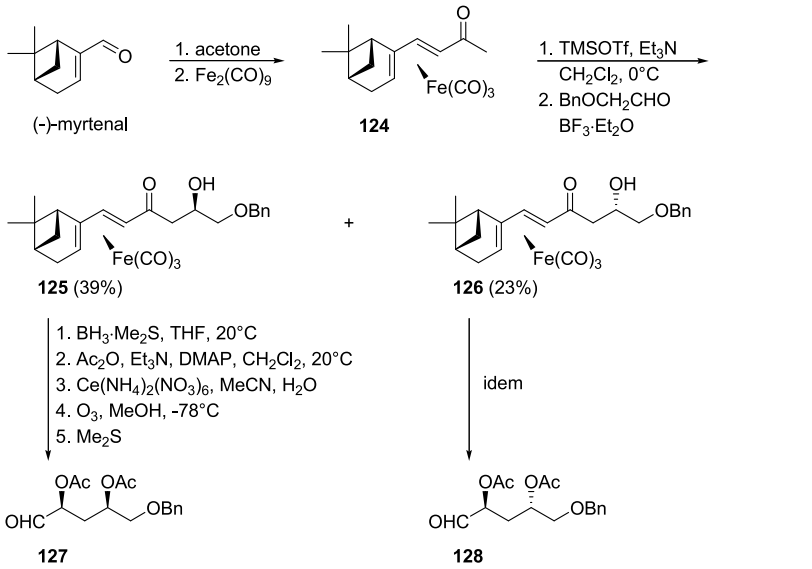
Scheme 53
Synthesis of a 4-deoxy-4-dimethylamino-D-ribonic acid derivative

the Hünig base (*i*-Pr$_2$NEt) (i-Pr = isopropyl) is used instead of the usual Et$_3$N. Enolization of imide **122**, followed by addition of tetramethylethylenediamine and then of aldehyde **121** gives rise to the *anti*-aldol **123** (60%) accompanied by 24% of other diastereomers (*Scheme 53*). Compound **123** has been used for amide formation with primary amines.

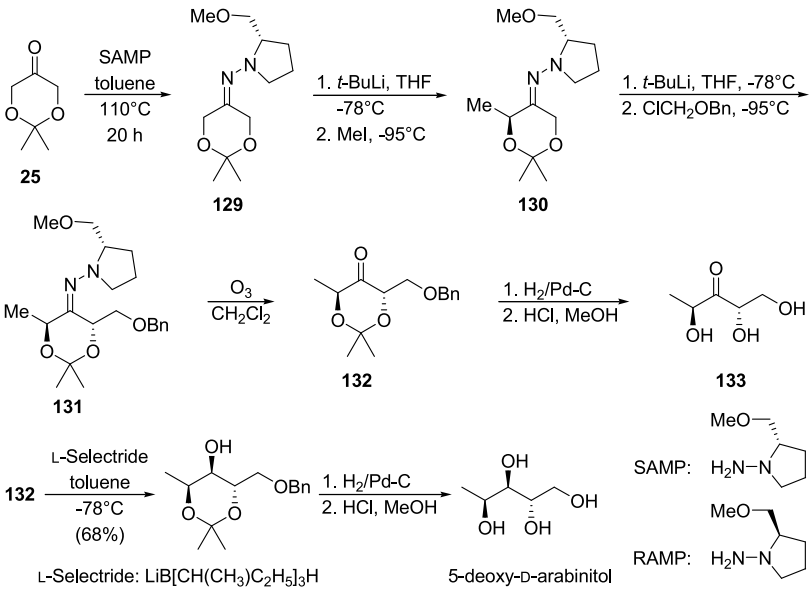
An enantioselective synthesis of 3-deoxypentoses from (−)-myrtenal has been proposed by Franck-Neumann and co-workers (*Scheme 54*) [251]. It features the Mukaiyama cross-aldolization of benzyloxyacetaldehyde and the tricarbonyliron complex **124** derived from the condensation of (−)-myrtenal with acetone. The diastereomeric aldols **125** and**126** are separated and converted (*Scheme 54*) into 3-deoxypentoses **127** and **128** [251].

Enders and Jegelka [252,253] have used 1,3-dioxan-5-one **25** to construct enantiomerically pure C$_5$- to C$_9$-deoxycarbohydrates. For example, reaction of **25** with SAMP [(*S*)-(−)-amino-2-(methoxymethyl)pyrrolidine] gives hydrazone **129** which is deprotonated and alkylated with methyl iodide to yield **130**. The monoalkylated hydrazone is then alkylated in the same manner with chloromethyl benzyl ether to form **131**. Cleavage of the hydrazone with ozone furnishes the protected ulose **132** (>98% de, >98% ee) which is deprotected to (−)-5-deoxy-L-*threo*-3-pentulose **133**. Reduction of **132** with L-Selectride and subsequent deprotection provides 5-deoxy-D-arabinitol (>95% de, >95% ee) (*Scheme 55*).

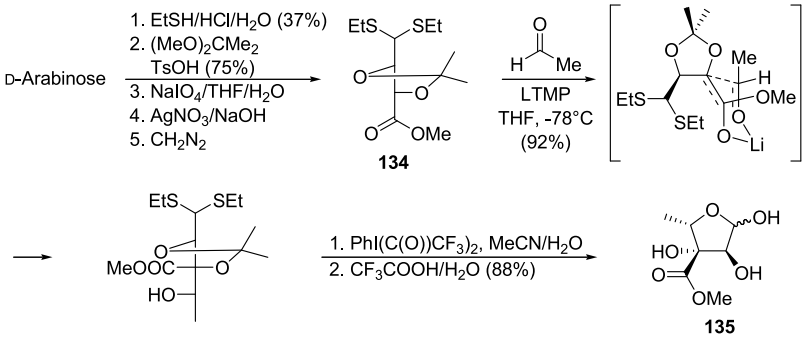
Seeberger and co-workers [254] have reacted ester **134** with acetaldehyde using a lithium salt of 2,2,4,4-tetramethylpiridine (LTMP; lithium 2,2,6,6-tetramethylpiperidide) as the base. This gives a 4:1 mixture of diastereomers, the major of which was treated with PhI(C(O)CF$_3$)$_2$ and then with aqueous CF$_3$COOH to generate L-aceric acid methyl ester **135** (*Scheme 56*).



Scheme 54
Franck-Neumann's synthesis of 3-deoxypentoses



Scheme 55
Enders' synthesis of 5-deoxy-D-arabinitol

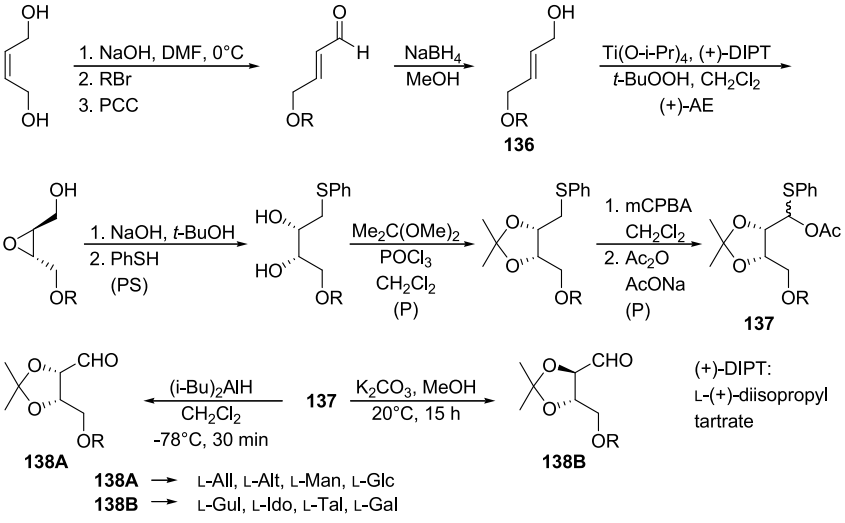


Scheme 56
Synthesis of aceric acid methyl ester

5.5.5 Aldehyde Olefination and Asymmetric Epoxidation

Wittig olefination of D- or L-glyceraldehyde acetonide with $Ph_3P{=}CHCHO$ gives, after reduction of the enal with diisobutylaluminum hydride, (*E*)-allylic alcohols that undergo Katsuki–Sharpless enantioselective epoxidation [255,256,257]. The method has been applied to prepare D-arabinitol (=D-lyxitol) and ribitol (a *meso* alditol) and D-arabinitol (or xylitol, another *meso* alditol) [258].

The Katsuki–Sharpless asymmetric epoxidation of (*E*)-allylic alcohols is the key-step in the total synthesis of all tetroses and hexoses developed by Sharpless and Masamune [259,260] and that are summarized in *Scheme 57* for the L-series. The epoxide obtained by oxidation of



Scheme 57
Sharpless and Masamune's syntheses of tetroses and hexoses

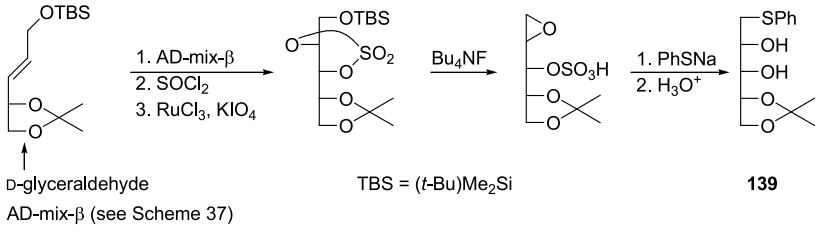
the allylic alcohol **136** undergoes a Payne rearrangement in the presence of NaOH, giving terminal epoxides that open regioselectively by PhSNa to give the corresponding phenylsulfide. After protection of the diol as acetonide, the sulfides are oxidized with metachloroperbenzoic acid into the corresponding sulfoxides that undergo Pummerer rearrangement on treatment with Ac_2O and AcONa liberating, after hydrolysis, the corresponding tetrose derivative **137**. Thus, (*Z*)-but-2-ene-1,4-diol can be converted into four L-tetroses (*Scheme 57*). Tetrose **138A** and **138B** can be reacted with $Ph_3PCHCHO$ giving (*E*)-hex-2-en-1-ol derivatives that undergo the same sequence of reactions as above (PS, P, P). This has permitted the synthesis of the eight L-hexoses. Starting with epoxidation of **136** with (−)-AE instead of (+)-AE, all D-tetrose and D-hexoses are obtained in the same way.

The methodology of Wittig–Horner–Emmons olefination to convert an aldehyde into the corresponding two-carbon chain elongated allylic alcohol and its subsequent asymmetric epoxidation has been used to prepare (+)-galactonojirimycin, (+)-nojirimycin, 1-deoxygalactonojirimycin, and 1-deoxynojirimycin, the method involves the regioselective opening of the epoxides with the azide anion [257,261,262,263,264,265,266,267].

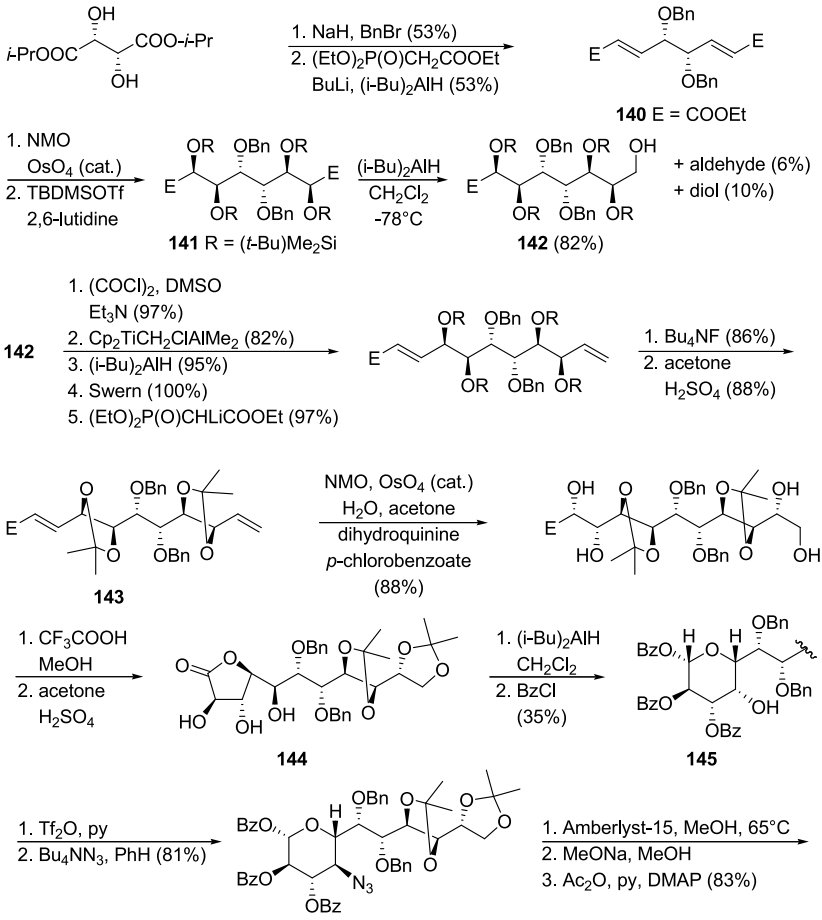
5.5.6 Aldehyde Olefination and Dihydroxylation

An example of the synthesis of a C_5-monosaccharide derivative (**139**) applying Sharpless asymmetric dihydroxylation is shown in *Scheme 58* [268,269,270,271].

Ikemoto and Schreiber [272,273] have prepared (−)-hikizimycin starting from L-(+)-tartaric acid for the hikosamine portion, and from D-glucose for the kanosamine part (for an alternative synthesis see [274]). The synthesis of a suitably protected form of hikosamine follows a two-directional chain strategy with terminus differentiation (*Scheme 59*) [273]. L-(+)-Diisopropyl tartrate, which will provide the C(6) and C(7) stereocenters of the undecose, is benzylated. In the same pot reduction with DIBAL-H and Wittig–Horner–Emmons double-chain elongation provides **140**. Double dihydroxylation of **140** follows Kishi's rule [275], giving a tetrol with high diastereoselectivity that is protected as silyl tetraether **141**. Desymmetrization of the diethyl octadioate **141** is possible with DIBAL-H in CH_2Cl_2 which generates alcohol **142** as the main product. This can be attributed to an entropy effect: once DIBAL-H, a dimeric reagent, has reacted with one of the two carboxylic moieties, a highly polar intermediate (aluminum alcoholate) is formed which blocks a large number of solvent molecules (CH_2Cl_2, "Napalm effect") and thus increases dramatically the mass of the system



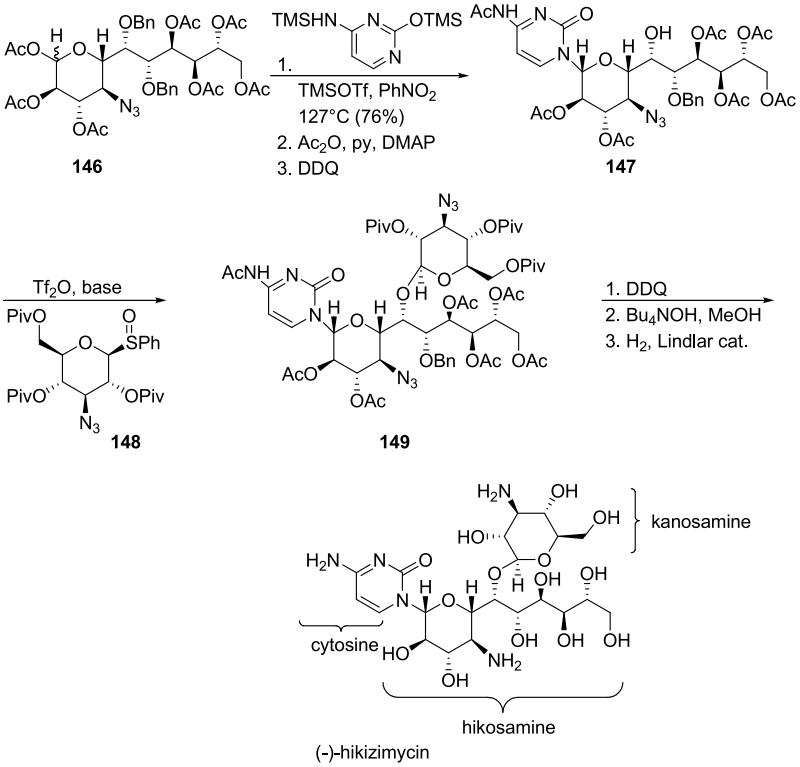
Scheme 58
Sharpless asymmetric dihydroxylation of an allylic silyl ether; configuration inversion via regioselective intramolecular displacement of cyclic sulfate



Scheme 59

Schreiber's synthesis of (−)-hikizimycin featuring a two-directional chain elongation strategy involving aldehyde olefination and face-selective dihydroxylations. **(see also on the following page)**

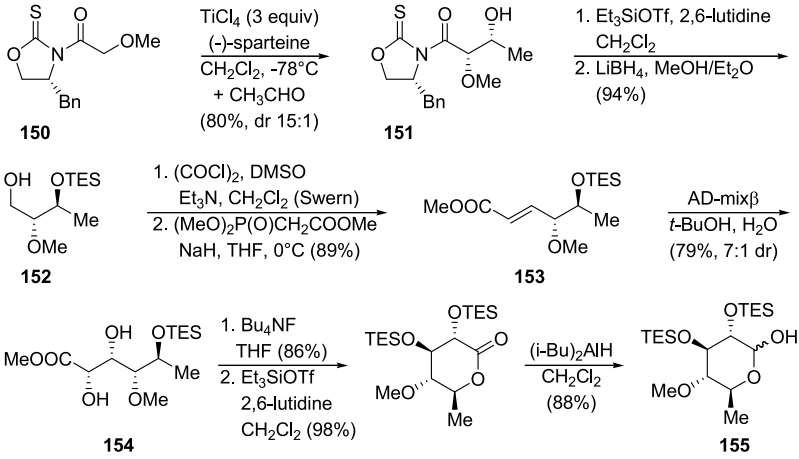
compared with the starting material. The reaction of the second carboxylic moiety is thus retarded because of a more negative entropy of condensation. (The larger the mass of the two reactants, the more negative is the entropy variation of their condensation, translational entropy. If the rate constant ratio of the two successive reductions were $k_1/k_2 = 2$, a maximum yield of 49% would have been obtained for **142**). Swern oxidation of **142**, followed by Tebbe vinylation, ester group reduction into a primary alcohol, then Swern oxidation into the corresponding aldehyde and Wittig–Horner–Emmons olefination generates a diene. Its four silyl ethers are exchanged for two acetonides giving **143**. Double hydroxylation in the presence of dihydroquinine *p*-chlorobenzoate leads to a tetrol with a good diastereoselectivity which gives γ-lactone **144** under acidic conditions. γ-Lactone**144** is then reduced into the corresponding furanose and converted into pyranoside **145**, with an unprotected 4-hydroxy group (steric hindrance). The latter is esterified as a triflate and then displaced with an azide anion.



Scheme 59 (continued)

This operation introduces the nitrogen moiety with the required configuration. Acetonide methanolysis followed by methanolysis of the benzoates and acetylation generates pyranoside **146** which undergoes Vorbrüggen's glycosidation with bis(trimethylsilyl)cytosine giving an intermediate that is acetylated and oxidized with dichlorodicyanobenzoquinone (DDQ). This leads to a site selective debenzylation of the 6-benzyloxy moiety giving **147**. Glycosidation of **147** with **148** under Kahne's conditions [276] gives **149**. Deprotection and hydrogenation of **149** furnishes (−)-hikizimycin.

Phenylalanine-derived oxazolidinone has been used in *Scheme 52* as a chiral auxiliary for asymmetric cross-aldolization (Evans-aldol reactions [277,278,279,280,281,282,283,284,285]). The 6-deoxy-L-glucose derivative **155** has been prepared by Crimmins and Long [286] starting with the condensation of acetaldehyde with the chlorotitanium enolate of *O*-methyl glycolyloxazolidinethione **150**. A 5:1 mixture is obtained from which pure **151** is isolated by a single crystallization. After alcohol silylation and subsequent reductive removal of the amide, alcohol **152** is obtained. Swern oxidation of **152** and subsequent Horner–Wadsworth–Emmons olefination provides ene-ester **153**. Sharpless asymmetric dihydroxylation provides diol **154** which was then converted into **155** (*Scheme 60*) (see also [287]).

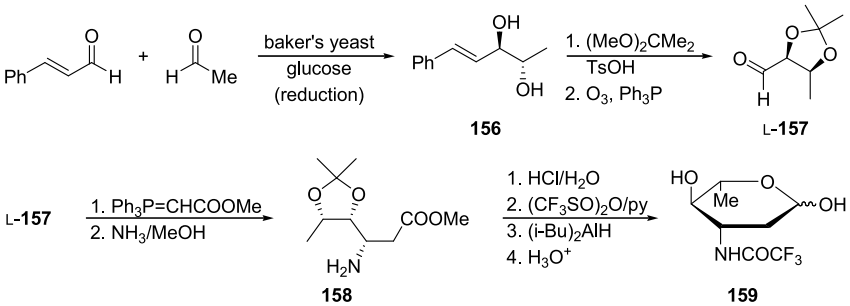


Scheme 60
Crimmins' synthesis of 6-deoxy-4-*O*-methyl-D-glucose

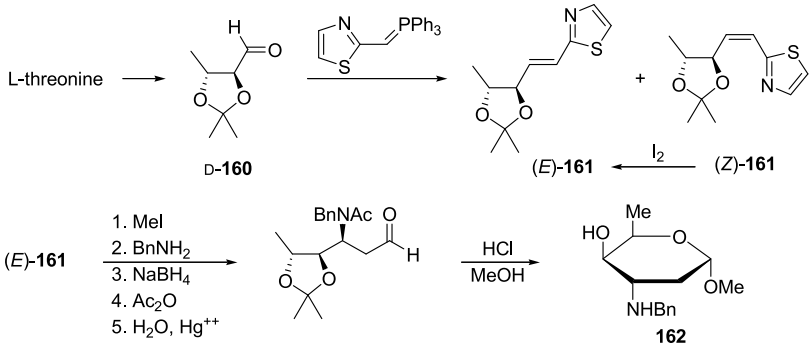
5.5.7 Aldehyde Olefination and Conjugate Addition

The first asymmetric total synthesis of acosamine and daunosamine starting from a nonsugar precursor was reported by Fuganti and co-workers [288,289,290,291]. They found that baker's yeast catalyzes the asymmetric pinacolic cross-coupling of cinnamaldehyde and ethenal giving *anti*-diol **156**. This diol is protected as an acetonide and submitted to ozonolysis giving L-**157**. Olefination of L-**157** with $Ph_3P{=}CHCOOMe$, followed by treatment with ammonia, provides **158** that is then converted into *N*-trifluoroacetylacosamine **158** (*Scheme 61*).

Methyl 3-*epi*-D-daunosaminide **162** has been derived from D-**160** via a Wittig-type olefination using (2-thiazolylmethylene)triphenylphosphorane (*Scheme 62*). A 1:1 mixture of (*E*)- and (*Z*)-alkenes is obtained which is isomerized in the presence of iodine into a 9:1 mixture of (*E*)-**161** and (*Z*)-**161**. Methylation of the thiazole moiety increases the electrophilicity of the alkene which then adds nucleophiles such as benzylamine. The adduct is treated with



Scheme 61
Fuganti's synthesis of *N*-trifluoroacetyl-L-acosamine



Scheme 62
Dondoni's synthesis of methyl *N*-benzyl-3-*epi*-D-daunosaminide

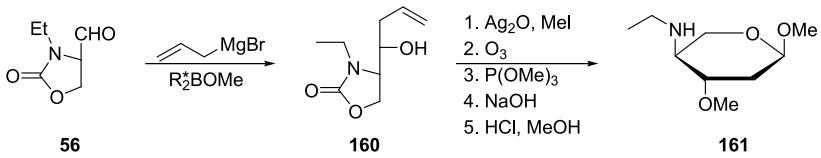
$NaBH_4$ to give a thiazolidine. Acetylation and mercury-mediated hydrolysis of the thiazolidine ring and subsequent acidic treatment in methanol yields the *N*-benzyl 3-*epi*-D-daunosaminide **162** [292].

5.5.8 Allylation and Subsequent Ozonolysis

The biologically important 2-deoxypentoses can be prepared readily by two-carbon chain elongation of 2,3-*O*-isopropylidene-D-glyceraldehyde following Roush's allylation method that relies on the highly diastereoselective additions of enantiomerically pure allylboronates derived from (*R*,*R*)- and (*S*,*S*)-tartaric acid. Similarly, the 2,6-dideoxyhexose derivatives have also been obtained by a similar route [293,294,295].

Addition of allylmagnesium bromide to serinal derivative **56** is *syn* selective in the presence of (−)-β-methoxydiisopinocampheylborane giving the *erythro* derivative **160** (*Scheme 62*). Methylation of the alcoholic moiety of **160**, followed by ozonolysis with reductive work-up, hydrolysis of the carbamate and acidic treatment forms the methyl glycoside **161** of the E-ring moiety of calicheamicin. A similar approach has been proposed by Roush starting from L-serinal derivative L-**65** [296] (*Scheme 63*).

The reaction of diallylzinc with 4-deoxy-2,3-*O*-isopropylidene-L-threose is *anti* stereoselective. After diol deprotection and ozonolysis 2-deoxy-L-fucose is obtained [297,298]. Similarly, allylation of (3*S*,4*S*)-3,4-isopropylidenedioxybutan-2-one leads to L-mycarose

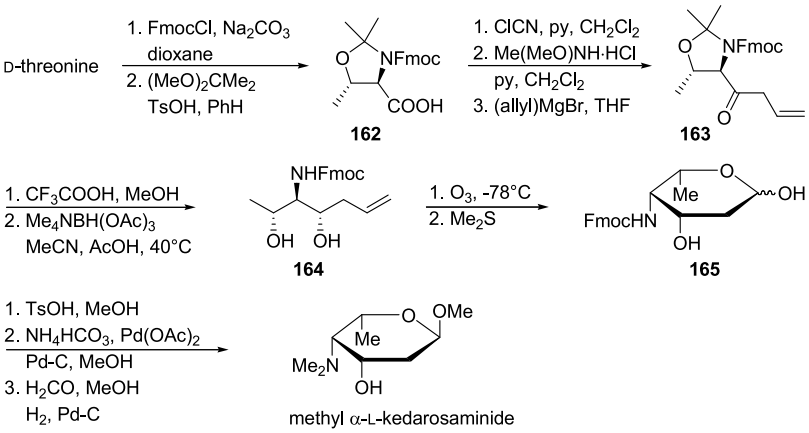


Scheme 63
Asymmetric allylation catalyzed by (−)-β-methoxydiisopinocamphenylborane

(2,6-dideoxy-3-*C*-methyl-L-ribohexopyranse). Using allylmagnesium bromide, L-boivinone and 2,3-dideoxy-L-*arabiono*-hexopyranose are prepared [299].

Two-carbon chain elongation of protected D-glyceraldehyde (*R*)-**37** and (*R*)-**99** can be realized via (*E*)-γ-alkoxyallylboronate additions, followed by alkene ozonolysis, furnishing D-arabinose derivatives [300].

The first noncarbohydrate-based asymmetric synthesis of kedarosamine uses the *N*,*O*-protected D-threonine **162**. It is converted into the corresponding Weinreb amide via the acyl chloride. Coupling with the allyl Grignard reagent provides **163**. The nonchelation controlled reduction of ketone **163** with $NaBH_4$ is *syn*-selective, whereas 1,2-chelation controlled reduction with $Zn(BH_4)_2$ is *anti*-selective. Deacetalization of **163** first, followed by reduction with the Evans' reagent, gives the *anti*-alcohol **164**. This product is then cleaved by ozonolysis, and subsequent ring closure to the corresponding pyranoses **165** occurs spontaneously. Fischer's glycosidation of **165** with methanol, followed by amine deprotection and methylation yields methyl α-L-kedarosaminide (◉ *Scheme 64*) [301].



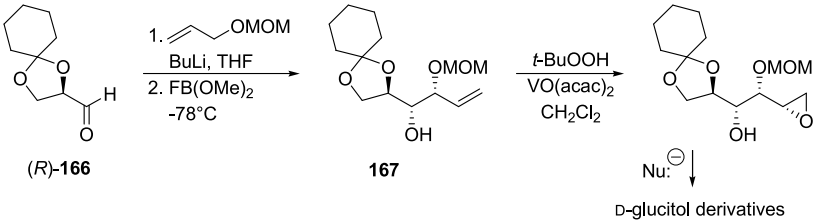
◘ Scheme 64
Kihlberg's synthesis of methyl α-L-kedarosaminide

5.6 Three-Carbon Chain Elongation

5.6.1 Allylmetal Additions

Roush's three-carbon-chain elongation method applied to 2,3-*O*-cyclohexylidene-D-glyceraldehyde (*R*)-**166** is an efficient approach to D-glucitol derivatives which relies on the *syn* selective epoxidation of homoallylic alcohol **167** (◉ *Scheme 65*) [300,302].

Addition of allyl bromide to *N*-benzyl-*N*-carbobenzoxy-*O*-*tert*-butyldimethyl-D-serinal in the presence of $SnCl_2$ and NaI gives preferentially the *anti*-adduct which was converted into enantiomerically pure 1,3-dideoxynojirimycin [303].



Scheme 65
Roush's conversion of D-glyceraldehyde into D-glucitol derivatives

5.6.2 Wittig–Horner–Emmons Olefination

Wittig–Horner–Emmons olefination of serinal derivative L-**65** with carbonylphosphorane **168** generates enone **169** which can be converted to (−)-nojirimycin via *anti*-dihydroxylation and reduction with $NaBH_4$ (*Scheme 66*), whereas reduction with Red-Al [Red-Al = $(MeOCH_2CH_2O)_2AlH_2Na$] leads to (−)-mannonojirimycin [304]. A similar analogue has been applied to generate galactostatine [305] and sialic acid analogues [306].

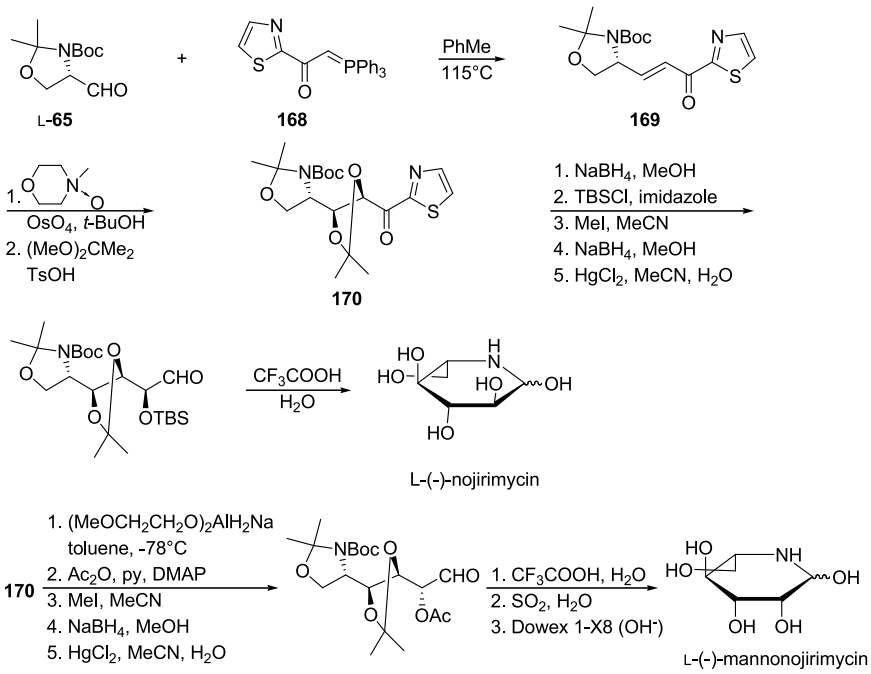
5.6.3 Aldol Reaction

Total syntheses of 3-deoxy-D-*ribo* and 3-deoxy-D-*arabino*-hexose have been realized via the cross-aldolization of 2,3-*O*-isopropylidene-D-glyceraldehyde (*R*)-**37** and 1,1-dimethoxyacetone (*Scheme 67*). The key step of the synthesis is the diastereoselective reduction of one of the aldols **171** via boron chelates. Treatment of **171** with triisobutylborane, and then with $NaBH_4$ gives *syn*-1,3- (**173**) and *anti*-1,3-diol **174** in a ratio 95:5. Acidic hydrolysis of **173** provides 3-deoxy-D-*ribo*-hexose. If aldol **171** is treated first with an equimolar amount of aluminum triisopropoxide, diol **174** is obtained in 62% yield (together with 15% of **173**). Compound **174** is converted then into 3-deoxy-D-*arabino*-hexose [307].

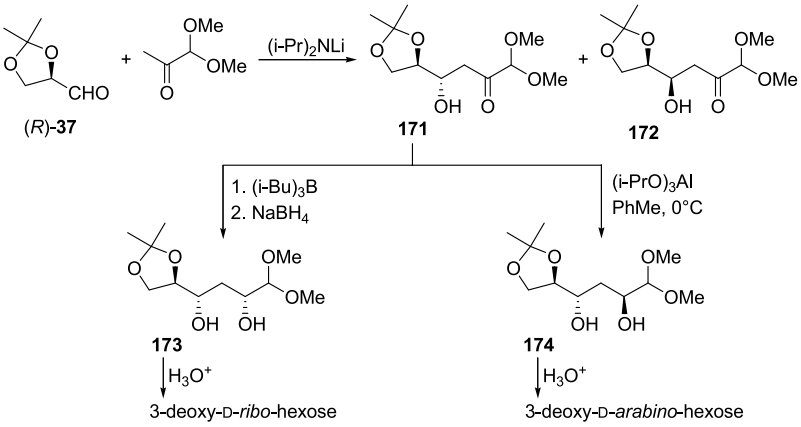
In a systematic study of the aldol reaction of methyl ketones Evans and co-workers [308,309] found that asymmetric induction in lithium enolates to α-alkoxy aldehydes is superior to that obtained from the corresponding enoxyboranes and enoxysilanes. The level of asymmetric induction in the lithium enolate additions is relatively insensitive to both the nature of the α-oxygen protective group and the steric hindrance of the enolate nucleophile. These additions are, however, sensitive to the nature of the β-alkyl substituent. For *anti*-α,β-dialkoxy-aldehydes, their reactions with methyl ketone enolates give mostly 3,4-*anti* products with consistently superior diastereoselectivity relative to the *syn*-α,β-dialkoxy-aldehydes. With α,β-isopropylidenedioxy-aldehydes the diastereoselectivity is the same regardless of the *syn*- vs. *anti*-α,β relative configuration. A Cornforth transition-state model (*Fig. 3*) is proposed to account for these observations in which the β-substituent dictates the position in space occupied by the α-oxygen protecting group (P1), which in turn governs aldehyde π-facial selectivity due to its proximity to the approaching enolate.

5.6.4 Other Methods of Three-Carbon Chain Elongation of Aldoses and Derivatives

Sames and Polt [310] have converted L-serine into *N*-methylfucosamine (*Scheme 68*). The method relies on the diastereoselective propenyllithium addition to the aldehyde derived from



Scheme 66
Dondoni's synthesis of azasugars



Scheme 67
Conversion of D-glyceraldehyde into 3-deoxy-D-hexoses

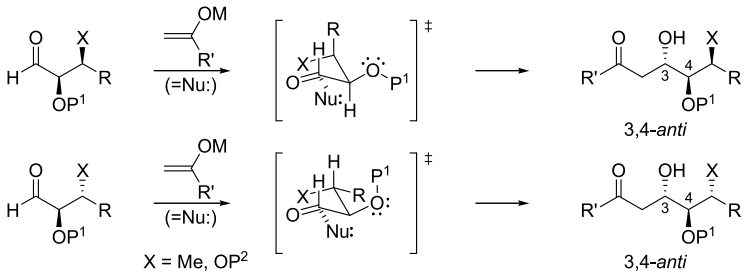
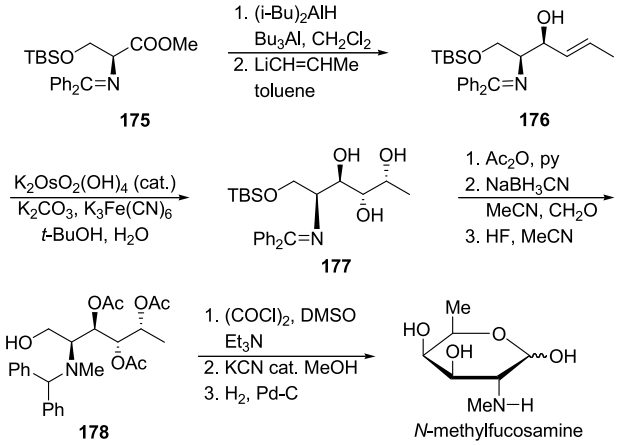


Figure 3
Preferred diastereoselectivities in aldol reaction



Scheme 68
Polt and Sames' synthesis of methylfucosamine

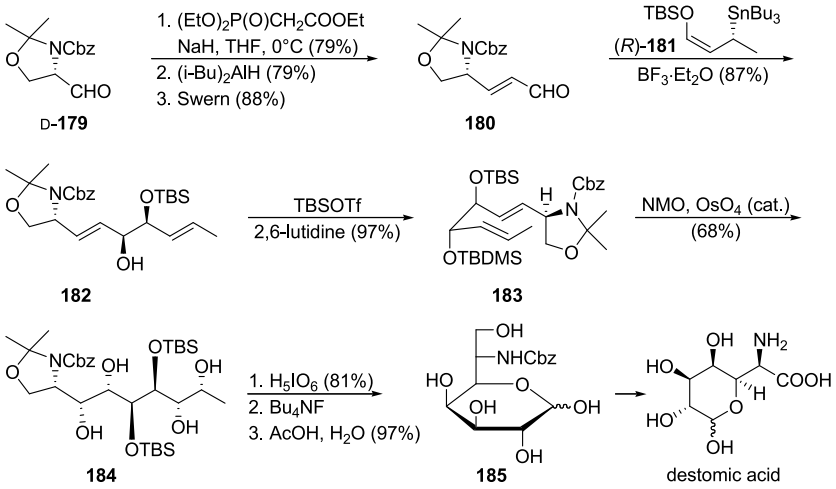
protected L-serine derivative **175** giving allylic alcohol **176**. Catalytic osmylation of **176** gives a 6:1 mixture of *anti*,*syn* **177** and *syn*,*syn*-aminotriols. Protection of the triol **177** as triacetate, reductive methylation, and desilylation provides **178**. Successive Swern oxidation, methanolysis, and benzhydryl group hydrogenolysis leads to *N*-methylfucosamine [311] (for other examples of three-carbon homologations: [312,313,314,315]).

5.7 Four-Carbon Chain Elongation

5.7.1 (But-2-en-1-yl) Metal Addition

The addition of crotylstannane derivatives to aldehydes are highly diastereoselective in the presence of $MgBr_2 \cdot OEt_2$ as the promoter [316,317,318].

The octitol derivative **184** has been prepared via BF_3 etherate-promoted addition of (*R*)-γ-OTBS allylic stannane (*R*)-**181** to enal **180** derived from L-serine (*Scheme 69*) [319]. This leads to a single *syn*-adduct **182** which is converted to the silyl ether **183**. This compound



Scheme 69
Marshall's allylation applied to the synthesis of destomic acid

adopts probably the Saito conformation which makes the double *cis*-dihydroxylation doubly diastereoselective with the formation of tetrol **184** as a major product. Selective oxidative cleavage of tetrol **184** affords a lactol. Desilylation followed by acetonide hydrolysis generates **185**, a precursor of destomic acid (see also: [320,321]).

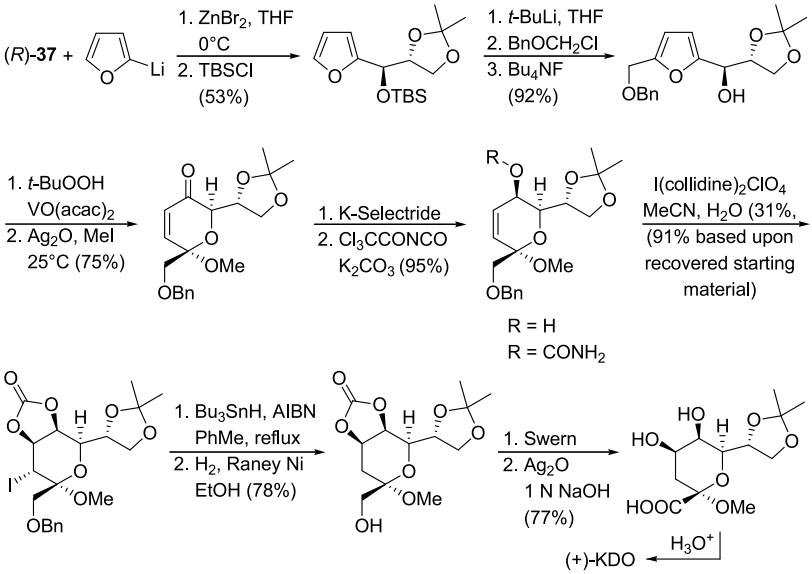
5.7.2 Nucleophilic Addition of α-Furyl Derivatives

A synthesis of (+)-KDO (*Scheme 70*) has been reported by Martin and Zinke [322] which relies on the highly stereoselective addition of 2-furyllithium to 2,3-*O*-isopropylidene-D-glyceraldehyde (*R*)-**37** [323].

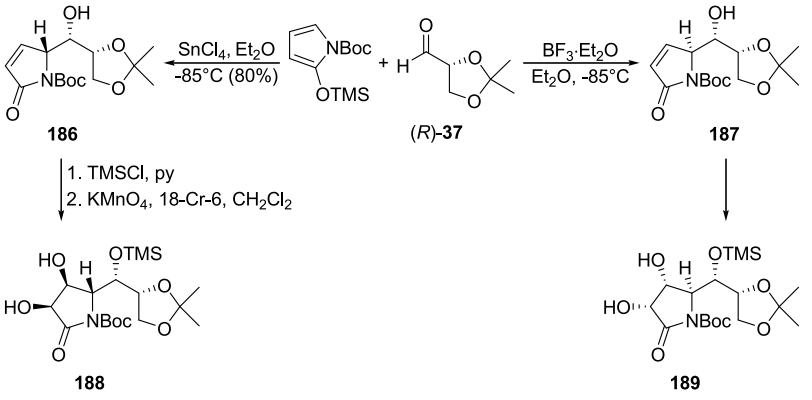
Condensation of 2,3-*O*-isopropylidene-D-glyceraldehyde-*N*-benzylimine with 2-(trimethylsilyloxy)furan in the presence of a Lewis acid generates a mixture of butenolides that has been converted into a D-*ribo*-*N*,*N*-diprotected derivative in 68% yield and then into enantiomerically pure 1,5-dideoxy-1,5-imino-D-*glycero*-D-*allo*-heptitol [324].

5.7.3 Hydroxyalkylation of Pyrrole Derivatives

The Lewis acid-promoted condensation of *N*-(*t*-butyloxycarbonyl)-2-(*t*-butyldimethylsiloxy) pyrrole with 2,3-*O*-isopropylidene-D-glyceraldehyde ((*R*)-**37**) generates pyrrolidones with high diastereoselectivity. For instance, with 1.5 equivalent of $SnCl_4$ in ether, the reaction gives crystalline D-*arabino*-configurated α,β-unsaturated γ-lactam **186** as the sole product, whereas, in the presence of one equivalent of BF_3 etherate, reversal stereochemistry is observed, resulting in predominant formation of crystalline D-*ribo*-epimer **187**, along with less than 20% of **186**. Almost quantitative epimerization **186** → **187** occurs on treatment with Et_3N in CH_2Cl_2in the presence of 4-dimethylaminopyridine (*Scheme 71*). Dihydroxylations of the



Scheme 70
Martin's asymmetric synthesis of (+)-KDO

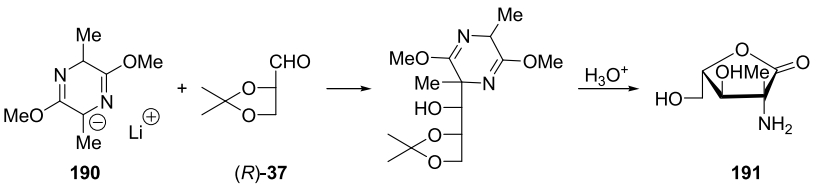


Scheme 71
Aminodeoxyaldonolactams through hydroxyalkylation of a pyrrole derivative

trimethylsilyl ethers of **186** and **187** generate the 4-amino-4-deoxy-heptono-1,4-lactam derivatives **188** and **189**, respectively [325,326]. Lactam hydrolysis of **188** with LiOH, followed by Malaprade diol cleavage with $NaIO_4$ and further oxidation and deprotection allows preparing 4-*epi*-polyoxamic acid [327]. Lactam **186** and its enantiomer derived from (*S*)-**37** have been converted into all four stereomers of *cis*-1,2-dihydroxypyrrolizidine [328]. Compounds **186** and **187** have been used also to prepare the *trans*-2,3-*cis*-3,4-dihydroxyprolines [329,330].

5.8 Synthesis of Branched-Chain Monosaccharides from C_3-Aldoses

Most of the methods presented for the de novo synthesis of "linear" monosaccharides can be used to prepare branched-chain sugars and analogs and some examples are given below. The branched-chain aminolactone **191** has been prepared by two-carbon chain elongation via addition of **190** to 2,3-*O*-isopropylidene-L-glyceraldehyde (*R*)-**37** (Scheme 72) [331].



Scheme 72
Synthesis of a branched-chain amino aldonolactone

Aldol reaction of (*S*)-2-benzyloxypropanal with the lithium enolate of methyl 2-methoxypropanoate gives a 7:2:1 mixture of β-hydroxyesters. The major product has been converted into L-cladinose, a saccharide moiety of erythromycin A [332]. The isomeric methyl 3,6-dideoxy-3-*C*-methylhexofuranosides have been derived from (*S*)-2-benzyloxypropanal via homoaldol reactions in three steps.

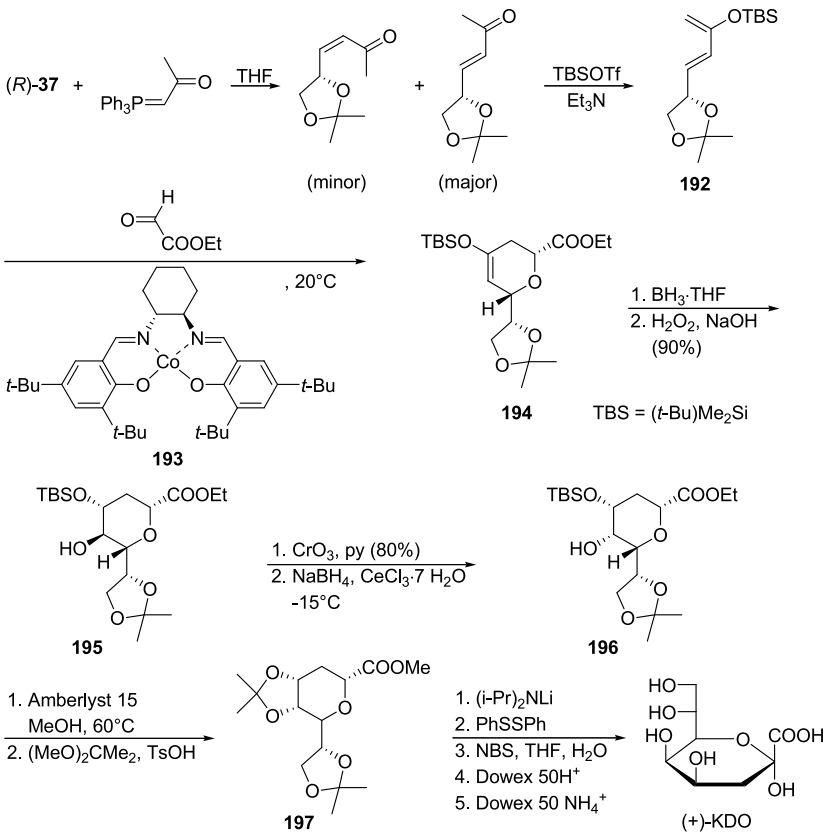
L-Arcanose and L-olimycose have been prepared in enantiomerically pure forms and with high stereoselectivity by Lewis-acid promoted addition of (*S*)-2-benzyloxypropanal to 1-trimethylsilyl-2,3-butadiene. Depending on the nature of the Lewis acid either the *syn* (with $TiCl_4$) or the *anti* adduct (with $BF_3 \cdot Et_2O$) can be obtained. Epoxidation with lateral control by the allylic alcohol moieties and standard reactions lead to the unprotected monosaccharides [334]. Total syntheses of 2,3-dideoxy-3-*C*-methyl-D-*manno*-heptose and of 2,3-dideoxy-2,3-di-*C*-methyl-D-*glycero*-D-*galacto*-heptose have been realized by addition of 2-(trimethylsiloxy)furan to 2,3-*O*-isopropylidene-D-glyceraldehyde ((*R*)-**37**) [335].

6 Hetero-Diels–Alder Additions

6.1 Achiral Aldehydes as Dienophiles

The cycloadditions of achiral oxy-substituted dienes with aldehydes in the presence of enantiomerically pure (+)-Eu(hfc)$_3$ [hfc = 3-(heptafluoropropylhydroxymethylene)-(+)-camphorate] shows only modest enantioselectivities. Similarly, modest diastereoselectivities are observed for the reactions of chiral oxy-substituted dienes with aldehydes in the presence of achiral Eu(fod)$_3$ (fod = 6,6,7,7,8,8,8-heptafluoro-2,2-dimethyl-3,5-octanedienoate). However, the combination of chiral dienes with chiral (+)-Eu(hfc)$_3$ catalyst shows interesting interactivities, resulting in some instances in diastereoselectivities of 97:3. This has permitted a total synthesis of L-glucose [336,337,338].

By applying a similar approach, Wu and co-workers [339] have observed a highly double-stereoselective hetero-Diels–Alder addition between diene **192** and ethyl glyoxylate catalyzed



Scheme 73
Wu's synthesis of (+)-KDO

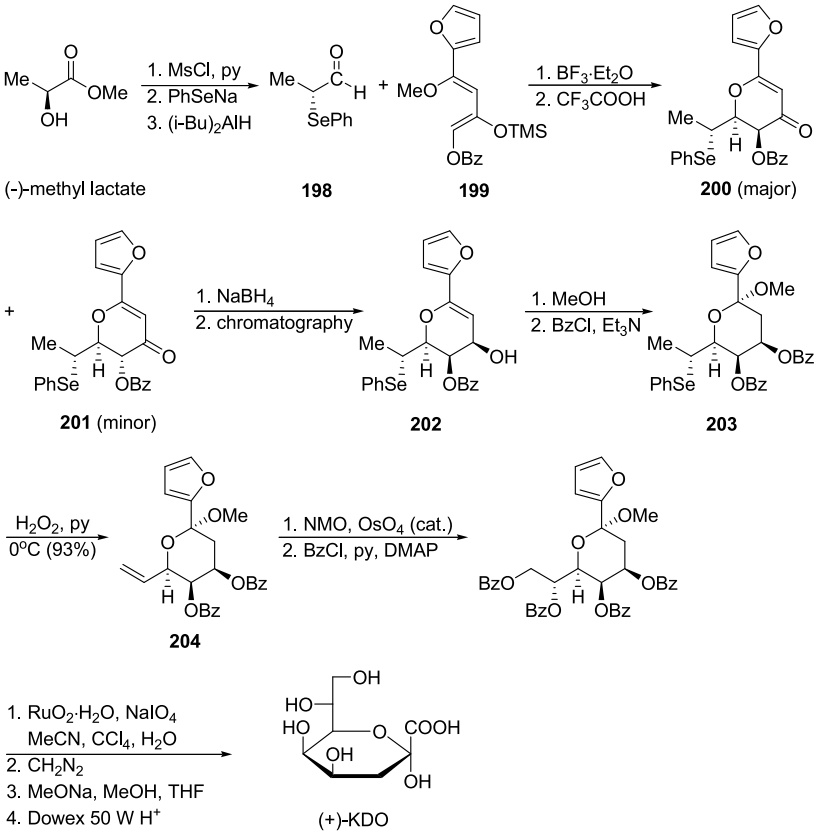
by the (salen)Co^{II} complex **193** (*Scheme 73*). The major adduct **194** is hydroborated to give alcohol **195** that is inverted via oxidation and subsequent reduction to **196**. Acid methanolysis followed by diol protection furnishes **197**. Quenching of the lithium enolate of **197** with phenyl disulfide and subsequent oxidation with *N*-bromosuccinimide forms, after deprotection, (+)-KDO (3-deoxy-D-*manno*-oct-2-ulosonic acid) [340].

The reaction of 5-nitrofuran-2-carbaldehyde and 2,3,4,6-tetra-*O*-acetyl-β-D-glucopyranosides of (*E*,*E*)-4-ethoxy-2-[(tertbutyl)dimethylsilyloxy]butadien-1-yl can be highly diastereoselective depending on the nature of the lanthanide Lewis acid used to promote the cycloaddition ($Yb(fod)_3$, $La(fod)_3$). The adducts so-obtained are readily converted into β-D-glucopyranosyl (1→4)-linked glycals [341]. 1-(*Z*)-Alkoxy-4-(*E*)-methoxybutadiene derivatives add to ethyl glyoxylate and diethyl oxomalonate under thermal and hyperbaric conditions. They provide dihydropyranic products with high stereoselectivities. These hetero-Diels–Alder adducts can be converted into racemic allose, mannose, and gulose derivatives [342] (for other hetero-Diels–Alder adducts of this aldehyde, see [343,344,345,346,347]).

6.2 Chiral Aldehydes as Dienophiles: Synthesis of Long-Chain Sugars

The first total synthesis of (+)-KDO was presented by Danishefsky and co-workers (*Scheme 74*) [348]. The hetero-Diels–Alder addition of α-selenoaldehyde **198** to the α-furyl-substituted diene **199** gives an adduct mixture which on treatment with CF_3COOH delivers a 5:1 mixture of *cis*/*trans* dihydropyrones **200** and **201**. Reduction of pure **200** generates **202** which adds methanol like a glycal; after benzoylation **203** is obtained. Oxidative elimination of the phenylseleno group gives alkene **204** which is dihydroxylated to a diol, the benzoylation of which affords a tetrabenzoate, the oxidation of it with RuO_4, followed by esterification with diazomethane, generates methyl KDO α-methyl glycoside. Deprotection leads to (+)-KDO.

The octosyl acids are isolated from *Streptomyces cacaoi*; they are part of a broader group of polyoxin antifungal nucleosides [349]. Danishefsky and co-workers [350] have reported a total synthesis of octosyl acid A featuring the hetero-Diels–Alder addition of (*E*)-1-methoxy-3-(trimethylsilyloxy)butadiene to a protected D-ribose-derived aldehyde.



Scheme 74
Danishefsky's total synthesis of (+)-KDO

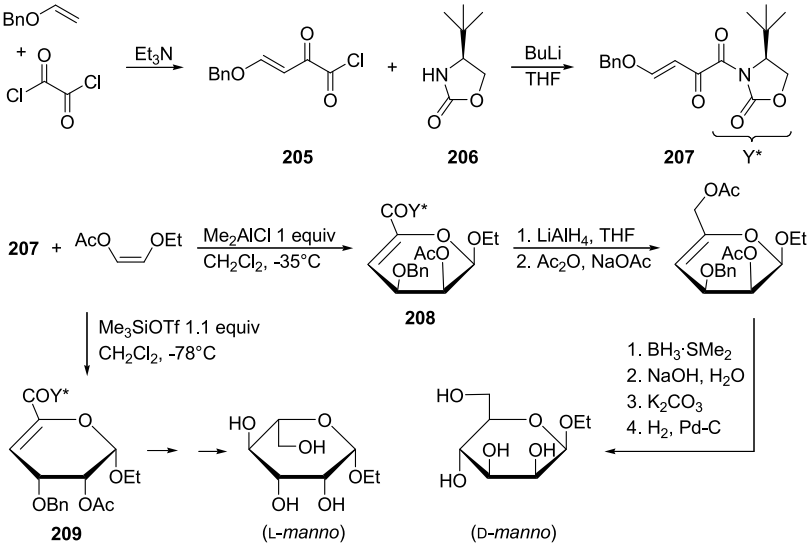
An elegant total synthesis of a semi-protected form of lincosamine has been realized by Marshall and Beaudoin [319]. Destomic acid (6-amino-6-deoxy-L-*glycero*-D-*galacto*-heptonic acid) has been derived in a similar way from a L-serinal derivative via hetero-Diels–Alder addition to 1-ethoxy-3-[(trimethylsilyl)oxy]-4-benzyloxy-1,3-butadiene. A similar method has been applied also to prepare a semi-protected form of anhydrogalantinic acid, a component of the antibiotic galantin I [351].

6.3 Hetero-Diels–Alder Addition of 1-Oxa-1,3-Dienes

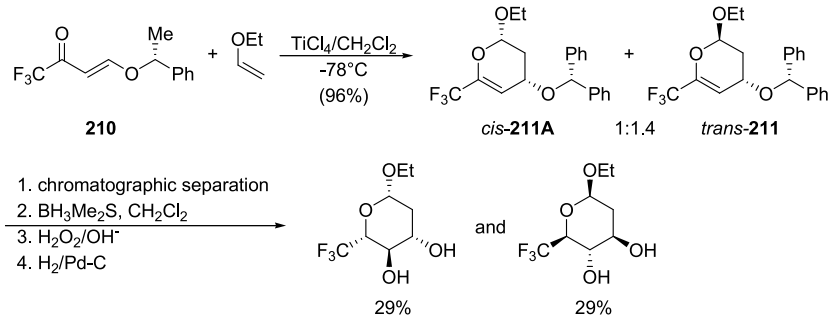
The [4+2]-cycloaddition of α,β-unsaturated aldehydes and ketones (1-oxa-1,3-dienes) to enol ethers (Diels–Alder addition with inverse electron demand) has been an attractive route for the synthesis of 3,4-dihydro-2H-pyrans [352,353,354,355,356,357,358] which can be converted into deoxy- and dideoxypyranosides [359,360,361,362,363,364,365,366,367].

The enantiopure 1-oxa-1,3-diene **207** is prepared by acylation of benzyl vinyl ether with oxalyl chloride; this generates acyl chloride **205** which acylates the lithium salt of 2-oxazolidinone **206**. In the presence of Me_2AlCl, diene **207** adds to (*Z*)-1-acetoxy-2-ethoxyethene giving mostly adduct **208**, whereas, when using Me_3SiOTf as a promoter of the hetero-Diels–Alder addition, diastereomer **209** is the major adduct (*Scheme 75*). Adducts **208** and **209** have been converted into ethyl β-D-mannopyranoside and ethyl-β-L-mannopyranoside, respectively [368,369]. L-Olivose has been prepared via a similar approach [370].

The Diels–Alder additions of enantiomerically pure oxadienes to (*Z*)-2-ethoxyvinyl acetate are highly diastereoselective and have been used to prepare derivatives of 4-deoxy-D-*lyxo*-hexose [357]. Larsen and co-workers [371] have prepared 2,6-dideoxy-6,6,6-trifluoro-*arabino*-



Scheme 75
Tietze's synthesis of L-hexoses



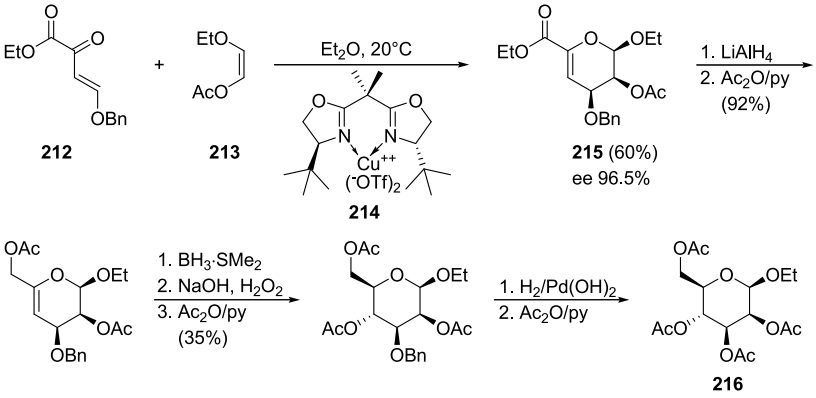
Scheme 76
Synthesis of 2,6-dideoxy-6,6,6-trifluoro-*arabino*-hexoses

hexoses via a $TiCl_4$-catalyzed hetero-Diels–Alder addition of (*E*)-1,1,1-trifluoro-4-[(1*R*)-1-phenylethoxy]but-3-en-2-one (**210**) to ethyl vinyl ether giving a mixture of adducts *cis*-**211** and *trans*-**211** (Scheme 76).

Alternatively monosaccharides have been prepared via cycloaddition of chiral enol ethers to oxadienes [372,373,374].

Independently, the groups of Evans [375,376,377] and Jørgensen [378] have shown that β,γ-unsaturated α-keto esters react with ethyl vinyl ether in the presence of enantiomerically pure bisoxazoline copper(II) complexes as catalysts leading to enantiomerically enriched dihydropyrans. For instance **212** and **213** in ether at 20 °C and in the presence of complex **214** add to give the *endo* adduct **215** in 60% yield and with 96.5% ee. The latter is then converted into ethyl β-D-*manno*-pyranoside tetraacetate **216** (Scheme 77) [379].

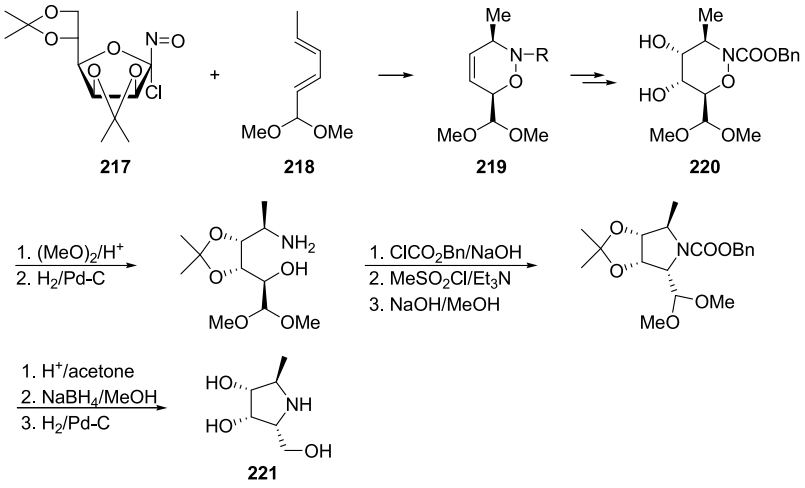
The first catalytic, highly enantioselective hetero-Diels–Alder reactions of thiabutadienes with an acyloxazolidinone dienophile using homochiral copper and nickel triflate and perchlorate bis(oxazoline) and bis(imine)complex catalysts to generate dihydrothiopyrans have been reported by Saito et al. [380].



Scheme 77
Jørgensen's asymmetric synthesis of an ethyl β-D-mannopyranoside derivative

6.4 Nitroso Dienophiles: Synthesis of Azasugars

Homochiral nitroso dienophiles can add to conjugated dienes in the hetero-Diels–Alder mode, sometimes with good diastereoselectivity [381,382,383,384]. For instance, the chloronitroso dienophile **217** derived from D-mannose [385] adds to diene **218** in MeOH/HC(OMe)$_3$ to give adduct **219** which has been converted into the 1,5,6-trideoxy-1,5-iminoalditols [383,386,387, 388] and into the potent α-L-fucosidase and α-galactosidase inhibitor **221** (Scheme 78) [389].

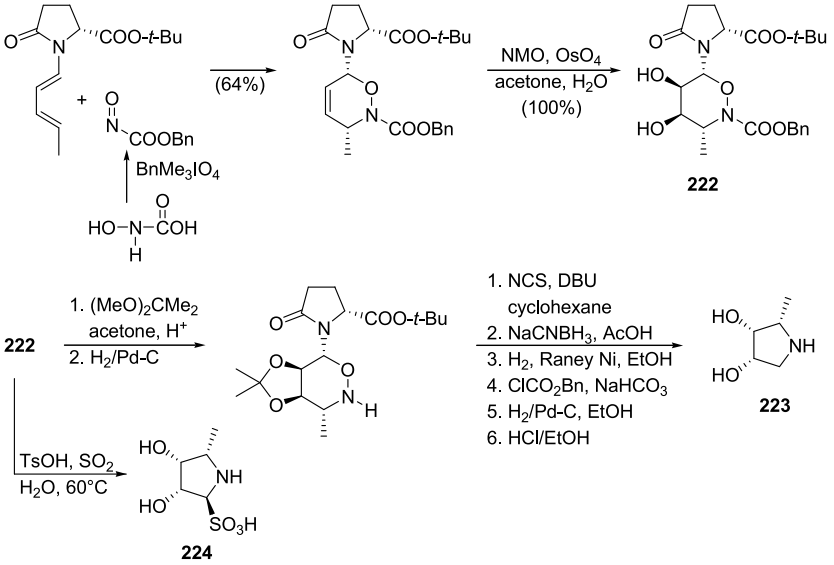


Scheme 78
Defoin's synthesis of 2,5-imino-2,5,6-trideoxy-D-altritol

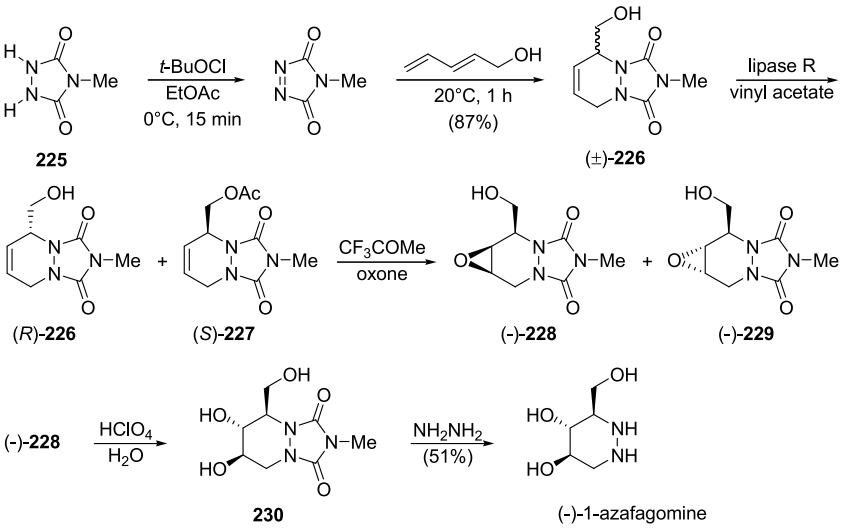
Good diastereoselectivity has been observed for the hetero-Diels–Alder addition of homochiral 1,3-dienes with achiral acyl-nitroso dienophiles. An example is shown in Scheme 79 for the total synthesis of 4-amino-4,5-dideoxy-L-lyxose derivatives **223** and **224** (potent inhibitors α-L-fucosidase), respectively [390] (see also [391,392,393]).

6.5 *N*-Methyltriazoline-3,5-Dione as the Dienophile: Synthesis of 1-Azafagomine

Methyl urazol **225** is oxidized into *N*-methyl triazoline-3,5-dione with *tert*-butyl hypochlorite. Without isolation it reacts with (*E*)-penta-2,4-diene-1-ol giving adduct (±)-**226** as a racemic mixture. On treatment of (±)-**226** with a lipase (R/Novozym 435) in vinyl acetate 38% of ester (*S*)-**227** (86% ee) and 29% of alcohol (*R*)-**226** (59% ee) is isolated. Epoxidation of (*S*)-**227** provides 13% of (−)-**228** and 68% of (−)-**229**. The latter epoxide is hydrolyzed under acidic conditions to provide **230** that reacts with hydrazine to give enantiomerically pure (−)-1-azafagomine (Scheme 80) [394]. This compound has a slow binding inhibitor of almond β-glucosidase (K_i = 0.33 μM). A similar sequence of reactions converts (*R*)-**226** into (+)-1-azafagomine. The latter is not an inhibitor of β-glucosidases [395].



Scheme 79
Defoin's synthesis of 4-amino-4,5-dideoxy-L-lyxose derivatives



Scheme 80
Bols' synthesis of 1-azafagomine

7 Cycloadditions of Furans

7.1 Diels–Alder Additions

Just and his group have pioneered the use of 7-oxabicyclo[2.2.1]hept-2-enes as starting materials in the synthesis of carbohydrates and analogs [396,397,398,399,400]. These bicyclic systems have the advantage to undergo highly face selective reactions. They are obtained simply by Diels–Alder addition of furan to alkene dienophiles [401,402,403,404,405,407].

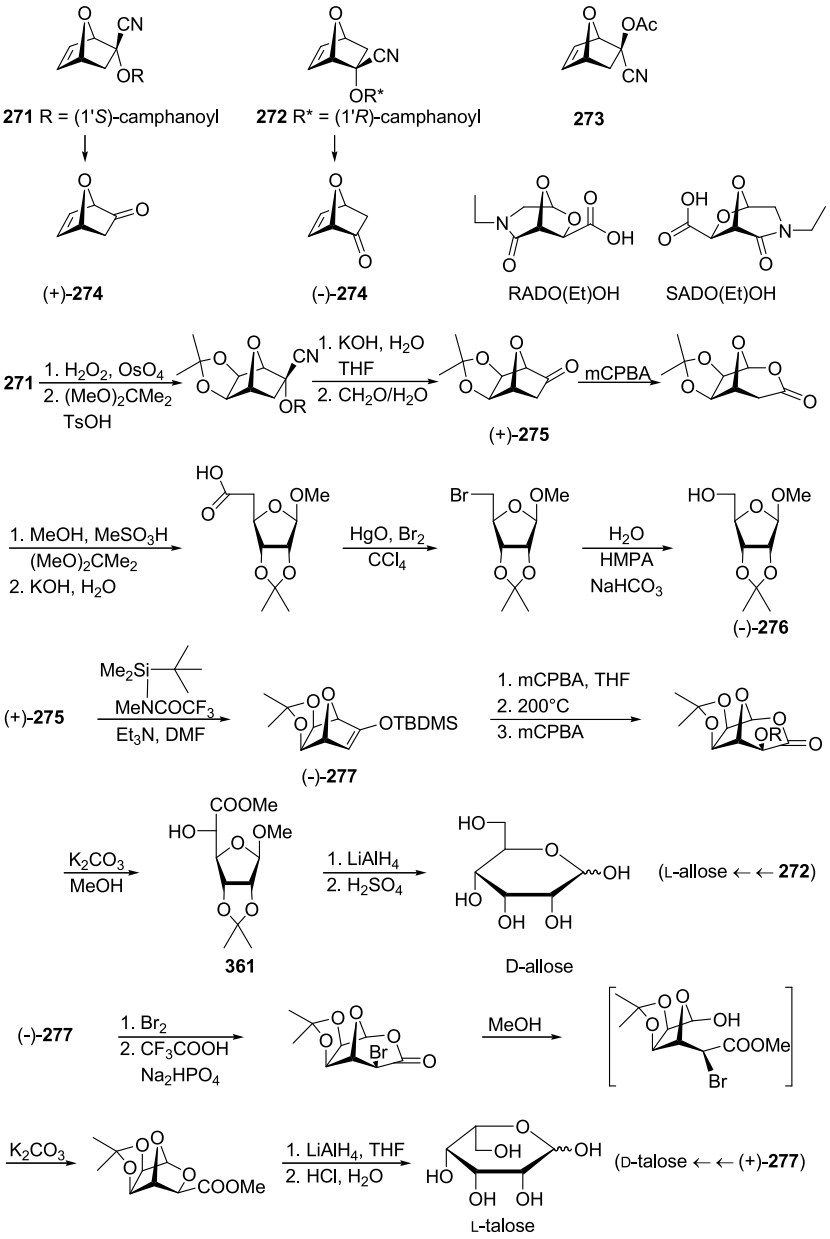
7.2 The "Naked Sugars of the First Generation"

The 1-cyanovinyl (1′*S*)-camphanate (derived from (1*S*)-camphanic acid and pyruvonitrile) adds to furan in the presence of ZnI_2 as the catalyst. After 7 days at room temperature a mixture of four possible diastereomeric Diels–Alder adducts is formed (95%) from which pure adduct **271** can be isolated by crystallization. Unreacted furan is recovered and the diastereomer mixture left from the crystallization is heated to give furan and 1-cyanovinyl (1′*S*)-camphanate that can be recycled to prepare more of the diastereomerically pure adduct **271** (the reversibility of the furan Diels–Alder addition is exploited here). Starting from (1*R*)-camphanic acid which is also commercially available, pure adduct **272** can be prepared in large quantities as readily [408]. Camphanic acid auxiliaries can be replaced by the chiral auxiliaries RADO(Et)OH or SADO(Et)OH derived from (*R*,*R*)-tartaric acid and (*S*,*S*)-tartaric acid [409,410]. Using 1-cyanovinyl acetate as the dienophile, a mixture of racemic adducts is obtained which is hydrolyzed into a mixture of cyanohydrines that can be resolved by complex formation with brucine. Reaction of the diastereomerically pure complex with Ac_2O provides the enantiomerically pure 7-oxanorbornene derivatives **273** [411] (for other enantiomerically enriched 7-oxabicyclo[2.2.1]heptane derivatives [412]).

Enantiomerically pure 7-oxanorbornenyl derivatives **271**, **272**, **273**, and their products of saponification (recovery of the chiral auxiliary in the aqueous phase), ketones (+)-**274** and (−)-**274**, are coined "naked sugars of the first generation" because they are chirons (= enantiomerically pure synthetic intermediates) like those derived from natural hexoses. They are enantiomerically pure like natural sugars, but with three unsubstituted (naked) carbon centers, the substitution of which follows highly stereoselective methods giving polysubstituted 7-oxabicyclo[2.2.1]heptane-2-ones that can be oxidized into the corresponding uronolactones as illustrated in *Scheme 81* for the total synthesis of the riboside derivative (−)-**276**, D-allose and L-talose [413].

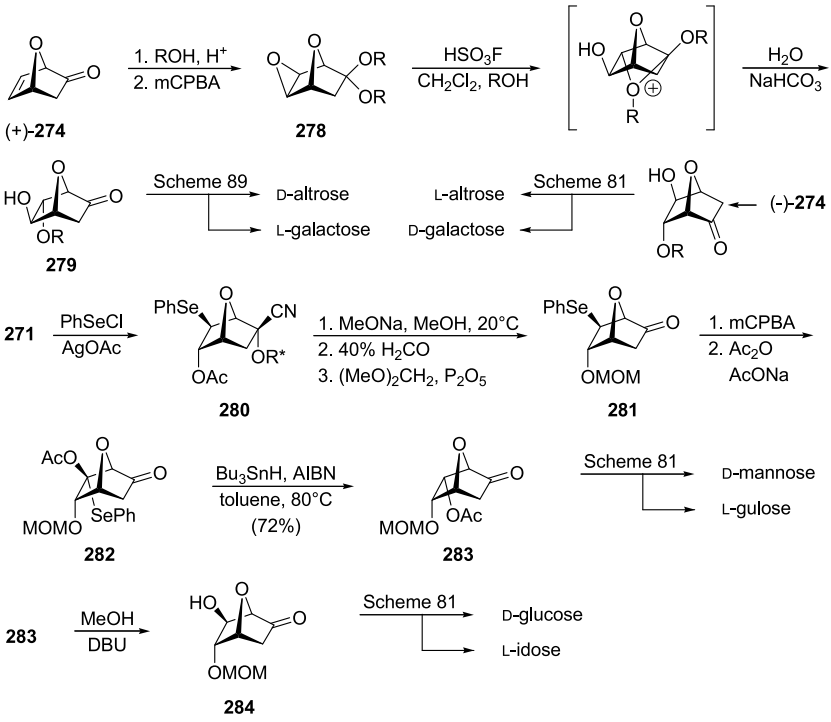
Acetals of (+)-**274** are epoxidized into **278** (*Scheme 82*). Acid promoted ring opening of the epoxides of **278** with participation of the *endo* OR group finally leads to the *trans*-5,6-dioxy-substituted 7-oxabicyclo[2.2.1]heptan-2-ones **279** [414,415,416]. The reaction sequences of *Scheme 81* applied to **279**, and to its enantiomer obtained from (−)-**277**, generate D- and L-altrose or L- and D-galactose, respectively.

Electrophilic addition of PhSeCl to alkene **271** is highly stereo- and regioselective (steric effects) and provides adduct **280** if carried out in the presence of an acetate nucleophile. Methanolysis, treatment with formaline and protection of the *endo* alcohol gives ketone **281**. Low-temperature oxidation of the selenide **281** generates a selenoxide which does not elimi-



Scheme 81

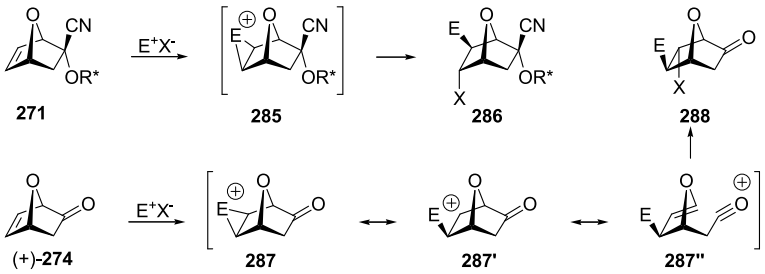
"Naked sugars of the first generation": asymmetric total synthesis of carbohydrates with early recovery of the chiral auxiliary



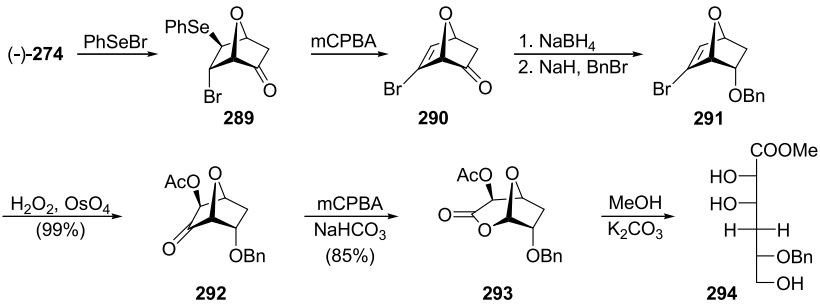
Scheme 82
Asymmetric synthesis of hexoses starting from furan

nate readily but has the time to undergo a seleno-Pummerer rearrangement in the presence of acetic anhydride and sodium acetate, leading to **282**. Reductive elimination of the seleno moiety with tin hydride forms **283** bearing two *endo* hydroxy groups protected in an orthogonal fashion (*Scheme 82*) [417]. Applying reactions of *Scheme 81* to **283** and to its enantiomer generate hexoses with a *manno* or *gulo* configuration. Methanolysis of **283** occurs with complete epimerization at C(6) (retro-aldol, aldolization) giving **284**, a precursor of D-glucose and L-idose derivatives (*Scheme 82*).

Under conditions of kinetic control, the cyano-esters (e. g. **271**) add soft electrophiles E^+X^- to generate the corresponding adducts **286** with high *exo* face selectivity and high regioselectivity. In these cases, the bridged ion intermediates **285** are attacked preferentially by the nucleophile X^- at C(5), the center that is the least sterically hindered and also the one which can support the highest partial positive charge [field effect of the electron-withdrawing CN and ester substituents at C(2)]. With the 7-oxabicyclo[2.2.1]hept-5-en-2-one ((+)-**274**, derived from **271**) electrophiles E^+X^- add with opposite regioselectivity and give the corresponding adducts **288**. The nucleophile's (X^-) preference for carbon center C(6) is attributed to the electron-releasing effect of the carbonyl function in positively charged intermediates **287** ↔ **287′** ↔ **287″** due to favorable n(CO) ↔ σ C(1,2) ↔ pC(6)$^+$ hyperconjugative interaction (frangomeric effect of the carbonyl n electron pairs) [418,419,420,421,422,423]. The



Scheme 83
The highly stereo- and regioselective electrophilic additions of the "naked sugars"



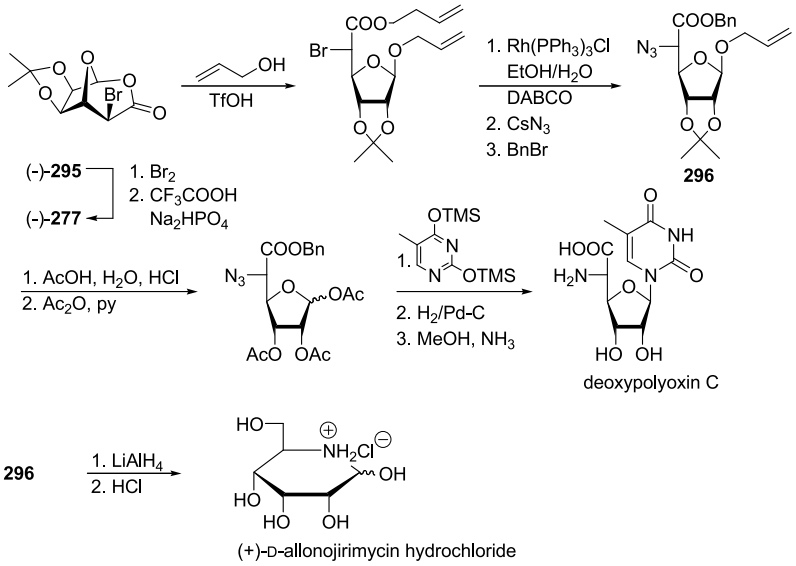
Scheme 84
Total synthesis of a 4-deoxypyranoside and derivatives

principle described here (Scheme 83) has been applied to convert the "naked sugars" into 2-, 3-, and 4-deoxyhexoses and derivatives in highly stereoselective manners [424,425,426], an example is summarized in Scheme 84.

Addition of PhSeBr to enone (−)-**274** gives adduct **289** which undergoes oxidative elimination of the phenylseleno group, to give bromoenone **290**. Ketone reduction and protection of the *endo* alcohol furnishes **291**, the double hydroxylation of its chloralkene unit generates **292** after acetylation. Baeyer–Villiger oxidation yields uronolactone **293** which has been converted into 3-deoxy-α-D-*arabino*-hexopyranoside and into methyl 4-deoxyhexonate **294** [426].

The "naked sugars of the first generation" (**271**, **272**, **273**, (+)-**274**, (−)-**274**) are useful chirons to prepare all kinds of rare sugar derivatives [427,428,429,430], as shown below, and of analogs such as *C*-linked disaccharides [431,432,433,434,435,436,437,438,439,440]. They have been converted into conduritols, cyclitols, and carbahexose derivatives [416,441]. The bicyclic ketones can also be converted into their enoxysilanes and then cleaved by ozonolysis. Depending on the work-up conditions 2,5-anhydrohexaldaric acid or 2,5-anhydrohexonic acid derivatives can be prepared readily, in both their enantiomerically pure forms [442,443,444, 445].

Light-induced aziridination of the benzyl acetal of (+)-**274** with $BocN_3$ (Boc = (*t*-Bu)OCO) gives an aziridine that converted into methyl 3-amino-3-deoxy-α-D-altropyranoside [446], and into 3-amino-3-deoxy-L-idose [447].

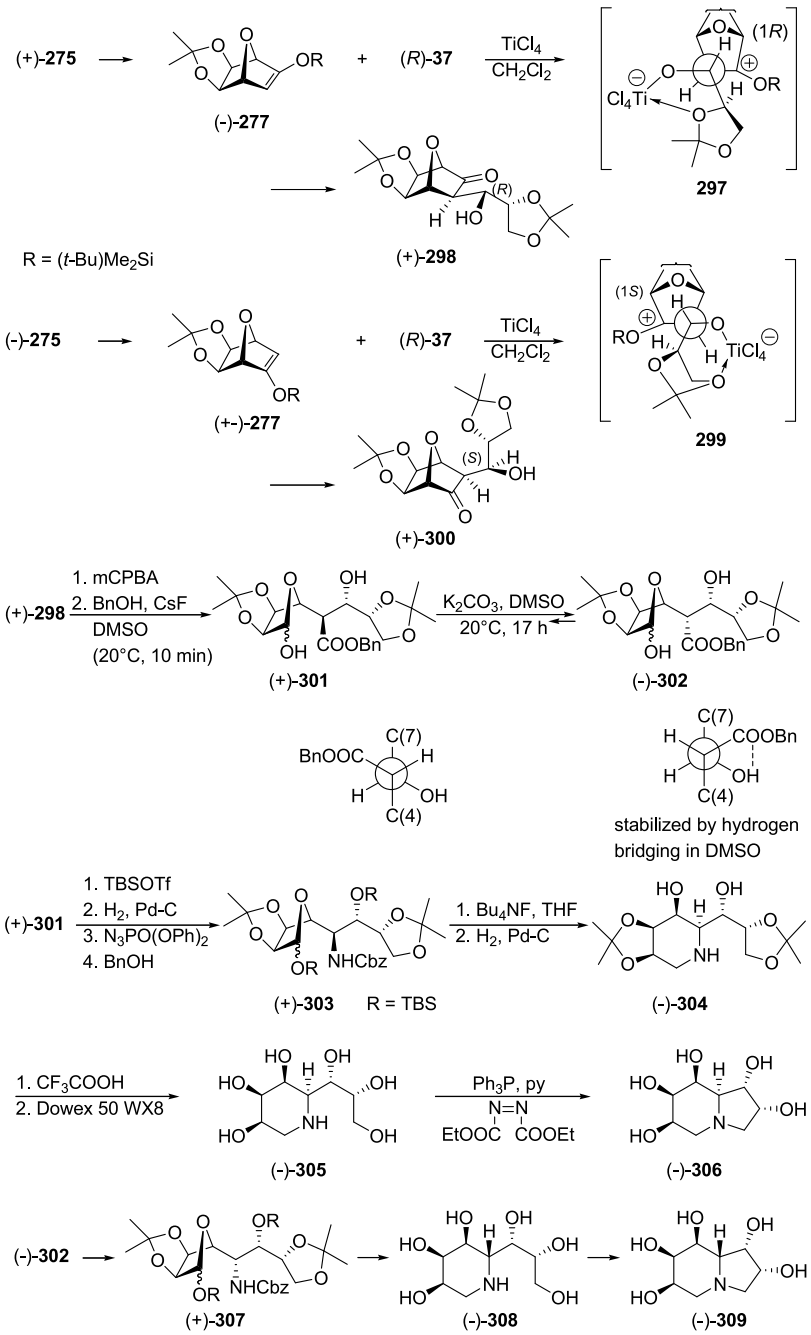


Scheme 85
Total synthesis of deoxypolyoxin C and of (+)-D-allonojirimycin

The "naked sugar" methodology has also provided deoxypolyoxin C (*Scheme 85*). The α-bromouronolactone (−)-**295** gives an allyl uronic ester on treatment with allyl alcohol and triflic acid. Selective hydrolysis of the allyl ester, followed by displacement of the bromide with the azide anion with retention of configuration generates an intermediate cesium carboxylate that reacts with benzyl bromide to give **296**. Conversion of **296** into a triacetate, followed by Vorbrüggen's glycosidation, reduction, and deprotection yields deoxypolyoxin C [448]. Intermediate **296** has been converted in two steps into (+)-D-allonojirimycin (*Scheme 85*) [449]. L-Daunosamine [427] and 2,3-dideoxy-D-*arabino*-hexose have also been derived from the "naked sugars" **271** and (+)-**274**, respectively [450,451].

7.3 Long-Chain Carbohydrates and Analogs

Mukaiyama cross-aldolizations of 2,3-*O*-isopropylidene-D-glyceraldehyde (*R*)-**37** with bicyclic ketones (+)-**275** and (−)-**275** are highly diastereoselective and lead to aldols (+)-**298** and (+)-**300**, respectively. For steric reasons only the *exo* face of the enoxysilanes (−)-**277** and (+)-**277** (see *Scheme 86*) can be attacked by the complexes [(*R*)-**37**] · $TiCl_4$. Depending on the configuration of the enoxysilanes, the (*R*)-**37** complex involving the 2-O or that involving the 3-O coordination with $TiCl_4$ reacts (transition states **297** vs. **299**, *Scheme 86*) to minimize steric repulsions. Baeyer–Villiger oxidation of aldol (+)-**301** generates the corresponding uronolactone that reacts with benzyl alcohol in DMSO and in the presence of cesium fluoride (base catalyst) to give the benzyl uronate (+)-**301** after 10 minutes at room temperature. On staying at this temperature in the presence of CsF, or of another base such



Scheme 86

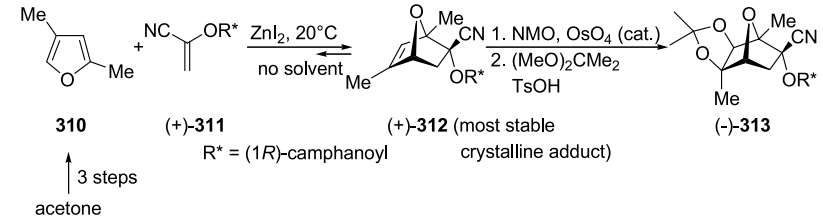
Cross-aldol reactions with "naked sugars": synthesis of long-chain carbohydrates, indolizidines, and analogs

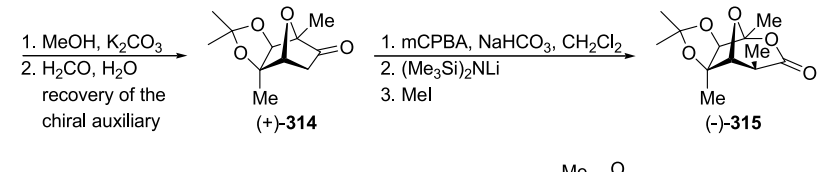
as K_2CO_3, (+)-**302** is epimerized into the more stable uronic ester (−)-**302**. Silylation of (+)-**301**, followed by hydrogenolysis of the benzyl ester and subsequent Curtius rearrangement (one-pot operation), provides the partially protected 5-amino-5-deoxy-D-*erythro*-D-*talo*-octofuranosides (+)-**303**. Desilylation of (+)-**303** and subsequent hydrogenolysis liberates the amine that engenders an imine with the aldose which is hydrogenated under the hydrogenolysis conditions, to give (−)-1,5-dideoxy-1,5-imino-2,3:7,8-di-*O*-isopropylidene-D-*erythro*-D-*talo*-octitol (−)-**304**. Acidic hydrolysis gives the unprotected 1,5-dideoxy-1,5-imino-octitol (−)-**305** which can be converted into the unprotected pentahydroxyindolizidine (−)-**306** in one step under Mitsunobu conditions. The same methods applied to benzyl uronate (−)-**302** leads to (+)-**307** that can be converted into (−)-1,5-dideoxy-1,5-imino-D-*erythro*-L-*allo*-octitol (−)-**308** and pentahydroxyindolizidine (−)-**309**. Following similar routes, aldol (+)-**300** has been converted into stereomeric 1,5-dideoxy-1,5-imino-alditols and pentahydroxyindolizidines. With the possibility to substitute the 7-oxabicyclo[2.2.1]heptan-2-one at C(5) and C(6) with two protected hydroxy groups in all possible configurations (see ❯ *Scheme 81*) or with other groups, the reaction sequence of ❯ *Scheme 86* makes possible the preparation of a large number of octose derivatives. This demonstrates the high efficiency and versatility of the "naked sugar" methodology for the de novo synthesis of unusual and complicated monosaccharides [452,453].

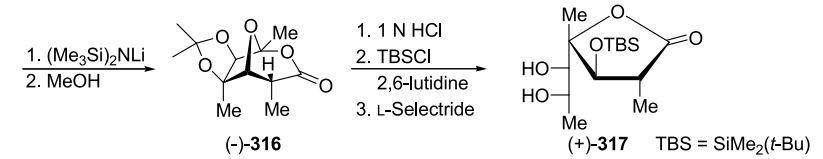
7.4 "Naked Sugars of the Second Generation": Synthesis of Doubly Branched-Chain Sugars

Doubly branched heptono-1,4-lactones as well as polypropionate fragments have been obtained from 2,4-dimethylfuran (**310**) via its Diels–Alder addition to 1-cyanovinyl-(1′*R*)-camphanate (+)-**311** [454]. Without solvent, the ZnI_2-catalyzed and reversible cycloaddition leads to a major crystalline diastereomeric adduct (+)-**312**. Double hydroxylation of the alkene moiety of (+)-**312**, followed by diol protection as an acetonide provides (−)-**313**. Methanolysis followed by treatment with formaline liberates ketone (+)-**314**. Baeyer–Villiger oxidation and subsequent α-methylation generates the *exo*-α-methyluronolactone (−)-**315**. Quenching of the lithium enolate of (−)-**315** with MeOH at –50 °C gives the *endo*-α-methyluronolactone (−)-**316**. Acidic hydrolysis of (−)-**316** and subsequent silylation and reduction forms (+)-**317** as the major heptono-1,4-lactone (❯ *Scheme 87*). Similarly, enantiomers of this doubly branched sugar can be prepared starting from adduct (−)-**318** obtained by addition of 2,4-dimethylfuran (**310**) to 1-cyanovinyl (1′*S*)-camphanate [(−)-**311**)]. After conversion of (−)-**318** into dimethyl acetal **319**, regio- and *exo*-face selective hydroboration and further transformations generate the doubly branched uronic acid (−)-**321** [455].

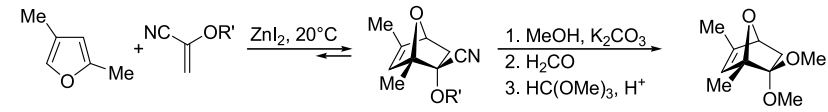
The method of thermodynamic diastereoselection (through diastereoselective crystallization of equilibrating adducts, see ❯ *Scheme 87*) has been applied to furan derivatives bearing chiral auxiliaries that can be recovered readily. For instance, the acetal of (2*S*,3*S*)-butane-2,3-diol and furfural is equilibrated in molten maleic anhydride with one major crystalline product [456]. In a similar way (1,*S*)-camphanate of furfuryl alcohol undergoes Diels–Alder addition in molten maleic anhydride giving one major crystalline adduct (+)-**322** [457] that has been converted into doubly branched carba-hexopyranoses and derivatives [458], and into the new 2,6-dideoxy-2,6-iminoheptitol **324** (❯ *Scheme 88*) [459].

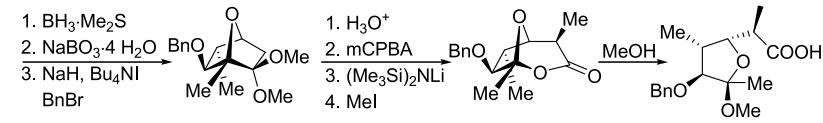






a





b

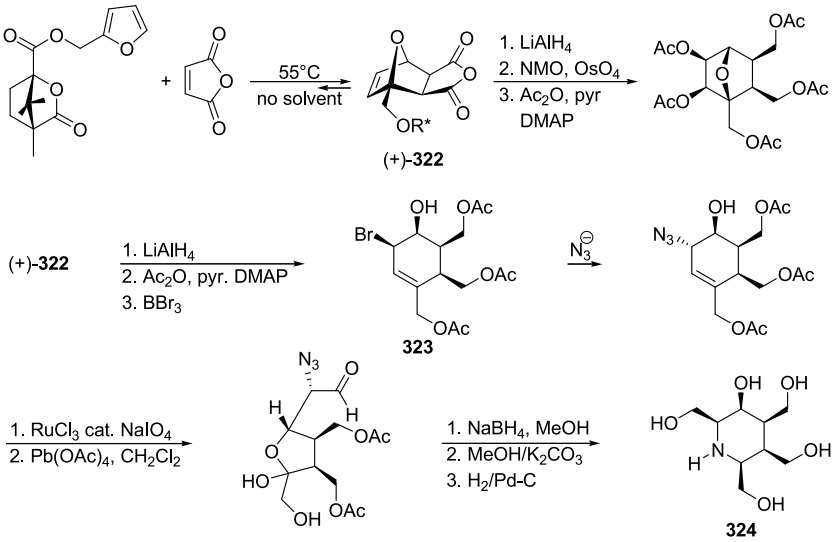
Scheme 87
Applications of the "naked sugars of the second generation" to the synthesis of branched-chain sugar derivatives

7.5 [4+3]-Cycloadditions of Furan

Apart from [4+2]-cycloadditions, furans undergo [4+3]-cycloadditions and dipolar [2+3]-cycloadditions. The reaction of nitrones to furan has been applied to prepare (+)-Norijimycin [460], aminopolyols [461,462,463] (+)- and (−)-1-deoxynojirimycin, (+)- and (−)-1-deoxy-idononojirimycin [464].

7.5.1 Synthesis of Semi-Protected 2,6-Anhydroheptitols

The [4+3]-cycloaddition of furan to the 1,3-dibromo-2-oxyallyl cation generates, after dehalogenation with the zinc/copper couple, 8-oxabicyclo[3.2.1]oct-6-en-3-one (**325**). It

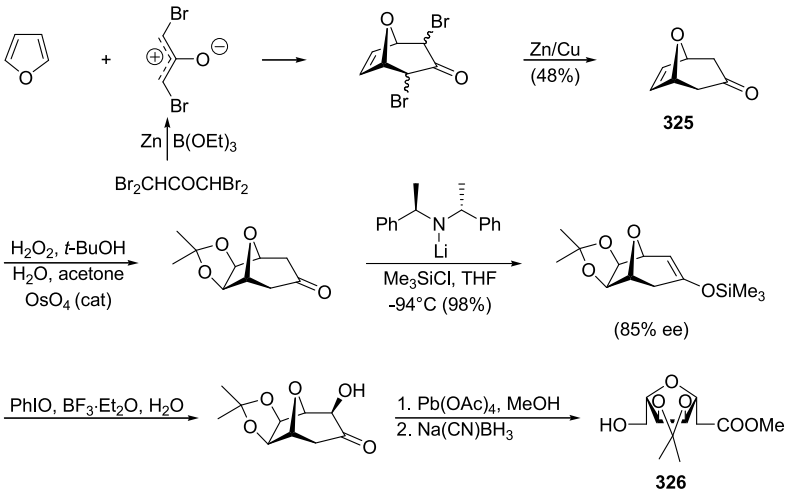


Scheme 88
Synthesis of a doubly branched iminoalditol

was converted into the semi-protected methyl 3,6-anhydro-2-deoxy-L-allo-heptonate **326** (*Scheme 89*) [465].

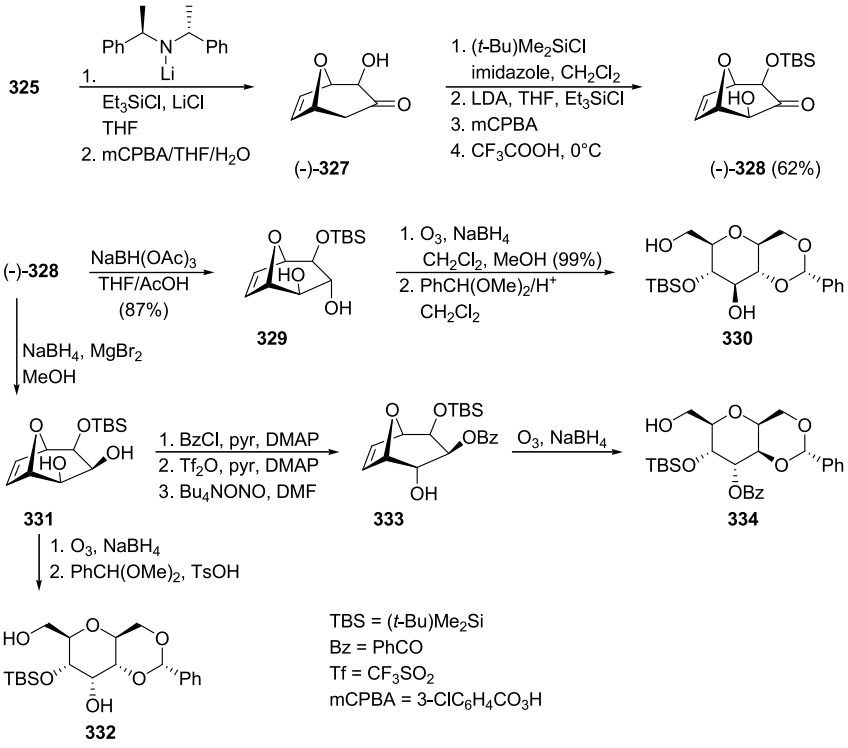
Hoffmann and co-workers [466,467] have developed general procedures that convert **325** into a number of 2,6-anhydroheptitols as enantiomerically pure isomers (*Scheme 90*). The treatment of **325** with (+)-bis[(*R*)-1-phenylethyl]amine, nBuLi, and Et_3SiCl/Et_3N (in situ quench) in THF gives the corresponding silyl enol ether that is oxidized with *meta*-chloroperbenzoic acid (mCPBA) in THF/H_2O into (−)-**327**. The β-hydroxyketone so-obtained is silylated and enolized on its turn with lithium diisopropylamide (LDA) in the presence of Et_3SiCl/Et_3N. The intermediate enol ether is then oxidized on its turn with mCPBA. After treatment with $CF_3COOH/THF/H_2O$ (−)-**328** is isolated in 62% overall yield. The reduction of (−)-**328** with $NaBH(OAc)_3$ gives *endo* alcohol **329**, whereas reduction with $NaBH_4$ and $MgBr_2$ in MeOH affords the exo alcohol **331**. Oxidative cleavage of the endocyclic alkene moieties of **329** and **331**, and subsequent reduction with $NaBH_4$ and acetalization with benzaldehyde dimethyl acetal provides alditols **330** and **332**, respectively. Selective monobenzoylation of **331**, followed by inversion of alcohol at C-2 provides **333** that is then converted into the 2,6-anhydroheptitol derivative **334** (*Scheme 90*).

Enantiomerically pure 3-oxo-8-oxabicyclo[3.2.1]oct-2-yl derivatives have been obtained by [4+3]-addition of furan with chiral 1,2-dioxyallyl cation engendered in situ by acid-catalyzed heterolysis of enantiomerically pure mixed acetals derived from 1,1-dimethoxyacetone and enantiomerically pure secondary benzyl alcohols [467]. For instance, mixed acetal **335** is enolized into the silyl enol ether **336**. In the presence of a catalytic amount of trimethyl silyl triflate, **336** generates a cationic intermediate that adds to furan at –95 °C giving adduct (−)-**337**. The latter can be enolized regioselectively and oxidized with mCPBA giving, after esterification, pivalate **338**. Reduction with $NaBH_4$ and subsequent alkene cleavage generates



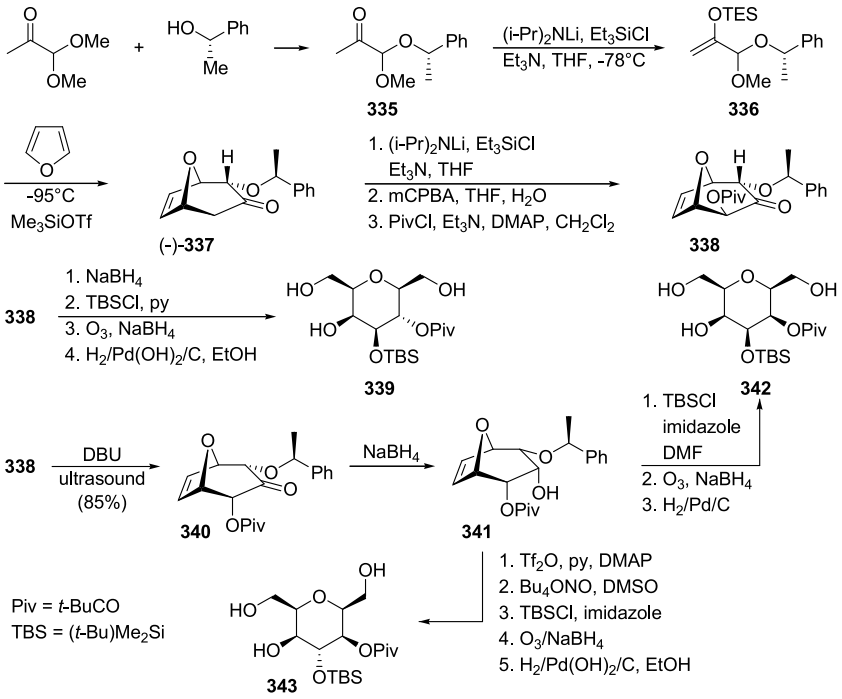
Scheme 89

Simpkins's synthesis of methyl 3,6-anhydro-2-deoxy-L-*allo*-heptanate



Scheme 90

Hoffmann's synthesis of 2,6-anhydroheptitols starting from 8-oxabicyclo[3.2.1]octan-3-one

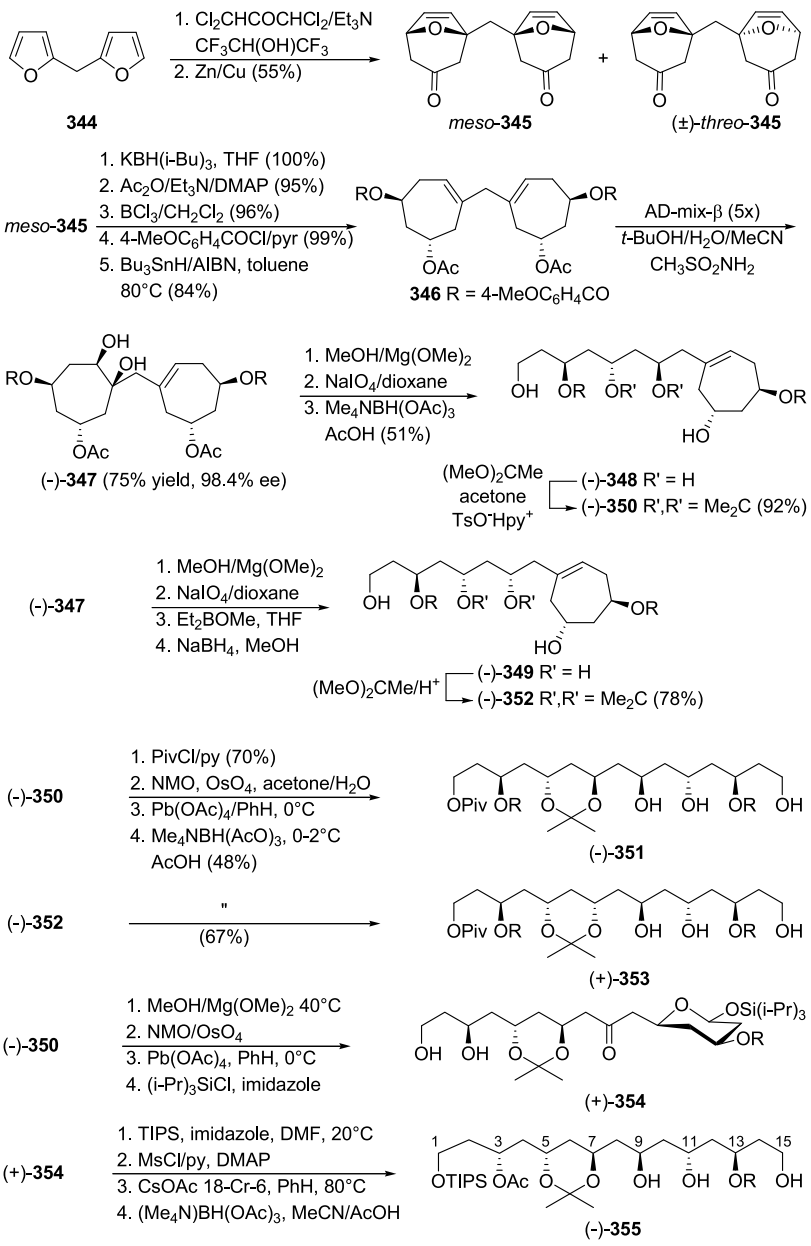


Scheme 91
Hoffmann's asymmetric [4+3]-cycloaddition of furan: total synthesis of 2,6-anhydroheptitols

339. Treatment of pivalate **338** with 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU) and ultrasound gives the 2-epimer **340**. Its reduction with $NaBH_4$ gives **341**. Subsequent alkene cleavage generates **342**. Inversion of the alcohol moiety C-3 of **342**, followed by alkene cleavage furnishes the 2,8-anhydroheptitol **343** (*Scheme 91*) [466,467].

7.6 Synthesis of Long-Chain Polyketides and Analogues

Polyketides include 1,3-polyols that can be seen as long-chain polydeoxyalditols. A noniterative and asymmetric synthesis of all possible stereoisomeric pentadecane-1,3,5,7,9,11,13,15-octols [468], of 3,5,7,9,11,13,15-heptahydroxypentadecanals [469] and corresponding 1,7-dioxaspiro[5.5]undecanes [470] has been developed applying the double [4+3]-cycloaddition of 1,1,3-trichloro-2-oxyallyl cation to 2,2′-methylenedifuran (**344**). Treatment of **344** with 2.6 equivalents of 1,1,3,3-tetrachloroacetone and Et_3N in $CF_3CH(OH)CF_3$ gives a mixture of hexachlorinated adducts that are not isolated but directly reduced with Zn/Cu in MeOH saturated with NH_4Cl. This produces a 45:55 mixture (55% based on **344**) of *meso*-**345** and (±)-*threo*-**345** that is separated by crystallization from furan first and then from H_2O. The expensive $CF_3CH(OH)CF_3$ solvent can be exchanged for toluene and the [4+3]-cycloaddition

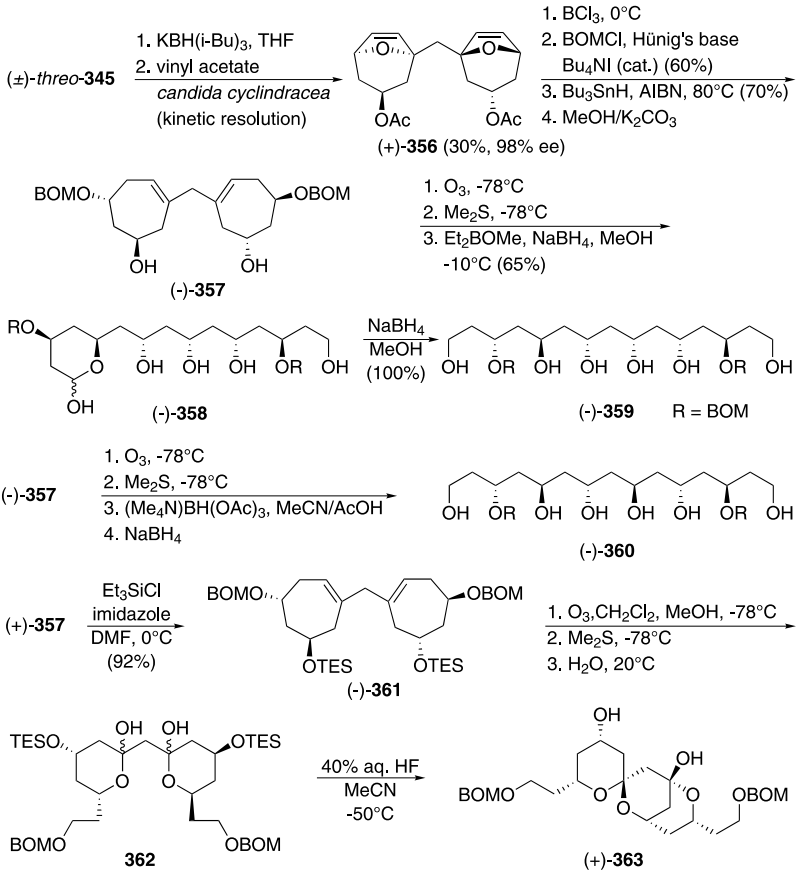


Scheme 92
Noniterative synthesis of long-chain 1,3-polyols

uses 1,3-dibromo-1,3-dichloroacetone with Et_2Zn. Under these conditions the same mixture of *meso*-**345** and *threo*-**345** is obtained in 40% yield [471].

Acetylation of diol *meso*-**345**, subsequent S_N2' opening with BCl_3, and esterification with *p*-methoxybenzoyl chloride/pyridine gives a dichloride that is dechlorinated into **346**.

Desymmetrization of **346** is realized by means of the Sharpless asymmetric dihydroxylation giving diol (−)-**347** as the major product (98.4% ee). After selective deacetylation of (−)-**347** with MeOH/$Mg(OMe)_2$, Malaprade oxidative cleavage of the vicinol diol, Evans' reduction [472] ($Me_4NBH(OAc)_3$) of the intermediate oxo-aldehyde delivers (−)-**348** in 75% overall yield, whereas Et_2BOMe and $NaBH_4$ (Narasaka's reduction [307,473]) furnishes (+)-**349**. Selective esterification of the primary alcohol moiety of (−)-**350** (acetonide of (−)-**348**) as the pivalate and subsequent oxidative cleavage of the cycloheptene ring and Evans' aldol reduction provides the semi-protected pentadecaneoctol derivative (−)-**351**. Similarly, stereoisomer (+)-**353** was derived from (−)-**352**. Two further stereoisomeric polyols



Scheme 93
Expeditious asymmetric synthesis of 3,5,7,9,11,13,15-heptahydroxypentadecanals and derivatives

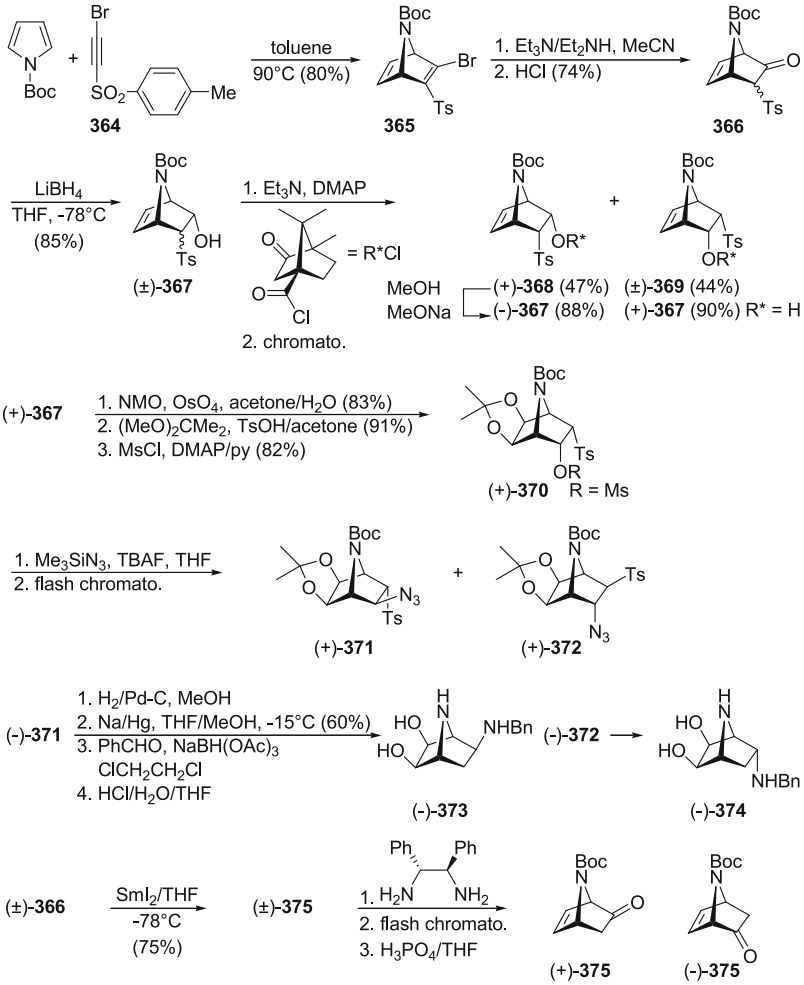
can be obtained by applying Narasaka's rather than Evans' reductive conditions. Starting with AD-mix-α, the enantiomers of these eight polyols are obtained with the same ease. Selective methanolysis of the acyclic *p*-methoxybenzoate (−)-**350** gives a diol that was then converted into (+)-**354**. Selective monosilylation of diol (+)-**354**, subsequent mesylation of the secondary alcohol and displacement with CsOAc generates, after Evans' reduction, the stereoisomeric semi-protected polyol (−)-**355**, or its 9-epimer under Narasaka's reduction conditions. Thus, combining the two enantiomeric Sharpless dihydroxylations to desymmetrize **346**, the Evans' and Narasaka's reduction techniques and the inversion of alcohol at C-3 (as shown for (+)-**354** → (−)-**355**) all possible stereoisomers of pentadecane-1,3,5,6,7,9,13,15-octols can be made quickly (*Scheme 92*). The method has been used to prepare the polyol subunit of the polyene macrolide antibiotic RK-397 [474]. It has been used also to prepare enantiomerically enriched (98%) 1-aminopentadecane-3,5,7,9,11,13,15-heptols and 1,15-diaminopentadecane-3,5,7,9,11,13-hexols and derivatives [475].

Racemic (±)-*threo*-**345** has been converted into diacetate (+)-**356** via enzymatic kinetic resolution of the corresponding racemic diol [476]. Ring opening of (+)-**356** with BCl_3, quenching of the bisborate as BOM ($PhCH_2OCH_2$bisacetal, reductive dechlorination and methanolysis of the diacetate provides diol (−)-**357**. Ozonolysis of diene (−)-**357** and subsequent Narasaka's reduction using a controlled amount of $NaBH_4$ allows the isolation of hemiacetal (−)-**358**. With an excess of $NaBH_4$, the semi-protected octol (−)-**359** is obtained. Applying Evans' conditions for reduction of the bis(β-hydroxyketone) resulting from the ozonolyis of diene (−)-**357** (−)-**360** is obtained (*Scheme 93*) [469]. With the silyl diether (−)-**361**, ozonolysis provides a mixture of hemiacetals **362** that was isomerized into spiroketal (+)-**363** (*Scheme 93*) [470].

8 Diels–Alder Addition of Pyrroles: "Naked Aza-Sugars"

Cycloaddition of 2-bromoethynyl *p*-tolyl sulfone (**364**) to *N*-Boc pyrrole gives 7-azanorbornadiene derivative **365** [477]. Treatment of **365** with Et_2NH and Et_3N, followed by hydrolysis with 10% aqueous HCl affords ketone **366** as a 1:1.2 mixture of *endo* vs. *exo* stereoisomers. Reduction of **366** with $LiBH_4$ provides exclusively the *endo*-alcohols (±)-**367** that were resolved by flash chromatography of their diastereomeric (1*S*,4*R*)-camphanates (+)-**368** and (−)-**369**. Methanolysis of these esters delivers (−)-**367** and (+)-**367**, respectively [478]. *Exo* selective dihydroxylation of alkene (+)-**367** and subsequent diol protection as an acetonide and *endo* alcohol mesylation provides (+)-**370** which was reacted with Me_3SiN_3 and Bu_4NF in THF. At 25° C a 5:1 mixture of (+)-**371** and (+)-**372** was obtained. These compounds are separated by flash column chromatography. Starting from (−)-**367** the enantiomers (−)-**371** and (−)-**372** are obtained [479]. Hydrogenation of (−)-**371**, desulfonylation with Na/Hg, reductive amination of benzaldehyde, and acidic removal of the Boc group provides (−)-**373**, an inhibitor of β-glycosidase from almonds (*Scheme 94*). Applying the same reaction sequence to (−)-**372** gives (−)-**374**. Desulfonylation of (±)-**366** with SmI_2 generates racemic *N*-Boc-7-azanorbornenone that can be resolved into (+)-**375** and (−)-**375** applying Alexakis optical resolution [480].

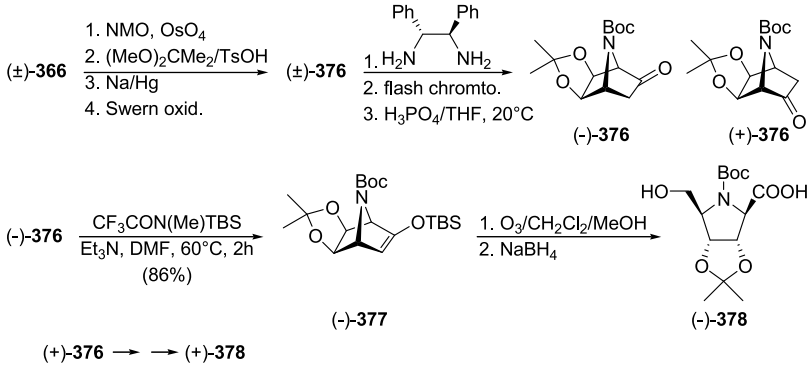
Similarly, ketone (±)-**376** obtained by dihydroxylation of **366**, followed by diol protection as acetonide, desulfonylation with Na/Hg, and Swern oxidation has been resolved by the Alex-



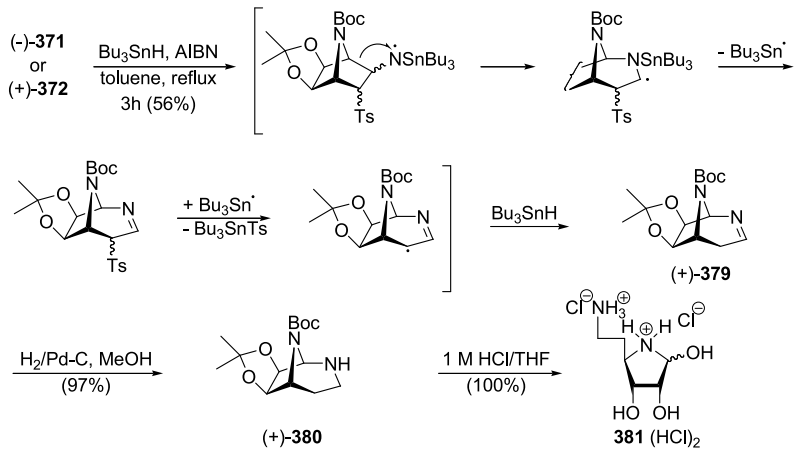
Scheme 94
Synthesis of enantiomerically pure 5-amino-7-aza-norbornane-2,3-diols and of the *N*-Boc-7-azanorbornenones ("Aza-naked sugars")

akis method [481] giving ketones (−)-**376** and (+)-**376**. The silyl enol ether (−)-**377** derived from (−)-**376** is ozonolyzed into the L-proline derivative (−)-**378** after reductive work-up with $NaBH_4$. Enantiomer (+)-**378** is obtained with the same ease from (+)-**376** (*Scheme 95*). These 2,5-dideoxy-2,5-imino-allonic acid derivatives have been used in the synthesis of tripeptides containing them [482].

Desulfonylation of (−)-**371** and (+)-**372** with Bu_3SnH and azoisobutyronitrile (AIBN) gives the 2,8-diazabicyclo[3,2,1]oct-2-ene derivative (+)-**379** as a major product. The proposed mechanism for this transformation is outlined in *Scheme 96*. Catalytic hydrogenation of (+)-**379** affords (+)-**380**. Upon treatment with aqueous HCl in THF, quantitative formation of **381** · HCl is observed. Starting from (+)-**371** and (−)-**372**, the enantiomer of **381** · $(HCl)_2$



Scheme 95
Synthesis of 2,5-dideoxy-2,5-imino-D- and L-allonic acid derivatives



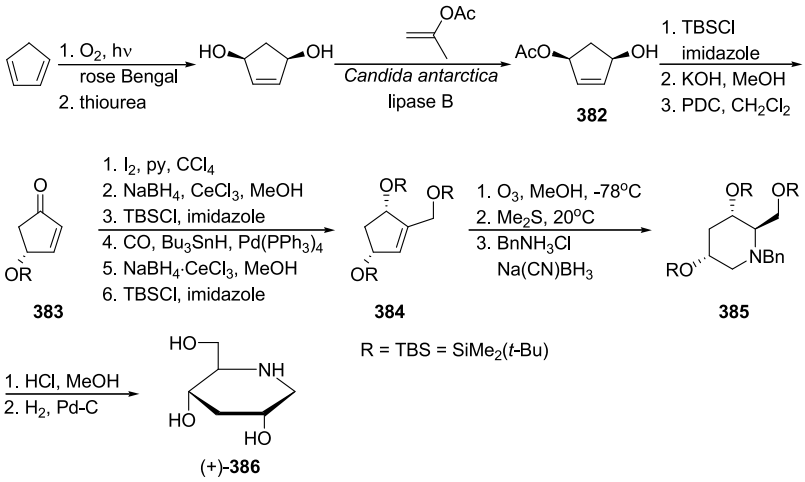
Scheme 96
Synthesis of 4,6-diamino-4,5,6-trideoxy-D-*ribo*-hexofuranose

is obtained in the same way. The 4,6-diamino-4,5,6-trideoxyhexoses are potent inhibitors of α-mannosidases [483].

9 Carbohydrates and Analogs From Achiral Polyenes

9.1 From Cyclopentadiene

Xylitol has been derived from the product of photo-oxidation of cyclopentadiene [484] which is (*Z*)-(4*RS*)-4,5-epoxypent-2-enal. An elegant total synthesis of acosamine and of daunosamine (13 steps, 15% overall yield) was developed at Roche starting from cyclopentadiene [485]. The key-step is the asymmetric monohydroboration of 5-methylcyclopentadiene with (−)-di-3-pinanylborane.



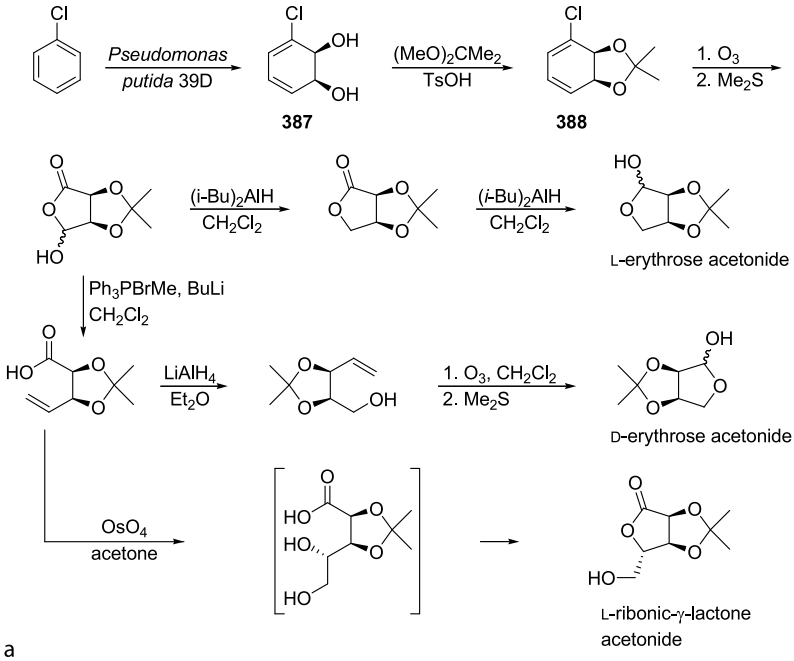
Scheme 97
Johnson's synthesis of 1,3-dideoxynojirimycin

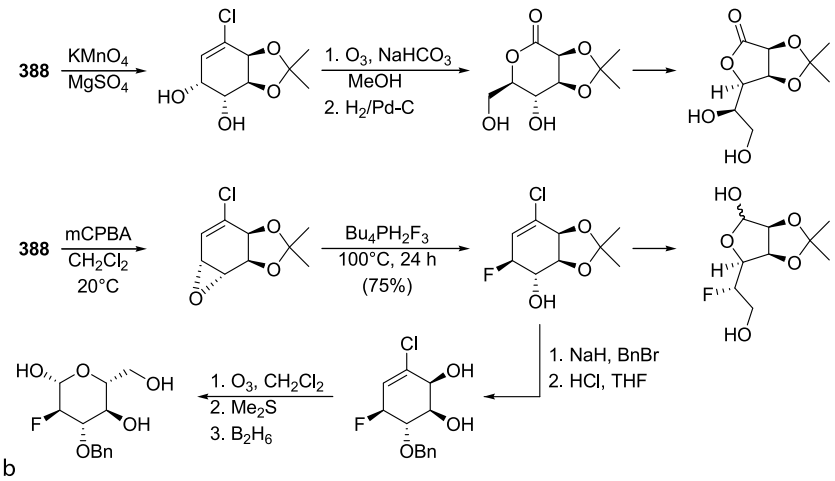
A total synthesis of 1,3-dideoxynojirimycin ((+)-**386**) starting from cyclopentadiene has been proposed by Johnson and co-workers [486]. Photooxidation of cyclopentadiene and reductive work-up with thiourea generates *cis*-cyclopent-2-ene-1,4-diol which is monoacylated with high enantioselectivity (>99% ee) with isoprenyl acetate and *Candida antarctica* lipase B (Novo Nordisk SP 435) to give **382**. After silylation of **382** and subsequent treatment with KOH and oxidation, enantiomerically pure enone **383** is obtained [487]. Treatment of **383** with iodine and pyridine leads to α-iodination [488]. Successive reduction under Luche conditions, alcohol silylation, carbonylation of the iodoalkene, reduction of the obtained enal, and alcohol silylation leads to **384**. Ozonolysis of **384** gives the corresponding keto-aldehyde which is then transformed into **385** via reductive amination with high diastereoselectivity (*syn*:*anti*> 20:1). Deprotection delivers (+)-**385** (*Scheme 97*). A similar approach has converted enantiomerically pure (2*R*,3*R*)-2,3-isopropylidenedioxycyclopent-4-en-1-one derived from cyclopentadiene [489,490] into (−)-1-deoxymannonojirimycin and (−)-1-deoxytalonojirimycin (1,5-dideoxy-1,5-imino-D-talitol) [491]. Enone **383** has been converted to (2*R*,3*S*)-2,3-bis[(*tert*-butyl)dimethylsilyloxy]cyclopent-4-en-1-one and then to (+)-1-deoxynojirimycin [492].
Mehta and co-workers have presented [493,494] a synthesis of racemic deoxynojirimycin analogs and isofagomin analogs, starting from the Diels–Alder adduct of vinyl acetate to 5,5-dimethoxy-1,2,3,4-tetrachlorocyclopentadiene.

9.2 From Benzene and Derivatives

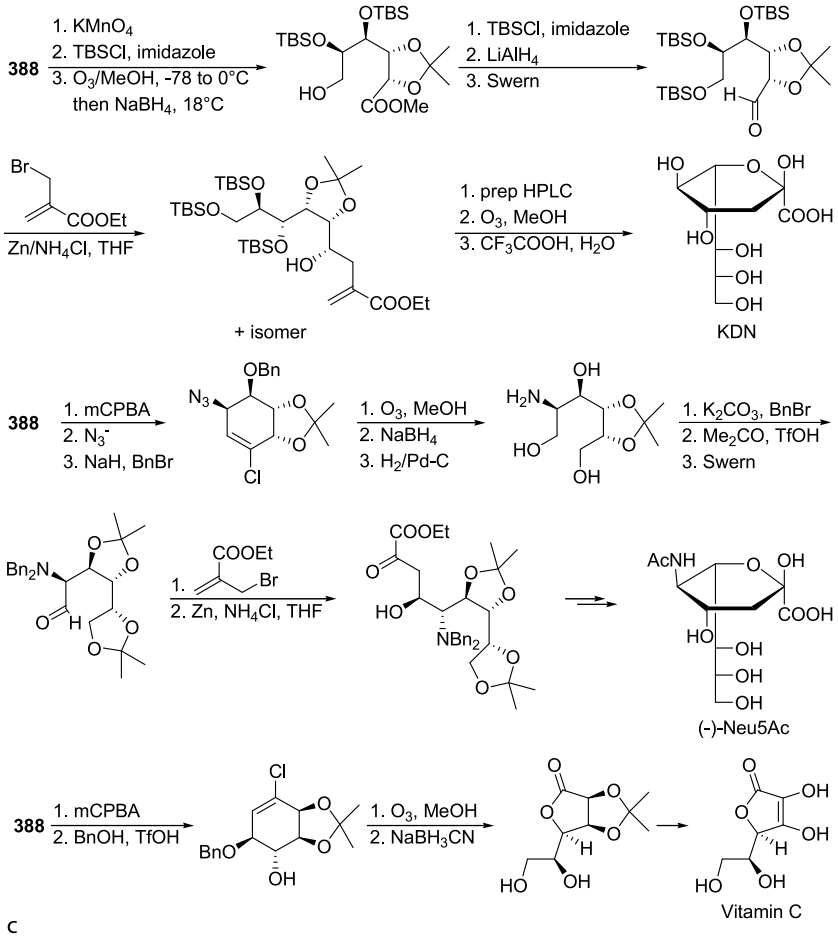
Dioxygenases present in the blocked mutants of *Pseudomonas putida*, a soil bacterium, degrades benzene and its derivatives into cyclohexa-3,5-diene-1,2-diols. With chlorobenzene, diol **387** is obtained with >99% ee. This compound is converted in a few chemical steps into

tetrose, pentose, and hexose derivatives (Scheme 98a) [495]. Further applications of this methodology are shown in Scheme 98b [496] and Scheme 98c [497]. A recombinant strain of *Escherichia coli* expressing naphthalene dioxygenase from *Pseudomonas* sp. NCIB





Scheme 98
Hudlicky's syntheses of tetrose, pentose and hexose derivatives. (see also on the following page)



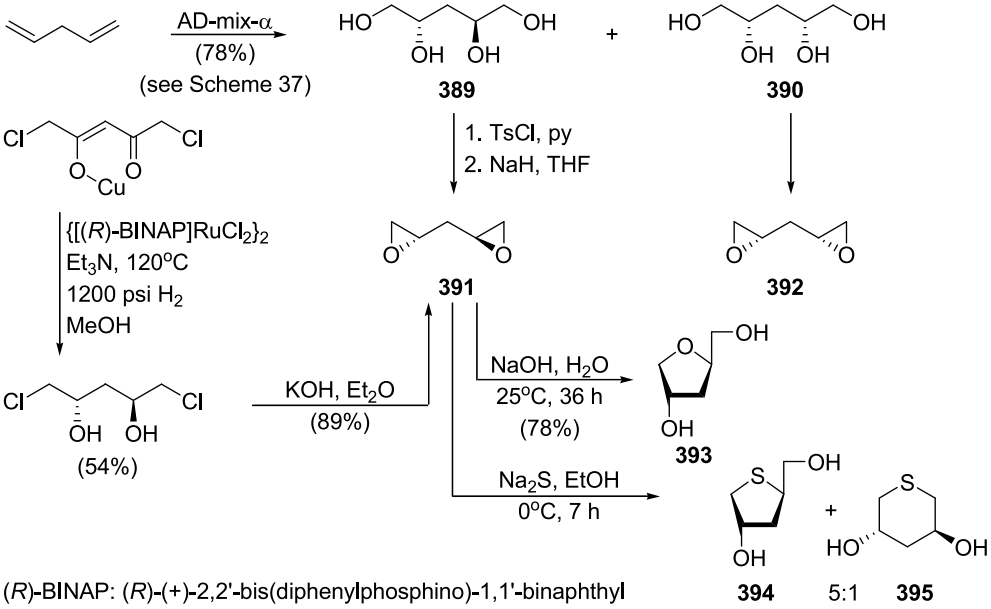
Scheme 98
(continued) Banwell's total syntheses of KDN, Neu5Ac, and Vitamin C

9816-4 has been found to oxidize specifically *N*-methyl-2-pyridone into the *cis*-5,6-dihydro-5,6-dihydroxy derivative [498].

9.3 From Cycloheptatriene and Cyclooctatetraene

Cycloheptatriene has been converted into L-glucose via *Pseudomonas cepacia* lipase-mediated desymmetrization of a *meso*-3-*O*-protected cyclohept-6-ene-1,3,5-triol using isopropenyl acetate as the solvent [499,500,501].

Cycloheptatrienone (tropone) has been converted into heptitol derivatives via the optical resolution of the (tropone)$Fe(CO)_3$ complex [502]. Racemic methyl β-allopyranoside and a 2*C*-branched analogue have been derived from cyclooctatetraene [503].



Scheme 99
Conversion of penta-1,4-diene into pentitol derivatives

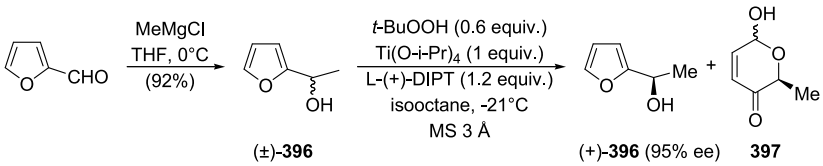
9.4 From Penta-1,4-Diene

When applied to penta-1,4-diene, the Sharpless asymmetric dihydroxylation forms a 1:1 mixture of (2*S*,4*S*)- and (2*S*,4*R*)-penta-1,2,4,5-tetrols **389** and **390** which can be converted to diepoxides **391** and **392**, respectively (*Scheme 99*) [504]. A stereo- and enantioselective synthesis of **391** is possible starting from 1,5-dichloropenta-2,4-diene applying Noyori's asymmetric hydrogenation [505]. Diepoxide **391** has been converted into alditol **393** and thioalditols **394** and **395**.

9.5 From Furfural

Furfural (furfuraldehyde) is a very inexpensive starting material obtained from the left-overs of agriculture (acidic distillation of straw and brans) [506]. It adds MeMgCl giving racemic 1-(2-furyl)ethanol ((±)-**396**) which has been resolved easily using either enzymatic methods [507,508,509] or metal-catalyzed kinetic resolution [510,511]. For instance, the reaction (±)-**396** with *t*-butyl hydroperoxide in the presence of catalytic amounts of $Ti(O\text{-}i\text{-}Pr)_4$ and L-(+)-di-isopropyl tartrate (L-(+)-DIPT) gives a mixture of optically active (+)-**396** and pyranone **397**. These compounds are separated readily and isolated in 32 and 53% yield, respectively (*Scheme 100*) [512,513].

Utilizing Novozyme 435 (immobilized on acrylic resin) to catalyze the transacetylation of (±)-**396** with isopropenyl acetate in $(i\text{-}Pr)_2O$, (−)-**396** is obtained in 36% yield and with an



Scheme 100
Oxidative kinetic resolution of 1-(2-furyl)ethanol

enantiomeric excess better than 95% (55% conversion rate). If the reaction is stopped at 45% conversion rate, ester **398** is obtained in > 97% ee. The same enzyme catalyzes the fast hydrolysis of **398** into (+)-**396** (*Scheme 101*) [514,515]. Oxidation of (+)-**396** with *N*-bromosuccinimide in THF/H_2O gives **399**. Subsequent oxidation with $CrO_3 \cdot NH_4Cl$ in CH_2Cl_2 furnishes **400** that is reduced into pure **401** under Luche's conditions ($NaBH_4/CeCl_3$) at –78 °C, and into a 2:1 mixture of **401** and **402** at 0 °C. Compound **402** is isolated and converted into pyranoses **403**–**405** as outlined in *Scheme 101* [514,515]. *C*-linked disaccharides have been derived from 1,4-di(2-furyl)butane-1,4-diol in a similar way [516].

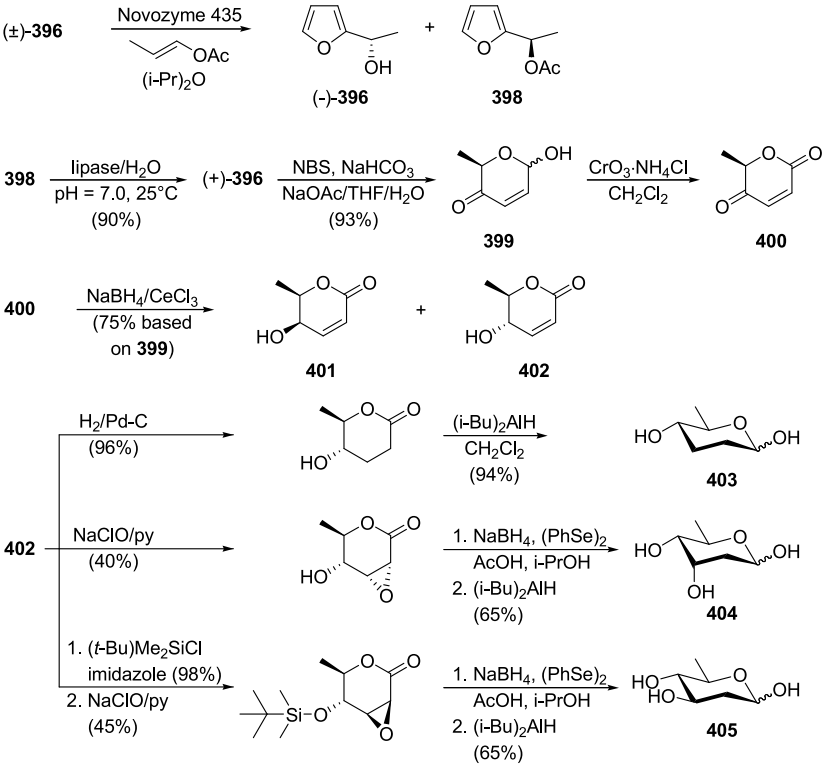
10 Kinetic Resolution of Racemic Allylic Electrophiles

The Katsuki–Sharpless asymmetric epoxidation of racemic diol (±)-**406** (obtained by allylation of (*E*)-crotonaldehyde) gives, after chromatographic separation, the *erythro*-epoxide (+)-**407** (33% yield, > 95% ee) and (−)-**406** (*Scheme 102*). These intermediates have been converted into (+)-D-olivose and (+)-D-ditoxose [517,518,519]. In a similar manner, asymmetric epoxidation of dienol (±)-**408** leads to (+)-**409**, which has been converted into (+)-oliose, and to (−)-**408**, which has been converted into (+)-cymarose [520,521].

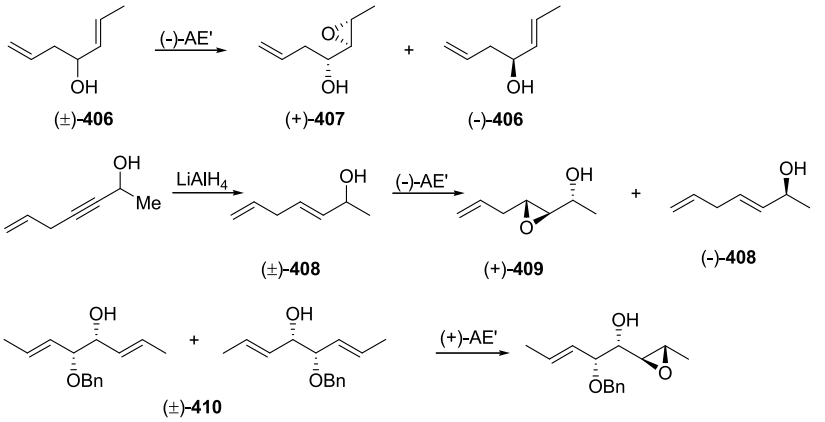
D- and L-Chalcose have been prepared in a similar way from the racemic mixtures of *threo*-dipropenylglycol (±)-**410** obtained by reductive dimerization of crotonaldehyde [522].

A simple, divergent, asymmetric synthesis of the four stereomers of the 3-amino-2,3,6-trideoxy-L-hexose family has been proposed by Dai and co-workers [523] which is based on the Katsuki–Sharpless asymmetric epoxidation of allylic alcohols (±)-**408**. *N*-Trifluoroacetyl-L-daunosamine, *N*-trifluoro acetyl-L-acosamine, *N*-benzoyl-D-acosamine, and *N*-benzoyl-D-nitrosamine have been derived from methyl sorbate via the methyl 4,5-epoxy-(*E*)-hex-2-enoates obtained via a chemoenzymatic method [524]. Application of the Katsuki–Sharpless enantioselective epoxidation to racemic mono-*O*-benzylated divinylglycol has allowed us to prepare enantiomerically pure L-*lyxo* and D-*lyxo*-pentoses and analogs [525,526,527,528].

Palladium-catalyzed asymmetric allylic alkylation [529] of suitable amines with two equivalents of racemic butadiene monoxide ((+)-**411**) allows for the expedient synthesis of *trans*- and *cis*-2,5-dihydropyrroles derivatives **416** and **417** that are versatile chirons towards the synthesis of a wide variety of iminosugars [530]. In the presence of 0.4% $[(allyl)PdCl)_2]$, 1.2% of enantiomerically pure diphosphine (*R*,*R*)-ligand, and Na_2CO_3 a 1:1 mixture of (±)-**411** and phthalimide reacted in CH_2Cl_2 at room temperature giving (*S*)-**412** in 99.6% yield and

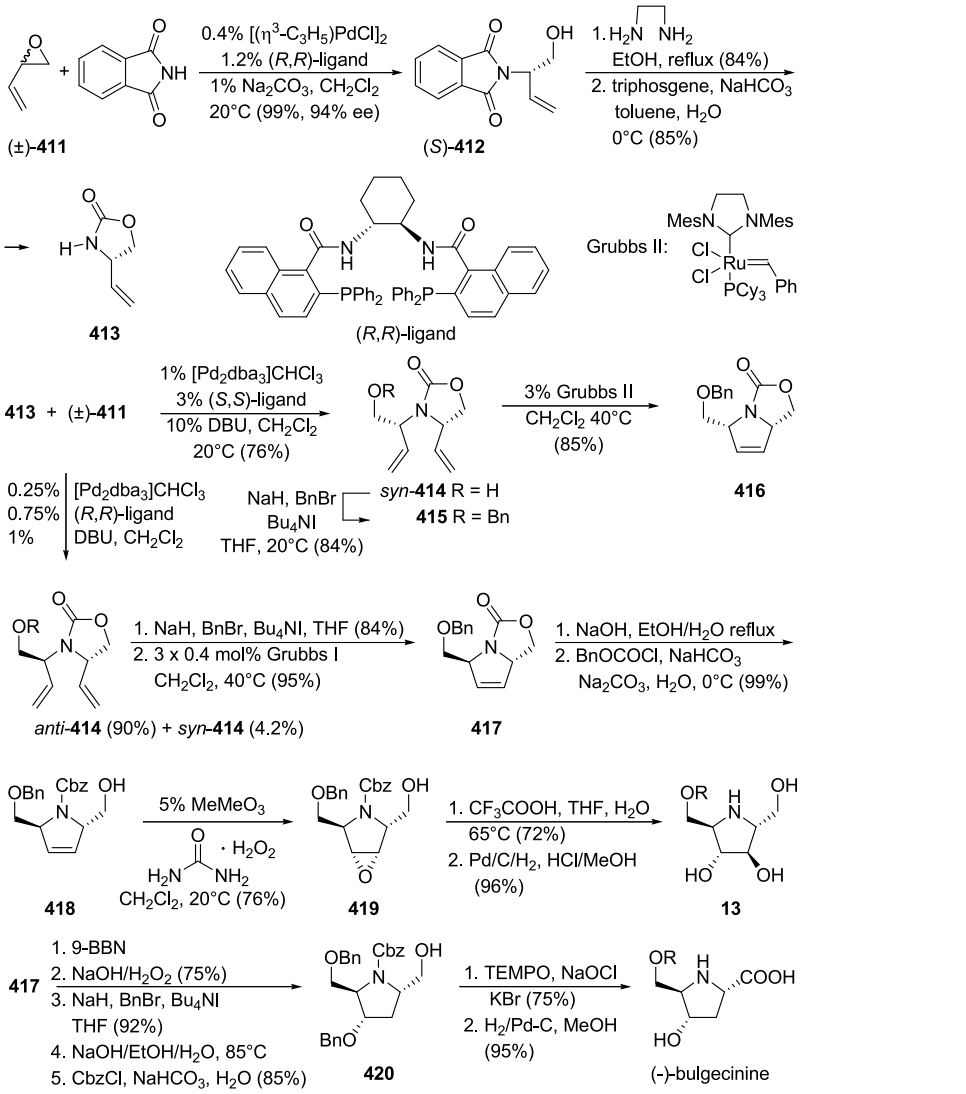


Scheme 101
Synthesis of trideoxy and dideoxyhexoses starting from furfural



(+)-AE': *t*-BuOOH, Ti(O-i-Pr)$_4$, (+)-diethyl tartrate, CH_2Cl_2, -20 °C
(-)-AE': *t*-BuOOH, Ti(O-i-Pr)$_4$, (-)-diethyl tartrate, CH_2Cl_2, -20 °C

Scheme 102
Kinetic resolution through Katsuki–Sharpless epoxidation of allylic alcohols



Scheme 103

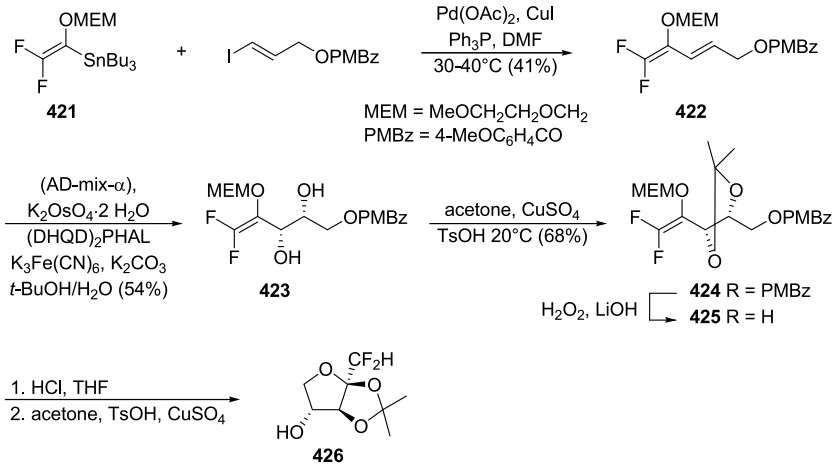
Trost's dynamic kinetic asymmetric amine allylation applied to the synthesis of 3,5-dideoxy-2,5-imino-D-mannitol and (−)-bulgecinine

94.4% ee. Under these conditions deracemization of (±)-**410** occurs and a dynamic kinetic asymmetric transformation is realized. Both enantiomers of epoxide (±)-**410** are converted into a single product of allylation (*S*)-**412**. Recrystallization from EtOH provides (*S*)-**412** with >99% ee. Unmasking of the amino group by aminolysis of (*S*)-**412**, followed by cyclization with triphosgene affords oxazolidinone **413**, the substrate for a second dynamic kinetic asymmetric reaction with epoxide (±)-**411** (*Scheme 103*). Thus, combining (±)-**411** with **413**

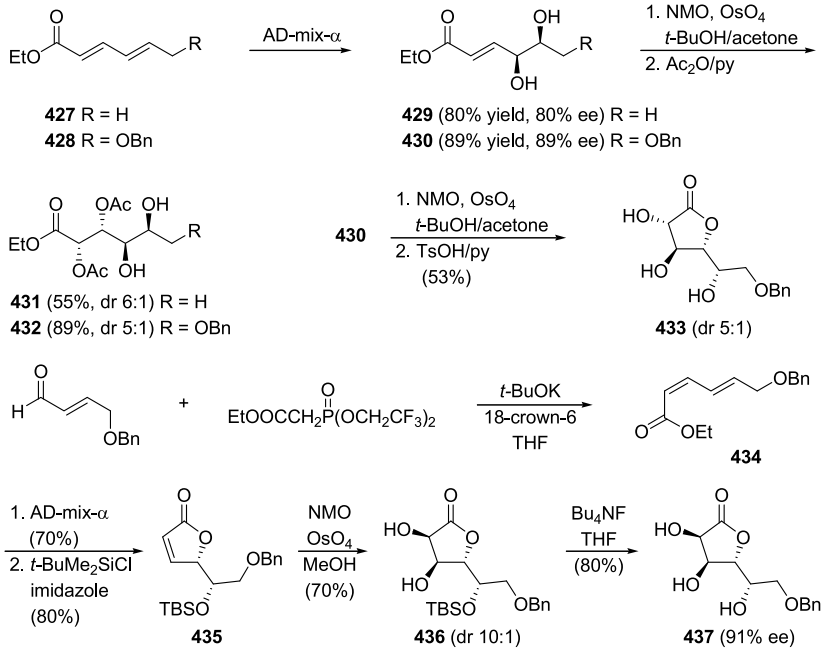
in the presence of 1% $[Pd_2dba_3]\cdot CHCl_3$ (dba = dibenzylideneacetone), 10% DBU, and 3% of diphosphine (*S*,*S*)-ligand, *syn*-**414** is obtained in 76% yield and a diastereomeric ratio (dr) (*syn* vs. *anti*) of 9:1. Ring-closing metathesis of the benzyl ether **415** (3 mol% Grubbs II catalyst) provides **416**. When using the (*R*,*R*)-ligand instead of the (*S*,*S*)-ligand in the second dynamic kinetic asymmetric reaction of (±)-**411** with **413**, *anti*-**414** is obtained in 90% yield (diastereoselectivity *anti* vs. *syn* 10:1). After benzylation and ring-closing methathesis *anti*-**414** produces **417**. Alkaline hydrolysis of carbamate **417** and subsequent amine protection delivers **418**. Its epoxidation required catalytic methyltrioxorhenium and urea·H_2O_2 giving a 76% yield in epoxide **419**. After regioselective epoxide ring hydrolysis and debenzylation iminosugar **13** is obtained (11 steps, 22% overall yield). Regio- and stereoselective hydroboration of **417** is possible using 9-borabicyclo[3.3.1]nonane (9-BBN). This leads to 75% yield of an alcohol that is benzylated. Subsequent carbamate hydrolysis and amine protection as benzylcarbamate provides **420**. Its primary alcohol moiety is oxidized into the corresponding carboxylic acid. Catalytic hydrogenolysis delivers (−)-bulgecinine (➲ *Scheme 103*) [530]. The same chemistry shown in ➲ *Scheme 103* using the (*S*,*S*)-ligand instead of the (*R*,*R*)-ligand in the first dynamic kinetic asymmetric allylation of phthalimide with racemic epoxide (±)-**411** allows us to prepare the enantiomers of **13** and (+)-bulgecinine with the same ease.

11 Enantioselective Sharpless Dihydroxylation

This extremely powerful method has already been presented in ➲ *Scheme 37* and selected examples of application are collected in this section. For instance, tetritol and tetrose derivatives have been prepared through asymmetric dihydroxylation of (*E*)-but-2-ene-1,4-diol [531] and 4-deoxy D- and L-threose are derived from benzene-1,2-dimethyl acetal of (*E*)-crotonaldehyde. Asymmetric dihydroxylation of the dimethyl acetal of 5-[(*tert*-butyldiphenylsilyl)oxy]-(*E*)-pent-3-enal generates a diol that is converted in three steps into 2-deoxyxylofuranosides with high enantiomeric excess [532]. Asymmetric dihydroxylation of 2-vinylfuran gives a diol that has been converted into (+)-isolevoglucosenone. This compound can be isomerized into (−)-levoglucosenone. The L-hexose derivatives (−)-isolevoglucosenone and (+)-levoglucosenone are obtained with the same ease [533]. Applying the same route, D- and L-mannose, have been obtained in five-step synthesis (39% overall yield) from furfural. Similarly, the same methodology has been extended to the preparation of D- and L-gulose and D- and L-talose (19% yield) [534,535,536] and to the synthesis of 2-deoxy and 2,3-dideoxyhexoses [537].
A route to difluorosugar **426** has also been developed. It includes a Stille coupling [538,539] of alkenetin compound **421** with (*E*)-3-iodoprop-2-en-1-yl 4-methoxybenzoate that generates diene **422**. Sharpless asymmetric dihydroxylation of **422** was chemoselective and provided diol **423** in 54% yield. The corresponding acetonide **424** was then saponified with H_2O_2/LiOH. After treatment with 12-M HCl in THF, the semi-protected 1-deoxy-1,1-difluoro-D-xylulose **426** was obtained (➲ *Scheme 104*) [540].
Ethyl-D-galactonates **431** and **432** have been prepared by two successive dihydroxylations of dienoates **427** and **428**, respectively (➲ *Scheme 105*). Intermediate diol **430** has been converted into L-galacto-γ-lactone **433** [541,542]. When using the (2*Z*,4*E*)-dienoate **434**, the same sequence of dihydroxylations provided L-*talo*-γ-lactone derivative **437** [543].



Scheme 104
Synthesis of 1-deoxy-1,1-difluoro-D-xylose



Scheme 105
O'Doherty's iterative dihydroxylation of dienoates

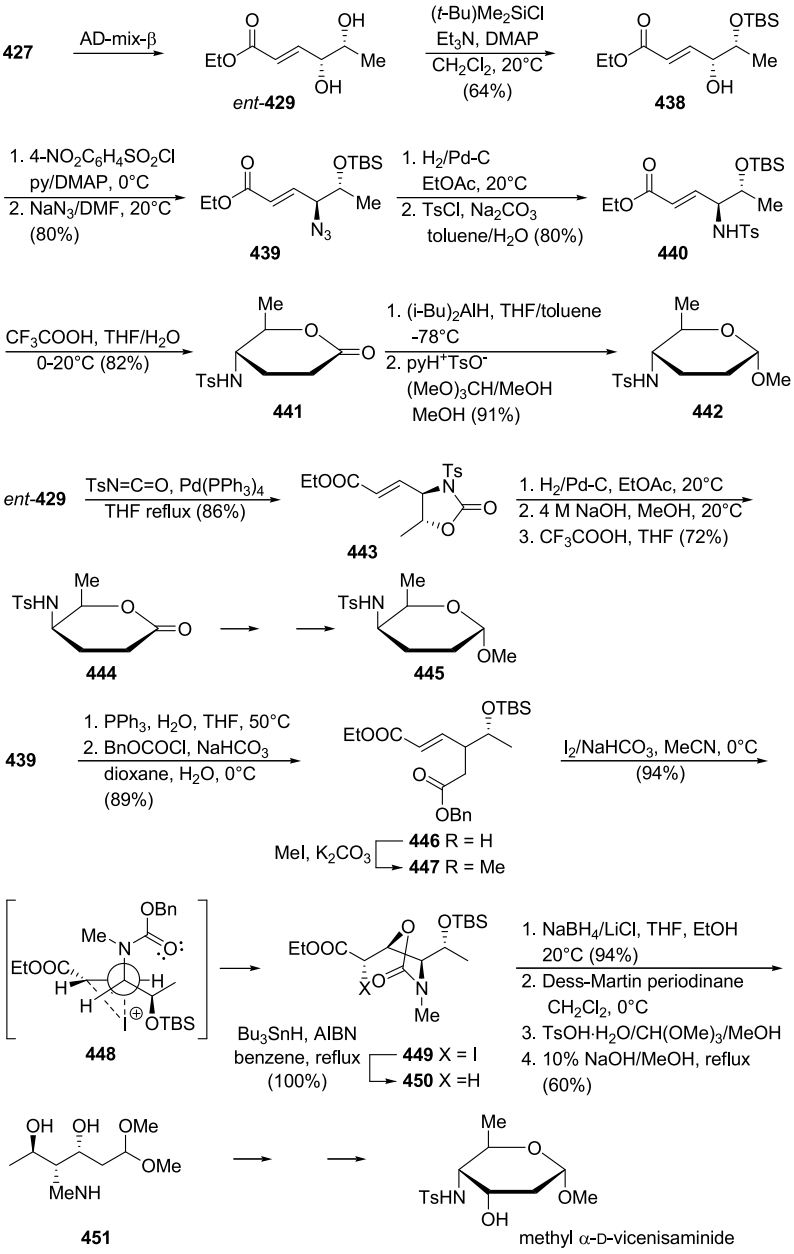
Matsushima and Kino [544] have obtained the enantiomer of **429** applying AD-mix-β-catalyzed dihydroxylation of **427**. Selective monosilylaltion with (*t*-Bu)Me_2SiCl affords **438** (93% ee). After esterification of the allylic alcohol **438** with *para*-nitrobenzenesulfonyl chloride and subsequent sulfonate displacement with NaN_3 (S_N2) azide **439** is isolated in

80% yield. Hydrogenation of azide **439** and subsequent tosylation of the intermediate amine produces sulfonamide **440**. Under acidic conditions **440** is desilylated and converted into δ-lactone **441**. Semi-reduction with $(i\text{-}Bu)_2AlH$ and Fischer glycosidation provides methyl *N*-Ts-α-D-tolyposaminide **442**. Diol of *ent*-**429** reacts with tosylisocyanate in the presence of a catalytic amount of $Pd(PPh_3)_4$ giving an *N*-tosylated isocyanate intermediate. A Pd-catalyzed allylation of the isocyanate generates the cyclic carbamate **443** (S_N with retention). Alkene hydrogenation and subsequent alkaline methanolysis affords δ-lactone **444** which has been converted into methyl 4-*epi*-*N*-Ts-α-D-tolyposaminide **445** as above for the conversion of **441** into **442**. Reduction of azide **439** with Ph_3P in aqueous THF produces the expected primary amine that is then converted into benzylcarbamate **446**. After *N*-methylation into **447**, iodocyclocarbamation provides the *trans*-product **449** in 94% yield (2% of *cis* isomer). The high stereoselectivity of this reaction can be interpreted in terms of transition state **448** which minimizes allylic strain (➲ *Scheme 106*). After reductive deiodination under radical conditions **450** is obtained. Reduction of the ethyl ester moiety of **450** into the corresponding primary alcohol, subsequent oxidation into an aldehyde with the Dess–Martin method, acetalization with methanol and cyclic carbamate hydrolysis produces **451**, a known precursor of methyl α-D-vicenisaminide [545].

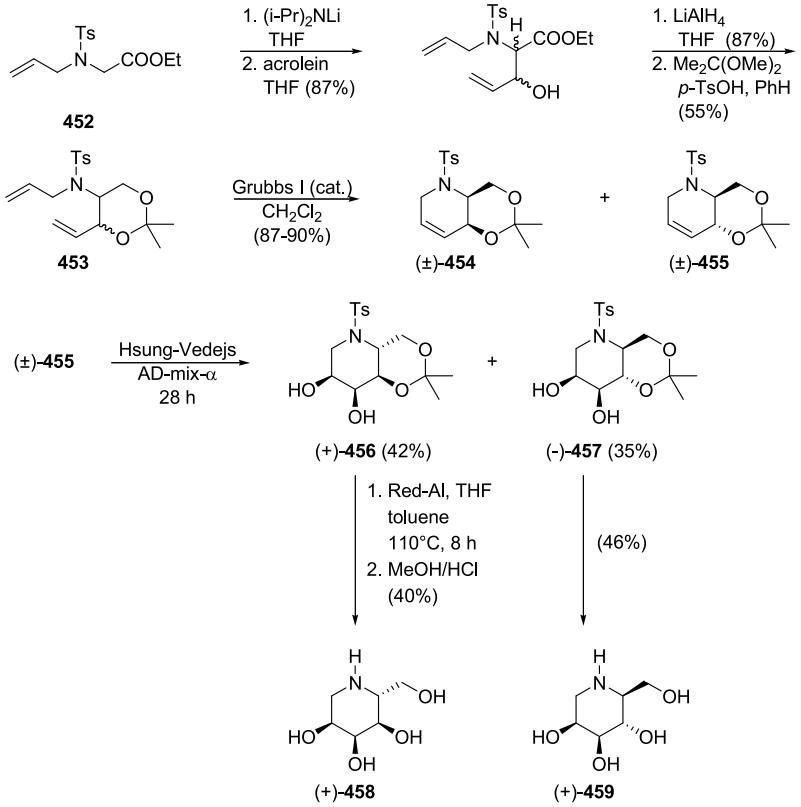
Both *cis*-((±)-**454**) and *trans*-2-substituted-1,2,3,6-tetrahydropyridin-3-ol ((±)-**455**) have been prepared via an aldol condensation of tosylamide **452** with acrolein and subsequent ring-closing metathesis catalyzed by Grubbs I catalyst (➲ *Scheme 107*) [546]. Asymmetric dihydroxylation of *trans*-acetonide (±)-**455** with the Hsung–Vedejs AD-mix-β [547,548] gives (+)-**456** and (−)-**457** in 42 and 35% yield, respectively. The tosyl groups and acetonides are then removed by Red-Al and HCl/MeOH to give 1,5-dideoxy-1,5-imino-D-allitol ((+)-**458**) and 1,5-dideoxy-1,5-imino-L-mannitol ((+)-**459**). Similarly, reaction of *trans*-acetonide **455** with Hsung–Vedejs AD-mix-α affords acetonides (−)-**456** and (+)-**457** in 50 and 40% yield, respectively. The latter can be converted, as before, into 1,5-dideoxy-1,5-imino-L-allitol ((−)-**458**) and 1,5-dideoxy-1,4-imino-D-mannitol ((−)-**459**) [547,548].

Lindström and co-workers [549] have presented an efficient asymmetric synthesis of the iminoalditol **464** (➲ *Scheme 108*). The method requires only four steps in water, without the use of protecting groups. (*E*,*E*)-1,6-Dibromohexa-2,4-diene (**460**) undergoes Sharpless asymmetric dihydroxylation with formation of diol **461** (70% yield, 97% ee). Upon heating in water at 50 °C, the allyl bromide is hydrolyzed chemoselectively giving triol **462**. Epoxidation of **462** with H_2O_2 in the presence of dinuclear peroxotungstate catalyst $K_2[W_2O_3(O_2)_4(H_2O)_2]$ [550] gave **463** in 99% yield and 92% diastereomeric excess. Ammonolysis of bromide **463** in aqueous ammonia was spontaneously followed by an intramolecular ring opening of the epoxide giving **464** (60% overall yield based on **460**).

Photooxidation of (*E*,*E*)-1,4-disubstituted butadienes produces the corresponding 3,6-dihydro-1,2-dioxines **465** that are dihydroxylated into diols **466**. Upon reduction (Zn/AcOH, or H_2/Pd-C, or H_2/Pt-C Mg/I_2 (cat.), or thioureas) racemic alditols **467** are obtained. Isomerization of the dioxines catalyzed by *N*,*N*′-bis(salicylidene)ethylenediaminocobalt (II) gives the corresponding racemic uloses **468** (➲ *Scheme 109*). The method has been used to prepare (±)-psicose, (±)-1-deoxypsicose (±)-6-deoxypsicose [551].

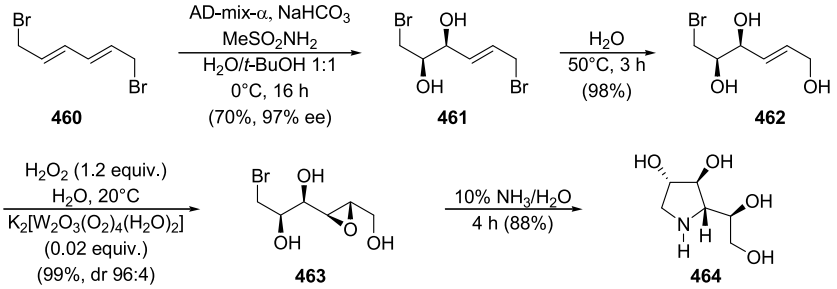


Scheme 106
Matsushima's synthesis of deoxyaminohexoses

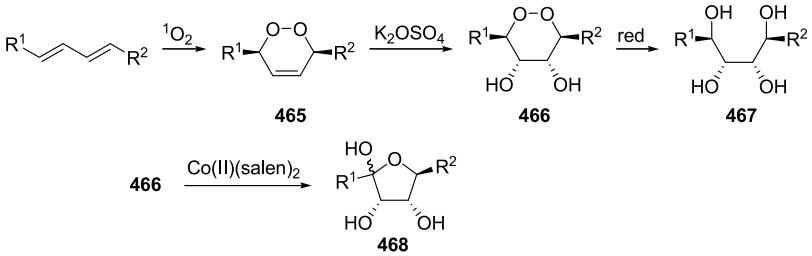


Hsung-Vedejs AD-mix-α: 1:3:3:2:0.1:0.05 alkene/$K_3Fe(CN)_6$/K_2CO_3/$MeSO_2NH_2$/$(DHQ)_2PHAL$/OsO_4
Grubbs I: $(Cy_3P)_2Ru(Cl)_2CHPh$

Scheme 107
Ring-closing metathesis and asymmetric dihydroxlyation: synthesis of iminoalditols



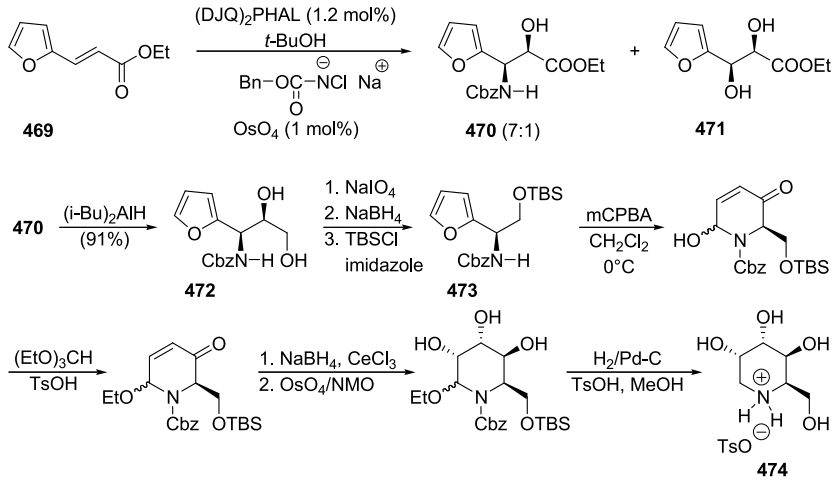
Scheme 108
Efficient asymmetric synthesis of an azasugar in water



Scheme 109
Taylor's synthesis of racemic alditols and uloses via photooxidation of butadiene derivatives

12 Enantioselective Sharpless Aminohydroxylation

The Sharpless asymmetric aminohydroxylation [552] of the electron-deficient 2-vinylfuran **469** gives a 7:1 mixture of semi-protected amino-alcohols **470** and **471** (41%). The major product **472** (ee > 86%) was reduced by diisobutylaluminum hydride giving a diol [553] which is then converted into the β-hydroxyfurylamine derivative **473**, an important synthetic building block for various biologically important compounds, including 1,5-dideoxy-1,5-imino-alditols such as **474** (*Scheme 110*). A less regioselective, but shorter way to intermediate **473** is the direct asymmetric aminohydroxylation of vinylfuran [554,555].



Scheme 110
Application of the Sharpless asymmetric aminohydroxylation and of the aza-Achmatowicz reaction to the synthesis of L-deoxymannonojimycin

13 Conclusion

For many years carbohydrates were very difficult synthetic targets because of their complexity arising from their stereochemistry and their multifunctionality. In parallel with the recent revolution in organic synthesis, a large number of complicated and rare monosaccharides have been prepared by total, asymmetric synthesis. Methods are available that allow one to reach both enantiomers of any natural or non-natural monosaccharides, including deoxyaminosugars, thiosugars, and azasugars, and this, quite often, in a few synthetic steps. Depending on the target, pure chemical procedures relying on asymmetric catalysis using either metallic or pure organic catalysts can be applied successfully, alone or in combination with chemoenzymatic methods.

References

1. Butlerow MA (1861) CR Séances Acad Sci 53:145
2. Butlerow MA (1861) Ann Chem 120:295
3. Jones JKN, Szarek WA (1973) In: ApSimon J (ed) Total Synthesis of Natural Products. Wiley-Interscience, New York, p 1
4. Zamojski A, Grynkiewicz G (1984) In: ApSimon J (ed) Total Synthesis of Natural Products. Wiley-Interscience, New York, p 141
5. McGarvey GJ, Kimura M, Oh T, Williams JM (1984) J Carbohydr Chem 3:125
6. Schmidt RR (1987) Pure Appl Chem 59:415
7. Zamojski A (1997) In: Hanessian S (ed) Prep. Carbohydr. Chem. Dekker, New York, p 615
8. Kirschning A, Jesberger M, Schoning K-U (2001) Synthesis 507
9. Vogel P (2001) In: Fraser-Reid B, Tatsuta K, Thiem J (eds) Encyclopedia of Glycosciences. Springer, Berlin, ch. 4.4, p 1023
10. Vogel P (2006) In: Levy DE, Fügedi P (eds) The Organic Chemistry of Sugars. CRC Taylor & Francis, Boca Raton, Fl, ch. 13, p 629
11. Loew O (1889) Dtsch Chem Ges 22:478
12. Loew O (1889) Dtsch Chem Ges 22:470
13. Fischer E, Passmore F (1890) Dtsch Chem Ges 23:370
14. Fischer E, Passmore F (1889) Dtsch Chem Ges 22:359
15. Mizuno T, Mori L, Shiomi N, Nakatsuji N (1970) Nippon Nogei Kagaku Kaishi, J Arg Chem Soc Jpn 44:324
16. Shigemasa Y, Nagae O, Sakazawa C, Nakashima R, Matsuura TA (1978) J Am Chem Soc 100:1309
17. Decker P (1973) Umschau 73:733
18. Socha RF, Weiss AH, Sakharov MM (1981) J Catal 67:207
19. Weiss AH, Socha RF, Likholobov VA, Sakharov MM (1980) Chemtech 10:643
20. Shigemasa Y, Sasaki T, Ueda N, Nakashima R (1984) Bull Chem Soc Jpn 57:2761
21. Matsumoto T, Yamamoto H, Inoue S (1984) J Am Chem Soc 106:4829
22. Matsumoto T, Inoue T (1983) J Chem Soc, Chem Commun 1983:171
23. Castells J, López-Calahorra F, Geijo F (1983) Carbohydr Res 116:197
24. Castells J, Geijo F, López-Calahorra F (1980) Tetrahedron Lett 21:4517
25. Shigemasa Y, Ueda T, Saimoto H (1989) J Carbohydr Chem 669
26. Tajima H, Niitsu T, Inoue H (1999) J Chem Eng Japan 32:776
27. Tajima H, Niitsu T, Inoue H (2000) J Chem Eng Japan 33:793
28. Tajima H, Inoue H, Ito M (2003) J Comput Chem Jpn 2:127
29. Tajima H, Tabata K, Niitsu T, Inoue H (2002) J Chem Eng Japan 35:564
30. Zhou Z-Z, Ji F-Q, Cao M, Yang G-F (2006) Adv Synth Catal 348:1826
31. Enders D, Kallfass U (2002) Angew Chem Int Ed 41:1743
32. Enders D, Balensiefer T (2004) Acc Chem Res 37:534
33. Stetter H, Kuhlmann H (2000) Angew Chem 30:2281
34. Stetter H, Kuhlmann H (1976) Chem Ber 109:2890

35. Müller D, Pitsch S, Kittaka A, Wagner E, Wintner CE, Eschenmoser A (1990) Helv Chim Acta 73:1410
36. Krishnamurthy R, Guntha S, Eschenmoser A (2000) Angew Chem Int Ed 39:2281
37. Okano T, Ito H, Konishi H, Kiji J (1986) Chem Lett 1731
38. Shevlin PB, Mcpherson DW, Melius P (1983) J Am Chem Soc 105:488
39. Flanagan G, Ahmed SN, Shevlin PB (1992) J Am Chem Soc 114:3892
40. Pestunova O, Simonov A, Snytnikov V, Stoyanovski V, Parmon V (2005) Adv Space Res 36:214
41. Weber AL (1992) J Mol Evol 35:1
42. Schwartz AW, De Graaf RM (1993) J Mol Evol 36:101
43. Pitsch S, Pombo-Villar E, Eschenmoser A (1994) Helv Chim Acta 77:2251
44. Pitsch S, Krishnamurthy R, Arrhenius G (2000) Helv Chim Acta 83:2398
45. Weber AL (2000) Origins Life Evol Biosphere 30:33
46. Ricardo A, Carrigan MA, Olcott AN, Benner SA (2004) Science 303:196
47. David S, Auge C, Gautheron C (1992) Adv Carbohydr Chem Biochem 49:175
48. Kajimoto T, Sugai T, Wong C-H (1993) Trends in Glycoscience and Glycotechnology 1993:193
49. Wong C-H, Halcomb RL, Ichikawa Y, Kajimoto T (1995) Angew Chem Int Ed Engl 34:412
50. Gijsen HJ, Qiao L, Fitz W, Wong C-H (1996) Chem Rev 96:443
51. Takayama S, McGarvey GJ, Wong C-H (1997) Chem Soc Rev 26:407
52. Machajewski TD, Wong C-H, Lerner RA (2000) Angew Chem Int Ed 39:1352
53. Silvestri MG, Desantis G, Mitchell M, Wong C-H (2003) Topics in Stereochemistry 23:267
54. Wymer N, Toone EJ (2000) Curr Opin Chem Biol 4:110
55. Bednarski MD, Waldmann HJ, Whitesides GM (1986) Tetrahedron Lett 27:5807
56. Durrwachter JR, Drueckhammer DG, Nizaki K, Sweers HM, Wong C-H (1986) J Am Chem Soc 108:7812
57. Durrwachter JR, Wong C-H (1988) J Org Chem 53:4175
58. Schultz M, Waldmann H, Kunz H, Vogt W (1990) Liebigs Ann Chem 1019
59. Lees WJ, Whitesides GM (1993) J Org Chem 58:1887
60. Henderson I, Laslo K, Wong C-H (1994) Tetrahedron Lett 35:359
61. Gijsen HJM, Wong C-H (1995) J Am Chem Soc 117:2947
62. Duncan R, Drueckhammer DG (1996) J Org Chem 61:438
63. Chenvert R, Lavoie M, Dasser M (1997) Can J Chem 75:68
64. Guanti G, Banfi L, Zannetti MT (2000) Tetrahedron Lett 41:3181
65. Schoevaaart R, Van Rantwijk F, Sheldon RA (2000) J Org Chem 65:6940
66. Fong S, Machajewski TD, Mak CC, Wong C-H (2000) Chem Biol 7:873
67. Guanti G, Zannetti MT, Banfi L, Riva R (2001) Adv Synth Catal 343:682
68. Hsu C-C, Hong Z, Wada M, Franke D, Wong C-H (2005) PNAS 102:9122
69. Navor Le Gautheron C, Ichikawa Y, Wong C-H (1991) J Am Chem Soc 113:7816
70. Lin CH, Sugai T, Halcomb RL, Ichikawa Y, Wong C-H (1992) J Am Chem Soc 114:10138
71. Von Itzstein M, Wu WY, Kok GB, Pegg MS, Dyason JC, Jin B, Van Phan T, Smythe ML, White HF, Oliver SW (1993) Nature 363:418
72. Wada M, Hsu C-C, Franke D, Mitchell M, Heine A, Wilson I, Wong C-H (2003) Bioorg Med Chem Lett 11:2091
73. Witzemann EJ (1914) J Am Chem Soc 36:2223
74. Morozov AA (1992) React Kinet Catal Lett 46:71
75. Morozov AA (1992) Chem Abstr 116:174545
76. Saimoto H, Kotani K, Shigemasa Y, Suzuki K, Harada KI (1989) Tetrahedron Lett 30:2553
77. Yamashita M, Wakao N, Nango M, Tsuda K (1992) J Polym Science(A):Polym Chem 30:2247
78. Yamashita M, Wakao N, Nango M, Tsuda K (1992) Chem Abstr 117:151580z
79. Ando T, Shioi S, Nakagawa M (1972) Bull Chem Soc Jpn 45:2611
80. Sonogashira K, Nakagawa M (1972) Bull Chem Soc Jpn 45:2616
81. Eyrisch O, Sinerius G, Fessner W-D (1993) Carbohydr Res 238:287
82. Franke D, Machajewski TD, Hsu C-C, Wong C-H (2003) J Org Chem 68:6828
83. Henderson I, Sharpless KB, Wong C-H (1994) J Am Chem Soc 116:558
84. Alajarin R, García-Junceda E, Wong C-H (1995) J Org Chem 60:4294
85. Ziegler T, Straub A, Effenberger F (1988) Angew Chem Int Ed Engl 27:716

86. Straub A, Effenberger F, Fisher P (1990) J Org Chem 55:3926
87. Effenberger F, Null V (1992) Liebigs Ann Chem 1211
88. Zhou P, Salleh H, Chan PCM, Lajoie G, Honek JF, Nambiar FP, Ward OP (1993) Carbohydr Res 239:155
89. Lemaire M, Valentin ML, Hecquet L, Demuynck C, Bulte J (1995) Tetrahedron:Asymmetry 6:67
90. Von der Osten CH, Sinskey AJ, Barbas CFI, Pederson RL, Wang Y-F, Wong C-H (1989) J Am Chem Soc 111:3924
91. Kajimoto T, Liu KKC, Pederson RL, Zhong Z, Ichikawa Y, Porco JA Jr, Wong C-H (1991) J Am Chem Soc 113:6187
92. Hung RR, Straub JA, Whitesides GM (1991) J Org Chem 56:3849
93. Liu KKC, Kajimoto T, Chen L, Zhong Z, Ichikawa Y, Wong C-H (1991) J Org Chem 56:6280
94. Kajimoto T, Chen D, Liu KKC, Wong C-H (1991) J Am Chem Soc 113:6678
95. Takaoka Y, Kajimoto T (1993) J Org Chem 58:4809
96. Miura T, Kajimoto T (2001) Chirality 13:577
97. Chou WC, Chen L, Fang JM, Wong C-H (1994) J Am Chem Soc 116:6191
98. Wong C-H, García-Junceda E, Chen L, Blanco O, Gijsen HJM, Steensma DH (1995) J Am Chem Soc 117:3333
99. Wagner J, Lerner RA, Barbas CFI (1995) Science 270:1797
100. Zhong G, Shabat D, List B, Anderson J, Sinha SC, Lerner RA, Barbas CFI (1998) Angew Chem Int Ed Engl 37:2481
101. Hoffmann T, Zhong GF, List B, Shabat D, Anderson J, Gramatikova S, Lerner RA, Barbas CF (1998) J Am Chem Soc 120:2768
102. Shabat D, List B, Lerner RA, Barbas CFI (1999) Tetrahedron Lett 40:1437
103. Hill RE, Sayer BG, Spenser ID (1989) J Am Chem Soc 111:1916
104. Sagner S, Eisenreich W, Fellermeier M, Latzel C, Bacher A, Zenk MH (1998) Tetrahedron Lett 39:2091
105. Piel J, Dontah J, Bandemer DK, Boland W (1998) Angew Chem Int Ed Engl 37:2478
106. Eder U, Sauer G, Weichert R (1971) Angew Chem Int Ed Engl 10:496
107. Hajos ZG, Parrish DR (1974) J Org Chem 39:1612
108. Cohen N (1976) Acc Chem Res 9:412
109. List B (2002) Tetrahedron 58:5573
110. Benaglia M, Puglisi A, Cozzi F (2003) Chem Rev 103:3401
111. Cozzi F (2006) Adv Synth Catal 348:1367
112. Yamada YMA, Yoshikawa N, Sasai H, Shibasaki M (1997) Angew Chem Int Ed Engl 36:1871
113. Nagakawa M, Nakao H, Watanabe KI (1985) J Am Chem Soc 122:12003
114. Trost BM, Ito H (2000) J Am Chem Soc 122:12003
115. List B, Lerner RA, Barbas CF (2000) J Am Chem Soc 122:2395
116. List B (2001) Synlett 1675
117. List B, Pojarliev P, Castello C (2001) Org Lett 3:573
118. Hoang L, Bahmanyar S, Houk KN, List B (2003) J Am Chem Soc 125:16
119. Allemann C, Gordillo R, Clemente FR, Cheong PHY, Houk KN (2004) Acc Chem Res 37:558
120. Bassan A, Zou WB, Reyes E, Himo F, Cordova A (2005) Angew Chem Int Ed 44:7028
121. Cordova A, Zou WB, Ibrahem I, Reyes E, Engqvist M, Liao WW (2005) Chem Commun 3586
122. Kano T, Takai J, Tokuda O, Maruoka K (2005) Angew Chem Int Ed 44:3055
123. Jarvo ER, Miller SJ (2002) Tetrahedron 58:2481
124. Dziedzic P, Zou WB, Hafren J, Cordova A (2006) Org Biomol Chem 4:38
125. Kofoed J, Nielsen J, Reymond JL (2003) Bioorg Med Chem Lett 13:2445
126. Ibrahem I, Zou WB, Xu YM, Cordova A (2006) Adv Synth Catal 348:211
127. Andreae MRM, Davis AP (2005) Tetrahedron:Asymmetry 16:2487
128. Sakthivel K, Notz W, Bui T, Barbas CF (2001) J Am Chem Soc 123:5260
129. Northrup AB, Mangion IK, Hettche F, MacMillan DWC (2004) Angew Chem Int Ed 43:2152
130. Davies SG, Sheppard RL, Smith AD, Thomson JE (2005) Chem Commun 3802
131. Cobb AJA, Shaw DM, Longbottom DA, Gold JB, Ley SV (2005) Org Biomol Chem 3:84
132. Hartikka A, Arvidsson PI (2004) Tetrahedron:Asymmetry 15:1831
133. Bellis E, Kokotos G (2005) Tetrahedron 61:8669
134. Nakadai M, Saito S, Yamamoto H (2002) Tetrahedron 58:8167
135. Lacoste E, Landais Y, Schenk K, Verlhac JB, Vincent JM (2004) Tetrahedron Lett 45:8035
136. Kazmaer U (2005) Angew Chem Int Ed 44:2186
137. Notz W, List B (2000) J Am Chem Soc 122:7386
138. Kolb HC, Van Nieuwenhze MS, Sharpless KB (1994) Chem Rev 94:2483

139. Cordova A, Notz W, Barbas CF (2002) Chem Commun 3024
140. Suri JT, Ramachary DB, Barbas CF (2005) Org Lett 7:1383
141. Enders D, Grondal C (2005) Angew Chem Int Ed 44:1210
142. Grondal C, Enders D (2005) Tetrahedron 62:329
143. Enders D, Grondal C (2005) Lett Org Chem 2:577
144. Ibrahem I, Cordova A (2005) Tetrahedron Lett 46:3363
145. Zou WB, Ibrahem I, Dziedzic P, Sunden H, Cordova A (2005) Chem Commun 4946
146. Northrup AB, MacMillan DWC (2002) J Am Chem Soc 124:6798
147. Northrup AB, Mangion IK, Hettche F, MacMillan DWC (2004) Angew Chem Int Ed 43:2152
148. Northrup AB, MacMillan DWC (2004) Science 305:1752
149. Mangion IK, MacMillan DWC (2005) J Am Chem Soc 127:3696
150. Storer RI, MacMillan DWC (2004) Tetrahedron 60:7705
151. Cordova A, Ibrahem I, Casas J, Sunden H, Engqvist M, Reyes E (2005) Chem Eur J 11:4772
152. Casas J, Engqvist M, Ibrahem I, Kaynak B, Cordova A (2005) Angew Chem Int Ed 44:1343
153. Reyes E, Cordova A (2005) Tetrahedron Lett 46:6605
154. Kofoed J, Reymond JL, Darbre T (2005) Org Biomol Chem 3:1850
155. Kofoed J, Machuqueiro M, Reymond JL, Darbre T (2004) Chem Commun 1540
156. Thayumanavan R, Tanaka F, Barbas CF (2004) Org Lett 6:3541
157. Cordova A, Notz W, Zhong GF, Betancort JM, Barbas CF (2002) J Am Chem Soc 124:1842
158. Cordova A, Watanabe S, Tanaka F, Notz W, Barbas CF (2002) J Am Chem Soc 124:1866
159. Notz W, Sakthivel K, Bui T, Zhong GF, Barbas CF (2001) Tetrahedron Lett 42:199
160. List B (2000) J Am Chem Soc 122:9336
161. List B, Pojarliev P, Biller WT, Martin HJ (2002) J Am Chem Soc 124:827
162. Enders D, Grondal C, Vrettou M, Raabe G (2005) Angew Chem Int Ed 44:4079
163. Ibrahem I, Zou WB, Casas J, Sunden H, Cordova A (2006) Tetrahedron 62:357
164. Hayashi Y, Tsuboi W, Ashimine I, Urushima T, Shoji M, Sakai K (2003) Angew Chem Int Ed 42:3677
165. Mukaiyama T, Shiina I, Kobayashi S (1990) Chem Lett 2201
166. Kobayashi S, Kawasuji T (1993) Synlett 911
167. Kiliani H (1885) Ber Deutsch Chem 18:3066
168. Fischer E (1889) Ber Deutsch Chem 22:2204
169. Lichtenthaler FW (1992) Angew Chem Int Ed Engl 31:1541
170. Soengas R, Izumori K, Simone MI, Watkin DJ, Skytte UP, Soetaert W, Fleet GWJ (2005) Tetrahedron Lett 46:5755
171. Hotchkiss D, Soengas R, Simone MI, van Ameijde J, Hunter S, Cowley AR, Fleet GWJ (2004) Tetrahedron Lett 45:9461
172. Roush WR, Hoong LK, Palmer MAJ, Straub JA, Palkowitz AD (1990) J Org Chem 55:4117
173. Roush WR, Ando K, Powers DB, Palkowitz AD, Halterman RL (1990) J Am Chem Soc 112:6339
174. Roush WR, Straub JA, Vannieuwenhze MS (1991) J Org Chem 56:1636
175. Roush WR, Lin XF, Straub JA (1991) J Org Chem 56:1649
176. Maruyama K, Ishihara Y, Yamamoto Y (1981) Tetrahedron Lett 22:4235
177. Marshall JA, Luke GP (1991) J Org Chem 56:483
178. Marshall JA (1996) Chem Rev 96:31
179. Jurczak J, Pikul S, Bauer T (1986) Tetrahedron 42:447
180. Perlin AS (1962) Methods in Carbohydrate Chemistry. Academic Press, New York, p 61
181. Hubschwerlen C (1986) Synthesis 962
182. Jager V, Wehner V (1989) Angew Chem Int Ed Engl 28:469
183. Bewsey JA (1977) Chem Ind 119
184. Seebach D, Hungerbuehler E (1980) In: Scheffold R (ed) Modern Synthetic Methods. Otto Salle Verlag, Verlag Sauerländer, Frankfurt am Main, p 91
185. Matteson DS, Peterson ML (1987) J Org Chem 52:5116
186. Bischofberger N, Waldmann H, Saito T, Simon ES, Lees W, Bednarski MD, Whitesides GM (1988) J Org Chem 53:3457
187. Ballou CE, Fischer HOL (1955) J Am Chem Soc 77:3329
188. Kerscher V, Kreiser W (1987) Tetrahedron Lett 28:531
189. Lewis CA, Sarlimbrene BR, Xu YJ, Miller SJ (2005) Org Lett 7:3021
190. Dewitt P, Misiti D, Zappia G (1989) Tetrahedron Lett 30:5505
191. Oi R, Sharpless KB (1992) Tetrahedron Lett 33:2095

192. Wuts PGM, Bigelow SS (1983) J Org Chem 48:3489
193. Lubell WD, Rapoport H (1987) J Am Chem Soc 109:236
194. Fehrentz JA, Castro B (1983) Synthesis 676
195. Nicolaou KC, Groneberg RD, Stylianides NA, Miyazaki T (1990) J Chem Soc, Chem Commun 1275
196. Sibi MP, Renhowe PA (1990) Tetrahedron Lett 31:7407
197. Dondoni A, Marra A (1997) In: Hanessian S (ed) Preparative Carbohydrate Chemistry. Marcel Dekker, New York, p 173
198. Dondoni A (1998) Synthesis 1681
199. Dondoni A, Marra A (1999) Chem Commun 2133
200. Dondoni A, Perrone D (1999) Tetrahedron Lett 40:9375
201. Dondoni A, Marra A (2000) Chem Rev 100:4395
202. Dondoni A (2000) Pure Appl Chem 72:1577
203. Dondoni A, Formaglio P, Marra A, Massi A (2001) Tetrahedron 57:7719
204. Dondoni A, Giovannini PP, Perrone D (2002) J Org Chem 67:7203
205. Dondoni A, Giovannini PP (2002) Synthesis 1701
206. Dondoni A, Marra A (2004) Chem Rev 104:2557
207. Dondoni A, Merino P (1995) Org Synth 21
208. Dondoni A, Marra A, Perrone D (1993) J Org Chem 58:275
209. Dondoni A, Fantin G, Fogagnolo M, Pedrini P (1990) J Org Chem 55:1439
210. Dondoni A, Perrone D, Merino P (1995) J Org Chem 60:8074
211. Dondoni A, Franco S, Junquera F, Merchan FL, Merino P, Tejero T, Bertolasi V (1995) Chem Eur J 1:505
212. Dondoni A, Franco S, Merchan FL, Merino P, Tejero T (1993) Tetrahedron Lett 34:5475
213. Merino P, Franco S, Merchan FL, Revuelta J, Tejero T (2002) Tetrahedron Lett 43:459
214. Kusakabe M, Sato F (1986) Chem Lett 1473
215. Benzingnguyen L, Perry MB (1978) J Org Chem 43:551
216. Liu FW, Yan L, Zhang JY, Liu HM (2006) Carbohydr Res 341:332
217. Chikashita H, Nikaya T, Itoh K (1993) Nat Prod Lett 2:183
218. Seebach D (1969) Synthesis 5:17
219. Page PCB, Vanniel MB, Prodger JC (1989) Tetrahedron 45:7643
220. Takahashi S, Kuzuhara H (1997) J Chem Soc, Perkin Trans 1: Org Bio-Org Chem 607
221. Bessieres B, Morin C (2003) J Org Chem 68:4100
222. Marco-Contelles J, de Opazo E, Arroyo N (2001) Tetrahedron 57:4729
223. Mlynarski J, Banaszek A (2003) Trends in Org Chem 10:51
224. Sisu E, Sollogoub M, Mallet JM, Sinay P (2002) Tetrahedron 58:10189
225. Yun M, Moon HR, Kim HO, Choi WJ, Kim YC, Park CS, Jeong LS (2005) Tetrahedron Lett 46:5903
226. Solladie G, Demailly G, Greck C (1985) J Org Chem 50:1552
227. Solladie G, Frechou C, Hutt J, Demailly G (1987) Bull Soc Chim Fr 827
228. Noe CR, Knollmuller M, Ettmayer P (1989) Liebigs Ann Chem 637
229. Wulff G, Hansen A (1986) Angew Chem Int Ed Engl 25:560
230. Mukaiyama T (1989) In: Horton D, Hawkins LD, McGarvey GJ (eds) Trends in Synthetic Carbohydrate Chemistry, ACS Symposium Series 386. American Chemical Society, Washington, DC, ch. 15, p 278
231. Mukaiyama T, Miwa T, Nakatsuka T (1982) Chem Lett 145
232. Isoda T, Akiyama R, Oyamada H, Kobayashi S (2006) Adv Synth Catal 348:1813
233. McGilvra JD, Unni AK, Modi K, Rawal VH (2006) Angew Chem Int Ed 45:6130
234. Desimoni G, Faita G, Jorgensen KA (2006) Chem Rev 106:3561
235. Jankowska J, Mlynarski J (2006) J Org Chem 71:1317
236. Itsuno S, Arima S, Haraguchi N (2005) Tetrahedron 61:12074
237. Yamamoto Y, Kirihata M, Ichimoto I, Ueda H (1985) Agric Biol Chem 49:1435
238. Wulff G, Hansen A (1987) Carbohydr Res 164:123
239. Banfi L, Cardani S, Potenza D, Scolastico C (1987) Tetrahedron 43:2317
240. Ulgheri F, Bacsa J, Nassimbeni L, Spanu P (2003) Tetrahedron Lett 44:671
241. Timmer MSM, Adibekian A, Seeberger PH (2005) Angew Chem Int Ed 44:7605
242. Guaragna A, Napolitana C, D'Alonzo D, Pedatella S, Palumbo G (2006) Org Lett 8:4863
243. Caputo R, De Nisco M, Festa P, Guaragna A, Palumbo G, Pedatella S (2004) J Org Chem 69:7033

244. Wehner V, Jager V (1990) Angew Chem Int Ed Engl 29:1169
245. Devant R, Mahler U, Braun M (1988) Chem Ber Rec 121:397
246. Graf S, Braun M (1993) Liebigs Ann Chem 1091
247. Braun M, Moritz J (1991) Synlett 750
248. Schollkopf U, Tiller T, Bardenhagen J (1988) Tetrahedron 44:5293
249. Rudge AJ, Collins I, Holmes AB, Baker R (1994) Angew Chem Int Ed Engl 33:2320
250. Evans DA, Gage JR, Leighton JL (1992) J Am Chem Soc 114:9434
251. FranckNeumann M, Bissinger P, Geoffroy P (1997) Tetrahedron Lett 38:4477
252. Enders D, Jegelka U (1993) Tetrahedron Lett 34:2453
253. Job A, Janeck CF, Bettray W, Peters R, Enders D (2002) Tetrahedron 58:2253
254. Timmer MSM, Stocker BL, Seeberger PH (2006) J Org Chem 71:8294
255. Katsuki T, Sharpless KB (1980) J Am Chem Soc 102:5974
256. Jorgensen KA (1989) Chem Rev 89:431
257. Sharpless KB (1986) Chem Ber 22:38
258. Katsuki T, Lee AWM, Ma P, Martin VS, Masamune S, Sharpless KB, Tuddenham D, Wlaker FJ (1982) J Org Chem 47:1373
259. Ma P, Martin VS, Masamune S, Sharpless KB, Viti SM (1982) J Org Chem 47:1378
260. Ko SY, Lee AWM, Masamune S, Reed LA, Sharpless KB, Walker FJ (1990) Tetrahedron 46:245
261. Aoyagi S, Fujimaki S, Yamazaki N, Kibayashi C (1991) J Org Chem 56:815
262. Iida H, Yamazaki N, Kibayashi C (1987) J Org Chem 52:3337
263. Jung ME, Gardiner JM (1992) Tetrahedron Lett 33:3841
264. Jung ME, Gardiner JM (1991) J Org Chem 56:2614
265. Matsushima Y, Nakayama T, Tohyama S, Eguchi T, Kakinuma K (2001) J Chem Soc, Perkin Trans 1569
266. Martin R, Moyano A, Pericas MA, Riera A (2000) Org Lett 2:93
267. Diaz Y, Bravo F, Castillon S (1999) J Org Chem 64:6508
268. Sharpless KB, Amberg W, Bennani YL, Crispino GA, Hartung J, Jeong KS, Kwong HL, Morikawa K, Wang ZM, Xu DQ, Zhang XL (1992) J Org Chem 57:2768
269. Becker H, Sharpless KB (1996) Angew Chem Int Ed Engl 35:448
270. Jorgensen M, Iversen EH, Madsen R (2001) J Org Chem 66:4625
271. Li GG, Chang HT, Sharpless KB (1996) Angew Chem Int Ed Engl 35:451
272. Ikemoto N, Schreiber SL (1992) J Am Chem Soc 114:2524
273. Poss CS, Schreiber SL (1994) Acc Chem Res 27:9
274. Furstner A, Wuchrer M (2005) Chem Eur J 12:76
275. Cha JK, Christ WJ, Kishi Y (1984) Tetrahedron 40:2247
276. Kahne D, Walker S, Cheng Y, Vanengen D (1989) J Am Chem Soc 111:6881
277. Evans DA, Bartroli J, Shih TL (1981) J Am Chem Soc 103:2127
278. Gage JR, Evans DA (1990) Org Synth 68:83
279. Davies SG, Sanganee HJ (1995) Tetrahedron: Asymmetry 6:671
280. Bull SD, Davies SG, Jones S, Polywka MEC, Prasad RS, Sanganee HJ (1998) Synlett 519
281. Gibson CL, Gillon K, Cook S (1998) Tetrahedron Lett 39:6733
282. Zhang YC, Phillips AJ, Sammakia T (2004) Org Lett 6:23
283. Evans DA, Ng HP, Clark JS, Rieger DL (1992) Tetrahedron 48:2127
284. Nerzstormes M, Thornton ER (1991) J Org Chem 56:2489
285. Phoon CW, Abell C (1998) Tetrahedron Lett 39:2655
286. Crimmins MT, Long A (2005) Org Lett 7:4157
287. Davies SG, Nicholson RL, Smith AD (2002) Synlett 1637
288. Fronza G, Fuganti C, Grasselli P (1980) J Chem Soc, Chem Commun 442
289. Fuganti C, Grasselli P, Pedrocchifantoni G (1981) Tetrahedron Lett 22:4017
290. Fuganti C, Grasselli P (1978) J Chem Soc, Chem Commun 299
291. Fronza G, Fuganti C, Grasselli P, Marinoni G (1979) Tetrahedron Lett 3883
292. Dondoni A, Fantin G, Fogagnolo M, Merino P (1990) Tetrahedron 46:6167
293. Roush WR, Walts AE, Hoong LK (1985) J Am Chem Soc 107:8186
294. Roush WR, Halterman RL (1986) J Am Chem Soc 108:294
295. Roush WR, Straub JA (1986) Tetrahedron Lett 27:3349
296. Roush WR, Hunt JA (1995) J Org Chem 60:798

297. Fronza G, Fuganti C, Grasselli P, Pedrocchi-fantoni G, Zirotti C (1982) Tetrahedron Lett 23:4143
298. Fronza G, Fuganti C, Grasselli P, Pedrocchifantoni G, Zirotti C (1984) Chem Lett 335
299. Williams DR, Klingler FD (1987) Tetrahedron Lett 28:869
300. Roush WR, Michaelides MR (1986) Tetrahedron Lett 27:3353
301. Vuljanic T, Kihlberg J, Somfai P (1994) Tetrahedron Lett 35:6937
302. Roush WR, Adam MA, Walts AE, Harris DJ (1986) J Am Chem Soc 108:3422
303. Gryko D, Jurczak J (1997) Tetrahedron Lett 38:8275
304. Dondoni A, Merino P, Perrone D (1993) Tetrahedron 49:2939
305. Dondoni A, Perrone D (1995) J Org Chem 60:4749
306. Dondoni A, Marra A, Boscarato A (1999) Chem Eur J 5:3562
307. Narasaka K, Pai FC (1984) Tetrahedron 40:2233
308. Evans DA, Cee VJ, Siska SJ (2006) J Am Chem Soc 128:9433
309. Evans DA, Glorius F, Burch JD (2005) Org Lett 7:3331
310. Sames D, Polt R (1994) J Org Chem 59:4596
311. Majewski M, Nowak P (2000) J Org Chem 65:5152
312. Stepowska H, Zamojski A (2001) Carbohydr Res 332:429
313. Li LS, Wu YL (2002) Tetrahedron 58:9049
314. Palmelund A, Madsen R (2005) J Org Chem 70:8248
315. Francisco CG, Gonzalez CC, Paz NR, Suarez E (2003) Org Lett 5:4171
316. Marshall JA, Seletsky BM, Coan PS (1994) J Org Chem 59:5139
317. Marshall JA, Seletsky BM, Luke GP (1994) J Org Chem 59:3413
318. Marshall JA (1996) Chem Rev 96:31
319. Marshall JA, Beaudoin S (1996) J Org Chem 61:581
320. Hashimoto H, Asano K, Fujii F, Yoshimura J (1982) Carbohydr Res 104:87
321. Dondoni A, Franco S, Merchan F, Merino P, Tejero T (1993) Synlett 78
322. Martin SF, Zinke PW (1989) J Am Chem Soc 111:2311
323. Casiraghi G, Zanardi F, Appendino G, Rassu G (2000) Chem Rev 100:1929
324. Rassu G, Pinna L, Spanu P, Culeddu N, Casiraghi G, Fava GG, Ferrari MB, Pelosi G (1992) Tetrahedron 48:727
325. Rassu G, Casiraghi G, Spanu P, Pinna L, Fava GG, Ferrari MB, Pelosi G (1992) Tetrahedron:Asymmetry 3:1035
326. Casiraghi G, Rassu G, Spanu P, Pinna L (1992) J Org Chem 57:3760
327. Casiraghi G, Rassu G, Spanu P, Pinna L (1994) Tetrahedron Lett 35:2423
328. Casiraghi G, Spanu P, Rassu G, Pinna L, Ulgheri F (1994) J Org Chem 59:2906
329. Zanardi F, Battistini L, Nespi M, Rassu G, Spanu P, Cornia M, Casiraghi G (1996) Tetrahedron:Asymmetry 7:1167
330. Rassu G, Zanardi F, Battistini L, Casiraghi G (2000) Chem Soc Rev 29:109
331. Depezay JC, Dureault A, Prange T (1984) Tetrahedron Lett 25:1459
332. Heathcock CH, Montgomery SH (1985) Tetrahedron Lett 26:1001
333. Hoppe D, Tarara G, Wilckens M, Jones PG, Schmidt D, Stezowski JJ (1987) Angew Chem Int Ed Engl 26:1034
334. Hatakeyama S, Sugawara K, Takano S (1991) Tetrahedron Lett 32:4513
335. Casiraghi G, Pinna L, Rassu G, Spanu P, Ulgheri F (1993) Tetrahedron:Asymmetry 4: 681
336. Bednarski M, Danishefsky S (1986) J Am Chem Soc 108:7060
337. Danishefsky S, Bednarski M, Izawa T, Maring C (1984) J Org Chem 49:2290
338. Danishefsky S, Bednarski M (1985) Tetrahedron Lett 26:3411
339. Hu YJ, Huang XD, Yao ZJ, Wu YL (1998) J Org Chem 63:2456
340. Lubineau A, Auge J, Lubin N (1993) Tetrahedron 49:4639
341. Cousins RPC, Pritchard RG, Raynor CM, Smith M, Stoodley RJ (2002) Tetrahedron Lett 43:489
342. Bataille C, Begin G, Guillam A, Lemiegre L, Lys C, Maddaluno J, Toupet L (2002) J Org Chem 67:8054
343. Osborn HMI, Coisson D (2004) Mini-Rev in Org Chem 1:41
344. Winkler JD, Oh K (2005) Org Lett 7:2421
345. Unni AK, Takenaka N, Yamamoto H, Rawal VH (2005) J Am Chem Soc 127:1336
346. Anada M, Washio T, Shimada N, Kitagaki S, Nakajima M, Shiro M, Hashimoto S (2004) Angew Chem Int Ed 43:2665

347. Furuno H, Hayano T, Kambara T, Sugimoto Y, Hanamoto T, Tanaka Y, Jin YZ, Kagawa T, Inanaga J (2003) Tetrahedron 59:10509
348. Danishefsky SJ, Deninno MP, Chen S (1988) J Am Chem Soc 110:3929
349. Isono K, Crain PF, Mccloskey JA (1975) J Am Chem Soc 97:943
350. Danishefsky SJ, Hungate R, Schulte G (1988) J Am Chem Soc 110:7434
351. Golebiowski A, Kozak J, Jurczak J (1991) J Org Chem 56:7344
352. Desimoni G, Tacconi G (1975) Chem Rev 75:651
353. Boger DL (1987), Hetero-Diels–Alder Methodology in Organic Synthesis edn. Academic Press, New York
354. Snider BB (1980) Acc Chem Res 13:426
355. Tietze LF, Beifuss U (1993) Angew Chem Int Ed Engl 32:131
356. Tietze LF, Kettschau G, Gewert JA, Schuffenhauer A (1998) Curr Org Chem 2:19
357. Tietze LF, Schneider C, Montenbruck A (1994) Angew Chem Int Ed Engl 33:980
358. de Meijere A, Leonov A, Heiner T, Noltemeyer M, Bes MT (2003) Eur J Org Chem 472
359. Ismail ZM, Hoffmann HMR (1982) Angew Chem Int Ed Engl 21:859
360. Elabed D, Jellal A, Santelli M (1984) Tetrahedron Lett 25:4503
361. Tietze LF, Voss E (1986) Tetrahedron Lett 27:6181
362. Apparao S, Maier ME, Schmidt RR (1987) Synthesis 900
363. Schmidt RR (1987) Pure Appl Chem 59:415
364. Tietze LF, Voss E, Harms K, Sheldrick GM (1985) Tetrahedron Lett 26:5273
365. Tietze LF, Hartfiel U (1990) Tetrahedron Lett 31:1697
366. Hayman CM, Larsen DS, Brooker S (1998) Austr J Chem 51:545
367. Boger DL, Robarge KD (1988) J Org Chem 53:5793
368. Tietze LF, Montenbruck A, Schneider C (1994) Synlett 509
369. Tietze LF, Schneider C, Grote A (1996) Chem Eur J 2:139
370. Schmidt RR, Maier M (1985) Tetrahedron Lett 26:2065
371. Hayman CM, Hanton LR, Larsen DS, Guthrie JM (1999) Austr J Chem 52:921
372. Dujardin G, Rossignol S, Brown E (1998) Synthesis 763
373. Dujardin G, Rossignol S, Brown E (1996) Tetrahedron Lett 37:4007
374. Liu HM, Zou DP, Zhang FY, Zhu WG, Peng T (2004) Eur J Org Chem 2103
375. Evans DA, Johnson JS (1998) J Am Chem Soc 120:4895
376. Evans DA, Olhava EJ, Johnson JS, Janey JM (1998) Angew Chem Int Ed Engl 37:3372
377. Evans DA, Johnson JS, Olhava EJ (2000) J Am Chem Soc 122:1635
378. Thorhauge J, Johannsen M, Jorgensen KA (1998) Angew Chem Int Ed Engl 37:2404
379. Audrain H, Thorhauge J, Hazell RG, Jorgensen KA (2000) J Org Chem 65:4487
380. Saito T, Takekawa K, Takahashi T (1999) Chem Commun 1001
381. Werbitzky O, Klier K, Felber H (1990) Liebigs Ann Chem 267
382. Braun H, Felber H, Kresze G, Schmidtchen FP, Prewo R, Vasella A (1993) Liebigs Ann Chem 261
383. Streith J, Defoin A (1996) Synlett 189
384. Defoin A, Joubert M, Heuchel JM, Strehler C, Streith J (2000) Synthesis 1719
385. Felber H, Kresze G, Prewo R, Vasella A (1986) Helv Chim Acta 69:1137
386. Defoin A, Sarazin H, Streith J (1995) Synlett 1187
387. Defoin A, Sarazin H, Streith J (1996) Helv Chim Acta 79:560
388. Defoin A, Sarazin H, Streith J (1997) Tetrahedron 53:13769
389. Sifferlen T, Defoin A, Streith J, Le Nouen D, Tarnus C, Dosbaa I, Foglietti MJ (2000) Tetrahedron 56:971
390. Joubert M, Defoin A, Tarnus C, Streith J (2000) Synlett 1366
391. Defoin A, Sarazin H, Sifferlen T, Strehler C, Streith J (1998) Helv Chim Acta 81:1417
392. Defoin A, Sifferlen T, Streith J, Dosbaa I, Foglietti MJ (1997) Tetrahedron:Asymmetry 8:363
393. Bach P, Bols M (1999) Tetrahedron Lett 40:3461
394. Liang XF, Bols M (1999) J Org Chem 64:8485
395. Ernholt BV, Thomsen IB, Lohse A, Plesner IW, Jensen KB, Hazell RG, Liang XF, Jakobsen A, Bols M (2000) Chem Eur J 6:278
396. Just G, Martel A (1973) Tetrahedron Lett 17:1517
397. Just G, Grozinger K (1975) Can J Chem 53:2701
398. Just G, Martel A, Grozinger K, Ramjeesingh M (1975) Can J Chem 53:131

399. Just G, Gronzinger K (1974) Tetrahedron Lett 15:4165
400. Just G, Ramjeesingh M, Liak TJ (1976) Can J Chem 54:2940
401. Just G, Lim MI (1977) Can J Chem 55:2993
402. Just G, Liak TJ, Lim MI, Potvin P, Tsantrizos YS (1980) Can J Chem 58:2024
403. Kozikowski AP, Floyd WC (1978) Tetrahedron Lett 1:19
404. Kowarski CR, Sarel S (1973) J Org Chem 38:117
405. Schmidt RR, Lieberknecht A (1978) Angew Chem Int Ed Engl 17:769
406. Sadeghi-Khomami A, Blake AJ, Wilson C, Thomas NR (2005) Org Lett 7:4891
407. Buser S, Vasella A (2005) Helv Chim Acta 88:3151
408. Vieira E, Vogel P (1983) Helv Chim Acta 66:1865
409. Reymond JL, Vogel P (1990) Tetrahedron: Asymmetry 1:729
410. Forster A, Kovac T, Mosimann H, Renaud P, Vogel P (1999) Tetrahedron:Asymmetry 10:567
411. Warm A, Vogel P (1987) Helv Chim Acta 70:690
412. Vogel P, Cossy J, Plumet J, Arjona O (1999) Tetrahedron 55:13521
413. Auberson Y, Vogel P (1989) Helv Chim Acta 72:278
414. Ledrian C, Vieira E, Vogel P (1989) Helv Chim Acta 72:338
415. Ledrian C, Vogel P (1987) Helv Chim Acta 70:1703
416. Ledrian C, Vionnet JP, Vogel P (1990) Helv Chim Acta 73:161
417. Emery F, Vogel P (1995) Synlett 420
418. Carrupt PA, Vogel P (1982) Tetrahedron Lett 23:2563
419. Carrupt PA, Vogel P (1984) Tetrahedron Lett 25:2879
420. Carrupt PA, Vogel P (1988) J Phys Org Chem 1:287
421. Carrupt PA, Vogel P (1989) Helv Chim Acta 72:1008
422. Carrupt PA, Vogel P (1990) J Org Chem 55:5696
423. Gerber P, Vogel P (2001) Ind J Chem Sec B-Org Chem Inc Med Chem 40:898
424. Gasparini F, Vogel P (1990) J Org Chem 55:2451
425. Fattori D, Deguchteneere E, Vogel P (1989) Tetrahedron Lett 30:7415
426. Fattori D, Vogel P (1992) Tetrahedron 48:10587
427. Warm A, Vogel P (1986) J Org Chem 51:5348
428. Vogel P, Fattori D, Gasparini F, Ledrian C (1990) Synlett 173
429. Vogel P (1990) Bull Soc Chim B 99:395
430. Vogel P (2000) Curr Org Chem 4:455
431. Bimwala RM, Vogel P (1992) J Org Chem 57:2076
432. Ferritto R, Vogel P (1994) Tetrahedron:Asymmetry 5:2077
433. Emery F, Vogel P (1995) J Org Chem 60:5843
434. Cossy J, Ranaivosata JL, Bellosta V, Ancerewicz J, Ferritto R, Vogel P (1995) J Org Chem 60:8351
435. Ferritto R, Vogel P (1996) Synlett 281
436. Jeanneret V, Meerpoel L, Vogel P (1997) Tetrahedron Lett 38:543
437. Baudat A, Vogel P (1997) J Org Chem 62:6252
438. Kraehenbuehl K, Picasso S, Vogel P (1998) Helv Chim Acta 81:1439
439. Pasquarello C, Picasso S, Demange R, Malissard M, Berger EG, Vogel P (2000) J Org Chem 65:4251
440. Gerber P, Vogel P (2001) Helv Chim Acta 84:1363
441. Allemann S, Vogel P (1994) Helv Chim Acta 77:1
442. Bimwala RM, Vogel P (1989) Helv Chim Acta 72:1825
443. Jeanneret V, Gasparini F, Pechy P, Vogel P (1992) Tetrahedron 48:10637
444. Pechy P, Gasparini F, Vogel P (1992) Synlett 676
445. Gasparini F, Vogel P (1989) Helv Chim Acta 72:271
446. Nativi C, Reymond JL, Vogel P (1989) Helv Chim Acta 72:882
447. Hunenberger P, Allemann S, Vogel P (1994) Carbohydr Res 257:175
448. Auberson Y, Vogel P (1990) Tetrahedron 46:7019
449. Auberson Y, Vogel P (1989) Angew Chem Int Ed Engl 28:1498
450. Deguchteneere E, Fattori D, Vogel P (1992) Tetrahedron 48:10603
451. Durgnat JM, Vogel P (1993) Helv Chim Acta 76:222
452. Jeganathan S, Vogel P (1991) J Org Chem 56:1133
453. Chen YW, Vogel P (1994) J Org Chem 59:2487
454. Kernen P, Vogel P (1993) Tetrahedron Lett 34:2473
455. Sevin AF, Vogel P (1994) J Org Chem 59:5920
456. Guidi A, TheurillatMoritz V, Vogel P, Pinkerton AA (1996) Tetrahedron:Asymmetry 7:3153

457. TheurillatMoritz V, Vogel P (1996) Tetrahedron:Asymmetry 7:3163
458. Jotterand N, Vogel P, Schenk K (1999) Helv Chim Acta 82:821
459. Jotterand N, Vogel P (1999) J Org Chem 64:8973
460. Vasella A, Voeffray R (1982) Helv Chim Acta 65:1134
461. Muller I, Jager V (1982) Tetrahedron Lett 23:4777
462. Jager V, Muller I (1985) Tetrahedron 41:3519
463. Muller R, Leibold T, Patzel M, Jager V (1994) Angew Chem Int Ed Engl 33:1295
464. Schaller C, Vogel P, Jager V (1998) Carbohydr Res 314:25
465. Cox PJ, Simpkins NS (1991) Synlett 321
466. Hoffmann HMR, Dunkel R, Mentzel M, Reuter H, Stark CBW (2001) Chem Eur J 7:4771
467. Stark CBW, Pierau S, Wartchow R, Hoffmann HMR (2000) Chem Eur J 6:684
468. Schwenter ME, Vogel P (2000) Chem Eur J 6:4091
469. Gerber-Lemaire S, Vogel P (2003) Eur J Org Chem 2959
470. Gerber-Lemaire S, Vogel P (2004) Eur J Org Chem 5040
471. Meilert KT, Schwenter ME, Shatz Y, Dubbaka SR, Vogel P (2003) J Org Chem 68:2964
472. Evans DA, Chapman KT, Carreira EM (1988) J Am Chem Soc 110:3560
473. Chen KM, Hardtmann GE, Prasad K, Repic O, Shapiro MJ (1987) Tetrahedron Lett 28:155
474. Gerber-Lemaire S, Carmona AT, Meilert KT, Vogel P (2006) Eur J Org Chem 891
475. Gerber-Lemaire S, Popowycz F, Glanzmann C, Vogel P (2002) Synthesis 1979
476. Csaky AG, Vogel P (2000) Tetrahedron:Asymmetry 11:4935
477. Chen ZM, Trudell ML (1996) Chem Rev 96:1179
478. Zhang CM, Ballay CJ, Trudell ML (1999) J Chem Soc, Perkin Trans 1675
479. Moreno-Vargas AJ, Schutz C, Scopelliti R, Vogel P (2003) J Org Chem 68:5632
480. Moreno-Vargas AJ, Vogel P (2003) Tetrahedron:Asymmetry 14:3173
481. Alexakis A, Frutos JC, Mangeney P (1993) Tetrahedron:Asymmetry 4:2431
482. Moreno-Vargas AJ, Robina I, Petricci E, Vogel P (2004) J Org Chem 69:4487
483. Moreno-Vargas AJ, Vogel P (2003) Tetrahedron Lett 44:5069
484. Holland D, Stoddart JF (1983) J Chem Soc, Perkin Trans 11553
485. Grethe G, Sereno J, Williams TH, Uskokovic MR (1983) J Org Chem 48:5315
486. Johnson CR, Golebiowski A, Braun MP, Sundram H (1994) Tetrahedron Lett 35:1833
487. Johnson CR, Braun MP (1993) J Am Chem Soc 115:11014
488. Johnson CR, Adams JP, Braun MP, Senanayake CBW, Wovkulich PM, Uskokovic MR (1992) Tetrahedron Lett 33:917
489. Johnson CR, Penning TD (1988) J Am Chem Soc 110:4726
490. Parry RJ, Haridas K, Dejong R, Johnson CR (1990) Tetrahedron Lett 31:7549
491. Johnson CR, Golebiowski A, Schoffers E, Sundram H, Braun MP (1995) Synlett 313
492. Johnson CR, Nerurkar BM, Golebiowski A, Sundram H, Esker JL (1995) J Chem Soc, Chem Commun 1139
493. Mehta G, Mohal N (2000) Tetrahedron Lett 41:5741
494. Mehta G, Mohal N (2000) Tetrahedron Lett 41:5747
495. Hudlicky T, Entwistle DA, Pitzer KK, Thorpe AJ (1996) Chem Rev 96:1195
496. Banwell M, De Savi C, Watson K (1998) J Chem Soc, Perkin Trans 12251
497. Banwell MG, Blakey S, Harfoot G, Longmore RW (1999) Austr J Chem 52:137
498. Modyanova L, Azerad R (2000) Tetrahedron Lett 41:3865
499. Johnson CR, Golebiowski A, Steensma DH (1992) J Am Chem Soc 114:9414
500. Johnson CR, Kozak J (1994) J Org Chem 59:2910
501. Johnson CR, Golebiowski A, Kozak J (1998) Carbohydr Res 309:331
502. Pearson AJ, Katiyar S (2000) Tetrahedron 56:2297
503. Mehta G, Pallavi K (2004) Tetrahedron Lett 45:3865
504. Jung ME, Kretschik O (1998) J Org Chem 63:2975
505. Rychnovsky SD, Griesgraber G, Zeller S, Skalitzky DJ (1991) J Org Chem 56:5161
506. Adams R, Voorhees V (1941) Org Synth Coll 2nd:280
507. Kaminska J, Gornicka I, Sikora M, Gora J (1996) Tetrahedron:Asymmetry 7:907
508. Ghanem A, Schurig V (2000) Chirality 13:118

509. Ghanem A (2003) Org Biomol Chem 1: 1282
510. Mandal SK, Sigman MS (2003) J Org Chem 68:7535
511. Akai S, Naka T, Omura S, Tanimoto K, Imanishi M, Takebe Y, Matsugi M, Kita Y (2002) Chem Eur J 8:4255
512. Kobayashi Y, Kusakabe M, Kitano Y, Sato F (1988) J Org Chem 53:1586
513. Kametani T, Tsubuki M, Tatsuzaki Y, Honda T (1990) J Chem Soc, Perkin Trans 1639
514. Zhu LZ, Talukdar A, Zhang GS, Kedenburg JP, Wang PG (2005) Synlett 1547
515. Kaminska JE, Smigelski K, Lobodzinska D, Gora J (2000) Tetrahedron:Asymmetry 11:1211
516. Nelson A (2004) New J Chem 28:771
517. Roush WR, Brown RJ, Dimare M (1983) J Org Chem 48:5083
518. Roush WR, Brown RJ (1982) J Org Chem 47:1371
519. Roush WR, Straub JA, Vannieuwenhze MS (1991) J Org Chem 56:1636
520. Roush WR, Brown RJ (1983) J Org Chem 48:5093
521. Nicolaou KC, Rodriguez RM, Mitchell HJ, Van Delft FL (1998) Angew Chem Int Ed Engl 37:1874
522. Kufner U, Schmidt RR (1986) Angew Chem Int Ed Engl 25:89
523. Dai LX, Lou BL, Zhang YZ (1988) J Am Chem Soc 110:5195
524. Ono M, Saotome C, Akita H (1997) Heterocycles 45:1257
525. Schmidt RR, Frische K (1988) Liebigs Ann Chem 209
526. Xu YM, Zhou WS (1996) Tetrahedron Lett 37:1461
527. Marshall JA, Tang Y (1994) J Org Chem 59:1457
528. Jager V, Hummer W (1990) Angew Chem Int Ed Engl 29:1171
529. Trost BM, Crawley ML (2003) Chem Rev 103:2921
530. Trost BM, Horne DB, Woltering MJ (2006) Chem Eur J 12:6607
531. Ko SY, Malik M (1993) Tetrahedron Lett 34:4675
532. Jung ME, Gardiner JM (1994) Tetrahedron Lett 35:6755
533. Taniguchi T, Ohnishi H, Ogasawara K (1996) Chem Commun 1477
534. Harris JM, Keranen MD, Nguyen H, Young VG, O'Doherty GA (2000) Carbohydr Res 328:17
535. Harris JM, O'Doherty GA (2000) Tetrahedron Lett 41:183
536. Harris JM, Keranen MD, O'Doherty GA (1999) J Org Chem 64:2982
537. Haukaas MH, O'Doherty GA (2002) Org Lett 4:1771
538. Farina V, Kapadia S, Krishnan B, Wang CJ, Liebeskind LS (1994) J Org Chem 59:5905
539. Hassan J, Sevignon M, Gozzi C, Schulz E, Lemaire M (2002) Chem Rev 102:1359
540. Cox LR, Deboos GA, Fullbrook JJ, Percy JM, Spencer NS, Tolley M (2003) Org Lett 5:337
541. Ahmed MM, Berry BP, Hunter TJ, Tomcik DJ, O'Doherty GA (2005) Org Lett 7:745
542. Ahmed MM, O'Doherty GA (2005) Tetrahedron Lett 46:3015
543. Ahmed MM, O'Doherty GA (2005) J Org Chem 70:10576
544. Matsushima Y, Kino J (2005) Tetrahedron Lett 46:8609
545. Matsushima Y, Nakayama T, Tohyama S, Eguchi T, Kakinuma K (2001) J Chem Soc, Perkin Trans 1569
546. Hong BC, Chen ZY, Nagarajan A, Kottani R, Chavan V, Chen WH, Jiang YF, Zhang SC, Liao JH, Sarshar S (2005) Carbohydr Res 340:2457
547. Zehnder LR, Wei LL, Hsung RP, Cole KP, McLaughlin MJ, Shen HC, Sklenicka HM, Wang JS, Zificsak CA (2001) Org Lett 3:2141
548. Vedejs E, Kruger AW (1999) J Org Chem 64:4790
549. Lindstrom UM, Ding R, Hidestal O (2005) Chem Commun 1773
550. Kamata K, Yamaguchi K, Hikichi S, Mizuno N (2003) Adv Synth Catal 345:1193
551. Robinson TV, Taylor DK, Tiekink ERT (2006) J Org Chem 71:7236
552. Li GG, Chang HT, Sharpless KB (1996) Angew Chem Int Ed Engl 35:451
553. Bushey ML, Haukaas MH, O'Doherty GA (1999) J Org Chem 64:2984
554. Yang CF, Xu YM, Liao LX, Zhou WS (1998) Tetrahedron Lett 39:9227
555. Ciufolini MA, Hermann CYW, Dong Q, Shimizu T, Swaminathan S, Xi N (1998) Synlett 105

4.4 Monosaccharides as Chiral Pools for the Synthesis of Complex Natural Compounds

Masaya Nakata
Department of Applied Chemistry, Faculty of Science and Technology, Keio University, Yokohama 223-8522, Japan
msynktxa@applc.keio.ac.jp

Abstract

A variety of complex natural compounds are synthesized by using monosaccharides as chiral starting materials. Exemplifying eight representative natural products, i. e., erythromycins, elaiophilin, herbimaycins, calbistrin A, lactacystin, tautomycin, FK-506, and halichondrin B, the methodology how monosaccharides are incorporated into natural product synthesis as chiral pools is described.

Keywords

Chiral pool; Natural product; Erythromycins; Elaiophilin; Herbimycins; Calbistrin A; Lactacystin; Tautomycin; FK-506; Halichondrin B

Abbreviations

AIBN	2,2′-azobisisobutyronitrile
All	allyl

In: *Glycoscience*. Fraser-Reid B, Tatsuta K, Thiem J (eds)
Chapter-DOI 10-1007/978-3-540-30429-6_21:

CSA	10-camphorsulfonic acid
DDQ	2,3-dichloro-5,6-dicyano-1,4-benzoquinone
DEAD	diethyl azodicarboxylate
DEIPS	diethylisopropylsilyl
DHQD	dihydroquinidine *p*-chlorobenzoate
DIAD	diisopropyl azodicarboxylate
DIBALH	diisobutylaluminum hydride
DMAP	4-(dimethylamino)pyridine
DME	1,2-dimethoxyethane
DMF	*N*,*N*-dimethylformamide
DMSO	dimethyl sulfoxide
HMPA	hexamethylphosphoramide
LDA	lithium diisopropylamide
LHMDS	lithium bis(trimethylsilyl)amide
***m*-CPBA**	3-chloroperoxybenzoic acid
MMTr	1-(bis(4-methoxyphenyl))-1-phenylmethyl
MOM	methoxymethyl
MP	4-methoxyphenyl
MPM	(4-methoxyphenyl)methyl
Ms	methanesulfonyl
NBS	*N*-bromosuccinimide
NIS	*N*-iodosuccinimide
NMO	4-methylmorpholine *N*-oxide
PCC	pyridinium chlorochromate
PDC	pyridinium dichromate
PPTS	pyridinium *p*-toluenesulfonate
TBAF	tetrabutylammonium fluoride
TBDMS	*t*-butyldimethylsilyl
TBDMSOTf	*t*-butyldimethylsilyl trifluoromethanesulfonate
TBDPSCl	*t*-butyldiphenylsilyl chloride
TFA	trifluoroacetic acid
THF	tetrahydrofuran
THP	tetrahydropyranyl
Tf	trifluoromethanesulfonyl
Tr	triphenylmethyl
Ts	*p*-toluenesulfonyl
TMEDA	N,N,N',N'-tetramethylethylenediamine
TMSCl	trimethylsilyl chloride

1 Introduction

Asymmetric total synthesis of natural products is one of the central subjects in the organic synthetic world. In order to reach this goal, there are two main strategies: the first uses chiral auxiliaries and the second uses chiral starting materials. The first most important task for comple-

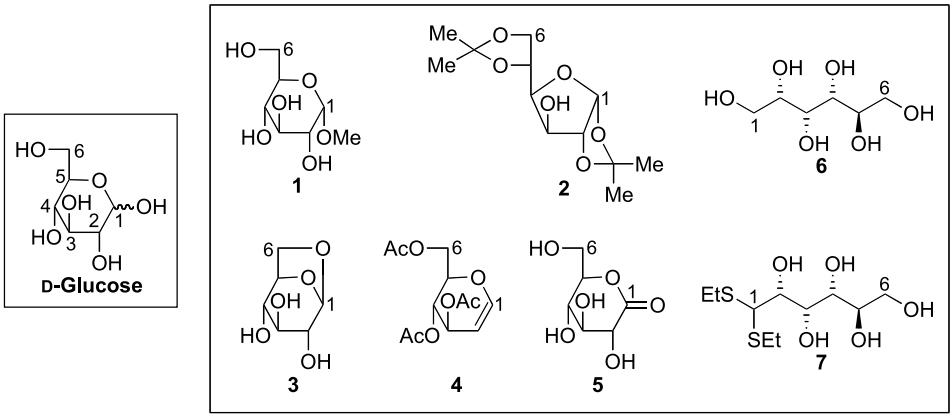


Figure 1
D-Glucose and its derivatives as chiral starting materials

tion of the total synthesis of complex natural products is reasonable disconnection of the target compound into the small starting materials by retrosynthetic analysis. Monosaccharides (carbohydrates) have been extensively used as fascinating starting materials in the enantiospecific synthesis of natural products. Some useful books on the carbohydrates-based chiral synthesis of natural products are available [1,2,3,4,5,6,7]. The advantages of using monosaccharides as a chiral pool are summarized as follows: (1) commercial availability; (2) low cost; (3) valuable carbon source (usually containing five or six carbon atoms); (4) enantiopure materials; (5) several definite chiral centers; (6) conformational rigidity which makes chemical transformations stereoselective.

As an example, several types of derivatives of D-glucose are available for starting materials (*Fig. 1*): methyl α-D-glucopyranoside (**1**), 1,2:5,6-di-*O*-isopropylidene-α-D-glucofuranose (**2**), levoglucosan (**3**), tri-*O*-acetyl-D-glucal (**4**), gluconolactone (**5**), D-glucidol (**6**), and the dithioacetal derivative **7**. They are all commercially available and also easily prepared in large quantities from D-glucose.

Readily available chiral starting materials other than monosaccharides are amino acids, hydroxy acids, hydroxy esters, and terpenes. Among them, hydroxy acids and esters are often used as alternatives for monosaccharides when one requires one or two chiral centers to construct natural products (*Fig. 2*). Tartaric acid and its esters have four carbon atoms and two chiral centers. Both enantiomers, **8** (2*R*,3*R*) and **9** (2*S*,3*S*), are commercially available. Malic acid has four carbon atoms and one chiral center; both enantiomers, **10** and **11**, are commercially available although (*R*)-malic acid (**11**) is more expensive. Lactic acid and its esters have three carbon atoms and one chiral center; both enantiomers, **12** and **13**, are commercially available [(*R*)-lactic acid (**13**) is more expensive]. Glyceraldehyde and its derivatives are among the most useful members of the chiral pool, and they have three carbon atoms and one chiral center. Both enantiomers, **14** (*R*) and **15** (*S*), are available from mannitol and L-ascorbic acid, respectively; therefore, they also belong to the family of monosaccharides, but are not included in this chapter. Other familiar starting materials are methyl (*R*)- and (*S*)-3-hydroxy-2-methylpropionates, **16** and **17**. They have one chiral center bearing a methyl group and two

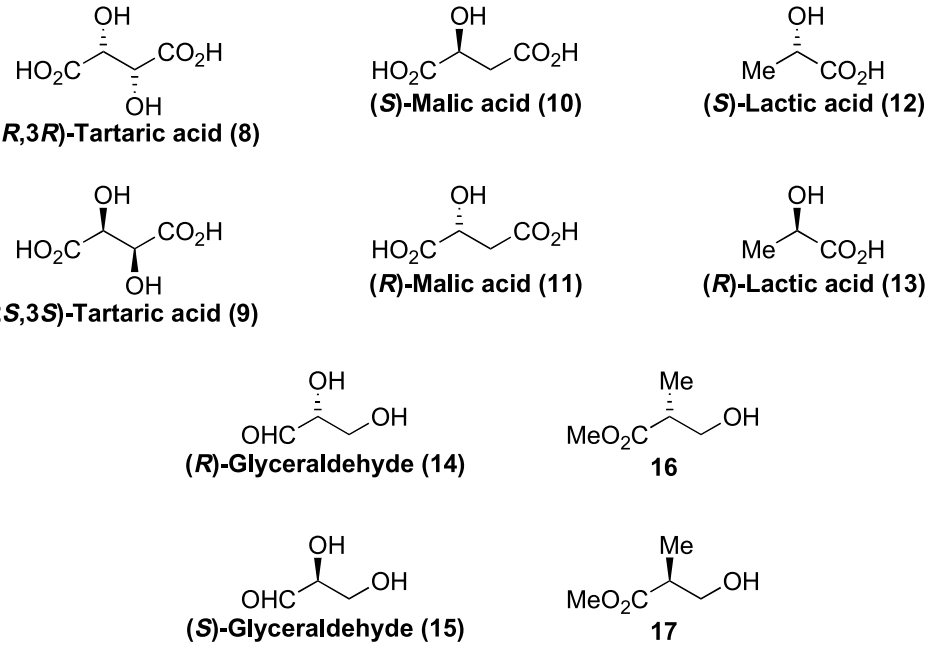


Figure 2
Hydroxy acids, esters, and aldehydes as chiral starting materials

differentiated functional groups: hydroxy and ester groups. Many natural products have been synthesized by assemblage of the synthetic intermediates prepared from monosaccharides and/or the above-mentioned chiral starting materials.

The chemical operations necessary for constructing the specified natural product from monosaccharides are the following: (1) differentiation of the hydroxy groups by selective protection; (2) oxidation of a hydroxy group; (3) deoxygenation of a hydroxy group; (4) inversion of the configuration of a hydroxy-bearing carbon; (5) addition of some nucleophiles to a carbonyl group, an epoxide, and an olefin; (6) conversion of cyclic monosaccharides to acyclic derivatives. The combination of these chemical operations makes possible the transformation of monosaccharides into the useful synthetic intermediates for completion of the total synthesis of natural products. Much useful information on individual chemical operations will be found in other chapters of this book. In this chapter, some recent examples of the synthesis of complex natural products employing monosaccharides as starting materials are described.

2 Synthesis of Natural Products

In natural product chemistry, there are three main categories of interest: isolation and structure determination, elucidation of natural phenomena, and synthesis. Among them, the synthetic community has been seeking new and efficient methods necessary for the synthesis of natural products and their analogs. In this section, we describe, exemplifying eight representative

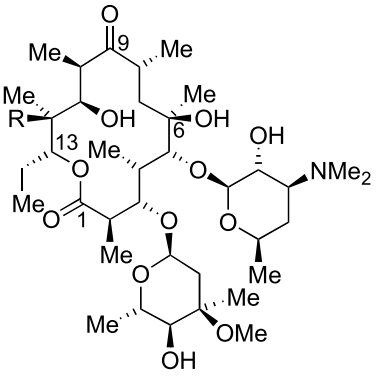
natural products, how the synthetic intermediates are prepared from monosaccharides. In the figures, the numbering system used for the natural products is employed. In the Schemes, the numbering system used for the monosaccharides is shown in parentheses in order to clarify how the starting monosaccharides are incorporated into the desired synthetic intermediates.

2.1 Erythromycins

Erythromycins, the representative and medicinally important macrolide antibiotics, have been widely studied and are still undoubtedly one of the most challenging target molecules for many synthetic organic chemists (● *Fig. 3*). Woodward and coworkers accomplished the first total synthesis of erythromycin A (**18**) in 1981 [8,9,10]. Corey and coworkers synthesized erythronolides A (**20**) and B (**21**), the aglycons of erythromycins A (**18**) and B (**19**), in 1978 [11,12] and 1979 [13].

Kinoshita, Nakata, and coworkers synthesized erythronolide A (**20**) (● *Fig. 4*) [14,15,16,17]. Erythronolide A (**20**) was divided into three segments, **22** (C1–C6), **23** (C7–C9), and **24** (C10–C13), of these fragments **22** and **24** were prepared from monosaccharides as chiral starting materials.

The C1–C6 segment **22** was prepared from methyl α-D-glucopyranoside (**1**) (● *Scheme 1*). The acetal function is a most popular protecting group of the hydroxy groups in monosaccharides. Benzylidenation of methyl α-D-glucopyranoside (**1**), which is also an acetal of D-glucose, with benzaldehyde in the presence of $ZnCl_2$ afforded the 4,6-*O*-benzylidene acetal **25**, whose C2-hydroxy group was selectively tosylated, and the resulting tosylate **26** was treated with a base to give epoxide **27**. Generally, epoxides are very useful synthetic intermediates because the rigidity of configurations makes regioselective ring-opening extremely effective. The nucleophilic attack of the methyl group to this epoxide **27** gave **28** as the sole product following the *trans*-diaxial opening rule. Inversion of the C2- and C3-configurations



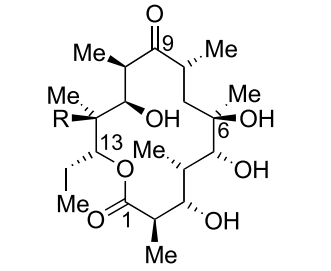


Figure 3
Erythromycins and their aglycons

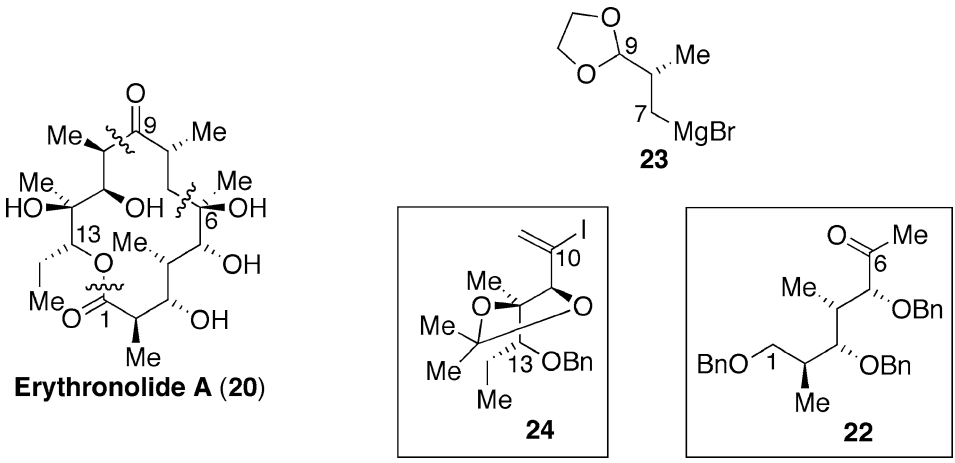


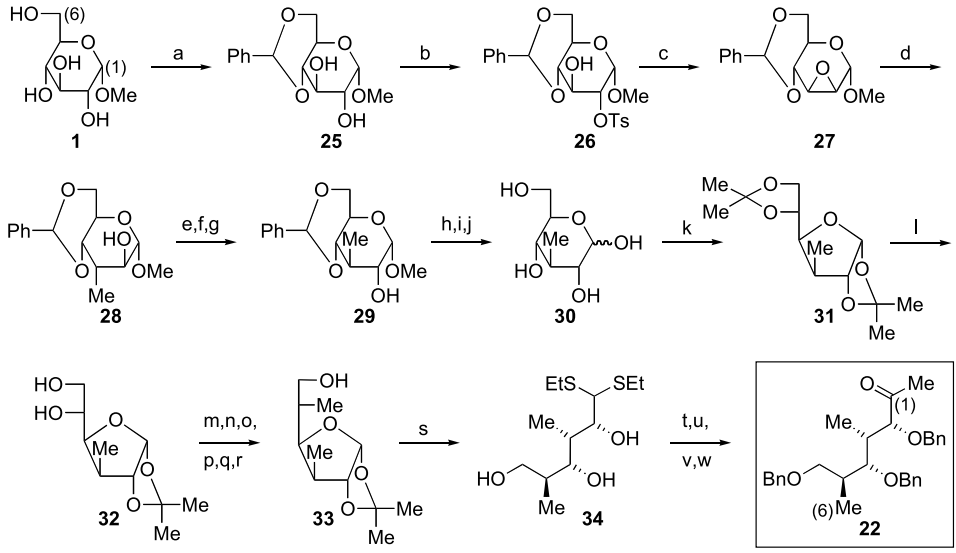
Figure 4
Segments of Kinoshita–Nakata synthesis of erythronolide A

of **28** was easily achieved by oxidation of the C2-hydroxy group, epimerization of the C3-position, and stereoselective reduction of the C2-carbonyl group, providing **29**. De-*O*-benzylidenation of **29** followed by acetolysis and hydrolysis afforded the free monosaccharide **30**. Pyranose-to-furanose conversion was realized by di-*O*-isopropylidenation of **30** with acetone in the presence of $FeCl_3$ to give the 1,2:5,6-di-*O*-isopropylidene derivative **31**. The 5,6-*O*-isopropylidene group in **31** was selectively deprotected with aqueous AcOH to afford **32**. The subsequent 6-step manipulation of **32** gave **33**. Conversion of cyclic monosaccharides to acyclic derivatives is one of the most important tasks when utilizing monosaccharides as a chiral pool in natural products synthesis. One such procedure is dialkyl dithioacetal opening. Treatment of **33** with ethanethiol in the presence of $BF_3 \cdot Et_2O$ provided **34**. Finally, the 4-step manipulation of **34** gave the C1–C6 segment **22** of erythronolide A (**20**).

The C10–C13 segment **24** was prepared from D-ribose (**35**) (*Scheme 2*). In this case, selective protection of the hydroxy groups was realized by isopropylidenation (from **35** to **36**). One of the other procedures for conversion of cyclic monosaccharides to acyclic derivatives is nucleophilic addition to the anomeric position in free monosaccharides. Grignard reagent, MeMgI, was added to **36** to provide **37** as the sole product. The subsequent manipulation of **37** to the C10–C13 segment **24**, which is not restricted in monosaccharides chemistry, is summarized in *Scheme 2*. After the completion of the synthesis of erythronolide A (**20**), Toshima, Nakata, Tatsuta, Kinoshita, and coworkers achieved the total synthesis of erythromycin A (**18**) by their own glycosidation method [18,19].

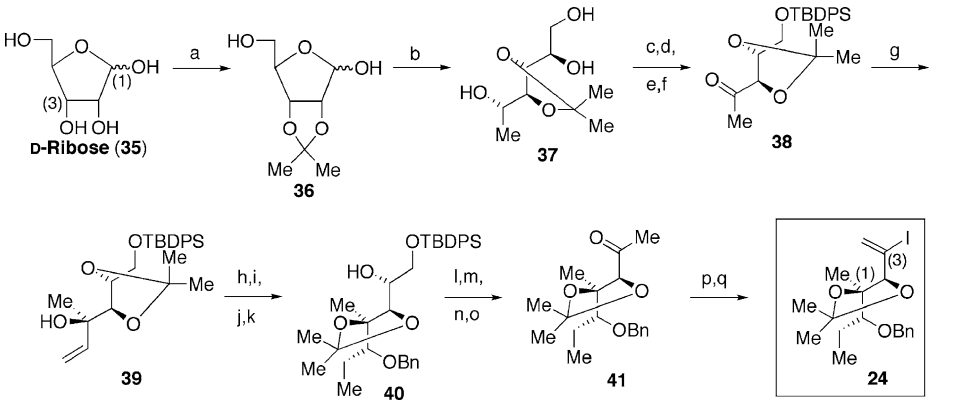
Kochetkov and coworkers synthesized erythronolides A (**20**) and B (**21**) [20,21]. They divided the target compounds into three segments, **42** (C1–C6), **43** (C7–C8), and **44** (C9–C13) or **45** (C9–C13) (*Fig. 5*).

The common C1–C6 segment **42** of erythronolides A (**20**) and B (**21**) was prepared from levoglucosan (**3**) (*Scheme 3*), which is also a very popular starting material. Selective di-*O*-tosylation of levoglucosan (**3**) gave the C2, C4-di-*O*-tosylate **46**, which was treated with a base to afford 3,4-epoxide **47**. Regioselective addition of MeMgCl to **47** in the presence of



Scheme 1

(a) PhCHO, $ZnCl_2$; (b) TsCl, py; (c) NaOMe/MeOH; (d) MeMgCl, ether; (e) DMSO, py, TFA, DCC, benzene; (f) Et_3N, DMF; (g) $LiAlH_4$, ether; (h) HCl, MeOH; (i) Ac_2O, H_2SO_4; (j) NaOH, MeOH; (k) acetone, $FeCl_3$; (l) aqueous AcOH, 60 °C; (m) TsCl, py; (n) NaOMe/MeOH, $CHCl_3$; (o) $LiAlH_4$, THF; (p) CrO_3, H_2SO_4, aqueous acetone; (q) Ph_3PCH_3Br, NaH, DMSO; (r) $BH_3 \cdot Me_2S$, CH_2Cl_2, then H_2O_2, aqueous NaOH; (s) EtSH, $BF_3 \cdot Et_2O$; (t) BnBr, NaH, DMF; (u) $HgCl_2$, HgO, aqueous acetone; (v) MeMgI, ether; (w) PCC, CH_2Cl_2



Scheme 2

(a) $Me_2C(OMe)_2$, H_2SO_4, acetone, 5 °C; (b) MeMgI, ether; (c) $NaIO_4$, aqueous acetone; (d) $LiAlH_4$, THF; (e) TBDPSCl, imidazole, DMF; (f) PCC, CH_2Cl_2; (g) CH_2=CHMgBr, THF; (h) O_3/O_2, CH_2Cl_2, −78 °C, then Me_2S; (i) EtMgBr, ether; (j) BnBr, NaH, THF; (k) $FeCl_3$, acetone; (l) TBAF, THF; (m) $NaIO_4$, aqueous acetone; (n) MeMgI, ether; (o) PCC, CH_2Cl_2; (p) $NH_2NH_2 \cdot H_2O$, Et_3N, EtOH, 70 °C; (q) I_2, tetramethylguanidine, toluene, 0 °C

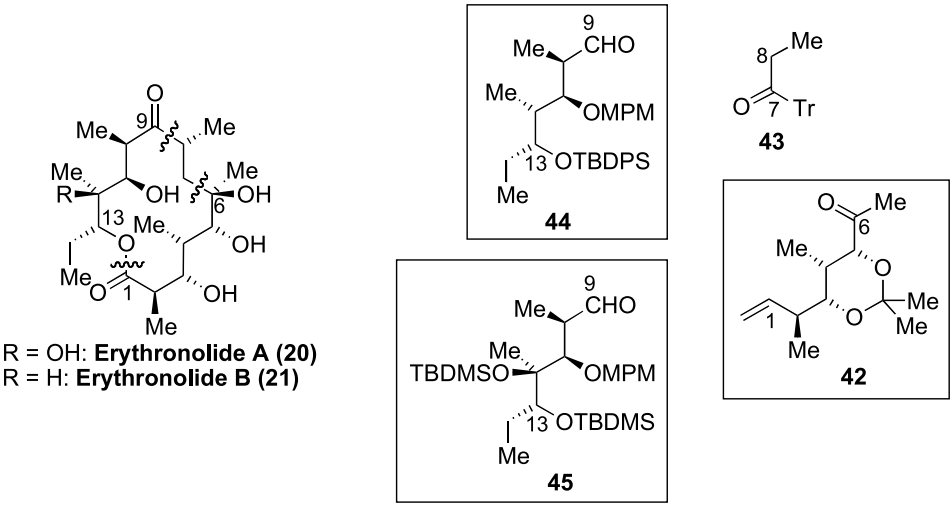
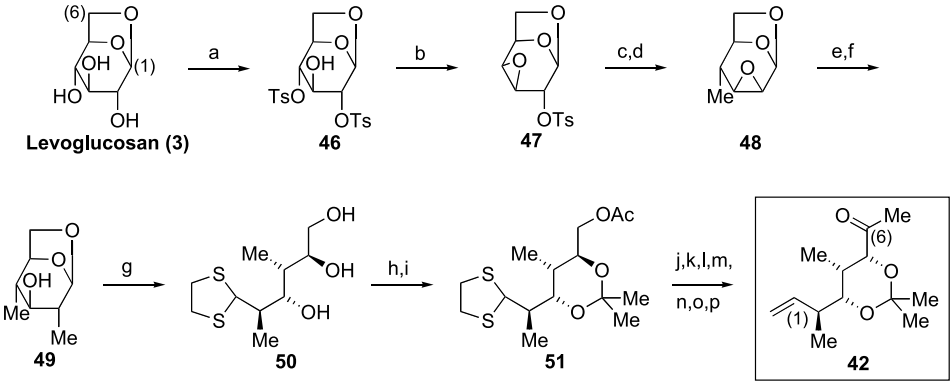


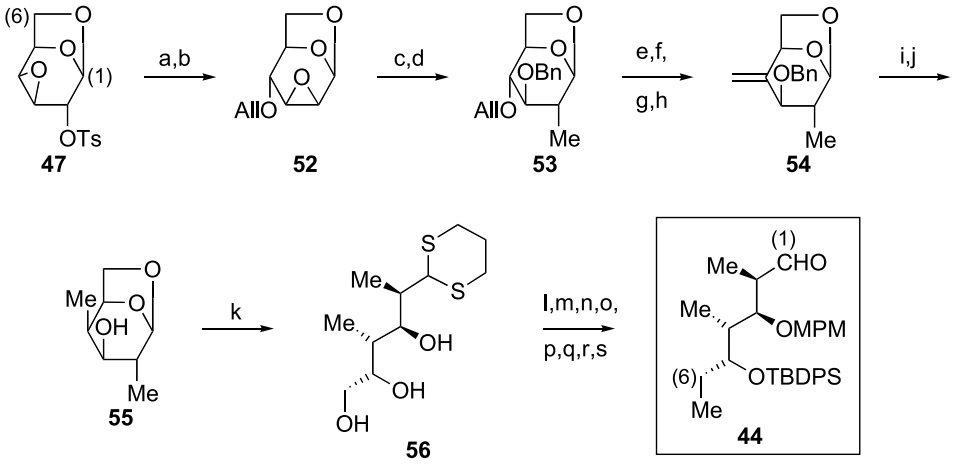
Figure 5
Segments of Kochetkov synthesis of erythronolides A and B



Scheme 3
(a) TsCl, py; (b) NaOMe/MeOH, $CHCl_3$, 0 °C; (c) MeMgCl, CuCl, THF, 0 °C; (d) NaOMe/MeOH, CH_2Cl_2, 0 °C; (e) DMSO, NaH; (f) Raney Ni, MeOH, heat; (g) $HS(CH_2)_2SH$, $BF_3 \cdot Et_2O$, CH_2Cl_2; (h) Ac_2O, py; (i) $Me_2C(OMe)_2$, TsOH, acetone; (j) $HgCl_2$, $CaCO_3$, aqueous MeCN; (k) $Ph_3P{=}CH_2$, benzene, heat; (l) NaOMe/MeOH; (m) Swern oxidation; (n) MeMgCl, ether; (o) Swern oxidation; (p) K_2CO_3, MeOH

CuCl provided the C3-alcohol as the sole, *trans*-diaxial opening product, which was treated with a base to afford **48**. The 2-step reaction for introducing the second methyl group to **48** (DMSO-anion addition and reductive desulfurization) gave the doubly methylated **49** [22]. Opening of the 1,6-anhydro ring was realized by treatment of **49** with 1,2-ethanedithiol and $BF_3 \cdot Et_2O$ to give **50**. This compound was transformed into the C1–C6 segment **42** by the procedure shown in *Scheme 3*.

The C9–C13 segment **44** of erythronolide B (**21**) was prepared (*Scheme 4*) from the synthetic intermediate **47** of the C1–C6 segment **42**. Instead of the methyl group introduction,



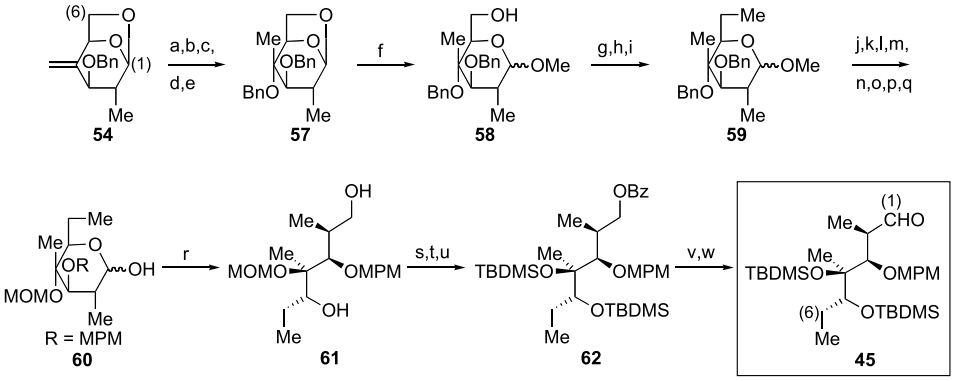
Scheme 4

(a) allyl alcohol, TsOH, benzene, heat; (b) NaOMe; (c) Me_2Mg, ether, heat; (d) BnBr, NaH, DMF; (e) *t*-BuOK, DMSO; (f) $HgCl_2$, aqueous acetone; (g) Swern oxidation; (h) $Ph_3P{=}CH_2$, benzene, heat; (i) $Cp_2Zr(H)Cl$, benzene, heat, then aqueous HCl; (j) H_2, Pd/C, MeOH; (k) $HS(CH_2)_3SH$, $BF_3 \cdot Et_2O$, CH_2Cl_2; (l) $Me_2C(OMe)_2$, TsOH, acetone; (m) MPMCl, NaH, DMF; (n) aqueous AcOH, heat; (o) TsCl, py; (p) K_2CO_3, MeOH; (q) MeMgCl, CuCl · Me_2S, THF; (r) TBDPSCl, Et_3N, CH_2Cl_2; (s) $HgCl_2$, $CdCO_3$, aqueous acetone

the *O*-function was introduced to epoxide **47** by acidic allyloxylation. After a base treatment, the resulting epoxide **52** was treated with Me_2Mg to provide the C2-methylated alcohol as the sole, *trans*-diaxial opening product. After benzylation, the resulting **53** was transformed into **54** by a 4-step manipulation. Stereoselective hydrozirconation of **54** followed by de-*O*-benzylation afforded the doubly methylated **55** [23,24,25]. Opening of the 1,6-anhydro ring was realized by treatment of **55** with 1,3-propanedithiol and $BF_3 \cdot Et_2O$ to afford **56**. Finally, the several-step manipulation of **56** gave the C9–C13 segment **44** of erythronolide B (**21**).

On the other hand, the C9–C13 segment **45** of erythronolide A (**20**) was prepared from **54** (*Scheme 5*), which is also a synthetic intermediate of the erythronolide B C9–C13 segment **44**. Using the double bond as a precursor of the tertiary alcohol portion, compound **54** was transformed into **57**. The 1,6-anhydro ring was once cleaved via methylglycosidation. After one-carbon elongation at the C6-position and several transformations, compound **58** was converted to the free monosaccharide **60** through **59**. Another procedure for the conversion of cyclic monosaccharides to acyclic derivatives is reduction of the anomeric carbon. $NaBH_4$ reduction of **60** gave the acyclic derivative **61**. A further 5-step manipulation provided the C9–C13 segment **45** of erythronolide A (**20**).

Kochetkov and coworkers synthesized erythronolides A (**20**) and B (**21**) via a different route [26,27,28]. There have been many reports on the syntheses of erythromycin. Among them, four examples are the Yonemitsu synthesis of erythronolide A (**20**) [29,30], featuring an extremely efficient macrocyclization (the modified Yamaguchi method); Martin's synthesis of erythromycin B (**19**) [31,32]; Evans synthesis of 6-deoxyerythronolide B [33,34], featuring the aldol-based assemblage of each synthetic segment; and the Carreira synthesis of erythronolide A (**20**), featuring the Mg-mediated cycloadditions of nitrile oxides [35].



Scheme 5
(a) OsO_4, NMO, aqueous acetone; (b) TsCl, py; (c) K_2CO_3, MeOH; (d) $LiAlH_4$, ether; (e) BnBr, NaH, DMF; (f) HCl, MeOH; (g) Swern oxidation; (h) $Ph_3P{=}CH_2$, benzene, heat; (i) $LiAlH_4$, $CoCl_2$, THF; (j) Raney Ni, EtOH, heat; (k) allyl alcohol, PPTS, heat; (l) Ac_2O, py, heat; (m) MOMCl, *i*-Pr_2NEt, CH_2Cl_2, heat; (n) NaOMe/MeOH; (o) MPMCl, NaH, DMF; (p) *t*-BuOK, DMSO; (q) $Hg(OAc)_2$, aqueous acetone; (r) $NaBH_4$, aqueous EtOH; (s) BzCl, py; (t) aqueous HCl, THF, 60 °C; (u) TBDMSOTf, Et_3N, CH_2Cl_2, heat; (v) NaOH, aqueous MeOH, heat; (w) Swern oxidation

2.2 Elaiophylin

Elaiophylin (**63**) belongs to a group of C_2-symmetrical 16-membered macrodiolides. Toshima, Tatsuta, and Kinoshita synthesized elaiophylin (**63**) by the assemblage of three segments, **64** (C1–C9), **65** (C10–C15), and **66** (monosaccharide moiety) [36,37] (*Fig. 6*).

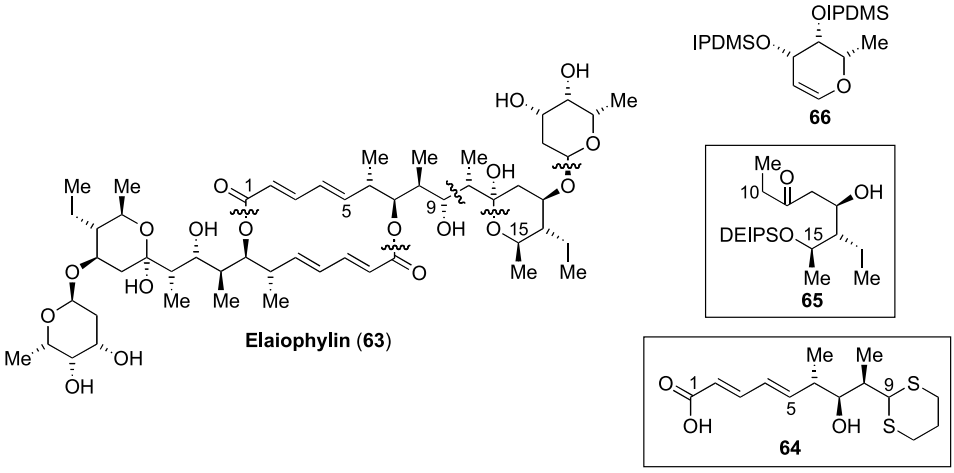
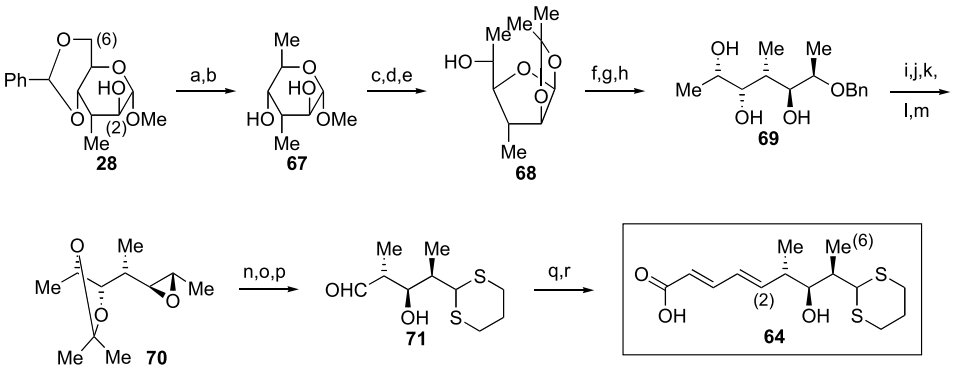


Figure 6
Segments of Toshima–Tatsuta–Kinoshita synthesis of elaiophylin

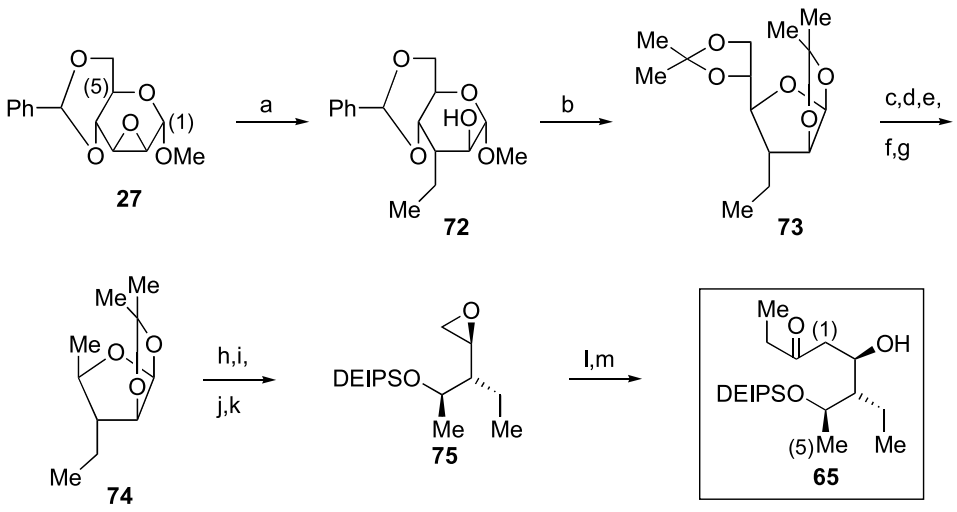


Scheme 6
(a) NBS, $BaCO_3$, CCl_4, heat; (b) $LiAlH_4$, THF, heat; (c) Ac_2O, H_2SO_4; (d) NaOH, aq MeOH; (e) acetone, $FeCl_3$; (f) BnBr, NaH, THF; (g) aqueous AcOH, 60 °C; (h) MeMgI, ether, heat; (i) $Me_2C(OMe)_2$, TsOH, DMF; (j) Ac_2O, DMAP, EtOAc; (k) H_2, Pd, MeOH; (l) MsCl, py; (m) NaOMe/MeOH, $CHCl_3$, 0 °C; (n) 2-lithio-1,3-dithiane, THF, 0 °C; (o) aqueous AcOH, 50 °C; (p) $Pb(OAc)_4$, KOAc, MeCN, −25 °C; (q) [(2*E*)-3-methoxycarbonyl-2-propenylidene]triphenylphosphorane, toluene, 80 °C; (r) LiOH, aqueous THF

The C1–C9 segment **64** was prepared from **28** (*Scheme 6*), which is the synthetic intermediate of erythronolide A (**20**). NBS-treatment of **28** followed by reduction gave the di-*C*-methylated compound **67**. Acetolysis of **67** followed by hydrolysis gave the free monosaccharide, which was subjected to isopropylidenation with acetone and $FeCl_3$ (pyranose-to-furanose conversion) to provide **68**. After benzylation of **68**, the 1,2-*O*-isopropylidene group was deprotected, and the resulting free monosaccharide was opened by the Grignard reaction (MeMgI) to give triol **69**. Triol **69** was transformed into the C1–C9 segment **64** by the route shown in *Scheme 6*.

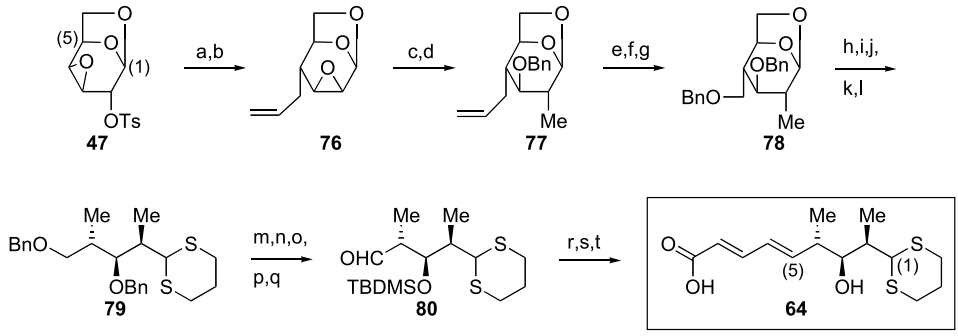
The C10–C15 segment **65** was prepared from epoxide **27** (*Scheme 7*), which is also a synthetic intermediate of erythronolide A (**20**). The nucleophilic attack of the ethyl group (EtMgCl) to epoxide **27** gave **72** as the sole product following the *trans*-diaxial opening rule. Pyranose-to-furanose conversion was realized by isopropylidenation with acetone and $BF_3 \cdot Et_2O$ to afford di-*O*-isopropylidene derivative **73**. The 5-step manipulation of **73** including selective de-*O*-isopropylidenation provided **74**. Subsequent 4-step manipulation of **74** including reductive opening of furanose to the acyclic derivative gave epoxide **75**. The 2-step manipulation of epoxide **75** provided the C10–C15 segment **65**.

Wakamatsu, Ban, Shibasaki, and coworkers have succeeded in the formal total synthesis of elaiophylin (**63**) [38] (*Scheme 8*). The C1–C9 segment **64** was prepared from epoxy-tosylate **47**, which is the synthetic intermediate of Kochetkov's erythronolide synthesis. This was regioselectively *C*-allylated with allylmagnesium chloride in the presence of CuI (*trans*-diaxial opening), and subsequent base treatment gave **76**. The methyl group was introduced regioselectively to **76** by MeMgCl in the presence of CuI (*trans*-diaxial opening), giving **77** after benzylation. After the 3-step conversion, the 1,6-anhydro ring of the resulting **78** was opened with 1,3-propanedithiol, $BF_3 \cdot Et_2O$, and TFA, and the resulting acyclic derivative was transformed to the C1–C9 segment **64** by the route shown in *Scheme 8*.



Scheme 7

(a) EtMgCl, ether; (b) acetone, $BF_3 \cdot Et_2O$; (c) aqueous AcOH; (d) $NaIO_4$, aqueous acetone; (e) $NaBH_4$, MeOH; (f) MsCl, py; (g) $LiAlH_4$, ether; (h) aqueous AcOH, 100 °C; (i) $LiAlH_4$, THF, 70 °C; (j) Ph_3P, DEAD, benzene, heat; (k) DEIPSCl, imidazole, CH_2Cl_2; (l) 2-ethyl-2-lithio-1,3-dithiane, THF, −20 °C; (m) $HgCl_2$, HgO, aqueous acetone, 0 °C



Scheme 8

(a) CH_2=$CHCH_2MgCl$, CuI, THF, 0 °C; (b) NaH, THF; (c) MeMgCl, CuI, ether, THF; (d) BnBr, NaH, Bu_4NI, THF; (e) $RhCl_3$, K_2CO_3, EtOH, heat; (f) O_3/O_2, CH_2Cl_2, −78 °C, then $NaBH_4$, MeOH, 0 °C; (g) BnBr, NaH, Bu_4NI, THF; (h) $HS(CH_2)_3SH$, $BF_3 \cdot Et_2O$, TFA, CH_2Cl_2, −30 °C; (i) $Pb(OAc)_4$, benzene, hexane, 0 °C; (j) $NaBH_4$, MeOH, 0 °C; (k) TsCl, DMAP, py, CH_2Cl_2; (l) $LiEt_3BH$, THF; (m) $HS(CH_2)_3SH$, $BF_3 \cdot Et_2O$, CH_2Cl_2; (n) PivCl, py, CH_2Cl_2, 0 °C; (o) TBDMSOTf, 2,6-lutidine, CH_2Cl_2, 0 °C; (p) $LiEt_3BH$, THF, −78 °C; (q) Swern oxidation; (r) [(2*E*)-3-methoxycarbonyl-2-propenylidene]triphenylphosphorane, toluene, 80 °C; (s) TBAF, benzoic acid, THF, heat; (t) LiOH, aqueous THF

Besides these two syntheses of elaiophylin (**63**), Seebach [39,40], Ziegler [41], Evans [42], and Paterson [43,44] have studied the synthesis of elaiophylin (**63**).

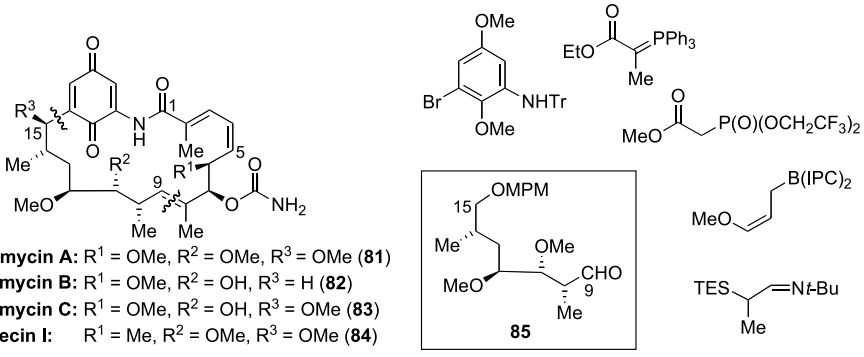
2.3 Herbimycins

Herbimycins A (**81**), B (**82**), and C (**83**), and macbecin I (**84**) belong to the benzoquinoid ansamycin antibiotics. Nakata, Tatsuta, and coworkers achieved the total synthesis of herbimycin A (**81**) [45,46] from the six segments as shown in *Fig. 7*.

The C9–C15 segment **85** was prepared from methyl α-D-mannopyranoside (**86**) (*Scheme 9*). Firstly, tritylation of **86** followed by selective mesylation gave the 2,3-di-*O*-mesylate **87**. This was treated with a base and the resulting epoxide **88** was opened with LiBr to give bromoepoxide **89**. Secondly, the regioselective opening of epoxide **89** with MeMgBr followed by a base treatment afforded **90**. Thirdly, the regioselective and reductive opening of epoxide **90** was realized by a borane reagent to provide **91**. The 6-step manipulation of **91** afforded **92**. After treatment of **92** with aqueous AcOH, the resulting free monosaccharide was opened by $NaBH_4$-reduction to provide **93**. The several-step manipulation of **93** gave the C9–C15 segment **85**.

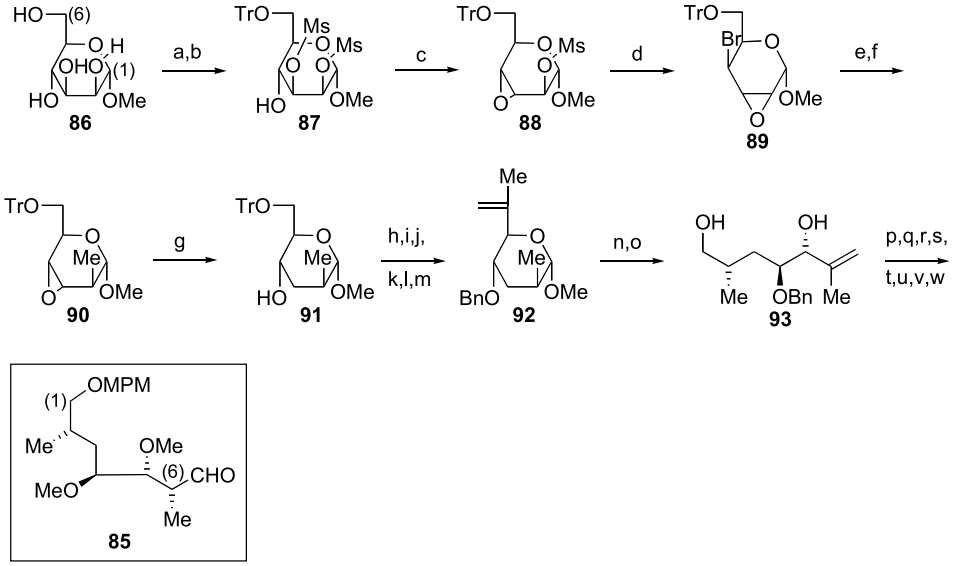
Kallmerten and coworkers synthesized the C5–C12 segment **94** of herbimycin A (**81**) (*Scheme 10*) [47]. The starting material was 1,2:5,6-di-*O*-isopropylidene-α-D-glucofuranose (**2**), which has only one free hydroxy group in its structure. Benzylation of the C3-hydroxy group of **2** followed by hydrolysis of both isopropylidene groups in **95** with acidic MeOH provided the methyl glucoside **96** in the form of its pyranoside. Tritylation and selective 2-*O*-methylation of **96** followed by C4-oxidation and stereoselective addition of Me_2CuLi afforded **97**. The 5-step transformation of **97** gave **98**. The Vasella-type reductive opening of **98** with Zn gave the free monosaccharide **99**, whose furanose structure was reductively opened with $LiBH_4$ to give **100**. Further 5-step manipulation of **100** afforded the C5–C12 segment **94**.

Panek and Carter also succeeded in the total synthesis of herbimycin A (**81**), employing their own crotylation methodology [48]. The structurally related macbecin I (**84**) was synthesized



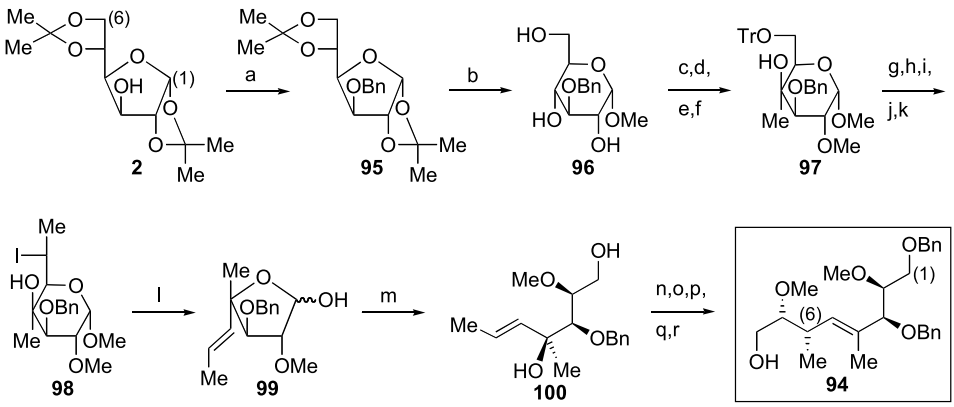
Herbimycin A: R^1 = OMe, R^2 = OMe, R^3 = OMe (**81**)
Herbimycin B: R^1 = OMe, R^2 = OH, R^3 = H (**82**)
Herbimycin C: R^1 = OMe, R^2 = OH, R^3 = OMe (**83**)
Macbecin I: R^1 = Me, R^2 = OMe, R^3 = OMe (**84**)

Figure 7
Segments of Nakata–Tatsuta synthesis of herbimycin A



Scheme 9

(a) TrCl, DMAP, py; (b) MsCl, py; (c) NaOMe/MeOH, dioxane; (d) LiBr, dioxane, 110 °C; (e) MeMgBr/ether, CH_2Cl_2; (f) NaH, DME, 0 °C; (g) $BH_3 \cdot Me_2S$, 2-methyl-2-butene, $NaBH_4$, THF, then H_2O_2, aqueous NaOH, 60 °C; (h) BnBr, NaH, DMF, 0 °C; (i) HCl/MeOH; (j) Swern oxidation; (k) MeMgBr/ether, THF, −78 °C; (l) PCC, CH_2Cl_2; (m) Ph_3PCH_3Br, BuLi, ether; (n) aqueous AcOH, 105 °C; (o) $NaBH_4$, EtOH; (p) $BH_3 \cdot Me_2S$, THF, then H_2O_2, aqueous NaOH, 60 °C; (q) $PhCH(OMe)_2$, CSA, CH_2Cl_2; (r) MPMCl, NaH, DMF; (s) H_2, Raney Ni, EtOH; (t) TBDPSCl, imidazole, DMF; (u) MeI, NaH, DMF; (v) TBAF, THF; (w) Dess–Martin oxidation



Scheme 10

(a) BnBr, NaH, DME; (b) H_2SO_4, MeOH; (c) TrCl, py; (d) MeI, NaH, THF; (e) SO_3, py, DMSO, Et_3N; (f) Me_2CuLi, MeLi, ether, −78 °C; (g) CSA, MeOH; (h) SO_3, py, DMSO, Et_3N; (i) MeMgBr, THF, −78 °C; (j) MsCl, py; (k) Bu_4NI, benzene; (l) Zn, EtOH; (m) $LiBH_4$, THF; (n) BnBr, NaH, DME; (o) 2-chloromethyl-4,5-dihydro-4,4-dimethyloxazole, KH, DME; (p) LDA, THF, −78 to 0 °C; (q) MeI, KH, DME; (r) TFA, H_2O, THF, then $LiAlH_4$

by Baker [49,50], Evans [51,52], and Panek [53,54]. The formal total synthesis of macbecin I was reported by Kallmerten [55,56] and Martin [57,58]. Synthetic studies of macbecin I were reported by Marshall [59].

2.4 Calbistrin A

Calbistrin A (**101**) is a cholesterol-lowering agent (*Fig. 8*). Tatsuta and coworkers have realized the total synthesis of calbistrin A (**101**) [60]. The key synthetic Diels–Alder precursor **102** for the decalin skeleton was prepared from aldehyde **103** and two other segments shown in *Fig. 8*.

The starting material for the octahydronaphthopyranone core is **91** (*Scheme 11*), which has been used in the total synthesis of herbimycin A (**81**). It was necessary to invert the C4-configuration for the crucial intramolecular Diels–Alder reaction at a later stage. Compound **91** was oxidized to ketone **104**, which was stereoselectively reduced with L-Selectride to provide

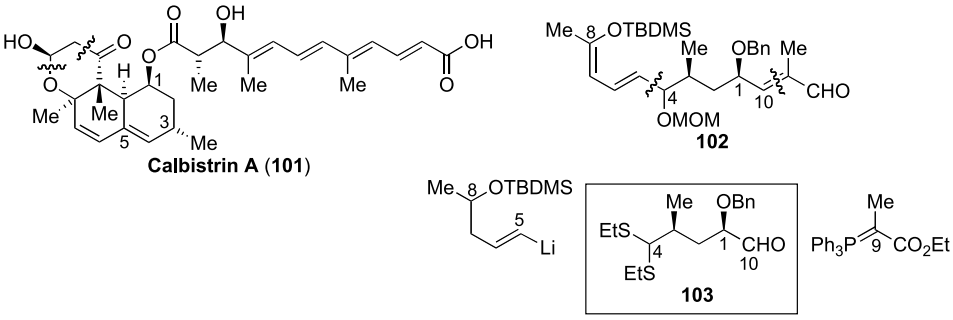
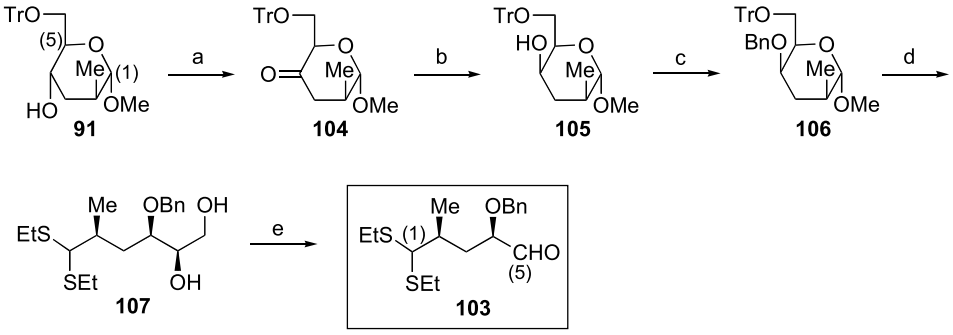


Figure 8
Segments of Tatsuta synthesis of calbistrin A



Scheme 11
(a) PDC, CH_2Cl_2; (b) L-Selectride, THF, −78 °C; (c) BnBr, NaH, DMF; (d) EtSH, $BF_3 \cdot Et_2O$; (e) $Pb(OAc)_4$, K_2CO_3, toluene

105. After *O*-benzylation of **105**, the resulting **106** was subjected to cyclic-to-acyclic conversion by treatment with ethanethiol and $BF_3 \cdot Et_2O$, giving diol **107**. This diol was cleaved with $Pb(OAc)_4$ to afford aldehyde **103**, which was transformed to **102** by several steps.

2.5 Lactacystin

Lactacystin (**108**) is the first nonprotein neurotrophic factor (*Fig. 9*). Chida and coworkers achieved the total synthesis of lactacystin (**108**), whose central core **109** was prepared from 1,2:5,6-di-*O*-isopropylidene-α-D-glucofuranose (**2**) [61,62].

1,2:5,6-Di-*O*-isopropylidene-α-D-glucofuranose (**2**) was oxidized with PCC, and the resulting ketone **110** was subjected to Wittig methylenation to afford **111** (*Scheme 12*). Hydrogenation of **111** stereoselectively gave **112**. After selective de-*O*-isopropylidenation of **112** with aqueous AcOH, the resulting diol was transformed into **114** through **113**. This was subjected

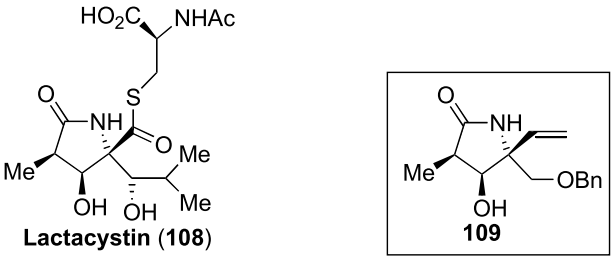
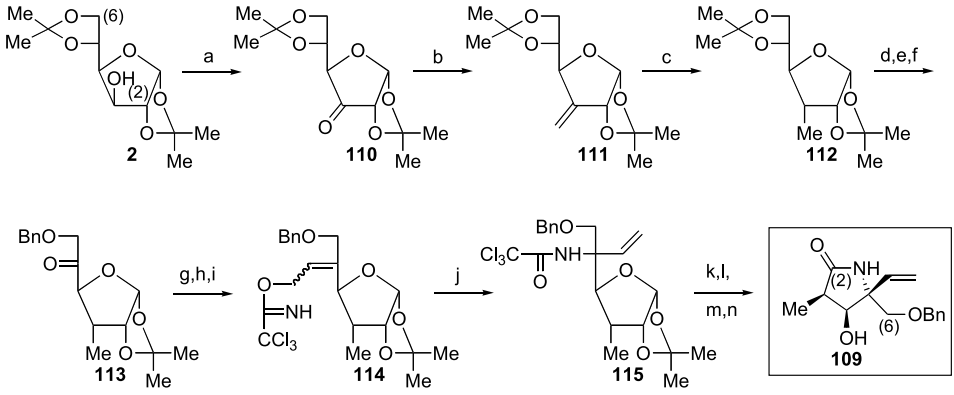


Figure 9
Segment of Chida synthesis of lactacystin

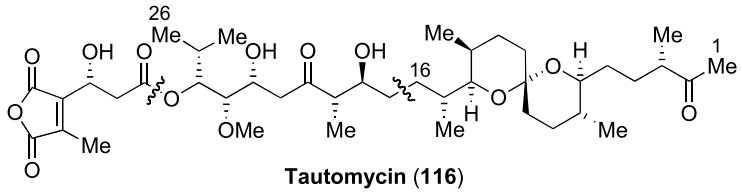


Scheme 12
(a) PCC, CH_2Cl_2; (b) Ph_3PCH_3Br, BuLi, THF; (c) H_2, Pd/C, MeOH; (d) aqueous AcOH, 60 °C; (e) Bu_2SnO, toluene, heat, then BnBr, CsF, DMF; (f) CrO_3, H_2SO_4, aqueous acetone, 0 °C; (g) $Ph_3P{=}CHCO_2Et$, toluene, 60 °C; (h) DIBALH, CH_2Cl_2, −15 °C; (i)CCl_3CN, NaH, ether; (j) toluene, sealed tube, 140 °C; (k) aqueous TFA, 0 °C; (l) $NaIO_4$, aqueous MeOH; (m) CrO_3, H_2SO_4, aqueous acetone, 5 °C; (n) $NaBH_4$, MeOH

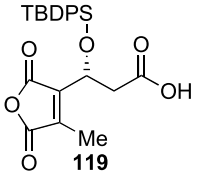
to the Overman rearrangement, affording **115**. Subsequent 4-step transformation of **115** gave the lactam portion **109** of lactacystin (**108**).
The groups of Corey [63,64,65,66], Omura and Smith [67,68], and Baldwin [69] have reported the total synthesis of lactacystin (**108**); the starting materials of their syntheses are amino acids. The groups of Panek [70], Hatakeyama [71], Kanai and Shibasaki [72], and Jacobsen [73] have also reported the total synthesis of lactacystin (**108**).

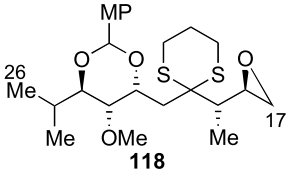
2.6 Tautomycin

Tautomycin (**116**) is an inhibitor of protein phosphatases (*Fig. 10*). The first total synthesis of tautomycin (**116**) was reported by Oikawa, Ichihara, and coworkers [74,75]. The second total synthesis of tautomycin (**116**) was reported by Isobe's group [76,77,78,79,80]. In Isobe's synthesis, tautomycin (**116**) was divided into three segments, **117** (C1–C16), **118** (C17–C26), and **119**, of which **117** was prepared from the monosaccharide derivatives.
The C1–C16 segment **117** was further divided into two subsegments, **120** and **121** (*Fig. 11*). The C1–C8 subsegment **120** was synthesized from levoglucosenone (**122**) (*Scheme 13*). Levoglucosenone (**122**) was subjected to conjugate addition of lithium methylcyanocuprate to give the axial methyl adduct **123**. Treatment of **123** with hydrazine and subsequent eliminative Wolff–Kishner reduction and acetylation afforded **124**. Formation of the ethyl glycoside of **124** followed by treatment of the acetylene derivative in the presence of $BF_3 \cdot Et_2O$ afforded the α-*C*-glycoside **125**. Epimerization of the anomeric center was achieved via the dicobalthexacarbonyl complex, giving **126**. Hydrosilylation of **126** followed by the 3-step conversion gave



Tautomycin (116)





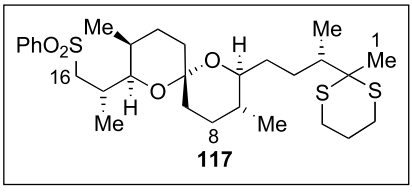


Figure 10
Segments of Isobe synthesis of tautomycin

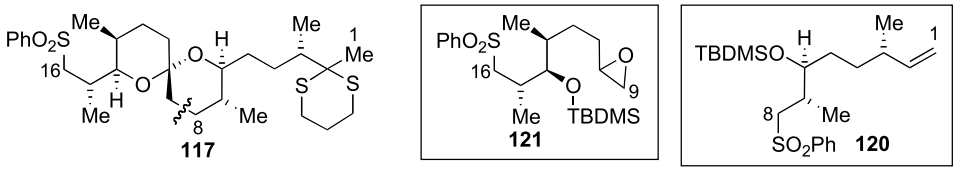
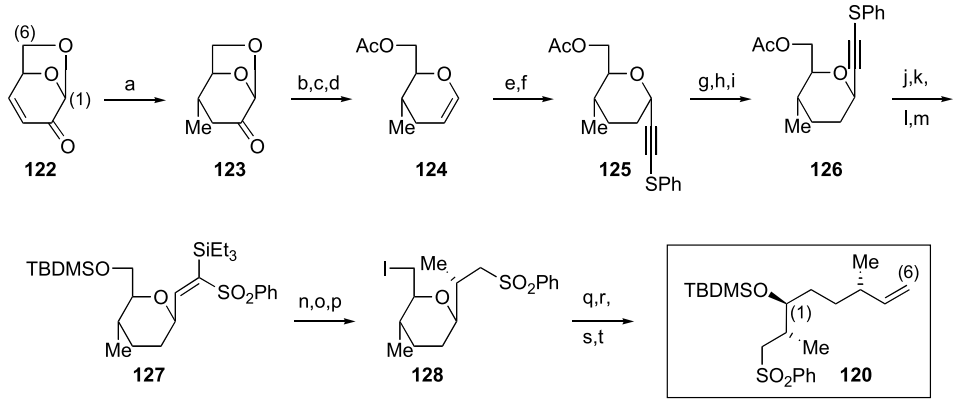
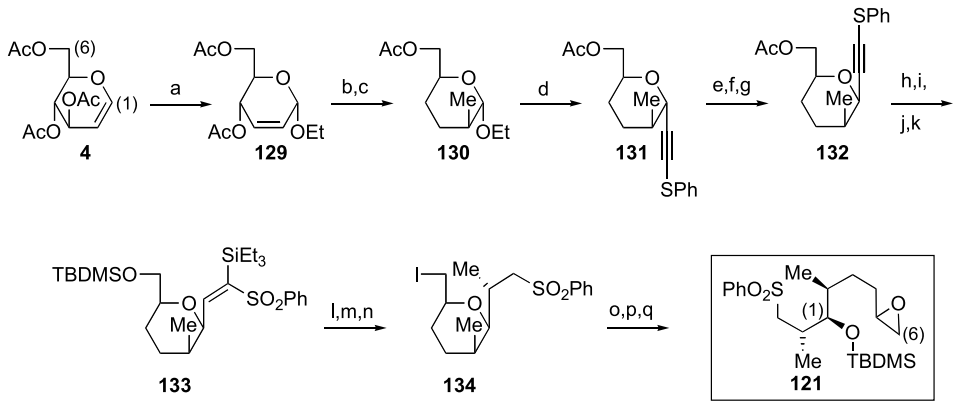


Figure 11
Subsegments of 117



Scheme 13
(a) CuCN, MeLi·LiBr, THF, 0 °C; (b) $H_2NNH_2 \cdot H_2O$, EtOH; (c) NaH, DMSO; (d) Ac_2O, py; (e) EtOH, TsOH, CH_2Cl_2; (f) $Me_3SiCCSPh$, $BF_3 \cdot Et_2O$, MeCN, 0 °C; (g) $Co_2(CO)_8$, CH_2Cl_2, ; (h) TfOH, CH_2Cl_2; (i) I_2, $NaHCO_3$, THF; (j) Et_3SiH, Na_2PtCl_6, Et_3N, BuOH, 110 °C; (k) NaOMe/MeOH, 0 °C; (l) TBDMSCl, imidazole, DMF; (m) *m*-CPBA, Na_2HPO_4, CH_2Cl_2, 0 °C; (n) MeLi · LiBr, hexane-ether, −78 °C; (o) TBAF, THF; (p) $(PhO)_3PMeI$, DMF; (q) Zn, py, EtOH, 70 °C; (r) Ph_3P, $PhCO_2H$, DEAD, THF; (s) NaOMe/MeOH; (t) TBDMSOTf, lutidine, CH_2Cl_2



Scheme 14
(a) EtOH, $BF_3 \cdot Et_2O$; (b) CuCN, MeLi · LiBr, ether, 0 °C; (c) H_2, Pt/C, EtOAc; (d) $Me_3SiCCSPh$, $BF_3 \cdot Et_2O$, MeCN, 0 °C; (e) $Co_2(CO)_8$, CH_2Cl_2, ; (f) TfOH, CH_2Cl_2; (g) I_2, $NaHCO_3$, THF; (h) Et_3SiH, Na_2PtCl_6, Et_3N, BuOH, 110 °C; (i) NaOMe/MeOH, 0 °C; (j) TBDMSCl, imidazole, DMF; (k) *m*-CPBA, Na_2HPO_4, CH_2Cl_2; (l) MeLi · LiBr, hexane-ether, −78 °C; (m) TBAF, THF; (n) $(PhO)_3PMeI$, DMF; (o) Zn, py, EtOH, 70 °C; (p) TBDMSOTf, lutidine, CH_2Cl_2; (q) *m*-CPBA, Na_2HPO_4, CH_2Cl_2

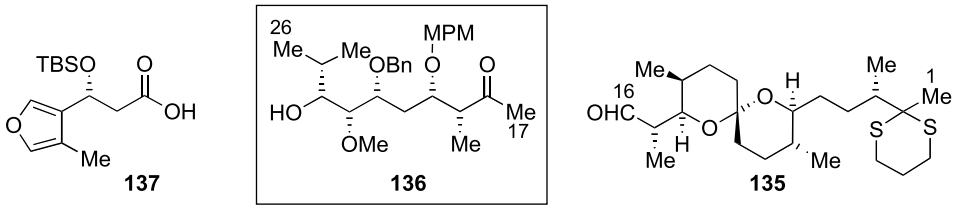


Figure 12
Segments of Shibasaki synthesis of tautomycin

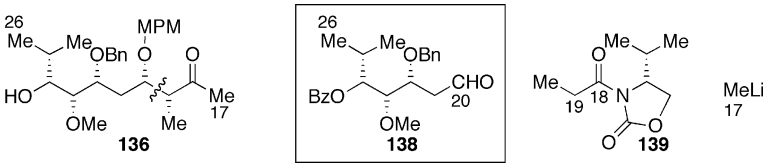


Figure 13
Subsegments of 136

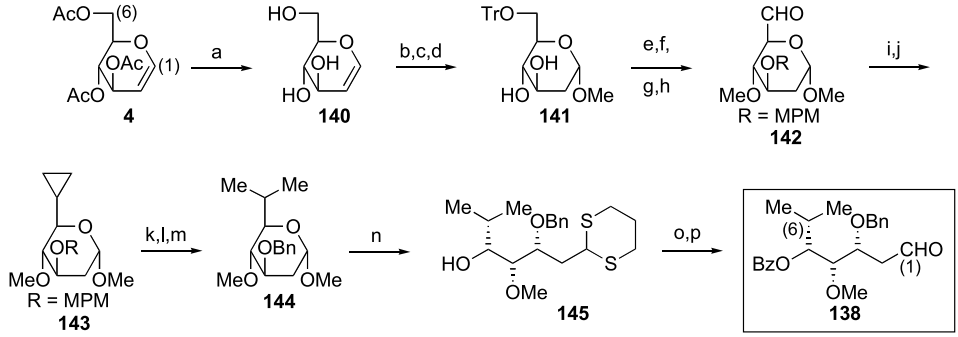
127. Diastereoselective conjugate addition of the MeLi · LiBr complex to **127** and the subsequent 2-step conversion gave **128**. The Vasella-type reductive ring-opening of **128** with Zn and the subsequent 3-step conversion gave the C1–C8 segment **120**.

The C9–C16 subsegment **121** was prepared from tri-*O*-acetyl-D-glucal (**4**) (*Scheme 14*). The Ferrier-type *O*-glycosidation of **4** with EtOH and $BF_3 \cdot Et_2O$ gave **129**. Treatment of **129** with lithium methylcyanocuprate followed by hydrogenation afforded **130**. Treatment of **130** with the acetylene derivative in the presence of $BF_3 \cdot Et_2O$ afforded the α-*C*-glycoside **131**. Epimerization of the anomeric center and the following transformation to the target was achieved via a conceptually similar route to the synthesis of the C1–C8 subsegment **120**, providing **121**.

Shibasaki and coworkers synthesized tautomycin (**116**) from three segments, **135** (C1–C16), **136** (C17–C26), and **137**, of which **136** was prepared from a monosaccharide derivative [81,82] (*Fig. 12*).

The C17–C26 segment **136** was further divided into three subsegments, **138**, **139**, and MeLi (*Fig. 13*), while the C20–C26 subsegment **138** was prepared from tri-*O*-acetyl-D-glucal (**4**) (*Scheme 15*). The acetyl groups in **4** were cleaved by methanolysis, and the resulting triol **140** was subjected to methoxymerculation and tritylation to afford **141**, which had been used as the starting material for the synthesis of *N*-methylmaysenine by Corey and coworkers [83]. Selective *O*-methoxyphenylmethylation of the C3-hydroxy group and the successive *O*-methylation, de-*O*-tritylation, and Swern oxidation afforded aldehyde **142**. The terminal isopropyl group was synthesized by cyclopropanation and reductive opening. Cyclic-to-acyclic conversion of the resulting **144** was realized by treatment with 1,3-propanedithiol and $BF_3 \cdot Et_2O$, giving **145**. Finally, benzoylation and dedithioacetalization of **145** afforded **138**.

Another total synthesis of tautomycin (**116**) by Chamberlin [84], a formal total synthesis by Marshall [85], and other synthetic studies have been reported [86,87].



Scheme 15

(a) NaOMe/MeOH; (b) $Hg(OAc)_2$, MeOH; (c) NaCl, MeOH, then $NaBH_4$, *i*-PrOH, 0 °C; (d) TrCl, py; (e) Bu_2SnO, toluene, heat, then MPMBr, CsF, DMF; (f) MeI, NaH, THF; (g) HCO_2H, ether, THF; (h) Swern oxidation; (i) Zn, CH_2I_2, Me_3Al, THF, 0 °C; (j) CH_2N_2, $Pd(OAc)_2$, ether, 0 °C; (k) DDQ, aqueous CH_2Cl_2; (l) H_2, PtO_2, AcOH; (m) BnBr, NaH, THF, DMF; (n) $HS(CH_2)_3SH$, $BF_3 \cdot Et_2O$, CH_2Cl_2; (o) BzCl, DMAP, py; (p) NBS, aqueous acetone, −23 °C

2.7 FK-506

FK-506 (**146**) is a clinically important immunosuppressive agent (*Fig. 14*). The first total synthesis of FK-506 (**146**) was achieved by the Merck group [88,89] using the Evans aldol technology. Danishefsky and coworkers also synthesized FK-506 (**146**) (formal total synthesis) from four segments, **147**, **148**, **149**, and **150** [90,91,92,93].

The C22–C27 segment **149** was prepared from tri-*O*-acetyl-D-galactal (**151**) (*Scheme 16*). Treatment of **151** with methanol in the presence of $SnCl_4$ afforded **152** (Ferrier-type *O*-glycosidation). Deacetylation of **152** followed by monosilylation provided **153**. Hydroxy-direct-

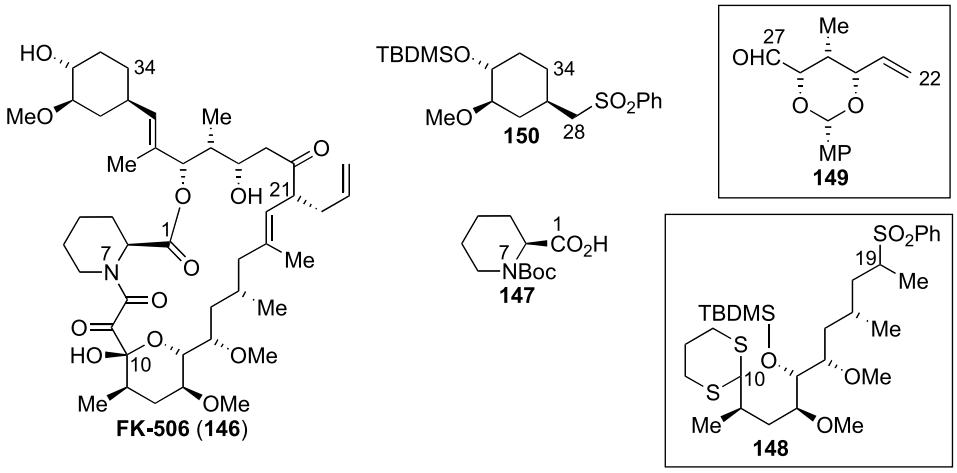
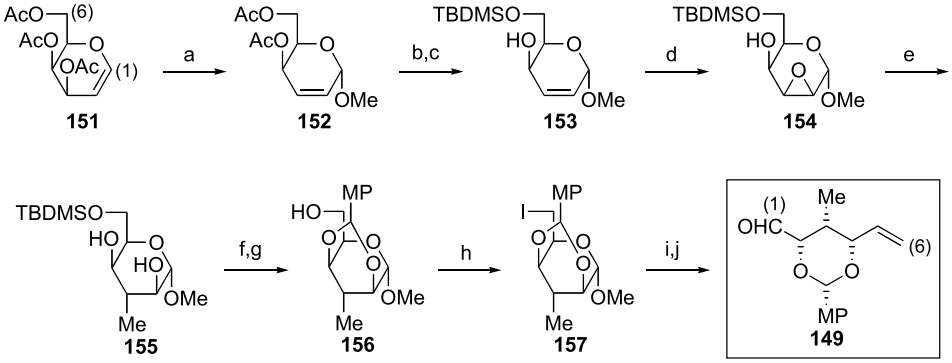


Figure 14

Segments of Danishefsky synthesis of FK-506



Scheme 16
(a) MeOH, $SnCl_4$, CH_2ClCH_2Cl; (b) NaOMe/MeOH; (c) TBDMSCl, Et_3N, DMAP, CH_2Cl_2; (d) *m*-CPBA, CH_2Cl_2, heat; (e) $Me_2CuCNLi_2$, ether, −78 to 0 °C; (f) $MPCH(OMe)_2$, CSA, benzene, heat; (g) TBAF, THF; (h) I_2, Ph_3P, py, benzene, 45 °C; (i) Zn, py, aqueous EtOH, heat; (j) PPTS, CH_2Cl_2

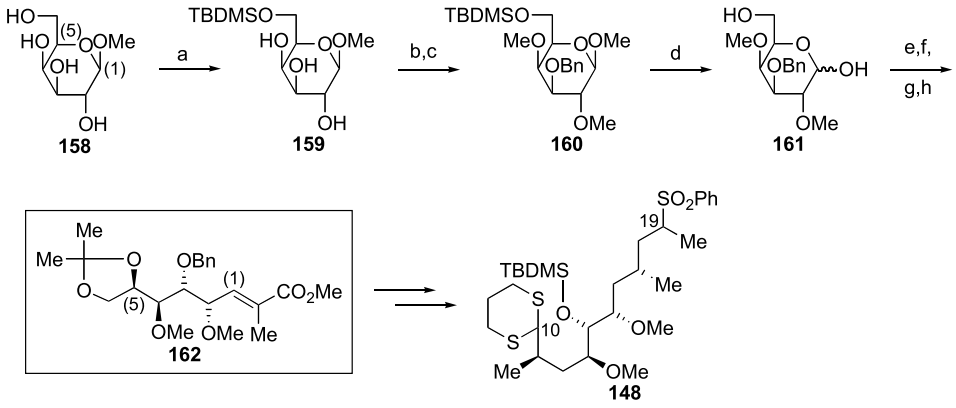
ed epoxidation of **153** with *m*-CPBA afforded **154** stereoselectively. Reaction of **154** with $Me_2CuCNLi_2$ afforded **155**, the product of *trans*-diaxial opening. Reaction of **155** with *p*-anisaldehyde dimethyl acetal in the presence of CSA followed by desilylation gave **156**. After iodination of **156**, the pyranose ring of the resulting **157** was opened through Vasella-type fragmentation and subsequent PPTS treatment for equilibrium of the methoxybenzylidene group to provide the C22–C27 segment **149**.

The C10–C19 segment **148** was prepared from methyl β-D-galactopyranoside (**158**) (*Scheme 17*). It was converted to its mono-TBDMS derivative **159**, and then via stannylation and monobenzylation to the C3-monobenzyl ether, which was methylated to afford **160**. Cleavage of both the silyl and methyl glycoside ethers of **160** afforded the free monosaccharide **161**. This was subjected to reductive ring-opening with $NaBH_4$. The resulting triol was converted to its isopropylidene derivative followed by Swern oxidation and Wittig olefination, giving **162**. This was converted to the C10–C19 segment **148** via several steps.

In the total synthesis of FK-506 (**146**), Schreiber and coworkers prepared the C10–C19 segment **163** from arabitol (**164**) via the two-directional chain synthesis strategy [94,95] (*Scheme 18*). A reaction with the Moffatt reagent was followed by acetate hydrolysis to provide a bis(epoxide) that was silylated to give **165**. Treatment of **165** with the lithium anion of ethoxyacetylene in the presence of $BF_3 \cdot Et_2O$ furnished, after workup with an acid, the bis(lactone) **166**. Methylation of the dianion of **166** proceeded with high diastereofacial selectivity to provide **167**. The corresponding benzyl ether was subjected to hydrolysis and exhaustive methylation to give the bis(methyl ester) **168**. This was transformed to the C10–C19 segment **163** as shown in *Scheme 18*.

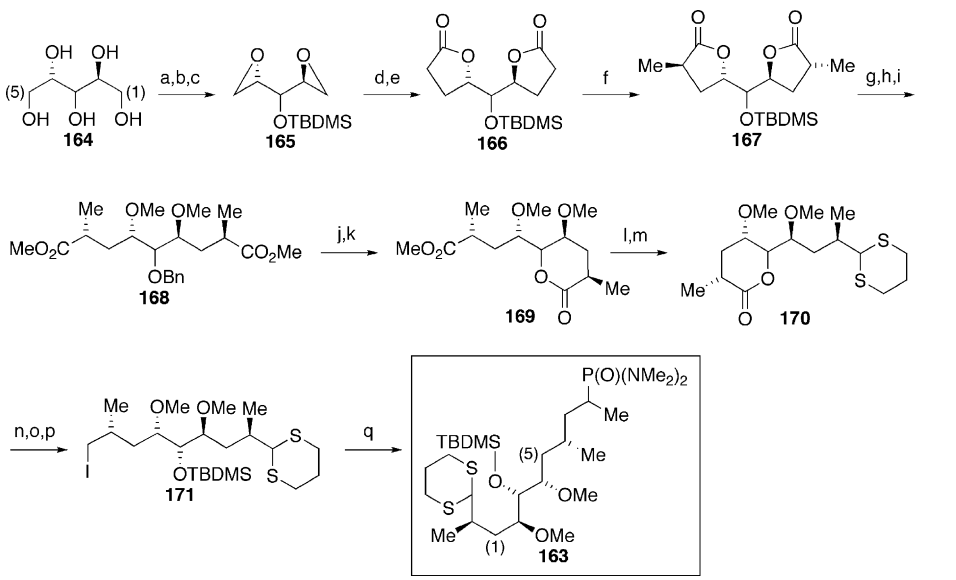
Ireland and coworkers in their total synthesis of FK-506 (**146**) synthesized the C8–C15 segment **172** and the C23–C26 segment **173** [96,97,98] (*Fig. 15*) from appropriate monosaccharide derivatives.

The C8–C15 segment **172** was prepared from **25** (*Scheme 19*). Di-*O*-tosylation of **25** followed by a base treatment provided *allo*-epoxide **174**. Treatment of **174** with Me_2CuLi



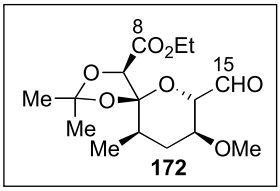
Scheme 17

(a) TBDMSCl, Et_3N, DMAP, CH_2Cl_2; (b) $(Bu_3Sn)_2O$, toluene, heat, then BnBr, Bu_4NI, 80 °C; (c) MeI, NaH, THF; (d) HCl, aqueous THF, heat; (e) $NaBH_4$, EtOH; (f) acetone, TsOH; (g) Swern oxidation; (h) $Ph_3P{=}C(Me)CO_2Me$, CH_2Cl_2



Scheme 18

(a) Acetoxyisobutyric acid chloride, MeCN; (b) NaOMe, THF; (c) TBDMSCl, NaH, THF; (d) ethoxyacetylene, BuLi, $BF_3 \cdot Et_2O$, THF, −78 °C; (e) $HgCl_2$, TsOH, EtOH, heat; (f) MeI, LDA, THF; (g) HF, aqueous MeCN; (h) $BnOC(=NH)CCl_3$, TfOH, CH_2Cl_2, cyclohexane; (i) NaOH, aqueous THF, MeOH, then MeI, NaH, DMF; (j) H_2, $Pd(OH)_2$, EtOAc; (k) PPTS, CH_2Cl_2; (l) L-Selectride, THF, −78 °C; (m) $HS(CH_2)_3SH$, $BF_3 \cdot Et_2O$, CH_2Cl_2; (n) $LiAlH_4$, THF, 0 °C; (o) I_2, Ph_3P, py, benzene ; (p) TBDMSOTf, Et_3N, CH_2Cl_2, 0 °C; (q) ethyl bis(dimethylamino)phosphonamide, BuLi, THF, −78 to 0 °C



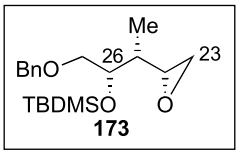
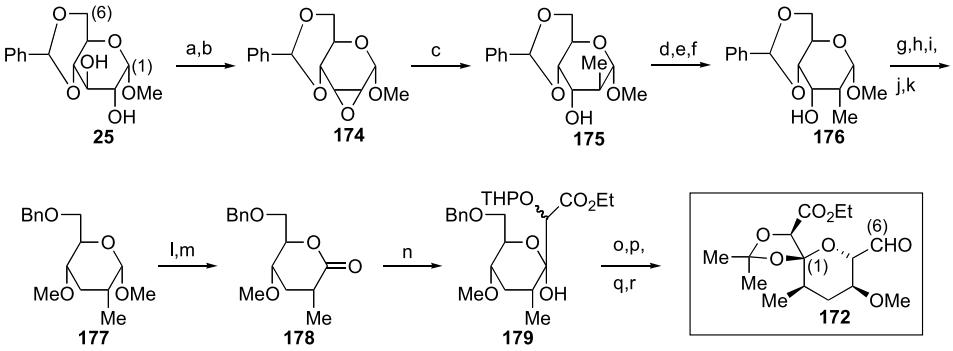


Figure 15
Segments of Ireland synthesis of FK-506

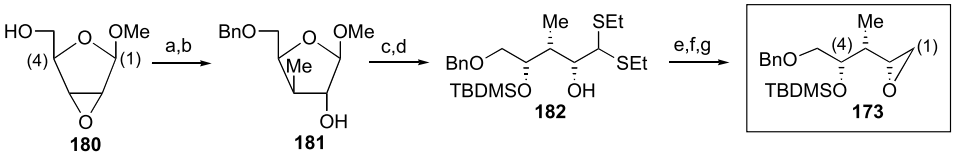


Scheme 19
(a) TsCl, py; (b) NaOMe/MeOH; (c) Me_2CuLi, ether, 0 °C; (d) DMSO, Ac_2O; (e) NaOMe/MeOH; (f) $NaBH_4$; (g) $Me_2NP(O)Cl_2$, BuLi, TMEDA, DME; (h) Me_2NH; (i) Li, $EtNH_2$, EtOH, THF, 10 °C; (j) BnBr, NaH, THF, 0 °C; (k) MeI, NaH, THF; (l) H_2SO_4, aqueous AcOH, 90 °C; (m) DMSO, Ac_2O; (n) $THPOCH_2CO_2Et$, LDA, THF, −78 °C; (o) PPTS, aqueous THF, 55 °C; (p) acetone, P_2O_5; (q) de-*O*-benzylation; (r) Dess–Martin oxidation

gave **175**, the *trans*-diaxial opening product. Oxidation of the C3-hydroxy group followed by epimerization of the C2-axial methyl group and stereoselective reduction with $NaBH_4$ provided **176**. After formation of the C3-phosphorodiamidate of **176**, both the C3 oxygen function and the benzylidene protecting group were removed by reduction with lithium. Selective benzylation followed by methylation gave **177**. The anomeric methyl glycoside was cleaved, and the resulting free monosaccharide was oxidized to afford lactone **178**. The lithium anion of tetrahydropyranyl ethyl glycolate was added to **178**, providing **179**. The 4-step manipulation of **179** gave the C8–C15 segment **172**.

The C23–C26 segment **173** was prepared from **180**, which was derived from ribose (*Scheme 20*). After *O*-benzylation of **180**, the resulting benzyl ether was treated with MeMgCl and $CuBr \cdot Me_2S$ to afford **181**. Thiol acetal formation followed by selective silylation provided **182**. The dithioacetal was cleaved and the resulting aldehyde was reduced with $NaBH_4$ to afford a diol, which was subjected to direct epoxidation to provide **173**.

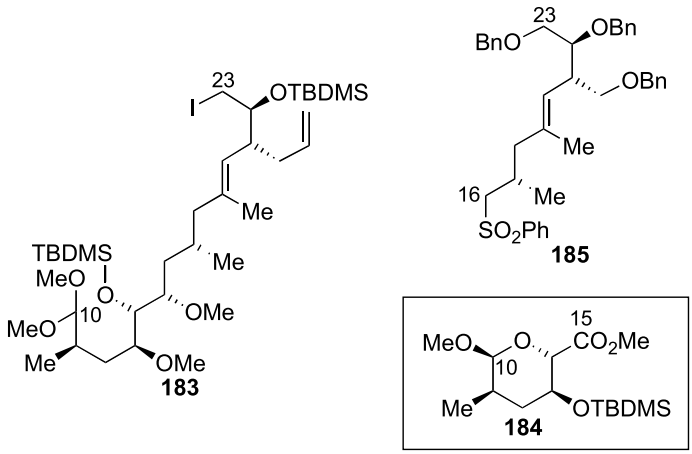
Smith and coworkers have used monosaccharide in their formal total synthesis of FK-506 (**146**) (*Fig. 16*) [99,100,101]. The C10–C23 segment **183** was prepared from two subsegments, **184** (C10–C15) and **185** (C16–C23), of which **184** was derived from the monosaccharide derivative.



180 181 182 173

Scheme 20

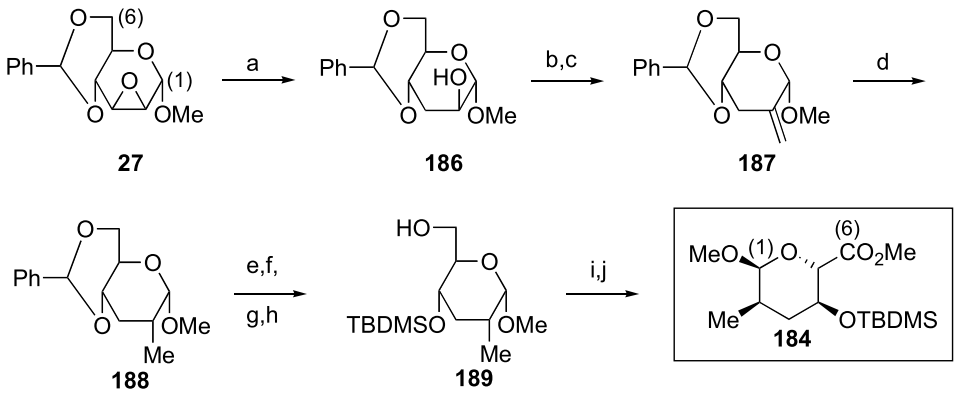
(a) *O*-benzylation; (b) MeMgCl, CuBr · Me_2S, THF; (c) EtSH, HCl; (d) TBDMSCl, imidazole, DMAP, DMF; (e) $HgCl_2$, $CaCO_3$, aqueous MeCN; (f) $NaBH_4$, aqueous THF; (g) Ts imidazole, NaH, THF



183 185 184

Figure 16

Segments of Smith synthesis of FK-506 and subsegment of 183



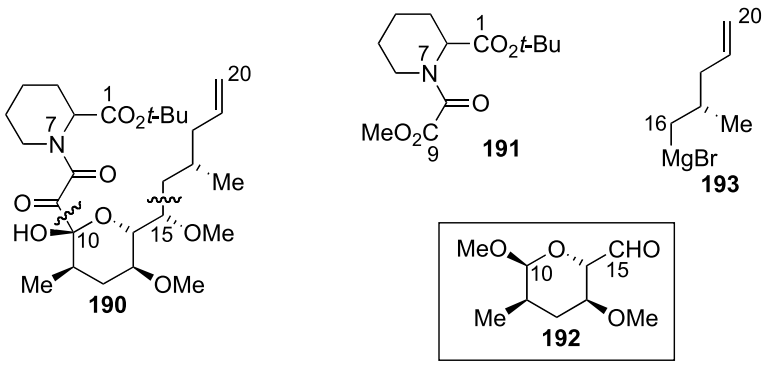
27 186 187 188 189 184

Scheme 21

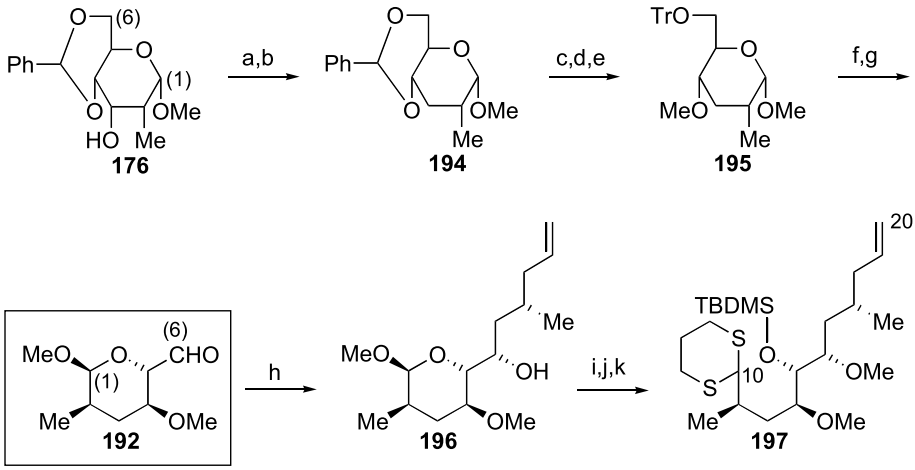
(a) $LiAlH_4$; (b) DMSO, Ac_2O; (c) Ph_3PCH_3Br, BuLi, DME; (d) H_2, Pd/C, EtOH; (e) H_2, $Pd(OH)_2$, EtOH; (f) $PhCO_2H$, Ph_3P, DIAD, THF; (g) TBDMSCl, imidazole, DMAP, DMF, 70 °C; (h) DIBALH, CH_2Cl_2, −78 °C; (i) $RuCl_3$, $NaIO_4$, MeCN, CCl_4, H_2O; (j) MeI, K_2CO_3, DMF

Epoxide **27** was reduced with $LiAlH_4$ to provide **186**, the *trans* diaxial-opening product (◉ *Scheme 21*). Oxidation of **186** followed by Wittig methylenation afforded **187** [102,103]. Stereoselective hydrogenation of **187** gave the equatorial methylated product **188**. Protecting group manipulation of **188** gave **189**, whose primary alcohol was oxidized to carboxylic acid and then esterified to afford **184**.

Rama Rao and coworkers have used a monosaccharide in their synthetic studies on FK-506 (**146**) (◉ *Fig. 17*) [104,105]. They divided the C1–C20 segment **190** into three subsegments, **191**, **192**, and **193**, of which the C10–C15 subsegment **192** was prepared from the monosaccharide derivative.



◘ Figure 17
Segments of Rama Rao synthesis of FK-506 and subsegments of 190



◘ Scheme 22
(a) NaH, CS_2, ether, heat, then MeI; (b) Bu_3SnH, AIBN, toluene, heat; (c) TsOH, MeOH; (d) TrCl, py, 60 °C; (e) MeI, NaH, THF, DMF, 0 °C; (f) Li, NH_3; (g) Swern oxidation; (h) 193, ether; (i) MeI, NaH, THF; (j) $HS(CH_2)_3SH$, $BF_3 \cdot Et_2O$, CH_2Cl_2, 0 °C; (k) TBDMSOTf, lutidine, CH_2Cl_2, 0 °C

The Ireland synthetic intermediate **176** was deoxygenated by the Barton procedure to provide **194** (◉ *Scheme 22*). Stepwise deprotection and protection of hydroxy groups in **194** afforded **195**. De-*O*-tritylation and Swern oxidation of **195** gave the C10–C15 segment **192**. As shown in ◉ *Scheme 22*, this aldehyde was transformed into the C10–C20 segment **197**.

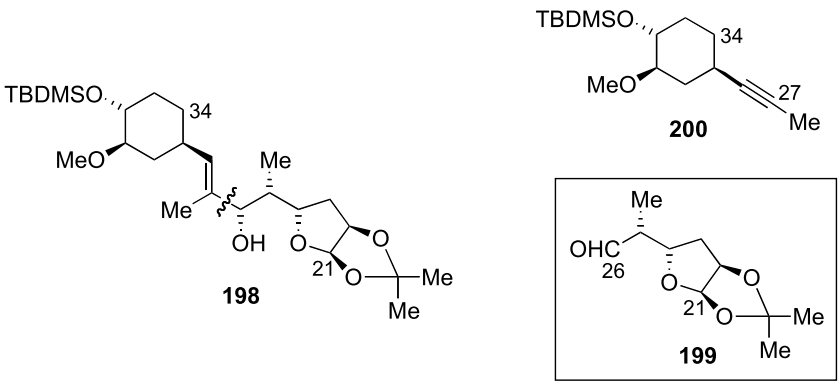
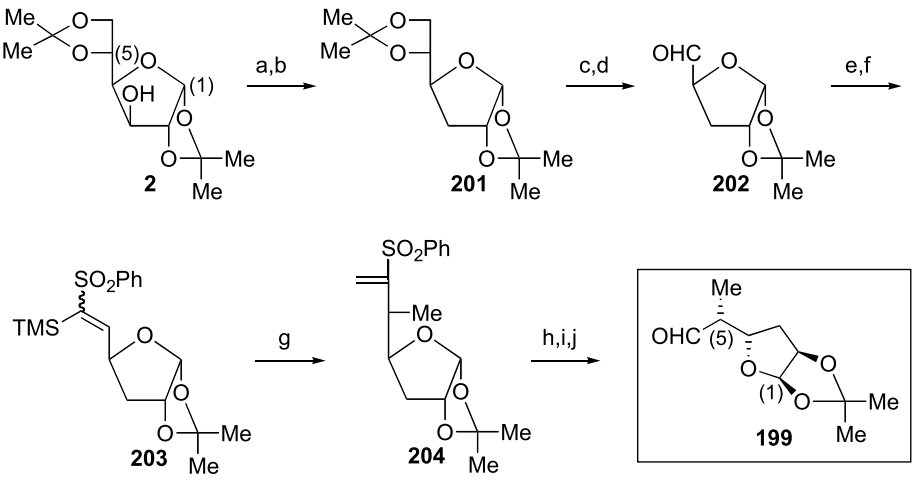


Figure 18
Segments of Maier synthesis of FK-506 and subsegments of 198



Scheme 23
(a) NaH, imidazole, CS_2, MeI, THF; (b) Bu_3SnH, toluene, heat; (c) H_2SO_4, aqueous MeOH; (d) $NaIO_4$, aqueous MeOH; (e) $PhS(TMS)_2CLi$; (f) *m*-CPBA, CH_2Cl_2, 0 °C; (g) MeLi, THF, −78 °C, then $H_2C{=}O$, −78 °C; (h) LDA, THF, −78 °C; (i) H_2, Lindlar catalyst, hexane; (j) O_3, CH_2Cl_2

Maier and coworkers focused on 1,2:5,6-di-*O*-isopropylidene-α-D-glucofuranose (**2**) in their synthetic studies on FK-506 (**146**) (*Fig. 18*) [106]. The C21–C34 segment **198** was prepared from **199** and **200**, of which **199** was prepared from **2**.

The Barton deoxygenation of **2** provided **201**, which was subjected to selective hydrolysis of the 5,6-*O*-isopropylidene group and oxidative scission of the resulting diol function to provide aldehyde **202** (*Scheme 23*). Reaction of aldehyde **202** with bis(trimehylsilyl)phenylthiomethyllithium afforded vinyl sulfide, which was oxidized with *m*-CPBA to the corresponding vinyl sulfone **203**. Addition of methyllithium to **203** followed by quenching of the resulting anion with formaldehyde provided vinyl sulfone **204** as a single isomer. Elimination of phenylsulfinic acid with LDA gave alkyne. Subsequent hydrogenation led to olefin which was then converted to the C21–C26 aldehyde **199**.

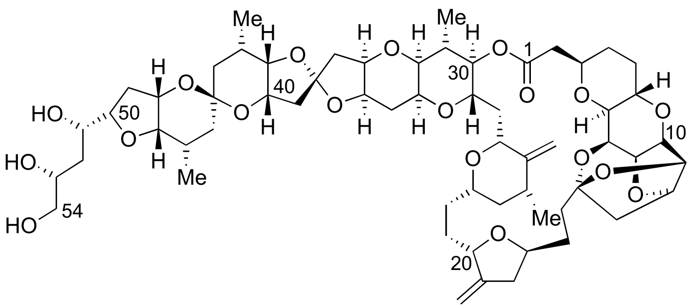
Formal total synthesis of FK-506 (**146**) was reported by Sih and coworkers [107].

2.8 Halichondrin B

Halichondrin B (**205**) has received much attention due to its extraordinary in vitro and in vivo antitumor activity (*Fig. 19*). Recently, structurally simplified macrocyclic ketone analogs of halichondrin B (**205**) were found to retain the potent cell growth inhibitory activity in vitro, stability in mouse serum, and in vivo efficacy of the natural product [108,109,110].

Kishi and coworkers have succeeded in the first total synthesis of halichondrin B (**205**) [111,112,113,114,115,116,117,118,119]. They divided halichondrin B (**205**) into four segments, **206–209** (*Fig. 20*), all of which were prepared from monosaccharides.

The C1–C13 segment **206** of halichondrin B (**205**) has been recently synthesized with improvements from **210** and **211**. The C1–C11 segment **210** was prepared from L-mannonic-γ-lactone (**212**) (*Scheme 24*). Di-*O*-cyclohexylidene protection of **212** followed by DIBALH reduction and Wittig reaction gave **213**. Stereoselective dihydroxylation and acetylation of **213** provided **214**. *C*-Glycosidation of **214** with methyl 3-trimethylsilylpent-4-enoate and $BF_3 \cdot Et_2O$ afforded α-glycoside **215** exclusively. This compound was transformed into **206** via the 7-step manipulation as shown in *Scheme 24*.



Halichondrin B (205)

Figure 19
Halichondrin B

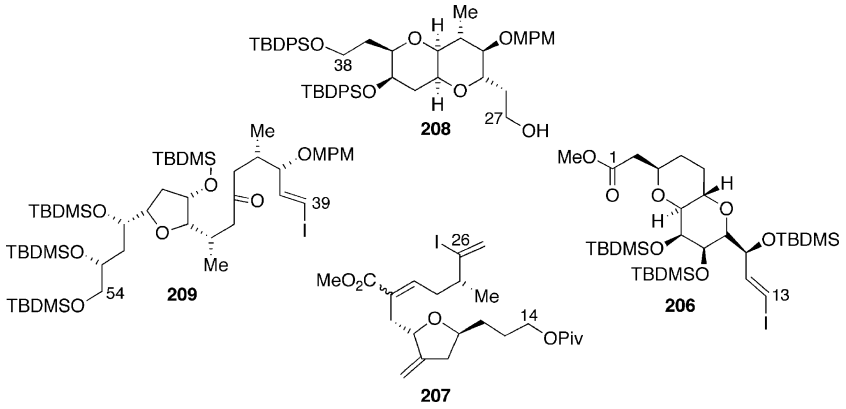
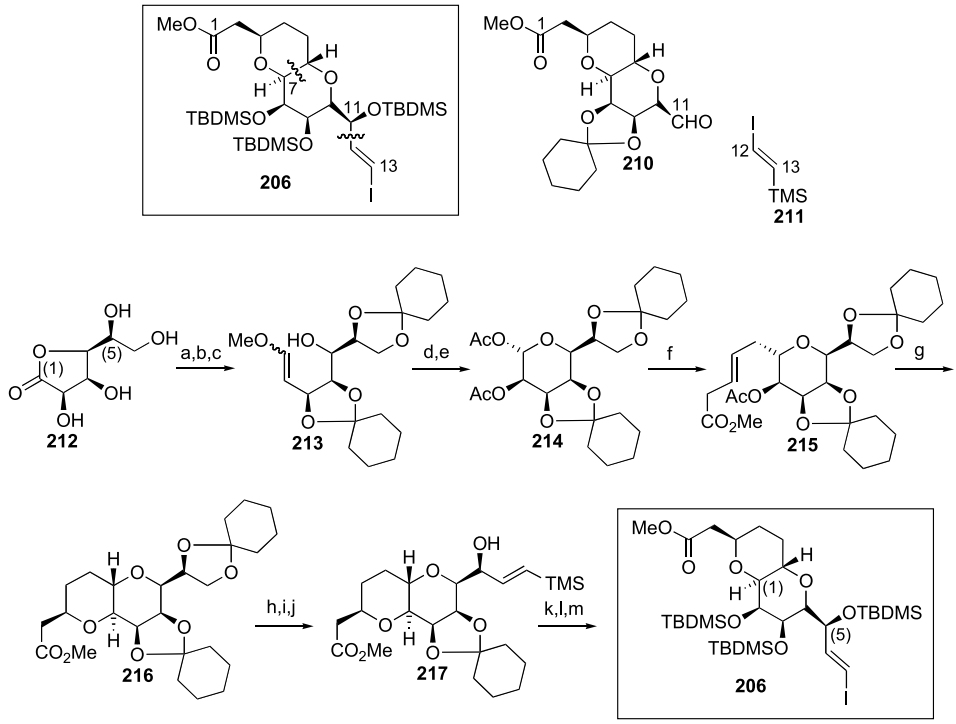
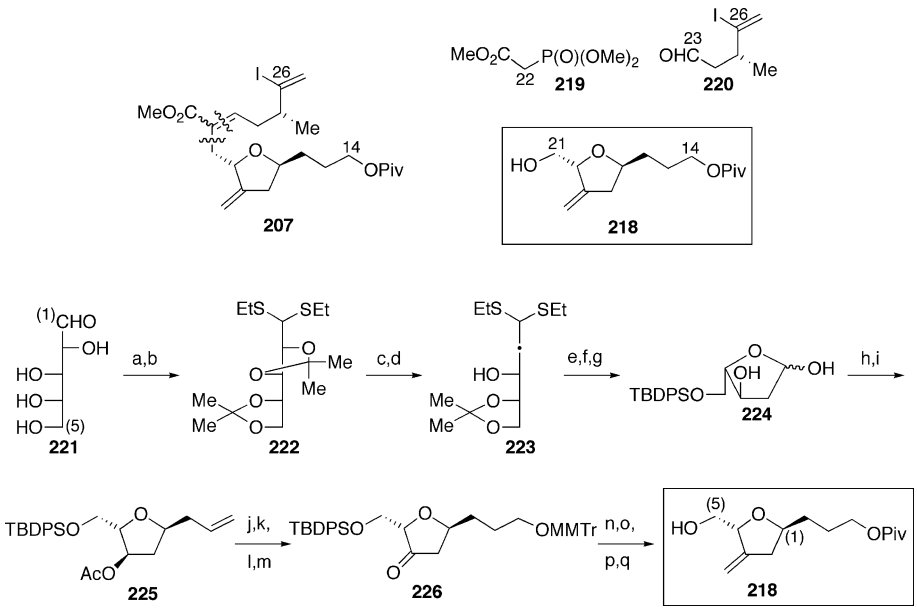


Figure 20
Segments of Kishi synthesis of halichondrin B



Scheme 24
(a) Cyclohexanone, H_2SO_4, toluene; (b) DIBALH, CH_2Cl_2, −78 °C; (c) $Ph_3P(Cl)CH_2OMe$, *t*-BuOK, THF, heat; (d) OsO_4, NMO, DHQD, aqueous acetone, −5 °C; (e) Ac_2O, DMAP, py; (f) methyl 3-trimethylsilylpent-4-enoate, $BF_3 \cdot Et_2O$, MeCN; (g) Triton-B(OMe), THF-EtOAc, 0 °C to rt; (h) aqueous AcOH, 80 °C; (i) $NaIO_4$, THF, pH 7 buffer; (j) 211, $NiCl_2$-$CrCl_2$, DMSO; (k) aqueous AcOH, TFA, 85 °C; (l) TBDMSOTf, lutidine, CH_2Cl_2; (m) NIS, MeCN-$ClCH_2CN$



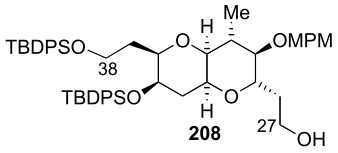
Scheme 25

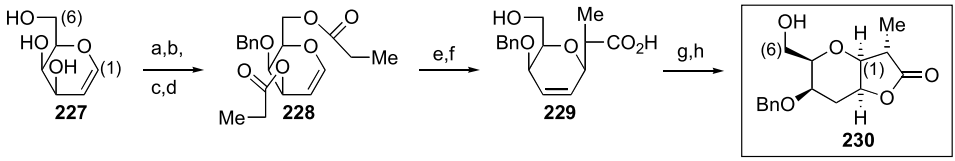
(a) EtSH, HCl; (b) acetone, acid; (c) *t*-BuOK, DMSO, THF; (d) $LiAlH_4$, THF; (e) aqueous AcOH; (f) TBDPSCl, imidazole; (g) I_2, $NaHCO_3$, aq acetone; (h) Ac_2O, py; (i) CH_2=$CHCH_2TMS$, $BF_3 \cdot Et_2O$, MeCN, 0 °C; (j) 9-BBN, then H_2O_2; (k) MMTrCl, Et_3N, CH_2Cl_2; (l) K_2CO_3, MeOH; (m) Swern oxidation; (n) MeOH, PPTS; (o) Tebbe reagent; (p) PivCl, py; (q) TBAF

The C14–C26 segment **207** was prepared from D-arabinose **221** (*Scheme 25*). After treatment of 2,3:4,5-di-*O*-isopropylidene-D-arabinose diethyl dithioacetal (**222**), prepared from D-arabinose **221** by dithioacetalization and isopropylidenation, with *t*-BuOK, the resulting ketene dithioacetal was reduced to give **223** [120]. De-*O*-isopropylidenation of **223** followed by selective silylation and dedithioacetalization provided furanose **224**. The diacetate prepared from **224** was subjected to *C*-glycosidation with allyl trimethylsilane and $BF_3 \cdot Et_2O$ to afford **225**. Further transformation of **225** gave the C14–C21 segment **218**. Sequential treatment of **218** with the C22 and the C23–C26 segments (**219** and **220**, respectively) provided the C14–C26 segment **207**.

The C27–C38 segment **208** was prepared from D-galactal **227** (*Scheme 26*). The silyl ether, prepared from **227**, was selectively benzylated, and the resulting C3-alcohol was desilylated and propanoylated to afford **228**. After the Ireland–Claisen rearrangement of **228**, carboxylic acid **229** was subjected to iodolactonization followed by reductive removal of iodine to give γ-lactone **230**. This was converted to the C27–C38 segment **208**.

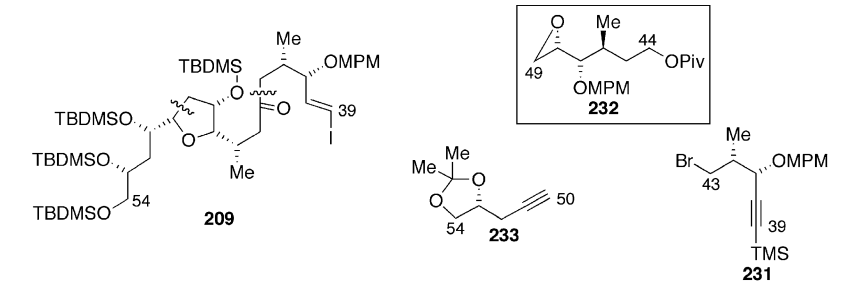
The C39–C 54 segment **209** was prepared from **231**, **232**, and **233**, of which the C44–C49 subsegment **232** was prepared from L-ascorbic acid **234** (*Scheme 27*). The α,β-unsaturated γ-lactone **235** was prepared from L-ascorbic acid **234** [121,122,123]. The conjugate addition of methylcuprate to **235** yielded the single stereoisomer **236**. Routine functional group manipulation allowed the transformation of **236** into epoxide **232**. Sequential elongation of **232** with

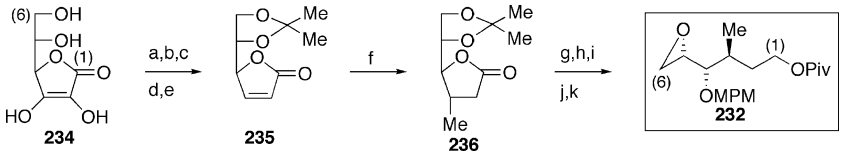




Scheme 26

(a) TBDMSCl, imidazole, DMF; (b) BnBr, NaH, THF, DMF; (c) TBAF, THF; (d) $(EtCO)_2O$, Et_3N, CH_2Cl_2; (e) LHMDS, TBDMSCl, HMPA, THF, −78 to 0 °C, followed by reflux, benzene; (f) NaOH, aqueous THF; (g) I_2, KI, $NaHCO_3$, H_2O; (h) Bu_3SnH, AIBN, benzene, heat





Scheme 27

(a) H_2, Pd/C, H_2O, 5.25 atm, 50 °C; (b) $Me_2C(OMe)_2$, $SnCl_2$, acetone, heat; (c) DMF dimethyl acetal, $CHCl_3$, heat; (d) MeI, 5 °C; (e) MeCN, heat; (f) Me_2CuLi, TMSCl, THF; (g) $LiAlH_4$, ether, 0 °C; (h) PivC, py, CH_2Cl_2; (i) MPMBr, KH, THF; (j) aqueous AcOH; (k) Ts imidazole, NaH, THF

the C50–C54 subsegment **233** and the C39–C43 subsegment **231** provided the C39–C54 segment **209**.

All segments obtained from these preparations were sequentially connected to complete the total synthesis of halichondrin B (**205**).

Salomon and coworkers described the synthetic studies of halichondrins [124,125,126,127]. They prepared three segments, **237**, **238**, and **239** (*Fig. 21*).

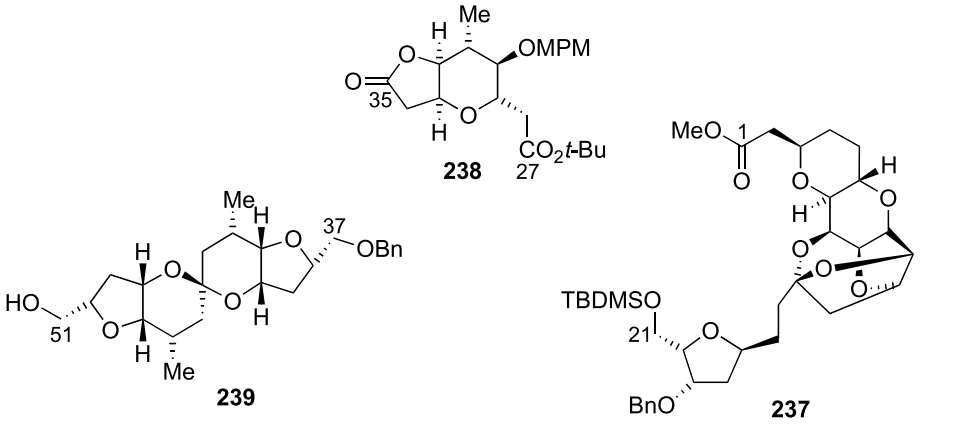


Figure 21
Segments of Salomon synthesis of halichondrin B

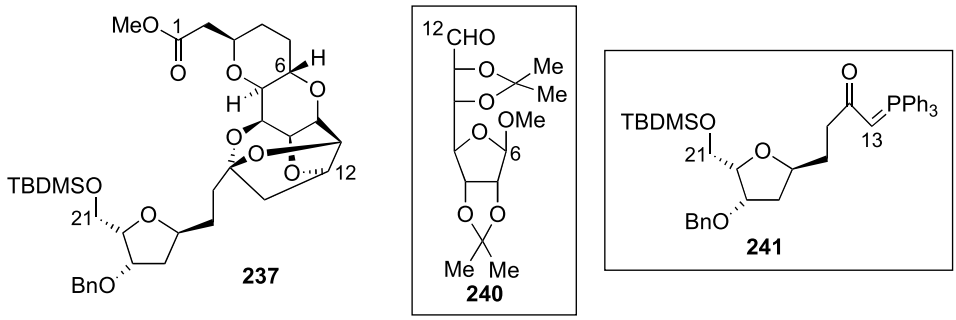
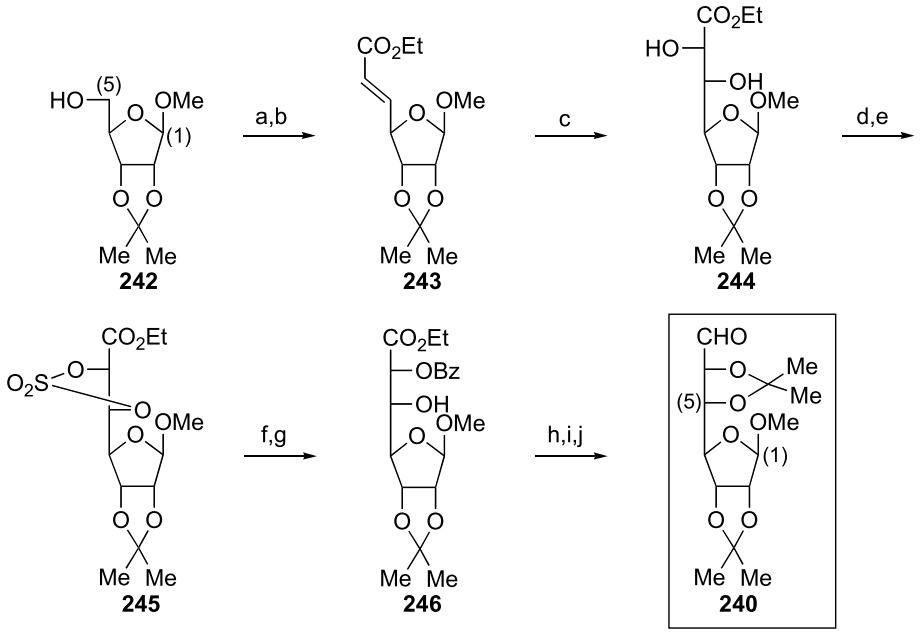


Figure 22
Subsegments of 237

The C1–C21 segment **237** was prepared from two main subsegments, **240** (C6–C12) and **241** (C13–C21) (*Fig. 22*).

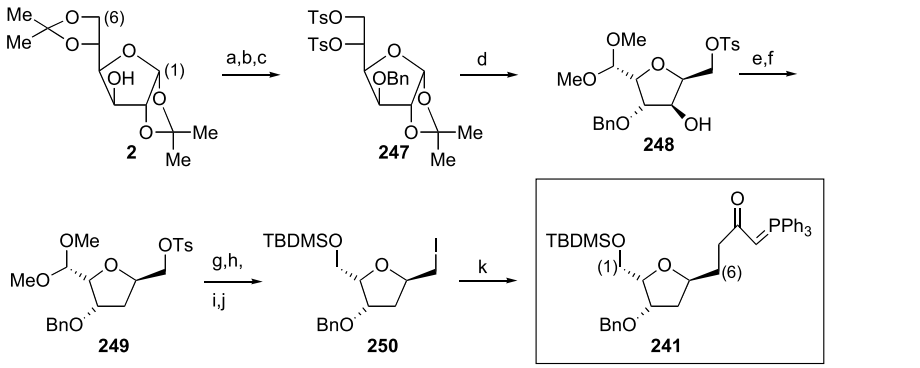
The commercially available ribofuranoside **242** was oxidized, and the resulting aldehyde was subjected to Wittig reaction to afford **243** (*Scheme 28*). OsO_4-catalyzed vicinal hydroxylation of **243** gave **244** as the major product. The configuration at C6 in **244** had to be selectively inverted. This was accomplished by a regioselective nucleophilic substitution of the derived cyclic sulfate **245**. Debenzoylation of benzoate **246**, isopropylidenation of the resulting diol, and partial reduction of the ester with DIBALH delivered aldehyde **240**, the C6–C12 subsegment.

The C13–C21 subsegment **241** was prepared from 1,2:5,6-di-*O*-isopropylidene-α-D-glucofuranose (**2**) (*Scheme 29*). Benzylation, selective deprotection, and tosylation of **2** gave **247**, which was converted to **248** by transacetalization and intramolecular *O*-alkylation. Barton's deoxygenation of **248** gave **249**. Deacetalization of **249** followed by the 3-step transformation provided **250**. Alkylation of **250** with α'-lithioacetonylidenetriphenylphosphorane gave **241**. These two subsegments, **240** and **241**, were coupled and transformed into **237**.



Scheme 28

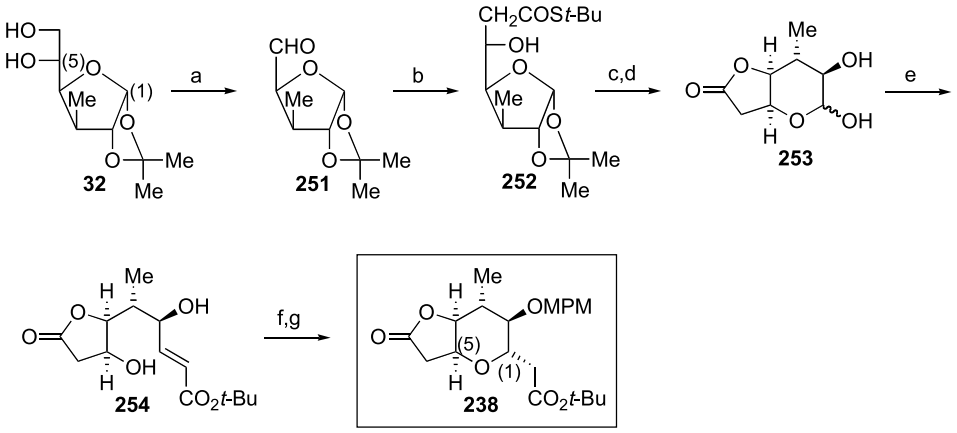
(a) Swern oxidation; (b) $Ph_3P{=}CHCO_2Et$; (c) OsO_4, NMO, DHQD; (d) Cl_2SO, Et_3N, CH_2Cl_2; (e) $RuCl_3$, $NaIO_4$, MeCN, CCl_4, H_2O; (f) Bu_4NPhCO_2, benzene; (g) H_2SO_4, aqueous THF; (h) $Ba(OMe)_2$, MeOH; (i) $Me_2C(OMe)_2$, PPTS, acetone; (j) DIBALH, CH_2Cl_2, −78 °C



Scheme 29

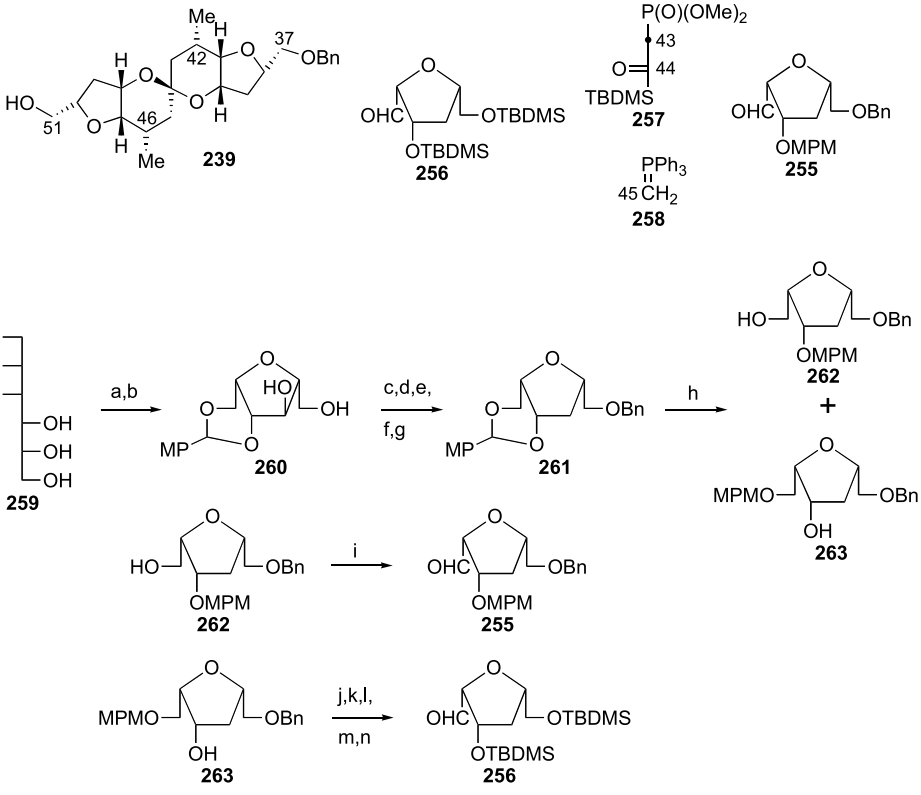
(a) BnBr, NaH, Bu_4NI; (b) aqueous AcOH, 34 °C; (c) TsCl, py; (d) HCl, MeOH; (e) PhO(C=S)Cl, DMAP; (f) Bu_3SnH, AIBN; (g) aqueous TFA; (h) $NaBH_4$; (i) NaI, 2-butanone; (j) TBDMSCl, imidazole; (k) $Ph_3P{=}CHC(O)CH_2Li$

The C27–C35 segment **238** was prepared from **32** (◉ *Scheme 30*), which is the synthetic intermediate of erythronolide A. Oxidative cleavage of **32** gave aldehyde **251**, which was subjected to aldol reaction with TBDMS-enol ether and $TiCl_4$ to provide **252**. Hydrolysis of thioester and deacetalization with TFA were accompanied by furanose-to-pyranose interconversion and



Scheme 30

(a) $NaIO_4$; (b) TBDMS enol ether of *t*-butyl thioacetate, $TiCl_4$, −78 °C; (c) NaOH, aqueous THF; (d) aqueous TFA; (e) $Ph_3P{=}CHCO_2t$-Bu; (f) NaH, THF; (g) MPM imidate, CF_3SO_3H



Scheme 31

(a) HCl; (b) *p*-methoxybenzaldehyde, MeOH; (c) TBDMSCl, imidazole; (d) NaH, imidazole, CS_2, MeI, THF; (e) Bu_3SnH, AIBN; (f) TBAF, THF; (g) BnBr, NaH; (h) DIBALH; (i) Swern oxidation; (j) TBDMSOTf, Et_3N; (k) H_2, Raney Ni, EtOH; (l) TBDMSCl, imidazole, CH_2Cl_2; (m) DDQ, aqueous CH_2Cl_2; (n) Swern oxidation

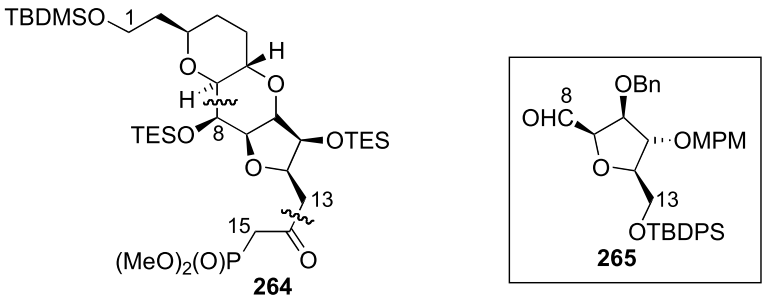
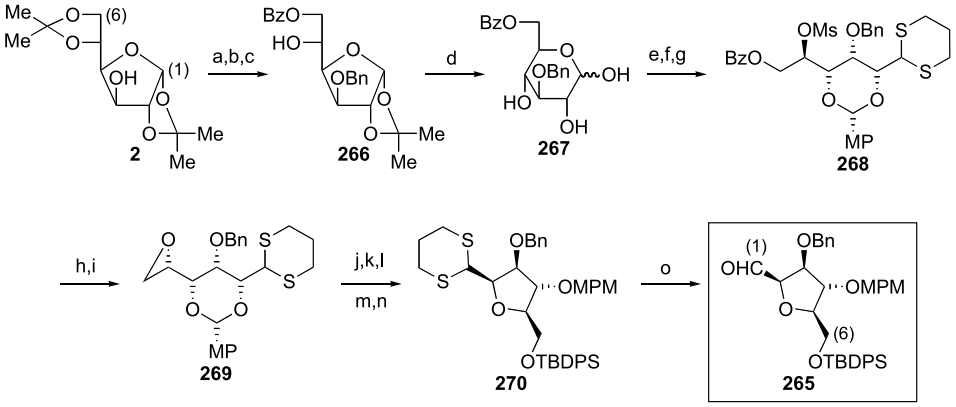


Figure 23
Segment of Horita–Yonemitsu synthesis of halichondrin B and subsegment of 264



Scheme 32
(a) BnCl, NaH, THF, DMSO; (b) H_2SO_4, aqueous MeOH; (c) BzCl, py, CH_2Cl_2, 0 °C; (d) HCl, aqueous THF, heat; (e) $HS(CH_2)_3SH$, $ZnCl_2$, MeCN, 60 °C; (f) $MPCH(OMe)_2$, CSA, CH_2Cl_2; (g) MsCl, Et_3N, CH_2Cl_2, 0 °C; (h) K_2CO_3, MeOH, CH_2Cl_2, 0 °C; (i) *t*-BuOK, THF; (j) AcOH, aqueous CH_2Cl_2; (k) TrCl, DMAP, Et_3N, CH_2Cl_2; (l) MPMCl, NaH, THF, DMSO; (m) CSA, MeOH, CH_2Cl_2; (n) TBDPSCl, imidazole, CH_2Cl_2; (o) HgO, $BF_3 \cdot Et_2O$, aqueous THF, 60 °C

lactonization to provide the *cis*-lactone **253**. Wittig olefination of **253** and heterocyclization of the intermediate α,β-unsaturated ester **254** followed by methoxybenzylation furnished the target lactone **238**.

The C37–C51 segment **239** was prepared from four subsegments, **255**–**258** (◉ *Scheme 31*). Acidic cyclization of D-mannitol (**259**) generated tetraol, which was selectively acetalized to give **260**. Hydrodehydroxylation of the derived monosilyl ether by the Barton protocol and replacement of the TBDMS group by a benzyl group afforded **261**. Reductive cleavage of **261** with DIBALH generated a 6:4 mixture of **262** and **263**; both compounds could be used in the synthesis of the C37–C51 segment **239**. Thus, **262** was oxidized under Swern conditions to give **255**. On the other hand, **263** was converted to **256** by the 5-step transformation. These two subsegments were connected step-by-step to afford the C37–C51 segment **239**.

Horita, Yonemitsu, and coworkers have described synthetic studies of halichondrin B (**205**) [128,129,130,131,132,133,134,135,136]. The C1–C15 segment **264** was prepared from **265** (*Fig. 23*)

Benzylation, selective hydrolysis, and benzoylation of 1,2:5,6-di-*O*-isopropylidene-α-D-glucofuranose (**2**) readily gave **266** (*Scheme 32*). Furanose-to-pyranose interconversion was realized by acidic hydrolysis of **266**, giving **267**. Opening of the pyranose ring with 1,3-propanedithiol was realized in the presence of $ZnCl_2$, providing an acyclic compound, which was treated with 4-methoxybenzaldehyde dimethyl acetal and CSA to give only a thermodynamically stable 1,3-dioxane compound. The remaining hydroxy group was then mesylated to give **268**, which was readily converted to epoxide **269** by removal of the benzoyl group followed by treatment of a base. The 5-step conversion of **269** gave **270**. Dedithioacetalization of **270** provided the C8–C13 subsegment **265**, which was converted to the C1–C15 segment **264** via a several-step manipulation.

Burke and coworkers have described the synthetic studies of halichondrin B (**205**) [137,138,139,140].

References

1. Vasella A (1980) Chiral building blocks in enantiomer synthesis. In: Scheffold R (ed) Modern Synthetic Methods. Otto Salle Verlag GmbH & Co., Frankfurt am Main, p 173
2. Hanessian S (1983) Total Synthesis of Natural Products: the 'Chiron' Approach. Pergamon Press, Oxford
3. Lichtenthaler FW (1992) Enantiopure building blocks from sugars and their utilization in natural product synthesis. In: Scheffold R (ed) Modern Synthetic Methods. VHCA, Basel, p 273
4. Collins PM, Ferrier RJ (1995) Monosaccharides. Wiley, Chichester
5. Bols M (1996) Carbohydrate Building Blocks. Wiley, New York
6. Tatsuta K (1997) J Syn Org Chem Jpn 55:970
7. Hanessian S (1997) (ed) Preparative Carbohydrate Chemistry. Marcel Dekker, New York
8. Woodward RB et al. (1981) J Am Chem Soc 103:3210
9. Woodward RB et al. (1981) J Am Chem Soc 103:3213
10. Woodward RB et al. (1981) J Am Chem Soc 103:3215
11. Corey EJ, Trybulski EJ, Melvin LS Jr., Nicolaou KC, Secrist JA, Lett R, Sheldrake PW, Falck JR, Brunelle DJ, Haslanger MF, Kim S, Yoo S (1978) J Am Chem Soc 100:4618
12. Corey EJ, Kim S, Yoo S, Nicolaou KC, Melvin LS Jr., Brunelle DJ, Falck JR, Trybulski EJ, Lett R, Scheldrake PW (1978) J Am Chem Soc 100:4620
13. Corey EJ, Hopkins PB, Kim S, Yoo S, Nambiar KP, Falck JR (1979) J Am Chem Soc 101:7131
14. Kinoshita M, Ohsawa N, Gomi S (1982) Carbohydr Res 109:5
15. Kinoshita M, Arai M, Tomooka K, Nakata M (1986) Tetrahedron Lett 27:1811
16. Kinoshita M, Arai M, Ohsawa N, Nakata M (1986) Tetrahedron Lett 27:1815
17. Nakata M, Arai M, Tomooka K, Ohsawa N, Kinoshita M (1989) Bull Chem Soc Jpn 62:2618
18. Toshima K, Mukaiyama S, Yoshida T, Tamai T, Tatsuta K (1991) Tetrahedron Lett 32:6155
19. Toshima K, Nozaki Y, Mukaiyama S, Tamai T, Nakata M, Tatsuta K, Kinoshita M (1995) J Am Chem Soc 117:3717
20. Sviridov AF, Ermolenko MS, Yashunsky DV, Borodkin VS, Kochetkov NK (1987) Tetrahedron Lett 28:3835
21. Kochetkov NK, Sviridov AF, Ermolemko MS, Yashunsky DV, Borodkin VS (1989) Tetrahedron 45:5109
22. Nakata M, Akiyama N, Kamata J, Kojima K, Masuda H, Kinoshita M, Tatsuta K (1990) Tetrahedron 46:4629
23. Kochetkov NK, Sviridov AF, Ermolenko MS (1981) Tetrahedron Lett 22:4315
24. Kochetkov NK, Sviridov AF, Ermolenko MS (1981) Tetrahedron Lett 22:4319

25. Kochetkov NK, Sviridov AF, Ermolenko MS, Yashunsky DV (1984) Tetrahedron Lett 25:1605
26. Sviridov AF, Ermolenko MS, Yashunsky DV, Borodkin VS, Kochetkov NK (1987) Tetrahedron Lett 28:3839
27. Sviridov AF, Borodkin VS, Ermolenko MS, Yashunsky DV, Kochetkov NK (1991) Tetrahedron 47:2291
28. Sviridov AF, Borodkin VS, Ermolenko MS, Yashunsky DV, Kochetkov NK (1991) Tetrahedron 47:2317
29. Hikota M, Tone H, Horita K, Yonemitsu O (1990) J Org Chem 55:7
30. Hikota M, Tone H, Horita K, Yonemitsu O (1990) Tetrahedron 46:4613
31. Martin SF, Hida T, Kym PR, Loft M, Hodgson A (1997) J Am Chem Soc 119:3193
32. Hergenrother PJ, Hodgson A, Judd AS, Lee WC, Martin SF (2003) Angew Chem Int Ed 42:3278
33. Evans DA, Kim AS (1997) Tetrahedron Lett 38:53
34. Evans DA, Kim AS, Metternich R, Novack VJ (1998) J Am Chem Soc 120:5921
35. Muri D, Lohse-Fraefel N, Carreira EM (2005) Angew Chem Int Ed 44:4036
36. Toshima K, Tatsuta K, Kinoshita M (1986) Tetrahedron Lett 27:4741
37. Toshima K, Tatsuta K, Kinoshita M (1988) Bull Chem Soc Jpn 61:2369
38. Nakamura H, Arata K, Wakamatsu T, Ban Y, Shibasaki M (1990) Chem Pharm Bull 38:2435
39. Seebach D, Chow H-F, Jackson RFW, Lawson K, Sutter MA, Thaisrivongs S, Zimmermann J (1985) J Am Chem Soc 107:5292
40. Seebach D, Chow H-F, Jackson RFW, Sutter MA, Thaisrivongs S, Zimmermann J (1986) Liebigs Ann Chem 1281
41. Ziegler FE, Tung JS (1991) J Org Chem 56:6530
42. Evans DA, Fitch DM (1997) J Org Chem 62:454
43. Paterson I, Man J (1997) Tetrahedron Lett 38:695
44. Paterson I, Lombart HG, Allerton C (1999) Org Lett 1:19
45. Nakata M, Osumi T, Ueno A, Kimura T, Tamai T, Tatsuta K (1991) Tetrahedron Lett 32:6015
46. Nakata M, Osumi T, Ueno A, Kimura T, Tamai T, Tatsuta K (1992) Bull Chem Soc Jpn 65:2974
47. Eshelman JE, Epps JL, Kallmerten J (1993) Tetrahedron Lett 34:749
48. Carter KD, Panek JS (2004) Org Lett 6:55
49. Baker R, Castro JL (1989) J Chem Soc Chem Commun 378
50. Baker R, Castro JL (1990) J Chem Soc Perkin Trans 1 47
51. Evans DA, Miller SJ, Ennis MD, Ornstein PL (1992) J Org Chem 57:1067
52. Evans DA, Miller SJ, Ennis MD (1993) J Org Chem 58:471
53. Panek JS, Xu F (1995) J Am Chem Soc 117:10587
54. Panek JS, Xu F, Rondon AC (1998) J Am Chem Soc 120:4113
55. Coutts SJ, Wittman MD, Kallmerten J (1990) Tetrahedron Lett 31:4301
56. Coutts SJ, Kallmerten J (1990) Tetrahedron Lett 31:4305
57. Martin SF, Dodge JA, Burgess LE, Hartmann M (1992) J Org Chem 57:1070
58. Martin SF, Dodge JA, Burgess LE, Limberakis C, Hartmann M (1996) Tetrahedron 52: 3229
59. Marshall JA, Sedrani R (1991) J Org Chem 56:5496
60. Tatsuta K, Itoh M, Hirama R, Araki N, Kitagawa M (1997) Tetrahedron Lett 38:583
61. Chida N, Takeoka J, Tsutsumi N, Ogawa S (1995) J Chem Soc Chem Commun 793
62. Chida N, Takeoka J, Ando K, Tsutsumi N, Ogawa S (1997) Tetrahedron 53:16287
63. Corey EJ, Reichard GA (1992) J Am Chem Soc 114:10677
64. Corey EJ, Reichard GA, Kania R (1993) Tetrahedron Lett 34:6977
65. Corey EJ, Li W, Nagamitsu T (1998) Angew Chem Int Ed Engl 37:1676
66. Corey EJ, Li W, Reichard GA (1998) J Am Chem Soc 120:2330
67. Sunazuka T, Nagamitsu T, Tatsuzaki K, Tanaka H, Omura S, Smith III AB (1993) J Am Chem Soc 115:5302
68. Nagamitsu T, Sunazuka T, Tanaka H, Omura S, Sprengeler PA, Smith III AB (1996) J Am Chem Soc 118:3584
69. Uno H, Baldwin JE, Russell AT (1994) J Am Chem Soc 116:2139
70. Panek JS, Masse CE (1999) Angew Chem Int Ed 38:1093
71. Ooi H, Ishibashi N, Iwabuchi Y, Ishihara J, Hatakeyama S (2004) J Org Chem 69:7765
72. Fukuda N, Sasaki K, Sastry TVRS, Kanai M, Shibasaki M (2006) J Org Chem 71:1220
73. Balskus EP, Jacobsen EN (2006) J Am Chem Soc 128:6810
74. Oikawa H, Oikawa M, Ueno T, Ichihara A (1994) Tetrahedron Lett 35:4809

75. Oikawa M, Ueno T, Oikawa H, Ichihara A (1995) J Org Chem 60:5048
76. Jiang Y, Ichikawa Y, Isobe M (1995) Synlett 285
77. Ichikawa Y, Tsuboi K, Jiang Y, Naganawa A, Isobe M (1995) Tetrahedron Lett 36:7101
78. Jiang Y, Ichikawa Y, Isobe M (1997) Tetrahedron 53:5103
79. Tsuboi K, Ichikawa Y, Naganawa A, Isobe M, Ubukata M, Isono K (1997) Tetrahedron 53:5083
80. Tsuboi K, Ichikawa Y, Jiang Y, Naganawa A, Isobe M (1997) Tetrahedron 53:5123
81. Nakamura S, Shibasaki M (1994) Tetrahedron Lett 35:4145
82. Shimizu S, Nakamura S, Nakada M, Shibasaki M (1996) Tetrahedron 52:13363
83. Corey EJ, Weigel LO, Chamberlin AR, Lipshutz B (1980) J Am Chem Soc 102:1439
84. Sheppeck JE, Liu W, Chamberlin AR (1997) J Org Chem 62:387
85. Marshall JA, Yanik MM (2001) J Org Chem 66:1373
86. Maurer KW, Armstrong RW (1996) J Org Chem 61:3106
87. Nagumo S, Arai T, Akita H (1997) Tetrahedron Lett 38:5165
88. Jones TK, Mills SG, Reamer RA, Askin D, Desmond R, Volante RP, Shinkai I (1989) J Am Chem Soc 111:1157
89. Jones TK, Reamer RA, Desmond R, Mills SG (1990) J Am Chem Soc 112:2998
90. Villalobos A, Danishefsky SJ (1989) J Org Chem 54:13
91. Linde II RG, Egbertson M, Coleman RS, Jones AB, Danishefsky SJ (1990) J Org Chem 55: 2771
92. Villalobos A, Danishefsky SJ (1990) J Org Chem 55:2776
93. Jones AB, Villalobos A, Linde II RG, Danishefsky SJ (1990) J Org Chem 55:2786
94. Schreiber SL, Sammakia T, Uehling DE (1989) J Org Chem 54:15
95. Nakatsuka M, Ragan JA, Sammakia T, Smith DB, Uehling DE, Schreiber SL (1990) J Am Chem Soc 112:5583
96. Ireland RE, Gleason JL, Gegnas LD, Highsmith TK (1996) J Org Chem 61:6856
97. Ireland RE, Liu L, Roper TD (1997) Tetrahedron 53:13221
98. Ireland RE, Liu L, Roper TD, Gleason JL (1997) Tetrahedron 53:13257
99. Smith III AB, Hale KJ (1989) Tetrahedron Lett 30:1037
100. Smith III AB, Hale KJ, Laakso LM, Chen K, Ri'era A (1989) Tetrahedron Lett 30:6963
101. Smith III AB, Chen K, Robinson DJ, Laakso LM, Hale KJ (1994) Tetrahedron Lett 35:4271
102. Rosenthal A, Catsoulacos P (1969) Can J Chem 47:2747
103. Jarosz S, Hicks DR, Fraser-Reid B (1982) J Org Chem 47:935
104. Rama Rao AV, Chakraborty TK, Reddy KL (1990) Tetrahedron Lett 31:1439
105. Rama Rao AV, Chakraborty TK, Reddy KL (1991) Tetrahedron Lett 32:1251
106. Maier ME, Haller B-U, Stumpf R, Fischer H (1993) Synlett 863
107. Gu R-L, Sih CJ (1990) Tetrahedron Lett 31:3283
108. Zheng W, Seletsky BM, Palme MH, Lydon PJ, Singer LA, Chase CE, Lemelin CA, Shen Y, Davis H, Tremblay L, Towle MJ, Salvato KA, Wels BF, Aalfs KA, Kishi Y, Littlefield BA, Yu MJ (2004) Bioorg Med Chem Lett 14:5551
109. Towle MJ, Salvato KA, Budrow J, Wels BF, Kuznetsov G, Aalfs KK, Welsh S, Zheng W, Seletsky BM, Palme MH, Habgood GJ, Singer LA, DiPietro LV, Wang Y, Chen JJ, Quincy DA, Davis A, Yoshimatsu K, Kishi Y, Yu MJ, Littlefield BA (2001) Cancer Res 61:1013
110. Kuznetsov G, Towle MJ, Cheng H, Kawamura T, TenDyke K, Liu D, Kishi Y, Yu MJ, Littlefield BA (2004) Cancer Res 64:5760
111. Aicher TD, Kishi Y (1987) Tetrahedron Lett 28:3463
112. Aicher TD, Buszek KR, Fang FG, Forsyth CJ, Jung SH, Kishi Y, Scola PM (1992) Tetrahedron Lett 33:1549
113. Buszek KR, Fang FG, Forsyth CJ, Jung SH, Kishi Y, Scola PM, Yoon SK (1992) Tetrahedron Lett 33:1553
114. Fang FG, Kishi Y, Matelich MC, Scola PM (1992) Tetrahedron Lett 33:1557
115. Aicher TD, Buszek KR, Fang FG, Forsyth CJ, Jung SH, Kishi Y, Matelich MC, Scola PM, Spero DM, Yoon SK (1992) J Am Chem Soc 114:3162
116. Kishi Y (1992) Pure Appl Chem 64:343
117. Duan JJW, Kishi Y (1993) Tetrahedron Lett 34:7541
118. Stamos DP, Kishi Y (1996) Tetarhedron Lett 37:8643
119. Stamos DP, Chen SS, Kishi Y (1997) J Org Chem 62:7552
120. Wong MYH, Gray GR (1978) J Am Chem Soc 100:3548

121. Vekemans JAJM, Boerekamp J, Godefroi EF, Chittenden GJF (1985) Recl Trav Chim Pays-Bas 104:266
122. Vekemans JAJM, de Bruyn RGM, Caris RCHM, Kokx AJPM, Konings JJHG, Godefroi EF, Chittenden GJF (1987) J Org Chem 52:1093
123. Vekemans JAJM, Franken GAM, Dapperens CWM, Godefroi EF, Chittenden GJF (1988) J Org Chem 53:627
124. Kim S, Salomon RG (1989) Tetrahedron Lett 30:6279
125. Cooper AJ, Salomon RG (1990) Tetrahedron Lett 31:3813
126. DiFranco E, Ravikumar VT, Salomon RG (1993) Tetrahedron Lett 34:3247
127. Cooper AJ, Pan W, Salomon RG (1993) Tetrahedron Lett 34:8193
128. Horita K, Hachiya S, Nagasawa M, Hikota M, Yonemitsu O (1994) Synlett 38
129. Horita K, Nagasawa M, Hachiya S, Yonemitsu O (1994) Synlett 40
130. Horita K, Sakurai Y, Nagasawa M, Hachiya S, Yonemitsu O (1994) Synlett 43
131. Horita K, Sakurai Y, Nagasawa M, Maeno K, Hachiya S, Yonemitsu O (1994) Synlett 46
132. Horita K, Hachiya S, Ogihara K, Yoshida Y, Nagasawa M, Yonemitsu O (1996) Heterocycles 42:99
133. Horita K, Hachiya S, Yamazaki T, Naitou T, Uenishi J, Yonemitsu O (1997) Chem Pharm Bull 45:1265
134. Horita K, Sakurai Y, Nagasawa M, Yonemitsu O (1997) Chem Pharm Bull 45:1558
135. Horita K, Nagasawa M, Hachiya S, Sakurai Y, Yamazaki T, Uenishi J, Yonemitsu O (1997) Tetrahedron Lett 38:8965
136. Horita K, Nagasawa M, Sakurai Y, Yonemitsu O (1998) Chem Pharm Bull 46:1199
137. Burke SD, Jung KW, Lambert WT, Phillips JR, Klovning JJ (2000) J Org Chem 65: 4070
138. Austad BC, Hart AC, Burke SD (2002) Tetrahedron 58:2011
139. Jiang L, Martinelli JR, Burke SD (2003) J Org Chem 68:1150
140. Lambert WT, Hanson GH, Benayoud F, Burke SD (2005) J Org Chem 70:9382

4.5 Monosaccharides as Scaffolds for the Synthesis of Novel Compounds

Paul V. Murphy[1], *Trinidad Velasco-Torrijos*[1]
[1] Centre for Synthesis and Chemical Biology, School of Chemistry and Chemical Biology, University College Dublin, Belfield, Dublin 4, Ireland
[2] Current address: Department of Chemistry, National University of Ireland Maynooth, Co. Kildare, Ireland
paul.v.murphy@ucd.ie, trinidad.velascotorrijos@nuim.ie

Abstract

This chapter focuses on monosaccharides and their derivatives as scaffolds for the synthesis of primarily bioactive compounds. Such carbohydrate derivatives have been designed to modulate mainly protein-protein and peptide-protein interactions although modulators of carbohydrate-protein and carbohydrate-nucleic acid interactions have also been of interest. The multiple hydroxyl groups that are present on saccharides have made pyranose, furanose and iminosugars ideal templates or scaffolds to which recognition or pharmacophoric groups can be grafted to generate novel compounds for medicinal chemistry. The synthesis of compounds for evaluations require strategies for regioselective reactions of saccharide hydroxyl groups and use of orthogonally stable protecting groups. Syntheses have been carried out on the solid phase and in solution. Also the use of uronic acids, amino sugars and sugar amino acids has facilitated the synthesis of peptidomimetics and prospecting libraries as they enable, through presence of

In: *Glycoscience*. Fraser-Reid B, Tatsuta K, Thiem J (eds)
Chapter-DOI 10-1007/978-3-540-30429-6_22:

amino or carboxylic acid groups, chemoselective approaches to be employed in solution and on solid phase. Sugar amino acids are readily incorporated, as peptide isosteres, to generate sugar-peptide hybrids or for the synthesis of novel carbopeptoids. The synthesis of new cyclic compounds, derived in part from saccharides, and their application as scaffolds is an emerging area and recent examples include spirocyclic compounds, benzodiazepine-saccharide hybrids and macrolide-saccharide hybrids. Potent bioactive saccharide derivatives have been identified that include enzyme inhibitors, somatostatin receptor ligands, integrin ligands, anti-viral compounds, shiga toxin inhibitors and cell growth inhibitors. Some saccharide derivatives have demonstrated improved cellular permeability when compared with peptides and are in clinical trials.

Keywords

Peptidomimetics; Sugar amino acid; Glycosamino acid; Combinatorial chemistry; High throughput screening; Iminosugars; Hybrid compounds; Orthogonal protecting groups; Bioactive compounds

Abbreviations

DMJ deoxymannojirimycin
GAA(s) glycosamino acid(s)
Glc glucose
GlcNAc *N*-acetyl-glucosamine
Gly glycine
LPA lysophosphatidic acid
Lys lysine
MMT monomethoxytrityl
MPM methoxyphenylmethyl
NeuAc neuraminic acid
NK neurokinin
Phe phenylalanine
Pro proline
RCM ring-closing metathesis
RGD Arg-Gly-Glu
SAA(s) sugar amino acid(s)
Ser serine
SST somatostatin
SSTR somatostatin receptor
Su succinimide
TBDPS *t*-butyldiphenylsilyl
TBS *t*-butyldimethylsilyl
TFA trifluoroacetic acid
Thr threonine
TIPS triisopropylsilyl
Trp tryptophan
Tyr tyrosine

1 Introduction

There is increasing activity in efforts to develop bioactive compounds which contain carbohydrate structures. In this regard the application of monosaccharides and derivatives as scaffolds has received attention. In some natural biopolymers a monosaccharide residue is considered to have an important scaffolding role, orienting the recognition or pharmacophoric groups in a defined orientation for binding to a receptor, while at the same time this saccharide residue may or may not make binding interactions with the receptor. For example, the tetrasaccharide sialyl Lewis X plays an important role in cell-cell adhesion that occurs during the immune response and inflammatory disease. The GlcNAc (*N*-acetyl-glucosamine) residue of sialyl Lewis X plays the role of a scaffold; the hydroxyl groups at C-3 and C-4 of GlcNAc facilitate the presentation of the fucose, galactose and sialic residues in a conformation so that these groups can make critical binding interactions with selectins and the glycosidic linkage at C-1 provides the mode by which the tetrasaccharide is anchored to the protein at the cell surface (◉ *Fig. 1*). The potential of sialyl Lewis X mimetics as therapeutics prompted research, most of which was published in the 1990s, towards the development of sialyl Lewis X mimetics (glycomimetics) and this area has been reviewed comprehensively [1].

That monosaccharides were also suitable for the synthesis of peptidomimetics [2] was first demonstrated by Nicolaou, Hirschmann and co-workers in 1989 who provided experimental validation that novel ligands (peptide β-turn mimetics) for peptide based on β-D-glucopyranose [3]. A conceptually similar approach suggesting the use of cyclohexane as scaffold had been initially proposed, although not investigated, by Farmer [4]. In this example, pharmacophoric groups, which correspond to the amino acid side chains of the tripeptide Phe-Trp-Lys were grafted to the pyranose, which acted as a replacement for the peptide backbone and projected the binding groups towards their respective receptor subsites (◉ *Fig. 1*). Since then, carbohydrate-based peptidomimetic development has become an active research area where the goal is to develop monosaccharide derivatives that bind to peptide receptors, with potentially better bioavailability, biostability and selectivity than endogenous or synthetic peptide ligands [5].

A cheminformatic or modelling approach has provided further indication of the potential of pyranosides as scaffolds. Meutermans and co-workers at Alchemia have investigated the structural diversity of pyranosides based on eight scaffolds (α- and β-anomers of D-glucopyranose, D-mannopyranose, D-allopyranose, and D-galactopyranose) and considered the geometric properties of these scaffolds when they display three unique substituents and compared them with tripeptide sequences in 200,000 proteins [6]. The outcome of this analysis showed that most peptides can be mimicked effectively with a subset of the pyranose scaffolds. One single scaffold [α-D-Glc or β-D-GlcNAc] provided access to many areas in the diversity space and the saccharides have the potential to access conformational space not available to peptides.

There has been significant research activity in efforts to synthesise novel bioactive molecules based on a wide range of (amino)sugar scaffolds [7], including L-sugars [6], iminosugars [8], sugar amino acids (SAAs) [9], and disaccharides [10]. Such compounds have been designed for the modulation of peptide-protein or protein-protein interactions, or for modulation of other biological recognition events including carbohydrate-protein and carbohydrate-nucleic acid interactions. A number of reviews concerning application of saccharides as scaffolds in combinatorial, bioorganic, and medicinal chemistry are available [5,11]. Many chemists have been

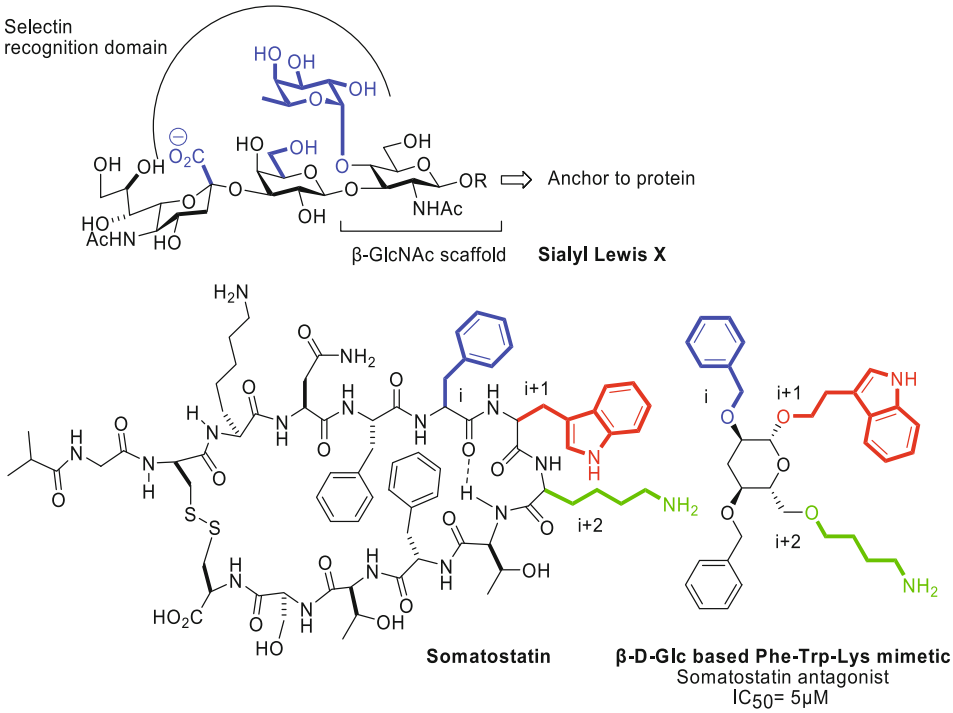


Figure 1
Compounds based on monosaccharide scaffolds: the roles of GlcNAc in presenting pharmacophoric groups (*grey*) of sialyl Lewis X to selectin receptors and β-D-glucopyranose in presenting pharmacophoric amino acid side chains (*grey*) of somatostatin

interested in the development of monosaccharides as scaffolds as they are chiral, they have multiple functional groups which facilitate the introduction of pharmacophoric groups and they have rigid conformation. The need for synthesis of carbon–carbon bonds is avoided as the hydroxyl groups can be converted into ethers. The introduction of SAAs, as intermediates for bioactive compound development, allows for chemoselective reactions to be applied on monosaccharide scaffolds, simplifying the synthesis of novel compounds for biological evaluation.

The purpose of this chapter is to provide an overview of the application of monosaccharides for the synthesis of novel compounds, and in particular we have focused on bioactive compounds that target protein-protein or peptide-protein interactions. We have not reviewed extensively the application of monosaccharides as scaffolds that modulate carbohydrate-protein interactions (glycomimetics or inhibitors of glycoprocessing enzymes) as these are considered elsewhere. Also it has not been possible to detail all of the work of those groups who have contributed to this area. The review is divided into two main sections: synthesis and application of monosaccharides that are SAAs and synthesis and application of monosaccharides that are not SAAs. The structures and applications of the monosaccharide building blocks are considered as are the synthesis and properties of target compounds.

2 Sugar Amino Acids

Sugar amino acids, also referred to as glycosamino acids (GAAs), are monosaccharide-based building blocks that feature both a carboxylic acid and an amine functional group. According to this definition, a great diversity of structures can be classified as SAAs or GAAs. Also included within this group of compounds are molecules featuring an azido functionality (sugar azido acids) as normally the corresponding amine can be obtained by reduction of the azide, and sometimes it is convenient to carry out synthetic transformations having the azide as a masked amino group. There has been a growing interest in SAAs and their application in synthetic chemistry and chemical biology in recent years [11e,f], [12]. On one hand, they keep the stereochemical and functional properties of carbohydrates, while on the other they can be incorporated into oligomeric structures, offering additional groups for further modifications. Examples of bioactive compounds of interest in this regard are sugar-peptide hybrids, either where the SAA is used as a peptide isostere or as a scaffold for the synthesis of non-peptide peptidomimetics or other libraries for lead compound discovery. The synthesis of foldamers (carbopeptoids) using peptide-coupling methodologies is another example of the use of SAAs. The presence of amino, azido, carboxyl and hydroxyl groups, present on the saccharide scaffolds, facilitates using chemoselective reactions which can minimise protecting group manipulations and simplify the synthesis of target compounds.

Since the introduction of SAAs by Kessler there have been numerous examples reported and representative examples of SAAs are shown in ❯ *Fig. 2*, ❯ *Fig. 3* and ❯ *Fig. 4*, depending on the functional groups being displayed on scaffolds derived from pyran [9,13,14,15,16,17,18,19,20,21,22,23], furan [22,24,25,26,27,28,29,20,31], oxetane [32], bicyclics and spirocyclics [33,34] and iminosugars [35,36,37,38], respectively. These compounds have found application as peptide isosteres, peptidomimetic building blocks, building blocks for peptoid synthesis, building blocks for synthesis of libraries for screening of glycosidases and for the synthesis of libraries for screening against a wide range of biological targets (prospecting libraries).

2.1 Synthesis of SAAs

The synthesis of a member of the SAA kit first developed by Kessler [9,13] is shown in ❯ *Scheme 1*. The reaction of glucurono-6,3-lactone with methanol in the presence of a base and subsequent acetylation with acetic anhydride and sodium acetate gave **41**. The anomeric acetate of **41** was displaced with azide when reacted with tin (IV) chloride and azidotrimethylsilane giving 1,2-*trans* glycoside **42**, with a β-configuration, which is controlled by the participating group at C-2. Subsequent reduction of the azide gave the amine **1**, which was incorporated, as a peptide isostere, into peptide-sugar hybrids.

Overhand, Overkleeft and their collaborators have synthesised a variety of SAAs [39]. The synthesis of Fmoc-protected **6** (❯ *Fig. 2*) is illustrated in ❯ *Scheme 2* [15]. Tri-*O*-acetyl-D-glucal was converted into the SAA precursor **43** in five steps. Formation of the trichloroacetimidate derivative **44** and a subsequent Overman rearrangement was used to introduce the amino group onto the pyran scaffold giving **45**. Hydrogenation of the olefin in **45** was accompanied by cleavage of the silyl ether to afford the primary alcohol which was oxidised with a catalytic amount of ruthenium (III) chloride in the presence of sodium periodate giving **46**. Subjection

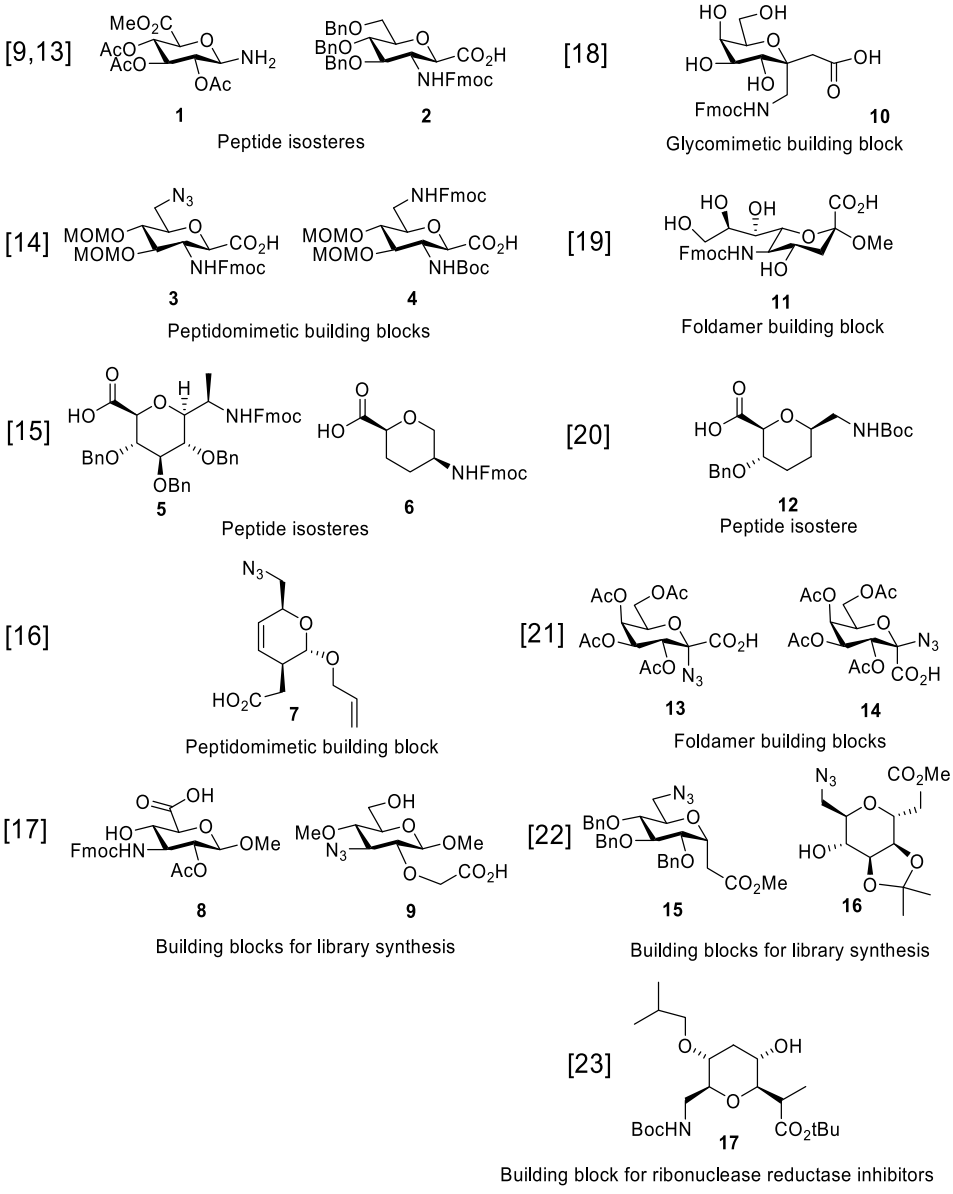
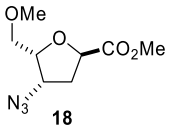


Figure 2
Examples of SAAs based on pyran scaffolds

of Boc-protected **46** to 50% TFA in dichloromethane and subsequent reaction with Fmoc-*O*-Su gave the desired amino acid **6**.

The Fleet research group developed methodologies for the preparation of α, β, γ and δ SAAs based on furan and oxetane scaffolds. The synthesis of γ and β SAAs **51** and **24**, is shown

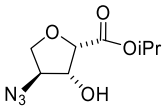
[24]



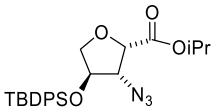
18

Building block for library synthesis

[29]



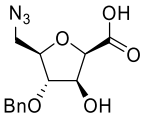
23



24

Foldamer building blocks

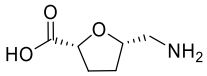
[25]



19

Peptide isostere

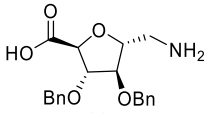
[30]



25

Peptide isostere

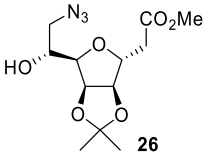
[26]



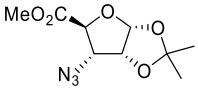
20

Peptide isostere

[22]



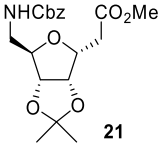
26



27

Building blocks for library synthesis

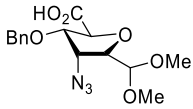
[27]



21

Building block for library synthesis

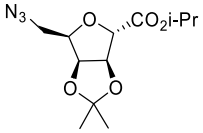
[31]



28

Peptidomimetic building blocks

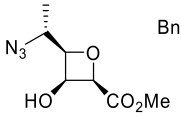
[28]



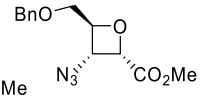
22

Foldamer building block

[32]



29



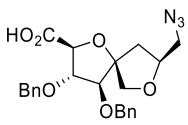
30

Foldamer building block and peptide isosteres

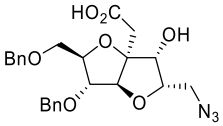
Figure 3
Examples of SAAs based on furan and oxetane scaffolds

in *Scheme 3* [29b]. The key triflate intermediate **48** was prepared via **47** in three steps from L-arabinose and this underwent ring contraction under acidic conditions in methanol, together with deprotection of the acetonide, to give the corresponding methyl ester which, was then subjected to acid-catalysed transesterification to form the stable isopropyl ester **49**. Treatment of diol **49** with a single equivalent of TBDPSCl gave a mixture of the monoprotected silyl ethers in the 3 and 4 positions **50** and **51**, respectively, with low regioselectivity. The 3-*O*-TBDPS-protected derivative **50** was converted via a triflate intermediate, to give γ-azido ester **51** with total inversion of the configuration at C-4. The β-azido ester **22** was prepared similarly from the 4-*O*-TBDPS-protected **52**.

[33]

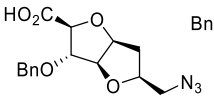


31

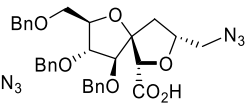


32

[33]



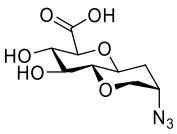
33



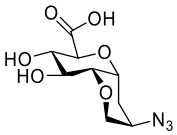
34

Dipeptide isosteres

[34]



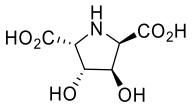
35



36

Peptidomimetic building blocks

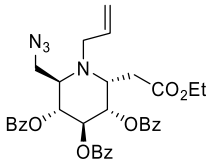
[35]



37

Peptidomimetic building block

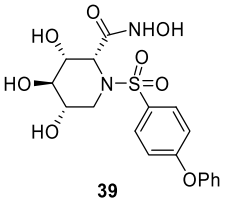
[36]



38

Building block for glycosidase library synthesis

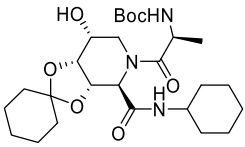
[37]



39

Matrix metalloproteinase inhibitor

[38]

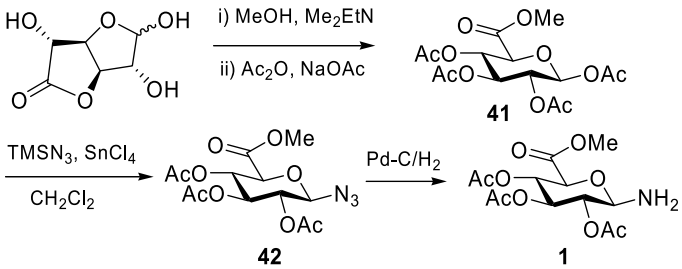


40

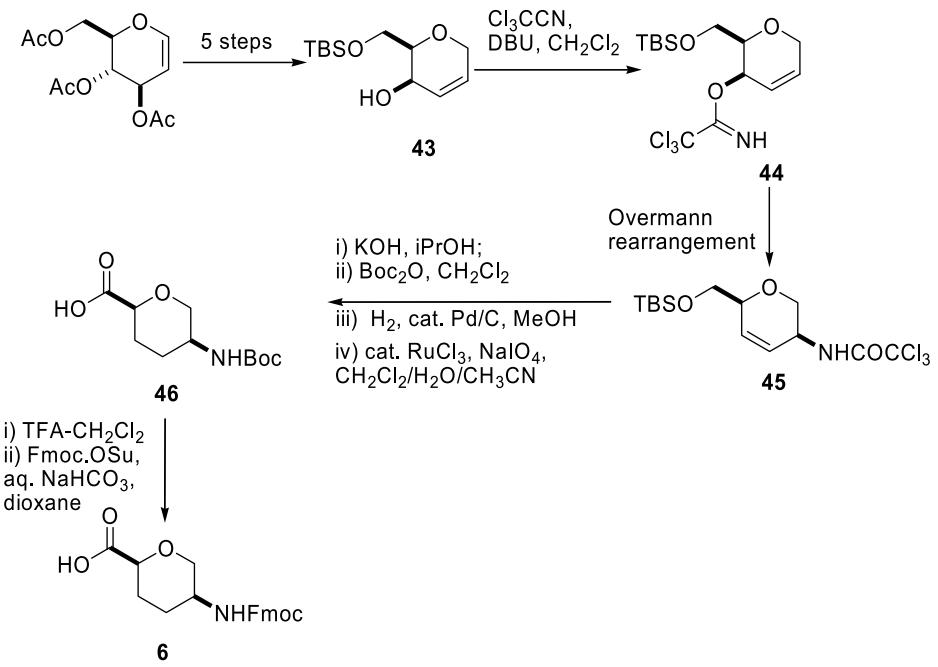
Combinatorial library member

Figure 4

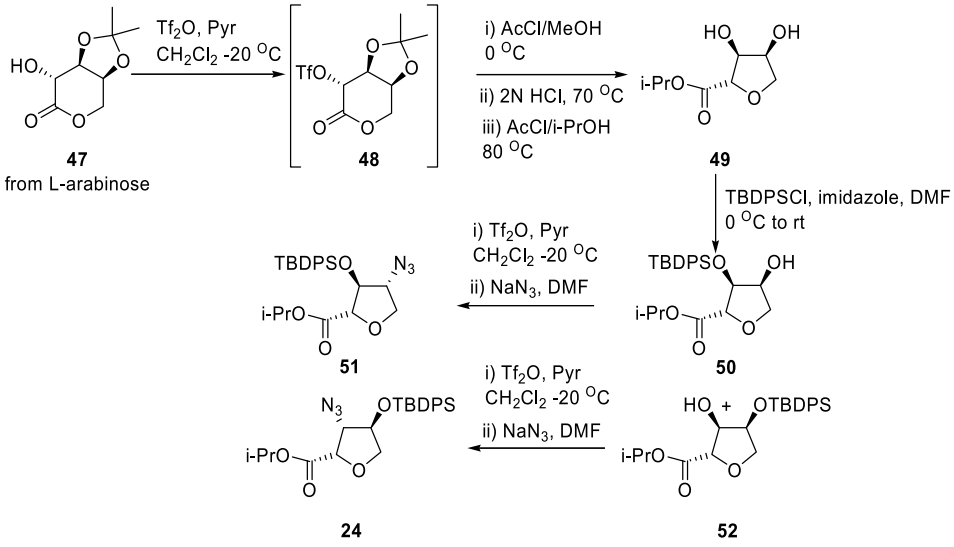
Examples of SAAs and derivatives based on fused bicyclic and spirocyclic saccharide derivatives and iminosugar scaffolds



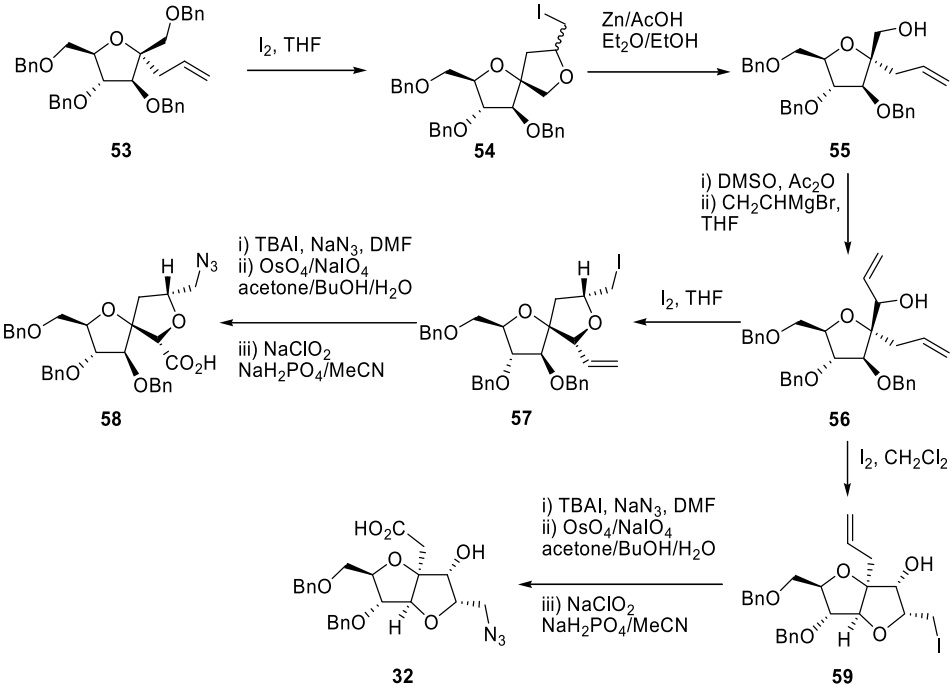
Scheme 1



Scheme 2



Scheme 3

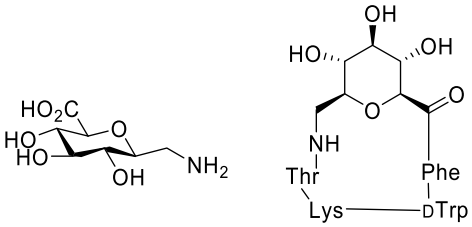


Scheme 4

Nicotra's group have synthesised SAAs based on fused bicyclic and spirocyclic compounds. The synthesis of **32** and **58** from the fructoside **53**, is shown in *Scheme 4* [33]. Treatment of fructoside **53** with I_2 resulted in a 5-exo cyclisation to give a diastereoisomeric mixture of spiro cyclic compounds **54**. The mixture was then reacted with zinc and acetic acid to give **55**, in which the double bond was regenerated with the selective deprotection of the benzyl group at C-1 having been achieved. Oxidation of the free hydroxyl to the corresponding aldehyde was followed by instalment of a second double bond through Grignard addition to give **55**. When this intermediate was treated with I_2 in THF, the reaction proceeded with attack of the free hydroxyl to give spirocyclic compound **57**. However, when the reaction of **56** with I_2 was carried out in dichloromethane as a solvent, the benzyloxy group at C-3 was involved in the cyclisation, this resulting in the formation of fused bicyclic compound **59**. Similar reaction sequences for both **57** and **58**, consisting of displacement of the iodide by sodium azide, followed by oxidative cleavage of the double bond, gave **58** and **32**, respectively.

2.2 Application of SAAs as Peptide Isosteres

The presence of amino and carboxylic acid groups conveniently protected allows for the coupling of the SAAs building blocks to form sugar-peptide hybrids. One or more amino acids of the peptidic backbone can be replaced by the sugar units, thus incorporating the conformation-



60 61

Figure 5
Kessler's SAA building block 60 was incorporated into cyclic peptide somatostatin analogue 61

al constraints inherent to the carbohydrate ring into the peptide secondary structure. Kessler and co-workers [9,13] have synthesised a collection of pyranose-based SAAs and incorporated such saccharides into linear and cyclic oligopeptides. They have studied by NMR and computational techniques the influence of the introduction of these building blocks on the conformational properties of the peptidomimetics. The different SAAs were incorporated into linear and cyclic oligopeptides. It was found that these compounds can be used as dipeptide isosteres [for example, of the natural Gly(glycine)-Gly dipeptide] and homoproline isosteres. They also mimic the linear, flexible β-turn and γ-turn peptide secondary structures. For example SAA **60** was incorporated into the cyclic peptide somatostatin (SST) analogue **61** (*Fig. 5*), which was a somatostatin mimetic in which the β-turn motif is mimicked by the SAA. The novel sugar-peptide hybrid had an IC_{50} of 0.15 µM for the inhibition of the release of growth hormone.

Kunwar, Chakraborty and co-workers have prepared a series of sugar-peptide hybrids as mimetics of a potent vasoactive intestinal peptide receptor inhibitor by incorporating furanoid SAAs into **62** (*Fig. 6*) [26]. The carbohydrate units act as dipeptide isosters of either Tyr(tyrosine)-Pro(proline) or Pro-Thr(threonine) present in the original structure. The anticancer activities of the resulting analogues were tested in a series of human cell-lines, showing similar or enhanced activities to the original peptide. The group of Overhand and Overkleeft has made a number of contributions. This has included the preparation of gramicidin S analogues containing arylated sugar amino acids such as **19** (*Fig. 3*) as replacements for D-Phe-Pro β-turn regions and the novel sugar-peptide hybrid, which proved to be as active as the parent compound as antibacterial agents and in lysing red blood cells [25]. The same group have also used SAAs to prepare 11 analogues of the *C*-terminal Ca_1a_2X motif found in the natural substrates of the prenyl transferases (PFT and PGGT-1) and a highly selective inhibitor **63** (*Fig. 6*) was identified [20]. These transferases catalyse the transfer of isoprenoid lipids from the corresponding pyrophosphates to cysteine residues at the *C*-termini of precursor proteins that are important in signal transduction. A number of inhibitors of these proteins are in clinical development. The methyl ester of the peptide mimetic was able to inhibit protein farnesylation in intact Met-18b-2 cells indicating that the methyl ester of **63** is cell permeable. Nicotra and collaborators used their bicyclic SAA **32** (*Fig. 4*) to synthesise a macrocyclic pseudopeptide incorporating the biologically relevant RGD sequence [40]; this compound inhibited adhesion of fetal bovine aortic endothelial GM7373 cells to their ligands: fibroblast growth factor, vitronectin and fibronectin indicating the compound is an antagonist

of the $\alpha_v\beta_3$ integrin receptor. Smith, Hirschmann, Cooperman and co-workers used **17** for the synthesis of the novel ribonuclease reductase inhibitor **64** (◉ *Fig. 6*) [23].
The Fleet research group has reported the preparation of a variety of δ-sugar azido esters based on oxetane-type scaffolds [32]. The sugar azido esters **65** (*cis*) and **66** (*trans*) can be regarded as building blocks for dipeptide isosteres of Gly-L-Ser(serine), although sugar-peptide conjugates based on these building blocks have not been prepared to date (◉ *Fig. 7*).
Several other groups have reported the use of SAAs as dipeptide isosteres of natural and non-natural occurring dipeptides [41].

2.3 Application of SAAs as Scaffolds for Synthesis of Pharmacophore Mapping or Prospecting Libraries

Sugar amino acids can be used for the synthesis of libraries of compounds for screening. Sofia and co-workers have made use of pyranose SAA scaffolds and combinatorial solid phase synthesis to generate pharmacophoric mapping libraries. The structure of their carbohydrate platforms is ideal to generate combinatorial molecular diversity and novel ligands for a variety of biological targets can be produced rapidly. They have prepared scaffolds containing three orthogonal functional groups. The chemoselective functionalisation of these groups reduces the need for protecting group strategies and facilitates the combinatorial introduction of the different pharmacophoric groups. In the example shown (◉ *Fig. 2*, ◉ *Scheme 5*) the pyranose scaffold **9** features a carboxylic acid, an azido and a free hydroxyl group. The scaffold was attached to a Rink amide resin derivative **67** by reaction of the carboxylic acid with the histidine free amino group to give solid supported scaffold **68**. It was then reacted with a variety of isocyanates to give carbamates at the 6 position and the azido group at C-3 was reduced using the Staudinger reaction to give amines which were coupled to carboxylic acids. A library of 48 compounds of type **69** was prepared. The same group defined a library synthesis based on Fmoc-protected SAAs [17].
The Fleet research group used furanose SAAs for the generation of solution-phase libraries [24]. The fully protected precursor **18** was prepared efficiently from L-gulonolactone. Subsequent reduction of the azide resulted in the free amine that was allowed to react with a range of isocyanates and isothiocyanates (as shown in the example in ◉ *Scheme 6*) to give the intermediate **70**. The transformation of the ester into an amide was achieved by treatment with free amine at elevated temperatures to give the final products of the library **71**. A 99-member library was prepared in this way and it has been subjected to biological screening for antibacterial activity.

2.4 Application of SAAs in Synthesis of Foldamers

There is increasing interest in the development of foldamers or polymeric structures that can adopt organised secondary structures like those of proteins, nucleic acids and some polysaccharides [42]. Sugar amino acid-based foldamer research has so far been primarily concerned with synthesis of polymers with secondary structural features. Such foldamers may find application as scaffolds for peptidomimetic development if they adopt turn, helical and strand structures observed for peptides or if derivatives can act as ligands for peptide receptors;

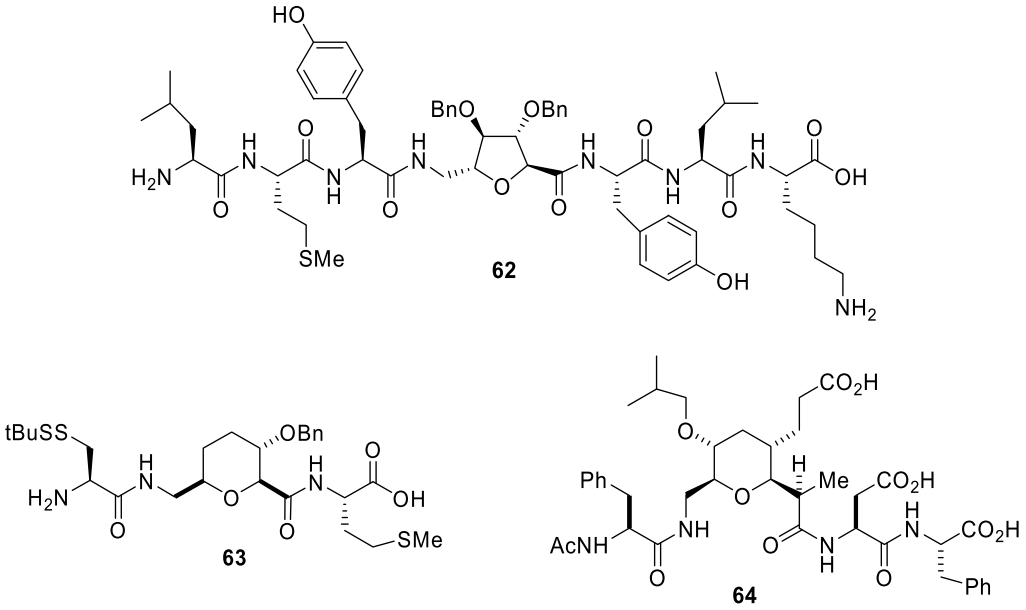


Figure 6
Vasoactive intestinal peptide receptor inhibitor 62, inhibitor of protein farnesylation 63 and ribonuclease reductase inhibitor 64 synthesised from SAAs

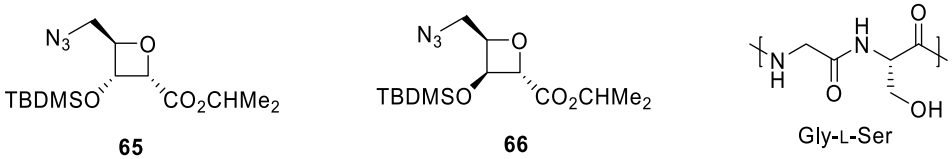
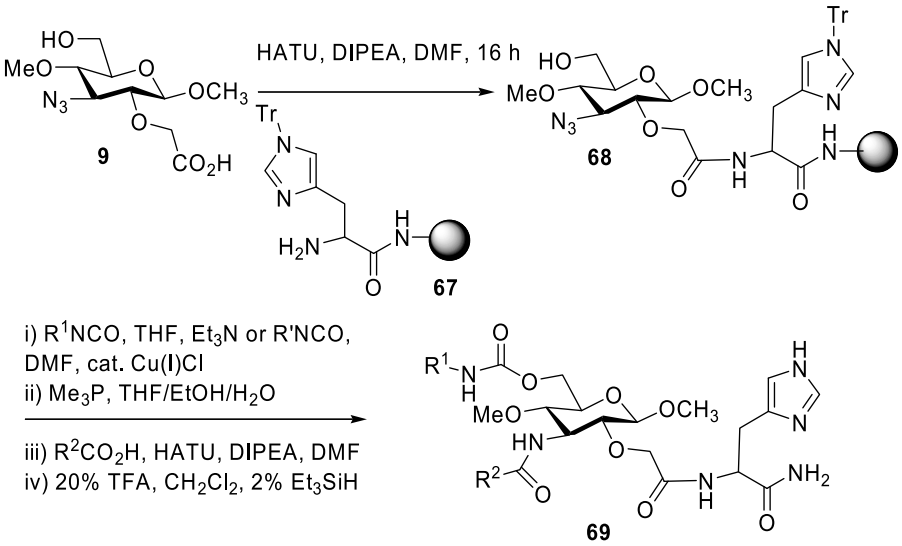
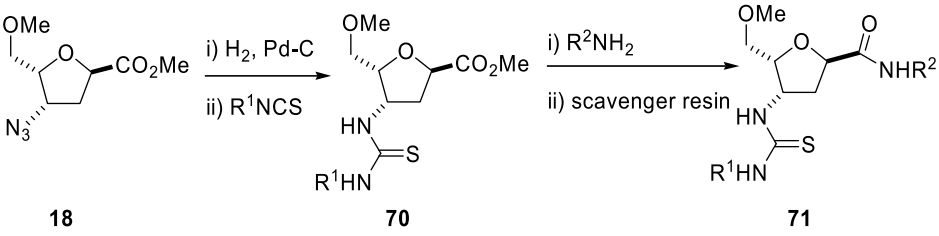


Figure 7
δ-Sugar azido esters based on oxetanes as Gly-L-Ser dipeptide isosteres



Scheme 5



Scheme 6

in such cases the additional functional groups present in SAAs could potentially be used to modify the polymer backbone to tailor the properties for binding to a target receptor. Gervay and co-workers have prepared amide-linked homo-oligomers derived from the naturally occurring SAA neuraminic acid (NeuAc) analogues [43]. They have used solid-phase peptide synthesis involving *N*-Fmoc protecting groups to generate a series of oligomers of general structure **72** (*Fig. 8*), ranging from dimeric to octameric compounds. They feature a lipid amide linker as a model for membrane-bound oligomers. The data obtained by NMR and CD

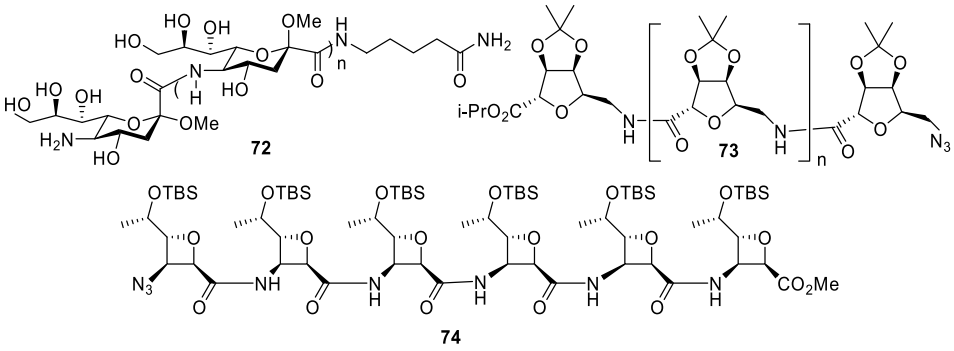


Figure 8
Foldamers prepared from SAAs

spectroscopy studies indicates that these compounds present defined secondary structures that may vary depending on their length. The corresponding homoligomers of *N*-Fmoc-protected sugar amino acids derived from α-*O*-methoxyneuraminic acid **11** (*Fig. 2*) and 2,3-dehydroneuraminic acids have been recently synthesised using solid phase methods [19]. Highly functionalised oligomers derived from other pyranosides have also been found to adopt well-defined structural conformations [44]. The Fleet research group has also reported foldamers **73** derived from furanoid sugar azido esters (*Fig. 8*). They have investigated the conformational properties in organic solvents of the corresponding tetramer and octamer. Using NMR, IR and CD spectroscopy and molecular dynamics studies, they have determined that the tetrameric compound does not adopt a hydrogen-bonded conformation whereas an octamer has a well-defined helical secondary structure stabilised by a 16-membered ring with inter-residue hydrogen bonds, similar to a π-helix [28]. Fleet and co-workers have also investigated the conformational properties of oligomers of oxetane SAAs prepared by peptide coupling techniques. The study of their secondary structure by NMR, both in $CDCl_3$ and C_6D_6, together with molecular mechanics analysis, revealed that hexamer **74** (*Fig. 8*), in which the hydroxy group is protected as a TBS ether, adopts well-defined helical structures stabilised by 10-membered hydrogen-bonded rings [32].

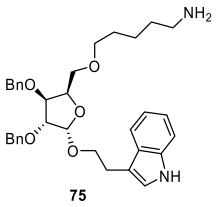
3 Bioactive Compounds Based on Monosaccharide Scaffolds that are not SAAs

Other pyranose, furanose and iminosugar derivatives that do not contain both amino and carboxylic acid functional groups have been of interest and are considered in this section.

3.1 Structures of Bioactive Compounds Designed for Specific Target Receptors

A selection of bioactive compounds based on furanose, pyranose and iminosugars [7,8,45,46,47,48,49,50,51,52,53,54,55,56,57,58,59,60,61,62,63,64,65,66,67,68,69] have been synthesised and some are shown in *Fig. 9* and *Fig. 10*. The application of pyranose scaffolds

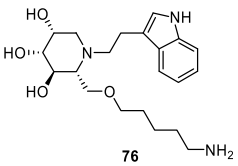
[46]



75

SSTR

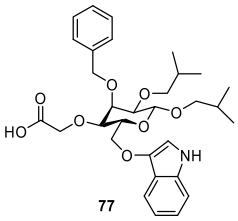
[47]



76

SSTR

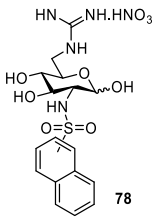
[48]



77

Endothelin antagonist

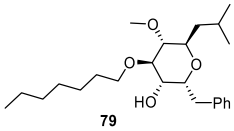
[49]



78

Thrombin inhibitor

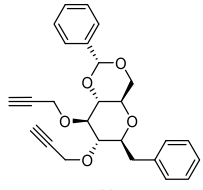
[50]



79

Multidrug-resistant-associated-protein inhibitor

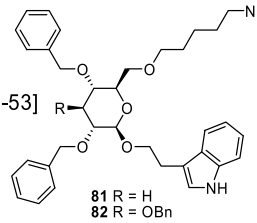
[51]



80

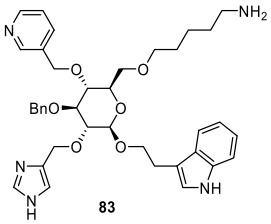
Cytomegalovirus

[52-53]



81 R = H
82 R = OBn

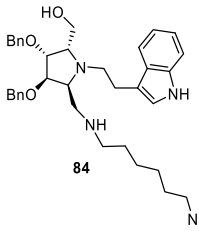
SSTR (R' = H) and NK-1R (R' = Ac)



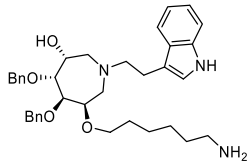
83

hSSTR-4 (53 nM)

[8]



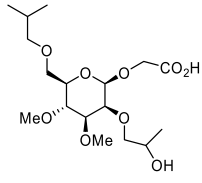
84



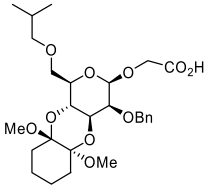
85

SSTR

[54]



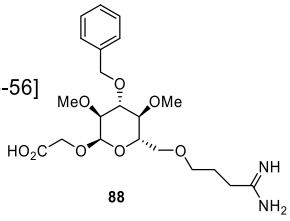
86



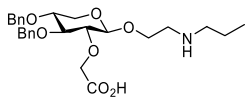
87

Integrins

[55-56]



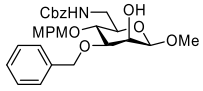
88



89

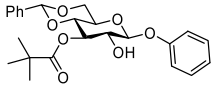
Integrins

[57]



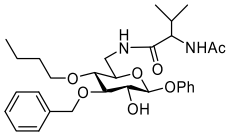
90

HIV-protease



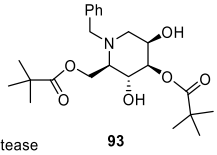
91

[58]



92

HIV-protease



93

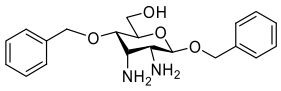
Figure 9
Bioactive compounds based on pyranose, furanose and iminosugar scaffolds that are not SAAs

for the discovery of novel peptidomimetics was first validated by initial work carried out by Nicolaou, Hirschmann and co-workers [3]. They designed and synthesised ligands such as **81** and **82** (*Fig. 1* and *Fig. 9*) as somatostatin mimetics based on compounds derived from β-D-glucopyranose. Somatostatin is a tetradecapeptide that regulates, through binding to its receptors (SSTR) a number of processes including the release of growth hormone and other pituitary hormones. It also plays a role in neuronal transmission. The side chains of the amino acids Phe, Trp and Lys were grafted to hydroxyl groups of the β-D-glucopyranose in an orientation designed to mimic the spatial presentation of the side chains of Phe (i), Trp (i + 1) and Lys (i + 2), respectively, of somatostatin (*Fig. 1*) and **81** and **82** were found to bind to the somatostatin receptor with IC_{50} values of 9.5 and 1.3 µM, respectively. Further refinement of the ligand structure was carried out by structural-activity relationship studies and different monosaccharides such as L-glucose and L-mannose were also investigated [7]. The introduction of different heteroaromatic moieties in the structure of the peptidomimetic led to the identification of a novel β-D-glucopyranoside-based ligand **83** (*Fig. 9*) with enhanced affinities (IC_{50} value of 53 nM) at human SSTR subtypes (SST4 and SST2) [3,52,53]. Other activities were observed for these peptidomimetics. Glucoside **81** is also a β2-adrenergic antagonist with an IC_{50} value of 3 µM and both **81** and **82** are ligands for the NK-1(neurokinin) receptor with IC_{50} values of 0.12 and 0.18 µM, respectively. Small changes in the structure of **82**, such as acetylation of the free amino group, resulted in enhanced activity as an NK-1 antagonist, with an IC_{50} value of 60 nM and complete loss of activity for SSTR. β-D-Glucopyranose was thus defined as a "privileged structure" since compounds based on this scaffold can bind to more than one type of receptor. This body of work from Hirschmann and collaborators illustrates the general approach that can be applied in the discovery of potent peptidomimetic compounds from carbohydrates and similar principles have been used by others in the design and synthesis of saccharide-based bioactive compounds and compounds that have been synthesised with a view to targeting particular receptors.

Somatostatin ligands based on iminosugar scaffolds have also been synthesised. Iminosugars are structural analogues of pyranosides, in which the ring oxygen atom is replaced by a nitrogen atom and iminosugar-based mimetics of somatostatin have been synthesised. In addition to somatostatin ligands [8,46,47], monosaccharide derivatives have been applied to the synthesis of cyclic endothelin antagonists [48], thrombin inhibitors [49], multidrug-resistant-associated protein inhibitors [50], cytomegalovirus inhibitors [51], integrin ligands [54,55,56], HIV protease inhibitors [57,58,59,60,64], fibrinogen receptor ligands [61], LPA agonists and antagonists [62], rhodopeptin mimetics [63], cysteine protease calpain inhibitors [65], RNA binding agents [66], inhibitors of the growth of human cell lines [67], multivalent carbohydrates [45] including shigatoxin inhibitors [68] and the synthesis of carbopeptides and carboproteins [69].

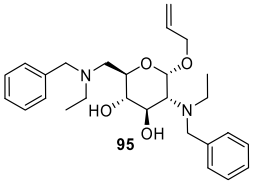
3.2 Selected Syntheses of Peptidomimetics Designed for Specific Target Receptors

The first synthesis of a carbohydrate peptidomimetic **82** was completed in eight steps from readily available tetra-*O*-acetyl-α-D-glucopyranosyl bromide **105**, which was first glycosidat-

[59] 

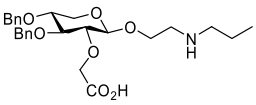
94

HIV-protease inhibitor

[60] 

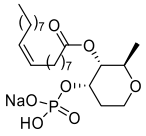
95

HIV-protease inhibitor

[61] 

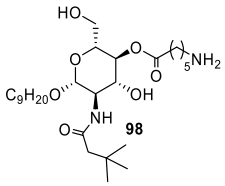
96

Fibrinogen receptor

[62] 

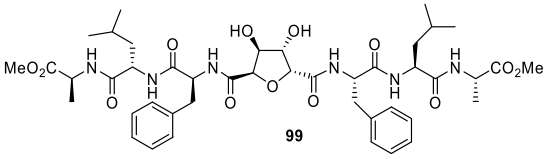
97

Lysophosphatidic acid receptor

[63] 

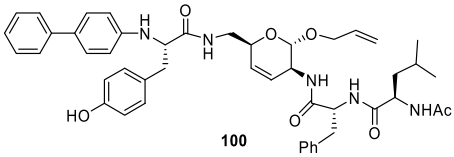
98

Rhodopeptin receptor

[64] 

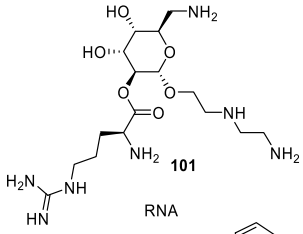
99

HIV-protease inhibitor

[65] 

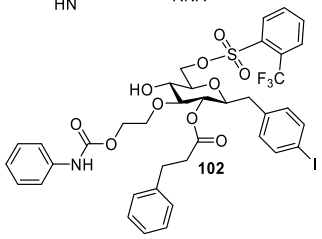
100

cysteine protease calpain inhibitor

[66] 

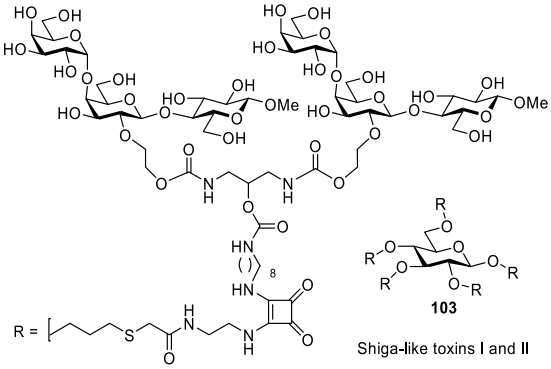
101

RNA

[67] 

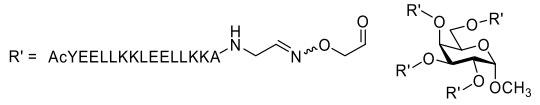
102

SH2 Domain, Cell Growth

[68] 

103

Shiga-like toxins I and II

[69] 

104

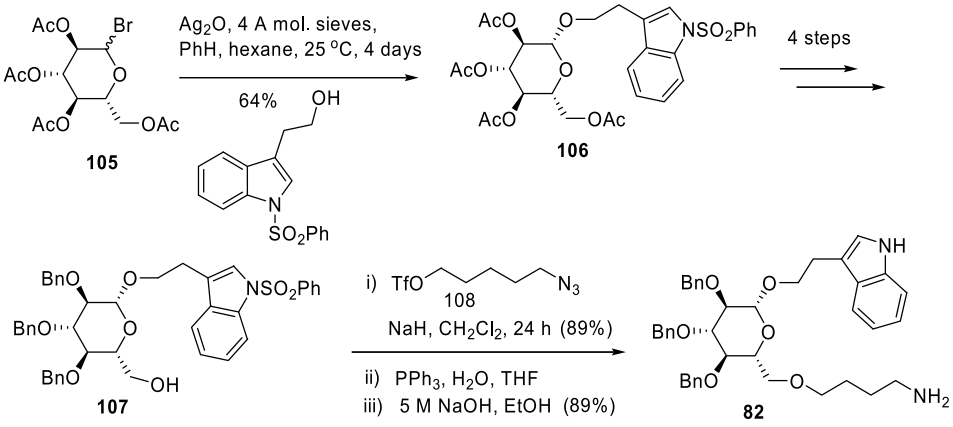
Carbopeptides and carboproteins

Figure 10
Bioactive compounds based on pyranose and furanose scaffolds that are not SAAs

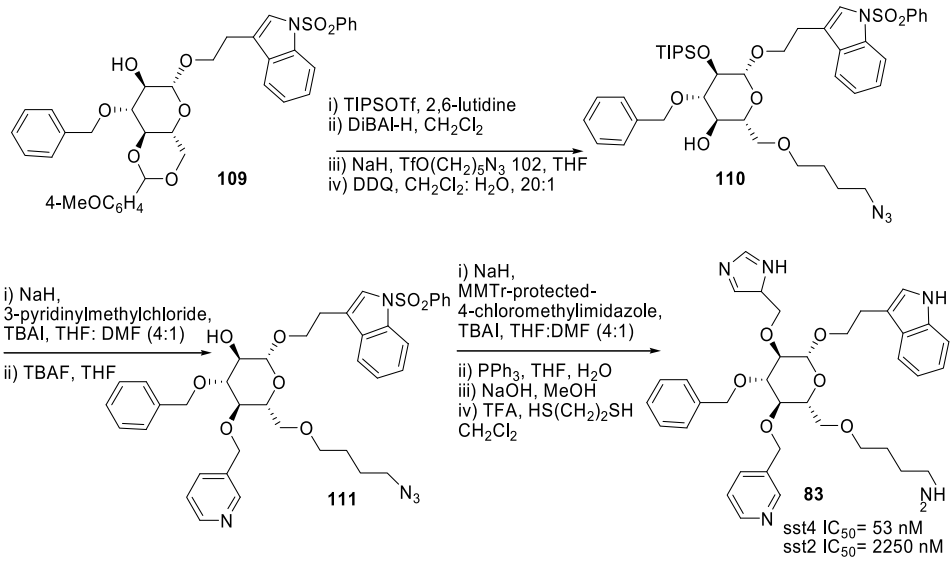
ed in the presence of a suitable alcohol acceptor using standard Koenigs–Knorr conditions to incorporate the tryptophan side chain at the anomeric carbon to give **106** (*Scheme 7*). A series of protecting group manipulations gave alcohol **107** which was subsequently reacted with triflate **108** to give an azide intermediate. The azide was then reduced in the presence of the benzyl ether protecting groups using triphenylphosphine and water, and the benzenesulfonamide was subsequently removed from the indole using aqueous sodium hydroxide, to give the β-turn mimetic **82**.

Hirschmann and co-workers have incorporated heterocyclic aromatic substituents onto the 2- and 4-OH of pyranosides structurally related to **82** to give one of the most potent saccharide-derived somatostatin ligands **83** developed to date. The synthesis of **83** is shown in *Scheme 8* from alcohol **109**. The TIPS protecting group was introduced at the 2-OH of **109** and regioselective cleavage of the methoxybenzylidene gave an intermediate where the 6-OH group is free and the 4-OH group is protected as the methoxyphenylmethyl (MPM) ether. Reaction of this intermediate with the triflate **108** and subsequent removal of the MPM ether gave **110**. The reaction of **110** with 3-picolyl chloride in the presence of sodium hydride and subsequent removal of the TIPS group gave **111**. Subsequent alkylation of **111** with *N*-monomethoxytrityl (MMTr)-protected 4-chloromethylimidazole followed by Staudinger reduction of the azide, removal of the indole sulfonamide and acidic removal of the MMT group in the presence of a dithiol scavenger gave **83**.

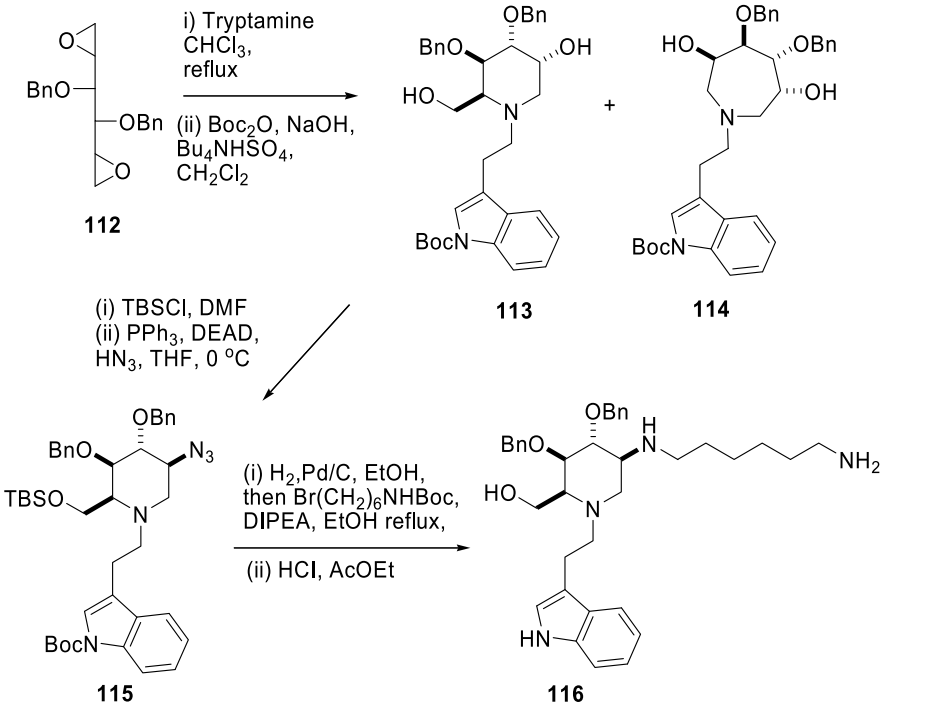
Depazay and co-workers have described the synthesis of a series of polyhydroxylated pyrrolidine, piperidine and azepanes derived from D-mannitol as novel mimetics of somatostatin [8]. The synthesis of one piperidine **116** is shown in *Scheme 9*. The authors used reaction of tryptamine with the D-mannitol derived bis-epoxide **112** followed by protection of the indole nitrogen with a Boc group to prepare the L-gulo-piperidine **113**, the azepane **114** also being formed. Selective protection of the primary alcohol followed by reaction with hydrazoic acid



Scheme 7



Scheme 8



Scheme 9

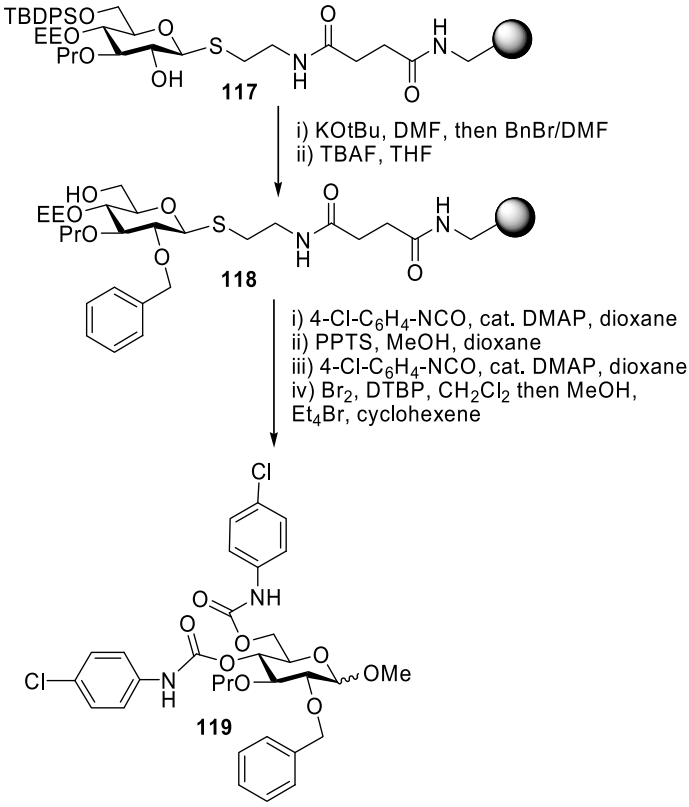
under Mitsunobu conditions gave the azide **115**. Reduction of the azide followed by introduction of the lysine side chain and subsequent removal of the protecting groups under acidic conditions gave **116**.

Gouin and Murphy have reported [47] the synthesis of the more hydrophilic dipeptide mimetic **76** (◉ *Fig. 9*) in 13 steps from 1-DMJ [70]. This DMJ derivative lacks the benzyl group in the side chain to mimic Phe, yet still displayed a K_i value of 26 μM in an unspecific SST receptor assay, and it showed preferential binding to hSSTR4 (48% at 1 μM) over hSSTR5.

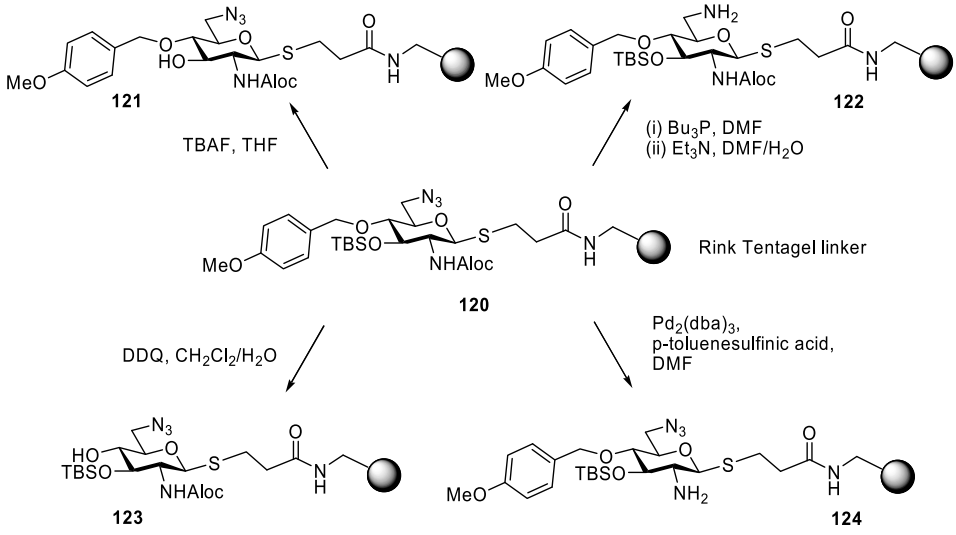
3.3 Strategies for Synthesis of Prospecting Libraries

The bioactive compounds outlined above display a wide range of bioactivities that have been observed for compounds based on monosaccharide scaffolds. Difficulties in the synthesis of novel compounds based on monosaccharides includes the ability to differentiate the hydroxyl groups from each other as they have similar reactivity and the strategies for synthesis therefore depend on the need for regioselective reactions and the use of orthogonally stable protecting groups. Selective deprotection must be available at all hydroxyl groups to be functionalised and all other protecting groups must remain unaffected during subsequent introduction of putative pharmacophoric groups. An additional complication for solid phase syntheses of such compounds is that one group is used to anchor the saccharide to the polymer support and must be stable to all chemical reagents used and then it needs to be cleaved without degrading the target compound. An alternative or complementary strategy can be the introduction of different kinds of functional groups (e. g. amine, carboxylic acid, azide, primary alcohol, secondary alcohol) to the scaffold. Then selective manipulation of these groups can be achieved on the basis of their chemoselective reactions, which forms the basis of the SAA concept discussed above. Such libraries can be focused towards an identification of a potent novel ligand for a specific target receptor or the library could be prepared with a view to screening for a range of targets. Successful syntheses of libraries on the solid phase that are not based on SAAs were published by the Kunz group [71,72]. They developed an orthogonal protection system for carbohydrate hydroxyl groups using the *tert*-butyldiphenylsilyl, acetate, allyl and ethoxyethyl protecting groups, as well as a special thioglycoside linker which doubled as a means to connect to a solid phase resin and as a glycosyl donor **117**. Each protecting group could be selectively removed and the hydroxyl group then derivatised to introduce a pharmacophoric side-chain. Substituents could be introduced at all positions of the glucopyranoside scaffold. The synthesis of a member is shown in ◉ *Scheme 10*. The 2-OH group of **117** was benzylated and then the silyl protecting group at C-6 was removed to give **118**. The 6-OH group was converted to a carbamate, the ethoxyethyl group then removed and a carbamate group then introduced at the 4 position. Finally the thioglycoside was converted to the glycosyl bromide, with concomitant removal of the saccharide from the resin, and halide-assisted glycosidation of this bromide with methanol gave **119** as a mixture of α- and β-glycosides.

More recently Kunz and co-workers described a synthesis of 2,6-dideoxy-2,6-diaminopyranoside libraries [73]. The scaffold **120** allows for selective deprotection of four positions in a manner that is independent of the sequence (◉ *Scheme 11*). Different functional groups have been incorporated at each position to give **121–4**, which can then be cleaved from the solid support. For example, the 2-*N*-allyloxycarbonyl protecting group of **120** can be selec-



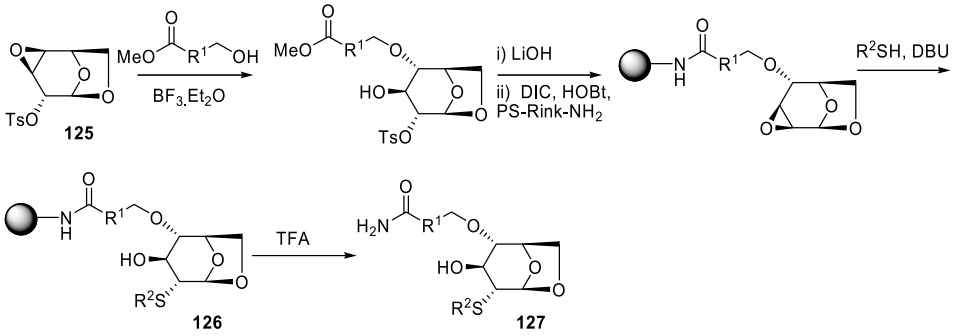
Scheme 10



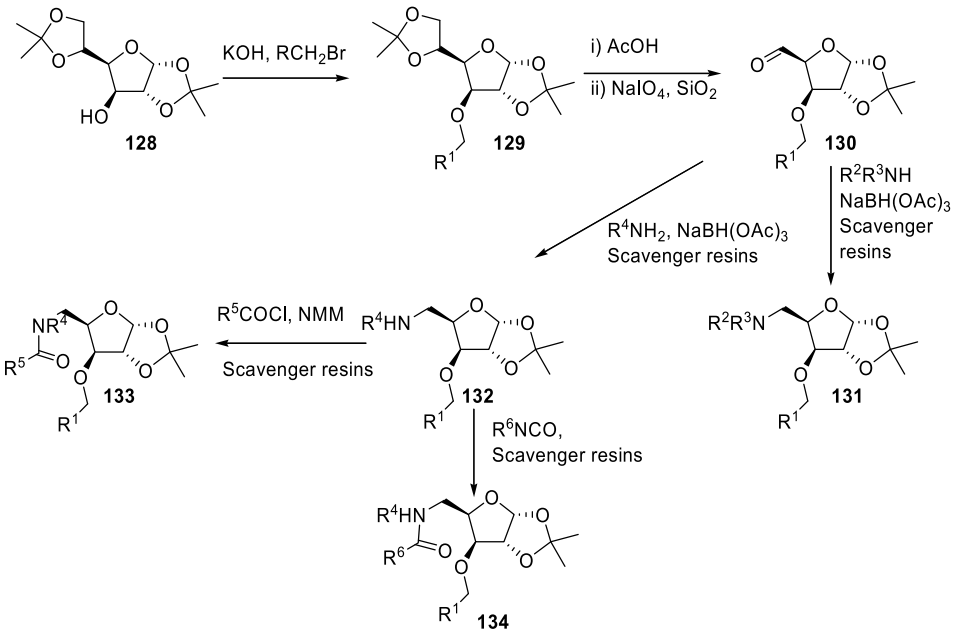
Scheme 11

tively removed by Pd (0) catalysis and then Fmoc-protected amino acids coupled to the resulting free amino group, subsequent cleavage from the Rink resin with trifluoroacetic acid in dichloromethane gives saccharide-amino acid conjugates.

Brill and co-workers used levoglucosane as a rigid scaffold to synthesise libraries both in solution [74] and solid phase [75]. The opening of epoxide **125** by different hydroxycarboxylic acid methyl esters catalysed by Lewis acid, followed by simultaneous new epoxide formation and ester hydrolysis gave the resulting free acid ready for attachment to the solid support. Epoxide opening reactions with alkoxides, amines and thiols were carried out on solid phase



Scheme 12



Scheme 13

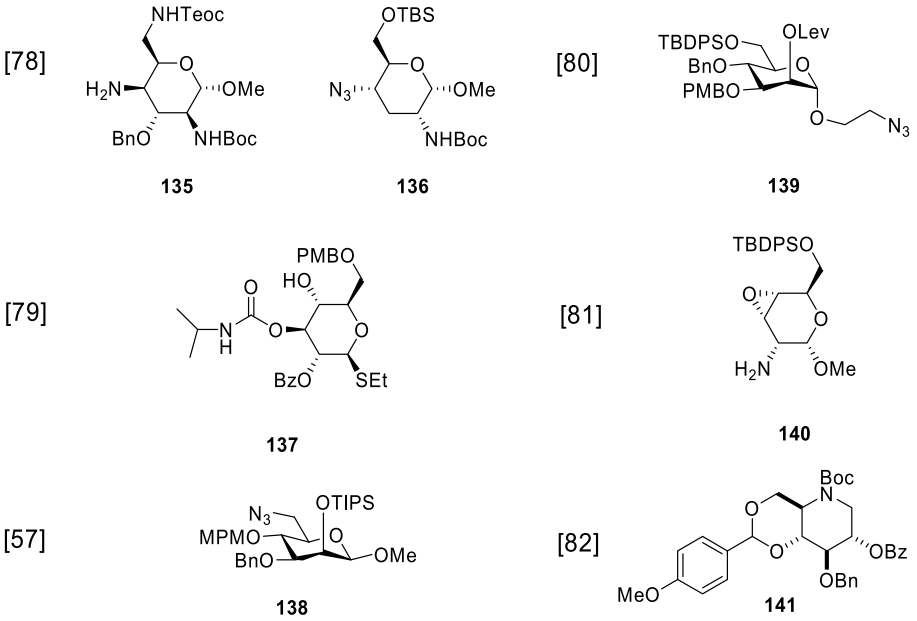


Figure 11
Examples of intermediates suitable for library synthesis

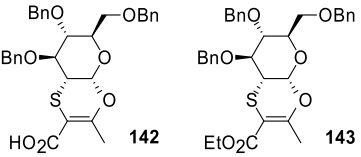
(shown for thiol giving **126**, *Scheme 12*). Cleavage from resin gave compounds such as **127**. Further derivatisation was achieved by palladium coupling reactions if these resulting compounds had aryl iodides.

Boldi and co-workers described the synthesis of a solution phase library based on furanose scaffolds [76]. Diacetone-D-glucose **128** was alkylated to give **129** and then converted in two steps to aldehyde **130**, which is the key intermediate for generation of the library. Reductive amination of **130** with secondary or primary amines gave derivatives of general formula **131** and **132**, respectively. Amines **132** could be further functionalised by reaction with various acid chlorides or isocyanates to yield compounds **133** and **134**, respectively (*Scheme 13*). In solution phase, reagents in excess were removed by using scavenger resins. Several libraries containing 5000 compounds were produced. The synthesis of the libraries on solid phase was not successful.

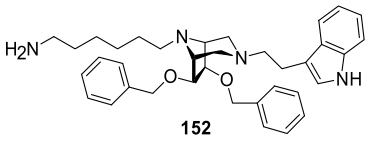
A selection of other orthogonally protected saccharide intermediates with potential for library synthesis is shown in *Fig. 11* [77].

3.4 Polycyclic Scaffolds and Bioactive Compounds Based on Monosaccharides

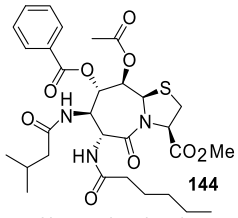
The generation of polycyclic systems from a monosaccharide (hybrid scaffolds) is a means of further increasing structural diversity [8,83,84,85,86,87,88,89,90,91,92,93,94,95,96]. Hirschmann, Smith and co-workers, for example, synthesised chimeric scaffolds such as **145** and **146** (*Fig. 12*) that are hybrids of β-D-glucopyranose and benzodiazepines and

[83] 

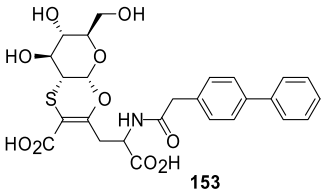
Tachykinin NK-1R antagonist

[8] 

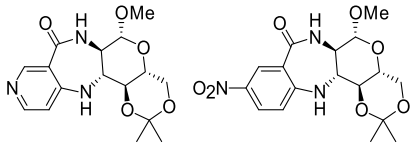
SSTR ligand

[84] 

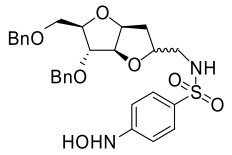
Hapalosin mimetic

[91] 

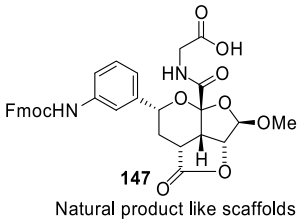
Matrix metalloproteinase inhibitor

[85] 

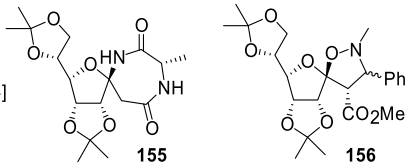
Sugar-benzodiazepines

[92] 

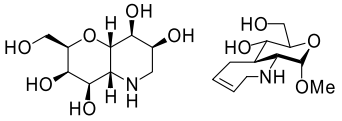
Ras protein activation inhibitor

[86] 

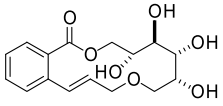
Natural product like scaffolds

[93-94] 

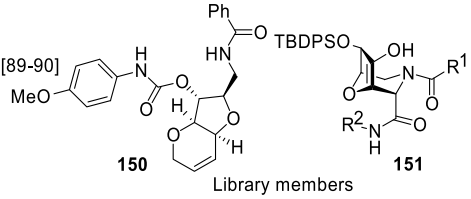
Spiro compounds

[87,88] 

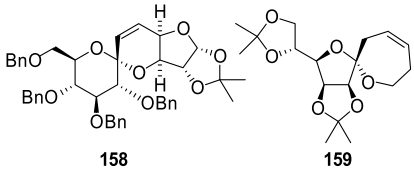
Pyranose-N-heterocycle hybrids

[95] 

Natural product like scaffolds

[89-90] 

Library members

[96] 

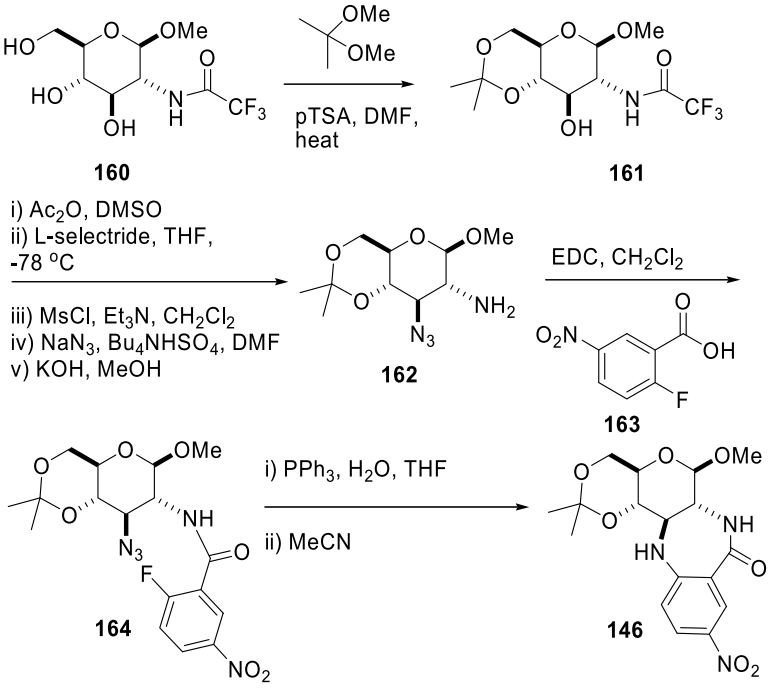
Spiro compounds

Figure 12
Examples of polycyclic scaffolds containing a monosaccharide component

rationalised the preparation of such compounds as providing a new molecular platform for exploration of receptor space, not accessible to pyranoside or benzodiazepine scaffolds when applied individually [85]. Other examples of such compounds are provided in *Fig. 12* of novel polycyclic systems that have been synthesised, some of which are bioactive and these include tachykinin NK-1 receptor antagonist [83], SSTR receptor ligands [8], hapalosin mimetics [84], matrix metalloproteinase inhibitor [91] and Ras protein activator inhibitors [92]. Also new scaffolds have been synthesised and these include natural product-like scaffolds [86,95], spirocyclic compounds [93,94,96,100], fused polycyclic compounds for development of libraries [89,90], hybrids of sugars and *N*-heterocycles [85,87,88].

3.5 Synthesis of Polycyclic Compounds Based on Monosaccharides that are not SAAs

The synthesis of **146** illustrates the general approach to the synthesis of the hybrids from β-D-glucopyranose and benzoheterodiazepines (*Scheme 14*). D-Glucosamine derivative **160** was converted to isopropylidene derivative **161**. The 3-OH group was formally replaced with retention of stereochemistry with an azide group. This was achieved by oxidation of the secondary alcohol followed by reduction with L-selectride giving an allosamine derivative. This was converted to the mesylate, substitution with azide followed by hydrolysis of trifluoroacetamide to give the amine **162**. Coupling of this amine to acid **163** gave **164**. The azido group



160 **161** **162** **163** **164** **146**

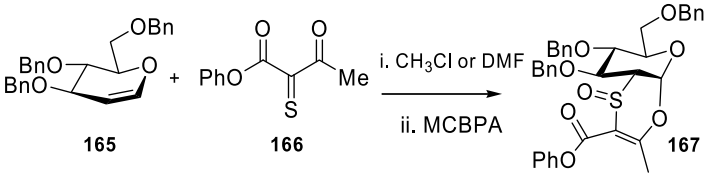
Scheme 14

was reduced using triphenylphosphine and water and cyclisation at high dilution gave the target compound **146** [85].

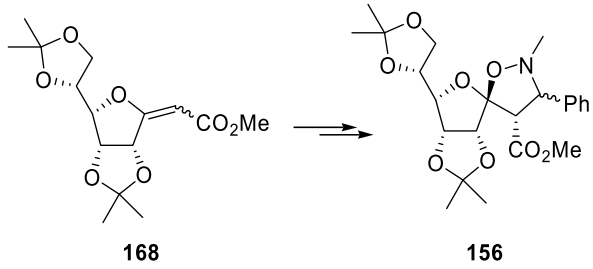
Cycloaddition reactions are also an expedient approach to the synthesis of polycyclic compounds. The synthesis of non-peptide ligands for the human tachykinin NK-2 receptor [83] was achieved by the chemo-, regio- and stereoselective [4+2] cycloaddition between glycals such as **165** and **166** followed by oxidation to provide sulfoxide **167** (● *Scheme 15*).

The cycloaddition of *exo*-glycals like **168** with nitrones and nitrile oxides provided spiroheterocylic compounds **156** (● *Scheme 16*) with good stereoselectivity and yields [93,94]. The *exo*-glycal intermediates proved also useful for the synthesis of spirocompounds based on a diazepine **155** (● *Fig. 12*).

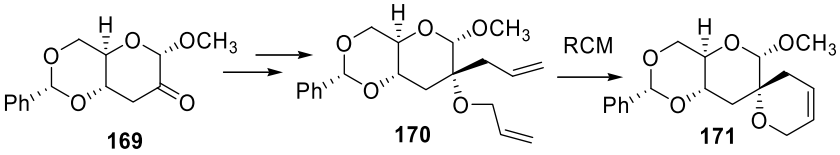
Radical cyclisations [97], Pauson–Kand cycloaddition [98] and palladium-catalysed coupling reactions [99] have been investigated to generate fused and spirocyclic carbohydrate scaffolds. Ring-closing metathesis (RCM) has recently become a major tool to produce cyclic compounds. Jenkins and co-workers have used RCM to generate five, six, seven, eight and nine-membered annelated fused sugar scaffolds and spirocyclic derivatives [100]. Reaction of ketones like **169** with various Grignard reagents of different chain length (four carbons in the



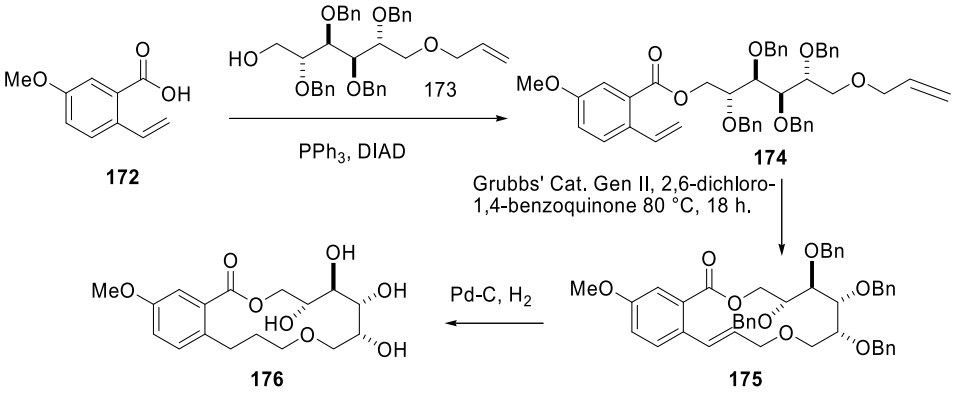
Scheme 15



Scheme 16



Scheme 17



Scheme 18

example shown in *Scheme 17*) led to olefins like **170**. Treatment with the Grubbs' catalyst gave the RCM products: in the example shown, spirofused ring derivative **171** was obtained. Using a similar methodology, the formation of a seven-membered fused ring and a seven-membered aza-sugar was also reported [88]. Leeuwenburgh, Overkleeft and co-workers have also taken advantage of the RCM methodology for the synthesis and elaboration of functionalised carbohydrate-derived spiroketals like **158** and **159** [96].

The benzomacrolactone (or resorcyclic acid lactone) structural motif can be considered a privileged or evolutionarily selected scaffold that codes properties required for binding to proteins and novel analogues thereof may provide a source of new bioactive compounds. Based on this premise the syntheses of novel polyhydroxylated oxa-macrolides, structural analogues of natural polyketide derived macrolides, have recently been described from aromatic precursors by Murphy and Matos, providing a basis for their development as scaffolds [95]. The syntheses were carried out from coupling of benzoic acids like **172** and appropriately protected D-mannitol or D-sorbitol derivatives like **173** by Mitsunobu reactions to give **174**, followed by RCM and hydrogenation to yield **175** and **176**, respectively (*Scheme 18*). Work towards application of such hybrids towards synthesis of bioactive compounds is currently underway.

4 Conclusion

There has been considerable progress in application of monosaccharide derivatives as scaffolds for bioactive molecule discovery and development and research has not been limited to discovery of compounds that target protein-carbohydrate interactions, but has led to carbohydrate derivatives that also modulate protein-protein, peptide-protein and carbohydrate-nucleic acid interactions. Progress has been facilitated by, and is continually dependant on, advances in synthetic carbohydrate chemistry both in solution and on polymer supports. Of particular interest are the development of orthogonal protecting group strategies, strategies for regioselective manipulation of saccharide hydroxyl groups and development of saccharide derivatives with a wide variety of functional groups which provides possibilities for carrying out chemose-

lective reactions on monosaccharide scaffolds. The synthesis of new cyclic compounds derived at least in part from saccharides, and their application as scaffolds is an emerging area in novel compound development. The introduction of sugar amino acids has simplified the synthesis of non-peptide peptidomimetics and libraries based on monosaccharides as well as leading to the synthesis of conformationally constrained sugar-peptide hybrids and carbohydrate foldamers and facilitates the investigation of properties of such novel derivatives. No drug molecule has been developed as of yet based on a monosaccharide scaffold. However, this goal is closer as potent ligands for receptors have been identified, it has been shown that some saccharide derivatives have improved cellular permeability over peptides and some compounds are in clinical trials.

Acknowledgements

The authors gratefully acknowledge funding received from Science Foundation Ireland and Astellas USA Foundation.

References

1. Simanek EE, McGarvey GJ, Jablonski JA, Wong CH (1998) Chem Rev 98:833
2. The first non-peptidal peptidomimetic was reported in 1986 by Belanger and Dufresne, which recognized the opiate receptor for which it was designed. See Belanger PC, Dufresne C (1986) Can J Chem 64:1514
3. Nicolaou KC, Salvino JM, Raynor K, Pietranico S, Reisine T, Freidinger RM, Hirschmann R (1990) In: Rivier JE, Marshall GR (eds) Peptides–Chemistry, Structure, and Biology: Proceedings of the 11th American Peptide Symposium. ESCOM, Leiden, p 881
4. Farmer PS (1980) In: Ariens EJ (ed) Drug Design, vol 10. Academic Press, New York, p 119
5. For some recent reviews see:
   a) Gentilucci L, Tolomelli A, Squassebia F (2006) Curr Med Chem 13:2449;
   b) Murphy PV, Dunne JL (2006) Curr Org Synth 3:403;
   c) Randolph JT, De Goey DA (2004) Curr Topics Med Chem 4:1079;
   d) Patch JA, Barron AE (2002) Curr Op Chem Biol 6:872
6. Thanh Le G, Abbenante G, Becker B, Grathwohl M, Halliday J, Tometzki G, Zuegg J, Meutermans W (2003) Drug Disc Today 8:701
7. Hirschmann R, Cichy-Knight MA, van Rijn RD, Sprengler PA, Spoors PG, Shakespere WC, Pietranico-Cole S, Barbosa J, Liu J, Yao W, Roher S, Smith AB III (1998) J Med Chem 41:1382
8. Le Merrer Y, Poitout L, Depezay JC, Dosbaa I, Geoffroy S, Foglietti MJ (1997) Bioorg Med Chem Lett 6:1667
9. Graf von Roedern E, Lohof E, Hessler G, Hoffmann M, Kessler H (1996) J Am Chem Soc 118:10156
10. For some recent syntheses of disaccharide libraries:
    a) Sofia MJ, Allanson N, Hatzenbuhler NT, Jain R, Kakarla R, Kogan N, Liang R, Liu DS, Silva DJ, Wang HM, Gange D, Anderson J, Chen A, Chi F, Dulina R, Huang BW, Kamau M, Wang CW, Baizman E, Branstrom, A, Bristol N, Goldman R, Han KH, Longley C, Midha S, Axelrod HR (1999) J Med Chem 42:3193;
    b) Baizman ER, Branstrom AA, Longley CB, Allanson N, Sofia MJ, Gange D, Goldman RC (2000) Microbiology-SGM 146:3129;
    c) Castoldi S, Cravini M, Micheli F, Piga E, Russo G, Seneci P, Lay L (2004) Eur J Org Chem 2853;
    d) Venot A, Swayze EE, Griffey RH, Boons GJ (2004) Chem Bio Chem 5:1228
11. a) Sofia MJ (1998) Mol Diversity 3:75;
    b) Sofia MJ (1998) Med Chem Res 8:362;
    c) Sofia MJ, Silva DJ (1999) Curr Opin Drug Discov Dev 2:365;
    d) Schweizer F, Hindsgaul O (1999) Curr Opin Chem Biol 3:29;

e) Gruner SAW, Locardi E, Lohof E, Kessler H (2002) Chem Rev 102:491;
f) Schweizer F (2002) Angew Chem Int Ed 41:231;
g) Marcaurelle LA, Seeberger PH (2002) Curr Opin Chem Biol 6:289;
h) Chakraborty TK, Srinivasu P, Tapadar S, Mohan BK (2004) J Chem Sci 116:187;
i) Schweizer F (2003) Trends Glycosci Glycotech 15:315;
j) Jensen KJ, Brask J (2005) Biopolymers 80:747;
k) Velter I, La Ferla B, Nicotra, F (2006) J Carbohydr Chem 25:97;
l) Becker B, Condie GC, Le GT, Meutermans W (2006) Mini-Reviews Med Chem 6:1299;
m) Meutermans W, Le GT, Becker B (2006) ChemMedChem 1:1164

12. Chakraborty TK, Srinivasu P, Tapadar S, Mohan BK (2005) Glycoconjugate J 22:83
13. a) von Roedern EG, Kessler H (1994) Angew Chem Int Ed Engl 33:670;
b) Stockle M, Voll G, Gunther R, Lohof E, Locardi E, Gruner S, Kessler H (2002) Org Lett 4:2501;
c) Locardi E, Stockle M, Gruner SAW, Kessler H.(2001) J Am Chem Soc 123:8189
14. Sicherl F, Wittmann V (2005) Angew Chem Int Ed 44:2096
15. a) Raunkjær M, El Oualid F, van der Marel GA, Overkleeft HS, Overhand M (2004) Org Lett 6:3167;
b) Kriek NMAJ, van der Hout E, Kelly P, van Meijgaarden KE, Geluk A, Ottenhoff THM, van der Marel GA, Overhand M, van Boom JH, Valentijn ARPM, Overkleeft HS (2003) Eur J Org Chem 2418
16. Montero A, Mann E, Herradon B (2004) Eur J Org Chem 3063
17. a) Sofia MJ, Hunter R, Chan TY, Vaughan A, Dulina R, Wang H, Gange D (1998) J Org Chem 63:2802;
b) Jain R, Kamau M, Wang C, Ippolito R, Wang H, Dulina R, Anderson J, Gange D, Sofia MJ (2003) Bioorg Med Chem Lett 13:2185
18. Schweizer F, Hindsgaulb O (2006) Carbohydr Res 341:730
19. Gregar TQ, Gervay-Hague J (2004) J Org Chem 69:1001
20. El Oualid F, Burm BEA, Leroy IM, Cohen LH, van Boom, van den Elst H, Overkleeft HS, van der Marel GA, Overhand M (2004) J Med Chem 47:3920
21. Czifrak K, Szilagyi P, Somsak L (2005) Tetrahedron: Asymm 16:127
22. McDevitt JP, Lansbury PT Jr (1996) J Am Chem Soc 118:3818
23. Smith AB, Sasho S, Barwis BA, Sprengeler P, Barbosa, J, Hirschmann R, Cooperman BS (1998) Bioorg Med Chem Lett 8:3133
24. a) Edwards AA, Ichihara O, Murfin S, Wilkes R, Whittaker M, Watkin DJ, Fleet GWJ (2004) J Comb Chem 6:230;
b) Watterson MP, Edwards AA, Leach JA, Smith MD, Ichihara O, Fleet GWJ (2003) Tetrahedron Lett 44:5853
25. a) Grotenbreg GM, Timmer MSM, Llamas-Saiz AL, Verdoes M, van der Marel GA, van Raaij MJ, Overkleeft HS, Overhand M (2004) J Am Chem Soc 126:3444;
b) Grotenbreg GM, Buizert AEM, Llamas-Saiz AL, Spalburg E, van Hooft PAV, de Neeling AJ, Noort D, van Raaij MJ, van der Marel GA, Overkleeft HS, Overhand M (2006) J Am Chem Soc 128:7559
26. a) Chakraborty TK, Srinivasu P, Bikshapathy E, Nagaraj R, Vairamani M, Kiran Kumar S, Kunwar AC (2003) J Org Chem 68:6257;
b) Chakraborty TK, Jayaprakash S, Diwan PV, Nagaraj R, Jampani SRB, Kunwar AC (1998) J Am Chem Soc 120:12962
27. Becattini B, Capozzi G, Falciani C, Menichetti S, Nativi C, Salvini A (2000) J Carbohydr Chem 19:653
28. a) Long DD, Hungerford NL, Smith MD, Brittain DEA, Marquess DG, Claridge TDW, Fleet GWJ (1999) Tetrahedron Lett 40: 2195;
b) Claridge TDW, Long DD, Hungerford NL, Aplin RT, Smith MD, Marquess DG, Fleet GWJ (1999) Tetrahedron Lett 40:2199;
c) Long DD, Stetz RJE, Nash RJ, Marquess DG, Lloyd JD, Winters AL, Asano N, Fleet GWJ (1999) J Chem Soc Perkin Trans 1:961;
d) Claridge TDW, Long DD, Baker CM, Odell B, Grant GH, Edwards AA, Tranter GE, Fleet GWJ, Smith MD (2005) J Org Chem 70:2082;
e) Mayes BA, Stetz RJE, Ansell CWG, Fleet GWJ (2004) Tetrahedron Lett 45:153;
f) Mayes BA, Simon L, Watkin DJ, Ansell CWG, Fleet GWJ (2004) Tetrahedron Lett 45:157;
g) Mayes BA, Cowley AR, Ansell CWG, Fleet GWJ (2004) Tetrahedron Lett 45:163

29. a) Sanjayan GJ, Stewart A, Hachisu S, Gonzalez R, Watterson MP, Fleet GWJ (2003) Tetrahedron Lett 44:5847;
b) Edwards AA, Sanjayan GJ, Hachisu S, Soengas R, Stewart A, Trantera GE, Fleet GWJ (2006) Tetrahedron 62:4110

30. a) Chakraborty TK, Reddy VR, Sudhakar G, Kumar SU, Reddy TJ, Kumar SK, Kunwar AC, Mathur M, Sharma R, Guptab N, Prasadb S (2004) Tetrahedron 60:8329;
b) Prasad S, Mathur A, Jaggi M, Sharma R, Gupta N, Reddy VR, Sudhakar G, Kumar SU, Kumar SK, Kunwar AC, Chakraborty TK (2005) J Peptide Res 66:75

31. Vera-Ayoso Y, Borrachero P, Cabrera-Escribano F, Gomez-Guillen M (2005) Tetrahedron: Asymm 16:889

32. a) Jenkinson (nee Barker) SF, Harris T, Fleet GWJ (2004) Tetrahedron Asymm 15:2667;
b) Johnson SW, Jenkinson (nee Barker) SF, Angus D, Jones JH, Fleet GWJ, Taillefumier C (2004) Tetrahedron: Asymm 15:2681;
c) Johnson SW, Jenkinson (nee Barker) SF, Perez-Victoria I, Edwards AA, Claridge TDW, Tranter GE, Fleet GWJ, Jones JH (2005) J Peptide Sci 11:517;
d) Johnson SW, Jenkinson (nee Barker) SF, Angus D, Perez-Victoria I, Claridge TDW, Fleet GWJ, Jones JH (2005) J Peptide Sci 11:303;
e) Johnson SW, Jenkinson (nee Barker) SF, Angus D, Jones JH, Watkin DJ, Fleet GWJ (2004) Tetrahedron: Asymm 15:3263;
f) Barker SF, Angus D, Taillefumier C, Probert MR, Watkin DJ, Watterson MP, Claridge TDW, Hungerford NL, Fleet GWJ (2001) Tetrahedron Lett 42:4247;
g) Claridge TDW, Goodman JM, Moreno A, Angus D, Barker SF, Taillefumier C, Watterson MP, Fleet GWJ (2001) Tetrahedron Lett 42: 4251

33. a) Forni E, Cipolla L, Caneva E, La Ferla B, Peri F, Nicotra F (2002) Tetrahedron Lett 43:1355;
b) Cipolla L, Forni E, JimPnez-Barbero J, Nicotra F (2002) Chem Eur J 8:3976;
c) Cipolla L, Peri F, La Ferla B, Redaelli C, Nicotra F (2005) Curr Org Synth 2:153

34. a) Risseeuw MDP, Grotenbreg GM, Witte MD, Tuin AW, Leeuwenburgh MA, Van der Marel GA, Overkleeft HS, Overhand M (2006) Eur J Org Chem 3877;
b) Grotenbreg GM, Tuin AW, Witte MD, Leeuwenburgh MA, van Boom JH, van der Marel GA, Overkleeft HS, Overhand M (2004) Synlett 5:904

35. a) Chakraborty TK, Srinivasu P, KiranKumar S, Kunwar AC (2002) J Org Chem 67:2093;
b) Chakraborty TK, Ghosh A, Nagaraj R, Ravi Sankar A, Kunwar AC (2001) Tetrahedron 57:9169

36. La Ferla B, Bugada P, Cipolla L, Peri F, Nicotra F (2004) Eur J Org Chem 2451

37. a) Moriyama H, Tsukida T, Inoue Y, Kondo H, Yoshino K, Nishimura S (2003) Bioorg Med Chem Lett 13:2737;
b) Moriyama H, Tsukida T, Inoue Y, Yokota K, Yoshino K, Kondo H, Miura N, Nishimura S (2004) J Med Chem 47:1930

38. a) Timmer MSM, Verhelst SHL, Grotenbreg GM, Overhand M, Overkleeft HS (2005) Pure Appl Chem 77:1173;
b) Timmer MSM, Risseeuw MDP, Verdoes M, Filipov DV, Plaisier JR, van der Marel GA, Overkleeft HS, van Boom JH (2005) Tetrahedron: Asymm 16:177

39. a) IJsselstijn M, Aguilera B, van der Marel G, van Boom JH, van Delft FL, Schoemaker HE Overkleeft HS, Rutjesa FPJT, Overhand (2004) Tetrahedron Lett 45:4379;
b) Grotenbreg GM, Christina AE, Buizert AEM, van der Marel GA, Overkleeft HS, Overhand M (2004) J Org Chem 69:8331;
c) Bongera KM, Wennekesa T, de Lavoir SVP, Esposito D, van den Berg RJBHN, Litjens REJN, van der Marel GA, Overkleeft HS (2006) QSAR Comb Sci 25:491;
d) Grotenbreg GM, Kronemeijer M, Timmer MSM, El Oualid F, van Well RM, Verdoes M, Spalburg E, van Hooft PAV, de Neeling AJ, Noort D, van Boom JH, van der Marel GA, Overkleeft HS, Overhand M (2004) J Org Chem 69:7851

40. Peri F, Bassetti R, Caneva E, De Giola I, La Ferla B, Presta M, Tanghetti E, Nicotra F (2002) J Chem Soc Perkin Trans 1:638

41. a) Durrat F, Xie J, Valery JM (2004) Tetrahedron Lett 45:1477;
b) Dondoni A, Marra A (2000) Chem Rev 100:4395;
c) Schrey A, Osterkamp F, Straudi A, Rickert C, Wagner H, Koert U, Herrschaft B, Harms K (1999) Eur J Org Chem 2977

42. Gellman SH (1998) Acc Chem Res 31:173

43. Szabo L, Smith BL, McReynolds KD, Parrill AL, Morris ER, Gervay J (1998) J Org Chem 63:1074
44. a) Suhara Y, Yamaguchi Y, Collins B, Schnaar RL, Yanagishita M, Hildreth JEK, Shimada I, Ichikawa Y (2002) Bioorg Med Chem 10:1999;
b) Suhara Y, Hildreth JEK, Ichikawa Y (1996) Tetrahedron Lett 37:1575
45. For other applications of monosaccharides as scaffolds in synthesis of multivalent compounds see: Patel A, Lindhorst TK (2001) J Org Chem 66:2674
46. Papageorgiou C, Haltiner R, Bruns C, Petcher TJ (1992) Bioorg Med Chem Lett 2:135
47. Gouin SG, Murphy PV (2005) J Org Chem 70:8527
48. Diguarher TL, Boudon A, Elwell C, Paterson DE, Billington DC (1996) Bioorg Med Chem Lett 16:1983
49. Wessel HP, Banner D, Gubernator K, Hilpert K, Muller K, Tschopp T (1997) Angew Chem Int Ed Engl 36:751
50. Dinh TG, Smith CD, Du X, Armstrong RWJ (1998) J Med Chem 41:981
51. Van Hoof S, Ruttens B, Hubrecht I, Smans G, Blom P, Sas B, van Hemel J, Vandenkerckhove J, Van der Eycken J (2006) Bioorg Med Chem Lett 16:1495
52. a) Hirschmann R, Nicolaou KC, Pietranico S, Salvino J, Leahy EM, Sprengeler PA, Furst G, Smith AB III, Strader CD, Cascieri MA, Candelore MR, Donaldson C, Vale W, Maechler L (1992) J Am Chem Soc 114:9217;
b) Hirschmann R, Nicolaou KC, Pietranico S, Leahy EM, Salvino J, Arison B, Cichy MA, Spoors PG, Shakepseare WC, Sprengler PA, Hamley P, Smith AB, Reisine T, Raynor K, Maechler L, Donaldson C, Vale W, Freidinger RM, Cascieri MR, Strader CD (1993) J Am Chem Soc 115:12550
53. Prasad V, Birzin ET, McVaugh CT, van Rijn RD, Rohrer SP, Chicci G, Underwood DJ, Thornton ER, Smith AB III, Hirschmann R (2003) J Med Chem 46:1858
54. a) Boer J, Gottschling D, Schuster A, Holzmann B, Kessler H (2001) Angew Chem Int Ed 40:3870;
b) Locardi E, Boer J, Modlinger A, Schuster A, Holzmann B, Kessler H (2003) J Med Chem 46:5752
55. Nicolaou KC, Trujillo JI, Chibale K (1997) Tetrahedron 53:8751
56. Moitessier N, Dufour S, Chretien F, Thiery JP, Maigret B, Chapleur Y (2001) Bioorg Med Chem 9:511
57. a) Murphy PV, O' Brien JL, Gorey-Feret LJ, Smith III AB (2003) Tetrahedron 59:2259;
b) Murphy PV, O' Brien JL, Gorey-Feret LJ, Smith III AB (2002) Bioorg Med Chem Lett 12:1763
58. a) Chery F, Murphy PV (2004) Tetrahedron Lett 45:2067;
b) Chery F, Cronin L, O' Brien JL, Murphy PV (2004) Tetrahedron 60:6597
59. Kurihara S, Tsumuraya T, Fujii I (1999) Bioorg Med Chem Lett 9I:1179
60. Cai J, Davison BE, Ganellin CR, Thaisrivongs S, Wibley KS (1997) Carbohydr Res 300:109
61. Moitessier N, Minoux H, Maigret B, Chretien F, Chapleur Y (1996) Lett Pept Sci 5:75
62. Tamaruya Y, Suzuki M, Kamura G, Kanai M, Hama K, Shimizu K, Aoki J, Arai H, Shibasaki M (2004) Angew Chem Int Ed 43:2834
63. a) Nakayama K, Kawato HC, Inagaki H, Ohta T (2001) Org Lett 3:3447;
b) Nakayama K, Kawato HC, Inagaki H, Ohta T (2001) Org Lett 3:3451
64. a) Chakraborty TK, Ghosh S, Jayaprakash S, Sarma JARP, Ravikanth V, Diwan PV, Nagaraj R, Kunwar AC (2000) J Org Chem 65: 6441;
b) Chakraborty TK, Ghosh S, Rao MHVR, Kunwar AC, Cho H, Ghosh AK (2000) Tetrahedron Lett 41:10121
65. Montero A, Mann E, Herradón B (2005) Tetrahedron Lett 46:401
66. Wong CH, Hendrix M, Manning DD, Rosenbohm C, Greenberg WA (1998) J Am Chem Soc 120:8319
67. Hanessian S, Saavedra OM, Xie F, Amboldi N, Battistini C (2000) Bioorg Med Chem Lett 10:439
68. Kitov PI, Sadowska JM, Mulvey G, Armstrong GD, Ling H, Pannu NS, Read RJ, Bundle DR (2000) Nature 403:669
69. a) Brask J, Jensen KJ (2000) Peptide Sci 6:290;
b) Brask J, Jensen KJ (2001) Bioorg Med Chem Lett 11:697
70. For the synthesis of glycosidase inhibitors based on iminosugar scaffolds see:
a) Chapman TM, Courtney S, Hay P, Davis BG (2003) Chem Eur J 9:3397;
b) Gerber-Lemaire S, Popowycz F, Rodrguez-Garcia E, Carmona Asenjo AT, Robina I, Vogel P (2002) Chem Bio Chem 5:466

71. Wunberg T, Kallus C, Opatz T, Henke S, Schmidt W, Kunz H (1998) Angew Chem Int Ed Engl 37:2503
72. a) Kallus C, Opatz T, Wunberg T, Schmidt W, Henke S, Kunz H (1999) Tetrahedron Lett 40:7783;
b) Opatz T, Kallus C, Wunberg T, Schmidt W, Henke S, Kunz H (2002) Carbohydr Res 337:2089;
c) Opatz T, Kallus C, Wunberg T, Schmidt W, Henke S, Kunz H (2003) Eur J Org Chem 1527;
d) Opatz T, Kallus C, Wunberg T, Kunz H (2004) Tetrahedron 60:8613;
e) Ohnsmann J, Madalinski M, Kunz H (2005) Chim Oggi 20
73. Huenger U, Ohnsmann J, Kunz H (2004) Angew Chem Int Ed 43:1104
74. Brill WKD, Tirefort D (1998) Tetrahedron Lett 39:787
75. a) Brill WKD, De Mesmaeker A, Wendeborn S (1998) Synlett 1085;
b) Wendeborn S, De Mesmaeker A, Brill WKD, Berteina S (2000) Acc Chem Res 33:215
76. Krueger EB, Hopkins TP, Keaney MT, Walters MA, Boldi AM (2002) J Comb Chem 4:229
77. For intermediates that contain orthogonal protecting groups and that have been used for oligosaccharide library synthesis see: Wong CH, Ye XS, Zhang Z (1998) J Am Chem Soc 120:7137
78. Moitessier N, Henry C, Aubert N, Chapleur Y (2005) Tetrahedron Lett 46:6191
79. Hirschmann R, Ducry L, Smith III AB (2000) J Org Chem 65:8307
80. Reina JJ, Rojo J (2006) Tetrahedron Lett 47:2475
81. Svejgaard L, Fuglsang H, Jensen PB, Kelly NM, Pedersen H, Andersen K, Ruhland T, Jensen KJ (2003) J Carbohydr Chem 22:179
82. Danieli E, Lalot J, Murphy P v (2007) Tetrahedron, 63 6827-34
83. a) Capozzi G, Giannini S, Menichetti S, Nativi C, Giolitti A, Patacchini R, Perrotta E, Altamura M, Maggi CA (2002) Bioorg Med Chem Lett 12:2263;
b) Maggi CA, Patacchini R, Rovero P, Giachetti AJ (1993) Pharmacol 13:23
84. Agoston K, Geyer A (2005) Chem Eur J 11:6407
85. a) Abrous L, Hynes Jr J, Friedrich SR, Smith AB III, Hirschmann R (2001) Org Lett 3:1089;
b) Abrous L, Jokiel PA, Friedrich SR, Hynes Jr J, Smith AB III, Hirschmann RJ (2004) Org Chem 69:280
86. a) Messer R, Schmitz A, Moesch L, Häner R (2004) J Org Chem 69:8558;
b) Messer R, Pelle X, Marzinzik AL, Lehmann H, Zimmermann J, Häner R (2005) Synlett 16:2441
87. Reddy BG, Vankar YD (2005) Angew Chem Int Ed 44:2001
88. Laventine DM, Jenkins PR, Cullis PM (2005) Tetrahedron Lett 46:2295
89. Timmer MSM, Verdoes M, Sliedregt LA, van der Marel GA, van Boom JH, Overkleeft HS (2003) J Org Chem 68:9406
90. Timmer MSM, Verhelst SHL, Grotenbreg GM, Overhand M, Overkleeft HS (2005) Pure Appl Chem 77:1173
91. Fragai M, Nativi C, Richichi B, Venturi C (2005) ChemBioChem 6:1345
92. a) Peri F, Airoldi C, Colombo S, Martegani E, van Neuren AS, Stein M, Marinzi C, Nicotra F (2005) ChemBioChem 6:1839;
b) Peri F,Airoldi F, Colombo S, Mari S, Jiménez-Barbero J, Martegani E, Nicotra F (2006) Eur J Org Chem 3707
93. Taillefumier C, Triegles S, Chapleur Y (2004) Tetrahedron 60:2213
94. Taillefumier C, Enderlin G, Chapleur Y (2005) Lett Org Chem 2:226
95. Matos MC, Murphy PV (2007) J Org Chem 72:1803–1806
96. van Hooft PAV, El Oualid F, Overkleeft HS, van der Marel GA, van Boom JH, Leeuwenburgh MA (2004) Org Biomol Chem 2:1395
97. For some examples see: a) Chiara JL, Garcia A, Cristobal-Lumbroso G (2005) J Org Chem 70:4142;
b) Lopez JC, Fraser-R B (1997) J Chem Soc Chem Commun 2251;
c) Wood AJ, Jenkins PR (1997) Tetrahedron Lett 38:1853
98. Leeuwenburgh MA, Appeldoom CCM, van Hooft PAV, Overkleeft HS, van der Marel GA, van Boom JH (2000) Eur J Org Chem 873
99. Sirou D, Bedjeguelal K (2000) Eur J Org Chem 4071
100. a) Holt DJ, Barker WD, Jenkins PR, Panda J, Ghosh S (2000) J Org Chem 65:482;
b) Holt DJ, Barker WD, Jenkins PR, Davies DL, Garratt S, Fawcett J, Russell DL, Ghosh S (1998) Angew Chem Int Ed Engl 37:3298

4.6 Monosaccharides as Chiral Auxiliaries and Ligands for Asymmetric Synthesis

Kiichiro Totani[1], *Kin-ichi Tadano**[2]
[1] RIKEN (The Institute of Physical and Chemical Research), 2-1 Hirosawa, Wako-shi, 351-0198 Saitama, Japan
[2] Department of Applied Chemistry, Keio University, Hiyoshi, Kohoku-ku, 223-8522 Yokohama, Japan
totani@riken.jp, tadano@applc.keio.ac.jp

Abstract

This chapter summarizes some recent achievements in asymmetric synthesis, which have been accomplished using sugar derivatives as chiral environments. The use of sugar derivatives as chiral auxiliaries or as chiral ligands of transition-metal catalysts for realizing the asymmetric version of synthetically important organic reactions, such as carbon–carbon bond-forming reactions, the hydrogenation of the carbon–carbon double bond and hydrocyanation, is the focus of this chapter. Prominent studies on these subjects by several research groups are especially emphasized.

In: *Glycoscience*. Fraser-Reid B, Tatsuta K, Thiem J (eds)
Chapter-DOI 10-1007/978-3-540-30429-6_23:

Keywords

Asymmetric synthesis; Chiral auxiliary; Sugar template; Chiral ligand; Transition-metal catalyst

Abbreviations

acac	acetylacetone
BINOL	1,1′-bi-2-naphthol
Bn	benzyl
Boc	*t*-butoxycarbonyl
BSA	*N*,*O*-bis(trimethylsilyl)acetamide
***i*Bu**	*iso*-butyl
***t*Bu**	*tert*-butyl
COD	1,5-cyclooctadiene
de	diastereomeric excess
DIBAL-H	diisobutylaluminum hydride
ee	enantiomeric excess
Et	ethyl
GABA	γ-aminobutyric acid
***c*Hex**	cyclohexyl
HMPA	hexamethylphosphoramide
LDA	lithium diisopropylamide
LiHMDS	lithium hexamethyldisilazide
Me	methyl
MesCO	2,4,6-trimethylbenzoyl
NaHMDS	sodium hexamethyldisilazide
Napht	naphthyl
NBS	*N*-bromosuccinimide
NMO	*N*-methylmorpholine-*N*-oxide
NMR	nuclear magnetic resonance
Ph	phenyl
Piv	*tert*-butylcarbonyl
***c*Pr**	cyclopropyl
***i*Pr**	*iso*-propyl
***n*Pr**	*n*-propyl
RMP	(*R*)-2-methoxymethylpyrrolidine
SMP	(*S*)-2-methoxymethylpyrrolidine
TBS	*tert*-butyldimethylsilyl
TFA	trifluoroacetic acid
THF	tetrahydrofuran
TIPS	triisopropylsilyl
TMS	trimethylsilyl

1 Introduction

The development of asymmetric synthesis is the most challenging subject in current organic synthesis, as various approaches have been demonstrated by the synthetic community in the world. The most frequently employed strategies for asymmetric synthesis are categorized into two approaches, i. e., the use of catalytic chiral ligands such as BINOLs, which coordinate a variety of transition-metal complexes (intermolecular asymmetric induction) [1,2,3] or the use of chiral auxiliaries derived from readily available natural products, such as L-amino acids (intramolecular asymmetric induction) [4,5,6,7,8,9,10,11]. Concerning the former approach, the use of chiral Lewis acids, prepared by a combination of a chiral ligand and a transition metal, has been most extensively explored for carbon–carbon or carbon–heteroatom bond formation in an asymmetric fashion. On the other hand, the use of chiral auxiliaries is not as generalized as the catalytic use of chiral ligands. The inevitable use of an equivalent amount of the chiral auxiliary seems to cause this rather unpopular view of the chiral auxiliary approach. Despite this fact, the use of a chiral spatial environment constituted by a chiral auxiliary is one of the most effective means for achieving a practical asymmetric synthesis. The use of the chiral auxiliary conceptually produces a diastereomeric mixture. In general, the separation of diastereomers is not very troublesome when the current sophisticated separation methods are used. In many cases, the observed diastereomeric ratio in an organic reaction is conventionally measured by spectroscopic means, such as NMR spectroscopy. These are advantages of the chiral auxiliary approach over other asymmetric approaches. This article focuses on the development of the use of sugars for asymmetric synthesis in the past 15 years. Compared to L-amino acids, natural alkaloids, or terpenoids, sugars have been less developed as a chiral environment [12,13,14,15]. On the other hand, sugars have been recognized as excellent molecular-recognizing substances, as exemplified by the fact that glycoproteins and oligosaccharides participate in the modulation of biological function [16].

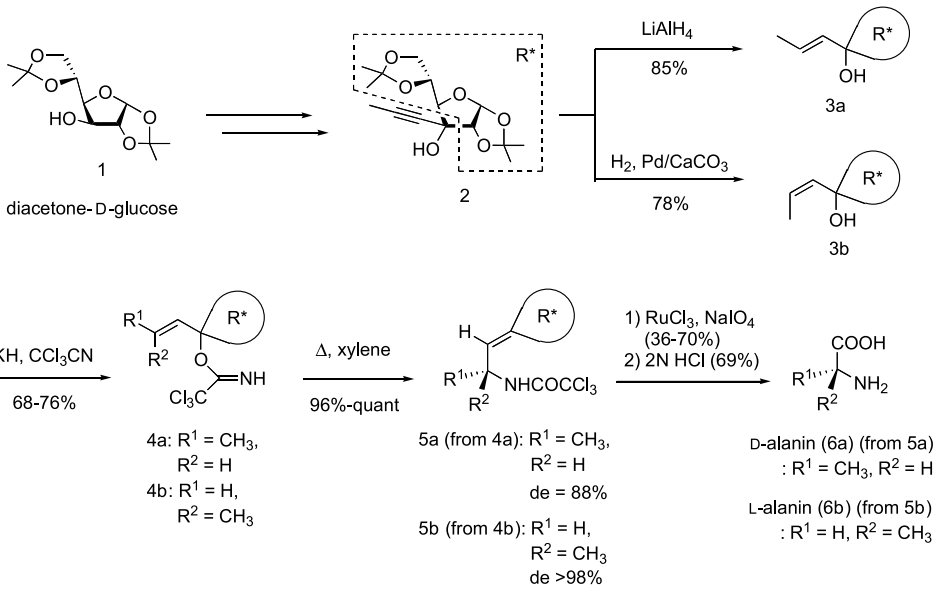
In this chapter, we summarize the above two asymmetric approaches achieved using sugar derivatives in these two decades in the order of “sugars as chiral auxiliaries” and then “sugars as chiral ligands for transition-metal-mediated organic synthesis”. In particular, the present work reviews and summarizes the work of prominent research groups whose main interest is asymmetric synthesis using sugar-based compounds. For the sake of brevity, less complete studies have not been included in this review.

2 Sugars as Chiral Auxiliaries

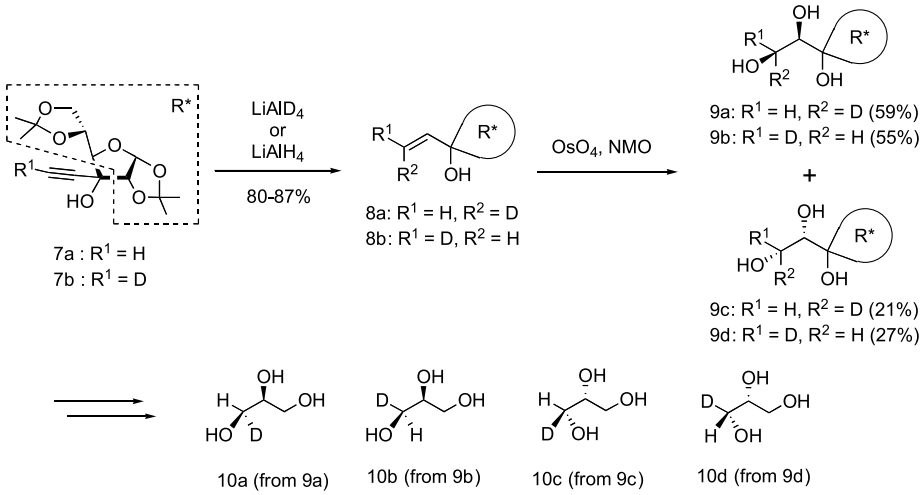
In the mid-1970s, Vasella reported the 1,3-dipolar cycloaddition reactions of nitrones incorporated into sugar templates [17,18]. The Vasella’s studies were regarded as seminal for the development of the stereoselective organic reactions achieved on sugar templates. Sugar-template-based stereoselective reactions have been actively investigated by a number of groups, especially in the past ten years [19,20,21,22,23,24,25,26,27,28,29,30,31,32]. The sugar-based templates utilized for asymmetric synthesis are mainly classified into five-membered glycofuranosidic frameworks or six-membered glycopyranosidic frameworks. Among glycofuranosidic templates, the utility of so-called diacetone-D-glucose i. e., 1,2:5,6-di-*O*-isopropylidene-α-D-glucofuranose, has been extensively investigated.

2.1 Studies by the Kakinuma Group

Kakinuma and co-workers have reported several results obtained using diacetone-D-glucose [33]. In the early 1990s, the Kakinuma group started to investigate the utility of diacetone-D-glucose as a useful chiral template. As an initial attempt, they used diacetone-D-glucose for the stereoselective synthesis of D- or L-alanine [34,35] (*Scheme 1*). Compound **2** was readily prepared from commercially available diacetone-D-glucose **1** by (1) oxidation to 3-ulose, (2) stereoselective addition of an ethynyl Grignard reagent, (3) temporary *O*-silylation, and (4) Li acetylide formation with BuLi and methylation with CH_3I. $LiAlH_4$ reduction of the triple bond in **2** introduced the *trans*-propenyl functionality to provide **3a** stereoselectively. On the other hand, hydrogenation of **2** in the presence of Lindlar catalyst provided the *cis*-propenyl derivative **3b**. Both **3a** and **3b** were converted to trichloroacetimidates **4a** and **4b**, respectively. The thermal Overman rearrangement (an aza version of the [3,3]-sigmatropic rearrangement) of **4a** proceeded highly stereoselectively to provide **5a** as a single rearranged product. Similarly, the Overman rearrangement of **4b** provided **5b** stereoselectively. Both stereochemistries of the double bond in substrates **4a** and **4b** completely reflected the newly created chiral center in **5a** and **5b**, respectively. Finally, the double bond in **5a** was cleaved by $RuCl_3/NaIO_4$ oxidation, followed by deprotection, providing D-alanine **6a**. Similarly, another rearranged product, **5b**, was transformed into L-alanine **6b**. Based on the kinetic and theoretical studies, the Kakinuma group proposed transition-state models to explain these highly selective chirality transfers observed in the Overman rearrangements of both **4a** and **4b**. Consequently, an enantiopure enantiomer of amino acid alanine as a tool for biological studies was prepared.



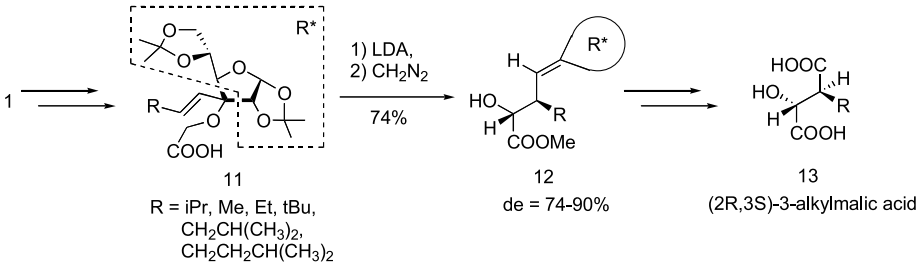
Scheme 1



Scheme 2

Using the diacetone-D-glucose derivatives **7a** or **7b** incorporating terminal hydrogen or deuterium labeled ethenyl functionality at C-3, respectively, four diastereomers of monodeuterated glycerols were synthesized [36] (*Scheme 2*). *trans*-Selective $LiAlD_4$ reduction of **7a** provided **8a**. $LiAlH_4$ reduction of the deuterium-labeled **7b** provides **8b**. *cis*-Diol unit introduction into **8a** and **8b** was carried out by OsO_4-mediated oxidation in the presence of *N*-methylmorpholine *N*-oxide. Two diols, **9a** and **9c**, were obtained from **8a** as a result of approximately 3:1 π-facial selectivity with OsO_4. Substrate **8b** provided **9b** and **9d** with analogous diastereoselectivity. From **9a–9d**, the four partially deuterated glycerols **10a–10d** were synthesized after detachment of the diacetone-D-glucofuranose unit via a 5-step reaction sequence.

The Kakinuma group explored the stereoselective dianion [2,3]-Wittig rearrangement of alkylated allyloxyacetic acids incorporated into C-3 of diacetone-D-glucose **1** for the synthesis of (2*R*,3*S*)-3-alkylmalic acids in conjunction with biochemical studies on the thermostable isopropylmalate dehydrogenase [37] (*Scheme 3*). The substrates for the [2,3]-Wittig rearrangement, compounds **11**, possessed a mono-substituted alkene moiety (R = *i*Pr, Me and others) and a (carboxy)methoxy group at C-3 of diacetone-D-glucose **1**. These substrates were synthesized from the corresponding 3-ulose by (1) addition of the respective alkyne (R = *i*Pr, Me, and others), (2) *trans*-selective triple-bond reduction with $LiAlH_4$, and (3) alkylation of the liberated 3-OH with bromoacetic acid. Treatment of **11** with LDA, followed by esterification of the resulting [2,3]-Wittig rearrangement product with CH_2N_2, provided **12** with a useful level of diastereoselectivity. Diastereomeric excess (de) of the *threo*-product (major) to the *erythro*-product (minor) varied in the range of 74 to 90%. Despite the size of the R group, each rearrangement proceeded smoothly to provide **12** efficiently. It should be emphasized that the alkyl group (= R) and the formed hydroxyl group in **12** are in an antiperiplanar relationship (*threo*-configuration) in every case. The desired stereochemically defined (2*R*,3*S*)-3-alkylmalic acids **13** (R = *i*Pr, Me, and others) were prepared from **12** by cleav-

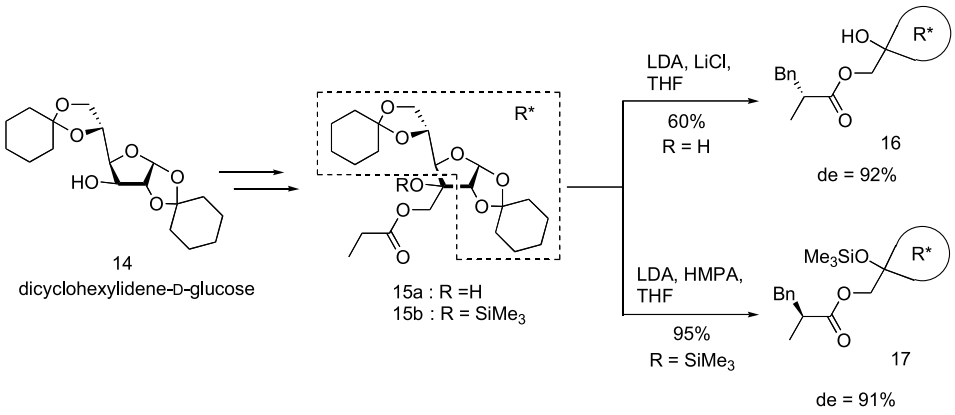


Scheme 3

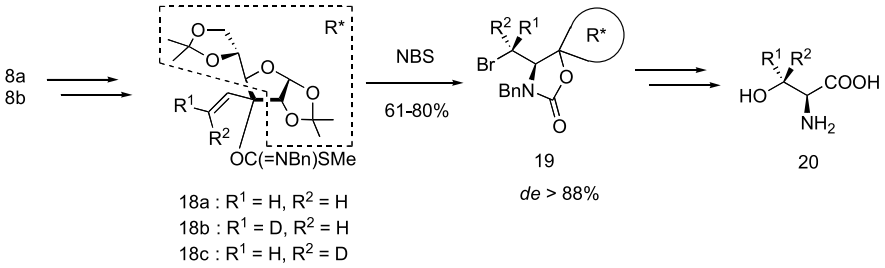
age of the carbon–carbon double bond after temporary protection of the hydroxyl group by ozonolysis, followed by further oxidative work-up with hydrogen peroxide and final deprotection.

The Kakinuma group explored the stereoselective α-alkylation of propionyl ester incorporated at C-3 of dicyclohexylidene-D-glucose **14** for investigating a remote asymmetric induction realized in the sugar template [38] (*Scheme 4*). The synthesis of propionate **15** for α-alkylation started from dicyclohexylidene-D-glucose **14** via the following reaction sequence: (1) preparation of 3-ulose by oxidation, (2) Wittig olefination with $Ph_3P{=}CH_2$, (3) OsO_4-NMO oxidation, and (4) selective acylation of the resulting diol. The substrate **15a** (R = H) was deprotonated using LDA as a base (with or without LiCl or HMPA). The resulting enolate was trapped with benzyl bromide to produce the α-benzylated propionate **16**. In the presence of LiCl, the benzylated product was obtained in 60% yield with a notably high diastereomeric excess of the (*R*)-isomer (de = 92%). Interestingly, the analogous benzylation using another substrate **15b** (R = TMS) produced α-benzylated propionate **17**, which possesses a newly introduced chiral carbon with an (*S*)-configuration in a high yield of 95%. The diastereomeric excess in this case was 91%. The occurrence of chelate formation in the transition-state governed the stereochemical course of the benzylation. Thus, the Z-enolate derived from **15a** (R = H) constituted a chelate form between the Li metal-cation and enolate-oxygen in the presence of LiCl. The formation of chelate was not possible in the case of **15b** (R = TMS). This difference in the chelate formation reflected the opposite chiralities of the newly created chiral carbon in **16** and **17**.

The potency of the diacetone-D-glucose framework as a chiral auxiliary was further demonstrated by the Kakinuma group through the stereoselective synthesis of β-deuterated L-amino acids such as (3*R*)- and (3*S*)-[3-$^2H_1$(D)]-L-serine [39] (*Scheme 5*). 3-*C*-[2-H or 2-$^2H_1$]-ethenyl-3-*O*-(*N*-benzyl)methylthioformimidoyl-1,2:5,6-di-*O*-isopropylidene-α-D-allofuranoses **18a–18c** were synthesized from the aforementioned 3-vinyl or β-deuterated vinyl derivatives of diacetone-D-glucose (**8a** or **8b** in *Scheme 2*), by treatment of benzyl isothiocyanate and iodomethane in the presence of NaH. The halonium ion-mediated intramolecular cyclization of **18a–18c** with NBS resulted in an oxazolinone (five-membered cyclic carbamates) formation highly stereoselectively to produce **19** as a major cyclization product. Furthermore, a deuterium was placed in product **19** stereoselectively (*R* or *S*). The nitrogen in the imine structure attacked exclusively the thus formed bromonium intermediate in a *five-exo-trig* cyclization mode to produce **19**. The bromomethylene group was a synthetically versatile



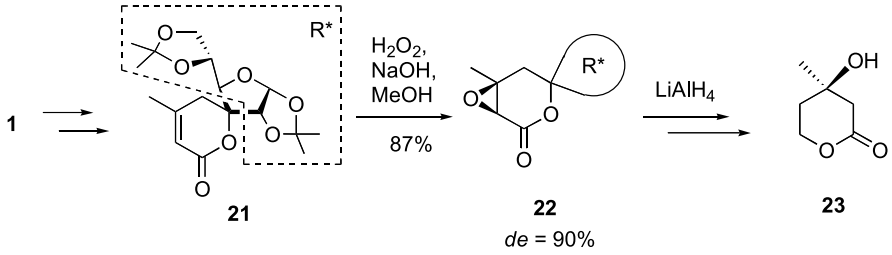
Scheme 4



Scheme 5

functionality. One example was the replacement of the bromo group by an acetoxy group. From **19**, stereochemically defined β-monodeuterated β-hydroxy-L-amino acids **20** were synthesized from the thus obtained acetate by oxazoline cleavage by alkaline hydrolysis followed by detachment of the sugar template.

Mevalonate is ubiquitous as the first biosynthetic precursor to various vertebrate and invertebrate hormones. To supply isotope-labeled mevalonate as a promising tool for biosynthetic studies, the Kakinuma group explored the highly enantioselective synthesis of (*R*)-mevalonolactone and partially deuterated (*R*)-mevalonolactone by means of the chirality-transcription strategy achieved using diacetone-D-glucose as a chiral template [40] (*Scheme 6*). The starting material of the mevalonate synthesis was compound **21**, which was in turn synthesized highly stereoselectively by the addition of the enolate derived from methyl 3-methylcrotonate (methyl senecioate) or its deuterated form to the 3-ulose of diacetone-D-glucose **1** in the presence of LDA, albeit in a modest yield of ~30%. The thus obtained unsaturated δ-lactone **21** was subjected to alkaline epoxidation with hydrogen peroxide to produce **22** as a result of the stereoselective attack (de = 90%) of the peroxide to the *Si*-face of the double bond. The epoxy-ring opening of the spirocyclic δ-lactone **22** with $LiAlH_4$ proceeded regioselectively. Acid hydrolysis of the acetals in the sugar template in the resulting 1,3-diol, followed by



Scheme 6

$NaIO_4$ oxidation eventually provided mevalonolactone **23** or deuterated mevelonolactone (not shown) efficiently.

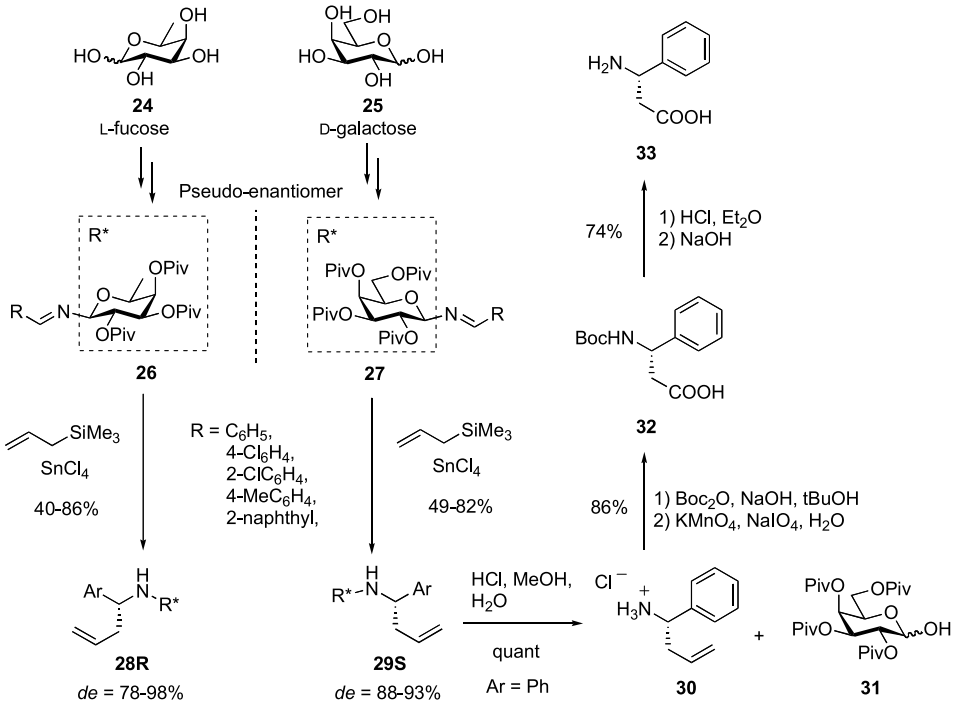
In another paper from the Kakinuma group, further result on the chirality's transfer realized using the diacetone-D-glucose template was described [41].

2.2 Studies by the Kunz Group

Since the mid-1980s, Kunz and co-workers started an exploration of asymmetric synthesis using sugar template-based chiral auxiliaries. Their early success on this topic is particularly shown by their demonstration of the utility of six-membered pento- and hexopyranose derivatives such as per-*O*-pivaloylated β-D-galactosylamine [42,43,44,45,46,47,48,49,50]. The collective studies of the Kunz group have been thoroughly reviewed [51,52].

The Kunz group has extensively explored the functionalized pyranose-based asymmetric carbon–carbon bond-forming reactions up to date. In the early 1990s, the following important results were disclosed from the Kunz group.

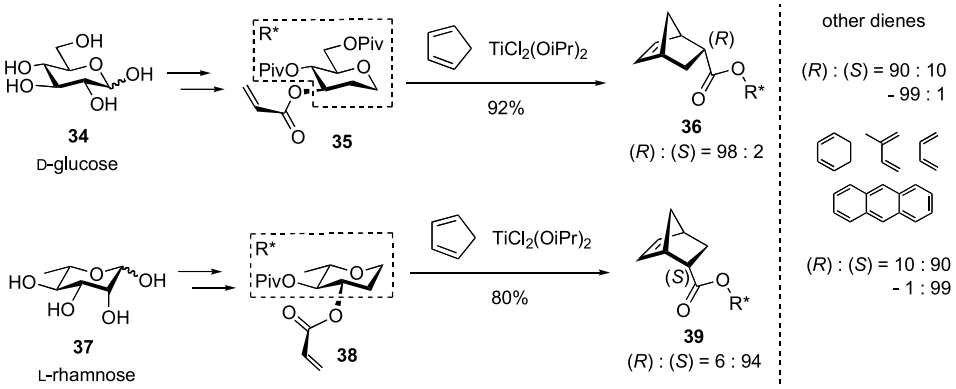
2,3,4-Tri-*O*-pivaloyl-α-L-fucopyranosylamine was prepared from L-fucose **24** and 2,3,4,6-tetra-*O*-pivaloyl-β-D-galactopyranosylamine was from D-galactose **25**. These glycosylamines were condensed with a variety of aromatic aldehydes to produce the corresponding glycosylimines **26** and **27**, respectively (*Scheme 7*). These glycosylimines were used for the diastereoselective synthesis of *N*-glycosyl-*N*-homoallylamines [53,54,55]. Therefore, per-*O*-acylated L-fucosylated arylimines **26** and D-galactosylated arylimines **27** were subjected to $SnCl_4$-mediated allylation using allyltrimethylsilane. These Lewis acid-promoted carbon–carbon bond formations proceeded with remarkable diastereoselectivity to provide the *N*-glycosyl-*N*-homoallylamines **28R** from **26** and **29S** from **27** in moderate-to-good chemical yields. Both the C-2 pivaloyloxy groups in **26** and **27** served as an effective shielding element in the course of the allylsilane approach toward imine carbon. Interestingly, the glycosylimines **26** and **27** are in a pseudo-enantiomeric relationship. Acid hydrolysis of the homoallylamine **29S** (Ar = Ph) detached the sugar template to provide α-allylated 3-butenylamine **30** (as its HCl salt), and the sugar template 2,3,4,6-tetra-*O*-pivaloyl-D-galactopyranose **31** was obtained. Protection of the amine functionality in **30**, followed by oxidative cleavage of the double bond using a combination of $KMnO_4$ and $NaIO_4$, produced (*S*)-*N*-Boc-β-phenyl-β-alanine **32**, from which (*S*)-(–)-β-phenyl-β-alanine **33** was obtained.



Scheme 7

The Kunz group explored the use of a sugar-based chiral auxiliary for stereoselective Diels–Alder reaction. For example, 3-*O*-acryloyl-1,5-anhydro-2-deoxy-4,6-di-*O*-pivaloyl-D-*arabino*-hexitol **35**, prepared from D-glucose **34**, was treated with cyclopentadiene in the presence of $TiCl_2(OiPr)_2$(3 mol equivalents to **35**) in CH_2Cl_2 at −30 °C [56] (*Scheme 8*). This Lewis acid-promoted [4+2] cycloaddition reaction proceeded with high diastereoselectivity to provide **36** possessing an (*R*)-configuration for the newly created stereogenic center. This Diels–Alder reaction using **35** as a dienophile was further explored with other dienes, such as cyclohexadiene, 2-methyl-1-3-butadiene, 1,3-butadiene and so on. In most cases, the cycloadditions took place smoothly in the presence of Lewis acid [$TiCl_4$, $TiCl_3(OPr)$], providing the adduct with remarkable diastereoselectivity. A similar Diels–Alder reaction was exemplified using 3-*O*-acryloyl-1,5-anhydro-2,6-deoxy-4-*O*-pivaloyl-L-*rhamno*-hexitol **38** as a dienophile and cyclopentadiene as a diene. The dienophile **38** was prepared from L-rhamnose **37**. The Lewis acid $TiCl_2(OiPr)_2$ was also effective in this case. More importantly, the adduct **39** possessed an (*S*)-configuration for the newly created stereogenic center. Therefore, the Diels–Alder adduct, i. e., 2-(carboxyl)bicyclo[2.2.1]hep-5-ene possessing either (*R*)- or (*S*)-chirality, can be obtained.

The aforementioned D-galactose-derived chiral imines **27** were used for the Strecker synthesis of D-α-amino nitriles and D-α-amino acids by the Kunz group. Various types of glycosylimines **27**, prepared from 2,3,4,6-tetra-*O*-pivaloyl-β-D-galacosylamine with aliphatic or aromatic aldehydes, were treated with TMSCN in the presence of $SnCl_4$, providing α-amino nitriles

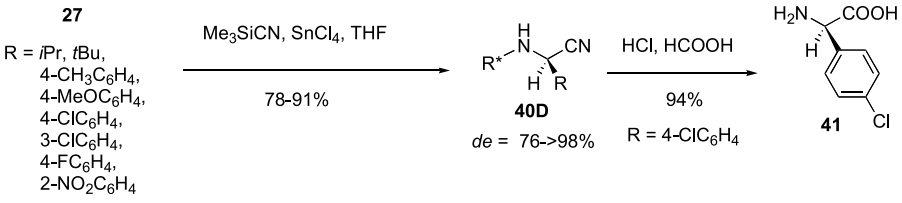


Scheme 8

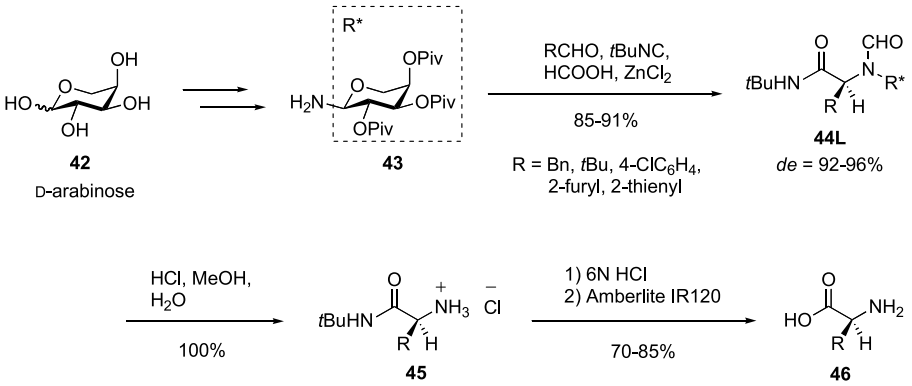
40D in good-to-excellent yields [57] (*Scheme 9*). Furthermore, these Strecker reactions provided α-amino nitriles **40D** with high diastereoselectivity in favor of the D-diastereomers. For a rational explanation of the observed high diastereoselectivity, a coordination of the imine nitrogen and the carbonyl oxygen of the 2-pivaloyl group by the metal of Lewis acid was proposed. This coordination fixed the conformation of the sugar template, to which the nitrile anion attacked from the rear side to avoid the steric hindrance caused by the bulky 2-pivaloyloxy group. Acid hydrolysis of the Strecker product, such as **40D** (R = 4-chlorophenyl), provided (*R*)-4-chloro-phenylglycine **41** in highly enantioenriched form.

The Kunz group also explored the synthesis of L-α-amino acids based on the sugar auxiliary protocol. For this purpose, the Ugi reaction was efficiently utilized [58]. As a sugar template, the Kunz group devised another glycosylamine, i. e., 2,3,4-tri-*O*-pivaloyl-α-D-arabinopyranosylamine **43**, a quasi-mirror image of the per-*O*-pivaloylated D-galactosylamine (*Scheme 10*). Glycosylamine **43**, prepared from D-arabinose **42**, was treated with a variety of aliphatic, aromatic, or heterocyclic aldehydes to produce the corresponding glycosylimines, which were further reacted with *tert*-butyl isocyanide and formic acid in the presence of $ZnCl_2$ at low temperatures. This four-component Ugi reaction provided *N*-formyl α-substituted α-L-amino acids as their amide forms **44L** in good-to-excellent yields.

The stereoselectivity of these four-component coupling reactions were remarkable in most cases. Similar Ugi reactions were studied using per-*O*-pivaloylated L-fucopyranosylamine as



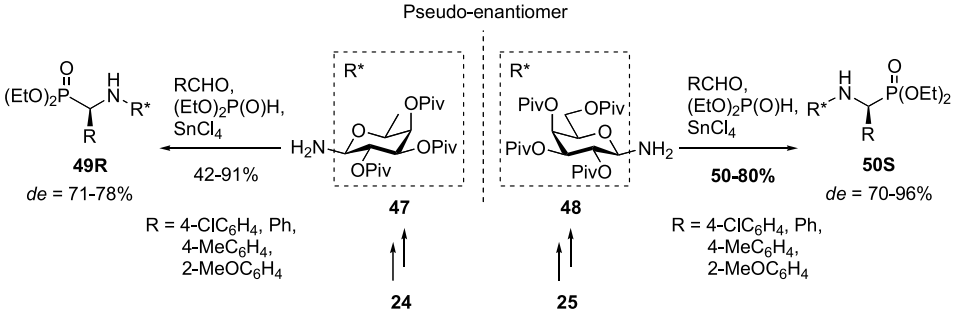
Scheme 9



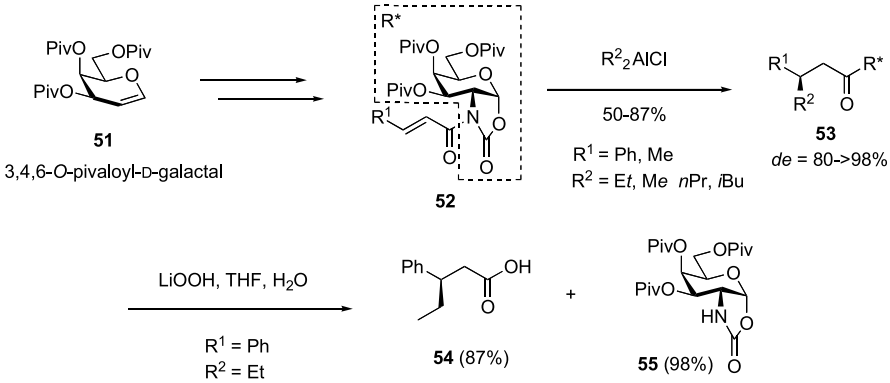
Scheme 10

a sugar template. In one case, *N*-fucosyl-(4-chlorophenyl)glycine amide was obtained in favor of the L-diastereomer with excellent stereoselectivity (not shown). The detachment of the sugar template from **44L** by acid hydrolysis provided L-α-amino acids as their *tert*-butylamides **45**, and further harsh hydrolysis of **45** provided L-α-amino acids **46**. Later, this sugar auxiliary-based Ugi-multicomponent synthesis was applied to a combinatorial approach realized on the solid phase by the Kunz group [59].

Using two glycosylamines **47** and **48**, prepared from L-fucose **24** and D-galactose **25**, respectively, the Kunz group synthesized α-aminophosphonic acid esters [60] (*Scheme 11*). The L-fucosylamine **47** reacted with a variety of aromatic aldehydes and $(EtO)_2P(O)H$ in the presence of $SnCl_4$ under one-pot reaction conditions. The *N*-glycosylated (*R*)-α-aminophosphonic acids **49R** incorporating the L-fucosyl residue were obtained as major products. The diastereoselectivity of diethylphosphite addition was in the range of 71–78%. In these cases, the Lewis acid served as an efficient stereocontrolling element through chelate formation between the imino nitrogen and the carbonyl group of 2-pivaloyl ester. On the other hand, the pseudo-enantiomeric D-galactosylamine **48** produced (*S*)-α-amino phosphonic acid derivatives **50S** preferentially under the same reaction conditions.



Scheme 11

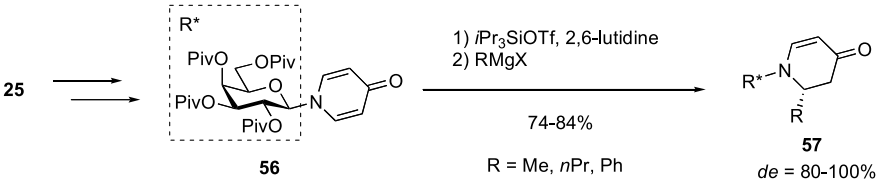


Scheme 12

The Kunz group has widely explored the utility of bicyclic oxazolidinones derived from 2-amino-2-deoxy sugars as chiral auxiliaries [61,62]. One example was a stereoselective radical 1,4-addition of organoaluminum chlorides to unsaturated carboxylic acid derivatives attached to the D-galactosamine-derived oxazolidinone (*Scheme 12*). 3,4,6-Tri-*O*-pivaloyl-1,2-dideoxy-α-D-galactopyraosido[1,2,5′,4′]oxazolidin-2′-one **55**, the sugar template for the 1,4-addition, was readily prepared from per-*O*-pivaloyled D-galactal **51** via azidonitration and hydrolysis, followed by treatment of the resulting 2-azido derivative with Ph_3P/CO_2. The resulting oxazolidinone derivative **55** was treated with MeMgBr for the deprotonation of NH of the oxazolidinone ring, followed by a reaction with an α,β-unsaturated acyl chloride such as cinnamoyl chloride to produce the *N*-functionalized oxazolidinone **52** (R^1 = Ph). Treatment of **52** with excess dialkylaluminum chloride (R^2 = Me, Et, Pr, *i*-Bu) in toluene or hexane at low temperature (−40 °C) produced the 1,4-adducts, i. e., β-alkylated carboxylic acid derivatives **53**, in good-to-excellent diastereoselectivities. In the case of Me_2AlCl, irradiation of the reaction mixture was required. In most cases, the diastereoselectivities in the 1,4-additions were remarkably high with a de of 80–98%. The removal of the sugar template from the adduct **53** (R^1 = Ph, R^2 = Et) was carried out by hydrolysis in the presence of LiOOH. As a result, (*R*)-3-phenylvaleric acid **54** was obtained in a highly enantioenriched form. In addition, the sugar template **55** was recovered almost quantitatively.

The Kunz group further demonstrated the asymmetric synthesis of β-alkylated α-halogenated (Cl or Br) carboxylic acid derivatives. For this purpose, the *N*-functionalized D-galactosamine-derived oxazolidinone **52** was treated with dialkylaluminum chloride, which was followed by addition of NXS (X = Cl, Br) to trap the intermediary enolate. Consequently, a variety of β-branched α-halo carboxylic acid derivatives were obtained with good diastereoselectivities [63].

Desymmetrization of prochiral compounds is one of the most convenient strategies for supplying enantioenriched chiral compounds. The Kunz group reported desymmetrization on the 4-pyridone tethering a sugar template at the nitrogen atom [64] (*Scheme 13*). The reaction of per-*O*-pivaloylated α-D-galactopyranosyl fluoride, prepared from D-galactose, with 4-trimethylsilyloxy-pyridine in the presence of Lewis acid ($TiCl_4$) provided *N*-(per-*O*-pivaloyl-

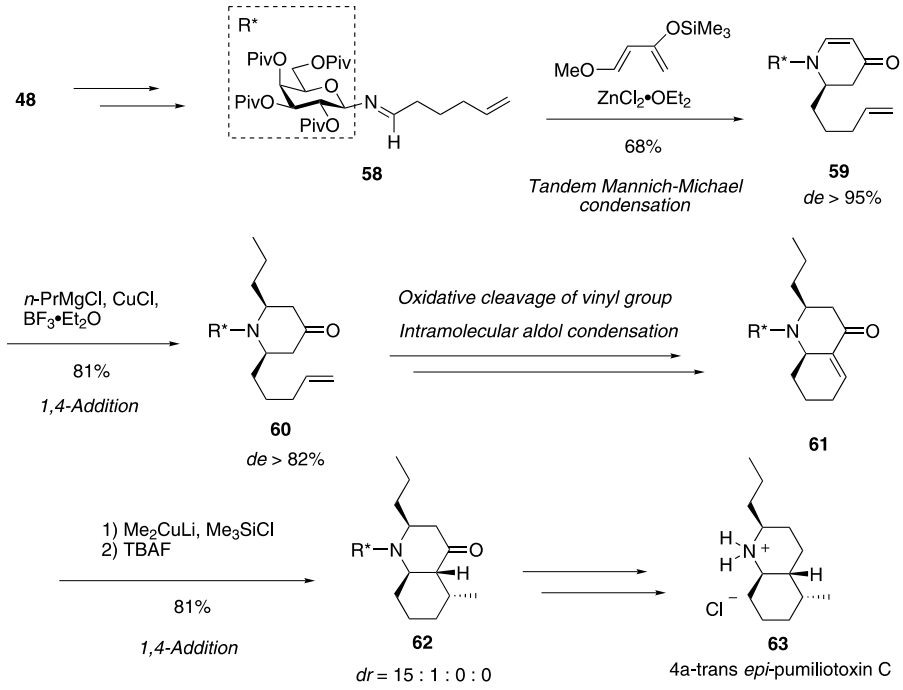


Scheme 13

β-D-galactopyranosyl)pyrid-4-one **56**. The thus obtained *N*-galactosyl pyridone **56** reacted with some Grignard reagents (R = Me, *n*-Pr, and Ph) in the presence of triisopropylsilyl trifluoromethanesulfonate and 2,6-lutidine for activation of the carbonyl. As a result, the 1,4-addition provided 2-alkyl-substituted dehydro-piperid-4-ones **57** in high yield and with good-to-excellent diastereoselectivities (de = > 80%). The sugar template directs the diastereoselectivity of the 1,4-additions at the planar 4-silyloxypyridinium intermediate; this constitutes a selective monofunctionalization and stereoselective attack at one of the two *Si*-faces of the pyridinium intermediate.

The utility of 2,3,4,6-tetra-*O*-pivaloyl-β-D-galactopyranosylamine **48** as a chiral auxiliary was further demonstrated by the Kunz group through a tandem Mannich–Michael reaction realized on the aldimine, such as **58** derived from **48** [65] (*Scheme 14*). The galactosylimine **58** was prepared by the reaction of the per-*O*-pivaloylated galactosylamine **48** and 5-hexenal. Treatment of **58** with (1-methoxy-3-trimethylsiloxy)butadiene (Danishefsky–Kitahara's diene) in the presence of $ZnCl_2OEt_2$ provided an *N*-galactosyl 2-alkylated 5,6-dehydro-piperidin-4-one **59** in excellent diastereoselectivity. This conversion of **58** to **59** included the initial Mannich reaction between the imine **58** and the siloxy diene, which was followed by the Michael addition of the intermediary amine to the resulting conjugate enone and subsequent removal of one mol of MeOH. For the construction of the 2,6-disubstituted piperid-4-one skeleton, the 1,4-addition to the enone moiety in **59** was explored. Neither Grignard reagents nor organocuprates reacted with **59**. However, a combination of the organocuprate, prepared from *n*-PrMgCl and CuCl, and boron trifluoride worked well for this purpose. Thus *cis*-2,6-dialkylated piperidinone **60**, which incorporates the sugar template at nitrogen, was obtained with excellent diastereoselectivity (> 10:1).

The Kunz group demonstrated the synthetic utility of this piperidinone **60** for the decahydroquinoline alkaloid synthesis. One example was the total synthesis of *trans*-annulated 4a-*epi*-pumiliotoxin C. Among more than 200 alkaloids isolated from a glandular secretion of a South American frog of the family *Dendrobates*, pumiliotoxin C is a representative of the biologically important *cis*-annulated decahydroquinolines. In an attempt to construct the pumiliotoxin skeleton, the piperidinone **60** was transformed as follows. Oxidative cleavage of the vinyl group in the C2-alkene moiety and base-mediated intramolecular aldol condensation of the resulting aldehyde efficiently produced an octahydroquinoline derivative **61**. The reaction of the bicyclic enone **61** with Me_2CuLi in the presence of chlorotrimethylsilane efficiently provided the 1,4-methyl adduct **62** after cleavage of the intermediary silyl enol ether. The ratio of the formed diastereo-mers was > 15:1 in favor of a *trans*-annulated decahydroquinoline **62**. This fact was in contrast to the previous result that the 1,4-addition of a cuprate to an

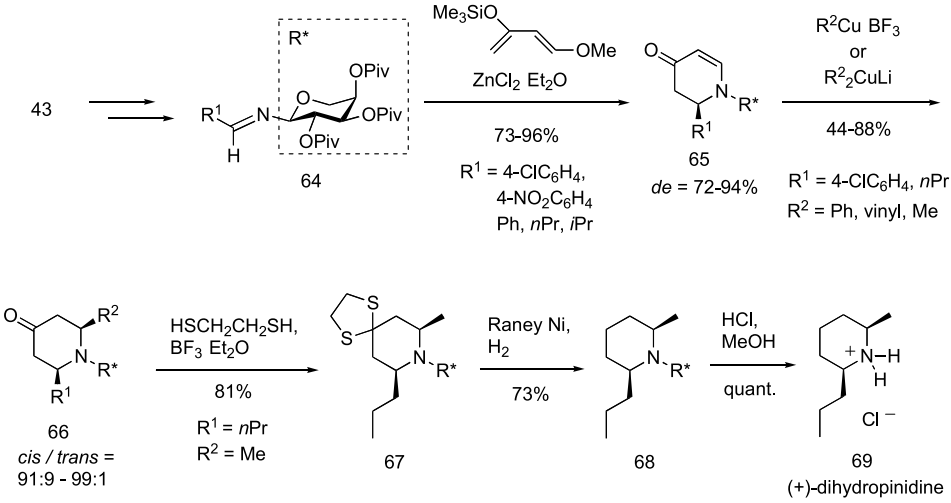


Scheme 14

N-phenoxycarbonyl bicyclic enone, which was structurally similar to **61**, exclusively provided a *cis*-annulated octahydroquinoline derivative. From the *trans*-annulated decahydroquinoline **62**, *trans*-4a-*epi*-pumiliotoxin C **63** was synthesized in a 5-step reaction sequence, including hydrogenolytic desulfurization of the dithiolane prepared by treatment of the carbonyl group in **62** with 1,2-ethanedithiol.

The Kunz group reported some glycosylamine-based chiral auxiliary approaches to the enantioselective synthesis of 2,6-*cis*-disubstituted piperidine alkaloids such as dihydropinidine [66], as well as that of the indolizidine alkaloids such as gephyrotoxin 167B [66] and *trans*-decahydroquinoline alkaloids such as *trans*-*epi*-pumiliotoxin C [67]. These approaches relied on the diastereofacial differentiation of the carbon nucleophile addition to the glycosyl aldimines through the steric, stereoelectronic, and complexing effects of sugars.

The per-*O*-pivaloylated D-arabinopyranosylamine **43** was used for asymmetric synthesis of the 2,6-*cis*-substituted piperidine alkaloid, (+)-dihydropinidine **69** by the Kunz group [68] (*Scheme 15*). The *N*-arabinosylimines **64**, prepared by the reaction of **43** with aromatic or aliphatic aldehydes, were reacted with Danishefsky–Kitahara's diene analogously to the reactions employed for **58**, as depicted in *Scheme 14*. Consequently, this zinc chloride-promoted tandem Mannich–Michael reaction provided a 5,6-didehydropiperidin-4-ones **65** in high yields and with high diastereoselectivity. It should be emphasized that *N*- arabinosyl dehydropiperidinones **65** possess the opposite configuration at C2 in the piperidinone ring, in contrast to the case of *N*-galactosyl dehydropiperidinones **59**. Furthermore, 1,4-addi-

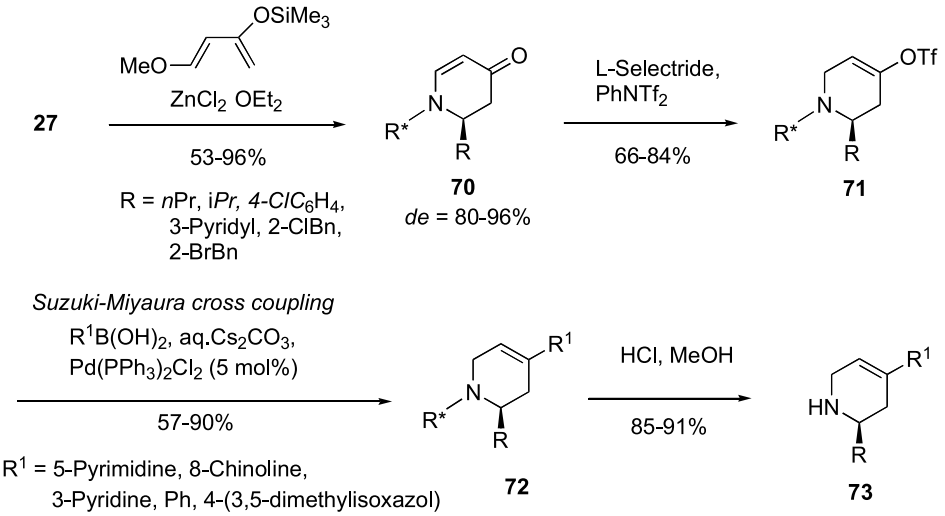


Scheme 15

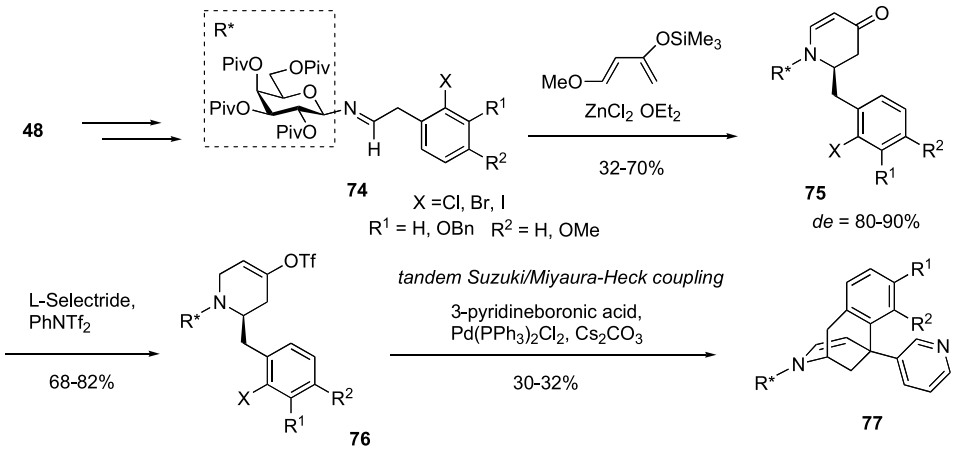
tions of organocopper/boron trifluoride complexes or organocuprates to **65** proceeded with excellent diastereoselectivity to produce 2,6-*cis*-dialkylated piperidin-4-ones **66** predominantly in moderate-to-good yields. A 4-piperidinone **66** ($R^1 = n$-Pr, R^2 = Me) was treated with ethanedithiol to provide dithiolane **67**. Hydrogenolytic desulfurization of the dithiolane in **67** with Raney nickel provided **68**. Removal of the D-arabinose moiety by acid hydrolysis eventually produced (+)-dihydropinidine **69** as its HCl salt in an enantiomerically pure form.

The palladium-catalyzed carbon–carbon coupling reaction (Suzuki–Miyaura coupling) was applied for the enantioselective synthesis of 2,4-disubstituted 4,5-dehydropiperidines by the Kunz group as an extension of the sugar-auxiliary-based dehydropiperidinone synthesis [69] (*Scheme 16*). As described in *Scheme 14*, the tandem Mannich–Michael reactions of the D-galactosylimines **27** (*Scheme 7*) with Danishfsky–Kitahara's diene provided 5,6-didehydropiperidin-4-ones **70** incorporating a variety of alkyl or aryl substituents at C-2. The thus formed piperidinones **70** were converted to the enol triflates **71** by conjugate hydride addition with L-Selectride, followed by trapping of the resulting enolate with $PhNTf_2$. The Suzuki–Miyaura coupling of the enol triflates **71** with a variety of alkyl or arylboronic acids in the presence of a catalytic amount of $Pd(PPh_3)_2Cl_2$ and Cs_2CO_3 produced the coupling products, i.e., 2-alkyl or -aryl-4-aryl (or hetero-aromatic) 4,5-dehydropiperidines **72**, in good yields. Acid hydrolysis of **72** to detach the D-galactopyranose moiety provided enantioenriched 2,4-disubstituted 4,5-dehydropiperidines **73**. The Kunz group applied the stereoselective piperidine synthesis to a combinatorial approach using a polymer-bound immobilized galactose auxiliary [70,71].

Using a similar synthetic concept as that described above (*Scheme 14* and *Scheme 16*), utility of the per-*O*-pivaloylated D-galactosylamine **48** as a chiral auxiliary was broadened to the benzomorphan synthesis by the Kunz group [72] (*Scheme 17* and *Scheme 18*). The

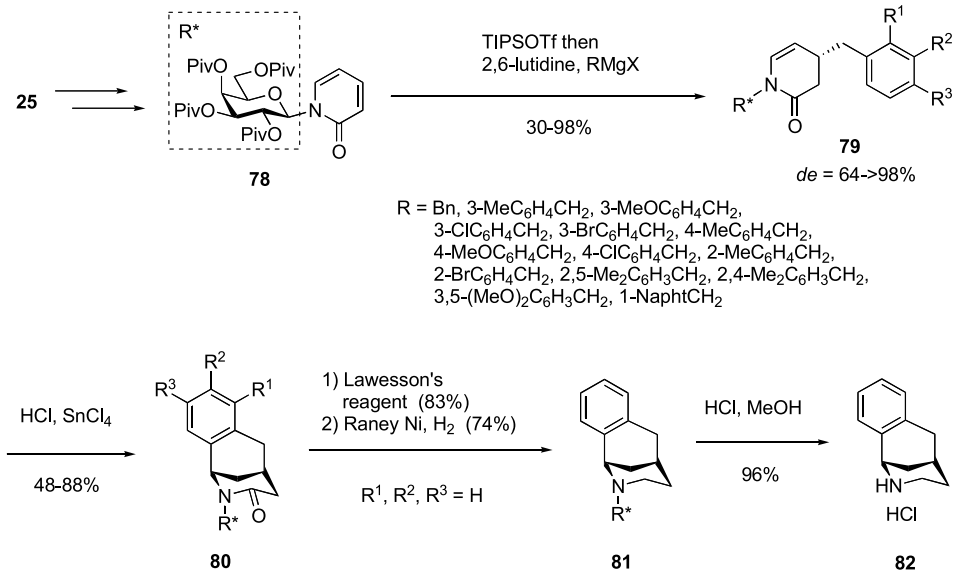


Scheme 16



Scheme 17

synthetic studies commenced with the imine **74**, prepared by the condensation of **48** with a variety of 2-halogenated phenylacetic aldehydes. The aromatic nuclei of some imines **74** possessed oxygen functionalities. The $ZnCl_2$-mediated tandem Mannich–Michael reactions of the aldimines **74** and Danishefsky–Kitahara's diene provided *N*-galactosyl 4-piperidinone derivatives **75** with useful levels of diastereoselectivity in all cases, although the yields were not necessarily high. The efficient shielding of the *Re*-face of the imine double bond by the bulky 2-pivaloyloxy group in the sugar template led to a highly stereoselective induction of the stereogenic center, i. e., an (*R*)-configuration, in the piperidinone ring. The 1,4-conjugate



Scheme 18

hydride addition to **75**, followed by trapping the resulting enolate with $PhNTf_2$, provided enol triflates **76**. The benzomorphan skeleton was constructed efficiently by a series of Suzuki–Miyaura coupling followed by Heck coupling reactions. Thus, the enol triflates **76** were subjected to palladium-catalyzed Suzuki–Miyaura coupling with 3-pyridineboronic acid, providing 2,4-disubstituted 4,5-didehydropiperidines (not shown). In the case of *o*-bromobenzyl derivatives **76** (X = Br), the reaction proceeded further after the initial Suzuki–Miyaura coupling with 3-pyridineboronic acid to construct a tricyclic benzomorphan framework. The products of these tandem reactions were **77**. The second intramolecular Heck reaction stereoselectively formed a new quaternary carbon center in the 4-position of the piperidine ring. The attack was controlled by the configuration of the benzyl substituent and thus occurred exclusively from the top face of the *endo* double bond.

An alternative strategy for the 7,8-benzomorphan synthesis was studied by the Kunz group starting from *N*-galactosyl 2-pyridone **78**, which in turn was prepared by *β*-selective *N*-glycosylation of 2-(trimethylsiloxy)pyridine with per-*O*-pivaloylated D-galactosyl fluoride in the presence of $TiCl_4$ at 70 °C (*Scheme 18*). The pyridone **78** was transformed into **79** by the regio- and stereoselective Grignard reaction with substituted benzylmagnesium halides. After activating with TIPSOTf, the Grignard addition provided **79** in moderate-to-good yields. Next the intramolecular aminoalkylation of **79** took place in the presence of a mixture of HCl and $SnCl_4$ to produce tricyclic benzazocinones **80** via the intermediary *N*-acyl iminium ion. The electrophilic attack at the phenyl ring occurred exclusively from the *cis* side. Further elaboration of the functional groups in **80** eventually provided **81**. This transformation included the reductive removal of the *δ*-lactam carbonyl via the thioamide, prepared by treatment of **80** with Lawesson's reagent followed by hydrogenolysis in the presence of Raney nickel and acid hydrolytic removal of the sugar template in the resulting **81** producing **82**.

The Kunz group used a variety of sugar derivatives as chiral auxiliaries for the [2+2] cycloadditions of ketenes and enol ethers, which produced eventually highly substituted cyclobutanol derivatives [73].

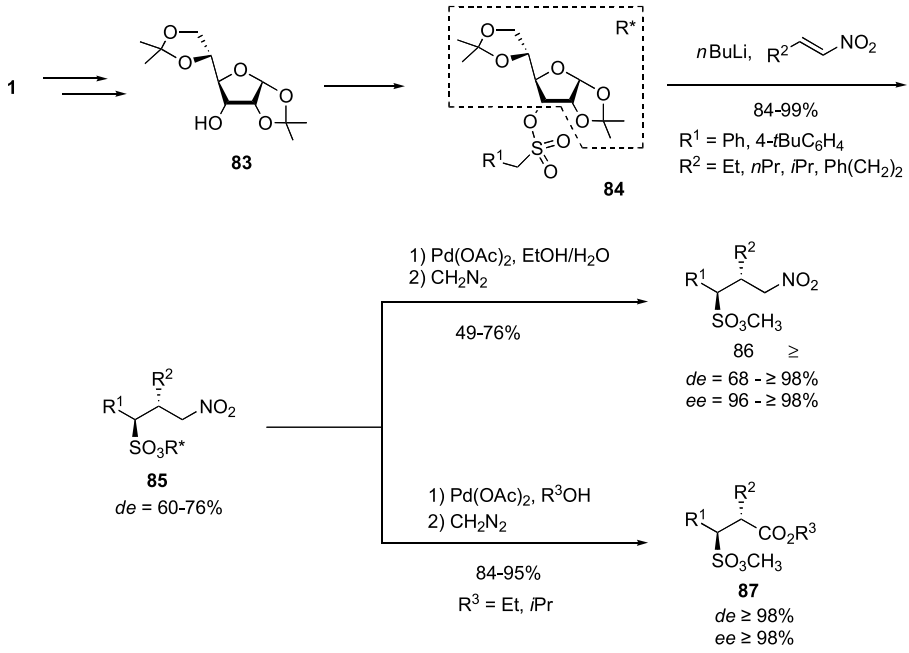
2.3 Studies by the Enders Group

Since the mid-1970s, Enders and co-workers have demonstrated the powerful abilities of (*S*)-2-methoxymethylpyrrolidine (SMP) and its (*R*)-enantiomer (RMP) as notable chiral auxiliaries. Numerous applications to symmetric synthesis using commercially available SMP and RMP have been appeared in the literatures [74,75,76]. In addition to the brilliant success of the worldwide-recognized chiral auxiliaries, the Enders group has also explored the utilization of sugar-based chiral auxiliaries since the start of the twenty-first century. The Enders group is predominantly concerned with diacetone-D-hexofuranoses as chiral templates.

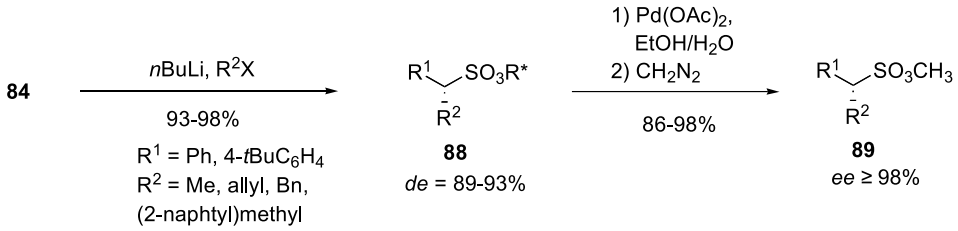
As earlier investigations on the sugar-based asymmetric synthesis, the Enders group utilized diacetone-D-allose, namely, 1,2:5,6-di-*O*-isopropylidene-α-D-allohexofuranose **83**, as a chiral auxiliary (*Scheme 19*). The Enders group first explored the diastereoselective and enantioselective synthesis of α,β-disubstituted γ-nitro methyl sulfonates using **83** to try to succeed in the asymmetric synthesis of analogues of homotaurine (3-aminopropanesulfonic acid) [77,78]. This asymmetric approach was performed starting from 3-*O*-(phenylmethyl)sulfonyl- (or -4-substituted phenylmethyl)sulfonyl-1,2:5,6-di-*O*-isopropylidene-α-D-allofuranoses **84**, which are readily available from **83**. Deprotonation of the phenylmethyl group with BuLi produced the lithiated sulfonate, which reacted with β-alkylated nitroalkenes. The Micheal reaction proceeded smoothly to produce the 1,4-adducts **85** in good-to-excellent yields. Various nitroalkenes were examined for the Michael reaction. In most cases, the results were satisfactory enough, and chiral 2-nitromethyl-1-phenyl-2-alkyl-1-sulfonic acid esters **85** were produced in moderate diastereoselectivities. Removal of the sugar auxiliary from the Michael adducts **85** were explored in detail by the Enders group. As a result, the corresponding methyl sulfonic esters **86**, in which the nitromethyl portion remained intact, were obtained in a virtually enantiopure form by treatment of **85** with catalytic $Pd(OAc)_2$ in aqueous EtOH, followed by methyl esterification of the resulting sulfonic acid with diazomethane. On the other hand, through treatment of **85** with a catalytic amount of $Pd(OAc)_2$ in net R^3OH (without water) under reflux, interestingly the β-alkoxycarbonyl methyl sulfonates **87** were obtained in excellent yields.

Using the 3-*O*-(phenylmethyl)sulfonyl derivatives **84**, the Enders group thoroughly explored stereoselective alkylation at the α-carbon of the (phenylmethyl) sulfonyl moiety. α-Alkylation with a variety of carbon electrophiles was examined using the lithiated phenylmethyl sulfonic acid ester **84** [79,80] (*Scheme 20*). The lithiated sulfonic acid esters derived from **84** were trapped with some alkyl halides, producing the α-alkylated products **88** in excellent yields with remarkably high diastereoselectivity after recrystallization. The sugar auxiliary was removed from **88** by treatment with $Pd(OAc)_2$ (catalytic) in aqueous EtOH followed by esterification of the resulting sulfonic acids, which provided α-alkylated phenylmethyl sulfonic acid methyl esters **89** in virtually enantiopure forms.

The Enders group examined the allylation of the α-lithiated phenylmethylsulfonyl D-allofuranose **84** [81,82] (*Scheme 21*). Thus, treatment of the lithiated **84** with allyl, crotyl, or



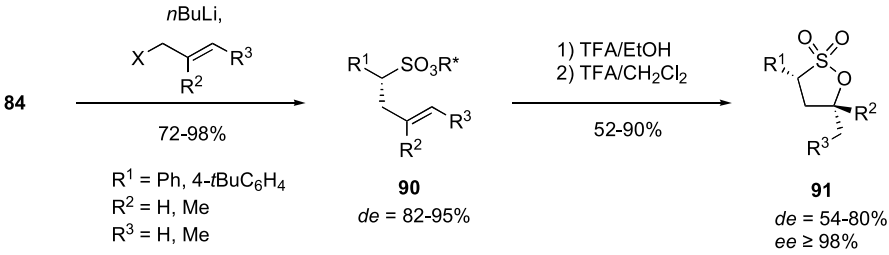
Scheme 19



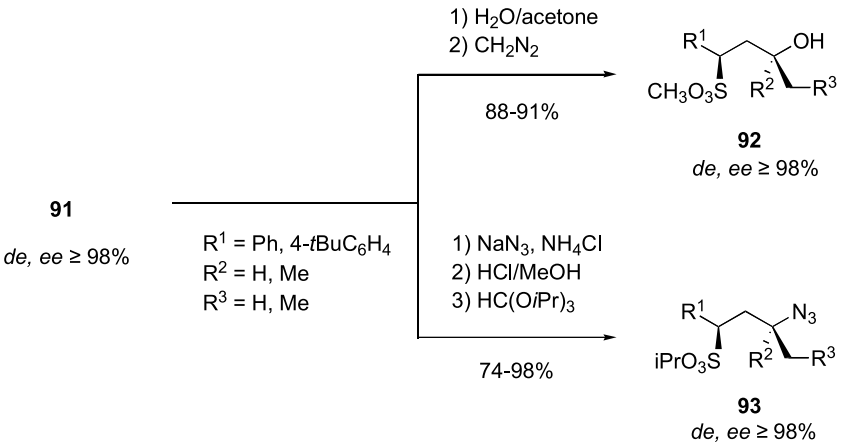
Scheme 20

methallyl chloride provided the allylated products **90** in good yields with high diastereoselectivity. Treatment of **90** with a catalytic amount of TFA in EtOH removed the sugar template to produce a sulfonic acid intermediates, which were treated with TFA in CH_2Cl_2 to produce α,γ-substituted γ-sulfones **91** (five-membered cyclic sulfonates) stereoselectively in an almost enantiopure form. To explain the stereochemical outcome of the sultone formation (the final cyclization step), it was most likely that a Markovnikov-type protonation of the olefin occurred under the acidic conditions and then a *five-exo*-cyclization proceeded through a more stable 1,3-*cis* relationship of the two bulkier substituents.

Ring opening of sultones **91** by hydrolysis in aqueous acetone produced acyclic γ-hydroxylated sulfonates **92** [83,84] (➲ *Scheme 22*). Furthermore, the S_N2 reaction of **91** with NaN_3



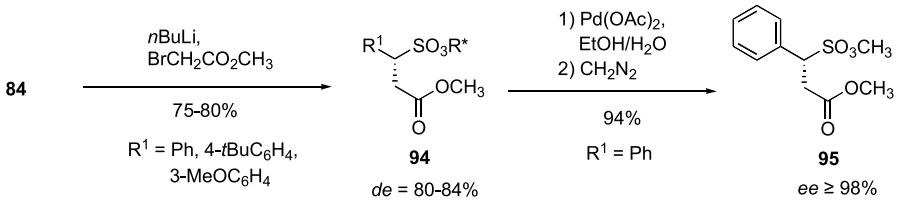
Scheme 21



Scheme 22

produced an azide-substituted product as a result of the ring opening. The sulfonic acid was then esterified to provide γ-azido sulfonic acid ester **93**.

As another example, the Enders group explored the introduction of an (methoxycarbonyl)methyl group at the α-carbon of the phenylmethyl group in **84** [85] (*Scheme 23*). Thus, the lithiated **84** reacted with methyl bromoacetate. As a result, phenylmethyl sulfonic acid esters **94** bearing a (methoxycarbonyl)methyl group at the α-position were produced in high yields with high diastereo-selectivities. A $Pd(OAc)_2$-catalyzed detachment of the sugar moiety provided **95**.



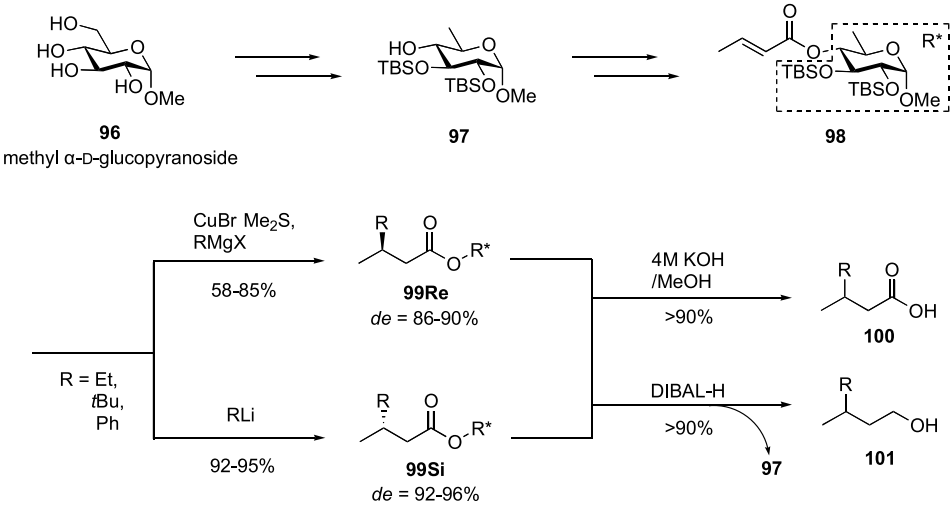
Scheme 23

2.4 Studies by the Tadano Group

Since the later half of 1990s, the Tadano and co-workers have designed and synthesized a variety of hexopyranoside-derived templates to verify their potential as efficient chiral auxiliaries for a variety of carbon–carbon bond-forming reactions. The Tadano group thoroughly explored the difference in the spatial environment constituted by each hexopyranoside configuration, which significantly affected the diastereoselectivity of the attempted organic reactions. It is generally recognized that, owing to the occurrence of anomerization, it is not easy to functionalize an anomeric position while strictly controlling the α- or β-configuration. The Tadano group synthesized a variety of chiral templates from D-glucose, D-mannose, or D-galactose, in which an unprotected hydroxyl group at C-2, -3, -4, or -6 was acylated to introduce the reaction site. These substrates were subjected to 1,4-addition reactions, alkylations, cycloadditions, and others. In each reaction, the Tadano group expected that the asymmetric induction would occur principally by the shielding effect of the bulkiness of the group neighboring the reaction site. Then, the sugar template would be removed by a basic, acidic, or reductive treatment. Consequently, a variety of α- and/or β-chiral carboxylic acids or alcohols could be obtained in the enantiomerically pure or enriched form. In addition, the used sugar templates would be recovered for recycling [86].

As the initial attempt for the asymmetric carbon–carbon bond-forming reactions using sugar templates, the Tadano group explored extensively the 1,4-conjugate additions to α,β-unsaturated esters, such as crotonyl or cinnamoyl esters incorporating at the C4 or C6 of methyl 6-iodo-, 6-deoxy-2,3-di-*O*-protected, 2,3,6-tri-*O*-protected or 2,3,4-tri-*O*-protected α-D-glucopyranosides. The 1,4-additions of the corresponding 2,3-*O*-benzyl and 2,3-*O*-acyl derivatives with a variety of organocopper reagents (RCu, R_2CuMgX) provided 1,4-adducts with moderate-to-high levels of diastereoselectivity. From these diastereomerically enriched 1,4-adducts obtained using the 4-*O*-crotonyl ester, enantioenriched β-alkylated butanoic acids were obtained by basic hydrolysis. [87]. The glycopyranoside-based diastereoselective 1,4-addition approach was further developed by the Tadano group using other hexopyranosides, such as α-D-*manno*- and α-D-galactopyranosides [88]. Some methyl 2,4,6-tri-*O*-protected 3-*O*-crotonyl-α-D-mannopyranosides provided 1,4-adducts with useful levels of diastereoselectivity in the reactions with ethyl and vinyl Grignard reagent-derived cuprates. Moreover, some methyl 2,4,6-tri-*O*-protected 3-*O*-crotonyl-α-D-galactopyranosides provided 1,4-adducts with high levels of diastereoselectivity in the reactions with ethyl, isopropyl or vinyl cuprates. A mechanistic explanation for the observed high diastereoselectivities achieved using these hexopyranoside-derived templates was proposed by the Tadano group using some transition-state models.

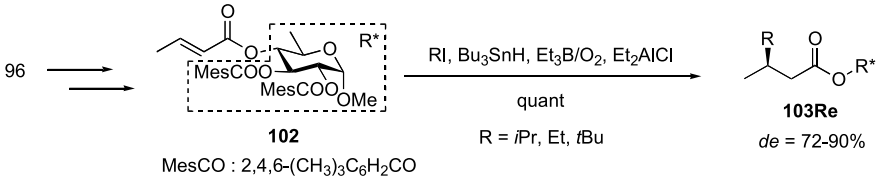
As the most effective chiral auxiliary for the 1,4-conjugate addition of organocuprates to an unsaturated ester incorporating into a hexopyranoside derivative, the Tadano group found methyl 6-deoxy-2,3-di-*O*-(*t*-butyldimethylsilyl)-α-D-glucopyranoside **97**, which was prepared from methyl α-D-glucopyranoside **96** [89] (*Scheme 24*). As a result, the 4-*O*-crotonyl derivative **98** of this sugar template **97** provided the 1,4-adduct in moderate-to-high yields with useful levels of diastereoselectivity in the reaction with alkyl (or phenyl) Grignard reagent-derived cuprates. The adducts **99Re** with an ethyl, *t*-butyl, or phenyl group at the β-carbon were preferentially obtained. The neighboring bulky 3-*O*-(*t*-butyldimethylsilyl) group effectively shielded the front side of the 4-*O*-crotonyl ester moiety. Thus, the attack of the organocopper



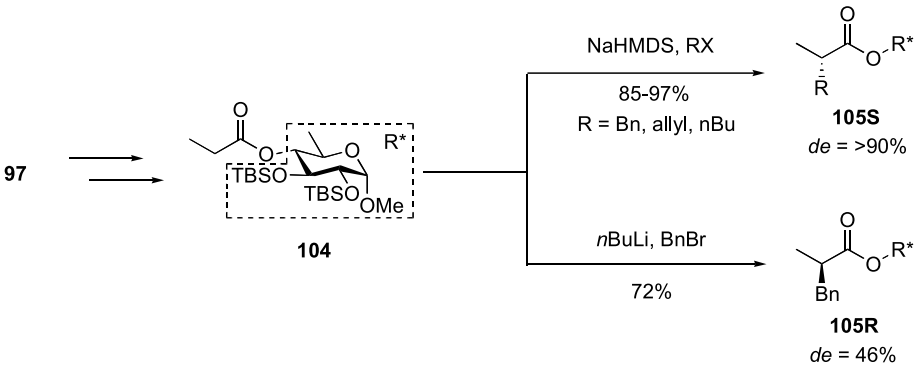
Scheme 24

reagent occurred predominantly from the less congested rear side. Furthermore, the conformation of the unsaturated ester moiety, i. e., *s-cis-syn-*, or *s-trans-syn*-conformations, governed the stereochemical outcome. Interestingly, the 1,4-additions of alkyllithiums(two alkyllithiums and phenyllithium) to **98** resulted in the formation of 1,4-adducts **99Si** (not 1,2-adducts) with the opposite configurations at the β-carbon. From 1,4-adducts **99Re** or **99Si**, β-alkyl (or phenyl) butanoic acids **100** or butanols **101** were synthesized in highly enantioenriched forms by saponification or hydride reduction, respectively.

The 1,4-additions of radical species to the 4-*O*-crotonyl derivative of methyl 6-deoxy-2,3-*O*-protected α-D-glucopyranosides were explored by the Tadano group [90] (*Scheme 25*). The radical reactions proceeded highly diastereoselectively using a traditional system of alkyl iodide, *n*-butyltin hydride and triethylborane/ oxygen in the presence of Lewis acid, such as $AlCl_3$, $BF_3{\cdot}Et_2O$, or Et_2AlCl, to produce a variety of β-alkylated butanoic acid esters tethering the sugar template. In particular, the 2,3-di-*O*-mesitoyl-type template (mesitoyl = 2,4,6-trimethylbenzoyl) **102** provided 1,4-adducts **103Re** (R = isopropyl, ethyl, *t*-butyl) most effectively.



Scheme 25

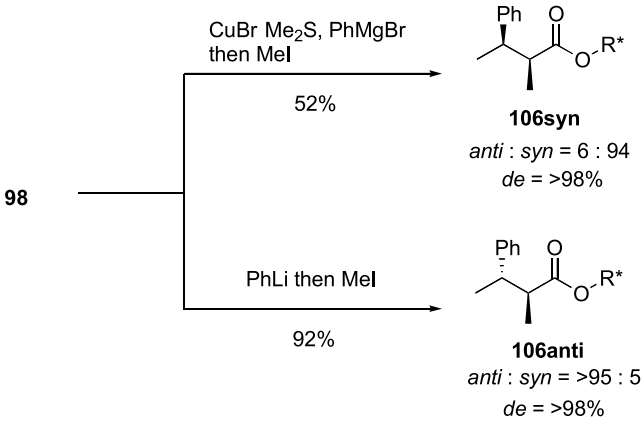


Scheme 26

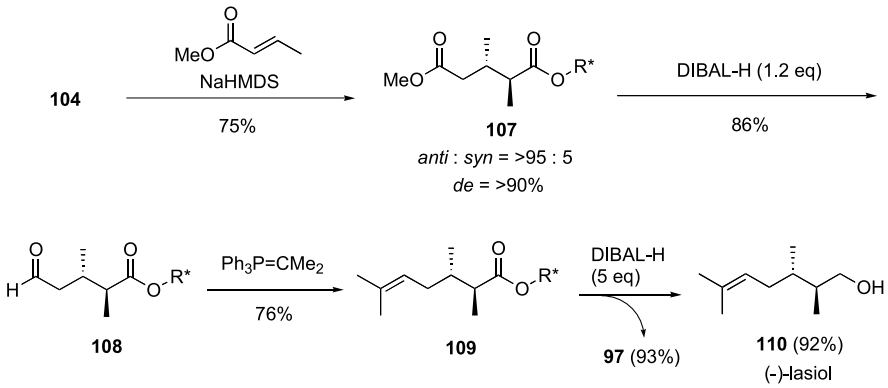
The utility of the 2,3-di-*O*-(*t*-butyldimethylsilyl)-type template was also evidenced through the α-alkylation of methyl 6-deoxy-2,3-di-*O*-(*t*-butyldimethylsilyl)-4-propionyl-α-D-glucopyranoside **104** [86] (*Scheme 26*). The propionyl ester **104** was deprotonated with a base, and the resulting enolate was trapped with benzyl, alkyl, or allyl halide to provide the α-benzylated, α-alkylated, or α-allylated products with excellent diastereoselectivity in all cases. Sodium hexamethyldisilazide (NaHMDS) was most effective base, which preferentially provided (*S*)-substituted products **105S**. The Tadano group speculated that the (*Z*)-enolate was formed preferentially from **104**, which was trapped by an electrophile, leading to the formation of (*S*)-isomers. The Tadano group also found that the deprotonation of **104** with *n*-BuLi (or LDA), followed by benzylation of the resulting enolate provided another α-alkylated product **105R**, albeit in a lower diastereoselectivity. Interestingly, the benzylated product **105R** possessed the (*R*)-configuration at the α-carbon opposite to **105S**.

The 1,4-conjugate addition of carbon nucleophiles to the 4-*O*-crotonyl ester **98** followed by α-alkylation in a one-pot process was realized with high *anti/syn*- and diastereoselectivities [89] (*Scheme 27*). Namely, the 1,4-addition of a phenyl-copper reagent, prepared from phenylmagnesium bromide and copper(I) bromide, to **98** followed by trapping of the intermediary enolate with iodomethane, provided α,β-dialkylated butanoic acid derivative **106syn** with high *anti/syn*-selectivity in favor of the *syn*-isomer. It should be emphasized that the diastereoselectivity (de) of this *syn*-adduct was almost complete. A further interesting finding was observed when using phenyllithium as a carbon nucleophile in place of the phenylcuprate. This one-pot reaction provided the *anti*-addition product **106anti** with high *anti/syn*-selectivity and diastereoselectivities.

As an application of the stereoselective α-alkylation realized using 4-*O*-propionyl derivative **104** to natural product synthesis, the Tadano group achieved the highly stereoselective total synthesis of (–)-Lasiol **110**, an acyclic monoterpene alcohol [91] (*Scheme 28*). The 1,4-addition of the enolate generated from **104** with NaHMDS to methyl crotonate proceeded highly stereoselectively to produce the 1,4-adduct **107** as a single diastereomer. The product **107** possessed an *anti*-stereochemistry for two vicinal methyl substituents at the α- and β-carbons of the ester functionality. The diastereoselectivity of the *anti*-isomers was also quite high (>90%). Natural lasiol was isolated as an insect sex-attracting pheromone, which was found



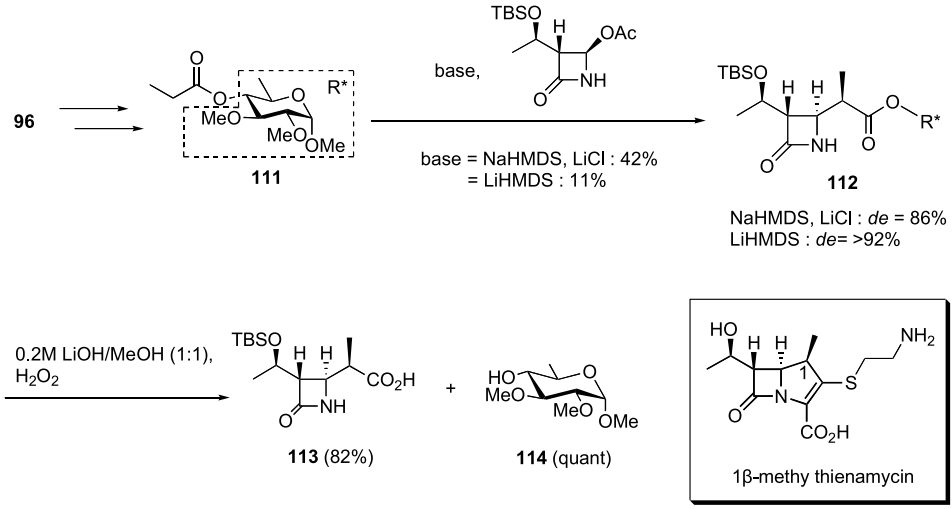
Scheme 27



Scheme 28

in the mandibular gland secretions of the male ant *Lasius meridiobalis*. The stereostructure of this pheromone was determined to be *erythro*-2,3,6-trimethyl-5-hepten-1-ol. The Tadano group synthesized (–)-lasiol conveniently from the adduct **107** as follows. Selective DIBAL-H reduction of the less-congested ester moiety in **107** efficiently produced the left-hand aldehyde **108**. A Wittig-olefination of **108** with $Ph_3P{=}C(Me)_2$ provided the 2-methylpropenylated product **109**. Reductive removal of the sugar template from **109** with excess DIBAL-H eventually provided (–)-lasiol **110**, and the sugar-template **97** was efficiently recovered.

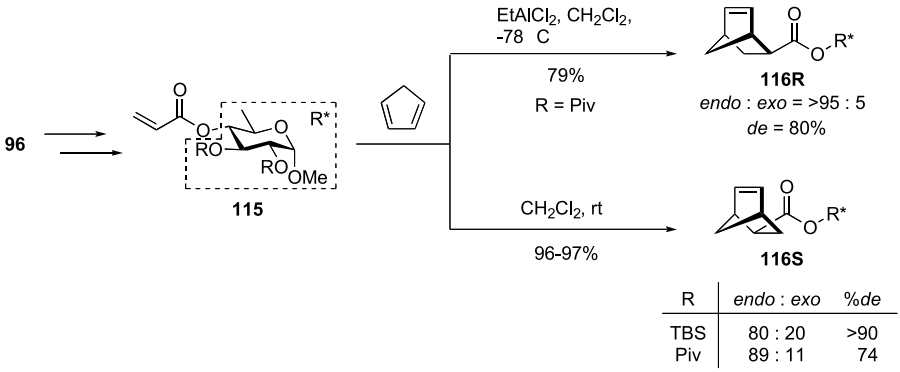
The Tadano group extended the use of the sugar-based chiral templates for the 1β-methyl carbapenem synthesis. As concerns the chemical synthesis of 1β-methyl carbapenems such as 1β-methyl thienamycin, the most common approach is a late-stage ring closure for bicyclic skeleton construction using a *C*-4 functionalized azetidin-2-one, such as **113**, namely (3*S*,4*S*)-3-[(*R*)-1-(*t*-butyldimethylsilyloxy)ethyl]-4-[(*R*)-1-carboxyethyl]azetidin-2-one, which may be constructed via the Mannich-like reaction of commercial (3*R*,4*R*)-4-acetoxy-3-[(*R*)-1-(*t*-



Scheme 29

butyldimethylsilyloxy)ethyl]azetidin-2-one and the enolates generated from chiral propionic acid esters (*Scheme 29*). After throughout searching of most effective sugar template, the Tadano group found the solution to the attempted stereoselective Mannich-like reaction [92]. As a result, the Mannich-like reaction of the enolate generated from methyl 6-deoxy-2,3-di-*O*-methyl-4-*O*-propionyl-α-D-glucopyranoside **111** and the azetidin-2-one provided the desired β-configured adduct **112** stereoselectively. The effective bases for the deprotonation of **111** were found to be LiHMDS or NaHMDS. From adduct **112**, the sugar template was readily removed without epimerization by alkaline hydrolysis in the presence of H_2O_2 to obtain the desired **113**, and the sugar template **114** was efficiently recovered.

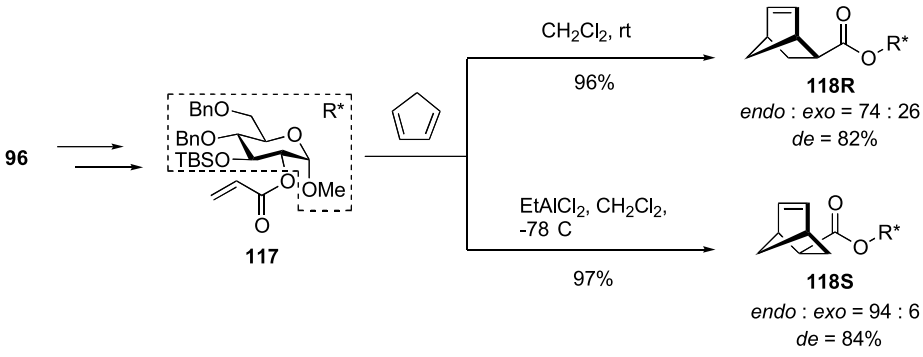
In addition to the aforementioned highly stereoselective asymmetric synthesis of the polar and radical organic reactions using a variety of sugar templates, the Tadano group extended this synthetic concept to pericyclic reactions, such as Diels–Alder reaction and 1,3-dipolar cycloaddition. Various *O*-protected methyl α-D-hexopyranosides (gluco-, *manno*- and galacto-series), in which an acryloyl ester was incorporated into an appropriate skeletal carbon, were prepared and used for thermal or Lewis acid-promoted Diels–Alder reactions [93,94]. As a representative case, the Diels–Alder reactions of the 4-*O*-acryloyl derivatives of methyl 6-deoxy-2,3-di-*O*-(*t*-butyldimethylsilyl)- and 2,3-di-*O*-pivaloyl-α-D-glucopyranoside **115** with cyclopentadiene were explored (*Scheme 30*). Under Lewis acid-promoted conditions ($EtAlCl_2$), the 2,3-di-*O*-pyvaloyl derivative **115** (R = Piv) provided one *endo*-adduct **116R** with excellent *endo*/*exo* selectivity and good π-facial selectivity (de = 80%). Under thermal conditions (rt), however, two 4-*O*-acryloyl esters **115** produced **116S** with good-to-excellent *endo*/*exo* selectivities (80:20 for 2,3-di-*O*-(*t*-butyldimethylsilyl) derivative; 89:11 for 2,3-*O*-pivaloyl derivative). Both 4-*O*-acryloyl esters **115** produced **116S** as well with good-to-excellent π-facial selectivities. From the two *endo*-adducts **116R** and **116S**, (2*S*)- and (2*R*)-enriched *endo*-5-norbornene-2-methanol were prepared by reductive removal of the sugar auxiliary, respectively.



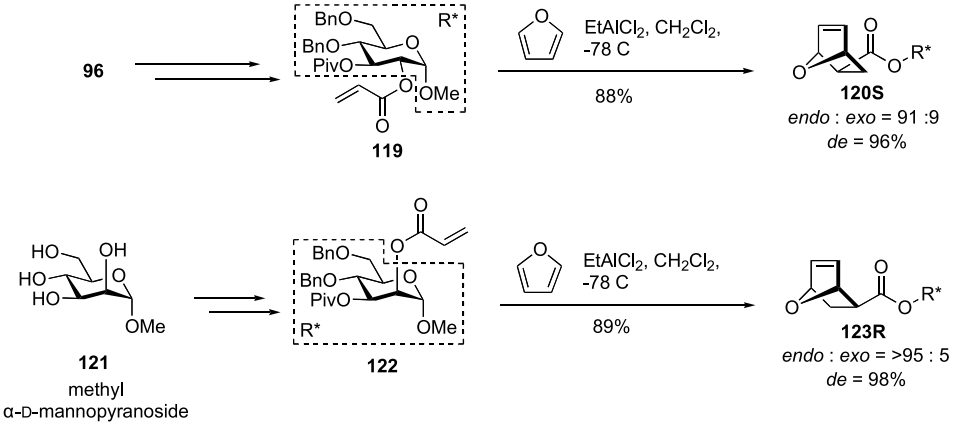
Scheme 30

As another example, the Diels–Alder reaction of 2-*O*-acryloyl derivative of methyl 4,6-di-*O*-benzyl-3-*O*-(*t*-butyldimethylsilyl)-α-D-glucopyranoside **117**, prepared from methyl α-D-glucopyranoside **96**, with cyclopentadiene, was explored under thermal or Lewis acid-mediated conditions [94] (*Scheme 31*). Under thermal (rt) conditions, the 2-*O*-acryloyl ester **117** provided one *endo*-adduct **118R** with moderate *endo/exo* selectivity (74:26) and good π-facial selectivity (de = 82%). Under Lewis acid-promoted conditions, another *endo*-adduct **118S** was a predominant product with good π-facial selectivity.

The Tadano group explored the $EtAlCl_2$-promoted Diels–Alder reactions with furan using other sugar templates [86] (*Scheme 32*). The D-gluco-type 2-*O*-acryloyl ester **119**, derived from **96**, provided an *endo*-adduct **120S** predominantly. The Diels–Alder reaction of D-*manno*-type 2-*O*-acryloyl derivative **122**, prepared from methyl α-D-mannopyranoside **121**, proceeded under the same $EtAlCl_2$-mediated conditions to provide an *endo*-adduct **123R**. Both reactions proceeded in the *endo*-mode preferentially. Most importantly, the diastereomeric excesses (π-facial selectivity) for the *endo*-adducts were magnificent, revealing that each *endo*-adduct was virtually a single diastereomer. The Diels–Alder substrates **119** and **122** differ from each other in the configuration at C-2. This difference dramatically influenced the



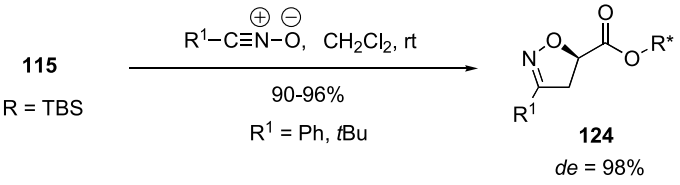
Scheme 31



Scheme 32

π-facial selectivity of the diene approach. The adducts **120S** and **123R** contain enantiomeric 7-oxabicyclo[2.2.1]hept-5-ene-2-carboxylic acid in their structures. 7-Oxabicyclo[2.2.1]hept-5-ene-2-carboxylic acid and related compounds are frequently used as chiral building blocks for natural products synthesis.

The 1,3-dipolar cycloaddition of nitrile oxide to an unsaturated ester is a useful synthetic strategy for the synthesis of heterocycles such as Δ^2-isooxazolines and α-hydroxy-γ-keto or γ-imino carboxylic acids. Thus, the 1,3-dipolar cyclo-addition of the 4-*O*-acryloyl derivative **115** (R = *t*-butyldimethylsilyl) with two nitrile oxides (R^1 = Ph or *t*-Bu) was explored by the Tadano group [95] (*Scheme 33*). In the case of benzonitrile oxide (R^1 = Ph), a functionalized Δ^2-isooxazoline **124** was obtained as a single isomer in excellent yield. Thus, the cycloaddition proceeded smoothly at room temperature with extreme stereoselectivity.



Scheme 33

3 Sugars as Chiral Ligands for Transition-Metal Catalysts

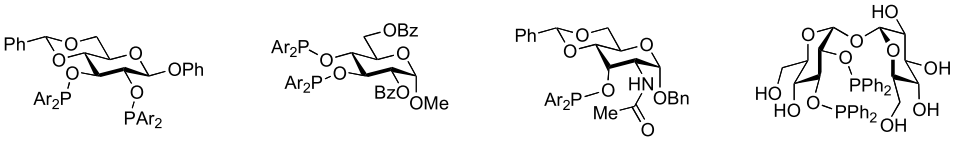
Historically the synthetic application of sugar derivatives as chiral ligands of transition-metal complexes to a variety of asymmetric reactions was initiated in the late 1970s [96,97,98,99]. The use of sugar-based chiral ligands as part of transition-metal complexes became more popular for synthetic organic chemists since the early 1990s and was thoroughly studied by several research groups [100,101,102,103,104,105,106,107,108,109,110,111,112,113,114]. Herein, some representative achievements on this subject by five research groups are summarized.

3.1 Studies by the RajanBabu Group

Since the early 1990s, the group headed by RajanBabu has achieved a variety of practical asymmetric syntheses, especially asymmetric hydrogenation, asymmetric hydrocyanation and asymmetric hydrovinylation of prochiral olefins using a variety of chiral transition-metal complexes incorporating sugar-based ligands. Several prominent papers published by the RajanBabu group since 1992 have received significant attention in the field of asymmetric synthesis [115,116,117,118,119,120,121,122,123,124,125,126,127,128,129]. Some representative examples of their studies are summarized below.

These researchers designed a variety of hexopyranose derivatives bearing mono-*O*-phosphinyl (OPR_2) or di-*O*-phosphinyl groups prepared by phosphinylation of appropriate hydroxyl group(s) in the starting mono- and disaccharide skeletons such as partially protected D-glucopyranosides. These ligands were readily prepared by treatment of appropriate sugar derivatives bearing unprotected hydroxyl group(s) with appropriate chlorodiarylphosphines. Usually, the designed sugar phophinites possess functionalized aryl (Ar) groups as substituents on the phosphorous element. Representative sugar phosphinites **125–129** developed by the RajanBabu group are shown in *Fig. 1*. Using these sugar-based phophinites as chiral ligands of transition-metal complexes (Ni, Rh, Pd), the RajanBabu group demonstrated catalytic asymmetric carbon–hydrogen and carbon–carbon bond-forming reactions. The transition-metal catalysts were in general prepared by mixing the phosphinite-type ligands and appropriate transition-metal salts such as $Ni(COD)_2$. Some of these transition-metal complexes are particularly valuable because of their water-soluble nature.

The asymmetric hydrocyanation of vinylarenes was explored by the RajanBabu group for the initial evaluation of the sugar phosphinite-type ligands as effective stereocontrolling elements. In 1992, they reported their initial results on the hydrocyanation of some vinylnaphthalenes such as 6-methoxy-2-vinylnaphthalene **130** with HCN in the presence of a catalytic amount (1.0–5.0 mol%) of Ni(0) complexes prepared from D-glucose-derived ligand **125** [115] (*Scheme 34*). The chemical yield and enantioselectivity (ee) of the hydrocyanation product **131** were both at levels of general use. To realize higher enantioselectivities in the hydrocyanation, they recognized that electron-withdrawing substituent(s), such as a CF_3 group, were requisite in the aryl groups of the phosphinites. The RajanBabu group also studied the mechanism of the hydrocyanation in depth through kinetic studies and spectral analysis of intermediary organometallic species [119]. The RajanBabu group devised other sugar phosphinites as the ligands for the hydrocyanation of vinylarenes [120]. In 2006, the RajanBabu group reported



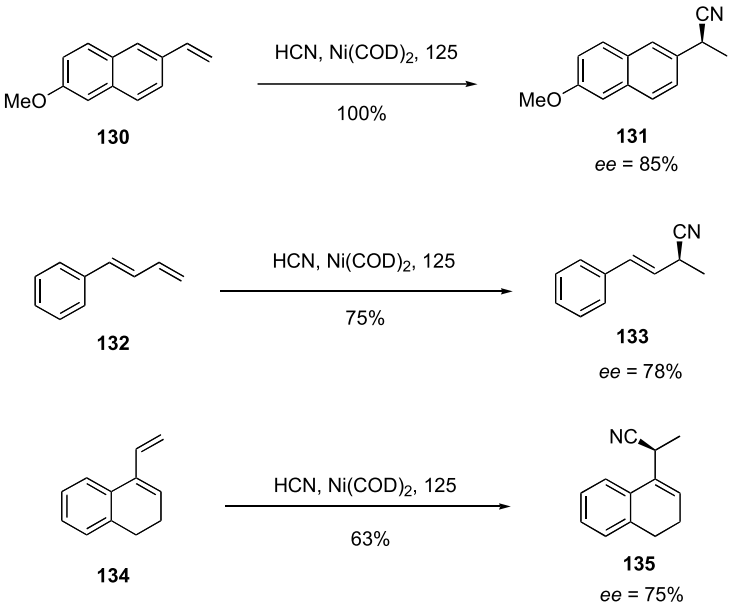
125 : Ar = 3,5-$(CF_3)_2$-C_6H_3
126 : Ar = 3,5-Me_2-C_6H_3

127
Ar = 3,5-Me_2-C_6H_3

128
Ar = 3,5-Me_2-C_6H_3

129

Figure 1
The sugar phosphinites designed by the RajanBabu group

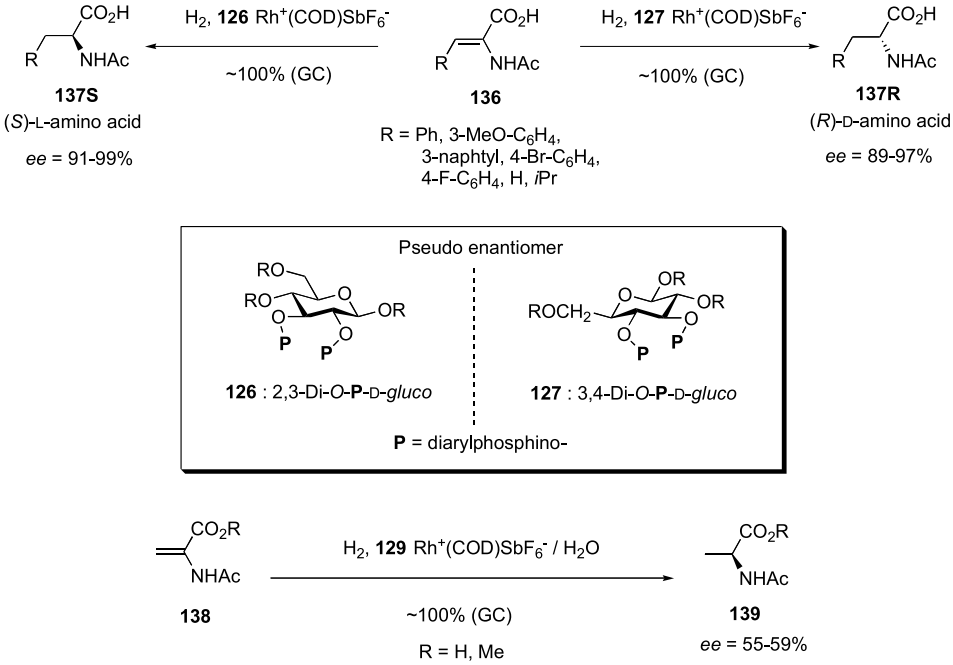


Scheme 34

the Ni(0)-catalyzed asymmetric hydrocyanation of 1,3-dienes using the ligand **125** [129]. As shown in *Scheme 34*, the asymmetric hydrocyanation of 1-phenyl-1,3-butadiene **132** proceeded regioselectively to provide the 1,2-adduct **133** exclusively in good chemical yield. The enantiomeric excess of the product was 78%. The similar hydrocyanation of 1-vinyl-3,4-dihydronaphthalene **134** using the ligand **125** provided **135** in somewhat lesser yield and enantioselectivity.

In addition, the RajanBabu group explored the Rh(I)-catalyzed asymmetric hydrogenation of a variety of dehydroamino acids, i. e., β-arylated α-acetamido-acrylic acid esters or acrylic acids **136**, using the sugar diphosphinite-based ligands, to prepare L- or D-amino acids [117,121] (*Scheme 35*). As in the cases of the asymmetric hydrocyanation of vinylarenes, sugar phosphinites, such as **126** and **127**, served as excellent stereocontrolling elements to produce either functionalized (*S*)- or (*R*)-phenylalanines with high-to-excellent enantioselectivity.

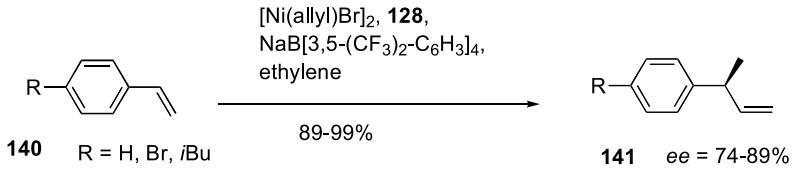
Interestingly, the sense of the chirality (*R* or *S*) of the produced phenylalanine derivatives **137** was dependent on the relative juxtaposition of the vicinal diphosphinites on a sugar backbone. Thus, the hydrogenation (H_2 under 30–40 psi) of **136** in the presence of a catalytic amount of Rh(I)(COD)SbF_6 complex with a 2,3-bis-*O*-[(diphenyl)phosphino]-D-glucopyranoside-type ligand **126** produced predominantly (*S*)-L-phenylalanine derivatives **137S**. Importantly, the electronic tuning of a ligand system has resulted in an enhancement of enantioselectivity. On the other hand, a 3,4-bis-*O*-[(diphenyl)phosphino]-D-glucopyranoside-type ligand **127** produced the corresponding (*R*)-D-amino acid **137R** with excellent enantioselection. Therefore, these sugar phosphinites **126** and **127** are in a pseudo-enantiomeric relationship as chiral ligands for the Rh(I)-catalyzed hydrogenation of dihydroamino acids. In sharp contrast to the



Scheme 35

hydrocyanation of vinylarenes, the ligands with electron-withdrawing groups in the aromatic rings, such as **125**, decreased in the enantioselectivity of hydrogenation. The asymmetric hydrogenation of dehydroamino acids was explored further using other D-fructofuranose-type diphosphinites with less enantioselectivity. As an extension of the sugar phosphinite-based asymmetric hydrogenation, the RajanBabu group successfully explored the hydrogenation of dimethyl itaconate [123]. Furthermore, disaccharide phosphinites, i. e., the α,α'-trehalose-derived diphosphinites **129**, were used as water-soluble catalysts for the Rh(I)-catalyzed hydrogenation of dehydroamino acids **138** to prepare (*S*)-alanine derivatives **139**; however, the observed enantioselectivity was unremarkable [122,125].

A D-allosamine-derived monophosphinite **128** served as a good chiral ligand for the asymmetric Ni(0)-catalyzed hydrovinylation of styrene derivatives **140** with ethylene as shown in *Scheme 36* [126,127,128]. Among some 2-acetamido-2-deoxy-3-*O*-diarylphosphinyl derivatives prepared from D-hexopyranoses, the ligand **128** revealed a useful level of enantios-



Scheme 36

electivity, which produced (*S*)-1-(methyl)allylbenzenes **141** with high enantioselectivity. One of the hydrovinylation products, i. e., 1-bromo-4-[(*S*)-1-methylallyl]benzene, was converted to (*R*)-ibuprofen, a pharmacologically important 2-arylpropionic acid.

3.2 Studies by the Diéguez/Ruiz/Claver Group

Since the late 1990s, the group consisting of Diéguez, Ruiz and Claver in Spain has explored asymmetric synthesis using sugar diphosphites [general formula = $OP(OR)_2$], sugar phosphine-phosphites, or sugar thioether-phosphinites as chiral ligands of transition-metal complexes. Representative ligands are shown in ❯ *Fig. 2*. Most of these sugar-derived ligands contain one/two phosphite moieties consisting of 2,2′-dihydroxy-1,1′-biphenyl or 2,2′-dihydroxy1-1,1′-binaphthyl moieties. Using these D-*aldo*-hexo- or pentofuranose derivatives as chiral ligands of a Cu, Rh, or Ni catalyst, asymmetric hydroformylation, hydrogenation, 1,4-addition, or hydrosilylation of prochiral olefins and asymmetric allylic alkylation was achieved [130,131,132,133,134,135,136,137,138,139,140,141,142,143,144,145,146,147,148,149,150,151,152,153]. The results obtained by this group are summarized below.

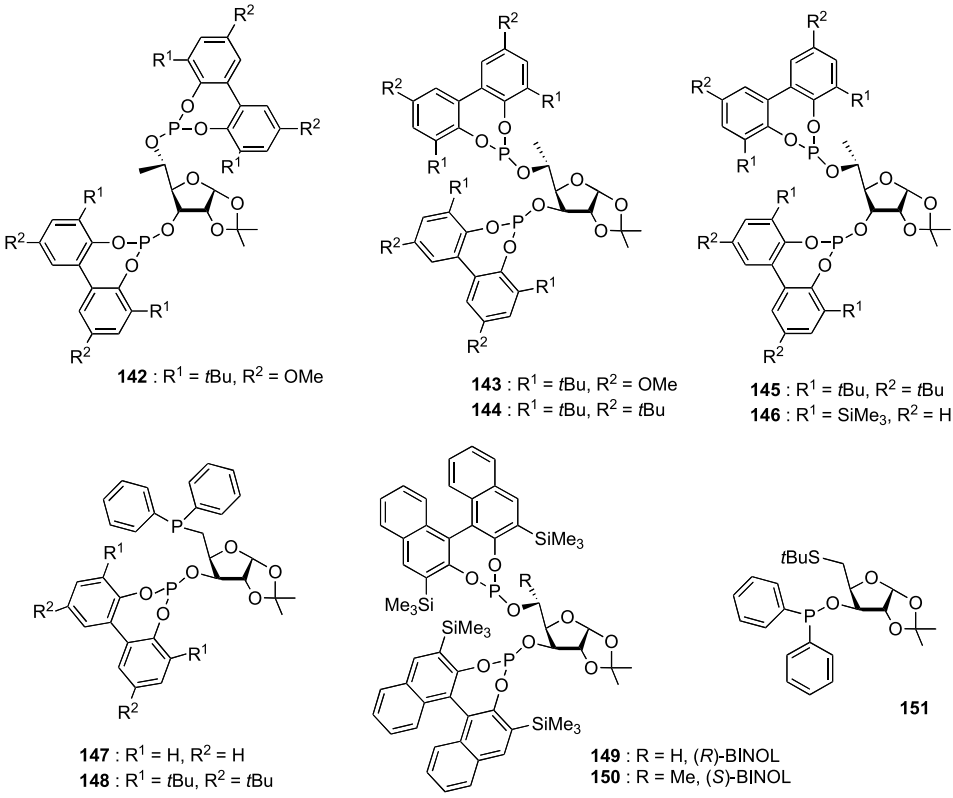
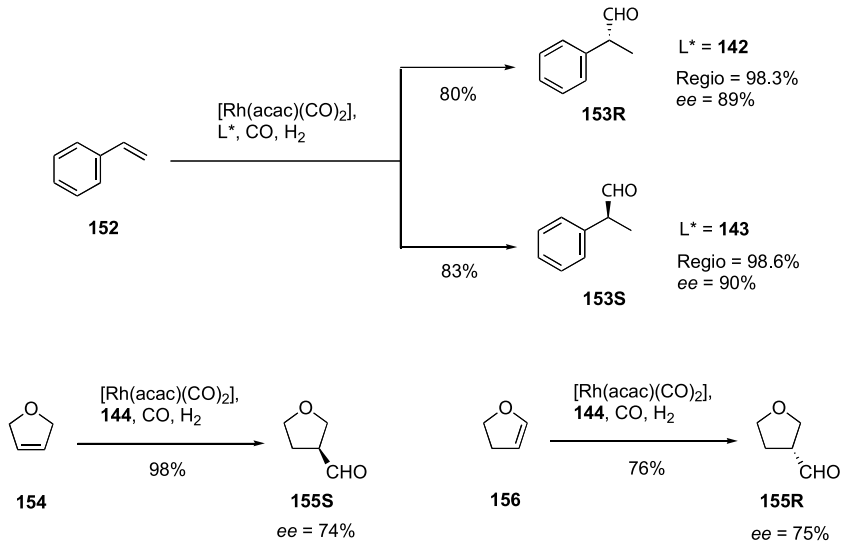


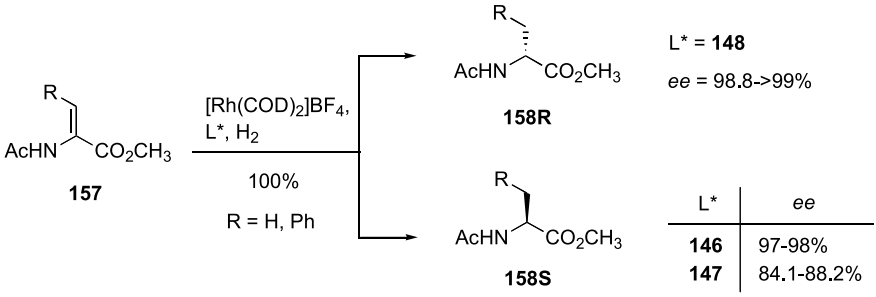
Figure 2
The sugar-derived chiral ligands designed by the Diéguez/Ruiz/Claver group

Asymmetric Rh-catalyzed hydroformylation of vinylarenes has attracted much attention as a convenient tool for obtaining enantiopure or -enriched functionalized aldehydes. The Diéguez/Ruiz/Claver group demonstrated the efficiency of the sugar-derived ligands as regio- and stereocontrolling elements for the Rh-catalyzed hydroformylation of vinylarenes or dihydrofurans [133,138,140,142,147]. Representative achievements reported by the Diéguez/Ruiz/Claver group on the asymmetric hydroformylation are shown in ❯ *Scheme 37*. Using ligand **142**, the hydroformylation of styrene **152** with CO/H_2 (10 bar) in the presence of an $Rh(acac)(CO)_2$ catalyst (0.1 mol%) in toluene provided (*R*)-2-phenylpropanal **153R** with 89% enantiomeric excess. The ligand **142** was prepared from D-glucose as follows. Deoxygenation of the C-6 hydroxyl group and stereochemical inversion of both C-3 and C-5 configurations of 1,2-*O*-isopropylidene-α-D-glucofuranose provided 6-deoxy-1,2-*O*-isopropylidene-L-talofuranose. Treatment of this sugar derivative with appropriate diarylated phosphorochloridite $[(ArO)_2PCl]$ in pyridine afforded the ligand **142**. Even more interestingly, a similar Rh-catalyzed hydroformylation of styrene in the presence of **143**, a 6-deoxy-D-glucofuranose-type ligand, produced (*S*)-2-phenylpropanal **153S** with a high ee of 90%. The configurations at C-3 and C-5 in the sugar part of the ligand, thus, direct the absolute configuration of the product of asymmetric hydroformylation.

The Diéguez/Ruiz/Claver group demonstrated another application of the sugar diphosphite-based asymmetric synthesis to the hydroformylation of two types of dihydrofurans. The Rh-catalyzed hydroformylation of 2,5-dihydrofuran **154** with CO/H_2 in the presence of ligand **144**, a modified **143**-type ligand with two bulky *tert*-butyl substituents in both biphenyl parts, produced (*S*)-3-formyltetrahydrofuran **155S** with moderate enantioselectivity (74% ee). When 2,3-dihydrofuran **156** was treated under the same hydroformylation conditions using the same



Scheme 37

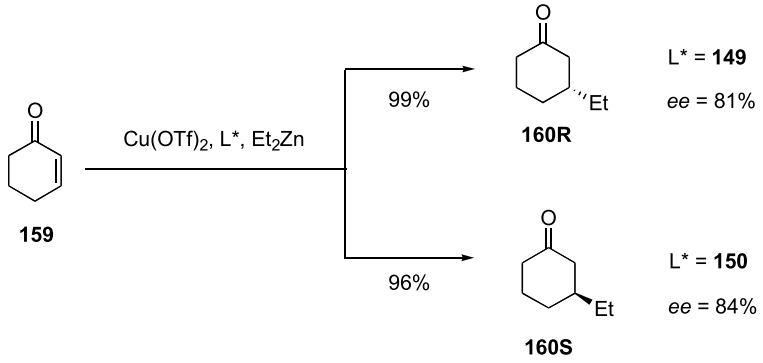


Scheme 38

ligand **144**, the opposite (*R*)-3-formyltetrahydrofuran **155R** was obtained with 75% ee. Thus, the sugar phosphite-based ligand **144** could discriminate the π-face in two substrates, **154** and **156**.

The utility of the phosphorus-containing chiral ligands has been expanded to the transition-metal-catalyzed asymmetric hydrogenation of prochiral olefins by the Diéguez/Ruiz/Claver group [132,134,137,143,146]. Some studies on this subject, which resulted in the achievement of high enantioselectivity, are shown in *Scheme 38*. Some *N*-acetyldehydroamino acid esters **157** were subjected to Rh(I)-catalyzed hydrogenation in the presence of sugar phosphine-phosphite-type ligands **147** and **148**. The ligands **147** and **148** were prepared from 1,2-*O*-isopropylidene-α-D-xylofuranose via the ring opening of the corresponding 3,5-anhydrosugar (a 3,5-oxetane derivative) by a PPh_2 anion and the subsequent phosphite formation of C-3 hydroxyl with an appropriate diarylphosphorochlorite. Using the catalyst consisted of the ligand **148** and $[Rh(COD)_2]BF_4$. *N*-acetylated (*R*)-alanine methyl ester or *N*-acetylated (*R*)-phenylalanine methyl ester **158R** was obtained quantitatively with excellent enantiomeric excess. On the contrary, the (*S*)-enantiomers of these amino acids **158S** were obtained in a highly enantioenriched form by switching the chiral ligand **148** to those derived from a sterically less congested sugar phosphine-phosphite or 1,2-*O*-isopropylidene-α-D-glucofuranose-derived diphosphite, i. e., **147** or **146**, respectively.

The asymmetric 1,4-addition of a carbon nucleophile, such as an organometallic species, to prochiral α,β-unsaturated esters (Michael reaction) is one of the most actively investigated areas in current organic synthesis. Consequently, synthetically relevant β-functionalized carbonyl compounds can be prepared in an enantiopure or -enriched form. For this purpose, several ligands for transition-metal catalysts have been developed so far by structural tuning of chiral compounds from natural or unnatural sources. Using some sugar diphosphite-type ligands, the Diéguez/Ruiz/Claver group investigated Cu(II)-catalyzed enantioselective 1,4-addition of organometallic species to α,β-unsaturated compounds [130,131,135,139,150]. Representative results on the Cu(II)-catalyzed 1,4-addition to cyclohex-2-enone **159** with sugar diphosphite-based catalysts are shown in *Scheme 39*. The nucleophilic addition of diethylzinc to **159** in the presence of a catalytic amount of $Cu(OTf)_2$ in CH_2Cl_2 at 0 °C proceeded almost exclusively in a 1,4-addition fashion. With the ligand **149**, this 1,4-addition produced (*R*)-3-ethylcyclohexanone **160R** quantitatively in 81% enantiomeric excess. The ligand **149** contains two (*R*)-1,1′-binaphthol skeletons, both carrying two bulky trimethylsilyl groups at

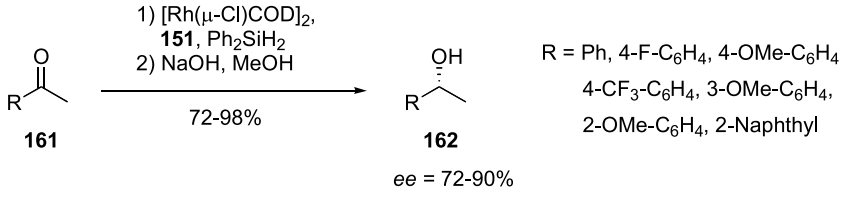


Scheme 39

C-2 and C-2′. The ligand **149** was prepared from 1,2-*O*-isopropylidene-α-D-xylofuranose with a methodology similar to that used for the preparation of **143** and **144**. By switching the chiral ligand **149** to **150**, the 3,5-diphosphites of 6-deoxy-1,2-*O*-isopropylidene-α-D-glucofuranose with two (*S*)-BINOLs, the 1,4-addition of diethylzinc to **159** produced the (*S*)-enantiomer **160S** with a remarkable enantiomeric excess of 84%.

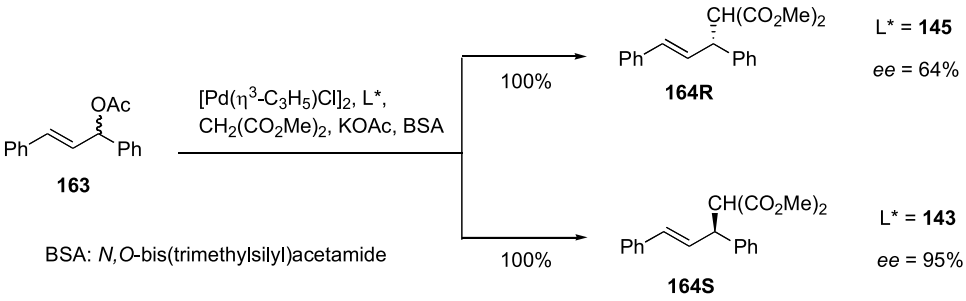
As a surrogate of the hydride-reduction of carbonyl compounds, the transition-metal-catalyzed hydrosilylation of ketones has been actively investigated in current organic synthesis. The Diéguez/Ruiz/Claver group has explored the Rh-catalyzed asymmetric hydrosilylation of ketones [141,151]. As shown in *Scheme 40*, a variety of acetophenones were subjected to $[Rh(\mu\text{-}Cl)COD]_2$-catalyzed hydrosilylation with Ph_2SiH_2 in the presence of a sugar-based chiral ligand. After considerable experimentation, the Diéguez/Ruiz/Claver group found that a novel sugar-based ligand, namely a D-xylofuranose-type thioether-phosphinite **151**, worked well as an effective stereocontrolling element for the aimed hydrosilylation. Using the sugar thioether-phosphinite **151**, various acetophenones **161** were converted to (*R*)-1-arylated ethanols **162** with moderate-to-high enantiomeric excesses after methanolysis of intermediary silyl ethers. A bulky group (*tert*-butyl) on the thioether moiety of the ligand was essential for the realization of the observed high enantioselectivity.

Besides the transition-metal-catalyzed asymmetric addition reactions to prochiral olefins, the substitution reaction of a carbon nucleophile to allylic esters has been investigated using a variety of chiral transition-metal catalysts. Using the aforementioned sugar diphosphites



Scheme 40

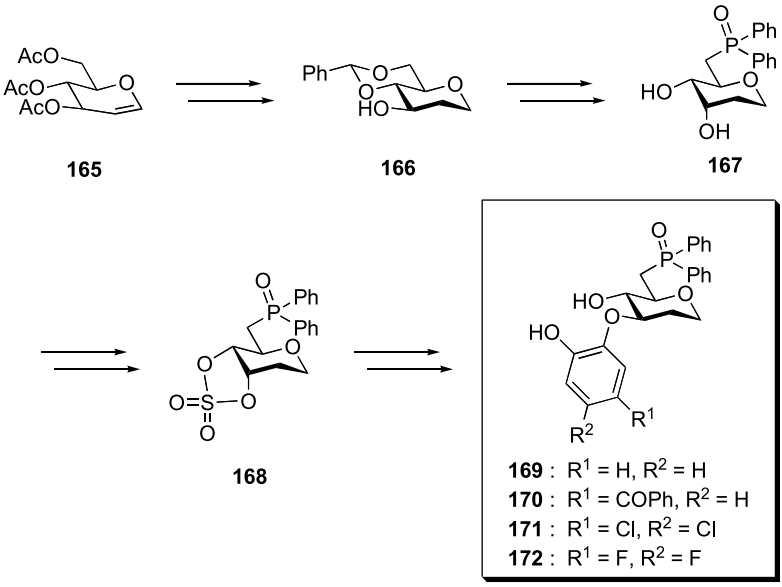
as effective and enantioselective ligands, the Diéguez/Ruiz/Claver group developed highly stereoselective Pd(0)-catalyzed allylic alkylations [136,148,149,152,153]. As shown in *Scheme 41*, racemic 1,3-diphenyl-3-acetoxyprop-1-ene **163** was treated with the anion of dimethyl malonate in the presence of a catalytic amount of $[Pd(\eta^3\text{-}C_3H_5)Cl]_2$ to produce the desired allylic substitution product **164**. The use of **145** resulted in the preferential formation of (*R*)-enantiomer **164R** with moderate stereoselectivity. On the other hand, the ligand **143** provided the (*S*)-enantiomer **164S** with a high enantiomeric excess of 95%.



Scheme 41

3.3 Studies by the Shibasaki/Kanai Group

Among the sophisticated approaches to the asymmetric synthesis catalyzed by a Lewis acid–Lewis base bifunctional system, the Shibasaki/Kanai group designed and used new sugar-based chiral ligands of metallic catalysts to achieve synthetically important carbon–carbon bond-forming reactions. In 2000, they reported a sugar phosphine oxide prepared from D-glucose, which demonstrated its possible use as a scaffold for the anticipated Lewis acid–Lewis base bifunctional system. Since then, the Shibasaki/Kanai group has actively explored the structural tuning of the original sugar-based ligand, which could result in the improvement of the enantioselectivity of the carbon–carbon bond-forming reactions. Using the modified ligands, the Shibasaki/Kanai group has reported a number of remarkable asymmetric carbon–carbon bond-forming reactions, such as the cyanosilylation of prochiral carbonyl compounds, the Strecker reaction of ketoimines, and the conjugate addition of cyanides to α,β-unsaturated *N*-acylpyrroles. The Shibasaki/Kanai group has utilized some enantiopure products obtained by the attempted asymmetric approaches for the synthesis of pharmacologically important natural products [154,155,156,157,158,159,160,161,162,163,164,165,166,167,168,169,170,171,172,173]. Some of the sugar-based ligands for transition-metal catalysts are shown in *Scheme 42*. The improved and practical synthesis of sugar phosphine oxides **169–172** was started from tri-*O*-acetyl-D-glucal **165**. Via a 4,6-*O*-benzylidene-2-deoxy-1,5-anhydro sugar derivative **166**, a 6-deoxy-6-diphenylphosphine oxide ($P(O)Ph_2$) derivative **167** was obtained by treatment of the 6-*O*-tosyl derivative with Ph_2PK and subsequent oxidation with H_2O_2, followed by formation of the 3-*O*-catechol ether using the 3,4-cyclic sulfate **168**. Some acyl- or halogen-substituted catechols were used for the last S_N2 ring-opening

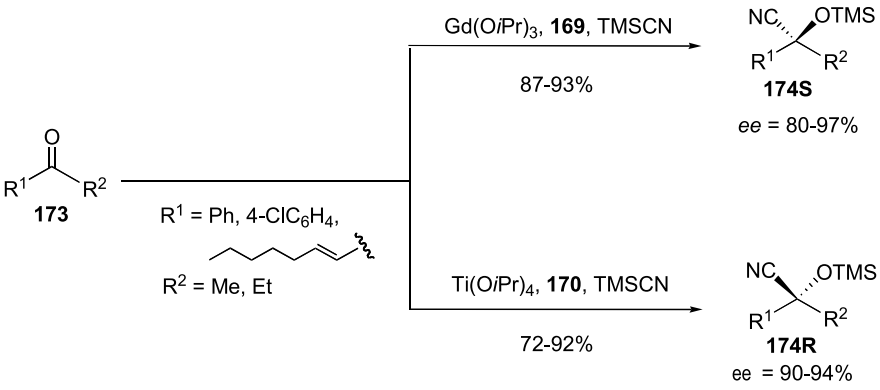


Scheme 42

of the cyclic sulfate **168** [159]. Recently, the synthesis of these ligands from racemic allyl oxirane and 3-butenol, was reported by the Shibasaki/Kanai group using the Trost's asymmetric dihydropyran formation as a key step [169]. Initially, the Shibasaki/Kanai group used the 2,3-diol-6-phosphine oxide derivative as a ligand for cyanosilylation. This type of ligand was less effective for the cyanosilylation of benzaldehyde with TMSCN in the presence of Et_2AlCl [154].

The enantioselective cyanosilylation of a variety of functionalized ketones has been explored using the above ligands in the presence of Al, Ti, Yb, Gd, Sm, La, or Zr-centered organometallic compounds [155,157,158,164]. Such studies included the treatment of aromatic and unsaturated ketones **173**, such as acetophenone, with TMSCN in the presence of a catalytic amount of the Lewis acid–Lewis base bifunctional catalysts prepared from $Gd(OiPr)_3$ and a sugar-based ligand, such as **169** or **170** (*Scheme 43*). Using the chiral catalyst prepared from $Gd(OiPr)_3$ and ligand **169**, trimethylsilylated cyanohydrines with (*S*)-configuration **174S** were obtained with high-to-excellent enantiomeric excesses. On the other hand, the chiral catalysts prepared from $Ti(OiPr)_4$ and other sugar phosphine oxide **170** produced the (*R*)-enantiomers **174R** with excellent enantiomeric excesses.

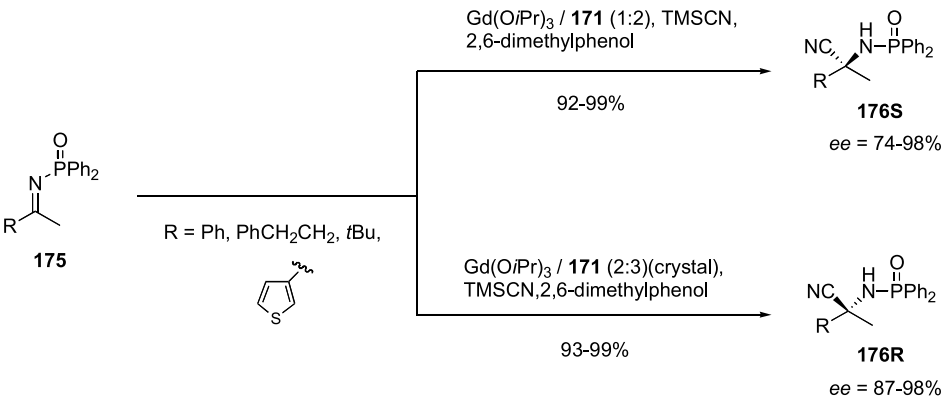
With the establishment of the highly stereoselective cyanosilylation of functionalized ketones using these sugar-based ligands, the Shibasaki/Kanai group applied the developed methodologies to the synthesis of medicinally and biologically important compounds, such as (1) the key intermediate for the (20*S*)-camptothecin family [158,160,164], (2) essential building blocks for the HIV protease inhibitor and bestatin [156], (3) a muscarinic receptor antagonist (*S*)-oxybutynin [161,168], (4) a novel antitumor agent and highly selective serine/threonine phosphatase PP2A inhibitor fostriecin [162], and (5) 2-[(2*R*)-arylmorpholin-2-yl]ethanol, a key intermediate of neurokinin receptor antagonists [163].



Scheme 43

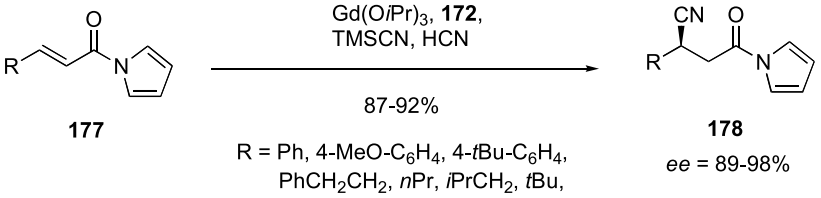
The Shibasaki/Kanai group has explored the catalytic enantioselective Strecker reaction of ketoimines for the extension of utility of the sugar-based ligands [165,166,167]. The products of the attempted Strecker reaction were converted to enantioenriched α-differentially disubstituted amino acids, synthetically valuable building blocks, by conventional functional group manipulation. As shown in *Scheme 44*, under optimized conditions using the Gd(O*i*Pr)$_3$/ligand **171** in a ratio of 1:2 or 2:3, the Strecker reaction of a variety of *N*-diphenylphosphinoyl ketoimines **175** with TMSCN proceeded with high-to- excellent enantioselectivity [173]. Interestingly, the ratio of Gd(O*i*Pr)$_3$ to the ligand governed the sense of enantioselectivity; thus, the 1:2 complex predominantly produced (*S*)-enantiomers **176S**. On the other hand, the 2:3 complex provided the (*R*)-enantiomers **176R** with high-to-excellent enantioselection.

A sugar-based ligand such as **172** served as a promising chiral ligand for the catalytic enantioselective conjugate addition of cyanide to α,β-unsaturated *N*-acylpyrroles as shown in



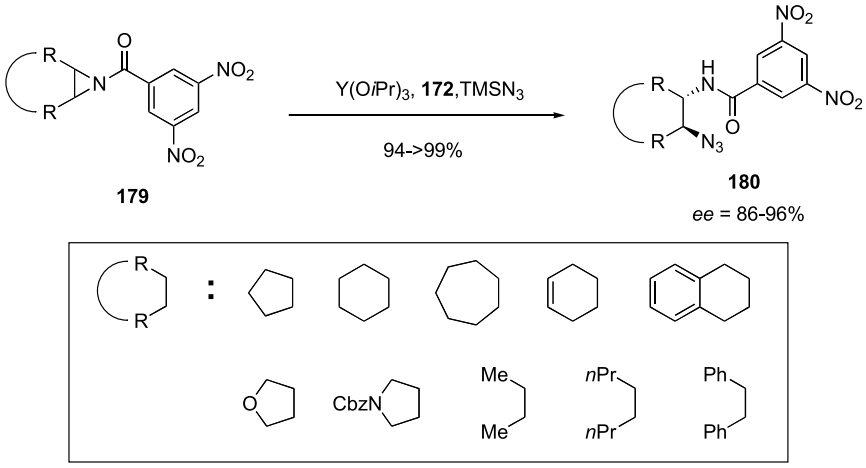
Scheme 44

Scheme 45 [170]. Thus, treatment of several α,β-unsaturated *N*-acylpyrroles **177** with TMSCN/HCN (0.5 or 1:2) in the presence of Gd(O*i*Pr)$_3$/**172** provided β-cyano-β-arylated or β-cyano-β-alkylated *N*-propionyl pyrroles **178** in good yields with excellent enantioselectivities. Some of the 1,4-adducts were efficiently converted to several pharmaceuticals, such as β-phenyl-GABA or (*ent*)-pregabalin.



Scheme 45

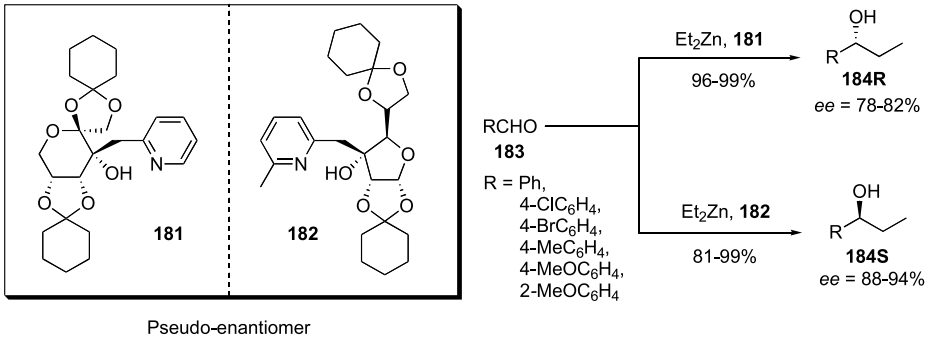
The potential of sugar-based ligand **172** as an enantiocontrolling element for asymmetric synthesis has been recently enhanced by the catalytic asymmetric ring opening of *meso*-aziridines. The results of the Y(O*i*Pr)$_3$/**172**-catalyzed ring opening of *meso*-aziridines **179** in the presence of $TMSN_3$ are shown in *Scheme 46* [172]. Similar to the aforementioned enantioselective reactions, the ring opening reactions proceeded smoothly with high enantioselectivity to produce azide compounds **180**. From one of the ring-opened products, an orally active anti-influenza drug, Tamiflu (oseltamivir phosphate), was prepared.



Scheme 46

3.4 Studies by the Zheng/Chen Group

As a ubiquitous chiral scaffold existing in nature, D-fructose is a readily available sugar, like as D-glucose. Since 2003, Zheng, Chen, and co-workers have investigated the utility of



Scheme 47

D-fructose-derived compounds as efficient chiral ligands for stereoselective carbon–carbon bond-forming reactions [174,175,176,177,178,179]. The remarkable results found by the Zheng/Chen group are summarized below.

The initial achievement was the highly enantioselective addition of diethylzinc to aldehydes, as shown in *Scheme 47* [174,175]. The Zheng/Chen group has developed several novel sugar-based ligands. Two representative new sugar-based ligands **181** and **182** are shown in *Scheme 47*. Ligand **181** was prepared from 1,2:4,5-di-*O*-cyclohexylidene-β-D-fructopyranose via the nucleophilic addition of the lithium anion generated from 2-methylpyridine to the corresponding 3-ulose. On the other hand, ligand **182** was prepared from 1,2:5,6-di-*O*-isopropylidene-α-D-glucofuranose in a similar way to that used for **181**, using the monolithio anion of 2,6-dimethylpyridine instead of 2-methylpyridine. Treatment of a variety of aromatic aldehydes **183** with diethylzinc in the presence of **181** produced (*R*)-adducts **184R** in high yields and with moderate-to-high enantiomeric excesses in all cases. By switching the ligand to **182**, the same addition conditions produced the (*S*)-enantiomers **184S** efficiently. Therefore, ligand **181** and **182** seem to be in a pseudo-enantiomeric relationship. The enantioselective additions of diethylzinc to aliphatic or cinnamyl aldehydes in the presence of **181** produced respective (*R*)-adducts in lower enantiomeric excesses.

In addition, the Zheng/Chen group has studied the usefulness of some sugar monophosphites as efficient chiral ligands for the catalytic asymmetric hydrogenation of prochiral olefins [176,177,178,179]. Some of the sugar monophosphites designed by the Zheng/Chen group are shown in *Fig. 3*. Ligands **185** were prepared from 1,2:4,5-di-*O*-cyclohexylidene-D-fructofuranose via stereoinversion of C-3, followed by phosphite formation at the C-3 hydroxy group with PCl_3 and then optically active BINOL. The ligand **185(*R*)** possesses (*R*)-BINOL, while **185(*S*)** possesses (*S*)-BINOL in their phosphite parts. Other sugar-based ligands, **186(*R*)** and **186(*S*)**, were prepared from 1,2:5,6-di-*O*-cyclohexylidene-α-D-glucofuranose by treatment with PCl_3 and then (*R*)- and (*S*)-BINOL, respectively. Finally, ligands **187(*R*)** and **187(*S*)** were prepared from 1,3:4,6-di-*O*-benzylidene-D-mannitol, a readily supplied sugar alcohol, via selective *O*-methylation and subsequent treatment with PCl_3 and then (*R*)- and (*S*)-BINOL, respectively.

Using these sugar-based ligands, the Zheng/Chen group has explored a variety of asymmetric hydrogenation of prochiral olefins. Three types of hydrogenation examined are

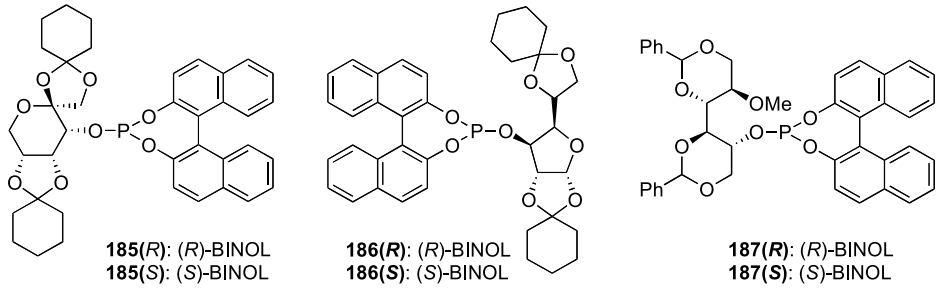


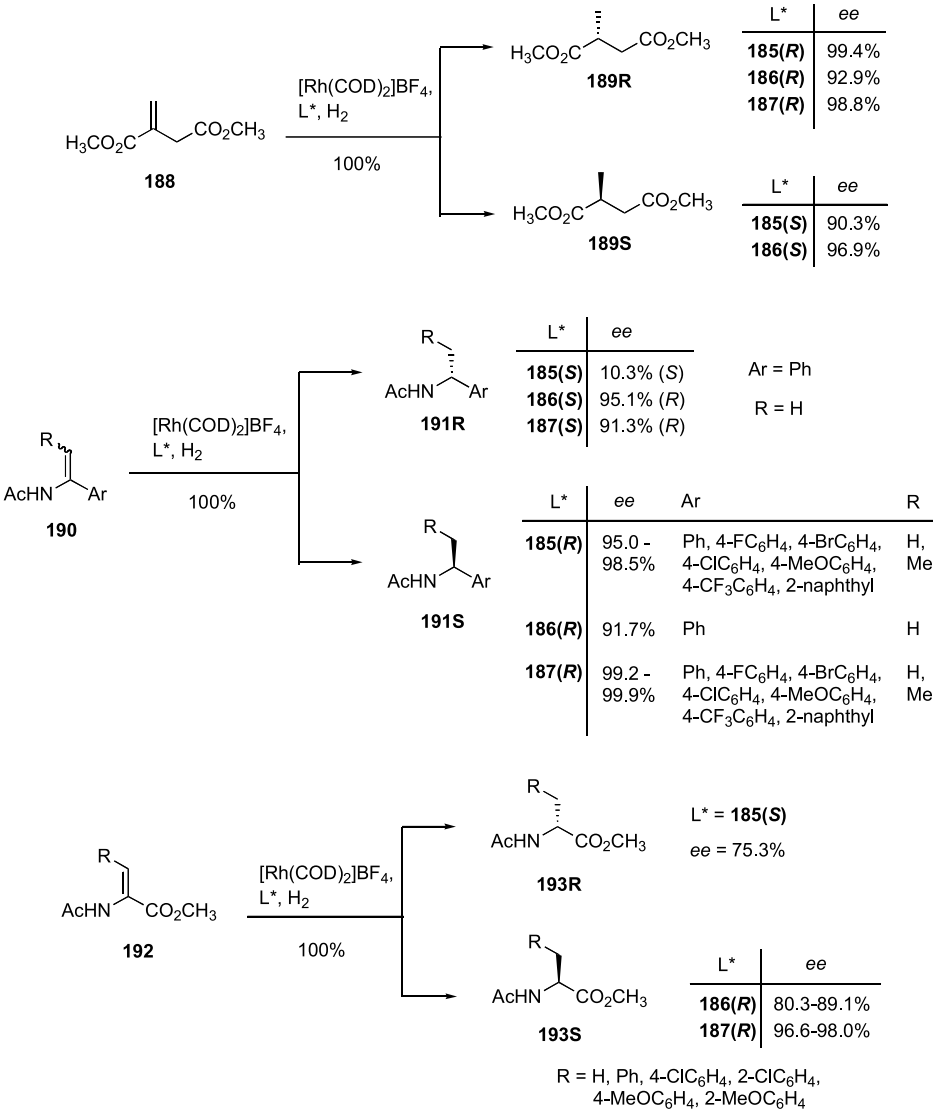
Figure 3
The sugar monophosphinites designed by the Zheng/Chen group

shown in *Scheme 48*. The hydrogenation of dimethyl itaconate **188** in the presence of a transition-metal complex consisting of $[Rh(COD)_2]BF_4$ and ligands carrying (*R*)-BINOL, i. e., **185(*R*)**, **186(*R*)**, or **187(*R*)** produced dimethyl (*R*)-2-methylsuccinate **189R** all quantitatively with excellent enantiomeric excesses. The hydrogenation of **188** in the presence of $[Rh(COD)_2]BF_4$/**185(*S*)** or **186(*S*)** quantitatively produced the (*S*)-enantiomer of **189** with more then 90% ee. Employing the same chiral catalysts **185**–**187**, the hydrogenation of 1,1-disubstituted α-arylenamides [-C=C(NHAc)Ar] **190** was investigated. Except for the case performed in the presence of **185(*S*)**, all ligands produced (*R*)-α-aryl-*N*-acetylalkylamines **191R** or (*S*)-α-aryl-*N*-acetylalkylamines **191S** with excellent enantiomeric excesses. The ligands incorporating (*S*)-BINOL produced **191R**, while those incorporating (*R*)-BINOL produced **191S**. These sugar-based ligands also served as excellent stereocontrolling elements for the Rh-catalyzed asymmetric hydrogenation of β-substituted or unsubstituted α-dehydroamino acids **192**. The results achieved using the same Rh-complex are as follows: (1) ligand **185(*S*)** produced **193R** with moderate enantioselectivity; (2) both **186(*R*)** and **187(*R*)** produced **193S** with moderate-to-excellent enantioselectivity.

3.5 Studies by the Davis Group

Recently, amino sugars such as D-glucosamine derivatives have been studied as sugar-based ligands for transition-metal-catalyzed asymmetric synthesis by Davis and co-workers. The results obtained by the Davis group are summarized below. The amino sugar-based ligands introduced by the Davis group are shown in *Fig. 4*. Ligands **194** to **197** were prepared from methyl 2-acetamino-4,6-*O*-benzylidene-2-deoxy-α- or β-glucopyranoside, readily available from D-glucosamine, via modification of the 2-amino group and stereoinversion of C3 (for **196**, a D-allosamine derivative). Using these ligands, asymmetric diethylzinc addition to aldehydes, asymmetric Reformatsky reactions, and asymmetric alkynylation of aldehydes were explored.

Using these ligands, the Davis group has reported the asymmetric additions of diethylzinc to aromatic aldehydes as the initial exploration of the asymmetric induction realized by amino sugar-based ligands [180]. As shown in *Scheme 49*, three aromatic aldehydes **198** were treated with diethylzinc in the presence of **194** or **195**. The addition products **199** were obtained in



Scheme 48

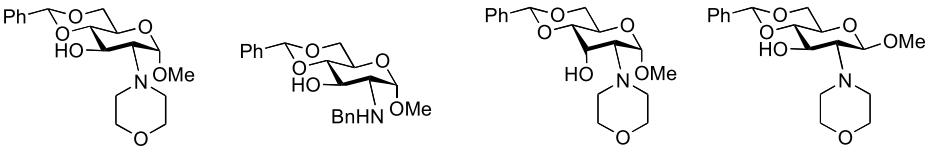
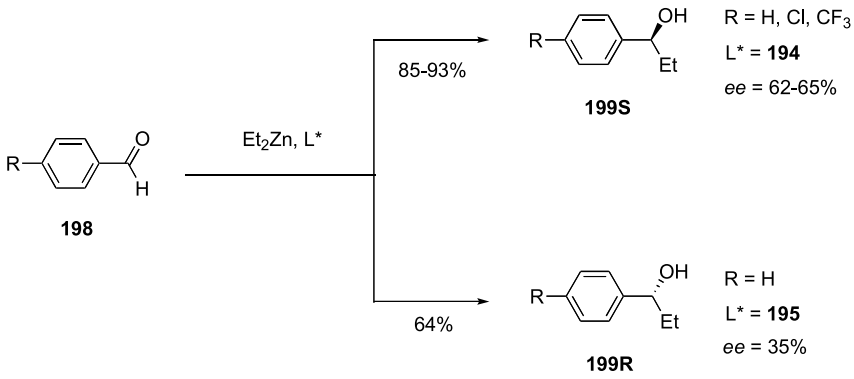


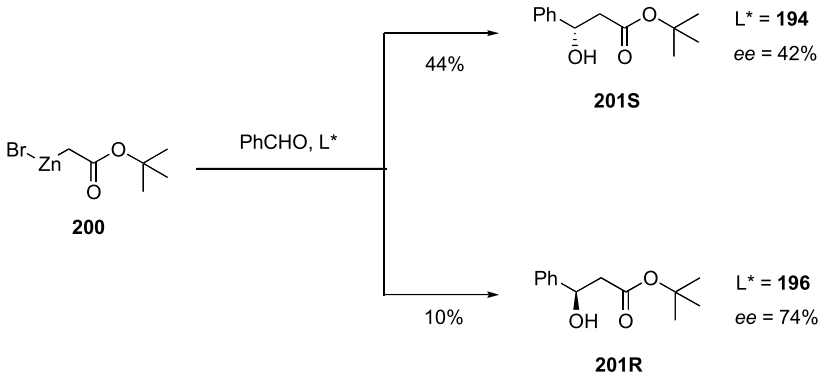
Figure 4



Scheme 49

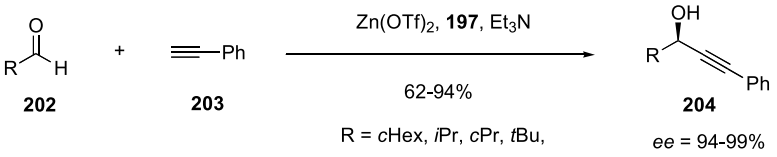
good yields. The enantiomeric excesses observed were not so remarkable in these cases. The major adducts **199S** possess an (*S*)-configuration. On the other hand, the analogous addition reactions using **195** produced the enantiomeric product **199R** as a major product, although neither the yield nor the enantiomeric excess was necessarily useful. The precise mechanistic consideration of the transition-states for these amino sugar-based diethylzinc additions to aldehydes was proposed by the Davis group.

The Reformatsky reaction is one of the promising surrogates of the aldol reaction. The Davis group developed the sugar-based asymmetric Reformatsky reaction for the synthesis of β-hydroxy esters, as shown in *Scheme 50* [181]. Thus, treatment of the *tert*-butyl bromoacetate-derived Reformatsky reagent **200** with benzaldehyde in the presence of **194** provided the desired alcohol **201S** with an (*S*)-configuration in a 42% enantiomeric excess. A similar Reformatsky reaction in the presence of **196** produced the enantiomeric (*R*)-isomer **201R** in improved enantioselectivity. Consequently, the D-glucosamine-type ligand **194** provided the (*S*)-aldol **201S** with a low level of enantioselectivity. On the other hand, the D-allosamine-type ligand **196** provided the (*R*)-aldol in 74% ee but in 10% yield.



Scheme 50

The asymmetric alkynylation of aldehydes is a current highlighted organic synthesis. The Davis group has explored the stereoselective alkynylation of aldehydes using transition-metal complexes incorporating amino sugar-based ligands [182]. As shown in *Scheme 51*, the alkynylation of phenylacetylene **203** with four types of aliphatic aldehydes in the presence of sugar-based ligands was investigated. The most efficient ligand was found to be **197**, a methyl β-D-glucosaminide-type ligand. Using this ligand, the $Zn(OTf)_2$-catalyzed asymmetric alkynylation of aldehydes **202** and phenylacetylene **203** produced propargylic alcohols **204** with excellent enantioselectivity.



Scheme 51

4 Concluding Remarks

As described above, a number of novel sugar-derived compounds as remarkable chiral scaffolds for asymmetric organic reactions have been developed in the past two decades. Various hexofuranose- and hexopyranose-derived templates have been devised as chiral auxiliaries or chiral ligands for transition-metal complexes. Using chiral scaffolds with potentially efficient sugar-derived templates, various asymmetric organic reactions have been realized. Representatives of the asymmetric synthesis achieved using sugar-based chiral auxiliaries are as follows: [3,3]-sigmatropic rearrangement, [2,3]-Wittig rearrangement, α-alkylation of esters, epoxidation of α,β-unsaturated γ-lactones, allylation of imines, 1,4-addition of α-lithiated sulfonates to nitroalkenes, tandem 1,4-addition/α-alkylation, 1,4-radical addition, Strecker reaction, Ugi reaction, Diels–Alder reaction, and 1,3-dipolar cycloaddition. Representatives of the asymmetric synthesis achieved using sugar-based chiral ligands for transition-metal catalysts are as follows: hydrogenation of dehydroamino acids or enamines, hydrocyanation, hydrovinylation, or hydroformylation of functionalized olefins, 1,4-addition of diethylzinc to cyclohexenone, hydrosilylation or cyanosilylation of ketones, allylic alkylation, Strecker reaction of imines, and Reformatsky reaction.

From the current scientific viewpoint, sugars are considered to be prominent biological-functional molecules and have been widely studied in the biosciences. In many cases, the role of sugars as an interface among a variety of biomolecules in living cells has been identified. Most of the synthetic concepts introduced in this chapter rely on the use of monosaccharides to facilitate the recognition of small molecules, such as organometallic species or low-molecular organic substances. It is to be expected that the scientific value of sugar-based asymmetric synthesis will be appreciated through further search for more efficient templates which constitute more effective chiral environments for stereoselective organic reactions.

References

1. Noyori R (1994) Asymmetric Catalysis in Organic Synthesis. Wiley, New York
2. Ojima I (2000) (ed.) Catalytic Asymmetric Synthesis. Wiley, New York
3. Jacobsen EN, Pfaltz A, Yamamoto H (1999) (eds.) Comprehensive Asymmetric Catalysts. Springer, Berlin Heidelberg New York
4. Evans DA, Bartroli J, Shin TL (1981) J Am Chem Soc 103:2127
5. Evans DA, Ennis MD, Mathre DJ (1982) J Am Chem Soc 104:1737
6. Evans DA (1984) In: Morrison JD (ed.) Asymmetric Synthesis. Academic Press, New York, 3:1
7. Lutomski KA, Meyers AI (1984) In: Morrison JD (ed.) Asymmetric Synthesis. Academic Press, New York, 3:213
8. Enders D (1984) In: Morrison JD (ed.) Asymmetric Synthesis. Academic Press, New York, 3:275
9. Evans DA, Britton TC, Ellman JA, Dorow RL (1990) J Am Chem Soc 112:4011
10. Oppolzer W (1990) Pure Appl Chem 62:1241
11. Groaning MD, Meyers AI (2000) Tetrahedron 56:9643
12. Cintas P (1991) Tetrahedron 47:6079
13. Seyden-Penne J (1995) Chiral Auxiliaries and Ligands in Asymmetric Synthesis. Wiley, New York
14. Hultin PG, Earle MA, Sudharshan M (1997) Tetrahedron 53:14823
15. Rück-Braun K, Kunz H (1999) Chiral Auxiliaries in Cycloaddition. Wiley-VCH, Weinheim
16. Dwek RA (1996) Chem Rev 96:683
17. Vasella A (1977) Helv Chim Acta 60:426
18. Vasella A (1977) Helv Chim Acta 60:1273
19. Schumacher R, Reissing H-U (1996) Synlett 1121
20. Tamura O, Mita N, Kusaka N, Suzuki H, Sakamoto M (1997) Tetrahedron Lett 38:429
21. Yoshida T, Chika J, Takei H (1998) Tetrahedron Lett 39:4305
22. Anaya J, Gero SD, Grande M, Hernando JIM, Laso NM (1999) Bioorg Med Chem 7:837
23. Enholm EJ, Gallagher ME, Jiang S, Batson WA (2000) Org Lett 2:3355
24. Enholm EJ, Cottone JS, Allais F (2001) Org Lett 3:145
25. Vega-Pérez JM, Vega M, Blanco E, Iglesias-Guerra F (2001) Tetrahedron: Asymm 12:3189
26. Ross GF, Herdtweck E, Ugi I (2002) Tetrahedron 58:6127
27. Yu H, Ballard CE, Boyle PD, Wang B (2002) Tetrahedron 58:7663
28. Garner P, Anderson JT, Cox PB, Klippenstein SJ, Leslie R, Scardovi N (2002) J Org Chem 67:6195
29. Cicchi S, Marradi M, Corsi M, Faggi C, Goti A (2003) Eur J Org Chem 4152
30. Huang L-L, Xu M-H, Lin G-Q (2005) J Org Chem 70:529
31. Kim HJ, Shin E, Chang J, Kim Y, Park YS (2005) Tetrahedron Lett 46:4115
32. Zhou G, Zheng W, Wang D, Zhang P, Pan Y (2006) Helv Chim Acta 89:520
33. Eguchi T, Kakinuma K (1997) J Syn Org Chem Jpn 55:814
34. Kakinuma K, Koudate T, Li H-Y, Eguchi T (1991) Tetrahedron Lett 32:5801
35. Eguchi T, Koudate T, Kakinuma K (1993) Tetrahedron 49:4527
36. Kakinuma K, Iihama Y, Takagi I, Ozawa K, Yamauchi N, Imamura N, Esumi Y, Uramoto M (1992) Tetrahedron 48:3763
37. Kakinuma K, Terasawa H, Li H-Y, Miyazaki K, Oshima T (1993) Biosci Biotech Biochem 57:1916
38. Kishida M, Eguchi T, Kakinuma K (1996) Tetrahedron Lett 37:2061
39. Maeda Y, Tago K, Eguchi T, Kakinuma K (1996) Biosci Biotech Biochem 60:1248
40. Kishida M, Yamauchi N, Sawada K, Ohashi Y, Eguchi T, Kakinuma K (1997) J Chem Soc Perkin Trans 1 891
41. Yamauchi N, Kishida M, Sawada K, Ohashi Y, Eguchi T, Kakinuma K (1998) Chem Lett 475
42. Kunz H, Sanger W (1987) Angew Chem Int Ed Engl 26:557
43. Kunz H, Pfrengle W (1988) J Am Chem Soc 110:651
44. Kunz H, Sager W, Pfrengle W, Schanzenbach D (1988) Tetrahedron Lett 29:4397
45. Kunz H, Pfrengle W (1988) Tetrahedron 44:5487
46. Kunz H, Mohr J (1988) J Chem Soc Chem Commun 1315
47. Kunz H, Pfrengle W (1989) Angew Chem Int Ed Engl 28:1067
48. Kunz H, Schanzenbach D (1989) Angew Chem Int Ed Engl 28:1068

49. Kunz H, Pfrengle W, Sager W (1989) Tetrahedron Lett 30:4109
50. Pfrengle W, Kunz H (1989) J Org Chem 54:4261
51. Kunz H, Rück K (1993) Angew Chem Int Ed Engl 32:336
52. Kunz H (1995) Pure Appl Chem 67:1627
53. Laschat S, Kunz H (1990) Synlett 51
54. Laschat S, Kunz H (1990) Synlett 629
55. Laschat S, Kunz H (1991) J Org Chem 56:5883
56. Stähle W, Kunz H (1991) Synlett 260
57. Kunz H, Sager W, Schanzenbach D, Decker M (1991) Liebigs Ann Chem 649
58. Kunz H, Pfrengle W, Rück K, Sanger W (1991) Synthesis 1039
59. Oertel K, Zech G, Kunz H (2000) Angew Chem Int Ed 39:1431
60. Laschat S, Kunz H (1992) Synthesis 90
61. Rück K, Kunz H (1992) Synlett 343
62. Rück K, Kunz H (1993) Synthesis 1018
63. Rück-Braun K, Stamm A, Engel S, Kunz H (1997) J Org Chem 62:967
64. Follmann M, Kunz H (1998) Synlett 989
65. Weymann M, Schultz-Kulara M, Kunz H (1998) Tetrahedron Lett 39:7835
66. Weyman M, Pfrengle W, Schollmeyer D, Kunz H (1997) Synthesis 1151
67. Weyman M, Schultz-Kukula M, Knauer S, Kunz H (2002) Monatsh Chem 133:57
68. Kranke B, Hebrault D, Schultz-Kukula M, Kunz H (2004) Synlett 671
69. Knauer S, Kunz H (2005) Tetrahedron: Asymm 529
70. Zech G, Kunz H (2003) Angew Chem Int Ed 42:787
71. Zech G, Kunz H (2004) Chem Eur J 10:4136
72. Klegraf E, Knauer S, Kunz H (2006) Angew Chem Int Ed 45:2623
73. Ganz I, Kunz H (1994) Synthesis 1353
74. Enders D, Heider K-J, Raabe G (1993) Angew Chem Int Ed Engl 32:598
75. Enders D, Papadopoulos K, Herdtweck E (1993) Tetrahedron 49:1821
76. Enders D, Klatt M (1996) Synthesis 1403
77. Enders D, Berner OM, Vignola N, Bats JW (2001) Chem Commun 2498
78. Enders D, Berner OM, Vignola N, Harnying W (2002) Synthesis 1945
79. Enders D, Vignola N, Berner OM, Bats JW (2002) Angew Chem Int Ed 41:109
80. Enders D, Vignola N, Berner OM, Harnying W (2005) Tetrahedron 61:3231
81. Enders D, Harnying W, Vignola N (2002) Synlett 1727
82. Enders D, Harnying W, Vignola N (2003) Eur J Org Chem 3939
83. Enders D, Harnying W, Raabe G (2004) Synthesis 590
84. Enders D, Harnying W (2004) Synthesis 2910
85. Enders D, Adelbrecht J-C, Harnying W (2005) Synthesis 2962
86. Totani, Takao K, Tadano K (2004) Synlett 2066
87. Totani K, Nagatsuka T, Takao K, Ohba S, Tadano K (1999) Org Lett 1:1447
88. Totani K, Nagatsuka T, Yamaguchi S, Takao K, Ohba S, Tadano K (2001) J Org Chem 66:5965
89. Totani K, Asano S, Takao K, Tadano K (2001) Synlett 1772
90. Munakata R, Totani K, Takao K, Tadano K (2000) Synlett 979
91. Asano S, Tamai T, Totani K, Takao K, Tadano K (2003) Synlett 2252
92. Sasaki D, Sawamoto D, Takao K, Tadano K, Okue M, Ajito K (2007) Heterocycles 72:103
93. Nagatsuka T, Yamaguchi S, Totani K, Takao K, Tadano K (2001) Synlett 481
94. Nagatsuka T, Yamaguchi S, Totani K, Takao K, Tadano K (2001) J Carbohydr Chem 20:519
95. Tamai, T, Asano S, Totani K, Takao K, Tadano K (2003) Synlett 1865
96. Cullen WR, Sugi Y (1978) Tetrahedron Lett 1635
97. Jackson R, Thompson DJJ (1978) J Organomet Chem 159:C29
98. Seike R (1979) Kinet Catak Lett 10:135
99. Sinou D, Descotes G (1980) Kinet Catal Lett 14:463
100. Inoguchi K, Sakuraba S, Achiwa K (1992) Synlett 169
101. Morimoto T, Chiba M, Achiwa K (1992) Chem Pharm Bull 40:2894
102. Kumar A, Oehme G, Roque JP, Schwarze M, Selke R (1994) Angew Chem Int Ed Engl 33:2197
103. Buisman GJH, Martin ME, Vos EJ, Klootwijk A, Kamer PCJ, van Leeuwen PWNM (1995) Tetrahedron: Asymm 6:719
104. Selke R, Holz J, Riepe A, Börner A (1998) Chem Eur J 4:769
105. Gläser B, Kunz H (1998) Synlett 53
106. Kadyrov R, Heller D, Selke R (1998) Tetrahedron: Asymm 9:329
107. Yonehara K, Hashizume T, Ohe K, Uemura S (1998) Bull Chem Soc Jpn 71:1967
108. Yonehara K, Hashizume T, Mori K, Ohe K, Uemura S (1999) J Org Chem 64:5593

109. Yonehara K, Ohe K, Uemura S (1999) J Org Chem 64:9381
110. Suarez A, Pizzano A, Fernandez I, Khiar N (2001) Tetrahedron: Asymm 12:633
111. Borriello C, Cucciolito ME, Panunzi A, Ruffo F, Saporito A (2003) Inorg Chem Commun 6:1081
112. Reetz MT, Goossen, LJ, Meiswinkel A, Paetzold J, Jensen JF (2003) Org Lett 5:3099
113. Borriello C, Litto RD, Panunzi A, Ruffo F (2004) Tetrahedron: Asymm 15:681
114. Khiar N, Suarez B, Valdivia V, Fernandez I (2005) Synlett 2963
115. RajanBabu TV, Casalnuovo AL (1992) J Am Chem Soc 114:6265
116. RajanBabu TV, Casalnuovo AL (1994) Pure Appl Chem 66:1535
117. RajanBabu TV, Ayers TA, Casalnuovo AL (1994) J Am Chem Soc 116:4101
118. RajanBabu TV, Ayers TA (1994) Tetrahedron Lett 35:4295
119. Casalnuovo AL, RajanBabu TV, Ayers TA, Warren TH (1994) J Am Chem Soc 116:9869
120. RajanBabu TV, Casalnuovo AL (1996) J Am Chem Soc 118:6325
121. RajanBabu TV, Ayers TA, Halliday GA, You KK, Calabrese JC (1997) J Org Chem 62:6012
122. Shin S, RajanBabu TV (1999) Org Lett 1:1229
123. RajanBabu TV, Radetich B, You KK, Ayers TA, Casalnuovo AL, Calabrese JC (1999) J Org Chem 64:3429
124. Clyne DS, Mermet-Bouvier YC, Nomura N, RajanBabu TV (1999) J Org Chem 64:7601
125. Rajanbabu TV, Yan, Y-Y, Shin S (2001) J Am Chem Soc 123:10207
126. Park H, RajanBabu TV (2002) J Am Chem Soc 124:734
127. RajanBabu TV (2003) Chem Rev 103:2845
128. Park H, Kumareswaran R, RajanBabu TV (2005) Tetrahedron 61:6352
129. Saha B, RajanBabu TV (2006) Org Lett 8:4657
130. Pàmies O, Net G, Ruiz A, Claver C (1999) Tetrahedron: Asymm 10:2007
131. Pàmies O, Net G, Ruiz A, Claver C, Woodward S (2000) Tetrahedron: Asymm 11:871
132. Pàmies O, Net G, Ruiz A, Claver C (2000) Tetrahedron: Asymm 11:1097
133. Diéguez M, Pàmies O, Ruiz A, Castillón S, Claver C (2000) Chem Commun1607
134. Pàmies O, Diéguez M, Net G, Ruiz A, Claver C (2000) Chem Commun 2383
135. Pàmies O, Diéguez M, Net G, Ruiz A, Claver C (2000) Tetrahedron: Asymm 11:4377
136. Diéguez, Jansat S, Gomez M, Ruiz A, Muller G, Claver C (2001) Chem Commun 1132
137. Pàmies O, Diéguez M, Net G, Ruiz A, Claver C (2001) J Org Chem 66:8364
138. Diéguez M, Pàmies O, Net G, Ruiz A, Claver C (2001) Tetrahedron: Asymm 12:651
139. Diéguez M, Ruiz A, Claver C (2001) Tetrahedron: Asymm 12:2895
140. Diéguez M, Pàmies O, Ruiz A, Castillón S, Claver C (2001) Chem Eur J 7:3086
141. Diéguez M, Pàmies O, Ruiz A, Claver C (2002) Tetrahedron: Asymm 13:83
142. Diéguez M, Pàmies O, Ruiz A, Claver C (2002) New J Chem 26:827
143. Diéguez M, Ruiz A, Claver C (2002) J Org Chem 67:3796
144. Diéguez M, Ruiz A, Claver C (2003) Dalton Trans 2957
145. Diéguez M, Pàmies O, Claver C (2004) Chem Rev 104:3189
146. Guimet E, Diéguez M, Ruiz A, Claver C (2004) Tetrahedron: Asymm 15:2247
147. Diéguez M, Pàmies O, Claver C (2005) Chem Commun 1221
148. Diéguez M, Pàmies O, Claver C (2005) Adv Synth Catal 347:1257
149. Mata Y, Diéguez M, Pàmies O, Claver C (2005) Adv Synth Catal 347:1943
150. Guimet E, Diéguez M, Ruiz A, Claver C (2005) Tetrahedron: Asymm 16:2161
151. Diéguez M, Pàmies O, Claver C (2005) Tetrahedron: Asymm 16:3877
152. Guimet E, Diéguez M, Ruiz A, Claver C (2005) Inorg Chim Acta 358:3824
153. Diéguez M, Pàmies O, Claver C (2006) J Organometal Chem 691:2257
154. Kanai M, Hamashima Y, Shibasaki M (2000) Tetrahedron Lett 41:2405
155. Hamashima Y, Kanai M, Shibasaki M (2000) J Am Chem Soc 122:7412
156. Manickam G, Nogami H, Kanai M, Gröger H, Shibasaki M (2001) Synlett 617
157. Hamashima Y, Kanai M, Shibasaki M (2001) Tetrahedron Lett 42:691
158. Yabu K, Masumoto S, Yamasaki S, Hamashima Y, Kanai M, Du W, Curran DP, Shibasaki M (2001) J Am Chem Soc 123:9908
159. Matsumoto S, Yabu K, Kanai M, Shibasaki M (2002) Tetrahedron Lett 43:2919
160. Yabu K, Masumoto S, Kanai M, Curran DP, Shibasaki M (2002) Tetrahedron Lett 43:2923
161. Masumoto S, Suzuki M, Kanai M, Shibasaki M (2002) Tetrahedron Lett 43:8647

162. Fujii K, Maki K, Kanai M, Shibasaki M (2003) Org Lett 5:733
163. Takamura M, Yabu K, Nishi T, Yanagisawa H, Kanai M, Shibasaki M (2003) Synlett 353
164. Yabu K, Matsumoto S, Kanai M, Du W, Curran DP, Shibasaki M (2003) Heterocycles 59:369
165. Masumoto S, Usuda H, Suzuki M, Kanai M, Shibasaki M (2003) J Am Chem Soc 125:5634
166. Kato N, Suzuki M, Kanai M, Shibasaki M (2004) Tetrahedron Lett 45:3147
167. Kato N, Suzuki M, Kanai M, Shibasaki M (2004) Tetrahedron Lett 45:3153
168. Masumoto S, Suzuki M, Kanai M, Shibasaki M (2004) Tetrahedron 60:10497
169. Kato N, Tomita, D, Maki K, Kanai, M, Shibasaki M (2004) J Org Chem 69:6128
170. Mita T, Sasaki K, Kanai M, Shibasaki M (2005) J Am Chem Soc 127:514
171. Kanai M, Kato N, Ichikawa E, Shibasaki M (2005) Synlett 1491
172. Fukuta Y, Mita T, Fukuda N, Kanai M, Shibasaki M (2006) J Am Chem Soc 128:6312
173. Kato N, Mita T, Kanai M, Therrien B, Kawano M, Yamaguchi K, Danjo H, Sei Y, Sato A, Furusho S, Shibasaki M (2006) J Am Chem Soc 128:6768
174. Huang H, Chen H, Hu X, Bai C, Zheng Z (2003) Tetrahedron: Asymm 14:297
175. Huang H, Zheng Z, Chen H, Bai C, Wang J (2003) Tetrahedron: Asymm 14:1285
176. Huang H, Zheng Z, Luo H, Bai C, Hu X, Chen H (2003) Org Lett 5:4137
177. Huang H, Zheng Z, Luo H, Bai C, Hu X, Chen H (2004) J Org Chem 69:2355
178. Huang H, Liu X, Chen S, Chen H, Zheng Z (2004) Tetrahedron: Asymm 15:2011
179. Huang H, Liu X, Chen H, Zheng Z (2005) Tetrahedron: Asymm 16:693
180. Emmerson DPG, Villard R, Mugnaini C, Bastsanov A, Howard JAK, Hems WP, Tooze RP, Davis BG (2003) Org Biomol Chem 1:3826
181. Emerson DPG, Hems WP, Davis BG (2005) Tetrahedron: Asymm 16:213
182. Emerson DPG, Hems WP, Davis BG (2006) Org Lett 8:207

4.7 Carbohydrate-Metal Complexes: Structural Chemistry of Stable Solution Species

Thorsten Allscher, *Peter Klüfers**, *Peter Mayer*
Department Chemie und Biochemie, Ludwig-Maximilians-Universität München, Butenandtstr. 9, 81377 München, Germany
kluef@cup.uni-muenchen.de

In: *Glycoscience*. Fraser-Reid B, Tatsuta K, Thiem J (eds)
Chapter-DOI 10-1007/978-3-540-30429-6_24:

Abstract

This review discusses the structural chemistry of metal complexes of carbohydrates and their derivatives with the focus on crystal structure and NMR data of stable solution species. There is evidence that the stability of these complexes is markedly increased when the carbohydrates operate as chelating polyolato ligands, this is, when they are multiply deprotonated polydentate ligands. The 1,2-diolato coordination mode resulting in five-membered chelate rings is most commonly observed. Small torsion angles within the 1,2-diolato moiety enable the chelation of small atoms while larger torsion angles are usually needed for the complexation of large atoms. Hence, as a consequence of the limited flexibility of the pyranose ring, pyranoidic 1,2-diols merely form less-stable complexes with small metal centers in marked contrast to furanoidic 1,2-diols. An additional contribution to the stability arises from hydrogen bonds, especially intramolecular ones, with the deprotonated ligator oxygen atoms acting as strong hydrogen-bond acceptors.

Keywords

Carbohydrate; Derivatives; Metal; Complex; Chelation; Solution structure; Crystal structure; NMR

Abbreviations

Ado	adenosine
AnEryt	anhydroerythritol
AnThre	anhydrothreitol
Arab	arabitol
Ara	arabinose
ax	axial
bpy	2,2′-bipyridyl
Bu	butyl
CD	cylodextrin
chxn	1,2-cyclohexanediamine
CIS	coordination-induced shift
Cp	cyclopentadienyl
Cyd	cytidine
DBU	1,8-diazabicyclo[5.4.0]undec-7-ene
dien	diethylenetriamine
dppp	1,3-bis(diphenylphosphino)propane
en	ethylenediamine
Eryt	erythritol
eq	equatorial
Et	ethyl

Fru fructose
Gal galactose
Gal1,6A$_2$ galactaric acid
Glc glucose
Glc1A gluconic acid
Glc1P glucose-1-phosphate
Glc1,6An 1,6-anhydro-β-D-glucose (levoglucosan)
Glc1,6A$_2$ glucaric acid
Guo guanosine
Ino inosine
Ins inositol
Lyx lyxose
Man mannose
Mann mannitol
Me methyl
Me$_3$tren tris(N,N',N''-trimethyl-2-aminoethyl)amine
py pyridine
Rib ribose
Ribt ribitol
Rul ribulose
Sorb sorbitol
Suc sucrose
tach 1,3,5-triaminocyclohexane
Tre α,α-trehalose
Thre threitol
tpb hydrido-tris(pyrazolyl)borate
tren tris(2-aminoethyl)amine
Urd uridine
Xyl xylose
Xylt xylitol
f furanose
p pyranose

1 Introduction

A central problem in the chemistry of carbohydrates is their polyfunctionality which usually leads to several possible coordination sites for metal atoms. As a consequence, in basic research the carbohydrates are often replaced by model ligands of reduced functionality in order to investigate the structure of carbohydrate–metal complexes. The most reliable structural information about solution and solid-state species is obtained by NMR spectroscopy and X-ray structure analysis of single crystals, respectively, both methods having developed significantly in the last years. Hence, we will focus on the structures of carbohydrate–metal complexes based on unambiguous data obtained by these two methods. The following carbohydrates and carbohydrate derivatives are included in this review: sugar alcohols, aldonic and

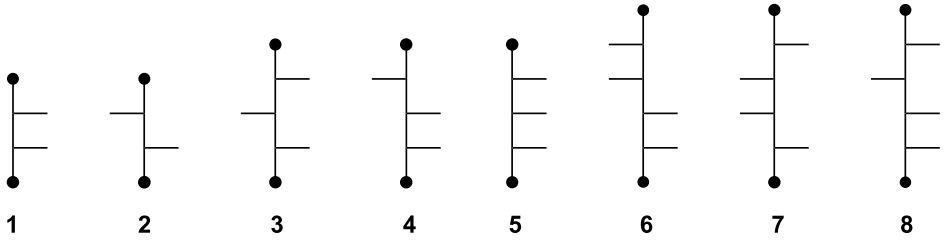
aldaric acids, anhydroerythritol, inositols, anhydro-sugars, glycosides, and reducing carbohydrates such as aldoses, ketoses, and glycuronates. Among the glycosides, the pyranosides and furanosides (mostly alkyl derivatives), nucleosides, non-reducing disaccharides, cyclodextrins, and polysaccharides are viewed. The nucleosides are the only non-*O*-glycosides considered in this review. Structures of simple addition products of carbohydrates and metal salts are not included either. Structures which should be included in this review, according to the requirements stated above, but have been reviewed elsewhere already, will mostly not be discussed again. Publications written in Chinese, Japanese, or Russian are not cited herein.

A summary of the historical development of reviews regarding carbohydrate–metal complexes can be found in a review written by Verchère et al. in 1998. Besides NMR and crystal structure data, they also consider other methods such as calorimetry and EXAFS spectroscopy [1]. Here a short listing of a selection of reviews written in the ensuing years: Complexes of the metals nickel(II), cobalt(III), and manganese(II,III) with *N*-glycosides derived from tris(2-aminoethyl)amine and aldoses have been reviewed by Yano [2] who has also written a summary on the structure and biological activity of complexes of transition metals with glycosylamines derived from sugar molecules and polyamines [3]. Nagy et al. discuss the coordination equilibria of carbohydrate–metal complexes in aqueous solution together with crystal structures [4,5] and summarize the application of EXAFS and XANES methods in the coordination chemistry of carbohydrates [6]. Solid-state and solution complexes with platinum-group metals are reviewed by Steinborn and Junicke [7]. The coordination chemistry of modified carbohydrates has been summarized by Alexeev et al. in two reviews [8,9]. The kinetics of proton-release reactions of aluminum(III) complexes with D-ribose and their stability has been reviewed by Petrou [10]. Chakravorty et al. describe heteroleptic vanadate chelate esters of monoionized diols and carbohydrates with tridentate auxiliary ligands [11]. The formation and the synthetic applications of carbohydrate-iron complexes possessing Fe–C bonds are discussed by Zamojski [12]. The metallation sites of nucleosides and nucleotides in ternary systems with polyamines has been summarized by Lomozik [13]. The metal complexation of oligo- and polysaccharides has also been reviewed. Metal-containing supramolecular catalytic systems may be achieved based on cyclodextrins and their derivatives (“cavitand” ligands) [14,15,16]. The metal-binding ability of chitosan under various conditions is discussed by Varma et al. [17].

In order to elucidate the coordination chemistry of individual carbohydrates, the following review will be carbohydrate-oriented instead of metal-oriented, that is, the structures of individual carbohydrates or, if only few structures are known, of a group of related carbohydrates are discussed en-block. We will start with complexes of simpler carbohydrates such as sugar alcohols or anhydroerythritol in order to derive the basics of metal coordination and will then proceed to typical carbohydrates.

2 Sugar Alcohols

The solution and solid state structures of the following sugar alcohols are discussed in separate sections: erythritol **1**, D-threitol **2**, xylitol **3**, D-arabitol **4**, ribitol **5**, D-mannitol **6**, dulcitol **7**, and D-sorbitol **8**.



Carbon numbering: C1 is situated at the top of each chain. ● = CH_2OH
Chiral sugar alcohols are depicted in their D-configuration.

2.1 Erythritol

Erythritol **1** is a tetraol with a central *erythro*-configured diol group. In combination with the terminal hydroxyl functions there are several conformations to act as a chelate ligand. In most of its characterized solid-state structures, erythritol is coordinated to two central atoms by acting as a bis-1,2-diolato ligand, at which the deprotonation of the hydroxyl groups is a consequence of the synthesis from alkaline solutions and/or complexation of more or less Lewis acidic central atoms. A complex of this type is obtained upon the reaction of erythritol with two equivalents of $[(NH_3)_2Pd(OH)_2]$ in aqueous solution which leads to the centrosymmetric structure of **9** (❯ *Fig. 1*) [18].

In **9**, the C_4 chain of erythritol adopts a zigzag conformation with a torsion angle of 180° which is as much forced by the inversion symmetry of the complex as is the torsion angle of the central *erythro*-diolato group (180°). The terminal oxygen atoms are oriented to the same side of the C_4 chain as their adjacent *erythro*-oxygen atoms. The torsion angle of the resulting chelating diolato group is 46.2°. The solutions mentioned above contain **9** as the main species which is confirmed by a downfield shift ("coordination induced shift", CIS) of about 12 ppm of the carbon atoms involved in the five-membered chelate ring. The asymmetric monometallated complex **11** which might be stabilized by an intramolecular γ hydrogen bond (the donor and acceptor oxygen atoms are separated by a C_3 chain), is detected as a byproduct whereas **11** becomes the main species and **9** a byproduct in equimolar solutions of erythritol

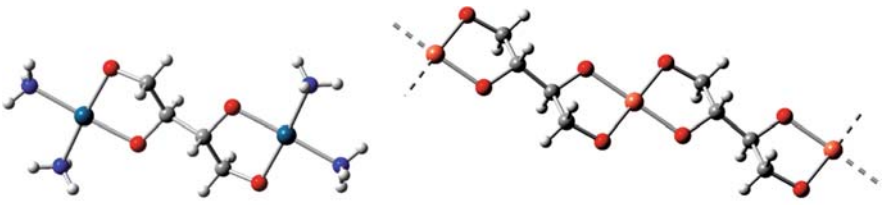
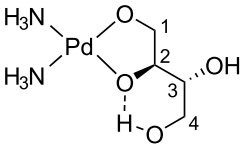
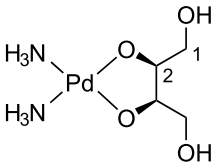


Figure 1
Molecular structure of $[\{(NH_3)_2Pd\}_2(ErytH_{-4})]$ 9 in crystals of $9 \cdot 2\ H_2O$ and a section of the linear coordination polymer built up by $[Cu(ErytH_{-4})]^{2-}$ 10 units in crystals of $Na_2$10 · Eryt · $12\ H_2O$

and $[(NH_3)_2Pd(OH)_2]$ (CIS of the chelating carbon atoms 1 and 2 in **11**: 10.9 and 9.8 ppm, respectively, non-chelating carbon atoms 3 and 4: 0.9 and 1.0 ppm, respectively). A minor species of these equimolar solutions is the symmetric monometallated complex **12** (C1 CIS: −0.9 ppm, C2 CIS: 10.0 ppm).



11



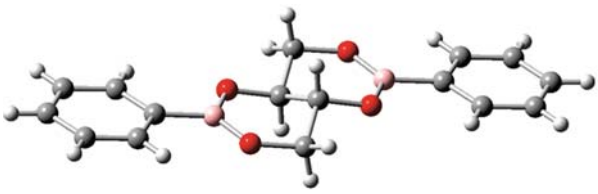
12

A similar species distribution is found for $[(en)Pd(OH)_2]$/erythritol solutions, and a crystal structure containing a binuclear complex analogous to **9** has been determined [19]. A solid-state structure of this type is also formed by ethylenediamine-copper(II) [20]. In strong alkaline aqueous solutions, the ethylenediamine-copper bond is cleaved and the formation of homoleptic complexes such as the linear coordination polymer **10** (◉ *Fig. 1*) is enabled [21].

The reaction of $(dppp)Pt(CO_3)$ with four equivalents of erythritol produces the (dppp)Pt derivative of **11** as the major species, while the corresponding *erythro*-derivative **12** is obtained in small yield only (11%) [22].

While palladium(II) and copper(II) build solely 1,2-diolato complexes with erythritol which leads to five-membered chelate rings, boron(III) is involved in the six-membered rings as in the bis(phenylboronic acid ester) **13** (◉ *Fig. 2*). Erythritol adopts a zigzag conformation similar to that in **9** but the oxygen atoms are grouped alternatively on opposite sites of the C_4-chain. The six-membered chelate rings in half-chair conformation provide a suitable bond pattern for a boron center involved in the delocalized π-bond system of the phenyl substituent.

In solution, **13** is merely a minor species which is identified by a significant difference in the chemical shifts of its two symmetrically independent erythritol-C atoms: the terminal carbon atom is shifted by 2.8 ppm downfield while the inner carbon atom which is involved in two chelate rings is shifted by 2.8 ppm upfield. Most of the phenylboronic acid is bound in five-membered chelate rings similar to **9** (^{13}C shifts: 5.3 ppm for terminal C atoms, 6.5 ppm for inner C atoms) [23].



13

14

◘ Figure 2
The crystal structure of $(PhB)_2(ErythH_{-4}$-$O^{1,3},O^{2,4})$ 13 and the molecular structure of $[Mo_2O_5(ErythH_{-4})]^{2-}$ 14 in crystals of $[N(n\text{-}Bu)_4]_2$14

In order to group all its hydroxyl functions on the same side of the C_4 chain, erythritol has to adopt a sickle conformation which is established as in the dimolybdate complex **14** (*Fig. 2*). The torsion angle of the C_4 chain is 49.5°, the torsion angle of the central *erythro*-configured diol function is 34.4°. The sickle conformation allows the asymmetric facial coordination of two molybdenum centers with one terminal and one inner alkoxido function in a bridging mode. A dimolybdate complex with only triply deprotonated erythritol has been crystallized from more acidic solutions. The structure largely resembles the one in **14**, but the remaining proton at the bridging terminal hydroxyl function causes markedly increased Mo–O distances of 2.50 and 2.53 Å (in **14**: 2.27 and 2.29 Å [24]. NMR experiments reveal that the dimolybdate structure is retained in solution. All carbon atoms show a more or less strong downfield shift ($\Delta\delta$ between 6.6 and 18.5 ppm) [25,26]. Similar shifts are obtained for a ditungstate complex ($\Delta\delta$ between 6.1 and 18.1 ppm) whose asymmetry could be additionally characterized by the existence of two different ^{183}W signals [27]. The sickle conformation is also adopted by doubly and triply deprotonated erythritol in a Fe_{14} oxocluster which is stabilized by calcium(II) ions (torsion angles: 49.0°, 47.2° (C_4 chain), 44.2°, 48.7° (*erythro*-diol)) [28].

2.2 Threitol

Threitol **2** is a tetraol with a central *threo*-configured diol group. While erythritol has to adopt a sickle conformation in order to locate all its oxygen atoms on the same side of the C_4 chain, threitol attains this in its zigzag conformation.

In **15** and **16** (*Fig. 3*), the zigzag conformation of L-threitol comes along with an approximately parallel alignment of the planes of carbon and oxygen atoms, respectively. The torsion angle of the C_4 chain is about 170° and the oxygen atoms occupy the corners of a rhomb.

Fourfoldly deprotonated L-threitol acts as a bis-1,3-diolato ligand in the V-shaped bis(phenylboronic acid ester) **15** (*Fig. 3*), while the *erythro*-configured erythritol leads to the planar ester **13** (*Fig. 2*). The significant CIS of the carbon atoms in **13** was also detected in solutions of **15** (2.1 ppm (terminal C), −5.9 (inner C)), however, 1,2-diolato species could not

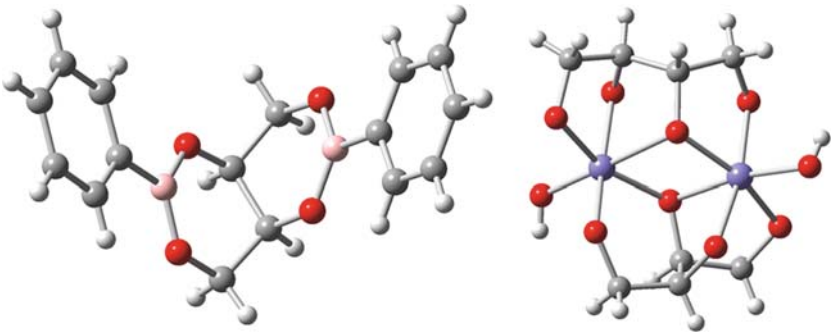
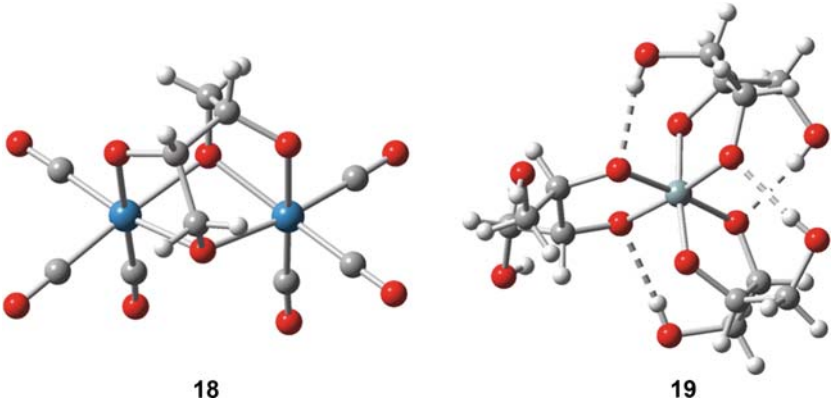


Figure 3
The crystal structure of $(PhB)_2$(L-ThreH$_{-4}$-$O^{1,3},O^{2,4}$) 15 and the molecular structure of $[Fe_2$(L-ThreH$_{-4})_2(OH)_2]^{4-}$ 16 in crystals of $Ba_2$16 · 12.5 H_2O

be ascertained [23]. A structure of the type of **15** was also obtained with the bis(3,4-dibromomethyl) derivative of phenylboronic acid [29]. Five-membered chelate rings are described in the homoleptic borate ester **17** with doubly deprotonated D- and L-threitol in which the deprotonation occurs at the chelating central *threo*-configured diol group. Depending on the involved enantiomeric pairs of threitol (**17** presents the (*R*,*R*)-(*R*,*R*) form), three different complexes are formed. The CIS of the *threo*-C atoms are reported with 3.3 and 3.4 ppm, the terminal C atoms show a smaller CIS of 0.8 and 1.5 ppm [30].

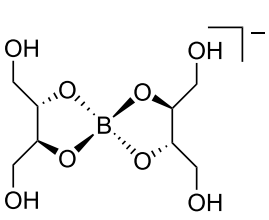
In the sandwich-type complex anions **16** (● *Fig. 3*), eight out of ten oxygen atoms for the formation of an edge-sharing $(FeO_5)_2$-dioctahedron are provided from two fourfoldly deprotonated L-threitol ligands which show manifold complexing abilities. One of the *threo*-configured oxygen atoms of the two ligands is in the μ_2 mode, the others bind to one iron(III) center only. Both of the threitolato ligands act in the *fac*-1,2,3-triolato mode; additionally one acts as a 1,2-diolato and one as a 1,3-diolato ligand. The O–O distances in the chelate rings vary between 2.65 Å (five-membered ring) and 2.95 Å (six-membered ring). The O–C–C–O torsion angles are between 35° and 50° [31].

A more bent zigzag chain with a C_4 torsion angle of about 104° is present in the dinuclear C_2-symmetric rhenium(I) complex anion **18** (● *Fig. 4*). The terminal oxygen atoms of triply deprotonated L-threitol are in a bridging mode whereas the terminal oxygen atoms bind to one center only. A very short intermolecular distance between the latter non-bridging oxygen atoms of just 2.36 Å indicates a strong hydrogen bond between them and thus only triply deprotonated L-threitol is present. The torsion angle of the central *threo*-configured diolato group is 133.4°, the terminal diolato groups form torsion angles of about 56°. This complex has also been found in aqueous solution as well as in acetonitrile at a Re:L-Thre ratio of 2. The latter solution contains a minor trinuclear species which becomes the main species upon the addition of methanol and a Re:L-Thre ratio of 3. Deduced from a trinuclear complex with triply deprotonated glycerol, triply deprotonated L-threitol is expected to be in a μ_3 mode [32].

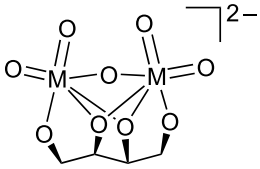


■ Figure 4
The molecular structures of $[Re_2(CO)_6(L\text{-}ThreH_{-3})]^-$ 18 in crystals of $NEt_4$18 · MeCN and $[\Lambda\text{-}Si(L\text{-}Thre2,3H_{-2})_3]^{2-}$ 19 in crystals of $Rb_2$19 · 3 H_2O. *Dashed lines*: hydrogen bonds

With rubidium as the counterion, a mononuclear, approximately C_2-symmetric homoleptic silicate complex anion **19** (*Fig. 4*) is formed by three L-threitolato ligands which are deprotonated at their central *threo*-configured chelating hydroxyl functions. The three polyols adopt two different conformations. Two ligands use their non-deprotonated hydroxyl functions to act as donors in intramolecular hydrogen bonds with deprotonated hydroxyl functions as acceptors. The acceptor functions of the two remaining deprotonated hydroxyl groups are used in hydrogen bonds to two water molecules. The third threitolato ligand adopts a conformation with a torsion angle of the C_4 chain of 95.1° which is significantly smaller than those observed for the other two ligands (144.6° and 136.2°, respectively). With cesium as counterion, a D_3-symmetric threitolato silicate anion $[\Lambda\text{-Si(D-Thre2,3H}_{-2})_3]^{2-}$ is obtained with all terminal hydroxyl groups being involved in intramolecular hydrogen bonds of the type in **19**. The torsion angles of the chelating *threo*-configured diolato group is in the range of 17–24° causing a bite of about 2.5 Å [33].

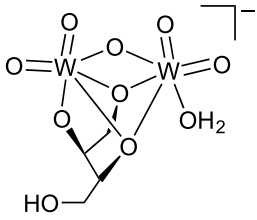


17



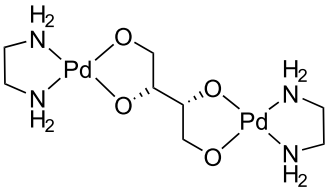
20

M = Mo,W

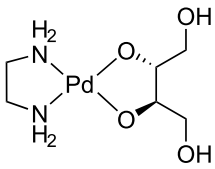


21

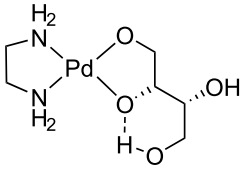
D- and L-threitol have been shown to form the C_2-symmetric dimolybdate **20** in which each molybdenum is facially coordinated by a 1,2,3-triolato group. The CIS of the *threo*-configured carbon atoms is reported with 9.8 ppm, the terminal carbon atoms have a CIS of 12.2 ppm [25,27]. This type of complexation is confirmed in a crystal structure with L-dithiothreitol as ligand [34]. The type of ditungstate complex formed with D- and L-threitol proves to be pH-dependent. At pH 7, the C_2-symmetric ditungstate **20** is formed (CIS *threo*-C: 9.2 ppm, terminal C: 10.5 ppm), while at pH 8–9 the asymmetric ditungstate **21** dominates. One terminal hydroxyl function is not involved in the complexation (CIS 1.8 ppm) while the other one now is in a bridging mode (CIS 9.5 ppm) together with one of the *threo*-configured hydroxyls (CIS 9.4 ppm). The second *threo*-configured hydroxyl is bound to one tungsten solely (CIS 13.1 ppm) [27,35].



22



23



24

$[Pd(en)(OH)_2]$ forms the solution species **22**–**24** with D-threitol with the species distribution strongly depending on the Pd:D-threitol molar ratio. In 1:1 solutions **23** is the main

species besides a minor fraction of **24**. Obviously the formation of chelate rings with a *threo*-configured diol group is more favored above an *erythro*-diol chelate formation such as in **12**, which was merely a minor species in the corresponding erythritol solutions. In 3:1 solutions the dimetallated complex **22** and the monometallated complex **23** coexist as main species [19]. Two monometallated products in the molar ratio of 3:2 are reported for the reaction of (dppp)Pt(CO_3) with four equivalents of *rac*-threitol. The major part corresponds to the (dppp)Pt-derivative of the*threo*-complex **23**, while in the minor part the terminal diol is coordinated similarly to **24** [22].

2.3 Xylitol

Xylitol **3** is a non-chiral, all-*threo*-configured pentitol which allows grouping of all its hydroxyl functions on the same side of the C_5 chain with a zigzag conformation. In the dinuclear cobalt(III) complex **25** (*Fig. 5*) two entirely deprotonated zigzag conformers provide a proper O_{10} set of an edge-shared Co_2O_{10} dioctahedron, whereas in the similar diferrate(III) **16** (*Fig. 3*) with the smaller all-*threo* configured L-threitol, two hydroxyls are required to complete the dioctahedron [36].

Fourfoldly deprotonated xylitol acts as a bis-1,2-diolato ligand in the dinuclear palladium(II) complex **26** (*Fig. 5*). The remaining terminal hydroxyl function is employed in a short intramolecular δ hydrogen bond. The enhanced stability thereby makes **26** the dominating species in [Pd(en)$(OH)_2$]/xylitol solutions of the molar ratio 3:1 and even 2:1, whereas in the corresponding 3:1-solution with D-threitol the dimetallated species and the monometallated species **12** are formed in approximately equal amounts. The main species in equimolar solutions of [Pd(en)$(OH)_2$] and xylitol is the monometallated, *threo*-chelate complex [Pd(en)(Xylt2,3H_{-2})] which is the hydroxymethyl homolog of the predominant [Pd(en)(Thre2,3H_{-2})] complex. The isomeric complex with a terminal chelate formation, [Pd(en)(Xylt1,2H_{-2})], is formed in small amounts only [19].

From solutions of (dppp)Pt(CO_3) and xylitol in a molar ratio of 1:4 two monometallated species are obtained. The formation of the 2,3-*threo* bound isomer (86%) is significantly preferred over the 1,2-terminal bound isomer (14%) [22].

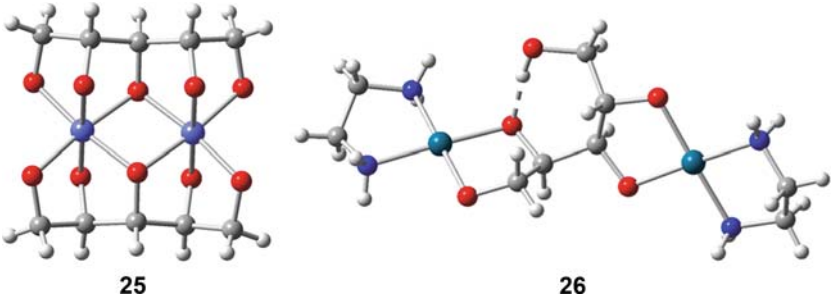


Figure 5
The molecular structures of $[Co_2(XyltH_{-5})_2]^{4-}$ 25 in crystals of $Li_5$25 · 8 H_2O and $[Pd_2(en)_2(Xylt1,2;3,4H_{-4})]$ 26 in crystals of 26 · 4 H_2O. *Dashed line*: hydrogen bond

Three doubly deprotonated xylitol molecules act as 1,2-diolato ligands in the hexacoordinate silicate $[Si(Xylt2,3H_{-2})_3]^{2-}$ **27** (Fig. 6) which is the hydroxymethyl derivative of the threitolato complex **19** (Fig. 4) and, as a consequence, shows the same stabilizing intramolecular hydrogen bonds [33].

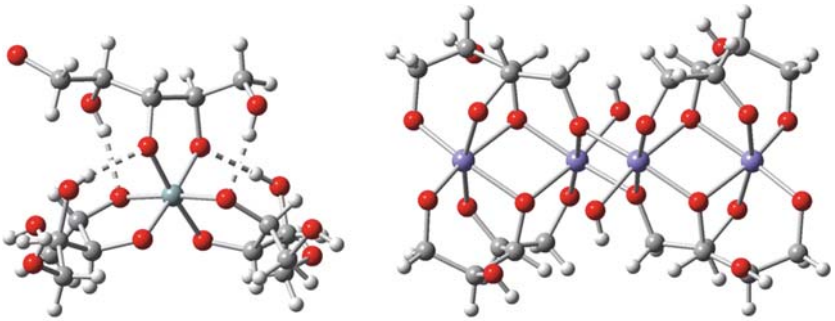
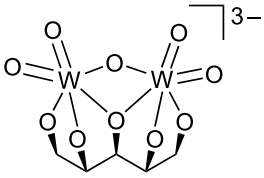


Figure 6
The molecular structures of $[Si(Xylt2,3H_{-2})_3]^{2-}$ 27 in crystals of Cs_2 27 · 2 H_2O and $[Fe_4(rac\text{-}Arab1,2,3,5H_{-4})_4(OH)_2]^{6-}$ 29 in crystals of $Sr_4$29(CO_3) · 33 H_2O. *Dashed lines*: hydrogen bonds

A hydroxymethyl derivative of the threitolato dimetallates **20** is also formed with the higher homolog xylitol, whereas the ditungstate proves to be stable at pH 7. At pH 8–9 xylitol acts as chelating ligand using its secondary hydroxyl functions similarly to **21**, while the primary hydroxyl functions are not employed for metal coordination.



28

In the pH range of 9–12 the ditungstate **28** is formed with an entirely deprotonated xylitol exhibiting the same coordination pattern as in **25** (Fig. 5): both metal centers are facially coordinated with the central alkoxido function in the μ_2 mode [25,26,27,35].

2.4 Arabitol

Arabitol **4** is a chiral pentitol bearing a *threo*- as well as an *erythro*-configured diol group. The centrosymmetric, tetranuclear ferrate(III) **29** (Fig. 6) consisting of four edge-sharing FeO_6-octahedrons is formed by four fourfoldly deprotonated *rac*-arabitols in a conformation which enables coordination to the iron(III) centers with the terminal and the *threo*-configured oxygen atoms. The coordination sphere of the inner two iron centers is completed by a hydroxide lig-

and each—a structural feature adopted by the closely related threitolato complex **16** (◉ *Fig. 3*) as well [31].

A comparison of the structures **29** (◉ *Fig. 6*) and the palladium(II) complex **30** (◉ *Fig. 7*) clarifies the flexibility of arabitol to adopt a conformation which satisfies the structural requirements of the metal centers: in **29**, arabitol acts as a 1,3-diolato ligand and as a 1,2,4 triolato ligand with C3 in a μ_2 mode to achieve an octahedral coordination geometry, while D-arabitol acts as a bis-1,2-diolato ligand in **30**. The diolato moieties are *erythro*-linked, hence **30** could be considered a hydroxymethyl-Pd(en) derivative of **9** (◉ *Fig. 1*), with the hydroxyl forming a stabilizing intramolecular γ hydrogen bond. This species is also predominant in [Pd(en)$(OH)_2$]/D-arabitol solutions with excess [Pd(en)$(OH)_2$] and still is a minor species in equimolar solutions. Those contain the *threo*-complex [Pd(en)(D-Arab2,3H_{-2})] as the only monometallated species [19].

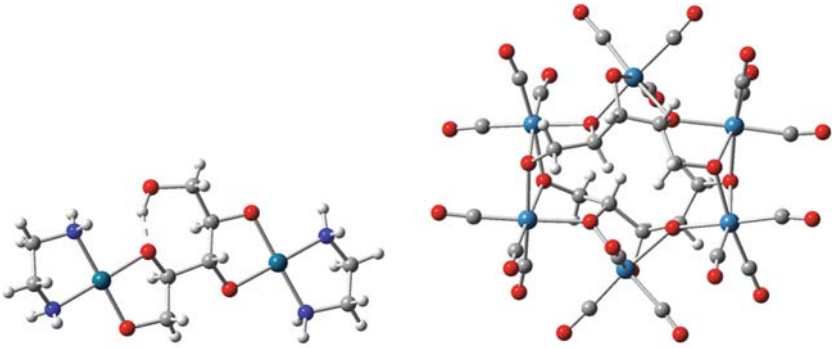


Figure 7
The molecular structures of $[Pd_2(en)_2$(D-Arab1,2;3,4H_{-4})] 30 in crystals of 30 · 7 H_2O and $[Re_6(CO)_{18}$(D-ArabH_{-5})$_2]^{4-}$ 31 in crystals of $(NEt_4)_2(DBUH)_2$31. *Dashed line*: hydrogen bond

Two entirely deprotonated D-arabitol ligands are employed to connect the $Re^{I}(CO)_3$ centers in the hexanuclear cluster **31** (◉ *Fig. 7*). All but the central alkoxido function at C3 are in a bridging mode [32].

The dimolybdates and ditungstates found with arabitol are closely related to the complex anions **14** and **21**. Two dimolybdates are formed with D-arabitol ligands which adopt a sickle conformation with their *erythro*-configured end as in **14** (◉ *Fig. 2*), this is the C_4 chain from C2 to C5, in each case. The isomers differ in the coordination pattern: one isomer has C2 and C4 in the μ_2 mode, the other one C3 and C5 [25]. These isomers are the main species with tungsten(VI) as well, but a minor species is additionally formed in which the *threo* end of D-arabitol, this is C1 to C3, is used for complexation closely related to **21** [26,37].

2.5 Ribitol (Adonitol)

Ribitol **5** is a completely *erythro*-configured pentitol. No crystal structure analyses of a ribitol complex have been reported so far. However, solution studies have been performed with palladium(II) and with molybdenum(VI) and tungsten(VI).

Studies with palladium(II) revealed that in [Pd(en)$(OH)_2$]/ribitol solutions with a molar ratio of 3:1 or 2:1 the dimetallated complex [Pd_2(en)$_2$(Ribt1,2;3,4H$_{-4}$)] is formed as main species. The remaining terminal hydroxyl function is employed in an intramolecular δ hydrogen bond to O2 causing the enhanced stability compared to the minor, bis(terminal) species [Pd_2(en)$_2$(Ribt1,2;4,5H$_{-4}$)]. The monometallated terminal complex [Pd(en)(Ribt1,2H$_{-2}$)] is the main species in equimolar solutions. The detection of the isomeric complex [Pd(en)(Ribt2,3H$_{-2}$)] as a minor species gives reason to state that terminal chelation in *erythro* polyols is more favorable than a chelation with two adjacent secondary hydroxyl functions [19]. This is confirmed by the observation that (dppp)Pt^{2+} forms the terminal 1,2-diolato complex with ribitol with marked excess compared to the 2,3-*erythro* bound isomer (molar ratio 83:17, respectively) [22].

Molybdenum(VI) and tungsten(VI) are found to form only one of two possible isomeric *erythro*-type dimetallates analogously to **14** (◉ *Fig. 2*) with fourfoldly deprotonated ribitol. One of the terminal hydroxyl functions is not part of the coordination site. Even the formed isomer proves to have low stability due to unfavorable steric interaction between the terminal carbon atoms of the sickle conformer [25,26,37].

2.6 Mannitol

D-Mannitol **6** is a chiral hexitol having two terminal *erythro*-configured diol groups and a *threo* link between them.

Aqueous solutions containing [Pd(en)$(OH)_2$] and D-mannitol in a molar ratio of 3:1 or 2:1 yield the dinuclear complex **32** (◉ *Fig. 8*) as main species. The structure corroborates earlier described preferences of Pd(en) in sugar alcohol complexes: the complexation occurs at two *erythro*-linked diol groups, one of which is *threo*-configured, the other one a terminal diol function. A δ hydrogen bond of the type O–H···O^- similar to those in **26** (◉ *Fig. 5*) and **30** (◉ *Fig. 7*) stabilizes the complex, while the remaining terminal hydroxyl function is not involved in intramolecular hydrogen bonding. In 3:1 solutions this hydrogen bond is cleaved to an observable extent as the minor, trimetallated species [Pd_3(en)$_3$(D-Mann1,2;3,4;5,6H$_{-6}$)] is formed. Besides **32** (◉ *Fig. 8*), two mononuclear, *threo*-coordinated complexes are observed in equimolar solutions: [Pd(en)(D-Mann1,2H$_{-2}$)] and [Pd(en)(D-Mann3,4H$_{-2}$)]; *erythro* coordination is not detected [19].

A more complex intramolecular hydrogen bond system is built in the platinum(II) complex **33** (◉ *Fig. 8*) in which the central *threo*-configured diol is the chelation site. A γ and a δ hydrogen bond, each starting at one of the terminal donor hydroxyl functions, end at one and the same alkoxido function. The other one is an acceptor in a more strained β hydrogen bond [22].

Affirmed by DFT calculations, the completion of an octahedral coordination sphere of trimethyl-platinum(II) in a dinuclear complex is accomplished with facial tridentately bonded D-mannitol (coordination mode: μ-κ^3O^1,O^2,O^4:κ^3O^3,O^5,O^6) [38].

In **34** (◉ *Fig. 9*), the silicon center is hexacoordinated through three central, *threo*-configured diol groups. The adjacent hydroxyl functions at C2 and C5 form stabilizing intramolecular hydrogen bonds similar to **19** (◉ *Fig. 4*) and **27** (◉ *Fig. 6*) [33].

Bonding through terminal- and *threo*-configured diol groups, which are *erythro*-linked, is observed in the tris(phenylboronic acid) ester **35** (◉ *Fig. 9*) while six-membered chelate rings

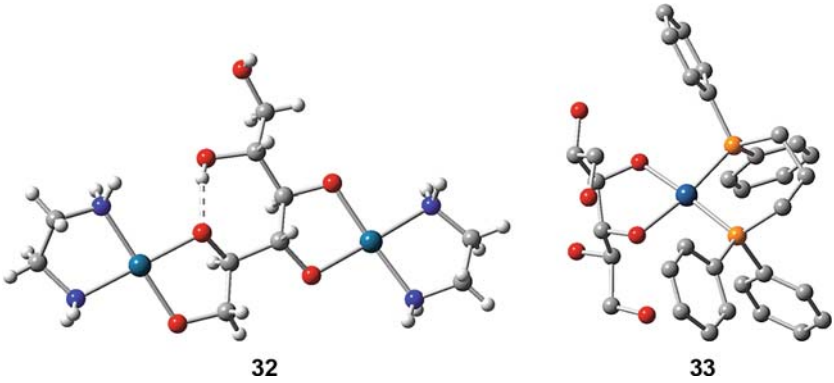


Figure 8
The molecular structures of $[Pd_2(en)_2(D\text{-}Mann1,2;3,4H_{-4})]$ 32 in crystals of 32 · 5 H_2O and (dppp)Pt(D-Mann3,4H_{-2}) 33 in crystals of 33 · CH_2Cl_2. *Dashed line*: hydrogen bond

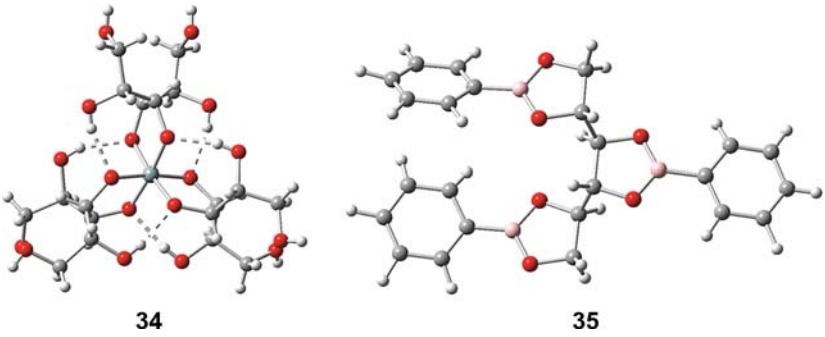


Figure 9
The molecular structure of $[\Lambda\text{-}Si(D\text{-}Mann3,4H_{-2})_3H_{-1}]^{3-}$ 34 in crystals of $Na_3$34 · 12 H_2O and the crystal structure of $(C_6H_5B)_3$(D-MannH_{-6}) 35. *Dashed lines*: hydrogen bonds

were formed with erythritol and L-threitol in **13** (● *Fig. 2*) and **15** (● *Fig. 3*), respectively [39]. Van Bekkum et al. suggest the formation of boron-mannitol monoester through bonding via the *erythro*-configured O2,O3 diol as well as via the *threo*-configured O3,O4 diol. The latter configuration is existent solely in the diester [30,40].

Two crystal structures of dimolybdates with different cations have been described. Both are of the type of **36** (● *Fig. 10*) and could be considered a dihydroxyethyl derivative of the erythritol complex **14** (● *Fig. 2*), whereas D-mannitol is only triply deprotonated in **36** similarly to an erythritol complex described above [24,41,42]. The structure of **36** (● *Fig. 10*) is retained in solution in two isomeric forms for molybdenum as well as for tungsten in a way similar to D-arabitol whose configuration matches the one of the coordination site of D-mannitol [25,26,37,43].

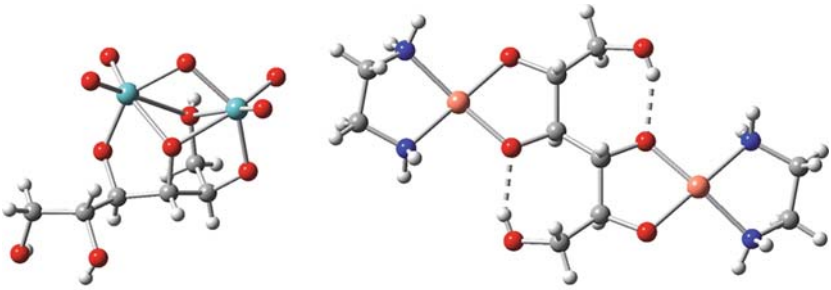


Figure 10
The molecular structures of $[Mo_2O_5(D\text{-Mann2,3,4H}_{-3})]$ 36 in crystals of Na36 · 2 H_2O and $[(en)_2Cu_2(Dulc2,3;4,5H_{-4})]$ 37 in crystals of 37 · 7 H_2O

2.7 Dulcitol (Galactitol)

Dulcitol **7** is a non-chiral hexitol with two *threo*-configured diols linked by a central *erythro*-configured diol group. To emphasize the analogies in the coordination patterns, dulcitol may be regarded as a 1,4-di(hydroxymethyl) derivative of erythritol **1**. Ethylenediamine-copper(II), for example, is found to form the dinuclear complex **37** (*Fig. 10*) which contains an erythritol substructure described above. The terminal hydroxymethyl groups form intramolecular δ hydrogen bonds resulting in an enhanced rigidity of the now *S*-shaped dulcitol [44]. This structural motif is found in several dulcitolato complexes of copper(II), nickel(II), and palladium(II). The strength of the established hydrogen bonds and thus the distances between the corresponding donor–acceptor pairs are basically a consequence of the Lewis acidity of the bonded metal-ligand fragment. In **37** the O–O distance is 2.53 Å while with the less Lewis-acidic $(tren)Ni^{2+}$ and $(Me_3tren)Ni^{2+}$ residues this distance is shortened to about 2.45 Å and 2.42 Å respectively [44]. In the copper-dulcitol compounds **37** (*Fig. 10*) as well as in the linear coordination polymer comprising the substructure **10** (*Fig. 1*) the donor–acceptor distances are about 2.62 Å [21].

Just a few species have been detected in solutions with ethylenediamine-palladium(II) at various Pd:dulcitol ratios. The only monometallated complex in equimolar solutions is $[Pd(en)(Dulc2,3H_{-2})]$ with Pd bound by a *threo*-configured diol moiety. Besides, a minor part is bonded in the dimetallated species $[Pd_2(en)_2(Dulc2,3;4,5H_{-4})]$, which is the main species in 2:1 and 3:1 solutions and whose solid-state structure resembles **37** (O–O distance 2.68 Å *Fig. 10*). A similar structure was obtained with ethylenediamine replaced by methylamine and isopropylidene-methyl-imine (O–O distance 2.54 Å. The fact that entirely deprotonated ducitol is obtained neither with $Cu(en)^{2+}$ nor with $Pd(en)^{2+}$ in solutions even with a metal:dulcitol ratio of 3 or higher, indicates the extraordinary stability of the *S*-conformation fixed by two intramolecular hydrogen bonds. The latter bonds are cleaved with the supporting ligand 1,3-bis(2′-dimethylaminoethyl)hexahydropyrimidine (tm-21:32-tet) which provides a suitable N_4 set to bind two palladium(II) centers which, in turn, could be bonded by a 1,2,3-bis-diolato moiety with the middle alkoxo function in a μ_2 mode. As a result a tetranuclear complex **38** (*Fig. 11*) with entirely deprotonated dulcitol is obtained [19].

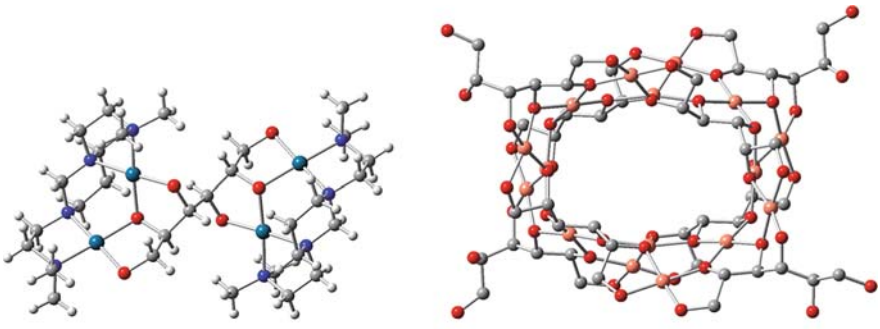


Figure 11
The molecular structures of $[Pd_4(tm\text{-}2,1{:}3,2\text{-}tet)_2(DulcH_{-6})]^{2+}$ 38 in crystals of $38Cl_2 \cdot 16\ H_2O$ and $[Cu_{16}(D\text{-}SorbH_{-6})_4(D\text{-}Sorb1,2,3,4H_{-4})_4]^{8-}$ 39 in crystals of $Li_8$39 $\cdot 46\ H_2O$

Coordination via the *threo*-configured diol at O2 and O3 is also favored by $(dppp)Pt^{2+}$. Species using the central *erythro*-diol or one of the terminal non-stereogenic diols have been detected in small amounts only [22].
Molybdenum and tungsten also use the central tetraol set of dulcitol, however, in a sickle conformation, to form dimolybdates respectively ditungstates with a coordination pattern similar to that found in the erythritol complex **14** (*Fig. 2*) [25,26,27,43].

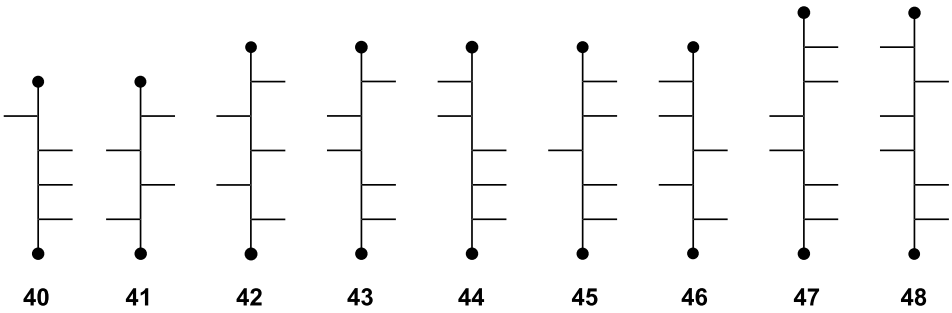
2.8 Sorbitol (Glucitol)

D-Sorbitol **8** is a hexitol with its diol functions from C2 to C5 configured with the sequence *threo*, *threo*, *erythro*. The resulting C_1 symmetry is possibly the reason for the rather small number of characterized complexes since the crystallization as well as the NMR spectroscopic characterization is hampered by the usually obtained mixture of species.
To the best of our knowledge, published crystal structure data are available merely for a toroidal, hexadecanuclear copper(II) complex **39** (*Fig. 11*) with eight multiply deprotonated D-sorbitol ligands. Half of them are completely deprotonated η^6,μ_5 hexaanions, the other half acts as μ_3 tetraanions with their *xylo*-configured C1–C4 subset. It was found that the magnetic coupling behavior is determined by the distribution of Cu–O–Cu angles and correlates with the ligand configurations of the chelating 1,2-diol entities [45].
Molybdenum is found to form four different dimolybdate(VI) complexes, all having proportions in the range of 21 to 32%. These are: two complexes with a central *threo*-configured diol of the type **20** with one of them using the C1–C4 subset, the other one using the C2–C5 subset, and two *erythro*-complexes of the type **14** (*Fig. 2*) with the C3–C6 subset in two reversed orientations [25,26].
Coordination to the *xylo*-configured site C2–C4 similar to **21** is mainly observed for tungsten. Two minor species are the *erythro*-complexes employing the C3–C6 subset as molybdenum [27,37].

Boronic acid is reported to form esters with D-sorbitol in alkaline solution. Van Bekkum et al. suggest the formation of a monoester involving the *erythro*-configured diol while a mono- and a diester may be formed over a *threo*-configured site [40]. Tridentate D-sorbitol is found to occupy three of four coordination sites of the boron center in a ferrocene derivative with its hydroxyl functions at C2, C3, and C5 [46].

2.9 L-Iditol and Higher Sugar Alcohols

Further NMR studies were performed of molybdate(VI) and tungstate(VI) complexes with the hexitols **40** and **41**, the heptitols **42**–**46**, and the octitols **47** and **48**.



Carbon numbering: C1 is situated at the top of each chain. ● = CH_2OH

An *erythro* complex similar to **14** (❯ *Fig. 2*) is established with the C2–C5 coordination site of D-altritol **40** [26]. L-Iditol **41** forms a C_2-symmetric dimolybdate(VI) complex in the pH range 5–10 using the all-*threo*-configured C2–C5 chelation site resulting in a structure similar to the threitol-complex **20**. A structure of the same type emerges with *meso*-glycero-ido-heptitol **42** with the all-*threo*-configured chelation site C2–C5. The corresponding ditungstate(VI) complexes are formed only at neutral pH ≤ 7. In the pH range 8–9, the hexitol **41** acts as a tridentate ligand with the chelation site C2–C4 to give a ditungstate(VI) of the type **21**, while the heptitol **42** acts as a pentadentate ligand in two isomers with a structure similar to **28**: the more stable isomer uses the C2–C6 chelation site, the other one uses the site C1–C5. In the pH range 9–12, the hexitol **40** also forms a complex similar to **28** [35].

D-glycero-D-galacto-heptitol **43** (perseitol) was first found to form a pair of isomers using the C2–C5 galacto site of chelation in dimolydate- and ditungstate complexes having a structure similar to **14** (❯ *Fig. 2*) [26,43]. A mixed bis-ditungstate complex was detected later. In this complex, one ditungstate is bound at the *erythro*-chelation site C4–C7, while the other ditungstate is coordinated through the *threo*-site C1–C3. The local structures at the two different sites resemble the structure types **14** (❯ *Fig. 2*) and **21**, respectively [37]. Bis-dimolybdate complexes have been observed with the octitols *meso*-erythro-*manno*-octitol **47** and D-erythro-L-gluco-octitol **48**, both comprising a C3–C8 perseitol **43** subset. In **47**, each of the two *eythro* chelation sites C1–C4 and C5–C8 is bound to a dimolybdate (for local structures compare with **14**, Fig 2). In contrast **48** binds one dimolydate in a *threo* complex at the chelation site C1–C4, the other is bound at the *erythro* chelation site C5–C8 (local structures **20** and **14**,

respectively). A less stable species with both ligands **47** and **48** uses the central *galacto* site C3–C6 to give *erythro* complexes (local structure compare with **14**, ◉ *Fig. 2*) [47].
D-Glycero-D-*manno*-heptitol **44** (volemitol) forms identical dimolydate and ditungstate complexes. The main species comprises the *arabino* chelation site C1–C4 with the local structure **14** (◉ *Fig. 2*) in two isomers. The same local structure is established in the other species which uses the C3–C6 chelation site. The reduced stability of the latter species is caused by the unfavorable steric interaction of the chelation site and the side chain [48].
Ditungstate complexes with the local structures **21** and **28** are observed with the ligands *meso*-glycero-gulo-heptitol **45** and D-glycero-L-gulo-heptitol **46**. The main species with tridentate **45** involves the central *xylo*-configured chelation site at C3–C5 resulting in a local structure **21** (at pH 9.9). A less stable complex is formed with the coordination site C2–C6 of pentadentate **45** (local structure **28**). Two isomers with the latter local structure **28** are obtained with **46** at pH 11.7. One uses the C3–C7 chelation site, the other binds ditungstate at the C2–C6 site. Local structure **21** is formed at pH 10.6, where **46** acts as tridentate ligand using its C4–C6 coordination site [49].

2.10 Regioselectivity and Stability

Though sugar alcohols provide various metal-binding sites, their complexation behavior is not as unpredictable as one might expect. Rules determining the chelation sites in ethylenediamine-palladium(II) complexes with discrete chelate rings, that is with neither chelating alkoxo function being in a bridging mode, seem to count for other metal complexes of these type as well [19]:

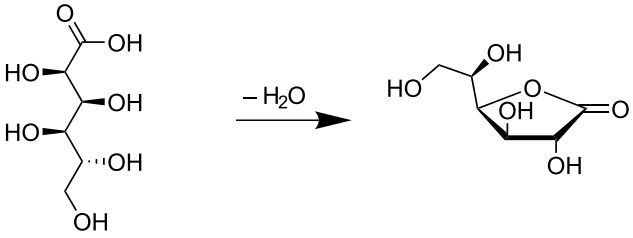
1 The stability of chelate rings formed by *threo*-configured, *erythro*-configured and terminal diols increases in the sequence: *erythro* < terminal < *threo*.
2 The *erythro*-linked adjacent chelating diol functions are more stable than *threo*-linked ones.
3 An additional stability gain emerges upon intramolecular hydrogen bonding between donor hydroxy and an acceptor alkoxido moiety. The most stable type of hydrogen bond is the δ type, that is, a C_4-chain separates the hydrogen bond donor from the acceptor.

The prediction of chelation sites of sugar alcohols in multinuclear metal complexes with alkoxido functions in a bridging mode is rather difficult due to the influence of further factors such as the metal size and oxidation state or the distance of the O atoms of the ligand. However, based on a wealth of NMR solution data and supported by some crystal structures, rules regarding the regioselectivity and stability of the well-investigated dimolydate- and ditungstate-complexes of sugar alcohols could be established [25,27,35,37,43,48,49,50].

3 Aldonic and Aldaric Acids

The oxidation of the primary hydroxyl function at C1 of a sugar alcohol to a carboxyl function leads to the corresponding aldonic acid. Further oxidation of the remaining terminal hydroxyl function of an aldonic acid to a carboxyl function leads to an aldaric acid. Both, aldonic and aldaric acids easily form lactones through intramolecular condensation reactions. As an exam-

ple, the formation of D-gluconic acid γ-lactone **50** (also known as D-glucono-1,4-lactone) starting from D-gluconic acid **49** is depicted in ● *Scheme 1*.

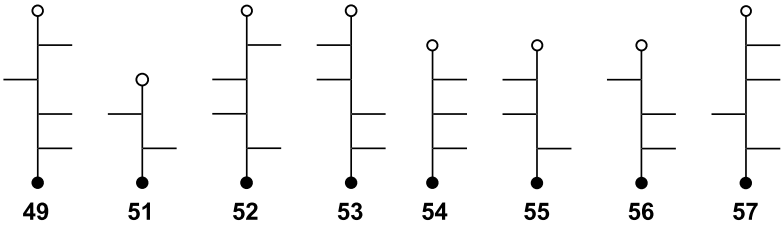


49 **50**

■ **Scheme 1**

3.1 Aldonic Acids

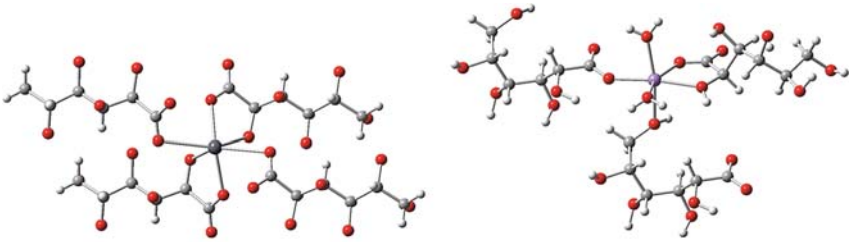
In most of the D-gluconic acid **49** complexes known from the literature, the metal centers are bonded through the α-hydroxycarboxylate moiety and five-membered chelate rings are formed. While the carboxyl function is always deprotonated in these structures, it depends on the Lewis acidity of the metal or the pH of the solution, whether the α-hydroxyl residue is deprotonated or not. No deprotonation of the latter occurs in the solid-state structures of lead(II)- and manganese(II) D-gluconate (**58** and **59**, respectively, ● *Fig. 12*). Furthermore D-gluconic acid binds to both metal centers by a single carboxylate oxygen atom and to the manganese central atom in **59** (● *Fig. 12*) by the terminal alcoholic hydroxyl function which leads to a coordination polymer [51,52]. $Co^{III}(en)_2$ is also chelated by the α-hydroxycarboxylato moieties of D-gluconic acid and L-mannonic acid, respectively [53]. Magnesium(II) is reported to be coordinated by the α-hydroxycarboxylato moiety and the adjacent hydroxyl function at C3 [54].



49 **51** **52** **53** **54** **55** **56** **57**

Carbon numbering: C1 is situated at the top of each chain. ● = CH_2OH ○ = COOH

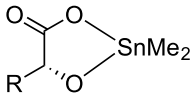
A stronger Lewis acid such as Me_2Sn^{2+} forces deprotonation of the α-hydroxyl function even at pH 5.5 (**60**). This coordination mode is also suggested for a copper(II)-D-gluconic acid complex [55]. In more alkaline solution (pH 8), the hydroxyl function at C4 is also deprotonated and Me_2Sn^{2+} is coordinated by the α-alkoxidocarboxylato moiety and O4 of tridentate



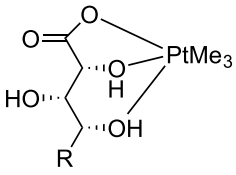
58 59

Figure 12
Coordination of the central metals in the crystal structures of Pb(D-Glc1A1H$_{-1}$)$_2$ 58 and Mn(D-Glc1A1H$_{-1}$)$_2$ · 2 H$_2$O 59

D-gluconate [56]. This coordination mode is also observed in a praseodymium(III) complex in the pH range of 8–10 and an equimolar ratio of metal and ligand. At pH 6–7 and a molar ratio of ligand:metal greater than 2:1, complete deprotonation of all the hydroxyl functions at C1, C2, and C4 is not achieved. Now the praseodymium(III) center is coordinated by two tridentate D-gluconate ligands with three of four coordinating alcoholic hydroxyl functions not being deprotonated [57]. A similar coordination pattern is reported for the trimethyl-platinum(IV) complex **61** which was obtained from aqueous solution. Just the carboxyl function is deprotonated while the coordinating hydroxyl groups at C2 and C4 remain non-deprotonated. A D-threonic acid **51** derivative has been shown to bind to an octahedrally coordinated platinum(IV) center through its α-hydroxycarboxylate moiety [58]. The metals manganese, cobalt, nickel, copper, cadmium, mercury, lead, and zinc in the oxidation state +II are also reported to be coordinated by the O1,2,4 set of D-gluconic acid [59].



60
R = (CH(OH))$_3$CH$_2$OH



61
R = CH(OH)CH$_2$OH

In the pH range of 5–10, the trivalent metals aluminum, gallium, and indium are coordinated by the deprotonated hydroxyl function of C1–C4 of tetradentate D-gluconic acid [60]. Coordination with involvement of the carboxylate moiety is also reported for bismuth(III) in acidic solution. In alkaline solution, however, D-gluconic acid acts as a tetradentate ligand using its deprotonated secondary hydroxyl functions only [61]. Recently the crystal structure of [Cu$_4${μ-(α-D-Glc-1P)}$_2$(μ-Glc1A1H$_{-1}$)$_2$(bpy)$_4$](NO$_3$)$_2$ has been reported in which the α-hydroxycarboylato moiety of gluconic acid is coordinated to one copper ion resulting in a five-membered chelate ring similar to **58** and **59** (*Fig. 12*). Additionally the carboxylato oxygen atom links two copper ions. A linear coordination polymer complex is obtained with D-glucarate [62].

While structural information on complexes of molybdenum(VI) and tungsten(VI) with sugar alcohols is available for solid-state as well as for solution species, crystal structures have not been determined so far for the corresponding complexes with aldonic and aldaric acids (compare ● *Sect. 3.2*), that is, the following structural description is based solely on NMR spectroscopic data.

At pH 5–7 and with an excess of ligand, tungsten(VI) is coordinated by the α-alkoxidocarboxylato moiety of **49** similarly to **60**. However, a dinuclear species chelated by the alcoholic hydroxyl set of C3–C6 is observed when there is an excess of tungsten. Since the central diol at C4 and C5 is *erythro*-configured, the local structure might be similar to **14** (● *Fig. 2*). Additionally, coordination through the alcoholic hydroxyl functions at C2, C3, C5, and C6 is reported [63].

The α-alkoxidocarboxylato moiety of **49** is the chelation site in peroxido-complexes of molybdenum and tungsten as well. Similar structures are obtained with D-galactonic acid and L-mannonic acid [64].

Molybdenum and tungsten are coordinated by the α-alkoxidocarboxylato moiety of **52** in solutions with a molar ratio of ligand:metal of 2:1 in a pH range of 3–9. The stability of these chelates decreases with increasing pH. At a molar ratio of ligand:metal of 1:2 and with maximal stability at about pH 7.5, dinuclear complexes structurally similar to those formed by galactitol become the main species. The metal coordination is now established by the secondary alkoxido functions with a local structure similar to **14** (● *Fig. 2*). Additional species are found for molybdenum(VI) in weakly acidic solutions with a molar ratio of molybdenum:ligand of 2:1. Entirely deprotonated, hexadentate D-galactonic acid is bonded to two molybdenum centers through the alkoxido functions of the C1–C3 subset and the C4–C6 subset, respectively. Two molybdenum atoms may be bonded by fivefold deprotonated D-galactonic acid which acts as a pentadentate ligand through its secondary alkoxido functions in a sickle conformation and the carboxylate function [65].

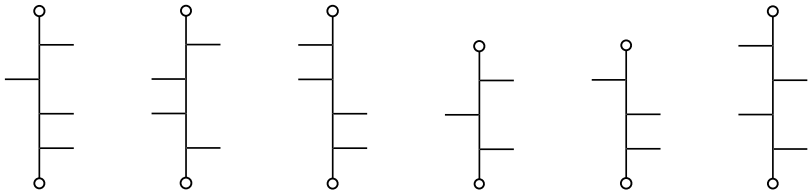
L-Mannonic acid (L-enantiomer of **53**) is found to form similar complexes with molybdenum(VI) described above for D-galactonic acid, but additional species are found for tungsten(VI). An excess of ligand preferentially leads to the formation of mononuclear complexes with the α-alkoxidocarboxylato moiety of **53** as the chelation site, especially in the pH range of 3–6, while di- and tetranuclear species are enriched in solutions with an excess of metal. Three dinuclear species structurally analogous to those formed with D-mannitol are observed in neutral and weakly alkaline solutions. Coordination of the alkoxido functions in the positions 3 to 6 in a sickle conformation leads to a dimetallate complex with a local structure similar to **14** (● *Fig. 2*), while a zigzag conformation of the C2–C5 subset results in a structure similar to **20**. With L-mannonic acid these complexes are also detected with tungsten(VI). Finally a complex using the C1–C4 subset in a sickle conformation is observed mainly with molybdenum in solutions of a pH in a range of 3–7 [66].

Several structures of borate esters of aldonic acids have been described based solely on NMR spectroscopic data. Borate mono- and diesters are preferentially formed at the *threo*-configured diol sites of D-gluconic acid. Less stable esters are formed by bonding at the *erythro*-configured diol at C4 and C5. Still less stable are six-membered borolane rings as a result of bonding via O2 and O4. The 3,4- and 5,6-positions are occupied by borate anions resulting in a diborate ester. D-mannonic acid forms borate monoesters and two diastereomeric diesters preferentially through the *threo*-configured diol at position 3 and 4, while the *erythro*-configured diol at C4

and C5 leads to less stable esters. Additionally the formation of a diborate ester is also reported. The two borate moieties are bonded at the *threo*-configured 3,4-position and at the terminal diol in 5,6-position. Borate monoester- and diester formation has also been investigated with D-ribonic acid **54**, D-lyxonic acid **55**, D-arabinonic acid **56**, D-gulonic acid **57**. Four borate esters of ribonic acid have been found but further interpretation was prevented due to the large number of ^{13}C NMR signals. The *threo*-3,4-diol site is preferred in borate esters of lyxonic acid and gulonic acid, whereas the *threo*-2,3-position proves to be the most stable one in borate esters of arabinonic acid [40,67].

3.2 Aldaric Acids

Reactions in approximately neutral aqueous solutions lead to three solid-state structures of complexes of D-glucaric acid **62** (also known as D-saccharic acid) with the central metals aluminum(III), zinc(II), and copper(II), all exhibiting similar structure features: (a) both α-hydroxycarboxylate moieties of D-glucaric acid act as chelation sites, (b) deprotonation occurs at the carboxylate group and, in the case of bonding to the stronger Lewis acid aluminum(III), additionally at both α-hydroxy functions, and (c) the central hydroxyl functions in 3,4-position are never involved in metal coordination. A dinuclear complex **69** (*Fig. 13*) is formed with aluminum(III) with one of the two α-alkoxido functions in a μ_2 mode [68], whereas coordination polymers are formed by the zinc(II) complex **68** (*Fig. 13*) and by the copper(II) complex, for which no crystallographic data are available [69,70].

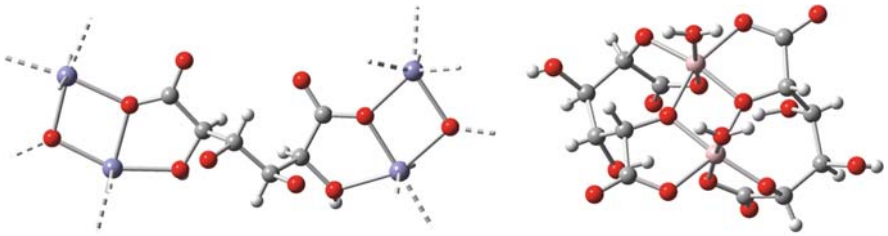


62 **63** **64** **65** **66** **67**

Carbon numbering: C1 is situated at the top of each chain. o = COOH

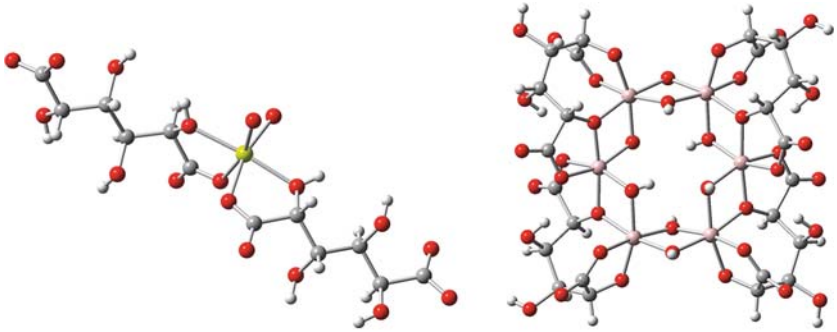
Both α-hydroxycarboxylate moieties of galactaric acid **63** (also known as mucic acid) act as a chelation site for magnesium centers in the solid-state structure **70** (*Fig. 14*), in which the galactarate is centrosymmetric and adopts a zigzag conformation [71]. Four galactarate molecules acting as bis(α-alkoxidocarboxylato) ligands and eight hydroxido ligands coordinate to six aluminum centers in a cyclic hexaaluminate **71** (*Fig. 14*). This structure is obtained from a reaction in strongly alkaline solutions and combines carboxylato-, hydroxido- and alkoxido bonding to the aluminum centers.

Copper(II), however, shows a different complexation behavior with galactaric acid in strong alkaline media. A coordination polymer **72** (*Fig. 15*) is formed by the copper atoms and the galactarate ions which act as *threo*-bis-diolato ligands. A polymer of the same type is formed with erythritol (compare **10**, *Fig. 1*) [72]. Carboxylato or hydroxido bonding is not observed in this copper complex. In weakly acidic solution with a pH of 4, however, copper(II) is chelated through an α-hydroxycarboxylate moiety of galactaric acid. Deprotonation



68 69

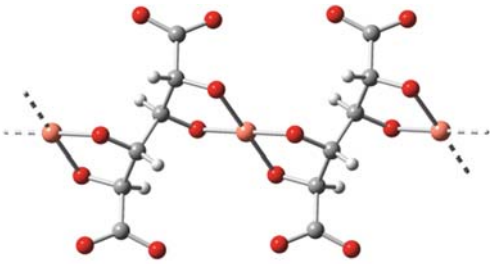
Figure 13
A part of the crystal structure of Zn(D-Glc1,6A$_2$1,6H$_{-2}$) 68 in crystals of 68 · 2 H$_2$O and the molecular structure of [{Al(D-Glc1,6A$_2$1,2;5,6H$_{-4}$)(H$_2$O)}$_2$]$^{2-}$ 69 in crystals of K$_2$69 · H$_2$O. *Dashed lines*: indicate Bonds to adjacent oxygen atoms



70 71

Figure 14
A section of the crystal structure of Mg(Gal1,6A$_2$1,6H$_{-2}$) · 2 H$_2$O 70 and the molecular structure of [Al$_6$(Gal1,6A$_2$1,2;5,6H$_{-4}$)$_4$(OH)$_8$]$^{6-}$ 71 in crystals of Na$_6$71 · 21 H$_2$O

of the α-hydroxyl function is reported to occur at pH 7.5 [73]. These observations indicate for copper(II) that carboxylato bonding becomes more and more unfavorable with increasing pH. The following structures of complexes with aldaric acids are solely based on NMR spectroscopic data. Several complexes with both α-alkoxidocarboxylato residues of glucaric acid **62** as chelation sites are discussed with molybdenum(VI) and tungsten(VI) as central metals. Both β-hydroxyl functions are deprotonated in a less acidic solution of pH 4.5–6 and a double molar excess of metal whereupon a dinuclear complex is formed. At a higher pH of about 6–8.5, coordination through the carboxylato residue becomes unfavorable and chelation through the alkoxido functions at C2–C5 leads to a *threo*-complex similar to that described above for D-sorbitol [74]. As with D-galactonic acid, two molybdenum(VI) atoms are chelated through the alkoxido functions of the C1–C3 subset and the C4–C6 subset, respectively, of entirely deprotonated galactaric acid **63** in weakly acidic solutions with a molar ratio of molybdenum:ligand of 4:1. Obviously the strong Lewis acidity of molybdenum(VI) enables the carboxylate oxygen atom of galactaric acid to take over the part of the terminal, less acidic



72

Figure 15
The solid-state structure of $[Cu(Gal1,6A_2H_{-6})]^{4-}$ 72 in crystals of $Na_4$72 · 12 H_2O. *Dashed lines* indicate bonds to adjacent oxygen atoms of the linear coordination polymer

alcoholic hydroxyl function in galactonic acid. The central diol functions are not involved in complexation in equimolar solution. However, in neutral solutions with an excess of molybdenum, the central diol function in combination with the adjacent alcoholic hydroxyl functions are arranged in a sickle conformation and form a dinuclear *erythro*-complex also found for galactitol [75]. Most of the complexes just described are also formed with tungsten(VI) as central metal [76].

Both, glucaric acid **62** and galactaric acid **63** form mono- and dinuclear peroxido complexes with molybdenum(VI and tungsten(VI). In all of them one or two α-alkoxidocarboxylato moieties are chelating the metal centers. The dinuclear species is predominantly formed in solutions with a molar ratio of metal:ligand of 2:1 [64].

The coordination behavior of molybdenum(VI) with D-mannaric acid **64** is similar to that with galactaric acid: two molybdenum atoms are bonded either by the two α-alkoxidocarboxylato moieties of fourfoldly deprotonated D-mannaric acid or by the alkoxido functions of the C1–C3 subset and the C4–C6 subset, respectively, of entirely deprotonated D-mannaric acid. Due to the *threo*-configuration of the central diol residue, a *threo* complex is established in neutral solutions with a molar ratio of Mo:ligand of 4:1 instead of an *erythro* complex which was observed with galactaric acid (compare structure **20** and **14** (◉ *Fig. 2*), respectively) [75]. The latter *threo* complex as well as a complex with the α-alkoxidocarboxylato moieties as chelation sites is also formed by tungsten(VI) [76].

In alkaline solutions of a pH of 11.5 containing calcium ions, aluminum(III) forms C_2-symmetric dinuclear sandwich-type complexes. D-glucaric acid **62** has to adopt a conformation which allows all coordinating alkoxido functions to point in the same direction. The aluminum centers are bonded by the *threo*-configured 2,3-site as well as by the *erythro*-configured 4,5-chelation site. Two *threo*-3,4-diolato moieties of two mannaric acid **64** ligands are the bonding sites of the tetrahedrally coordinated aluminum center. Tetrahedrally coordinated aluminum is also present in the bis-diolato complexes with xylaric acid **65** and D-arabinaric acid **66**, respectively; both complexes are formed in calcium-containing solutions. Xylaric acid is bonded to aluminum through its C2–C3 *threo*-diol function, at which the central hydroxyl function at C3 is in the μ_2 mode and is bonded to a calcium ion as well. A C_2-symmetric dinuclear complex is formed with D-arabinaric acid. The aluminum centers are coordinated to a C1–C2 α-hydroxycarboxylate moiety from one ligand and to a C3–C4

erythro-diol function from the other ligand. The complex formation is supported by calcium ions which are bonded at several sites of the complexes just described [77].
Borate mono- and diesters are preferentially formed of the *threo*-3,4-diol functions of glucaric acid **62** [40,67,78]. Several cations have been shown to have different effects on the borate-ester equilibria. Monovalent ions such as Na^{I}, K^{I}, and Ag^{I} hardly show any effect, whereas divalent cations with moderate polarizing abilities (Mg^{II}, Ca^{II}, Co^{II}, Ni^{II}, Sr^{II}, Cd^{II}, Ba^{II}) increase the diester formation. Strongly polarizing cations such as Al^{III}, Cu^{II}, Zn^{II}, Pr^{III}, and Pb^{II} induce dissociation of the borate diesters [79]. Monoesters and diesters with galactaric acid **63** are formed mostly through the *threo*-configured diols at 2,3-position. Diborate ester formation is also observed in which the two borate anions are bonded at the equivalent *threo*-configured diols in 2,3- and 4,5-position [40,67].

3.3 Regioselectivity and Stability

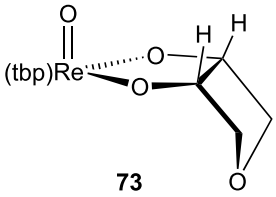
The electron-withdrawing effect of the carboxyl group markedly enhances the acidity of the adjacent α-hydroxyl group. As a consequence, the α-hydroxycarboxylate moiety becomes an attractive ligand mainly in acidic and neutral solutions in which sugar alcohols form complexes of low stability only, especially with metals of weak Lewis acidity. In alkaline solutions, however, the less acidic alcoholic hydroxyl functions are mostly preferred in complex formation. For these solutions the rules referring to the regioselectivity and stability of sugar-alcohol complexes may be valid for aldonic and aldaric acids as well.

4 Anhydroerythritol

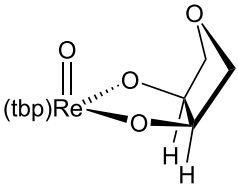
Anhydroerythritol (AnEryt) or *cis*-oxolane-3,4-diol is a particularly well-suited ligand to investigate the coordination chemistry of carbohydrates as it can be considered to be a molecular cutout of important glycosides such as the nucleosides. On the one hand, its acidity is about the same as that of the glycosides, on the other hand, AnEryt is easily obtainable, configurationally stable in the entire pH range of interest, and chemically stable under alkaline conditions. AnEryt shares the latter two properties, which are critical for the glycoses, with the glycosides. Contrary to even glycosides, as a *meso* compound AnEryt gives rise to particularly simple ^{13}C NMR spectra. Finally, AnEryt is achiral, a property that is found to be beneficial for the crystallization of its complexes.
Usually the diolato κ^2O,O'-AnErytH$_{-2}$ ligand forms five-membered chelate rings in the coordination compounds, the central atoms spanning a size range from boron(III) to lead(II). The unusual versatility of the AnErytH$_{-2}$ ligand is the result of the variability of a furanoidic ligand's bite, AnEryt being the prototype of this kind of ligand.
Coordination of central metal atoms with auxiliary ligands so that only two free binding sites remain, leads, in most cases, in combination with the deprotonation of the diol functions of AnEryt, to a predictable binding mode of κ^2O,O' as in the cases of cobalt(III), $\{RuNO\}^6$, oxorhenium(V) and palladium(II).
Four of the six binding sites of the octahedrally coordinated cobalt(III) in [(en)$_2$Co(AnErytH$_{-2}$)] are occupied by ethylenediamine ligands, κ^2O,O'-AnErytH$_{-2}$ binds over the remaining two sites [80]. Caused by the *cis* conformation of the two diol functions of

AnEryt, the five-membered chelate ring is flat. An intramolecular hydrogen bond between the O1 atom of the AnEryt ring and an en ligand is responsible for the conformation of the complex (the conformation of the oxolane ring is $^{C4}T_{O1}$, distorted in the direction of E_{O1}). The Co–O bond length averages 1.88 Å.

Comparable conditions can be found for the structure of [*mer*-(dien)(NO)Ru(AnErytH$_{-2}$)] $BPh_4 \cdot 2\ H_2O$ where four of six coordination sites were occupied by NO and *mer*-dien so that κ^2O,O'-AnErytH$_{-2}$ is once again the ligand for the remaining two binding sites with a O–C–C–O angle of −22.1° [81]. Diolate-Ru(NO) bonding is governed by the *trans*-influence of the nitrosyl ligand. The Ru–O(*trans*) distance is with 1.94 Å markedly shortened compared with the Ru–O(*cis*) distance with 2.05 Å, the Ru–O(*trans*)–C angle is with 117.9° always more obtuse than the Ru–O(*cis*)–C angle with 112.1° and, most remarkably, considering the usual features of polyolato-metal structures, the O(*trans*) center does not act as a hydrogen-bond acceptor. On the other hand, O(*cis*) is a hydrogen-bond acceptor, as is usual. [(tpb)ReO(AnErytH$_{-2}$)] is also an octahedrally coordinated complex and built from equal amounts of two isomers, which differ in the orientation of the oxolanediolato ligand with respect to the (tbp)ReO fragment [82] (**73**, **74**).



73



74

As known from the structure of [(en)$_2$Co(AnErytH$_{-2}$)], the O1 atom of the AnEryt ring in the *anti* isomer **73** of [(tpb)ReO(AnErytH$_{-2}$)] builds a hydrogen bond with the tpb ligand. No intramolecular hydrogen bonding can be found in the *syn* isomer **74**, the Re–O bond averages 1.93 Å in both cases. In contrast, the O–C–C–O angles of AnErytH$_{-2}$ differ in the case of *syn* and *anti* coordination, in compound **73** the angle amounts to −1.6° and in **74** to −19.3°.

In the dinuclear anion $[Re_2(CO)_6(\mu\text{-OMe})_2(\mu\text{-AnErytH}_{-1})]^-$ **75**, AnEryt differs from the currently known coordination pattern [32]. A single monodentate κO-μ-AnErytH$_{-1}$ ligand takes one of the three alkoxido positions in the usual hexacarbonyl-trialkoxo-dirhenate(I) anion, leaving the other two for methoxido groups, as shown in ❯ *Fig. 16*. The second diol function forms an intramolecular hydrogen bond to the deprotonated diol function. The distances of the coordinated O atom to the two Re atoms average 2.17 Å and the O–C–C–O angle of AnErytH$_{-1}$ is 42.6°.

For the square-planarly coordinated palladium, the use of an auxiliary ligand leads to the expected κ^2O,O'-coordination of one AnErytH$_{-2}$. There are two known characterized complexes with [(bpy)Pd(AnErytH$_{-2}$)] $\cdot$ 6.5 H_2O and [(en)Pd(AnErytH$_{-2}$)] $\cdot$ 4 H_2O [83,84]. Two binding sites of the square-planarly coordinated palladium are occupied by bpy or en and two sites bind to the deprotonated AnErytH$_{-2}$. In both cases the Pd–O bond length averages 1.99 Å With 43.0°, the O–C–C–O angle of [(bpy)Pd(AnErytH$_{-2}$)] $\cdot$ 6.5 H_2O and 38.7°, the O–C–C–O angle of [(en)Pd(AnErytH$_{-2}$)] $\cdot$ 4 H_2O, the angles are closer to that of the κO-μ-AnErytH$_{-1}$ ligand than to the angles of the compounds introduced so far.

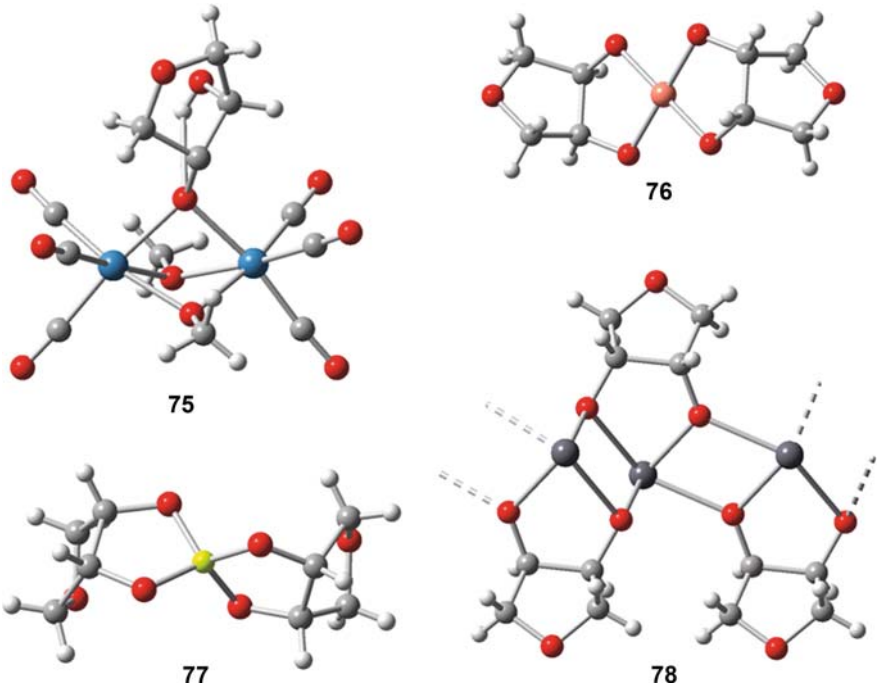


Figure 16
X-ray structures of the anionic $[Re_2(CO)_6(\mu\text{-}OMe)_2(\mu\text{-}AnErythH_{-1})]^-$ 75 in [K([18]crown-6)]75, $[Cu(AnErythH_{-2})_2]^{2-}$ 76 in $A_2$76 **· 4 H_2O (A = Li, K, Na), $[B(AnErythH_{-2})_2]^-$ 77 in K77 · H_2O and a section of the linear coordination polymer built by $Pb(AnErythH_{-2})$ 78 in 78 · H_2O**

When the central metal atom is free of auxiliary ligands there are more possibilities for coordination of AnEryt. Keeping to the square-planar coordination Cu(II) builds complexes with two AnEryt, as shown in ❯ *Fig. 16*. Three structures of the type $A_2[Cu(AnErythH_{-2})_2] \cdot 4\ H_2O$ (A_2**76** $\cdot 4\ H_2O$) with A = Li, Na and K [85,86] have been determined. The structures of Na and K on the one hand and Li on the other hand differ in the coordination of the alkali ions to the deprotonated diol functions of the $Cu(AnErythH_{-2})$ substructure. The Cu–O bond length averages 1.93 Å and the O–C–C–O angle of $AnErythH_{-2}$ grows in dependence on the cation from 33.7 (Li) to 35.4° (K).

Comparable coordination schemes with two $AnErythH_{-2}$ ligands exist for tetrahedrally coordinated metal centers, too. Boron(III), beryllium(II) and lead(II) build structures where the metal center is tetrahedrally coordinated with two $AnErythH_{-2}$ ligands [87,88], for example the structure of $K[B(AnErythH_{-2})_2] \cdot 2\ H_2O$ (K**76** $\cdot 2\ H_2O$) in ❯ *Fig. 16*.

According to the increasing atom size, the bond length increases from the B–O bond length with 1.47 Å over the Be–O bond length with 1.63 Å to an Pb–O bond length between 2.25 Å to 2.39 Å The boron compound shows two similar O–C–C–O angles of 12.0°, the beryllium compound an angle of 26.8° and a second of 0.8°.

As shown in ❯ *Fig. 16*, lead(II) forms not only monomers of $[Pb(AnErythH_{-2})_2]^{2-}$ but also polymers of $Pb(AnErythH_{-2}) \cdot H_2O$ (**78** $\cdot H_2O$), which are chains of Pb(II) coordinated by

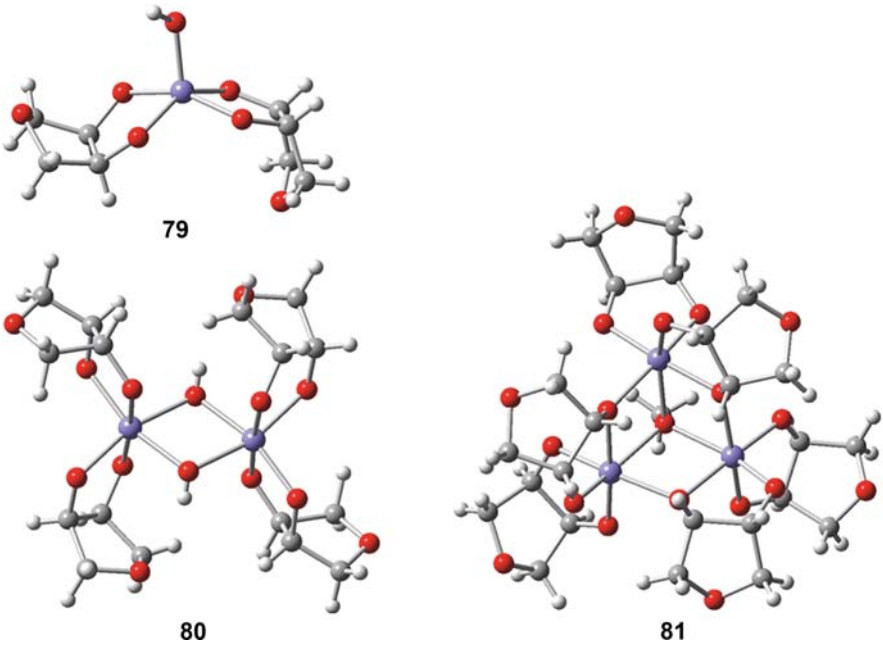


Figure 17
X-ray structures of 79 in $Na_2[Fe(AnErythH_{-2})_2(OH)] \cdot 0.5\ NaNO_3 \cdot 3.5\ H_2O$, 80 in $Ba_2[Fe_2(AnErythH_{-2})_4(\mu\text{-}OH)_2] \cdot 12\ H_2O$, and 81 in $Na_4[Fe_3(AnErythH_{-2})_6(OMe)] \cdot 2.5\ NaNO_3$

four bridging *O*-atoms [89]. In contrast to the monomeric $[Pb(AnErythH_{-2})_2]^{2-}$, the Pb–O distances range between 2.24 and 2.47 Å the O–C–C–O angle of $AnErythH_{-2}$ averages 20°. $[Fe(AnErythH_{-2})_2(OH)]^{2-}$ **79**, $[Fe_2(AnErythH_{-2})_4(\mu\text{-}OH)_2]^{4-}$ **80** and $[Fe_3(AnErythH_{-2})_6(OMe)]^{4-}$ **81** are three anionic compounds with doubly deprotonated $AnErythH_{-2}$ ligands and iron(III) [90]. The three compounds are shown in ❯ *Fig. 17*. In the case of **79** there are two isomers caused of *syn* and *anti* orientation of AnEryt ligands. ❯ *Figure 17* only shows the *syn/anti* isomer, although the structure of the *anti/anti* isomer is also known. In **79**, iron(III) is square-pyramidally coordinated by two $AnErythH_{-2}$ and one hydroxido ligand. Nevertheless the ideal geometry varies at one coordinate of the Berry-pseudorotation at about 16.6% for the *syn/anti* isomer and at 27.6% for the *anti/anti* isomer. **80** is a $(\mu\text{-}OH)_2$ bridged dimer of **79** where both iron centers now are coordinated sixfold. Species **81** is a three-centered tetraanion where three of six diolato ligands bind to one of the iron atoms, the other three diolato ligands build $\eta^1{:}\eta^1,\mu_2$-bridges to a central μ_3-methoxidotriiron unit.

The distances between the iron centers and the O atoms vary depending on the compound. In **79**, the O atoms of the AnEryt ligands have distances between 1.93 and 1.98 Å to the iron center, in **80** the distances range from 1.96 to 2.02 Å and in compound **81** from 1.94 to 2.09 Å (❯ *Fig. 17*).

Similarly to iron(III), there are known mono-, di- and trinuclear compounds with silicium. Almost identical to **79** are the structures of $[M(AnErythH_{-2})_2(OH)]^-$ with M = Si **82** and Ge [91]. In $[SiPh(AnErythH_{-2})_2]^-$ [92] **83** a phenyl substituent occupies the apical position of the pyramidal coordination sphere instead of a hydroxyl residue (❯ *Fig. 18*).

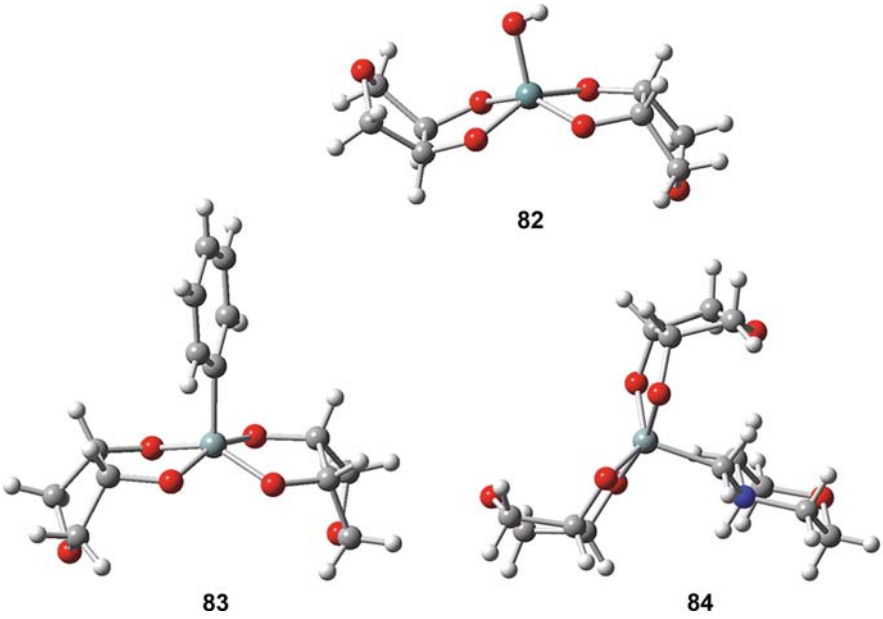


Figure 18
X-ray structures of 82 in $Li[Si(AnErytH_{-2})_2(OH)] \cdot H_2O$, 83 in $K[SiPh(AnErytH_{-2})_2] \cdot 1.5$ MeOH, and 84 in $[(AnErytH_{-2})_2SiCH_2NHC_4H_8O]$

The anionic complex **82** crystallizes with different alkali cations which results in different conformations of the Si center in the crystal structure. In the case of **83**, *anti/anti* and *syn/anti* coordinated silicium centers exist in the crystal structure in equal parts.
The Si-O($AnErytH_{-2}$) distances average 1.70 Å (Li) and 1.73 Å (Na, K, Cs), the $AnErytH_{-2}$ O–C–C–O angles are 14.1 and 9.7° for sodium and two times −3.5° for cesium.
In the case of silicium not only the various possibilities of coordination are of interest, the hydrolytic stability in neutral aqueous solution and the resulting biochemical relevance is of much greater importance. One approach in this direction is the introduction of a positive charge next to the Si center [93].
As can be observed in *Fig. 18*, the Si coordination polyhedra of $[(AnErytH_{-2})_2SiCH_2NHC_4H_8O]$ **84** and additionally $[(AnErytH_{-2})_2SiCH_2NHC_4H_8O] \cdot H_2O$ in the crystal are strongly distorted trigonal bipyramids, in which each bidentate ligand spans one axial (O1, O3) and one equatorial site (O2, O4). The Si–O distances of the SiO_4C skeletons of **84** and **84** $\cdot H_2O$ range from 1.68 to 1.76 Å and the axial Si–O bonds (1.73–1.76 Å are significantly longer than the equatorial ones (1.68–1.71 Å. As seen in *Fig. 18*, the *syn/anti* geometry is observed in the crystals of **84**.
Taking into account the different possible orientations of the two oxolane rings, relative to the apical ligand, for the three compounds **82**, **83** and **84**, the existence of *syn/syn* isomers in addition to *syn/anti* and *anti/anti* isomers must be considered. Although there is some confusion in the literature [92,93,94,95,96,97] about the correct interpretation of the NMR data of silicate complexes, there is no doubt that the three above-mentioned compounds exist in solution. In *Fig. 19*, ^{29}Si and ^{13}C (DEPT-135) NMR spectra of **82** and **83** are shown.

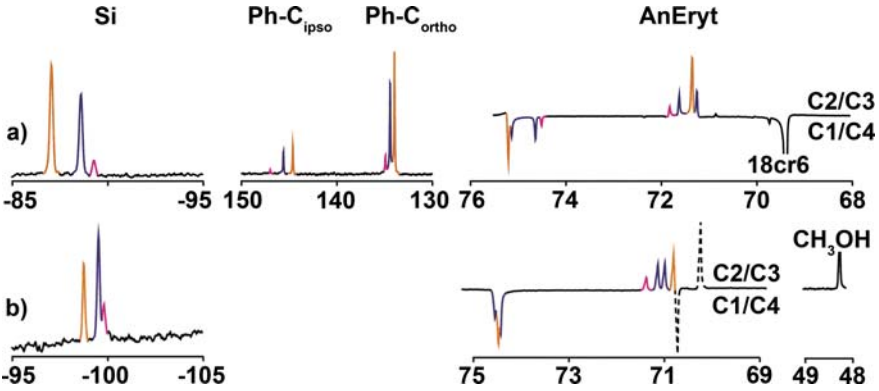


Figure 19
^{29}Si and ^{13}C (DEPT-135) NMR spectra of pentacoordinate AnEryt–silicate species prepared from the diol, $SiPh(OMe)_3$ or $Si(OMe)_4$, and base at the molar ratio given in a) and b). Color code: orange: anti/anti; violet: syn/anti; pink: syn/syn isomer of the respective $[Si(R)(AnErytH_{-2})_2]^-$ ion. a) R = Ph; base: KOMe/[18]crown-6; solvent: methanol; molar ratio: 2:1:2; total Si concentration: 0.38 mol kg^{-1}. AnEryt region in DEPT mode referenced to the [18]crown-6 signal. b) R = OH; base: LiOH; solvent: water; molar ratio: 3:1:1; total Si concentration: 0.54 mol kg^{-1}. The methanol signal stems from hydrolysis of the $Si(OMe)_4$ starting material and may be used as a reference for DEPT assignment. *Dashed line*: free AnEryt [92]

The ^{29}Si NMR spectrum in ❯ *Fig. 19* a) exhibits the signals of two major and one minor pentacoordinate species. According to the species in crystal structure, the two main signals are assigned to the *anti/anti* and*syn/anti* isomers, the minor peak to the *syn/syn* isomer.
In light of this assumption, interpretation of the ^{13}C NMR spectrum is straightforward. The ^{13}C DEPT-135 NMR spectrum is shown in ❯ *Fig. 19* a) on the right side. Both, the *anti/anti* and the *syn/syn* isomer have apparent C_{2v} symmetry and should thus give rise to one signal for C2/C3 and one signal for C1/C4. The *syn/anti* isomer on the other hand, with its apparent C_s symmetry, will show twice the number of signals. A special feature of the phenylsilicate spectra will gain significance in more complicated cases. ❯ *Figure 19* a) shows the ^{13}C NMR signals of the phenyl residue in addition to the diol signals. It should be noted that both the signals stemming from the *ipso*-carbon atom as well as the signals from the two *ortho*-carbon atoms mirror the ^{29}Si NMR spectrum.
The corresponding spectrum of an aqueous diolatosilicate solution is shown in ❯ *Fig. 19* b). The same rules could be applied to the ^{29}Si and ^{13}C NMR spectra of **84** in solution. Here the two main and one minor signals in the ^{29}Si NMR spectrum are found, too [97].
Contrary to prior interpretation [93], hydrolysis products dominate the NMR spectra of aqueous solutions of **84** at a pH value near to neutrality, particularly at a higher dilution, so the introduction of a positive charge next to the Si center does not prevent hydrolysis.
The properties of four-coordinate Si centers bonded to an alkylenedioxy substituent derived from a furanoidic diol also were investigated with the two isomeric oxolane-3,4-diols AnEryt and L-anhydrothreitol [92]. Thus, substitution of silicium with two phenyl residues leads to $Ph_2Si(AnErytH_{-2})$ **85** in the case of AnEryt. The molecular structure is that of a monomer. The five-membered chelate ring is almost planar, with a diol torsion angle close to 0°. Such geometrical parameters cannot be met by L-AnThre. However, this diol provides another example

that the inability to form a chelate ring must not be confused with a lack of reactivity. Thus, the solid-state structure of $Ph_2Si(\text{L-AnThreH}_{-2})$ **86** is not that of a monomeric chelate. Instead, an unstrained molecule with all bonding angles close to their ideal values is observed in a trimeric structure (*Fig. 20*). Although 0° torsion is outside the range of achievable diol torsion angles, the great flexibility of furanoidic rings is obvious from the L-AnThre structure as well. To build the trimer, the diol torsion angles span almost the entire available 100° range of a furanoidic diol, with the actual values being between 82° and 165°.

According to the above-mentioned tetrahedrally coordinated compounds silicium could also build complexes with two AnErytH_{-2} ligands, shown in *Fig. 20*, $Si(\text{AnErytH}_{-2})_2$ **87** [92]. The average Si–O bond length amounts to 1.63 Å and is comparable to the bond length of $[Be(\text{AnErytH}_{-2})_2]^{2-}$, the O–C–C–O torsion angles of the AnErytH_{-2} ligands are −0.8 and −1.2° which differ from the Be compound noticeably. Furthermore there is a dimer form of **87** [92], also shown in **88** in *Fig. 20*. Here the Si–O bond length increases up to 1.72 Å and the O–C–C–O angles to −2.8 and −16.8°.

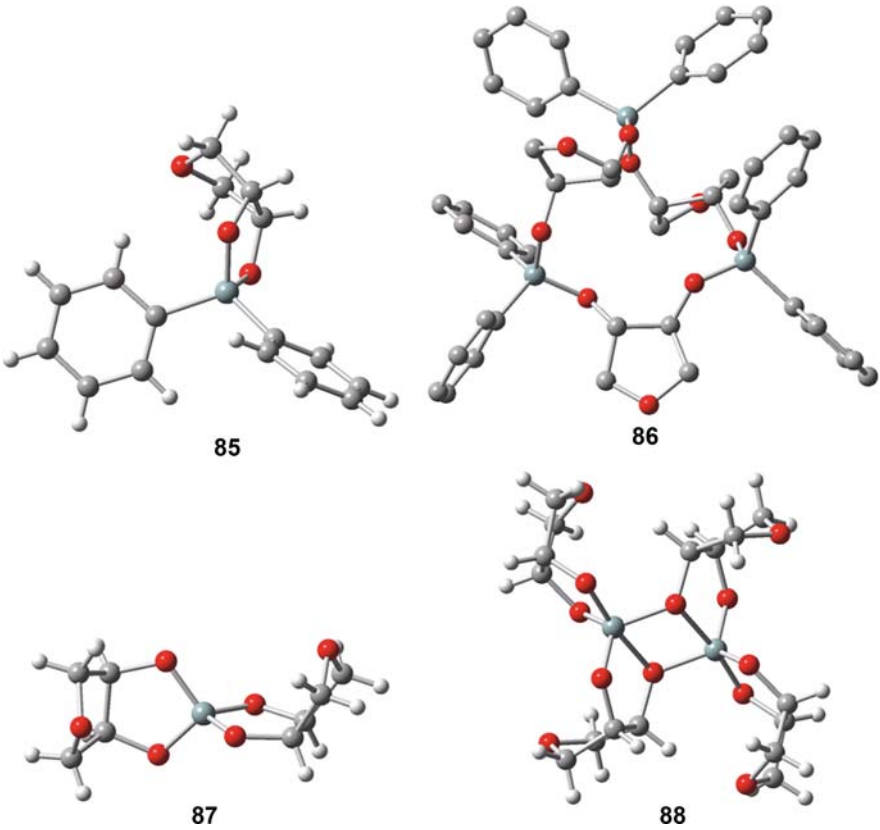


Figure 20
X-ray structures of $Ph_2Si(\text{AnErytH}_{-2})$ 85, $Ph_2Si(\text{L-AnThreH}_{-2})$ 86, $Si(\text{AnErytH}_{-2})_2$ 87, and $[Si(\text{AnErytH}_{-2})_2]_2$ 88

In summary it may be said that AnEryt has a special position compared to the other polyols. The fact that AnEryt models a furanoside and not a pyranoside is of particular importance for dealing with small central atoms. Because of the smaller O–C–C–O torsion angle in the range from 0 to approximately 40°, the steric demand is smaller than that of pyranoses as could be shown for cobalt(III) complexes [80]. A *cis*-furanoidic diol thus appears to be a better choice than open-chain diols and, particularly, pyranoidic diol functions. The reason for the conformational flexibility of a furanose ring is its balance of strain. While open-chain diols experience Pitzer strain when twisted into an eclipsed conformation, and pyranoses are burdened with ring strain on being twisted towards 0° torsion, a furanose is characterized by a balance of the various types of strain over a pronounced range of torsions.

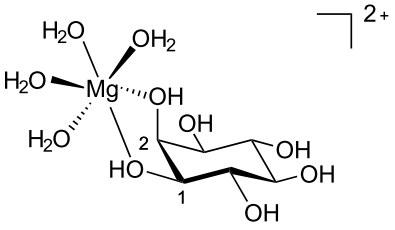
5 Inositols and Anhydro-Sugars

5.1 Inositols

There are nine various inositols or cyclohexane-1,2,3,4,5,6-hexaols, including one pair of isomers. The *myo*-inositol is the most widespread isomer of inositol and, excluding *allo*-, *cis*-, *epi*- and *muco*-inositol, all other inositols are of natural occurrence.

Myo-inositol is also the inositol, which is most investigated in connection with metal coordination. First of all, there are three known structures with magnesium, praseodymium and neodymium with the specific feature that the metal atoms coordinate to *myo*-inositol without deprotonation of the affected hydroxyl groups.

The magnesium atom in $[Mg(myo\text{-Ins})(H_2O)_4]Cl_2$ (**89**Cl_2) [98] is coordinated to two *cis*-vicinal hydroxyl groups and four water molecules in an octahedral geometry. As a consequence of cation binding, the cyclohexane ring geometry in the vicinity of the coordination site is distorted. Deviations of nearly 15° from the norm occur in some torsion angles.



89

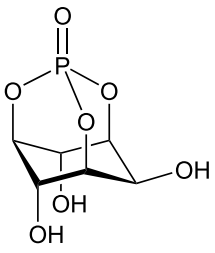
The intramolecular $O_{eq}1$ to $O_{ax}2$ distance is shortened 0.33 Å and that for $O_{eq}3$ to $O_{ax}2$, 0.11 Å from a normal value of 2.89 Å [99]. The decrease in the C–C–O angles for the C1–C2 bond could be an effect of the magnesium ion attraction to the hydroxyl groups. The torsion angles for bond C1–C2 differ by as much as 14° from those of *myo*-inositol.

Praseodymium and neodymium built almost identical structures with *myo*-inositol. In the crystal structures of $[Pr(myo\text{-Ins})]Cl_3 \cdot 9\ H_2O$ [100] and $[Nd(myo\text{-Ins})]Cl_3 \cdot 9\ H_2O$ [101], by trivalent lanthanide cations complexed *myo*-inositol, each Pr or Nb is coordinated to nine oxygen atoms, two from the inositol (two adjacent hydroxyl groups) and seven from water molecules

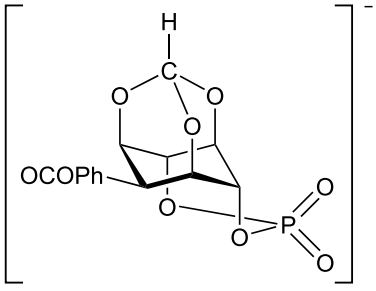
in a tricapped trigonal prism geometry, with Pr–O distances ranging from 2.47 to 2.69 Å and Nb–O distances ranging from 2.45 to 2.68 Å. The other two water molecules are hydrogen-bonded. No direct contacts exist between Pr/Nd and Cl.

These metal–hydroxyl interactions generally produce small conformational changes in the sugar at the metal-binding sites. Comparing uncomplexed *myo*-inositol with the Nd-inositol complex, the C–C, C–O distances and bond angles have only small changes.

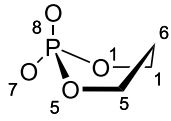
In the complex of [(PO)(*myo*-InsH$_{-3}$)] **90** [102], the cyclohexane ring adopts the expected chair conformation, with the 2-, 4- and 6-hydroxyl groups being axial, equatorial and axial, respectively. There is an intramolecular hydrogen bond between the O6 atom as donor and O5 as acceptor. The three P–O single bonds, of average length 1.58 Å are all equivalent within significant error, as are the P–O–P bond angles with an average of 104.5°.



90



91a



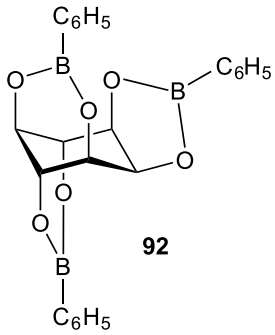
91b

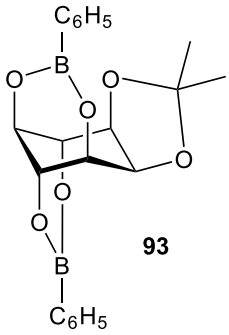
The six-membered 1,3,2-dioxaphosphorinane ring in the anionic 2-*O*-benzoyl-1,3,5-*O*-methylidyne-*myo*-inositol-4,6-cyclophosphate **91a** [103], however, has a boat rather than a chair conformation. The P–O bond distances involving the inositol residue are normal but longer than the mean value of 1.58 Å found by Neidle et al. [102] in **90**. Conspicuously, in **91b**, the six-membered 1,3,2-dioxaphosphorinane ring adopts a boat conformation, with P and C6 on the same side at distances of 0.61 and 0.68 Å respectively, from the mean plane through C1, C5, O1 and O5. This is different from the situation found in **90**, where all three six-membered 1,3,2-dioxaphosphorinane rings have chair conformations.

Furthermore, from *myo*-inositol and 1,2-*O*-isopropylidene-*myo*-inositol two tetracyclic phenylboronic esters [$B_3(C_6H_5)_3$(*myo*-InsH$_{-6}$)] **92** and [$B_2(C_6H_5)_2(C_3H_6)$(*myo*-InsH$_{-6}$)] **93** have been prepared [104], whereas **92** was determined by crystal structure, the structure of **93** was proved by ^{11}B and ^{13}C NMR spectroscopy. Compound **92** is a tetracyclic derivative of the less stable conformer of inositol (five axial hydroxyl groups and one equatorial) with two dioxaboroline rings at opposite faces of the six-membered ring and a dioxaborolidine ring bridging the C1 and C2 atoms at axial and equatorial positions. A similar structure was found for **93** with the difference that, bridging C1 and C2, there is a dioxolane ring. The boron atoms are planar with their attached atoms and stabilized by retrocoordination between the boron and oxygen and carbon atoms, respectively. The two phenyl rings that are in the same face of the molecule are essentially parallel, with a dihedral angle between planes of 28.3 ± 0.8°.

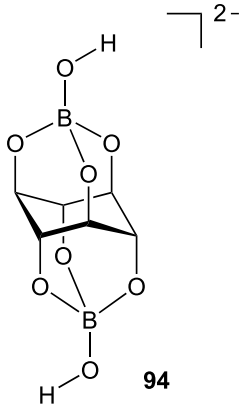
With *scyllo*-inositol boron builds a complex $Na_2[B_2(OH)_2$(*scyllo*-InsH$_{-6}$)] · 9 H_2O (Na_2**94** · 9 H_2O) [105], where an all-*cis*-*scyllo*-inositol builds a hexaoxadiamantane-like structure with two boron centers. The *scyllo*-inositol diborate anion itself comprises a rigid cage structure,

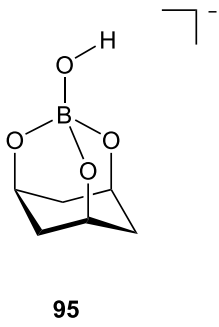
the C atoms in the inositol ring are in a chair conformation with normal distances and angles. The two B atoms, situated on either side of the ring, are each bonded axially to alternate ring C atoms via three O atoms with B–O bond distances of 1.48 to 1.51 Å and C–O distances of 1.42 to 1.45 Å. Completing the tetrahedral coordination, each B atom is bonded also to an OH at 1.42 Å.

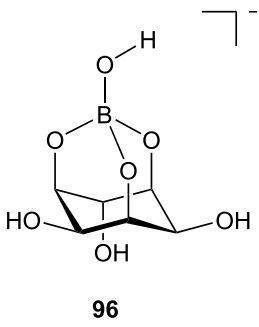




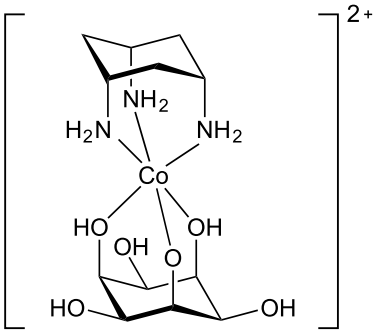
Corresponding to structure **94**, in aqueous alkaline solutions the existence of the two anionic tridentate borate esters [B(OH)(all-*cis*-1,3,5-triol-chxnH$_{-3}$)]$^-$ **95** and [B(OH)(*epi*-InsH$_{-3}$)]$^-$ **96** is proved by ^{11}B NMR spectroscopy [40]. The borate ester of all-*cis*-cyclohexane-1,3,5-triol has a chemical shift of −18.1 ppm, the ester of *epi*-inositol a chemical shift of −19.4 ppm.



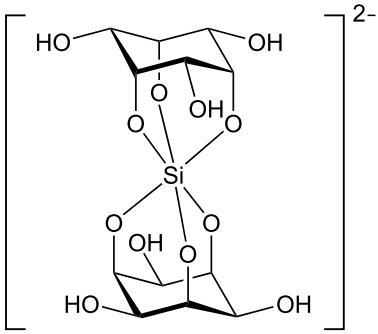




In the structure of [Co(tach)(*cis*-InsH$_{-1}$)](NO$_3$)$_2$ (**97**(NO$_3$)$_2$) [106], cobalt(III) is bound to the three amino groups of all-*cis*-1,3,5-triaminocyclohexane and to three axial oxygen donors of *cis*-inositol. The structure shows that only one of the coordinated hydroxyl groups of *cis*-inositol is deprotonated. Consequently, the Co–O$_H$ distance of 1.90 Å is significantly shorter than the corresponding Co–OH distances of 1.94 Å. Moreover, the C–O bond lengths of the coordinated alcoholic groups are slightly longer than the corresponding bond distance of the coordinated alkoxido group. In turn, the Co–N distance, which is *trans* to the alkoxido group, is significantly elongated, indicating a stronger *trans* influence of RO$^-$ compared with ROH.



97



98

Without tach as auxiliary ligand $Cs_2[Si(cis\text{-}InsH_{-3})_2]$ (Cs_2**98**) [92] shows a bis-tridentate 1,3-diolate binding-*cis*-inositol complex. In contrast to the finding that a pyranoidic 1,2,3-triol like $^1C_4$-Me-β-D-Rib*p*, with its large torsional angles, is not so efficient an Si-chelator as a furanoidic diol, the significant enrichment of aqueous alkaline silicate solutions by hexacoordinate species on addition of *cis*-inositol is remarkable.

Besides these, to one metal atom coordinated *cis*-inositol complexes, there are two related fields of research. One is the complex coordination of metal atoms to the polyaminoalcohols, especially 1,3,5-triamino-1,3,5-trideoxy-*cis*-inositol, since the amino group can either act as additional donors for metal binding or serve as internal bases, facilitating the deprotonation and coordination of the hydroxyl groups [107].

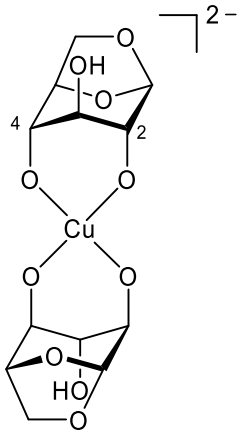
The other field of research deals with *cis*-inositol-coated polyoxometalate clusters [108,109, 110]. $[OFe_6\{(cis\text{-}Ins)_6\text{-}21H\}]^{5-}$ and $[Ta_7O_{12}(cis\text{-}InsH_{-3})_6]^{7-}$ build polyoxometalate cores which are coordinated to *cis*-inositol on the outer binding sites. In the case of $[OFe_6\{(cis\text{-}Ins)_6H_{-21}\}]^{5-}$ there are five different modes of coordination of *cis*-inostiol to the Fe atoms, so that there is no preferred coordination mode recognizable. In the case of $[Ta_7O_{12}(cis\text{-}InsH_{-3})_6]^{7-}$, *cis*-inositol binds over the three *cis*-standing hydroxyl groups to the Ta atoms.

5.2 1,6-Anhydro-β-D-Glucose (Levoglucosan)

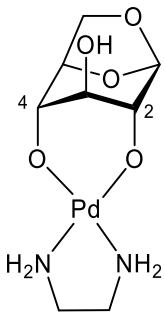
Currently there are three known structures of metal-coordinated 1,6-anhydro-β-D-glucose (Glc1,6An; levoglucosan). The cuprate anion of $Li_2[Cu(Glc1,6An2,4H_{-2})_2] \cdot 8\ H_2O$ is a homoleptic mononuclear 1,3-polyolato(2-) complex.

Generally, in terms of flexibility a chelate ligand is of limited variability, particularly when the ligator atoms are attached to a further cyclic fragment and above all, to a six-membered ring in its energetically favored chair conformation. In the coordination chemistry of carbohydrates, this general statement may be quantified for 1,2-diol fragments incorporated in pyranose rings. In cuprate(II) or cobalt(II) complexes of β-D-galactose, α-D-mannose and β-D-xylose, the mean O–O distance of the ligating diol groups is 2.67 $\pm$ 0.01 Å. This corresponds to a 0.19 Å reduction of the respective mean distance of 2.85 $\pm$ 0.02 Å in the uncomplexed methyl glycopyranosides.

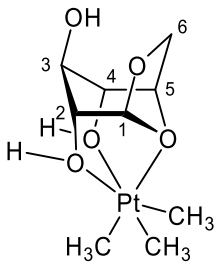
Assuming that a bicyclic pyranoidic carbohydrate will also be of restricted flexibility, the twofold reduction of the ligator atom distance (0.38 Å on coordination of cupric ions by 1,6-anhydro-β-D-glucose, compared with the above-mentioned 1,2-diolates, was unexpected due to the large O–O separation of 3.30 Å in the free anhydro sugar [111].



99



100



101

In the mononuclear cuprate(II) anion **99** (4 + 2)-coordinated Cu^{II} forms six-membered rings with doubly deprotonated levoglucosan ligands at a shorter Cu–O distance (Cu–O2, 1.95 Å Cu–O4, 1.99 Å. Two longer contacts are established with the O5 atoms of the pyranose rings. The chelate bonding mode is enabled by the astonishingly enhanced flexibility of the bicyclic ligand compared with simple pyranoses [112].

[(en)Pd(Glc1,6An2,4H_{-2})] · H_2O (**100** · H_2O) is formed in a reaction from Pd-en and 1,6-anhydro-β-D-glucose [84]. Levoglucosan again can act as a 1,3-diolato ligand by bonding through O2 and O4. The distances between Pd–O2 and Pd–O4 amount to 2.04 Å and 2.05 Å the angle between O2–Pd–O4 to 97.0°. In fact, **100** is a six-membered-ring chelate complex and the molecular structure of **100** again shows flexibility of the bicyclic ligand, which undergoes distortion of the pyranose chair towards a boat conformation. In this distortion, greater flattening of the pyranose ring at C3 corresponds to an increasing O2–O4 distance. Conversely, in **100** the respective hydroxyl groups occupy a more axial position, and the 'bite' is smaller than in the free anhydro sugar. This can be compared with the even smaller O–O distance in Li_2**99** · 4 H_2O.

In [$PtMe_3$(Glc1,6An)]BF_4 (**101**BF_4) 1,6-anhydro-β-D-glucose acts as a neutral tridentate ligand which is coordinated through two hydroxyl groups (O2, O4) and an acetal oxygen atom (O5) of the glucopyranose ring. Thus, two five-membered rings and one six-membered 1,3,2-dioxaplatinacyclohexane ring are formed exhibiting a distorted half-chair and chair conformation, respectively. The cyclic system is not free from angle strain which is revealed particularly by the O–Pt–O angles. Two of them are distinctly smaller than 90° (O2–Pt–O5, 73.6°; O4–Pt–O5, 75.6°), whereas the C–Pt–C angles remain nearly orthogonal (88.1–89.9°). The two Pt–O bonds to the hydroxyl groups are almost equal (Pt–O2/O4, 2.25 Å. Contrarily, the Pt–O bond to the acetal oxygen is significantly longer (Pt–O5, 2.29 Å. This may be

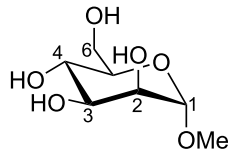
explained in terms of the very low donor capability of the acetal oxygen atom and of the high *trans* influence of the methyl ligand [113].

6 Glycosides

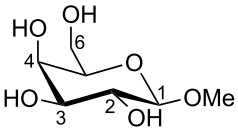
Glycosides are obtained by conversion of the cyclic semiacetal form of aldoses and ketoses to an acetal. We will focus on the metal complexes and esters of the following glycosides: alkyl pyranosides, alkyl furanosides, non-reducing disaccharides, polysaccharides, and finally the nucleosides as a member of the group of *N*-glycosides.

6.1 Pyranosides

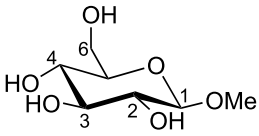
In the following, solid-state and solution structures of the pyranosides **102**–**107** and their protected derivatives will be described.



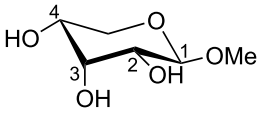
102



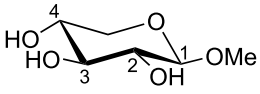
103



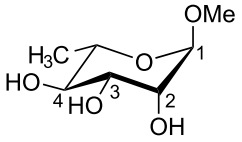
104



105



106



107

Methyl-α-D-mannopyranoside (Me-α-D-Man*p*) **102** is one of the structurally best-investigated compounds among the pyranosides. The reaction of equimolar amounts of $Cu(OH)_2$, ethylenediamine, and **102** in aqueous solution leads to a blue solution and crystals of the heteroleptic copper(II) complex [(en)Cu(Me-α-D-Man*p*2,3H$_{-2}$)] · 2 H_2O. The crystal-structure determination reveals that the copper center is coordinated in a square planar geometry (compare **37**, **39**, **72**) with two of the four coordination sites occupied by the *cis*-$O2_{ax}$,$O3_{eq}$ diolato moiety of **102** in a $^4C_1$-conformation (ax = axial, eq = equatorial) [20]. The homoleptic cuprate **108** (❯ *Fig. 21*) is obtained by a reaction of $Cu(OH)_2$ with the double-molar amount of **102** in strong alkaline solution. The coordination pattern of the carbohydrate is similar to that observed in the just described heteroleptic complex: both pyranosides are in a $^4C_1$-conformation and act as *cis*-$O2_{ax}$,$O3_{eq}$ diolato ligands. The preference of the *cis*-$O2_{ax}$,$O3_{eq}$ chelation site to the *trans*-$O3_{eq}$,$O4_{eq}$ site may be explained by the acidity of the hydroxyl functions (the acidity decreases in the sequence O2-H > O3-H > O4-H) as well as by the difference of the corresponding O–C–C–O torsion angles in **108** and in free methyl-α-D-mannopyranoside, respectively. The torsion angles of the chelating diolato moieties in **108** are 47.5° and 47.8°, respectively, which is about 8° less than the $O2_{ax}$–C–C–$O3_{eq}$

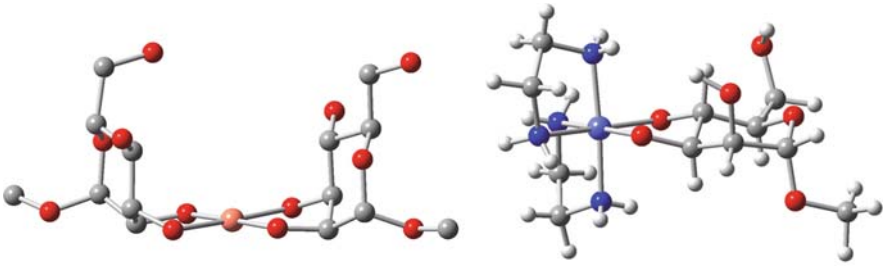


Figure 21
The molecular structures of [Cu(Me-α-D-Man*p*2,3H$_{-2}$)(Me-α-D-Man*p*2,3,4H$_{-3}$)]$^{3-}$ 108 in crystals of $Li_3$108 · 5 H_2O and Δ-[(en)$_2$Co(Me-α-D-Man*p*3,4H$_{-2}$)]$^+$ 109 in crystals of 109ClO_4 · $NaClO_4$ · 2 H_2O (hydrogen atoms omitted for clarity in 108)

torsion angle and about 20° less than the $O3_{eq}$–C–C–$O4_{eq}$ torsion angle in free $^4C_1$-methyl-α-D-mannopyranoside [114]. Obviously the diol residue causing less conformative strain of the pyranoside is chosen as the chelation site [85].

Besides the conformative strain and acidity of the hydroxyl functions, however, steric factors may also influence the selection of a chelation site. In the complex cation **109** (● *Fig. 21*), the cobalt(III) center is coordinated by the *trans*-$O3_{eq}$,$O4_{eq}$ diolato moiety of **102**. The latter forms a torsion angle of 53.9° which is just about 1.8° less than the *cis*-$O2_{ax}$,$O3_{eq}$ torsion angle of free $^4C_1$-methyl-α-D-mannopyranoside. Thus only little deformation of the pyranoside would be necessary with a *cis*-$O2_{ax}$,$O3_{eq}$ chelation site. However, a comparison of the structures **108** and **109** (● *Fig. 21*) reveals that the latter site is unfavorable in **109** for steric reasons. The involvement of an axially bonded oxygen atom in a five-membered chelate ring results in an approximately perpendicular orientation of the planes of the chelate ring and the pyranose ring (compare **108**), while the planes are coplanar with two equatorially bonded oxygen ligator atoms (compare **109**). The perpendicular orientation is hampered in **109** by amino groups of ethylenediamine below and above the CoO_2N_2 plane, while it is enabled in **108** in which the coordination sites below and above the copper center are not occupied [80].

A three-dimensional, triply connected network is established in **110** (● *Fig. 22*) with tridentate $^4C_1$-methyl-α-D-mannopyranoside **102** and bismuth(III) (the figure shows the symmetrically independent bismuth atoms and mannopyranosides). Each of the carbohydrates acts as triolato ligand and is bonded in five-membered chelate rings formed by the *cis*-$O2_{ax}$,$O3_{eq}$ as well as by the *trans*-$O3_{eq}$,$O4_{eq}$ diolato moieties to two bismuth centers with O3 in the μ_2 mode. The torsion angles of the *cis*-diols (48.3°, 49.4°) are smaller than those of the *trans*-diols (55.4°, 57.4°) as is found in **108** (● *Fig. 21*) and in free methyl-α-D-mannopyranoside. The two independent bismuth atoms bonded to the *trans*-$O3_{eq}$,$O4_{eq}$ diolato moieties are located on sites which possess site symmetry 3 and are the branching points of the network [115].

Moreover, several vanadate chelate esters with O4- and O6-protected derivatives of $^4C_1$-methyl-α-D-mannopyranoside have been structurally characterized. In all these esters the remaining unprotected hydroxyl functions of the *cis*-$O2_{ax}$,$O3_{eq}$ diol residue act as the chelation site. The dinuclear vanadate chelate ester **111** (● *Fig. 22*) is obtained with methyl-4,6-*O*-benzylidene-α-D-mannopyranoside. The two VO_2^{2+}-centers are bridged by two O2-alkoxido

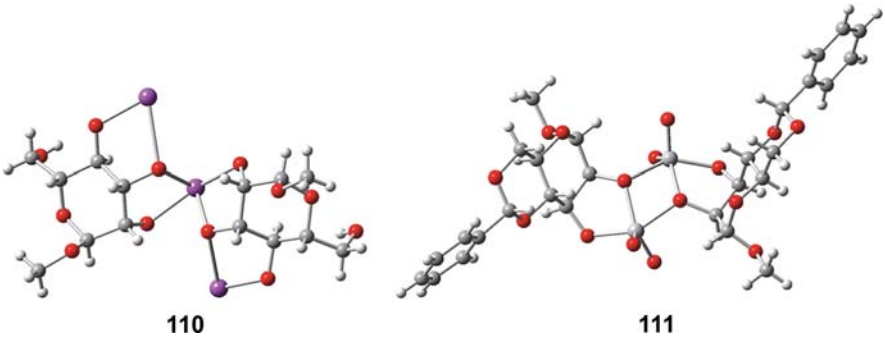


Figure 22
The molecular structures of $[Bi_{5/3}(Me\text{-}\alpha\text{-}D\text{-}Manp2,3,4H_{-3})_2]^-$ 110 in crystals of $Na_2$110(OH) · 12 H_2O and $[\{VO_2(Me\text{-}4,6\text{-}O\text{-}benzylidene\text{-}\alpha\text{-}D\text{-}Manp2,3H_{-2})\}_2]^{2-}$ 111 in crystals of $[NBu_4]_2$111

functions and have a distorted trigonal-bipyramidal geometry. As a consequence of the small crystal radius of 0.60 Å of a penta-coordinated vanadium(V) [116] the torsion angles of the chelating *cis*-$O2_{ax}$,$O3_{eq}$ diolato functions are just 39.5° and 40.1° with O2–O3 distances of 2.49 Å and 2.50 Å respectively [117]. Only one of the two hydroxyl functions of the chelating *cis*-$O2_{ax}$,$O3_{eq}$ diol residue is deprotonated in the vanadate esters of methyl-4,6-di-*O*-methyl-α-D-mannopyranoside [118,119,120], and methyl-4,6-*O*-benzylidene-α-D-mannopyranoside [121], namely the one located in the *trans*-position to the *N*-ligator atom of the auxiliary ligands bonded tridentately to the vanadium(V) centers. A comprehensive review concerning this type of vanadate chelate esters has been written by Chakravorty et al. [11].

Methyl-β-D-galactopyranoside (Me-β-D-Gal*p*) **103** in its $^4C_1$-conformation comprises two adjacent diol functions: a *trans*-$O2_{eq}$,$O3_{eq}$ diol and a *cis*-$O3_{eq}$,$O4_{ax}$ diol. Analogously to methyl-α-D-mannopyranoside in **109** (*Fig. 21*), methyl-β-D-galactopyranoside acts as chelating ligand in a bis(ethylenediamine)cobalt(III) complex through its *trans*-$O2_{eq}$,$O3_{eq}$ diolato moiety for sterical reasons [80]. A torsion angle of just 32.3° (55.6° in free Me-β-D-Gal*p* [122]) is formed by the chelating *cis*-$O3_{eq}$,$O4_{ax}$ diolato group of methyl-β-D-galactopyranoside in the rhenium(V) complex **112** (*Fig. 23*) while the torsion angle of the *trans*-$O2_{eq}$,$O3_{eq}$ diol is enlarged to 75.7° (62.9 in free Me-β-D-Gal*p*). A comparably strained pyranoside is present in methyl 3,4-*O*-isopropylidene-β-D-Gal*p* [123]. NMR-experiments reveal that the oxolane ring is oriented towards the rhenium-bonded oxo group (*syn*-isomer) in solution as well, while the *anti*-isomer is detectable as intermediate only [82].

The *cis*-$O3_{eq}$,$O4_{ax}$ diol group of 2,6-*O*-protected derivatives of methyl-β-D-galactopyranoside is also reported to act as a chelation site in a vanadate ester [120], and in the phosphinite **113** (*Fig. 23*) which incorporates a five-membered dioxophospholane ring [124].

Glucopyranosides include two almost equivalent *trans*-configured diol functions: the $O2_{eq}$, $O3_{eq}$ site and the $O3_{eq}$,$O4_{eq}$ site. This may cause the formation of mixtures of metal derivatives and be the reason for very limited structural information on metal complexes of unprotected glucopyranosides. Methyl-β-D-glucopyranoside (Me-β-D-Glc*p*) **104** is found to be coordinated to a (tren)Ni^{2+}-center (tren = tris(2-aminoethylamine)) by its $O3_{eq}$,$O4_{eq}$ diolato moiety in the solid state [125]. More structures are characterized for 4,6-*O*-protected glucopyranosides since their complexation behavior is more calculable. The dinuclear titanium(IV)

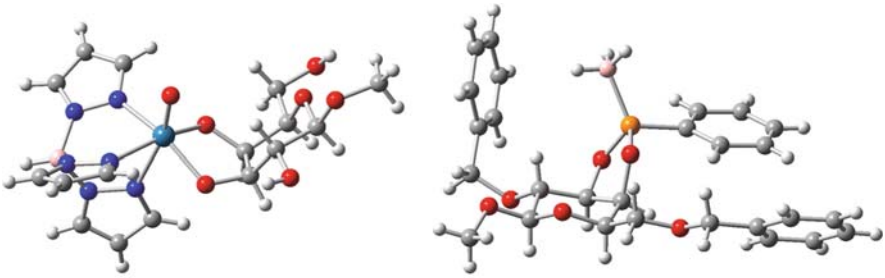


Figure 23
The crystal structures of [(tpb)ReO(Me-β-D-Gal*p*3,4H$_{-2}$)] 112 and methyl 2,6-di-*O*-benzyl-3,4-*O*-phenylphosphinediyl-β-D-Gal*p* (P–B)borane 113

complex **114** (*Fig. 24*) is formed with the 4,6-*O*-benzylidene derivative of **104** (a complex similar to **114** is obtained with the 4,6-*O*-naphthylmethylidene derivative of **104**), in which the latter acts as a bidentate but not as a chelating ligand [126]. Obviously the titanium(IV) center (crystal radius 0.56 Å [116]) is too small to be chelated by a *trans*-configured diol of a (gluco)pyranoside, while the next-in-size homolog zirconium(IV) (crystal radius for hexacoordinated Zr^{IV}: 0.86 Å [116]) is chelated in a dinuclear complex with the carbohydrate used in **114** [127]. Pentacoordinated tin(IV) (crystal radius 0.76 Å [116]) is of a size suitable to be chelated by such a ligand as well. In the dinuclear complex **115** (*Fig. 24*) which has a constitution comparable to **111** (*Fig. 22*), the *trans*-O2$_{eq}$,O3$_{eq}$ diolato chelation site is coordinated to a (n-Bu)$_2$Sn^{2+} center while a second (n-Bu)$_2$Sn^{2+} center is bonded by the bridging O3-alkoxido function [128]. NMR studies reveal that the chelation of pentacoordinated vanadium(V) (crystal radius 0.60 Å [116]) with the *cis*-configured O2$_{ax}$,O3$_{eq}$ diolato site of mannopyranoside is preferred over the *trans*-configured O2$_{eq}$,O3$_{eq}$ diolato site of glucopyranoside [117].

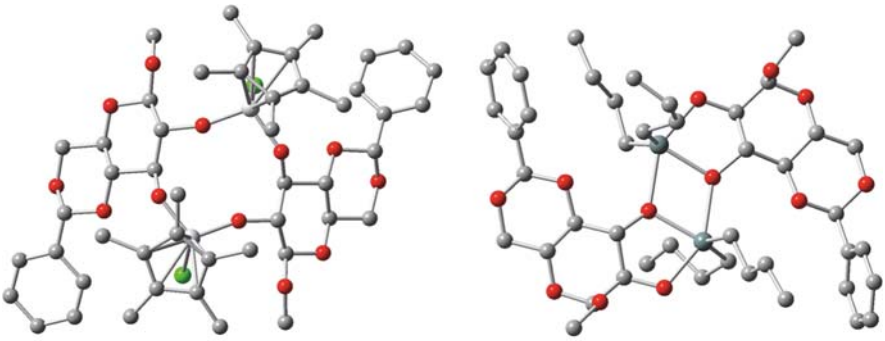


Figure 24
The crystal structures of [(Cp*TiCl)-μ-(methyl 4,6-*O*-benzylidene-β-D-Glc*p*2,3H$_{-2}$)]$_2$ 114 and [(Sn(n-Bu)$_2$)-μ-(methyl 4,6-*O*-benzylidene-α-D-Glc*p*2,3H$_{-2}$)]$_2$ 115 (hydrogen atoms omitted for clarity)

A homoleptic cuprate similar to **108** (◉ *Fig. 21*) is formed with the $^{1}C_{4}$-conformer of methyl α-L-rhamnopyranoside **107** with the *cis*-$O2_{ax}$,$O3_{eq}$ diolato moiety acting as the chelation site [129]. However, chelation with the *trans*-$O2_{eq}$,$O3_{eq}$ diolato moiety of methyl-β-D-xylopyranoside (Me-β-D-Xyl*p*) **106** is observed in the heteroleptic copper(II) complex **116** (◉ *Fig. 25*), in which a water molecule is bonded to the now pentacoordinated copper center with a distance of 2.26 Å [130]. According to Shannon, the crystal radii of penta- and tetracoordinated (in a square planar geometry) copper(II) are 0.79 Å and 0.71 Å respectively [116]. A transformation of the more stable $^{4}C_{1}$-conformer of free methyl-β-D-ribopyranoside (Me-β-D-Rib*p*) **105** to the $^{1}C_{4}$-conformer is observed in the trinuclear rhenate **117** (◉ *Fig. 25*). A suitable arrangement of oxygen ligator atoms is provided by the all-*cis*-$O2_{ax}$,$O3_{eq}$,$O4_{ax}$ triolato moiety which is bonded to the Re_3-core with all alkoxido functions in the μ_2 mode [32].

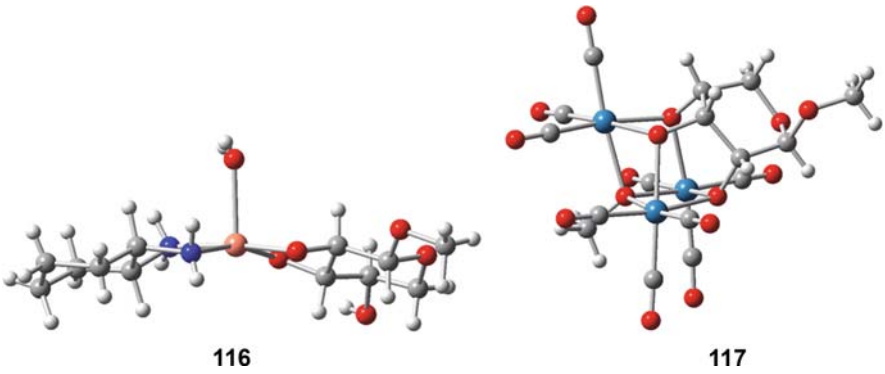
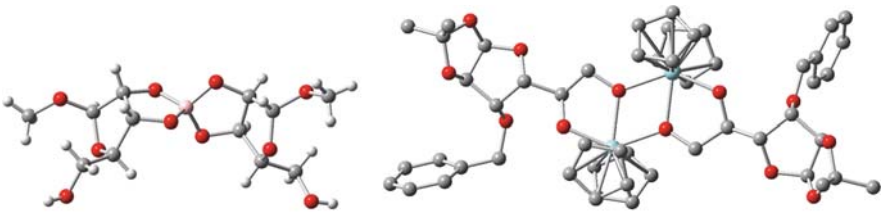


Figure 25
The molecular structures of [(R,R-chxn)(H_2O)Cu(Me-β-D-Xyl*p*2,3H$_{-2}$)] 116 in crystals of 116 · H_2O and [Re_3(CO)$_9$-(μ_3-OMe)(μ_3-$^{1}C_{4}$-Me-β-D-Rib*p*2,3,4H$_{-3}$)]$^{-}$ 117 in crystals of (DBUH)117

6.2 Furanosides

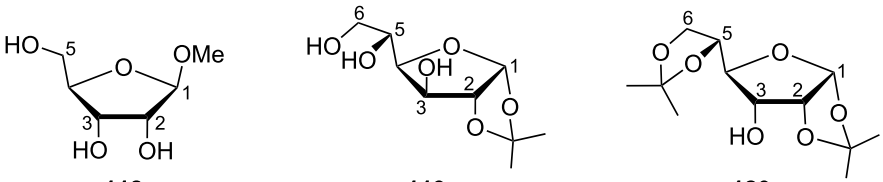
Methyl-β-D-ribofuranoside (Me-β-D-Rib*f*) **118** comprises a *cis*-O2,O3 diol group and often shows a complexation behavior like its non-chiral analog anhydroerythritol, the complexes of which have been described above. A bis(diolato) borate similar to **121** (◉ *Fig. 26*), for example, is also formed with anhydroerythritol. The boron center is tetrahedrally coordinated by two chelating *cis*-O2,O3 diolato groups which form torsion angles of 12.8° and 9.2° with O2–O3 distances of 2.32 Å and 2.30 Å respectively. These rather small angles and distances are accessible by furanoidic diol groups but not by pyranoidic diol groups as a consequence of the small boron atom. Solutions with a molar ratio of B:**118** of 1:2 contain two diastereomeric forms of **121** (◉ *Fig. 26*) as well as monodiolato-dihydroxo borate [87]. In the ruthenium complex cation [(*mer*-dien)(NO)Ru(Me-β-D-Rib*f*2,3H$_{-2}$)]$^{+}$, the O3 and O2 ligator atoms of the diolato moiety occupy two equatorial sites trans to NO and the central nitrogen atom of dien, respectively. The *trans*-influence of the NO group results in a Ru–O distance which is



121 122

Figure 26
The molecular structures of [B(Me-β-D-Ribf2,3H$_{-2}$)]$^-$ 121 in crystals of Na121 · 2 H_2O and [{Zr(Cp)$_2$(3-*O*-benzyl-1,2-*O*-isopropylidene-α-D-Glcf5,6H$_{-2}$)}$_2$] 122 in crystals of 122 · C_7H_8. The hydrogen atoms have been omitted for clarity in 122

about 0.15 Å shorter than the other Ru–O distance. Two species have been identified in the mother liquor, both showing a ^{13}C NMR downfield shift of about 15 ppm for the carbon atoms of the chelating diolato residue [81]. Three isomers already described for silicate complexes of anhydroerythritol have also been identified in solutions of [SiR(Me-β-D-Ribf2,3H$_{-2}$)2]$^-$ (R = Ph, OMe): *anti/anti*, *syn/anti*, and *syn/syn*. A downfield shift of 1.4–2.5 ppm is observed for the carbon atoms of the chelating diolato moiety [92].

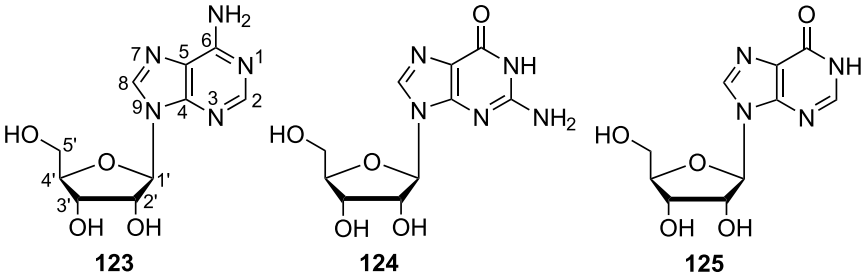


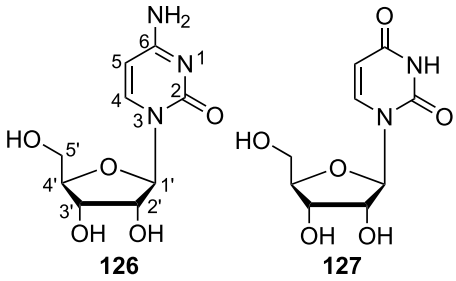
118 119 120

The 5-methyl-protected derivative of **118** has been found to act as a chelating ligand in vanadate esters in which only one of the two coordinating hydroxyl groups of the *cis*-diol moiety is deprotonated [119,120]. 1,2-*O*-isopropylidene-α-D-glucofuranose **119** is bonded to a $Ru_3(CO)_8^{2-}$ residue as a tetradentate ligand, namely through the hydroxyl function at C6, two μ_2-alkoxido functions at C3 and C5, respectively, and the oxygen atom of the furanose ring [131]. The three hydroxyl oxygen atoms of **119** act as facially coordinating sites in an octahedral platinum(IV) complex, which is, interestingly, also obtained when the 5,6-*O*-isopropylidene derivative of **119** is used as the starting material. The same complexation behavior is shown by 1,2:5,6-di-*O*-isopropylidene-α-D-allofuranose **120** [132,133]. These and many more furanoside and other carbohydrate complexes with platinum-group metals have been described already in a comprehensive review written by Steinborn and Junicke [7].
The two remaining hydroxyl functions at C5 and C6 of the 3-*O*-benzyl derivative of **119** are deprotonated and act as the chelation site in the dinuclear zirconium(IV) complex **122** (*Fig. 26*) in which the less-stressed terminal alkoxido functions are in a bridging mode in the solid state as well as in solution [134].

6.3 Nucleosides

An *N*-glycosidic linkage between β-D-ribose and the nucleobases adenine, guanine, hypoxanthine, cytosine, and uracile leads to the ribonucleosides adenosine (ado) **123**, guanosine (guo) **124**, inosine (ino) **125**, cytidine (cyd) **126**, and uridine (urd) **127**, respectively. In the following, complexes and esters using the chelating sites of the ribose moiety are discussed.





In the 1970s structural investigations of osmium(VI) and uranium(VI) complexes with adenosine **123** and its 5′-phosphate derivatives emphasized the chelation ability of the furanoidic *cis*-diol group. In the crystal structure of the mononuclear heteroleptic complex $[OsO_2(py)_2(Ado2',3'H_{-2})]$ (py = pyridine), adenosine acts as diolato ligand with its ligator atoms O2′ and O3′ being *trans* to the nitrogen atoms of the two coordinating pyridine molecules. The reported π stacking of the adenine bases is a structural feature typical for solid-state nucleoside complexes [135]. With 2,2′-bipyridyl instead of the two pyridine molecules, osmium(VI) complexes similar to that described above are also obtained in neutral aqueous solution with 5′-phosphate derivatives of adenosine **123** and uridine **127**, respectively [136]. Di- and tetranuclear UO^{2+} complexes are formed at pH 11 with adenosine-5′-monophosphate with O3′ of the chelating diolato moiety bridging between two uranium centers (compare dinuclear VO^{2+} complex **128**, ❯ *Fig. 27*) while at pH 7.5 the uranium centers are coordinated by phosphate oxygen atoms only [137]. The same behavior is also shown by adenosine-5′-di- and -triphosphate with the exception that no dinuclear species is observed with the diphosphate [138]. The dimeric dinuclear solid-state structure of the vanadium(V)-adenosine complex anion **128** (❯ *Fig. 27*) has been found to be maintained as the main species besides several minor products in a solution of pH 6.5 [139]. The dimeric structure has been suggested by most of the groups working on the investigation of the solution structures of vanadium

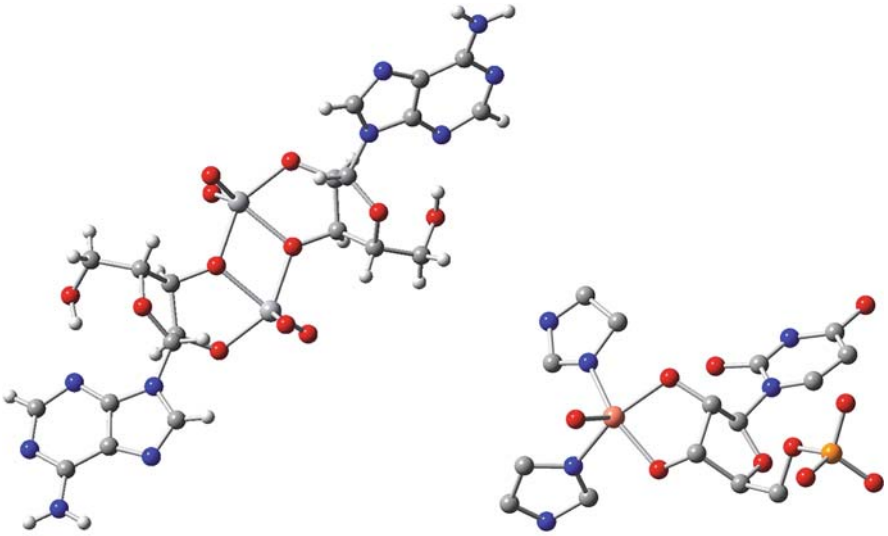


Figure 27
The molecular structures of $[\{VO_2(Ado2',3'H_{-2})\}_2]^{2-}$ 128 in crystals of $[NEt_4]_2$128 · 4.74 H_2O and [Cu(5′-UMP2′,3′H$_{-2}$)(im)$_2$(H_2O)] 129 in crystals of 129 · 4 H_2O (hydrogen atoms omitted for clarity). Abbreviations: UMP = uridine monophosphate, im = imidazole

complexes with the nucleosides **123**–**127** as well as their 5′-derivatives in the years before the detection of **128** (◉ *Fig. 27*) [117,140,141,142,143,144,145,146].
Dinuclear subunits similar to **128** are also found in the octanuclear copper-uridine complex anion $[Cu_8(Urd3,2',3'H_{-3})_8]^{8-}$, in which two square-planar coordinated copper atoms are linked by two O2′ ligator atoms of two triply deprotonated uridine moieties. The coordination sphere of the copper centers is completed by a nitrogen atom (N1) of the nucleobase of uridine [147]. The mononuclear copper complex **129** (◉ *Fig. 27*) is formed with uridine-5′-monophosphate in neutral aqueous solution in which the diol is deprotonated as a result of the chelation at the copper center while the non-coordinating nucleobase is not deprotonated under these conditions [148]. 3-Acetamidophenylboronic acid reacts with the diol group of adenosine to a cyclic monoester with tricoordinate boron [149], whereas the reaction of boric acid and guanosine in strong alkaline aqueous solution leads to reaction products similar to those described above for the boron-anhydroerythritol system: two stereoisomers of bis(diolato)borates (compare **121**, ◉ *Fig. 26*) exist in addition to monodiolato-dihydroxo borate ester [87]. The coordination behavior of $CpMo^{2+}$ towards the ribonucleosides **123**, **124**, **126**, and **127** and the corresponding 5′-monophosphates has been shown to be pH-dependent. At a pH of 9, the molybdocene centers are chelated exclusively by the diolato moiety, while in neutral solution a coordination of the nucleobase and phosphate oxgen atoms is also observed [150].

The reaction of cytidine **126** with [*mer*-(dien)(NO)$RuCl_2$]PF_6 in an aqueous solution of a pH of 12.5 leads to two isomers of diolato complexes similar to those described above with **118** while guanosine and uridine—both providing deprotonated amide functions under these conditions—are bonded to the ruthenium center by their nucleobases [81].

Homoleptic bis(diolato) bismuthates and antimonates are formed with adenosine as well as with guanosine in alkaline solutions with a pH of at least 11. The coordination pattern of all the complexes resembles the one in the bis(adenosinato)antimonate **130** (◉ *Fig. 28*): two nucleosides act as diolato ligands and occupy four of five coordination sites of the metal centers which are coordinated in a distorted trigonal-bipyramidal geometry with the stereochemically active lone pair in an equatorial orientation. Being identical in all the structures, the packing of the guanosinato complex anions is stabilized by two intermolecular hydrogen bonds of the type N2–H···N3 between two adjacent guanine base pairs. A CIS of 3–6 ppm of the carbon atoms C2′ and C3′ involved in the five-membered chelate ring formed by bismuth(III) and antimony(III) with adenosine **123**, guanosine **124**, and inosine **125** indicates that the coordination mode is maintained in alkaline solution [151,152].

The formation of several penta- or hexacoordinated silicate complexes is observed on the addition of adenosine, guanosine, and cytidine to aqueous alkaline silicate solutions. Hexacoordination is favored by an increase of the pH and high molar ratios of furanoidic cis-diol:silicon [94]. Hexacoordinate silicon is present in the homoleptic tris(diolato)silicate **131** (◉ *Fig. 28*) with cytidine **126** as chelating ligand. Pentacoordinate methoxy- or phenyl-bis(nucleosidato)silicates of the type $[SiR(Nuc2',3'H_{-2})2]^-$ (R = Ph, OMe) are obtained with the nucleosides (Nuc) adenosine and cytidine in methoxide-containing methanolic solutions. Different *syn/anti*-isomers similar to those observed with anhydroerythritol or methyl-β-D-ribofuranoside can be identified by ^{29}Si- and ^{13}C NMR data [92].

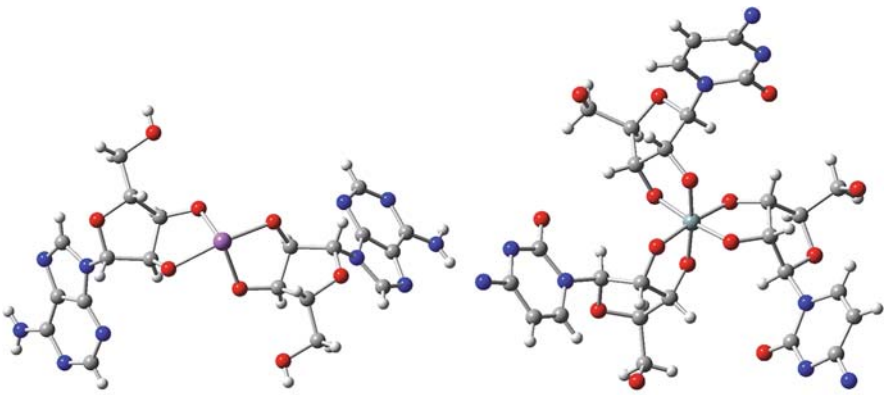
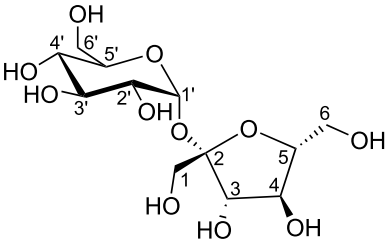


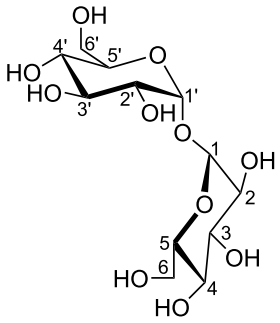
Figure 28
The molecular structures of $[Sb(Ado2',3'H_{-2})]^-$ 130 in crystals of $Na130 \cdot H_2O$ and $[\Lambda\text{-}Si(Cyd2',3'H_{-2})_3]^{2-}$ 131 in crystals of $Cs_2131 \cdot 21.5\ H_2O$

6.4 Non-Reducing Disaccharides

Metal complexes of non-reducing disaccharides such as sucrose (β-D-Fru*f*-(2↔1)-α-D-Glc*p*) **132** or α,α-trehalose (α-D-Glc*p*-(1↔1)-α-D-Glc*p*) **133** not only deserve interest in their own right but also as model compounds for the corresponding complexes with oligo- and polysaccharides. Thus the action of the coordinating cellulose solvents [Pd(en)$(OH)_2$] (Pd-en), [Cu(en)$(OH)_2$] (Cu-en), [Ni(tren)$(OH)_2$] (Ni-tren), and [Cd(tren)$(OH)_2$] (Cd-tren) on the polysaccharide were derived from crystal structures with **132** and **133** (compare ❯ *Sect. 6.6*).



132



133

In the dinuclear complex **134** (❯ *Fig. 29*), two Pd^{II}(en) moieties are chelated by the Glc*p*-O3′,O4′- and the Fru*f*-O1,O3-diolato moieties of sucrose **132**, that is, in five- and six-membered chelates, respectively. The selection of the Glc*p*-O3′,O4′ site instead of the similar Glc*p*-O2′,O3′ site is reasonable since an intramolecular hydrogen bond of the type O2′–H···$O1^-$ stabilizes the sucrose residue in a conformation which is also found for the free carbohydrate (there with a hydrogen bond of the type O1–H···O2′ in the reversed direction) [84,153]. A similar coordination and hydrogen-bond pattern is found with Cu^{II}(en) [153]. However, only a mononuclear complex with the Ni^{II}(tren) residue bonded to the Glc*p*-O2′,O3′ diolato site is obtained in a reaction of Ni-tren and sucrose in molar ratio of 2:1. In this complex, the intramolecular hydrogen bond O1–H···O2′ which is also found in free sucrose is retained [125,153]. A dicationic Δ-configured complex is formed by sucrose and bis(phenanthroline)cobalt(III) in which the metal center is involved in an eight-membered ring with the Glc*p*-O2′- and the Fru*f*-O1 chelation sites. NMR data reveal that O2′ is deprotonated and they support the structure of the Δ-configured complex by the observation of an NOE involving the sucrose and phenanthroline residues [154].

Two crystal structures of metal complexes of α,α-trehalose **133** are reported. Two Cd^{II}(tren) residues are chelated by the Glc*p*-O2,O3 and the Glc*p*-O2′,O3′ diolato moieties, respectively, in the dinuclear complex **135** (❯ *Fig. 29*). As in free α,α-trehalose in the crystalline state, direct intramolecular hydrogen bonds are not found due to conformational restraints, but two sequences of the type $O2^-$···H–O^w···H–O6′ with a linking water molecule H_2O^w are observed (the reversed direction is found in free α,α-trehalose) [153]. Only one of the two Glc*p*-O2,O3-chelation sites of α,α-trehalose is chosen in a mononuclear complex with Ni-Me_3tren, the *N*,*N*′,*N*″-trimethyl analog of Ni-tren, in which no support by an intramolecular

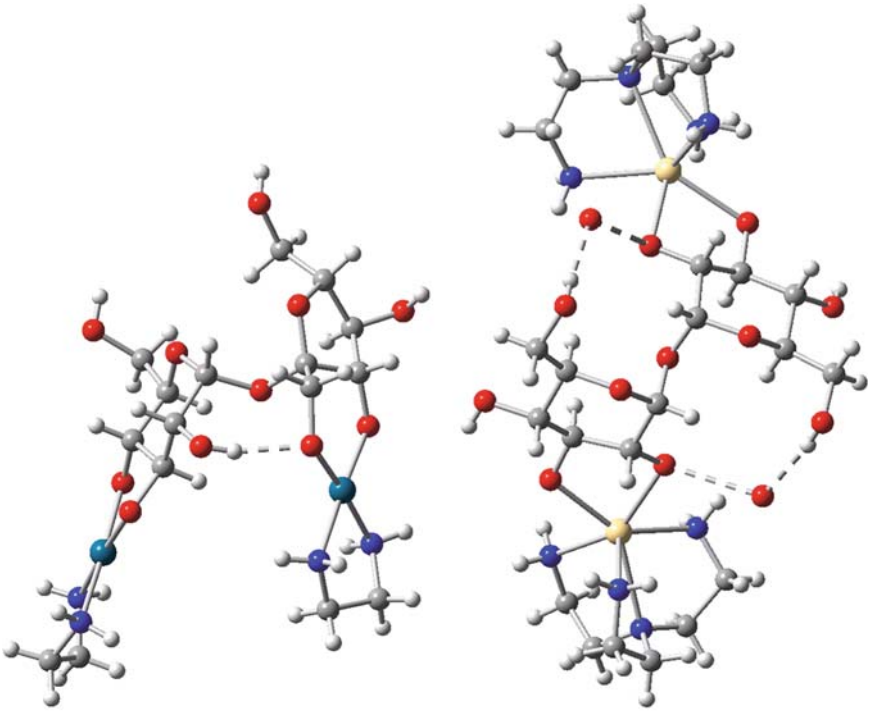


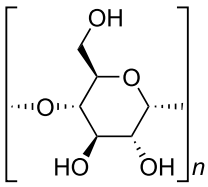
Figure 29
The molecular structures of $[(en)_2Pd_2(Suc1,3,3',4'H_{-4})]$ 134 in crystals of 134 · 11 H_2O and $[(tren)_2Cd_2(\alpha,\alpha\text{-}Tre2,3,2',3'H_{-4})] \cdot 2\ H_2O$ 135 in crystals of 135 · 15 H_2O. *Dashed lines*: hydrogen bonds

hydrogen bond is observed. It seems that in this case the chelation is governed by the enhanced acidity of the Glc*p*-O2,O3 site compared to the Glc*p*-O3,O4 site [125].

6.5 Cyclodextrins

Cyclodextrins (CD) are cyclic torus-shaped oligosaccharides with n α-(1↔4)-glycosidic-linked glucose molecules ($n = 6$: α-CD, $n = 7$: β-CD, $n = 8$: γ-CD) (**136**). All the Glc*p*-O2,O3 diol functions are oriented to the same side of the torus resulting in a regularly arranged coordination site pattern. This enables the formation of sandwich-type complexes in which the metal centers are enclosed between the bisdiol moieties of two CD molecules (double torus). The number of metals being chelated by the diol functions depends on factors such as the Lewis acidity of the metal or the size of the alkali counterion. Six bismuth(III) centers are bonded by six bisdiolato moieties provided by two α-cyclodextrin molecules in the electrically neutral C_2-symmetric structure **137** (*Fig. 30*, a = view through the double torus, b = view onto the double torus) though a molar ratio of bismuth:α-CD of 3:2 was present during the reaction in aqueous sodium hydroxide. This indicates that the reaction is governed less by stoichiometry than by molecular recognition, and, as a consequence, **137** (*Fig. 30*) may

be considered a supramolecular assembly. The Lewis acidity of the bismuth(III) atoms is sufficient to force complete deprotonation of the bisdiol moieties (compare the structures with adenosine, guanosine or methyl-α-D-mannopyranoside, for example). This leads to almost square O_4 patterns between adjacent bisdiolato moieties which are suited to coordinate sodium ions located within the double torus (**138**). The coordination sphere of the six sodium ions is completed by six water molecules, respectively, which are arranged in a homodromic hexaqua cycle (**137a**, ❯ *Fig. 30*). A similar structure is obtained with oxo-vanadium(IV) instead of bismuth(III) in which the oxo group bonded to the vanadium atom replaces the stereochemically active lone pair of the bismuth(III) centers [155]. Complete deprotonation of two γ-CD tori is also observed in the hexadecanuclear lead(II) complex $[Pb_{16}(\gamma\text{-CDH}_{-16})_2]$ which has been crystallized from a reaction mixture containing excessive γ-CD. Eight lead atoms each are located within the double torus (compare the sodium sites in **137**, ❯ *Fig. 30*) and outside the double torus (compare the bismuth sites in **137**, ❯ *Fig. 30*), respectively. The resulting lead-alkoxido substructure is comparable to that observed in **78** (❯ *Fig. 16*) [156].

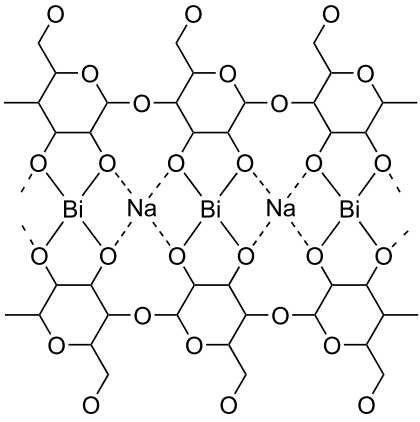


136

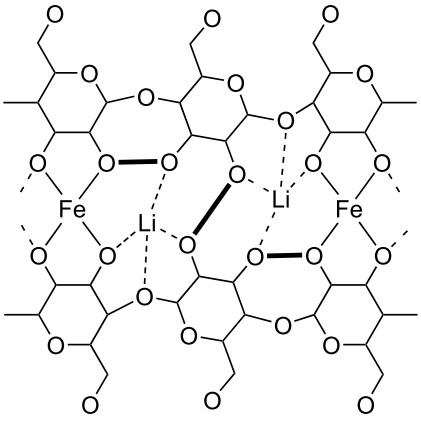
bold lines: hydrogen bonds
dashed lines: alkali ion-oxygen bonds

glucose-atom-numbering scheme on the top left side

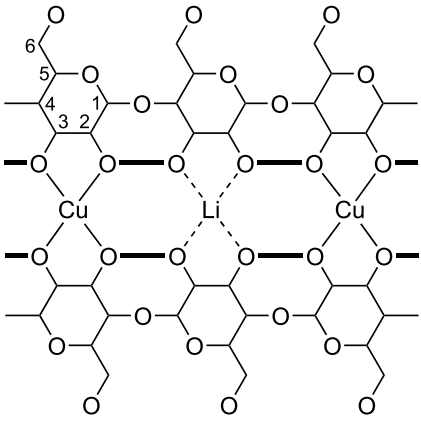
hydrogen atoms omitted for clarity



138



139



140

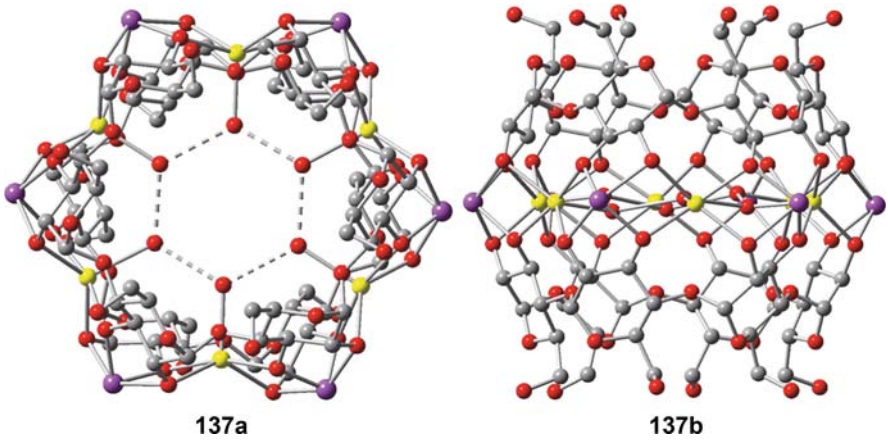
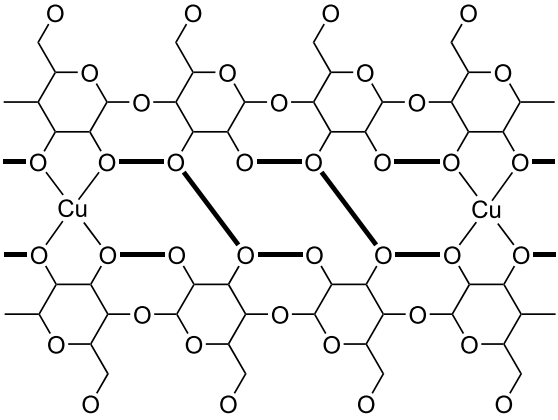


Figure 30
The molecular structure of $[Na_6(H_2O)_6Bi_6(\alpha\text{-CDH}_{-12})_2]$ 137 in crystals of 137 · 47 H_2O in a view through the double torus (137a) and in a view onto the double torus (137b). Bi atoms (violet), Na atoms (yellow). Hydrogen atoms omitted for clarity. *Dashed lines*: hydrogen bonds

Less Lewis acidic metals such as iron and manganese in their oxidation states +II cannot afford complete deprotonation but three of them are bonded by every second bisdiolato moiety in an α-CD double torus with the formula $[Li_6(H_2O)_6M_3(H_2O)_3(\alpha\text{-CDH}_{-12})_2]^{5-}$ (M = Mn, Fe). The fact that Fe^{II} or Mn^{II} are chelated by diolato functions of cyclodextrins might be surprising since no diolato complexes with simpler carbohydrates are known, but is comprehensible on a closer look at the structure **139**. The increased acceptor functions of the deprotonated diolato sites are stabilized by hydrogen bonds as well as by Li cation contacts. One of the adjacent not iron- or manganese-bonded glucose units is tilted towards a chelating glucose unit resulting in a shortening of the O2′–O3′ hydrogen bonds (O2′–O3′ distance about 2.5 Å in **139**). This way two optimal sites for the coordination of two lithium ions are generated between two bisdiolato moieties. Each of the six lithium ions occupying these sites bears an additional water molecule which, in turn, is arranged in a homodromic hexaqua cycle comparable to that in **137a** (*Fig. 30*) [155]. A similar coordination and hydrogen bond pattern is established in a sandwich-type tetracuprate with β-CD, $[Li_7(H_2O)_7Cu_4(\beta\text{-CDH}_{-11.5})_2]^{4-}$. However, one lithium ion is bonded to the O_4 site between two adjacent bisdiolato-cuprates comparable to the sodium atoms in **137** (*Fig. 30*) [157].

Trinuclear sandwich-type cuprates of the type $[A_3(H_2O)_3Cu_3(\alpha\text{-CDH}_{-6})_2]^{6-}$ are formed with lithium and sodium as counterions A, whereas dinuclear $[Cu_2(\alpha\text{-CDH}_{-4})_2]^{4-}$ are isolated as potassium and rubidium salts. The lithium- and sodium cuprates contain the substructure **140** in which the copper atoms are coordinated by two bisdiolato moieties in a slightly distorted square planar geometry and the alkali ions mainly occupy the bisdiol sites located between two cuprates. Stabilizing hydrogen bonds of the type O–H···O^- between non-deprotonated diol moieties and the adjacent diolato moieties are formed. Each of the alkali ions bears a water molecule as a fifth ligand. These bonds are directed towards the outside of the cylindic tricuprates (compare the in-torus bonded aqua ligands in **137**, *Fig. 30*). In the less-

deprotonated dicuprates, however, no incorporation of the larger, less-polarizing alkali metals potassium and rubidium is observed. These compounds are stabilized by short cooperative hydrogen bonds depicted in **141**. The anhydroglucose units not bearing a copper atom have to be tilted slightly to enable inter-cyclodextrin hydrogen bonds of the type O3–H···O3 [158].



bold lines: hydrogen bonds

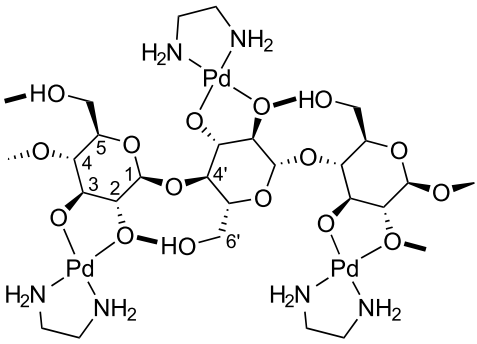
hydrogen atoms omitted for clarity

141

$Co^{III}(en)_2$ and Co^{III}(cyclen) (cyclen = 1,4,7,10-tetraazacyclododecane) residues are coordinated by a diolato moiety of α-CD or β-CD in an aqueous alkaline solution (compare $Co^{III}(en)_2$ complexation by a pyranosidic diolato moiety in **109**, ❯ *Fig. 21*). This coordination mode could be confirmed by a large downfield shift of the the carbon atoms involved in the five-membered diolato-chelate ring [159].

6.6 Polysaccharides

The disaccharide structures described above give first evidence of how cellulose, a polymer formed by (1↔4)-glycosidic linking of β-D-glucose molecules, is present in coordinating solvents like Pd-en, Cu-en, Ni-tren, or Cd-tren. Supporting information could be gained by ^{13}C NMR data obtained from Pd-en solutions of DP 40 cellulose (cellulose with an average degree of polymerization of 40). A downfield coordination-induced shift of about 10 ppm of the carbon atoms C2 and C3, which is also observed for the 1,2-diolate complexation of sucrose (compare 6.4), and unsplit main signals of all of the glucose carbon atoms strongly indicate the formation of entirely palladated and molecularly dispersed cellulose **142**. The formation of intramolecular hydrogen bonds of the type O6′–H···$O2^-$ leads to a significant increase in chain stiffening [84,160]. The same coordination mode is suggested for Cu-en and cuoxam, an aqueous solution of copper(II)-hydroxide in ammonia, as well as for Ni-tren and Cd-tren [161,162].

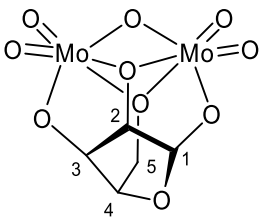


142

7 Reducing Carbohydrates

Although there are many data for model compounds for carbohydrates, relatively little is known about the reducing carbohydrates as ligands in metal complexes. One of the reasons for this may be the high reactivity of the reducing sugars. Furthermore their ease of oxidation and their instability in alkaline solution must be accounted for, and most importantly the monosaccharides are mixtures in solution. When a coordinating metal moiety is exposed to glucose, for example, it is hardly predictable whether the sugar will react through its α- or β-furanose, α- or β- pyranose or a less stable form. Nevertheless, the number of structurally characterized coordination compounds with reducing carbohydrates is increasing.

From the reaction between ammonium molybdate and the pentose xylose, a complex composed of a dimolybdenum core and one sugar moiety **143** [163] could be obtained.



143

As Bilik [164,165,166] has demonstrated Mo(VI) will catalyze epimerization of pentose sugars at C2. So it is not surprising that with a reducing sugar, the first structure analysis with a transition metal complex shows lyxose as the monosaccharide in complex. The metal center is dinuclear with a triple oxygen bridge linking the two molybdenum atoms without remarkable Mo–O distances. The dinuclear metal center fixes the lyxose in an unusual furanose form, the β-D-lyxofuranose. The pseudo-rotation formulae have been applied to this form of lyxose. A phase angle of 247.7° together with a normal amplitude of pucker (reflecting the degree of non-planarity of the ring) of 35.6° is calculated. The conformation of lyxose is thus midway between the discrete descriptions ^{C4}E and $^{C4}T_{O1}$.

As the pentose sugar is able to provide an O_4 rhomb, structure analyses have shown that the O_5 pattern provided by β-mannofuranose turned out to be an ideal building block of M_2O_{10} double octahedra of binuclear metallates(III).

Thus, a series of isotypic sugar complexes of trivalent metals with the formula $Ba_2[M_2(\beta\text{-D-Man}f\text{H}_{-5})_2] \cdot n\ H_2O$ (M = Fe, V, Cr, Al, Ga, Mn; n ≈ 13), as shown for Mn^{III} in ❯ *Fig. 31*, has been developed [167,168].

The compounds contain a homoleptic sugar-metal complex as anion, which includes no hydroxido or oxido ligands. Instead, all the protons of the five hydroxyl groups of mannose are split off to form a pentaanionic polyolato ligand. The polyolate is derived from the β-furanose form of mannose, which is insignificant in solution [169], but is the only form with all the hydroxyl groups on one side of the ring. According to the already described epimerization of xylose towards lyxose in the presence of oxidomolybdenum(VI), D-fructose and D-glucose isomerize to the sugar ligand with the highest possible denticity, the β-mannofuranose [167].

The environments of the two metal atoms in **144** are not equivalent. M1 forms five- and six-membered chelates with the open-chain part of the ligand, whereas M2 is surrounded by two oxolanetriolato fragments in a distorted E_2 conformation. $[Fe_2(rac\text{-Man}f\text{H}_{-5})_2]^{4-}$ **145** [31] with iron as the central metal is also known. In contrast to **144**, in the C_i symmetric **145** the two Fe centers are equivalent. Besides the structure of $[Mn^{III}{}_2(\beta\text{-D-Man}f\text{H}_{-5})_2]^{4-}$ there is an additional structure with two different oxidation states, $[Mn^{IV}Mn^{III}(\beta\text{-D-Man}f\text{H}_{-5})_2]^{3-}$ [168]. The unexpected oxidation state +VI for one of the central atoms was unambiguously assigned and became apparent for the mean Mn–O bond length in the octahedral MnO_6 coordination. For Mn^{III}, the mean Mn–O distance is 2.01 ± 0.13 Å and for Mn^{VI}, it is 1.89 ± 0.03 Å.

In the structures, introduced so far, an O_4 rhomb or O_5 patterns turned out to be ideal building blocks for carbohydrate-metal complexes. The most important monosaccharide, D-glucose, in any of its hemiacetal forms does not exceed the simple O2 diol pattern (a consequence is the epimerization reaction of glucose to β-mannofuranose in the presence of trivalent metal ions).

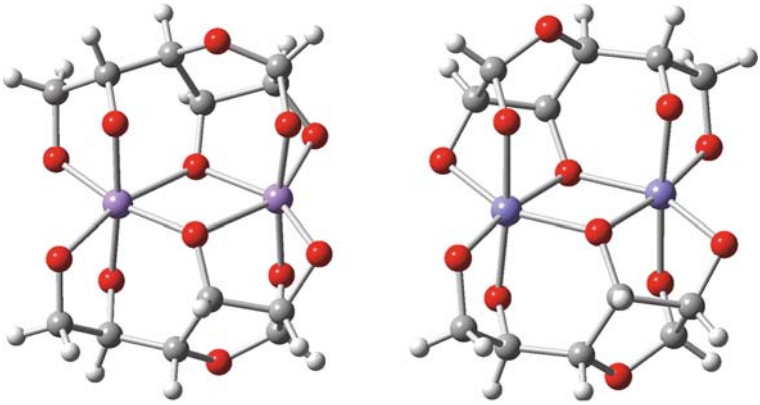
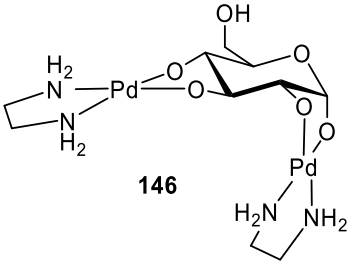


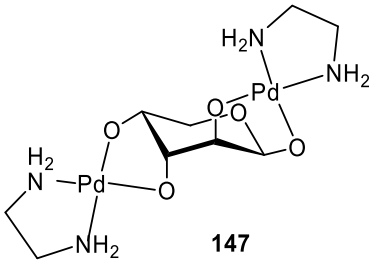
Figure 31
Molecular structures of the tetraanions $[Mn^{III}{}_2(\beta\text{-D-Man}f\text{H}_{-5})_2]^{4-}$ 144 and $[Fe_2(rac\text{-Man}f\text{H}_{-5})_2]^{4-}$ 145 in crystals of $Ba_2[Mn^{III}{}_2(\beta\text{-D-Man}f\text{H}_{-5})_2] \cdot 13\ H_2O$ and $Ba_2[Fe_2(rac\text{-Man}f\text{H}_{-5})_2] \cdot 13\ H_2O$

This steric restriction is taken into account by exclusively providing the (en)Pd^{II} moiety, which is a diol chelator.

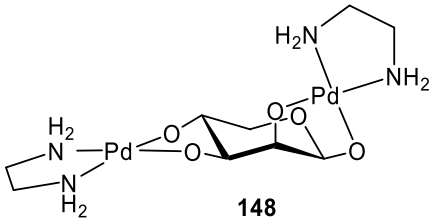
Although palladium is a good oxidant, oxidation of the aldose could be avoided to such an extent that even crystallization became possible. The first crystal structure of a metal derivative of glucose, [(en)$_2$Pd$_2$(α-D-Glc*p*1,2;3,4H$_{-4}$)] · 7 H_2O (**146** · 7 H_2O) was obtained from a solution with a 3:1 Pd:D-glucose molar ratio [170]. **146** shows the O1–O4 deprotonated α-D-glucopyranose tetraanion, which is coordinated as a bis(chelate) ligand to two palladium(II) central atoms with the Pd–O distance averages at 2.00 Å, the O–C–C–O angles are 43 and −50°. ^{13}C NMR data give evidence for the existence of an additional species [(en)$_2$Pd$_2$(β-D-Glc*p*1,2;3,4H$_{-4}$)] in solution. Reduction of the Pd:D-glucose molar ratio to 1:1 results in a mixture of monometallated glucoses. The 1,2-metallated species are of particular significance, owing to the maximum acidity of the hydroxy group at C1.



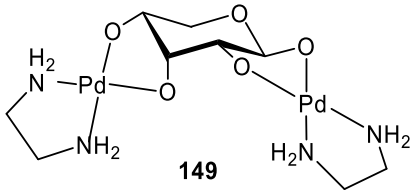
146



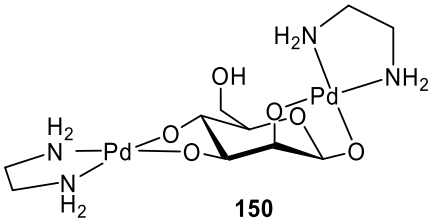
147



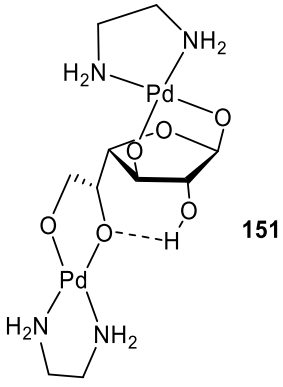
148



149



150



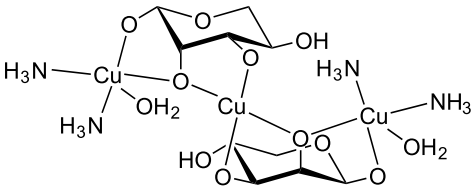
151

In this way several Pd_2 complexes 2[171,172], namely $[(en)_2Pd_2(\beta\text{-D-Ara}p1,2;3,4H_{-4})] \cdot 5\,H_2O$ (**147** $\cdot 5\,H_2O$), $[(en)_2Pd_2(\beta\text{-D-Lyx}p1,2;3,4H_{-4})] \cdot 7\,H_2O$ (**148** $\cdot 7\,H_2O$), $[(en)_2Pd_2(\beta\text{-D-Rib}p1,2;3,4H_{-4})] \cdot 6.5\,H_2O$ (**149** $\cdot 6.5\,H_2O$), $[(en)_2Pd_2(\beta\text{-D-Man}p1,2;3,4H_{-4})] \cdot 9.4\,H_2O$ (**150** $\cdot 9.4\,H_2O$), which are the main species in solution, and $[(en)_2Pd_2(\beta\text{-D-Gal}f1,3;5,6H_{-4})] \cdot 5\,H_2O \cdot C_2H_5OH$ (**151** $\cdot 5\,H_2O \cdot C_2H_5OH$) could be obtained.

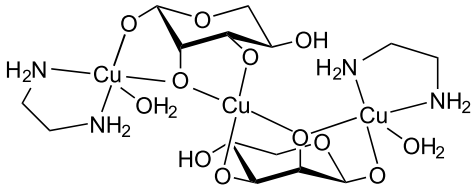
With the exception of **151** one of the two Pd(en) ligands binds to the O1 and O2 and the other to the O3 and O4 diol function of the β-D-pyranose form of the carbohydrate in all compounds. Further, the Pd–O distances and the O–C–C–O angles are in comparable ranges, too. **151**, however, shows a furanose form of D-galactose. Contrary to the other pyranose structures with Pd(en) moieties, a strong intramolecular hydrogen bond of the O–H···O^- type is established with O2-H as the donor and O5 as the acceptor. One of the palladium atoms is coordinated by a 1,3-diolato(2-) ligand. Though this bonding mode is unusual, it is not unique, as shown for levoglucosan which binds to palladium(II) in a similar way. The second Pd(en) ligand binds to the remaining diol fragment with O5 and O6.

^{13}C NMR data show, that there exist in solution more palladium–carbohydrate species than could be obtained in crystals. Besides the above-mentioned compounds there are signals for the following species identified by the palladium induced shift of about 8 to 15 ppm: $[(en)_2Pd_2(\alpha\text{-D-Ara}p1,2;3,4H_{-4})]$, $[(en)Pd(\beta\text{-D-Lyx}p1,2H_{-2})]$, $[(en)Pd(\beta\text{-D-Lyx}p2,3H_{-2})]$, $[(en)_2Pd_2(\alpha\text{-D-Rib}p1,2;3,4H_{-4})]$, $[(en)_2Pd_2(\alpha\text{-D-Xyl}p1,2;3,4H_{-4})]$, $[(en)_2Pd_2(\beta\text{-D-Xyl}p1,2;3,4H_{-4})]$, $[(en)_2Pd_2(\alpha\text{-D-Gal}p1,2;3,4H_{-4})]$ and $[(en)_2Pd_2(\beta\text{-D-Gal}p1,2;3,4H_{-4})]$. The dimetallation of the carbohydrates influences the distribution of the different forms of sugars in solution in contrast to the forms in equilibrium without coordination. Formation of dimetallated pyranoses requires two diol functions that are not *trans*-diaxial. Hence, for lyxopyranose and its homologue mannopyranose, only the β-anomer in its $^4C_1$ conformation can be dimetallated. The major dimetallated species in solutions of ribose and galactose can be predicted as well, since the most stable form of the free sugars in aqueous solution (β-pyranose) provides two diol functions that are well suited for palladium bonding without rearrangement. The same appears to hold true first glance for the xylose/glucose pair. However, the actual anomer distribution deviates from that of the free sugars by a shift of the concentrations towards the α-anomer. The bonding modes of dimetallated arabinopyranoses cannot be predicted. Suitable bisdiol conformations are adopted for the β-anomer both in its $^1C_4$ and $^4C_1$ conformations (two axial substituents in each case) and in its $^1C_4$-α-anomer, the latter and the $^4C_1$-β-anomer being handicapped by an anomeric effect when the findings in the xylose/glucose case are generalized. In fact, the spectra show the β-form is preferred.

While (en)Pd fragments bind to diol groups without loss of the ethylenediamine ligand (nitrogen ligation is clearly favored over alkoxide bonding), at cupric centers the prediction of a particular structure is complicated by the ability of diolato ligands to substitute nitrogen ligands. In crystals of $[(NH_3)_4Cu_3(\beta\text{-D-Lyx}p1,2,3H_{-3})_2(H_2O)_2] \cdot 4\,H_2O$ [172] (**152** $\cdot 4\,H_2O$), two bridging lyxose ligands adopt the β-pyranose form in C2 symmetrical, trinuclear molecules; they are each threefold deprotonated along the *cis-cis* sequence of the hydroxy groups with C1, C2, and C3.



152

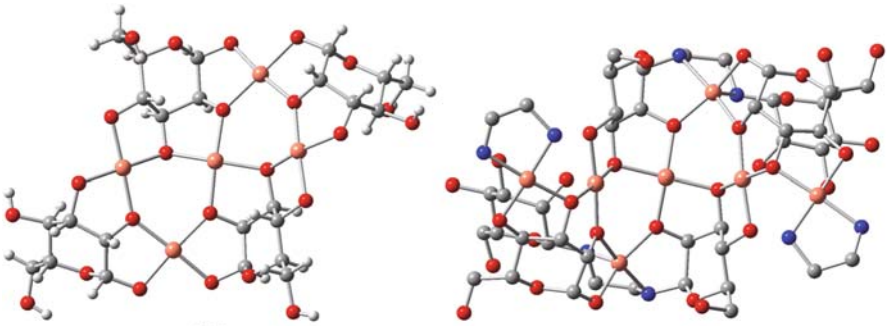


153

The coordination chemistry of the cupric centers is normal, that is, the bis-diolato-bonded Cu1 atom is square-planar coordinated with no further ligands, whereas Cu2, which is coordinated by a diolato ligand derived from the anomeric and the epimeric hydroxy groups, is surrounded by a square pyramid consisting of the two alkoxy oxygens, two ammine ligands and a water ligand at a longer distance. Intramolecular hydrogen bonds are established from N2–H donors and O3 acceptors. The ability of a polyolato ligand to replace ammine ligands is a prerequisite for the formation of **152**. The molecular structure of $[(en)_2Cu_3(\beta\text{-D-Lyx}p1,2,3H_{-3})_2(H_2O)_2] \cdot 4\ H_2O$ [172] (**153** $\cdot$ 4 H_2O) resembles that in the ammine complex, including the detail of intramolecular hydrogen bonding from amine donors to alkoxide acceptors. The structure determination demonstrates that the stronger ethylenediamine ligand is also replaced by the lyxose anions.

More examples for the higher reactivity of copper centers, in relation to the preference of diolato ligands to substitute nitrogen ligands, are given in structures **154** and **155** [173], shown in ❯ *Fig. 32*.

Complex **154** is an anionic homoleptic complex in which the five square-plane-coordinated copper(II) atoms are exclusively bound to the deprotonated β-D-mannopyranose in $^4C_1$ conformation. The increased redox stability therefore is not a consequence of the spatial separation of the oxidizing Cu(II) atoms and the half-acetal function of the monosaccharide. However, it is a consequence of the involvement of the O atoms, at the anomer C atoms, in the coordination of copper, as it is for the homoeptic lyxose complexes.



154 155

Figure 32

Structure of $[Cu_5(\beta\text{-D-Man}p)_4H_{-13}]^{3-}$ (154) in K_3 154 $\cdot$ α-D-Man$p$ $\cdot$ 16.5 H_2O and $[(en)_2Cu_7(\beta\text{-D-Man}p1,2,3,4H_{-4})_2(L2,3,4H_{-3})_2]$ (155) in 155 $\cdot$ 26.6 H_2O with L = *N*,*N*'-bis(β-D-mannopyranosyl)-ethylendiamine

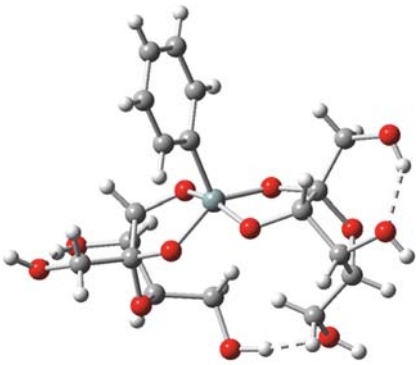
In contrast to the other aldoses, the copper atoms find an especially favorable coordination environment for lyxose and mannose. Cu(II), indeed, has the ability to build stable complexes with the pyranoid *trans*-diol fragments, shown for the complexed 3,4-diol functions, but the more flat chelate five-membered rings, which are built with the with pyranoid *cis*-diol fragments, seemed to be more favorable for copper centers.

The *cis,cis*-orientation of the triol function with O1, O2 and O3 probably leads to an increased complex stability in the environment of the anomeric center and, as a consequence, to decreased tendency for reduction of copper.

A second compound, **155**, contains a reaction product with ethylenediamin and two moieties of mannose. Although the structure of the pentacuprate is quite complex, the composition of **155** could be derived from the simple educts. Besides oxidation, the building of *N*-glycosides is another reaction of the aldoses.

For silicon, the metal-coordinating properties, as found in AnEryt, should, in principle, be shared by each of the monosaccharides. A particular monosaccharide is expected to act as a good ligand if its *cis*-furanose form is of considerable stability, so that the stability constant of the complex is not charged with the isomerization energy of the ligand. In fact, almost all of the monosaccharides enrich alkaline aqueous silicate solutions with five- and six-coordinate silicon species, to some extent. However, the determination of the compositions and structures of the involved species is largely complicated by the mere number of the species in equilibrium. This has recently been demonstrated for D-ribose solutions by Kinrade et al., who detected a vast amount of various five- and six-coordinate species in such solutions [94]. For monosaccharides other than D-ribose, the situation is not any better.

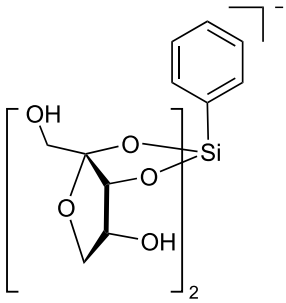
To avoid these problems, the hydroxide ligand is substituted by an inert phenyl residue. In this way a series of phenylsilicates was crystallized. Five complexes have been isolated and characterized with two ketoses and three aldopentoses [174]. The silicon central atom in [K([18]crown-6)]**156** · MeOH is part of two chelate rings, with the ligands being O^2,O^3-β-D-fructofuranose dianions (*Fig. 33*).



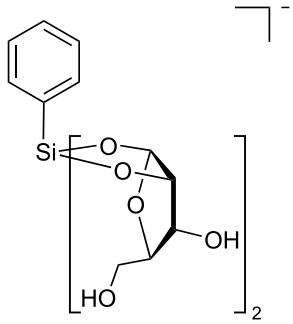
156

Figure 33
Structure of the anion [PhSi(β-D-Fru*f* 2,3H$_{-2}$)$_2$]$^-$ 156 in the crystal structure of [K([18]crown-6)][PhSi(β-D-Fru*f* 2,3H$_{-2}$)$_2$] · MeOH

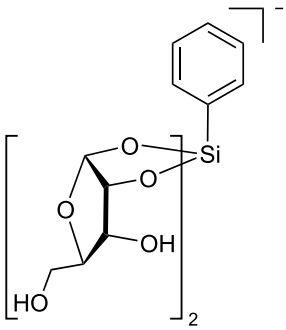
The β-furanose isomer is best suited for silicon ligation because it exhibits a torsion angle close to 0° for the most acidic diol function, thus assuring a flat chelate ring. The same structural principles are also found in the anions [PhSi(a-D-Rul*f*2,3H$_{-2}$)$_2$]$^-$ **157**, [PhSi(β-D-Ara*f*1,2H$_{-2}$)$_2$]$^-$ **158**, [PhSi-(a-D-Rib*f*1,2H$_{-2}$)$_2$]$^-$ **159** and [PhSi(a-D-Xyl*f*1,2H$_{-2}$)$_2$]$^-$ **160**.



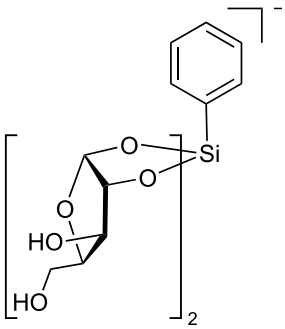
157



158



159



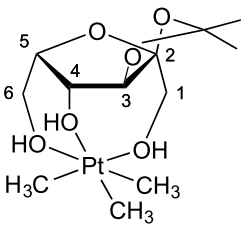
160

Continuing with the furanoses, in **161** [133] 2,3-*O*-isopropylidene-α-L-sorbofuranose acts as a neutral tridentate ligand which is coordinated by three hydroxyl groups (κ^3O^1,*O*^{4},*O*^{6} coordination).

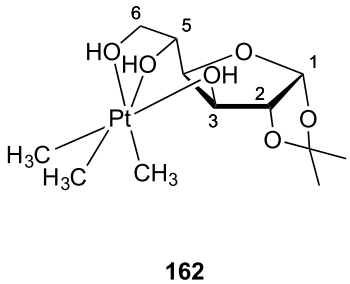
The six-, seven-, and eight-membered 1,3,2-dioxaplatinacyclohexane rings exhibit boat, chair, and distorted chair conformations, respectively. The cyclic system is not free of bond angle strain, as indicated by the O–Pt–O angles in particular. One of these angles is distinctly smaller than 90° (O4–Pt–O6 = 82.5° vs. O1–Pt–O6 = 95.3° and O1–Pt–O4 = 91.2°. Two Pt–O bonds (O1, O4) average 2.25 Å and are equivalent to those in **162**. In contrast, the Pt–O6 bond with 2.16 Å is significantly shorter. The furanose ring assumes an envelope conformation in which C4 is displaced from the C5,O4,C2,C3 plane by 0.37 Å. A similar distortion of 0.57 Å but in the opposite direction is observed in the furanose ring of 1,2-*O*-isopropylidene-α-D-glucofuranose in **162** [133].

In **163** [133], 2,3;4,5-di-*O*-isopropylidene-β-D-fructopyranose acts as a neutral bidentate ligand which is coordinated by the hydroxyl group (O1) and the acetal oxygen atom (O6) of the pyranose ring (κ^2O^1,*O*^{6} coordination). The octahedral coordination of the platinum is com-

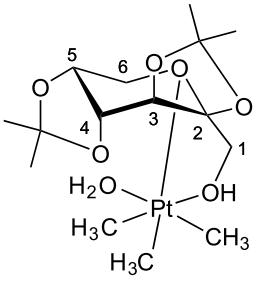
pleted by an aqua ligand. The five-membered 1,3,2-dioxaplatina cyclohexane ring exhibits a half-chair conformation. This ring is not free of bond angle strain, which is revealed by the small O1–Pt–O6 angle (75.3°) and the Pt–O bonds average 2.24 Å. The pyranose ring of the free ligand 2,3;4,5-di-*O*-isopropylidene-β-D-fructopyranose exhibits a $^2S_0$ conformation in the solid state [175]. In contrast, in the crystal structure of **163** [133], the twisted boat conformation of the pyranose ring is distorted due to complexation with platinum. This distortion is caused by changes in the dihedral angles in the sugar ring; in particular, the dihedral angle C3–C4–C5–C6 is reduced by 7° and C4–C5–C6–O6 is increased by 6°. Another structurally characterized β-D-fructopyranose is $[B_2(C_6H_5)_2(\beta\text{-D-Fru}p2,3;4,5H_{-4})] \cdot$ acetone [176], shown in *Fig. 34*.



161

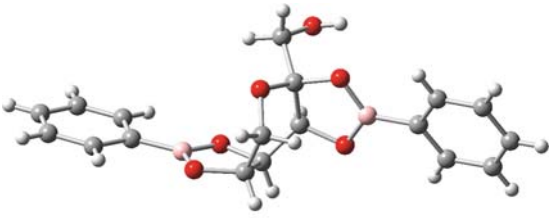


162



163

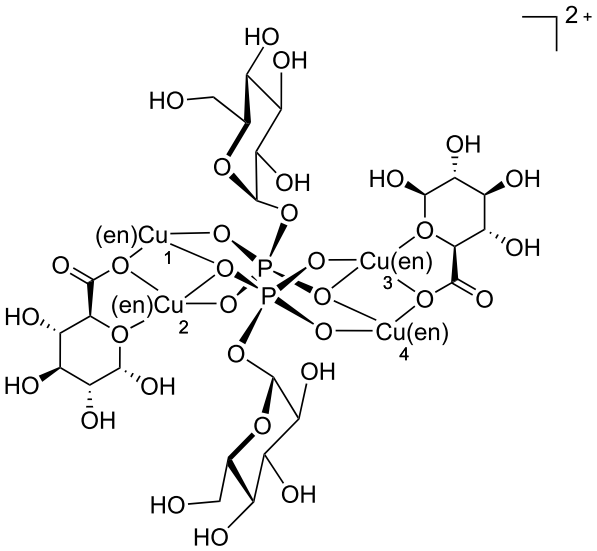
In **164**, phenylboronic acid forms five-membered cyclic boronates with the 2,3 and 4,5 diol function of the β-D-fructopyranose.



164

Figure 34
Stucture of $[B_2(C_6H_5)_2(\beta\text{-D-Fru}p2,3;4,5H_{-4})]$ (164) in crystals of $[B_2(C_6H_5)_2(\beta\text{-D-Fru}p2,3;4,5H_{-4})] \cdot$ acetone

In **165** [62], two D-glucoronate anions are introduced onto a tetracopper(II) scaffold of $[Cu_4\{\mu\text{-}(\alpha\text{-D-Glc-1P})\}_2(bpy)_4]^{4+}$.



165

Of the two D-glucoronate ligands, one adopts the α-D-pyranose form, bridging the Cu1 and Cu2 ions, and the other adopts the β-D-pyranose form, connecting the Cu3 and Cu4 atoms. The C6 carboxylate group of each D-glucoronate connects the two Cu atoms in a η^1,η^1 monoatom bridging mode. The five-membered chelation by axial coordination is completed of the C5 cyclic O atom. The metal binding of the pyranose ring-O atom of D-glucoronate may accelerate interconversion between the α and β anomers through an intermediate open-chain form in solution.

Besides these complexes characterized by crystal-structure analysis, there is an extensive number of compounds containing reducing sugars, investigated only by various spectroscopic methods. Examples include borate esters of carbohydrates that have been investigated in depth by ^{11}B and ^{13}C NMR spectroscopy [177,178]. Furthermore, there is a great deal of literature that provides NMR spectroscopic data for molybdenum and tungsten complexes of carbohydrates as reviewed in [1]. However, it has remained as an ambitious goal for future work to support the NMR-derived conclusions by solid-state work. Hence, the above-mentioned oxidomolybdenum-lyxose complex, which demands an O_4 rhomb pattern for coordination, provides the only crystal-structure analysis in this substance class.

8 Concluding Remarks

While structural information about metal complexes of reducing carbohydrates was only sparsely available 10 years ago, a considerable increase has been observed in recent years.

In our opinion, this development was facilitated mainly by two key factors: the technical progress of all analytical methods, particularly in the fields of NMR spectroscopy and X-ray diffractometry, and the plenty of structural data meanwhile available for metal complexes of model compounds of carbohydrates. The basic research on the structural chemistry of the latter complexes followed by a transfer of the thereby gained knowledge in stability and regioselectivity of metal coordination into reducing carbohydrates has proved to be very successful. By this way, the improvement of existing and the development of new applications of metal complexes of carbohydrates, which provide a cheap and renewable feedstock, is merely a matter of time.

References

1. Verchere JF, Chapelle S, Xin F, Crans DC (1998) Prog Inorg Chem 47:837
2. Yano S (1998) Chin J Polym Sci 16:193
3. Yano S, Mikata Y (2002) Bull Chem Soc Jpn 75:2097
4. Gyurcsik B, Nagy L (2000) Coord Chem Rev 203:81
5. Nagy L, Szorcsik A (2002) J Inorg Biochem 89:1
6. Nagy L, Yamaguchi T, Yoshida K (2003) Struct Chem 14:77
7. Steinborn D, Junicke H (2000) Chem Rev 100:4283
8. Alexeev Y, Vasilchenko IS, Kharisov BI, Blanco LM, Garnovsk AD, II, Zhdanov Y (2004) J Coord Chem 57:1447
9. Zhdanov YA, Alekseev YE (2002) Russ Chem Rev 71:969
10. Petrou AL (2002) Coord Chem Rev 228:153
11. Baruah B, Das S, Chakravorty A (2003) Coord Chem Rev 237:135
12. Zamojski A, Jarosz S (2003) Curr Org Chem 7:1
13. Lomozik L, Gasowska A, Bregier-Jarzebowska R, Jastrzab R (2005) Coord Chem Rev 249:2335
14. Karakhanov EE, Maksimov AL, Runova EA, Kardasheva YS, Terenina MV, Buchneva TS, Guchkova AY (2003) Macromol Symp 204:159
15. Engeldinger E, Armspach D, Matt D (2003) Chem Rev 103:4147
16. Jeunesse C, Armspach D, Matt D (2005) Chem Commun 5603
17. Varma AJ, Deshpande SV, Kennedy JF (2004) Carbohydr Polym 55:77
18. Kästele X, Klüfers P, Kunte T (2001) Z Anorg Allg Chem 627:2042
19. Allscher T, Kästele X, Kettenbach G, Klüfers P, Kunte T (2007) Chem Asian J 2:1037
20. Habermann N, Klaassen M, Klüfers P (1993) Carbohydr Res 241:9
21. Klüfers P, Schuhmacher J (1994) Angew Chem 106:1839
22. Andrews MA, Voss EJ, Gould GL, Klooster WT, Koetzle TF (1994) J Am Chem Soc 116:5730
23. Klüfers P, Labisch O (2003) Z Anorg Allg Chem 629:1441
24. Ma L, Liu S, Zubieta J (1989) Polyhedron 8:1571
25. Chapelle S, Verchere JF, Sauvage JP (1990) Polyhedron 9:1225
26. Matulova M, Bilik V, Alfoldi J (1989) Chem Pap 43:403
27. Chapelle S, Verchere JF (1992) Inorg Chem 31:648
28. Burger J, Klüfers P (1997) Angew Chem Int Ed Engl 36:776
29. Ishi I, Nakashima K, Shinkai S, Araki K (1998) Tetrahedron 54:8679
30. Munoz A, Lamande L (1992) Carbohydr Res 225:113
31. Burger J, Klüfers P (1997) Z Anorg Allg Chem 623:1547
32. Hinrichs M, Hofbauer FR, Klüfers P (2006) Chem Eur J 12:4675
33. Benner K, Klüfers P, Vogt M (2003) Angew Chem Int Ed 42:1058
34. Burgmayer SJN, Stiefel EI (1988) Inorg Chem 27:2518
35. Chapelle S, Sauvage JP, Koell P, Verchere JF (1995) Inorg Chem 34:918
36. Burger J, Klüfers P (1998) Z Anorg Allg Chem 624:359
37. Chapelle S, Sauvage JP, Verchere JF (1994) Inorg Chem 33:1966
38. Junicke H, Steinborn D (2003) Inorg Chim Acta 346:129

39. Gupta A, Kirfel A, Will G, Wulff G (1977) Acta Crystallogr Sect B: Struct Sci B33:637
40. Van Duin M, Peters JA, Kieboom APG, Van Bekkum H (1985) Tetrahedron 41:3411
41. Hedman B (1977) Acta Crystallogr Sect B: Struct Sci B33:3077
42. Godfrey JE, Waters JM (1975) Cryst Struct Comm 4:5
43. Chapelle S, Verchere JF (1991) Carbohydr Res 211:279
44. Herdin S, Klüfers P, Kunte T, Piotrowski H (2004) Z Anorg Allg Chem 630:701
45. Klüfers P, Schuhmacher J (1995) Angew Chem Int Ed Engl 34:2119
46. Norrild JC (2001) J Chem Soc Perkin Trans 1719
47. Matulova M, Hricoviniova Z (2002) Carbohydr Res 337:1745
48. Chapelle S, Verchere JF (1995) Carbohydr Res 266:161
49. Chapelle S, Koll P, Verchere JF (1998) Carbohydr Res 306:27
50. Chapelle S, Sauvage JP, Koell P, Verchere JF (1995) Inorg Chem 34:918
51. Lis T (1979) Acta Crystallogr Sect B: Struct Crystallogr Cryst Chem B35:1699
52. Lis T (1984) Acta Crystallogr Sect C: Cryst Struct Commun C40:374
53. Tsubomura T, Yano S, Yoshikawa S (1988) Bull Chem Soc Jpn 61:3497
54. Tajmir-Riahi HA (1990) J Inorg Biochem 39:33
55. Gajda T, Gyurcsik B, Jakusch T, Burger K, Henry B, Delpuech JJ (1998) Inorg Chim Acta 275–276:130
56. Szorcsik A, Nagy L, Gyurcsik B, Vanko G, Kraemer R, Vertes A, Yamaguchi T, Yoshida K (2004) J Radioanal Nucl Chem 260:459
57. Giroux S, Rubini P, Henry B, Aury S (2000) Polyhedron 19:1567
58. Junicke H, Arendt Y, Steinborn D (2000) Inorg Chim Acta 304:224
59. Escandar GM, Salas Peregrin JM, Gonzalez Sierra M, Martino D, Santoro M, Frutos A, Garcia SI, Labadie G, Sala LF (1996) Polyhedron 15:2251
60. Escandar GM, Olivieri AC, Gonzalez-Sierra M, Frutos AA, Sala LF (1995) J Chem Soc Dalton Trans 799
61. Frutos AA, Sala LF, Escandar GM, Devillers M, Peregrin JMS, Sierra MG (1999) Polyhedron 18:989
62. Kato M, Sah AK, Tanase T, Mikuriya M (2006) Inorg Chem 45:6646
63. Llopis E, Ramirez JA, Cervilla A (1986) Transition Met Chem 11:489
64. Ramos ML, Caldeira MM, Gil VMS (2000) J Chem Soc Dalton Trans 2099
65. Ramos ML, Caldeira MM, Gil VMS (1997) Carbohydr Res 297:191
66. Ramos ML, Calderia MM, Gil VMS (1997) Carbohydr Res 299:209
67. Van Duin M, Peters JA, Kieboom APG, Van Bekkum H (1986) Rec Trav Chim Pays-Bas 105:488
68. Lakatos A, Bertani R, Kiss T, Venzo A, Casarin M, Benetollo F, Ganis P, Favretto D (2004) Chem Eur J 10:1281
69. Ferrier F, Avezou A, Terzian G, Benlian D (1998) J Mol Struct 442:281
70. Abrahams BF, Moylan M, Orchard SD, Robson R (2003) Angew Chem Int Ed 42:1848
71. Sheldrick B, Mackie W (1989) Acta Crystallogr Sect C: Cryst Struct Commun C45:1072
72. Klüfers P, Kramer G, Piotrowski H, Senker J (2002) Z Naturforsch B: Chem Sci 57:1446
73. Saladini M, Candini M, Iacopino D, Menabue L (1999) Inorg Chim Acta 292:189
74. Ramos ML, Caldeira MM, Gil VMS (1991) Inorg Chim Acta 180:219
75. Ramos ML, Caldeira M, Gil VMS, Van Bekkum H, Peters JA, Van Bekkum H, Peters JA (1994) Polyhedron 13:1825
76. Ramos ML, Caldeira MM, Gil VMS, Van Bekkum H, Peters JA (1994) J Coord Chem 33:319
77. Venema FR, Peters JA, Van Bekkum H (1993) Rec Trav Chim Pays-Bas 112:445
78. Van Duin M, Peters JA, Kieboom APG, Van Bekkum H (1987) J Chem Soc Perkin Trans 2473
79. Van Duin M, Peters JA, Kieboom APG, Van Bekkum H (1987) J Chem Soc Dalton Trans 2051
80. Burger J, Klüfers P (1995) Chem Ber 128:75
81. Barth M, Kästele X, Klüfers P (2005) Eur J Inorg Chem1353
82. Klüfers P, Krotz O, Ossberger M (2002) Eur J Inorg Chem1919
83. Achternbosch M, Klüfers P (1994) Acta Crystallogr Sect C: Cryst Struct Commun C50:175
84. Ahlrichs R, Ballauff M, Eichkorn K, Hanemann O, Kettenbach G, Klüfers P (1998) Chem Eur J 4:835
85. Habermann N, Jung G, Klaassen M, Klüfers P (1992) Chem Ber 125:809

86. Klaassen M, Klüfers P (1993) Z Anorg Allg Chem 619:661
87. Benner K, Klüfers P (2000) Carbohydr Res 327:287
88. Klüfers P, Mayer P, Schuhmacher J (1995) Z Anorg Allg Chem 621:1373
89. Klüfers P, Schuhmacher J (1995) Z Anorg Allg Chem 621:19
90. Burger J, Klüfers P (1996) Z Anorg Allg Chem 622:1740
91. Benner K, Klüfers P, Schuhmacher J (1999) Z Anorg Allg Chem 625:541
92. Kästele X, Klüfers P, Kopp F, Schuhmacher J, Vogt M (2005) Chem Eur J 11:6326
93. Tacke R, Bertermann R, Burschka C, Dragota S (2005) Angew Chem Int Ed 44:5292
94. Kinrade SD, Deguns EW, Gillson AM, Knight CTG (2003) Dalton Trans 3713
95. Kinrade SD, Balec RJ, Schach AS, Wang J, Knight CTG (2004) Dalton Trans 3241
96. Lambert JB, Lu G, Singer SR, Kolb VM (2004) J Am Chem Soc 126:9611
97. Kästele X, Klüfers P, Tacke R (2006) Angew Chem Int Ed 45:3212
98. Blank G (1973) Acta Crystallogr Sect B: Struct Crystallogr Cryst Chem 29:1677
99. Rabinowitz IN, Kraut J (1964) Acta Cryst 17:159
100. Yang L, Wang Z, Zhao Y, Tian W, Xu Y, Weng S, Wu J (2000) Carbohydr Res 329:847
101. Yang L, Tao D, Sun Y, Jin X, Zhao Y, Yang Z, Weng S, Wu J, Xu G (2001) J Mol Struct 560:105
102. Neidle S, Gaffney PRJ, Reese CB (1998) Acta Crystallogr Sect C: Cryst Struct Commun C54:1191
103. Kumara Swamy KC, Kumaraswamy S (2001) Acta Crystallogr Sect C: Cryst Struct Commun C57:1147
104. Salazar-Pereda V, Martinez-Martinez L, Flores-Parra A, Rosales-Hoz MdJ, riza-Castolo A, Contreras R (1994) Heteroat Chem 5:139
105. Grainger CT (1981) Acta Crystallogr Sect B: Struct Crystallogr Cryst Chem B37:563
106. Hausherr-Primo L, Hegetschweiler K, Ruegger H, Odier L, Hancock RD, Schmalle HW, Gramlich V (1994) J Chem Soc Dalton Trans 1689
107. Hegetschweiler K (1997) Bol Soc Chil Quim 42:257
108. Hegetschweiler K, Hausherr-Primo L, Koppenol WH, Gramlich V, Odier L, Meyer W, Winkler H, Trautwein AX (1995) Angew Chem Int Ed Engl 34:2242
109. Hegetschweiler K, Raber T, Reiss GJ, Frank W, Worle M, Currao A, Nesper R, Kradolfer T (1997) Angew Chem Int Ed Engl 36:1964
110. Morgenstern B, Sander J, Huch V, Hegetschweiler K (2001) Inorg Chem 40:5307
111. Park YJ, Kim HS, Jeffrey GA (1971) Acta Crystallogr Sect B: Struct Crystallogr Cryst Chem 27:220
112. Gack C, Klüfers P (1996) Acta Crystallogr Sect C: Cryst Struct Commun C52:2972
113. Junicke H, Bruhn C, Stroehl D, Kluge R, Steinborn D (1998) Inorg Chem 37:4603
114. Jeffrey GA, McMullan RK, Takagi S (1977) Acta Crystallogr Sect B: Struct Crystallogr Cryst Chem B33:728
115. Klüfers P, Mayer P (1998) Acta Crystallogr Sect C: Cryst Struct Commun C54:583
116. Shannon RD (1976) Acta Crystallogr Sect A: Cryst Phys Diffr Theor Gen Cryst A32:751
117. Zhang B, Zhang S, Wang K (1996) J Chem Soc Dalton Trans 3257
118. Rajak KK, Rath SP, Mondal S, Chakravorty A (1999) Indian J Chem Sect A: Inorg Bio-inorg Phys Theor Anal Chem 38A:405
119. Rajak KK, Rath SP, Mondal S, Chakravorty A (1999) J Chem Soc Dalton Trans 2537
120. Rajak KK, Rath SP, Mondal S, Chakravorty A (1999) Inorg Chem 38:3283
121. Rajak KK, Baruah B, Rath SP, Chakravorty A (2000) Inorg Chem 39:1598
122. Takagi S, Jeffrey GA (1979) Acta Crystallogr Sect B: Struct Crystallogr Cryst Chem B35:902
123. Barili PL, Catelani G, Fabrizi G, Lamba D (1993) Carbohydr Res 243:165
124. Johansson MJ, Bergh A, Larsson K (2004) Acta Crystallogr Sect C: Cryst Struct Commun C60:o312
125. Herdin S, Kettenbach G, Klüfers P (2004) Z Naturforsch B: Chem Sci 59:134
126. Kuentzer D, Jessen L, Heck J (2005) Chem Commun 5653
127. Jessen L, Haupt ETK, Heck J (2001) Chem Eur J 7:3791
128. Cameron TS, Bakshi PK, Thangarasa R, Grindley TB (1992) Can J Chem 70:1623
129. Klaassen M, Klüfers P (1994) Acta Crystallogr Sect C: Cryst Struct Commun C50:686
130. Klaassen M, Klüfers P (1994) Z Anorg Allg Chem 620:1631
131. Bhadrui S, Sapre N, Khwaja H, Jones PG (1992) J Organomet Chem 426:C12
132. Steinborn D, Junicke H, Bruhn C (1997) Angew Chem Int Ed Engl 36:2686

133. Junicke H, Bruhn C, Kluge R, Serianni AS, Steinborn D (1999) J Am Chem Soc 121:6232
134. Meyer zu Berstenhorst B, Erker G, Kehr G, Froehlich R (2006) Dalton Trans 3200
135. Conn JF, Kim JJ, Suddath FL, Blattmann P, Rich A (1974) J Am Chem Soc 96:7152
136. Daniel FB, Behrman EJ (1976) Biochemistry 15:565
137. Feldman I, Rich KE (1970) J Am Chem Soc 92:4559
138. Feldman I, Rich KE, Agarwal RT (1970) J Am Chem Soc 92:6818
139. Angus-Dunne SJ, Batchelor RJ, Tracey AS, Einstein FWB (1995) J Am Chem Soc 117:5292
140. Sakurai H, Goda T, Shimomura S, Yoshimura T (1982) Nucleic Acids Symp Ser 11:253
141. Geraldes CFGC, Castro MM (1989) J Inorg Biochem 35:79
142. Richter J, Rehder D (1991) Z Naturforsch B: Chem Sci 46:1613
143. Crans DC, Harnung SE, Larsen E, Shin PK, Theisen LA, Trabjerg I (1991) Acta Chem Scand 45:456
144. Tracey AS, Leon-Lai CH (1991) Inorg Chem 30:3200
145. Zhang X, Tracey AS (1992) Acta Chem Scand 46:1170
146. Richter J, Rehder D, Wyns L, Haikal A (1995) Inorg Chim Acta 238:155
147. Galy J, Mosset A, Grenthe I, Puigdomenech I, Sjoeberg B, Hulten F (1987) J Am Chem Soc 109:380
148. Begum NS, Manohar H (1992) Polyhedron 11:2823
149. Cai SX, Keana JFW (1991) Bioconjug Chem 2:317
150. Erxleben A, Yovkova L (2006) Inorg Chim Acta 359:2350
151. Klüfers P, Mayer P (1997) Z Anorg Allg Chem 623:1496
152. Klüfers P, Mayer P (2007) Z Anorg Allg Chem 633:903
153. Kettenbach G, Klüfers P, Mayer P (1997) Macromol Symp 120:291
154. Parada J, Bunel S, Ibarra C, Larrazabal G, Moraga E, Gillitt ND, Bunton CA (2001) Carbohydr Res 333:185
155. Geisselmann A, Klüfers P, Kropfgans C, Mayer P, Piotrowski H (2005) Angew Chem Int Ed 44:924
156. Klüfers P, Schuhmacher J (1994) Angew Chem 106:1925
157. Fuchs R, Habermann N, Klüfers P (1993) Angew Chem 105:895
158. Klüfers P, Piotrowski H, Uhlendorf J (1997) Chem Eur J 3:601
159. Yamanari K, Nakamichi M, Shimura Y (1989) Inorg Chem 28:248
160. Burger J, Kettenbach G, Klüfers P (1995) Macromol Symp 99:113
161. Saalwaechter K, Burchard W, Klüfers P, Kettenbach G, Mayer P, Klemm D, Dugarmaa S (2000) Macromolecules 33:4094
162. Burchard W, Habermann N, Klüfers P, Seger B, Wilhelm U (1994) Angew Chem 106:936
163. Taylor GE, Waters JM (1981) Tetrahedron Lett 22:1277
164. Bilik V, Petrus L, Farkas V (1975) Chem Zvesti 29:690
165. Bilik V, Voelter W, Bayer E (1972) Ann Chem 759:189
166. Bilik V (1972) Chem Zvesti 26:372
167. Burger J, Gack C, Klüfers P (1996) Angew Chem Int Ed Engl 34:2647
168. Geisselmann A, Klüfers P, Pilawa B (1998) Angew Chem Int Ed Engl 37:1119
169. Wilbur DJ, Williams C, Allerhand A (1977) J Am Chem Soc 99:5450
170. Klüfers P, Kunte T (2001) Angew Chem Int Ed 40:4210
171. Klüfers P, Kunte T (2003) Chem Eur J 9:2013
172. Klüfers P, Kunte T (2002) Eur J Inorg Chem 1285
173. Klüfers P, Kunte T (2004) Z Anorg Allg Chem 630:553
174. Klüfers P, Kopp F, Vogt M (2004) Chem Eur J 10:4538
175. Lis T, Weichsel A (1987) Acta Crystallogr Sect C: Cryst Struct Commun C43:1954
176. Draffin SP, Duggan PJ, Fallon GD (2004) Acta Crystallogr Sect E: Struct Rep Online E60:o1520
177. Chapelle S, Verchere JF (1988) Tetrahedron 44:4469
178. van den BR, Peters JA, van BH (1994) Carbohydr Res 253:1

Part 5
Oligosaccharides

5.1 Oligosaccharides: Occurrence, Significance, and Properties

Zbigniew J. Witczak
Department of Pharmaceutical Sciences, Nesbitt School of Pharmacy, Wilkes University, Wilkes-Barre, PA 18766, USA
zbigniew.witczak@wilkes.edu

Abstract

This chapter is exclusively devoted to the occurrence, significance and physical and randomly selected chemical properties of various group of disaccharides, trisaccharides, tetrasaccharides, pentasaccharides and hexasaccharides.

Keywords

Disaccharides; *C*-Disaccharides; Thio-disaccharides; Oligosaccharides

Abbreviations

CD circular dichroism
ROA Raman optical activity

In: *Glycoscience*. Fraser-Reid B, Tatsuta K, Thiem J (eds)
Chapter-DOI 10-1007/978-3-540-30429-6_25:

1 Introduction

Naturally occurring oligosaccharides including disaccharides and trisaccharides have attracted enormous attention in the last few years. Multidisciplinary efforts of biochemists and medicinal chemists have led to new developments [1,2,3,4,5] in the field, which are based primarily on the phenomenal physical properties such as stability, crystallinity, and solubility. Chemical properties include selective reactivity towards various reagents such as oxidative and reducing agents and this has opened up new possibilities for functionalization and derivatization of value-added oligo- and polysaccharides.

This chapter is entirely devoted to the selected physical and chemical properties of those derivatives, which should always be considered as critical factors for any preparative manipulation or functionalization and conversion into desired target derivatives possessing highly practical application as pharmaceuticals, food ingredients, and industrial intermediates.

2 Occurrence of Oligosaccharides

Oligosaccharides occur in the living system as free compounds as well as being linked to many other plant constituents, such as phenols, saponins, steroids, flavonoids, as their corresponding glycosides.

Many natural oligosaccharide antibiotics, for example aminoglycoside antibiotics or macrolide antibiotics, are composed of specific monosaccharides such as deoxy-sugars and amino-sugars as well as cyclitols including streptamine. These monosaccharide components are linked via naturally occurring and important linkages such as (1→4)-, (1→2)-, (1→3)-, (1→6)-, and (1→1)-. The variety of natural sources of oligosaccharides is presented in ❯ *Table 1*.

3 Significance of Oligosaccharides: Industrial Utilization and Application

Among the disaccharides it is sucrose that reaches the status of commodity as it is a cheap industrial raw material. Any other di- or tri- saccharides are rarely used as chemical intermediates. In ❯ *Table 2* a list of existing and potential products through derivatization of sucrose is presented. With regards to high-value products, compounds with pharmacological activity can also be made from sucrose, for example, the anti-ulcerative drug Sucraflate (aluminum sucrose octasulfate) which reduces stomach acidity.

Lactose and maltose, readily available in large quantities from whey and starch, have some application in the form of their reduction products maltitol and lactitol, which are sweetening agents.

4 Disaccharides

Among the disaccharides the most important are sucrose, maltose, and lactose as cheap and valuable sources of renewable materials for industrial applications. Stability and reactivity are important factors for the selective functionalization and future application in the production

Table 1
Name, structure, and occurrence of selected natural oligosaccharides (Compiled from Carbohydrates (1987) Collins PM (ed) Chapman and Hall Chemistry)

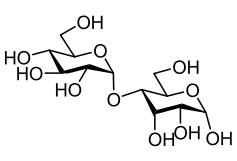
Name	Structure	Occurrence
Disaccharides		
Cellobiose	β-D-Glc*p*-(1→4)-D-Glc*p*	Cellulose unit
Digilanidobiose	β-D-Glc*p*-(1→4)-D-Digitoxose	Cardiac glycosides of Lanta
Gentiobiose	β-D-Glc*p*-(1→6)-D-Glc*p*	Amygdalin, crocin glycosides
Isomaltose	α-D-Glc*p*-(1→6)-D-Glc*p*	Amylopectin unit of starch
Kojibiose	α-D-Glc*p*-(1→2)-D-Glc*p*	Free sugar in honey
Laminaribiose	β-D-Glc*p*-(1→3)-D-Glc*p*	Free sugar in honey
Lactose	β-D-Gal*p*-(1→4)-D-Glc*p*	Free sugar in milk
Lactulose	β-D-Gal*p*-(1→4)-D-Fru*f*	Degradation product of lactose
Lycobiose	4-β-D-Glc*p*-(1→4)-D-Gal*p*	Sugar component of glycoside tomatin
Maltose	α-D-Glc*p*-(1→4)-D-Glc*p*	Free sugar of malt, beer, fruits, honey, unit of starch
Maltulose	α-D-Glc*p*-(1→4)-D- Fru*f*	Free sugar in malt, beer, honey
Mannobiose	β-D-Man*p*-(1→4)-D-Man*p*	Unit of the guaran polysaccharide
Melibiose	α-D-Gal*p*-(1→6)-D-Glc*p*	Free sugar in Cocoa beans, degradation product of raffinose
Neohesperidose	α-L-Rha*p*-(1→2)-D-Glc*p*	Sugar component of flavonoid glycosides, neohesperidin, naringin
Neotrehalose	α-D-Glc*p*-(1→1)-β-D-Glc*p*	Free sugar in koji extract
Nigerose	α-D-Glc*p*-(1→3)-D-Glc*p*	Free sugar in honey, beer
Palatinose	α-D-Glc*p*-(1→6)-D-Fru*f*	Microbiological conversion of sucrose
Primverose	β-D-Xyl*p*-(1→6)-D-Glc*p*	Free sugar in carob fruits, glycoside of *Primula officinalis*
Robinobiose	α-L-Rha*p*-(1→6)-D-Gal*p*	Sugar component of Robinin
Rutinose	α-L-Rha*p*-(1→6)-D-Glc*p*	Sugar component of flavonoid glycosides, hesperidin
Scillabiose	β-D-Glc*p*-(1→4)-l-Rha*p*	Sugar component of glycoside *Scilla maritime*
Solabiose	β-D-Glc*p*-(1→3)-β-D-Gal*p*	Sugar component of glycoside β-Solanin
Sophorose	β-D-Glc*p*-(1→2)-D-Glc*p*	Sugar component of glycoside *Sophora Japonica L* and Stevioside
Strophanthobiose	β-D-Glc*p*-(1→4)-D-Cymarose	Sugar component of K-strophantin
Sucrose	β-D-Fru*f*-(2→1)-D-Glc*p*	Free sugar in sugar cane, sugar beet, plants, fruits
Trehalose	α-D-Glc*p*-(1→1)-α-D-Glc*p*	Free sugar in mushrooms and in the blood of a few insects and of grasshoppers
Turanose	α-D-Glc*p*-(1→3)-D-Fru*p(f)*	Free sugar in honey
Vicianose	α-L-Arab*p*-(1→6)-D-Glc*p*	Sugar component of glycoside Viciaanin
Trisaccharides		
Centose	α-D-Glc*p*-(1→4)-α-D-Glc*p*-(1→2)-Glc*p*	Free sugar in honey
Erlose	α-D-Glc*p*-(1→4)-α-D-Glc*p*-(1→2)-β-Fru*f*	Free sugar in honey

Table 1 (continued)

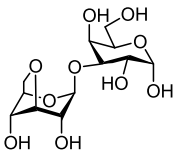
Name	Structure	Occurrence
Gentianose	β-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→2)-β-Fru*f*	Free sugar in the rhizomes of Gentiana
Isomaltotriose	α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→6)-D-Glc*p*	Free sugar in honey
Isopanose	α-D-Glc*p*-(1→4)-α-D-Glc*p*-(1→6)-D-Glc*p*	Free sugar in honey
1-Kestose	α-D-Glc*p*-(1→2)-β-D-Fru*f*-(1→2)-β-D-Fru*f*	Free sugar in honey
6-Kestose	α-D-Glc*p*-(1→2)-β-D-Fru*f*-(6→2)-β-D-Fru*f*	Free sugar in honey
Lycotriose	β-D-Glc*p*-(1→2)-β-D-Glc*p*-(1→4)-Gal*p*	Sugar component of glycoside Tomatin
Maltotriose	α-D-Glc*p*-(1→4)-α-D-Glc*p*-(1→4)-D-Glc*p*	Degradation product of amylose
Manninotriose	β-D-Gal*p*-(1→6)-β-D-Gal*p*-(1→6)-α-D-Glc*p*	Free sugar in Manna, degradation product of stachyose
Melezitose	α-D-Glc*p*-(1→3)-β-D-Fru*f*-(2→1)-α-D-Glc*p*	Free sugar in honey, manna, exudates of lime and pine
Panose	α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→4)-D-Glc*p*	Free sugar in honey, degradation product of amylopectin
Planteose	*O*-β-D-Gal*p*-(1→6)-β-D-Fru*f*-(2→1)-α-D-Glc*p*	Free sugar in seeds of Plantago varieties
Raffinose	α-D-Gal*p*-(1→6)-α-D-Glc*p*-(1→2)-β-Fru*f*	Free sugar in plants and leguminose seeds: soja beans, mung beans
Solatriose	β-D-Glc*p*-(1→3)-β-D-Gal*p*-(1→2)-α-L-Rha*p*	Sugar component of glycoside solanin
Strophantotriose	β-D-Glc*p*-(1→6)-β-D-Glc*p*-(1→4)-D-Cymarose	Sugar component of h-strophanthoside
Umbilliferose	α-D-Gal*p*-(1→2)-α-D-Glc*p*-(1→2)-β-Fru*f*	Free sugar in the roots of Umbilliferae
Tetrasaccharides		
Lychnose	α-D-Gal*p*-(1→6)-α-D-Gal*p*(1→2)-β-D-Fru*f*-(1→1)-α-D-Gal*p*	Free sugar in rhizomes of *Lychnis doica*
Maltotetraose	α-D-Glc*p*-(1→4)-α-D-Glc*p*-(1→4)-D-Glc*p*-(1→4)-D-Glc*p*	Free sugar in starch hydrolysates
Stachyose	α-D-Gal*p*-(1→6)-α-D-Gal*p*(1→6)-α-D-Glc*p*(1→2)-β-D-Fru*f*	Free sugar in rhizomes of the Japanese artichoke
Pentasaccharides		
Verbascose	α-D-Gal*p*-(1→6)-α-D-Gal*p*(1→6)-α-D-Gal*p*-α-D-(1→6)-Glc*p*-(1→2)-β-D-Fru*f*	Free sugar in the rhizomes of wool plant, and seeds of cow peas, winged and lima beans
Hexasaccharides		
Ajugose	[α-D-Gal*p*-(1→6)]$_4$-α-D-Glc*p*(1→2)-β-D-Fru*f*	Free sugar in the roots of *Verbascum thapsiformae* and *Ajuga nipponensis*

Table 2
Commercial and potential products from sucrose

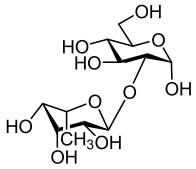
D-Glucose + D-Fructose	
D-Sorbitol (D-Glucitol) +D-Mannitol	
Products of fermentation or enzymatic degradation	Acids (gluconic, itaconic, levulinic, citric)
	Ethanol
	Antibiotics
	Favor enhancers
	Gums (Dextrans, fructans, xanthan)
	Polyhydroxybutyrate and analogs
	Oligosaccharides (Neosugars, isomaltulose)
Sucrose tricarboxylic acid	
Sucrose dialdehyde	
Sucrose anhydrides	
Sucrose polyesters (***Olestra***)	
Derivatives:	Carboxylic acid esters
	Sulfonic acid esters
	Ethers
	Polyurethans
	Deoxychloro derivatives (***Splenda***)
	Hydroxymethylfurfural



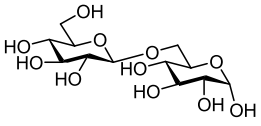
Mannobiose



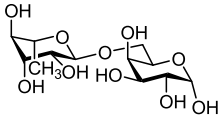
Neoagarobiose



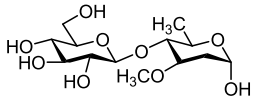
Neohesperidose



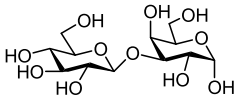
Primeverose



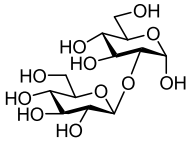
Robinobiose



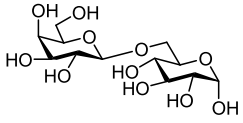
Strophantobiose



Solabiose



Sophorose



Vicianose

Scheme 1

of value-added and consumer products [6,7] such as "Olestra" and "Sucralose". Owing to the rapid development of modern analytical method, namely mass spectrometry and NMR spectroscopy, an immense number of oligo- and polysaccharides have been characterized in detail. Taking into account the fact that even a small number of different monosaccharides can be combined to form a huge number of distinct oligosaccharides and that each oligosaccharide can also adopt several conformations, it seems certain that they have many varied physicochemical properties. This section summarizes some of the important disaccharides with the specific elements of their chemical properties and a selection of the protecting groups that were developed for the synthesis of oligosaccharides and complex carbohydrates. Some of the representative disaccharides are shown in *Scheme 1*.

4.1 Chemical Properties, Reactivity, and Protecting Groups

Disaccharides contain a high concentration of polar functional groups and in many cases the preferred strategy for the synthesis depends on the different levels of reactivity of primary and secondary hydroxyl groups. More complex sugar moieties are found in oligosaccharide-containing antibiotics. For example amino, nitro, deoxy-sugars, branched-chain and thio-sugars are found as constituents of complex glycosylated derivatives such as anthracyclines, calichemicins, aureolic acid, orthosomycins and most importantly, aminoglycosides. The physicochemical properties of these disaccharide components of complex oligosaccharides are listed in *Table 3*.

Several of the complex oligosaccharides (including macrolide and aminoglycoside antibiotics) and glucosaminoglycans (low molecular weight heparin and heparan) have become very valuable drugs of critically important commercial value. This is one of the reasons that synthesis of these important oligosaccharides and their specifically functionalized analogs is one of the main focuses of new developments in the fields of glycomimetics and particularly the chemical aspects of glycobiology.

Synthetic approaches to these classes of higher oligosaccharides normally demand a regiospecific protecting group in order to leave only those primary or secondary hydroxyl groups free that are to be further functionalized. This particular strategy may give enhanced (or targeted) hydroxyl group reactivity because of the increased accessibility of the targeted hydroxyl group to be functionalized.

One of the first and most important protecting groups to emerge for facilitating selective transformations was the isopropylidene acetal and today, acetals remain the principal means of preserving 1,2- and 1,3-diols. A number of reviews dealing with the various aspects of the chemical reactivity [56,57,58,59] as well as providing extensive accounts on stability, selective formation [57], and selective cleavage [59] have been published.

Acetals, including benzylidene, cycloxehylidene, and cyclopentylidene, reveal virtually unlimited stability to basic conditions, which is an important chemical property that is crucial for any synthetic strategy requiring those conditions. On the other hand all the acetals are quite sensitive towards acid. This specific chemical property derives from participation of a lone pair on the adjacent oxygen atom in the cleavage of a resonance stabilized and protonated intermediate, known as an oxonium ion.

Table 3
Physicochemical data of selected disaccharides (Compiled from Carbohydrates (1987) Collins PM (ed) Chapman and Hall Chemistry)

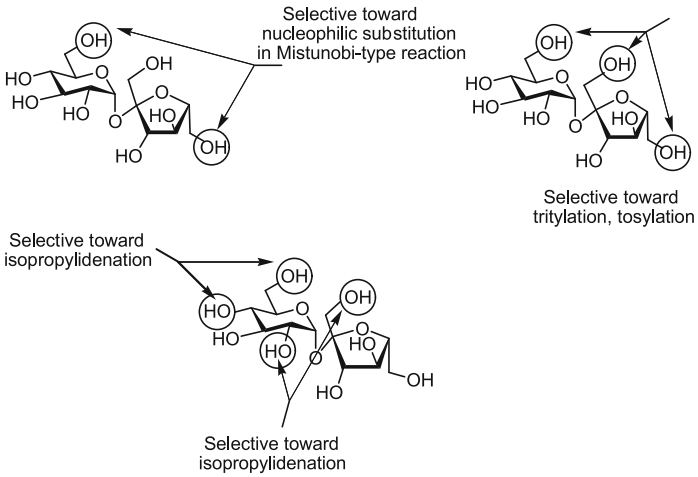
Compound		M.p. (°C)	[a]D (H_2O)	Refs.
Isomaltose	α-D-Glc*p*-(1→6)-D-Glc*p*	120	+120°	[8,9]
Isoprimeverose	α-D-Glc*p*-(1→6)-D-Glc*p*	200–201	+151.3°→+121.3°	[10,11]
Cellobiose	β-D-Glc*p*-(1→4)-D-Glc*p*	225	+24°→+35.2°	[12,13,14]
Gentiobiose	β-D-Glc*p*-(1→4)-D-Glc*p*	195	+32°→+9.6°	[15]
Laminaribiose	β-D-Glc*p*-(1→3)-D-Glc*p*	205	+18°(c 2.0)	[16]
Nigerose	β-D-Glc*p*-(1→4)-D-Glc*p*	156	+120	[17,18]
Sophorose	β-D-Glc*p*-(1→4)-D-Glc*p*	196–198	+33°→+19°	[19,20,21,22]
Kojibiose	β-D-Glc*p*-(1→4)-D-Glc*p*	187	+162°→137°	[23,24,25]
Trehalose	β-D-Glc*p*-(1→1)-D-Glc*p*			
	α, α	97	+178.3°	[26]
	α, β	210–220	+95°	[26]
	β, β	135–140	–40.2°	[26]
Sucrose	β-D-Glc*p*-(2→1)-D-Glc*p*	*Oct-Ac* 69	+59.6°	[27]
Turanose	α-D-Glc*p*-(1→3)-D-Fru*p(f)*	157	+22°→+75.3°	[28]
Maltulose	β-D-Glc*p*-(1→4)-D-Glc*p*	113–115	+58°→+64°	[29]
Palatinose	β-D-Gal*p*-(1→4)-D-Glc*p*	120	+120°	[30]
Primeverose	β-D-Xyl*p*-(1→6)-D-Glc*p*	194–197	+23.8°→–3.4°(c 2.5)	[31]
Leucrose	α-D-Glc*p*-(1→5)-D-Fru*p*	161–163	–8.2°→+7.6°(c 4.0)	[32,33]
Lycobiose	β-D-Gal*p*-(1→4)-D-Gal*p*	246–247 dec	+70°→41.5°(c 1.0)	[34]
Lactose	β-D-Gal*p*-(1→4)-D-Glc*p*	202 (monohyd.)		[35,36]
	α-form	252 °C (anhydr.), (monohydrate)	+85°→52.6°(c 8.0)	
	β-form	252(anhydr.)	+55.4°(c 4.0)	
Lacto-*N*-biose	β-D-Gal*p*-(1→2)-D-GlnAc*p*	197	+22°→+7°24 h, (c 1.0)	[37]
Lactulose	β-D-Glc*p*-(1→4)-D-Glc*p*	168–171	–51.4°(c 4.0)	[38]
Neolactose	β-D-Gal*p*-(1→4)-D-Alt*p*	190 (dec.)	+34.6°→+35.5°	[39]
Neohesperidose	α-L-Rhm*p*-(1→2)-D-Glc*p*	191–192	–53.3→–3.9°(c 3.0)	[40]
Neoagarobise	3,6-anhydro-α-L-Gal*p*-(1→3)-D-Gal*p*	207–208	+34.4°→+20.3°	[41]
Nigerose (Sakebiose)	α-D-Glc*p*-(1→3)-D-Glc*p*	156	+125°→+138°(c 2.0)	[42]
Melibiose	α-D-Galp-(1→6)-D-Glc*p*	85 (dihydrate), 179–181(hydr.)	+110.5°→126.5°(+143°)	[43]
Planteobiose (Melibiulose)	α-D-Gal*p*-(1→6)-D-Fru*f*	85	+55.5°	[44]
Robinobiose	β-L-Rhm*p*-(1→6)-D-Gal*p*		0°→+2.7°	[47]
Rufinose	β-D-Glc*p*-(1→4)-D-Glc*p*	120	+120°	[48]
Rutinose	β-L-Rhm*p*-(1→6)-D-Glc*p*	189–192 dec.	+3.2°→–0.8°	[49]
Sambubiose	β-D-Xyl*p*-(1→2)-D-Glcp	202–203	+32°→+17°(c 1.0)	[50]

Table 3 (continued)
Physicochemical data of selected disaccharides (Compiled from Carbohydrates (1987) Collins PM (ed) Chapman and Hall Chemistry)

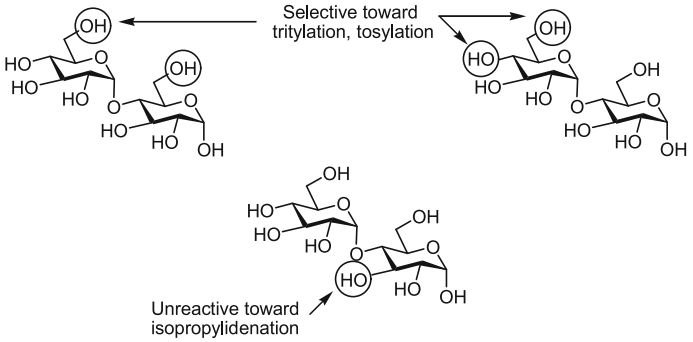
Compound		M.p. (°C)	[a]D (H_2O)	Refs.
Strophantobiose	β-D-Glc*p*-(1→4)-D-Glc*p*	144–146	+33.8°(c 2.1)	[51]
Mannobiose	β-D-Glc*p*-(1→4)-D-Man*p*	193–194	–7.7°→–2.2°	[52,53]
Vicianose	β-D-Glc*p*-(1→4)-D-Glc*p*	210	+56.5°	[54]
Solabiose	β-D-Glc*p*-(1→4)-D-Glc*p*	203–205	+42°(c 0.8)	[55]

Since the various protecting groups, for example *O*-protection by acetalization, acylation, allylation, and benzylation, can be removed under completely different reaction conditions (acidic/basic, reductive) it is relatively easy to obtain the desired substitution pattern via a suitable and relatively stable protection group strategy. Extensive protection group strategies have been based on the different level of reactivity of primary and secondary hydroxyls. Classical examples of disaccharide functionalization through selective protection/deprotection are sucrose isopropylidenation [60], selective tritylation, or tosylation [61]. Particularly important and practically useful is functionalization of sucrose at 6 and 6′ during Mitsunobu-type nucleophilic substitution as demonstrated during the clever synthesis of 1-deoxynojirimycin by de Raadt and Stutz [62]. This very important approach allows selective substitution of two of the three primary hydroxyl groups (*Scheme 2*).

The totally different steric environments of the two primary hydroxyl groups in maltose allows the selective functionalization of the C-6′ position due to its greater steric hindrance relative to the C-6 position. Thus, maltose can be selectively tosylated and tritylated exclusively on the C-6′ position. Also, configurational factors create unusual selectivity during acetylation or benzylation of maltose, confirming that the C-3 position is the least reactive secondary group and remains free and unacetylated under the reaction conditions.



Scheme 2



Scheme 3
Selective functionalization of maltose

Thus, the differential reactivity of the hydroxyl groups at various positions can be a considerable factor for the preparation of suitably functionalized synthetic intermediates. The general rules of conformation as it influences reactivity should be followed in any synthetic strategy.
Firstly, equatorial hydroxyl groups are more reactive and much easier to functionalize through conventional esterification and etherification. New functional *equatorial* groups are more susceptible to cleavage and to attack by hydrolytic agents than those on *axial* positions.
Secondly, the primary hydroxyl groups on C-6, being essentially exocyclic, are far more reactive than any of the ring/chain hydroxyls and can be functionalized further without the necessity of blocking other functional groups.
Thirdly, the presence of particularly bulky functional groups can render adjacent secondary hydroxyl groups relatively unreactive even when they are in equatorial orientations. The previously mentioned unreactivity of the C-3 position during maltose acetylation is one of the classical examples of such a peculiar rule.
Moreover, the electronegativity of the ring oxygen makes substitution at the neighboring hydroxyl more likely than at either position C-3 or position C-4 as is observed during benzylation of methyl α-D-maltoside or even methyl α-D-glucoside. This is a general effect of the ring oxygen during the reaction carried out at the C-2 hydroxyl group. On the other hand, it can also be predicted that electron deficiency will promote the efficiency of the reaction, the C-3 hydroxyl group is more reactive than the hydroxyls at C-3 and C-4. This is most likely due to preferential reaction at C-2 followed by steric hindrance at C-3 if C-4 is axial and functionalized as well.
Neighboring group participation is also another important factor for predicting the reactivity of secondary hydroxyl groups, particularly at the C-2 position. Under basic conditions, the C-2 hydroxyl tends to be more acidic than the C-3 hydroxyl and this may be advantageously exploited in certain cases such as partial benzylation under phase-transfer catalysis. The latter reaction conditions also contribute to the relatively good selectivity for substitution at a primary hydroxyl group in preference to a secondary one at either C-3 or C-4.
The protection of the anomeric hydroxyl group is of primary importance for the selective functionalization of other remaining secondary hydroxyls and must be treated separately from other protections. It can be considered as either protection of the acetal hydroxyl group or protection of the acetal, C-1, carbon.

In the synthesis of the precursor of a glycosyl donor, the anomeric hydroxyl group is usually protected as an acetal. Thus, the protective group must be removed in order to introduce the leaving group at the anomeric position in this case.

The introduction of thioglycosides to carbohydrate chemistry reversed the idea of protection of the anomeric hydroxyl group. Thioglycosides as protective groups of acetal carbon are generally stable under Lewis acidic conditions used to activate other leaving groups in glycosylation reactions. Most importantly, they can be selectively activated under oxidative conditions to give an active glycosylating agent, the oxocarbonium ion. Thioglycosides therefore are extremely useful not only as protective groups of acetal carbons but also as excellent leaving groups [63].

Recently, a new intramolecular-type approach for the regioselective debenzylation or protection of a secondary hydroxyl group in disaccharides has been reported by Madsen and Bols [64].

4.2 ^{1}H and ^{13}C NMR Spectroscopy

The specific correlations of functional groups with the chemical shifts of carbon and associated protons are always important tools for structure elucidation of simple sugars as well as complex oligosaccharides. The conformational behavior [65,66] of a particular disaccharide and oligosaccharide in solution is determined not only by intramolecular interactions but also by potential interactions that exist between the oligosaccharide and its environment. In particular, the nature of the solvent or more generally the environment in which the compound finds itself, can profoundly affect its geometry. Two such important factors, namely the fully solvated state (i. e. the molecule free in solution) and the molecule "solvated" by a macromolecular receptor (i. e. the molecule bound to the active site of the protein) contribute to the preferred conformation.

Nonselective spin-lattice relaxation rates (s^{-1})* for the anomeric protons of selected oligosaccharides are listed in *Table 4*.

An extensive NMR study of various 1→4-linked disaccharides with D- and L- monosaccharides mixed in disaccharides has been reported [66]. Each group of disaccharides exhibits specific changes in chemical shifts due to glycosidations of the C-4 hydroxy group, which are different between the groups. The differences in chemical shifts were correlated to the preferred conformations of the disaccharides and used to predict the ^{1}H and ^{13}C NMR chemical shifts of 1→4-linked oligo- and polysaccharides.

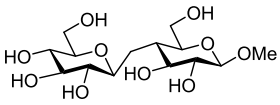
Table 4
Nonselective spin-lattice relaxation rates (s^{-1})* for the anomeric protons of selected oligosaccharides

	Glycose residue			Glycosyl group		
Compound	H-1α	H-1β	H-1α/H-1β	H-1α	H-1β	H-1α/H-1β
Maltotriose	0.59	1.37	2.3	1.87		3.2
Raffinose				1.16		
Stachyose				1.59		

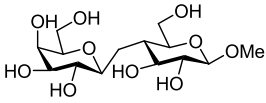
*Compiled from Hall LD, Preston CM (1973) *Carbohydr Res* 29:322; ibid (1976) 49:3

4.3 *C*-Disaccharides

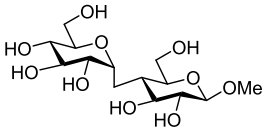
C-disaccharides [67] are special classes of disaccharides with a nonhydrolyzable C–C bridge at various positions. *C*-disaccharides are extremely stable entities that will not be fragmented by biological or chemical hydrolysis [68,69]. Perhaps the major difference between *C*- and *O*-glycosides is found within chemical reactivities. The major difference is the absence of anomeric effects for *C*-disaccharides and incapability of forming hydrogen bonds. Another large difference between physical constants is in the dipole moments. Interestingly, *O*- and *C*-disaccharides exhibit similar coupling constants in their ^{1}H NMR spectra.



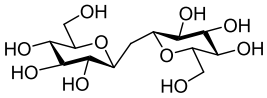
C-Cellobioside



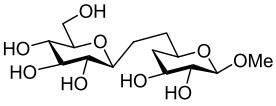
C-Lactoside



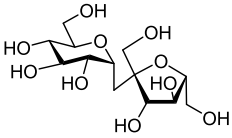
C-Maltoside



C-Trehalose



C-Gentiobioside



C-Sucrose

Scheme 4

Table 5
Physicochemical data of selected *C*-disaccharides[a]

Compound		M.p. (°C)	[α]D (H_2O)	Refs.
(1–1)-C-disaccharides				
C-trehalose	α, α	225	+81.8 (c 1.59)	[72]
	α, β		+33.0 (c 1.02)	
	α, β		+30.7 (c 0.73)	[72]
(1–2)-C-disaccharides				
C-bis-galactose		185		[69]
C-sucrose				
C-sucrose octa-Ac-		103–105	+21.6 (c 0.3 MeOH)	[73]
(1–4)-C-disaccharides				
Methyl *C*-cellobioside			+63.2 (MeOH)	[74]
Methyl *C*-maltoside			53.3 (MeOH)	[74]
(1–6)-C-disaccharides				
Methyl *C*-gentiobioside				[75]

[a]Compiled from original references

C-Disaccharides have become important target molecules because they are potential reversible inhibitors of glycosidases and disaccharidases [69]. Some of the important representatives of this class of derivatives are listed in ❯ *Scheme 4*.

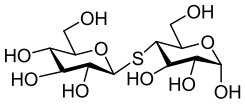
Interestingly, conformational studies of synthesized *C*-disaccharides [70,71,72,73,74,75,76] still furnish new evidence regarding the unpredictability or at least difficulty to properly verify or predict conformation based on steric effect. As an example, the NMR study of the conformation of *C*-glycofuranosides and *C*-glycopyranosides reported by Brakata et al. [77] clearly demonstrated that the $^{1}H_{1'-2'}$ coupling constants were limited in their abilities to define conformations. Other analogs of *C*-disaccharides including a special class of *aza*-*C*-disaccharides, which can be regarded as being closest in structure to the natural sequences, have also been synthesized. Selected representatives of this important class of nonhydrolyzable disaccharides are listed in ❯ *Table 5*.

4.4 Thiodisaccharides

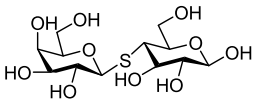
One way to increase the stability of the anomeric linkage is to replace the inter-glycosidic oxygen atom with the sulfur to furnish thio-disaccharides or with a methylene group to furnish a *C*-disaccharide. There has been some debate over the validity of this concept since the conformation may not be the same as that of the parent *O*-saccharide [78]. Indeed, the conformation of thio-disaccharides including (1→2-thiodisacharides) based on the molecular model, clearly confirms the stereogeometry and the difference to the parent *O*-saccharide [79]. All the measurements are similar to the calculation for the sulfur bridge of other thio-disaccharides and are substantially different to that of the oxygen counterpart as previously postulated in the literature [80,81].

Despite these differences, thio-disaccharides with sulfur as the connecting bridge of two sugar moieties are popular as convenient probes for enzyme-inhibition studies [82,83,84,85]. These biological and physicochemical properties are attributed to the different electronic properties of the sulfur (electronegativity) and oxygen atoms. The specific chemical character of the sulfur, particularly the size and the electron density of the sulfur atom, which is more dispersed, creates significant differences. Moreover, the C–S bond is longer (ca. 1.8 Å) and the C-S-C angle (ca. 95–100°) is smaller than the corresponding oxygen-containing structures. One thus can apply rational design rather than empirical experimentation in the optimization of synthetic targets, i. e. thio-disaccharides. When comparing the chemical properties of oxygen and sulfur analogs, two other factors must be considered. Firstly, their ability to undergo a wide variety of reduction and oxidation reactions makes them convenient intermediates in many synthetic approaches to other important disaccharides such as deoxy or branched-chain analogs. Secondly, sulfur derivatives such as thiols can be oxidized easily to form disulfides. These two chemical characteristics of thio-disaccharides are primary factors in considering any synthetic strategies and are the potential biochemical mechanisms for forming natural products containing sulfur. Selected representatives of this family of nonhydrolyzable thio-disaccharides are presented in ❯ *Scheme 5*.

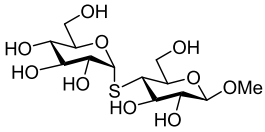
Conformational equilibria of thio-oligosaccharides revealed that the most important thiodisaccharide (4-thio-α-maltoside) has a conformation similar to that of the parent maltosyl residue and that the thioglycoside linkage presents a high degree of flexibility. Interestingly, the C–S



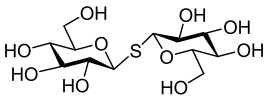
4-Thiocellobiose



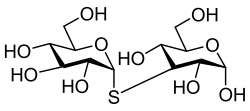
4-Thiolactose



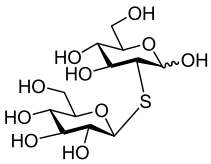
4-Thiomaltoside



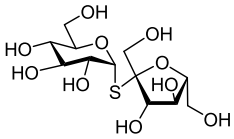
1-Thiotrehalose



3-Thionigerose



2-Thiosophorose



1-Thiosucrose

Scheme 5

Table 6
Physicochemical data of selected thiodisaccharides[a]

Compound	M.p. (°C)	[α]D (H_2O)	Ref.
(1–1)-thiodisaccharides			
1-thiotrehalose	175 + 133		[88]
(1–2)-thiodisaccharides			
2-thiokojibiose	165	+113°	[90]
2-thiosophorose	173–175	–8°+ → –13.1°	[91]
3′-deoxy-thiosophorose		–10.26 → –16.2°	
(1–3)-thiodisaccharides			
(1–4)-thiodisaccharides			
thiocellobiose	175	–16°	[88]

[a]Compiled from original references

bond in 4-thio-α-maltoside is longer than the corresponding C–O bond by ca. 0.4 Å. Among the first thiodisaccharides, thio-trehalose was synthesized in 1914. Other physical properties of new important analogs of thio-disaccharides are listed in *Table 6*.

5 Physical Properties

Specific optical rotation, circular dichroism (CD), and the newly developed exciton chirality method [96] for functionalized carbohydrates bearing specific chromophores are particularly

useful for the determination of absolute configuration not only of monosaccharides but also more complex carbohydrates including disaccharides.

Among other physical properties of the above-mentioned classes of disaccharides, one of the particularly important properties is Raman optical activity (ROA).

Carbohydrates (disaccharides) in aqueous solution are favorable examples for ROA studies furnishing an informative band structure over a wide range of the vibrational spectrum. Their complex and highly coupled normal modes generate strong ROA bands that produce patterns characteristic of the various types of structural units. ROA data can be used to confidently specify the central component of carbohydrate stereochemistry. For disaccharides [92] ROA can provide the nature and conformation of the glycosidic link, and can also probe extended secondary structures of oligosaccharides [93].

The conformation of di-, oligo-, and polysaccharides center on the glycosidic link, for which several distinct ROA signatures have been identified. The most important and critical one is a couplet centered at $\sim 917\,\mathrm{cm}^{-1}$, positive at lower and negative at higher wavenumbers, which originates in C-O-C stretch modes of the common α-(1→4) glycosidic link present in maltose.

5.1 Polarimetry

The determination of sugars by polarimetry is carried out preferably with analytically pure derivatives in higher concentration. Mono-, di-, and smaller oligosaccharides are optically active as a result of the presence of their chiral centers and rotate the plane of the polarized light. The highly specific rotation of disaccharides is dependent not only on the wavelength of the light and temperature, but also to a small extent on the concentration as shown by two common examples of simple disaccharides such as maltose and sucrose.

Maltose behaves as a normal reducing sugar, showing mutarotation and existing finally in aqueous solution as an equilibrated mixture of the α-form and the β-form. Modern ideas about connections between optical rotation and structure are reported in the literature [105].

6 Oligosaccharides

Oligosaccharides such as those found in plant material (see *Table 1*) or on surfaces of cells are assembled through the actions of specific enzymes. The enzymes involved in these biochemical processes belong to various classes of glycosyl transferases and glycosidases. The stability of glycosidic linkages as well as the sequence of assembling monosaccharide components as designed by nature are extremely difficult to reproduce in the laboratory. This is one of the main reasons that laboratory oligosaccharide synthesis starting from basic monosaccharide units through specifically protected di- and trisaccharide fragments is developing so rapidly [106,107,108,109,110,111,112]. The synthetic approaches to oligosaccharides normally demands regiospecific protection group (of disaccharides) schemes which will leave only those hydroxyl groups free that are to be further functionalized. Alternatively, certain other hydroxyl groups may be left unprotected. Simultaneously, however, such a strategy requires regioselective functionalization through glycosylation by attaching another larger fragment of sugar moiety, as was developed in the synthesis of tri- and tetrasaccharides. Oligosac-

charides bound to proteins, as complex glycoprotein carbohydrate groups, are classified into *N*- and *O*-glycoprotein based on the presence of functional groups at different positions. Specific sequences of the core fragments can be determined by ^{13}C-NMR methods [113]. Moreover, combinations of homo- and hetero-nuclear shift correlated 2D NMR techniques have been used to assign ^{1}H and ^{13}C signals of oligosaccharides related to arabinoxylan [114]. ^{13}C-NMR chemical shifts of oligosaccharides have been compiled and tabulated according to their linkages [115]. ^{1}H-NMR methods were used to determine the branching patterns of ceramide pentadecasaccharides with distinctive differences between the parent spectrum and that of partially degraded products [116].

6.1 Trisaccharides

The most common trisaccharides were isolated from natural sources and were identified as partial components of complex oligosaccharides and are listed below in ❯ *Table 7*.

7 Complex Oligosaccharides

Complex oligosaccharides as complex carbohydrates are classified into specific classes based on the presence and sequence of various multiple sugar moieties linked via functional groups at totally different positions. One of the specific functionalities of tetra-, penta-, and hexasaccharides is their water solubility, highly dependant on the presence of a number of *O*-glycosylated and *N*-acetamido moieties of individual fragments. One well-known example of this classical functionality is heparin pentasaccharide. The physical properties of the most important selected complex oligosaccharides, such as tetra-, penta-, and hexasaccharides, are listed in ❯ *Table 8*.

7.1 Tetrasaccharides

Few tetrasaccharides are well known; they have been isolated from natural sources such as plant and animal tissues or through fermentation and enzymatic synthesis. The biological, chemical, or physical properties of an oligosaccharide are largely determined by what is exposed on the outer surface. Most interactions with other molecules will occur at the surface of the molecules.

The conformational studies and physical properties of tetrasaccharides along with representative examples of disaccharides and *O*- and *N*-oligosaccharides have been reported by Meyer [65].

7.2 Pentasaccharides

Among the most common pentasaccharides, the most important is verbascose:*O*-α-D-Gal*p*-(1→6)-*O*-α-D-Gal*p*(1→6)-*O*-α-D-Gal*p*-*O*-α-D-(1→6)-Glc*p*-(1↔2)-α-D-Fru*f* isolated from plant material, particularly from the rhizomes of wool plant and the seeds of cow peas (see also ❯ *Table 1*).

Table 7
Physicochemical data of selected trisaccharides[a]

Compound		M.p. (°C)	[α]D (H_2O)	Refs.
Maltotriose				
α-D-Glc*p*-(1→4)-α-D-Glc*p*-(1→4)-D-Glc*p*	*per-Ac*	134–136	+160°	[117]
Panose				
α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→4)-D-Glc*p*		223 dec	+163°→ +155°(c 4.0)	[118]
Isopanose				
α-D-Glc*p*-(1→4)-α-D-Glc*p*-(1→6)-D-Glc*p*		120	+120°	[119]
Isomaltotriose				
α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→6)-D-Glc*p*		120	+120°	[120]
Centose				
α-D-Glc*p*-(1→4)-α-D-Glc*p*-(1→2)-D-Glc*p*		120	+120°	[121]
Gentiotriose				
β-D-Glc*p*-(1→6)-β-D-Glc*p*-(1→6)-β-D-Glc*p*	*per-Ac*	212	–6.5°(c 2.0)	[122]
Gentianose				
β-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→2)-β-D-Fru*f*		220	+30.8°(c 0.5)	[123]
1-Kestose (Isokestose)				
α-D-Glc*p*-(1→2)-β-D-Fru*f*-(1→2)-β-D-Fru*f*		82–88	+29.2°	[124]
6-Kestose				
α-D-Glc*p*-(1→2)-β-D-Fru*f*-(6→2)-β-D-Fru*f*		145	+27.3°(c 2.19)	[125]
Kojitriose				
α-D-Glc*p*-(1→2)-α-D-Glc*p*-(1→2)-β-D-Glc*p*		228–230 dec	+150°→156°(c 1.7)	[126]
Lactulosucrose				
α-D-Gal*p*-(1→6)-α-D-Glc*p*-(1→2)-β-D-Fru*f*		120	44.1 (c 2.0)	[127]
Melizitose				
α-D-Glc*p*-(1→3)-β-D-Fru*f*-(2→1)-α-D-Glc*p*		153–154	+88.2°	[128]
Raffinose				
α-D-Gal*p*-(1→6)-α-D-Glc*p*-(1→2)-β-D-Fru*f*		118	+123°(c 2.0)	[129]
Umbeliferose				
α-D-Gal*p*-(1→2)-α-D-Glc*p*-(1→2)-β-D-Fru*f*		153–154	+88.2°	[130]
Planteose		131–132		
β-D-Glc*p*-(1→6)-β-D-Fru*f*-(2→1)-α-D-Glc*p*		123 (dihydrate)	+129.0°(c 4.0)	[132]
Solatriose				
β-D-Glc*p*-(1→3)-β-D-Glc*p*-(1→2)-α-L-Rhm*p*		145–160	+33°→ −4.4°	[133]

[a]Compiled from Carbohydrates (1987) Collins PM (ed) Chapman and Hall Chemistry

Table 8
Physicochemical data of selected tetra-, penta-, and hexasaccharides[a]

Compound	M.p. (°C)	[α]D (H_2O)	Refs.
Tetrasaccharides			
Maltotetraose			
α-D-Glc*p*-(1→4)-α-D-Glc*p*-(1→4)-α-D-Glc*p*-(1→4)-D-Glc*p*	153–154	+165.5°	[133]
Stachyose			
α-D-Glc*p*-(1→4)-α-D-Glc*p*-(1→4)-α-D-Glc*p*-(1→4)-D-Glc*p*	101–105	+131.3	[134]
Lychnose			
α-D-Glc*p*-(1→4)-α-D-Glc*p*-(1→4)-α-D-Glc*p*-(1→4)-D-Glc*p*	153–154	+88.2°	[135]
Pentasaccharides			
Cyclopentaose			
α-D-Glc*p*-[(1→4)-α-D-Glc*p*-(1→4)]$_3$-α-D-Glc*p*	266–268	+8–11°	[136]
Verbascose			
α-D-Gal-(1→6)-α-D-Gal*p*-(1→6)-α-D-Gal*p*-α-D-(1→6)-Glc*p*-(1→2)-β-D-Fru*f*			
Hexasaccharides			
Ajugose			
α-D-Gal*p*-[(1→6)-α-D-Gal*p*-(1→4)]$_3$-α-D-Glc*p*-(1→2)-D-Fru*f*	270	+190.8°	[137]
Cellohexaose			
α-D-Glc*p*-[(1→4)-α-D-Glc*p*-(1→4)]$_4$-α-D-Glc*p*-(1→4)-D-Glc*p*	275–278	+10°	[138]

[a]Compiled from original references

7.3 Hexasaccharides

The most common representatives of this class of oligosaccharides are ajugose and cellohexaose, also isolated from plant material, the roots of *Verbascum thapsiformae* and *Ajuga nipponensis*.

References

1. Khan R (1984) Adv Carbohydr Chem Biochem 33:84
2. Sharon N (1975) Complex Carbohydrates. Their Chemistry, Biosynthesis and Functions. Addison-Wesley, Reading, MA
3. Fraser-Reid B, Madsen R, Campbell AS, Roberts CS, Merritt JR (1999) In: Hecht SM (ed) Bioorganic Chemistry Carbohydrates. Oxford University Press, New York, p 244
4. Nicolaou KC, Bockovich NJ (1999) Chemical Synthesis of Complex Carbohydrates. In: Hecht SM (ed) Bioorganic Chemistry Carbohydrates. Oxford University Press, New York, p 134
5. Sears P, Wong CH (1999) Angew Chem Int Ed 38:2300
6. Glueck CJ, Hastings MM, Allen RD (1982) Am J Clin Nutr 35:1352
7. Hough L, Phadnis SP (1976) Nature 263:800; Hough L, Phadnis SP, Tarelli E (1975) Carbohydr Res 44: 37
8. Lindberg B (1949) Acta Chem Scand 3:1355; Takasawa A (1954) J Antibiot Ser A 7:51

9. Takiura K (1973) Chem Pharm Bull 21:523; Berry JM (1974) Carbohydr Res 38:339
10. Zemplen G, Bognar R (1939) Ber 72:1160; Perila O, Bishop CT (1961) Can J Chem 39:815
11. LeDizet P (1972) Carbohydr Res 24:505
12. Rendelman JA (1972) Carbohydr Res 21:235
13. Takeo K (1983) Carbohydr Res 121:163
14. Stefanovic V (1961) Chem Ber 94:2359; Talley EA (1963) Methods Carbohydr Chem 2:337
15. King RR, Bishop CT (1975) Can J Chem 53:1970
16. Takeo K (1980) Carbohydr Res 86:15 1; ibid (1981) 93:157
17. Lemieux RU, James K, Nagabushan TL (1973) Can J Chem 51:42
18. Usui T (1973) J Chem Soc Perkin Trans 1, 2425
19. Clancy MJ (1960) J Chem Soc 4213
20. Ganfield MCW (1975) J Biol Chem 250:702
21. Yamauchi F, Aso K (1961) Nature 189:753
22. Birch G (1963) Adv Carbohydr Chem 18:201
23. Birch G, Richardson AC (1970) J Chem Soc C 749
24. Defaye J, Driguez H, Henrissat H, Gelas, J, Bar-Guilloux E (1978) Carbohydr Res 63:41
25. Lee CK (1976) Carbohydr Res 50:152
26. Cook SJ, Khan R, Brown JM (1984) J Carbohydr Chem 3:343
27. Oscarson S, Schgelmeble FW (2000) J Am Chem Soc 122:8869; Iley DE, Fraser-Reid B (1975) J Am Chem Soc 97:2563; Lemieux RU, Huber B (1963) J Am Chem Soc 75:4418; Khan R (1976) Adv Carbohydr Chem Biochem 33:235
28. Hudson CS (1946) Adv Carbohydr Chem 2:1
29. Hough L, Jones JKN, Richards EL (1953) J Chem Soc 2005
30. Avigad G (1959) Biochem J 73:587
31. Begbie R (1966) Carbohydr Res 2:272
32. Stodola FH, Koepsell HJ, Sharpe ES (1952) J Am Chem Soc 74:3202; ibid (1956) 78:2314
33. Thiem J, Kleeberg M, Schwengers D (1989) Alimenta 2:23
34. Sloneker JM, Orentas DG, Knutson CA, Watson PR, Jeanes A (1968) Can J Chem 46:3353
35. Takeo K (1983) Carbohydr Res 121:163
36. Curtis EJC, Jones JKN (1959) Can J Chem 37:358
37. Lemieux RU, Driguez H (1975) J Am Chem Soc 97:4063
38. Hicks KB (1980) Carbohydr Res 82:393; Adachi S (1969) Carbohydr Res 9:242
39. Richtmyer H (1935) J Am Chem Soc 57:1716; ibid (1936) 58:2534
40. Horowitz RM (1963) Tetrahedron 19:773
41. Vattuone MA (1975) Carbohydr Res 39:164
42. Charlton W (1927) J Chem Soc 1527
43. Picket A, Vogel H (1926) Helv Chim Acta 9:806
44. French D, Wild GH, Young B, Jones WJ (1953) J Am Chem Soc 75:709
45. Wickstrom A, Courtuis JE, Le Dizet D, Archambault A (1958) Compt Rend 246:1624
46. Kamiya S (1967) Agric Biol Chem 31:261
47. Gorin PAJ, Perlin AS (1959) Can J Chem 37:1930
48. Gorin PAJ, Perlin AS (1959) Can J Chem 37:1921
49. Bognar R (1967) Carbohydr Res 5:241
50. Erbing B, Lindberg B (1969) Acta Chim Scand 23:2213
51. Lichti H, von Wartburg A (1961) Helv Chim Acta 44:238
52. Klages F (1938) Justus Liebigs Ann Chem 535:175
53. Gorin PAJ (1975) Carbohydr Res 39:3
54. Helferich B, Rauch S (1928) Justus Liebigs Ann Chem 465:168; Whistler RL, Corbet WH (1955) J Am Chem Soc 77:3822
55. Flowers HM (1967) Carbohydr Res 4:312; Gorin PAJ, Spencer JFT (1961) Can J Chem 39:2282
56. De Belder AN (1977) Adv Carbohydr Chem Biochem 34:179
57. Clode DM (1979) Chem Rev 79:491
58. Gelas J (1982) Adv Carbohydr Chem Biochem 39:71
59. Haines AH (1982) Adv Carbohydr Chem Biochem 39:13
60. Khan R (1985) Carbohydr Res 16:261
61. Koizumi K, Utamura T (1974) Carbohydr Res 33:127
62. de Raadt A, Stutz AE (1992) Tetrahedron Lett 189
63. Norberg T (1996) Glycosylation Properties and Reactivity of Thioglycosides, Sulfoxides and other *S*-glycosides: Current Scope and Future Prospects. In: Khan SH, O'Neill RA (eds) Modern Methods in Carbohydrate Synthesis. Harwood Academic Publishers, Amsterdam, p 82
64. Madsen J, Bols M (1998) Angew Chem Int Ed Engl 37:3177
65. Meyer B (1990) Topics Curr Chem 154:143
66. Backman I, Erbing B, Jansson P, Kenne L (1988) J Chem Soc Perkin Trans I 889
67. Du Y, Lindhart RJ (1998) Tetrahedron 54:9913
68. Beau JM, Gallagher T (1997) Topics Curr Chem 187:1
69. Nicotra F (1997) Topics Curr Chem 187:55

70. Levy DE, Tang C (1995) The Chemistry of *C*-Glycosides. Elsevier Science, Oxford, vol 13, p 4
71. Postema MHD (1995) C-Glycoside Synthesis. CRC Press, Boca Raton, FL, p 346
72. Martin OR, Lai W (1993) J Org Chem 58:176
73. Dyer UC, Kishi Y (1988) J Org Chem 53:3383; O'Leary DJ, Kishi Y (1993) J Org Chem 58:308
74. Carcano M, Nicotra F, Panza L, Russo G (1989) J Chem Soc Chem Commun 642; Lay L, Nicotra F, Pangrazio C, Panza L, Russo G (1994) J Chem Soc Perkin Trans 1, 333
75. Sutherlin DP, Armstrong RW (1996) J Am Chem Soc 118:9802
76. Postema M, Calimente D, Liu L, Behrmann T (2000) J Org Chem 65:6061
77. Barkata M, Farr RN, Chaguir B, Massiot G, Lavaud C, Anderson WR, Sinou D, Daves GD Jr (1993) J Org Chem 58:2992
78. Johns BA, Pan YT, Elbein AD, Johnson CR (1997) J Am Chem Soc 119:4856
79. Johnson CR, Johns BA (1997) Tetrahedron Lett 38:7977
80. Asensio JL, Cañada FJ, Garciia-Herrero A, Murillo MT, Fernandez-Mayorales A, John BA, Kozak J, Zhu Z, Johnson CR, Jimenez-Barbero J (1999) J Am Chem Soc 121:11318
81. Leewenburgh MA, Picasso S, Overkleft HS, van Der Marel GA, Vogel P, van Boom JH (1999) Eur J Org Chem 1185
82. Marquis C, Picasso S, Vogel P (1999) Synthesis 1441
83. Vogel P (2000) J Chem Soc Chem Commun 642
84. Weatherman RV, Kiesling LL (1996) J Org Chem 61:534
85. Witczak ZJ, Chen H, Kaplon P (2000) Tetrahedron Asymmetry 11: 519
86. Geyer A, Hummel G, Eisele T, Reinehardt S, Schmidt RR (1996) Chem Eur J 2:981
87. Espinosa JF, Cañada FJ, Asensio JL, Dietrich H, Martin-Lomas M, Schmidt RR, Jiménez-Barbero J (1996) Angew Chem Int Ed Engl 35:303
88. Defaye J, Gulliot JM (1994) Carbohydr Res 253:185
89. Blanc-Muesser M, Defaye J, Driguez H (1978) Carbohydr Res 67:305
90. Blanc-Muesser M, Defaye J, Driguez H (1982) J Chem Soc Perkin Trans 1, 15
91. Hamacher K (1984) Carbohydr Res 128:291
92. Chretien F, Di Cesare P, Gross B (1988) J Chem Soc Perkin Trans 1, 3297
93. Defaye J, Gelas J (1991) In: Atta-ur-Rahman (ed) Studies in Natural Products Chemistry, vol 8E. Elsevier, Amsterdam, pp 315–357
94. Driguez H (1997) Thiooligosaccharides in Glycobiology. In Driguez H, Thiem J (eds) Topics in Current Chemistry, vol 187 Glycoscience Synthesis of Substrate Analogs and Mimetics, Springer, Berlin, Heidelberg, New York, p 85
95. Witczak ZJ (1999) Curr Med Chem 5:125: Dey PM, Witczak ZJ (2003) Mini Review in Med Chem 3:271; Robina I, Vogel P, Witczak ZJ (2001) Curr Org Chem 5:1177; Witczak ZJ, Culhane JM (2005) Appl Microbiol Biotechnol 69:237 and references cited therein
96. Witczak ZJ, Sun J, Mielguj R (1995) Bioorg Med Chem Lett 5:2169
97. Witczak ZJ, Chhabra R, Chen H, Xie XQ (1997) Carbohydr Res 301:167
98. Witczak ZJ, Chhabra R, Boryczewski D (2000) J Carbohydr Chem 19:543
99. Becker T, Thimm B, Thiem J (1996) J Carbohydr Chem 15:1179
100. Zawisza A, Kryczka B, Lhoste P, Porwanski S, Sinou D (2000) J Carbohydr Chem 19:795
101. Mehta, S, Andrews JS, Johnston BD, Pinto BM (1994) J Am Chem Soc 116:1569
102. Harada N, Nakanishi K (1983) Circular Dichroism Spectroscopy: Exciton Coupling in Organic Stereochemistry. University Science Books, Mill Valley, CA; Harada N, Nakanishi K (1969) J Am Chem Soc 91:3889
103. Bell AF, Hecht L, Barron LD (1995) Spectrochimica Acta A51:1367
104. Bell AF, Hecht L, Barron LD (1994) J Am Chem Soc 116:5155
105. Koto S, Uchida T, Zen S (1973) Bull Chem Soc Jpn 46:2520
106. Lonn H (1985) Carbohydr Res 139:105
107. Mootoo DR, Konradsson P, Udodong U, Fraser-Reid B (1988) J Am Chem Soc 110:5583
108. Veeneman GH, van Boom JH (1990) Tetrahedron Lett 31:275
109. Yamada H, Harada T, Takahashi T (1994) J Am Chem Soc 116:7919
110. Kanie O, Ito Y, Ogawa T (1994) J Am Chem Soc 116:12073
111. Grice P, Ley SV, Pietruszka J, Priepke HWM, Walther EPE (1995) Synlett 781
112. Baressi F, Hisgaul O (1995) J Carbohydr Chem 14:1043
113. Guertsen R, Cote F, Hahn MG, Boons GJ (1999) J Org Chem 64:7828
114. Wang W, Kong F (1999) J Org Chem 64:5091
115. Bock K, Pedersen C, Pedersen H (1984) Adv

Carbohydr Chem Biochem 42:193
116. Petrakova E, Krupova I, Schraml J, Hirsch J (1991) Collect Czech Chem Commun 56:1300
117. Thompson A (1952) J Am Chem Soc 74:3612
118. Pan SC, Nicolson LW, Kolachaw C (1951) J Am Chem Soc 73:2547; Peat S, Tuevey JR, Evans JM (1959) J Chem Soc 3223; Wolfrom ML (1966) J Chem Soc Chem Commun 2
119. Aspinall GO, Caincross IM, Ross KM (1963) J Chem Soc 1721
120. Pazur JH, French D (1952) J Biol Chem 196:265
121. Whistler RL (1954) J Am Chem Soc 76:1071
122. Lindberg B (1954) Acta Chem Scand 8:985
123. Samuelsom O (1972) Methods Carbohydr Chem 6:66
124. Bacon JSD (1953) J Chem Soc 2528
125. Hammer H (1968) Acta Chem Scand 22:197; Albon N (1953) J Chem Soc 24
126. Takeo K (1981) Carbohydr Res 88:158
127. Suzuki H (1964) Arch Biochem Biophys 105: 339
128. Leitch GC (1927) J Chem Soc 588
129. Haworth NW (1923) J Chem Soc 3125; French D (1954) Adv Carbohydr Chem Biochem 9:149
130. Swendsen AB (1956) Acta Chem Scand 10:1500
131. Briggs LH (1963) J Chem Soc 2848
132. Takeo K (1983) Carbohydr Res 121:328
133. Whistler RL (1954) J Am Chem Soc 76:1671
134. Wolfrom ML (1952) J Am Chem Soc 74:6299
135. Aspinall GO, Ross KM (1963) J Chem Soc 1021
136. Michell AJ (1970) Carbohydr Res 12:453
137. Murakami S (1941) Acta Phytochim 12:97
138. Wolfrom ML (1952) J Am Chen Soc 74:5331

5.2 Sucrose and Related Oligosaccharides

*Gillian Eggleston**
United States Department of Agriculture, SRRC-ARS-USDA,
1100 Robert E. Lee Boulevard, New Orleans,
LA 70124, USA
gillian@srrc.ars.usda.gov

Abstract

Sucrose (α-D-glucopyranosyl-(1↔2)-β-D-fructofuranoside) is the most common low-molecular-weight sugar found in the plant kingdom. It is ubiquitously known as common table sugar and primarily produced industrially from sugarcane (*Saccharum officinarum*) and sugar beet (*Beta vulgaris*); the basics of the industrial manufacture of sucrose are outlined in this chapter. Commercial sucrose has a very high purity (>99.9%) making it one of the purest organic substances produced on an industrial scale. Value-addition to sucrose via chemical and biotechnological reactions is becoming more important for the diversification of the sugar industry to maintain the industries' competitiveness in a world increasingly turning to a bio-based economy. The basis for the chemical reactivity of sucrose is the eight hydroxyl groups present on the

In: *Glycoscience*. Fraser-Reid B, Tatsuta K, Thiem J (eds)
Chapter-DOI 10-1007/978-3-540-30429-6_26:

molecule, although, sucrose chemical reactivity is regarded as difficult. Increasing use of enzymatic biotechnological techniques to derivatize sucrose is expected, to add special functionalities to sucrose products like biodegradability, biocompatibility, and non-toxicity. Analysis of sucrose by colorimetric, enzymatic, oxidation-reduction and chromatography methods are discussed. Oligosaccharides related to sucrose are outlined in detail and include sucrose-based plant, honey and *in vitro* oligosaccharides.

Keywords

Industrial sugar production; Sucrose physico-chemical properties; Sucrose analysis; Sucrose value-added products; Sucrose oligosaccharides

Abbreviations

AOAC	Association of Official Agricultural Chemists
ATP	adenosine triphosphate
FOS	fructooligosaccharides
FT-IR	Fourier-transform infrared
GC	gas chromatography
G-6-P	glucose-6-phosphate
G6P-DH	glucose-6-phosphate dehydrogenase
HPLC	high performance liquid chromatography
HK	hexokinase
HMF	5-hydroxymethyl-2-furaldehyde
IC	ion chromatography
ICUMSA	International Commission for Uniform Methods in Sugar Analysis
NAD	nicotinamide-adenine dinucleotide
NADH	nicotinamide-adenine dinucleotide reduced form
NIR	near-infrared
PTB	Physikalisch-Technische Bundesanstalt
SAIB	sucrose acetate isobutyrate

1 Introduction

Sucrose (α-D-glucopyranosyl-(1↔2)-β-D-fructofuranoside) is the most common low molecular weight sugar found in the plant kingdom: it is a major product of the photosynthetic activity of leaves and a major transportable source of organic carbon - found in the phloem saps of all plants. Sucrose serves as an essential reserve and energy source for non-photosynthetic tissues, and provides carbon for the synthesis of storage compounds, including starch and less frequently fructosyloligosaccharides (fructans) [1]. For a complete description of the biosynthesis and degradation of sucrose in plants the reader is referred to the review by Madore [1].
Sucrose is ubiquitously known as common table sugar and the sugar industry is actually the "sucrose" industry. Sucrose is primarily produced industrially from sugarcane (*Saccharum officinarum*) and sugarbeet (*Beta vulgaris*), but is also produced from sugar palm (*Arenga pinnata*), and sugar maple (*Acer saccharum*) for smaller, specialty markets. On a worldwide

basis, household sucrose consumption accounts for approximately 55% of the total consumption.

Like for other carbohydrates, the human body utilizes sucrose as a fuel to obtain heat and form adenosine triphosphate (ATP), the basic molecule of the energy system for all organisms.

2 Industrial Sucrose Production

Commercially available sucrose has very high purity (> 99.9%) making it one of the purest organic substances produced on an industrial scale. To obtain such a pure product, from both sugarcane and sugarbeet, rather complex isolation and purification processes are followed. Industrial sucrose production is essentially a series of separations of non-sucrose compounds (usually termed non-sugars or impurities) from sucrose, and the chemistry of the sequential process units is designed for maximum removal of non-sugars with minimum destruction of sucrose. Sugarcane processing often occurs in two stages. Firstly, the juice is extracted from sugarcane (sucrose yields range between 10–15% weight of sugarcane) and converted to inedible raw sugar (~97–98% pure sucrose; golden yellow/brown crystals) at factories (mills) in tropical and sub-tropical areas where sugarcane is grown. Secondly, after raw sugar has been transported to a refinery, it is refined using very similar unit processes used in raw sugar manufacture, to the familiar white, refined edible sugar (> 99.9% sucrose). In some tropical areas of the world, particularly Asia, plantation white, mill white, or direct white sugar (> 99% sucrose with more color than white, refined sugar) is produced directly from sugarcane. There have been recent developments in some countries to produce high quality raw sugar (VHP sugar and Blanco Directo sugar) directly from sugarcane juice at the factory. In comparison to sugarcane, sugarbeets are grown in temperate areas and are processed directly into white sugar (> 99.7% sucrose) at nearby factories.

2.1 Raw Sugar Manufacture from Sugarcane

A brief outline of the typical unit processes in the manufacture of raw sugar from sugarcane at the factory is shown in *Fig. 1*. Harvested sugarcane (either whole-stalks or billets, depending on the country) is delivered to the factory; the chemical composition and processing characteristics of the delivered sugarcane are dependent on the variety [2]. Juice is extracted from the sugarcane either by tandem milling or, less frequently, diffusion. Imbibition water is applied to one or more of the later mills in the tandem and recirculated to improve extraction. Industrial mill extraction can vary between ~89–97%, depending on many factors including the tandem design. Diffusion, in comparison to milling, often gives greater extraction of sucrose (~93–98%), but also of non-sucrose impurities that increases the loss of sucrose to molasses; more imbibition water is also required in diffusion compared to tandem milling. Extracted juice is then purified in the clarification unit process. Clarification processes can vary, but typically hot lime clarification gives the best results [3]. Juice is first heated up to ~93 °C to allow natural floc formation from the coagulation of colloids and macromolecules, including polysaccharides and proteins. Lime (as milk of lime or calcium saccharate) is added to neutralize organic acids in the juice and to form a heavy precipitate, primarily of calcium phosphate bridged flocs. After flashing at ~102 °C to remove interfering bubbles, this precipitate is settled with the aid

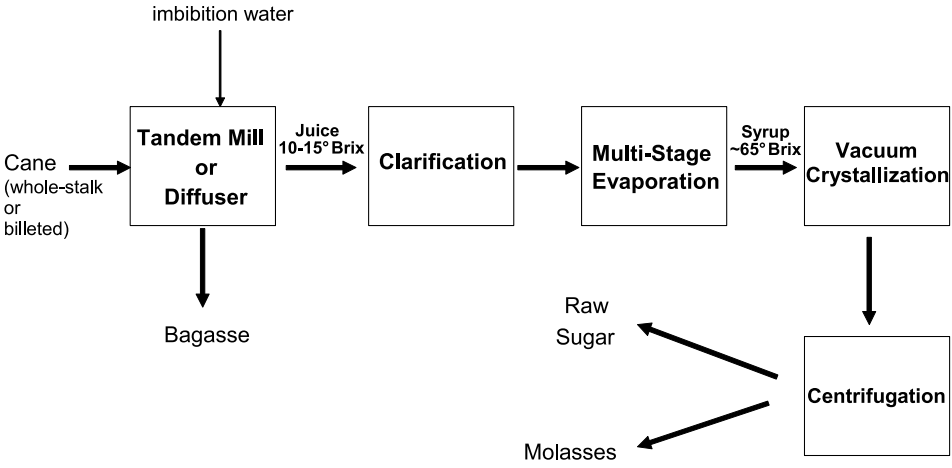


Figure 1
Basic scheme of the raw sugar manufacturing process in a sugarcane factory

of polyelectrolyte flocculants in a large clarification tank, and impurities are precipitated out. The clarification process transforms dark, cloudy sugarcane juice into a clear (clarified) juice (13–18% dissolved solids). This clarified juice is then concentrated through a series of multiple-effect evaporators to syrup of ~65% dissolved solids. By far the most common used type of evaporator in raw sugar manufacture is the simple Robert's-type conventional calandria, rising film evaporator [4], although occasionally more sophisticated evaporators such as Kestner (long tube, rising film) and plate (falling film) evaporators are utilized.

Syrup from the evaporation unit process is then further concentrated, under vacuum at lower temperature than in evaporation to minimize the chemical degradation of sucrose, and crystallized (*Fig. 1*). The vacuum pans are seeded with finely ground sucrose to allow larger sucrose crystals to form. A mixture of sucrose crystals and mother liquor (massecuites) is produced that is then separated in centrifuges, and the mother liquor is re-concentrated and re-crystallized to give two more crops or strikes of crystals. The final liquor is the by-product molasses. Crystallization is often regarded as an industrial "art" and is the major purification step in the manufacture of sugar [5]. The raw sugar obtained is washed (sprayed with water during centrifugation), dried on belts, and then stored in bulk in large warehouses. Raw sugar is transported to the refinery by truck, rail car, or ship.

2.2 Refined Sugar Production from Raw Sugar

Unlike sugarcane factories that are limited to operation during the sugarcane harvest season, raw sugar refineries operate year round. Raw sugar contains ~1–3% non-sugars and the aim of refining is to remove them from the sucrose; typical mean compositions of raw and refined sugars are listed in *Table 1*. Because the production of refined sugar from raw sugar at the refinery is also a series of separations of non-sugars from sucrose (*Fig. 2*), it consists of many similar processes as in raw sugar manufacture (*Fig. 1*).

Table 1
Typical analysis of raw and refined sugar from sugarcane. From [6]

Component	Raw sugar	Refined sugar
Sucrose (%)	97.73	99.95
Invert (%)	0.56	0.006
Ash (%)	0.45	0.007
Organic non-sugars (%)	0.62	0.014
Water (%)	0.64	0.023
Color (ICUMSA units)	2000–5000	10–50

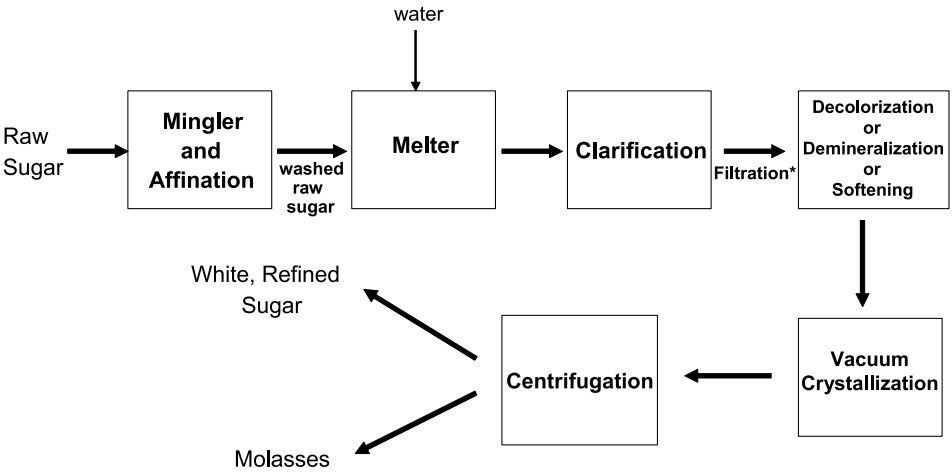


Figure 2
Basic scheme of the white, refined sugar manufacturing process in a raw sugar refinery. Clarification can be phosphatation or carbonation, and decolorization can be with granular carbon, bone char, or ion-exchange resins. *Filtration is not required if phosphatation clarification is utilized

The first stage of the refining process is affination where the raw sugar is mixed with syrup (magma) and then centrifuged and washed with water, to remove the molasses layer around the raw sugar crystals. The affined (or affinated) sugar is dissolved (melted) in water to create a melt liquor of 68–73 °Brix (% dissolved solids). The melt liquor is then clarified by either a phosphatation or carbonatation (carbonation) process [7], followed by decolorization. To obtain refined sugars with low levels of color, 80–90% of color must be removed from the yellow, clarified liquor. Traditionally, bone char and granular activated carbon adsorbents have been used, but many modern refineries now use ion-exchange resins. Bone char and desalting ion-exchange resins remove inorganic salts as well as color, and all adsorbents remove some turbid particles [6]. Multi-stage crystallization is the final purification process at the refinery to produce white, refined sugar which is sold to household and industrial markets; industrial users purchase it preferably in bulk or as liquid sugar in tank trucks.

2.3 Refined Sugar Production from Sugarbeet

Production of refined sugar from sugarbeets has some similarities to refined cane sugar production, as both are a series of process units aimed at separating and removing impurities from the sucrose. However, dissimilarities exist as sugarbeet is a tuberous root and sugarcane a grass. Moreover, because sugarbeet juice contains less invert sugar (glucose and fructose) but more nitrogen-containing compounds than sugarcane juice, sugar beet processing occurs at alkaline pH's compared to the more acidic pH's in sugarcane processing. This is deliberate, as the complete decomposition of invert sugar in beet sugar manufacture is desirable to reduce Maillard color forming reactions between amino acids and invert sugars. In comparison, in sugarcane processing, it is considered an advantage to maintain a small concentration of invert sugar, because it has the net effect of lowering sucrose losses to final molasses [8].

Sugarbeets are harvested defoliated and delivered to the factory. Often sugarbeets are harvested at a rate in excess of factory processing capacity and excess sugarbeets are stored in long-term storage piles on factory or remote grounds. Sugarbeets from the storage piles are introduced to the factory, washed, and sliced into "V" shaped cossettes. Cossettes are added to a counter-current diffuser which may be of a horizontal drum, sloped vat, or vertical cylindrical tower design [9]. Sucrose and impurities are extracted with hot water; approximately 98% sucrose is extracted in the counter-current diffusion operation. Diffusion juice contains ~12% sucrose and 2% soluble impurities on sugarbeet weight, and is heated to ~85 °C before it is purified with a double-carbonatation clarification process. Milk of lime and carbon dioxide are added during clarification. In some sugarbeet factories, sulfur dioxide is added to filtered, clarified juice to minimize color formation during subsequent processing. The resulting clarified "thin" juice is then concentrated from ~14% to ~60–65% °Brix ("thick" juice) across multiple-effect evaporators (Robert's-type and Kestner evaporator designs are favored), then triple-crystallized and centrifuged to produce white, refined sugar (>99.7% purity). In some sugarbeet factories additional purification steps are employed, such as softening, demineralization, or color removal with ion-exchange resins or carbon adsorbents. Additionally, sucrose and betaine are also recovered from beet molasses with chromatography, a process that is much easier in sugarbeet than sugarcane processing.

For more detailed information on the industrial production of sucrose from sugarcane and sugarbeet, the reader is referred to other comprehensive texts [7,10,11,12,13].

2.4 By-Products of Industrial Sucrose Production

The major by-products of industrial sucrose production are bagasse, beet pulp, and molasses. By volume, bagasse from cane sucrose manufacture, is the most important by-product and is the primary source of fuel for the generation of steam and electricity (cogeneration) in sugarcane factories. If there is a surplus of bagasse it can be used to produce value-added products including paper and chipboard and, more recently, there has been much research on the production of cellulosic bioethanol from treated bagasse although, this is not a technological or commercial reality yet. Beet pulp is a source of animal feed as wet pulp, pressed pulp silage, dried pulp, with or without added molasses [14]. Molasses is currently the most valuable by-product of sugar manufacture. It exists in a range of grades: edible molasses (high test

molasses), cane and beet molasses, and refinery molasses. It is used as an animal feed additive, in the industrial production of rum and other beverage alcohols, and bakers' yeast and citric acid [14]. Molasses is also used as an organic feedstock in various other fermentation processes.

3 Chemical and Physical Properties of Sucrose

Sucrose occurs in different physical forms (crystalline, amorphous, and liquid). Sucrose is utilized primarily because of three functional properties: (1) nutritional value, (2) sweetness, and (3) bulking ability, although it imparts many more functional properties that are listed in *Table 2*.

Sucrose consists of two monosaccharides, glucose and fructose, joined by a glycosidic bond between carbon atom 1 of the glucose unit and carbon atom 2 of the fructose unit (*Fig. 3*). Since it contains no free anomeric carbon atom (*Fig. 3*), it is a non-reducing sugar. Some basic chemical, physical, and physico-chemical properties of pure sucrose are listed in *Table 3*.

Structure-function studies on sucrose have been accomplished. In 1956, Lemieux and Huber [16] reported the first chemical synthesis of sucrose by the reaction of 1,2-anhydro-

Table 2
Relative differences in the characteristics of chemical methods to measure sucrose in technological sugars. Adapted from [34]

Property						
Nutritional:	Nutritional value					
Sensory:	Sweetness	Taste/aroma	Texture	Appearance (color/ luster)		
Physical:	Crystallization	Viscosity	Osmotic pressure (water activity)	Hygroscopicity	Dry substance/ consistency	Grain size/ distribution
Microbial:	Preservation	Fermentation				
Chemical:	Inversion	Caramelization	Antioxidant effect			

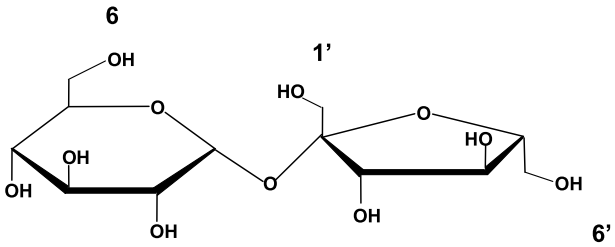


Figure 3
Sucrose structure showing the three primary hydroxyl groups (6, 1′, 6′)

Table 3
Basic chemical and physical properties of pure sucrose[a]

Property	Value and unit	Note
Molecular formula	$C_{12}H_{22}O_{11}$	
Molecular mass	342.30 g/mol	
Flavor character	Pure, sweet, odorless	
Crystal shape	Monoclinic hemihedral	
Density and phase	1.587 g/cm^3, solid	
Solubility in water	2.115 g/100 ml	At 25 °C
Melting point	186 ± 4 °C	
Chiral rotation $[\alpha]_D$	+66.47 to +66.49	At 25 °C
Refractive index	1.5376	
Thermal conductivity	0.592 W/(m.K)	20% solution at 70 °C
	0.450 W/(m.K)	60% solution at 70 °C
Specific heat capacity	1243.9 J/(kg.K)	At 25 °C
Viscosity	1.9 cP	20% solution at 20 °C
	58.0 cP	60% solution at 20 °C

[a]For more detailed information see [7] and [15]

α-D-glycopyranose triacetate with syrupy 1,3,4,6-tetra-*O*-acetyl-D-fructose. The molecular geometry realized in the crystal is characterized by two intramolecular hydrogen bonds between the glucose and fructose molecules [17] that are shown in *Fig. 4*. In aqueous solution, the intramolecular bonds in sucrose remain [18], with an insertion of a water molecule between the glucosyl-2-OH and fructosyl-1-OH [19], which is also shown in *Fig. 4*.

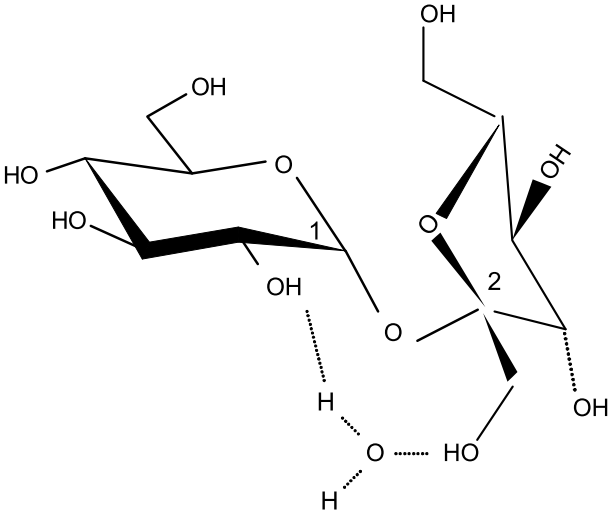


Figure 4
Sucrose structure in aqueous solution. The two sugar units are similarly disposed towards each other, caused by insertion of a water molecule between the glucosyl-2-OH and fructosyl-1-OH groups and the water bridge is fixed by hydrogen bonding. From [20]

4 Analysis of Sucrose

It is not the role of this chapter to give a detailed review of the different analytical techniques used to analyze pure sucrose, and sucrose in technological products. For this the reader is referred to comprehensive reviews by Robyt and White [21], Lescure [22], Eggleston and Côté [23], and Folkes and Jordan [24]. Furthermore, scientific associations exist such as the Association of Official Agricultural Chemists (AOAC) and the International Commission for Uniform Methods in Sugar Analysis (ICUMSA), which provide validated official methods for the analysis of sucrose in different products. ICUMSA provides official methods for use in the sugar industry trade.

Physical and chemical techniques can be used to analyze pure sucrose and sucrose in technological products. Physical spectroscopic techniques, not usually found in routine, industrial laboratories are used to identify sucrose. These include vibrational spectroscopy that is capable of rapidly generating a "fingerprint" of the product. Fourier-transform infrared (FT-IR) spectroscopy can provide information on the anomeric bond and hydroxyl groups [25,26] and, in spite of a wide range of water absorption in the near-infrared (NIR) region, it can be used to accurately identify pure sucrose [27]. Laser-Raman spectroscopy has been used to study the structure of concentrated sucrose solutions and a table is available with characteristic frequency bands and corresponding intensities [28]. Proton NMR spectroscopy is a powerful tool to identify and distinguish sucrose. The ^{1}H NMR spectrum of sucrose reveals the anomer proton bound to the carbon located between the two oxygen atoms (see ❯ *Fig. 3*), and protons bound to other carbons of the rings or the hydroxyl methyl groups of sucrose [29]. ^{1}H NMR spectroscopy, in combination with pattern recognition techniques, has been used to detect the adulteration of orange juices with beet invert sugar [30]. Fourier transform ^{1}H NMR has been used to determine the sucrose content of beet press juice [31] based on time decay of water resonance in the presence of Cr (III). The ^{13}C NMR spectrum of sucrose has also been elucidated [27,32].

Other physical techniques used to measure sucrose in technological products are polarimetry (optical rotation) and NIR spectroscopy. Polarimetry is very important in the sugar industry and is still ubiquitously used for sugarcane or sugarbeet payments to growers, for trade, and custom controls. For this reason, it has been necessary to establish an industrial operation method and measuring instrument scale, i. e., to directly obtain the concentration of sucrose "pol" in 26 g sugar/100 mL water at 20 °C, which corresponds to 100°Z (latest 1986 scale agreed by ICUMSA). This scale was obtained from extremely accurate measurements performed by the Physikalisch-Technische Bundesanstalt (PTB) in Germany and by the Bureau of Standards in the USA [22]. 100°Z corresponds to the optical rotation of pure sucrose at the wavelength of the green rays of the isotope ^{198}Hg ($\lambda = 546.2271$ nm in vacuo; 20.00 °C; 200-mm path length). ICUMSA has also approved the measurement of sucrose "pol" at 880 nm, which allows for more accurate measurements of dark factory juices [33]; dual wavelength polarimeters at 546 and 880 nm are readily available. The drawback of using "pol" to measure sucrose, however, is that it represents a cumulative measurement of all optically active compounds in the solution. Eggleston et al. [34] reported that "pol" can *over-estimate* sucrose and, consequently, *underestimate* sucrose losses across industrial process units. Thus, "pol" values in technological products are precise but often inaccurate. Double polarization methods [35,36,37] are available to remedy the presence of other optically active substances

in the sugar solution. A second polarization of the solution is incorporated after sucrose is selectively hydrolyzed to form an equimolar mixture of glucose and fructose (invert). Sucrose is then calculated by dividing the change in "pol" before and after hydrolysis with an algebraic constant [22]. Although double polarization methods are convenient for industrial laboratory use, in practice selective hydrolysis with acids or enzymes is very difficult. In recent years, NIR spectroscopy (750 to 2500 nm) has been of continued interest to the world-wide sugar industry [38] to measure sucrose, and other sugar quality parameters such as fiber and % dissolved solids, because of its ease of use and speed. However, this technique is dependent on the instrumentation used and generated calibrations. For sucrose analysis the calibration is based on "pol" values, so is only as accurate as polarimetry.

The physical methods described above tend to be more accurate for pure sucrose solutions than complex industrial sugar solutions; for the latter, chemical methods are more accurate. Chemical methods are broadly divided into those based on (i) colorimetric reactions, (ii) enzymatic reactions, (iii) oxidation-reduction properties, or (iv) chromatography separations.

4.1 Colorimetric Methods

For colorimetric methods, most reactions are conducted in the presence of sulfuric acid. Dehydration reactions occur after hydrolysis by warm drying, leading to the formation of deoxyulose and subsequent 5-hdroxymethyl-2-furaldehyde (HMF) which is easily measured at 280 nm. The Molisch reaction is well-known [21,22]: the addition of α-naphthol and concentrated sulfuric acid to the solution causes the formation of a violet-red ring that is used in qualitative determinations. Anthrone can be used as a chromophore at 585 nm, but has been mainly replaced by a simpler phenol-sulfuric acid method using a 480 nm wavelength; both of these methods are quantitative. It must be noted, however, that these colorimetric methods are often tedious and not highly selective for sucrose as they are applicable to most sugars [21,22]. *Table 4* lists the relative differences of colorimetric methods with other chemical methods.

4.2 Enzymatic Methods

Enzymatic methods to determine sucrose are more specific and increasingly being used [22]. Sucrose is readily hydrolyzed with β-fructofuranosidase (invertase) at pH 4.5. Fructose can then be quantitatively determined by phosphorylating with adenosine-5′-triphosphate (ATP) using hexokinase (HK), then converting fructose-6-phosphate to glucose-6-phosphate (G-6-P) by phosphoglucose isomerase. Glucose-6-phosphate is then oxidized to 6-phosphogluconate in the presence of nicotinamide-adenine dinucleotide (NAD) by glucose-6-phosphate dehydrogenase (G6P-DH); NADH is formed which is easily measured at 340 nm. Glucose can also be quantitatively determined by phosphorylation with ATP and HK to G-6-P, then oxidized to 6-phospho-gluconate in the presence of NAD and G6P-DH, with measurement of formed NADH. Glucose can also be determined by oxidation with glucose oxidase to form hydrogen peroxide, which is then reacted with a dye in the presence of peroxidase. The colored product is measured spectrophometrically. Kits with immobilized or non-immobilized enzymes are available for such enzymatic methods. However, enzymes are relatively expensive and the

Table 4
Relative differences in the characteristics of chemical methods to measure sucrose in technological sugars. Adapted from [34]

Type of Method	Method	Technical proficiency required	High cost		Labor intensive	Rapid	Potentially dangerous	Measures specific sugars
			Consumables	Equipment				
Colorimetric	Phenol-sulfuric method		X	X	X		X	
Oxidation-reduction	Lane and Eynon methods				X		X	
Oxidation-reduction	Luff Schoorl method				X		X	
Chromatography	HPLC	X		X		X		X
Chromatography	IC[b]	X		X				X
Chromatography	GC	X	X	X	X		X	X
Enzymatic	Enzymatic		X			X		X

[a]International Commission for Uniform Methods in Sugar Analysis (ICUMSA) methods [b]Also known as High Performance Anion Exchange Chromatography (HPAEC)

enzymes have a limited shelf-life (◉ *Table 4*). Special instrumentation is required if immobilized enzymes are used.

4.3 Oxidation-Reduction Methods

Oxidation-reduction methods to determine "% apparent sucrose" are based on the reducing properties of the free carbonyl group of sucrose hydrolysis products, glucose and fructose. Sucrose, therefore, has first to be hydrolyzed by acid. The reducing sugars are reacted with copper solutions in alkaline medium. The most widely used copper reduction methods are Lane and Eynon and Luff Schoorl titration methods [39]. Although copper reduction methods are precise [39] and require no expensive equipment or high degree of technical proficiency, they can be labor intensive and potentially dangerous (◉ *Table 4*). Furthermore, they tend to *over-estimate* sucrose because of over-titrations and because all reducing substances present are measured [39].

4.4 Chromatographic Methods

The most accurate techniques for the analysis of sucrose are those based on chromatographic separations. Gas chromatography (GC) or gas-liquid chromatography is considered the most accurate technique (K. Schaffler, Sugar Milling Research Institute, South Africa, personal communication) and based on the sample being vaporized at high temperatures, injected onto a column and transported through, and separated on, a column by the flow of inert gas. Because

sucrose and other sugars do not sublime at high temperatures, derivatization is required. Silylation [40] and silylation-oximation derivatization [40] are common. As GC is time consuming (*Table 4*), other chromatographic techniques have become more popular, particularly high performance liquid chromatography (HPLC) with cation-exchange columns (sodium or potassium forms) or aminopropyl-silica columns both with refractive index detection, and ion chromatography (IC) with anion-exchange columns and pulsed amperometric detection (a.k.a. HPAEC-PAD). The advantage of IC is that it is very sensitive, but can be relatively unstable. Eggleston and Côté [23] recently reviewed the use of chromatographic techniques for sucrose and other oligosaccharides.

5 Value-Added Products from Sucrose

Value-added products from sucrose that meet existing needs can increase the demand, value, and consumption of sucrose. However, only a small percentage of the sugar produced in the world is used in non-food applications, with ~1.7% at present in the US [41]. This is unfortunate as much research effort and funds have been expended on the identification and development of value-added products from sucrose. Part of the reason for such little impact of this research is that the scientists inventing the products have not fully considered the market, and do not have the business acumen to sell such products to industry. More involvement by industry, particularly at the conception phase, would help to gain more impact. This is becoming increasingly more important, as value-addition to sucrose and diversification of the sugar industry are becoming a greater necessity in the maintenance of the industries' competitiveness in a world increasingly turning to sustainable agriculturally derived chemicals to replace petrochemically derived chemicals (bio-based economy).

Sucrose is a source for many value-added products because of its chemical and enzymatic reactivity. The basis for the chemical reactivity of sucrose is the eight hydroxyl groups present on the molecule (*Fig. 3*). However, sucrose chemical reactivity is regarded as difficult [20,42] because the selective introduction of other functional groups into sucrose is usually possible only in non-aqueous solvents or by using a cost-intensive protecting group chemistry, and the sucrose molecule is susceptible to degradation. Three of the hydroxyl groups are in the primary positions ($-CH_2OH$) and five in secondary positions (–CHOH) (*Fig. 3*). Generally, the three primary hydroxyls have greater reactivity but they often prove a hindrance as they are difficult to react exclusively [20,42]. The synthesis of an enormous number of sucrose derivatives is possible; substitution with just one group type could theoretically give 255 different compounds! Moreover, the alcohol group can be derivatized to form esters, ethers, and substitution derivatives [44]. Sucrose can be readily degraded by acids, oxidizing agents, alkalis, and catalytic hydrogen to compounds of lower molecular weight.

Sucrose is also an exceptional molecule for enzymatic synthesis reactions [41] and such biotechnological methods are finding increasing use for the selective derivitization, functionalization, and rearrangement of the sucrose molecule [42]. Sucrose often acts as a donor molecule for enzymatic transfer reactions to form oligosaccharides and polysaccharides. The advantages of biotechnological methods over chemical methods are that the reactions occur in water and without expensive protecting group chemistry [42], the field of biotechnology is rapidly advancing, particularly in the industrial production of enzymes that can be read-

ily utilized, they are environmentally friendly (waste is biodegradable), and more efficient conversions occur. Furthermore, the use of biotechnological methods will allow the addition of special functionalities to the sucrose molecule like biodegradability, biocompatibility, and non-toxicity, which are becoming important in the emerging bio-economy. Products formed from chemical and enzymatic reactions will be discussed in a later section of this chapter.

5.1 Current Commercial Value-Added Products from Sucrose

The several value-added commodity chemicals that are currently commercially produced from sucrose feedstocks include sucrose esters and fat substitutes, sweeteners, polyurethane foams, organic acids, polyols, and polysaccharides and oligosaccharides [20,41,42,43,44]. The eight hydroxyl groups in sucrose (*Fig. 3*) can be reacted with many fatty acid groups to form sucrose esters. Sucrose esters are ideal non-ionic surfactants; they are non-toxic and biodegradable. Sucrose esters used in the food industry include sucrose stearate, palmitate, laurate, behenate, oleate, and erucate; the quality and application depends on the percentage monoester content [43,45]. Sucrose can also be reacted with fatty acids to form a liquid sucrose polyester (hexa-through octa-substituted fatty acid esters of sucrose), a dietary fat-substitute that is marketed as Olestra^{TM} or Olean^{TM}. Olestra has properties similar to liquid vegetable oil with no calorific value and is often used in snack foods [41]. Sucrose acetate isobutyrate (SAIB) is a commercial mixture of sucrose esters that is used in the industrial, food, cosmetic, and pharmaceutical industries [43]. Sucrose octanoate esters are also currently being registered as "green" biopesticides and may have a bright and worthwhile future [46]. With the recent biotechnological advances in the production of new and better industrial enzymes [47], there is considerable interest in the improved production of sucrose esters (normally chemically synthesized) with enzymes [44,48].

Selective chlorination of sucrose, by which three of sucrose's hydroxyl groups are substituted with chlorine atoms, produces sucralose (1,6-dichloro-1,6-dideoxy-β-D-fructofuranosyl-(2↔1)-4-chloro-4-deoxy-α-galactopyranoside) sweetener sold as Splenda^{TM} [49]. Sucralose is approximately 650 times sweeter than sucrose and can be used in baking. Its synthesis involves treatment of a partially acetylated sucrose with sulfuryl chloride; the initially formed chlorosulfate esters act as leaving groups and undergo nucleophilic displacement by chlorine [20]. Sucralose was first approved for use in Canada in 1991; as of 2006, it has been approved in over 60 countries.

Other commercial sweeteners formed enzymatically from sucrose include isomaltulose (α-D-glucopyranosyl-(1→6)-D-fructose) or Palatinose^{TM}, and isomalt (a mixture of α-D-glucopyranosyl-(1→6)-D-mannitol and α-D-glucopyranosyl-(1→6)-D-glucitol) or Palatinit^{TM}, both produced on a large scale by Palatinit GmbH, a subsidiary of Sudzucker AG. Mitsui Sugar C. of Japan and Cargill Corp. of the US also produce isomaltulose. These enzymatically produced sweeteners are marketed as natural, non-cariogenic, low-glycemic index sweeteners [44].

Polyurethane polymers used in numerous industries including the paint, cotton, rubber, and metal industries [43] are mostly manufactured by the reaction of a polyol produced from sucrose with di-isocyanate. The polyol is produced from the reaction of sucrose with an alkylene oxide (such as ethylene, propylene, or butylenes oxide). Resulting polyurethane foams

can be used for packaging and insulation. Commercial products include Veranol™ (Dow Polyurethanes) and Pluracol™ 975 Rigid Polyol (BASF Chemicals) [43].

Currently, there are a few commodity chemicals produced by the fermentation of sucrose feedstocks. These include organic acids, polyols and, on a smaller scale because of the cost of production limiting the potential market, polysaccharides and oligosaccharides [43]. Advantages of fermentations (whole-cell or enzymatic) over traditional chemical synthesis are lower energy costs, higher yields, product specificity, and environmental friendliness. Some important value-added products of sucrose manufactured from fermentation reactions, are shown in *Fig. 5*. For a full description of all these products, the reader is referred to the recent reviews by Walford and Morel du Boil [43] and Côté [44]. Citric acid (E330) is an approved flavor compound, preservative, and buffer used in many industries. It is produced from a submerged fermentation of *Aspergillus niger* with sucrose or molasses (sugarcane or sugarbeet) as the substrate [50]. L-Ascorbic acid or Vitamin C is now increasingly being produced by fermentation as well, using sucrose or glucose as the substrate [51]. Lactic acid is utilized in numerous industries; it can be chemically synthesized to produce a racemic mixture (D and L forms) or produced as either the D or L form by fermentations. Sucrose, molasses, and corn syrup have been the preferred substrate feedstocks [52] with the latter being currently preferred. Sugar alcohols and polyols have traditionally been manufactured by the hydrogenation of sugars, including sucrose, using a nickel catalyst. Production of commercial sugar alcohols is now tending to enzymatic fermentation. Sucrose is also used as a feedstock for the fermentative production of a wide variety of microbial polysaccharides and oligosaccharides. Some microbial polysaccharides and oligosaccharides are synthesized uniquely from sucrose and depend on the action of glucansucrases. Examples of such products include dextran (α-(1→6)-linked-D-glucans) used in the pharmaceutical industry, and fructooligosaccharides (FOS) marketed as prebiotic nutraceuticals in the food industry (*Fig. 5*).

5.2 Possible Future Value-Added Products from Sucrose

The large-scale and economic diversification of sucrose into industries other than the food industry, is not a reality. However, this may change in certain countries as the concept and use of renewable carbohydrate feedstocks for biotransformations is growing rapidly. This is because carbohydrates are the most abundant organic materials in nature, they are renewable, and because of surging costs of traditional petroleum and natural gas feedstock in the chemical industry. Sucrose fits well into the renewable carbohydrate feedstock concept because of its availability, high purity, and because sugarcane and sugarbeet are amongst the plants giving the highest yields of carbohydrates per hectare. Products of sugarcane and sugarbeet, including sucrose are, therefore, ideal raw materials for the production of new biomaterials such as bioethanol, although commercialization will depend mostly on economic factors, such as government subsidies. Some countries are already producing bioethanol from sugarcane (and to a lesser extent sugarbeet) juice and/or molasses, particularly Brazil, but much more world-wide production is expected in the next 20 years. Other bioproducts will likely include biopolymers as viable alternatives to common plastic polymers. For example, lactic acid production as an intermediate for polymer and plasticizer manufacture could be from sucrose feedstocks [43]. Other example biopolymers from sucrose, are sucrose hydrogels (sucro acrylate gels) that are

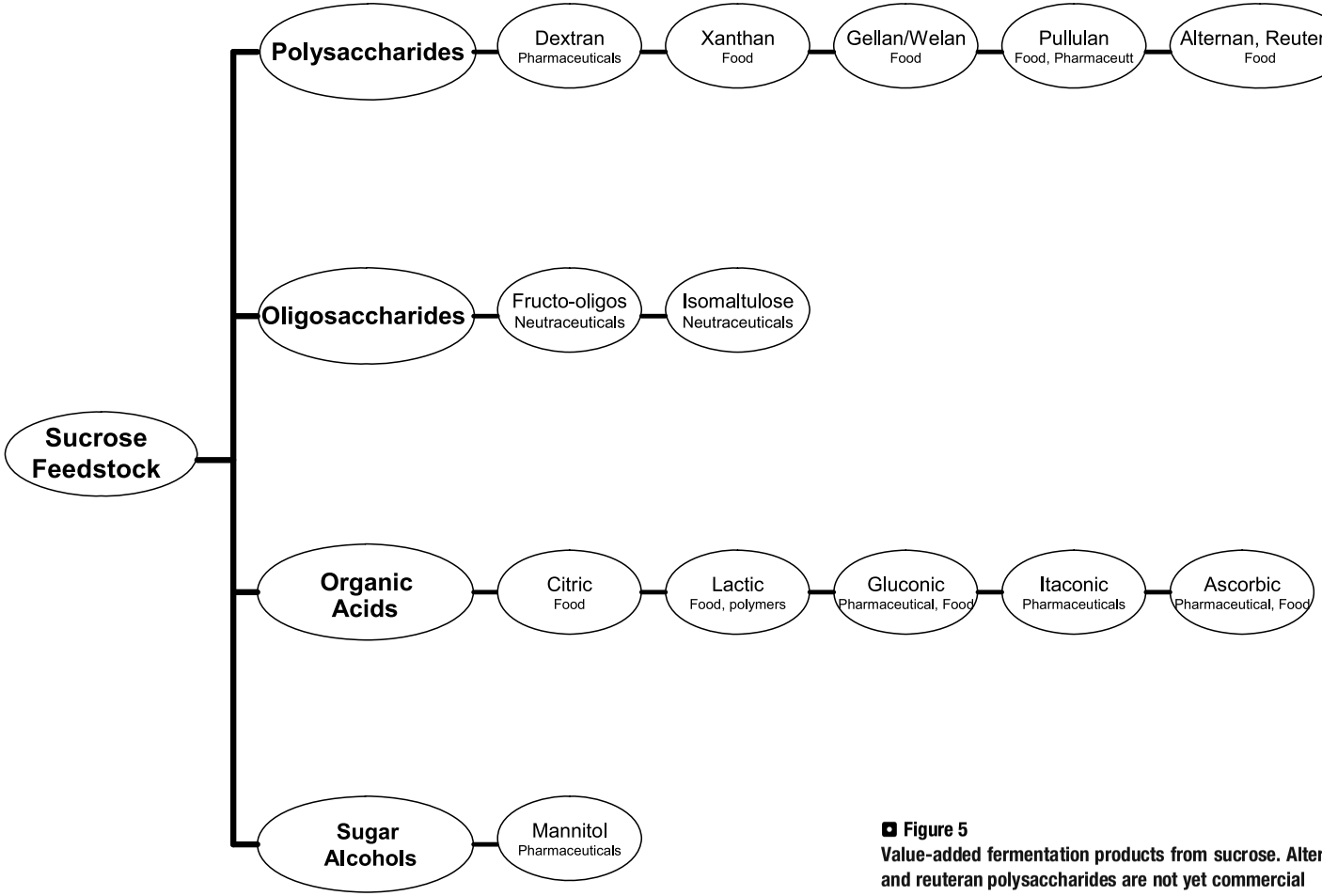


Figure 5
Value-added fermentation products from sucrose. Alternan and reuteran polysaccharides are not yet commercial

super-porous and fast-swelling. Sucrogels can be made in any size and shape and could find many industrial applications [41,53,54]. Epoxy sucrose-based products that can replace petrochemically derived diepoxide used in epoxy resins [55] also have potential, but so far have not been commercialized. With recent and expected biotechnology advances in the production of industrial enzymes [47], more enzymatically formed value-added products from sucrose are envisaged, similar to those discussed in the above section.

6 Oligosaccharides Related to Sucrose

Sucrose oligosaccharides, including isomers of sucrose are found in nature and others can be manufactured enzymatically. The following is a description of the most important sucrose related oligosaccharides.

Oligosaccharides can be grouped into two distinct classes: primary and secondary oligosaccharides [56]. Primary oligosaccharides are synthesized *in vivo* or *in vitro* from a mono- or oligosaccharide and a glycosyl donor by the action of a glycosyl transferase. Sucrose, and primary oligosaccharides based on sucrose, are the most ubiquitous plant primary oligosaccharides. In comparison, secondary oligosaccharides are formed *in vivo* or *in vitro* by hydrolysis of higher oligosaccharides, polysaccharides, glycoproteins, and glycolipids [23]. More detailed descriptions of secondary oligosaccharides can be found in other reviews by Côté [44], Eggleston and Côté [23], and Doner and Hicks [57].

6.1 Sucrose-Based Plant Oligosaccharides

Primary oligosaccharides based on sucrose are formed by the transfer of galactose (D-galactopyranosyl), glucose (D-glucopyranosyl), or fructose (D-fructofuranosyl) residues to sucrose with a glycosyl transferase. Series of families of homologous sucrose-related plant oligosaccharides exist, of which the "raffinose" and "kestose" series are the best known and most widespread [23]. The raffinose series of sucrose-related oligosaccharides is illustrated in ❯ *Fig. 6*. The trisaccharide raffinose is formed by the addition of galactose to the glucose moiety in the sucrose molecule via an α-(1→6) glycosidic linkage. The series is formed by the continual attachment of a galactose residue via an α-(1→6) linkage to the galactose moiety of the previous oligosaccharide in the series (❯ *Fig. 6*) and continues up to a nonasaccharide. These oligosaccharides act as important carbohydrate reserves in the vegetative storage organs and seeds of many plants, and also play an important role in sugar translocation [23,58]. They are found in soy products and cottonseed meal.

The kestose or fructan series of sucrose-related oligosaccharides is found in at least 12% of vascular plant species [23]. Kestoses contain one glucose moiety per oligosaccharide chain, derived from the parent sucrose molecule by the action of fructosyltransferases. The simplest kestose, monofructosyl-sucrose is a trisaccharide that exists as three isomers: 1-kestose (1^F-β-D-fructofuranosylsucrose), 6-kestose (6^F-β-D-fructofuranosylsucrose), and neo-kestose (6^G-β-D-fructofuranosylsucrose). These isomers form the basis of three major kestose subseries with different linkage patterns [23].

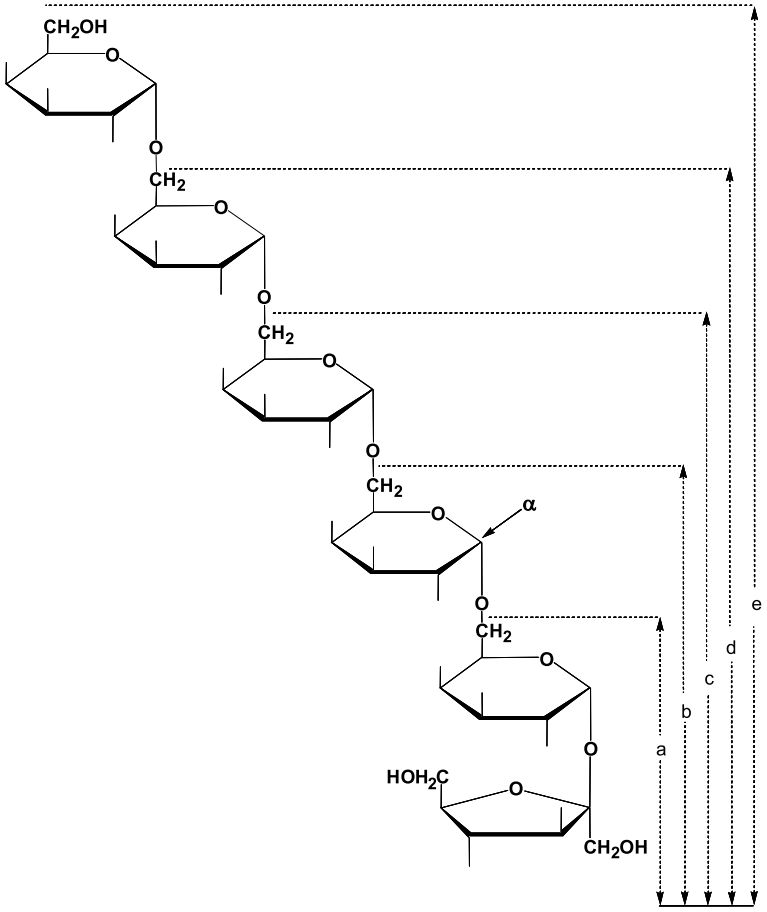


Figure 6
Raffinose series of oligosaccharides, (a) sucrose, (b) raffinose, (c) stachyose, (d) verbascose, (e) augnose

Planteose is a trisaccharide comprising sucrose and galactose (*O*-α-D-galactopyranosyl-(1→6)-*O*-β-D-fructofuranosyl-(2→1)-α-D-glucopyranoside) that is found in certain plants. It acts as a short-term storage carbohydrate in *Actinidia* leaves [59].

6.2 Sucrose-Based Honey Oligosaccharides

Honey represents the greatest source of oligosaccharides from plant material attacked by insects, and the sucrose content can range from 1 to 15% [59,60]. There are two types of honey: floral and honeydew. Floral honey originates from the nectar of flowers. Honeydew is obtained by the honeybee indirectly from sweet syrups excreted by various hemipterous insects feeding on tree sap. Honey oligosaccharides are presumed to arise from the action

of glycosidases from the bees [44], and have been the subject of comprehensive reviews by Doner [61] and Doner and Hicks [57]. Glycosidases catalyze transglycosylation reactions with sucrose, as well as condensation reactions. α-D-glucopyranosidases and β-D-fructofuranosidases present in honey cause not only the hydrolysis of sucrose in nectar, but also the formation of a variety of glucose and fructose containing oligosaccharides [44]. The most common oligosaccharides in honey include the glucose disaccharides: maltose (α-l→4 linked), kojibiose (α-1→2 linked), isomaltose (α-1→6 linked), and α,α-trehalose (α1↔α1 linked). Marked quantities of sucrose-related oligosaccharide isomers found in honey are: isomaltulose (α-D-glucopyranosyl-(1→6)-D-fructose), leucrose (α-D-glucopyranosyl-(1→5)-D-fructose), maltulose (α-D-glucopyranosyl-α-1→4-D-fructse), and turanose (α-D-glucopyranosyl-1→3-D-fructose). Glycosylated sucrose oligosaccharides are also found in honey and include erlose, melezitose, theanderose, and kestoses, which are discussed in more detail here.

Erlose (α-D-glucopyranosyl-(1→4)-α-D-glucopyranosyl-(1↔2)-β-D-fructofuranoside) is produced by the action of honey invertase on sucrose. Melezitose is an unusual oligosaccharide found in honey. Its structure is α-D-glucopyranosyl-(1→3)-β-D-fructofuranosyl-(2↔1)-α-D-glucopyranoside, which is illustrated in ❯ *Fig. 7*. Melezitose also occurs in honeydew plant exudates and various types of "manna" (white, crystalline substances) that periodically cover trees, particularly larch and Douglas fir trees [62]. Melezitose is produced from the action of aphids and other insects in the sap of plants [63]. Although melezitose has been consumed by humans for many years, little is known of its nutritional or health benefits [44].

Theanderose(α-D-glucopyranosyl-(1→6)-α-D-glucopyranosyl-(1↔2)-β-D-fructo-furanoside) is also referred to as isomaltosucrose. It is synthesized in honey from the transferase action of β-D-fructofuranosidase on sucrose. It can also be found in certain plants, including sugarcane, and has been used as a marker to differentiate sugarcane refined sugar from sugar beet refined sugar [64]. It has also been proposed as a non-cariogenic sweetener [65]. Kestoses are discussed in the previous section.

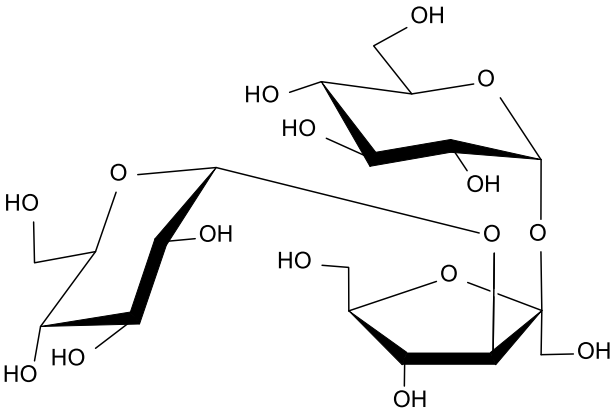


Figure 7
Structure of melezitose

6.3 Sucrose-Based In Vitro Oligosaccharides

Short chain fructooligosaccharides (scFOS; same as the kestose group discussed in the section above) can be prepared in vitro from the reaction of a microbial fructosyltransferase on sucrose. Up to 3–5 degrees of polymerization can be formed. These products have been studied extensively as prebiotic nutraceuticals [44].

Leucrose (α-D-glucopyranosyl-(1→5)-D-fructopyranose) is formed by dextransucrase-catalyzed transfer of glucosyl units to fructose acceptors. Dextransucrase occurs naturally in *Leuconostoc mesenteroides* lactic acid bacteria. The nominal reaction of dextransucrase is to incorporate the α-D-glucopyranosyl units from sucrose into dextran, an α(1→6)-linked D-glucan. Fructose is liberated in this reaction, and this fructose is used to produce leucrose in a side acceptor reaction. Leucrose is used as a sweetener, being approximately half as sweet as sucrose, and is non-cariogenic, resistant to acid hydrolysis, and releases glucose into the bloodstream only slowly [44]. It is manufactured commercially using immobilized dextransucrase, with Pfeifer and Langen of Germany the major producer. Other sucrose-based oligosaccharide sweeteners are xylosucrose (α-D-xylopyranosyl-(1↔2)-β-D-fructofuranoside) and lactosucrose (β-D-galactopyranosyl-(1→4)-α-D-glucopyranosyl-(1↔2)-β-D-fructofuranoside).

Xylosucrose is formed by the action of levansucrase on sucrose: D-fructofuranosyl units from sucrose are transferred to a xylose acceptor [66]. Xylosucrose can also be produced by the transferase action of β-D-fructofuranosidase. Lactosucrose has been studied [44] as a dietary ingredient for its prebiotic effect, and is produced commercially in Japan. It can be synthesized enzymatically via the fructosyltransferase reaction of either levansucrase or β-D-fructofuranosidase, with the latter reaction used in commercial production.

Oligosaccharides are also synthesized during the acid-catalyzed hydrolysis of sucrose, particularly under conditions of weak acid and high sucrose concentrations [67]. Oligosaccharides formed include kestoses [67] and IS1 (*O*-α-D-glucopyranosyl-(1→3)-D-fructose), IS2 (*O*-β-D-fructofuranosyl-(2→6)-D-glucose), and IS3 (*O*-α-D-glucopyranosyl-β-D-glucopyranoside) oligosaccharides. The mechanism for the formation of IS1-3 oligosaccharides was recently published by Thavarajah and Low [68].

References

1. Madore MA Glycoscience. Springer, Berlin Heidelberg New York, ch 5.4, p 1661
2. Legendre BL (1998) In: Clarke M, Godshall MA (eds) Chemistry and Processing of Sugarbeet and Sugarcane. Elsevier, Amsterdam, ch 12, p 176
3. Eggleston G, Monge A, Ogier B (2003) J Food Proc Preserv 26:433–454
4. Eggleston G, Monge A (2007) J Food Proc Preserv 31:52–72
5. Vaccari G, Mantovani G (1995) In: Mathlouthi M, Reiser P (eds) Sucrose Properties and Applications. Blackie Academic and Professional, London, ch 3
6. Clarke M (1988) In: Clarke M, Godshall MA (eds) Chemistry and Processing of Sugarbeet and Sugarcane. Elsevier, Amsterdam, ch 11, p 162
7. Van der Poel PW, Schiweck H, Schwartz T (1998) (eds) Sugar Technology. Beet and Cane Sugar Manufacture. Verlag Dr. Albert Bartens KG, Berlin, Germany
8. Eggleston G, Amorim H (2006) Int Sugar J 108:271–282
9. Bichsel SE (1988) In: Clarke M, Godshall MA (eds) Chemistry and Processing of Sugarbeet and Sugarcane. Elsevier, Amsterdam
10. Chen JCP, Chou CC (1993) (eds) Cane Sugar Handbook, 12th edn. Wiley, New York

11. Honig P (1953) (ed) Principles of Sugar Technology. Elsevier, New York
12. Hugot E (1972) (ed) Handbook of Cane Sugar Engineering, 2nd edn. Elsevier, Amsterdam
13. McGinnis RA (1982) (ed) Beet-Sugar Technology, 3rd edn. Beet Sugar Development Foundation, Fort Collins, Colorado
14. Bruhns G, Riffer R, Van Bekkum H, Lammers H, Schiweck H, Heitz F, Mauch W (1998) In: Van der Poel PW, Schiweck H, Schwartz T (eds) Sugar Technology. Beet and Cane Sugar Manufacture. Verlag Dr. Albert Bartens KG, Berlin, Germany, ch 1, p 37
15. Mathlouthi M, Reiser P (1995) (eds) Sucrose Properties and Applications. Blackie Academic, Glasgow
16. Lemieux RU, Huber G (1956)
17. Brown GM, Levy HA (1973) Acta Crystallogr, Sect B 29:790–797
18. Bock K, Lemieux RU (1982) Carbohyd Res 100:63–74
19. Lichtenthaler FW, Immel S (1995) Liebigs Ann Chem 1925
20. Lichtenthaler FW, Peters S (2004) Chimie 7:65–90
21. Robyt J, White BJ (1990) (eds) Biochemical Techniques. Theory and Practice. Waveland Press, Illinois, USA
22. Lescure JP (1995) In: Mathlouthi M, Reiser P (eds) Sucrose Properties and Applications. Blackie Academic, Glasgow, ch 7, p 155
23. Eggleston G, Côté G (2003) (eds) Oligosaccharides in Food and Agriculture. ACS Symposium Series 849, Oxford Univ Press, ch 1, pp 1–64
24. Folkes DJ, Jordan MA (2006) Food Sci Technol 159:1–40
25. Barker SA, Bourne EJ, Stacey M, Whiffen DH (1954) J Chem Soc, 171–176
26. Schallenberger R, Birch GG (1975) Sugar Chemistry. AVI Publishing Company, Westport, USA
27. Mathlouthi M, Colli AL, Koenig JL (1986) Carbohydr Res 147:1–10
28. Mathlouthi M, Luu DV (1980) Carbohydr Res 81:203–212
29. Lemieux RU, Stevens JD (1966) Can J Chem 44(3):249–262
30. Vogels JTWE, Terwel L, Tas AC, Van Der Berg F, Dukel F, Van Der Greef J (1996) J Agric Food Chem 44:175–180
31. Lowman DW, Maciel GE (1979) Anal Chem 51(1):85–90
32. Grabka J (1993) Ind Alim Agric 110:714–719
33. Anon (1994) The Braunschweig method for polarization of white sugar by polarimetry – Official. Methods GS2/3-1. ICUMSA Methods Book, Bartens, Germany
34. Eggleston G, Vercellotti JR, Edye L, Clarke M (1996) J Carbohydr Chem 15:1–94
35. Herzfield A (1988) Zuckerind 38:699
36. Dutton J (1979) In: Schneider F (ed) Sugar Analysis. ICUMSA publication, Peterborough, UK, p 30
37. Anon (1942) Polarimetry, saccharimetry and the sugars. NBS Circular C440. Washington, DC, USA, p 155
38. Meyer JH (1998) Proc Conf Sugar Proc Res, Savannah, USA, p 144
39. Eggleston G (2007) GS4 Molasses Report. Proc. 25th Session of ICUMSA. Bartens, Germany 88:40–48
40. Schaffler KJ, Morel du Boil PG (1984) Sugar Technol Rev 11:95–185
41. Godshall MA (2001) Int Sugar J 103:378–384
42. Roper H (2002) Starch/Starke 54:89–99
43. Walford SN, Morel du Boil PG (2006) Proc S Afr Sug Technol Ass 80:39–611
44. Côté GL (2007) In: Rastall RA (ed) Novel Enzyme Technology for Food Applications. Woodhead Publishing, Cambridge, UK (in press)
45. Queneau Y, Fitremann S, Trombotto CR (2004) Chimie 7
46. Chortyk OT, Pomonis JG, Johnson AW (1996) J Agric Food Chem 44:1551–1557
47. Eggleston G (2007) In: Eggleston G, Vercellotti JR (eds) Advances in the Applications of Industrial Enzymes on Carbohydrate Materials. ACS Symp Series, Oxford Univ Press, Oxford, ch 1, pp. 1–16
48. Plou FJ, Cruces MA, Ferrer M, Fuentes G, Pastor E, Bernable M, Christensen M, Comelles F, Parra JL, Ballesteros A (2002) J Biotechnol 96:55–66
49. Hough L (1989) Int Sugar J 91:231–31, 35, 37
50. Pazouki M, Felse PA, Sinha J, Panda T (2000) Bioproc Eng 22:353–361
51. Hancock RD, Viola R (2002) Trends Biotechnol 20:299–305
52. Narayanan N, Roychoudhury PK, Srivastava A (2004) Elec J Biotech 7:167–179
53. Patil NS, Dordick JS, Rethwisch DG (1996) Biomaterials 17: 2343–2350
54. Chen J, Park K (2000) Carb Polymers 41:259–268
55. Sachinvala ND, Winsor DL, Parikh DV, Solhjoo HH, Blanchard EJ, Bertoniere NR (1998) Proc Beltwide Cotton Conf 1:784

56. Kandler O, Hopf H. (1980) In: Preiss J (ed) The biochemistry of plants. Academic Press, New York, vol 3, pp 221–270
57. Doner LW, Hicks KB (1983) In: Lineback DR (ed) Food Carbohydrates. AVI Publish. Co., Westport, Connecticut, pp 74–112
58. Jeremias K (1969) Ber Deut Bot Ges 1:87
59. Klages KU, Boldingh HL, Clooney JM, MacRae EA (2004) Funct Plant Biol 31:1205–1214
60. Hendry G (1987) New Phytol 106:201–216
61. Siddiqui IR (1977) Adv Carbohydr Chem Biochem 30:371–444
62. Doner LW (1977) J Sci Food Agric 28:443–456
63. Hudson CS (1946) Adv Carbohydr Chem 2:1–36
64. Bacon JSD, Dickinson B (1957) Biochem J 66:289–297
65. Morel du Boil PM (1997) Int Sugar J 99:102–106
66. Kitahata S, Fujita K (1993) In: Nakakuki T (ed) Oligosaccharides – Production, Properties and Applications. Gordan and Breach Science, Switzerland, p 158
67. Avigad G, Feingold DS, Hestrin S (1956) Biochim Biophys Acta 20:129–134
68. Richard GN (1988) In: Clarke M, Godshall MA (eds) Chemistry and Processing of Sugarbeet and Sugarcane. Elsevier, Amsterdam, p 253
69. Thavarajah P, Low NH (2006) J Agric Food Chem 54:2754–2760

5.3 Oligosaccharides in Food and Agriculture

Michelle E. Collins, *Robert A. Rastall**
Department of Food Biosciences, The University of Reading,
Reading RG6 6AP, UK
r.a.rastall@reading.ac.uk

Abstract

Oligosaccharides are an integral part of the daily diet for humans and animals. They are primarily used for their nutritional properties, however they are currently receiving much attention due to their physiological effect on the microflora of the gastrointestinal tract. Galacto-oligosaccharides and the fructan-type oligosaccharides, namely FOS and inulin are well established as beneficial to the host and are classified as prebiotic based on data from clinical studies. These compounds dominate this sector of the market, although there are oligosaccharides emerging which have produced very interesting in vitro results in terms of prebiotic status and human trials are required to strengthen the claim. Such compounds include pectic oligosaccharides, gluco-oligosaccharides, gentio-oligosaccharides, kojio-oligosaccharides, and alternan oligosaccharides. The raw materials for production of these prebiotic compounds are derived from natural sources such as plants but also from by products of the food processing industry. In addition to being prebiotic these compounds can be incorporated into foodstuffs due to the physiochemical properties they possess.

Keywords

Prebiotics; Efficacy; Applications; Properties

In: *Glycoscience*. Fraser-Reid B, Tatsuta K, Thiem J (eds)
Chapter-DOI 10-1007/978-3-540-30429-6_27:

Abbreviations

DGGE	denaturing gradient gel electrophoresis
FAB-MS	fast atom bombardment mass spectrometry
NMR	nuclear magnetic resonance
qRT-PCR	quantitative real time polymerase chain reaction
SCFA	short chain fatty acids

1 Introduction

Carbohydrates are an intrinsic part of the daily diet for mammals. In nature they exist in many forms, however it is oligosaccharides which are of interest here. The production of oligosaccharides for use in food (or as part of the diet) is either by direct extraction from plants [1], controlled hydrolysis of polysaccharides [2], or by enzymatic synthesis using enzymes which possess hydrolytic or transglycosylation activity [3,4]. Oligosaccharides are composed of 2 to 20 monosaccharide units joined together by glycosidic bonds and exist as a linear or branched arrangement.

Awareness of the health benefits of certain food-grade oligosaccharides has increased in recent years and has received particular interest in Japan followed by the United States and Europe.

Table 1
Prebiotic oligosaccharides

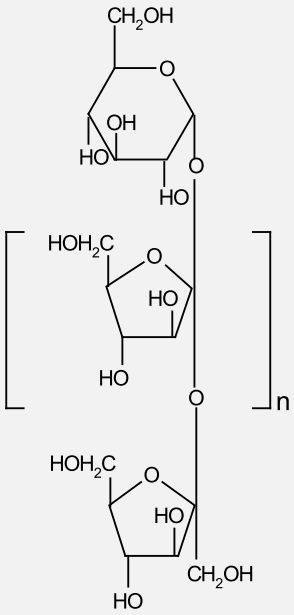
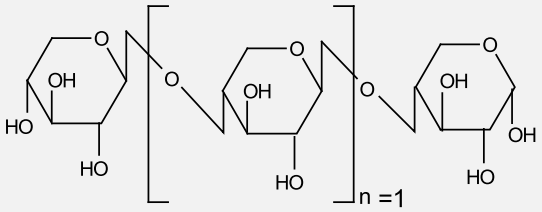
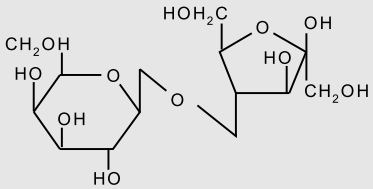
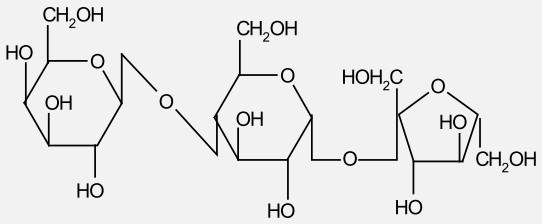
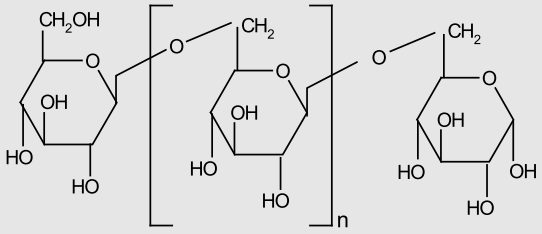
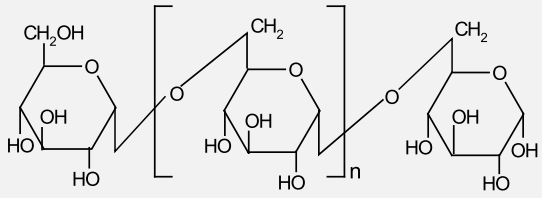
Prebiotic oligosaccharides	Structure
Fructo-oligosaccharides	

Table 1
(continued)

Prebiotic oligosaccharides	Structure
Xylo-oligosaccharides	
Lactulose	
Lactosucrose	
Gentio-oligosaccharides	
Isomalto-oligosaccharides	

Such oligosaccharides include the fructan-type oligosaccharides; inulin and fructose oligosaccharides (FOS), galacto-oligosaccharides (GOS), isomalto-oligosaccharides (IMO), lactulose, gentio-oligosaccharides, soybean oligosaccharides (SOS), xylo-oligosaccharides (XOS), and lactuosucrose (*Table 1*). Some occur naturally in several foods such as leek, asparagus,

chicory, Jerusalem artichoke, garlic, onion, wheat, banana, and oats, as well as soybean. However, these foods contain only trace levels of oligosaccharides, so developments in functional foods have taken the approach of removing the active ingredients from such sources and adding them to more frequently consumed products [5]. The oligosaccharides listed are classified as prebiotic. Defined by Gibson and Roberfroid [6], a prebiotic is "a non-digestible food ingredient that beneficially affects the host by selectively stimulating the growth and / or activity of one or a limited number of bacteria in the colon that have the potential to improve host health." A prebiotic dose of 5 g day^{-1} should be sufficient to elicit a positive effect upon the gut microbiota (in some exceptional cases this may be nearer to 8 g day^{-1}). Doses below 5 g day^{-1} are not likely to be effective as the overall dietary carbohydrate load into the gut will mask any prebiotic effect. A possible side effect of prebiotic intake is intestinal discomfort from gas production. However, the beneficial bacteria cannot produce gas as part of their metabolic process. Therefore, at a rational dose, of up to 20 g day^{-1}, gas distension should not occur. If gas is being generated then the carbohydrate is not acting as an authentic prebiotic. This is perhaps because dosage is too high and the prebiotic effect is being compromised, i. e. bacteria other than the target organisms are becoming involved in the fermentation [6].

It is apparent that the indigenous microflora can dictate the well being of the host in synergy with diet. However, exercise also has a role in health and longevity. The microflora of the gastrointestinal tract are thought to comprise around 95% of the total cells in the body [7]. In mammals there exists variability in bacterial numbers and populations between the stomach, small intestine, and colon. The total bacterial count in gastric contents is usually below 10^3 g^{-1}, with numbers in the small intestine ranging from about 10^4 ml^{-1} contents to about 10^6–10^7 at the terminal ileum. The colon is the most heavily populated area of the gastrointestinal tract (GIT), with numbers typically in the region of 10^{12} g $contents^{-1}$ [8]. The majority of bacteria are non-sporulating anaerobes and can be categorized as being either beneficial or potentially pathogenic due to their metabolic activities and fermentation end products. The resident GIT organisms ferment substances that cannot be digested by the host in the small intestine. These include resistant starch, non-digestible carbohydrates, oligosaccharides, proteins, and mucins. The two main types of fermentation that are carried out in the gut are saccharolytic and proteolytic. Saccharolytic fermentation is more favorable than a proteolytic fermentation due to the metabolic end products that are formed. The main products of saccharolytic fermentation are the short chain fatty acids, acetate, propionate, and butyrate. All contribute towards the host's daily energy requirements. Acetate is metabolized in systemic areas like muscle and used to generate ATP, while propionate is transported to the liver. Butyrate is an important source of energy for the colonocytes and has anti-tumor properties. The end products of proteolytic fermentation on the other hand, include toxic metabolites (such as phenolic compounds, amines, and ammonia) some of which are carcinogens [9].

Bifidobacteria and lactobacilli are considered to be beneficial bacteria. They positively influence the host and are associated with immunostimulation, improved digestion, and absorption, syntheses of vitamins, inhibiting the growth of potential pathogens, reducing cholesterol, and lowering gas production. These bacteria are classified as probiotic [10] and as they are indigenous to the host they can be selectively promoted by the consumption of prebiotics.

Of the prebiotics listed some have received extensive investigation whereas efficacy data on others may not be so abundant. The fructan-type oligosaccharides, GOS and IMO will be discussed in further detail here. Pectic oligosaccharides would be considered as emerging prebi-

otics. Although their effect in vivo has not been fully investigated, results from in vitro studies are interesting and exciting and thus this oligosaccharide will also be discussed. Reviews on the remaining oligosaccharides are available [11].

2 Fructan-Type Oligosaccharides

Inulin, levan, graminan, phlein, and kestoses are classified as fructans [12]. Inulin exists either as a linear, branched, or cyclic fructan whereas the structure of the remaining fructans are linear and branched (*Table 2*) [13].

Fructans occur naturally, are predominantly found in plants, fungi, and are produced extracellularly by bacteria such as *Streptococcus mutans* [13]. Plants are the most likely source of inulin for the production of material to be incorporated into foodstuffs due to its abundance in plants and also from a safety perspective. Inulin is found in the tubers and roots of the plant family *Compositae* which includes aster, dandelion, dahlias, comos, burdock, goldenrod, chicory, lettuce, and Jerusalem artichoke [14,15]. Van Loo et al. [16] identified the quantity of inulin in various plants as well as their degree of polymerization (*Table 3*). An extensive review of occurrence and distribution of fructans in nature is provided [16,17,18].

Niness [19] describes the commercial production of inulin from chicory roots where the inulin is extracted using processes similar to those employed for the extraction of sucrose from sugar beet. Chicory inulin is a linear β-(2→1)-linked fructan and is composed of a mixture of oligosaccharides in which the DP (degree of polymerization) varies from between 2 to 65 monosaccharide units with an average DP of 12. The final spray dried product consists of 6–10% sugars constituting glucose, fructose, and sucrose [12,19].

Fructo-oligosaccharides are produced by the enzymatic hydrolysis of inulin with an endoinulinase (EC 3.2.1.7) or by the transfructosylation activity of a β-fructofuranosidase (EC 3.2.1.26 and EC 3.2.1.80) using sucrose as the substrate. The latter method can utilize a batch system using soluble enzymes or a continuous system using enzymes or whole cells immobilized, for example, onto a suitable ion-exchange resin [20]. For synthesis of FOS β-fructofuranosidase (β-FFase) from Aspergillus species, *A. niger* and *A. japonicus* are often used due to their high fructosyltransferase activity. As well as being used as soluble enzymes their enzymes have been immobilized onto various matrixes such as gluten [21], oxirane containing polymeric beads [22], and ceramic membranes [23]. Cheng et al. [24] immobilized β-FFase from *A. japonicus* by entrapment in calcium alginate gel in a packed bed reactor. After 25 days

Table 2
Fructan structures

Name	Linkage	Structure	Origin
Inulin	β (2–1)	Linear, branched, cyclic	Plant, bacteria, fungi
Levan	β (2–6)	Linear, branched	Plant, bacteria, fungi
Phlein	β (2–6)	Linear, branched	Plant
Graminan	β (2–1) and β (2–6)	Linear, branched	Plant
Kestoses	β (2–1) and β (2–6)	Linear, branched	Plant

Table 3
Inulin content and chain length in plants

Natural food	% Fructan on fresh	DP
Onion (*Allium cepa*)	1.1–7.5 (Avg. 3.6)	1–12
Jerusalem Artichoke (*Helianthus tuberosus*)	17–20.25	2–19 = 74%, 19–40 = 20%, >4 = 6%
Globe artichoke (*Cynara scolymus*)	1.8% (fresh), 1.7% (cooked)	> 5 = 95%
Chicory (*Cichorium intybus*)	15.2–20.5 (Avg. 16.2)	2–19 = 55%, 19–40 = 28%, >40 = 17%
Asparagus (*Asparagus officinalis*)	2.6 (shoots), 0.1 (fresh plant)	
Leek (*Allium ampeloprasum*)	2.9 (fresh leaves)	12
Garlic (*Allium sativum*)	12.98	> 5 = 75%, Avg. DP = 15
Salsify	4.2	> 5 = 75%
Banana (*Musa cavendishii*)	0.7	< 5 = 100%
Wheat (*Triticum aestivum*)	1.17 (white flour)	7–8
Rye (*Secale cereale*)	0.6%	
Barley (*Hordeum vulgare*)	22.1 (young kernals), 1.1 (mature kernals)	
Dandelion (*Taraxacum officinale*)	12.8 (leaves)	

DP = degree of polymerization

17% of enzyme activity was lost, however, appreciable amounts of FOS were formed. Park et al. [25] report on the production of neo-FOS using immobilized whole cells of *Penicillium citrinum* as an alternative to immobilizing the actual enzyme. Continuous production of FOS was successfully maintained in a packed bed reactor for 50 days thereby indicating the potential for use in an industrial scale. Others have demonstrated the effect of a mixed enzyme system for FOS production. Sheu et al. [26] used a β-fructofuranosidase from *A. japonicus* and a glucose oxidase enzyme. During the synthesis reaction glucose is released as a by-product which is subsequently inhibitory to the activity of β-fructofuranosidase. However, the presence of glucose oxidase converted the residual glucose to gluconic acid which was then precipitated to calcium gluconate. The final mixture contained >90% w w^{-1} FOS on a dry weight basis with small amounts of glucose, sucrose, and calcium gluconate. As for all synthesis reactions the source of enzyme, the pH, the reaction temperature and substrate concentration will influence the yield and type of product. These oligosaccharides are composed of 1-kestose (GF2), nystose (GF3), and 1F-fructofuranosyl nystose (GF4), in which fructosyl residues are bound at the 2-OH position of the fructofuranosyl moiety of sucrose (GF) [19,27].

Inulin and FOS are commercially produced in Japan and Europe (*Table 4*) and have widespread applications as food ingredients. However, in the context of this chapter the application of inulin and FOS will be discussed as a functional food ingredient beneficial for the gastrointestinal tract of humans and animals.

Stimulating the growth of the beneficial bacteria resident in the colon is a prerequisite for any oligosaccharide to be considered as a prebiotic. In vitro studies using pure or mixed cultures and fructan oligosaccharides as the sole carbon source provide initial results to support

Table 4
Commercial fructan-containing products

Product	Type	Source	Content / DP	Fructose glucose sucrose
Beneo™ ST	Powder	Chicory	Inulin >90%, average DP ≥ 10	Fructose and glucose ≤4%, sucrose ≤8%
Beneo™ HP	Powder	Chicory	Inulin > 99.5%, average DP ≥23	≤ 0.5%
Beneo™ oligofructose	Powder	Partial enzymatic hydrolysis of chicory inulin	Oligofructose ≥93.2%, DP 2–8	≤ 6.8%
Actilight 950S and 950P	Syrup	Sugar beet	GF2 37% ±6%, GF3 53% ±6%, GF4 10% ±6%	5% ±2%
Meioligo	Syrup	Sucrose	GF2 25–30%, GF3 10–15%, GF4 5–10%	25–50%

Beneo™ products are manufactured by Orafti (Belgium)
Actilight products are manufactured by Beghin-Meiji Industries (France)
Meioligo™ products are manufactured by Meiji Seika Kaisha (Japan)
DP = degree of polymerization
GF2 = 1-kestose
GF3 = nystose
GF4 = fructofuranosylnystose

the prebiotic effect of these compounds. Gibson and Wang [28] observed that the growth of *Bifidobacterium infantis* was better on oligofructose compared to the effect on seven other bifidobacteria strains. Observations were based on changes in specific growth rates, culture pH values, and enzyme production. Others have also observed the preference of bifidobacteria for oligofructose compared to inulin, for example *B. animalis* was more efficient at fermenting oligofructose than inulin [29]. Marx et al. [30] investigated the effect of FOS derived from levan on the growth of *Bifidobacterium adolescentis*, *B. longum*, *B. breve*, and *B. pseudocatenulatum*. Of these bifidobacteria species the growth of *B. adolescentis* was better, additionally high concentrations of organic acids were produced by this strain. The FOS derived from hydrolysis of inulin is β-(2→1)-linked whereas levanoligosaccharides are β-(2→6)-linked. This difference in structure could ultimately affect their fermentability by the indigenous microflora, the chemical nature of the FOS and the degree of depolymerization will also influence the extent of fermentation. The DP of levanoligosaccharides is approximately 20 whereas that of FOS from inulin has a DP of 2–8. Yamamoto et al. [31] used FOS from levan which had a DP of 37; no increase in growth was observed; therefore in this study it could be the molecular mass which was hindering the metabolism of FOS rather than the source of FOS. The ability of various lactobacillus species to ferment short chain purified FOS was determined in pure culture studies [32]. Of the 16 strains selected 12 of these utilized FOS and *Lactobacillus rhamnosus* GG, a well studied and commercially used probiotic, did not. The effect on bifidobacteria species was also investigated and *B. bifidum* did not grow; this observation does not corroborate with Rossi et al. [33] who confirmed the growth of various types of this species on Raftilose P95. As bifidobacteria and lactobacillus species are reported to exert beneficial effects on the host [34]

these strains are of particular interest in in vitro studies to determine if they grow and how well they grow on various substrates which may be potential prebiotics. Rossi et al. [33] grew pure bifidobacteria cultures in a basal media containing FOS or inulin; of the 55 strains tested only seven utilized inulin whilst all strains utilized FOS. The DP of FOS was 2–10 whereas that of inulin was 25; as previously mentioned the longer chain may influence the extent of fermentability. Roberfroid et al. [29] reported results from in vitro studies whereby molecules with a DP greater than 10 are not fermented as quickly as those molecules with a DP of 10 or less.

Pure culture studies are not sufficient to confirm the efficacy of a potential prebiotic. Batch cultures [35], single-stage continuous cultures [36], or a three-stage continuous culture [37], the latter representing the human colon, further validate the claim. However, the ultimate confirmation comes from the results obtained from human clinical trials. Various trials have been performed where inulin or FOS has been incorporated into the diet and the main outcomes from these studies are described in the following text

A balanced multiple crossover trial incorporated Raftilose P95 at two concentrations (5 g day^{-1} and 10 g day^{-1}) into the diet of healthy men [38]. Fecal samples were analyzed for changes in pH, SCFA, and weight. However, analysis of the fecal microflora were not determined at the beginning and the end of the study. Catabolism of the FOS was confirmed as no FOS could be recovered in the feces or urine. Breath H_2 excretion increased on consumption of the supplements and higher levels were detected in individuals consuming 10 g day^{-1} of FOS, therefore suggesting the effect is dose dependent. Flatulence also increased at the higher dose. SCFA in feces could not be detected, which is not surprising as it is more probable they were absorbed by the colonic mucosa [39]. Bouhnik et al. [40] investigated the effect of short FOS (SC-FOS: 44% GF_2, 46% GF_3, and 10% DF_4) in healthy individuals. The suggested dose for FOS was identified as 10 g day^{-1} due to the mild side effects of excess flatulence and bloating. Levels of bifidobacteria increased as the concentration of FOS in the diet increased. Bifidobacteria express high levels of β-fructosidase which hydrolyze terminal, non-reducing β-D-fructose residues in fructans. Traditional microbiological tools were utilized to ascertain the prebiotic effect. Kruse et al. [41] investigated the effect of inulin in healthy individuals. A dose of 22–34 g day^{-1} of inulin was administered. Fluorescent in situ hybridization (FISH) was used to enumerate changes in the bacterial population and bifidobacteria numbers significantly increased. There was a moderate increase in flatulence, bloating, and lipids in blood and levels of SCFA remained unchanged. Until now the studies have used inulin or FOS as an additive to diets. Tuohy et al. [42] performed a trial with biscuits, containing partially hydrolyzed guar gum and FOS. Thirty-one volunteers were recruited and changes in bacterial populations of feces were determined by FISH. Although the overall results from the study suggest a bifidogenic effect compared to the placebo and control, no differences in bifidobacteria levels were identified in those individuals who had high levels of bifidobacteria at the beginning of the study. Thus, the health benefits of prebiotics may be more applicable to those suffering from IBS, elderly patients who are in hospital and at risk of nosocomial infections and antibiotic acquired diarrhea. Infants are a segment of the population who could also benefit from prebiotics at an early age. It is well established that babies who are breast fed have higher levels of bifidobacteria compared to those on a formula diet. This implies that breast fed babies may be less susceptible to enteric infections compared to their counterparts on infant formula. Thus, the supplementation of formula with a prebiotic is an attractive approach to

overcome these problems. Moro et al. [43] fed infants a FOS-GOS formula containing either 0.4 g day^{-1} or 0.8 g day^{-1} FOS. A significant increase in bifidobacteria was evident at the end of the study and was dose related. Lactobacillus also increased, pH of feces decreased and stools were softer as the dose of FOS-GOS increased. Guesry et al. [44] studied FOS as the only supplement in a formula for term infants. Dose was 3 g day^{-1}, almost ten times higher than that administered by Moro et al. [43]. No bifidogenic effect was observed and no side effects were recorded. The gold standard for a clinical trial is a double blind placebo controlled randomized study. The trials discussed above did not always follow this standard. Volunteers for studies are often healthy and consuming a Western diet. It is apparent that the parameters assessed at the beginning, middle or end of a trial do include (1) changes in bacterial populations using either traditional or molecular methods, (2) pH of feces and SCFA composition, (3) stool characteristics e. g. weight and frequency, (4) side effects e. g. bloating, flatulence, excessive breath hydrogen. However, the methods used to assess these parameters are not always consistent and it is particularly important when enumerating changes in the bacterial population. Molecular methods such as FISH, DGGE, and qRT-PCR are superior to traditional microbiological methods. Additional functionalities of fructan-type oligosaccharides are reported [45].

3 Galacto-Oligosaccharides

Lactose (4-O-β-D-galactopyranosyl-D-glucopyranose) is a simple sugar naturally found in milk and also in whey. Disposal of whey, a by-product of cheese manufacturing, used to be an environmental problem due to its high biological oxygen demand. However, whey is now used as a valuable source of lactose. A variety of products can be derived from lactose [46]. Hydrolysis yields the monosaccharides D-glucose and D-galactose which are sweeter and more soluble than lactose. Additionally, they are digested and absorbed in the small intestine, a desirable attribute, as 70% of the world's population are reported to be lactose intolerant, i. e., deficient in the enzyme β-galactosidase. Chemical treatment of lactose produces lactulose, lactitol, or lactobionic acid formed by isomerization, reduction, or oxidation reactions, respectively [47]. Although these products are reported to have beneficial effects on the host [48], it is the synthesis of galacto-oligosaccharides (GOS) from lactose which is of particular interest to this chapter as they are purported to be prebiotic. GOS are produced by galactosyl transfer catalyzed by the transgalactosylation activity of β-galactosidases. These enzymes (EC 3.2.1.23 β-D galactoside galactohydrolase) are isolated and purified from various organisms including *Aspergillus oryzae* [49], *A. niger* [50], *Bacillus circulans* [51], *Kluyveromyces lactis* [52], *K. fragilis* [53], *Streptococcus thermophilus* [54], *Escherichia coli* [55], *Lactobacillus* spp. [56], and *Bifidobacteria* spp. [57].The source of the enzyme will influence its optimal pH and temperature. Fungal β-galactosidases have an optimal pH between 2.5 and 4.5, whereas the optima for yeast and bacterial β-galactosidases are 6–7 and 6.5–7.5, respectively. Factors influencing the production and yield of GOS include pH, temperature, enzyme concentration, and processing conditions. For efficient transgalactosylation activity the initial concentration of lactose is the most significant factor [58]. The higher the concentration the higher the yield of GOS as the amount of available water is low. In reactions where low levels of lactose are used (e. g. 5% w v^{-1}), hydrolysis occurs due to the large amounts of water present as it acts

as an acceptor. GOS are found naturally in human milk and to a smaller extent in bovine milk [59]. These compounds mainly consist of galactosyl galactose and galactosyl glucose and are classified as prebiotic as they are not digested by the β-galactosidase from the human small intestine, and they are reported to selectively stimulate the beneficial bacteria of the colon.

Structures of GOS products isolated from the transgalactosylation reaction mixture have been elucidated. Gel permeation chromatography and charcoal chromatography are tools used to separate the different oligomers, whereas methylation analysis, FAB-MS and NMR are used to confirm the structure. Yanahira et al. [60] confirmed that the predominant linkage of GOS disaccharide products produced using the β-galactosidase from *B. circulans* was β-(1→6), however Usui et al. [61] reported the formation of β-(1→4) linked disaccharides. β-Galactosidases extracted from various bifidobacteria species predominantly produced β-(1→6) oligomers [57] whereas those from *Lactobacillus reuteri* produce both β-(1→6) and β-(1→3) linked oligomers [56]. These structures confirm that the enzyme source influences the structure of the oligosaccharides (*Table 5*).

Table 5
Oligosaccharides formed by transgalactosylation activity of β-galactosidase

Structure	Reference
Disaccharides	
β-D-Gal*p*(1→2)-D-Glc	[49]
β-D-Gal*p*(1→3)-D-Glc	[49,53]
β-D-Gal*p*(1→6)-D-Glc	[49,52,53,56]
β-D-Gal*p*(1→3)-D-Gal	[49,56]
β-D-Gal*p*(1→6)-D-Gal	[49,52,53,54,56]
Trisaccharides	
β-D-Gal*p*(1→6)-β-D-Gal*p*(1→4)-D-Glc	[52]
β-D-Gal*p*(1→4)-β-D-Gal*p*(1→4)-D-Glc	
β-D-Gal*p*(1→3)-β-D-Gal*p*(1→4)-D-Glc	[49]
β-D-Gal*p*(1→6)-β-D-Gal*p*(1→6)-D-Glc	[52]
β-D-Gal*p*(1→6)-[β-D-Gal*p*(1→4)]-D-Glc	[49,56]
β-D-Gal*p*(1→3)-[β-D-Gal*p*(1→2)]-D-Glc	[56]
β-D-Gal*p*(1→3)-[β-D-Gal*p*(1→6)]-D-Gal	[52]
β-D-Gal*p*-(1→6) ⟩ D- Glc β-D-Gal*p*-(1→4)	[49]
Tetrasaccharides	
β-D-Gal*p* (1→6)-β-D-Gal*p* (1→6)-β-D-Gal*p*-(1→4)-D-Glc	[49]
Pentasaccharides	
β-D-Gal*p* (1→6)-β-D-Gal*p* (1→6)-β-D-Gal*p*-(→-6)-β-D-Gal*p* (1→4)-D-Glc	[49]

The prebiotic effect of GOS may be more pronounced if GOS is synthesized using β-galactosidases from probiotic bacteria. Rabiu et al. [57] used bifidobacterial β-galactosidases to produce GOS and subsequently investigated the effect of these products on the growth rate of the corresponding bifidobacteria strain. As expected the growth rate was highest when the corresponding oligosaccharides were used. This approach could be investigated in further in vitro studies as well as in vivo. To confirm the selectivity of GOS for certain probiotic strains, species specific primers or probes would be required. Commercial GOS currently available include Oligomate 55® and Vivinal produced by Yakult (Japan) and Borculodomo Ingredients (The Netherlands), respectively. The prebiotic efficacy of these products has been investigated in in vitro, animal, infant, and human studies [62]. Of particular interest to those purchasing GOS for incorporation into foods are the data from clinical trials, but unfortunately the results from these investigations are mixed. Early studies [63] indicated an increase in bifidobacteria at the end of the feeding period (43 mg kg^{-1} day^{-1}). This effect was further confirmed in two studies by Ito et al., who administered GOS at 15 g day^{-1} [64] and 2.5 g day^{-1} [65]. In the latter study the initial levels of bifidobacteria in the volunteers were low and significantly increased during the consumption of GOS, but numbers returned to starting levels after discontinuing the treatment. A longer feeding study was carried out by Bouhnik et al. [66]. Volunteers received 10 g GOS day^{-1} for 21 days and again, a significant increase in bifidobacteria levels was observed. Conversely, Teuri et al. [67] did not detect a change in bifidobacteria levels when individuals consumed yogurt containing 15 g GOS on a daily basis. The methods used to enumerate the changes in the bacterial population were traditional microbiology techniques and these do have their limitations, whereas the introduction of molecular tools is advantageous (68). Human studies which suggest GOS did not have a prebiotic effect were performed by Alles et al. [69], Alander et al. [70], and Satokari et al. [71] even though the dosage of GOS was comparable to the earlier in vivo studies. These results indicate the necessity for further clinical trials, possibly introducing a more standardized approach as well as stringent analytical conditions, thus enabling a direct comparison between studies. Additionally, the current data is generated from healthy individuals and not those who would be more likely to benefit from such treatment, such as the elderly or those suffering from acute or chronic gastrointestinal disorders.

4 Isomalto-Oligosaccharides

Starch, cellulose, and murein are three of the most abundant organic compounds found on earth. Starch is of particular interest as it is the starting material for production of IMO. Starch exists as a mixture of amylose and amylopectin. Amylose, is a linear polysaccharide comprised of D-glucose units linked α-(1→4) to each other and amylopectin is a branched polysaccharide composed of D-glucose units linked α-(1→4), with 5% of linkages α-(1→6) [72]. The starch component of many plants is a mixture of amylose and amylopectin, although the starches from waxy varieties of maize, rice, sorghum, and barley are completely composed of amylopectin.

In the manufacture of IMO the starch is initially hydrolyzed by α-amylase (EC 3.2.1.1) to malto-oligosaccharides followed by treatment with pullulanase (EC 3.2.1.41) and β-amylase (EC 3.2.1.2) and finally α-glucosidase is added. The α-glucosidase (EC 3.2.1.20) catalyzes

a transglucosylation reaction whereby the α-(1→4) malto-oligosaccharides are converted into α-(1→6) IMO [73]. Isomalt 900, a commercial IMO product, has been produced this way and the mixture consists of isomaltose (α-D-Glc*p*-(1→6)-D-Glc), isomaltotriose (α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→6)-D-Glc) and panose (α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→4)-D-Glc). By definition, IMO are "glycosyl saccharides containing one or more α-1,6 glucosidic linkages with or without α-1,4 glucosidic linkages" [74]. Non-commercial IMO are also produced, the starting material is dextran rather than starch [2,75]. These typically contain a much higher percentage of α-(1→6) linkages, often to the exclusion of other types.

α-Amylases for use in IMO production have been isolated and purified from *Bacillus subtilis* [76] and *B. stearothermophilus* [77]. α-D-Glucosidases possessing hydrolytic and transglucosylation activity are mainly isolated from Aspergillus species including *A. niger* [78] and *A. carbonarious* [79,80]. *Aureobasidium pullulans* also produces an α-glucosidase [81]. As in the manufacture of many oligosaccharides, the source of enzyme, substrate concentration and processing conditions influence the yield of product. A product of high purity demands a higher price. Pan and Lee [74] have used yeast cells to ferment residual glucose from an IMO mix.

The prebiotic properties of IMO have been reported. Kohmoto et al. [73] used pure cultures of bifidobacteria species. The growth rate increased for all except *Bifidobacterium bifidum*. Additionally this strain did not ferment the individual substrates; panose, isomaltose, and isomaltotriose. Olano-Martin et al. [82] used a non-commercial IMO mix (oligodextrans) in a three-stage continuous system to monitor the in vitro effect in the proximal (vessel 1), transverse (vessel 2), and distal (vessel 3) regions of the colon. Bifidobacteria and lactobacilli were stimulated, with highest levels detected in vessel 1. However, bacteroides species predominated the bacterial populations in all three vessels. High concentrations of the SCFA butyric acid were produced. This acid is an end fermentation product of eubacteria and clostridia and its production is favored due to its effects on colonocytes [39]. Clinical trials substantiated the prebiotic claim associated with IMO's. Irrespective of dose (6–20 g day^{-1}) of the commercial IMO product, Isomalt 900, an increase in bifidobacteria was observed; however, increases were only significant when the dose was ≥15 g day^{-1} [73,83,84]. Although a prebiotic effect was observed it is unlikely that the disaccharide (isomaltose) contributes to this effect as it was probably digested in the small intestine.

IMO's occur naturally in some fermented Japanese foods, e. g. soybean sauce and sake. Levels in these products may be considerably lower than those used in the human studies, thus consumption of Isomalt 900 would contribute to a healthier microflora and thus enhance host well being.

5 Pectic Oligosaccharides

Pectin, a waste product of the fruit juice industry, has many applications in the food ingredients sector. It forms gels in the presence of Ca^{2+} ions and its solubility is determined by its molecular weight and degree of esterification as well as solution pH and temperature. Pectin can be incorporated into jams, jellies, low calorie foods as a fat or sugar replacer, and frozen foods [85]. Found in the cell wall of all plants, concentrations are higher in the fruits rather than the rest of the plant. Some fruits contain higher amounts of pectin than others; for

instance apple pulp contains 10–15% w w^{-1} pectin whereas orange and lemon peel contain 20–30% w w^{-1} [72]. The reader is referred to Chap. 6.4 for a discussion of pectin structure. The American Association of Cereal Chemists define dietary fiber as "the edible parts of plants or analogous carbohydrates that are resistant to digestion and absorption in the human small intestine with complete or partial fermentation in the large intestine" [86]. Pectins are classified as dietary fiber. The enzymes associated with degradation of pectin into its corresponding oligosaccharides are mainly endo-type pectate lyases (EC 4.2.2.10) as well as pectate hydrolases (EC 3.2.1.15). Studies have identified this enzyme from bacteria indigenous to the gastrointestinal tract and thus they exhibit the ability to degrade pectin [87,88,89]. Intestinal bacteria include *Bacteroides thetaiotaomicron* which produce pectate lyase and pectate hydrolase and the *Clostridium butyricum–Clostridium beijerinckii* group produces endo-type pectate lyase [90]. Endo type pectate lyases are found in human feces. These bacteria digest pectin to 4–5 unsaturated digalacturonic acid, which is subsequently fermented to short chain fatty acids, namely acetic. Such acids are absorbed across the intestinal wall and used as an energy source. *C. butyricum* has been safely used as a probiotic for many years in Eastern Asia [91] and are reported to produce high levels of SCFAs in the gut lumen and in vitro [92]. Pectin is desirable as a food ingredient—it has excellent gelling properties and is degraded by certain indigenous bacteria. However, the by-products from the preparation of pectin could also be incorporated into the diet. Manufacture of pectic oligosaccharides from citrus or apple pectins by enzymatic hydrolysis (1) or steam injection [93] are desirable and inexpensive approaches to producing a food ingredient with added functionality.

Pectic oligosaccharides are good candidates to be considered for prebiotic testing as they are derived from pectin which arrives at the colon intact. This attribute could also be considered for pectin hydrolysis products, but no results are available yet to confirm this. There are few studies which have investigated the prebiotic potential of POS, and although they are in vitro, results are promising. Olano-Martin et al. [94] were the first to compare the effect of pectin and POS on the growth of pure cultures of various species indigenous to the gastrointestinal tract. Selection criteria for potential prebiotics include "selective stimulation of the beneficial bacteria" and pure culture studies would demonstrate this well. The pectin used was high methyl (60–66%) citrus pectin (HMP) and low methyl (8%) apple pectin (LMP). Their corresponding oligosaccharides were produced in a membrane reactor by controlled enzymatic hydrolysis of pectin using an endo-polygalacturonase [1]. All bacteria (bifidobacteria sp., lactobacillus sp., bacteroides sp., clostridium sp., and *E. coli* sp.) tested grew well on the LMP and the oligosaccharides derived from LMP. A follow-up study was performed using pH-controlled fecal batch cultures [94]. A bifidogenic effect was observed for pectins and their products, as bifidobacteria numbers significantly increased as the fermentation progressed until 24 hours, after which numbers started to decline. It is interesting to note that bacteroides numbers remained constant when POS was used as the carbon source but increased when LMP was used, suggesting the degree of esterification influences the fermentation by gut bacteria. As bacteroides did not change in the presence of POS it strengthens the prebiotic claim—failure to sustain the growth of potentially detrimental bacteria while promoting the growth of beneficial bacteria. Manderson et al. [95] determined the prebiotic properties of POS derived from orange albedo using in vitro batch cultures. HPAEC-PAD analysis confirmed that the degree of polymerization was probably greater than the POS used by Olano-Martin et al. [94]. However, an increase in bifidobacteria numbers was also observed, thus supporting observations of previous studies.

Additionally, *Eubacterium rectale* levels increased with a concomitant increase in butyric acid. Presence of this SCFA is desirable due to its health promoting properties in the gastrointestinal region [39]. POS produced by enzymatic treatment of bergamot peel were evaluated for prebiotic properties in fecal batch cultures [96]. The degree of polymerization was between 3 and 7 and the carbohydrate composition was markedly different from the POS derived from orange albedo. Although differences in composition and structure compared to other POS, the POS from bergamot peel increased the numbers of bifidobacteria and lactobacilli, whereas levels of clostridia decreased, thus corroborating with results from previous studies. A beneficial effect from in vitro studies was observed.

However, to further substantiate the claim of the prebiotic efficacy of POS, more rigorous in vitro investigations are required and ultimately in vivo studies will validate the claim. The disadvantages of in vitro methods are the absence of synergistic, antagonistic, and/or competitive effects as well as the absence of an immune system. Pectic oligosaccharides are an exciting new development in prebiotics as they can be cheaply manufactured from agricultural waste materials. As well as the reported prebiotic effect, POS have additional functionalities including repression of liver lipid accumulation in rats [97], action as anti-fungal phytoalexin-elicitors in plants [98], activity as anti-bacterial agents [99], protection of colonocytes against *E. coli* verocytotoxins [100] and stimulation of apoptosis in human colonic adenocarcinoma cells [101].

6 Novel Prebiotics

Despite the range of prebiotic oligosaccharides on the market (described above), there is much interest in developing novel forms of oligosaccharides as prebiotics. In most cases, these novel forms have only been evaluated using in vitro systems and are awaiting human data to confirm their efficacy.

6.1 Gentio-Oligosaccharides

Gentio-oligosaccharides (GeOS) are a commercial product manufactured by Nippon Shokuhin Kako in Japan. They contain mainly β-(1→6)-linked D-glucose units. They are claimed to be prebiotic but there is very little information available. One study has shown that they have prebiotic potential in vitro [102]. In 24 hour fecal batch culture [103], GeOS were compared to fructo-oligosaccharides (FOS) and the GeOS resulted in the largest increases in bifidobacteria, lactobacilli, and total bacteria. FOS, however, were more selective as they did not stimulate the growth of non-probiotic bacterial groups. GeOS also produced the highest levels of SCFA and the lowest gas production.

6.2 Gluco-Oligosaccharides

Gluco-oligosaccharides (GlcOS) can be made by the action of an enzyme, dextran dextrinase produced by *Gluconobacter oxydans* on maltodextrins. They have been made in whole-cell bioreactors [104] and evaluated in fecal batch cultures [105] and in three-stage gut models [106]. Mountzouris et al. [104] used methylation analysis to determine the linkages of

the GlcOS as being a mixture of 1→6 and 1→4 linked with some 1→4→6 branching. GlcOS increased numbers of bifidobacteria and lactobacilli within 24 hour in batch culture. Bacteroides, clostridial and eubacterial populations were slightly decreased by 48 hour. The same materials have also been studied in a three-stage continuous model of the human colon. GlcOS resulted in increases in numbers of bifidobacteria and lactobacilli in all three vessels, representing the proximal, transverse, and distal colonic areas. The prebiotic indices of the glucooligosaccharides were 2.29, 4.23, and 2.74 in V1, V2, and V3, respectively. Palframan et al. [107] established the quantitative index based on the in vitro fermentation of carbohydrates by key bacterial groups. Furthermore, the PI is a relationship between changes in the "beneficial" and "undesirable" elements within the microflora, all of them related to their starting levels.

6.3 Alternan Oligosaccharides

The bacterium *Leuconostoc mesenteroides* produces a range of enzymes which produce dextran-type polysaccharides from sucrose [108]. *L. mesenteroides* NRRL B-21297 produces an alternansucrase enzyme which can transfer glucose from sucrose onto acceptor oligosaccharides such as maltose. In this way a series of α-(1→6) and α-(1→3)-linked oligosaccharides can be produced. After size fractionation, oligosaccharide fractions with average DP values of 3, 4, 5, 5.7, 6.7, and 7.4 were obtained and tested using pure cultures or a fecal batch culture system. In pure culture, most of the bacteria tested failed to grow well on DP6.7 and DP7.4 fractions and grew best on DP3 [109]. In mixed fecal culture, DP3 resulted in the highest prebiotic effect, followed by DP4 and DP6.7. DP7.4 was not selectively fermented [110].

6.4 Functionally Enhanced Prebiotics

As we study the fermentation properties of an increasing range of oligosaccharides we are starting to unravel the structure-function relationships in these molecules [111]. This, coupled with advances in carbohydrate bioengineering, will allow the knowledge-based design of prebiotics with specific functional enhancements.

7 Oligosaccharides in Animal Feed

Extensive data are available on the prebiotic efficacy of oligosaccharides from in vitro models representing the human colon and from human clinical trials. Whether results from these studies can be extended to animals needs to be determined. The physiology and microflora of the gastrointestinal tract will vary from animal to animal and in most cases there may be limited or no knowledge of the microecology.

As antibiotics at subtherapeutic levels are banned in many countries, the addition of prebiotics to animal feed is becomes increasingly attractive. This is more to improve the well being of the animal nutritionally rather than improve or sustain a healthy microflora. Withdrawal of antibiotics from feed may leave livestock more susceptible to enteric infections. Pigs, calves, poultry, and rabbits often suffer from Clostridia-associated infections, whereas broilers are prone to Campylobacter and Salmonella infections, pigs are also at risk from pathogenic

E. coli infections. If a prebiotic can be added to the daily feed and it subsequently stimulates the beneficial bacteria in the colon at the expense of pathogens, it is most probable that these bacteria, namely bifidobacteria and lactobacilli, adhere to sites on epithelial cells and ultimately prevent the binding of pathogenic bacteria [112]. In diets containing reasonable quantities of carbohydrates saccharolytic fermentation will prevail, thus the pH of the GIT will remain stable and subsequently reduce the onset of Clostridia infections, as a more alkaline pH is required by species of this genus. Inulin added to rabbit feed was fermented in the cecum, produced SCFA, and reduced the risk of clostridiosis [113].

Literature on the effect of prebiotics in livestock animals is currently not widely available, although this may change with time as the necessity for addition of prebiotics to feed is increasing. Tzortzis et al. [114] investigated the efficacy of a novel GOS mixture in a parallel continuous randomized pig trial. Pigs were fed one of the following diets; (1) a commercial diet, (2) a diet supplemented with 1.6% w w^{-1} GOS, (3) a diet supplemented with 4% w w^{-1} GOS or (4) a diet supplemented with 1.6% w w^{-1} inulin. The feeding period lasted 34 days after which samples were collected for enumeration of bacteria, SCFA, and pH. A bifidogenic effect was found in those pigs consuming a diet supplemented with 4% w w^{-1} GOS, and an increase in acetate and a decrease in pH was observed. These findings are similar to those reported by Smiricky-Tjardes et al. [115]. However, in vitro studies suggest the microflora of swine feces produce lower levels of acetate when GOS is added to the fermentation, whereas higher levels of acetate were produced when raffinose, stachyose, or xylo-oligosaccharides were used as the carbon source [116].

Lactulose (β-D-Gal*p*-(1→4)-D-Fru) is also considered to be prebiotic in humans at sublaxative doses [48]. Decreased levels of *Clostridium perfringens* were associated with the inclusion of lactulose into the diet of perioparturient sows [117]. Interestingly, counts remained low 10 days after the supplementation ended. Furthermore, high antibody levels of IgG to *C. perfringens* phospholipase C developed and these were passed to the piglets via the colostrum.

Enteric diseases in livestock (pigs, poultry sheep, calves, and others) are becoming a concern more so today than in previous years due to the increased resistance of certain pathogens to antibiotics [118] and also due to the banning of subtherapeutic antibiotic administration in Europe. This ban could eventually extend to the USA [119]. As previously mentioned prebiotics are an alternative and if these are combined with probiotics, thus resulting in a synbiotic preparation, a very desirable substitute will be available. However, extensive testing may be required in animal models other than rats due to the complex structure of the GIT and diverse microflora. Some animals may have shorter transit times compared to others, and the pH of the region may vary somewhat. Overcoming these hurdles and others will result in a healthier livestock for human consumption.

8 Physical Properties of Oligosaccharides and their Application in Foods

In addition to exerting a beneficial effect on host health, prebiotics can be incorporated into foods as they possess various physiochemical properties. Playne and Crittenden [48] identified desirable attributes for these oligosaccharides (*Table 6*). Currently the Japanese are the market leaders as regards prebiotic food stuffs, but America and Europe are slowly catching up.

Table 6
General properties of oligosaccharides

Reduced sweetness (typically 0.3–0.6 times that of sucrose)
Not digested by brush border enzymes of the small intestine, pancreatic juices and enzymes
Selectively stimulate the indigenous beneficial bacteria of the gastrointestinal tract
Modify the viscosity and freezing points of foods
Influence emulsification, gel formation, and gel binding
Exhibit bacteriostatic properties
Favorably alter the color of foods
Possess properties similar to dietary fiber
Low caloric value
Act as an anti-caries agent
Stable at high temperatures and at various pH values
Act as a humectant and control moisture

This realization of a gap in the market for such products has stemmed from the emphasis on health and longevity and all sections of the population can be targeted, from preterm infants to the elderly. The prebiotics dominating worldwide markets are inulin, FOS, and GOS. Suppliers of the fructan-type oligosaccharides include Orafti, Sensus, Cosucra, Beghin Meiji, and BioNeutra Inc. Yakult, Borculodomo Ingredients, Snow Brand Milk Products, and Nissin Sugar are suppliers of GOS. Foods in which these prebiotics can be added to are bakery products, breakfast cereals, table spreads, frozen desserts, and dairy products. The latter food group have seen the addition of inulin and FOS to yogurts and as well as exhibiting a prebiotic effect, the oligosaccharides can influence the mouthfeel and sweetness of the product. In bread, GOS is added to prevent starch retrogradation, additionally the GOS is not fermented by the yeast and therefore the bread has an acceptable taste and texture. GOS and FOS have also been incorporated into infant formula by Nutricia and Milupa who market Omneo and Conformil products, respectively. The physiochemical properties of inulin are well documented and are widely added to foods [120]. A white odorless product with a bland taste, inulin is often added to bread mixes and baked goods to retain moisture and replace sugar. Inulin has also been used as a fat replacer where high performance inulin shows better functionality than the standard chicory inulin [121].

To summarize, the established prebiotics have a dual functionality and it is likely this could also be extended to the less established and emerging prebiotics.

References

1. Olano-Martin E, Mountzouris KC, Gibson GR, Rastall RA (2001) J Food Sci 66:966
2. Mountzouris KC, Gilmour SG, Rastall RA (2002) J Food Sci 67:1767
3. Barreteau H, Delattre C, Michaud P (2006) Food Technol Biotechnol 44:323
4. Monsan P, Paul F (1995) FEMS Microbiol Reviews 16:187
5. Gibson GR, Berry Ottaway P, Rastall RA (2000) Prebiotics: New Developments in Functional Foods. Chandos Publishing Limited, Oxford
6. Gibson GR, Roberfroid MB (1995) J Nutr

125:1401

7. Gibson GR, Willems A, Reading S, Collins MD (1996) Proc Nutr Soc 55:899
8. Finegold SM, Sutter VL, Mathisen GE (1983) Normal indigenous intestinal flora. In: Hentges DJ (ed) Human Intestinal Microflora in Health and Disease. Academic Press, London, p. 3
9. Conway PL (1995) Microbial ecology of the human large intestine. In: Gibson GR, Macfarlane GT (eds) Human Colonic Bacteria: Role in Nutrition, Physiology and Pathology. CRC Press, Boca Raton, FL, p. 1
10. Rastall RA, Gibson GE, Gill HS, Guamer F, Klaenhammer TR, Pot B, Reid G, Rowland IR, Sanders ME (2005) FEMS Microbiol Ecol 52:145
11. Tuohy KM, Rouzaud GCM, Brück WM, Gibson GR (2005) Curr Pharm Design 11:75
12. Suzuki M, Chatterton NJ (eds) (1993) Science and Technology of fructans. CRC Press, Boca Raton, Fl
13. Roberfroid M (2005) Inulin Type Fructans: Functional Food Ingredients. CRC Press, Boca Raton, Fl
14. Schutz K, Muks E, Carle R, Schieber A (2006) Biomed Chromatogr 20:1295
15. Incoll LD, Bonnett GD (1993) The occurrence of fructan in food plants. In: Fuchs A (ed) Inulin and inulin containing crops. Elsevier Science, Amsterdam, p. 309
16. Van Loo J, Coussement P, De Leenheer L, Hoebregs H, Smits G (1995) Critic Rev Food Sci Nutr 35:525
17. Carpita NC, Kanabus J, Housley TL (1989) J Plant Physiol 134:162
18. Hendry GAF, Wallace RK (1993) The origin, distribution and evolutionary significance of fructans. In: Suzuki M, Chatterton NJ (eds) Science and Technology of Fructans. CRC Press, Boca Raton, Fl, p. 119
19. Niness KR (1999) J Nutr 129:1402
20. Kono T (1993) Fructooligosaccharides. In: Nakakuki T (ed) Oligosaccharides. Gordon and Breach Science Publishers, New York, p. 50
21. Chien CS, Lee WC, Lin TJ (2001) Enzyme Microb Technol 29:252
22. Lee WC, Chiang CJ, Tsai PY (1999) Ind Eng Chem Res 38:2564
23. Nishizawa K, Nakajima M, Nabetani H (2000) Biotechnol Bioeng 68:92
24. Cheng CY, Duan KJ, Sheu DC, Liu CT, Li SY (1996) J Chem Technol Biotechnol 66:135
25. Park MC, Lim JS, Kim JC, Park SW, Kim SW (2005) Biotechnol Lett 27:127
26. Sheu DC, Lio PJ, Chen ST, Lin CT, Duan KJ (2001) Biotechnol Lett 23:1499
27. Yun JW (1996) Enzyme Microb Technol 19:107
28. Gibson GR, Wang X (1994) Food Micrbiol 11:491
29. Roberfroid MB, Van Loo JAE, Gibson GR (1998) J Nutr 128:11
30. Marx SP, Winkler S, Hartmeier W (2000) FEMS Microbiol Lett 182:163
31. Yamamoto Y, Takahashi Y, Kawano M, Lizuka M, Matsumoto T, Saeki S, Yamaguchi H (1999) J Nutr Biochem 10:13
32. Kaplan H, Hutkins RW (2000) Appl Environ Microbiol 66:2682
33. Rossi M, Corradini C, Amaretti A, Nicolini M, Pompei A, Zanoni S, Matteuzzi D (2005) Appl Environ Microbiol 71:6150
34. Simon GL, Gorbach SL (1984) Gastroenterolgy 86:174
35. Rumbey CJ, Rowland IR (1992) Crit Rev Food Sci Nutr 34:229
36. Gibson GR, Wang X (1994) FEMS Microbiol Lett 118:121
37. Gibson GR, Wang X (1994) J Appl Bacteriology 77:412
38. Alles MS, Hautvasti JGA, Nagengast FM, Hartemink R, Van Laere KMJ, Jansen JBMJ (1996) Brtish J Nutr 76:211
39. Cummings JH (1984) Proce Nutr Soc 43:35
40. Bouhnik Y, Vahedi K, Achour L, Attar A, Salfati J, Pochart P, Marteau P, Flourie B, Bornet F, Rambaud J (1999) J Nutr 129:113
41. Kruse HP, Kleessen B, Blaut M (1999) Brtish J Nutr 82:375
42. Tuohy KM, Kolida S, Lustenberger AM, Gibson GR (2001) British J Nutr 86:341
43. Moro G, Minoli I, Mosca M, Fanaro S, Jelinek J, Stahl B, Boehm G (2002) J Pediatric Gastro Nutr 34:291
44. Guesry PR, Bodanski H, Tomsit E (2000) J Pediatr Gastroenterol Nutr 31:252
45. Scholz-Ahrens KE, Ade P, Marten B, Weber P, Timm W, Asil Y, Glüer CC, Schrezenmein J (2007) J Nutr 137:S838
46. Thelwall LAW (1985) Developments in the chemistry and chemical modification of lactose. In: Fox PE (ed) Developments in Dairy Chem-

istry 3: Lactose and Minor Constituents. Elsevier, London, pp 35–67
47. Holsinger VH (1997) Physical and chemical properties of lactose. In: Fox PE (ed) Advanced Dairy Chemistry 3: Lactose, Water, Salts and Vitamins. Chapman and Hall, London
48. Playne MJ, Crittenden R (1996) Bulletin of IDF 313:10
49. Toba T, Adachi S (1978) Dairy Sci 61:33
50. Toba T, Tomita Y, Itoh T, Adachi S (1981) Dairy Sci 64:185
51. Yanahira S, Kobayashi T, Suguri T, Nakakoshi M, Miura S, Ishikawa H, Nakajima I (1995) Biosci Biotechnol Biochem 59:1021
52. Asp NG, Burvall A, Dahlqvist A, Hallgren P (1980) Food Chem 5:147
53. Pazur JH, Tipton CL, Budovich T, Marsh JM (1958) J Am Chem Sci 80:119
54. Greenberg NA, Mahoney RR (183) 10:195
55. Huberm RE, Kurz G, Wallenfels K (1976) Biochemistry 15:1994
56. Splechtna B, Nguyen TH, Steinbock M, Kulbe KD, Lorenz W, Haltrich D (2006) J Agric Food Chem 54:4999
57. Rabiu BA, Jay AJ, Gibson GR, Rastall RA (2001) Appl Environ Micro 67:2526
58. Boon MA, Janssen AEM, van't Riet K (2000) Enzyme Microb Tech 26:271
59. Kunz C, Rudloff S (1993) Acta Paediatrica 82:903
60. Yanahira S, Kobayashi T, Suguri T, Nakakoshi M, Miura S, Ishikawa H, Nakajimab I (1995) Biosci Biotech Biochem 59:1021
61. Usui T, Kubota S, Ohi H (1993) Carbohydr Res 244:315
62. Rastall RA (2006) Galactooligosaccharides as prebiotic food ingredients. In: Gibson GR, Rastall RA (eds) Prebiotics Development and Application. Wiley, Chichester, West Sussex, p. 101
63. Tanaka R, Takayama H, Morotomi M, Kuroshima T, Ueyama S, Matsumoto K, Kuroda A, Mutai M (1983) Bifidobacteria Micro 2:17
64. Ito M, Deguchi Y, Miyamori A, Matsumoto K, Kikuchi H, Kobayashi Y, Yajima T, Kan T (1993) Micro Ecol Health Dis 3:285
65. Ito M, Kimura M, Deguchi Y, Miyamori-Watabe A, Yajima T, Kan T (1993) J Nutr Sci Vita 39:279
66. Bouhnik Y, Flourie B, D'Agay-Abensour L, Pochart P, Gramet G, Durand M, Rambaud JC (1997) J Nutr 127:444
67. Teuri U, Korpela R, Saxelin M, Montonen L, Salminen S (1998) J Nutr Sci Vita 44:465
68. McCartney AL (2002) British J Nutr 88:S29
69. Alles MS, Hartemink R, Meyboom S, Harryvan JL, Van Laere KMJ, Nagengast FM, Hautvast JGA (1999) Am J Clin Nutr 69:980
70. Alander M, Mättö J, Kneifel W, Johansson M, Kögler B, Crittenden R, Mattila-Sandholm T, Saarela M (2001) Int Dairy J 11:817
71. Satokari RM, Vaughan EE, Akkermans ADL, Saarela M, de Vos W (2001) Sys Appl Micro 24:227
72. Robyt JF (1998) Essentials of Carbohydrate Chemistry. Springer, Berlin Heidelberg New York
73. Kohmoto T, Fukui F, Takaku H, Machida Y, Arai M, Mitsuoka T (1988) Bifidobact Microflora 7:61
74. Pan Y, Lee C (2005) Biotech Bioeng 89:797
75. Goulas AK, Fisher DA, Grimble GK, Grandison AS, Rastall RA (2004) Enzyme Microb Tech 35:327
76. Kuriki T, Yanase M, Takata H, Takesada Y, Imanaka T, Okada S (1993) Appl Environ Microbiol 59:953
77. Lee HS, Auh JH, Yoon HG, Kim MJ, Park JH, Hong SS, Kang MH, Kim TJ, Moon TW, Kim JW, Park KH (2002) J Agric Food Chem 50:2812
78. Kurimoto M, Nishimoto T, Nakada T, Chaen H, Fukuda S, Tsujisaka Y (1997) Biosci Biotechnol Biochem 61:699
79. Duan KJ, Sheu DC, Lin MT, Hsueh HC (1994) Biotechnol Lett 16:1151
80. Sheu DC, Huang CI, Duan KJ (1997) Biotechnol Tech 11:287
81. Yun JW, Lee MG, Song SK (1994) Biotechnol Lett 16:1145
82. Olano-Martin E, Mountzouris KC, Gibson GR, Rastall RA (2000) British J Nutr 83:247
83. Kohomoto T, Fukui F, Takaku H, Mitsuoka T (1991) Agric Biol Chem 55:2157
84. Kaneko T, Yokoyama A, Suzuki M (1995) Biosci Biotechnol Biochem 59:1190
85. Thakur BR, Singh RK, Handa AK (1997) Crit Rev Food Sci Nutr 37:47
86. American Association of Cereal Chemists Report (2001) Cereal Foods World 46:112
87. MacCarthy RE, Kotarski SF, Slayers AA (1985) J Bacteriol 161:493
88. Jensen NS, Canale-Parola E (1985) Appl Environ Microbiol 50:172
89. Matsuura Y (1991) Agric Biol Chem 55:885

90. Nakajima N, Ishihara K, Tanabe M, Matsubara K, Matsuura Y (1999) J Biosci Bioeng 88:33
91. Fuller R (1989) J Appl Bact 66:365
92. Araki Y, Andoh A, Fujiyama Y, Takizama J, Takizawa W, Bamba T (2002) Int J Mol Med 9:53
93. Fishman ML, Walker PN, Chau HK, Hotchkiss AT (2003) Biomacromolecules 4:880
94. Olano-Martin E, Gibso GR, Rastall RA (2002) J Appl Microbiol 93:505
95. Manderson K, Pinart M, Tuohy KM, Grace WE, Hotchkiss AT, Widmer W, Yadhav MP, Gibson GR, Rastall RA (2005) Appl Environ Microbiol 71:8383
96. Mandalari G, Nueno Palop C, Tuohy K, Gibson GR, Bennett RN, Waldron KW, Bisigano G, Narbad A, Faulds CB (2007) Appl Microbiol Biotechnol 73:1173
97. Yamaguchi F, Shimizu N, Hatanaka C (1994) Biosci Biotech Biochem 58:679
98. Bishop PD, Pearce G, Byrant JE, Ryan CA (1984) J Biol Chem 259:13172
99. Iwasaki K, Inoue M, Matsubara Y (1998) Biosci Biotech Biochem 62:262
100. Olano-Martin E, Williams MR, Gibson GR, Rastall RA (2003) FEMS Microbiol Lett 218:101
101. Olano-Martin E, Rimbach GH, Gibson GR, Rastall RA (2003) Anticancer Res 23:341
102. Rycroft CE, Jones MR, Gibson GR, Rastall RA (2001) Lett App Microbiol 32:156
103. Rycroft CE, Jones MR, Gibson GR, Rastall RA (2001) J Appl Microbiol 91:878
104. Mountzouris KC, Gilmour SG, Jay AJ, Rastall RA (1999) J Appl Microbiol 87:546
105. Wichienchot S, Prasertsan P, Hongpattarakere T, Gibson GR, Rastall RA (2006) Curr Issues Intest Microbiol 7:13
106. Wichienchot S, Prasertsan P, Hongpattarakere T, Gibson GR, Rastall RA (2006) Curr Issues Intest Microbiol 7:7
107. Palframan R, Gibson GR, Rastall RA (2003) Curr Issues Intest Microbiol 4:71
108. Côté GL, Holt SM, Miller-Fosmore C (2003) ACS Symposium Series February 18, 1
109. Holt SM, Miller-Fosmore CM, Côté GL (2005) Lett App Microbiol 40:385
110. Sanz ML, Côté GL, Gibson GR, Rastall RA (2005) J Agric Food Chem 53:5911
111. Sanz ML, Gibson GR, Rastall RA (2005) J Agric Food Chem 53:5192
112. Crittenden R (2006) Emerging prebiotic carbohydrates. In: Gibson GR, Rastall RA (eds) Prebiotics Development and Applications. Wiley, Chichester, UK
113. Maertens L, Aerts J, De Boever J (2004) World Rabbit Sci 12:235
114. Tzortzis G, Athanasios KG, Gee JM, Gibson GR (2005) J Nutr 135:1726
115. Smiricky-Tjardes MR, Grieshop CM, Flickinger EA, Bauer LL, Fahey GC (2003) Am Soc Animal Sci 81:2535
116. Smiricky-Tjardes MR, Flickinger EA, Grieshop CM, Bauer LL, Murphy MR, Fahey GC (2003) Am Soc Animal Sci 81:2505
117. Krueger M, Schroedl W, Isik K, Lange W, Hagemann L (2002) Eur J Nutr 41,26
118. Levy SB (2002) J Antimicrob Chemoth 49:25
119. Patterson JA, Burkholder KM (2003) Poultry Sci 82:627
120. Franck A, Coussement PAA (1997) Food Ingred Anal Int October, 8
121. Franck A (2002) Br J Nutr 87:287

5.4 Combinatorial Methods in Oligosaccharide Synthesis

Katsunori Tanaka, *Yukari Fujimoto*, *Shin-ichi Tanaka*, *Yasutaka Mori*, *Koichi Fukase**
Department of Chemistry, Graduate School of Science, Osaka University, Toyonaka, Osaka 560–0043, Japan
koichi@chem.sci.osaka-u.ac.jp

Abstract

This chapter will focus on recent achievements in the combinatorial approach to glycobiology. Significant advances in polymer-supported synthesis as well as chemoenzymic methods have led to the assembly of complex oligosaccharides in a combinatorial fashion even on desired positions of target peptides or proteins. The successful application of oligosaccharide libraries and carbohydrate mimetics to glycobiology such as elucidation of the immunobiological function of glycoconjugates or the discovery of carbohydrate/protein interactions on the basis of microarrays will also be discussed.

Keywords

Combinatorial chemistry; Library; Microarray; Solid-phase synthesis; Automated synthesis; Polymer-supported solution synthesis; Tag-assisted solution synthesis; High-throughput synthesis

In: *Glycoscience*. Fraser-Reid B, Tatsuta K, Thiem J (eds)
Chapter-DOI 10-1007/978-3-540-30429-6_28:

Abbreviations

Bn	benzyl
Bz	benzoyl
DIPEA	diisopropylethylamine
DMF	dimethylformamide
EDTA	ethylenediaminetetraacetic acid
Fmoc	9-fluorenylmethoxycarbonyl
HBTU	2-(1*H*-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate
HOBt	1-hydroxybenzotriazole
MALDI-TOF-MS	matrix-assisted laser desorption ionization mass spectrometry
MPEG	polyethylene glycol monomethyl ether
MS	molecular sieves
NIS	*N*-iodosuccinimide
PEG	polyethylene glycol
TBAHS	tetra-*n*-butylammonium hydrogen sulfate
TBDMS	*tert*-butyldimethylsilyl
TBDPS	*tert*-butyldiphenylsilyl
TFA	trifluoroacetic acid
TMSOTf	trimethylsilyl triflate
Troc	2,2,2-trichloroethoxycarbonyl
UDP	uridine diphosphate

1 Introduction

Combinatorial chemistry (CC) is acknowledged as a technology for synthesizing compound libraries and screening them for desirable properties. It has become established in drug discovery and various drug leads have been found and optimized by CC. In addition, many biological processes have been analyzed by using bioactive compounds found by CC. Carbohydrate-based and carbohydrate-related libraries are a growing trend in CC, because various oligosaccharides and glycoconjugates play important roles in cellular recognition, adhesion, cell differentiation/proliferation, tumor metastasis, inflammation, pathogen/host recognition, and so on. In recent years, both solution and solid phase strategies for oligosaccharide synthesis have been developed to prepare oligosaccharide libraries for biological investigation. The use of carbohydrates as scaffolds in the generation of combinatorial libraries has also been reported. In this chapter, we focus on the following two categories, (1) the oligosaccharide library including libraries of analogues and modified oligosaccharides and (2) the library of carbohydrate mimetics. Their use for glycobiology is also described.

2 The Oligosaccharide Library Including Libraries of Analogues and Modified Oligosaccharides

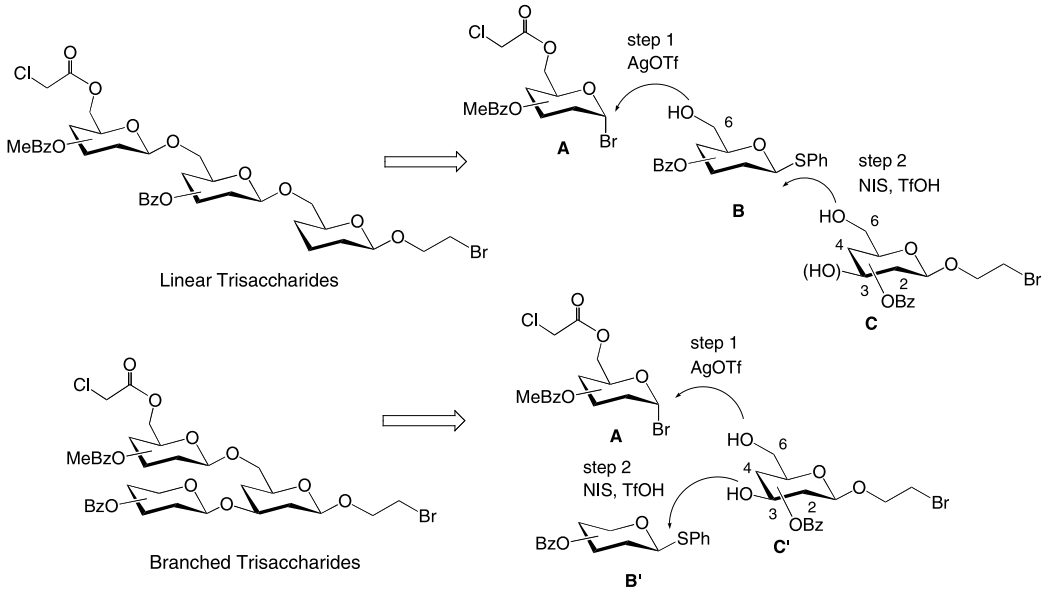
Usually, natural oligosaccharides and glycoconjugates exist in heterogeneous mixtures that are difficult to isolate in pure form. Even if the purification is possible, only small amounts are obtained in general. Not all aspects of glycobiology, now often termed glycomics, are well understood in comparison to genomics and proteomics. One of the reasons is that oligosaccharides and glycoconjugates are much more difficult to synthesize than nucleic acids, peptides, and proteins. The recent development of oligosaccharide and glycoconjugate synthesis, however, has enabled the preparation of complex natural oligosaccharides as well as chemically modified oligosaccharide structures. Synthetic oligosaccharides and glycoconjugates are now used to study their roles in biologically important processes such as inflammation, cell–cell recognition, immunological response, metastasis, and fertilization. In this section, preparation of oligosaccharide and glycoconjugate libraries and their use for glycobiology are described [1,2,3,4]. The combinatorial synthesis of oligosaccharides especially on the solid-phase, is also reviewed in Chap. 5.5, "Polymer-Supported and Tag-Assisted Methods in Oligosaccharide Synthesis."

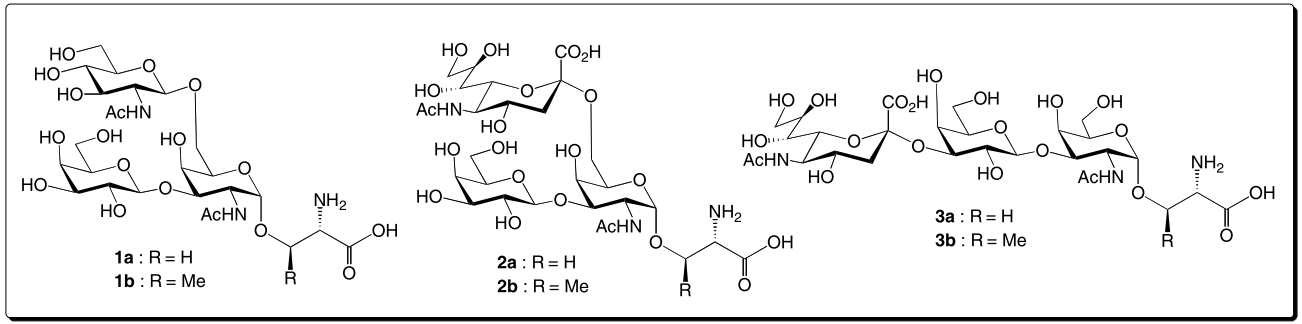
2.1 Synthesis of Oligosaccharides Libraries

A one-pot solution synthesis of oligosaccharides, first realized by Wong and co-workers [5], greatly expanded the possibility of automation for library synthesis [6]. They have prepared more than 100 differently protected *p*-methylphenyl thioglycosides and their reactivity data for glycosylation, i. e., based on armed and disarmed concepts, were programmed on the computer with OptiMer. The program chooses the best reactants for the one-pot synthesis of the target oligosaccharide, namely, providing a glycosylation sequence to the target by reacting from the most to the least reactive thioglycosides. One-pot synthesis has been applied to a large number of oligosaccharides including Globo-H hexasaccharide, known as a cancer antigen [7]. Since protecting group manipulation is not necessary for the present one-pot synthesis, the method has opened up the path to automated library synthesis with a great diversity in structures.

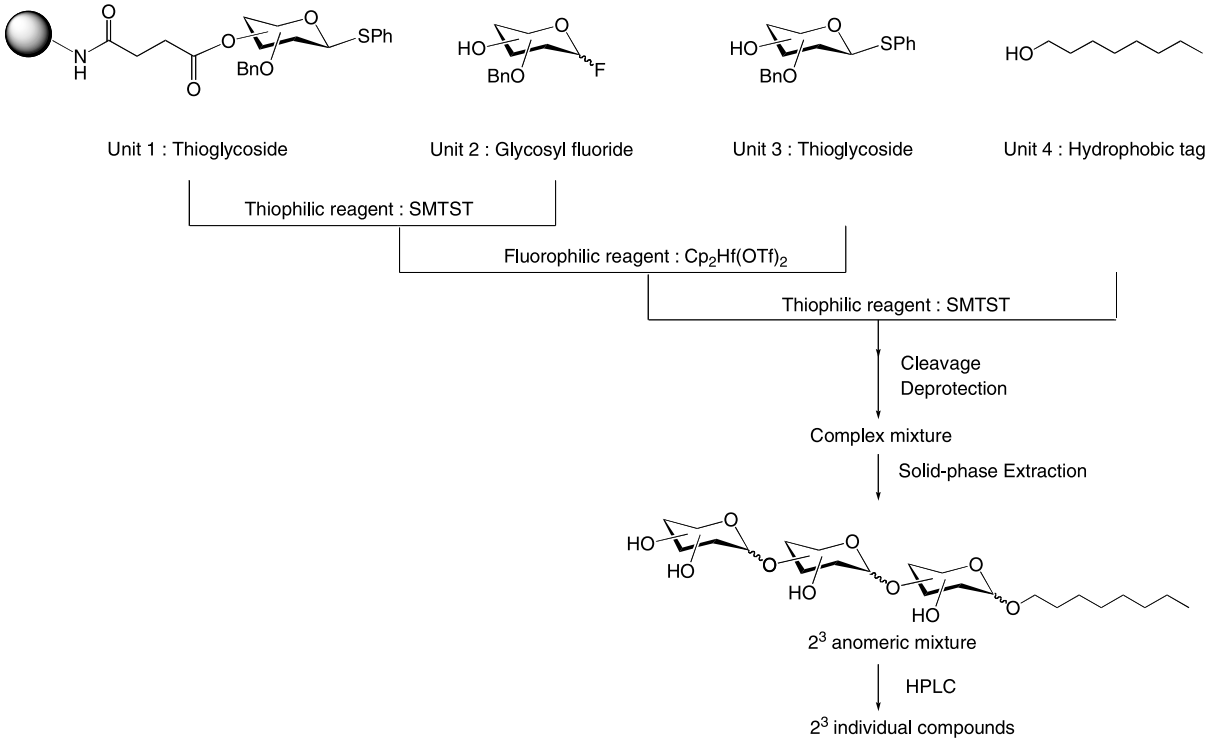
Takahashi and co-workers have also been investigating solution-phase one-pot glycosylation, which led to both linear- and branched oligosaccharides, by using a combination of different leaving groups, e. g., bromide or phenylthio groups to be activated by AgOTf and NIS-TfOH, respectively. They have prepared a 72-trisaccharide library consisting of glucosides, galactosides, and mannosides with the Quest 210^{TM} manual synthesizer [8] as well as the library of H-type 1 and 2 trisaccharides [9] (*Scheme 1*). Their one-pot protocol has also been applied to the synthesis of the glycosyl amino acid derivatives **1–3** [10,11].

Hindsgaul and co-workers have reported the library synthesis of oligosaccharides based on the random glycosylation strategy [12,13,14]. In this method, the lactosyl or glucosamyl acceptors with all free hydroxyls were glycosylated "randomly" with the appropriate donors, such as fucosyl trichloroacetimide, to give glycoside isomers with all possible anomeric configurations and linkage positions. Kanie and co-workers have further improved this idea of orthogonal oligosaccharide synthesis, in preparing a library consisting of a complete set of Fuc-Gal-Glu isomers, the partial structures of blood type antigens, tumor antigens, or *H. Pylori* ligands (*Scheme 2*) [15,16,17]. For the first orthogonal glycosylation, they


Linear Trisaccharides
step 1
AgOTf
step 2
NIS, TfOH
A
B
C
Branched Trisaccharides
A
B'
C'


1a : R = H
1b : R = Me
2a : R = H
2b : R = Me
3a : R = H
3b : R = Me

Scheme 1


Unit 1 : Thioglycoside
Unit 2 : Glycosyl fluoride
Unit 3 : Thioglycoside
Unit 4 : Hydrophobic tag
Thiophilic reagent : SMTST
Fluorophilic reagent : $Cp_2Hf(OTf)_2$
Thiophilic reagent : SMTST
Cleavage
Deprotection
Complex mixture
Solid-phase Extraction
2^3 anomeric mixture
HPLC
2^3 individual compounds

Scheme 2

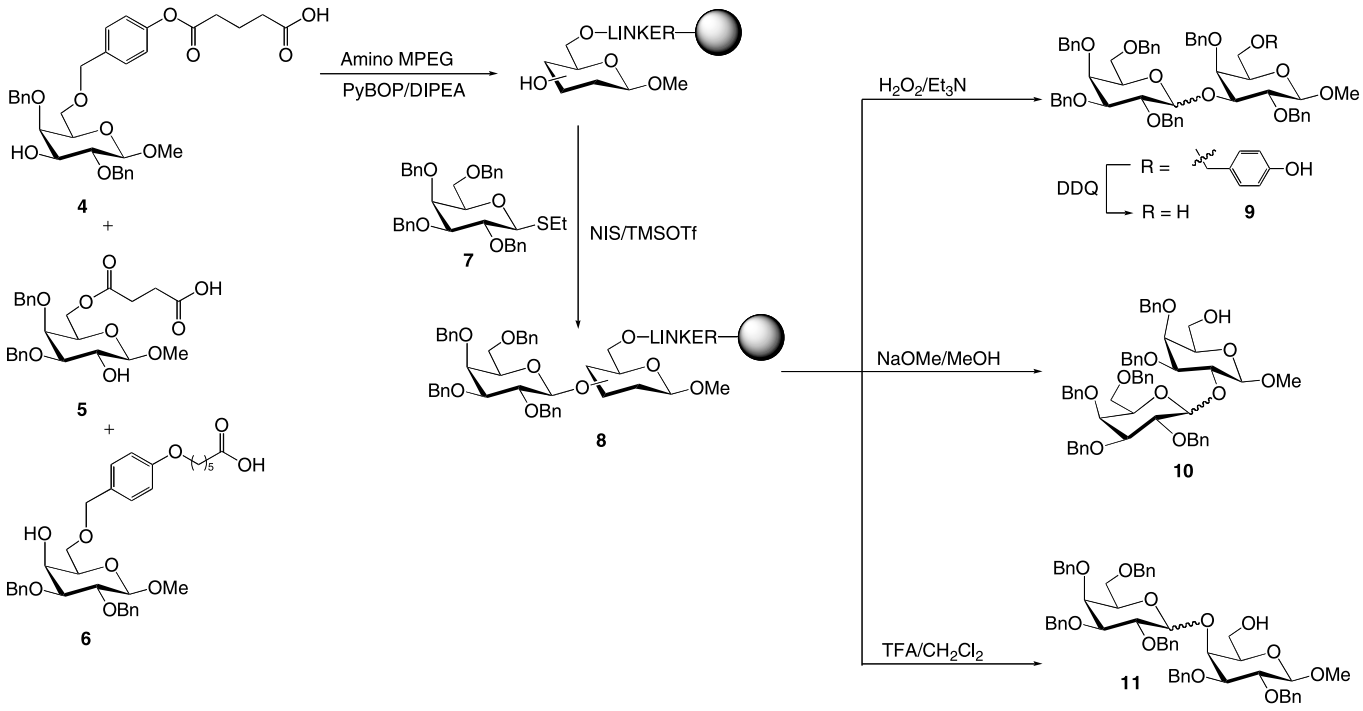
used perbenzylated phenylthiofucopyranoside loaded on the polymer-support TentaGel as the donor, and tri-*O*-benzyl-galactopyranosyl fluorides bearing single hydroxyls either at the 2-, 3-, 4-, or 6-position as the acceptors. The glycosylation was performed by using SMTST (dimethyl(methylthio)sulfonium trifluoromethanesulfonate) as a selective activator of the phenylthio leaving group in the presence of fluoride. The fluoride leaving group at the reducing end of the resulting disaccharides was then activated by $Cp_2Hf(OTf)_2$ to effect the second orthogonal glycosylation with tri-*O*-benzylated phenylthioglucopyranoside. After *n*-octanol as a hydrophobic tag was introduced via glycosylation, the trisaccharides were released from the resin, and the global deprotection of the *O*-benzyl groups successfully provided a Fuc-Gal-Glu library with a complete set of regio- and stereoisomers. These isomers were separated by HPLC with a C18 column with the help of the hydrophobic octanol tag. During the present orthogonal synthesis, they also achieved a stationary solid-phase reaction, without the need for mixing or flowing of the reaction mixtures.

Solid- or soluble polymer-supported synthesis using more than one linker that is cleaved under different conditions has also been applied for library synthesis. Takahashi and co-workers developed a strategy for the combinatorial solid-phase synthesis of the tree-type carbohydrate clusters using three types of orthogonally cleavable linkers (Chap. 5.5, Fig. 1), namely, the photo-cleavable 2-nitrobenzyl-type linker (orthogonal cleavage site 1), base-cleavable ester linker (orthogonal cleavage site 2), and cobalt complexation/acid-cleavable alkyne linker (orthogonal cleavage site 3) [18]. The combined use of three types of spacers, two types of carbohydrates, and two types of orthogonally cleavable linkers provided 12 members of the carbohydrate cluster library (six dimers and six tetramers).

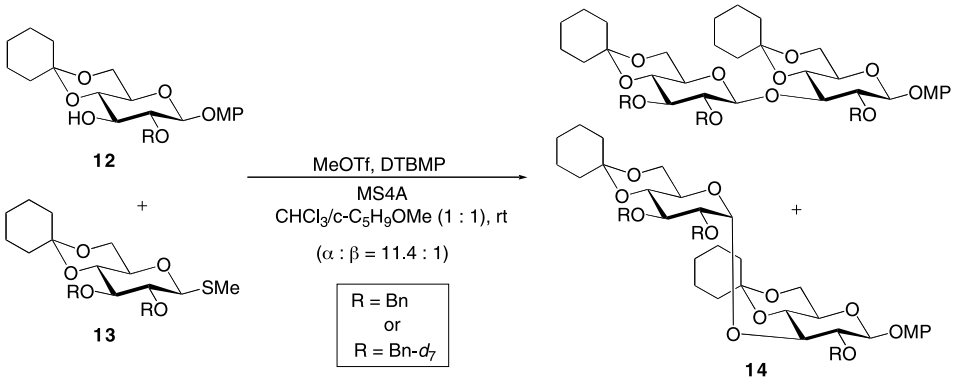
Boons and co-workers introduced the three galactosyl acceptors **4–6** with the free hydroxyls at different positions on soluble polymeric supports (amino-modified MPEG) via three different linkers (*Scheme 3*) [19]. After the reaction with thiogalactosyl donor **7** using NIS/TMSOTf in the mixed solvents of CH_2Cl_2-diethyl ether, the resulting mixtures of three disaccharides on the identical supports were released in turn, in response to the specific reaction conditions. Thus, the treatment of **8** firstly with H_2O_2 and Et_3N liberated **9** from the MPEG-bound oligosaccharides **10** and **11**, which were collected by the precipitation from diethyl ether. Subsequent treatment of the polymer-supports with NaOMe in MeOH gave **10** and polymer-bound **11**, which was finally released by the treatment with TFA in CH_2Cl_2.

2.2 Screening Method for Glycosylation

Ito and co-workers have developed a high-throughput screening method for glycosylation using Bn-d_7 as the protecting group of the hydroxyls (*Scheme 4*) [20]. Since the ^{1}H NMR signals of benzylic methylene protons overlap with those of anomeric protons at 4–5 ppm, the use of deuterated benzyl groups allows the anomeric selectivity to be determined easily by NMR. In addition, the incorporation of Bn-d_7, either on the acceptor or donor, increases the molecular weight by 7 as compared to the normal Bn group, and therefore, in the presence of the non-labeled donor or acceptor, the MS spectrum of the reaction mixture provides a quantitative estimate of the product yields and substrate recovery. They have screened over 240 conditions for α(1–3) linkage formation between the glucose residues **12** and **13**, and optimum conditions for **14** giving the quantitative yield and high selectivity were found.


Amino MPEG
PyBOP/DIPEA
NIS/TMSOTf
H_2O_2/Et_3N
DDQ
R =
R = H
NaOMe/MeOH
TFA/CH_2Cl_2
4
5
6
7
8
9
10
11

Scheme 3

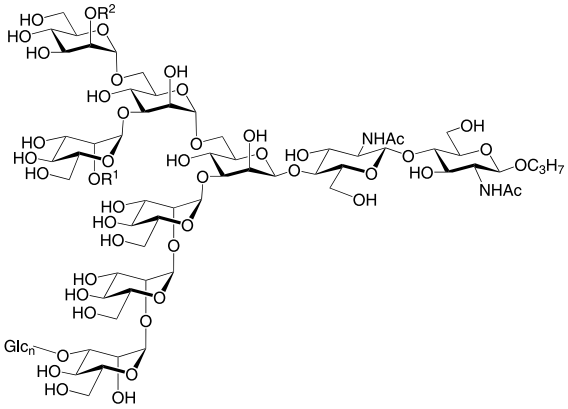


Scheme 4

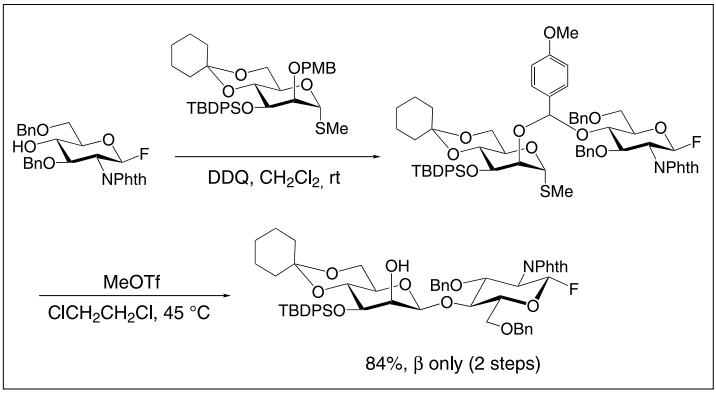
2.3 Combinatorial Synthesis of Complex Oligosaccharides and Glycopeptides

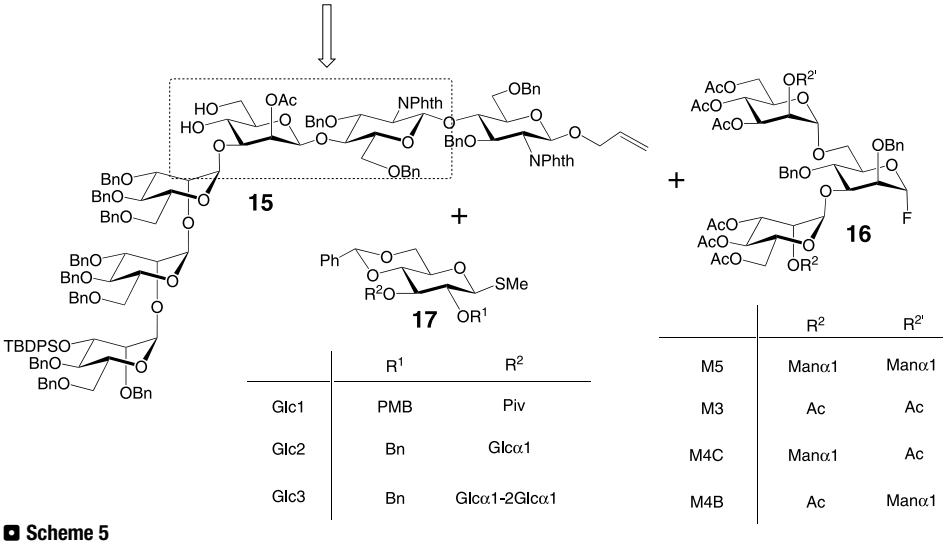
A recent study has revealed the fundamental roles of the endoplasmic reticulum (ER) related high-mannose-type glycans in protein quality control, such as folding, transport, and degradation. Ito and co-workers have achieved the convergent and comprehensive synthesis of these *N*-linked high mannose-type glycans (*Scheme 5*) [21,22,23,24]. Their strategy is based on (i) synthesis of an invariant hexasaccharide fragment **15** via intramolecular aglycon delivery developed in this laboratory, (ii) synthesis of the oligomannose **16**, (iii) preparation of the oligoglucose fragments **17**, and (iv) systematic union of the fragments **15–17** to the desired high mannose-type glycans. During the synthesis of these complex oligosaccharides, they successfully used the pentafluoropropionyl or trifluoroacetyl groups as a temporary hydroxyl protection [25], while a new concept of glycosylation and deprotection of TBDPS under the frozen conditions has also been developed [26]. The synthesized oligosaccharides and their non-covalent conjugates with dihydrofolate reductase (DHFR) will be used to investigate the interaction with the lectins inside and on the ER membrane, e. g., $Man_8GlcNAc_2$ and α-$Glc_1Man_8GlcNAc_2$ for the interaction with MLP (mannosidase-like protein) during the protein degradation process, or $Glc_2Man_9GlcNAc_2$ and $Glc_3Man_8GlcNAc_2$, after the introduction of the dolichyl pyrophosphate moiety, for the interaction with OST (oligosaccharyltransferase) during the translocation process across the ER membrane. Actually, they have succeeded in observing the interaction of α-$Glc_1Man_9GlcNAc_2$ with CRT (calreticulin) by NMR [22], an ER-resident lectin-chaperon that specifically recognizes the monoglucosylated glycans as the initial checking point for proper folding of the proteins.

While the biological roles of high-mannose-type *N*-glycans have been gradually elucidated, partly due to the synthetic achievements such as Ito's work mentioned above, those of a variety of more complex *N*-glycans, such as complex-type *N*-glycans, still remained to be clarified. New methods on the solid-supports, together with the accumulation of the information on the acceptor/donor reactivity in the solution-phase, led to the synthesis of even more complex *N*-glycans on the solid-supports. Schmidt and co-workers have achieved solid phase synthesis



	R^1	R^2	Glc_n
M9	Man	Manα1	H
M7	H	H	H
M8C	Man	H	H
M8B	H	Manα1	H
G3M9	Man	Manα1	Glc_3
G2M9	Man	Manα1	Glc_2
G1M9	Man	Manα1	Glc_1
G1M7	H	H	Glc_1
G1M8C	Man	H	Glc_1
G1M8B	H	Manα1	Glc_1





	R^1	R^2
Glc1	PMB	Piv
Glc2	Bn	Glcα1
Glc3	Bn	Glcα1-2Glcα1

	R^2	$R^{2'}$
M5	Manα1	Manα1
M3	Ac	Ac
M4C	Manα1	Ac
M4B	Ac	Manα1

Scheme 5

of a small library of *N*-glycans (described in detail in Chap. 5.5, Scheme 10) [27]. They used the hydroxymethylbenzyl benzoate linker attached to the Merrifield resin. Glycosylation on the resin was successfully performed using the monosaccharide units of the trichloroacetimidate, which allow chain extension, branching, and chain termination for the preparation of a variety of *N*-glycans, while the tedious β-mannosyl linkage was constructed in advance in solution. Their solid-phase synthesis involves (i) glycosylation under the TMSOTf catalyst, (ii) selective removal of the temporary protecting groups for the sugar extension on the solid-phase, namely, Fmoc by Et_3N and the phenoxyacetyl group by 0.5 equivalents of NaOMe in $MeOH/CH_2Cl_2$, and (iii) product cleavage from the resin by 4.0 equivalents of NaOMe in $MeOH/CH_2Cl_2$ and then *O*-acetylation for convenient product isolation. These precise designs of linker, protecting group manipulations, and stereoselective glycosylation on the resin enabled the successful preparation of a small library of 17 *N*-glycan structures (*Fig. 1*).

Seeberger and co-workers have also achieved a solution-phase synthesis of several linear and branched high-mannose oligosaccharides **18–23** in order to investigate the detailed interaction of the HIV viral surface envelope glycoprotein, gp120 with the anti-HIV protein, cyanovirin-*N* (*Fig. 2*) [28]. On the basis of the microarray analysis, these oligosaccharides were also found to show high specificity to the microvirin (MVN), which shares 33% identity with the cyanovirin-*N*; MVN might be involved in cell–cell recognition and cell–cell attachment of *Microcystis* [29]. The automated solid-phase synthesis of high-mannose oligosaccharides has also been realized by the same group, which is described in detail in Chap. 5.5.

The malaria parasite *Plasmodium falciparum* accounts for about 100 million clinical cases and the deaths of more than two million people annually due to inflammation caused by the malaria toxin. The inflammatory cascade of malaria is triggered by glycosylphosphatidylinositol (GPI) on the cell surface and the GPI is responsible for much of malaria's morbidity and mortality. Seeberger and co-workers have synthesized the malaria GPI glycan of the sequence NH_2–CH_2–CH_2–O–PO_3–(Manα1,2)6Manα(1,2)Manα(1,6)Manα1,4)Glc$NH_2$$\alpha$1,6)myo-inositol-1,2-cyclic-phosphate and its related structures (*Fig. 3*) [30,31]. They employed a flexible synthetic strategy that relies on a late-stage coupling between oligomannosides of various lengths and the pseudo-disaccharide glycosyl acceptor to readily access various malarial GPI structures. Mice that were treated with chemically synthesized GPI attached to the protein were substantially protected from death by malaria. Recipients were substantially protected against malarial acidosis, pulmonary edema, cerebral syndrome, and fatality. Anti-GPI antibodies neutralized proinflammatory activity by *P. falciparum* in vitro. These glycans are currently under development as anti-toxin malaria vaccines.

They succeeded in synthesizing these hexasaccharide glycans by using a combination of automated solid-phase methods and solution-phase fragment coupling in a matter of days, compared with several weeks for a comparable solution-phase synthesis [32]. They also synthesized fully lipidated GPI anchors of *P. falciparum* by employing a highly convergent strategy [33]. This strategy utilized three orthogonal protecting groups, which can be chemoselectively deprotected and functionalized in the late stage of the synthesis. The synthetic GPI is currently used as a molecular probe for the study of malarial pathogenesis and aspects of fundamental immunology.

Proteoglycans are major components of the extracellular matrix in animal cells. They are composed of core proteins and glycosaminoglycan polysaccharides. Heparin and heparan sulfate are the most complex glycosaminoglycans, a family of molecules that also includes chon-

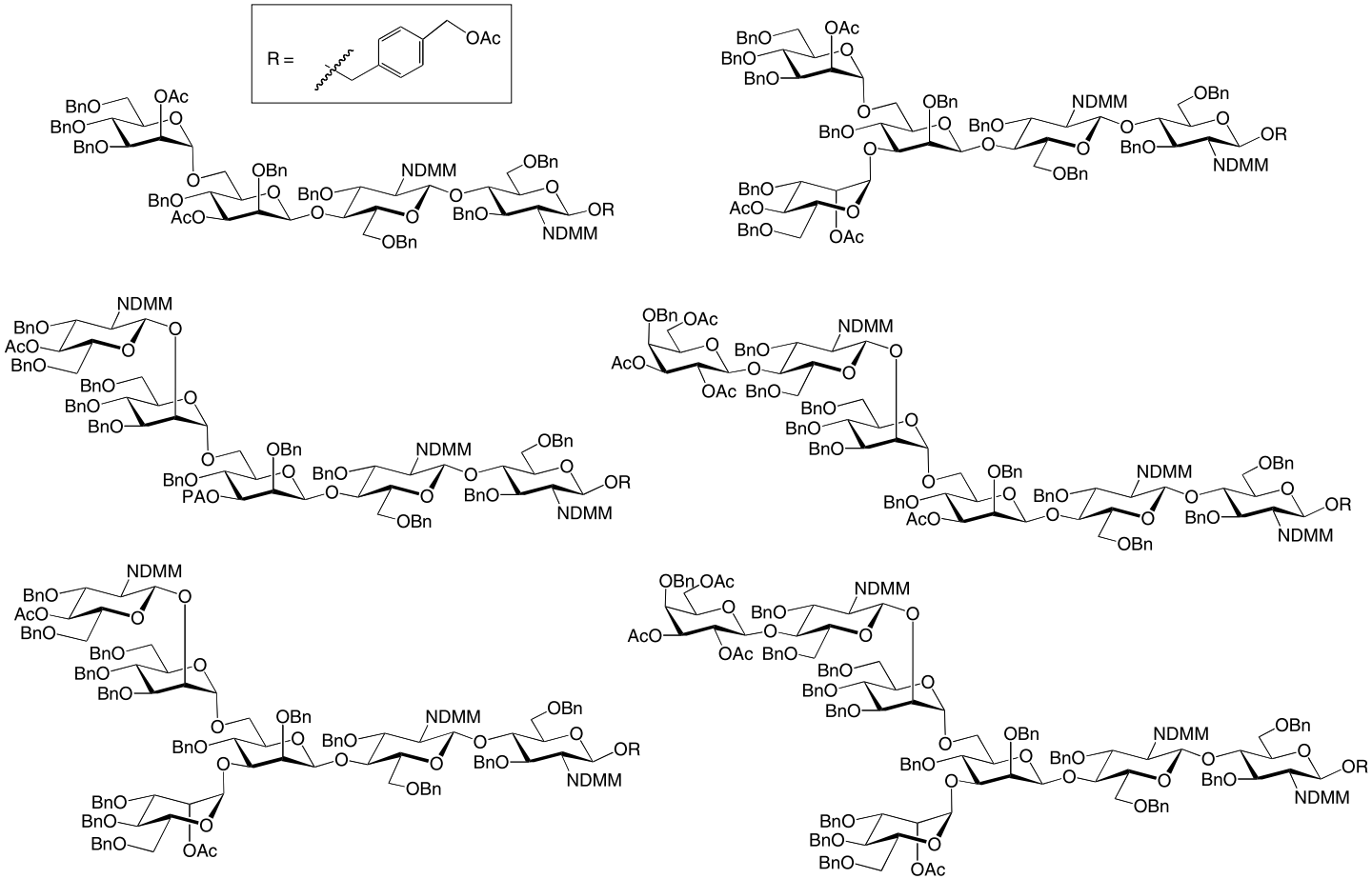


Figure 1
***N*-glycan library by Schmidt and co-workers**

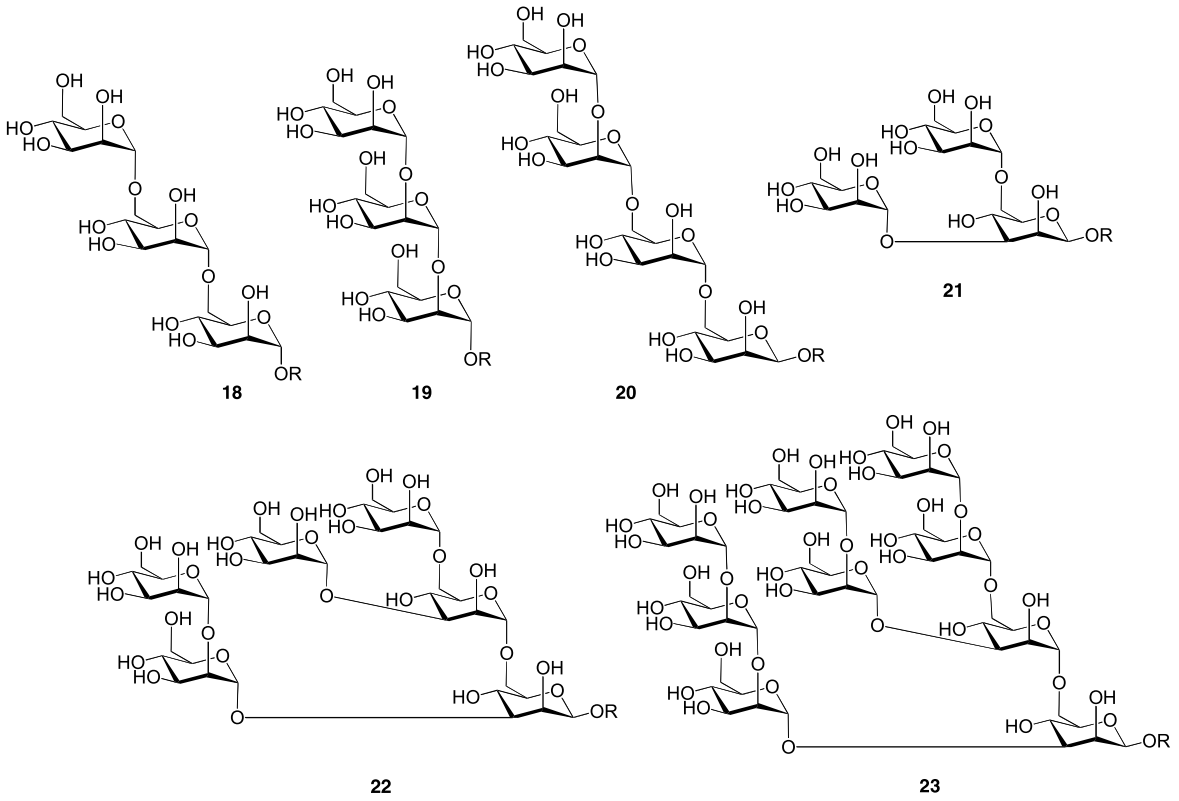


Figure 2
Linear and branched high-mannose oligosaccharides synthesized by Seeberger and co-workers

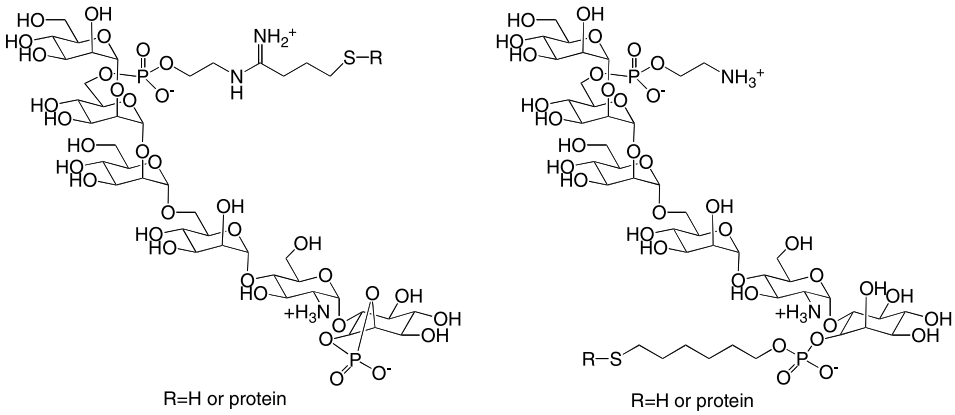


Figure 3
Structures of malaria vaccine candidates

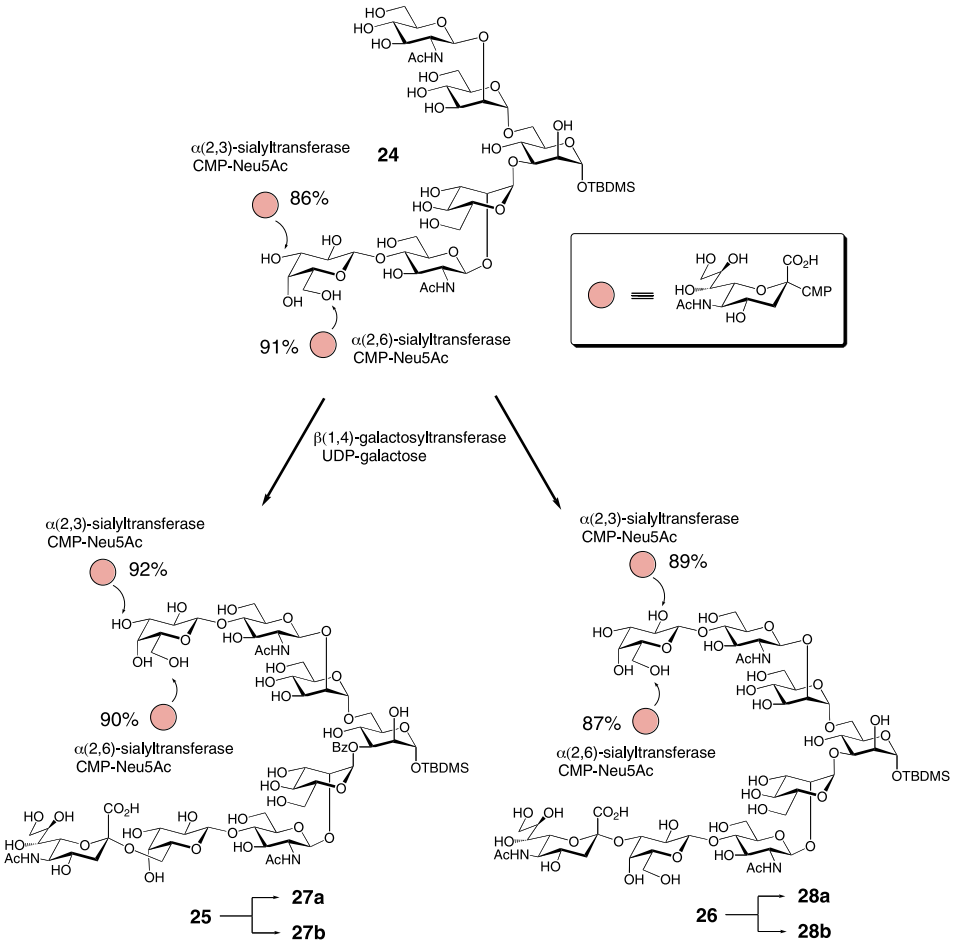
droitin sulfate, keratin sulfate, and dermatan sulfate. Glycosaminoglycan has important biological functions by binding to different growth factors, enzymes, morphogens, cell adhesion molecules, and cytokines. Heparin, a heterogeneous sulfated glycosaminoglycan found only in the mast cell, has been widely used as an injectable anticoagulant. Heparin consists of a repeating disaccharide unit composed of a uronic acid 1,4-linked to a D-glucosamine. Heparan sulfate is widely distributed in different cell types and tissues. Heparan sulfate contains less *O*-sulfates but is more heterogeneous than heparin. The major uronic acid in heparin is L-iduronic acid whereas that in heparan sulfate is the D-glucuronic acid unit. Since heparin and heparan sulfate are very heterogeneous molecules and have many biological functions that depend on specific sequences with particular modifications and sulfations, defined heparin and heparan oligosaccharides have been required to investigate their precise biological functions. The purification of oligosaccharides after enzymatic digestion or enzymatic reconstruction of glycosaminoglycan chains has provided libraries of oligomers that have been used for biological investigations [34,35,36,37,38,39,40,41,42,43,44,45].

Chemical synthesis of heparin fragments and analogues has been used for biological studies and to define structure-activity relationships [46]. For example, the heparin-antithrombin III interaction [47,48,49,50,51,52,53], which is responsible for the anticoagulant activity of heparin, has been studied in detail by using synthetic fragments. The interaction of heparin with FGF-1, FGF-2, platelets [54,55,56,57,58,59,60,61,62,63,64], and the Herpes Simplex virus, has also been studied by using synthetic specimens.

A modular assembly approach has been developed for the synthesis of a broad range of heparin/heparan sulfate oligosaccharide fragments [65,66]. Polymer-supported syntheses of heparin oligosaccharides have also been developed toward the preparation of libraries as described in Chap. 5.5 [67,68,69].

Since the chemical synthesis of oligosaccharides requires many synthetic steps including protection and deprotection procedures, the enzymatic approach has attracted much attention for the rapid synthesis of oligosaccharides. In addition, the perfect regio- and stereoselectivities of enzymatic methods with glycosyltransferases are quite attractive. Several trans-

ferases such as β(1,4)-galactosyltransferase, α(1,3)-fucosyltransferase, and α-sialyltransferase have been used for the enzymatic synthesis. Ito and co-workers have successfully synthesized α(2–3)- and α(2–6)-sialylated biantennary glycans **24–28**, the branched positions of typical complex-type *N*-glycans, based on the polymer-supported strategy and enzymic glycosylation (*Scheme 6*) [70]. The common hexasaccharide **24** was prepared by the combination of (i) MPEG (monomethyl polyethylene glycol; *M*w 750)-supported synthesis and (ii) capture-release protocol using the Boc-cystein loaded on Merrifield resin, developed in their laboratory (see Chap. 5.5, Scheme 13). Hexasaccharide **24** was diverged to monosialylated **25** and **26** and then to disialylated **27** and **28** by the enzymic glycosylation using the commercially available α(2,3)- or α(2,6)-sialyltransferase and β(1,4)-galactosyltransferase together with the appropriate sugar nucleotides.



Scheme 6

Nishimura and co-workers took good advantage of the combination of chemical and enzymatic synthesis, especially by using α(2,3)- or α(2,6)-sialyltransferases from porcine and rat livers, and a variety of the biologically relevant glycopeptides, glycosphingolipids, and glycomimetics has successfully been prepared [71,72]. Their library-directed synthesis based on this approach involves, e. g., (a) sialyl Lewis X-based glycoclusters on the β-cyclodextrin which strongly inhibit protein–carbohydrate interaction, such as that of E-selectin and SLeXn-BSA [73,74,75], (b) mucin-like glycoprotein mimics of the antifreeze proteins [76,77,78,79], (c) triantennary-type sialolacto- and sialolactosamine polymers which exhibit much higher anti-influenza virus activity than monovalent-type glycopolymers [80], (d) conical-type sialolacto-peptides as the inhibitors of influenza hemagglutinin [81], and (e) glycoconjugates of insulins with dendric sialolactosamines, which exhibit a prolonged glucose-lowering effect in hyperglycemic mice [82]. Recently, they developed the catch-and-release strategy between solid-phase and water-soluble polymer supports, the so-called "polymer blotting method", which allows for the rapid and efficient synthesis of glycopeptides [83,84]. The method involves (i) conventional solid-phase synthesis of glycopeptides that contain mono-, di-, and tri-saccharides on Thr or Ser residues and Blase (glutamic acid specific protease)-sensitive amino acid sequence, (ii) removal of the protecting groups and release from the resin, (iii) attachment of the released glycopeptides to the water-soluble polymers via an oxime anchor, (iv) sugar elongations using glycosyltransferases, and (v) cleavage of the desired glycopeptides from the soluble supports by Blase-mediated hydrolysis. They applied three glycosyltransferases, namely, β(1,4)-galactosyltransferase, α(2,3)-(*O*)-sialyltransferase, and α(2,3)-(*N*)-sialyltransferase, and successfully prepared six kinds of parallel and 36 kinds of combinatorial libraries of MUC1 mucin glycopeptides (*Fig. 4*) [83]. The chemoenzymatic synthesis of glycopeptides and proteins, such as HIV-1 V3 glycopeptides and CD52 glycoproteins, have also been reported by Wang and co-workers [85,86].

Kajihara and co-workers have reported the chemoenzymic library synthesis of *N*-linked glycans, by adapting the branch specific exo-glycosidases digestion of the monosialyloligosaccharides obtained from the biantennary complex-type sialylundecasaccharide **29** (from egg yolk) by acid hydrolysis of one molecule of Neu5Ac (*Fig. 5*) [87,88,89,90,91,92]. The isolation of two mono-Neu5Ac derivatives by ODS HPLC column was efficiently achieved by applying the Fmoc and benzyl protecting groups, and the resulting pure mono-Neu5Ac oligosaccharides were gradually disintegrated into 24 oligosaccharide derivatives by α-D-neuraminidase, β-D-galactosidase, *N*-acetyl-β-D-glucosaminidase, and α-D-mannosidase, respectively. Oligosaccharides thus obtained were successfully conjugated to the aspartic acid residue, and *N*-linked glycoproteins, having sialyl-undeca- and asialo-nonasaccharides, were synthesized on the solid-support, HMPA-PEGA resin [87].

A totally synthetic approach to glycopeptides as well as glycoproteins, on the basis of glycal-based oligosaccharide synthesis without the use of transferases, has been thoroughly investigated by Danishefsky and co-workers. Their recent successful works feature, e. g., (a) preparation of *N*-linked glycopeptides presenting the H-type 2 human blood group determinants [93,94], (b) synthesis of normal and transformed PSA (prostate specific antigen) glycopeptides in an attempt to produce the PSA antibody for cancer diagnosis purposes [95,96], (c) synthesis of high-mannose- and hybrid-type *N*-glycopeptides as well as their dimerized peptides of the envelope protein gp120 in pursuit of carbohydrate-based HIV vaccines [97,98,99], and (d) coupling of the glycopeptides toward total synthesis of the

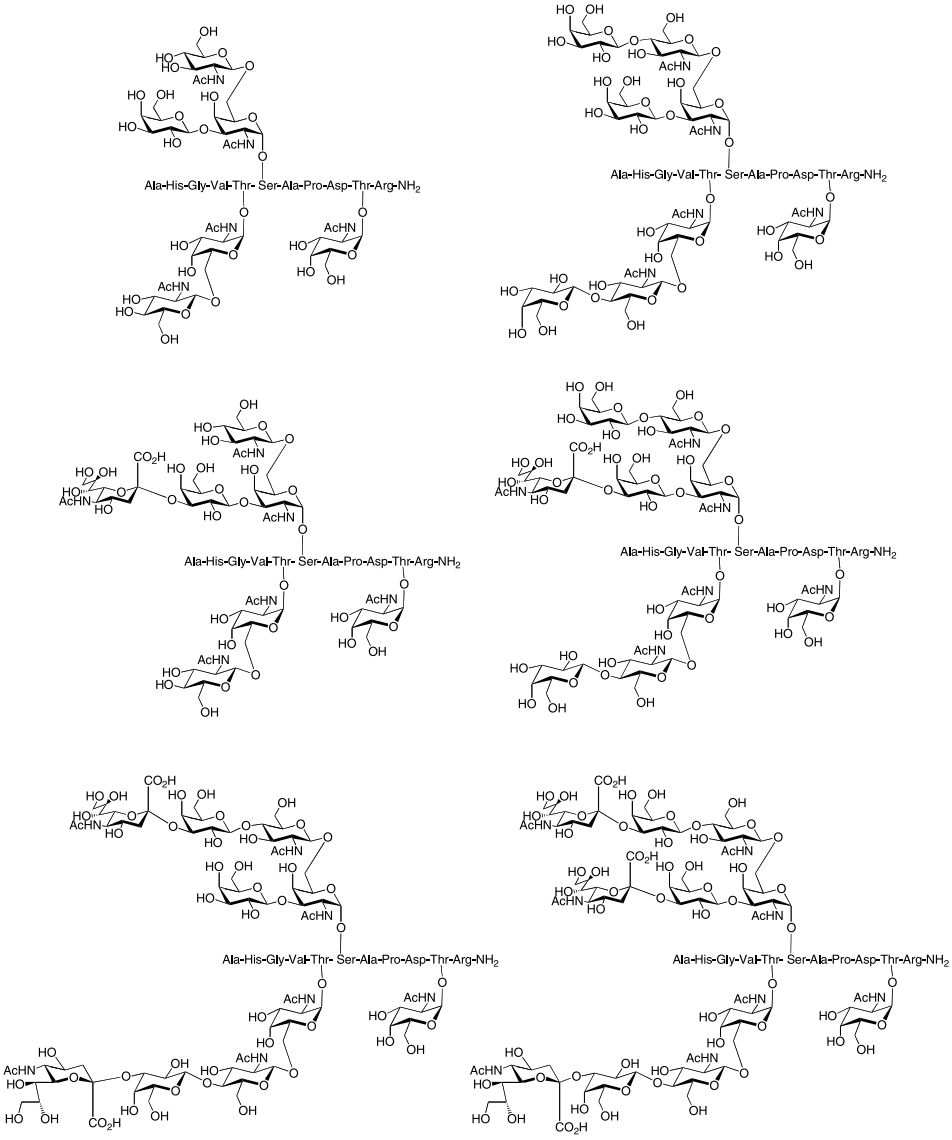


Figure 4
Chemoenzymic synthesis of glycopeptide library by Nishimura and co-workers

glycoproteins, such as EPO (erythropoietin), based on the new chemical ligation methods developed in the same laboratory [100,101,102,103]. Furthermore, on the preparation of glycopeptide-based cancer vaccines, they have displayed different cancer-related carbohydrate antigens on a single peptide backbone, and examined to induce immunological responses (*Fig. 6*) [104,105,106,107,108,109,110,111,112,113,114,115]. Among several vaccines

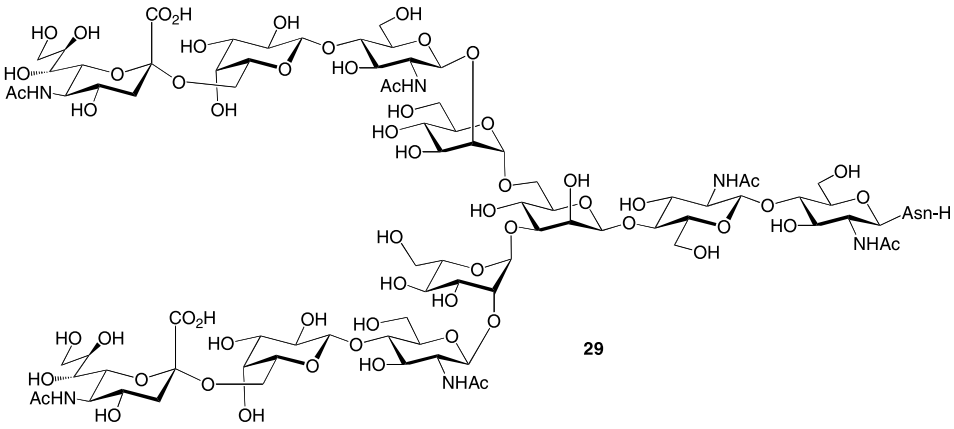


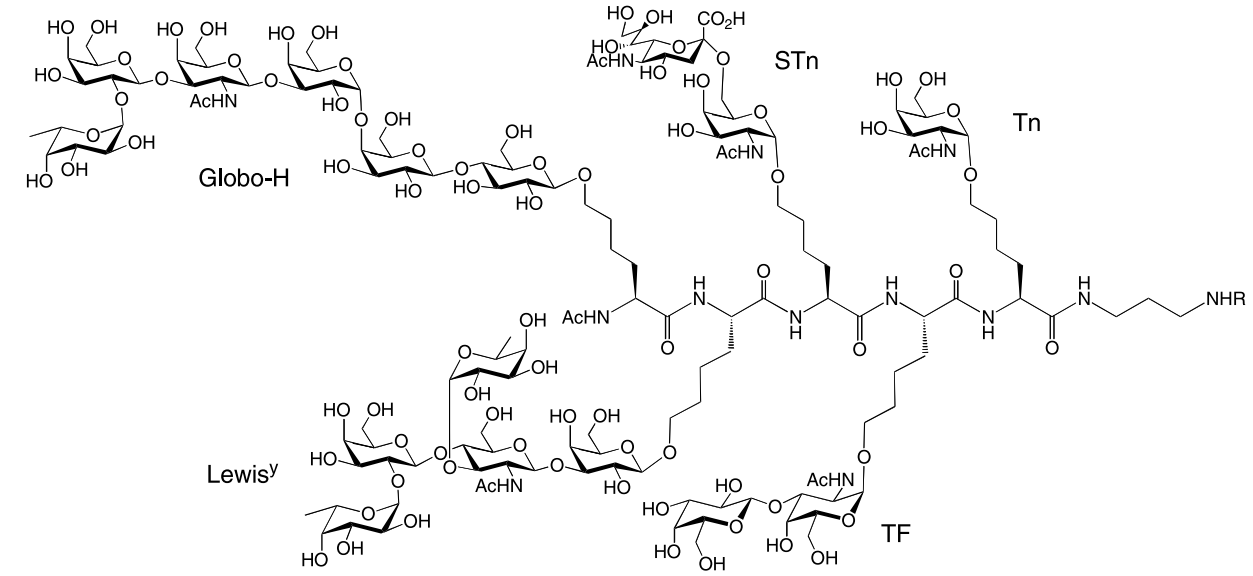
Figure 5
Biantennary complex-type sialylundecasaccharide from egg yolk

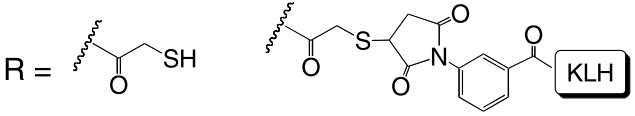
prepared, the glycopeptide **30** was conjugated to the carrier molecules, i. e., KLH (keyhole limpet hemocyanin) **31** and Pam_3Cys **32**. The cumulative data suggest that the immunological properties of the individual antigen in **30**, namely, Globo-H, Le^y, STn, TF, and Tn antigens, each of which is overexpressed on prostate and breast cancer cell surfaces, are preserved [104]. Thus, the highly elaborate vaccines constitute a promising category for clinical applications.

Nakahara and co-workers have investigated the solid-phase synthesis of the prototype glycopeptidyl fragments of serglycine, a proteoglycan with the repeating peptide sequence of L-seryl-L-glycine, by the condensation of the glycosyl amino acid unit on Sieber amide resin (*Scheme 7*) [116]. They have developed the allyl ester-type linker, which can be cleaved by a $Pd(PPh_3)_4$ catalyst, leading to the efficient preparation of various lengths of [*O*-(2,3,4-tri-*O*-acethyl-D-xylosyl)-L-seryl-L-glycine]$_n$ (n = 1–8). The same linker on the Sieber resin was also applied to the solid-phase synthesis of mucin-type glycopeptide [117] and human leukosialin segment [118]. The Ig domain of emmprin has also been synthesized by the same group, through the successful preparation of the thioester fragment with *N*-linked core pentasaccharide by a modified Fmoc solid-phase method [119].

Meldal and co-workers have investigated direct glycosylation of a variety of glycosyl trichloroacetimidates with the amino acid side chain of serine, threonine, and tyrosine in peptides on solid-supports, pursuing the efficient synthesis of a glycopeptide library (*Fig. 7*) [120]. The peptides were linked to either $POEPOP_{1500}$ (polyoxyethylene-polyoxypropylene) or $SPOCC_{1500}$ (super-permeable organic combinatorial chemistry) resins via a photolabile linker, which facilitates glycosylation analysis by both MALDI-TOF mass spectrometry and nanoprobe MAS (magnetic angle spinning) NMR spectroscopy. On the basis of these analytical systems, nearly quantitative yields were concluded for the reactions of acceptor peptides on the solid-phase with glucose, galactose, mannose, and fucose trichloroacetimidates **33–37**.

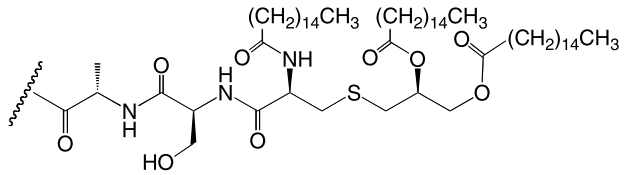
Schmidt and co-workers have investigated the synthesis of a series of *S*-linked glycopeptides, which are expected to be more stable to glycosidases than the naturally occurring *N*- and





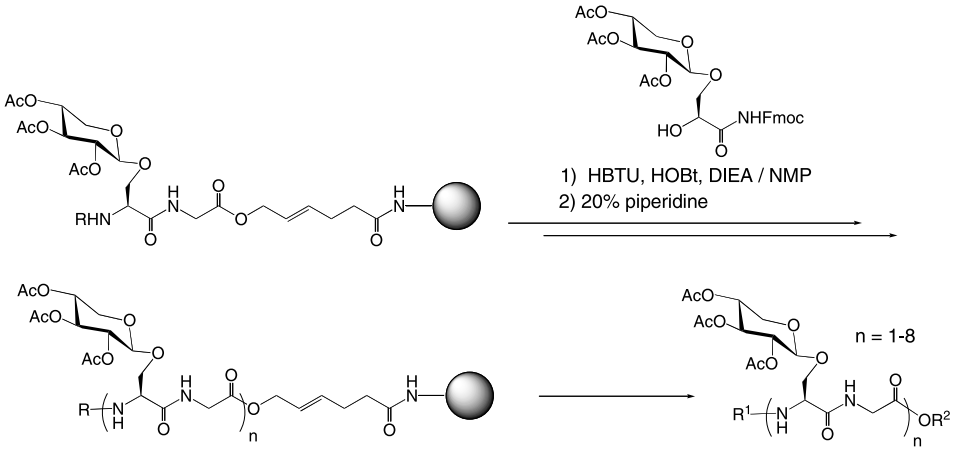
30

KLH conjugate: **31**



Pam_3Cys conjugate: **32**

Figure 6
Highly elaborate cancer vaccines by Danishefsky and co-workers



Scheme 7

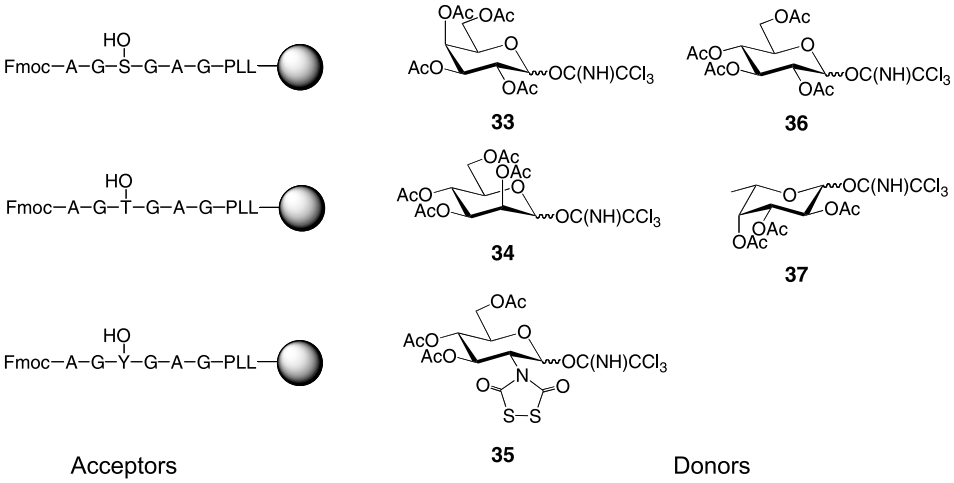
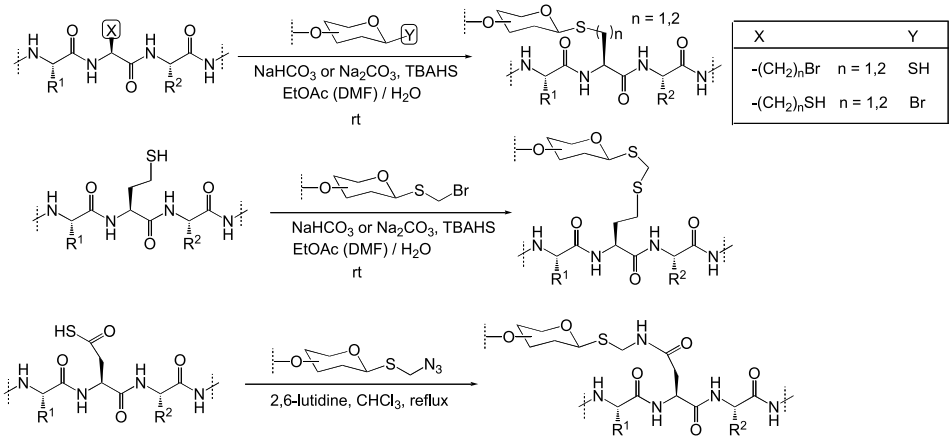


Figure 7
Direct glycosylation strategy for a glycopeptide library by Meldal and co-workers

O-linkages with the peptide backbone (◉ *Scheme 8*) [121,122,123,124]. They have developed three conjugation methods of sugar moieties to the peptides; (i) direct *S*-glycosylation of cystein or homocystein residues with glycosyl halides or glycosylthiomethyl bromides, (ii) *S*-alkylation of the 1-thiosugars with β-bromoalanine or γ-bromohomoalanine residues, and (iii) reaction of glycosylthiomethyl azides with asparatate or glutamate thio acids. These reactions can be performed in water and provide the *S*-linked glycopeptides in excellent yields. Bertozzi and co-workers prepared *N*-linked glycopeptide analogs that replace the glycosidic linkages extending from the core pentasaccharide with thiomethylamides, being amenable to rapid assembly of the complex-type *N*-linked glycopeptide mimetics (◉ *Fig. 8*) [125,126].



Scheme 8

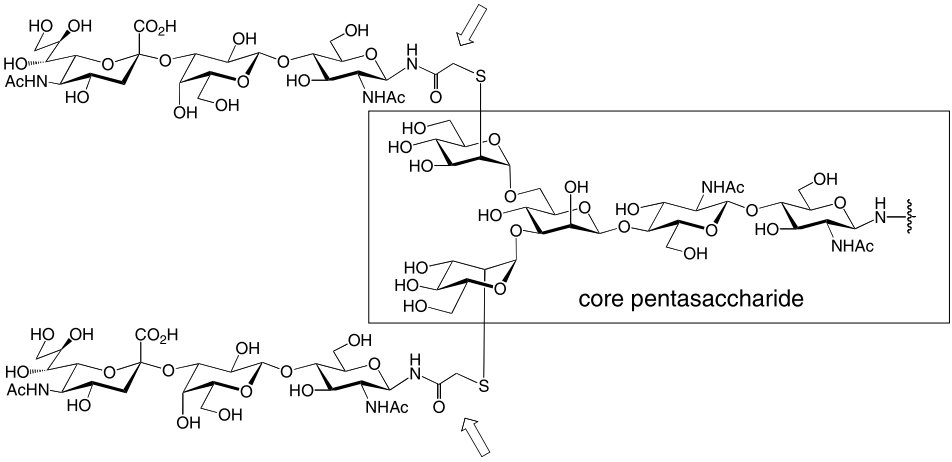
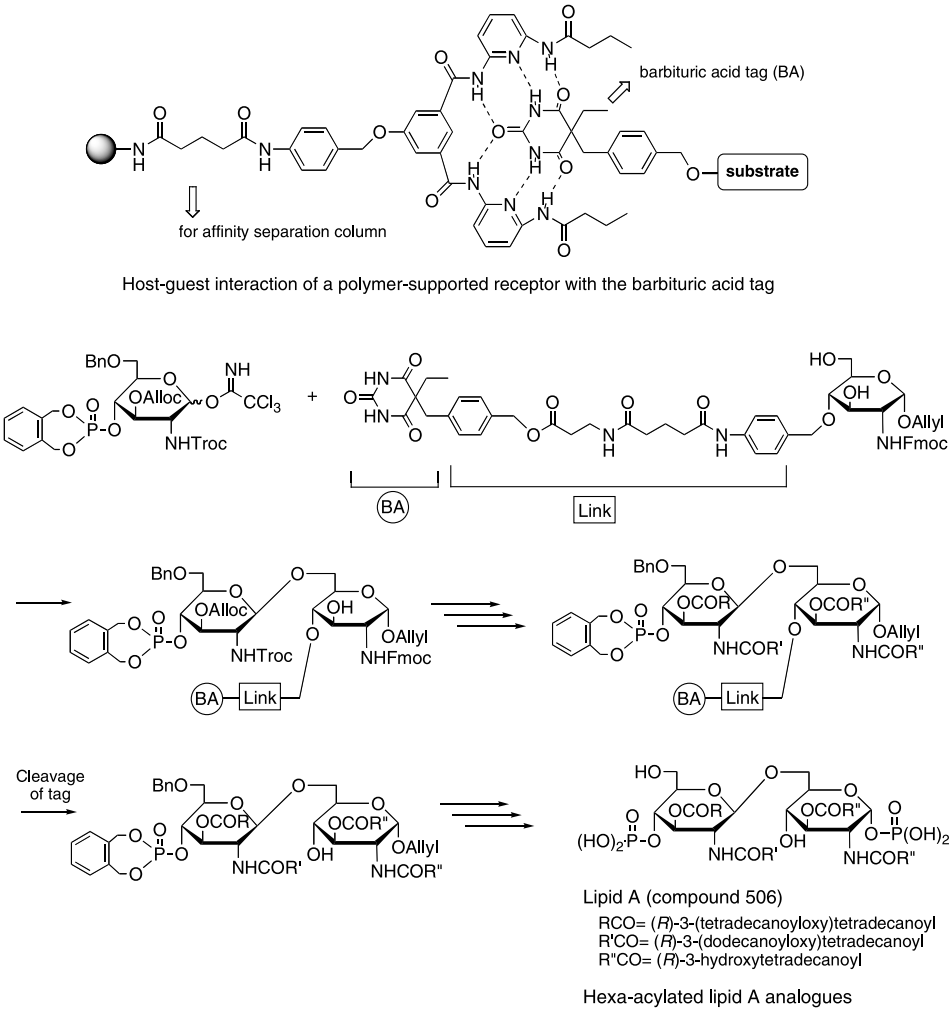


Figure 8
Complex-type *N*-linked glycopeptide mimetics by Bertozzi and co-workers

2.4 Combinatorial Synthesis of Glycoconjugates for Elucidation of Immunobiological Function

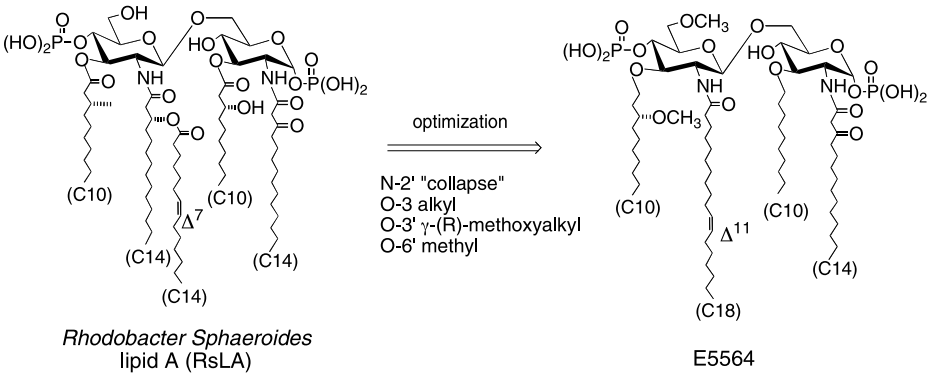
The concept of CC has also been applied to the investigation of the biological function of glycoconjugates including glycolipids and glycopeptides. An example is the synthesis of bacterial lipopolysaccharide (LPS) and its partial structures including terminal glycolipid moiety lipid A, which have potent immunostimulatory activity to stimulate the innate immune system. Fukase and Kusumoto prepared the library of lipid A analogues to elucidate the structural requirements for the immunostimulatory and antagonistic activity with the diversity on



Scheme 9

the fatty acids and the acidic groups of lipid A [127,128]. They prepared lipid A analogues which have hexaacyl groups with different acylating patterns, utilizing a barbituric acid tag for affinity separation, based on their "synthesis based on affinity separation (SAS)" method (*Scheme 9*) [129,130]. The principle of the SAS method and other applications using ether-type tags are explained in Chap. 5.5. With this method a focused library of lipid A analogues was efficiently obtained, which showed how structural differences at the acyl part of lipid A affect biological activity.

A library of acidic amino-acid-substituted monosaccharide lipid A analogues was also prepared for elucidation of the structural basis to biological function [131]. The library of lipid A analogues was also used for the crystallization of the coreceptor MD-2, which constitutes the receptor complex with Toll-like receptor 4 (TLR4). The antagonistic tetraacylated lipid A (the



Scheme 10

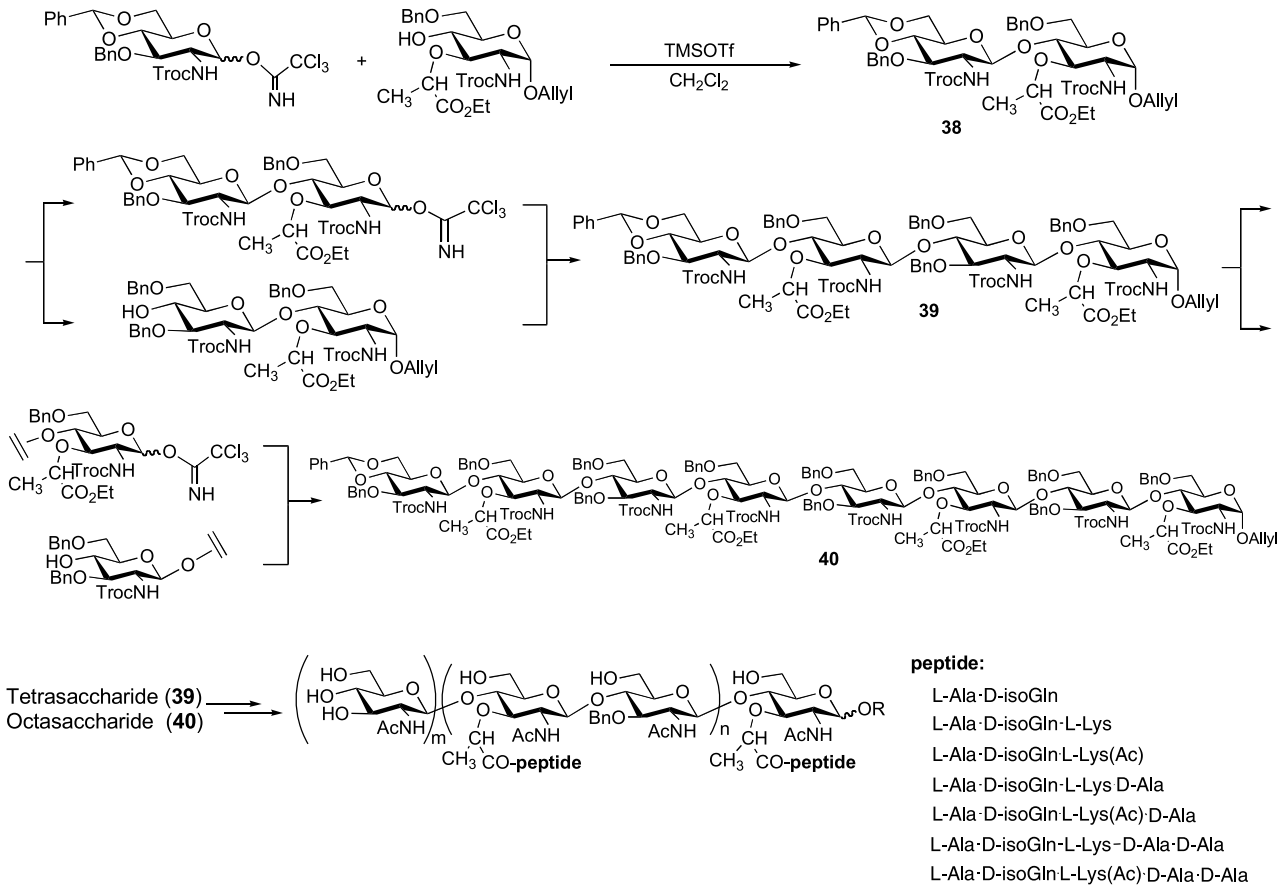
biological precursor lipid IVa, compound 406) is bound to MD-2 and effectively led the protein to be crystallized to observe the first crystal structure by X-ray analysis [132].

Since the potent immunostimulation of LPS often causes sepsis when an overwhelming infection is present, the antagonist of LPS is a candidate for a therapeutic agent; an effective drug has so far not been available. Eisai's group have developed E5531 [133] and E5564 from a library of a series of antagonists based on natural antagonistic lipid As from *Rhodobacter capsulatus* (RcLA) and *Rhodobacter sphaeroides* (RsLA), to improve the potency of the biological activity and the stability of the compounds [134]. As for E5564, the structural modification from the ester to ether linkage at the 3- and 3'-positions, methylation at the 6-position and also "collapsing" the unsaturated acyloxy group at the 2'-position made the molecule more stable and also more synthetically accessible (*Scheme 10*). The resultant analogue, E5564 demonstrated enhanced antagonistic activity over E5531, and did not have detectable agonistic activities. E5564 is now at the stage of phase III clinical trials.

One of the glycolipids that has drawn much interest is α-D-galactopyranosylceramide (α-GalCer), which has potent immunostimulatory and anti-tumor activities, found by Kirin Brew's group. They developed analogues with enhanced activity [135,136,137]. α-GalCer acts as a specific ligand presented by CD1d to a receptor of natural killer T-cells (NKT cells) to lead the immunostimulation, and Wong and co-workers have synthesized analogues for further investigations on biological function [138,139,140,141]. Franck and co-workers have also sought improved compounds with the modification of the structure to that of *C*-glycoside, and succeeded in obtaining analogues with much enhanced activity [142,143,144].

Rademann and co-workers have applied the hydrophobically assisted switching phase (HASP) synthesis for a glycolipid library [145]. The HASP method is further explained in Chap. 5.5. They used the method for the synthesis of immunostimulative rhamnolipids to observe proinflammatory activities in human mononuclear cells.

Fukase and co-workers have synthesized peptidoglycan (PG) partial structures in order to investigate their mechanism of action in immunostimulation [146,147]. Since peptidoglycan and its fragments derived from natural sources may be contaminated with other immunostimulating components, the precise mechanism of action of PG was evaluated by using structurally defined synthetic specimens. By using an efficient synthetic strategy, mono-, di-, tetra-, and octasaccharide fragments of PG were synthesized in good yields (*Scheme 11*).


TMSOTf
CH_2Cl_2
38
39
40
Tetrasaccharide (39)
Octasaccharide (40)
CO-peptide
peptide:
L-Ala·D-isoGln
L-Ala·D-isoGln·L-Lys
L-Ala·D-isoGln·L-Lys(Ac)
L-Ala·D-isoGln·L-Lys D-Ala
L-Ala·D-isoGln·L-Lys(Ac)·D-Ala
L-Ala·D-isoGln-L-Lys-D-Ala·D-Ala
L-Ala·D-isoGln·L-Lys(Ac)·D-Ala·D-Ala

Scheme 11

They revealed that not only glycan but also peptide were essential for expression of the immunostimulating activity [148]. Toll-like receptor 2 (TLR2), which is one of the candidates for PG receptor, was not stimulated by the series of synthetic PG partial structures. On the other hand, intracellular receptor Nod2 recognized the partial structures containing the muramyl (MurNAc) dipeptide (MDP) moiety. These results indicated that Nod2 is the intracellular receptor of muropeptides derived from peptidoglycan and the minimal ligand of Nod2 is MDP [149].

Synthetic PG partial structures have also been used for elucidation of the structure recognized by peptidoglycan recognition proteins (PGRPs). PGRPs constitute a large group of proteins found in insects and mammals that bind to bacterial PG. Dziarski et al. reported that human PGRP-L is *N*-acetylmuramoyl-L-alanine amidase, which hydrolyzes the amide bond between MurNAc and L-Ala of bacterial PG. The minimum PG fragment hydrolyzed by PGRP-L is MurNAc-tripeptide.

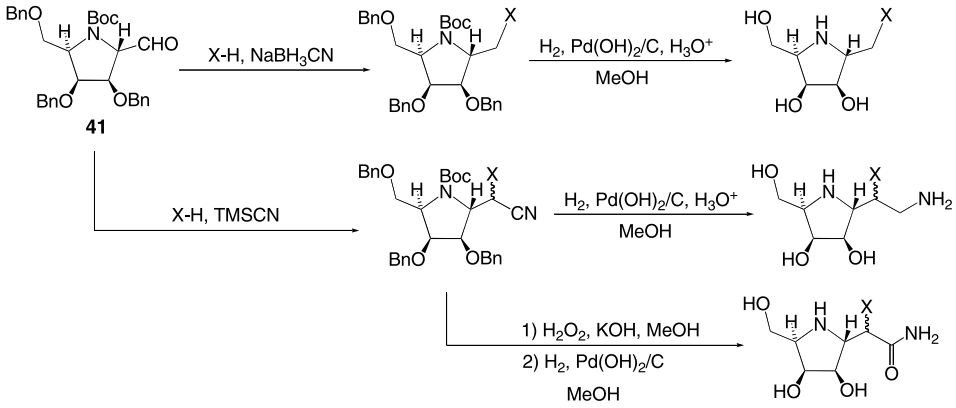
Ezekowitz and co-workers have revealed that human PGRP-S binds to and inhibits the growth of bacteria. PGRP-S co-localized with lysozyme in neutrophil extracellular traps and the antimicrobial effect of PGRP-S was synergistic with lysozyme [150]. PGRP-S bound to the tetrasaccharide having two tripeptide chains and the tetrasaccharide having two tetrapeptide chains, suggesting that the third amino acid in the peptide chain determines the binding affinity to PGRP-S.

Melanin synthesis of the arthropods is essential for defense and development. The melanization cascade is activated by bacterial PG or fungal β-1,3-glucan. Lee and co-workers reported that the tetrasaccharide tetrapeptide functioned as a competitive inhibitor of the natural PG-induced melanization reaction [151]. By using a tetrasaccharide tetrapeptide-coupled column, the *Tenebrio molitor* PG recognition protein (Tm-PGRP) was purified without activation of the prophenoloxidase.

2.5 Combinatorial Approach to Glycoenzyme Inhibitors and Lectin Ligands

A potential usefulness of the library as a source of inhibitors of glycoenzymes has been clearly exemplified by Kanie and co-workers (*Scheme 12*) [152]. A combinatorial approach based on the five-membered iminocyclitols with a *galacto*-configuration has been investigated; a small library of 27 iminocyclitol derivatives was prepared by either reductive amination or Strecker reaction of the aldehyde **41**. They have found that the introduction of a C10-alkyl group remarkably enhanced the inhibitory activity against α-mannosidase up to 10 μM, compared with its parent iminocyclitol and deoxymannojirimycin. Furthermore, the compounds bearing the phenylethyl group showed an extremely strong inhibitory activity against α-galactosaminidase (*Ki* value: 29.4 nM).

Schmidt and co-workers have been investigating a series of potent inhibitors of α(2,6)-sialyltransferase (*Fig. 9*). They developed the transition analog **42** [153], based on the proposed mechanism of the sialyl transfer, that involves the partial dissociation of the CMP and the formation of the planer oxocarbenium ion structure in the transition state. They also found that the planer neuraminyl moiety in **42** can be replaced by the aromatic groups, leading to the readily accessible aromatic inhibitors of α(2,6)-sialyltransferase from rat liver. Further library



Scheme 12

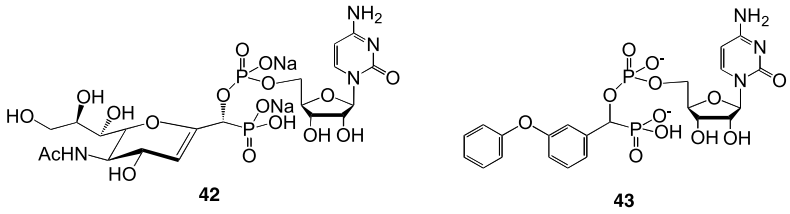
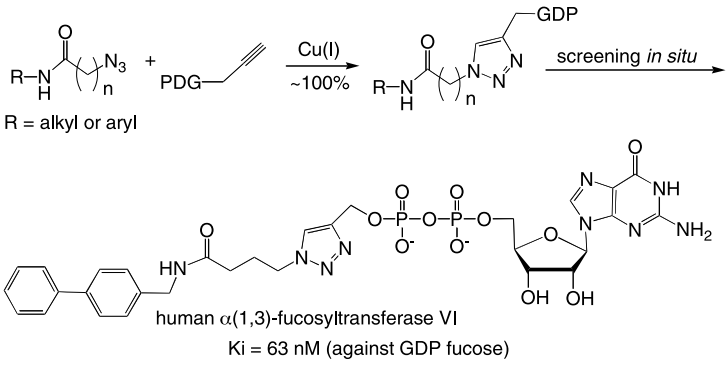


Figure 9
α(2,6)-Sialyltransferase inhibitors by Schmidt and co-workers

synthesis led to the development of **43** [154], which mimics a planer anomeric carbon, a large distance between the anomeric carbon and the CMP group, and two negative charges in the transition state, showed a *Ki* value down to 70 nM.

A potent inhibitor against human α(1,3)-fucosyltransferase VI was identified from a GDP-triazole library of 85 compounds, which was produced by the Cu(I)-catalyzed [2 + 3] cycloaddition reaction between azide and acetylene, followed by in situ screening without product isolation (*Scheme 13*) [155]. Kinetic evaluation of the purified inhibitor (*Scheme 13*) showed that it is a competitive inhibitor against GDP-fucose with $Ki = 62$ nM, which would make this compound the first nanomolar and most potent inhibitor of Fuc-Ts.

The polypeptide *N*-acetyl-α-galactosaminyltransferases (ppGalNAcTs) initiate mucin-type *O*-linked glycosylation and therefore play pivotal roles in cell–cell communication and protection of tissues. Bertozzi and co-workers screened a 1338-member uridine-based library to identify small molecule inhibitors of ppGalNAcTs (*Fig. 10*) [156]. Using a high-throughput enzyme-linked lectin assay (ELLA), two inhibitors of murine ppGalNAcTs-1 (*Ki* approximately 8 μM) were identified that also inhibit several other members of the family. The compounds did not inhibit other mammalian glycosyltransferases or nucleotide sugar utilizing enzymes, suggesting selectivity for the ppGalNAcTs. Treatment of cells with the compounds abrogated mucin-type *O*-linked glycosylation but not *N*-linked glycosylation and also induced apoptosis. These uridine analogs represent the first generation of chemical tools to study the functions of mucin-type *O*-linked glycosylation.



Scheme 13

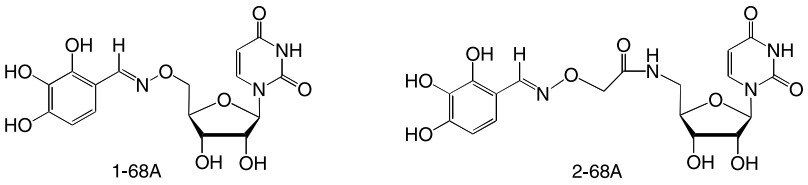
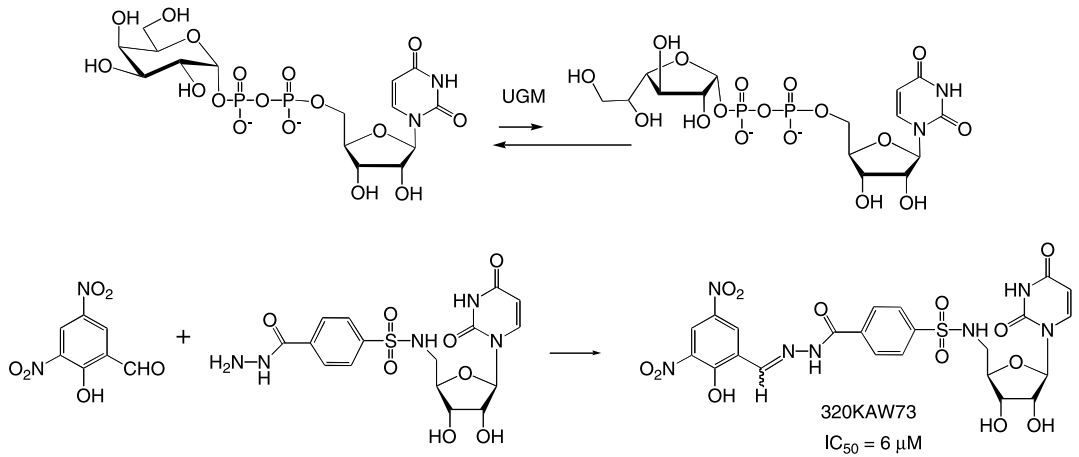


Figure 10
Uridine-based inhibitors of ppGalNAcTs by Bertozzi and co-workers

The flavoenzyme UDP-galactopyranose mutase (UGM) plays a key role in the cell wall biosynthesis of many pathogens, including *Mycobacterium tuberculosis*. McNeil and co-workers developed a microtiter plate assay for UGM (*Scheme 14*) [157]. The assay is based on the release of tritiated formaldehyde from UDP-galactofuranose but not UDP-galactopyranose by periodate and was used to identify a uridine-based enzyme inhibitor from a chemical library. The potent inhibitor 320KAW73 ($IC_{50} = 6\,\mu M$) was identified.

Kiessling and co-workers screened 16,000 compounds in a fluorescence polarization assay by using a synthetic fluorescent ligand; effective inhibitors of UGM were identified (*Fig. 11*) [158].

Kiso and co-workers have succeeded in the total synthesis of a series of gangliosides and sialic acid-containing glycosphingolipids as well as their derivatives, and revealed their biological functions at the molecular level. Among them, they have found that the cholinergic neuron-specific Cho-1 (α-series) gangliosides are high affinity ligands for neural siglecs (sialic acid-binding Ig-like lectins) (*Fig. 12*). α(2–3)-Linked sialic acid at Gal (II^3) and α(2–6)-linked sialic acid at GalNAc(III^6) greatly enhance the siglec-mediated cell adhesion, and especially, GQ1bα showed a remarkably high affinity to myelin-associated glycoprotein (MAG, siglec-4a). On the basis of the interesting finding through structure-activity studies that these internal sialic acids can be substituted by other anionic groups, they have synthesized three sulfated GM1b analogs **44–46**, as the structural mimics of GQ1bα [159,160]. Surprisingly, **46** bearing the β(1–4)linkage between GalNAc(III) and Gal(IV), being different from the natural β(1–3)



Scheme14

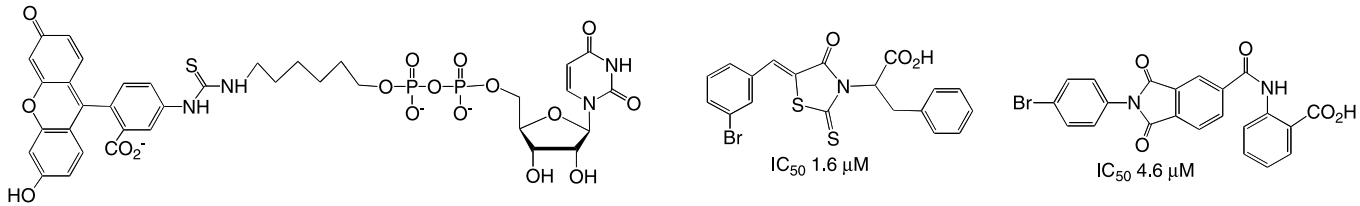


Figure 11
UGM Inhibitors by Kiessling andco-workers

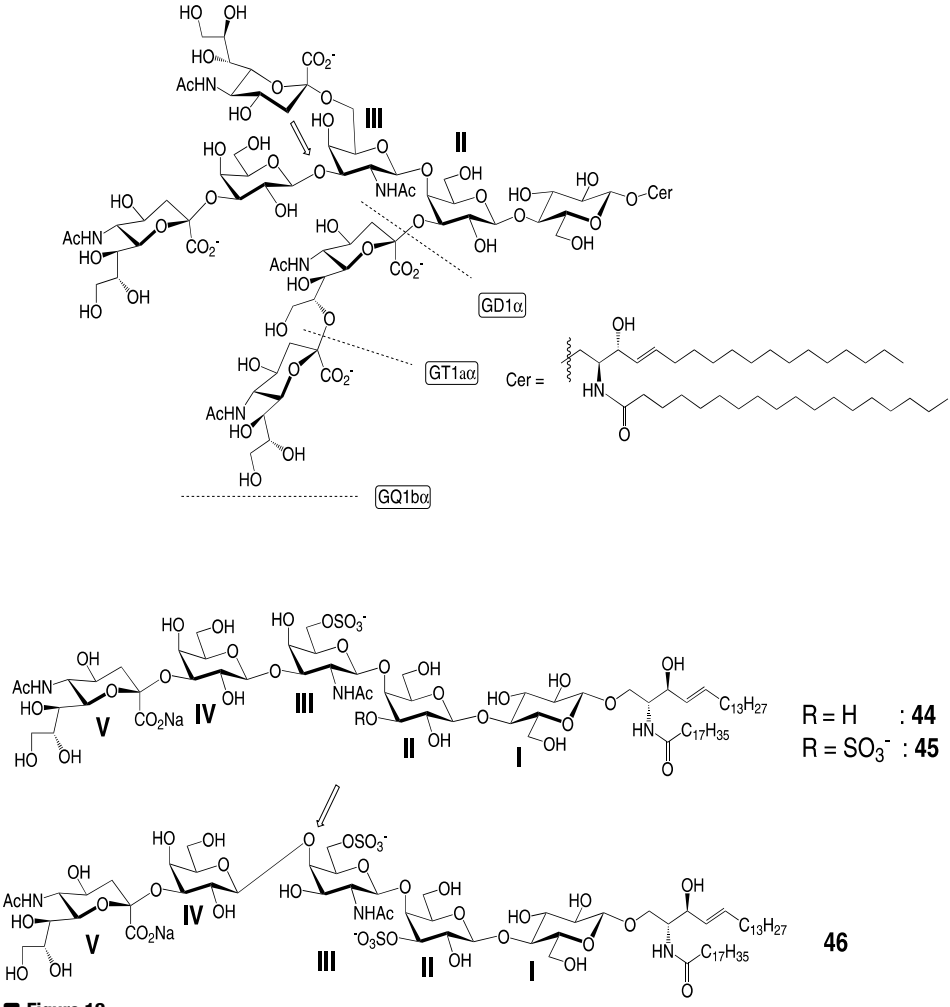


Figure 12
Cholinergic neuron-specific Cho-1 (α-series) gangliosides and sulfated GM1b analogs by Kiso and co-workers

linkage in GQ1bα (see *arrows* in *Fig. 12*), exhibited the most potent MAG binding affinity tested to date. The derivative **46** was also found to be the most potent inhibitor of NADase activity of leukocyte cell surface antigen CD38, inconsistent with the requirement of the internal α(2–3)-linked sialic acid in GQ1bα for NADase inhibition. The structures of **46** and **45** calculated using CAChe software revealed that both sulfate groups and the terminal sialic acid in **46** make the negative charge cluster, which might be easily recognized by MBP-CD38, while the terminal sialic acid in **45** protrudes from the anion cluster of sulfates. Future investigations on the synthetic sulfated gangliosides will open new perspectives on their structure-function relationships.

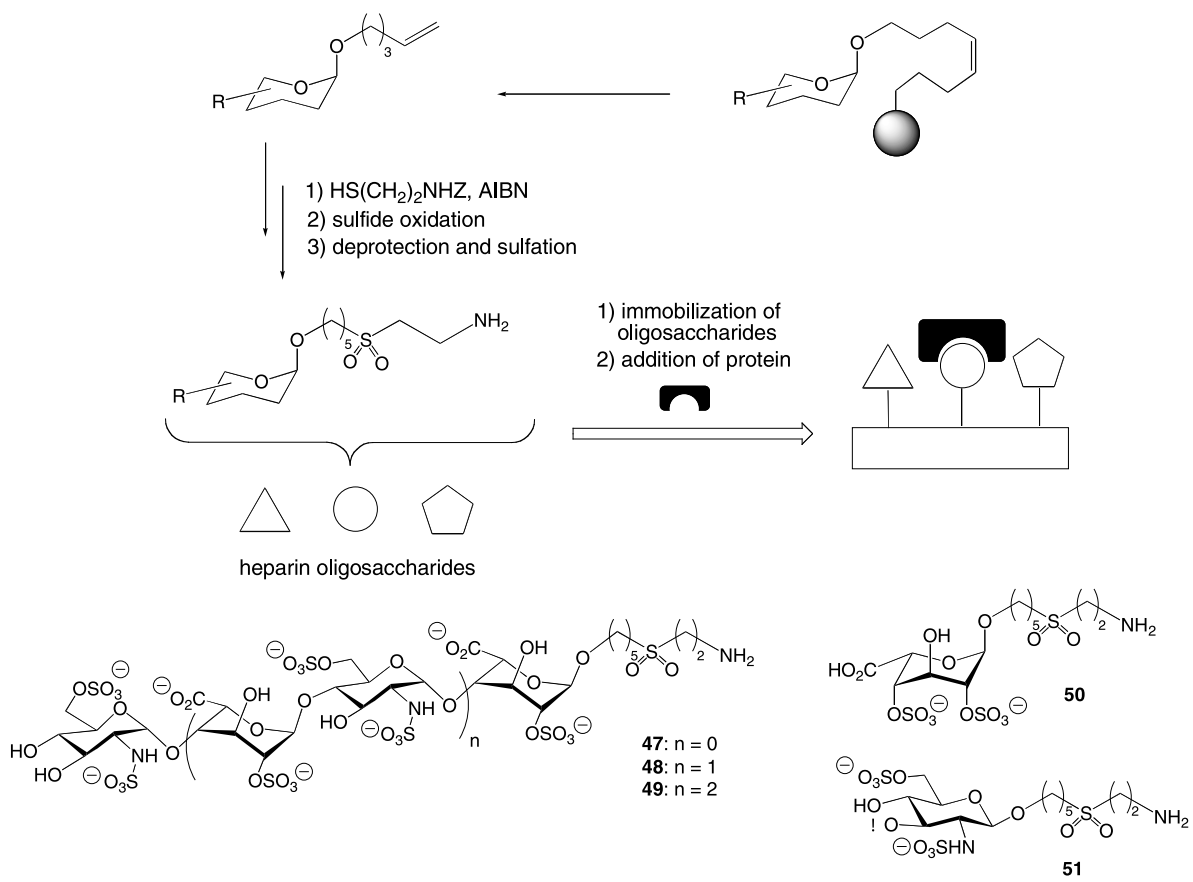
2.6 Application of the Oligosaccharide Library to a Microarray

Seeberger and co-workers examined the interaction between heparin oligosaccharides and fibroblast growth factors (FGF-1 and FGF-2) on the microchip, which are involved in developmental and physiological processes, including cell proliferation, differentiation, morphogenesis, and angiogenesis (*Scheme 15*) [161]. The 2-aminoethylsulfonyl linker in **47–51** was introduced by the reaction of the pentenyl moiety of the heparin derivatives, being generated during their automated synthesis, with 2-(benzyloxycarbonylamino)-1-ethanethiol in the presence of AIBN followed by air oxidation of sulfide and deprotection of the benzyloxycarbony group. After the amine-functionalized **47–51** were covalently attached to the amine-reactive CodeLink slides that are coated with a hydrophilic polymer with *N*-hydroxysuccinimidyl ester, the FGF-binding assay was performed using the anti-human FGF polyclonal antibodies and a secondary antibody labeled with an Alexa 546 dye. The basic FGF-2 strongly bound to **48** and **49** based on the highest fluorescence intensities, while monosaccharide constituents **50** and **51** were not bound. Similar results were obtained for the acidic FGF-1, but interestingly, they found that **50** exhibited a fluorescent intensity comparable to those of **48** and **49**; the presence of an unnatural 2,4-*O*-sulfation pattern might be responsible for this interaction.

Seeberger and co-workers have also investigated the glycan-based interaction of HIV-1 envelope glycoproteins on the microarray pursuing vaccines and agents that interfere with HIV events [162]. Screening of the seven high mannose-type oligosaccharides on the slides, prepared by total synthesis, revealed established the five relevant gp-120 binding proteins, including the dendric cell lectin DC-SIGN (Dendric Cell-Specific Intercellular Adhesion Molecule-1-Grabbing Nonintegrin), the HIV-inactivating protein scytovirin, and the human antibody 2G12. The present microarray analysis also included the synthetic neoglycoproteins, which gave additional information on the gp-120 interaction, regarding the importance of the peptide backbone to which the high mannose oligosaccharides are attached.

Similarly, Wong and co-workers have studied carbohydrate specificity to the anti-HIV-1 antibody, 2G12 using a carbohydrate microarray and ELISA (solution-phase Enzyme-Linked Immunosorbent Assay) [163]. The azide group introduced on the reducing end of the oligosaccharides [164] were covalently attached to a microplate via a Cu(I)-mediated 1,3-dipolar cycloaddition reaction (Meldal/Sharpless click reaction) through a disulfide linker [165]; the use of the cleavable S-S linker enabled the quantification of the oligosaccharides immobilized on the slides by mass spectrometry or the phenol-sulfuric acid test (*Fig. 13*). The study revealed that the antibody 2G12 bound both D1 and D3 arms of the $Man_9GlcNAc_2$ structure, thus providing important information for the design of HIV vaccines.

They also developed a detection method for oligosaccharides on the slides by desorption/ionization on silicon mass spectrometry (DIOS-MS) without the matrix [166]. An amine-containing antigen N3 minor was prepared by their one-pot synthesis and it was printed on the surface of modified porous silicon with a photocleavable linker. The linker can be cleaved by a laser ($\lambda = 337$ nm) during the mass spectrometric measurements, thereby enabling direct characterization of the released and ionized carbohydrates on the porous silicon with a minuscule amount of sample. The new method will greatly expand the high-throughput analysis of a large glycan array, which will be prepared by their programmed one-pot oligosaccharide synthesis based on the use of designer thioglycoside building blocks with defined relative reactivity values (RRVs) that are collected in the Optimer database (*Sect. 2.1*).


1) HS(CH2)2NHZ, AIBN
2) sulfide oxidation
3) deprotection and sulfation
1) immobilization of oligosaccharides
2) addition of protein
heparin oligosaccharides
47: n = 0
48: n = 1
49: n = 2
50
51

Scheme 15

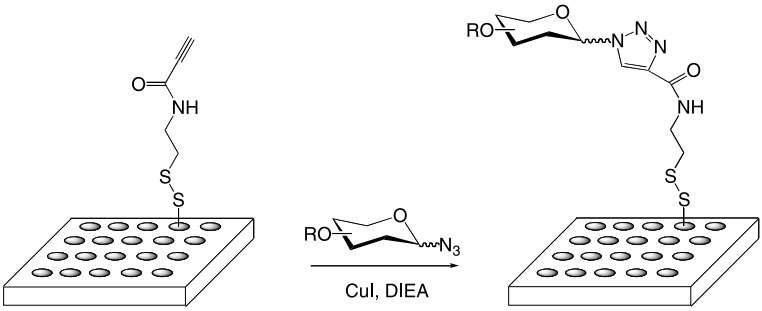


Figure 13
Covalent attachment to a microplate by Wong and co-workers

Schmidt and co-workers have also immobilized the mannose derivatives to the commercially available glass slides under reductive amination or acylation conditions [167]. Their study showed that the ConA interaction was in accordance with those obtained in aqueous solution and other immobilization methods, and non-specific protein binding to the functionalized glass slides was negligible. The same carbohydrate microarray plate could also be tested for other lectins, therefore the "precious" immobilized carbohydrates can be reused.
Fluorous-based non-covalent immobilization on the slide surface has also been reported recently [168]. Fluorous-tags containing oligosaccharides, which were prepared by the fluorous-based synthesis, were printed on the fluorous-coated glass slides, and the slides were screened by fluorescein-labeled lectins as a model of carbohydrate–protein interactions. Although the method avoids the extra derivatization steps for the immobilization on the slides, the lack of applicability of the current fluorous-based synthesis to the biologically relevant oligosaccharides, i. e., complex glycans consisting of more than four sugars, as well as weak fluorous/fluorous interactions being problematic during the washing steps on the microchips, still remains to be improved.

2.7 Libraries Using Carbohydrate Scaffolds and Miscellaneous

Another important subject in combinatorial chemistry of carbohydrates is the use of carbohydrates as scaffolds. So far various libraries using carbohydrates as scaffolds have been developed and many bioactive compounds have been found towards various targets including the somatostatin receptors (SSTR), integrins, HIV-1 protease, matrix metalloproteinases (MMP), multidrug resistance-associated protein (MRP), and as RNA binders [169]. Montesarchio and co-workers have also prepared cyclic phosphate-linked oligosaccharides in order to investigate their conformational properties [170], as well as the sucrose-oligonucleotide hybrids [171], in an attempt to increase chemical and enzymic stability of the oligonucleotides without interfering with the duplex formation and with the ability of G-rich sequences to adopt a quadruplex structure. Takahashi and co-workers have reported a parallel synthesis of a enediyne-oligosaccharide library, using silyl-linked solid-support [172]. Combinatorial approaches to the glycopolymers and glycoclusters, based on the asparaginyl template [173], ε-sugar amino acid templates [174], and cyclo-β-tetrapeptides [175], have also been reported.

3 Conclusions

In this chapter, we focused on the synthesis of oligosaccharide libraries and carbohydrate mimetic libraries and their use for glycobiology. Recently, oligosaccharide synthesis including solid-phase synthesis has been dramatically improved and various complex oligosaccharides and glycoconjugates have been synthesized for investigation of their biological functions. Many efforts will be continued to establish efficient synthesis of oligosaccharides. The library approach based on chemical synthesis will become a generally accepted method for elucidation of oligosaccharide functions, since chemical synthesis can provide a series of compounds in pure states without contamination by other bioactive components. In addition, oligosaccharide and glycoconjugate libraries will lead to an understanding of the biological importance of carbohydrate structural diversity with the advent of glycobiology.

References

1. Fukase K (2001) In: Fraser-Reid BO, Tatsuta K, Thiem J (eds) Glycoscience: Chemistry and Chemical Biology I-III, Springer, Berlin, Chap 5.3.5
2. Khersonsky SM, Ho CM, Garcia MF, Chang Y-T (2003) Cur Topics Med Chem 3:617
3. Baytas SN, Linhardt RJ (2004) Mini-Rev Org Chem 1:27
4. Werz DB, Seeberger PH (2005) Chem Eur J 11:3194
5. Douglas NL, Ley SV, Lücking U, Warriner SL (1998) J Chem Soc, Perkin Trans 1:51
6. Zhang Z, Ollmann IR, Ye X-S, Wischnat R, Baasov T, Wong C-H (1999) J Am Chem Soc 121:734
7. Burkhart F, Zhang Z, Wacowich-Sgarbi S, Wong C-H (2001) Angew Chem Int Ed 40:1274
8. Takahashi T, Adachi M, Matsuda A, Doi T (2000) Tetrahedron Lett 41:2599
9. Tanaka H, Matoba N, Takahashi T (2005) Chem Lett 34:400
10. Tanaka H, Adachi M, Takahashi T (2004) Tetrahedron Lett 45:1433
11. Tanaka H, Adachi M, Takahashi T (2005) Chem Eur J 11:849
12. Kanie O, Barresi F, Ding Y, Labbe J, Otter A, Forsberg LS, Ernst B, Hindsgaul O (1995) Angew Chem Int Ed Engl 34:2720
13. Ding Y, Labbe J, Kanie O, Hindsgaul O (1996) Bioorg Med Chem 4:683
14. Yu B, Li B, Xing G, Hui Y (2001) J Comb Chem 3:404
15. Ohtsuka I, Ako T, Kato R, Daikoku S, Koroghi S, Kanemitsu T, Kanie O (2006) Carbohydr Res 341:1476
16. Kanie O, Ohtsuka I, Ako T, Daikoku S, Kanie Y, Kato R (2006) Angew Chem Int Ed 45:3851
17. Ako T, Daikoku S, Ohtsuka I, Kato R, Kanie O (2006) Chem Asian J 1:798
18. Amaya T, Tanaka H, Takahashi T (2004) Synlett: 497
19. Elsayed GA, Zhu T, Boons G-J (2002) Tetrahedron Lett 43:4691
20. Ishiwata A, Ito Y (2005) Tetrahedron Lett 46:3521
21. Matsuo I, Ito Y (2003) Carbohydr Res 338:2163
22. Matsuo I, Wada M, Manabe S, Yamaguchi Y, Otaka K, Kato K, Ito Y (2003) J Am Chem Soc 125:3402
23. Matsuo I, Kashiwagi T, Totani K, Ito Y (2005) Tetrahedron Lett 46:4197
24. Matsuo I, Totani K, Tatami A, Ito Y (2006) Tetrahedron 62:8262
25. Takatani M, Matsuo I, Ito Y (2003) Carbohydr Res 338:1073
26. Takatani M, Nakano J, Arai MA, Ishiwata A, Ohta H, Ito Y (2004) Tetrahedron Lett 45:3929
27. Jonke S, Liu K-G, Schmidt RR (2006) Chem Eur J 12:1274
28. Ratner DM, Plante OJ, Seeberger PH (2002) Eur J Org Chem 5:826
29. Kehr J-C, Zilliges Y, Springer A, Disney MD, Ratner DD, Bouchier C, Seeberger PH, de Marsac NT, Dittmann E (2006) Mol Microbiol 59:893
30. Schofield L, Hewitt MC, Evans K, Siomos MA, Seeberger PH (2002) Nature 418:785
31. Kwon Y-U, Soucy RL, Snyder DA, Seeberger PH (2005) Chem Eur J 11:2493

32. Hewitt MC, Snyder DA, Seeberger PH (2002) J Am Chem Soc 124:13434
33. Liu X, Kwon Y-U, Seeberger PH (2005) J Am Chem Soc 127:5004
34. Hileman RE, Smith AE, Toida T, Linhardt RJ (1997) Glycobiology 7:231
35. Toida T, Hileman RE, Smith AE, Vlahova PI, Linhardt RJ (1996) J Biol Chem 271:32040
36. Rice KG, Linhardt RJ (1989) Carbohydr Res 190:219
37. Ashikari-Hada S, Habuchi H, Kariya Y, Itoh N, Reddi AH, Kimata K (2004) J Biol Chem 279:12346
38. Pye DA, Vives RR, Turnbull JE, Hyde P, Gallagher JT (1998) J Biol Chem 4:22936
39. Ito Y, Hikino M, Yajima Y, Mikami T, Sirko S, von Holst A, Faissner A, Fukui S, Sugahara K (2005) Glycobiology 15:593
40. Yates EA, Guimond SE, Turnbull JE (2004) J Med Chem 47:277
41. Jemth P, Smeds E, Do AT, Habuchi H, Kimata K, Lindahl U, Kusche-Gullberg M (2003) J Biol Chem 278:24371
42. Takagaki K, Munakata H, Majima M, Kakizaki I, Endo M (2000) J Biochem 2000:695
43. Takagaki K, Munakata H, Kakizaki I, Iwafune M, Itabashi T, Endo M (2002) J Biol Chem 277:8882
44. Pye DA, Vives RR, Hyde P, Gallagher JT (2000) Glycobiology 10:1183
45. Guimond SE, Turnbull JE (1999) Curr Biol 9:1343
46. Noti C, Seeberger PH (2005) Chem Biol 12:731
47. Sinay PJ, Jacquinet JC, Petitou M, Duchaussoy P, Lederman I, Choay J, Torri G (1984) Carbohydr Res 132:C5
48. Petitou M, Duchaussoy P, Lederman I, Choay J, Sinay P, Jacquinet JC, Torri G (1986) Carbohydr Res 147:221
49. Petitou M, Duchaussoy P, Lederman I, Choay J, Jacquinet JC, Sinay P, Torri G (1987) Carbohydr Res 167:67
50. van Boeckel CAA, Beetz T, van Aelst SF (1988) Tetrahedron Lett 29:803
51. Grootenhuis PD, Westerduin P, Meuleman D, Petitou M, van Boeckel CAA (1995) Nat Struct Biol 2:736
52. van Boeckel CAA, Petitou M (1993) Angew Chem Int Ed Engl 32:1671
53. Petitou M, van Boeckel CAA (2004) Angew Chem Int Ed 43:3118
54. Tabeur C, Mallet JM, Bono F, Herbert JM, Petitou M, Sinay P (1999) Bioorg Med Chem 7:2003
55. Nilsson M, Svahn CM, Westman J (1993) Carbohydr Res 246:161
56. Westman J, Nilsson M, Ornitz DM, Svahn CM (1995) J Carbohydr Chem 14:95
57. Kovensky J, Duchaussoy P, Petitou M, Sinay P (1996) Tetrahedron: Asymmetry 7:3119
58. Poletti L, Fleischer M, Vogel C, Guerrini M, Torri G, Lay L (2001) Eur J Org Chem: 2727
59. de Paz JL, Angulo J, Lassaletta JM, Nieto PM, Redondo-Horcajo M, Lozano RM, Gimenez-Gallego G, Martin-Lomas M (2001) ChemBiochem 2:673
60. Angulo J, Ojeda R, de Paz JL, Lucas R, Nieto PM, Lozano RM, Redondo-Horcajo M, Gimenez-Gallego G, Martin-Lomas M (2004) ChemBiochem 5:55
61. Suda Y, Bird K, Shiyama T, Koshida S, Marques D, Fukase K, Sobel M, Kusumoto S (1996) Tetrahedron Lett 37:1053
62. Koshida S, Suda Y, Sobel M, Ormsby J, Kusumoto S (1999) Bioorg Med Chem 9:3127
63. Sobel M, Fish WR, Toma N, Luo S, Bird K, Mori K, Kusumoto S, Blystone SD, Suda Y (2001) J Vasc Surg 33:587
64. Koshida S, Suda Y, Sobel M, Kusumoto S (2001) Tetrahedron Lett 42:1289
65. Haller M, Boons GJ (2001) J Chem Soc Perkin Trans 1:814
66. Orgueira HA, Bartolozzi A, Schell P, Litjens RE, Palmacci ER, Seeberger PH (2003) Chem Eur J 9:140
67. Dreef-Tromp CM, Willems HAM, Westerduin P, van Veelen P, van Boeckel CAA (1997) Bioorg Med Chem Lett 7:1175
68. Ojeda R, de Paz JL, Martin-Lomas M (2003) Chem Commun: 2486
69. Ojeda R, Terenti O, de Paz JL, Martin-Lomas M (2004) Glycoconj J 21:179
70. Hanashima S, Manabe S, Ito Y (2005) Angew Chem Int Ed 44:4218
71. Furuike T, Yamada K, Ohta T, Monde K, Nishimura S-I (2003) Tetrahedron 59:5105
72. Naruchi K, Hamamoto T, Kurogochi M, Hinou H, Shimizu H, Matsushita T, Fujitani N, Kondo H, Nishimura S-I (2006) J Org Chem 71:9609
73. Furuike T, Sadamoto R, Niikura K, Monde K, Sakairi N, Nishimura S-I (2005) Tetrahedron 61:1737
74. André S, Kaltner H, Furuike T, Nishimura S-I, Gabius H-J (2004) Bioconjugate Chem 15:87
75. Sallas F, Niikura K, Nishimura S-I (2004) Chem Commun: 596

76. Tachibana Y, Matsubara N, Nakajima F, Tsuda T, Tsuda S, Monde K, Nishimura S-I (2002) Tetrahedron 58:10213
77. Matsushita T, Hinou H, Fumoto M, Kurogochi M, Fujitani N, Shimizu H, Nishimura S-I (2006) J Org Chem 71:3051
78. Tachibana Y, Monde K, Nishimura S-I (2004) Macromolecules 37:6771
79. Matsumoto S, Matsushita M, Morita T, Kamachi H, Tsukiyama S, Furukawa Y, Koshida S, Tachibana Y, Nishimura S-I, Todo S (2006) Cryobiology 52:90
80. Furuike T, Aiba S, Suzuki T, Takahashi T, Suzuki Y, Yamada K, Nishimura S-I (2000) J Chem Soc, Perkin Trans 1:3000
81. Ohta T, Miura N, Fujitani N, Nakajima F, Niikura K, Sadamoto R, Guo C-T, Suzuki T, Suzuki Y, Monde K et al (2003) Angew Chem Int Ed 42:5186
82. Sato M, Furuike T, Sadamoto R, Fujitani N, Nakahara T, Niikura K, Monde K, Kondo H, Nishimura S-I (2004) J Am Chem Soc 126:14013
83. Fumoto M, Hinou H, Ohta T, Ito T, Yamada K, Takimoto A, Kondo H, Shimizu H, Inazu T, Nakahara Y et al (2005) J Am Chem Soc 127:11804
84. Fumoto M, Hinou H, Matsushita T, Kurogochi M, Ohta T, Ito T, Yamada K, Takimoto A, Kondo H, Inazu T et al (2005) Angew Chem Int Ed 44:2534
85. Li H, Singh S, Zeng Y, Song H, Wang L-X (2005) Bioorg Med Chem Lett 15:895
86. Li H, Li B, Song H, Breydo L, Baskakov IV, Wang L-X (2005) J Org Chem 70:9990
87. Kajihara Y, Suzuki Y, Yamamoto N, Sasaki K, Sakakibara T, Juneja LR (2004) Chem Eur J 10:971
88. Fukae K, Yamamoto N, Hatakeyama Y, Kajihara Y (2004) Glycoconjugate J 21:243
89. Yamamoto N, Ohmori Y, Sakakibara T, Sasaki K, Juneja LR, Kajihara Y (2003) Angew Chem Int Ed 42:2537
90. Kajihara Y, Yamamoto N, Miyazaki T, Sato H (2005) Curr Med Chem 12:527
91. Yamamoto N, Takayanagi A, Sakakibara T, Dawson PE, Kajihara Y (2006) Tetrahedron Lett 47:1341
92. Kajihara Y, Yoshihara A, Hirano K, Yamamoto N (2006) Carbohydr Res 341:1333
93. Wang Z-G, Warren JD, Dudkin VY, Zhang X, Iserloh U, Visser M, Eckhardt M, Seeberger PH (2006) Tetrahedron 62:4954
94. Wang Z-G, Zhang X, Visser M, Live D, Zatorski A, Iserloh U, Lloyd KO, Danishefsky SJ (2001) Angew Chem Int Ed 40:1728
95. Dudkin VY, Miller JS, Danishefsky SJ (2004) J Am Chem Soc 126:736
96. Dudkin VY, Orlova M, Geng X, Mandal M, Olson WC, Danishefsky SJ (2004) J Am Chem Soc 126:9560
97. Wang Z-G, Zhang X, Live D, Danishefsky SJ (2000) Angew Chem Int Ed 39:3652
98. Mandal M, Dudkin VY, Geng X, Danishefsky SJ (2004) Angew Chem Int Ed 43:2557
99. Geng X, Dudkin VY, Mandal M, Danishefsky SJ (2004) Angew Chem Int Ed 43:2562
100. Miller JS, Dudkin VY, Lyon GJ, Muir TW, Danishefsky SJ (2003) Angew Chem Int Ed 42:431
101. Warren JD, Miller JS, Keding SJ, Danishefsky SJ (2004) J Am Chem Soc 126:6576
102. Chen J, Chen G, Wu B, Wan Q, Tan Z, Hua Z, Danishefsky SJ (2006) Tetrahedron Lett 47:8013
103. Wu B, Chen J, Warren JD, Chen G, Hua Z, Danishefsky SJ (2006) Angew Chem Int Ed 45:4116
104. Ragupathi G, Koide F, Livingston PO, Cho YS, Endo A, Wan Q, Spassova MK, Keding SJ, Allen J, Ouerfelli O et al (2006) J Am Chem Soc 128:2715
105. Glunz PW, Hintermann S, Williams LJ, Schwarz JB, Kuduk SD, Kudryashov V, Lloyd KO, Danishefsky SJ (2000) J Am Chem Soc 122:7273
106. Sabbatini PJ, Kudryashov V, Ragupathi G, Danishefsky SJ, Livingston PO, Bornmann W, Spassova M, Zatorski A, Spriggs D, Aghajanian C et al (2000) Int J Cancer 87:79
107. Allen JR, Allen JG, Zhang X-F, Williams LJ, Zatorski A, Ragupathi G, Livingston PO, Danishefsky SJ (2000) Chem Eur J 6:1366
108. Allen JR, Harris CR, Danishefsky SJ (2001) J Am Chem Soc 123:1890
109. Ragupathi G, Deshpande PP, Coltart DM, Kim HM, Williams LJ, Danishefsky SJ, Livingston PO (2002) Int J Cancer 99:207
110. Coltart DM, Royyuru AK, Williams LJ, Glunz PW, Sames D, Kuduk SD, Schwarz JB, Chen X-T, Danishefsky SJ, Live DH (2002) J Am Chem Soc 124:9833
111. Biswas K, Coltart DM, Danishefsky SJ (2002) Tetrahedron Lett 43:6107
112. Keding SJ, Endo A, Danishefsky SJ (2003) Tetrahedron 59:7023
113. Ragupathi G, Koide F, Sathyan N, Kagan E, Spassova M, Bornmann W, Gregor P, Reis CA, Clausen H, Danishefsky SJ et al (2003) Cancer Immunol Immunother 52:608

114. Cho YS, Wan Q, Danishefsky SJ (2005) Bioorg Med Chem 13:5259
115. Spassova MK, Bornmann WG, Ragupathi G, Sukenick G, Livingston PO, Danishefsky SJ (2005) J Org Chem 70:3383
116. Nakahahara Y, Ando S, Ito Y, Hojo H, Nakahara Y (2001) Biosci Biotechnol Biochem 65:1358
117. Ichiyanagi T, Takatani M, Sakamoto K, Nakahara Y, Ito Y, Hojo H, Nakahara Y (2002) Tetrahedron Lett 43:3297
118. Takano Y, Habiro M, Someya M, Hojo H, Nakahara Y (2002) Tetrahedron Lett 43:8395
119. Hojo H, Haginoya E, Matsumoto Y, Nakahara Y, Nabeshima K, Toole BP, Watanabe Y (2003) Tetrahedron Lett 44:2961
120. Halkes KM, Gotfredsen CH, Grotli M, Miranda LP, Duus JO, Meldal M (2001) Chem Eur J 7:3584
121. Zhu X, Pachamuthu K, Schmidt RR (2003) J Org Chem 68:5641
122. Zhu X, Schmidt RR (2003) Tetrahedron Lett 44:6063
123. Zhu X, Schmidt RR (2004) Chem Eur J 10:875
124. Zhu X, Pachamuthu K, Schmidt RR (2004) Org Lett 6:1083
125. Marcaurelle LA, Bertozzi CR (2001) J Am Chem Soc 123:1587
126. Pratt MR, Bertozzi CR (2003) J Am Chem Soc 125:6149
127. Kusumoto S, Fukase K, Fukase Y, Kataoka M, Yoshizaki H, Sato K, Oikawa M, Suda Y (2003) J Endotoxin Res 9:361
128. Fujimoto Y, Adachi Y, Akamatsu M, Fukase Y, Kataoka M, Suda Y, Fukase K, Kusumoto S (2005) J Endotoxin Res 11:341
129. Zhang S-Q, Fukase K, Izumi M, Fukase Y, Kusumoto S (2001) Synlett: 590
130. Fukase Y, Zhang S-Q, Iseki K, Oikawa M, Fukase K, Kusumoto S (2001) Synlett: 1693
131. Akamatsu M, Fujimoto Y, Kataoka M, Suda Y, Kusumoto S, Fukase K (2006) Bioorg Med Chem 14:6759
132. Ohto U, Fukase K, Miyake K, Satow Y (2007) Science (Washington, DC) (in press)
133. Christ WJ, Asano O, Robidoux ALC, Perez M, Wang Y, Dubuc GR, Gavin WE, Hawkins LD, McGuinness PD et al (1995) Science (Washington, DC) 268:80
134. Rossignol DP, Hawkins LD, Christ WJ, Kobayashi S, Kawata T, Lynn M, Yamatsu I, Kishi Y (1999) In: Brade H (ed) Endotoxin in Health and Disease. CRC Press, Boca Raton, FL, p. 699
135. Morita M, Motoki K, Akimoto K, Natori T, Sakai T, Sawa E, Yamaji K, Koezuka Y, Kobayashi E, Fukushima H (1995) J Med Chem 38:2176
136. Uchimura A, Shimizu T, Morita M, Ueno H, Motoki K, Fukushima H, Natori T, Koezuka Y (1997) Bioorg Med Chem 5:2245
137. Sakai T, Morita M, Matsunaga N, Akimoto K, Yokoyama T, Iijima H, Koezuka Y (1999) Bioorg Med Chem Lett 9:697
138. Fujio M, Wu D, Garcia-Navarro R, Ho DD, Tsuji M, Wong C-H (2006) J Am Chem Soc 128:9022
139. Fan G-T, Pan Y-s, Lu K-C, Cheng Y-P, Lin W-C, Lin S, Lin C-H, Wong C-H, Fang J-M, Lin C-C (2005) Tetrahedron 61:1855
140. Plettenburg O, Bodmer-Narkevitch V, Wong C-H (2002) J Org Chem 67:4559
141. Xing G-W, Wu D, Poles MA, Horowitz A, Tsuji M, Ho DD, Wong C-H (2005) Bioorg Med Chem 13:2907
142. Yang G, Schmieg J, Tsuji M, Franck RW (2004) Angew Chem Int Ed 43:3818
143. Chen G, Chien M, Tsuji M, Franck RW (2006) Chem Bio Chem 7:1017
144. Franck RW, Tsuji M (2006) Acc Chem Res 39:692
145. Bauer J, Brandenburg K, Zaehringer U, Rademann J (2006) Chem Eur J 12:7116
146. Inamura S, Fujimoto Y, Kawasaki A, Shiokawa Z, Woelk E, Heine H, Lindner B, Inohara N, Kusumoto S, Fukase K (2006) Org Biomol Chem 4:232
147. Fujimoto Y, Inamura S, Kawasaki A, Shiokawa Z, Shimoyama A, Hashimoto T, Kusumoto S, Fukase K (2007) J Endotoxin Res 13:189
148. Inohara N, Ogura Y, Fontalba A, Gutierrez O, Pons F, Crespo J, Fukase K, Inamura S, Kusumoto S, Hashimoto M et al (2003) J Biol Chem 278:5509
149. Wang Z-M, Li X, Cocklin RR, Wang M, Wang M, Fukase K, Inamura S, Kusumoto S, Gupta D, Dziarski R (2003) J Biol Chem 278:49044
150. Cho JH, Fraser IP, Fukase K, Kusumoto S, Fujimoto Y, Stahl GL, Ezekowitz RA (2005) Blood 106:2551
151. Park JW, Je B-R, Piao S, Inamura S, Fujimoto Y, Fukase K, Kusumoto S, Söderhäll K, Ha N-C, Lee BL (2006) J Biol Chem 281:7747
152. Saotome C, Wong C-H, Kanie O (2001) Chem Biol 8:1061
153. Schworer R, Schumidt RR (2002) J Am Chem Soc 124:1632
154. Skropeta D, Schworer R, Haag T, Schumidt RR (2004) Glycoconjugate J 21:205

155. Lee LV, Mitchell ML, Huang S-J, Fokin VV, Sharpless KB, Wong C-H (2003) J Am Chem Soc 125:9588
156. Hang HC, Yu C, Ten Hagen KG, Tian E, Winans KA, Tabak LA, Bertozzi CR (2004) Chem Biol 11:337
157. Scherman MS, Winans KA, Stern RJ, Jones V, Bertozzi CR, McNeil MR (2003) Antimicrob Agents Chemother 47:378
158. Soltero-Higgin M, Carlson EE, Phillips JH, Kiessling LL (2004) J Am Chem Soc 126:10532
159. Ito H, Ishida H, Collins BE, Fromholt SE, Schnaar RL, Kiso M (2003) Carbohydr Res 338:1621
160. Hara-Yokoyama M, Ito H, Ueno-Noto K, Takano K, Ishida H, Kiso M (2003) Bioorg Med Chem Lett 13:3441
161. de Paz JL, Noti C, Seeberger PH (2006) J Am Chem Soc 128:2766
162. Adams EW, Ratner DM, Bokesch HR, McMahon JB, O Keefe BR, Seeberger PH (2004) Chem Biol 11:875
163. Calarese DA, Lee HK, Huang CY, Best MD, Astronomo RD, Stanfield RL, Katinger H, Burton DR, Wong CH, Wilson IA (2005) Proc Natl Acad Sci U S A 102:13372
164. Nyffeler PT, Liang CH, Koeller KM, Wong CH (2002) J Am Chem Soc 124:10773
165. Bryan MC, Fazio F, Lee HK, Huang CY, Chang A, Best MD, Calarese DA, Blixt O, Paulson JC, Burton DR et al (2004) J Am Chem Soc 126:8640
166. Lee J-C, Wu C-Y, Apon JV, Siuzdak G, Wong C-H (2006) Angew Chem Int Ed 45:2753
167. Biskup MB, Iler JUM, Weingart R, Schumidt RR (2005) Chem Bio Chem 6:1007
168. Ko KS, Jaipuri FA, Pohl NL (2005) J Am Chem Soc 127:13162
169. Becker B, Condie GC, Le GT, Meutermans W (2006) Mini-Rev Med Chem 6:1299
170. Di Fabio G, Randazzo A, D'Onofrio J, Ausin C, Pedroso E, Grandas A, De Napoli L, Montesarchio D (2006) J Org Chem 71:3395
171. Adinolfi M, De Napoli L, Di Fabio G, Iadonisi A, Montesarchio D (2004) Org Biomol Chem 2:1879
172. Matsuda A, Doi T, Tanaka H, Takahashi T (2001) Synlett 7:1101
173. Ziegler T, Roseling D, Subramanian LR (2002) Tetrahedron: Asymmetry 13:911
174. van Well RM, Marinelli L, Erkelens K, van der Marel GA, Lavecchia A, Overkleeft HS, van Boom JH, Kessler H, Overhand M (2003) Eur J Org Chem 2003:2303
175. Virta P, Karskela M, Lonnberg H (2006) J Org Chem 71:1989

5.5 Polymer-Supported and Tag-Assisted Methods in Oligosaccharide Synthesis

Katsunori Tanaka, *Koichi Fukase*
Department of Chemistry, Graduate School of Science, Osaka University, Osaka 560-0043, Japan
koichi@chem.sci.osaka-u.ac.jp

Abstract

Recent advances in polymer-supported synthesis of oligosaccharides, i. e., solid-phase synthesis, polymer-supported synthesis, and tag-assisted solution synthesis, are overviewed. A variety of new methods on polymer-supports, combined with accumulation of information on oligosaccharides synthesis in solution, has led to the assembly of complex oligosaccharide structures in the solid-phase, and even to a programmed automation system. A continuous flow synthesis of oligosaccharides using a microreactor, a newly appearing protocol for glycosylation analysis, high-throughput- and combinatorial synthesis, and even for production-scale synthesis, will also be described.

Keywords

Solid-phase synthesis; Polymer-supported synthesis; Polymer-supported solution synthesis; Tag-assisted solution synthesis; Combinatorial chemistry; High-throughput synthesis; Library; Microreactor

In: *Glycoscience*. Fraser-Reid B, Tatsuta K, Thiem J (eds)
Chapter-DOI 10-1007/978-3-540-30429-6_29:

Abbreviations

Bn	benzyl
Bz	benzoyl
CAN	ammonium cerium(IV) nitrate
ClAc	chloroacetyl
DIPEA	diisopropylethylamine
DMF	dimethylformamide
DMM	dimethylmaleoyl
DMTSB	dimethyl(methylthio)sulfonium tetrafluoroborate
DMTST	dimethyl(methylthio)sulfonium triflate
DTBMP	2,6-di-*tert*-butyl-4-methylpyridine
DTBP	di-*tert*-butylpyridine
EDTA	ethylenediaminetetraacetic acid
Fmoc	9-fluorenylmethoxycarbonyl
HBTU	2-(1*H*-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate
HEPES	*N*-(2-hydroxyethyl)-piperazine-*N*′-(2-ethanesulfonic acid)
HOBt	1-hydroxybenzotriazole
Lev	levulinoyl
MALDI-TOF-MS	matrix-assisted laser desorption ionization mass spectrometry
MES	2-morpholinoethanesulfonic acid
MPEG	polyethylene glycol monomethyl ether
MS	molecular sieves
NIPAm	*N*-*i*-propylacrylamide
NIS	*N*-iodosuccinimide
PEG	polyethylene glycol
Pht	phthalyl
Piv	pivaloyl
PS-PEG	polystyrene-polyethylene glycol
PyBOP	(1*H*-benzotriazole-1-yloxy) tripyrrolidinophosphonium hexafluorophosphate
TBDMS	*tert*-butyldimethylsilyl
TCA	trichloroacetyl
TFA	trifluoroacetic acid
TMSOTf	trimethylsilyl triflate
Troc	2,2,2-trichloroethoxycarbonyl
UDP	uridine diphosphate

1 Introduction

Synthesis of oligosaccharides and glycoconjugates has played an important role in elucidation of their biological functions. However, substantial time and efforts are often needed, since it requires multiple step transformations involving iterative deprotection-glycosylation

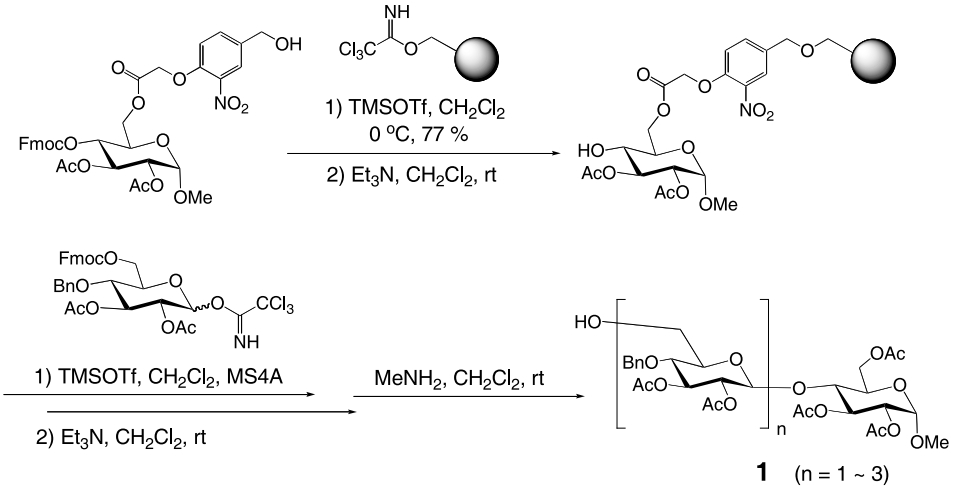
procedures. Extensive efforts therefore have been made to establish an efficient synthesis of oligosaccharides by means of a solid-phase method and a polymer-supported solution method. In solid-phase synthesis, the starting material binds to the polymer support, which allows the product to be isolated by filtration after each reaction. Finally, the product is cleaved from the resin. The operation is simple and rapid and hence solid-phase synthesis enables rapid preparation of a large number of compounds and thus increases the speed of development of a lot of functional molecules including new drugs and new materials. However, there have been several critical issues for solid-phase oligosaccharide synthesis, e. g., the synthetic intermediate cannot be purified. Each glycosylation and deprotection therefore must proceed in a high yield. Recent progress in glycosylation and protective group manipulation has solved these problems and made solid-phase synthesis of oligosaccharides practical [1]. In this chapter, the authors focus not only on solid-phase synthesis but also polymer-supported solution synthesis as well as tag-supported synthesis, both of which have been proved useful for rapid preparation of oligosaccharides.

2 Solid-Phase Methods for Synthesis of Oligosaccharides

2.1 New Linkers and Protection Groups for Solid-Phase Synthesis of Oligosaccharides

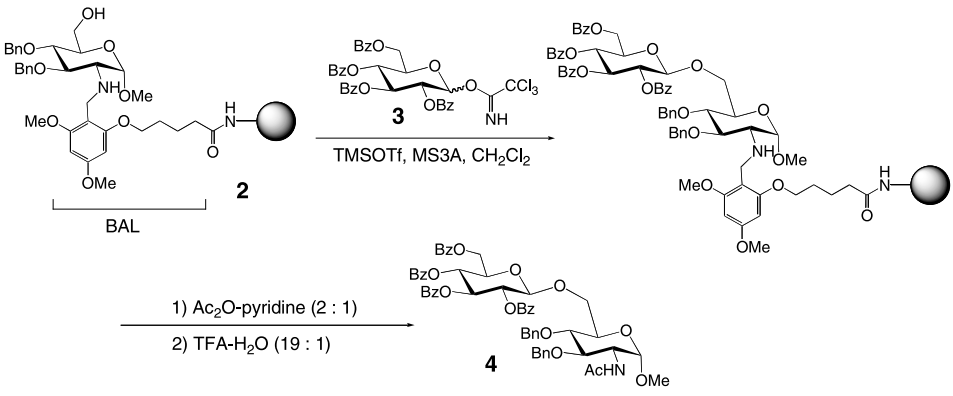
The choice of linkers is critically important for both solid-phase and polymer-supported solution synthesis. Selective cleavage of the linkers, leaving other protecting groups intact, and then purification and final deprotection is a rational strategy for oligosaccharides synthesis. This is distinctly different from the solid-phase synthesis of peptides, in which all the side-chain protecting groups and linkers are removed simultaneously at the end of the synthesis on the solid-supports. In the case of oligosaccharides, the purification in protected forms is generally more preferable owing to the hydrophilic properties of completely deprotected oligosaccharides. Traditionally, the linkers have been directly attached to the anomeric positions as thioglycoside, 1-*O*-acyl, photolabile *o*-nitrobenzyl glycoside, *p*-alkoxybenzyl glycoside, base-labile 9-hydroxymethylfluorene, or silylethyl glycoside [1]. On the other hand, ester-type, silyl ether-type, or *p*-acylaminobenzyl-type linkers have also been used to obtain the anomerically pure oligosaccharides after cleavage from the resins. More recently, unique linkers, which are bifunctional, temporary, and cleavable in a completely chemoselective manner through recently developed transition metal-catalyzed reactions, have been developed and successfully applied to the solid-phase synthesis of oligosaccharides. The development of such linkers enabled quite challenging oligosaccharide assembly on solid-supports, such as *N*-glycans or glycosylphosphatidylinositol of complex and branched structures, even using automated systems, which will be described in *Sect. 2.2*.

Schmidt and co-workers have developed the *o*-nitro-phenoxyacetate linker, which can be readily and quantitatively cleaved by treatment with methylamine under mild conditions (*Scheme 1*) [2]. By repeating the sequences of glycosylation with glucosyl trichloroacetimidate in the presence of TMSOTf, and Et_3N-induced deprotection of the Fmoc group, an efficient solid-phase synthesis of di- to pentasaccharides **1** containing glucose β(1–4) and β(1–6) linkages has been achieved.

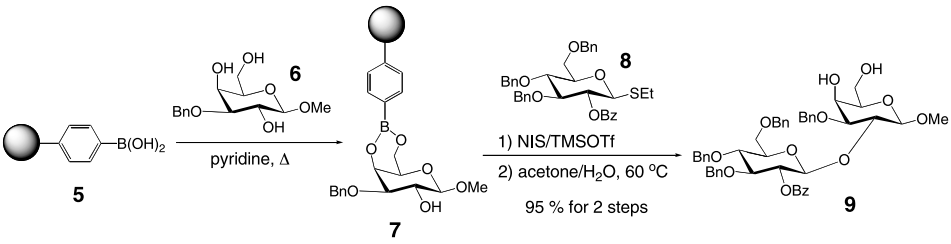


Scheme 1

Jensen and co-workers have developed the tris(alkoxy)benzylamine linker (BAL) on a polystyrene support, which is introduced to aminosugars by reductive amination of the corresponding aldehyde in the presence of $NaBH_3CN$ in DMF-AcOH (*Scheme 2*) [3,4]. This linker proved to be stable to concentrated TFA and other Lewis acids, but was readily cleaved by TFA-H_2O (19:1) after *N*-acetylation. Thus, polystyrene-bound glucosamine derivative **2** was glycosylated with glucosyl trichloroacetimidate **3** using TMSOTf as an activator. After the acetylation of the C2-amino group by Ac_2O in pyridine, β(1–6)-linked disaccharide **4** was liberated from the solid support in 82% yield by TFA treatment. The high stability of the linker under strongly acidic conditions allows for the use of excess Lewis acid for glycosylations on the solid-phase.



Scheme 2



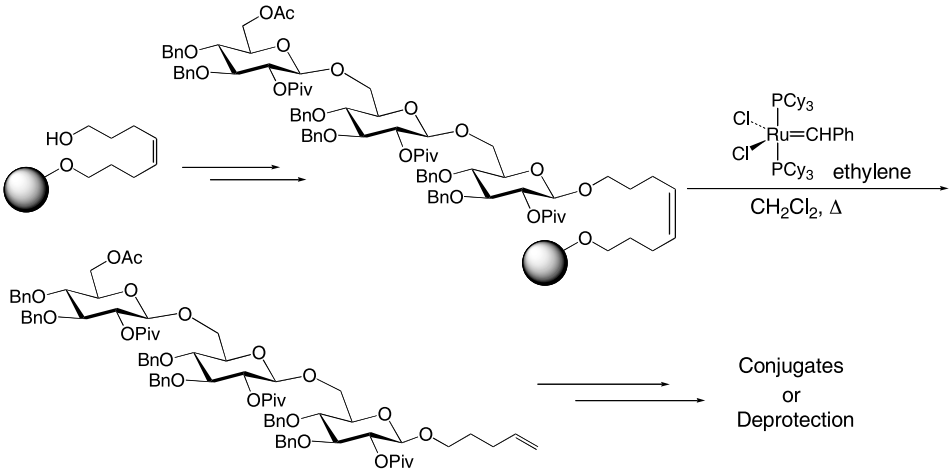
Scheme 3

The *N*-1-(4,4-dimethyl-2,6-dioxocyclohexylidene)ethyl linker has also been introduced between the C2-amino group of the glucosamine and a solid-support [5]. This linker was stable during glycosylation with thioglycoside in the presence of methyl trifluoromethane sulfonate, but readily cleaved by hydrazine, primary amines, or even ammonia.

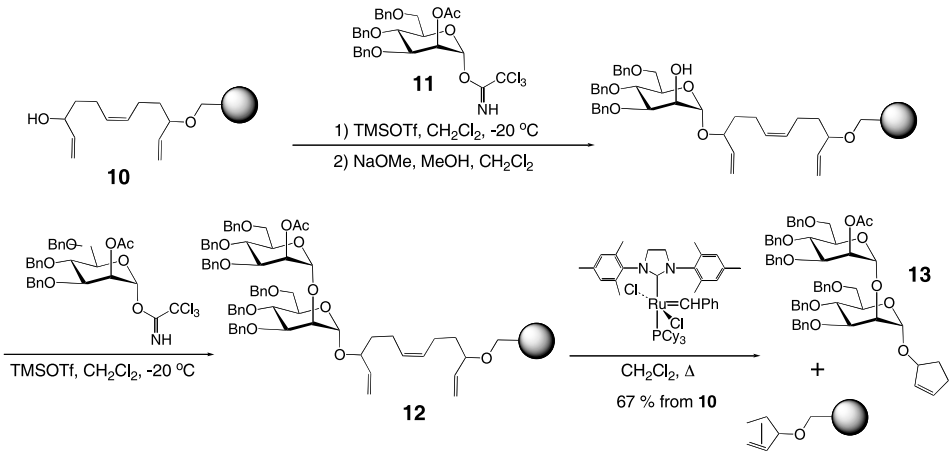
Boons and co-workers used polystyrylboronic acid as a temporary polymeric support for oligosaccharide synthesis (*Scheme 3*) [6,7]. Polystyrylboronic acid **5**, easily prepared from the corresponding bromide by treatment with trimethyl borate and *n*-BuLi in high loading capacity, was introduced on the 4,6-*O*-positions of the acceptor **6** by heating in pyridine. The immobilized boronate acceptor **7** was glycosylated with thioglycoside donor **8** by using NIS/TMSOTf as the promoter to give the β(1–2)-linked disaccharide **9** in 95% yield, after in situ cleavage from the solid support by heating in acetone/water. The polystyrylboronic acid was recovered and re-used for subsequent loading to the acceptors and glycosylation.

Recent exciting progress in the development of efficient and stable metathesis catalysts and their use in metathesis reactions both in solution-phase and solid-phase, have prompted several groups to investigate the metathesis-based linker for solid-supported synthesis. Olefin-based linkers have attracted a great deal of attention for solid-phase oligosaccharide synthesis since the isolated olefins are inert under a wide range of acid-catalyzed glycosylation and deprotection conditions, and at the end of the synthesis, these linkers can be cleaved under mild and olefin-selective conditions by both intramolecular and intermolecular metathesis. Seeberger and co-workers have applied the octanediol linker, which is readily prepared from cycloocta-1,5-diene, to solid-phase oligosaccharide synthesis, especially to programmed automated synthesis (see *Sect. 2.2*). The octanediol linker is cleaved from the resin by cross (intermolecular) metathesis with ethylene gas in the presence of the Grubbs' first generation catalyst to give *n*-pentenyl glycosides (*Scheme 4*) [8]. The *n*-pentenyl moiety thus generated can be used as a glycosylating donor or an anchor to connect the synthesized oligosaccharides to proteins, labels, or the surface of the solid support during preparation of oligosaccharides microarrays [9,10].

Van der Marel and co-workers have designed a self-cleavable linker via tandem ring-closing metathesis (*Scheme 5*) [11]. Their triene linker on Merrifield resin **10** was glycosylated twice with mannosyl trichloroacetimidate **11** in the presence of TMSOTf in CH_2Cl_2 at $-20\,^{\circ}C$ to provide the α-linked disaccharide **12**, which was treated with Grubbs' second generation ruthenium catalyst to liberate the cyclopent-2-enyl mannnoside **13** in 67% yield from **10** (average 90% for each step). A cyclopent-2-enyl moiety was then isomerized to the corresponding vinyl ether glycoside by treatment with $(PPh_3)_3RhCl$ and DBU in EtOH, which was hydrolyzed by



Scheme 4

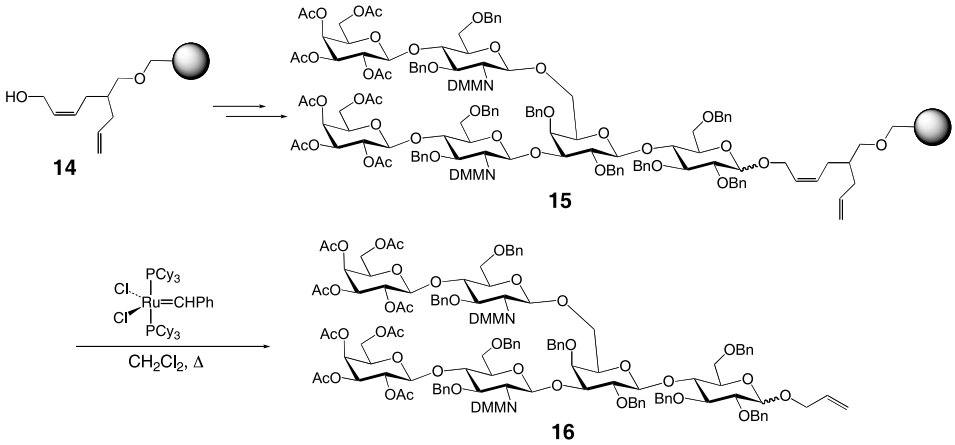


Scheme 5

I_2 in THF-H_2O. The self-cleavable linker **10**, which provides the thermodynamically stable cyclopentene both on solid supports and on the products, prevents catalyst deactivation due to covalent attachment of the ruthenium catalyst to either the solid-support or the products via C=C double bonds during the metathesis reaction.

Schmidt and co-workers have also independently developed the self-cleavable diene linker **14** and solid-phase synthesis of the branched hexasaccharide **16** has been achieved (*Scheme 6*) [12]. The ring-closing metathesis of **15** provided the allyl glycoside **16** from which the allyl protecting group was readily removed by Ir-mediated isomerization followed by hydrolysis.

Fukase and co-workers have developed two types of alkyne linkers (*Scheme 7*) [13]. A propargyl glycoside-type linker between a sugar residue and a solid support was formed

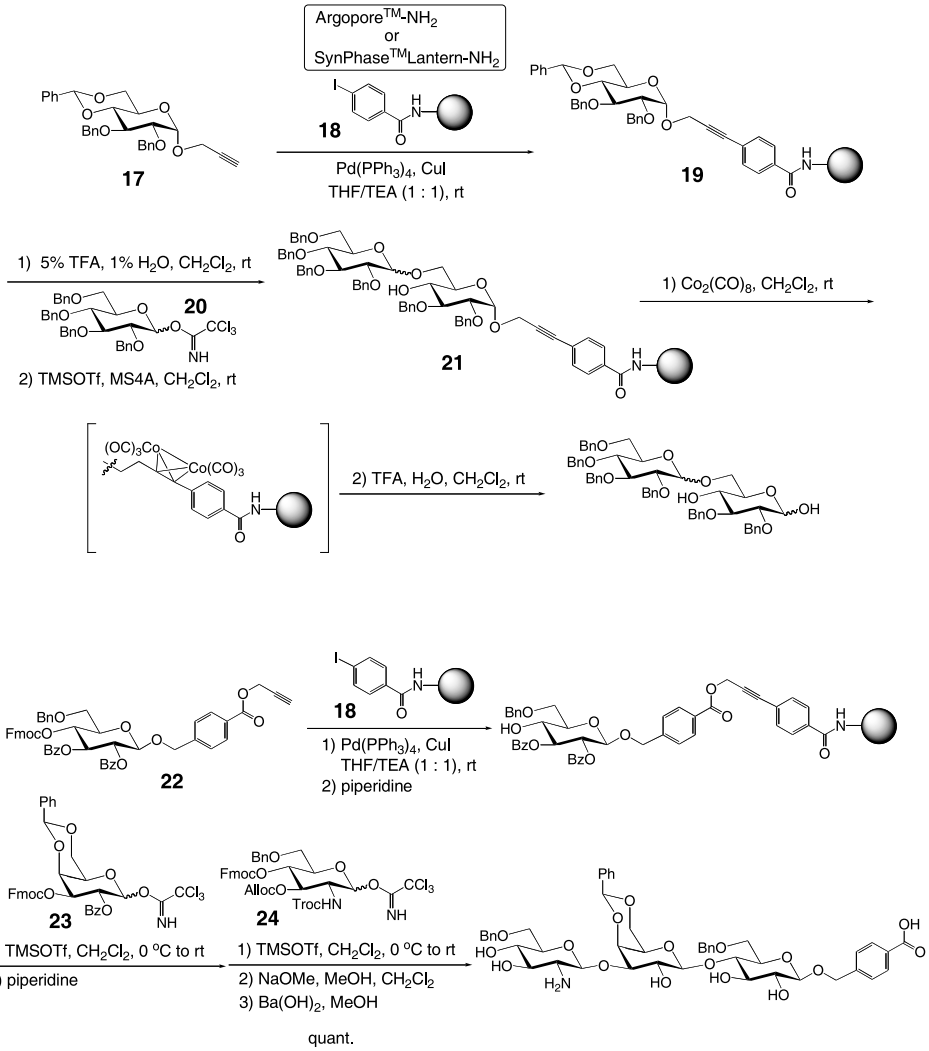


Scheme 6

by the Sonogashira reaction; the propargyl glycoside **17** was treated with polymer-supported iodobenzene **18** in the presence of $Pd(PPh_3)_4$ and CuI in THF/Et_3N to provide **19**. After the 4,6-*O*-benzylidene group was removed by treatment with TFA, the liberated C6-hydroxyl was glycosylated with glucose trichloroacetimidate **20** using TMSOTf as a promoter to give the disaccharide **21**. The alkyne-based linker is thus stable during the glycosylation or the TFA treatment, but can be readily cleaved with TFA after the conversion to the corresponding alkyne-cobalt complex by reacting with $Co_2(CO)_8$. Alternatively, a propargyl ester-type linker was also generated by the same Sonogashira reaction of a 4-(propargyloxycarbonyl)benzyl glycoside **22** with the polymer-supported iodobenzene **18**. After a sequence of glycosylation reactions with imidates **23** and **24**, the trisaccharide was released from the resin by treatment with TFA of the corresponding alkyne-cobalt complex. The latter ester-type linker is generally advantageous since the carboxybenzyl glycoside, liberated from the resin under mildly acidic conditions, is readily purified and the linker can also be cleaved under nucleophilic conditions (NaOMe in MeOH/CH_2Cl_2). Noteworthy is that the Sonogashira reaction was found to proceed only at spatially reactive sites on the solid support to which the reagents can access readily, so that the subsequent reactions including glycosylation and Fmoc deprotection on the solid-phase proceeded smoothly and quantitatively, resulting in high total yields of the desired oligosaccharides.

The peptide-based linker has also been investigated by Warriner and co-workers (*Scheme 8*) [14]. Since the solid-phase synthesis of peptide is well established, both glycosyl acceptors and donors were initially conjugated to the 3-hydroxyproline template, and introduced on the Aminomethyl Novagel™ resin using standard Fmoc peptide coupling protocol (a sequence of amide formation by using HOBt, HBTU, and DIPEA, followed by Fmoc deprotection by piperidine). They examined the stereoselectivities of glycosylation between acceptors and donors that were loaded to a peptide template, and 12 members of the disaccharide library were constructed.

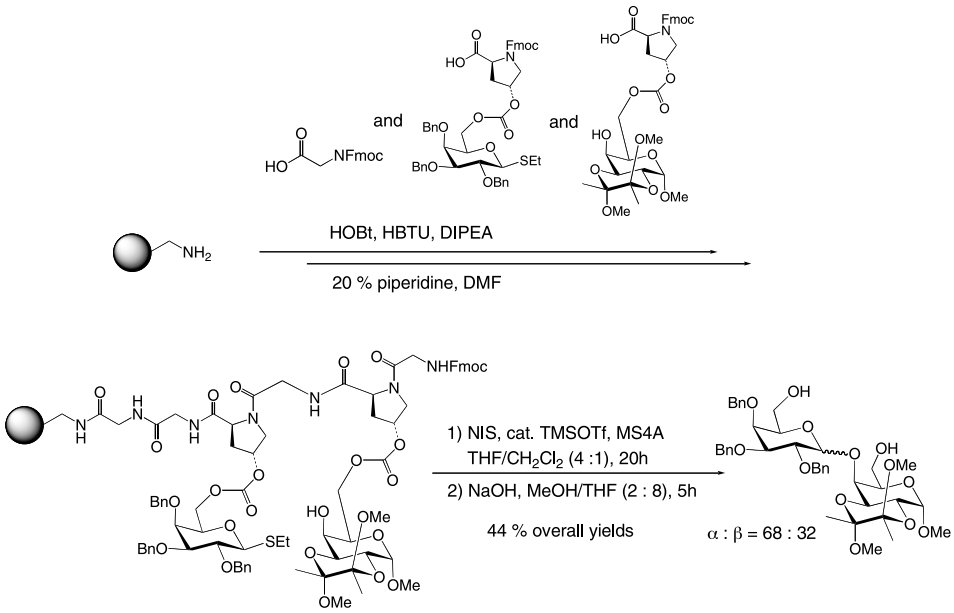
Based on a library-directed approach, Takahashi and co-workers developed a synthetic strategy for the combinatorial solid-phase synthesis of tree-type carbohydrate clusters using three types



Scheme 7

of orthogonally cleavable linkers, namely, a photo-cleavable 2-nitrobenzyl-type linker (orthogonal cleavage site 1), a base-cleavable ester linker (orthogonal cleavage site 2), and a cobalt complexation/acid-cleavable alkyne linker (orthogonal cleavage site 3) (*Fig. 1*) [15]. The combined use of three types of spacers, two types of carbohydrates, and two types of orthogonally cleavable linkers provided 12 members of a carbohydrate cluster library (6 dimers and 6 tetramers).

In addition to the importance of linkers described above, the precise choice of protecting groups is also an important factor for oligosaccharide assembly to be successful on solid-supports. Utilization of protecting groups that are stable under glycosylation conditions but can



Scheme 8

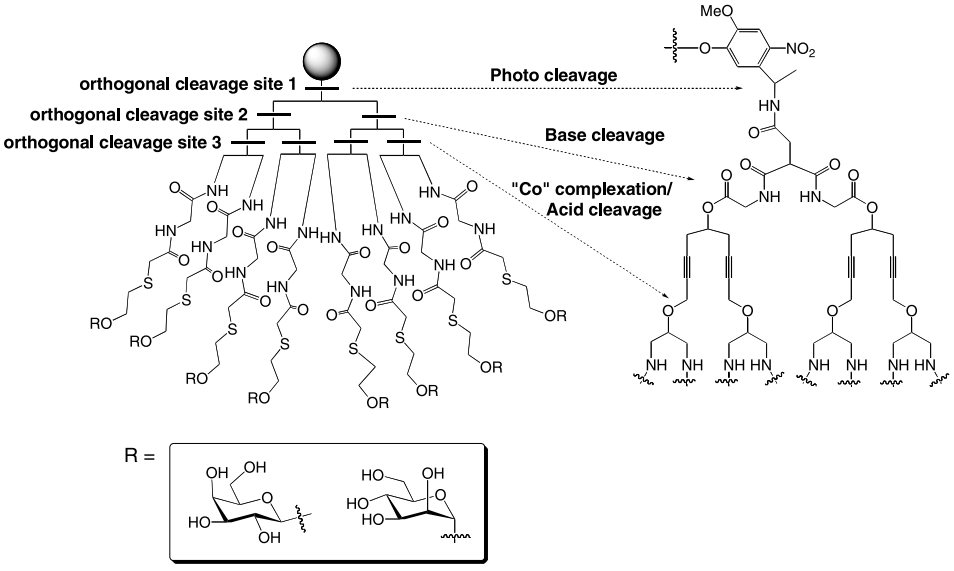
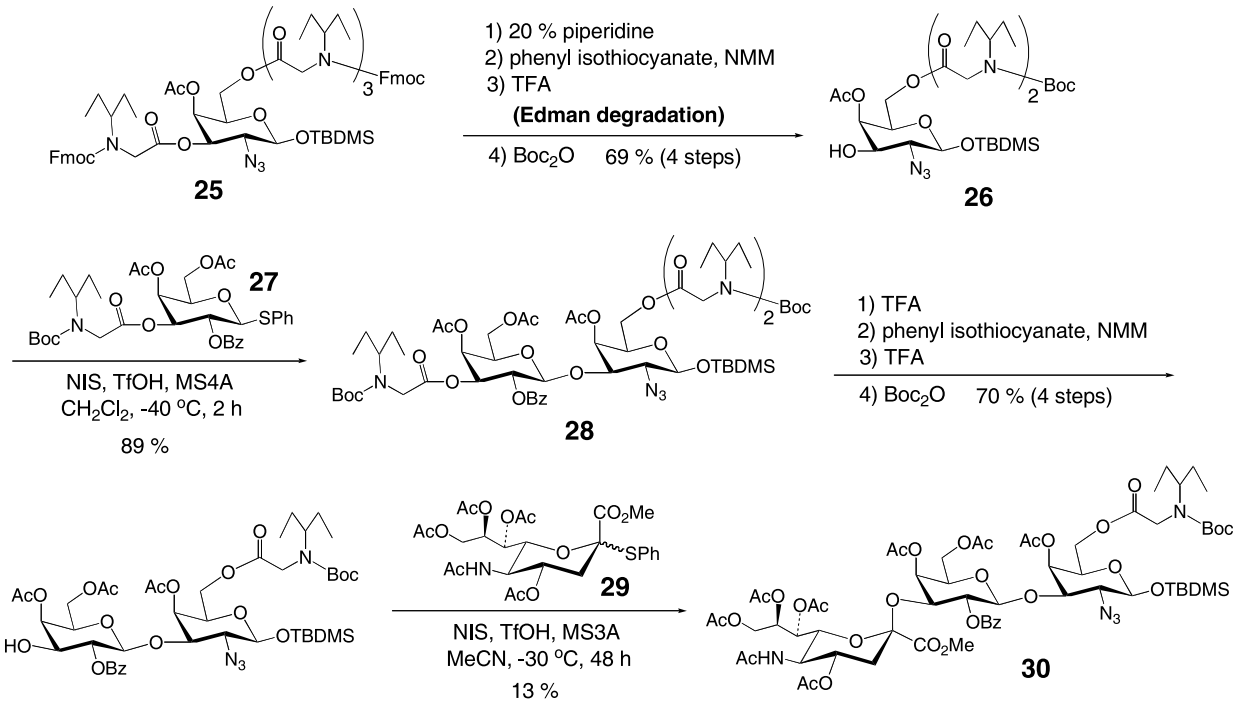


Figure 1
Combinatorial solid-phase synthesis of tree-type carbohydrate clusters


1) 20 % piperidine
2) phenyl isothiocyanate, NMM
3) TFA
(Edman degradation)
4) Boc2O 69 % (4 steps)
25
26
27
28
1) TFA
2) phenyl isothiocyanate, NMM
3) TFA
4) Boc2O 70 % (4 steps)
NIS, TfOH, MS4A
CH2Cl2, -40 °C, 2 h
89 %
29
30
NIS, TfOH, MS3A
MeCN, -30 °C, 48 h
13 %

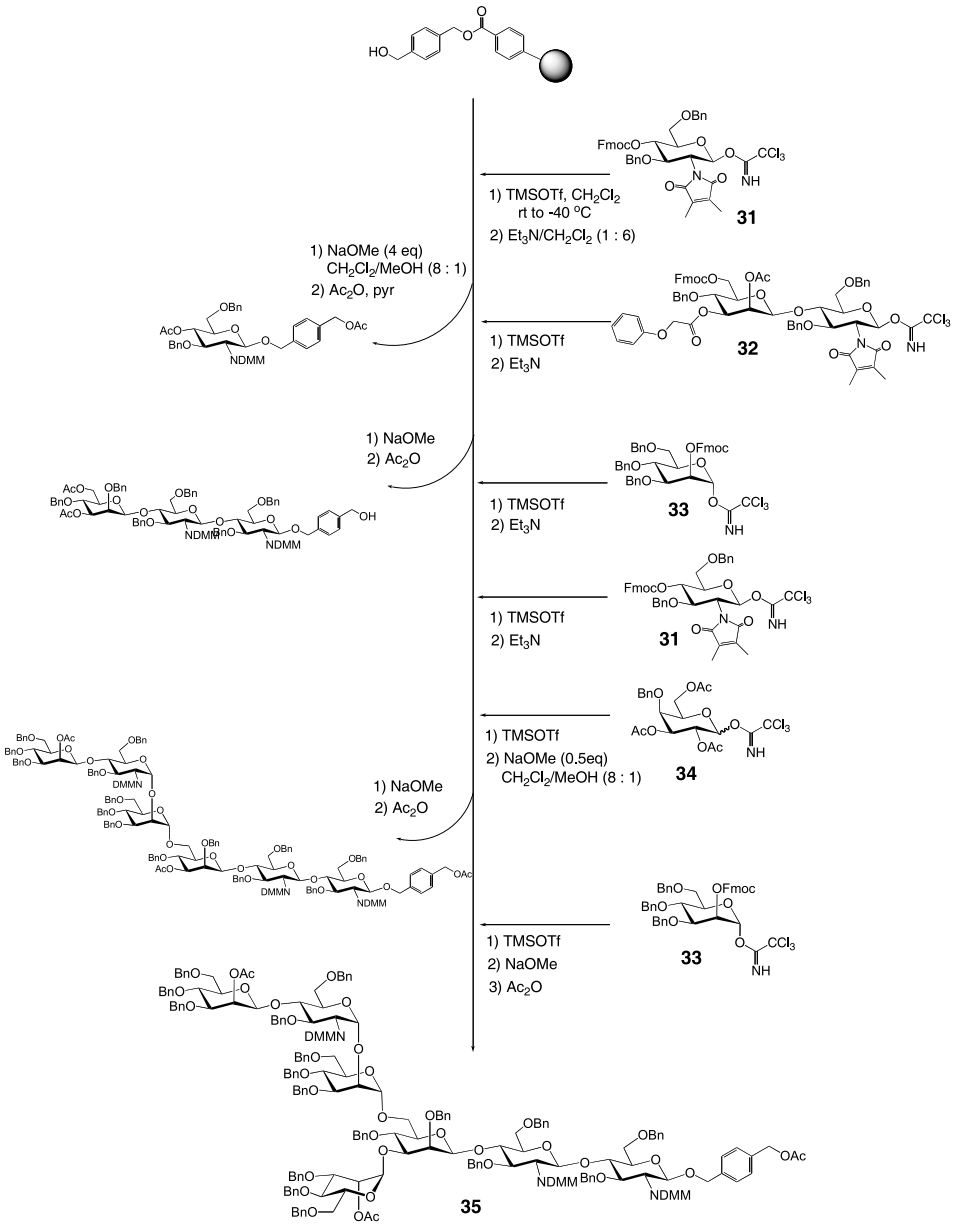
Scheme 9

be readily removed under mild conditions and that are compatible to other protecting groups and linkers is highly desirable for combinatorial and high-throughput oligosaccharides synthesis. The excellent orchestration of protecting group manipulation, leading to the construction of complex oligosaccharide structures will be discussed in *Sect. 2.2*. It is also preferable if the protecting group serves as a tag to identify the compounds as well as to indicate their protection states. Kasumi and co-workers have reported unique protecting group manipulations, by using (N-α-1-ethylpropylglycine)$_n$ [(EPG)$_n$], for both solution- and solid-phase synthesis of oligosaccharides (*Scheme 9*) [16]. The hydroxyl groups of oligosaccharides were protected by an (EPG)$_n$ group with a different degree of polymerization, which can be uniquely identified. The (EPG)$_n$ groups were then deprotected successively from the highest to the lowest degree of polymerization by repeating the Edman degradation cycle. Thus, the galactose derivative **25** with (EPG)$_3$ at the C6-hydroxyl and (EPG)$_1$ at the C4-hydroxyl, was subjected to the Edman degradation, i. e., by (i) Fmoc deprotection with 20% piperidine, (ii) phenyl isothiocyanate treatment, (iii) degradation by TFA, and (iv) Boc protection of the newly appearing amino group to provide the acceptor **26** having (EPG)$_2$ at the C6-hydroxyl, in 69% yield over four steps. This acceptor was glycosylated with thiogalactose **27** bearing one EPG at the C3-position by using NIS and TfOH as the activators to give **28** in 89% yield. Repetition of the Edman degradation cycle liberated the C3′-hydroxyl in 70% yield, which was further sialylated with thioglycoside **29** to give **30** in 13% yield. They applied this protecting group manipulation to solid-phase oligosaccharide synthesis.

2.2 Solid-Phase Synthesis of Complex Oligosaccharides

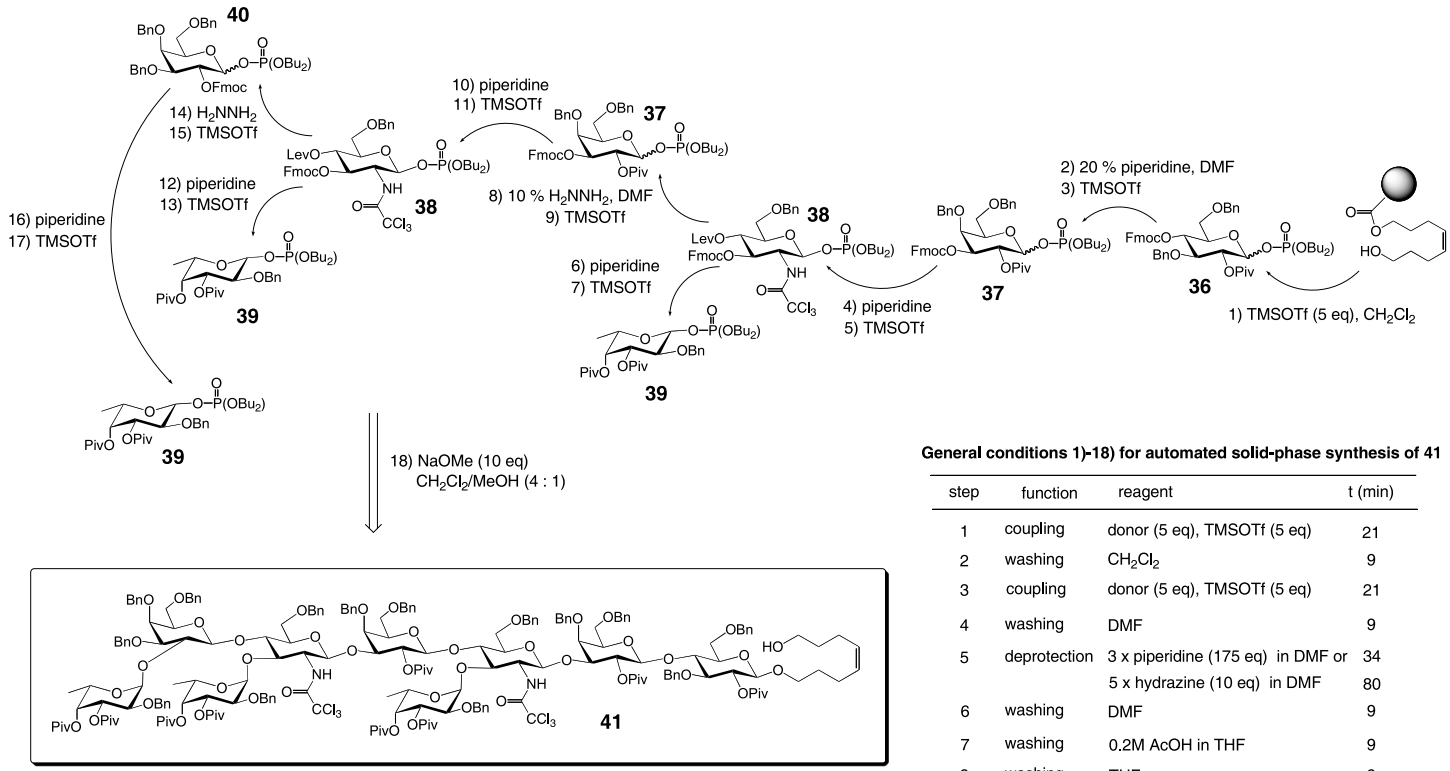
New methods for solid-supports, together with the accumulation of information on acceptor/donor reactivity in the solution-phase, have led to the synthesis of complex oligosaccharides on solid-supports [17,18,19,20,21]. Schmidt and co-workers have achieved solid-phase synthesis of a small library of N-glycans (*Scheme 10*) [22,23,24]. They used the hydroxymethylbenzyl benzoate linker attached to the Merrifield resin. Glycosylation on the resin was successfully performed using the glycosyl trichloroacetimidates [25], such as glucosamine **31**, mannosylβ(1–4)glucosamine **32**, mannose **33**, and galactose **34** which allow chain extension, branching, and chain termination for the synthesis of heptasaccharide N-glycan **35**; the tedious β-mannosyl linkage in **32** was constructed in advance in solution phase. O-Benzyl and O-benzoyl groups were used as permanent, while O-Fmoc and O-phenoxyacetyl groups were used as temporary protecting groups for the sugar extension on the solid-phase. The N-dimethylmaleoyl group was selected for the protection of glucosamine **31**, since not only the neighboring effects ensure β-selective glycosylation, but also this protecting group confers high reactivity on glucosamine **31** as an acceptor on the polystyrene [26]. Their solid-phase synthesis involves (i) glycosylation under the TMSOTf catalyst, (ii) selective removal of the temporary protecting groups, Fmoc by Et_3N and the phenoxyacetyl group by 0.5 equivalents of NaOMe in MeOH/CH_2Cl_2, and (iii) cleavage of products from the resin by 4.0 equivalents of NaOMe in MeOH/CH_2Cl_2 and then O-acetylation for convenient product isolation. These precise designs of linker, protecting group manipulations, and stereoselective glycosylation on the resin enabled the successful preparation of a small library of 17 N-glycan structures.

The assembly of oligosaccharides on solid-supports gave a quite exciting and promising opportunity for automation. Seeberger and co-workers have designed the first automated oligosac-



Scheme 10

charide synthesizer for the solid-phase (*Scheme 11*) [27,28,29,30,31,32,33,34,35,36]. A peptide synthesizer was re-engineered to function as an automated oligosaccharide synthesizer and to provide easy access to structures as large as dodecasaccharides about 20-times faster than previously reported manual methods. Even oligosaccharides with branched struc-



General conditions 1)-18) for automated solid-phase synthesis of 41

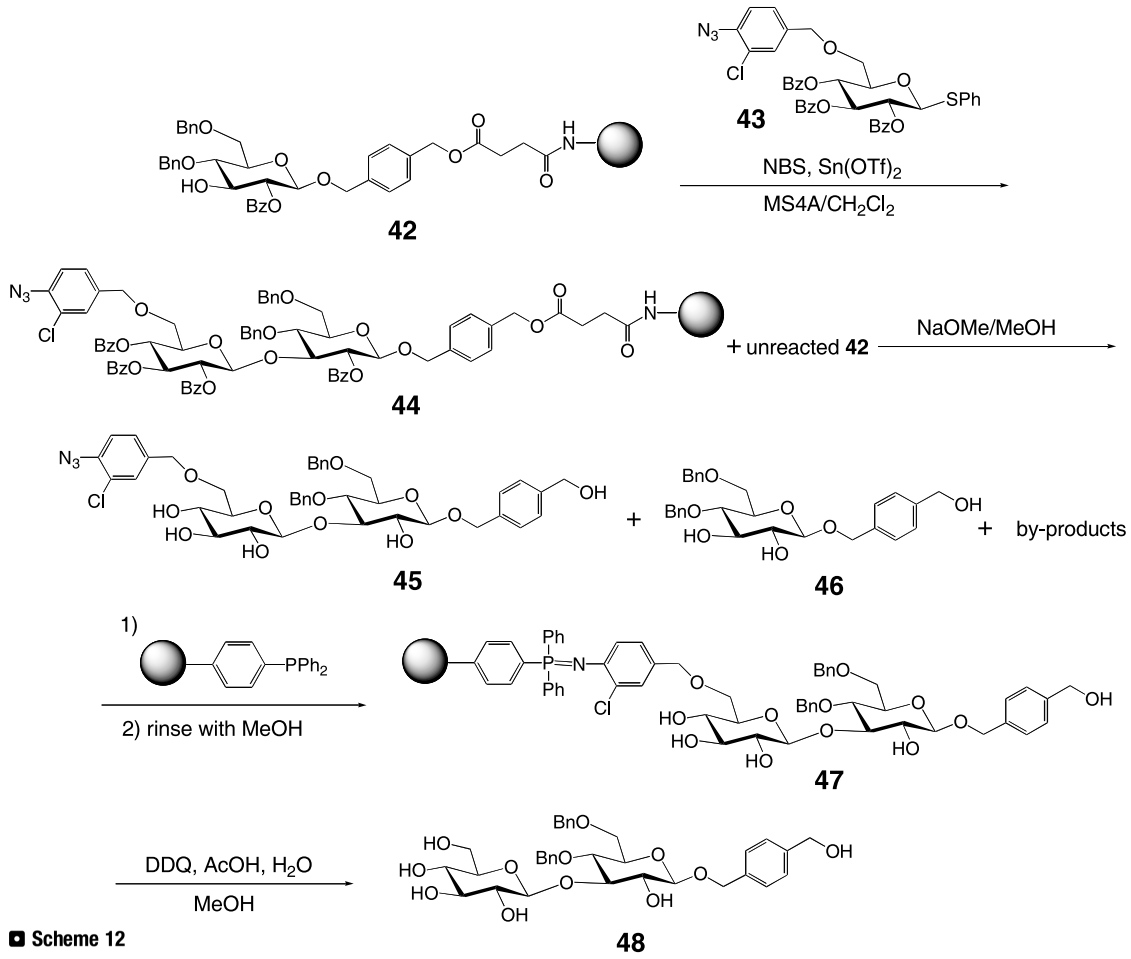
step	function	reagent	t (min)
1	coupling	donor (5 eq), TMSOTf (5 eq)	21
2	washing	CH_2Cl_2	9
3	coupling	donor (5 eq), TMSOTf (5 eq)	21
4	washing	DMF	9
5	deprotection	3 x piperidine (175 eq) in DMF or 5 x hydrazine (10 eq) in DMF	34 80
6	washing	DMF	9
7	washing	0.2M AcOH in THF	9
8	washing	THF	9
9	washing	CH_2Cl_2	9

Scheme 11

tures are now accessible and a series of oligosaccharides of biological relevance, i. e., glycosylphosphatidylinositol (GPI) [37], *N*-glycan core pentasaccharide [38], or Lewis blood group oligosaccharides [39], have been prepared. The solid-phase synthesizer utilizes the glycosyl phosphate building blocks and the octanediol linker attached to the resin which allow for a simple automated coupling/deprotection cycle for oligosaccharide assembly. The released oligosaccharides from solid-supports were applied for biological purposes, i. e., bioconjugation or a microarray [9,10]. An example of the automated synthesis of Lewis blood group oligosaccharide **41** is shown in *Scheme 11* [39]. Five monomer units **36–40**, which allow for the stereoselective construction of five different glycosidic bonds in **41**, were designed for the automated synthesis of **41**. Especially, fucosyl phosphate **39** was selected based on the solution study; this donor showed excellent selectivity in constructing the difficult α(1–2)-*cis* and α(1–3)-*cis* fucose linkages common to the Lewis antigens. The octanediol linker was attached to the resin by an ester linkage that is cleaved by nucleophilic bases such as NaOMe. Fmoc was used as the temporary protecting group for hydroxyls not only because it is quite stable under the glycosylation conditions and easily cleaved by the non-nucleophilic amines, but also it facilitates the qualitative evaluation of the efficiency for each glycosylation and deprotection cycle by monitoring the UV/Vis of the dibenzofurubene derivative after deprotection; this monitoring method is especially valuable for the automated synthesis. For the branching linkage of **38**, levulinoyl ester, being cleaved by hydrazine treatment, was chosen as another temporary protection group for elongation of sugar-chains. General automated cycles involve (i) glycosylation by activating the phosphate donors with TMSOTf in CH_2Cl_2 at $-15\,^{\circ}C$ and (ii) selective removal of temporary protecting groups, Fmoc by 20% piperidine in DMF, and the levulinoyl group by 10% hydrazine in DMF, as shown in *Scheme 11*. If the glycosylation was not completed by the dibenzofurubene UV analysis, another cycle of glycosylation was performed, i. e., double glycosylation. Importantly, a series of washing steps was performed after each deprotection to swell the resin and purge any remaining basic materials for the next glycosylation cycles. At the end of the synthesis, the product was cleaved from the resin by treatment with 10 equivalents of NaOMe in MeOH/CH_2Cl_2, and by this automated method, **41** was produced in 6.5% overall yield in as short as 23 hours. Further improvements of the automated synthesis might allow everyone, even nonchemists, to prepare complex and biologically relevant oligosaccharides, similar to the peptide and nucleotide synthesizers, which will speed up the elucidation of their biological functions as well as clinical applications of oligosaccharide-based vaccines.

2.3 Solid-Phase Methods for Purification of Synthesized Oligosaccharides

The isolation of products is sometimes a tedious and time-consuming procedure. Especially in the area of combinatorial synthesis of oligosaccharides either in solution or on solid-phase, a simple and effective isolation method is necessary because many compounds are handled at the same time. For instance, chromatography is a quite popular method for isolation of the products of organic synthesis. However, it is somewhat tedious and consumes a considerable amount of solvents, and optimization of the separation conditions sometimes requires much time. Therefore, techniques for polymer-supported purification such as “catch-and-release”


43
NBS, Sn(OTf)2
MS4A/CH2Cl2
42
+ unreacted 42
NaOMe/MeOH
44
45
+
+ by-products
46
1)
2) rinse with MeOH
47
DDQ, AcOH, H2O
MeOH
48

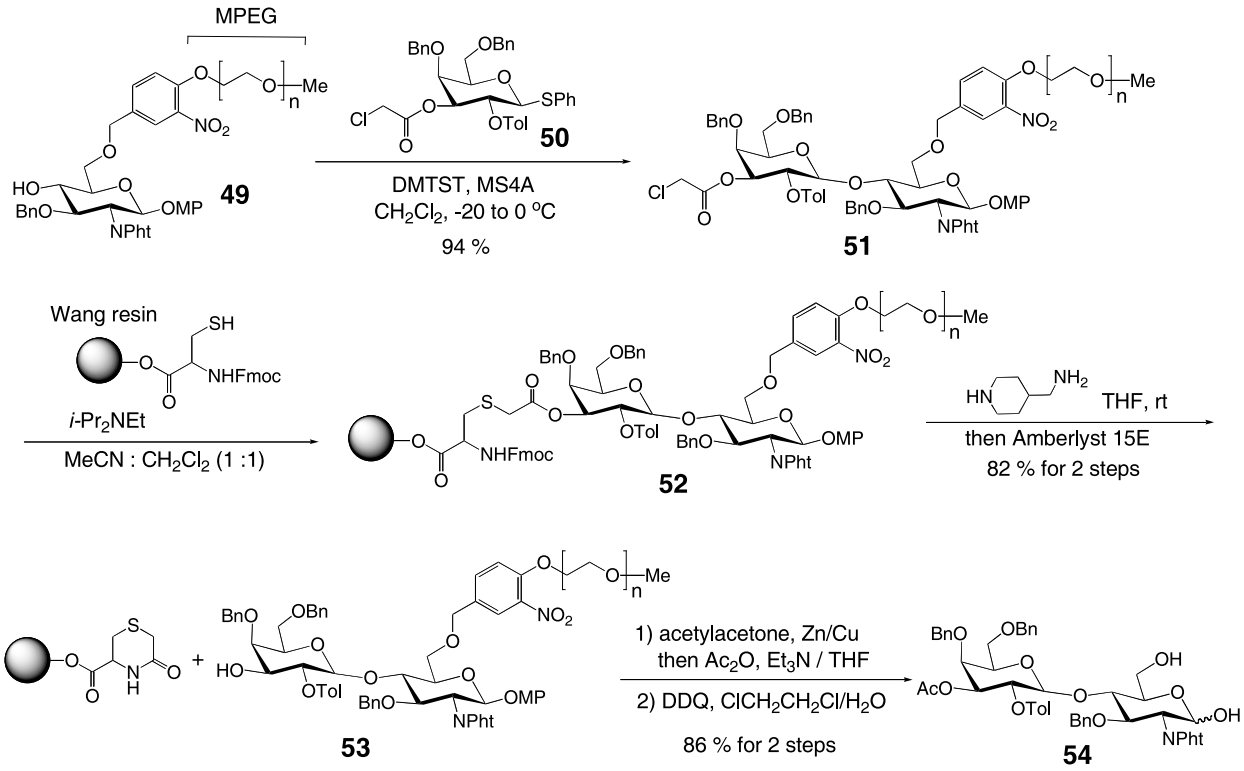
Scheme 12

purification methods, polymer-bound reagents, and scavenger resins, have been developed. Fukase and co-workers have developed the 4-azido-3-chlorobenzyl (ClAzb) group not only as a safety catch-type protection group of hydroxyls but also as a tag for catch-and-release purification (*Scheme 12*) [40,41]. A desired compound possessing the ClAzb group can be selectively caught by a solid-supported phosphine by the specific reaction between the azido function and the phosphine, and the product on the polymer-support was separated from the other compounds simply by rinsing. The desired compound was then released by treatment with DDQ (*Scheme 12*). A glucose acceptor bound to a macroporous polystyrene (ArgoPore^{TM}-NH_2) **42** was glycosylated with thioglycoside **43** using NBS and $Sn(OTf)_2$ in the presence of MS4A to give a mixture of disaccharide **44** and unreacted **42**. The resin was treated with NaOMe in MeOH to cleave the linker, and a mixture of disaccharide **45**, monosaccharide **46**, and other by-products was obtained. The mixture was treated with triphenylphosphine-(polyethyleneglycol-polystyrene-copolymer) resin in MeOH and the resulting solid-supported disaccharide **47** was treated with DDQ. With this procedure, a highly pure disaccharide **48** was obtained in 38% yield from **42**.

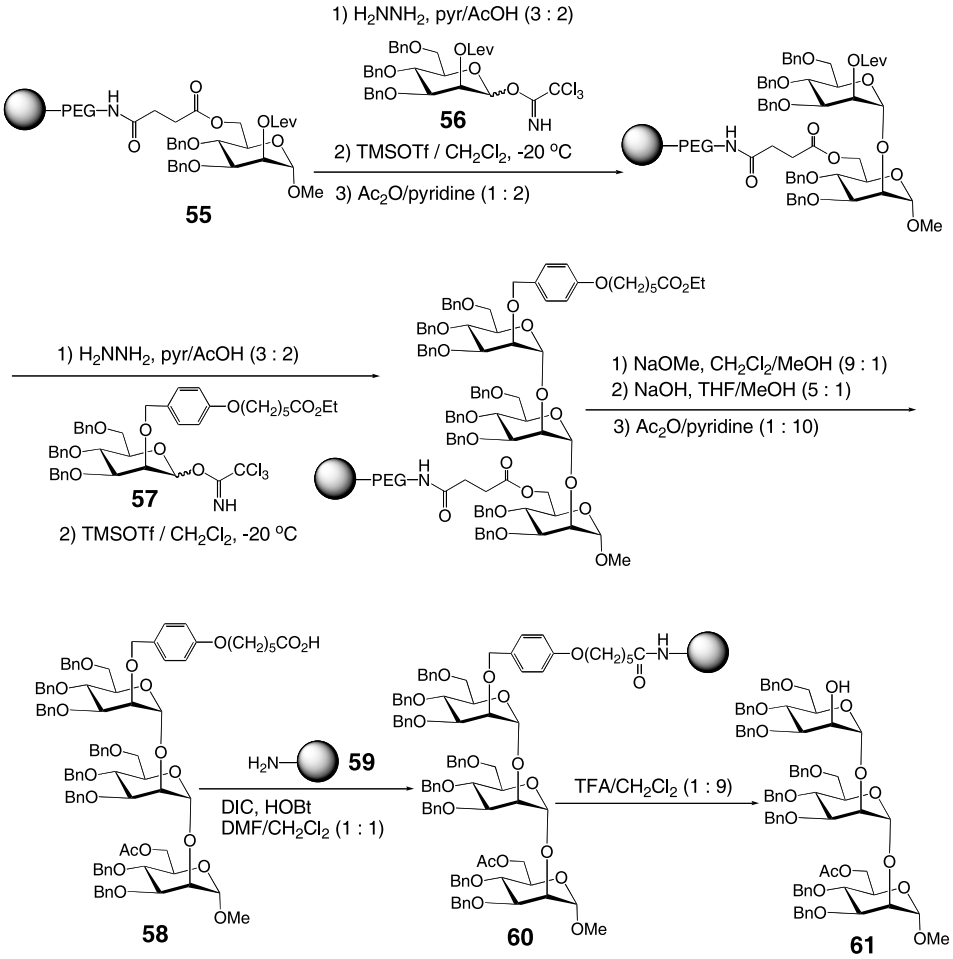
Ito and co-workers have combined soluble polymer-supported synthesis with the catch-and-release protocol using the chloroacetyl (ClAc) tag (*Scheme 13*) [42,43,44]. A glucosamine acceptor bound to a low molecular weight poly(ethylene)glycol (PEG) **49** was glycosylated with a thiogalactoside donor **50** in the presence of dimethyl(methylthio)sulfonium triflate (DMTST) in CH_2Cl_2 to provide the disaccharide **51** in 94% yield. **51** was then successfully captured by the Fmoc-cystein-loaded Wang resin at the ClAc tag in the presence of diisopropylethylamine. After washing the resin, **52** was released from the resin by Fmoc-deprotection with 4-(aminomethyl)piperidine followed by spontaneous intramolecular cyclization to afford **53** with excellent purity (82% for two steps). The reduction of the nitro group of the PEG linker to amine by reaction with acetylacetone and Zn/Cu, followed by acetylation and DDQ oxidation gave the disaccharide **54** in 86% yield for two steps.

The idea of "catch-and-release" was applied to the solid-phase synthesis of α-linked trimannose **61** by Guo and co-workers (*Scheme 14*) [45]. The mannose acceptor **55**, loaded on the PEG-grafted polystyrene via an amide linkage, was glycosylated twice with the mannosyl trichloroacetimidate donors by using TMSOTf as an activator in CH_2Cl_2, i. e., firstly with **56** and secondly with **57** bearing the ester (precursor of carboxylate) group as a tag. After cleavage from the resin and hydrolysis of the ester-tag under basic conditions (NaOMe, CH_2Cl_2/MeOH, and then NaOH, THF/MeOH), a trimannose derivative **58** was efficiently caught out from mono-, di-, and other by-products by the amino-functionalized "fishing" resin **59** by acylation. The pure compound **61** was then released from the solid-support by the treatment of **60** with TFA.

On the other hand, an efficient polymer-assisted method for deprotection of the protected oligosaccharides has been demonstrated by Takahashi and co-workers (*Scheme 15*) [46]. The deprotection of the synthesized oligosaccharides, including the cleavage of various *O*-protecting groups and the replacement of *N*-protecting groups with an *N*-acetyl group is sometimes troublesome under the standard conditions. Furthermore, the complete deprotection of the synthesized oligosaccharides frequently requires careful selection of the reaction solvents to prevent the partially deprotected intermediates from precipitating. They first loaded the synthesized oligosaccharide **62** to the aminomethyl ArgoPore through a tetrahydropyranyl (THP) linker. The protected oligosaccharide on polymer-support **63** thus obtained was subject-


MPEG
49
50
DMTST, MS4A
CH_2Cl_2, -20 to 0 °C
94 %
51
Wang resin
i-Pr_2NEt
MeCN : CH_2Cl_2 (1 :1)
52
THF, rt
then Amberlyst 15E
82 % for 2 steps
53
1) acetylacetone, Zn/Cu
then Ac_2O, Et_3N / THF
2) DDQ, $ClCH_2CH_2Cl/H_2O$
86 % for 2 steps
54

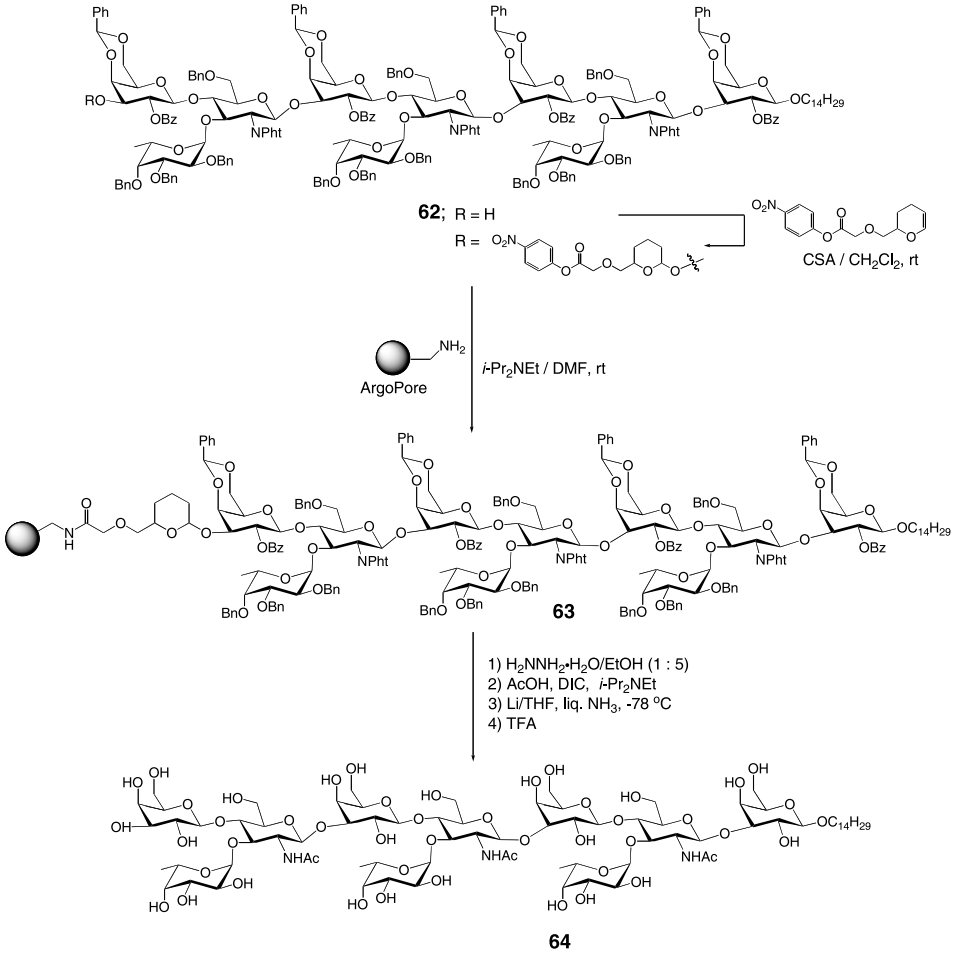
Scheme 13



Scheme 14

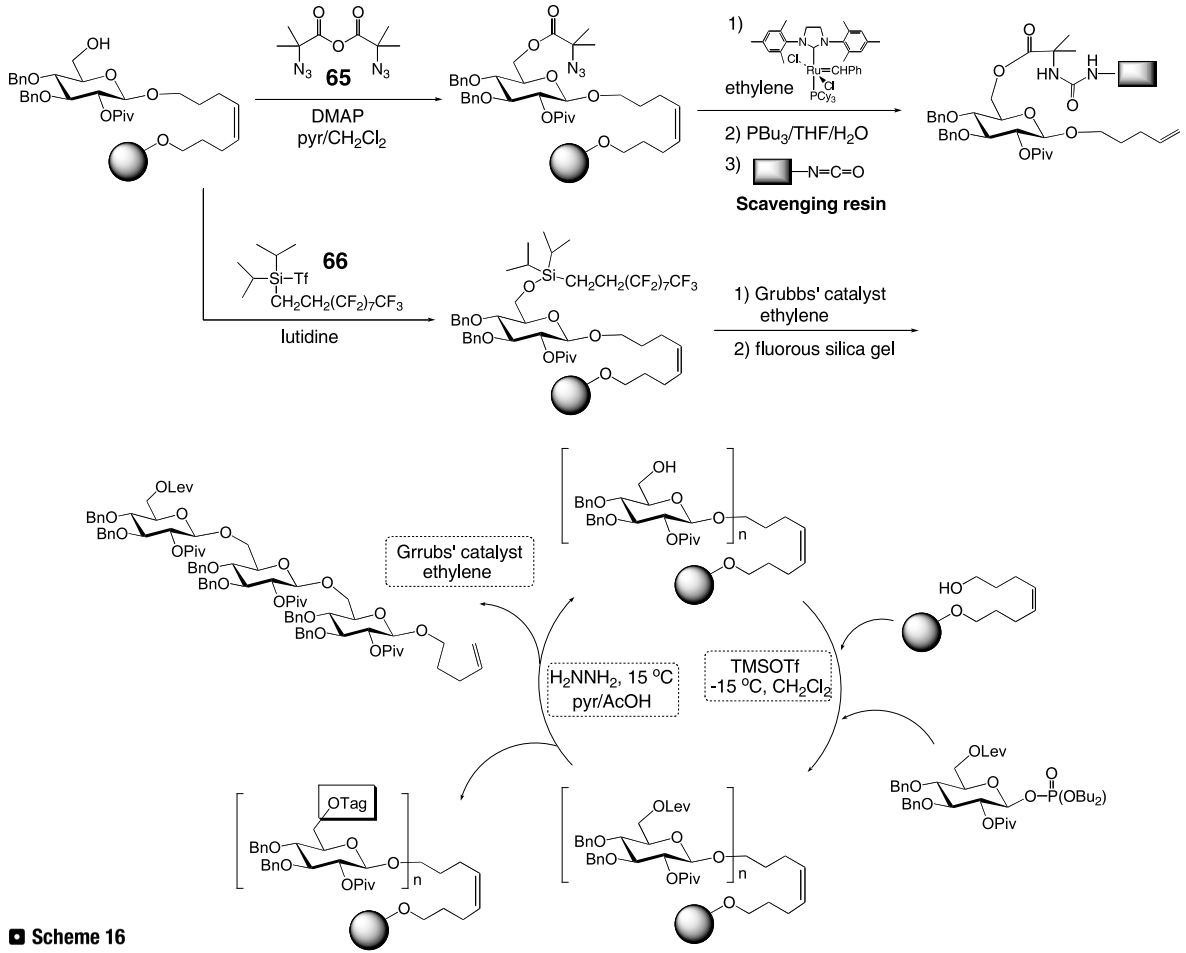
ed to (i) deprotection of *N*-phthalimide groups by hydrazine, (ii) acetylation of the resulting free amino groups, (iii) removal of the benzyl groups by Birch reduction, and (iv) deprotection of 4,6-*O*-benzylidene acetal groups and simultaneous cleavage from the THP linker by TFA, to provide the fully deprotected trimeric Lewis X epitope **64** in 58% overall yield from **62**. Although loading and release of the synthesized oligosaccharides might be accompanied by slight loss of the materials, this method allows the practical library synthesis of complex oligosaccharides.

The polymer-supported methods for removal of the unreacted acceptors during the glycosylation process also provide rapid access to automated synthesis and the parallel library synthesis of oligosaccharides. Ley and co-workers applied the polymer-supported tosyl chloride (PS-TsCl) to remove the hydroxyl-containing contaminants, including the acceptors [47].



Scheme 15

The unreacted acceptor was also sequestered and recovered by Dondoni and Massi based on catch-and-release protocols [48]. By this method, the hydroxyls of the unreacted acceptors were reacted with the trichloroacetyl isocyanate under neutral conditions, and the resulting trichloroacetyl carbamates (urethanes) were extracted by the highly basic and non-nucleophilic polymer-supported BEMP (2-*tert*-butylimino-2-diethylamino-1,3-dimethylperhydro-1,3,2-diazaphosphorine on polystyrene) by making use of the strong ionic interactions. The pure acceptor was released from the resin by KOH treatment and used again for glycosylation. Seeberger and co-workers have successfully used two capping methods to remove the unreacted acceptors during their automated synthesis of oligosaccharides (*Scheme 16*, and see *Sect. 2.2*) [49]. Both 2-azido-2-methylpropionic acid anhydride **65** (A-tag) and (heptadecafluorodecyl)diisopropylsilyl triflate **66** (F-tag) were applied to cap the hydroxyls of unre-


65
DMAP
pyr/CH2Cl2
ethylene
2) PBu3/THF/H2O
3)
Scavenging resin
66
lutidine
1) Grubbs' catalyst
ethylene
2) fluorous silica gel
Grrubs' catalyst
ethylene
H2NNH2, 15 °C
pyr/AcOH
TMSOTf
-15 °C, CH2Cl2

Scheme 16

acted glycosyl acceptors during each incomplete glycosylation step. At the end of the synthesis, the tagged compounds were efficiently removed after the cleavage from the resin by cross metathesis with ethylene. For the A-tag, the azide moiety was reduced by tributyl phosphine and the resulting amino-derivative was fished out by an isocyanate silica gel scavenging resin. On the other hand, the F-tagged compounds were removed by column chromatography packed with tridecafluoro($Si(CH_2CH_2C_6F_{13})_3$)-functionalized silica gel. By this method, they successfully simplified the purification process for the automated synthesis of β-linked (1–6)-triglucose and α-linked (1–2)-mannose.

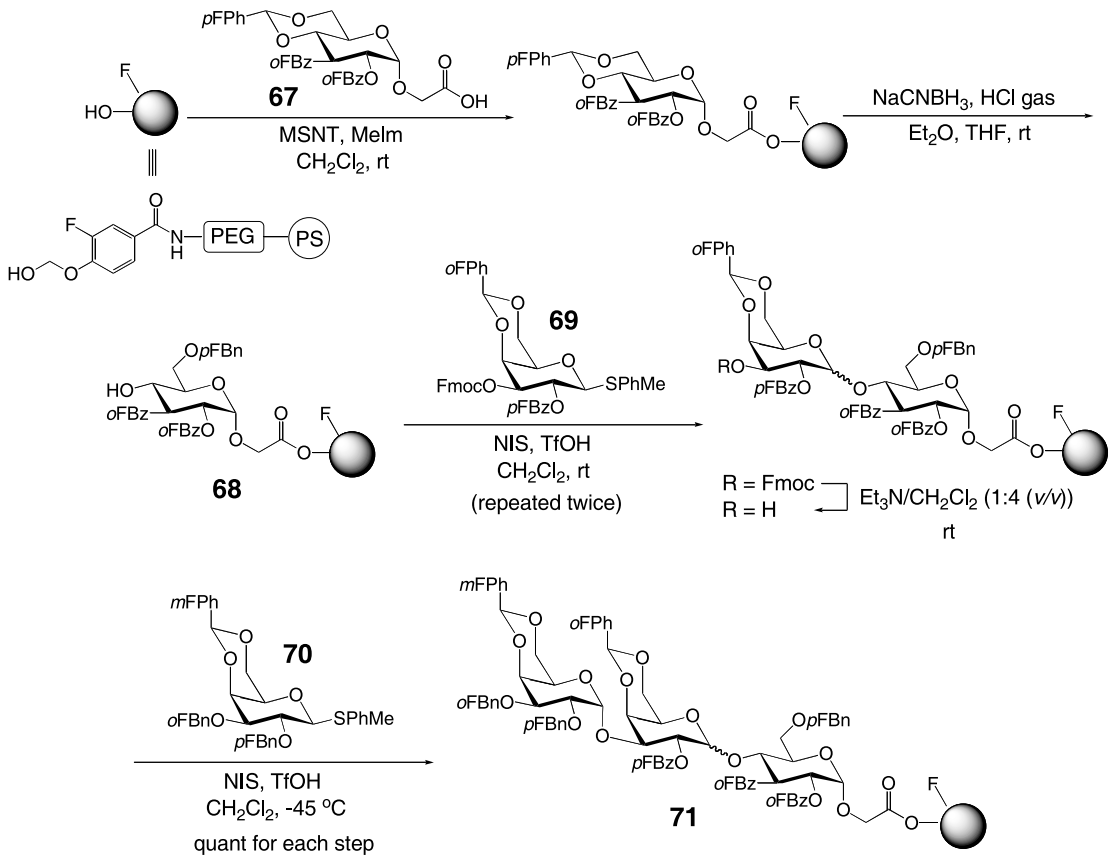
2.4 Monitoring of Solid-Phase Reactions

Real-time monitoring of the polymer-supported reaction is nearly impossible by simple and routine methods such as TLC and HPLC. In the case of solid-phase peptide synthesis, the end of the reaction can be readily checked by the ninhydrin test. In the solid-phase synthesis of oligosaccharides, a convenient method for checking the hydroxyl group on the polymer support has not been developed yet. In most cases, the reaction is checked after small-scale cleavage of the products. If the Fmoc protecting group is used, the efficiency in glycosylation and deprotection on the resins can be evaluated by UV absorption of dibenzofurubene derivatives when it is deprotected by the mild bases, such as piperidine or Et_3N.

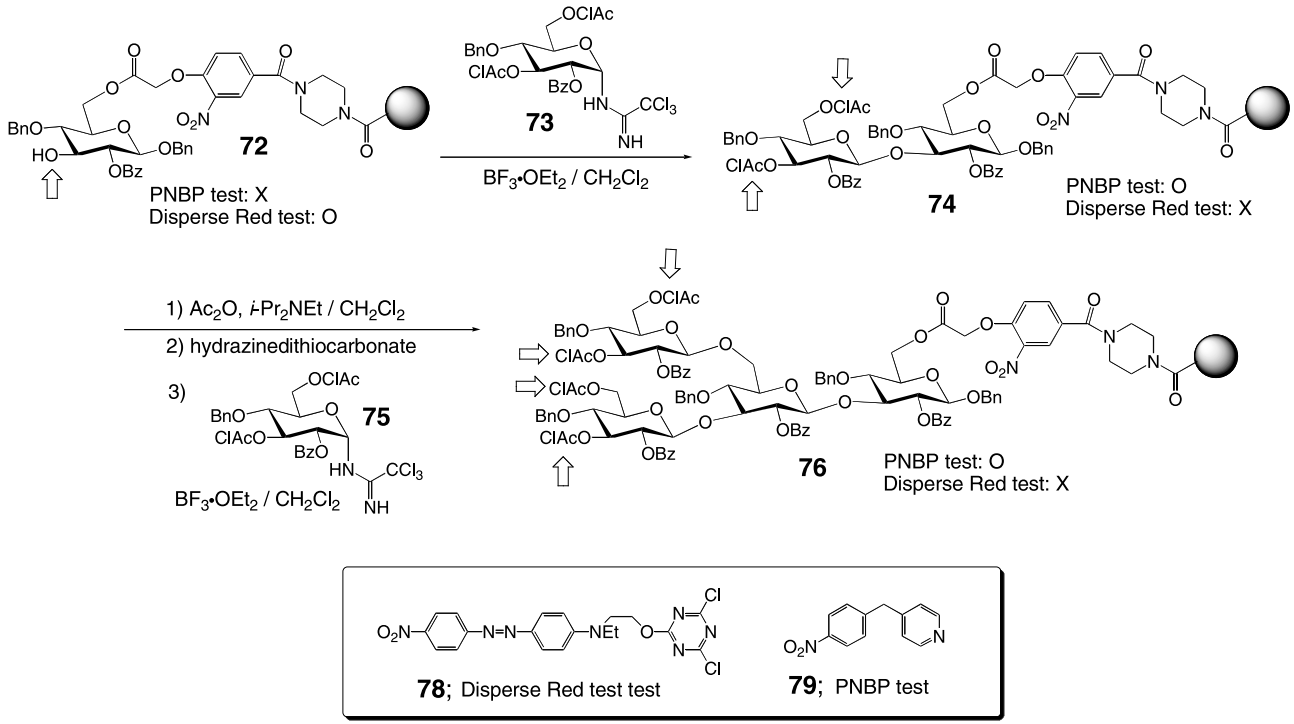
Recent advances in NMR have proven to be effective for direct and nondestructive monitoring of the progress in oligosaccharide synthesis. Wong and Kanie used the inverse gated decoupling technique of ^{13}C NMR to monitor the glycosylations on the TentaGel resin, in which two ^{13}C-enriched carbonyl tags were introduced both on the glycine linker as an internal standard and on the acetyl protecting group of the introducing glycosyl donors [50]. Sialy Lewis X branched tetrasaccharide was synthesized by this monitoring method.

In a similar way, Kihlberg and co-workers have applied the ^{19}F-tag both on the linker and the protecting groups, i. e., *p*-fluorinated benzyl ethers, benzoates, and benzylidene acetals of the glycosyl reagents (❯ *Scheme 17*) [51,52,53]. The gel-phase ^{19}F NMR allowed highly sensitive, simple, and quantitative analysis of the oligosaccharide assembly via continuous glycosylation and deprotection process on the resin. ArgoGel resin with a 3-fluoro-4-hydroxymethyloxybenzamide linker was esterified with glucosyl derivative **67**, which was protected with two *o*-fluorobenzoates and 4,6-*O*-*p*-fluorobenzylidene acetal. The loading yield was easily evaluated based on the integration of the resonances from the fluorine atoms in the linker ($\delta = -134.4$ ppm), the 4,6-*O*-*p*-F-benzylidene acetal group (–113.2 ppm), and two *o*-F-benzoyl groups (–109.0 and -110.1 ppm). Followed by the reductive opening of the 4,6-*O*-*p*-F-benzylidene acetal group by $NaCNBH_3$ in the presence of HCl gas, the acceptor **68** was glycosylated with thiogalactoside **69** and subsequently with **70**, each bearing the fluorine tag in different positions of the 4,6-*O*-benzylidene acetal groups, i. e., either ortho- ($\delta = -120.9$ ppm) or meta-positions ($\delta = -113.9$ ppm) of the aromatic groups. The sequence of reactions provided the trisaccharide **71**, α-Gal epitope (Galα(1–3)Galβ(1–4)Glc), which is responsible for hyperacute rejection in xenotransplantation of porcine organs. Each step was evaluated to proceed quantitatively based on this ^{19}F NMR analysis.

Ito and co-workers have developed a practical and convenient method for the real-time monitoring of solid-phase oligosaccharide synthesis based on the “on-resin color tests”,


pFPh
oFBz
oFBzO
67
OH
MSNT, MeIm
CH2Cl2, rt
NaCNBH3, HCl gas
Et2O, THF, rt
HO
F
PEG
PS
68
69
oFPh
FmocO
SPhMe
pFBzO
OpFBn
RO
NIS, TfOH
CH2Cl2, rt
(repeated twice)
R = Fmoc
R = H
Et3N/CH2Cl2 (1:4 (v/v))
rt
mFPh
70
oFBnO
pFBnO
NIS, TfOH
CH2Cl2, -45 °C
quant for each step
71

Scheme 17


72
PNBP test: X
Disperse Red test: O
73
BF3•OEt2 / CH2Cl2
74
PNBP test: O
Disperse Red test: X
1) Ac2O, i-Pr2NEt / CH2Cl2
2) hydrazinedithiocarbonate
3)
75
BF3•OEt2 / CH2Cl2
76
PNBP test: O
Disperse Red test: X
78; Disperse Red test test
79; PNBP test

Scheme 18

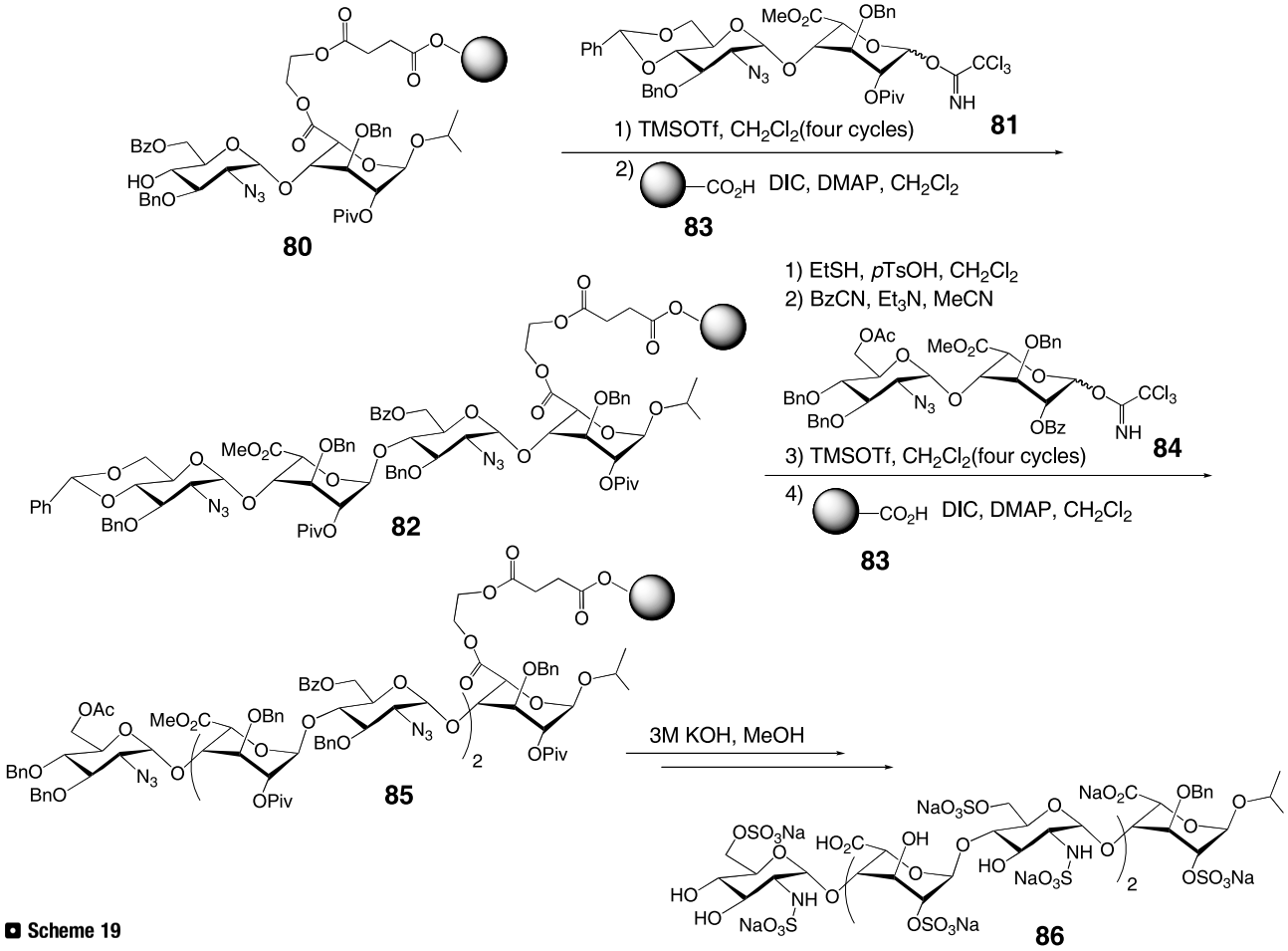
which enable side-by-side detection of glycoside bond formation and acceptor consumption (◉ *Scheme 18*) [54]. They used two color test systems: the modified Taddei's test with the cyanuric chloride-Disperse Red conjugate **78** was applied for detection of the presence of nucleophilic free hydroxyl and amino groups on the resin, whereas the chloroacetyl (ClAc) group was used as a "temporary" hydroxyl protecting group, which can be detected by a red color generated by the action of *p*-nitrobenzylpyridine (PNBP) **79** under basic conditions. On the basis of these monitoring methods, the repeating units of the immuno-active oligosaccharide schizophyllan was synthesized (◉ *Scheme 18*). The TentaGel-loaded acceptor **72**, which is negative for the PNBP test and positive for the Disperse Red test, was glycosylated with 3- and 6-*O*-ClAc-protected trichloroacetimidate **73** in the presence of BF_3OEt_2 in CH_2Cl_2 to provide β(1–3)-linked disaccharide **74**. The completion of the reaction was easily checked by the positive response for the PNBP test and the negative response for the Disperse Red test. After the selective deprotection of the ClAc group by hydrazinedithiocarbonate treatment, the free hydroxyls at 3′- and 6′-positions were subjected to further glycosylation with glucosyl trichloroacetimidate **75** to give resin-bound tetrasaccharide **76**, which was easily detected by the deep red color for the PNBP test.

3 Polymer-Supported and Tag-Assisted Oligosaccharide Synthesis in Solution

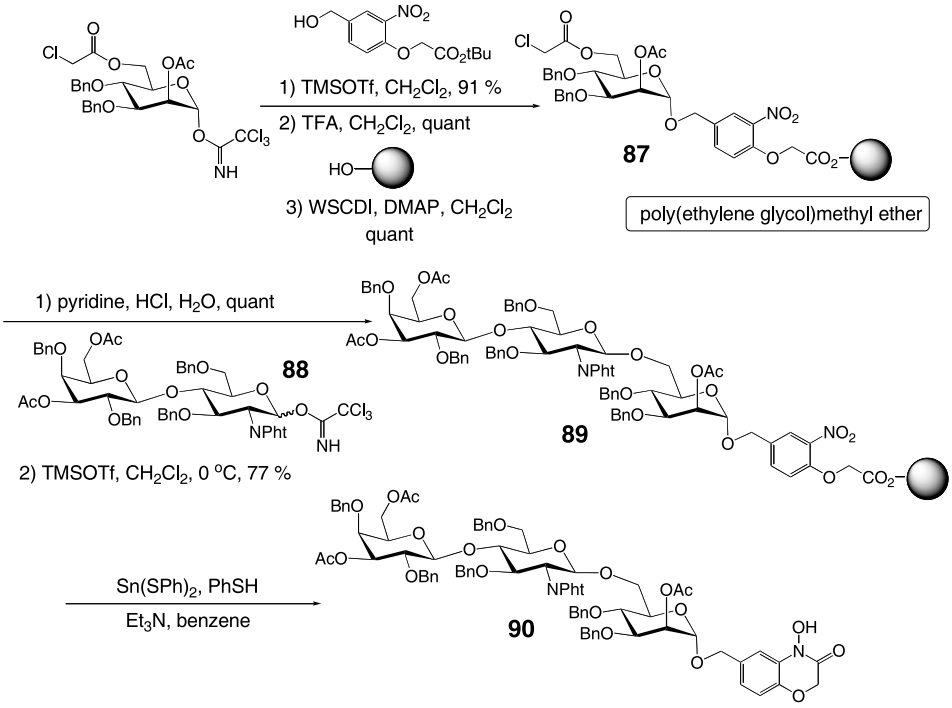
A phase tag strategy has been developed for high-throughput synthesis as a hybrid system that combines the merits of solid-phase synthesis (easy separation) and solution-phase synthesis (homogeneous reaction conditions). In this strategy, a compound having a tag group is easily separated from untagged molecules. Various phase tags such as soluble polymers, fluorous, hydrophobic, and basic tags have been reported.

3.1 Polymer-Supported Synthesis of Oligosaccharides

Polymer-supported solution synthesis has several advantages over solid-phase synthesis. The reaction can be carried out in solution and the solubility of the reactants allows reaction kinetics and anomeric stereocontrol similar to those observed in solution chemistry. Furthermore, in some cases, the process of the reaction can be easily checked by the conventional TLC or MALDI-TOF-MS. Out of a large number of polymer-supported syntheses so far reported, polyethylene glycol monomethyl ether (MPEG) has been widely used for oligosaccharide synthesis [55,56]. Polyethylene glycol is soluble in a variety of solvents such as CH_2Cl_2 and the protic solvents but insoluble in diethyl ether. The polymer-bound products are therefore isolated by precipitating the polyethylene glycol supports using ether. Martin-Lomas and co-workers have succeeded in the synthesis of heparin-like oligosaccharide **86** using both MPEG as a soluble polymer support and Merrifield-type resin (polyethylene glycol-grafted polystyrene) as a solid-support used for scavenging the unreacted acceptors (◉ *Scheme 19*) [57,58]. The disaccharide acceptor **80** bound to MPEG through a succinic ester linkage at the uronic acid carboxylate, was glycosylated with the trichloroacetimidate donor 81 acceptor **81** in the presence of TMSOTf to provide the α(1–4)-linked tetrasaccharide **82** with complete stereoselectivity. This glycosylation step was followed by capping with the Merrifield-type resin **83** functional-


80
81
1) TMSOTf, CH2Cl2(four cycles)
2)
CO2H DIC, DMAP, CH2Cl2
83
82
1) EtSH, pTsOH, CH2Cl2
2) BzCN, Et3N, MeCN
84
3) TMSOTf, CH2Cl2(four cycles)
4)
CO2H DIC, DMAP, CH2Cl2
83
85
3M KOH, MeOH
86

Scheme 19



Scheme 20

ized with carboxylic acid in order to esterify the unreacted C4-hydroxyl group of the acceptor **80**. The outcome of the α-stereoselectivity and the ensuring glycosylation could directly be monitored by NMR. Removal of the benzylidene acetal group of **82** by EtSH in the presence of *p*-TsOH, benzoylation of the C6-hydroxyl, a second glycosylation with **84**, and capping of the unreacted acceptor **82** with scavenger resin **83** provided the key intermediate **85**, which was successfully transformed to hexasaccharide **86**.

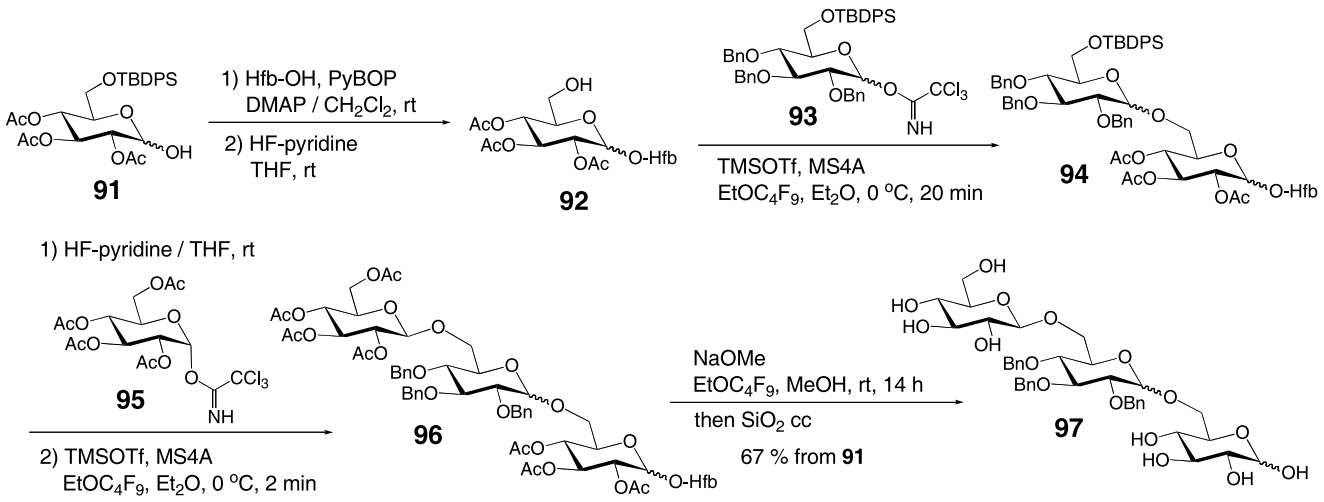
Ito and co-workers have reported the MPEG-supported synthesis of the trisaccharide **90**, the outer branching fragment of the tetraantennary complex-type glycan chain, by applying the self-cleavable Wang resin-type linker (*Scheme 20*) [59]. The mannosyl acceptor **87** loaded on the MPEG was glycosylated with the lactosamine-derived trichloroacetimidate **88** by using TMSOTf, providing β-glycoside **89** in 77% yield. MPEG-bound trisaccharide **89** was then treated with $Sn(SPh)_2$ in the presence of thiophenol to release **90** from the soluble-support via selective reduction of the nitro group followed by spontaneous intramolecular cyclization.

A simple formylacetal (CH_2) has also been used as a linker for oligosaccharide synthesis on MPEG. This small linker allows the hindered hydroxyls, such as the 4-hydroxy group of glucose, to be installed on the MPEG [60]. The formylacetal linker is stable under the acidic conditions mostly used for glycosylation, but it is readily cleaved by Lewis acids (TMSI or $Ce(OTf)_x$) or trifluoroacetic acid. A simplified work-up and purification procedure was also reported by use of the solid acid catalyst.

3.2 Tag-Assisted Synthesis of Oligosaccharides

Tag-assisted methods in solution-phase, other than MPEG supports, have also emerged as an attractive strategy for combinatorial and/or high-throughput synthesis of oligosaccharides. A highly fluorinated compound is soluble in fluorous solvents such as perfluorohexane, and is readily separated from nonfluorinated compounds through simple fluorous-organic solvent partitioning. Therefore, the fluorous synthesis is an attractive alternative to solid-phase synthesis, directed towards combinatorial chemistry and parallel synthesis. Although several fluorous protecting groups are reported for this purpose, their application is limited to the synthesis of simple small molecules because the fluorine content in each fluorous protecting group is low [61]. Inazu and co-workers have developed fluorous tags with high fluorine content and applied them to oligosaccharide synthesis (*Scheme 21*) [62,63]. They initially investigated the Bfp-OH tag (bisfluorous chain type propanoyl), which was introduced at three of four hydroxyls on the glycosyl acceptors for the synthesis of a simple tetrasaccharide as well as the Gb3 oligosaccharide. Later, the Hfb-OH tag (hexakisfluorous chain type butanoyl) with more fluorous chains was designed in order to enhance the efficiency in fluorous extraction; a rapid synthesis of trisaccharide **97** was achieved by introducing only one fluorous tag (*Scheme 21*). The Hfb tag was first attached to the anomeric hydroxyl of the glucose derivative **91** by using PyBOP and DMAP, and the deprotection of the TBDPS group by HF-pyridine in THF provided the glucose acceptor **92**. Compound **92** was then glycosylated with glucosyl trichloroacetimidate **93** by using TMSOTf in a mixed solvent of $EtOC_4F_9$ and Et_2O to provide disaccharide **94**. The fluorous tag-containing **94** was extracted with the fluorous solvent FC-72 by partitioning the product mixtures between FC-72 and an organic solvent. No additional purification, such as silica gel chromatography, was carried out. After removing the TBDPS group of **94** with HF-pyridine treatment in THF, the second glycosylation was performed with **95** to provide trisaccharide **96** after the partitioning by the fluorous/organic solvents. Finally, the Hfb group of **96** was removed by the treatment with NaOMe in $EtOC_4F_9$/MeOH to afford the crude **97**, which was extracted into the MeOH layer from the partitioning mixture of CF-72 and MeOH. By this method, pure **97** was obtained in 67% yield from **91** after silica gel chromatography, while the methyl ester derivative of the Hfb-OH tag, which was extracted into the CF-72 layer during the partitioning process, was collected and reused after hydrolysis with NaOH.

Fukase and co-workers have developed a concept termed "Synthesis based on Affinity Separation (SAS)", in which the desired tagged compound is separated from the reaction mixture by solid-phase extraction using specific molecular recognition (*Scheme 22*) [64,65,66,67]. They first employed the interaction between a crown ether (32-crown-10) as a tag on the desired compounds and the ammonium ion for SAS. After each reaction cycle, the reaction mixture was applied to the column loaded with aminomethylated polystyrene resin [trifluoroacetic acid (TFA) form]. By using nonpolar eluents such as CH_2Cl_2 and toluene, the tagged compound was selectively absorbed on the column, whereas other untagged impurities were washed off. Subsequent desorption by CH_2Cl_2-Et_3N or CH_2Cl_2-MeOH (1:1) afforded the desired compound with high purity. However, since the crown ether tag needed to be prepared by the tedious synthesis, the commercially available short-chain polyethylene glycol (PEG) was used as the second-generation tag. Since Triton X-100 is a detergent having a PEG chain and a hydrophobic moiety, the tagged compounds show good solubility in many organic sol-


OTBDPS
AcO
AcO
OAc
OH
91
1) Hfb-OH, PyBOP
DMAP / CH2Cl2, rt
2) HF-pyridine
THF, rt
OH
O-Hfb
92
OTBDPS
BnO
BnO
OBn
CCl3
NH
93
TMSOTf, MS4A
EtOC4F9, Et2O, 0 °C, 20 min
94
1) HF-pyridine / THF, rt
95
2) TMSOTf, MS4A
EtOC4F9, Et2O, 0 °C, 2 min
96
NaOMe
EtOC4F9, MeOH, rt, 14 h
then SiO2 cc
67 % from 91
97
HO
Bfp-OH
C8F17
F17C8
Hfb-OH

Scheme 21

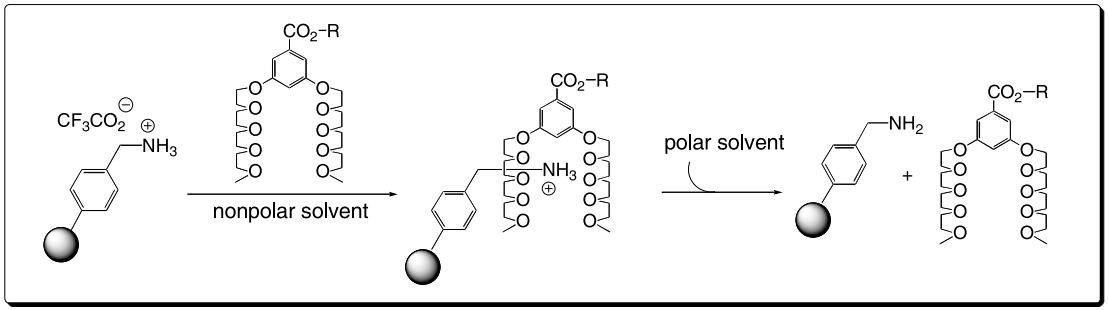
vents. However, the chain length of Triton X-100 is heterogeneous and Triton X-100 with the shorter chain length only showed weak binding. A podand-type ether with a pseudo-benzo-31-crown-10 structure, introduced as the third generation tag, was much easier to synthesize and found to show high affinity to the ammonium ion on the solid-support [67]. The present SAS method has been successfully applied to oligosaccharide synthesis (*Scheme 22*). The glycosylation between acceptor **98**, having the tag moiety via an acylaminobenzyl linker, and the *N*-Troc glucosamine trichloroacetimidate **99** was effected by using TMSOTf. After the affinity separation, the resulting disaccharide **100** was further subjected to glycosylation with an excess amount of the donor **99**. Although the affinity separation could not remove a small amount of the unreacted disaccharide **100** from the desired trisaccharide **101** (since both bear the tag moieties), subsequent separation by silica-gel column chromatography afforded pure **101** in 61% yield from **100**. Cleavage of the acylaminobenzyl linker in **101** was achieved by treatment with excess DDQ followed by the affinity column separation to provide the pure trisaccharide **102** in 82% yield.

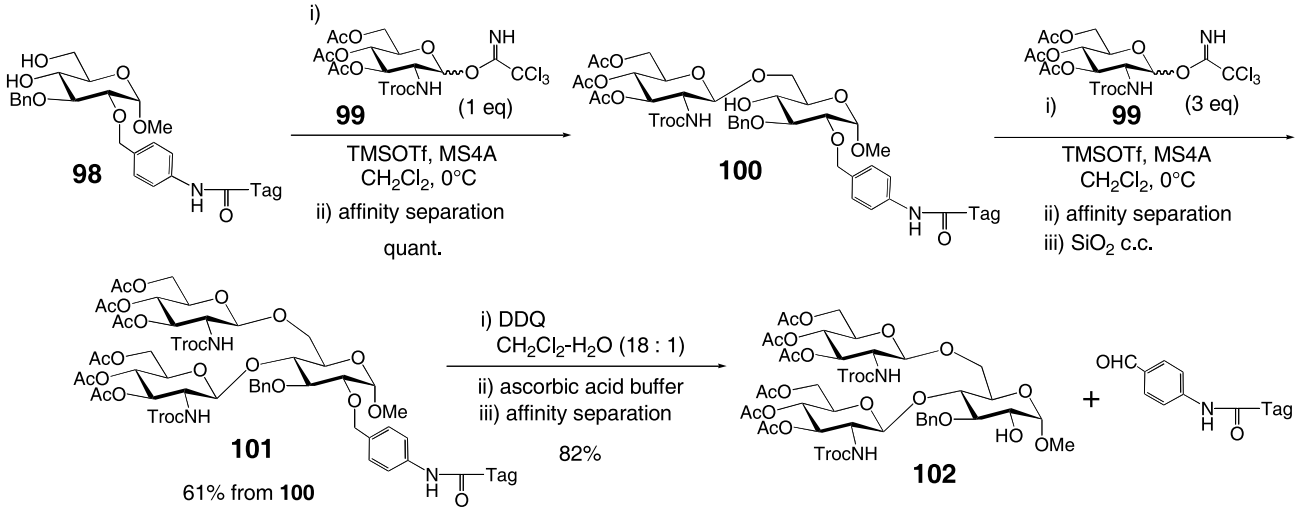
Alternatively, Rademann and co-workers have combined solid-phase and solution-phase reactions for the synthesis of oligosaccharides by using the hydrophobic tag, a concept called "Hydrophobically Assisted Switching Synthesis (HASP)" (*Scheme 23*) [68,69]. They applied a sufficiently long, double C18-hydrocarbon chain **103** as a fully reversible tag for the reactions both in solution and on the solid-support. The hydrophobic tag was introduced to the anomeric position of L-rhamnose, and α(1–2)-selective glycosylation was conducted first in solution (CH_2Cl_2) with rhamnosyl trichloroacetimidate **104** in the presence of TMSOTf as a Lewis acid (quantitative yield). The C-18 silica support was then added to the reaction mixture, solvents were evaporated, and all untagged materials were easily removed by washing with MeOH/H_2O. After the 2-*O*-phenoxyacetyl group was deprotected on the solid-support by the reaction with NaOMe in MeOH/H_2O (92%) followed by the washing step, a sufficiently pure compound was released from the solid-support by addition of $MgSO_4$ and eluting by CH_2Cl_2. After the HASP cycle was repeated a few times, the 3,4-*O*-butane-2,3-diacetal protecting group was removed by TFA. Finally, hydrogenation provided α(1–2)-liked oligorhamnose **105**, with an average yield of 94% per step. Each reaction could be directly followed by TLC and MS for this HASP method.

3.3 Polymer-Supported Enzymatic Synthesis of Oligosaccharides

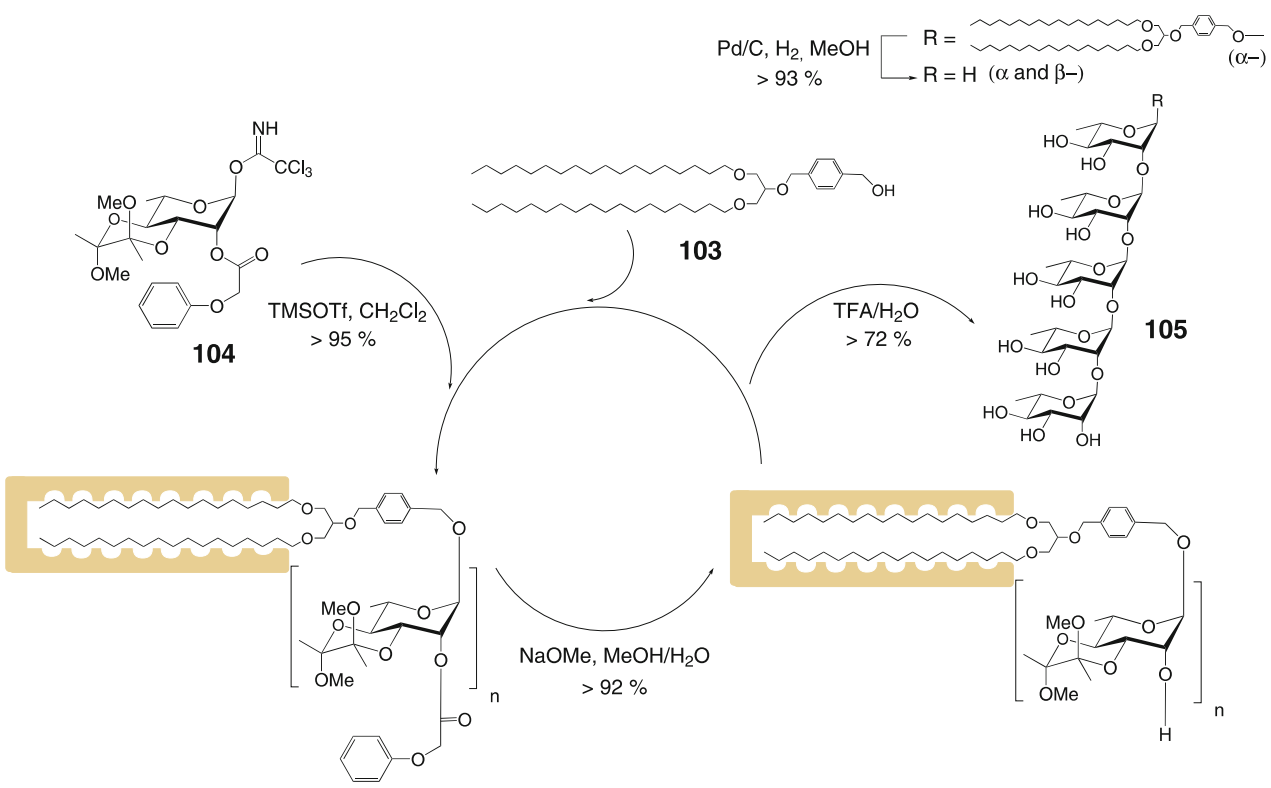
Since the chemical synthesis of oligosaccharides requires many synthetic steps including protection and deprotection procedures, the enzymatic approach has attracted much attention for the rapid synthesis of oligosaccharides. In addition, the perfect regio- and stereoselectivities of enzymatic methods with glycosyltransferases are quite attractive. Several transferases such as β(1,4)-galactosyltransferase, α(1,3)-fucosyltransferase, and α-sialyltransferase have been used for polymer-supported enzymatic synthesis [70,71,72,73,74,75]. The selection of the polymer support is very important for the polymer-supported enzymatic synthesis of oligosaccharides.

Nishimura and co-workers have developed the catch-and-release strategy between solid-phase and water-soluble polymer supports, the so called "polymer blotting method", which allows for the rapid and efficient synthesis of glycopeptides [76]. The method involves (i) conven-

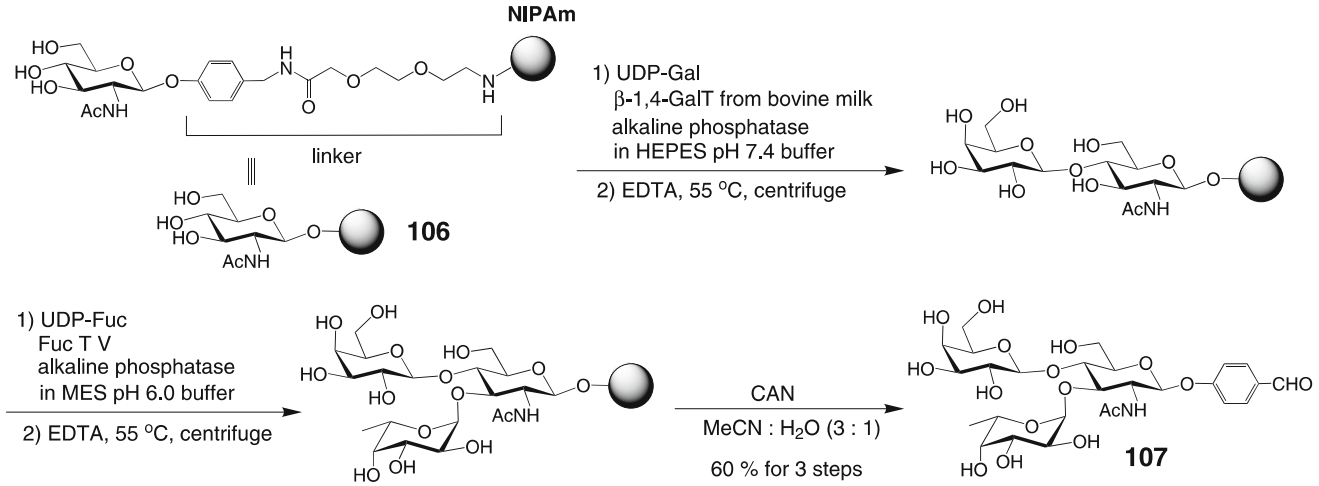

CF3CO2
NH3
CO2–R
nonpolar solvent
polar solvent
NH2
+


98
i)
99
(1 eq)
TMSOTf, MS4A
CH2Cl2, 0°C
ii) affinity separation
quant.
100
i)
99
(3 eq)
TMSOTf, MS4A
CH2Cl2, 0°C
ii) affinity separation
iii) SiO2 c.c.
101
61% from 100
i) DDQ
CH2Cl2-H2O (18 : 1)
ii) ascorbic acid buffer
iii) affinity separation
82%
102
Tag

Scheme 22


Pd/C, H2, MeOH
> 93 %
R =
R = H (α and β–)
(α–)
NH
CCl3
MeO
OMe
104
TMSOTf, CH2Cl2
> 95 %
103
OH
TFA/H2O
> 72 %
105
HO
OH
R
MeO
OMe
n
NaOMe, MeOH/H2O
> 92 %
H

Scheme 23


NIPAm
linker
106
1) UDP-Gal
β-1,4-GalT from bovine milk
alkaline phosphatase
in HEPES pH 7.4 buffer
2) EDTA, 55 °C, centrifuge
1) UDP-Fuc
Fuc T V
alkaline phosphatase
in MES pH 6.0 buffer
2) EDTA, 55 °C, centrifuge
CAN
MeCN : H2O (3 : 1)
60 % for 3 steps
107

Scheme 24

tional solid-phase synthesis of glycopeptides that contain mono-, di-, and tri-saccharides on Thr or Ser residues and Blase (glutamic acid specific protease)-sensitive amino acid sequence, (ii) removal of the protecting groups and release from the resin, (iii) attachment of the released glycopeptides to the water-soluble polymers via an oxime anchor, (iv) sugar elongations using glycosyltransferases, and (v) cleavage of the desired glycopeptides from the soluble supports by Blase mediated hydrolysis. They have applied three glycosyltransferases, namely, β(1,4)-galactosyltransferase, α(2,3)-(*O*)-sialyltransferase, and α(2,3)-(*N*)-sialyltransferase, and successfully prepared six kinds of parallel and 36 kinds of combinatorial libraries of MUC1 mucin glycopeptides. Applications of the method to other glycopeptides have also been reported (see Chap. 5.4).

Wong and co-workers investigated enzymatic oligosaccharide synthesis on a thermo-responsible polymer support (*Scheme 24*) [77]. The copolymers of *N*-*i*-propylacrylamide (NIPAm) and functionalized monomers are thermo-responsible and exhibit inverse temperature-dependent solubility in water; they are soluble in cold water but become insoluble and precipitate from the solution when the temperature is higher than their lower critical solution temperature. The NIPAm polymer-support greatly simplified the purification process for oligosaccharide synthesis in water; carbohydrate acceptors attached to the NIPAm support were used for enzymatic glycosylation, and the glycosylation products can be isolated simply by thermal precipitation and centrifugation. The availability of direct NMR analysis also makes the method attractive. Galactosylation of the immobilized GlcNAc **106** was performed by using 100 mU bovine milk β(1,4)-galactosyltransferase with 1 equivalent of UDP-galactose donor in *N*-(2-hydroxyethyl)-piperazine-*N*′-(2-ethanesulfonic acid) (HEPES) buffer at pH 7.4. Followed by the precipitation of the product by heating the solution to 55 °C, it was further fucosylated using the bovine α(1,3)-fucosyl transferase and GDP-fucose donor. After cleavage from the NIPAm support by CAN treatment, the trisaccharide Le^X **107** was obtained in 60% yield for three steps without chromatographic separation of the intermediates. They also have immobilized several enzymes on the same NIPAm polymers, such as subtilisin BPN′, which were found to show comparable activities to their soluble forms. For example, after the NIPAm-subtilisin conjugate was used for proteolysis of a modified form of the glycoprotein ribonuclease (RNase B), the immobilized enzyme can be easily recovered for reuse by gentle heating and precipitation. The methods are expected to be applied to glycosyltransferases and this thermo-responsible polymer-supported procedure will greatly facilitate enzyme-catalyzed organic synthesis.

3.4 Microfluidic Methods for Oligosaccharides Synthesis

Continuous flow synthesis using a microfluidic system has emerged as an innovative technology in organic synthesis, i. e., from small-scale optimization of the reaction conditions, high-throughput synthesis, to large-scale production. This method has various advantages in terms of reaction control, such as fast mixing, effective temperature control, and precise residence time control. Since the reaction can be performed under the flow process, the method is readily applicable to library construction by preparing the stock solutions of glycosyl acceptors and donors and flowing each by each through syringe- or HPLC-pumps. On the other hand, once the reaction conditions are optimized for small-scale operation, the same condi-

tions are directly applied to large-scale synthesis, i. e., for process synthesis on an industrial scale, by using several microreactors under the flow process. Seeberger and co-workers reported the first application of the microfluidic system to optimize the glycosylation on an analytical scale [78]. The reaction performed in a microreactor was monitored by HPLC, which is directly connected to the microreactor. The products composition was analyzed by changing the concentration, reaction time, and temperature; the optimal conditions for α-mannosylation and the different reactivity due to the protection patterns of acceptors have been determined.

Fukase and co-workers have combined a microreactor and their SAS protocol (Synthesis based on Affinity Separation, ❯ *Sect. 3.2*) in pursuit of high-throughput oligosaccharide synthesis (❯ *Fig. 2*) [67]. They used a combination of an IMM micromixer and a stainless tube reactor, equipped with a stainless column as an affinity separation unit. The glycosylation was carried out in the micromixer by combining the CH_2Cl_2 solution of acceptor **108** and donor **109** with the solution of TMSOTf in CH_2Cl_2 under continuous flow. The mixture was directly introduced to the stainless steal affinity column and the flow channel was switched to wash the column with CH_2Cl_2. The desired disaccharide **110** was trapped in the column, but then eluted with CH_2Cl_2-MeOH (1:1); the product **110** was obtained in 91% yield.

A successful α-sialylation under microfluidic conditions has also been reported by the same group; the reactivity was totally different from that observed under the batch conditions [79]. The highly reactive sialyl donor developed in this laboratory, bearing the *N*5-phthalyl group and the *N*-phenyl trifluoroacetimidate leaving group [80], could not be used for large-scale α-sialylation with a galactose acceptor in a batch apparatus. Because of inadequate mixing and inefficient heat transfer under the batch conditions, the exposure of the reactive sialyl donor to Lewis acid mainly leads to glycal formation. They have circumvented these problems by using microfluidic systems and realized quantitative α-sialylation with perfect α-selectivity, even at a large-scale. The increase in yield was also observed with the lactose acceptors under the microfluidic α-sialylation.

These leading examples will not only open up a new path to high-throughput and combinatorial library syntheses, favorably by combination with the programmed automation of glycosylation-analysis sequences, but also provide new insights into the reactivity of glycosylation. Once an oligosaccharide of biological relevance has been found from the microfluidic reactions, it will be provided at an industrial-scale by the same apparatus [81].

4 Conclusions

Synthetic oligosaccharides and glycoconjugates often play a decisive role in elucidation of their biofunctions, since chemical synthesis can provide homogeneous preparations without contamination of other bioactive compounds. For the clinical application, the use of synthetic specimens is also important to avoid side-effects due to possible contamination. Although oligosaccharide synthesis has seen dramatic developments as described above, a high level of technical expertise is still required. Further improvement of oligosaccharide synthesis is expected to establish general and efficient methods, which will speed up elucidation of the biological functions of oligosaccharides as well as clinical applications of oligosaccharide-based drugs.

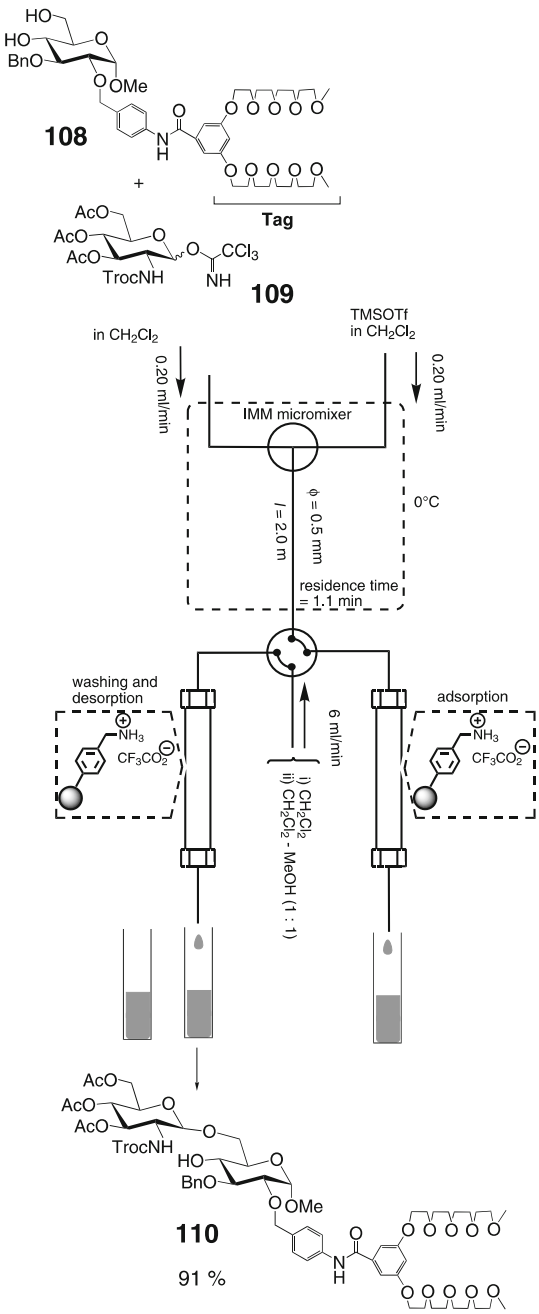


Figure 2
Continuous flow glycosylation using a microreactor and SAS

References

1. Fukase K (2001) Combinatorial and solid-phase methods in oligosaccharide synthesis. In: Fraser-Reid B, Tatsuta K, Thiem J (eds) Glycoscience II. Springer, Berlin Heidelberg New York, p 1621
2. Wu X, Grathwohl M, Schmidt RR (2001) Org Lett 3:747
3. Tolborg JF, Jensen KJ (2000) Chem Commun:147
4. Petersen L, Jensen KJ (2001) J Chem Soc, Perkin Trans 1:2175
5. Drinnan N, West ML, Broadhurst M, Kellam B, Toth I (2001) Tetrahedron Lett 42:1159
6. Belogi G, Zhu T, Boons GJ (2000) Tetrahedron Lett 41:6965
7. Belogi G, Zhu T, Boons GJ (2000) Tetrahedron Lett 41:6969
8. Melean LG, Haase WC, Seeberger PH (2000) Tetrahedron Lett 41:4329
9. de Paz JL, Noti C, Seeberger PH (2006) J Am Chem Soc 128:2766
10. de Paz JL, Seeberger PH (2006) QSAR Combi Sci 25:1027
11. Timmer MSM, Codee JDC, Overkleeft HS, van Boom JH, van der Marel GA (2004) Synlett 2155
12. Knerr L, Schmidt RR (2000) Eur J Org Chem:2803
13. Izumi M, Fukase K, Kusumoto S (2002) Synlett:1409
14. Burt J, Dean T, Warriner S (2004) Chem Commun:454
15. Amaya T, Tanaka H, Takahashi T (2004) Synlett:497
16. Komba S, Kitaoka M, Kasumi T (2005) Eur J Org Chem:5313
17. Takahashi T, Okano A, Amaya T, Tanaka H, Doi T (2002) Synlett:911
18. Lopez-Prados J, Felix C, Reichardt NC, de Paz JL, Morales EQ, Martin-Lamos JL (2005) Org Biomol Chem 3:764
19. Kanie O, Ohtsuka I, Ako T, Daikoku S, Kanie Y, Kato R (2006) Angew Chem Int Ed 45:3851
20. Roussel F, Knerr L, Schmidt RR (2001) Eur J Org Chem:2067
21. Roussel F, Takhi M, Schmidt RR (2001) J Org Chem 66:8540
22. Jonke S, Liu KG, Schmidt RR (2006) Chem Eur J 12:1274
23. Wu X, Grathwohl M, Schmidt RR (2002) Angew Chem Int Ed 41:4489
24. Wu X, Schmidt RR (2004) J Org Chem 69:1853
25. Roussel F, Knerr L, Grathwohl M, Schmidt RR (2000) Org Lett 2:3043
26. Wu X, Schmidt RR (2004) Eur J Org Chem 2826
27. Hewitt MC, Seeberger PH (2001) J Org Chem 66:4233
28. Love KR, Andrade RB, Seeberger PH (2001) J Org Chem 66:8165
29. Plante OJ, Palmacci ER, Andrade RB, Seeberger PH (2001) J Am Chem Soc 123:9545
30. Plante OJ, Palmacci ER, Seeberger PH (2001) Science 291:1523
31. Bartolozzi A, Seeberger PH (2001) Curr Opin Struct Biol 11:587
32. Hewitt MC, Seeberger PH (2001) Org Lett 3:3699
33. Marcaurelle LA, Seeberger PH (2002) Curr Opin Struct Biol 6:289
34. Palmacci ER, Plante OJ, Hewitt MC, Seeberger PH (2003) Helv Chim Acta 86:3975
35. Seeberger PH (2003) Chem Commun:1115
36. Seeberger PH, Werz DB (2005) Nature 4:751
37. Hewitt MC, Snyder DA, Seeberger PH (2002) J Am Chem Soc 124:13434
38. Ratner DM, Swanson ER, Seeberger PH (2003) Org Lett 5:4713
39. Love KR, Seeberger PH (2004) Angew Chem Int Ed 43:602
40. Egusa K, Fukase K, Nakai Y, Kusumoto S (2000) Synlett:27
41. Egusa K, Kusumoto S, Fukase K (2001) Synlett:777
42. Ando H, Manabe S, Nakahara Y, Ito Y (2001) Angew Chem Int Ed 40:4725
43. Manabe S, Ito Y (2001) Chem Pharm Bull 49:1234
44. Hanashima S, Manabe S, Ito Y (2003) Synlett 979
45. Wu J, Guo Z (2006) J Org Chem 71:7067
46. Tanaka H, Ishida T, Matoba N, Tsukamoto H, Yamada H, Takahashi T (2006) Angew Chem Int Ed 45:6349
47. MacCoss RN, Brennan PE, Ley SV (2003) Org Biomol Chem 1:2029
48. Dondoni A, Marra A, Massi A (2005) Angew Chem Int Ed 44:1672
49. Palmacci ER, Hewitt MC, Seeberger PH (2001) Angew Chem Int Ed 40:4533
50. Kanemitsu T, Wong CH, Kaine O (2002) J Am Chem Soc 124:3591
51. Mogemark M, Elofsson M, Kihlberg J (2001) Org Lett 3:1463

52. Mogemark M, Elofsson M, Kihlberg J (2002) ChemBioChem 12:1266
53. Mogemark M, Elofsson M, Kihlberg J (2003) J Org Chem 68:7281
54. Manabe S, Ito Y (2002) J Am Chem Soc 124:12638
55. Geurtsen R, Boons GJ (2002) Eur J Org Chem:1473
56. Quiclet-Sire B, Wilczewska A, Zard SZ (2000) Tetrahedron Lett 41:5673
57. Ojeda R, de Paz JL, Martin-Lomas M (2003) Chem Commun:2486
58. Ojeda R, Terenti O, de Paz JL, Martin-Lomas M (2004) Glycoconjugate J 21:179
59. Manabe S, Nakahara Y, Ito Y (2000) Synlett:1241
60. Oikawa M, Tanaka T, Kusumoto S, Sasaki M (2004) Tetrahedron Lett 45:787
61. Manzoni L (2003) Chem Commun:2930
62. Miura T, Goto K, Hosaka D, Inazu T (2003) Angew Chem Int Ed 42:2047
63. Miura T, Inazu T (2003) Tetrahedron Lett 44:1819
64. Zhang SQ, Fukase K, Kusumoto S (1999) Tetrahedron Lett 40:7479
65. Zhang SQ, Fukase K, Izumi M, Fukase Y, Kusumoto S (2001) Synlett:590
66. Fukase Y, Zhang SQ, Iseki K, Oikawa M, Fukase K, Kusumoto S (2001) Synlett:1693
67. Fukase K, Takashina M, Hori Y, Tanaka D, Tanaka K, Kusumoto S (2005) Synlett:2342
68. Bauer J, Rademann J (2005) J Am Chem Soc 127:7296
69. Bauer J, Brandenburg K, Zahringer U, Rademann J (2006) Chem Eur J 12:7116
70. Flitsch SL (2000) Curr Opin Chem Biol 4:619
71. Nishiguchi S, Yamada K, Fuji Y, Shibatani S, Toda A, Nishimura S (2001) Chem Commun 1944
72. Bezay N, Dudziak G, Liese A, Kunz H (2001) Angew Chem Int Ed 40:2292
73. Haneda K, Inazu T, Mizuno M, Iguchi R, Tanabe H, Fujimori K, Yamamoto K, Kumagai H, Tsumori K, Munekata E (2001) Biochim Biophys Acta 1526:242
74. Tolborg JF, Peterson L, Jensen KJ, Mayer C, Jakeman DL, Warren RAJ, Withers SG (2002) J Org Chem 67:4143
75. Naruchi K, Hamamoto T, Kurogochi M, Hinou H, Shimizu H, Matsushita T, Fujitani N, Kondo H, Nishimura SI (2006) J Org Chem 71:9609
76. Fumoto M, Hinou H, Ohta T, Ito T, Yamada K, Takimoto A, Kondo H, Shimizu H, Inazu T, Nakahara Y, Nishimura S (2005) J Am Chem Soc 127:11804
77. Huang X, Witte KL, Bergbreiter DE, Wong CH (2001) Adv Synth Catal 343:675
78. Ratner DM, Murphy ER, Jhunjhhunwala M, Snyder DA, Jensen KF, Seeberger PH (2005) Chem Commun:578
79. Tanaka Si, Goi T, Tanaka K, Fukase K (2007) J Carbohydr Chem 26:369
80. Tanaka K, Goi T, Fukase K (2005) Synlett 2958
81. Tanaka K, Fukase K (2007) Synlett:164

5.6 Stereoselective Synthesis of *β-manno*-Glycosides

Akihiro Ishiwata, *Yukishige Ito*
RIKEN (The Institute of Physical and Chemical Research),
Saitama 351–0198, Japan
yukito@riken.jp

Abstract

Among the various types of *O*-glycosides with biological relevance, the β-glycoside of D-mannose (*β-manno*-glycoside) has been considered as one of the most challenging targets from a synthetic point of view. The majority of synthetic approaches to *β-manno*-glycoside can be put into two categories (the direct glycosylation and the glycosylation-inversion approaches). Additionally, a variety of intriguing approaches have been investigated with substantial success (alkylative glycosylation of 1,2-stannylene acetal, reductive manipulation of orthoester, intramolecular aglycon delivery, and enzymatic glycosylation). In this chapter, progress in the conceptually demanding β-mannosylation technology will be discussed, including current state of the art, with particular focus upon applications related to the synthesis of glycoprotein-related oligosaccharides.

Keywords

Mannose; Glycoprotein; Glycosylation; Stereoselective

Abbreviations

All allyl
Bn benzyl

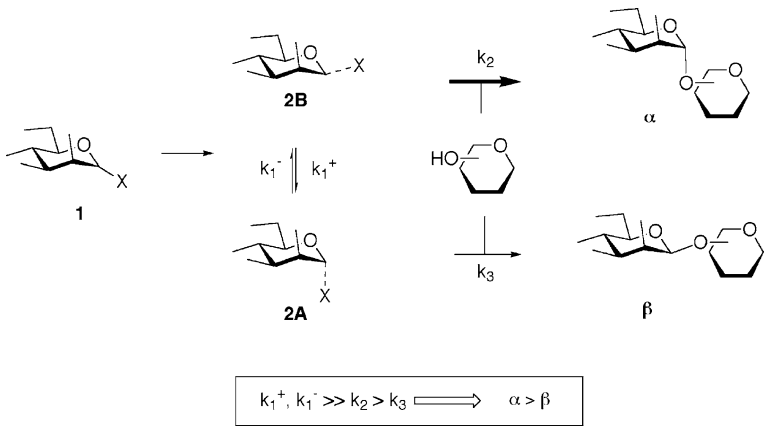
In: *Glycoscience*. Fraser-Reid B, Tatsuta K, Thiem J (eds)
Chapter-DOI 10-1007/978-3-540-30429-6_30:

MP *p*-methoxyphenyl
Phth phthaloyl
TCA trichloroacetyl
TBDPS *t*-butyldiphenylsilyl

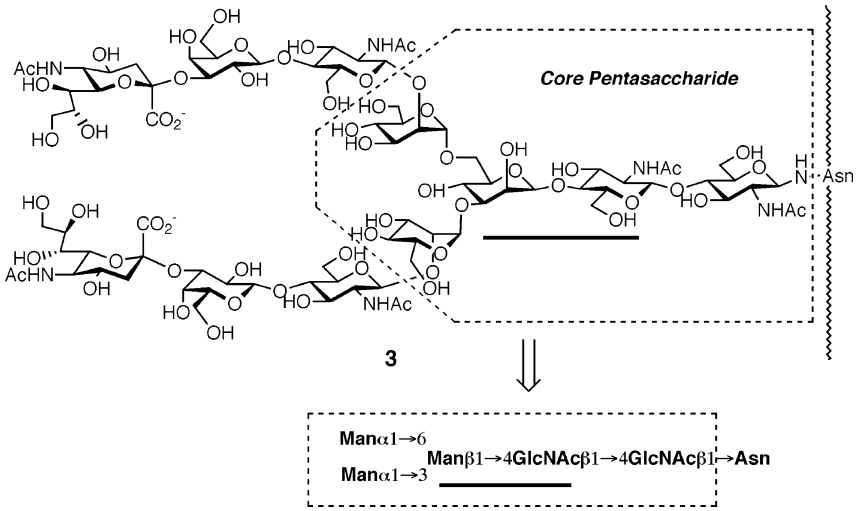
1 Introduction

In the last three decades or so, various types of so-called modern glycosylation reactions have been developed, which can be clearly distinguished from the classical methods [1]. These reactions utilize glycosyl donors **1** [X = F, $OC(NH)CCl_3$, SR, S(O)Ph, $OP(OR)_2$, etc.], which are relatively stable but can be activated into highly reactive glycosylating entities, under mild reaction conditions. For example, glycosyl fluorides, which used to attract little attention in synthetic carbohydrate chemistry, have proved to be powerful glycosyl donors, once exposed to appropriate activation conditions [2]. Likewise, thioglycosides [3] and trichloroacetimidates [4] have gained solid reputations as highly versatile glycosyl donors in oligosaccharide synthesis. The combination of these methodologies allows us to design synthetic routes to even more highly complex oligosaccharide structures [5]. However, these methodologies are still not free of stereochemical problems in a general sense.

Among the various types of *O*-glycosides with biological relevance, the β-glycoside of D-mannose (β-*manno*-glycoside) has been considered as one of the most challenging targets from a synthetic point of view [6]. This difficulty arises from its unique stereochemical array at the *C*-1/2 positions which is addressed in the simplified reaction pathway of *O*-glycosylation depicted in ❯ *Scheme 1*. Starting from the glycosyl donor **1**, expulsion of the leaving group **X** generates an oxocarbenium ion-like species, which usually exists as a rapidly equilibrating mixture of α- (**2A**) and β- (**2B**) ion pairs. Since the pseudo-axially oriented **2A** is more stabilized by the anomeric effect [6,7], **2B** preferentially reacts with an aglycon to give the α-glycoside as a major product, provided that the equilibration between **2A** and **2B** is rapid



Scheme 1

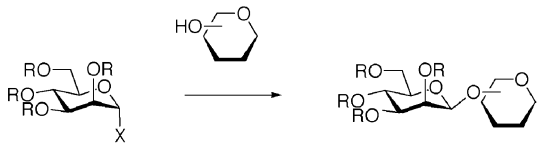


Scheme 2

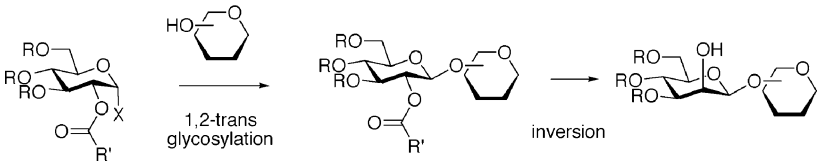
enough. Additionally, β-*manno*-glycoside has a 1,2-*cis* relative stereochemistry, formation of which cannot be effected by neighboring group participation of the *C*-2 substituent.

The biological importance of β-*manno*-glycoside is obvious from its widespread occurrence in glycoproteins [8]. Namely, all types of asparagine (Asn)-linked glycoprotein oligosaccharides have a so-called "core" pentasaccharide which includes β-configurated mannose (β-Man) linked to the 4-position of *N*-acetylglucosamine (GlcNAc), as can be seen in the undecasaccharide **3**, which is a typical complex type oligosaccharide (*Scheme 2*). Therefore, in order to pursue synthetic approaches to glycoproteins and related molecules, the stereocontrolled synthesis of β-Man is a fundamental requirement.

Direct Glycosylation Approach



Glycosylation-Inversion Approach



Scheme 3

The majority of synthetic approaches to β-*manno*-glycoside can be put into one of two categories (Scheme 3). Schematically, the most straightforward is the direct glycosylation approach using the mannosyl donor. Since the formation of β-*manno*-glycoside is disfavored under normal conditions (see above), a promoter possessing a special property in combination with a certain type of leaving group is required for its synthesis.

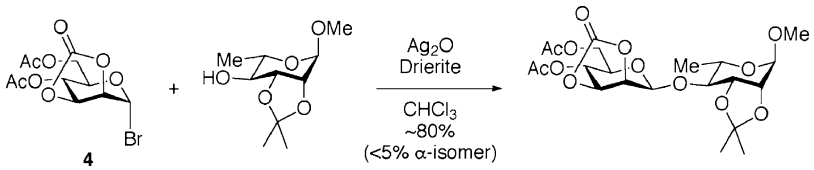
In a general sense, the more predictable tactic has been the glycosylation-inversion approach, which consists of β-gluco(1,2-*trans*)-glycosylation followed by the inversion of *C*-2 stereochemistry through oxidation-reduction or by an S_N2-type reaction.

Additionally, a variety of intriguing approaches have been investigated with substantial success [9,10,11,12,13].

In this chapter, progress in the conceptually demanding β-mannosylation technology will be discussed, including current state of the art, with particular focus upon applications related to the synthesis of glycoprotein-related oligosaccharides.

2 Direct Glycosylation Approach

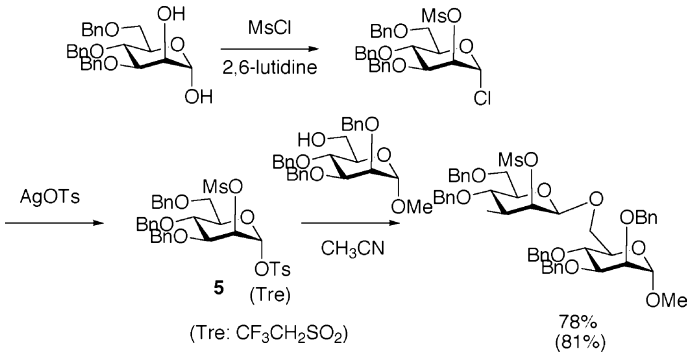
Direct β-selective glycosylation was achieved as early as 1961 by Gorin and Perlinupon, using the bromide **4** [14]. In this case, silver oxide was used as the promoter and the presence of the 2,3-di-*O*-carbonyl group seems to be the critical feature for a successful β-selective reaction. The same donor was used in a more efficient manner by Bebault and Dutton to prepare β-Man1→4Rha, which corresponds to the disaccharide segment of a constituent of the *Salmonella* type D_2 and E cell surfaces [15] (Scheme 4).



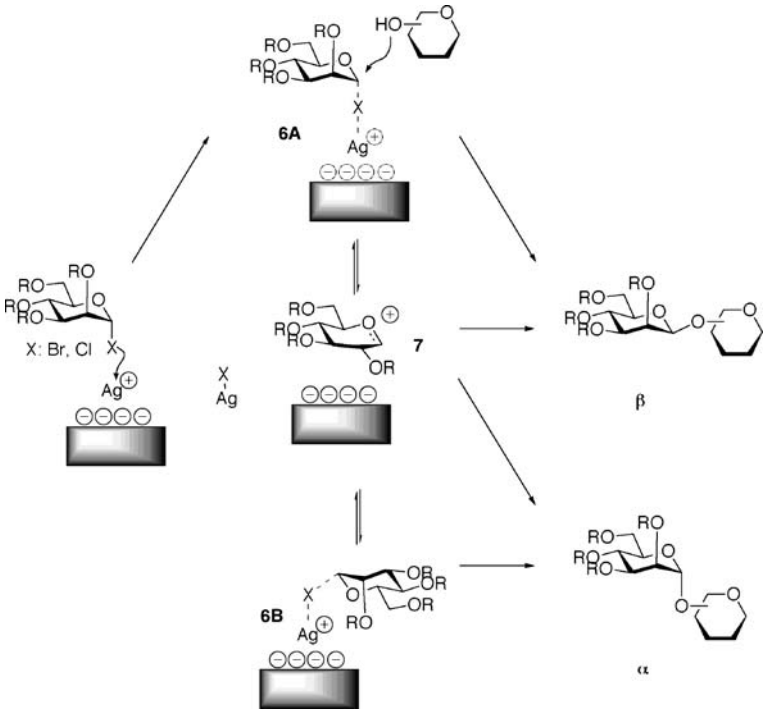
Scheme 4

S_N2-type conditions were investigated by Srivastava and Shuerch [16], who used the 2-*O*-mesyl-protected glycosyl sulfonate **5** (Scheme 5). The reactions were performed in acetonitrile without any promoter. The presence of the strongly electron-withdrawing mesyl group at *C*-2 seems to be essential. The corresponding 2-*O*-benzoate gave an α-linked product exclusively, presumably due to the neighboring effect of the benzoyl group. Although this method gave quite satisfactory results when applied to reactive acceptors, its further applicability to more challenging structures has not been demonstrated. A potential drawback from the practical point of view would be the conditions required for deprotection of the mesyl group (Na/liq NH_3).

As a more generally useful methodology, the direct β-mannosylation can be effected by a combination of a mannosyl halide with an "insoluble" silver salt [6]. The selectivity can



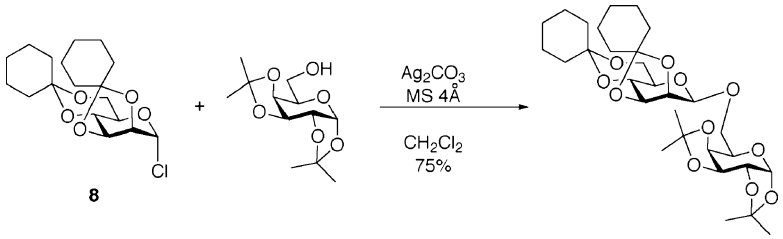
Scheme 5



Scheme 6

be explained by the mechanism depicted in *Scheme 6*. Thus, the surface-bound, activated bromide **6A** is attacked by the hydroxy group in an S_N2-like fashion. Under these conditions, stereochemical scrambling of the anomeric center to give **6B** via the free carbenium ion **7** can be minimized and inversion of configuration predominates in the glycosylation reaction.

For example, Garegg and Iversen demonstrated that Ag_2CO_3 was effective for the activation of mannosyl chloride **8** having 2,3–4,6-di-*O*-cyclohexylidene protection [17]. Reaction with



Scheme 7

a primary alcohol gave β-glycoside in high yield (*Scheme 7*). However, only poor results were obtained when this strategy was applied to the glycosylation of a secondary alcohol. Since the Ag_2CO_3-mediated glycosylation inevitably generates water as it proceeds, this reagent may not be optimal as a promoter.

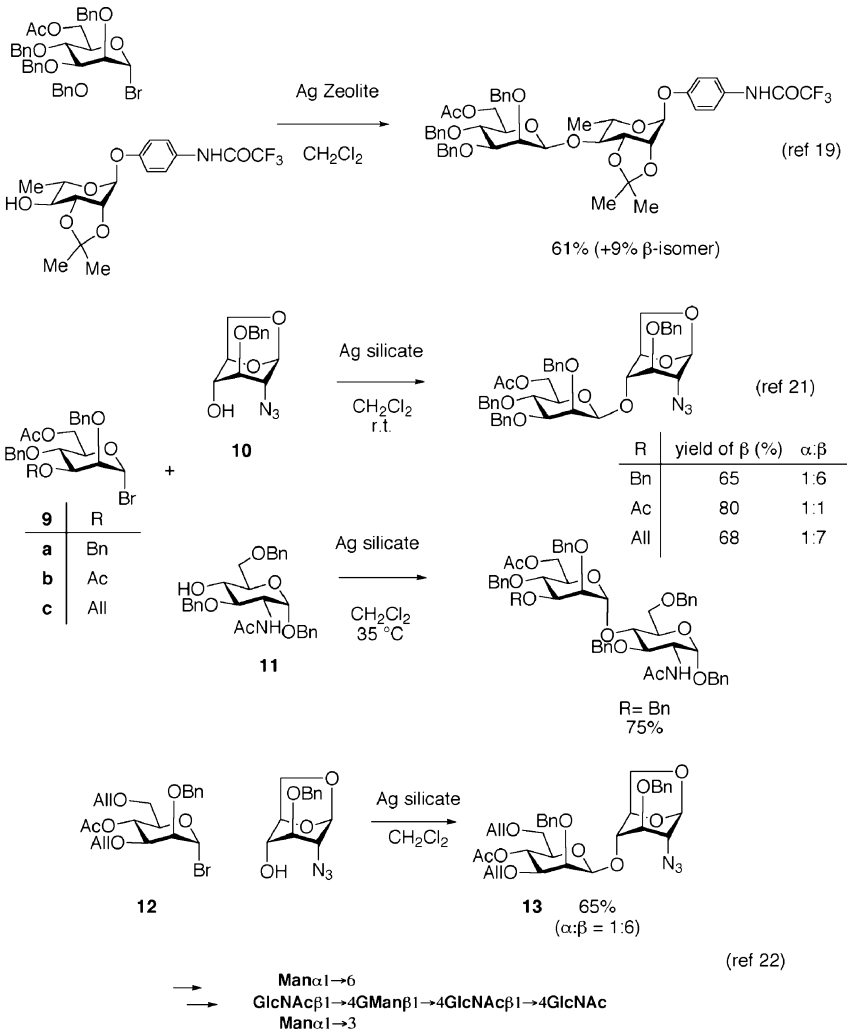
More widely used promoters for this purpose are silver zeolite and silver silicate, which were developed by Paulsen [18] and Garegg [19], respectively. Although not being commercially available, these reagents can be conveniently prepared from silver nitrate and proved to be quite powerful. More recently, van Boeckel developed the use of silver silica-alumina, which was reported to be more reactive than other insoluble salts and to have an enhanced cation capacity [20]. Representative examples of successful β-mannosylation by using insoluble salts are depicted in *Scheme 8* [19,21,22].

As can be expected from the proposed mechanism (*Scheme 6*), the stereoselectivity of insoluble silver salt-mediated reactions is quite sensitive to the steric hindrance of the acceptor. In those cases where the OH group is not sufficiently reactive, the surface-bound glycosyl donor is liberated into the solution to give the more reactive oxocarbenium ion **7**, which reacts in a manner that favors the formation of the undesired α-glycoside. Alternatively, reassociation with the negatively charged surface gives a mixture of α- and β-ion pairs (**6A** and **6B**) and subsequent reaction with the alcohol proceeds in a non-selective manner.

This aspect was very clearly demonstrated by a series of reactions performed by Paulsen et al. [19,20,21,22,23,24,25], in the course of their synthetic studies on Asn-linked oligosaccharides (*Scheme 8*). When using the bromide **9a**, the high selectivity (α:β = 1:7) obtained by reaction with **10** completely disappeared when **11** was used as an acceptor. It should be noted that, having a fixed $^1C_4$ conformation, the OH group of **10** has a reduced steric hindrance compared to that in **11**. The slightly modified donor **12** afforded a similar result (α:β = 1:6) and the β-linked disaccharide product **13** was converted into the bisecting GlcNAc-containing hexasaccharide.

Ogawa and coworkers have also extensively described synthetic studies on glycoprotein oligosaccharides and mannose-containing glycosphingolipids using similar reactions [26,27,28,29] (*Scheme 9*). For instance, the reaction of **14** with disaccharide **15** afforded **16** (α:β = 1.1:1), which was used as the key intermediate in their first chemical synthesis of a sialic acid-containing complex-type undecasaccharide [30].

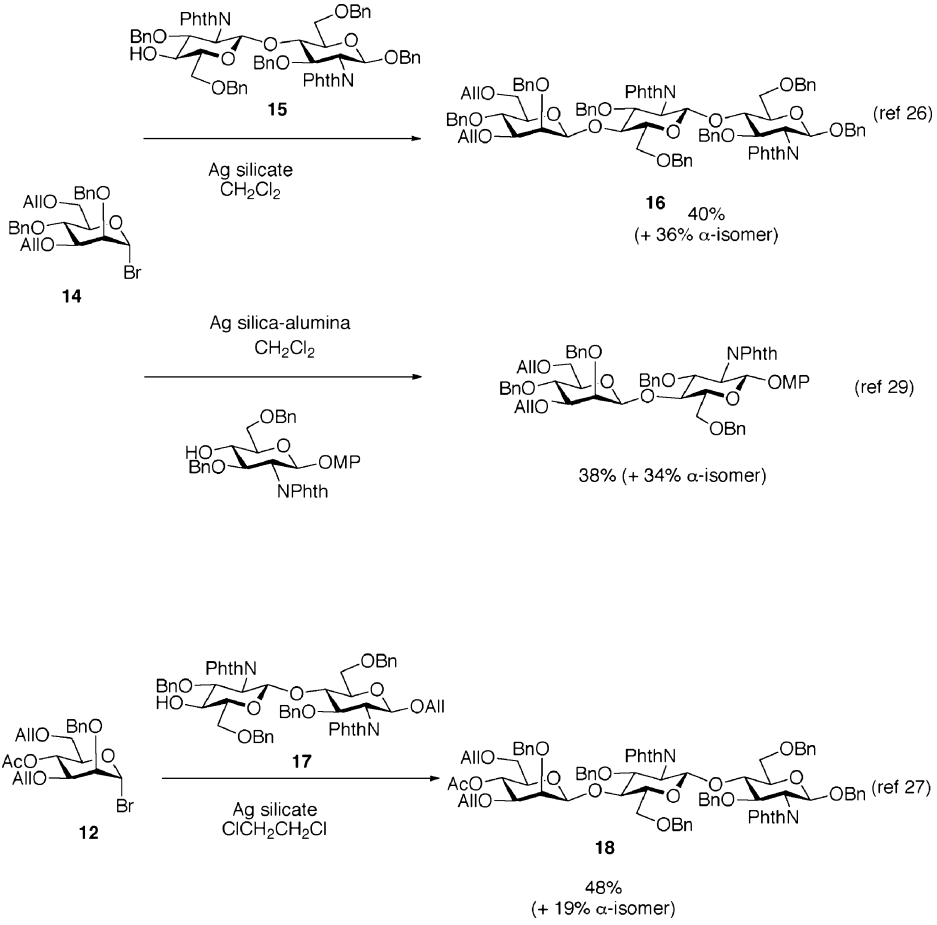
The stereoselectivity of the insoluble silver salt-mediated glycosylation is also affected by the protecting groups on the glycosyl donor. It is well known that the reactivity of the glycosyl donor is enhanced as the number of ether-type protecting groups (i. e. benzyl, allyl,



Scheme 8

etc.) increases, while the ester-type (i. e. acetyl, benzoyl, etc.) group deactivates the glycosyl donor [6,31]. Since the successful β-mannosylation requires a rapid nucleophilic displacement (see *Scheme 6*), it can be anticipated that a mannosyl donor having an acyl protecting group should be less β-selective. In fact, the mannosyl bromide **9b** having two acetyl groups was shown to give nearly equal amounts of α- and β-linked disaccharides, in sharp contrast to the reaction using **9a** [21] (see *Scheme 8*).

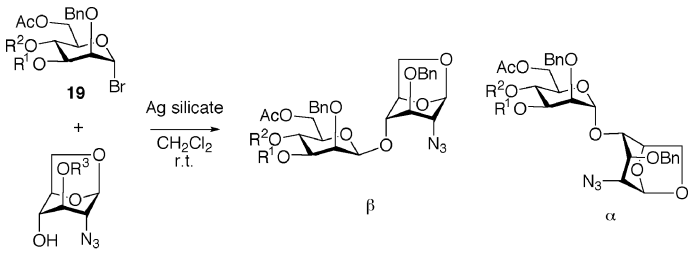
The effect of the electron-withdrawing acyl group on the stereoselectivity seems to be position-dependent. Systematic studies by van Boeckel and coworkers revealed an interesting effect of the acyl group in the variously protected mannosyl bromide **19** [32] (*Scheme 10*). Compared



Scheme 9

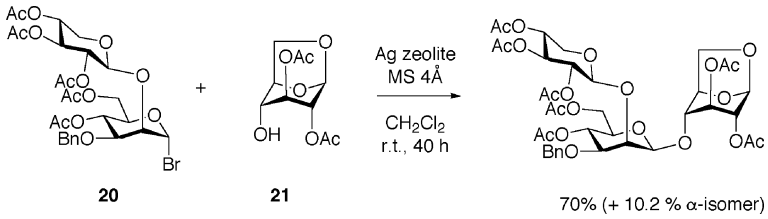
to the tri-*O*-benzyl-protected bromide (entry 2), β-selectivity increased from 6:1 to 9:1 when an additional acetyl group was introduced at the 4-position (entry 4), while a 3-*O*-acetylated donor gave no stereoselectivity (entry 1). This effect was also confirmed by Ogawa et al., who demonstrated that the 4-*O*-Ac carrying bromide **12** gives a significant level of β-selectivity (β:α = 2.5:1), even when applied to the sterically hindered acceptor **17** [27] (see *Scheme 9*). The resultant trisaccharide **18** was converted into the octasaccharide which has a bisecting GlcNAc as well as a fucose residue linked to the innermost GlcNAc.

An even more impressive result was reported by Takeda and coworkers [33] (*Scheme 11*). In their synthetic studies on the spermatozoa-derived glycolipid, the disaccharide donor **20** was coupled with **21** by the action of silver silicate. The reaction proceeded in a remarkably selective manner (α:β = 1:7) to afford the desired disaccharide in 70% yield. This is one of a very few examples of the direct formation of a β-*manno*-glycoside using an oligosaccharide

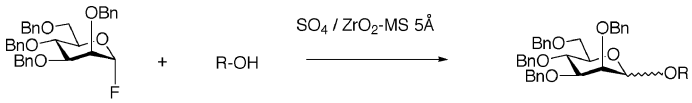


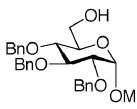
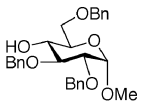
entry	1	2	3	4	5	6
R^1	Ac	Bn	TCA	Bn	Ac	Bn
R^2	Bn	Bn	Bn	Ac	TCA	All
R^3	Bn	Bn	Bn	Bn	Bn	Ac
β:α	1:1	6:1	1:1.2	9:1	>10:1	5.1:1
yield (%)	80	76	40	45	20	67

Scheme 10



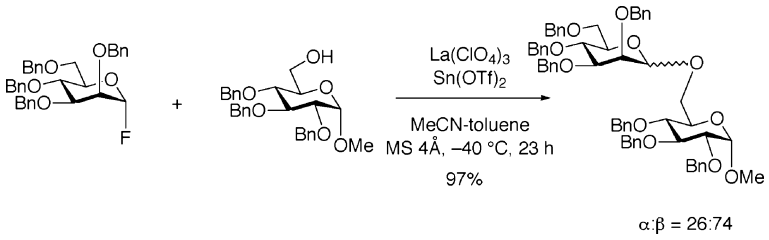
Scheme 11



R-OH	Solvent	yield (%)	β:α
	Et_2O	84	73:27
	MeCN	88	3:97
	Et_2O	55	44:56
	MeCN	75	2:98

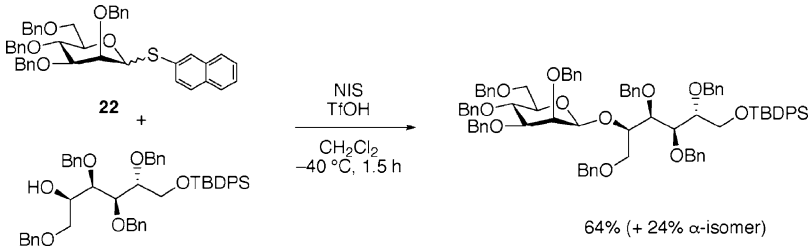
Scheme 12

donor and nicely underscores the beneficial effect of the 4-*O*-Ac group. In this particular case, silver zeolite gave significantly better results as compared to silver silicate (α:β = 1:2.7).
The insoluble silver salt method was also applied to the β-selective syntheses of gluco, galacto, 2-deoxy-2-azido, and 2-deoxy glycosides [34,35,36].
Toshima and coworkers reported a novel type of insoluble promoter (SO_4-ZrO_2) which is effective for the modestly selective synthesis of β-*manno*-glycoside from the fluoride [37] (● *Scheme 12*). Interestingly, the β-selectivity observed in diethyl ether was completely lost when the reaction was performed in acetonitrile which gives the α-glycoside with high selectivity. It should be noticed that the usually observed directing effects of diethyl ether and acetonitrile are reversed in this particular system.
A partially successful β-mannosylation with a fluoride was also reported by Shibasaki and coworkers [38]. As depicted in ● *Scheme 13*, $La(ClO_4)_3$-$Sn(OTf)_2$ was effective, promoting the reaction with a primary alcohol in a β-selective manner (α:β = ~1:3).



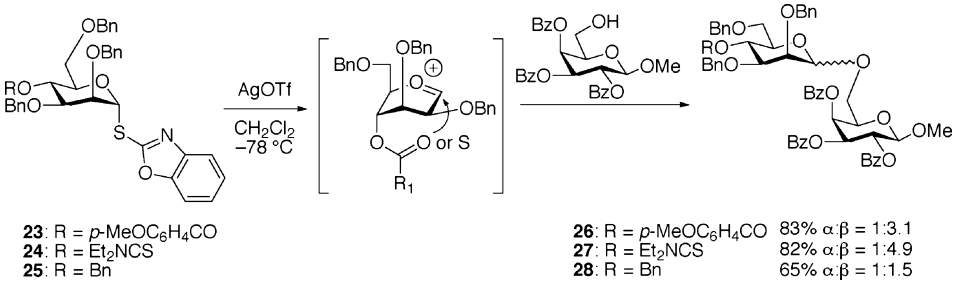
● Scheme 13

Tatsuta and Yasuda reported the interesting use of the naphthylthio glycoside **22** which gives the β-manno isomer stereoselectively, by the action of NIS-TfOH [39] (● *Scheme 14*). The origin of this selectivity still awaits clarification by further systematic investigations.



● Scheme 14

Demchenko et al. developed 1-*S*-benzoxazolyl (SBox) mannoside **23**–**25** as potent donors for the synthesis of β-mannoside [40]. They demonstrated that an acyl or thiocarbamoyl group at the *C*-4 position gave a favorable effect, possibly by participating to the *C*-1 position. Nucleophilic attachment of an acceptor occurs predominantly from the β-face [40] (● *Scheme 15*).



Scheme 15

These direct glycosylations with non-conventional donors have so far met with only limited success. However, subtle modification of the reaction conditions might further enhance their efficiency to a generally useful level, as can be seen in the case of sulfoxide-based technology developed by Crich and coworkers.

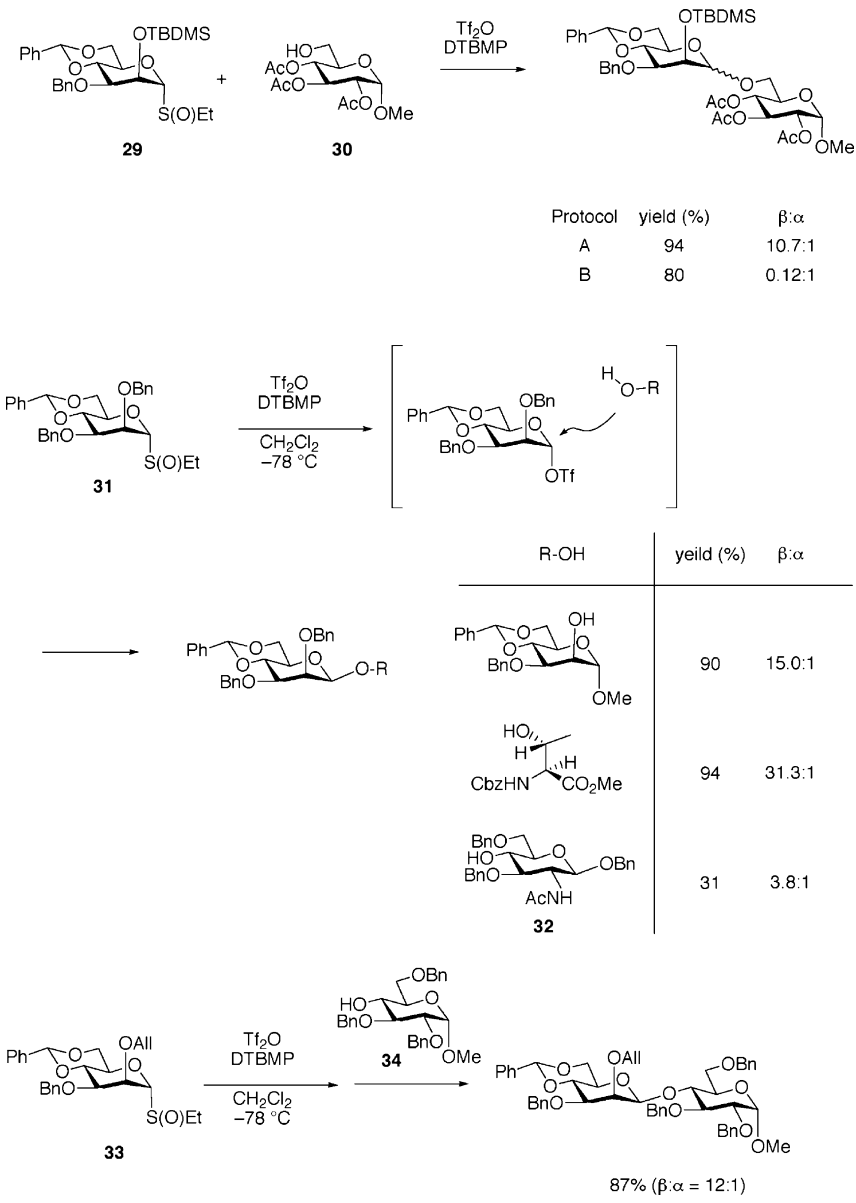
Glycosyl sulfoxide was originally developed by Kahne and proved to be an extremely powerful glycosyl donor that can be activated at low temperature by Tf_2O [41]. Initial attempts by Crich and Sun to apply this methodology to β-mannosylation were performed by using **29** in diethyl ether-benzene at −78 °C [42] (*Scheme 16*). When the acceptor **30** was added to a solution of the sulfoxide preactivated by Tf_2O and 2,6-di-*tert*-butyl-4-methylpyridine (DTBMP), the predominant formation of the β-isomer was observed (Protocol **A**). On the other hand, addition of Tf_2O to the mixture of **29**, **30**, and DTBMP afforded the α-glycoside as the major product (Protocol **B**). A higher β-selectivity was obtained when **31**, having a less bulky substituent at the 2-position, was used as a donor in dichloromethane [43] (*Scheme 16*).

The authors hypothesized the in situ formation of α-glycosyl triflate which reacts with the glycosyl acceptor in an S_N2 manner to give the β-glycoside. Although the reaction with the 4-hydroxy group of *N*-acetylglucosamine **32** proceeded with lower efficiency, the glucose-derived acceptor **34**, which is also known to be a hindered aglycon, was very successfully reacted with **33** to afford the corresponding disaccharide with an excellent β:α ratio [44]. Further transformation into the trisaccharide component of *Hyriopsis schlegelii* glycosphingolipid has been achieved.

Furthermore, these authors were also able to perform the β-mannosylations with the thioglycoside **35** under the action of PhSOTf [45] or 1-benzenesulfinyl piperidine (BSP)-Tf_2O [46] (*Scheme 17*). In the cases of both glycosyl sulfoxide and thioglycoside, protection of the 4- and 6-positions by a cyclic acetal is critical. The corresponding perbenzylated donors gave poor β-selectivity [47].

Mechanistic analysis based on the kinetic isotope effect suggested that the reaction proceeds through an oxocarbenium ion-like intermediate, such as the contact ion pair **37**, although the covalent linked α-triflate **36** was detected by NMR [48] (*Scheme 18*). They systematically investigated the effects of protective groups [49,50,51,52,53,54], including 2-*O*-propargyl ether (**38**) [52,53], and 3-*O*-1-naphthylpropargyl ether (**40**) [54] (*Scheme 19*).

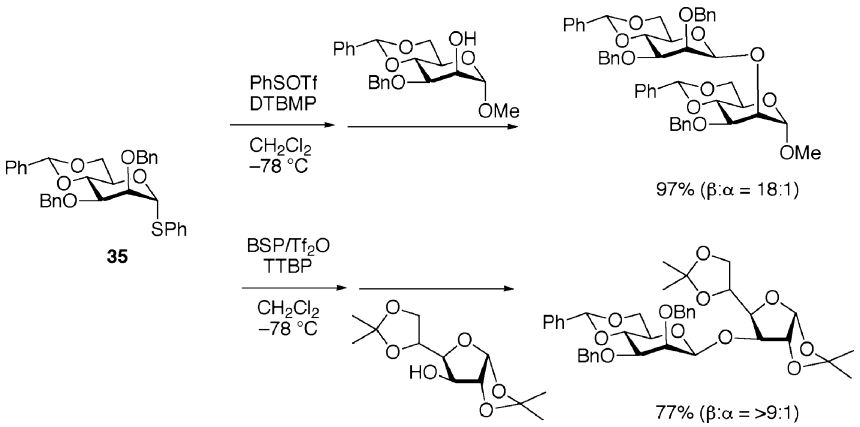
The practicality of direct glycosylation methodology was demonstrated in the context of the synthesis of β-mannan [55,56] and β-rhamnoside (6-deoxy-β-mannoside) [57,58,59]. In the latter case, a construction of the β-mannoside by the sulfoxide method was followed by



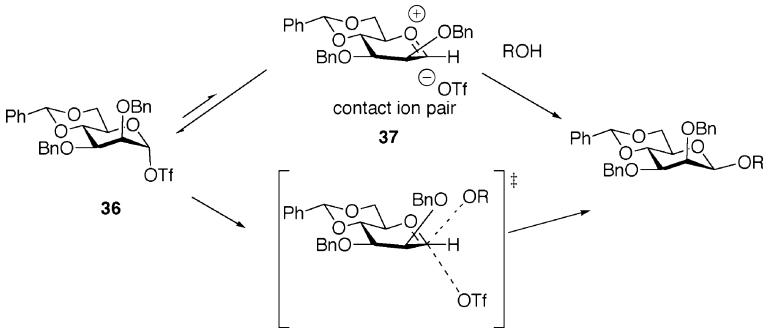
Scheme 16

reductive radical fragmentation of the 4,6-cyclic acetals. The methodology was also applied L-glycero-β-D-mannoheptopyranoside [60,61].

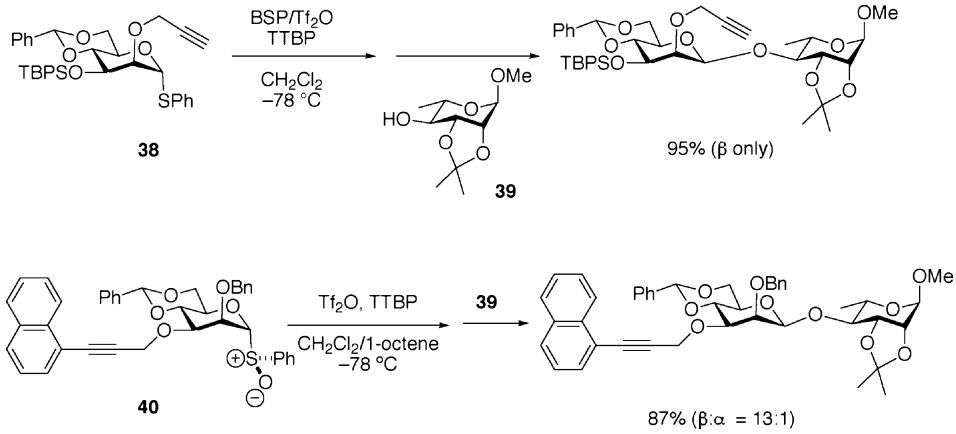
An intriguing system using mannosyl dimethylphosphinothioate **41** was developed by Inazu and coworkers [62] (*Scheme 20*). When reacted with glucose-derived acceptor **42** in the



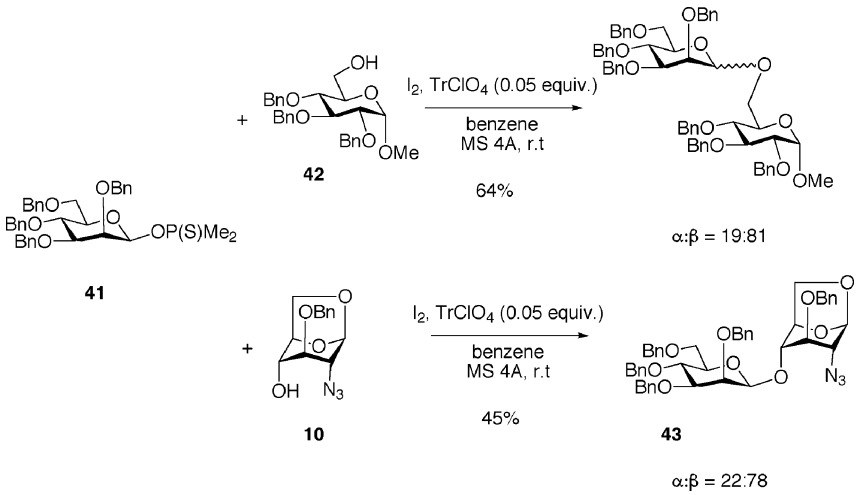
Scheme 17



Scheme 18



Scheme 19



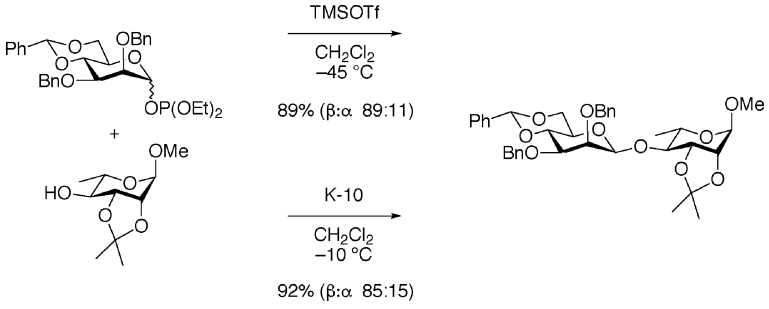
Scheme 20

presence of a catalytic amount of $TrClO_4$ together with 1 equiv. of I_2, the β-glycoside was obtained as the major isomer in a ratio of 81:19. Masked β-Man1→4GlcNAc **43** can also be prepared with a substantial level of selectivity in the same manner.

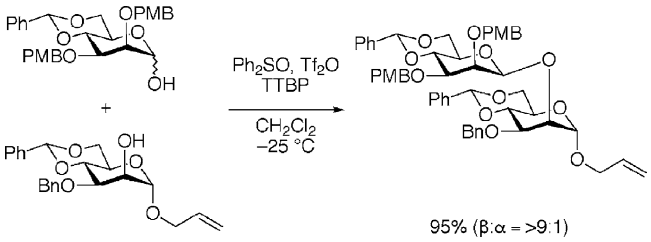
Hashimoto and coworkers successfully applied their phosphite method to selective β-mannosylation [63,64,65] (*Scheme 21*). Similar to the case of Crich's mannosylation, cyclic protection at the 4- and 6-positions is essential. Mannosyl phosphate and phosphoramidate were also investigated [65,66]. Toshima et al. used the glycosyl phosphites activated with a solid acid catalyst to give β-mannoside predominantly [67,68].

Dehydrative glycosylation, which utilizes reducing sugar as a donor, was developed by Gin et al. [69]. Seeberger et al. demonstrated that this reaction can be applied to the formation of β-mannosidic linkages [70] (*Scheme 22*).

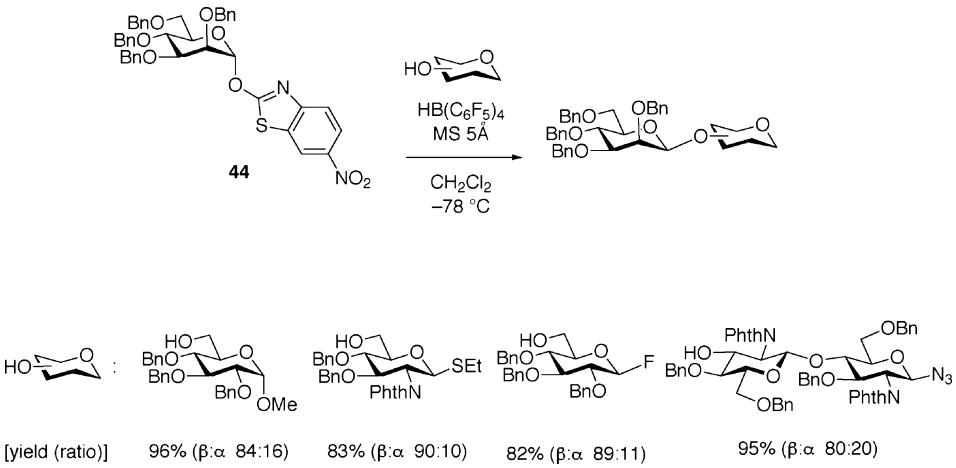
Mukaiyama et al. examined 6-nitrothiazolyl mannoside **44** as a glycosyl donor [71,72] (*Scheme 23*). They reported that in the presence of a catalytic amount of $HB(C_6F_5)_4$,



Scheme 21



Scheme 22

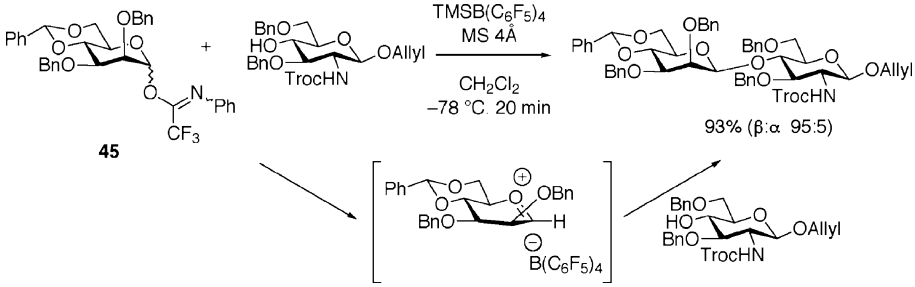


Scheme 23

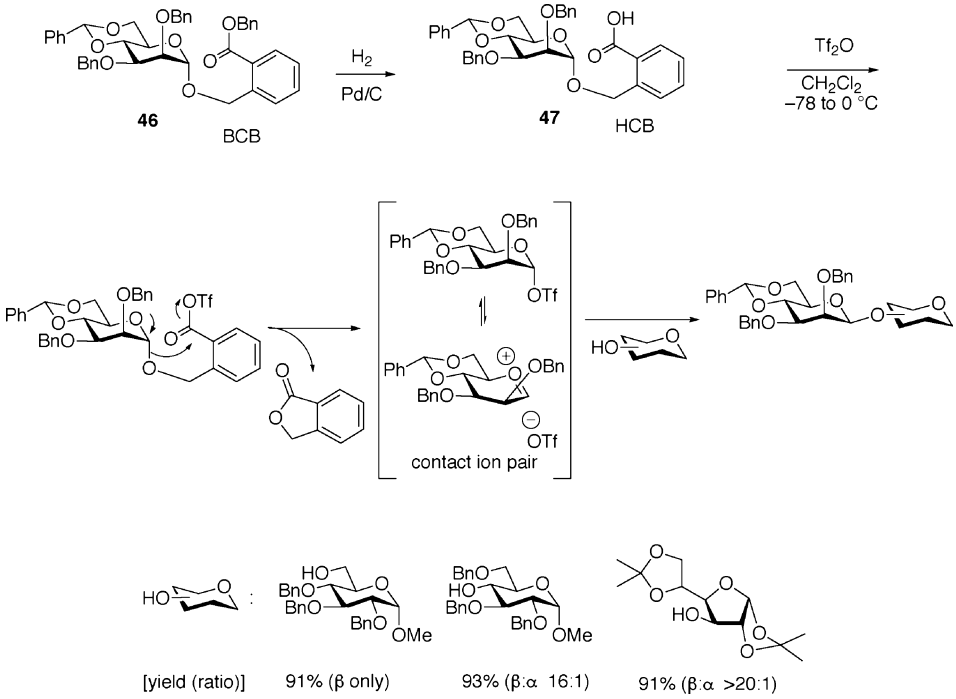
this donor gave β-mannoside in a substantially selective manner. This method was also applied to the synthesis of the core pentasaccharide of *N*-glycan [73].

Trichloroacetimidate developed by Schmidt has been recognized as one of the most powerful glycosyl donors that has ever been developed. They examined its suitability for β-mannosylation [74]. It was demonstrated that preferential β-mannoside formation was observed, when 4,6-*O*-benzylidene protection was introduced to the donor. More recently, Fukase et al. found that the β-mannoside was obtained with high stereoselectivity from the mannosyl-*N*-phenyltrifluoroacetimidate **45** by the action of a catalytic amount of $TMSB(C_6F_5)_4$, which was prepared from $AgB(C_6F_5)_4$ and TMSCl [75] (*Scheme 24*). The counter anion, $B(C_6F_5)_4^-$, would shield the α-face of the glycosyl cation, possibly due to the steric repulsion with the OBn group at the 2-position. They reported that the corresponding trichloroacetimidate gave lower selectivity.

2-(Hydroxycarbonyl)benzyl (HCB) mannoside **47**, developed by Kim et al. was shown to be a potent donor for the synthesis of β-mannosides [76]. It was generated from 2-(benzyloxycarbonyl)benzyl (BCB) derivative **46** through catalytic hydrogenolysis (*Scheme 25*).



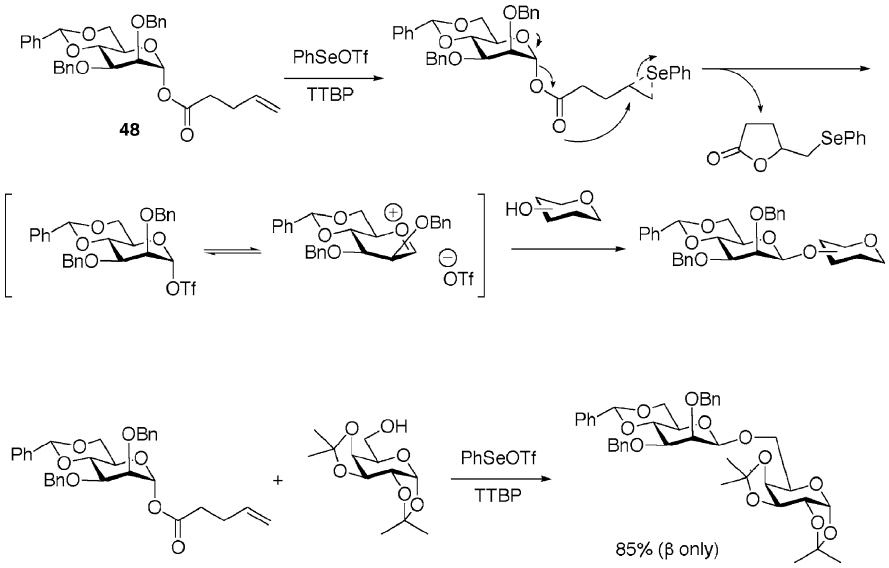
Scheme 24



Scheme 25

In a similar manner, syntheses of other types of 1,2-*cis*-glycosides, such as β-arabinofuranoside and α-galactofuranoside were achieved with high stereoselectivity [77,78].

Kunz et al. and Fraser-Reid et al. explored glycosyl pentenoate as a novel type of donor [79,80]. Kim et al. applied it to the highly stereoselective synthesis of β-mannosides [81]. In this case, mannosyl pentenoate **48** was treated with PhSeOTf [82] in the presence of 2,4,6-tri-tert-butylpyrimidine (TTBP) (*Scheme 26*).



Scheme 26

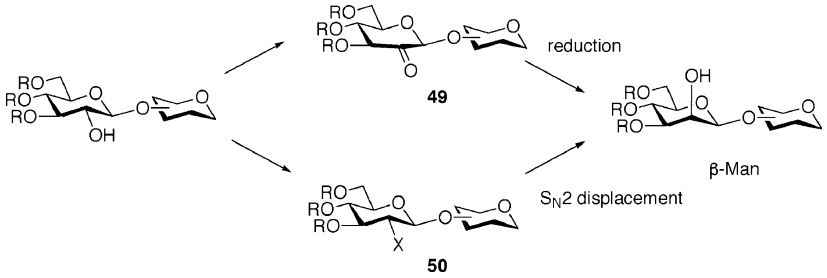
3 *C*-2 Inversion Approach

3.1 Reduction of 2-Oxo Sugars

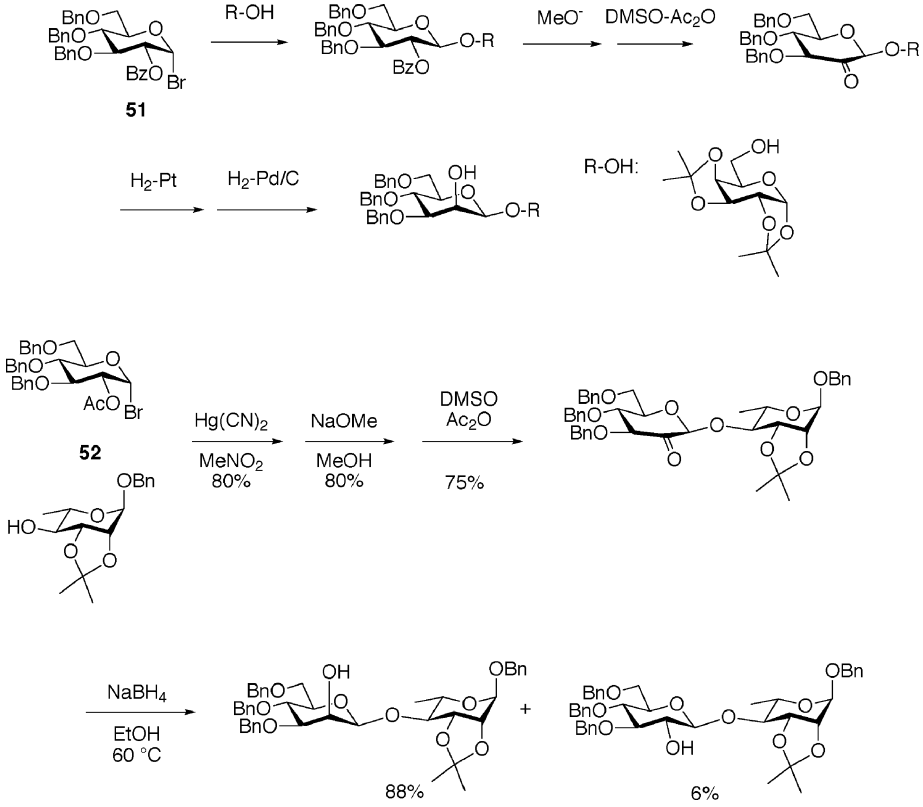
Since the β-gluco glycoside can be readily synthesized stereoselectively [6], the use of an initial glycosylation product having this configuration as a precursor of β-*manno*-glycoside is a rational tactic. Two possibilities can be considered to achieve the overall transformation of a gluco- into a *manno*-isomer (*Scheme 27*). The first option is oxidation-reduction in which the 2-uroside **49** is the key intermediate. The drawback of this approach would be that the stereoselectivity of the reduction step is not always perfect and the product often contains the β-D-gluco-isomer as a minor component. An alternative approach consists of derivatization of the 2-OH into certain leaving groups (X in **50**) that is to be subjected to an S_N2-type displacement by an *O*-nucleophile.

Whichever tactic is employed, the protecting groups should be strategically chosen so that the C2 position can be selectively manipulated.

Historically, the oxidation-reduction protocol was the only reliable method to prepare the β-*manno*-glycoside until the insoluble silver salt method was established. In 1972, Ekborg et al. reported the use of the 2-*O*-benzoylated glucosyl bromide **51** for this purpose [83] (*Scheme 28*). Glycosylation with diacetone galactose and selective deprotection of the 2′-position followed by DMSO oxidation afforded the 2-uroside which was reduced by hydrogenation over Pt. The use of the 1,2-*O*-orthoester for the same purpose was also reported [84].



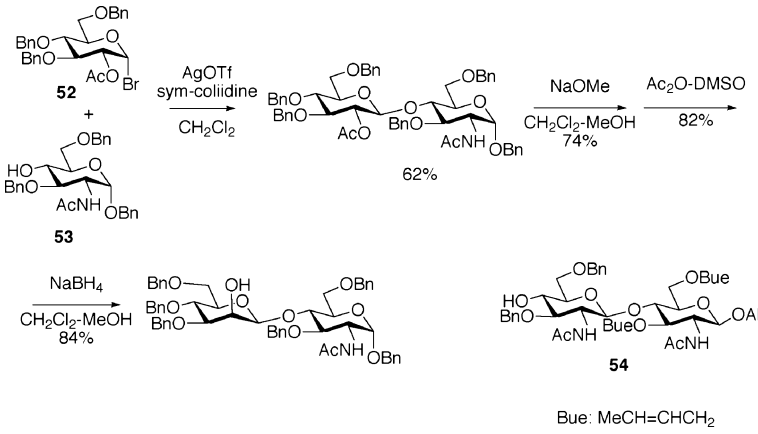
Scheme 27



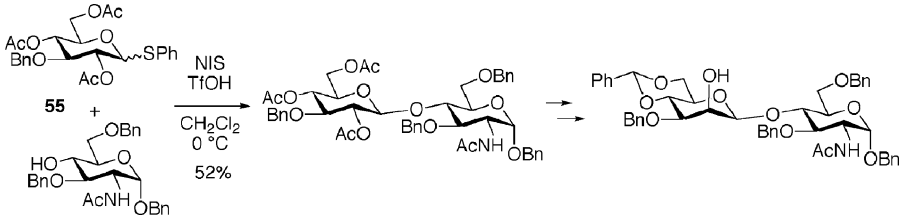
Scheme 28

Subsequently, Kochetkov used the 2-*O*-acetylated **52** and demonstrated that stereoselective reduction can be realized conveniently by $NaBH_4$ [85] (*Scheme 20*).

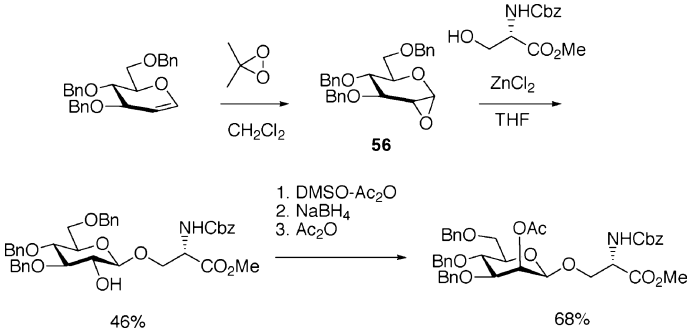
A similar strategy was used by Jeanloz and coworkers in their synthesis of the trisaccharide unit (β-Man1→4-β-GlcNAc1→4GlcNAc) of an Asn-linked oligosaccharide [86,87] (*Scheme 29*). Coupling of **52** with the mono- (**53**) and the disaccharide (**54**) acceptor afford-



Scheme 29



Scheme 30



Scheme 31

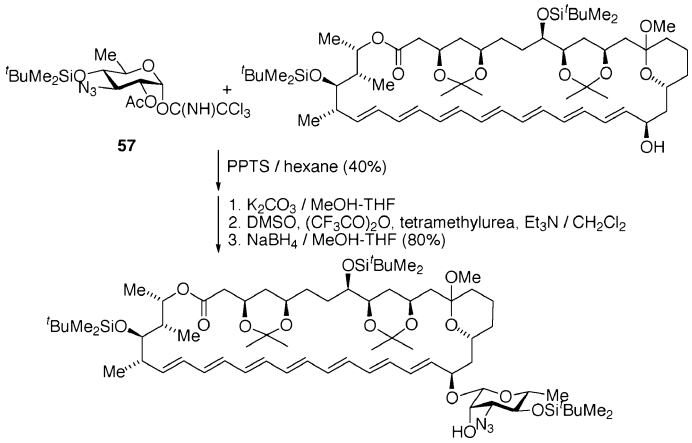
ed the corresponding β-glycosides. Subsequent oxidation and borohydride reduction afforded the β-*manno*-glycosides.

As a more recent example, Khan and Matta utilized the thioglycoside **55** [88] (*Scheme 30*). Glycosylation with the GlcNAc component was effected by NIS-TfOH.

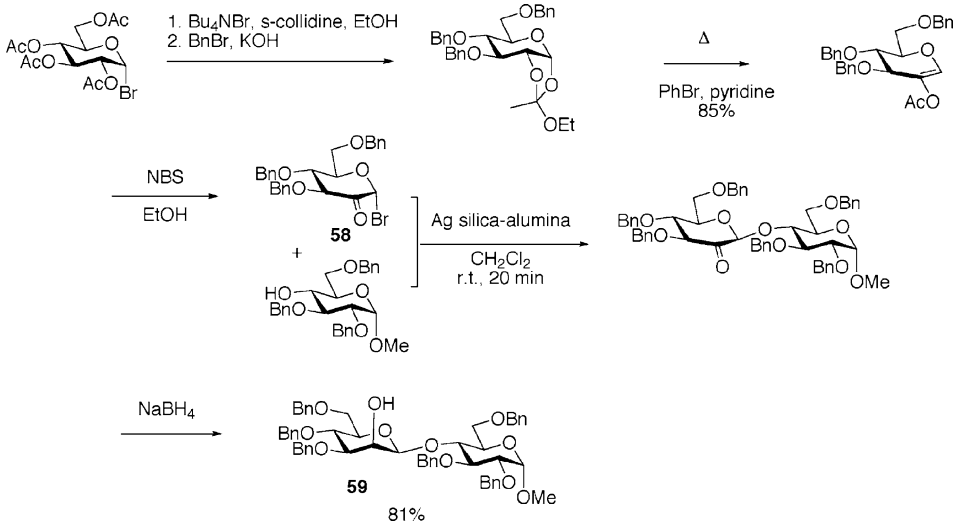
In combination with the oxidation-reduction protocol, the glycal-based oligosaccharide assembling strategy [89] which has been extensively developed by Danishefsky gives an expedient

synthetic route to β-*manno*-glycoside [90] (› *Scheme 31*). The glucal-derived 1,2-anhydro sugar **56** was coupled with a glycosyl acceptor to afford the β-glycoside that has a free 2-OH group. Subsequent transformation to the β-*manno*-glycoside was performed in a standard manner.

Fraser-Reid et al. have reported the utilization of the *n*-pentenyl orthobenzoate for the synthesis of octa-β-1,2-mannan through the repetition of β-selective glucosylation followed by the oxidation-reduction protocol [91].



Scheme 32



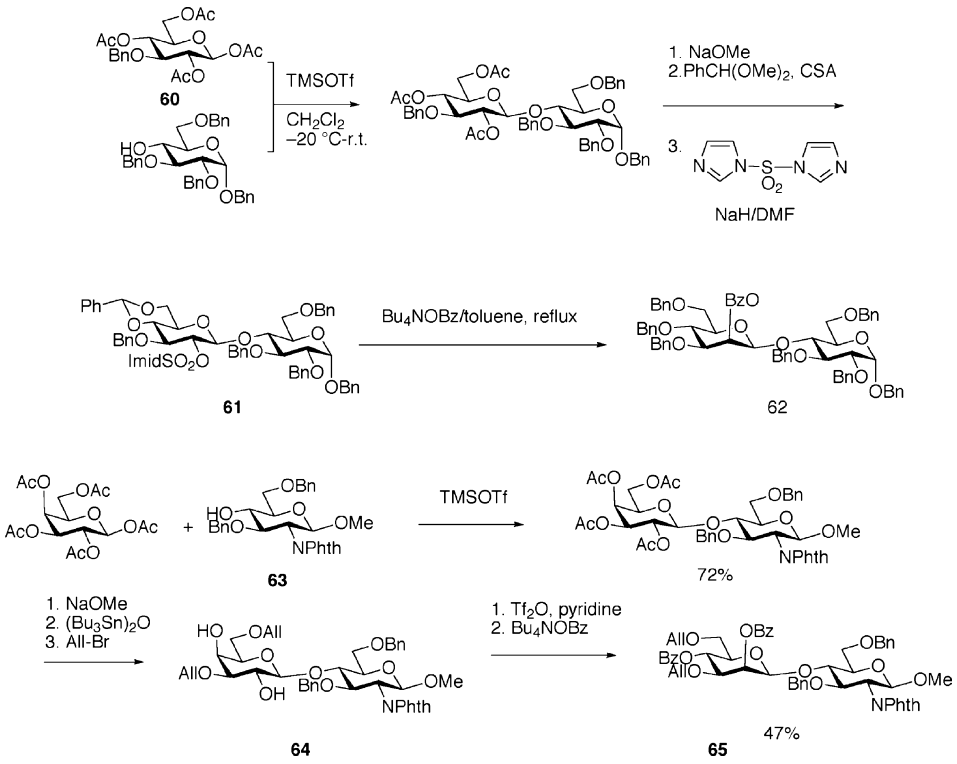
Scheme 33

In their landmark total synthesis of polyene macrolide Amphotericin B [92], Nicolaou and coworkers designed a mycosamine precursor **57** and coupled it with a highly functionalized aglycon (◉ *Scheme 32*). Inversion of the C2-OH followed by the azide reduction and deprotections competed the total synthesis.

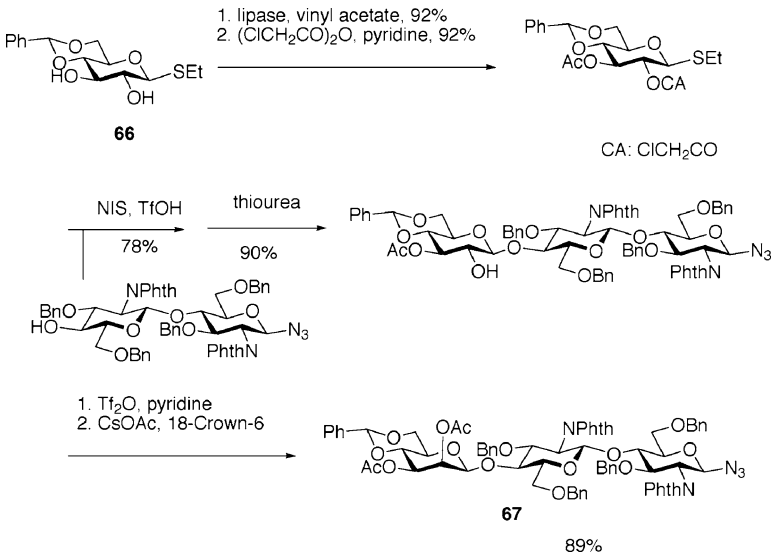
A smart variant of the oxidation-reduction approach was developed by Lichtenthaler and coworkers, who used 2-urosyl bromide as a donor [93,94,95]. On account of the α-halo ketone structure, an S_N2-like reaction at the anomeric position is facilitated. The required bromide **58** can be very conveniently prepared from acetobromoglucose (◉ *Scheme 33*). As expected, Ag salt-promoted glycosylation afforded the β-glycosides that can be reduced selectively to give the β-mannoside **59**.

3.2 S_N2-Type Inversion of the 2-Position

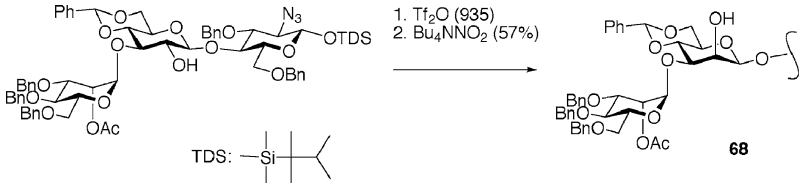
Inversion of the *C*-2 stereochemistry can be performed by S_N2-type displacement, after installing a proper leaving group at this position. In their series of publications, David and coworkers used the imidazyl and the triflate groups for this purpose in two different ways [96,97] (◉ *Scheme 34*). They used the 3-*O*-benzylated glucosyl acetate **60** as a donor



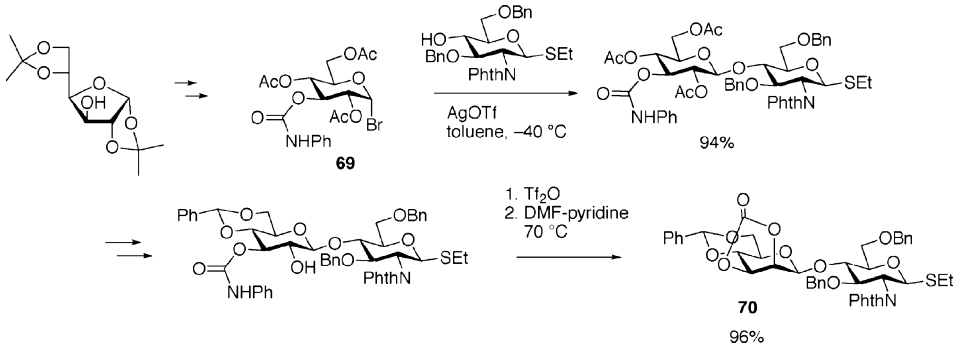
◘ **Scheme 34**



Scheme 35



Scheme 36



Scheme 37

which was coupled with an acceptor by the action of trimethylsilyl triflate (TMSOTf). Subsequent derivatization into the 4,6-*O*-benzylidene-protected 2-*O*-imidazylate **61**, followed by nucleophilic displacement with Bu_4NOBz afforded **62** in high overall yield. Alternatively, glycosylation of **63** was performed with galactose pentaacetate. Subsequent regioselective protection was achieved by using the tin oxide method to give **64**, which was transformed into the di-*O*-triflate. Treatment with Bu_4NOBz gave doubly inverted product **65**.

The double stereochemical inversion of β-Gal to β-Man, demonstrated by Sato et al., can be performed more conveniently by using a 3,6-di-*O*-pivalyl-protected 2,4-di-*O*-triflate that was treated with CsOAc to afford the β-mannoside [98].

Ajisaka et al. has developed a remarkably efficient synthetic route to the core trisaccharide of Asn-linked oligosaccharide, by the combined use of enzyme-assisted regioselective protection and glycosylation-inversion protocols [99]. Thus, starting from the 4,6-*O*-benzylidene-protected thioglycoside **66**, transformation into the 3-*O*-acetate was achieved in high yield by using lipase from *Pseudomonas fluorescens* (Amano AK^{TM}) in the presence of vinyl acetate. Chloroacetylation, glycosylation with a chitobiose derivative, deprotection of the 2-position and triflation and S_N2 inversion with CsOAc in the presence of 18-crown-6 gave **67** (*Scheme 35*).

Schmidt reported the versatile strategy for the synthesis of complex type, Asn-linked oligosaccharides in which the core trisaccharide **68** was prepared as depicted in *Scheme 36*. In their case, inversion of the *C*-2 position was performed by using Bu_4NNO_2 as a nucleophile [100].

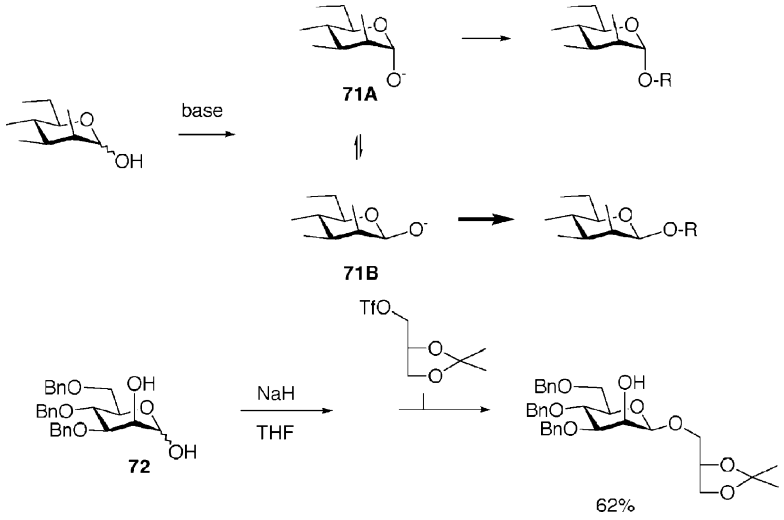
Gunther and Kunz developed a highly efficient, intramolecular S_N2-type reaction for the conversion of β-glucoside into a β-mannoside [101,102] (*Scheme 37*). Starting with glucosyl donor **69** equipped with a 3-*O*-phenylcarbamoyl group, the 1,2-*trans* glycoside was prepared. Further transformation into the corresponding 2-*O*-triflate was followed by heating to 70 °C in DMF-pyridine, which afforded the β-mannoside **70** after acidic work-up. Application of this highly efficient protocol to various types of Asn-linked oligosaccharides was extensively investigated by Unverzagt [103,104,105,106,107,108,109].

4 Other Approaches

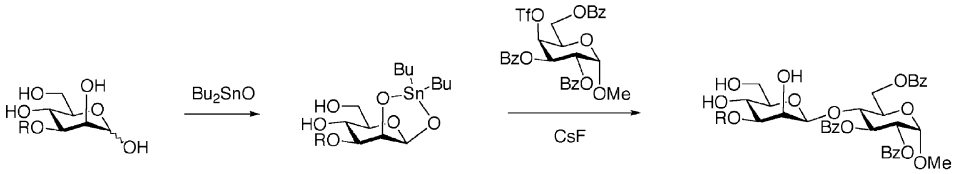
4.1 *O*-Alkylation Method

Stereochemical control in *O*-alkylative β-glycosylation relies on the kinetic anomeric effect [110]; namely, that the oxyanion derived from a reducing sugar is more reactive in its β-oriented form **71B** than the corresponding α-anion **71A** which should be more abundant. This tactic has met with partial success in cases of reactive substrate [111] (*Scheme 38*). For example, reaction of **72** with a primary triflate in THF in the presence of NaH afforded the β-glycoside in 62% yield.

Srivastava and Schuerch [112] and Dessings et al. [113]. reported on the potential utility of 1,2-*O*-stannylene acetals in β-*manno*-glycoside synthesis. Recently, Hodosi and Kovác established a highly efficient β-mannosylation process involving alkylation of a 1,2-*O*-stannylene acetal with the triflate derived from an aglycon [114] (*Scheme 39*). A remarkable feature of this method is its extreme simplicity. Even free mannose can be used as a precursor of



Scheme 38



73	R
a	H
b	Bn

74	R	solvent	yield (%)
a	H	DMSO	59
b	Bn	DMF	67

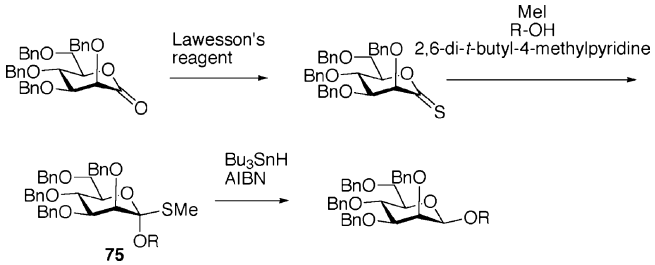
Scheme 39

stannylene acetal which was reacted with galactose 4-*O*-triflate to give the β1→4-linked disaccharide **74a**. However, a significantly higher yield was obtained by using the 3-*O*-protected mannose **73b**, because stannylene acetal migration (from 1,2- to 2,3-positions) can be blocked in this way [115]. Undoubtedly, this method can be seen as one of the most practical and efficient methods for the synthesis of β-*manno*-glycoside.

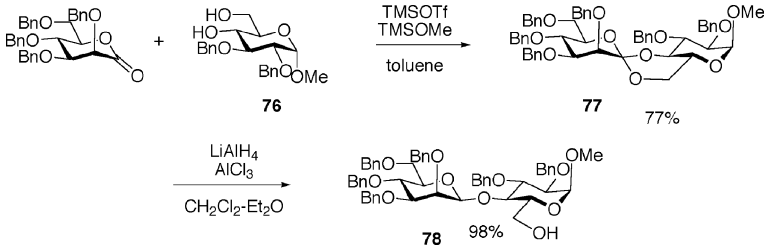
4.2 Reduction of the Anomeric Position

Reductions at anomeric positions, both by ionic and radical mechanisms, deliver hydrogen from the axial direction. Kahne prepared the hemithio orthoester **75** from thiolactone and subjected it to tin hydride reduction, which resulted in predominant formation of the β-glycoside [116] (*Scheme 40*).

A similar concept was applied for the synthesis of 2-deoxy-β-glycoside where Barton reductive decarboxylation was utilized as the key transformation [117].



Scheme 40



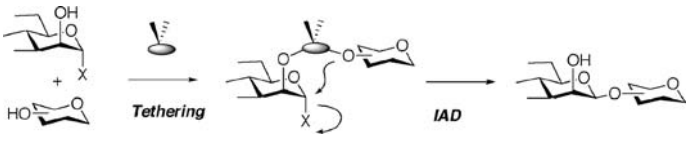
Scheme 41

Iimori et al. reported the metal hydride reduction of glycosylidene acetals (anomeric orthoesters), which constitutes an efficient synthetic strategy well-suited for the β-Man1→4Glc (and probably β-Man1→4GlcNAc) linkage (*Scheme 41*). The required compound **77** was prepared from the lactone by treatment with TMSOTf and TMSOMe in the presence of diol **76** [118,119]. Reductive opening of the orthoester was effected by $LiAlH_4$-$AlCl_3$ to afford **78** almost quantitatively. The reaction was remarkably stereo-and regioselective, and the formation of either the corresponding α-isomer or the 1→6-linked disaccharide was not detected [120,121].

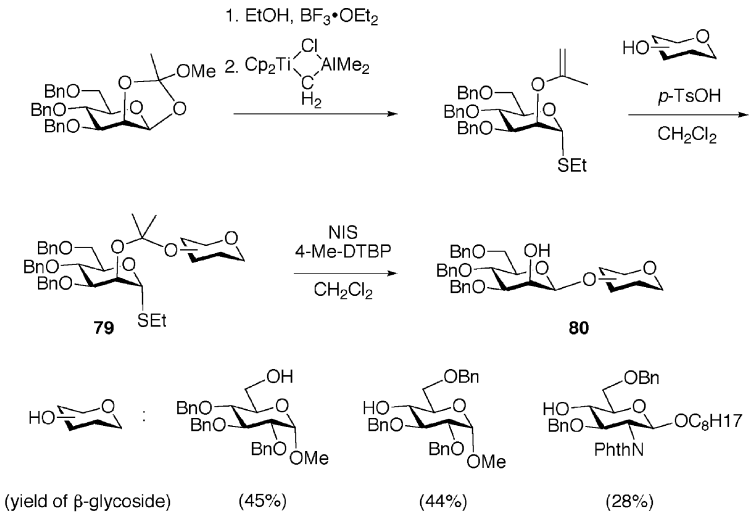
4.3 Intramolecular Glycosylation

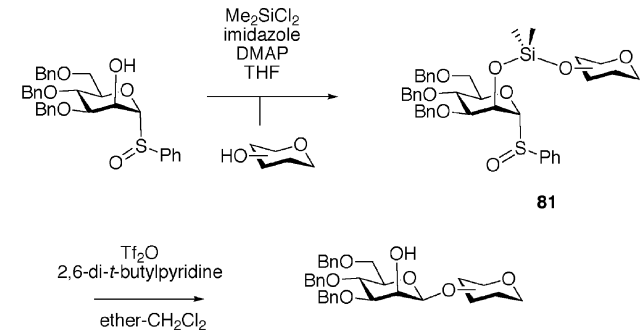
Among the various types of intramolecular glycoside bond formation [122], a strategy called intramolecular aglycon delivery (IAD) which was first proposed by Barresi and Hindsgaul [123] has been demonstrated to be well suited for β-manno glycosylation. Because of the stereochemical constraint, this approach should give the β-*manno*-glycoside without any stereochemical ambiguity (*Scheme 42*). The authors prepared the mixed isopropylidene acetal **79** from the corresponding isopropenyl ether and alcohol. Subsequent treatment with NIS gave the β-mannoside **80** [123,124,125] (*Scheme 43*). Shortly after Hindsgaul's report, Stork and coworkers reported the use of the silaketal-like intermediate **81** for the same purpose [126,127] (*Scheme 43*).

More recently, Fairbanks et al. explored 2-*O*-isopropenyl and 2-*O*-1-(4-methoxyphenyl)vinyl ether **83**, which are obtainable from **82** [128,129,130] (*Scheme 44*). They also reported the

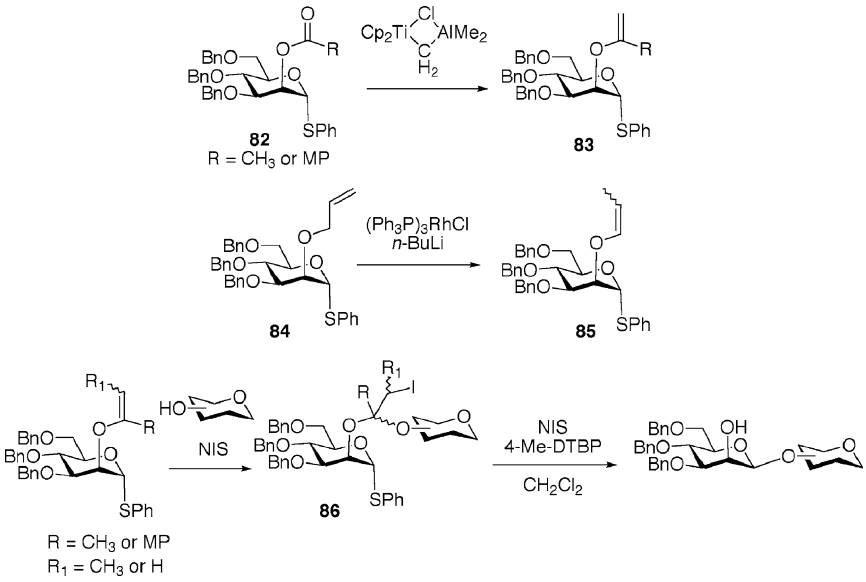


Scheme 42





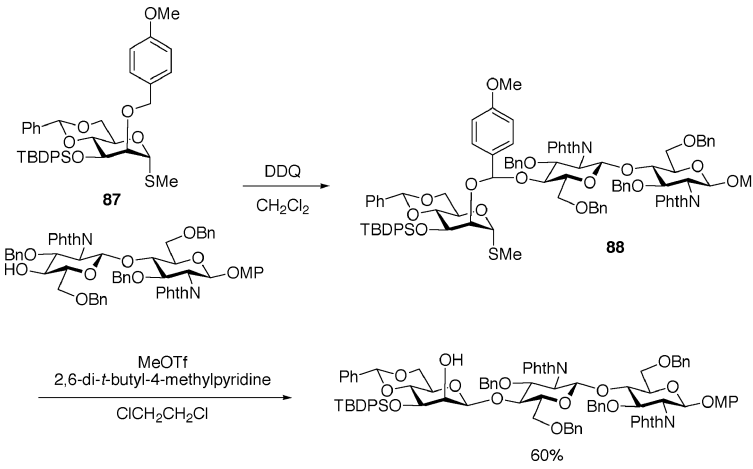
Scheme 43



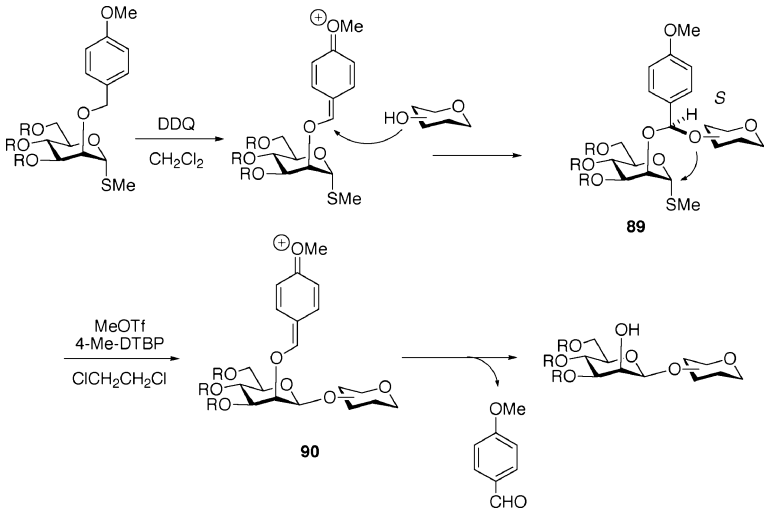
Scheme 44

use of vinyl and propenyl ether (**85**) [131,132], which were prepared by Ir-catalyzed transvinylation and [131] isomerization of allyl ether **84**, respectively. The mixed acetals **86** were synthesized from the vinyl derivatives by the iodoetherification with NIS.

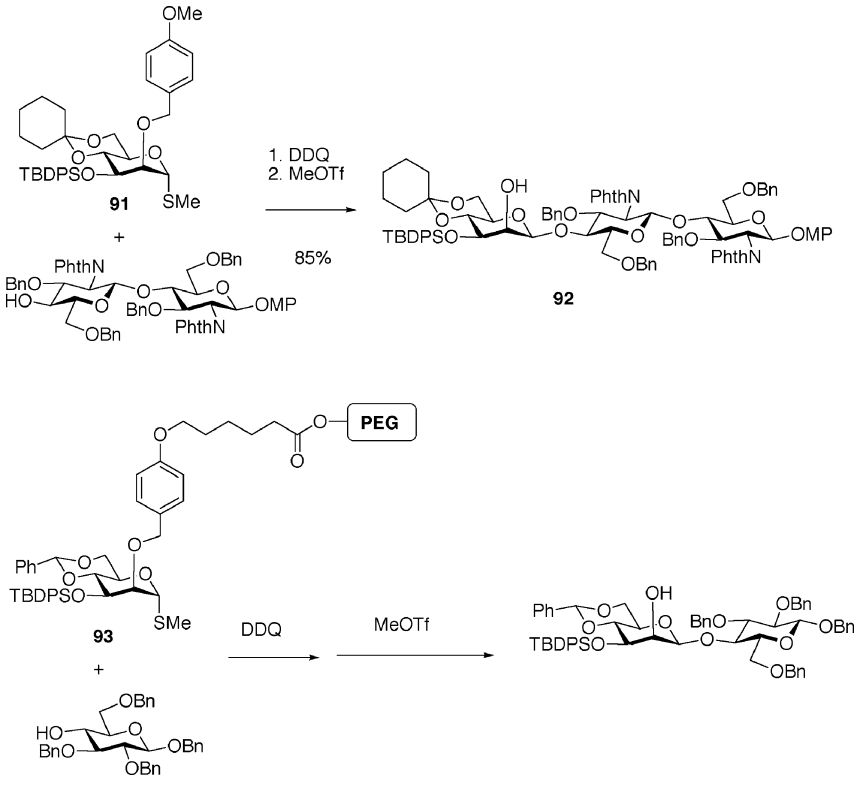
Ito and Ogawa developed an alternative type of IAD, in which the mannosyl donor equipped with a *p*-methoxybenzyl (PMB) group at the C2 position, compound **87**, was used as a donor [133] (*Scheme 45*). Treatment of **87** with the aglycon in the presence of DDQ



Scheme 45



Scheme 46



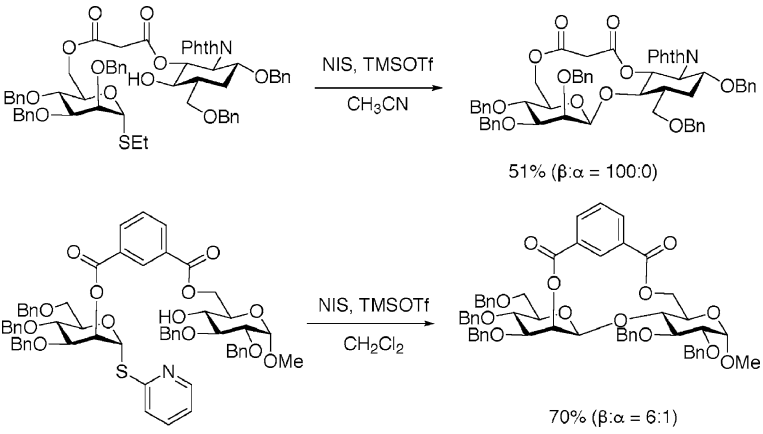
Scheme 47

afforded the mixed acetal **88**. Subsequent activation of the mannose anomeric position with the MeOTf-triggered IAD to afford the β-mannoside as a single stereoisomer in 60% yield. The stereochemistry of the mixed acetal **89** was assigned to be *S* based on NMR and computational analyses [134]. The mechanistic studies with monitoring of the reaction by NMR suggested that the non-hydrolytic pathway that converts the metastable intermediate **90** to product was operative [135] (*Scheme 46*). This PMB-assisted IAD strategy was easily applicable to the condensation of oligosaccharide fragments and a convergent synthesis of the core structure found in Asn-linked oligosaccharide was achieved in a fully stereocontrolled manner [136,137,138].

The efficiency of the PMB-assisted β-mannosylation protocol was enhanced (*Scheme 47*), when the 4,6-*O*-cyclohexylidene group was introduced. For example, compound **91** gave trisaccharide **92** in 85% yield [139]. PMB-assisted IAD has been shown to be highly practical for the construction of complex-type [140,141,142,143] as well as high mannose-type *N*-glycans [144,145,146,147].

In addition, a polymer-supported version of the intramolecular aglycon delivery was explored [148]. In this system, mannosyl thioglycoside **93** bound to polyethylene glycol (PEG) via a *p*-alkoxybenzyl linker was subjected to the two-step sequence and the resultant β-*manno*-glycoside is specifically released into the non-polymeric phase, while most of the by-products remain bound to the polymer.

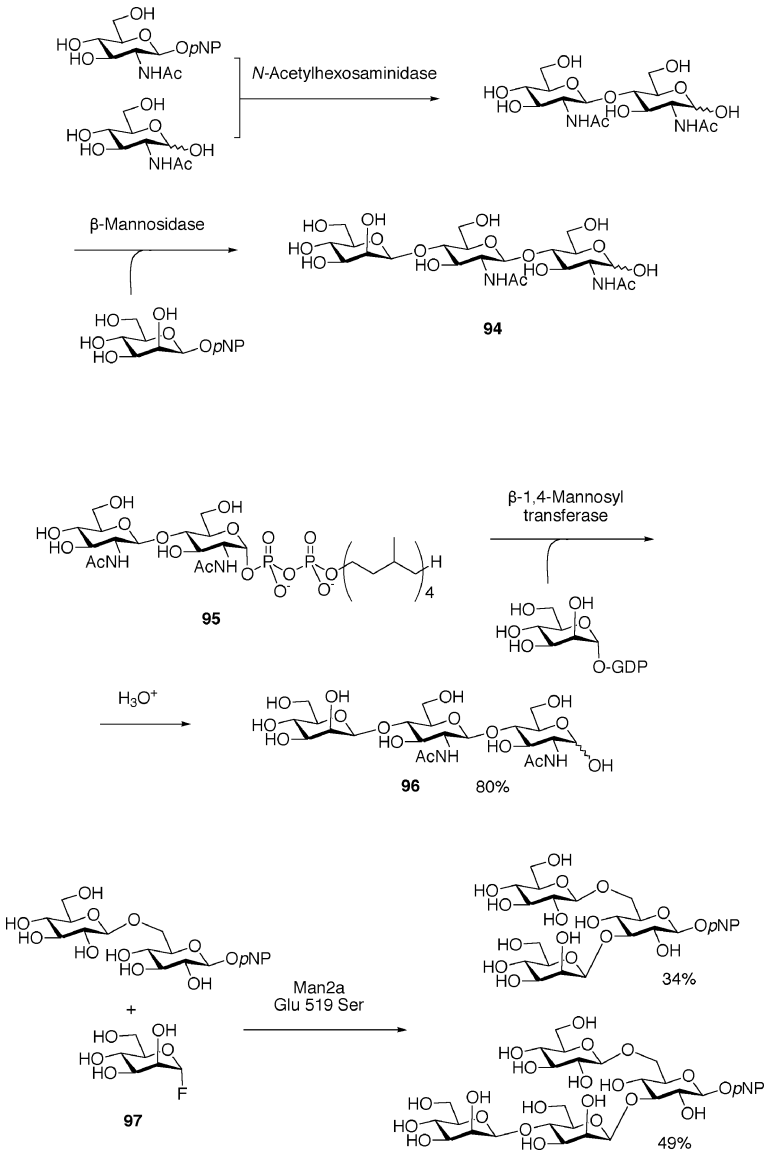
Other types of bridged strategy were developed with promise. For example, Ziegler [149,150] reported the use of a diester linker that connected the donor and the acceptor components (*Scheme 48*). A related strategy was developed by Schmidt [151].



Scheme 48

4.4 Enzymatic Synthesis of β-Manno Glycoside

Enzymatic glycosylations are rapidly attaining practical utility in complex oligosaccharide synthesis [152,153,154]. Recently, Crout and coworkers developed a highly efficient synthesis



Scheme 49

of β-Man1→4-β-GlcNAc1→4GlcNAc, a core trisaccharide of Asn-linked glycans, by using β-mannosidase [155]. Thus, enzymatically prepared chitobiose was treated with β-mannosidase from *A. oryzae*, in the presence of *p*-nitrophenyl mannoside as a glycosyl donor to afford trisaccharide **94** in 26% yield (*Scheme 49*). Since no protection/deprotection is required in the whole sequence, this procedure should be highly valuable for routine preparation of β-Man1→4GlcNAc-containing oligosaccharides.

On the other hand, an efficient expression system of recombinant yeast β1→4-mannosyl transferase in its soluble form was developed by Flitsch and coworkers [156]. Using this enzyme preparation, they performed milligram-scale β-mannosyltransferase-catalyzed glycosylation. It should be noted that easily accessible phytanol [157] carrying chitobiose **95** can be used successfully as an acceptor substrate in place of chitobiosyl dolichol phosphate (*Scheme 49*). After acidic cleavage of the glycosyl phosphodiester linkage, the β-mannosyl chitobiose **96** was isolated in 80% overall yield [156]. For larger-scale preparations, a GDP-mannose regeneration system [158], which was established for α1→2-mannosyl transferase catalyzed glycosylation, may well be applied.

The glycosynthase methodology has been successfully extended to the enzymatic synthesis of β-mannoside. The Glc519Ser mutant of the retaining β-mannosidase (Man2a) proved to afford a catalytic conversion of the various acceptors with the α-mannosyl fluoride **97** as a readily accessible donor [159] (*Scheme 49*). The chitobiose and GlcNAc derivatives, however, would not function as the acceptor.

5 Conclusion

As discussed in this chapter, a number of elegant approaches have been investigated to solve the problem of β-*manno*-glycoside synthesis. In particular, the early improvements in direct glycosylation [42,43,44,45,47,62], glycosylation-inversion [93,94,95,99,101,102,103,104,105] alkylative glycosylation of 1,2-stannylene acetal [114,115], reductive manipulation of orthoester [118,120], intramolecular aglycon delivery [123,124,125,126,127,133,136,137,138,139,148] and enzymatic glycosylation [154,155] are quite dramatic. Recent progress and promising results from a variety of strategies in the last decade, especially in direct glycosylation [46,52,53,54,65,68,73,75,76,81] and intramolecular aglycon delivery [128,131,132,143,147], suggest that practically the requirements for β-mannoside production will be more easily satisfied. By using these methods, β-mannosylation of a wide range of acceptor substrates can be performed with nearly complete stereochemical control.

References

1. Toshima K, Tatsuta K (1993) Chem Rev 93:1503
2. Shimizu M, Tojo H, Yokoyama M (1998) Synthesis 799 and literature cited therein
3. Garegg PJ (1997) Adv Carbohydr Chem Biochem 52:179
4. Schmidt R (1986) Angew Chem Int Ed Engl 25:212
5. Matsuzaki Y, Ito Y, Nakahara Y, Ogawa T (1993) Tetrahedron Lett 34:1061
6. Paulsen H (1982) Angew Chem Int Ed Engl 21:155
7. Tvaroska I, Bleha T (1989) Adv Carbohydr Chem Biochem 47:45
8. Kobata A (1993) Acc Chem Res 26:319
9. Gridley JJ, Osborn MI (2000) J Chem Soc Perkin Trans 1 1471
10. Davis BG (2000) J Chem Soc Perkin Trans 1:2137
11. Jung K-H, Müller M, Schmidt RR (2000) Chem Rev 100:4423
12. Demchenko AV (2003) Synlett 1225
13. Fairbanks AJ (2003) Synlett 1945
14. Gorin PAJ, Perlin AS (1961) Can J Chem 39: 2474
15. Bebaut GM, Dutton GGS (1974) Carbohydr Res 37:309
16. Srivastava VK, Schuerch C (1981) J Org Chem 46:1121

17. Garegg PJ, Iversen T (1979) Carbohydr Res 70:C13
18. Paulsen H, Lockhoff O (1981) Chem Ber 114: 3102
19. Garegg PJ, Ossowski P (1983) Acta Chem Scand B37:249
20. van Boeckel CAA, Beetz T (1987) Recl Trav Chim Pays-Bas 106:596
21. Paulsen H, Lebuhn R (1983) Leibigs Ann Chem:1047
22. Paulsen H, Heume M, Györgydeak Z, Lebuhn R (1985) Carbohydr Res 144:57
23. Paulsen H, Lebuhn R, Lockhoff O (1982) Carbohydr Res 103:C7
24. Paulsen H, Lebuhn R (1984) Carbohydr Res 130:85
25. Paulsen H (1990) Angew Chem Int Ed Engl 29:823
26. Ogawa T, Kitajima T, Nukada T (1983) Carbohydr Res 123:C5
27. Yamazaki F, Kitajima T, Nukada T, Ito Y, Ogawa T (1990) Carbohydr Res 201:15
28. Koike K, Mori M, Ito Y, Nakahara Y, Ogawa T (1990) Agrc Biol Chem 54:2931
29. Matsuo I, Nakahara YU, Ito Y, Nukada T, Nakahara Y, Ogawa T (1995) Bioorg Med Chem 11:1455
30. Ogawa T, Sugimoto M, Kitajima T (1986) Tetrahedron Lett 27:5739
31. Mootoo DR, Konradsson P, Udodong U, Fraser-Reid B (1988) J Am Chem Soc 110:5583
32. van Boeckel CAA, Beetz T, van Aelst SF (1984) Tetrahedron 40:4097
33. Kanie O, Takeda T, Hada N, Ogihara Y (1991) J Carbohydr Chem 10:561
34. van Boeckel CAA, Beetz T (1985) Recl Trav Chem Pays-Bas 104:171
35. van Boeckel CAA, Beetz T (1985) Recl Trav Chem Pays-Bas 104:174
36. Garegg PJ, Köpper S, Ossowski P, Thiem J (1986) J Carbohydr Chem 5:59
37. Toshima K, Kasumi K, Matsumura S (1998) Synlett 643
38. Kim W-S, Sasai H, Shibasaki M (1996) Tetrahedron Lett 37:7797
39. Tatsuta K, Yasuda S (1996) Tetrahedron Lett 37:2453
40. De Meo C, Kamat MN, Demchenko AV (2005) Eur J Org Chem 4:706
41. Kahne D, Walker S, Cheng Y, Engen DV (1989) J Am Chem Soc 111:6881
42. Crich D, Sun S (1996) J Org Chem 61:4506
43. Crich D, Sun S (1997) J Org Chem 62:1198
44. Crich D, Sun S (1998) Tetrahedron Lett 39: 1681
45. Crich D, Sun S (1998) J Am Chem Soc Chem 120:435
46. Crich D, Smith M (2001) J Am Chem Soc Chem 123:9015
47. Crich D, Sun S (1997) J Am Chem Soc 119:11217
48. Crich D, Chandrasekera NS (2004) Angew Chem Int Ed 43:5386
49. Crich D, Cai W (1999) J Org Chem 71:8473
50. Crich D, de la Mora M, Vinod AU (2003) J Org Chem 68:9532
51. Crich D, Vinogradova O (2006) J Org Chem 71:8473
52. Crich D, Jayalath P (2005) Org Lett 7:2277
53. Crich D, Jayalath P, Hutton TK (2006) J Org Chem 71:3064
54. Crich D, Wu B (2006) Org Lett 8:4879
55. Crich D, Banerjee A, Yao Q (2004) J Am Chem Soc 126:14930
56. Crich D, Li W, Li H (2004) J Am Chem Soc 126:15081
57. Crich D, Yao Q (2003) Org Lett 5:2189
58. Crich D, Yao Q (2004) J Am Chem Soc 126: 8232
59. Crich D, Bowers AA (2006) Org Lett 8:4327
60. Crich D, Banerjee A (2005) Org Lett 7:1395
61. Crich D, Banerjee A (2006) J Am Chem Soc 128:8078
62. Yamanoi T, Nakamura K, Takeyama H, Yanagihara K, Inazu T (1994) Bull Chem Soc Jpn 67:1359
63. Sato S, Sakamoto H, Nakajima M, Hashimoto S (1998) Abstracts of 20th Japan Carbohydrate Symposium, p 98
64. Tsuda T, Sato S, Nakamura S, Hashimoto S (2003) Heterocycles 59:509
65. Tsuda T, Arihara R, Sato S, Koshiba M, Nakamura S, Hashimoto S (2005) Tetrahedron 60:10719
66. Plante OJ, Palmacci ER, Seeberger PH (2000) Org Lett 2:3841
67. Nagai H, Matsumura S, Toshima K (2003) Carbohydr Res 338:1531
68. Nagai H, Sasaki K, Matsumura S, Toshima K (2005) Carbohydr Res 340:337
69. Garcia BA, Poole JL, Gin DY (1997) J Am Chem Soc 119:7597
70. Codée JDC, Hossain LH, Seeberger PH (2005) Org Lett 7:3251
71. Hashihayata T, Mandai H, Mukaiyama T (2004) Bull Chem Soc Jpn 77:169

72. Mandai H, Mukaiyama T (2005) Chem Lett 34:702
73. Mandai H, Mukaiyama T (2006) Bull Chem Soc Jpn 79:479
74. Weingart R, Schmidt RR (2000) Tetrahedron Lett 41:8753
75. Tanaka S-I, Takashima M, Tokimoto H, Fujimoto Y, Tanaka K, Fukase K (2005) Synlett 2325
76. Kim KS, Kim JH, Lee YJ, Lee YJ, Park J (2001) J Am Chem Soc 123:8477
77. Lee YJ, Lee K, Jung EH, Jeon HB, Kim KS (2005) Org Lett 7:3263
78. Lee YJ, Lee BY, Jeon HB, Kim KS (2006) Org Lett 8:3971
79. Kunz H, Wernig P, Schultz M (1990) Synlett 631
80. Lopez JC, Fraser-Reid B (1991) J Chem Soc Chem Commun 159
81. Baek JY, Choi TJ, Jeon HB, Kim KS (2006) Angew Chem Int Ed 45:7436
82. Ito Y, Ogawa T (1988) Tetrahedron Lett 29:1061
83. Ekborg G, Lindberg B, Lönngren J (1972) Acta Chim Scand 26:3287
84. Borén H, Ekborg G, Eklind K, Garegg PJ, Pilotti Å, Swahn CG (1973) Acta Chem Scand 27:2639
85. Kochetkov NK, Dmitriev BA, Malysheva NN, Chernyak AY, Klimov EM, Bayamova NE, Toggov VI (1975) Carbohydr Res 45:283
86. Warren CD, Augé C, Laver ML, Suzuki S, Power D, Jeanloz R (1980) Carbohydr Res 82:71
87. Augé C, Warren CD, Jeanloz R, Kiso M, Anderson L (1980) Carbohydr Res 82:85
88. Khan SH, Matta K (1995) Carbohydr Res 278:351
89. Danishefsky SJ, Hu S, Cirillo PF, Eckhardt M, Seeberger PH (1997) Chem Eur J 3:1617 and literature cited therein
90. Liu KK, Danishefsky S (1994) J Org Chem 59:1892
91. Mathew F, Mach M, Hazen KC, Fraser-Reid B (2003) Synlett 1319
92. Nicolaou KC, Daines RA, Chakraborty TK, Ogawa Y (1987) J Am Chem Soc 109:2821
93. Lichtenthaler FW, Schneider-Adams T, Immel S (1994) J Org Chem 59:6735
94. Lichtenthaler FW, Schneider-Adams T, Immel S (1994) J Org Chem 59:6728
95. Lichtenthaler FW, Kläres U, Szurmai Z, Werner B (1998) Carbohydr Res 305:293
96. David S, Malleron A, Dini C (1989) Carbohydr Res 188:193
97. Alais J, David S (1990) Carbohydr Res 201:69
98. Sato K-I, Yoshitomo A, Takai Y (1997) Bull Chem Soc Jpn 70:885
99. Matsuo I, Isomura M, Walton R, Ajisaka K (1996) Tetrahedron Lett 37:8795
100. Weiler S, Schmidt RR (1998) Tetrahedron Lett 39:2299
101. Kunz H, Günther W (1988) Angew Chem Int Ed Engl 27:1086
102. Günther W, Kunz H (1992) Carbohydr Res 228:217
103. Unverzagt C (1994) Angew Chem Int Ed Engl 33:1102
104. Unverzagt C (1996) Angew Chem Int Ed Engl 35:2350
105. Seifert J, Unverzagt C (1996) Tetrahedron Lett 37:6527
106. Unverzagt C (2003) Chem Eur J 9:1369
107. Seifert J, Unverzagt C (1997) Tetrahedron Lett 38:7857
108. Prahl I, Unverzagt C (2000) Tetrahedron Lett 41:10189
109. Weiss H, Unverzagt C (2003) Angew Chem Int Ed 42:4261
110. Schmidt RR, Michel J, Roos M (1984) Liebigs Ann Chem 1343
111. Tamura J-I, Schmidt RR (1995) J Carbohydr Chem 14:895
112. Srivastava VK, Schuerch C (1979) Tetrahedro Lett 35:3269
113. Dessinges A, Olesker A, Lukacs G (1984) Carbohydr Res 126:C6
114. Hodosi G, Kovác (1997) J Am Chem Soc 119:2335
115. Hodosi G, Kovác (1998) Carbohydr Res 308:63
116. Kahne D, Yang D, Jin J, Miller R, Paguaga E (1988) J Am Chem Soc 110:8716
117. Crich D, Hermann (1993) Tetrahedron Lett 34:3385
118. Ohtake H, Iimori T, Ikegami S (1997) Tetrahedron Lett 38:3413
119. Ohtake H, Ichiba N, Shiro M, Ikegami S (2000) J Org Chem 65:8164
120. Iimori T, Ohtake H, Ikegami S (1997) Tetrahedron Lett 38:3415
121. Ohtake H, Ichiba N, Ikegami S (2000) J Org Chem 65:8171
122. Huchel U, Schmidt RR (1998) Tetrahedron Lett 39:7693 and literature cited therein
123. Barresi F, Hindsgaul O (1991) J Am Chem Soc 113:9376
124. Barresi F, Hindsgaul O (1992) Synlett 759
125. Barresi F, Hindsgaul O (1994) Can J Chem 72:1447
126. Stork G, Kim G (1992) J Am Chem Soc 114:1087

127. Stork G, La Clair JL (1996) J Am Chem Soc 118:247
128. Ennis SC, Fairbanks AJ, Tennant-Eyles RJ, Yeates HS (1999) Synlett 1387
129. Ennis SC, Fairbanks AJ, Slinn CA, Tennant-Eyles RJ, Yeates HS (2001) Tetrahedron 57:4221
130. Fairbanks AJ (1999) Synlett 1945
131. Chayajarus K, Chambers DJ, Chughtai MJ, Fairbanks AJ (2004) Org Lett 6:3797
132. Aloui M, Chambers DJ, Cumpstey I, Fairbanks AJ, Redgrave AJ, Seward MP (2002) Chem Eur J 8:2608
133. Ito Y, Ogawa T (1994) Angew Chem Int Ed Engl 33:1765
134. Lergenmüller M, Nukada T, Kuramochi K, Dan A, Ogawa T, Ito Y (1999) Eur J Org Chem 1367
135. Ito Y, Ando H, Wada M, Kawai T, Ohnishi Y, Nakahara Y (2001) Tetrahedron 57:4123
136. Dan A, Ito Y, Ogawa T (1995) J Org Chem 60:4680
137. Dan A, Ito Y, Ogawa T (1995) Tetrahedron Lett 36:7487
138. Dan A, Lergenmüller M, Amano M, Nakahara Y, Ogawa T, Ito Y (1998) Chem Eur J 4:2181
139. Ito Y, Ohnishi Y, Ogawa T, Nakahara Y (1998) Synlett 1102
140. Seifert J, Lergenmüller M, Ito Y (2000) Angew Chem Int Ed 39:531
141. Ohnishi Y, Ando H, Kawai T, Nakahara Y, Ito Y (2000) Carbohydr Res 328:263
142. Nakano J, Ohta H, Ito Y (2006) Bioorg Med Chem Lett 16:928
143. Nakano J, Ishiwata A, Ohta H, Ito Y (2007) Carbohydr Res: 342:675
144. Matsuo I, Wada M, Manabe S, Yamaguchi Y, Otake K, Kato K, Ito Y (2003) J Am Chem Soc 125:3402
145. Matsuo I, Ito Y (2003) Carbohydr Res 338:2163
146. Matsuo I, Kashiwagi T, Totani K, Ito Y (2005) Tetrahedron Lett 46:4197
147. Matsuo I, Totani K, Tatami A, Ito Y (2006) Tetrahedron 62:8262
148. Ito Y, Ogawa T (1997) J Am Chem Soc 119:5562
149. Ziegler T, Lemanski G (1998) Angew Chem Int Ed Engl 37:3129
150. Lemanski G, Ziegler T (2000) Eur J Org Chem 181
151. Abdel-Rahman AA, Ashry ESHE, Schmidt RR (2002) Carbohydr Res 337:195
152. Gijsen HJM, Qiao L, Fitz W, Wong C-H (1996) Chem Rev 96:443
153. Ichikawa Y (1997) Enzymatic synthesis of oligosaccharides and glycopeptides. In: Large DG, Warren CD (eds) Glycopeptides and related compounds. Marcel Dekker, New York, p 79
154. Watt G, Lowden PAS, Flitsch SL (1997) Curr Opin Struct Biol (1997) 7:652
155. Singh S, Scigelova M, Crout DHG (1996) Chem Commun 993
156. Watt GM, Revers L, Webberley MC, Wilson IBH, Flitsch SL (1997) Angew Chem Int Ed Engl 36:2354
157. Flisch SL, Pinches HL, Taylor JP, Turner NJ (1992) J Chem Soc Perkin Trans 1:2087
158. Ichikawa Y, Wang R, Wong C-H (1994) Methods Enzymol 247:107
159. Mashiru O, Zechel DL, Stoll D, Mohammadzadeh T, Warren AJ, Withers SG (2001) Angew Chem Int Ed 40:417

5.7 Selective α-Sialylation

Hiromune Ando[1], *Makoto Kiso*[2]
[1] Division of Instrumental Analysis, Life Science Research Center, Gifu University, Gifu-shi, Gifu 501–1193, Japan
[2] Department of Applied Bioorganic Chemistry, Faculty of Applied Biological Sciences, Gifu University, Gifu-shi, Gifu 501–1193, Japan
hando@gifu-u.ac.jp, kiso@cc.gifu-u.ac.jp

In: *Glycoscience*. Fraser-Reid B, Tatsuta K, Thiem J (eds)
Chapter-DOI 10-1007/978-3-540-30429-6_31:

Abstract

Chemical construction of a α-glycoside of sialic acid is a demanding subject due to the structural disadvantages of sialic acid: the carboxylate group at C1, the deoxy moiety adjacent to the anomeric center, and the glycerol branch from C6. This chapter will provide an overview of sialic acid chemistry and show how the disadvantages can be surmounted and it will also cover the cutting-edge methodology used for α-selective sialylation. In this chapter, the forty-year history of sialylation is compartmentalized into three periods. The first period represents the use of 2-halogeno sialy donors. The second period covers the use of C3-appended sialyl donors and a set of thioglycoside donors, including also the acetonitrile effect, and third is the period of C5-modified sialy donors. While following this division, the chemistry of α-selective sialylation is discussed, and some examples of the stereoselective syntheses of sialyl oligosaccharides are described. Furthermore, the synthesis of the homopolymer of sialic acid and artificial sialyl *S*- and *C*-glycosides is discussed.

Keywords

Sialylation; Glycosylation; Stereoselectivity; Glycosyl donor; Glycosyl acceptor; Sialic acid; Neuraminic acid; Solvent effect; Auxiliary

Abbreviations

AIBN 2,2-azobisisobutylonitrile
BOM benzyloxymethyl
CAc chloroacetyl
DMTST dimethyl(methylthio)sulfonium triflate
DTBC 2,6-di-*tert*-butyl-4-cresol
DTBP 2,6-di-*tert*-butylpyridine
EDCI *N*-ethyl-*N*′-(3-dimethylaminopropyl)carbodiimide
HOBt 1-hydroxy-1*H*-benzotriazole
LDA lithium diisopropylamide
Lev 4-oxopentanoyl
Oct octyl
SE 2-(trimethylsilyl)ethyl
SEM 2-(trimethylsilyl)ethoxymethyl
TDS dimethylthexylsilyl
THF tetrahydrofuran
TTBP 2,4,6-tri-*tert*-butylpyrimidine

1 Introduction

The sialic acids are a large family of mono-saccharides derived from 3-deoxy-non-2-ulosonic acid. The three major sialic acids found in nature are *N*-acetyl (Neu5Ac) **1** and *N*-glycolyl (Neu5Gc) **2** derivatives of neuraminic acid (5-amino-3,5-dideoxy-D-*glycero*-D-*galacto*-non-2-ulosonic acid) and KDN (3-deoxy-D-*glycero*-D-*galacto*-non-2-ulosonic acid) **3** (● *Fig. 1*) [1,2,3].

The original discovery of sialic acids by Klenk and Blix dates back to the 1930s. Blix isolated sialic acid from submaxillary mucin, named after the Greek word sialos (*saliva*), and Klenk isolated the neuraminic acid derivative from brain glycolipids (neuro + amine + acid). Until the discovery of KDN by Inoue from fish eggs, sialic acid had been defined as neuraminic acid and its derivatives. To date the remarkable structural diversity of sialic acids in nature is well known with the family comprising over 50 structurally distinct compounds and it is also known that the occurrence of sialic acid is limited to the deuterostome lineage of animals (vertebrates and a few higher invertebrates such as starfish) and certain types of bacteria [2,3]. Neu5Ac is biosynthesized via condensation of ManNAc-6-P with pyruvate, while KDN is biosynthesized via condensation of Man-6-P with pyruvate. Furthermore, the α-carbon of the *N*-acetyl group within CMP-Neu5Ac is oxidized to produce CMP-NeuGc catalyzed by CMP-Neu5Ac hydrolase. The hydroxyls of these backbone compounds of "sialic acids" can be further modified by acetylation, lactylation, phospholylation, sulfonation, or methylation [2,3].

Within glycan chains conjugated with lipid or protein scaffolds, sialic acids typically occupy the distal end of glycan chains through α-(2–6) and/or α-(2–3)-linkage with galactose, *N*-acetyl-galactosamine, glucose, or *N*-acetyl-glucosamine, and through α-(2–8), α-(2–9), or α-(2–4)-linkages with another neuraminic acid residue, forming disialic residues. Moreover, homopolymers of sialic acid through α-(2–8) or α-(2–9)-linkage are also formed in nature. CMP-sialic acid **4** is only an exception of the molecule where anomeric configuration of sialic acid is β (● *Fig. 2*) [3]. Unusually, the hydroxyl of the glycolyl moiety of Neu5Gc is glycosylated with other Neu5Gc in sequence, thereby forming a linear polymer of Neu5Gc [4].

Because of the outmost positioning, sialic acid can participate in carbohydrate–protein interaction that mediates a variety of important biological processes such as cell growth, cell differentiation, cell adhesion, cellular immune response, fertilization, oncogenesis, viral infection, and so on. Therefore one can imagine that the elucidation of key biological roles of sialic acid will provide clear insight into dynamism of carbohydrate-mediated biological processes and diseases.

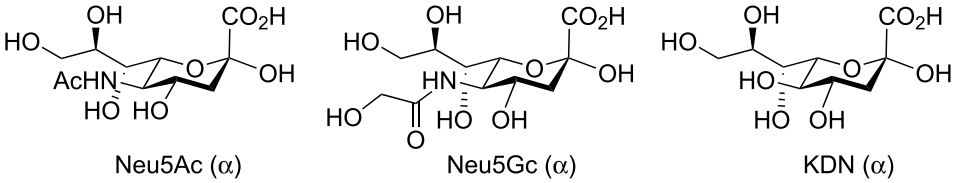
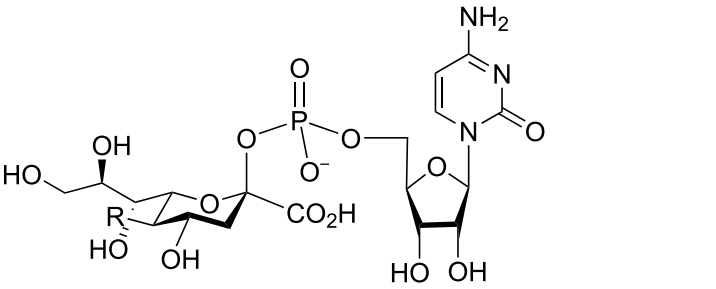


Figure 1
Typical sialic acids



CMP-sialic acid (R = NHAc, NHGc, OH)

4

Figure 2
β-Glycoside of sialic acid in nature

In this context, the production of a large quantity of fine sialo-glycans and its conjugates with lipid or protein (peptide) has been a critical subject.

α-Glycoside of sialic acid can be synthesized by the reaction of oxocarbenium ion generated from a sialic acid donor with the hydroxyl of the glycosyl partner. However, the issue of coupling yield and stereoselectivity is much more complicated by the special structural features of sialic acid. First, the electron-withdrawing carboxyl group on the anomeric center makes the tertiary oxocarbenium ion intermediate unstable and susceptible to 2,3-elimination in collaboration with its 3-deoxy structure. Second, because of the deoxy structure, no neighboring functionality at the C3 position is available, assisting with the formation of the alternative thermodynamically more stable β-glycoside. Also the glycerol moiety branching from the C6 position is probably steric hindrance to the anomeric carbon. Therefore, the difficulty of α-sialylation has been centered on the carbohydrate chemistry as well as that of β-mannosylation [5,6].

Back in 1965, Meindl and Tuppy published the paper "Über synthetische Ketoside der *N*-acetyl-D-neuraminsäure"; to our knowledge, it is the first report on sialoside synthesis [7]. In 1971, Khorlin was the first to synthesize the interlinkage between Neu5Ac and glucose or galactose [8]. They prepared a 2-chloro derivative of Neu5Ac as a glycosyl donor, which was reacted with glycosyl acceptors in the presence of Ag_2CO_3 in $CHCl_3$.

Since the outset of the history of sialylation reaction, a huge number of literatures on the synthesis of α-sialoside have been published [9,10,11,12]. In view of type of sialyl donors used in the glycosylation, the 40-years history can be roughly divided into three periods as shown in *Fig. 3*.

The 2-halogeno donor is placed at the origin of the first period from the 1960s to the early 1980s. At the beginning of the second period (the middle of the 1980s), the stream was divided into two branches: one is the lineage of the thioglycoside donor and the other is the C3-appended donor. A thioglycoside derivative of Neu5Ac turned out to be effective for α-selective glycosidation with the assistance of the nitrile solvent effect, and the leaving group was replaced with other varieties. On the other hand, the C3-appended donor incorporates a neighboring functionality at the C3 position in order to direct the α-selective coupling, being independent

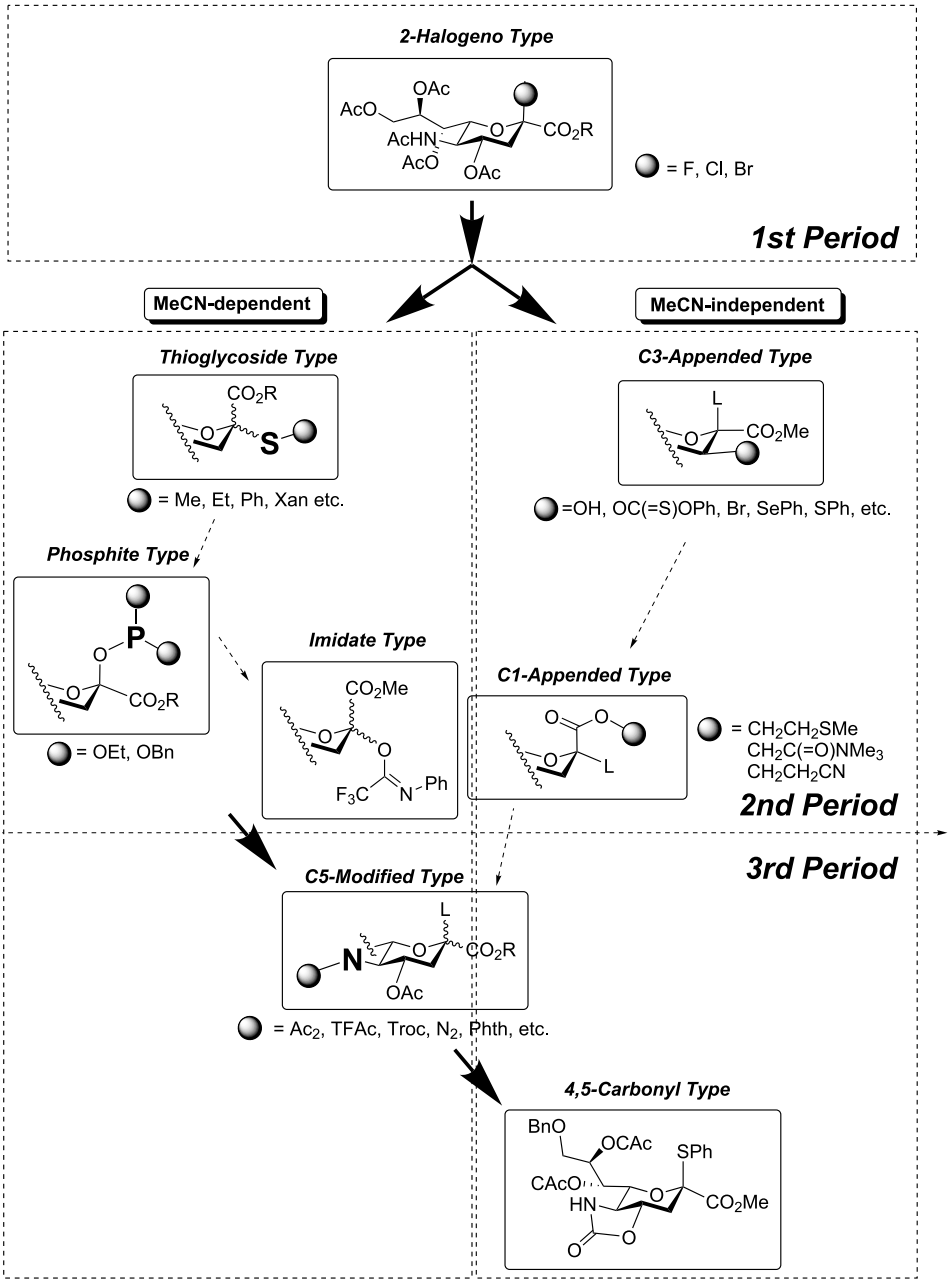


Figure 3
Division of the history

of the solvent effect. The principle was also applied to the design of the C1-appended type of donors.

In 1996, a new type of thioglycoside donor, the *N*,*N*-diacetyl sialyl donor, was devised by Boons and co-workers. This is the beginning of the third period. The *N*,*N*-diacetyl sialyl donor was followed by other analogous C5-modified sialyl donors. Very recently, Tanaka and his co-workers have developed the 5-*N*-4-*O*-oxazolidinone sialyl donor, which can provide α-sialoside without assistance of the nitrile solvent effect despite the absence of an auxiliary at the C3 or C1 positions.

This chapter will outline the chemistry of sialylation and give examples of effective sialylation and the synthesis of complicated sialyl oligosaccharides.

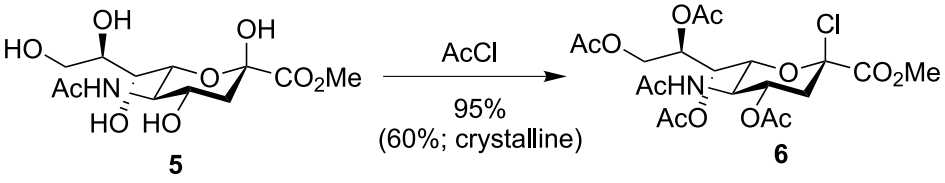
2 The First Period

2.1 2-Chloro Sialyl Donor

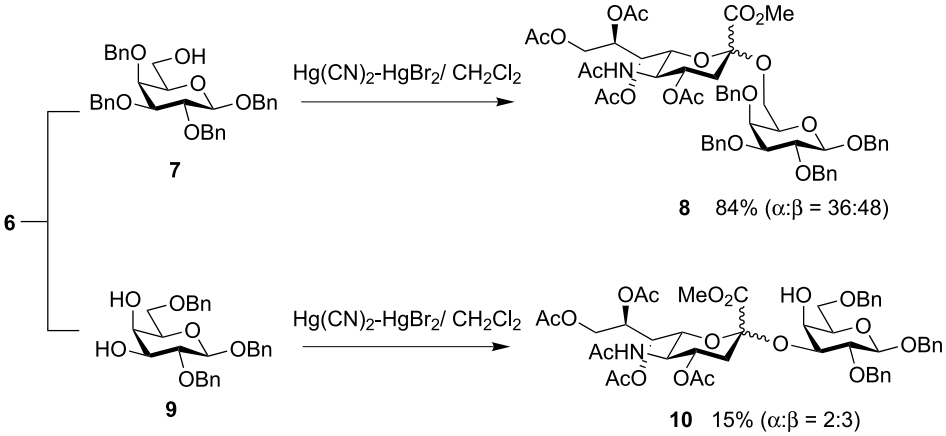
The 2-halogeno derivative of Neu5Ac was first used for the preparation of α-sialosides with the combination of a Koenigs–Knorr type activation system. The 2-chloro derivative has been more widely utilized until the emergence of a second generation of sialyl donor in the 1980s. The 2-chloro derivative is readily available from the methyl ester of Neu5Ac by the action of $TiCl_4$ [13], HCl/AcCl [14,15], AcCl/MeOH [16], or AcCl [17]. For example, the reaction of methyl ester **5** with AcCl produces 2-chloride **6** in 95% yield, the crystalline form of which is stable for several weeks when stored in a freezer (*Scheme 1*) [17]. There are a lot of methods available for the activation of the 2-chloride donor; Ag_2CO_3 (Koenigs–Knorr method) [7,8], $Hg(CN)_2$-$HgBr_2$ (Helferich modification) [7,13], insoluble silver salts (silver polymaleate [18], silver salicylate [19], silver zeolite [20], or silver mercaptoethanesulfonate [21]), soft Lewis acids (AgOTf [22,23], $Sn(OTf)_2$ [23] etc) or I_2 [24].

The stereoselectivity and yield of glycosylation varies to a wide extent, depending on the structure of the glycosylation partner and solvent. As shown in *Scheme 2*, the 2-chloride donor **6** was coupled with the primary hydroxyl of galactoside **7** in 84% yield ($\alpha/\beta = 36/48$), while with the secondary C3 hydroxyl of **9** the yield of α-glycoside was extremely diminished (15%, $\alpha/\beta = 2/3$) [13,25].

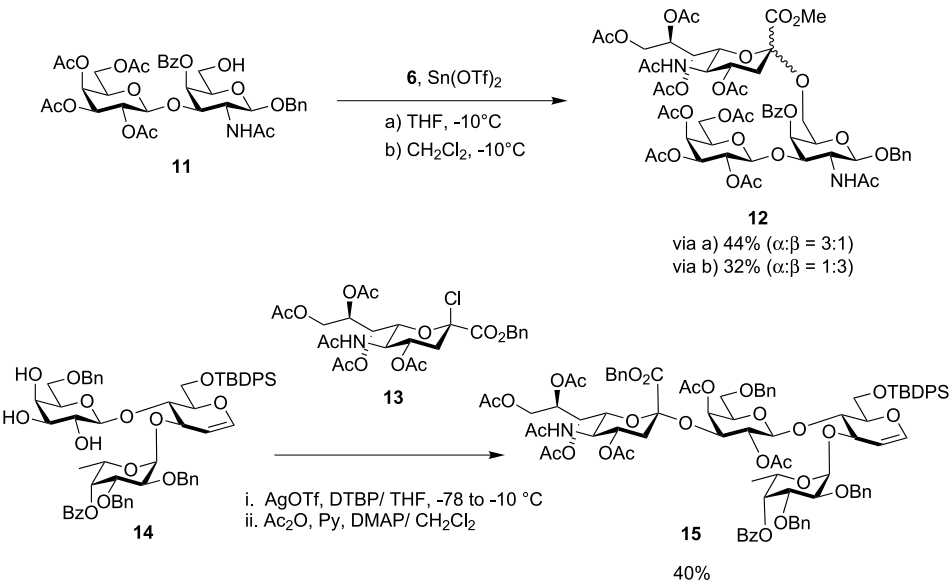
Auge and co-workers have found that the use of THF as a reaction solvent enhances the α-selectivity during the coupling reaction of **6** and disaccharide acceptor **11** effected by $Sn(OTf)_2$ (*Scheme 3*) [23]. This "THF solvent effect" was also successfully applied to the delivery of sialyl Lewis X tetrasaccharide by Danishefski's group; donor **13** was regioselec-



Scheme 1



Scheme 2

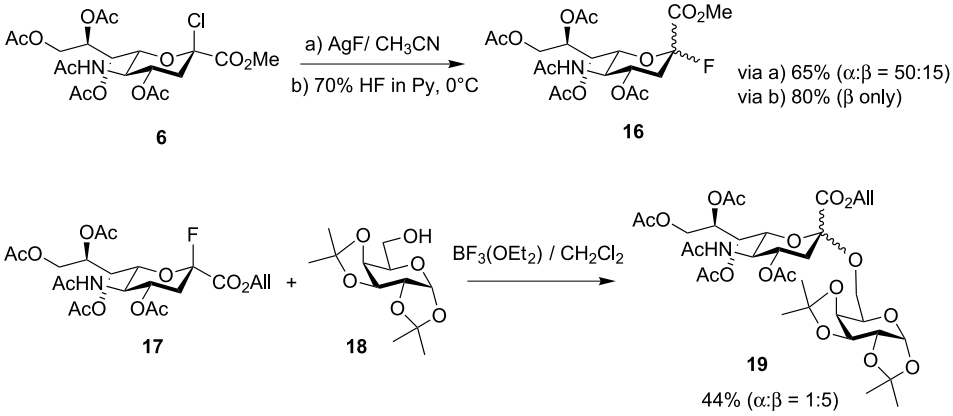


Scheme 3

tively glycosidated with the C3 hydroxyl group of trisaccharide acceptor **14** in the presence of AgOTf and DTBP at −78 to −10 °C, and the subsequent acetylation gave **15** in 40% yield [26].

2.2 2-Bromo and Fluoro Sialyl Donors

The 2-bromo derivative of Neu5Ac is also accessible from the corresponding 2-*O*-acetyl derivative by treatment with $TiBr_4$ or other procedures [13,16]. However, the high reactivity of the 2-bromo derivative causes preferential formation of the 2,3-en byproduct during condensation, and is also prone to decomposition during storage in a freezer. On the other hand, 2-fluoro derivative **16**, which can be prepared from the 2-chloro derivative **6** by the reaction with AgF in MeCN or 70% HF in pyridine, is more stable than 2-chloride and 2-bromide [27]. However, 2-fluoro donor **17**, upon activation by $BF_3(OEt_2)$, was glycosidated with galactosyl acceptor **18** to produce the β-sialyl outcome preferentially (● *Scheme 4*) [28].



■ **Scheme 4**

3 The Second Period

As mentioned above, α-anomeric selectivity during the glycosidation of 2-halogeno type donors lies in the medium range. In the late 1980s, the respective groups of Goto and Hasegawa both provided breakthroughs in relation to α-selective sialylation. Goto's group incorporated an equatorial hydroxyl at the C3 position (*trans* to α-glycoside) of Neu5Ac, which was anticipated to show nucleophilic participation with the contiguous anomeric carbon during condensation, directing α-sialoside formation. In the laboratory of Hasegawa's group, the methylthioglycoside of Neu5Ac was synthesized for the first time, and they have attempted the glycosidation of the 2-methylsulfenyl donor in MeCN. Fortunately, their attempts have resulted in novel effective α-sialylation methods.

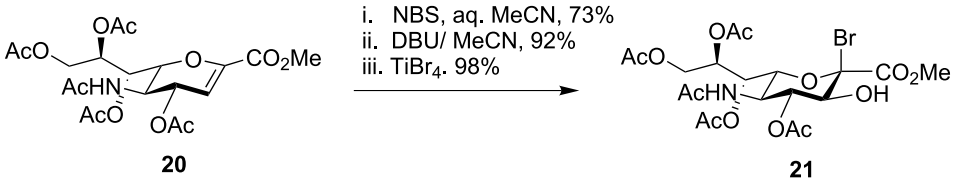
3.1 C3-Appended Sialyl Donors

Goto's group first synthesized the 2-bromo-3-hydroxy derivative of Neu5Ac **21** from the corresponding 2,3-en derivative **20** via bromohydrination, epoxidation, and 2-bromination (*Scheme 5*) [29].

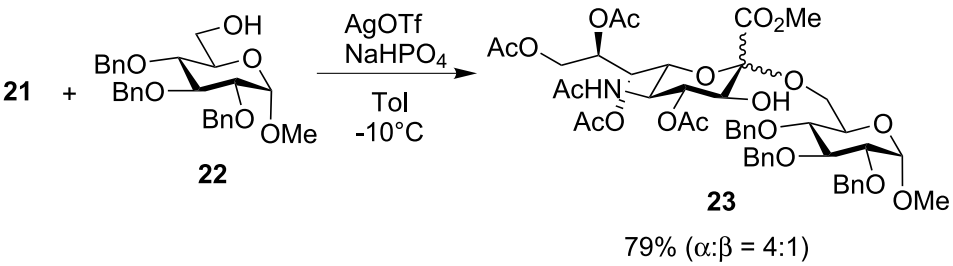
The initial glycosidation of the 3-hydroxy derivative **21** with glucosyl acceptor **22** in the presence of AgOTf and $NaHPO_4$ in toluene afforded Neu5Ac-α-(2–6)-Glc disaccharide **23** in 63% yield accompanied with a β-isomer (16%) (*Scheme 6*) [29,30].

This result suggests that the higher stereoselectivity is endowed with nucleophilic participation of the C3 auxiliary, and the high yield of the coupling is due to the prevention of 2,3-elimination also by the presence of the C3 hydroxyl (*Fig. 4*).

On the basis of this result, the C3 hydroxyl was replaced with phenylselenyl, phenysulfenyl, phenoxythiocarbonyl groups or bromine, and the C2 bromine was also replaced with fluorine, chlorine, or the diethylphosphite group. Schmidt's group demonstrated that, when the C3



Scheme 5



Scheme 6

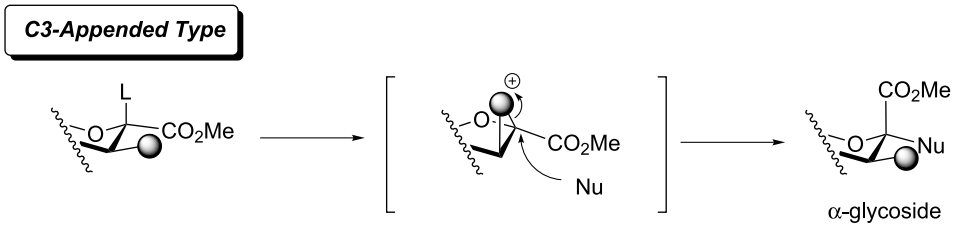
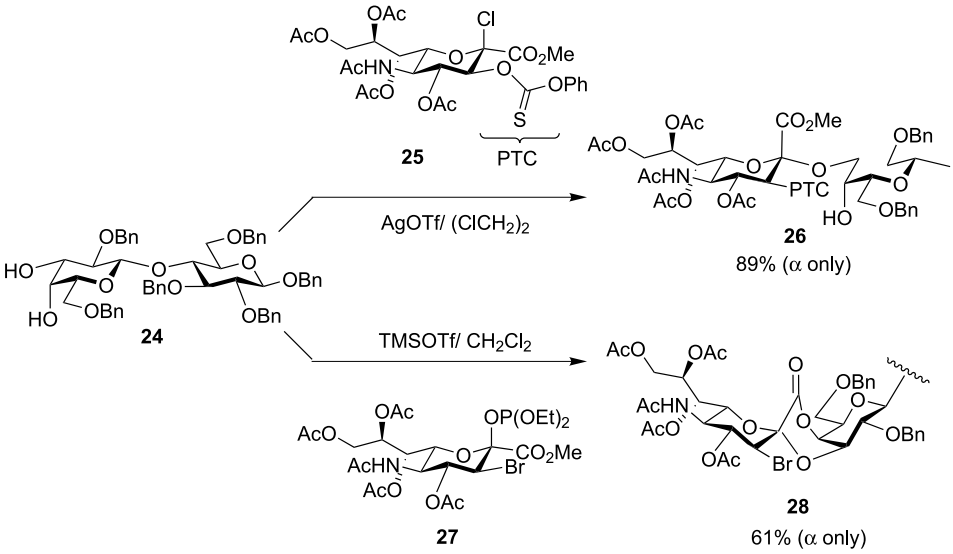


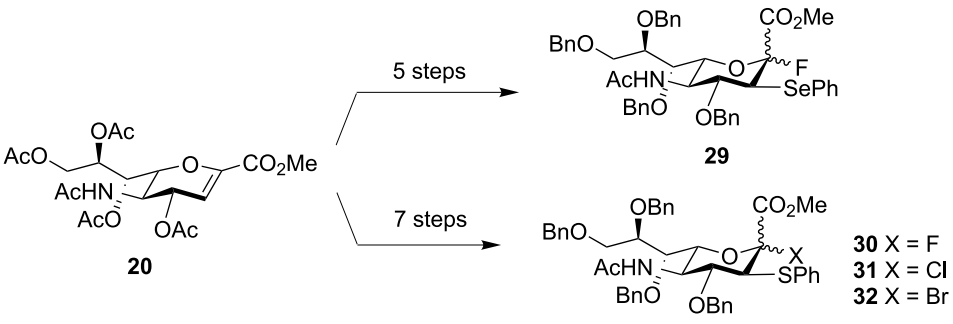
Figure 4
Neighboring participation of C3 auxiliary to direct α-sialylation

hydroxyl was masked with a phenoxythiocarbonyl group in the 2-chloro analog of **21**, the corresponding donor **25** was incorporated at the secondary C3′ hydroxyl of lactosyl acceptor **24** to produce **26** in high yield as a single stereoisomer (*Scheme 7*) [31]. Schmidt's group also reported that 2-phosphite-3-bromide derivative **27** served as an effective glycosyl donor for α-sialylation in TMSOTf-catalyzed glycosidation with **24**, producing 2′,4- lactonated trisaccharide **28** in 61% yield [31].

Phenylselenyl and phenylsulfenyl groups are expected to direct α-sialylation more effectively via predictable episelenium or episulfenium ion intermediates, due to their high nucleophilicity. Ogawa and Ito first developed the 3-phenylselenyl derivative of Neu5Ac **29** having fluorine as a leaving group, which was prepared from 2,3-en compound **20** through the reaction sequence in five steps (*Scheme 8*) [32].



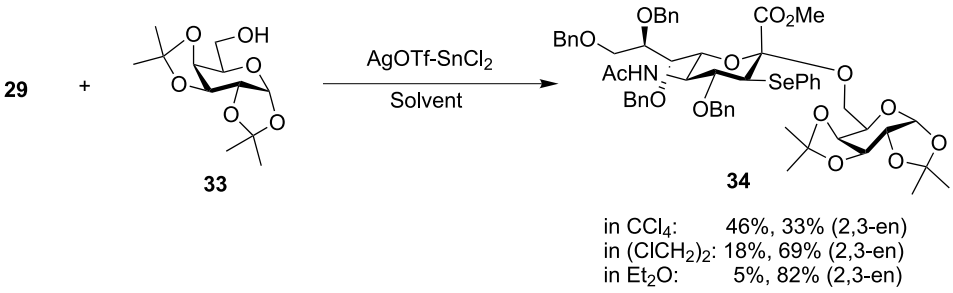
Scheme 7



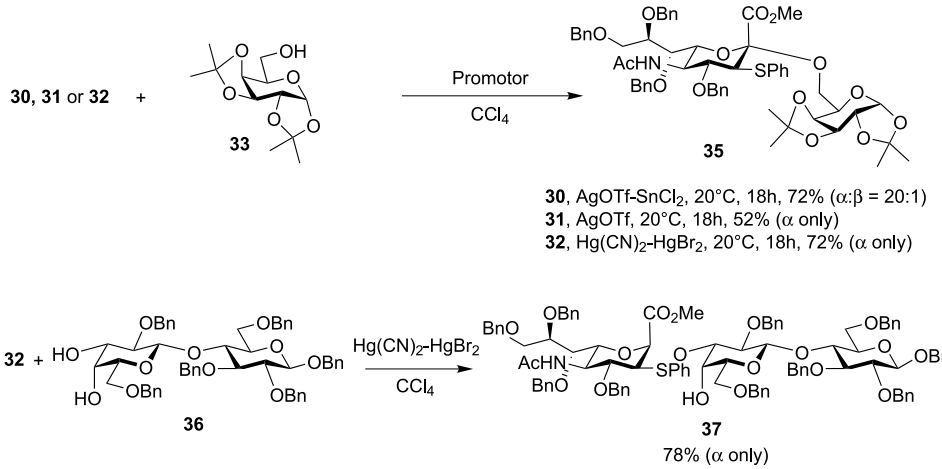
Scheme 8

The coupling reaction with diacetone galactosyl acceptor **33** and the donor **29** was examined in various solvents (Scheme 9) [32]. In all solvents the reaction constantly provided perfect α-selectivity, whereas the coupling yield differed greatly. The best result was obtained when the reaction was carried out in CCl_4 to produce **34** in 46% yield, which was accompanied with the competitive formation of 2,3-en derivative (33%).

This issue was resolved by the replacement of the phenylselenyl group with the phenylsulfenyl group. Among the 3-phenylsulfenyl sialyl donors (**30**, **31**, and **32**) having fluorine, chlorine, and bromine as leaving groups, the 2-bromo derivative **32** turned out to be superior to others (Scheme 10) [33]. The glycan part of ganglioside GM3 **37** was produced by the reaction of **32** with the 3′,4′-dihydroxy lactosyl acceptor **36** in 78% yield, α-exclusively.



Scheme 9



Scheme 10

3.2 C1-Appended Sialyl Donors

Incorporation of an auxiliary at the C1 position, which is anticipated to show long-range participation with the β-face of the anomeric center as depicted in ❯ *Fig. 5*, was originally examined by Takahashi and co-workers. They incorporated the 2-methylthioethyl group through an ester linkage at the C1 position of 2,4,7,8,9-penta-*O*-acetyl derivative **38**, and the 2-acetoxy group was substituted with the methylsulfenyl group (❯ *Scheme 11*) [34]. Similarly, Gin's group synthesized *N*,*N*-dimethylglycolylamide ester derivatives **40** and **41** from **38** [35,36]. Ito's group showed that the cyanoethyl ester derivative **43** was also readily accessible from 2-methylester derivatives **42** via demethylation with high nucleophilic Ph_3SiSH in the presence of Cs_2CO_3 and DTBC, and DCC-mediated dehydrative coupling with 1-cyanoethanol in one pot [37].

In the case of the use of 1-methylthioethyl derivative **39**, 1,2-dimethoxyethane as the solvent was critical for α-selectivity. The optimized coupling with 3,4-diol galactosyl acceptor **44** by iodonium promotion produced disaccharide **45** in the yield of 21% with high anomeric selectivity (❯ *Scheme 12*) [34]. The 2-dimethyl glycolylamide functionality also directed α-selectivity during coupling reaction in CH_2Cl_2. The best result was obtained when 2-hydroxy derivative **41** was activated by the action of (*p*-nitrophenyl)phenyl sulfoxide and Tf_2O in CH_2Cl_2 at −78 °C in the presence of galactose derivative **46**, producing disaccharide **47** in 57% yield [36].

On the other hand, the use of 1-cyanoethyl ester derivative **43** for the glycosylation of **48** resulted in β-exclusive formation of Neu5Ac-(2–6)-Glc sequence **49** in high yield, while in MeCN the production of α-glycoside was preferred [37].

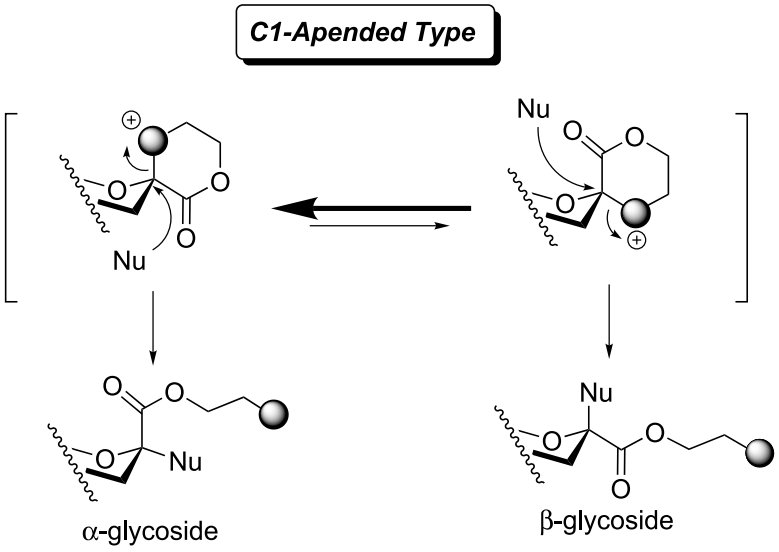
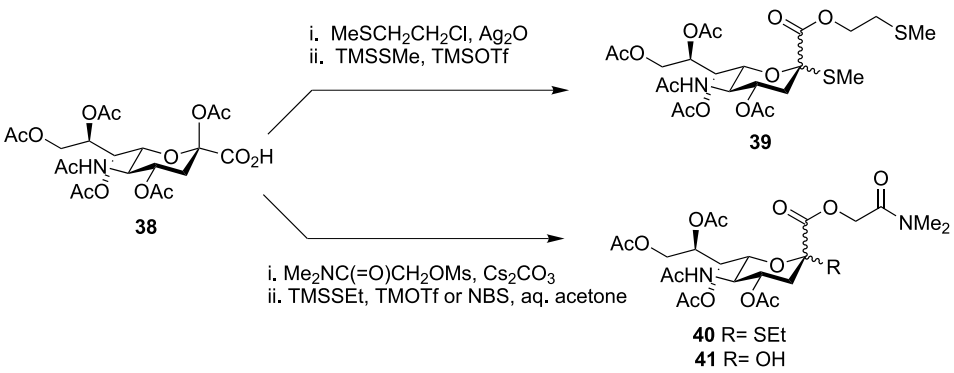
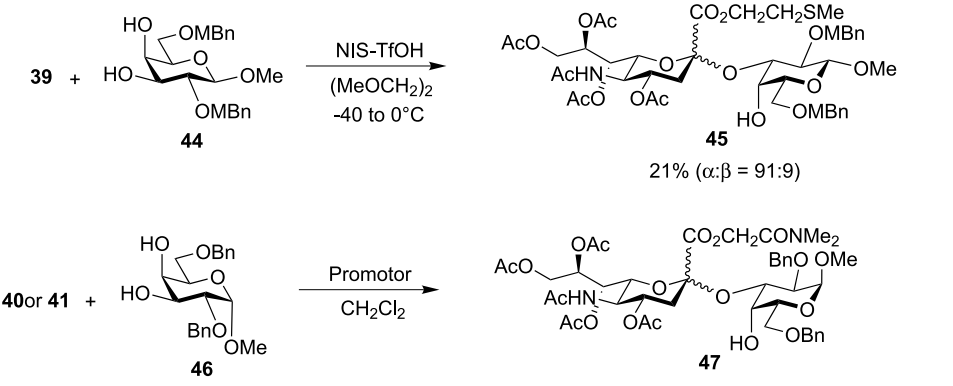
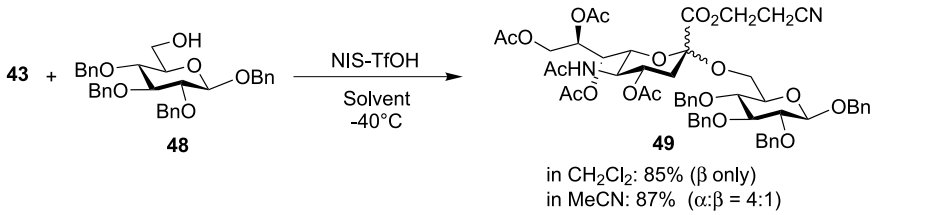


Figure 5
Long-range participation of C1 auxiliary to direct α-sialylation



Scheme 11



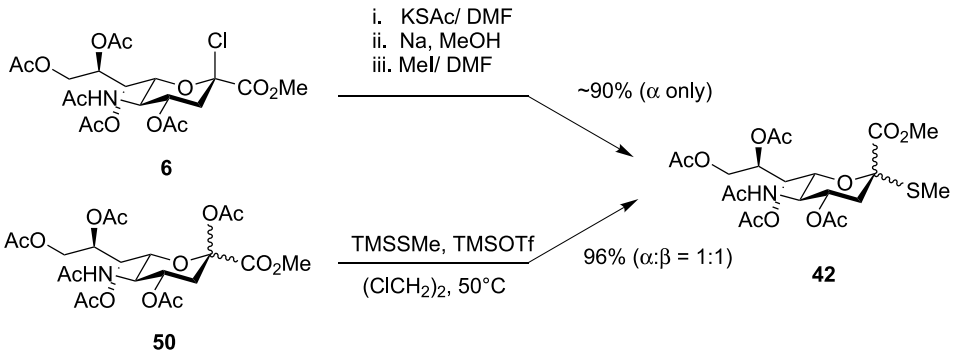


Scheme 12

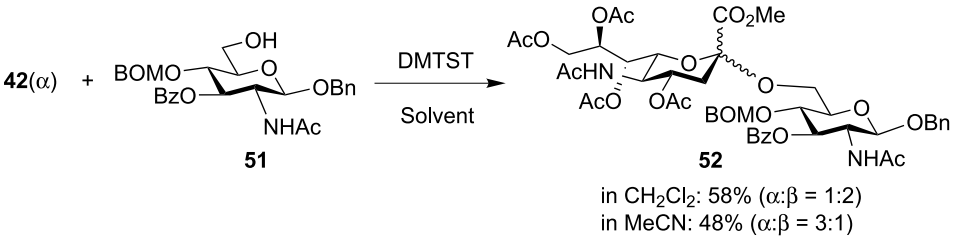
3.3 2-Thioglycoside Sialyl Donors

The 2-α-methylsulfenyl derivative of Neu5Ac **42** was originally synthesized from the corresponding 2-β-chloro derivative **6** through the reaction sequence in three steps: nucleophilic substitution by KSAc, de-*S*-acetylation by NaOMe, and methylation of the resulting sodium salt of glycosyl thiolate with MeI (*Scheme 13*) [38]. The treatment of 2-acetoxy derivative **50** with TMSSMe and TMSOTf also produced **42** in high yield as a mixture of stereoisomers [39]. The methyl group was further replaced with many other groups such as ethyl [40], phenyl [40], substituted phenyl [41,42], (ethoxy)dithiocarbonyl [40], phenyltetrazoyl [43], benzooxazoyl [44], benzothiazoyl [45], benzimidazoyl [45], thiazoyl [44], and lauryl [46] groups. These relatives can also be prepared by the S_N2-like replacement of the 2-chloride with the corresponding thiolate anion, or by the Lewis acid-catalyzed replacement of the 2-acetoxy group with the corresponding thiol reagent.

In the first paper on the glycosidation of the methylthio donor written by Hasegawa's group, **42** was reacted with the C6 hydroxyl of glucosamine derivative **23** upon the activation by DMTST in CH_2Cl_2 or MeCN (*Scheme 14*) [38]. As a result, glycosylation in MeCN proved α-selective to yield Neu5Ac-(2–6)-GlcNAc disaccharide **52** in 48% yield as a mixture of α- and β-isomers ($\alpha/\beta = 3/1$), whereas in CH_2Cl_2 disaccharide was β-preferentially produced in 58% yield ($\alpha/\beta = 1/2$). This result indicated that nitrile solvent assistance can effectively direct the formation of α-glycoside of sialic acid.



Scheme 13



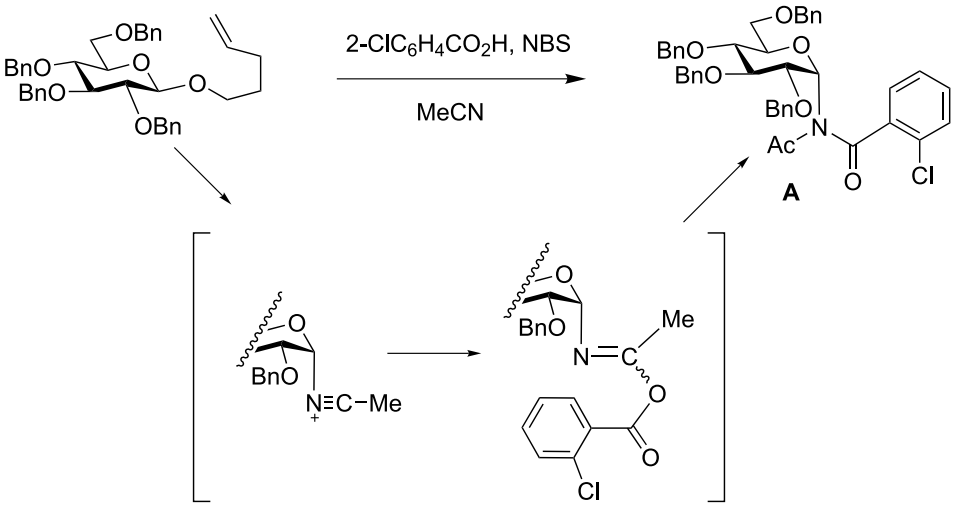
Scheme 14

3.3.1 Acetonitrile Participation in Sialylation

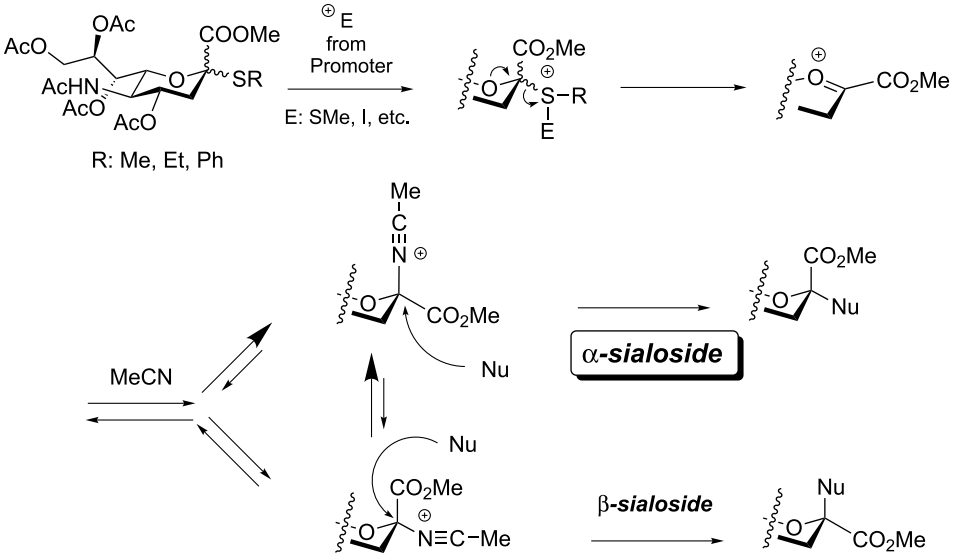
It has been reported that acetonitrile participation directs equatorial glycoside formation during the condensation of a glycosyl donor having non-participating functionality adjacent to the anomeric carbon [47,48,49, 50]. To account for this phenomenon, the intermediacy of the axial-oriented nitrilium ion was advocated by Pavia et al. [51] and Lemieux and Ratcliffe [52]. In contrast, Sinay and Pougny [53] and Schmidt and Michel [54] proposed the β-counterpart because of the so-called reverse anomeric effect [55]. In 1990, Ratcliffe and Fraiser-Reid successfully quenched the nitrilium intermediacy with 2-chlorobenzoic acid to convert into imide (**A**), and determined its α-configuration by NMR spectroscopy (● *Scheme 15*) [56], and later the absence of the reverse anomeric effect was concluded by Perrin's group [57,58]. On the basis of this evidence, the α-selective sialylation in MeCN can be rationalized by the following plausible mechanism (● *Scheme 16*) [59].

A thiophilic electrophile (+SMe) from DMTST reacts with a lone pair on sulfur to afford a cationic sulfonium species, which is an excellent leaving group. Because of the participation of a lone pair on the ring oxygen, the C–S bond is cleaved, forming an oxocarbenium ion intermediate, with which MeCN reacts to form an axial-oriented nitrilium intermediate under kinetic control. This then undergoes S_N2 nucleophilic substitution with a hydroxyl of the glycosyl accepter to yield α-sialoside.

Since other thiophilic activation systems (DMTST [48], NIS-TfOH [60,61], PhSeOTf [62], NBS-Bu_4NOTf or Ph_2IOTf [63], MeSBr-AgOTf [64]) are also compatible with the acetonitrile assistance, the combined use of a thioglycoside donor and acetonitrile (or other nitrile solvents) became a powerful option for α-sialoside synthesis, being parallel with C3-appended donors.



● **Scheme 15**



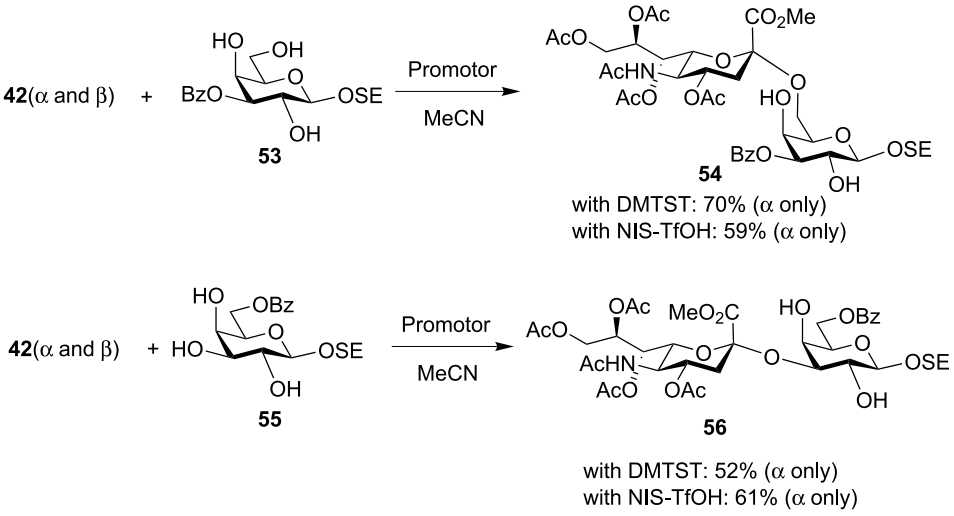
Scheme 16

3.3.2 Thioglycosides and their Relatives

As mentioned above, 2-methylthioglycoside derivative **42** can be activated by DMTST, NIS-TfOH, PhSeOTf, NBS-Bu_4NOTf, or MeSBr-AgOTf. Among the promoter systems, the excellent affinity of DMTST and NIS-TfOH was exemplified by a large number of successful α-sialylations with the assistance of MeCN [65,66,67,68,69,70]. It is of note that DMTST was preferred for the sialylation with the primary hydroxyl group of a sugar acceptor while the NIS-TfOH system was preferred for the secondary hydroxyl (*Scheme 17*) [59]. As shown in *Fig. 6*, the 2-methylthio derivative was successfully used for the glycosylation of hydroxyl within large oligosaccharide acceptor **57** [71].

2-Phenylsulfenyl derivative **58** is easily accessible from the corresponding 2-chloro derivative **6** or 2-acetoxy derivative **50** (*Scheme 18*) [40,42].

Because the glycosyl donor property of **58** is comparable to that of 2-methylsulfenyl derivative **42**, the use of costly **42** was replaced with the inexpensive **58** in recent examples. Usually, the iodonium promoter system provides better yield and α-selectivity for the glycosidation of a 2-phenylsulfenyl donor. The examples are shown in *Scheme 19*. The direct sialylation of the primary hydroxyl of trisaccharide **61** with sialyl donor **58**, which was promoted by NIS-TfOH in MeCN, gave tetrasaccharide in 90% yield as an anomeric mixture ($\alpha/\beta = 74/16$) [72]. Furthermore, 2-phenylsulfenyl donor **58**, upon activation by NIS-TMSOTf, was incorporated into the hindered secondary hydroxyl within tetrasaccharide **62** in 40% as α-glycoside [73] (*Fig. 7*). The phenylthioglycoside derivatives of α-(2–8)-linked dimeric, trimeric, and tetrameric sialic acid as glycosyl donors were also effectively applied to the synthesis of the oligo-sialyl glycoconjugates [74,75,76,77], for example, ganglioside GQ1b [78] and GT3 [79].



Scheme 17

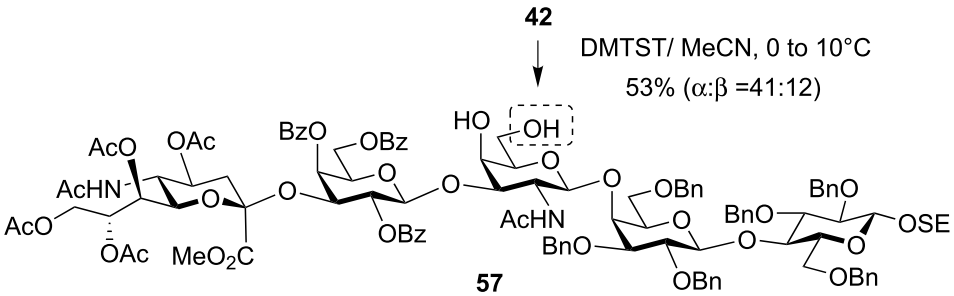
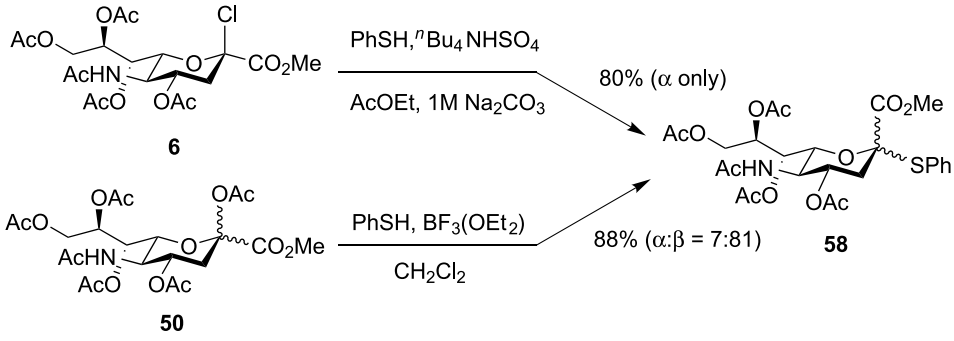
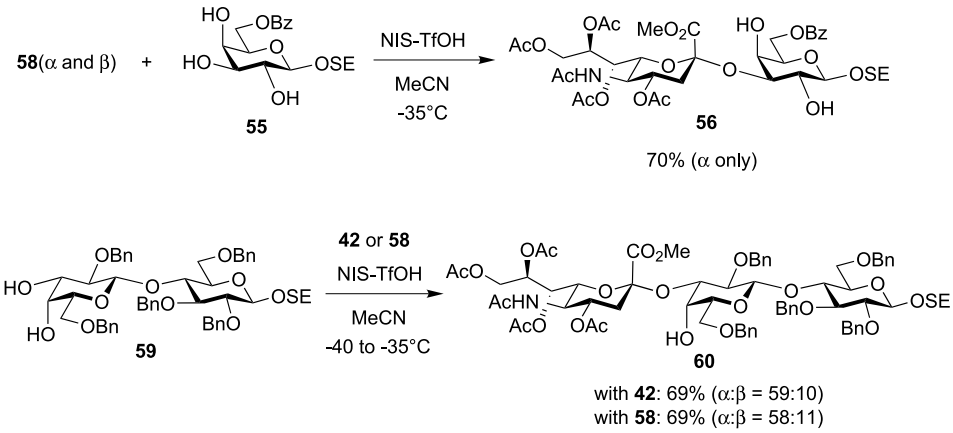


Figure 6
Sialylation of pentasaccharide acceptor with sialyl donor 42



Scheme 18



Scheme 19

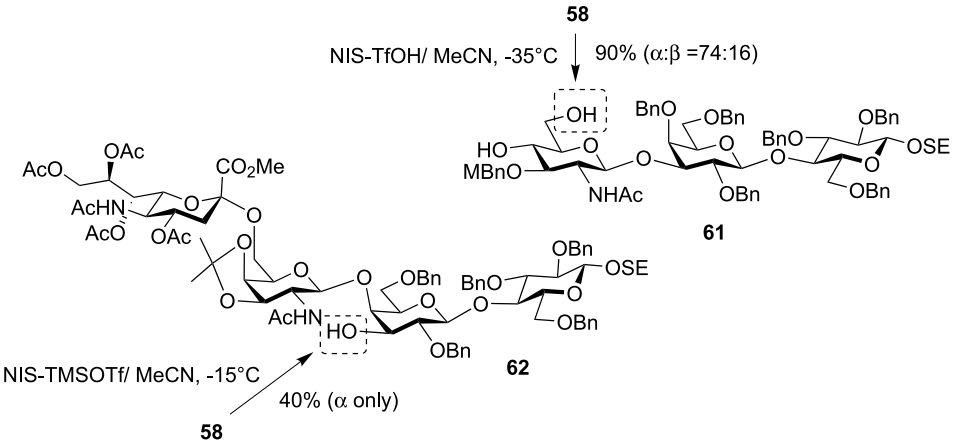
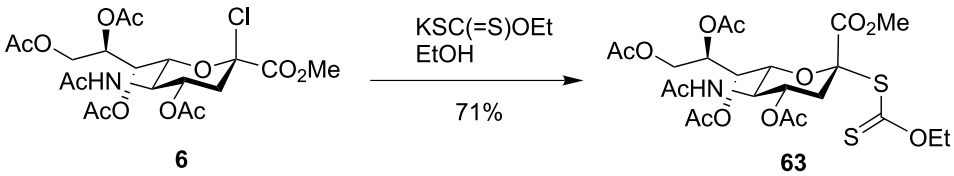


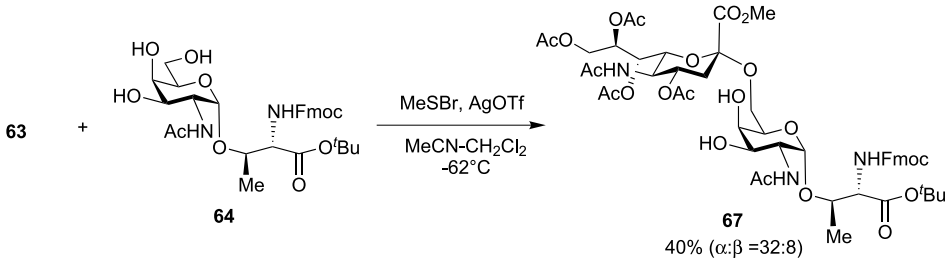
Figure 7
Sialylation of oligosaccharide acceptors with sialyl donor 58

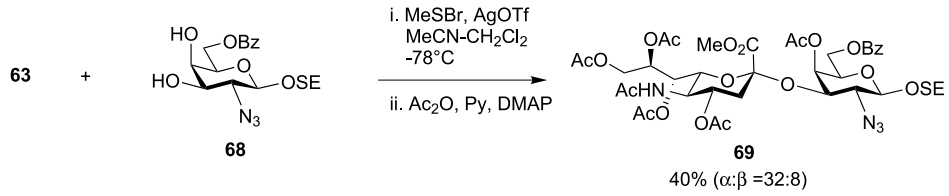
A 2-(ethoxy)dithiocarbonate (xanthate) derivative of Neu5Ac **63** has been synthesized from the 2-chloro derivative **6** by reaction with potassium ethoxydithiocarbonate in EtOH (*Scheme 20*) [40]. This compound was obtained in a stable crystalline form, which has a good shelf-life.

The 2-xanthate derivative can also be activated by DMTST and NIS-TfOH in MeCN, but the coupling yield and α-selectivity were moderate in many cases. This problem has been solved by the introduction of MeSOTf as a glycosylation promoter by Bilberg and Lönn, which can be generated in situ by the reaction of methylsulfenyl bromide with AgOTf



Scheme 20





Scheme 21

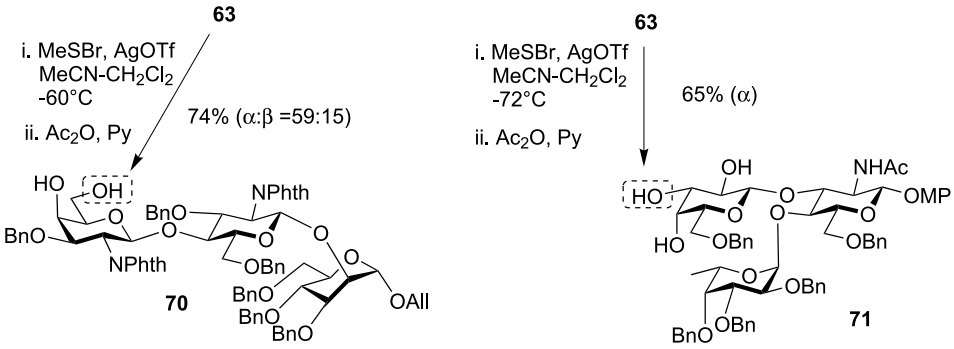
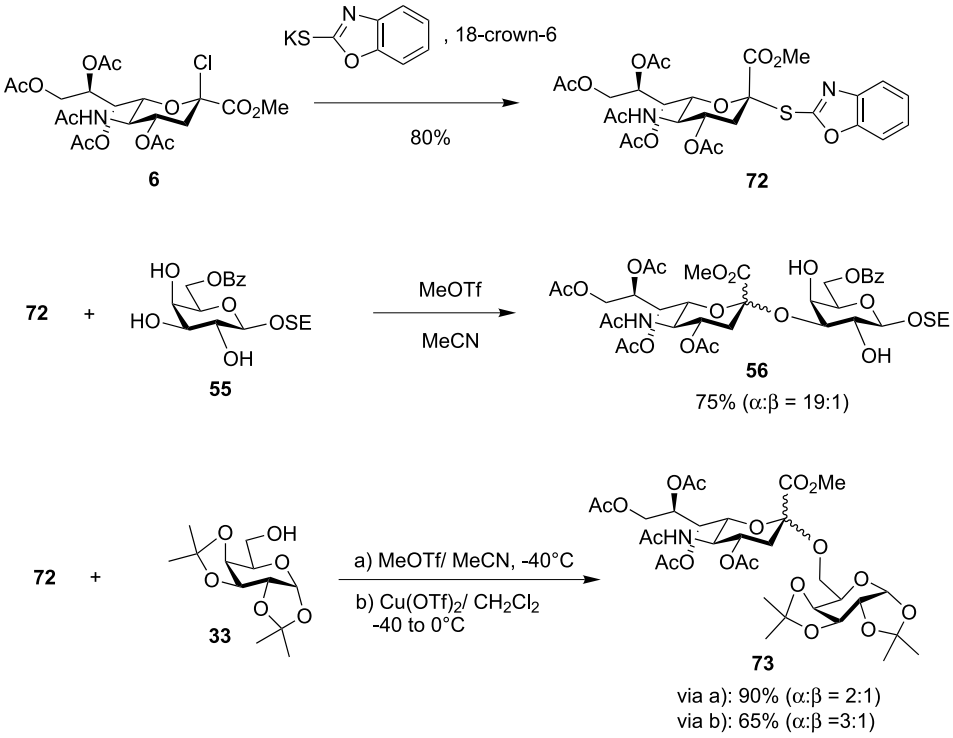


Figure 8
Sialylation of trisaccharide acceptors with sialyl donor 63

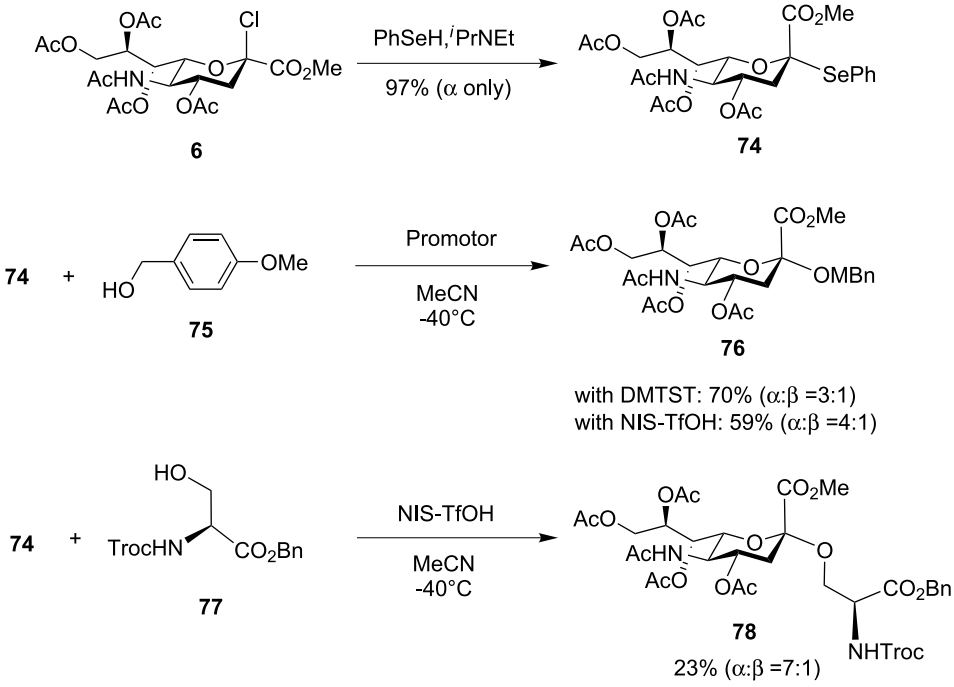


Scheme 22

(*Scheme 21*) [80,81,82]. Bilberg and Lönn also found the significance of solvent for glycoside formation; the use of a mixture of MeCN-CH_2Cl_2 (3/2) afforded the best result. The optimized activation system extended the usage of 2-xanthate derivative **63** to the assembly of a variety of sialyl oligosaccharides as shown in *Fig. 8* [83,84].

Recently, De Meo and Parker introduced 2-benzooxazoylsulfenyl (SBox) derivative **72** [44]. The compound was straightforwardly prepared by the treatment of the 2-chloride derivative **6** with potassium 2-mercaptobenzoxazolate and 18-crown-6 in acetone in 80% yield (*Scheme 22*). For the activation of the 2-SBox donor **72**, MeOTf proved to be the best option. Glycosylation of the donor **72** with 6-*O*-benzoyl-galactosyl acceptor **73** in MeCN at −40 °C yielded α-(2–3)-sialyl galactoside **56** in 71% yield, α-selectively. However, with the more reactive primary hydroxyl group of galactose acceptor **33**, the glycosidation was less α-selective (90%, $\alpha/\beta = 2/1$). Interestingly, when this coupling was promoted by $Cu(OTf)_2$ in CH_2Cl_2 at −40 °C to 0 °C, α-selectivity was improved ($\alpha/\beta = 3/1$).

The 2-phenylselenyl derivative of Neu5Ac **74**, which can be prepared by the treatment of 2-β-chloride **6** with selenophenol and Hünig's base, was also introduced as a glycosyl donor [85]. This compound can be activated by DMTST or NIS-TfOH in MeCN at −40 °C to afford α-sialoside. In general, the yield and α-selectivity were not constantly high (*Scheme 23*).



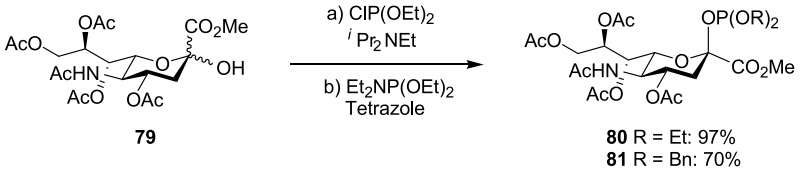
Scheme 23

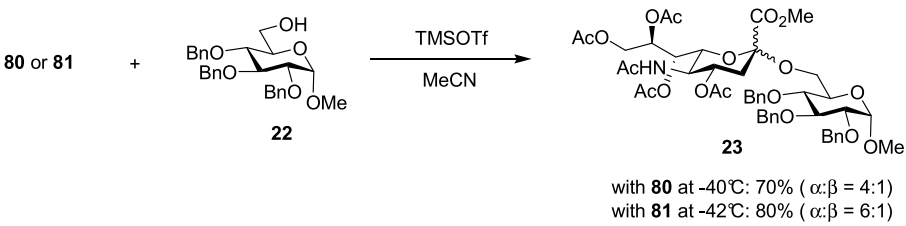
3.4 2-Phosphite and 2-Imidate Sialyl Donors

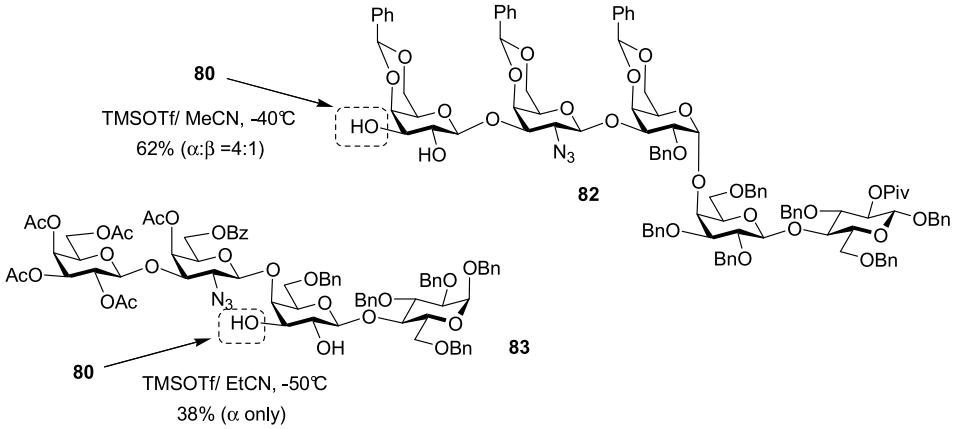
3.4.1 2-Phosphite Derivatives

Schmidt and Wong have developed 2-diethyl and 2-dibenzyl phosphite derivatives of Neu5Ac **80** [86] and **81** [87], individually. Both of the 2-phosphite derivatives can be prepared from a 2-hydroxy derivative of Neu5Ac **79** by the reaction with $ClP(OEt)_2$ in the presence of Hünig's base, and by the reaction of *N*,*N*-diethyl phosphoro amidite and tetrazole, respectively (*Scheme 24*).

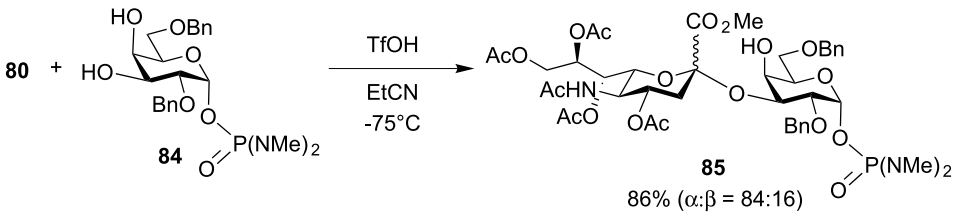
Because of the fact that *N*,*N*-diethyl phosphoro amidite is not commercially available, there is limited application of dibenzyl phophite **81** to oligosaccharide synthesis. On the other hand, the power of diethyl phophite **80** as a useful alternative to the other Neu5Ac donors was exemplified by its incorporation within many linear and convergent syntheses of the complicated sialyl oligosaccharides. The glycosidation of 2-phosphite derivative **80** can be catalyzed by 10 to 20 mol % of acid [TMSOTf, $Sn(OTf)_2$, etc]. This feature has allowed design of synthetic strategies that involve the use of acid-labile protecting groups within glycosylation units [88,89]. Moreover, Hashimoto's group performed, within the synthesis of ganglioside GM3, the orthogonal coupling of 2-diethyl phosphite of Neu5Ac **80** with 1-phosphoroamidite galactose derivative **84** by using TfOH as a glycosylation catalyst (*Scheme 25*) [90].







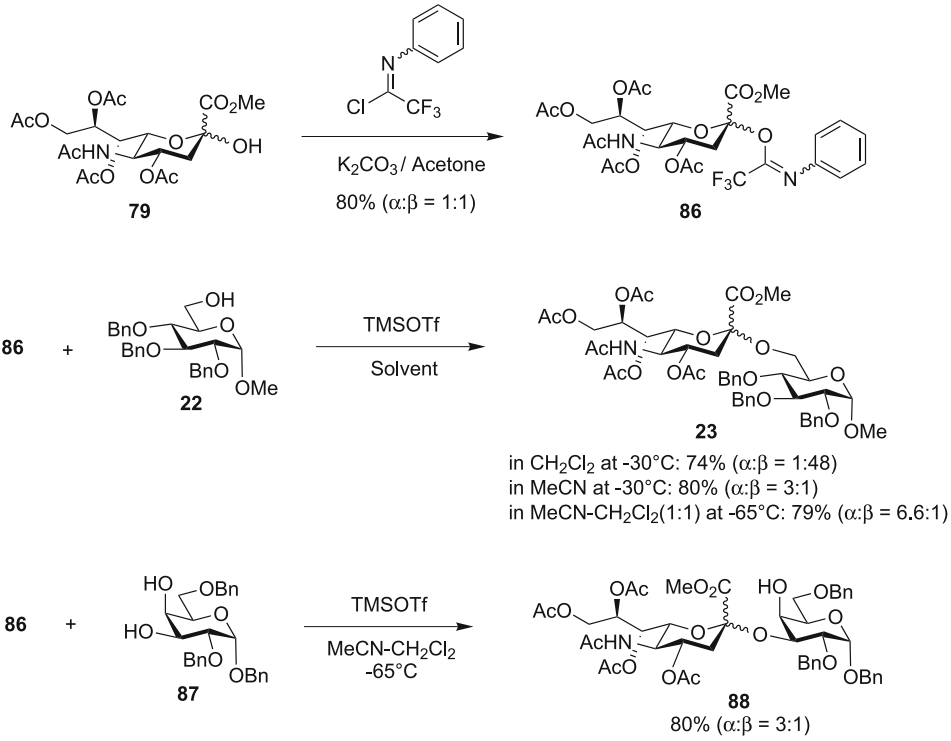
Scheme 24



Scheme 25

3.4.2 2-Imidate Derivative

It is known that the attractive trichloroacetimidate glycosyl donor invented by Schmidt's group [91] can not be incorporated into sialylation, mostly due to the poor stability of the corresponding 2-trichloroacetimidate derivative. Recently, Yu's group successfully incorporated a 2-trifluoroacetimidate derivative of Neu5Ac into α-sialoside synthesis [92]. The reaction of a 2-hydroxy derivative of Neu5Ac **79** with *N*-phenyltrifluoroacetimidoyl chloride and K_2CO_3 in acetone produced the corresponding 2-trifluoroacetimidate derivative **86** in 80% yield as a mixture of α- and β-isomers (α/β = 1/1) (*Scheme 26*). In the coupling reaction with the C6 hydroxyl of glucosyl acceptor **22**, the use of a mixture of MeCN-CH_2Cl_2 (1/1) as solvent provided high α-selectivity, thereby yielding disaccharide **23** in 79% yield as a stereo mixture (α/β = 6.6/1). The efficacy of 2-trifluoroacetimidate derivative **86** as a sialyl donor was also demonstrated by conducting the glycosylation of 3,4-diol galactosyl acceptor **87** in MeCN again, producing disaccharide **88** in 80% yield as a stereo mixture (α/β = 3/1).



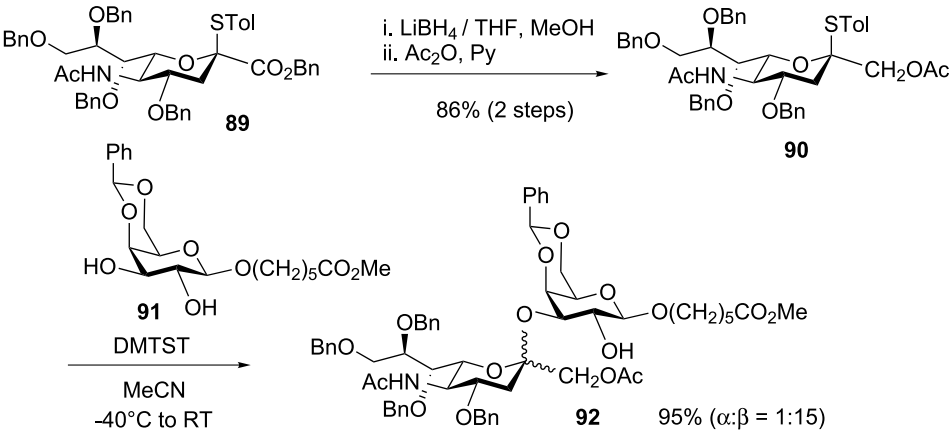
Scheme 26

4 The Third Period

Looking at the array of Neu5Ac donors mentioned above, it is of note that the type of leaving group, its activation system, and solvent dominantly affect the reactivity of the glycosyl donor in terms of the efficiency of glycosidation and the α-anomeric selectivity. Furthermore, it is known that the feasibility of glycosidation depends on the unpredictable "match-mismatch" relationship between coupling partners. This issue makes sialyl glycan synthesis much more complicated. In this context, many researchers are still eager to develop a universal method for α-sialylation.

4.1 Effect of Reduction of the C1 Carbonyl Group on the Sialy Donor

In work on nitrile solvent-dependent α-sialylation, the importance of the C1 carbonyl group for α-selectivity was suggested by Wong's group [93]. Thus, they converted the carbonyl group of Neu5Ac derivative **89** into a hydroxymethyl group to obtain the Neu5Ac derivative **90** (*Scheme 27*). The DMTST-promoted glycosylation of the derivative **90** with 4,6-benzylidenated galactosyl acceptor **91** afforded the mixture of α- and β-sialylated products **92** in 90% yield with the ratio of 1/15. Although they also examined glycosylation by altering acceptors and promoters, β-sialosides were preferentially produced in all events.



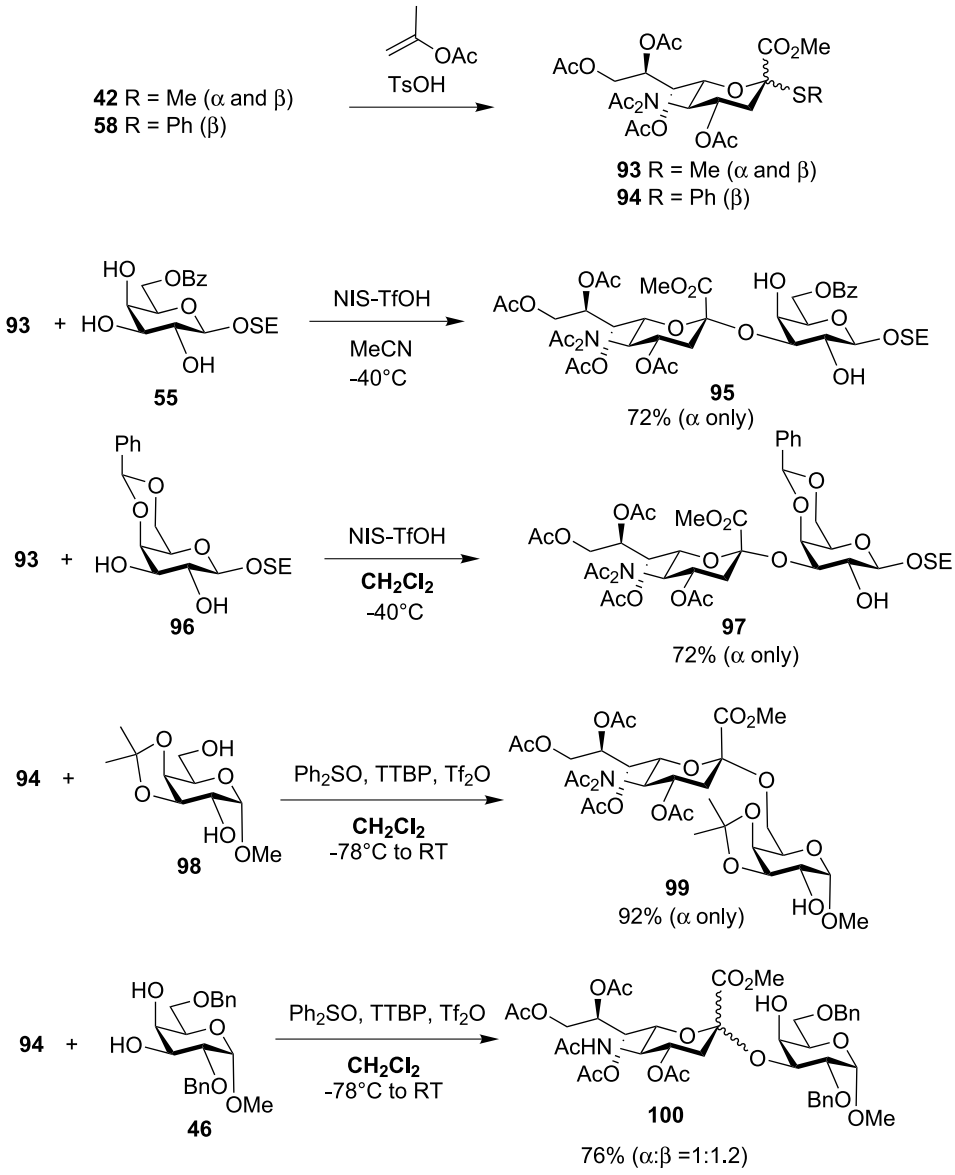
Scheme 27

4.2 C5-Modified Sialyl Donors

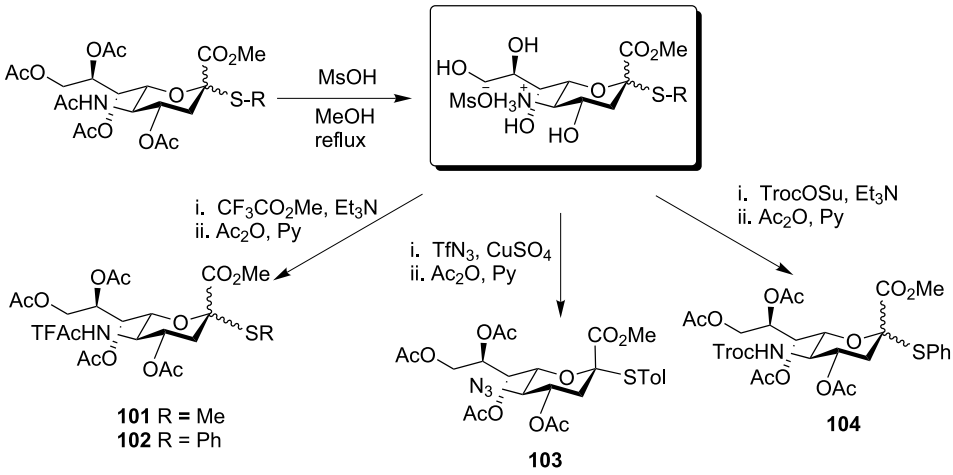
4.2.1 *N,N*-Diacetyl Derivatives

On the other hand, Boons and co-workers demonstrated that the conversion of the C5 *N*-acetyl group into a *N*,*N*-diacetyl group raised the coupling efficiency and α-selectivity during sialylation [94,95].

The initial purpose of the acetylation of the C5 acetamido group is to prevent *N*-methylation in the course of MeOTf-promoted glycosidation using methylthioglycoside of Neu5Ac-α-(2–3)-Gal as a glycosyl donor. Without the acetyl protection, the glycosidation with a glucosyl acceptor produced *N*-acetyl-*N*-methyl trisaccharide in 70% yield. *N*,*N*-Diacetyl



Scheme 28



Scheme 29

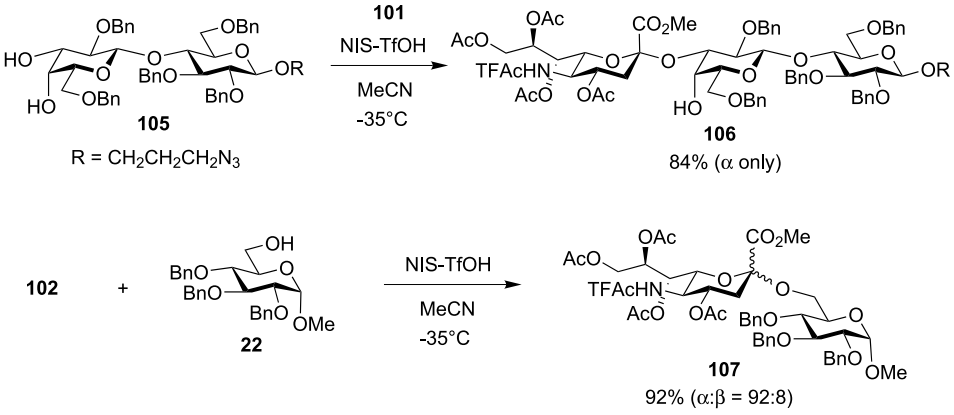
derivative **93** can be prepared by the acid-catalyzed acetylation of the *N*-acetyl derivative **42** with isopropenyl acetate at 65 °C in 99% yield (*Scheme 28*). They found that the glycosidation of the *N*,*N*-diacetyl sialy donor **93** with the suitably protected galactosyl acceptor **55** by activation with NIS-TfOH in MeCN was swift and terminated within 5 min, and the sialyl outcome **95** was produced as a single α-isomer in high yield (72%). More surprisingly, the sialylation of 4,6-benzylidenated galactosyl acceptor **96** in CH_2Cl_2 also exhibited perfect α-selectivity, producing disaccharide **97** in 72% yield. Upon the action by diphenyl sulfoxide, Tf_2O, and TTBP, the 2-phenylsulfenyl derivative **94** was condensed with the C6 hydroxyl of galactose derivative **98** in CH_2Cl_2 to produce α-glycoside **99** solely [96]. However, the α-selectivity differed depending on the structure of the coupling partner.

The new finding of the *remote functionality effect* on the sialyl donor property has given rise to a variety of C5-modified sialyl donors **101**–**104** (*Scheme 29*).

4.2.2 *N*-Trifluoroacetyl Derivatives

Originally, the *N*-trifluoroacetyl (TFAc) derivative of sialic acid was introduced by Bovins and co-workers as a 2-bromide glycosyl donor, which was reacted with benzyl alcohol in the presence of sym-collidine to give a 1/2 mixture of α- and β-glycosides in 30% yield [16]. In our laboratory, the 2-phenylsulfenyl derivative of *N*-TFAc sialic acid has been used for the synthesis of de-*N*-acetyl-6-sulfo-sialyl Lewis X [97]. However, Boons and co-workers were again first to describe the high potential of the *N*-TFAc sialyl donor [98].

The glycosidation of the methylthioglycoside of *N*-TFAc sialic acid derivative **101** with 3′,4′-dihydroxy lactoside derivative **105** in the presence of NIS-TfOH at −35 °C in MeCN produced sole α-sialoside with improved yield (84%), which also terminated within 5 min (*Scheme 30*). Also, 2-phenylsulfenyl derivative **102** exhibited excellent glycosyl donor properties in the coupling reaction with the 6-hydroxy derivative of glucoside **22**, producing disaccharide **107** with high α-selectivity. Furthermore, the highly reactive sialyl donor



Scheme 30

was also successfully applied to the synthesis of α-(2–8)-linked disialic acid (discussed in *Section 5.1.2*).

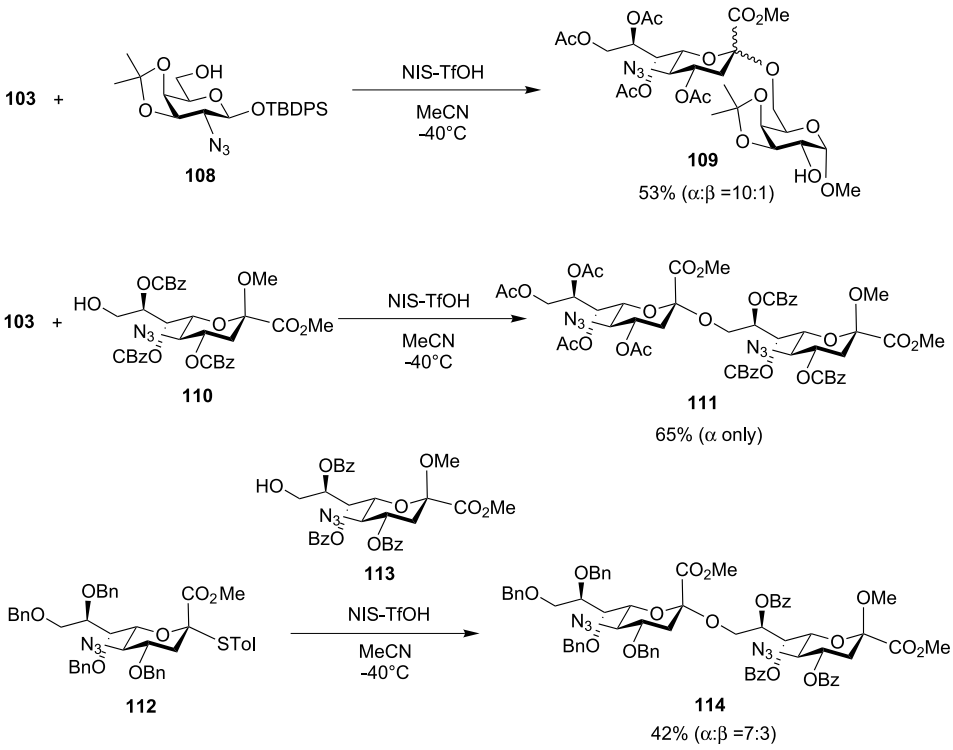
4.2.3 5-Azido Derivatives

It was also reported that the substitution of the C5 acetamido group with an azido group enhances the α-selectivity during sialylation in MeCN, especially with primary sugar hydroxyls [99,100]. The azido group can be introduced at the C5 position by diazo transfer reaction of the corresponding 5-amino derivative with TfN_3 to afford the C5 azido derivative, which was successively acetylated (vide supra) [99]. As shown in *Scheme 31*, the power of the 5-azido donor **103** was successfully applied to the synthesis of α-(2–9)-linked disialic acid, where 9-hydroxy-5-azido derivative **110** served as the glycosylation partner. When the acetyl groups within **103** were replaced with benzyl groups, however, the yield of α-sialyl product was extremely diminished [101].

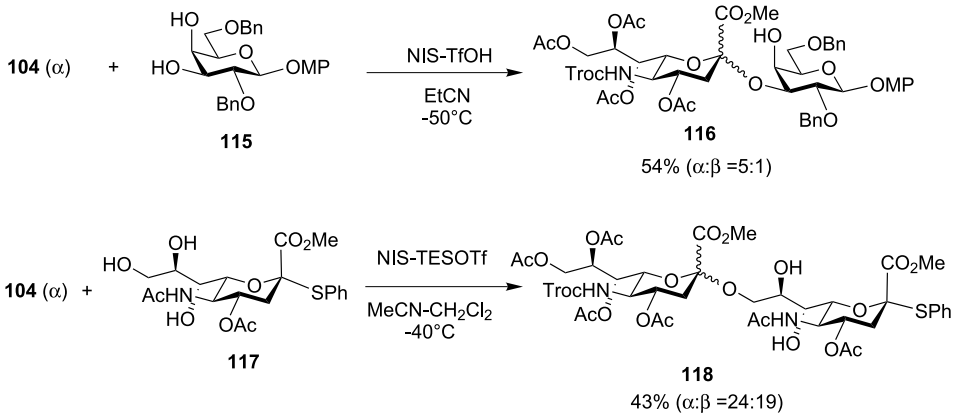
4.2.4 *N*-2,2,2-Trichloroethoxycarbonyl Derivatives

The 2,2,2-trichloroethoxycarbamoyl (*N*-Troc) group at the C5 position also serves as an enhancer of the reactivity of a sialic acid donor during glycosidation. The synthesis of the *N*-Troc phenylthioglycoside donor was first introduced by Wu and co-workers; it was incorporated within the synthesis of a α-(2–5)-linked oligomer of Neu5Gc [102]. There is easy access to *N*-Troc derivative **104** from the corresponding 5-amino derivative as shown in *Scheme 29*. The high potential of α- and β-phenylthioglycoside derivatives **104** were explored by our group [103,104] and Takahashi's group [105,106].

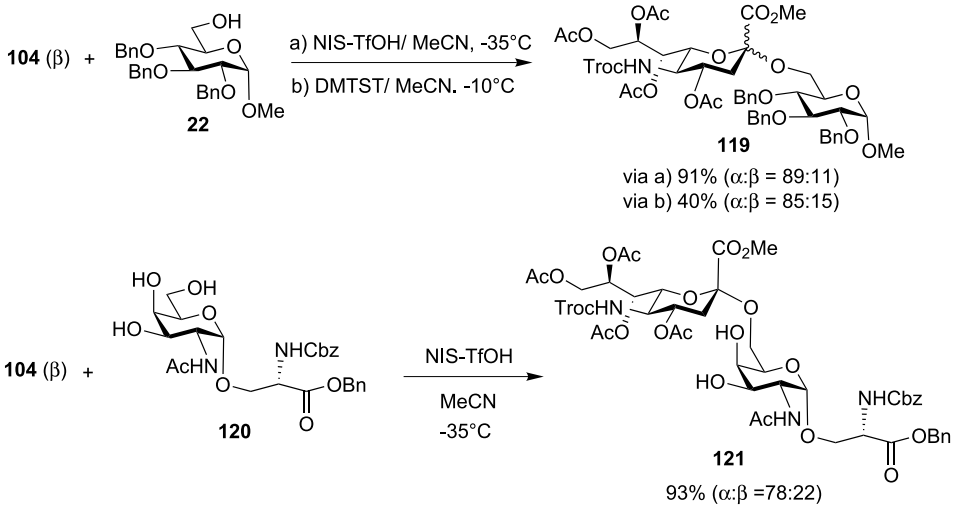
α-Sialylation with the *N*-Troc sialyl donor in MeCN is more effective by comparison with the corresponding *N*-acetyl donor. The glycosylation of 3,4-dihydroxy galactosyl acceptor **115** with *N*-Troc donor **104** in the presence of NIS-TfOH in EtCN at −50 °C terminated within 20 min to produce α-sialoside in 54% yield with β-isomer (10%), whereas the reaction with the *N*-acetyl donor was sluggish at −25 °C in MeCN to afford α-sialoside in 23%



Scheme 31



Scheme 32

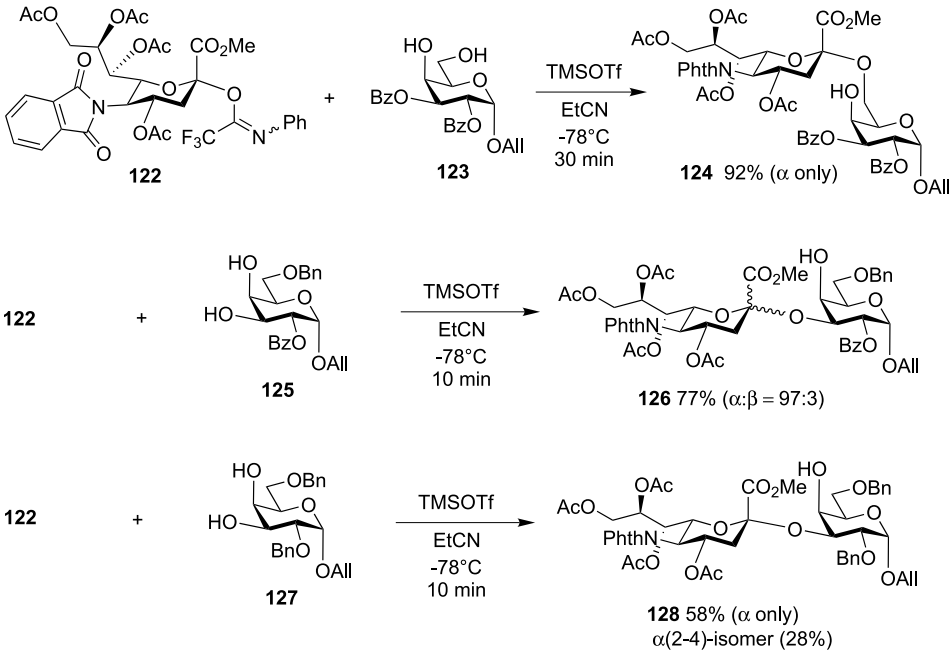


Scheme 33

yield (*Scheme 32*). Because of the distinct difference of the reactivity between *N*-Troc and *N*-acetyl donors, an "armed-disarmed"-like coupling reaction could be performed between the potent phenylthioglycoside of *N*-Troc sialic acid **104** and the less reactive phenylthioglycoside of *N*-acetyl sialic acid **117**, producing (2–9)-linked disialoside **118** in 43% yield ($\alpha/\beta = 24/19$). In Takahashi's group, β-phenylthioglycoside of *N*-Troc sialic acid proved to be a prominent glycosyl donor during work on glycosidation with a variety of C5-functionalized sialyl donors. The NIS-TfOH system has better affinity to the *N*-Troc sialyl donor **104** than DMTST does. An example is shown in *Scheme 33*. Furthermore, in the coupling reaction of *N*-Troc donor **104** and the primary hydroxyl of galactosyl serine derivative **120**, the quantity of molecular sieves (3A) and the concentration of substrates turned out to influence the efficiency of the reaction. Under optimized conditions (molecular sieve 3 Å; 0.5 g mmol^{-1}), the glycosylation of **120** with **104** in MeCN (5 mL mmol^{-1}) fashioned sialyl Tn antigen glycan **121** in 84% yield. They also have incorporated the *N*-Troc donor into a one-pot synthesis of sialyl T antigen frame, successfully.

4.2.5 *N*-Phthaloyl Derivative

Recently, a *N*-phthalimido-2-trifluoroacetimidate derivative of sialic acid **122** has been introduced as an efficient glycosyl donor (*Scheme 34*) [107]. It could be glycosidated with C6 hydroxyl of galactosyl acceptors **123** in MeCN catalyzed by TMSOTf to exclusively produce α-sialyl outcome **124**. Similarly, the glycosidation with *p*-nitrophenyl 2,6-di-*O*-benzylated galactose derivative **125** produced α-sialoside **126** in 72% yield along with a slight quantity of β-isomer. However, when the benzoyl group at the C2 position was replaced with a benzyl group, the sialylation also took place at the C4 hydroxyl to produce **128** (58%) along with $\alpha(2 \rightarrow 4)$-isomer (28%).

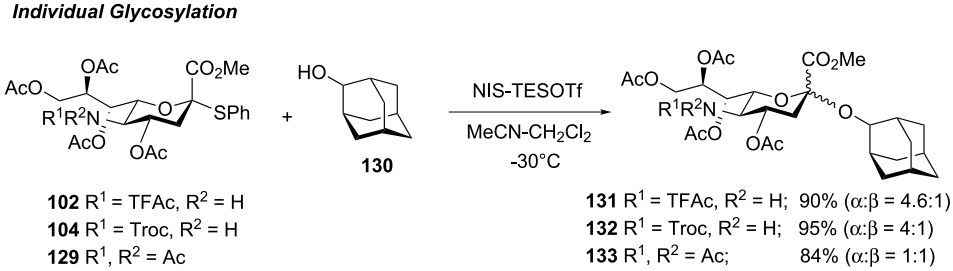


Scheme 34

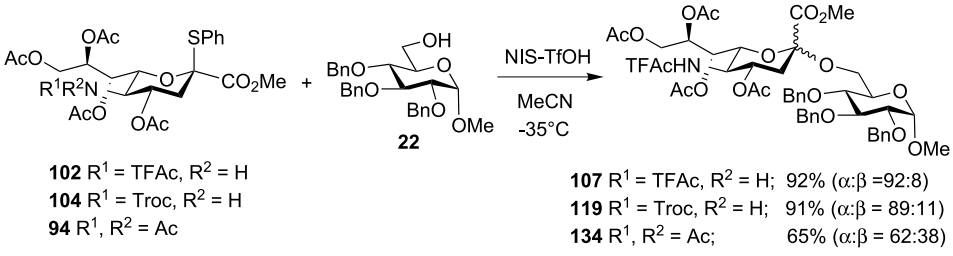
4.2.6 Comparison of Glycosyl Donor Properties

When making a comparison of sialyl donor properties, one should consider that it is difficult to conclude the general order between various sialyl donors in terms of glycosidation efficiency, because the efficiency differs depending on the match-mismatch relationships between glycosyl donor and acceptor. Accordingly, the relationship is related to the relative energy level of the transition state of the substrate complex, which comprises an oxocarbenium ion, acceptor, anionic species derived from the glycosylation promoter (triflate anion, etc), solvent molecule, and so on. Therefore, one can conclude the individual order between sialyl donors in terms of glycosidation efficiency for each reaction system.

In order to compare the glycosylation efficiency of C5-modified sialyl donors, our group has conducted competitive coupling reactions (*Scheme 35*) [103]. The individual glycosylation of *N*,*N*-diacetyl, *N*-TFAc, and *N*-Troc donors (**129**, **102**, and **104**) with 2-adamantanol **130** promoted by NIS-TfOH in MeCN-CH_2Cl_2 (5/1) provided sialyl glycosides **133** (84%, $\alpha/\beta = 1/1$), **131** (90%, $\alpha/\beta = 4.6/1$), **132** (95%, $\alpha/\beta = 4/1$), respectively. In all events, reactions terminated within 10 min. Next, competitive coupling reactions of **102**, **104**, and **129** under similar conditions afforded *N*-Troc sialoside **132** (40%, $\alpha/\beta = 4/1$), *N*,*N*-diacetyl sialoside **133** (22%, $\alpha/\beta = 1/1$), and *N*-TFAc sialoside **131** (15%, $\alpha/\beta = 4/1$). This result suggests the hierarchy of *α-sialylation efficiency* is *N*-Troc donor **104** > *N*-TFAc donor **102** > *N*,*N*-diacetyl donor **122** *in the reaction system*.



Scheme 35



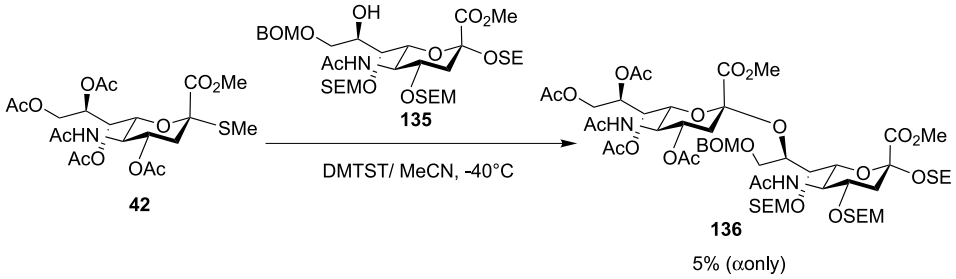
Scheme 36

As depicted in *Scheme 36* [106], the results of glycosylations of the primary hydroxyl of glucosyl acceptor **22** with *N*-Troc, *N*-TFAc, and *N*,*N*-diacetyl sialyl donors (**104**, **102**, and **94**) promoted by NIS-TfOH in MeCN at −35 °C also showed a similar hierarchy of the α-sialylation efficiency between these donors; *N*-Troc sialoside **119** was obtained in 91% ($\alpha/\beta = 89/11$), *N*-TFAc sialoside **107** in 92% ($\alpha/\beta = 92/8$), and *N*,*N*-diacetyl sialoside **134** in 65% yield ($\alpha/\beta = 62/38$), respectively. On the basis of these results, it can be concluded that the efficiency of α-sialylation with a *N*,*N*-diacetyl sialyl donor is less constant than the corresponding *N*-Troc and *N*-TFAc derivatives, depending on the reaction system.

5 Synthesis of an Oligomer of Sialic Acid

5.1 α-(2–8)-Linked Disialic Acid

The synthesis of α-(2–8)-linked disialic acid is complicated because of the remarkably diminished reactivity of the C8 hydroxyl of Neu5Ac. The difficulty was exemplified by the unsuccessful results of the reaction shown in *Scheme 37* [38]. Schmidt has hypothesized that the



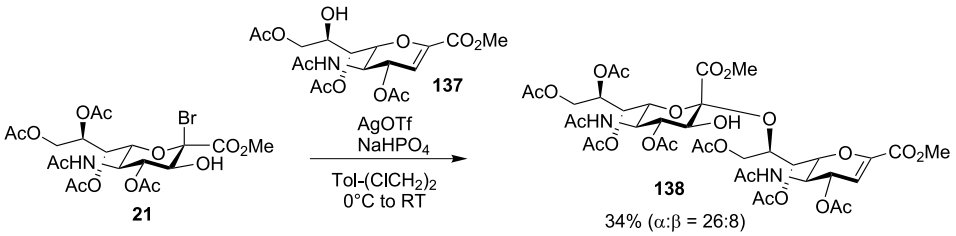
Scheme 37

low nucleophilicity is due to the complicated interaction with the C5 acetamido group, ring oxygen, C1 carboxylate group, or glycosidic C2-OR bond [108,109].

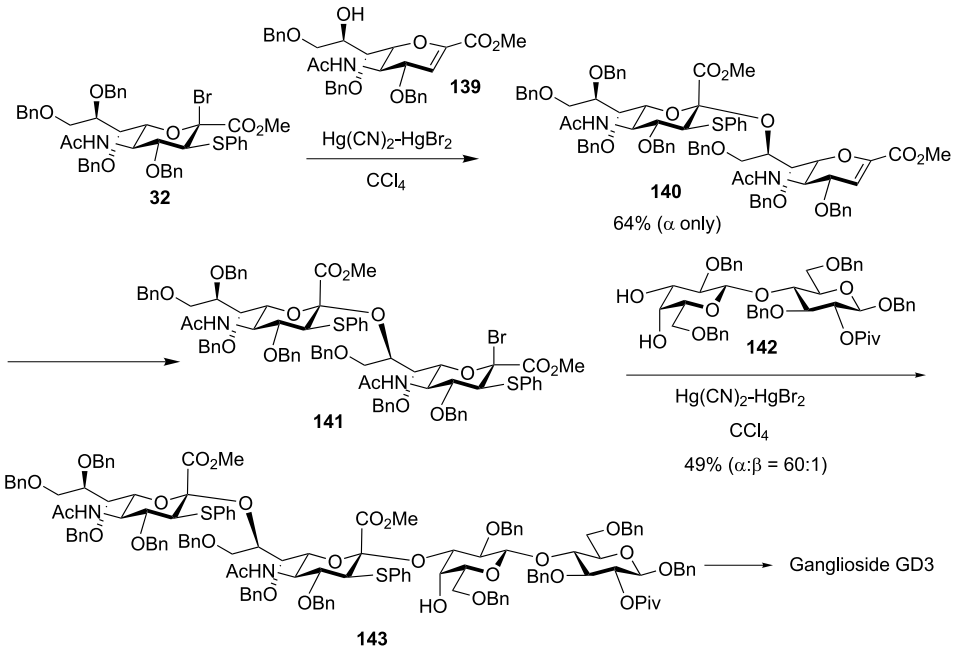
5.1.1 The Use of C3-Appended Sialyl Donors

The first accomplishment of α-(2–8)-disialic acid synthesis came with the combined use of the 3-hydroxy derivative of Neu5Ac **21** and 2,3-en derivative of Neu5Ac **137**, which was expected to prevent the intramolecular hydrogen bonding between the C8 hydroxyl and C1-carboxylate group or C2-OR bond. The coupling reaction was promoted by AgOTf in the presence of $NaHPO_4$ to afford the dimer **138** in 34% yield as a 3/1 mixture of α/β anomers (*Scheme 38*) [110].

Other C3-appended-type donors were also used for the production of α-(2–8)-disialic acid derivatives with the combination of 2,3-en-type glycosyl acceptors [31,111]. Ogawa's group employed the 3-phenylsulfenyl donor **32** and the 4,7,9-tri-*O*-benzylated 2,3-en acceptor **139** in the first total synthesis of ganglioside GD3. The coupling reaction provided single α-isomer of disialic acid **140** in 64% yield. The dimer was further converted to the 3-phenylsulfenyl disialyl derivative **141**, which was glycosidated with C3′ hydroxyl within lactoside derivative **142** in high yield, producing GD3 glycan frame **143** (*Scheme 39*) [110]. The coupling of the C3-appended donor with the Neu5Ac acceptor in $^2C_5$ chair form was also performed by several groups [112,113].



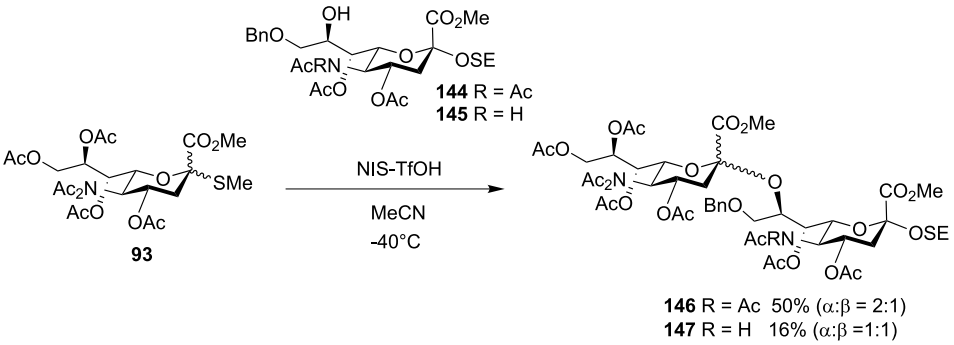
Scheme 38



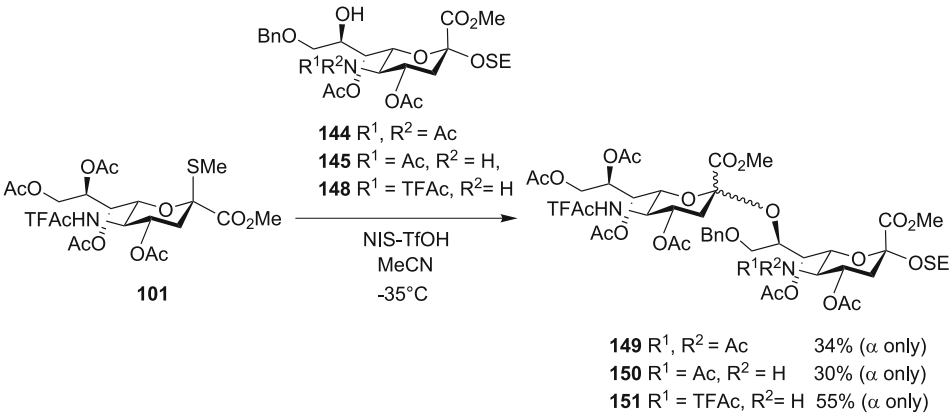
Scheme 39

5.1.2 The Use of C5-Modified Sialyl Units

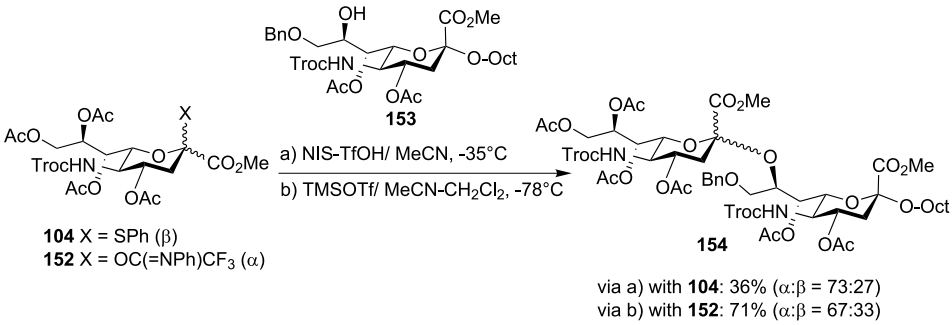
Similar to the reactivity tuning of sialyl donors mentioned above, it was found that the modification of the C5 acetamido group or its substitution by other functionalities also raises the reactivity of the C8 hydroxyl of sialic acid. Boons and co-workers originally synthesized *N*,*N*-diacetyl sialyl acceptor **144**, which was applied to the sialylation with *N*,*N*-diacetyl sialyl donor **93** in MeCN at −40 °C. In comparison to *N*-acetyl sialyl acceptor **145**, the sialylation produced α-(2–8)-disialic acid **146** in raised yield (50%, $\alpha/\beta = 2/1$) (*Scheme 40*) [94,95].



Scheme 40



Scheme 41



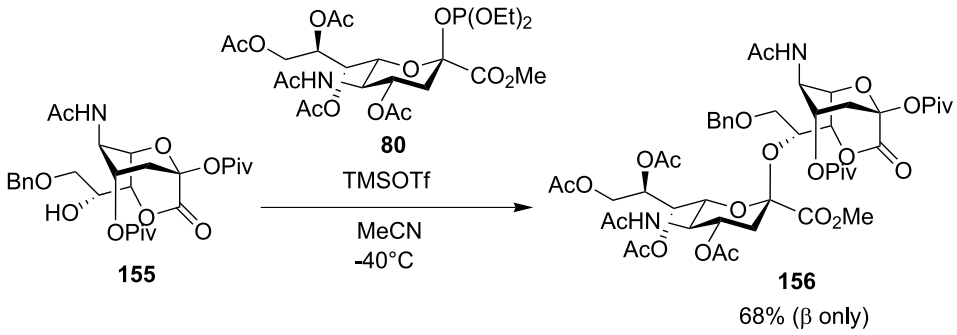
Scheme 42

Boons and co-workers next examined the coupling reaction between *N*-TFAc sialyl donor **101** and acceptor **148** (*Scheme 41*) [98]. As a result, the reaction provided the corresponding dimer **151** in 55% yield as a single isomer. This method enabled the highly stereoselective synthesis of GD3 ganglioside.

Similarly, the coupling reaction between the *N*-Troc sialyl donor and acceptor proved effective. 2-Trifluoroacetimidate-5-*N*-Troc donor **152** was reacted with *N*-Troc sialyl acceptor **153** in the presence of a catalytic amount of TMSOTf at −78 °C in MeCN-CH_2Cl_2 to afford the dimer **154** in 71% yield as a stereo mixture ($\alpha/\beta = 67/33$) (*Scheme 42*) [114].

5.1.3 Conformational Change of the Sialyl Acceptor

In order to prevent the steric hampering of the C5 acetamido group via intramolecular hydrogen bonding, Schmidt's group has converted the $^2C_5$ conformation into $^5C_2$ by making a 1,7-lacton bridge, where the C8 hydroxyl is far from the acetamido group. This improvement of steric circumstance around the C8 hydroxyl resulted in an increase in the yield of glycosylation product, but β-anomer **156** was solely produced (*Scheme 43*) [108].



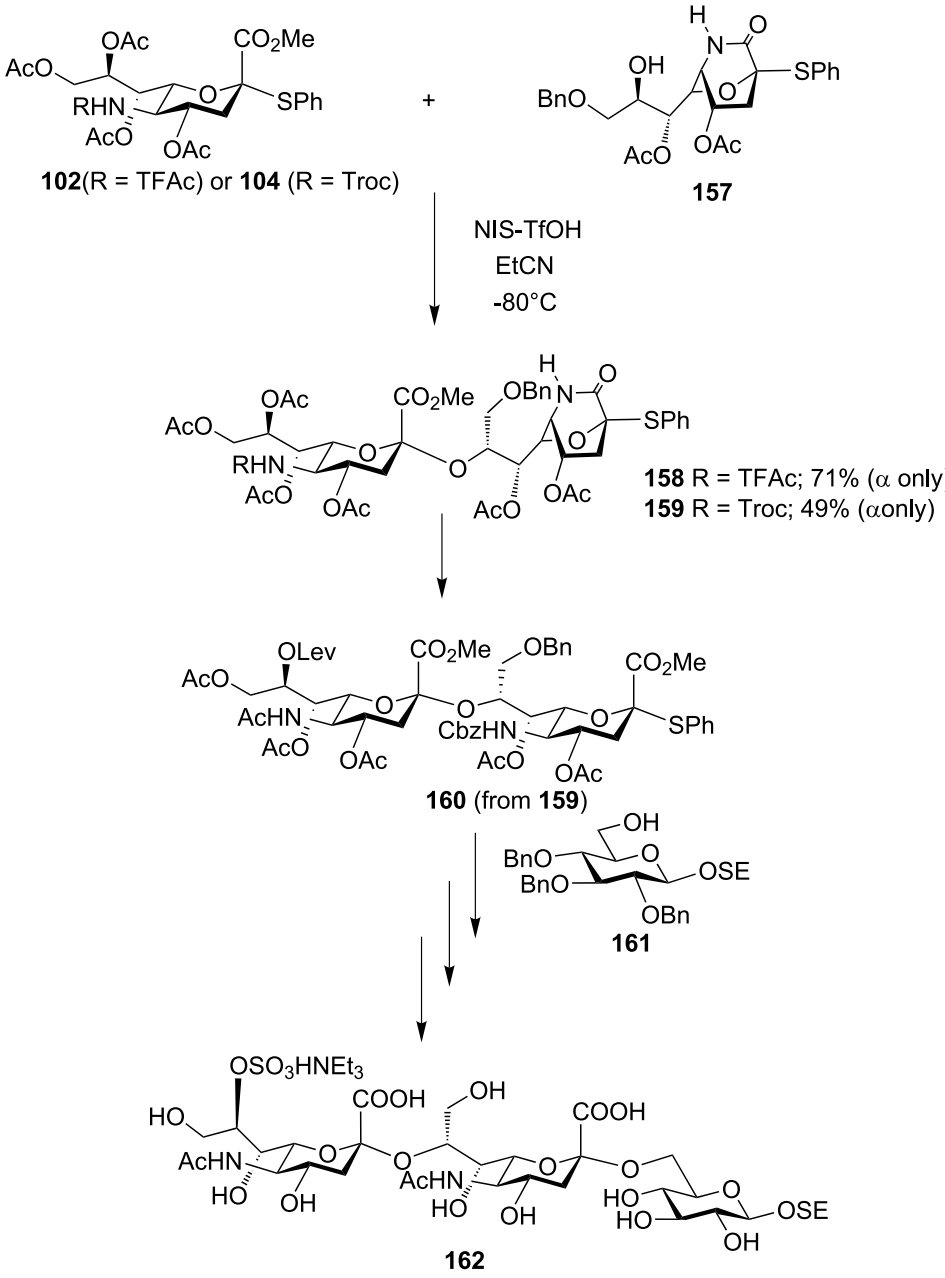
Scheme 43

On the other hand, 1,5-lactamized sialic acid derivative **157**, which was developed by our group, served as an excellent coupling partner for α-sialylation (*Scheme 44*) [115]. Thus, the glycosylation of the C8 hydroxyl of the 1,5-lactamized derivative **157** having a phenylsulfenyl group at C2 with *N*-TFAc and *N*-Troc sialyl donors **102** and **104**, which was promoted by NIS-TfOH in EtCN at −80 °C, delivered α-(2–8)-linked disialic acid **158** and **159** in 71% and 49% yields with complete stereoselectivity, leaving the phenylsulfenyl group at the bridgehead carbon unaffected. For the further glycosidation, the lactam ring was opened to retrieve the $^2C_5$ conformation, and the resulting disialyl donor **160** was enabled to glycosidate with the C6 hydroxyl of glucosyl acceptor **161** to produce the Neu-α-(2–8)-Neu-α-(2–6)-Glc sequence, which led to the glycan structure of Hp-s6 ganglioside **162** having a 8-sulfonyl-Neu5Ac residue.

5.2 α-(2–8)-Linked Oligosialic Acid

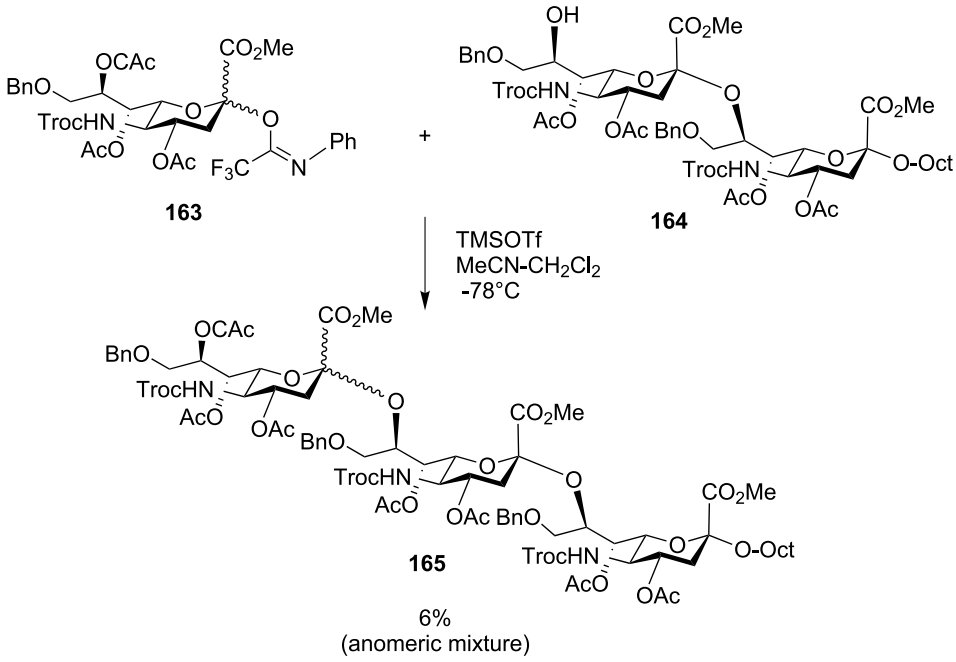
Despite the development of the exquisite methods for α-(2–8)-linked disialic acid synthesis discussed above, the assembly of the α-(2–8)-linked oligomer has been far from a success. Recently, Tanaka disclosed the unpredictable difficulty of the oligomer synthesis. As mentioned above, the coupling of *N*-Troc-protected sialic acid donor and acceptor successfully delivered the dimer of sialic acid. However, the *N*-Troc donor **163** was hardly incorporated into the *N*-Troc disialyl acceptor **164**: the glycosylation produced the stereo mixture of trisialic acid **165** in 6% yield (*Scheme 45*) [114].

On the basis of this lesson, Tanaka has redesigned the strategy of α-(2–8)-linked oligosialic acid synthesis [116]. As shown in *Scheme 46*, the 5-*N*-4-*O*-oxazolidinone moiety was introduced into the sialyl donor and acceptor. Surprisingly, the coupling reaction of the sialy donor **166** and acceptor **167** promoted by NIS-TfOH provided solely α-(2–8)-linked dimer **168** in 86% yield without the assistance of MeCN. Also in the case of glycosylation of the C6 hydroxyl of the glucoside derivative in CH_2Cl_2, α-sialoside was again the sole product. Furthermore, the sialylation of the C8 hydroxyl within the corresponding dimer **169** and trimer **171** acceptors with the 4,5-oxazolidinone sialyl donor **166** proved α-selective and high-yielding. Thereby, the α-(2–8)-linked tetrasialic acid has been synthesized for the first time.



Glycan part of Hp-s6

Scheme 44



Scheme 45

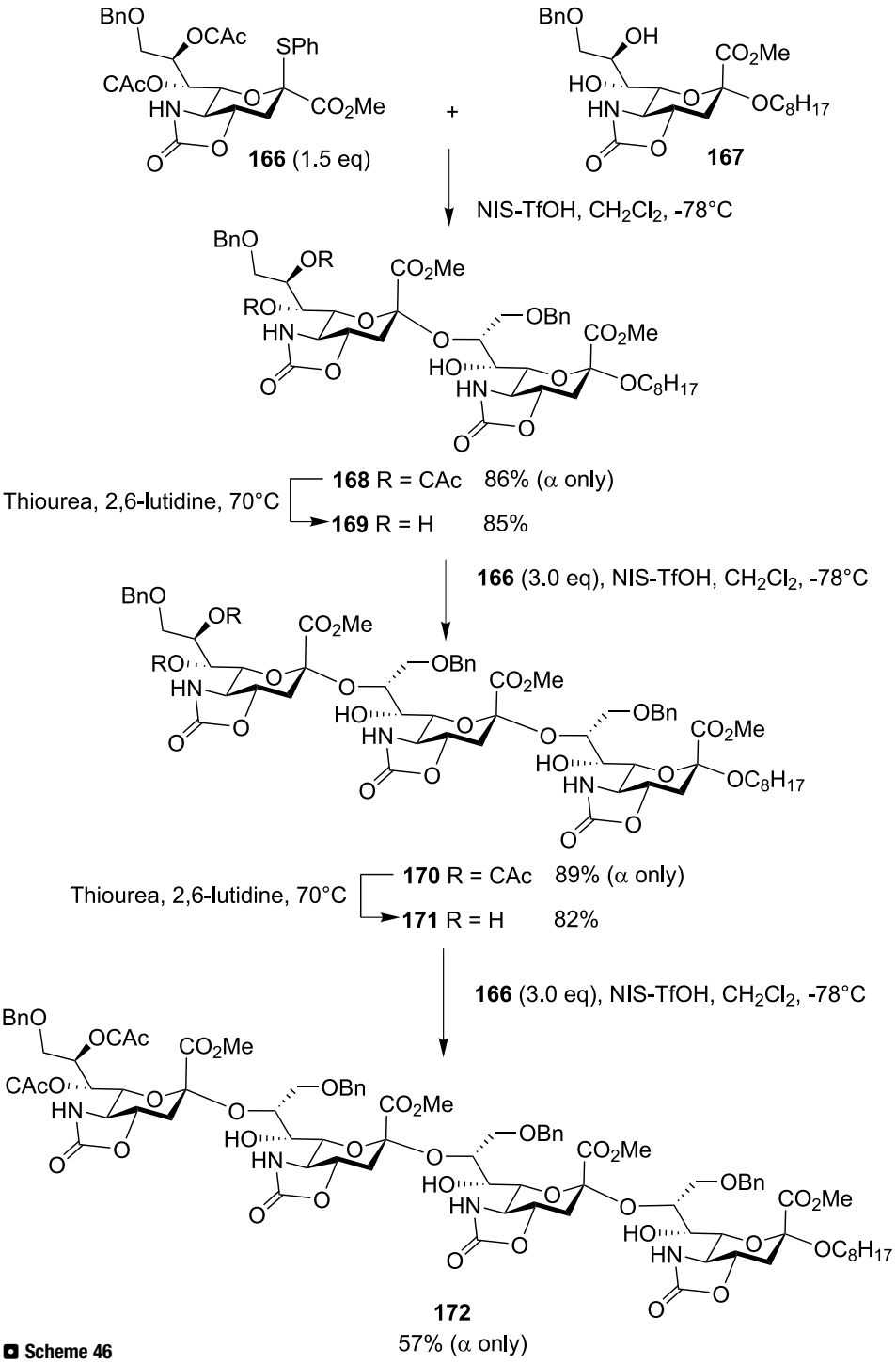
5.3 α-(2–9)-Linked Oligosialic Acid

α-(2–9)-Linkage between sialic acids can be assembled by straightforward chemistry, due to the high reactivity of the C9 primary hydroxyl of Neu5Ac. Initial synthesis has been accomplished by the coupling reaction of the 2-chloro sialyl donor and 9-hydroxy derivative of Neu5Ac in the presence of $Hg(CN)_2$-$HgBr_2$ [117]. To date there have been several reports on α-(2–9)-linked dimer synthesis, where the thioglycoside donor, C3-appended donor, or C5-modified donor was successfully used [95,99,110,116,118]. For a recent example, the assembly of α-(2–9)-linked pentasialic acid by Lin's group is shown in *Scheme 47* [119].

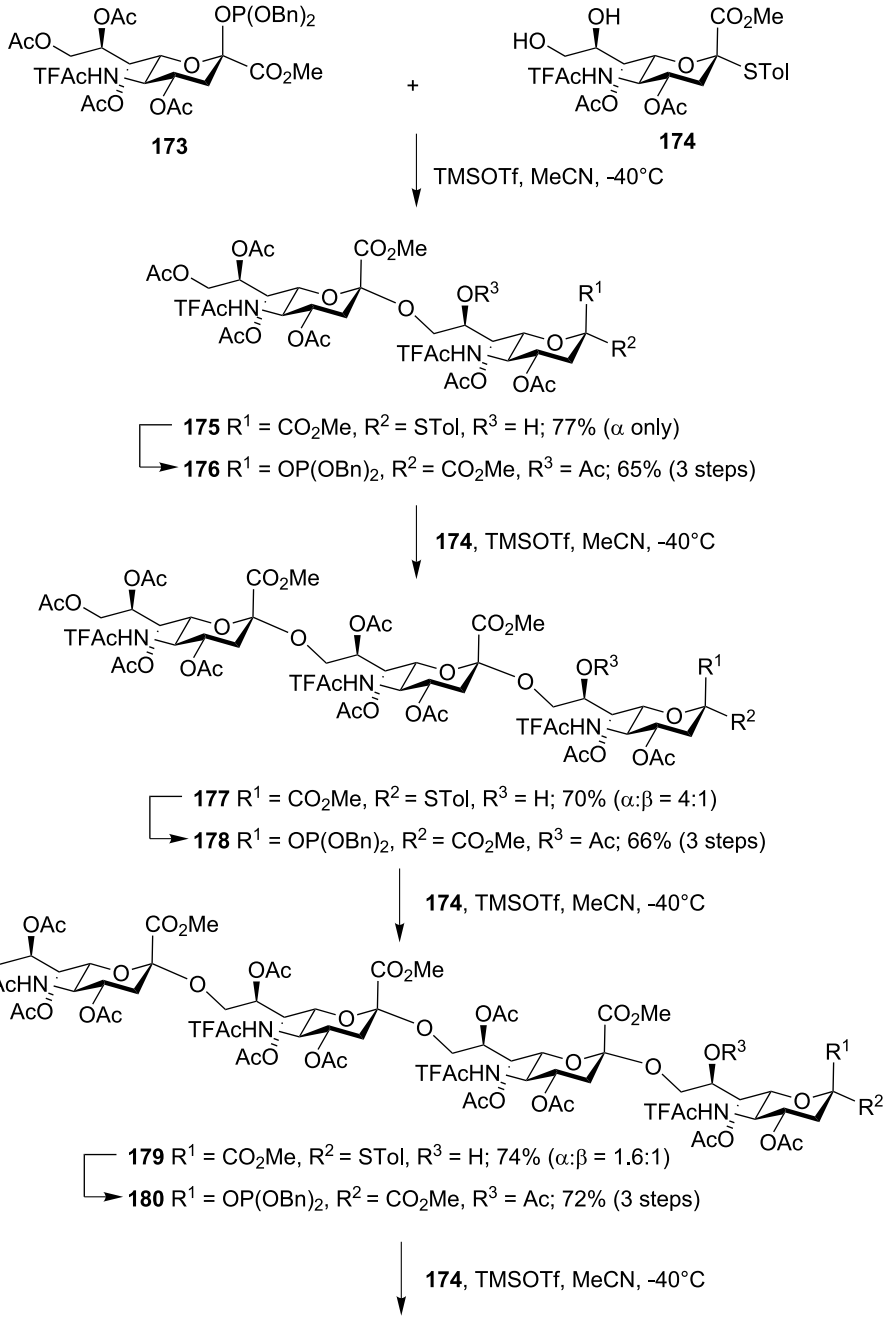
5.4 α-(2–4) and α-(2–5)-Linked Disialic Acid

Recently, various new gangliosides have been isolated from marine animals, some of which have a α-(2–4)-linked disialic acid residue within their glycan structures [120,121]. The construction of the α-(2–4)-linkage is not straightforward, due to the low reactivity of the C4 hydroxyl of Neu5Ac.

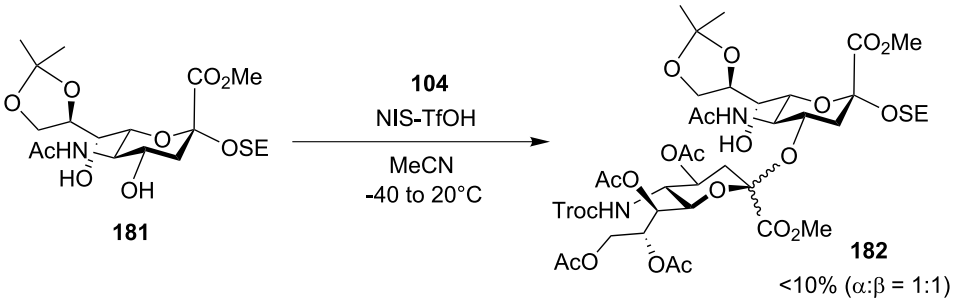
An initial attempt of sialylation of the C4 hydroxyl of the Neu5Ac derivative **181** with *N*-Troc sialyl donor **104** in our laboratory resulted in the production of a stereo mixture of disialic acid **182** in less than 10% yield (*Scheme 48*) [115].


166 (1.5 eq)
+
167
NIS-TfOH, CH2Cl2, -78°C
168 R = CAc 86% (α only)
Thiourea, 2,6-lutidine, 70°C
169 R = H 85%
166 (3.0 eq), NIS-TfOH, CH2Cl2, -78°C
170 R = CAc 89% (α only)
Thiourea, 2,6-lutidine, 70°C
171 R = H 82%
166 (3.0 eq), NIS-TfOH, CH2Cl2, -78°C
172
57% (α only)

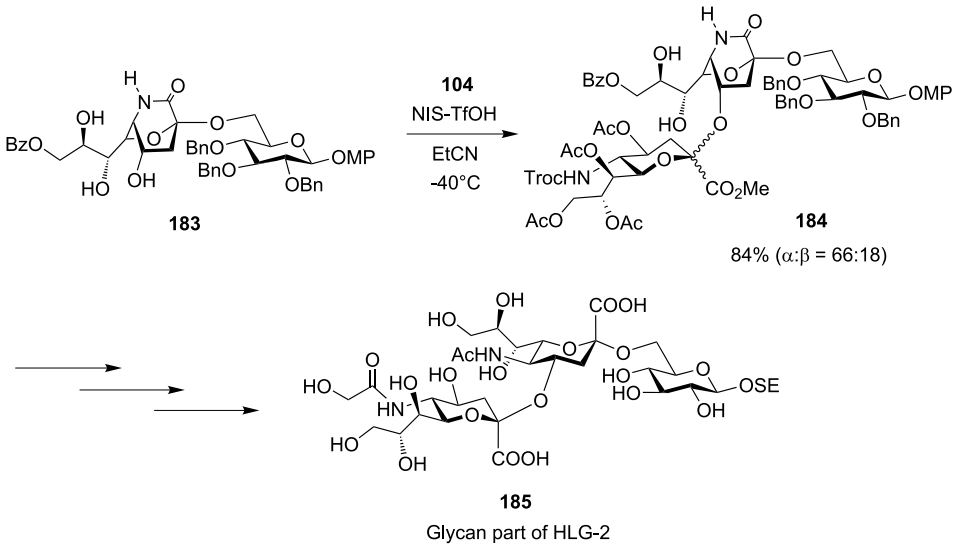
Scheme 46


OAc
OP(OBn)2
AcO
TFAcHN
AcO
OAc
CO2Me
173
+
OH
CO2Me
HO
TFAcHN
AcO
OAc
STol
174
TMSOTf, MeCN, -40°C
OAc
CO2Me
AcO
TFAcHN
AcO
OAc
OR3
R1
TFAcHN
AcO
OAc
R2
175 R1 = CO2Me, R2 = STol, R3 = H; 77% (α only)
176 R1 = OP(OBn)2, R2 = CO2Me, R3 = Ac; 65% (3 steps)
174, TMSOTf, MeCN, -40°C
177 R1 = CO2Me, R2 = STol, R3 = H; 70% (α:β = 4:1)
178 R1 = OP(OBn)2, R2 = CO2Me, R3 = Ac; 66% (3 steps)
174, TMSOTf, MeCN, -40°C
179 R1 = CO2Me, R2 = STol, R3 = H; 74% (α:β = 1.6:1)
180 R1 = OP(OBn)2, R2 = CO2Me, R3 = Ac; 72% (3 steps)
174, TMSOTf, MeCN, -40°C
Pentamer: 45% (α:β = 1:1.3)

Scheme 47



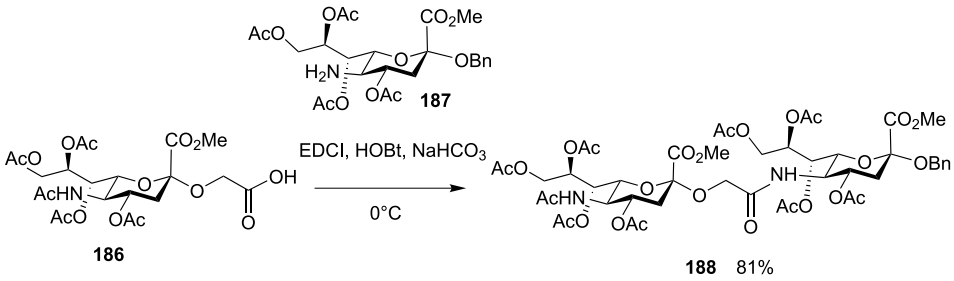
Scheme 48



Scheme 49

The breakthrough again came with the introduction of a 1,5-lactam bridge into the sialic acid acceptor (*Scheme 49*) [115]. The reaction of the 1,5-lactamized sialoside acceptor **183** with *N*-Troc donor **104** gave Neu-α-(2–4)-Neu-α-(2–6)-Glc trisaccharide **184** with improved efficiency (84%, $\alpha/\beta = 66/18$). The trisaccharide was further converted to Neu5Gc-α-(2–4)-Neu5Ac-α-(2–6)-Glc glycan frame **185** of the new ganglioside HLG-2. Taking account of this result, it can be hypothesized that the low reactivity of the C4 hydroxyl of Neu5Ac is probably due to steric hampering by the C5 acetamido group, similar to the case of the C8 hydroxyl.

Recently, *O*-glycolyl-linked oligo-sialic acid was also identified as a component of glycoproteins associated with the egg jelly coat of sea urchins [4]. Wu's group first reported the assembly of α-(2–5)-linked dimeric sialic acid by the condensation of sialyl glycolyl acid derivative **186** with 5-amino-sialic acid **187** in the presence of EDCI and HOBt (*Scheme 50*) [102].

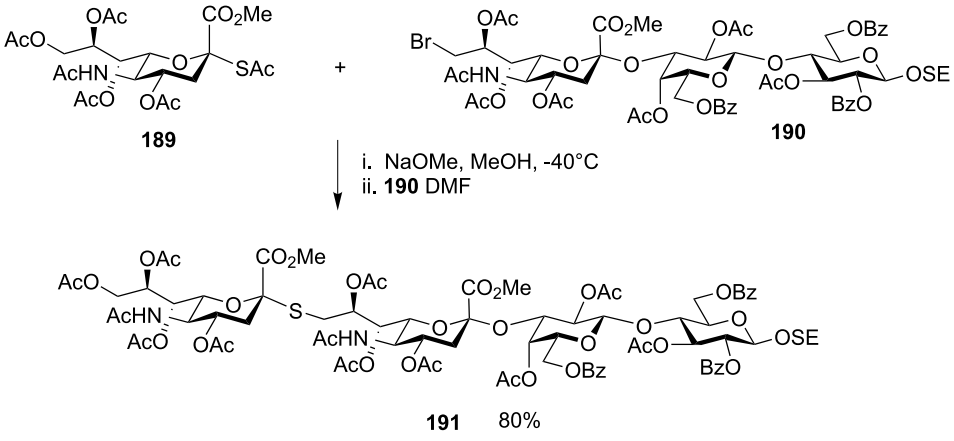


Scheme 50

Moreover, this approach was effective for the synthesis of a α-(2–5)-linked NeuGc oligomer as large as octasaccharide [122]. Other groups also reported the synthesis of the α-(2–5)-linked dimer by similar approaches [123,124].

6 Synthesis of Sialyl *S*- and *C*-Glycosides

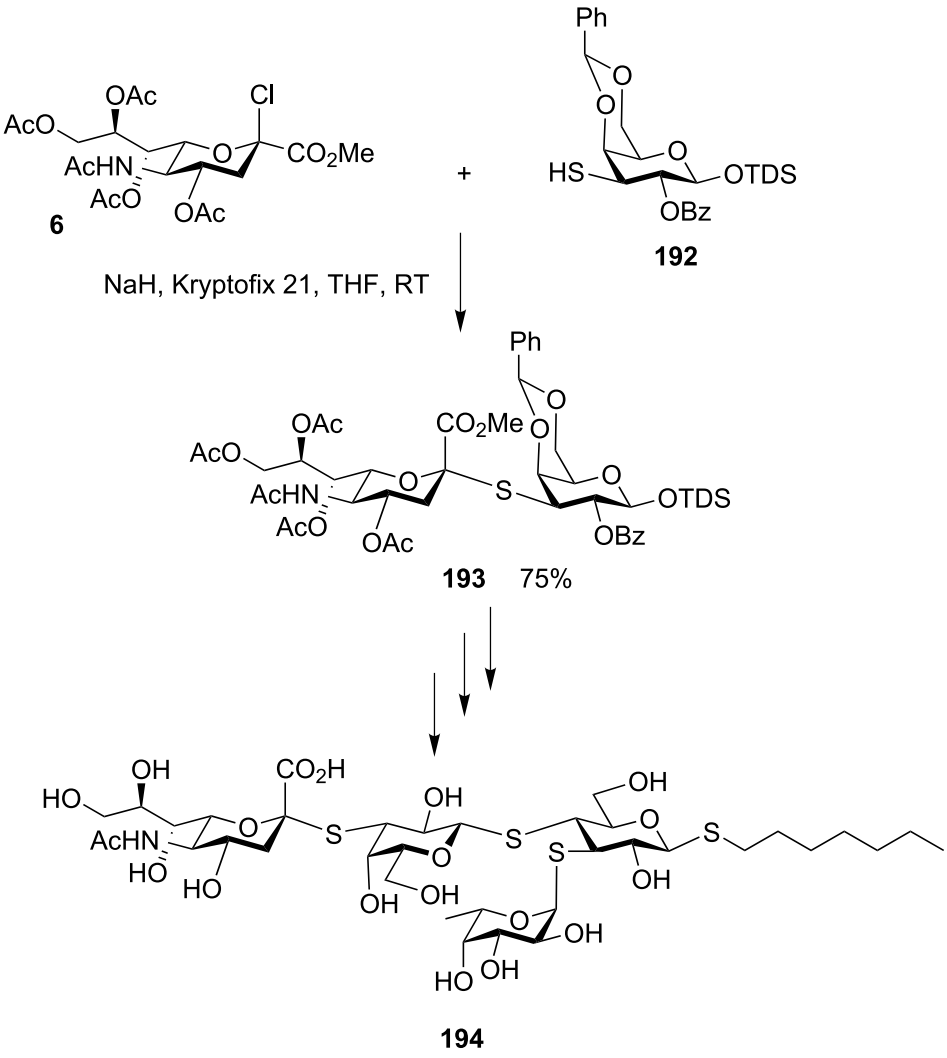
Replacement of the anomeric oxygen of the glycoside bond by sulfur, carbon, or nitrogen is an important technique of carbohydrate mimicking. Such replacement endows the carbohydrate with resistance to enzyme-mediated hydrolysis with the upkeep of the original biological property. *S*- and *C*-linked sialyl glycosides have been targeted for the purpose of the application to biological study, carbohydrate-based drug development, vaccination [125] and so on.



Scheme 51

6.1 *S*-Glycoside

Thioglycoside between sialic acid and other sugars can be synthesized easily by the S_N2 reaction of the glycosylthiolate anion with a halo or alkylsulfonyl sugar counterpart [126]. Mostly, the α-sialyl thiolate anion, which was generated by the treatment of α-thioacetate of Neu5Ac with NaOMe, was subsequently used for the reaction with a secondary halogeno- or alkylsulfonyl-carbon to furnish thioglycoside with complete retention of α-configuration (*Scheme 51*) [127].



Scheme 52

The S_N2 replacement of the 2-β-chloro derivative of Neu5Ac with a thiol group incorporated into the sugar derivative also successfully produces sialyl thioglycoside. In recent examples, the latter approach was incorporated into sialyl thiooligosaccharide synthesis (Scheme 52) [128].

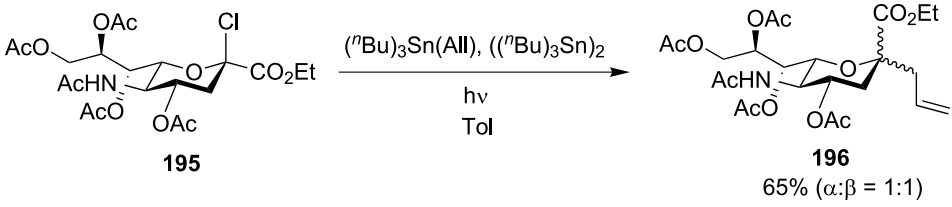
6.2 *C*-Glycosides

Since sialyl *C*-glycoside formation results in a quaternary carbon atom, it is synthetically a difficult subject. In addition, as mentioned earlier, due to the electron-withdrawing carboxylate group, the generation of the carbenium ion is also problematic.

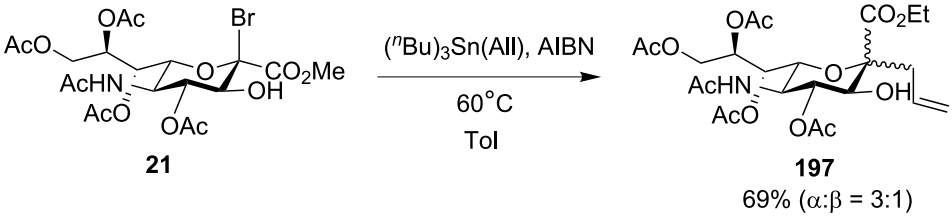
In 1991, the syntheses of sialyl *C*-glycosides were introduced by Bednarski et al., Paulsen et al., and Vasella et al. Bednarski and Paulsen exploited a 2-halogeno derivative of Neu5Ac as a precursor of the anomeric radical, which was anticipated to be stabilized by the "capto-dative effect" of the C1 carboxylate group and ring oxygen. Bednarski reacted the 2-chloro derivative **195** with allyltributyltin and a catalytic amount of bis(tributyltin), and photolized for 18 h to obtain *C*-glycoside **196** as a 1/1 of anomeric mixture in 65% yield (Scheme 53) [129].

To improve the diastereoselectivity of *C*-glycosidation, Paulsen's group exploited 2-bromo-3-hydroxy derivative **21**. As a result, the reaction of **21** with allyltributyltin and AIBN at 60 °C provided **197** (69%), favoring α-glycoside in the ratio of 3/1 (Scheme 54) [130].

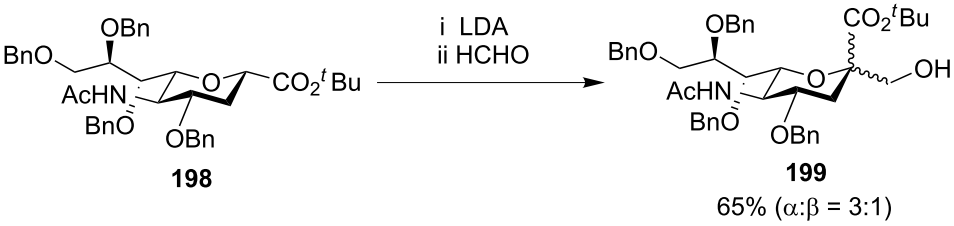
Unlike the radical approaches mentioned above, Vasella exploited a carbanion approach. Thus, the anomeric carbanion was anticipated to be stabilized by the C1-carboxylate group. A 1-deoxy derivative of Neu5Ac **198** was treated with LDA, and the resulting anion was reacted with formaldehyde to produce a 3/1 mixture of α- and β-isomers of 2-hydroxymethyl product **199** (Scheme 55) [131].



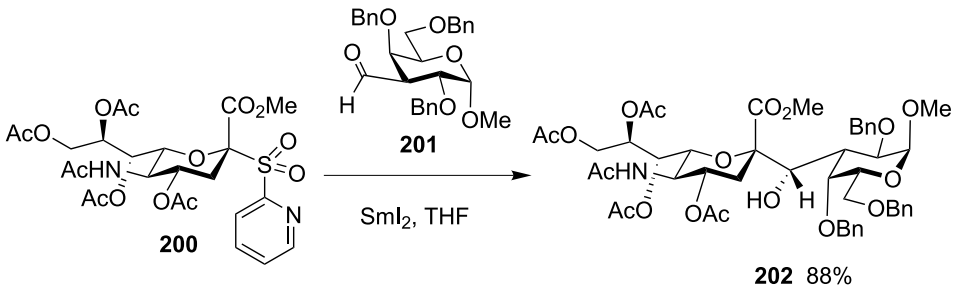
Scheme 53



Scheme 54



Scheme 55



Scheme 56

Also exploiting the anomeric anion approach, Lindhart and co-workers have generated sialyl samarium (III) in situ as an anomeric anion equivalent, which was reacted with carbonyl compounds under Barbier conditions. Thus, a 2-α-pyridyl sulfone derivative of Neu5Ac **200** was reacted with *C*-formyl sugar **201** in the presence of SmI_2 in THF, producing *C*-disaccharide **202** in 88% yield with complete diastereoselectivity (*Scheme 56*) [132]. This approach was successfully applied to the synthesis of *C*-glycoside mimics of biologically relevant sialyl glycans, such as GM3 [133] and sialyl Tn antigen [134]. Later, the 2-sulfone derivative was replaced with 2-pyridylsulfide and 2-acetate derivatives by Beau's group [135,136]. Both derivatives were also capable of serving as precursors for an anomeric anion in samarium-mediated Reformatsky-type coupling reactions, producing *C*-sialyl glycosides. Both approaches developed by Linhart and Beau lead to hydroxymethylene-bridged Neu5Ac-monosaccharide sequences. The hydroxyl group has to be cleaved to complete the replacement of interglycosidic oxygen with carbon.

Schmidt's group has disclosed a different strategy which leads to methylene-bridged *C*-sialoside [137]. They designed a strategy based on the electrophilic cyclization of an open chain precursor of Neu5Ac using phenylselenyl triflate as a cyclization reagent, followed by the generation of the carboxylate moiety and introduction of the nitrogen. This strategy was successfully applied to the synthesis of methylene-bridged Neu5Ac-α-(2–3)-Gal *C*-disaccharide.

7 Closing Remarks

As discussed in this chapter, a large variety of sialic acid donors have been developed, addressing anomeric selectivity during glycosylation. Because of the nucleophilic participation of the auxiliary mounted at C3 with anomeric carbon, the C3-appended derivatives of Neu5Ac have realized nearly complete α-selectivity during the glycosylation of a wide spectrum of glycosyl acceptors. On the other hand, C1-auxiliaries were not generally effective in directing α-selective coupling reactions. The combined use of the nitrile solvent effect and 2-methylsulfenyl, phenylsulfenyl, and xanthate derivatives of Neu5Ac offered facile α-sialylation, the efficacy and practicality of which have been exemplified by a number of syntheses of sialy oligosaccharides. The discovery of a C5 functionality effect on the reactivity of the sialyl donor and sialyl acceptor has brought about the innovation of sialylation chemistry. Furthermore, it has led to the tour de force of the first synthesis of α-(2–8)-linked tetrasialic acid. However, the chemistry underlying the remote functionality effect is yet to be understood.

References

1. Schauer R (1982) Sialic acids: chemistry, metabolism and function. Cell biology monographs, vol 10. Springer, Berlin Heidelberg New York
2. Varki A (1999) In: Varki A, Cummings R, Esko J, Freeze H, Hart G, Marth J (eds) Essentials of Glycobiology. Cold Spring Harbor Laboratory Press, New York, p 195
3. Angata T, Varki A (2002) Chem Rev 102:439
4. Kitazume S, Kitajima K, Inoue S, Troy FA II, Cho JW, Lennarz WJ, Inoue Y (1994) J Biol Chem 269:22712
5. Paulsen H (1990) Angew Chem Int Ed Engl 29:823
6. Ito Y, Ohnishi Y (2001) In: Fraser-Reid BO, Tatsuta K, Thiem J (eds) Glycoscience I–III. Springer, Berlin Heidelberg New York, p 1589
7. Meindl P, Tuppy H (1965) Monatsh Chem 96:802
8. Khorlin AY, Privalova IM, Brystrova IB (1971) Carbohydr Res 19:272
9. Okamoto K, Goto T (1990) Tetrahedron 46:89
10. DeNinno MP (1991) Synthesis: 583
11. Hasegawa A, Kiso M (1997) In: Hanessian S (ed) Preparative Carbohydrate Chemisty. Marcel Dekker, New York, p 357
12. Boons GJ, Demchenko AV (2000) Chem Rev 100:4539
13. Paulsen H, Tietz H (1984) Carbohydr Res 125:47
14. Kuhn R, Lutz P, MacDonald DL (1966) Chem Ber 99:611
15. Furuhata K, Komiyama K, Ogura H, Hata T (1991) Chem Pharm Bull 39:255
16. Bryamova NE, Tuzikov AB, Bovin NV (1992) Carbohydr Res 237:161
17. Ogura H, Furuhata K (1986) Carbohydr Res 158:37
18. Eschenfelder V, Brossmer R (1980) Carbohydr Res 78:190
19. van der Vleugel DJM, van Heeswijk WAR, Vliegenthart JFG (1982) Carbohydr Res 102:121
20. Thomas RL, Sarkar AK, Kohata K, Abbas SA, Matta KL (1990) Tetrahedron Lett 32:2825
21. Dekany G, Wright K, Ward P, Toth I (1996) J Carbohydr Chem 15:383
22. Sato S, Furuhata K, Itoh M, Shitori Y, Ogura H (1988) Chem Pharm Bull 36:914
23. Lubineau A, Auge C, Bouxom B, Gautheron C (1992) J Carbohydr Chem 11:59
24. Kartha KPR, Aloui M, Field RA (1996) Tetrahedron Lett 37:8807
25. Ogawa T, Sugimoto M (1985) Carbohydr Res 135:c5
26. Danishefsky SJ, Gervay J, Peterson JM, McDonald FE, Koseki K, Oriyama T, Griffith DA (1992) J Am Chem Soc 114:8329
27. Sharma MN, Eby R (1984) Carbohydr Res 127:201
28. Kunz H, Waldmann H, Klinkhammer U (1988) Helv Chim Acta 71:1868
29. Okamoto K, Kondo T, Goto T (1986) Tetrahedron Lett 27:5233

30. Okamoto K, Kondo T, Goto T (1987) Tetrahedron 43:5919
31. Castro-Palomino JC, Tsvetkov YE, Schmidt RR (1998) J Am Chem Soc 120:5434
32. Ito Y, Ogawa T (1987) Tetrahedron Lett 28:6221
33. Ito Y, Ogawa T (1988) Tetrahedron Lett 29:3987
34. Takahashi T, Tsukamoto H, Yamada H (1997) Tetrahedron Lett 38:8223
35. Haberman JM, Gin DY (2001) Org Lett 3:1665
36. Haberman JM, Gin DY (2003) Org Lett 5:2539
37. Ishiwata A, Ito Y (2003) Synlett: 1339
38. Kanie O, Kiso M, Hasegawa A (1988) J Carbohydr Chem 7:501
39. Hasegawa A, Ohki H, Nagahama T, Ishida H, Kiso M (1991) Carbohydr Res 212:277
40. Marra A, Sinaÿ P (1989) Carbohydr Res 187:35
41. Kirchner E, Thiem F, Dernick R, Heukeshoven J, Thiem J (1988) J Carbohydr Chem 7:453
42. Cao S, Meunier SJ, Andersson FO, Letellier M, Roy R (1994) Tetrahedron: Asymm 5:2303
43. Takeda K, Tsuboyama K, Torii K, Furuhata K, Sato N, Ogura H (1990) Carbohydr Res 203:57
44. De Meo C, Parker O (2005) Tetrahedron: Asymm 16:303
45. Ikeda K, Aizawa M, Sato K, Sato M (2006) Bioorg Med Chem Lett 16:2618
46. Matsuoka K, Onaga T, Mori T, Sakamoto JI, Koyama T, Sakairi N, Hatano K, Terunuma D (2004) Tetrahedron Lett 45:9383
47. Schmidt RR, Rücker E (1980) Tetrahedron Lett 21:1421
48. Hashimoto S, Hayashi M, Noyori R (1984) Tetrahedron Lett 25:1379
49. Andersson F, Fügedi P, Garegg PJ, Nashed M (1986) Tetrahedron Lett 27:3919
50. Ito Y, Ogawa Y (1987) Tetrahedron Lett 28:4701
51. Pavia AA, Ung-Chhun SN, Durand JL (1981) J Org Chem 46:3158
52. Lemieux RU, Ratcliffe RM (1979) Can J Chem 57:1244
53. Sinaÿ P, Pougny JR (1976) Tetrahedron Lett 4073
54. Schmidt RR, Michel J (1985) J Carbohydr Chem 4:141
55. Lemieux RU, Morgan AR (1965) Can J Chem 43:2205
56. Ratcliffe AJ, Fraser-Reid B (1990) J Chem Soc Perkin Trans 1:747
57. Perrin CL, Armstrong KB (1993) J Am Chem Soc 115:6825
58. Fabian MA, Perrin CL, Sinnott ML (1994) J Am Chem Soc 116:8398
59. Hasegawa A, Nagahama T, Ohki H, Hotta K, Ishida H, Kiso M (1991) J Carbohydr Chem 10:493
60. Veeneman GH, van Leeuwen SH, van Boom JH (1990) Tetrahedron Lett 31:1331
61. Konradsson P, Udodong UE, Frase-Reid B (1990) Tetrahedron Lett 31:4313
62. Ito Y, Ogawa T (1988) Tetrahedron Lett 29:1061
63. Fukase K, Hasuoka A, Kusumoto S (1993) Tetrahedron Lett 34:2187
64. Stahl W, Sprengard U, Kretzscmar G, Kunz H (1994) Angew Chem Int Ed Engl 33:2096
65. Murase T, Ishida H, Kiso M, Hasegawa A (1988) Carbohydr Res 184:c1
66. Murase T, Kameyama A, Kartha KPR, Ishida H, Kiso M, Hasegawa A (1989) J Carbohydr Chem 8:265
67. Murase T, Ishida H, Kiso M, Hasegawa A (1988) Carbohydr Res 188:71
68. Hasegawa A, Ogawa H, Ishida H, Kiso M (1990) J Carbohydr Chem 9:393
69. Hasegawa A, Adachi K, Yoshida M, Kiso M (1992) J Carbohydr Chem 11:95
70. Kiso M, Ando K, Furi Y, Ishida H, Hasegawa A (1993) J Carbohydr Chem 12:985
71. Prabhanjan H, Aoyama K, Kiso M, Hasegawa A (1992) Carbohydr Res 233:87
72. Ando T, Ishida H, Kiso M (2001) J Carbohydr Chem 20:425
73. Ito H, Ishida H, Kiso M (1998) Carbohydr Res 306:581
74. Hasegawa A, Ishida H, Kiso M (1993) J Carbohydr Chem 12:371
75. Ishida H, Ohta Y, Tsukada Y, Kiso M, Hasegawa A (1993) Carbohydr Res 246:75
76. Hotta K, Kawase T, Ishida H, Kiso M, Hasegawa A (1995) J Carbohydr Chem 14:961
77. Ando H, Ishida H, Kiso M (1999) J Carbohydr Chem 18:603
78. Ishida HK, Ishida H, Kiso M, Hasegawa A (1994) Tetrahedron: Asymm 5:2493
79. Ando H, Ishida H, Kiso M (1997) Carbohydr Res 300:207
80. Bilberg W, Lonn H (1991) Tetrahedron Lett 32:7453
81. Liebe B, Kunz H (1997) Angew Chem Int Ed Engl 36:618
82. Wilstermann M, Kononov LO, Nilsson U, Ray AS, Magnusson G (1995) J Am Chem Soc 117:4742
83. van Seeventer PB, Kerékgyártó J, van Dorst JALM, Halkes KM, Kamerling JP, Vliegenthart JFG (1997) Carbohydr Res 300:127

84. Ellervik U, Magnusson G (1998) J Org Chem 63:9314
85. Ikeda K, Sugiyama Y, Tanaka K, Sato M (2002) Bioorg Med Chem Lett 12:2309
86. Martin TJ, Schmidt RR (1992) Tetrahedron Lett 33:6123
87. Kondo H, Ichikawa Y, Wong CH (1992) J Am Chem Soc 114:8748
88. Lassaletta JM, Schmidt RR (1995) Tetrahedron Lett 36:4209
89. Stauch T, Greilich U, Schmidt RR (1995) Liebigs Ann Chem: 2101
90. Sakamoto H, Nakamura S, Tsuda T, Hashimoto S (2000) Tetrahedron Lett 41:7691
91. Schmidt RR (1986) Angew Chem Int Ed Engl 25:212
92. Cai S, Yu B (2003) Org Lett 5:3827
93. Ye XS, Huang X, Wong CH (2001) Chem Commun: 974
94. Demchenko AV, Boons GJ (1998) Tetrahedron Lett 39:3065
95. Demchenko AV, Boons GJ (1999) Chem Eur J 5:1278
96. Crich D, Li W (2006) Org Lett 8:959
97. Komba S, Galustian C, Ishida H, Feizi T, Kannagi R, Kiso M (1999) Angew Chem Int Ed 38:1131
98. De Meo C, Demchenko AV, Boons GJ (2001) J Org Chem 66:5490
99. Yu CS, Niikura K, Lin CC, Wong CH (2001) Angew Chem Int Ed 40:2900
100. Mukaiyama T, Mandai H, Jona H (2002) Chem Lett: 1182
101. Lu KC, Tseng SY, Lin CC (2002) Carbohydr Res 337:755
102. Ren CT, Chen CS, Wu SH (2002) J Org Chem 67:1376
103. Ando H, Koike Y, Ishida H, Kiso M (2003) Tetrahedron Lett 44:6883
104. Fuse T, Ando H, Imamura A, Sawada N, Ishida H, Kiso M, Ando T, Li SC, Li YT (2006) Glycoconj J 23:329
105. Adachi M, Tanaka H, Takahashi T (2004) Synlett: 603
106. Tanaka H, Adachi M, Takahashi T (2005) Chem Eur J 11:849
107. Tanaka K, Goi T, Fukase K (2005) Synlett: 2958
108. Tsvetkov Y, Schmidt RR (1994) Tetrahedron Lett 35:8583
109. Castro-Palomino JC, Tsvetkov YE, Schneider R, Schmidt RR (1997) Tetrahedron Lett 38:6837
110. Okamoto K, Kondo T, Goto T (1986) Tetrahedron Lett 27:5229
111. Ito Y, Numata M, Sugimoto M, Ogawa (1989) J Am Chem Soc 111:8508
112. Hossain N, Magnusson G (1999) Tetrahedron Lett 40:2217
113. Castro-Palomino JC, Simon B, Speer O, Leist M, Schmidt RR (2001) Chem Eur J 7:2178
114. Tanaka H, Nishiura Y, Adachi M, Takahashi T (2006) Heterocycles 67:107
115. Ando H, Koike Y, Koizumi S, Ishida H, Kiso M (2005) Angew Chem Int Ed 44:6759
116. Tanaka H, Nishiura Y, Takahashi T (2006) J Am Chem Soc 128:7124
117. Ogawa T, Sugimoto M (1984) Carbohydr Res 128:c1
118. Hasegawa A, Ogawa M, Ishida H, Kiso M (1990) J Carbohydr Chem 9:393
119. Lin CC, Huang KT, Lin CC (2005) Org Lett 7:4169
120. Yamada K, Harada Y, Miyamoto T, Isobe R, Higuchi R (2000) Chem Pharm Bull 48:157
121. Yamada K, Matsubara R, Kaneko M, Miyamoto T, Higuchi R (2001) Chem Pharm Bull 49:447
122. Ren CT, Chen CS, Yu YP, Tsai YF, Lin PY, Chen YJ, Zou W, Wu SH (2003) Chem Eur J 9:1085
123. Fan GT, Lee CC, Lin CC, Fang JM (2002) J Org Chem 67:7565
124. McAuliffe JC, Rabuka D, Hindsgaul O (2002) Org Lett 4:3067
125. Bundle DR, Rich JR, Jacques S, Yu HN, Nitz M, Ling CC (2005) Angew Chem In Ed 44:7725
126. Pachamuthu K, Schmidt RR (2006) Chem Rev 106:160
127. Ishida H, Kiso M, Hasegawa A (1994) Methods Enzymol 242:183
128. Eisele T, Toepfer A, Kretzschmar G, Schmidt RR (1996) Tetrahedron Lett 37:1389
129. Nagy JO, Bednarski MD (1991) Tetrahedron Lett 32:3953
130. Paulsen H, Matschulat P (1991) Liebigs Ann Chem: 487
131. Wallimann K, Vasella A (1991) Helv Chim Acta 74:1520
132. Vlahov IR, Vlahova PI, Linhardt R (1997) J Am Chem Soc 119:1480
133. Du Y, Linhardt RJ (1998) Carbohydr Res 308:161
134. Kuberan B, Sikkander SA, Tomiyama H, Linhardt RJ (2003) Angew Chem Int Ed 42:2073
135. Abdallah Z, Doisneau G, Beau JM (2003) Angew Chem Int Ed 42:5209
136. Malapelle A, Abdallah Z, Doisneau G, Beau JM (2006) Angew Chem Int Ed 45:6016
137. Notz W, Hartel C, Waldscheck B, Schmidt RR (2001) J Org Chem 66:4250

5.8 Enzymatic Glycosylation by Transferases

Ola Blixt[*1], *Nahid Razi*[2]
[1] Department of Cellular and Molecular Medicine, Faculty of Health Sciences, University of Copenhagen, 2200 Copenhagen N, Denmark
[2] Department of Molecular Biology, Glycan Array Synthesis Core D, Consortium for Functional Glycomics, The Scripps Research Institute, La Jolla, CA 92037, USA
olablixt@imbg.ku.dk, nrazi@scripps.edu

Abstract

Glycosyltransferases are important biological catalysts in cellular systems generating complex cell surface glycans involved in adhesion and signaling processes. Recent advances in glycoscience have increased the demands to access significant amount of glycans representing the glycome. Glycosyltransferases are now playing a key role for in vitro synthesis of oligosaccharides and the bacterial genome are increasingly utilized for cloning and over expression of active transferases in glycosylation reactions. This chapter highlights the recent progress towards preparative synthesis of oligosaccharides representing terminal sequences of glycoproteins and glycolipids using recombinant transferases. Transferases are also being explored

In: *Glycoscience*. Fraser-Reid B, Tatsuta K, Thiem J (eds)
Chapter-DOI 10-1007/978-3-540-30429-6_32:

in the context of solid-phase synthesis, immobilized on resins and over expression in vivo by engineered bacteria.

Keywords

Glycosyltransferases; Chemoenzymatic synthesis; Oligosaccharide; Glycoconjugate; Carbohydrate synthesis

Abbreviations

α 1,4-GlcNAc	α1,4-*N*-acetylglucosamine
CFG	Consortium for Functional Glycomics
CHO	Chinese hamster ovary
FUTs	fucosyltransferases
GT	glycosyltransferase
LN	*N*-acetyllactosamine
PAPS	phosphoadenosine phosphosulfate
pLN	poly-*N*-acetyllactosamine
STs	sialyltransferases

1 Introduction

Carbohydrates are diverse biological structures that occur naturally in conjugation with proteins (glycoproteins) or lipids (glycolipids) and are involved in different cellular functions. The contribution of glycans in cellular recognition, adhesion, motility, and signaling processes, have been well documented [1,2]. Therefore, abnormalities in biosynthesis or presentation of carbohydrates can effectively lead to functional irregularity and diseased states. However, in almost every instance this pursuit is limited by the unavailability of carbohydrate compounds for in vitro studies, due to the enormous complexity of carbohydrate synthesis.

Extensive efforts have been made to overcome the challenging tasks of carbohydrate synthesis. In a glycosylation reaction, both donors (nucleotide-activated monosaccharides) and acceptors (which receive the activated monosaccharide) contain many similar functional groups that must be differentiated and protected to accomplish a desirable glycosidic bond. The product must then be selectively deprotected for the next round of reactions. The complexity of protecting-group manipulation increases with every additional glycosidic linkage (*Fig. 1*). Furthermore, the two possible isomers concerning the geometry of the anomeric carbon atom of each monosaccharide, i. e. alpha and beta isomers, introduces another level of complication, namely "stereoselectivity," to the field of carbohydrate chemistry. Multiple steps of protection-deprotection, as well as, a careful design of reactions, are required to construct a regio- and stereoselective glycosidic bond in a chemical approach. Thus, traditional organic chemistry with solution phase synthesis is not well-suited with difficulties involved in carbohydrate synthesis [3,4]. Because of this problem, there is currently no single step-wise synthetic approach that is applicable to the synthesis of all oligosaccharides or even, just the 15 million possible tetrasaccharides that can be assembled from the nine common monosaccharides found in humans. Nonetheless, in the past few decades, the work of many research groups has started to open up new paths to oligosaccharide and glycoconjugate synthesis. Coupling techniques with

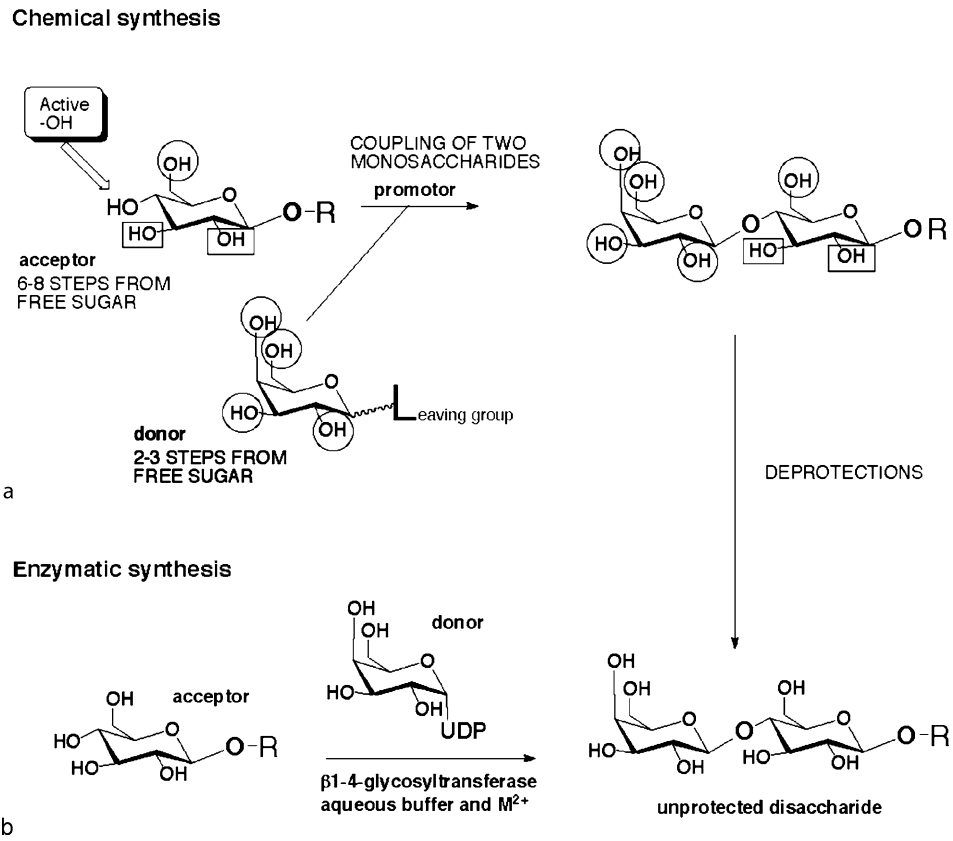


Figure 1
Generalization of chemical and enzymatic synthesis of a carbohydrate disaccharide

better yields and stereoselectivity have been worked out, and new protecting group chemistries have also become available. Alternatives to traditional chemical synthesis, such as solid-phase synthesis and one-pot reactivity based glycosylations [5,6,7] and synthetic automations [5] have been developed to prevail over the synthetic constraints. These elegant methods dramatically shorten and simplify chemical synthesis and, although with differing degrees of difficulty, make it possible to synthesize essentially any glycosidic linkage. Despite such improvements, the construction of a defined glycosidic bond is still one of the most time consuming and challenging tasks facing synthetic chemists, organic synthesis is not realistic for the large-scale pharmaceutical preparations [8].

A practical alternative to traditional, non-biological, organic synthesis is the utility of enzymes as catalysts for oligosaccharide synthesis in chemoenzymatic methods. Through the use of enzymes, the impractical, or sometimes impossible, synthetic manipulations of complex carbohydrates can be performed in an environmentally benign manner. The enzymes can be used as the sole catalyst in a reaction, in combination with other enzymes, or with non-biological reagents. The chiral nature of enzymes results in the formation of defined stereo- and regio

specific products with remarkable rate acceleration of the reaction (typically 10^5 to 10^8). Most enzymes operate at room temperature, under neutral aqueous conditions, and in the absence of substrate functional-group protection. The application of enzymes in carbohydrate synthesis offers promising opportunities for the development of industrial, chemical, and pharmaceutical processes [9,10].

In this chapter, we discuss the recent developments in glycosyltransferase (GT) preparations and their use in chemoenzymatic synthesis of oligosaccharides. Special attention will be given to the practical preparative strategy of the multi-step synthesis of complex carbohydrates using a "tool box" of functional recombinant enzymes developed in different expression systems. A library of glycans with representative structures found on glycoproteins and glycolipids have been generated in preparative scale and enabled distribution of these useful compounds to multiple users in the community to be used towards advancing the field of glycomics (http://www.functionalglycomics.com).

2 Oligosaccharide Synthesis with Enzymes

The coupling of monosaccharides to form oligosaccharides via specific, glycosidic linkages can be catalyzed by glycosidases or GTs [10].

2.1 Glycosidases

Glycosidases, enzymes that in vivo function to cleave oligosaccharides and polysaccharides via glycosyl transfer to water, can form glycosidic linkages under conditions in which a carbohydrate hydroxyl moiety acts as a more efficient nucleophile than water. Glycosidases were used well ahead of GTs for the preparative synthesis of sugars and are, most frequently, used to synthesize oligosaccarides via "transglycosylation" [11] and condensation [12]. Glycosidases are stable, easy to isolate and generally, more available than GTs. Glycosidases, however, have suffered from two major drawbacks: poor regio-selectivities and low chemical yields for the construction of glycosidic bonds. The low production yield is due to the fact that the glycosidases promote not only glycosylation, but also the reverse reaction, namely, hydrolysis of the resulting product. Efforts have been made to circumvent the limitations of glycosidases by mutagenizing methods that abolish glycosyl hydrolase activity and create novel catalysts, such as "glycosylsynthases" [13] (see also Chap. 5.9).

2.2 Glycosyltransferases

GTs, on the other hand, offer the significant advantages of increased yield and specificity relative to glycosidase-based syntheses. They do so, however, at the expense of limited substrate specificity, relatively high cost of the glycosyl nucleotide donor, and limited enzyme availability. GTs are responsible for the biosynthesis of oligosaccharides and are divided into two classes according to the type of glycosyl donor they accept. One group follows the Leloir pathway, which is the utilization of activated glycosyl esters of nucleoside mono- or diphosphates as glycosyl donors. This group was named after the Argentinean biochemist that discovered and investigated its properties and functions. The other group, the non-Leloir transferases typically

uses glycosyl phosphates as glycosyl donors (i. e. monosaccharides). Non-Leloir transferases have been used preferentially in synthesizing nucleosides and nucleoside analogs, but also in a variety of oligosaccharides and polysaccharides [14]. Leloir-type GTs, generally known as GTs, are responsible for the synthesis of most *N*- and *O*-linked glycoproteins and glycoconjugates in mammals and display higher stereo- and regioselectivities than both the glycosidase and their glycosyl phosphate-transferring relatives. The activated nucleotide-monosaccharides (glycosyl donors) are the substrates for the GTs in mammals. Each glycosyl donor is accepted by various GTs mediating the transfer of the glycosyl group onto different acceptors. In vivo, these enzymes are generally specific for the given glycosyl donor, the type of linkage to be formed, the structure of the acceptor and a specific hydroxyl group of the acceptor. Such observed specificity is required to permit the synthesis of information-containing biomolecules without a template. Nevertheless, GTs display some flexibility with respect to their substrate requirements under in vitro conditions, i. e. when the substrates are provided in high concentration.

Utilizing GTs in the chemo-enzymatic synthesis of oligosaccharides has in the past, been limited by different restriction factors such as: availability of the enzymes, availability of the activated monosaccharide building blocks, inhibition of glycosylation pathway by the released nucleotide phosphate and the enzyme substrate specificity. To overcome the inhibition effects of by-products and consumption of expensive sugar-nucleotides, elegant sugar-nucleotide regeneration systems have been explored but have shown limited use due to it's rather complex compositions [15].

GTs have been utilized in different contexts to synthesize oligosaccharide moieties on natural glycoconjugates, glycoconjugate mimetics and a variety of un-natural aglycons and analogs. The use of isolated enzymes or engineered whole cells allows the synthesis of oligosaccharides in both, in vitro and in vivo methods. GTs have been mostly utilized as catalysts in aqueous solution reactions ranging from milligram to kilogram scale reactions [16,17]. In order to take advantage of these biocatalysts for preparative synthesis of oligosaccharides two prerequisites have to be fulfilled: the recombinant GTs and the sugar nucleotides have to be available in preparative scales.

Advances in molecular biology have unquestionably diminished most of the restrictions in enzymatic synthesis. Flourishing cloning and isolating GTs from different sources, particularly bacteria [18,19,20], constructing mutated and fusion enzymes [21] with more favorable properties for synthesis than their parent proteins, as well as, bacterial engineering [22], have continuously broadened the GTs utilization in carbohydrate synthesis.

2.2.1 Recombinant Glycosyltransferase Expression

Over 200 GTs appear to exist in mammalian tissues [23]. Extraction and purification of GTs from tissues are neither practical, nor responsive to the demands of large amounts of the enzymes required for enzyme characterization or oligosaccaride synthesis. Since the first GT gene, β4-galactosyltransferase (β4Gal-T), which was cloned two decades ago [24,25], many GT genes from different natural sources have been cloned and most of these enzymes are shown to be suitable for oligosaccharide synthesis [23,26]. In addition, these enzymes have been used for non-natural substrates to synthesize versatile oligosaccharide analogs on a preparative scale for drug discovery in glycobiology research [27,28]. A wide variety

of expression systems, such as bacterial, yeast, insect and mammalian—each with its own advantages and disadvantages—have been used to produce sufficient amounts of GTs for enzymatic characterization and oligosaccharide synthesis. However, large-scale production of eukaryotic enzymes for industrial purposes is still a challenge facing the researchers in the field. Bacterial expression has the advantage of being simple to manipulate, but the disadvantage of lacking the eukaryotic post-translational machinery. This often results in the desired protein being locked in inclusion bodies. Nevertheless, some mammalian GTs have been successfully expressed in *Euscheria coli* and refolded to generate active protein [29,30,31,32]. Insect systems, such as sf-9 and High five, and mammalian systems, such as Chinese hamster ovary (CHO) and monkey and (COS) celllines have been most widely used for expression of GTs; the examples are too numerous to count. However, background endogenous activities, as well as the more complex nature of the laboratory tasks, hamper the large-scale production of mammalian enzymes in these systems. Fungi and yeasts have the advantage of being simpler to cultivate than mammalian cells, while still being eukaryotic. Large-scale expression of GTs was successfully reported in *Aspergillus niger* and yeast species such as the budding yeast *Saccharomyces cerevisiae* and the methylotrophic yeast *Pichia pastoris* [33,34,35]. GT expressions in yeast have been very successful. Recent genetic manipulations of yeast have led to constructing transformed yeast cells that produce GTs that either remain anchored to the cell surface so that the intact yeast cells are being used as an enzyme source [36,37], or are secreted into the culture. The whole yeast cells, used as an enzyme source, can easily be separated from the reaction products and can be used repetitively which makes the technique favorable and attractive.

A large collaborative effort was launched within the Consortium for Functional Glycomics (CFG) to centralize many of the cloned and expressed GTs and over expressing them as active enzymes. These enzymes are expressed as referenced in *Table 1* and continuously being explored for preparative synthesis of oligosaccharides and glycoconjugates as described below.

2.2.2 Sugar Nucleotide Generation Systems

Although several of the recombinant enzymes are readily available for synthesis, their corresponding sugar-nucleotides are not accessible as reagents and are very expensive. These drawbacks were recognized early by Wong and co-workers who introduced the in-situ full cycle regeneration system for sugar-nucleotides [38,39] (*Fig. 2*). Recently, Wang and co-workers extended the concept of sugar-nucleotide generation using the bacterial biosynthetic pathway [40]. Despite these improvements, the full-cycle regeneration approach has not been widely applicable due to its rather complex layout. Therefore, synthesis of milligram amounts of oligosaccharides that utilize stoichiometric amounts of commercial sugar-nucleotides have generally been the preferred choice in many laboratories. Several examples of the synthesis of neo-glycoproteins and glycolipids of well-defined carbohydrate structures have also been recorded in the literature [9,41,42,43,44]. However, with an increasing academic and pharmaceutical interest in the glycomics arena an urgent need for the accessibility of glycoconjugates has been recognized and this reflects the attention given to the large-scale synthesis of this group of compounds.

Recent advances in cloning and recombinant expression of bacterial GTs and their accessory enzymes such as sugar-nucleotide synthases have enabled further scale-up synthesis of

Table 1
Overexpressed GTs for preparative synthesis and conditions used for assays and synthesis

Entry	Complete name (abbreviation)	Source	Expression cell	Activity (U/L)	Assay Acceptor	Buffer/pH (100–50 mM)	Sugar – nucleotide	Additives 10 mM/1 mM	Assay development	Used in synthesis
a	β1-3GlcNAc T(*lgtA*)	N. *meningitidis*	AD202	114U/L	LacNAcβsp	Caco,/7.5	UDP-^{3}HGN	$MnCl_2$		Cell lysate
b	β1-6GlcNAc T(Core-2)	Human	CHO	2U/L	GalNAc-αOBn	Caco,/7.0	UDP-^{3}HGN			Protein C
c	α1-3GalT (BraGalT1)	Bovine	AD202	375U/L	Lactose	Caco,/7.5	UDP-^{3}HGal	$MnCl_2$	Dowex/Cl^-	Ni-Agarose
d	α1-3GalT (GTB)	Human			LacNAcβsp	Caco,/7.0	UDP-^{3}HGal	$MnCl_2$/DTT		SP-Seph.
e	β1-3GalT(GalT5)	Human	Sf-9	15 U/L	GlcNAcβ sp	Mes, /6.0	UDP-^{3}HGal	$MnCl_2$	Dowex/Cl^-	Concentrate
f	β1,3GalT(cgtB)	C. jejuni	AD202	0.2 U/L	GlcNAcβOMe	Mes, /6.0	UDP-^{3}HGal	$MnCl_2$/DTT	Dowex/Cl^-	Lysate
g	α1-4GalT/UDPGalE (lgtC)	N. *meningitidis*	AD202	66U/L	Lactose??	Tris,/7.5	UDP-^{3}HGlc	$MnCl_2$/DTT	Dowex/Cl^-	Lysate
h	β1-4GalT/UDPGalE (lgtB)	N. *meningitidis*	AD202	72U/L	GlcNAcβOMe	Tris,/7.5	UDP-^{3}HGlc	$MnCl_2$	Dowex/Cl^-	Lysate
i	α1-3GalNAc (GTA)	Human			LacNAcβsp	Caco,/ 7.0	UDP^{3}HGalN	$MnCl_2$/DTT		SP-Seph.
j	β1-3GalNACT/ UDPGalNAcE (lgtD)	P. shigelloides	AD202	40U/L	Galα4Lac/Lac	Hepes/7.5	UDP-^{3}HGN	$MnCl_2$/DTT	Dowex/Cl^-	Ni-Agarose
k	β1-4GalNAcT (cgtA)	C *jejuni*	AD202	166U/L	SiaLactose	Hepes/7	UDP-^{3}HGN	$MnCl_2$	Dowex/PO_4	Lysate
l	β1-4GalNAcT (mutant)	Bovine	BL21(D3)	40U/L	GlcNAc-spacer	Tris/8	UDP-^{3}HGN		Dowex/Cl^-	
m	α1-2FucT (FUT2)	Human	Sf-9	10U/L	Lactose	Tris/7.5	GDP-^{14}CFuc	$MnCl_2$	Dowex/Cl^-	Concentrate
n	α1-3/4FucT (FUT3)	Human	Sf-9	15U/L	Lac/LacNAc	Tris/7.5	GDP-^{14}CFuc	$MnCl_2$	Dowex/Cl^-	Concentrate
o	α1-3FucT (FUT4)	Human	Sf-9	8U/L	Lac/LacNAc	Tris/7.5	GDP-^{14}CFuc	$MnCl_2$	Dowex/Cl^-	Concentrate
p	α1-3FucT (FUT5)	Human	Aspar.niger	19U/L	Lac/LacNAc	Tris/7.5	GDP-^{14}CFuc	$MnCl_2$	Dowex/Cl^-	NH4 precip.
q	α1-3FucT (FUT6)	Human	Sf-9	25U/L	Lac/LacNAc	Tris/7.5	GDP-^{14}CFuc	$MnCl_2$	Dowex/Cl^-	Concentrate
r	α1-3FucT(FUT7)	Human	A. Niger	18 U/L	SiaLacNAc	Tris/7.5	GDP-^{14}CFuc	$MnCl_2$	Dowex/Cl^-	NH4 Precip.

Table 1
(continued)

Entry	Complete name (abbreviation)	Source	Expression cell	Activity (U/L)	Assay Acceptor	Buffer/pH (100–50 mM)	Sugar – nucleotide	Additives 10 mM/1 mM	Assay development	Used in synthesis
s	α2-3SiaT/CMP Synt (ST3)	N. *meningitidis*	AD202	124U/L	Lactose	Caco/6.5	CMP-^{14}CSia		Dowex/PO_4	PEG precip.
t	α2-3SiaT (rST3GalIII)	Rat	Aspar.niger	38U/L	Lactose	Caco/6.5	CMP-^{14}CSia	$MnCl_2$	Dowex/PO_4	NH4 precip.
u	α2-3SiaT/α2-8SiaT (cstII)	C. jejuni	AD202	50U/L?	Lactose	Hepes/7.5	CMP-^{14}CSia	$MnCl_2$	Dowex/PO_4	lysate
v	α2-3SiaT(pST3GalI)	Porcine	Sf-9	15U/L	Lactose	Caco/6.5	CMP-^{14}CSia		Dowex/PO_4	CDP affinity
w	α2-6SiaT (hST6GalI)	Human	Hi Five	14U/L	Lactose	Caco/6.5	CMP-^{14}CSia	NaCl	Dowex/PO_4	SP ion exch.
x	α2-6SiaT(chST6GalNAcI)	Chicken	Hi Five	20U/L	Asialofetuin	Caco/7.5	CMP-^{14}CSia	NaCl	Seph. G-50	CDP affinity
y	UDP-GlcNAc/GalNac-4-E	Rat	AD202	1100U/L	UDPGNAc/LN	Hepes/7.5	UDP-^{3}HGlc	$MnCl_2$/NAD	Dowex/Cl^-	Lysate
z	UDP Glc/GalE (GalE)	P.aeruginosa	AD202	205 U/L	UDPGlc/ lactose	Hepes/7.5	UDP-^{3}HGlc	$MnCl_2$/NAD	Dowex/Cl^-	Lysate
å	KSGal6ST	Human	CHO	1-2U/L	Di-LacNAc	Immidazol/6.5	PAPS	$CaCl_2$	Dowex/PO_4	Concentrate
ä	KSGal6ST	Human	CHO	1-2U/L	GlcNAcβ3-LN	Immidazol/6.5	PAPS	$CaCl_2$	Dowex/PO_4	Concentrate
ö	α1-4GnT	Human	CHO	1U/L	LacNAc	Caco/7.5	UDP-^{3}HGN	$MnCl_2$	Dowex/Cl^-	Concentrate
A	β1-3GalT	Drosophila	Sf9	2U/L	GalNAc-pNP	Caco/ 6.5	UDP-Gal	MnCl2	C18	Cell lysate

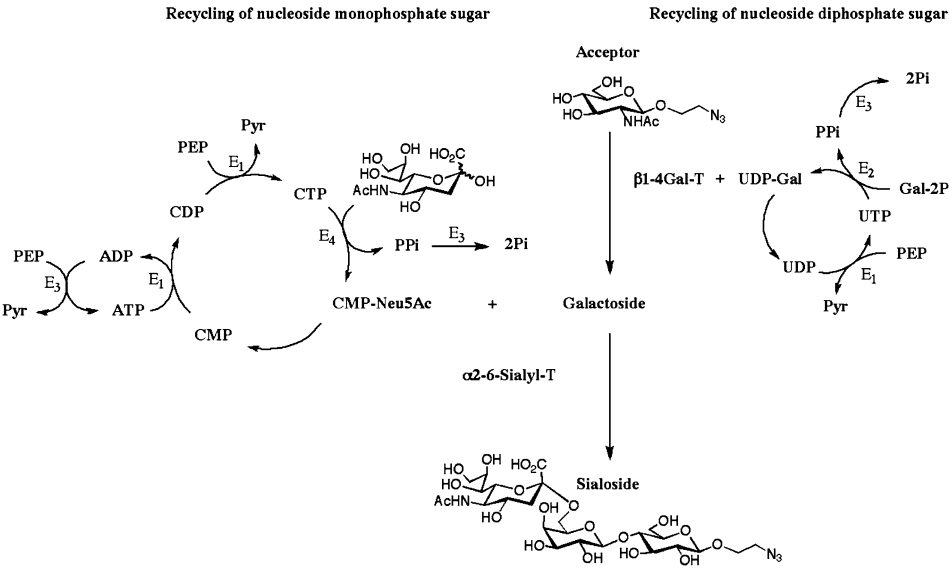


Figure 2
Enzymatic glycosylation using sugar nucleotide recycling systems. E_1 = pyruvate kinase, E_2 = sugar nucleotide pyrophosphorylase, E_3 = pyrophosphatase, E_4 = sugar nucleotide synthase, E_5 = myokinase, PEP = phosphoenol pyruvate, Pyr = pyruvate

oligosaccharides [20,45]. Several of these sugar-nucleotide synthases and other aldolases can efficiently generate the donor substrates from inexpensive precursors in bulk quantities. For example, utilizing rat UDP-*N*-acetylglucosamine-4′epimerase (GalNAcE) that converts less expensive UDP-GlcNAc to UDP-GalNAc in-situ has drastically reduced the cost of multi-gram synthesis of galactosides [46]. Moreover, the accessory enzyme can also be conveniently combined into one-pot reactions without the need of a full re-cycling system all performed in scaleable amounts (*Fig. 3*).

3 Enzymatic Glycosylation in Solution

Enzymatic synthesis of oligosaccharides can be performed with the general procedure as follows: Reaction mixtures containing acceptor substrates (20–50 mM) and the corresponding sugar-nucleotide donor substrate (1–3 mole equivalents) are dissolved in an enzyme specific buffer including additives (see *Table 1* for details). Reaction is initiated by adding enzymes (typically 1–5 Units / mmole acceptor) to the reaction mixture after which it is slowly agitated at room temperature or 37 °C for 24–48 h. Ion exchange and size exclusion chromatography are the common procedures for product isolation. However, serious purification problems can be encountered when scaling up to multi-gram reactions. Additives, such as nucleotide sugars, buffer salts, and crude enzymes are among the factors that interfere with efficient purification with conventional chromatographic techniques used at the mg-scale. In general, the selected

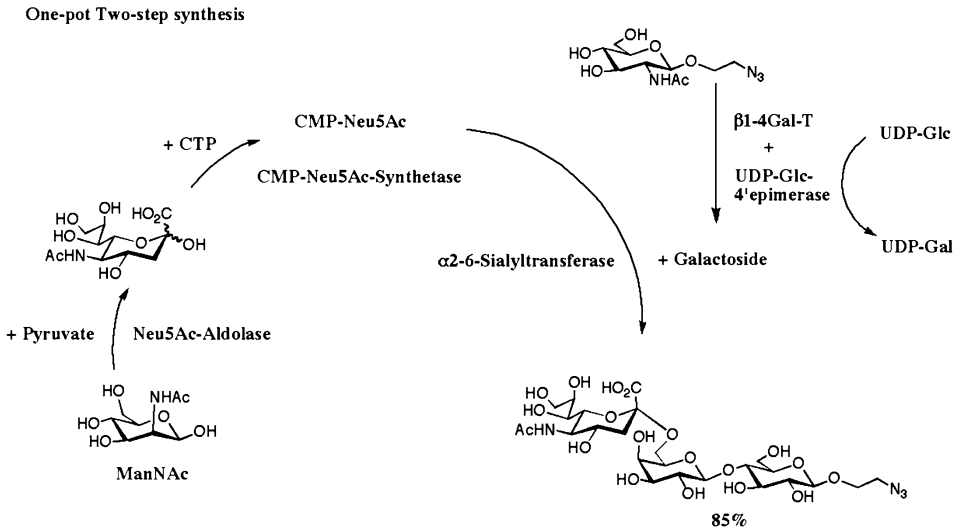


Figure 3
Enzymatic synthesis using partial in-situ generation of sugar nucleotides

isolation strategy depends on the scale and the compounds being prepared. Typical yields are 50–90% with a purity of 90–95% [47]. Various oligosaccharide sequences representing terminal structures of *N*-glycans [48,49] and glycolipids [46,50] have been successfully synthesized in gram-scales, using the simplified one-pot approach.

3.1 Poly-*N*-Acetyllactosamine Synthesis

Poly-*N*-acetyllactosamine (pLN), a unique carbohydrate structure composed of *N*-acetyllactosamine (LN) repeats, is known to represent the backbone of many glycan structures that participate in different cell functions [51,52] as well as being a tumor-associated antigen [53,54,55,56]. pLNs are present in two forms of linkages, Galβ1-3GlcNAc and Galβ1- 4GlcNAc, which constitute the type I pLN and type II pLN, respectively. Because of the practical limitations, most of these structures are not available by isolation. Therefore, the preparation of lactosamine derivatives is crucial for studying their structures and precise interactions with other biomolecules. Numerous chemical syntheses, including chemoenzymatic efforts [57,58,59,60] for pLN have been developed [61,62,63,64,65]. Recently, we further improved the synthesis of pLNs by taking advantage of the relaxed substrate specificity of the recombinant bacterial enzymes β1,4-galactosyltransferase (β4GalT) (entry h, *Table 1*) and β1,3-*N*-acetylglucosaminyltransferase (β3GlcNAcT) (entry a, *Table 1*) [66,67] which, by concerted action, generated repeating LN units. A similar strategy was used, albeit utilizing a β1,3galactosyltransferase (GalT-5) (entry e, *Table 1*) to make type 1 structures. These structures were further modified with different recombinant fucosyltransferases (FUTs) and sialyltransferases (STs) to obtain a diverse collection of pLN derivatives (*Fig. 4* and *Fig. 5*) [49]. Type 1 and type 2 pLN repeats are seen either separately each in one chain,

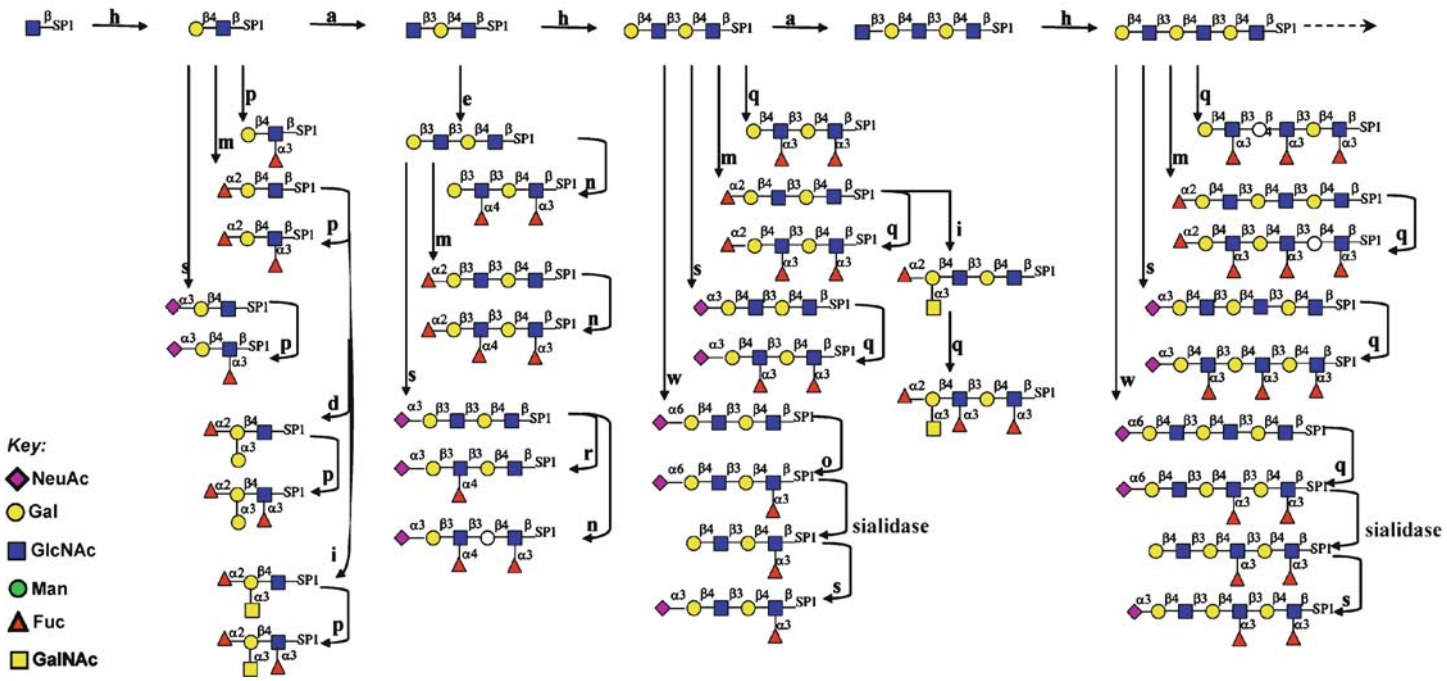


Figure 4
Enzymatic synthesis of poly-*N*-acetyllactosamines of type 2 chains

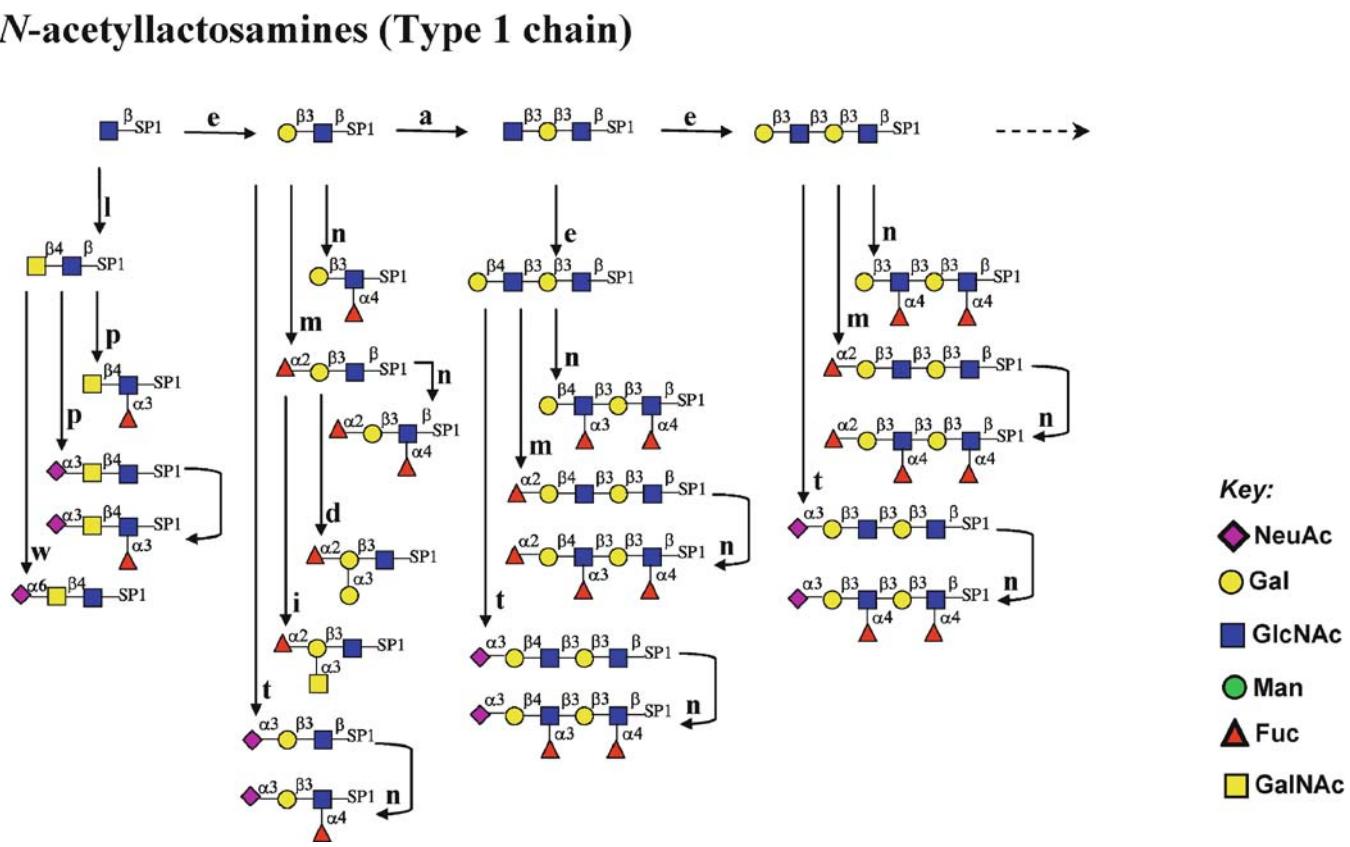


Figure 5
Enzymatic synthesis of poly-*N*-acetyllactosamines of type 1 chains

or in combination within the same pLN chain. A wide array of sialylated and/or fucosylated derivatives of type 1, type 2, and combinations of both types are found in nature.

3.2 Ganglioside Mimetic Synthesis

Gangliosides are diverse sets of sialylated glycolipids that are integral components of most cell membranes, and particularly abundant in neuronal tissues. They are implicated as tumor-associated antigens; active immunization with gangliosides has been shown to suppress melanoma growth [68,69,70,71,72,73]. Several elegant chemical approaches for ganglioside synthesis have been reported over decades by various groups [74,75,76]. The crucial step to synthesize the gangliosides is the introduction of multiple units of sialic acids onto these structures. The enzymatic approach in generating multi-sialylated glycans has significantly facilitated the time demanding and tedious procedures of ganglioside synthesis [46]. A key enzyme in this reaction is the bifunctional α-(2→3/8)-sialyltransferase (α3/8SiaT) [77,78] that incorporates multiple sialic acid units onto glycans (poly-sialic acidk). The number of the sialic acid units is regulated by the controlled supply of the sugar nucleotide donor, CMP-Neu5Ac. By using a chemoenzymatic approach, a series of ganglioside mimetics, built on a chemically made 2-azidoethyl lactoside, utilizing a combination of relevant enzymes, were synthesized (*Fig. 6*) [46].

3.3 Globo-Series of Glycolipid Synthesis

The P blood group antigens are glycan structures displayed by membrane-associated glycosphingolipids present on red blood cells and on other tissues [79,80]. The P1 antigen is formed by the addition of a galactose in an α1,4-linkage to the paragloboside by the P1 α1,4galactosyltransferase, forming the pentasaccharide Galα1–4Galβ1–4GlcNAcβ1–3Galβ1–4Glc [81]. The physiological functions of the P blood group antigens are not known, but these molecules have been implicated in the pathophysiology of urinary tract infections and parvovirus infections (*Fig. 7*) [82].
Using a recombinant fusion construct of the key enzyme α1,4galactosyltransferase-GalE [81] (entry g) and a β1,3-*N*-acetylgalactosaminyltransferase [83] (entry j *Table 1*), in combination with the recombinant GalNAc epimerase (entry y *Table 1*) [46] we generated various structures of the globoside series (published at http://www.functionalglycomics.org).

3.4 *O*-Linked Glycans Synthesis

The most common *O*-linked carbohydrates are based on core structures represented by the Tn-(GalNAcα1–1Thr/Ser) and T- (Galβ1–3GalNAcα1–1Thr/Ser)-antigens (*Fig. 8*). Sialylated versions of these *O*-antigens are expressed at low levels by many normal tissues, but become highly expressed in many types of human malignancies including colon, breast, pancreas, ovary, stomach, and lung adenocarcinomas, as well as, myelogenous leukemias [84,85,86]. *O*-glycans based on the Core-2 epitope (Galβ1–3[GlcNAcβ1–6]GalNAcα1–1Thr/Ser), on the other hand, are expressed in normal tissues. Synthesis of various compounds of *O*-linked structures using the key enzymes of human core-2-β1–6GlcNAcT (C2GnT), chicken GalNAcα2–

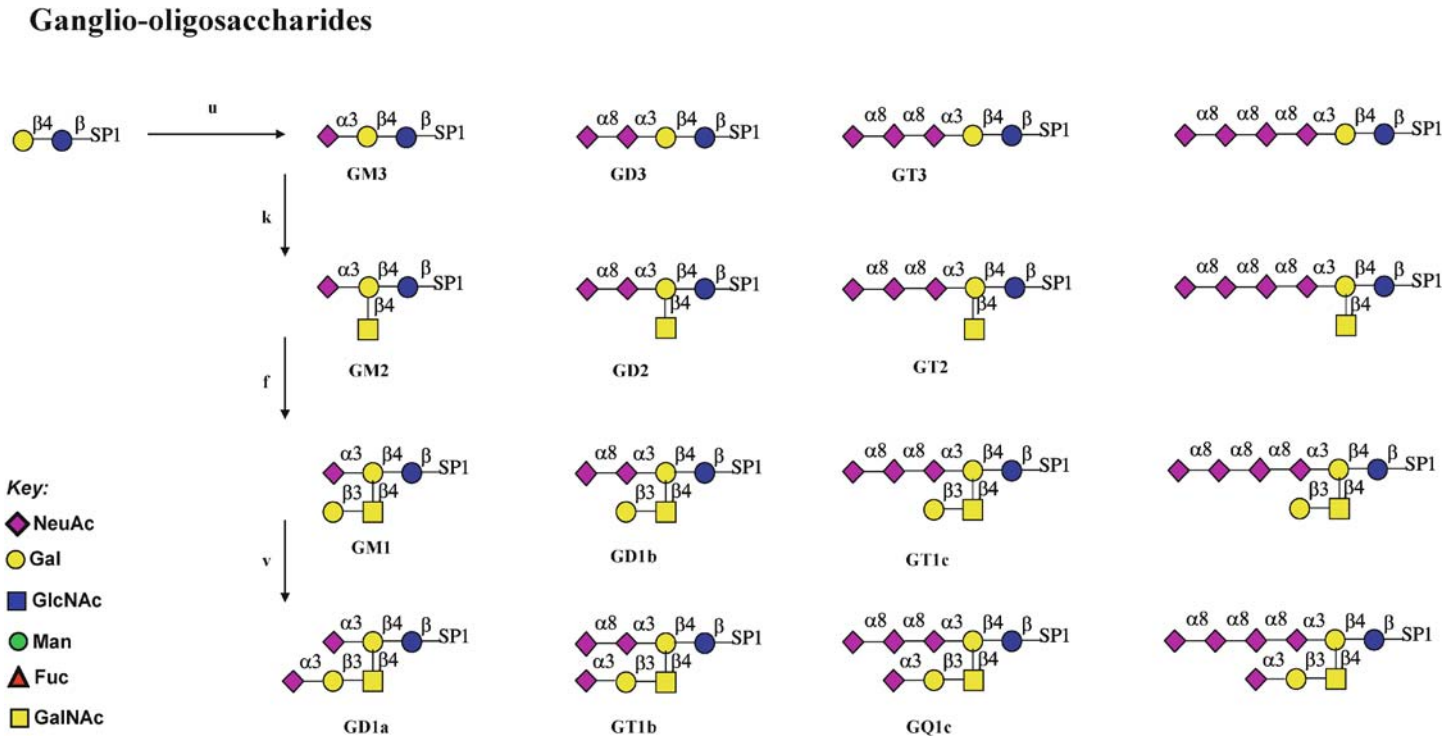


Figure 6
Enzymatic synthesis of ganglio-oligosaccharides

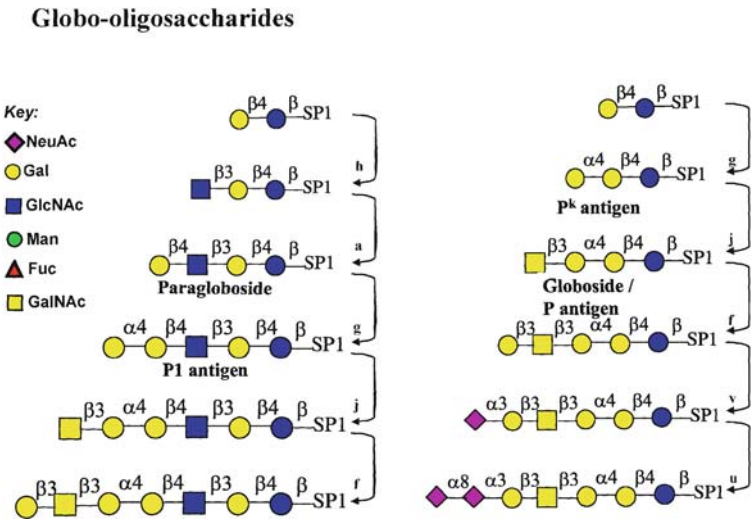


Figure 7
Enzymatic synthesis of globo-oligosaccharides

6-sialyltrasferase (chST6GalNAc-I), and pig Galα2–3-sialyltrasferase (ST3Gal-I) on the threonine-containing aglycon acceptors (*Fig. 8*) have been accomplished [87] (also published at http://www.functionalglycomics.com). Recently, elegant enzymatic elongations of peptides using recombinant polypeptide *N*-acetylgalactosaminyltransferases [88,89,90] were performed and demonstrate the power of enzymes for a fully enzymatic executed synthesis of complex *O*-linked glycoproteins with well-defined carbohydrate structures.

3.5 α-1,4-*N*-Acetylglucosamine-Capped Glycan Synthesis

Glycoproteins carrying GlcNAcα1,4Galβ-R at non-reducing terminals are defined structures of particular interest, they are found on the class III gastrointestinal mucins. Class III mucins exist in deeper portions of the mucosa and are secreted from the gastrointestinal gland mucous cells, including mucous neck cells, cardiac gland cells, and pyloric gland cells [91]. *Helicobacter pylori* rarely colonize the deeper portions of gastric mucosa, where the gland mucous cells produce mucins having terminal α1,4-*N*-acetylglucosamine (α1,4-GlcNAc) attached to *O*-glycans. It has been proposed that α1,4-GlcNAc-capped *O*-glycans have protective properties against *H. pylori*. Thus, the unique *O*-glycans in gastric mucin appeared to function as a natural antibiotic, protecting the host from H. *pylori* infection [92].

The human α1,4-*N*-acetylglucosaminyltransferase, which has been cloned [91] and overexpressed in stable CHO cell lines, was used for preparative synthesis. This enzyme catalyzes the transfer of GlcNAc, with an α-linkage to C4 of the non-reducing terminal galactose in beta linkage, α1,4-GlcNAc-capped sequences of *N*- and *O*-glycans were synthesized in milligram amounts (published at http://www.functionalglycomics.com).

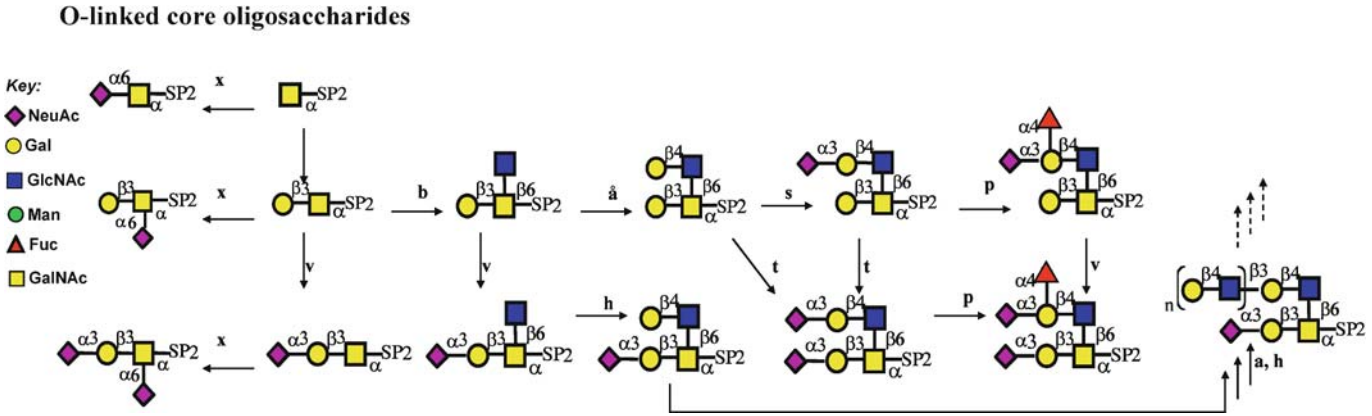


Figure 8
Enzymatic synthesis of *O*-glycans

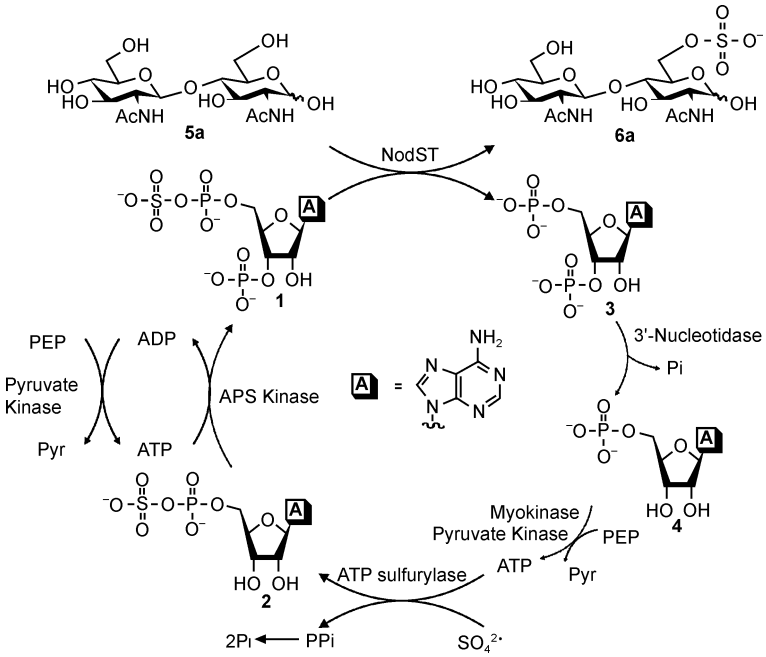


Figure 9
Multi-enzyme PAPS regeneration (from reference [39] with permission from publisher)

3.6 Synthesis of Sulfated Glycans

Sulfated glycans have generated interest due to their roles in specific cell signaling and biological functions both in normal and disease processes. Just as the kinases that mediate the diverse phosphorylation events within the cytosol have received great attention in recent years, the Golgy-associated sulfotransferases that impart specific sulfation modifications are now stimulating considerable interest [93,94,95]. Sulfated oligosaccharides can be chemically synthesized and followed by enzymatic elongations [96,97,98]. A number of sulfated disaccharides of Galβ1–3GlcNAc, Galβ1–4GlcNAc and Galb1–4Glc, sulfated at various positions, were chemically synthesized and further elongated by enzymatic approach with ST and FUT [99] from *Table 1* (see http://www.functionalglycomics.com). In addition, sulfotransferases of Keratan Sulfate 6-*O*-galactosylsulfotrasferase (KSGal6ST) [100,101] and corneal 6-*O*-glucosaminylsulfotransferase (GlcNAc6ST) [102] are promising enzymes for sulfation of poly-LacNAc derivatives. However, the preparative enzymatic sulfation is untenable because of the high cost and instability of phospho adenosin phospho sulfate (PAPS), which is the universal donor of sulfate for most sulfotransferases. Furthermore, the released PAP (3′-phosphoadenosine-5′-phosphate), from the PAPS desulfation in sulfotransferase reactions, is a potent inhibitor of the sulfotransferase activity. To overcome the PAPS limitations in enzymatic sulfation with sulfotransferases, different PAPS regenerating methods have been developed (*Fig. 9*) [39].

3.7 Synthesis of *N*-Linked Glycans

Isolation and purification of natural glycans is complicated and tedious. Nevertheless, it is important to include them in existing and future glycan libraries. Several recent reports demonstrate that isolated or chemically synthesized [103,104] *N*-glycans can be efficiently modified with various glycosyltransferases [105,106,107,108]. The accumulated enzyme library at the CFG (● *Table 1*) is being used for modifications of *N*-glycans as exemplified in ● *Fig. 10* (http://www.functionalglycomics.com).

4 Enzymatic Glycosylation on Solid-Phase

Enzymatic solution-phase synthesis still requires several tedious chromatographic purifications between the enzymatic elongation steps. Enzymatic elongation of oligosaccharides on a solid-phase support was first successfully demonstrated on a controlled-pore glass [109], which has helped avoid the tedious purification process. The key issue in the solid-phase strategy is to establish a proper solid-phase support that is compatible with both organic and aqueous solvents and swells well in an aqueous buffer, so that the enzymes could access the substrates easily. The strategy also requires a proper linker to anchor the oligosaccharide, and the linker should be selectively cleavable to release the product (● *Fig. 11*) [110]. The tetrasaccharide sialyl-Lewis^x was synthesized on an agarose matrix in three consecutive enzymatic steps, followed by a final size exclusion chromatography purification step in 57% yield [111].
A combined solid- and soluble-phase synthesis was used in the preparation of PSGL-1 *N*-terminal glycopeptide that contains a sulfated tyrosine and an α-*O*-linked sialyl Lewis^X structure [112]. In this method, a glycopeptide with a short disaccharide was assembled by conventional solid-phase peptide synthesis. After the glycopeptide was sulfated and deprotected, the resultant free glycopeptide was treated with a series of glycosyltransferases including β-1,4-galactosyltransferase, α-2,3-sialyltransferase, and α-1,3-fucosyltransferase in the presence of corresponding glycosyl donors to afford the target glycopeptide. The product of each enzymatic reaction was purified by size exclusion and ion-exchange columns.

5 Immobilized Glycosyltransferases

The use of immobilized glycosyltransferases for synthesis offers the advantage to use the same enzyme molecule over a long (● *Fig. 12*) period of time. The extent of the reaction can be controlled and there are automation capabilities as well for continuous operations. Auge and co-workers pioneered this multi-enzyme immobilization technique and demonstrated efficient synthesis of blood group I epitopes [113]. Similar work was also reported for both immobilization of enzymes [114] and whole enzyme producing cells [40]. The enzyme-displayed yeast cell technique using *S. cerevisiae* has been used to immobilize enzymes on the same cell surface to promote one-pot sequential synthesis of oligosaccharides [115]. Sequential synthesis of sLe^X by co-expressing catalytic domains of two mammalian glycosyltransferases (α2,3 sialylT and α1,3 fucosylT VII) fused with yeast cell wall proteins was achieved [116]. In attempts to elucidate the functions of sugar chains as bioactive substances, as well as, advancing the techniques for glycan synthesis and carbohydrate analysis, the National Institute of

Figure 10
Examples of enzymatic remodeling of *N*-glycans isolated from egg yolk

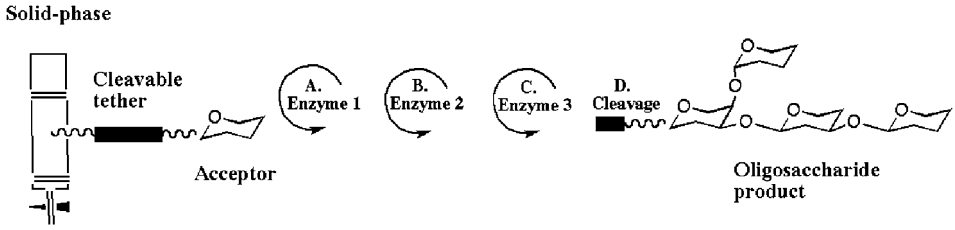


Figure 11
Chemoenzymatic solid-phase oligosaccharide synthesis

Advanced Industrial Science and Technology in Japan has recently launched a project titled "Construction of a glycogene library" [117]. This program aims at identifying and cloning all the possible human glyco genes, particularly glycosyltransferases. The enzyme-displayed yeast cell technique has been employed to develop a library of enzymes expressed on the yeast cell wall [23,118] (*Fig. 13*).

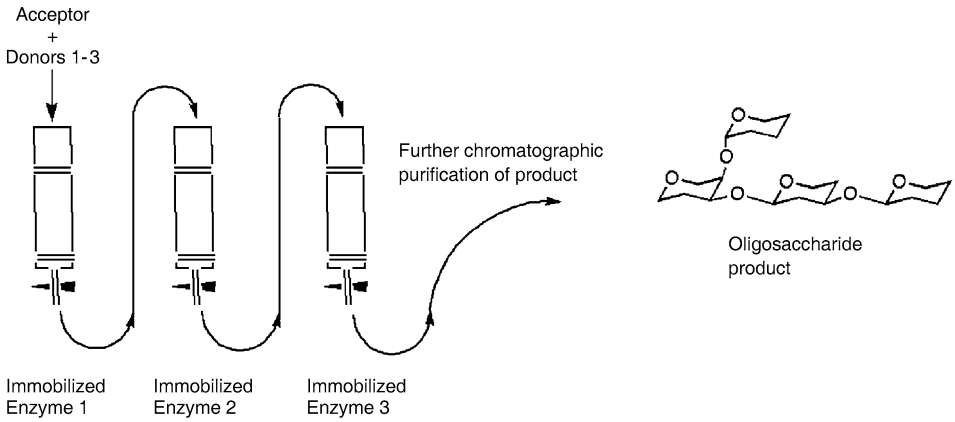


Figure 12
A serial-sequence of immobilized glycosyltransferases for oligosaccharide production

6 In-situ Glycosylation Using Engineered Bacteria

A very elegant and innovative approach for large-scale production of oligosaccharides was developed by Koizumi and co-workers. Expensive sugar-nucleotides were regenerated in-situ by coupling metabolically engineered *E. coli* bacteria [119,120,121]. Starting from inexpensive and industrially available precursors, the microorganisms produce high levels of nucleoside 5′-triphosphates and thereby, overcome the drawbacks of enzymatic methods that require expensive phosphophenolpyruvate and nucleoside 5-phosphates. Multi-gram scale synthesis of di- and trisaccharides such as LacNAc, globi-triose, and sialyllactose have been accomplished [22,122].

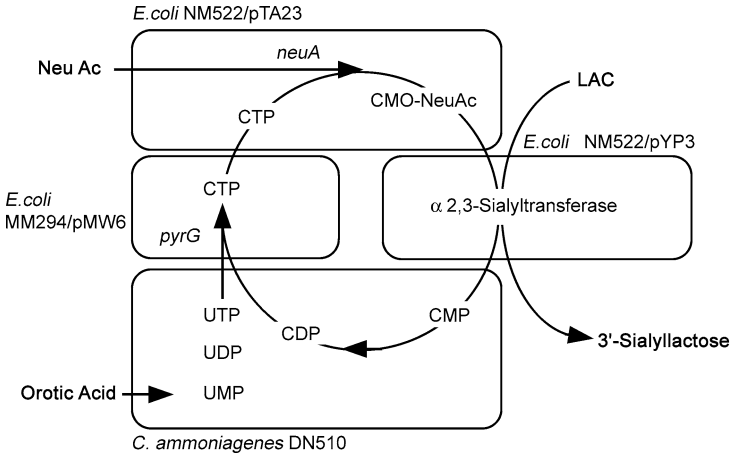


Figure 13
Production system of CMP-Neu5Ac and 3′-sialyllactose (adapted from reference [22] with permission from publisher)

7 Future Outlook

Sugar-chain engineering is one of the key techniques of biotechnology. As knowledge of their physiological roles increases, research on carbohydrate-based therapeutics will continue to accelerate and draw more attention to oligosaccharide synthesis. Once viewed as a hydrophilic heterogeneous nuisance to be removed from cell surfaces or secreted proteins of interest, oligosaccharides are now being synthesized on proteins and displayed on different platforms to mimic the repertoire of the cell surface glycans.

Continuous improvement of chemoenzymatic glycan synthesis is imperative. The combined efforts of the recent development of chemical advances and the exploration of new and more flexible GTs will continue to make advances towards synthesis of the glycome. Some of the GTs show remarkable flexibility and broad tolerance towards non-natural residues of the acceptor substrates and the sugar nucleotide donors. For example, the replacement by aromatic or heteroaromatic moieties, charged residues, sulfonamides or bulky polar monosaccharides, are well tolerated by enzymes and offers an efficient approach to the creation of oligosaccharide analog libraries that can be used in drug developments by designing enzyme inhibitors and ligand antagonists [123]. Although an increasing number of glycosyltransferases are commercially available, they are still prohibitively expensive, particularly for synthesis on a preparative scale. Nonetheless, glycosyltransferases have shown promise as reagents for solid phase synthesis that may facilitate the development of automated solid-phase glycan synthesis, similar to what has revolutionalized the protein and nucleic acid era.

The cutting-edge technique, known as glycan array technology, has emerged as a new set of tools to facilitate the studies of carbohydrate–protein interactions [124]. A principle challenge in cell biology is to define the paradigm of carbohydrate protein interactions in many biological processes. Glycan array technology offers a solution to the high-throughput analysis and profiling of carbohydrate-related interactions, as well as, the identification of optimal

sugar ligands and inhibitors [125,126]. Expanding the glycan array with synthetic oligosaccharide structures that represent the complete pattern of the natural glycans is a focus for the future.

Acknowledgement

This work was funded by NIGMS and The Consortium for Functional Glycomics GM62116. Our thanks go to Ms. Davina Chavez for proof reading of the manuscript.

References

1. Varki A (1993) Glycobiology 3:97
2. Dwek RA (1996) Chem Rev 96:683
3. Garegg PJ (2004) Adv Carbohydr Chem Biochem 59:69
4. Nicolaou KC, Mitchell HJ (2001) Angew Chem Int Ed 40:1576
5. Bartolozzi A, Seeberger PH (2001) Curr Opin Struct Biol 11:587
6. Tanaka H, Adachi M, Tsukamoto H, Ikeda T, Yamada H, Takahashi T (2002) Organic Letters 4:4213
7. Angeloni S, Ridet JL, Kusy N, Gao H, Crevoisier F, Guinchard S, Kochhar S, Sigrist H, Sprenger N (2005) Glycobiology 15:31
8. Boons G-J, Demchenko AV (2000) Chem Rev (Washington, D. C.) 100:4539
9. Hanson S, Best M, Bryan MC, Wong CH (2004) Trends Biochem Sci 29:656
10. Wymer N, Toone EJ (2000) Curr Opin Chem Biol 4:110
11. Murata T, Usui T (1997) Biosci Biotechnol Biochem 61:1059
12. Crout DH, Vic G (1998) Curr Opin Chem Biol 2:98
13. Mackenzie LF, Wang QP, Warren RAJ, Withers SG (1998) J Am Chem Soc 120:5583
14. Seibel J, Jordening HJ, Buchholz K (2006) Biocatal Biotransfor 24:311
15. Koeller KM, Wong CH (2000) Glycobiology 10:1157
16. Johnson KF (1999) Glycoconjugate J 16:141
17. Kamath VP, Yeske RE, Gregson JM, Ratcliffe RM, Fang YR, Palcic MM (2004) Carbohydr Res 339:1141
18. Wakarchuk W, Martin A, Jennings MP, Moxon ER, Richards JC (1996) J Biol Chem 271:19166
19. Wakarchuk WW, Cunningham A, Watson DC, Young M (1998) Prot Eng 11:295
20. Gilbert M, Bayer R, Cunningham A-M, DeFrees S, Gao Y, Watson DC, Young NM, Wakarchuk WW (1998) Nat Biotechnol 16:769
21. Gilbert M, Brisson J-R, Karwaski M-F, Michniewicz J, Cunningham A-M, Wu Y, Young NM, Wakarchuk WW (2000) J Biol Chem 275:3896
22. Endo T, Koizumi S, Tabata K, Ozaki A (2000) Appl Microbiol Biotechnol 53:257
23. Shimma Y, Saito F, Oosawa F, Jigami Y (2006) Appl Environ Microbiol 72:7003
24. Shaper NL, Shaper JH, Meuth JL, Fox JL, Chang H, Kirsch IR, Hollis GF (1986) Proc Natl Acad Sci USA 83:1573
25. Narimatsu H, Sinha S, Brew K, Okayama H, Qasba PK (1986) Proc Natl Acad Sci USA 83:4720
26. Kikuchi N, Narimatsu H (2006) Biochim Biophys Acta 1760:578
27. Blixt O, Paulson JC (2003) Adv Synth Catal 345:687
28. Langenhan JM, Griffith BR, Thorson JS (2005) J Nat Prod 68:1696
29. Hidari KI, Horie N, Murata T, Miyamoto D, Suzuki T, Usui T, Suzuki Y (2005) Glycoconj J 22:1
30. Boeggeman EE, Balaji PV, Sethi N, Masibay AS, Qasba PK (1993) Protein Eng 6:779
31. Gilbert M, Cunningham AM, Watson DC, Martin A, Richards JC, Wakarchuk WW (1997) Eur J Biochem 249:187
32. Gilbert M, Watson DC, Cunningham AM, Jennings MP, Young NM, Wakarchuk WW (1996) J Biol Chem 271:28271
33. Borsig L, Ivanov SX, Herrmann GF, Kragl U, Wandrey C, Berger EG (1995) Biochem Biophys Res Commun 210:14
34. Bencurova M, Rendic D, Fabini G, Kopecky EM, Altmann F, Wilson IB (2003) Biochimie 85:413

35. Malissard M, Zeng S, Berger EG (1999) Glycoconj J 16:125
36. Herrmann GF, Elling L, Krezdorn CH, Kleene R, Berger EG, Wandrey C (1995) Bioorg Med Chem Lett 5:673
37. Lubineau A, Le Narvor C, Auge C, Gallet PF, Petit JM, Julien R (1998) J Mol Catal B-Enzymatic 5:229
38. Ichikawa Y, Wang R, Wong CH (1994) Methods Enzymol 247:107
39. Burkart MD, Izumi M, Chapman E, Lin CH, Wong CH (2000) J Org Chem 65:5565
40. Nahalka J, Liu Z, Chen X, Wang PG (2003) Chemistry 9:372
41. Koeller KM, Wong CH (2000) Chem Rev 100:4465
42. Sears P, Wong CH (2001) Science 291:2344
43. Hahn ME, Muir TW (2005) Trends Biochem Sci 30:26
44. Macmillan D, Bertozzi CR (2004) Angew Chem Int Ed 43:1355
45. Blixt O, Brown J, Schur MJ, Wakarchuk W, Paulson JC (2001) J Org Chem 66:2442
46. Blixt O, Vasiliu D, Allin K, Jacobsen N, Warnock D, Razi N, Paulson JC, Bernatchez S, Gilbert M, Wakarchuk W (2005) Carbohydr Res 340:1963
47. Blixt O, Razi N (2006) Methods Enzymol 415:137
48. Blixt O, Collins BE, van den Nieuwenhof IM, Crocker PR, Paulson JC (2003) J Biol Chem 278:31007
49. Vasiliu D, Razi N, Zhang Y, Jacobsen N, Allin K, Liu X, Hoffmann J, Bohorov O, Blixt O (2006) Carbohydr Res 341:1447
50. ITO Y, Paulson JC (1993) J Am Chem Soc 115:1603
51. Niemela R, Natunen J, Majuri ML, Maaheimo H, Helin J, Lowe JB, Renkonen O, Renkonen R (1998) J Biol Chem 273:4021
52. Ujita M, McAuliffe J, Hindsgaul O, Sasaki K, Fukuda MN, Fukuda M (1999) J Biol Chem 274:16717
53. Nicolaou KC, Caulfield TJ, Kataoka H, Stylianides NA (1990) J Am Chem Soc 112:3693
54. Nicolaou KC, Hummel CW, Iwabuchi Y (1992) J Am Chem Soc 114:3126
55. Hakomori S (2001) The Molecular Immunology of complex carbohydrates-2. Kluwer Academic/Plenum Publishers, Amsterdam, p 369
56. Leppaenen A, Penttilae L, Renkonen O, McEver RP, Cummings RD (2002) J Biol Chem 277:39749
57. Bårström M, Bengtsson M, Blixt O, Norberg T (2000) Carbohydr Res 328:525
58. Koeller KM, Wong CH (2000) Chemistry 6:1243
59. Bintein F, Auge C, Lubineau A (2003) Carbohydr Res 338:1163
60. Zeng X, Uzawa H (2005) Carbohydr Res 340:2469
61. Srivastava G, Hindsgaul O (1991) J Carbohydr Chem 10:927
62. Aly MRE, Ibrahim E-SI, El-Ashry ESH, Schmidt RR (2000) Eur J Org Chem: 319
63. Misra AK, Fukuda M, Hindsaul O (2001) Bioorg Med Chem Lett 11:2667
64. Mong TK-K, Huang C-Y, Wong C-H (2003) J Org Chem 68:2135
65. Buskas T, Li Y, Boons G-J (2005) Chem Eur J 11:5457
66. Blixt O, van Die I, Norberg T, van den Eijnden DH (1999) Glycobiology 9:1061
67. Blixt O, Brown J, Schur M, Wakarchuk W, Paulson JC (2001) J Org Chem 66:2442
68. Kannagi R, Izawa M, Koike T, Miyazaki K, Kimura N (2004) Cancer Sci 95:377
69. Gagnon M, Saragovi HU (2002) Expert Opin Ther Pat 12:1215
70. Svennerholm L (2001) Adv Gen 44:33
71. Schnaar RL (2000) In: Carbohydrates in Chemistry and Biology. Wiley-VCH, Weinheim, vol 4, p 1013
72. Ravindranath MH, Gonzales AM, Nishimoto K, Tam W-Y, Soh D, Morton DL (2000) Indian J Experiment Biol 38:301
73. Hakomori S-I, Zhang Y (1997) Chem Biol 4:97
74. Cheshev PE, Khatuntseva EA, Tsvetkov Iu E, Shashkov AS, Nifant'ev NE (2004) Bioorg Khim 30:68
75. Castro-Palomino JC, Simon B, Speer O, Leist M, Schmidt RR (2001) Chem-Eur J 7:2178
76. Hasegawa A (1992) Kagaku, Zokan (Kyoto, Japan), vol 122, p 21
77. Chiu CP, Watts AG, Lairson LL, Gilbert M, Lim D, Wakarchuk WW, Withers SG, Strynadka NC (2004) Nat Struct Mol Biol 11:163
78. Gilbert M, Karwaski MF, Bernatchez S, Young NM, Taboada E, Michniewicz J, Cunningham AM, Wakarchuk WW (2002) J Biol Chem 277:327
79. Hellberg A, Poole J, Olsson ML (2002) J Biol Chem 277:29455
80. Ziegler T, Jacobsohn N, Funfstuck R (2004) Int J Antimicrob Agents 24(1):70

81. Zhang J, Kowal P, Chen X, Wang PG (2003) Org Biomol Chem 1:3048
82. Yang Z, Bergstrom J, Karlsson KA (1994) J Biol Chem 269:14620
83. Shao J, Zhang J, Kowal P, Lu Y, Wang PG (2003) Chem Commun (Camb): 1422
84. Dabelsteen E (1996) J Pathol 179:358
85. Itzkowitz SH, Yuan M, Montgomery CK, Kjeldsen T, Takahashi HK, Bigbee WL (1989) Cancer Res 49:197
86. Mori S, Suzushima H, Nishikawa K, Miyake H, Yonemura Y, Tsuji N, Kawaguchi T, Asou N, Kawakita M, Takatsuki K (1995) Acta Haematologica 94:32
87. Blixt O, Allin K, Pereira L, Datta A, Paulson JC (2002) J Am Chem Soc 124:5739
88. Bennett EP, Hassan H, Mandel U, Mirgorodskaya E, Roepstorff P, Burchell J, Taylor-Papadimitriou J, Hollingsworth MA, Merkx G, van Kessel AG, Eiberg H, Steffensen R, Clausen H (1998) J Biol Chem 273:30472
89. Kato K, Takeuchi H, Kanoh A, Mandel U, Hassan H, Clausen H, Irimura T (2001) Glycobiology 11:821
90. Sorensen AL, Reis CA, Tarp MA, Mandel U, Ramachandran K, Sankaranarayanan V, Schwientek T, Graham R, Taylor-Papadimitriou J, Hollingsworth MA, Burchell J, Clausen H (2006) Glycobiology 16:96
91. Nakayama J, Yeh JC, Misra AK, Ito S, Katsuyama T, Fukuda M (1999) Proc Natl Acad Sci USA 96:8991
92. Aspholm-Hurtig M, Dailide G, Lahmann M, Kalia A, Ilver D, Roche N, Vikstrom S, Sjostrom R, Linden S, Backstrom A, Lundberg C, Arnqvist A, Mahdavi J, Nilsson UJ, Velapatino B, Gilman RH, Gerhard M, Alarcon T, Lopez-Brea M, Nakazawa T, Fox JG, Correa P, Dominguez-Bello MG, Perez-Perez GI, Blaser MJ, Normark S, Carlstedt I, Oscarson S, Teneberg S, Berg DE, Boren T (2004) Science 305:519
93. Habuchi O (2000) Seikagaku 72:427
94. Hemmerich S, Rosen SD (2000) Glycobiology 10:849
95. Fukuda M, Hiraoka N, Akama TO, Fukuda MN (2001) J Biol Chem 276:47747
96. Xia J, Alderfer JL, Srikrishnan T, Chandrasekaran EV, Matta KL (2002) Bioorg Med Chem 10:3673
97. Jain RK, Piskorz CF, Chandrasekaran EV, Matta KL (1998) Glycoconj J 15:951
98. Kuhns W, Jain RK, Matta KL, Paulsen H, Baker MA, Geyer R, Brockhausen I (1995) Glycobiology 5:689
99. Chandrasekaran EV, Jain RK, Larsen RD, Wlasichuk K, DiCioccio RA, Matta KL (1996) Biochemistry 35:8925
100. Fukuta M, Inazawa J, Torii T, Tsuzuki K, Shimada E, Habuchi O (1997) J Biol Chem 272:32321
101. Torii T, Fukuta M, Habuchi O (2000) Glycobiology 10:203
102. Uchimura K, Muramatsu H, Kadomatsu K, Fan QW, Kurosawa N, Mitsuoka C, Kannagi R, Habuchi O, Muramatsu T (1998) J Biol Chem 273:22577
103. Takatani M, Nakama T, Kubo K, Manabe S, Nakahara Y, Ito Y, Nakahara Y (2000) Glycoconj J 17:361
104. Mezzato S, Schaffrath M, Unverzagt C (2005) Angew Chem Int Ed 44:1650
105. Seko A, Koketsu M, Nishizono M, Enoki Y, Ibrahim HR, Juneja LR, Kim M, Yamamoto T (1997) Biochim Biophys Acta 1335:23
106. Kajihara Y, Yamamoto N, Miyazaki T, Sato H (2005) Curr Med Chem 12:527
107. Andre S, Kojima S, Gundel G, Russwurm R, Schratt X, Unverzagt C, Gabius HJ (2006) Biochim Biophys Acta 1760:768
108. Unverzagt C (1996) Angew Chem Int Ed Engl 35:2350
109. Halcomb RL, Huang HM, Wong CH (1994) J Am Chem Soc 116:11315
110. Blixt O, Norberg T (1997) J Carbohyd Chem 16(2):143
111. Blixt O, Norberg T (1998) J Org Chem 63:2705
112. Koeller KM, Smith ME, Wong CH (2000) Bioorg Med Chem 8:1017
113. Auge C, Fernandez-Fernandez R, Gautheron C (1990) Carbohydr Res 200:257
114. Nishiguchi S, Yamada K, Fuji Y, Shibatani S, Toda A, Nishimura SI, Nishimura SI (2001) Chem Comm: 1944
115. Abe H, Shimma Y, Jigami Y (2003) Glycobiology 13:87
116. Salo H, Sievi E, Suntio T, Mecklin M, Mattila P, Renkonen R, Makarow M (2005) FEMS Yeast Res 5:341
117. Narimatsu H (2004) Glycoconj J 21:17
118. Shimma Y, Jigami Y (2004) Glycoconj J 21:75
119. Koizumi S, Endo T, Tabata K, Ozaki A (1998) Nat Biotechnol 16:847
120. Endo T, Koizumi S, Tabata K, Ozaki A (2000) Glycobiology 10:1101

121. Endo T, Koizumi S (2000) Curr Opin StructBiol 10:536
122. Endo T, Koizumi S, Tabata K, Kakita S, Ozaki A (1999) Carbohydr Res 316:179
123. Ernst B, Oehrlein R (1999) Glycoconj J 16:161
124. Culf AS, Cuperlovic-Culf M, Ouellette RJ (2006) Omics 10:289
125. Paulson JC, Blixt O, Collins BE (2006) Nat Chem Biol 2:238
126. Blixt O, Head S, Mondala T, Scanlan C, Huflejt ME, Alvarez R, Bryan MC, Fazio F, Calarese D, Stevens J, Razi N, Stevens DJ, Skehel JJ, van Die I, Burton DR, Wilson IA, Cummings R, Bovin N, Wong CH, Paulson JC (2004) Proc Natl Acad Sci USA 101:17033

5.9 Enzymatic Glycosylation by Glycohydrolases and Glycosynthases

Julian Thimm, *Joachim Thiem*
Department of Chemistry, Faculty of Science,
University of Hamburg, 20146 Hamburg, Germany
joachim.thiem@chemie.uni-hamburg.de

Abstract

Chemoenzymatic applications of glycosidases and glycosynthases employing transglycosylation reactions are outlined. Addressed are the absolute stereospecificity as well as the generally poor regioselectivity of glycohydrolases. A dominant part of this chapter is concerned with the use of exo-glycosidases. Manifold reactions with different activated galactopyranosides are reported for the synthesis of a variety of terminally β-galactosylated oligosaccharides. By choice of the appropriate enzyme also the regiochemistry can be largely controlled. Some preparative transformations are reported for α-fucosidases, β-mannosidases, β-*N*-acetylhexosaminidases and—quite important—α-sialidases. Another part of this chapter describes the favorite use of *trans*-sialidases for important novel synthetic endeavors. Further, the chemistry with endo-glycosidases, in particular with regard to glycoprotein construction is mentioned. The last section briefly reports on recent excellent results employing glycosynthases for construction of complex heterooligosaccharides.

In: *Glycoscience*. Fraser-Reid B, Tatsuta K, Thiem J (eds)
Chapter-DOI 10-1007/978-3-540-30429-6_33:

Keywords

Exo-glycosidases; Endo-glycosidases; Glycosynthases; Transglycosylation; Reverse hydrolysis; Regio- and stereoselectivity

Abbreviations

DMF dimethyl formamide
DMSO dimethyl sulfoxide
HPLC high performacnce liquid chromatography
MAG myelin-associated glycoprotein
MPEG polyethyleneglycole ω-monomethylether
TSia *trans*-sialidase

1 Introduction

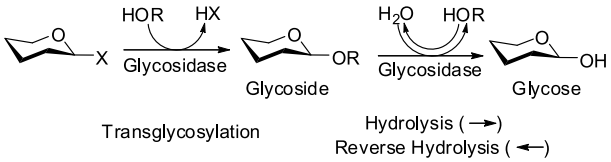
The chemoenzymatic assembly of complex oligosaccharides and glycoconjugates enjoys the application of all enzymes involved in carbohydrate metabolism and catabolism. In comparison to chemical glycosylation enzymatic glycosylation is of advantage in several cases, such as

- Unreliable chemical procedures for glycosylation.
- Harsh methods leading to decomposition and generally low yields.
- The wrong anomers are formed by chemical glycosylation, e. g. 1,2-*trans*-glycosides instead of the desired α-galactopyranosides or β-mannopyranosides.
- Heavy metal catalysts are employed, which is not acceptable for products with potential use in the pharmaceutical area or in nutrition due to toxicity and high price.

Most prominent is the use of glycosyltransferases in synthesis (cf. Chap. 5.8), however, the appropriate application of "glycosylhydrolases" (glycosidases) in the synthetic direction is of considerable interest. This subject has been of growing interest for preparative carbohydrate chemistry for about the last 15 years and thus has been reviewed several times [1,2,3,4,5,6,7,8,9,10,11,12].

2 Exo-Glycosidases

Exo-Glycosidases are involved in the trimming processes of glycoprotein biosynthesis [13]. In general, they are used to cleave the terminal non-reducing saccharide unit of an oligosaccharide (Gly-OR), and do show a high specificity for that unit. As to the interglycosidic linkage they are less specific, and the aglycon (R) may even vary widely. Thus, an activated intermediate will be formed and with water hydrolysis will occur. The process can be inverted and then reverse hydrolysis may be used to synthesize lower oligosaccharides. This is a thermodynamically controlled reaction which is favored, for example, by increase of concentration of the starting materials (Gly-OH and ROH) or enhanced temperature.



Scheme 1

Alternatively, transglycosylation may be employed which requires an activated glycosyl donor component (Gly-X) carrying a good anomeric leaving group (X). In this case, a kinetically controlled product formation is at hand, and in order to minimize the competing hydrolytic cleavage before equilibrium is reached the reactions should be quenched at the maximum of product formation. Generally, yields in transglycosylation are considerably higher than in reverse hydrolysis, and therefore the former is largely preferred (*Scheme 1*).

Glycosidases are low-priced enzymes, and they also require cheap donor substrates. They are quite robust to handle and show absolute stereoselectivity in their substrate recognition. The main drawbacks for their use in glycoside synthesis are generally low yields and little regioselectivity. This can be partly controlled by choice of the appropriate glycosidase with more pronounced regioselectivity or transglycosylations by dosing of reactants and use of co-solvents.

Another eminent advantage of using glycosidases in glycosylation is their good stereoselectivity. These enzymes are considered to be "retaining enzymes." There are, however, a few exceptions of inverting glycosidases which lead to products with an inverted anomeric configuration by using glycosyl fluorides or, for example, glycosyl pyridinium ions.

In contrast to glycosyltransferases, glycosidases are able to glycosylate many "xeno-substrates" with primary or secondary hydroxy groups. Even though the activity of water as acceptor in the reaction mixture is usually one to two orders higher than the activity of the alcohol to be glycosylated, the yields of the glycosides are often much higher than would be expected from simple thermodynamic calculations. This is considered to be caused by the enhanced affinity of the alcoholic substrates compared to water for binding to the "aglycon site" of the glycosidase. A possible influence of the groups adjacent to the respective hydroxy group is often speculated, and electron-rich moieties such as double bonds in the allylic position, conjugated systems of double bonds, aromatic or heteroaromatic systems generally support glycosylation. Aliphatic alcohols are also suitable acceptors for glycosyl moieties transferred by glycosidases. Amphiphatic properties and the presence of nitrogen in the acceptor molecule usually enhance transglycosylation yields.

A broad variety of substrates were efficiently glycosylated by glycosidases such as aliphatic and alicyclic alcohols [14], phenols, oximes, steroids and terpenes [15], amino acids [16], alkaloids [17] and many other substances.

The substrate to be glycosylated by a glycosidase should be at least partly soluble in water. The solubility can be enhanced by addition of water-miscible solvents, for example, acetonitrile, dimethylformamide, dimethylsulfoxide, dioxane, or *tert*-butyl alcohol. Concentrations of co-solvents up to 30% are usually well tolerated by most glycosidases. In contrast to lipases, lowering the water activity by co-solvents usually does not increase yields, the improved solubility of the acceptor and of the donor is the only decisive factor. There are, however, examples

of lowering the water activity by, for example, addition of various salts [18] that improved the glycosylation yields.

There are few examples of enzymatic glycosylation of simple alcohols in biphasic systems with organic solvents or by polyethyleneglycol-modified glycosidases in organic solvents [19]. In general glycosidases show rather poor regioselectivities, and with respect to their original biofunction this is expected. Nevertheless, many exhibit preferential cleavage of distinct glycosidic linkages and thus the corresponding transglycosylation effects. Should the acceptor substrate present more than one hydroxyl group for glycosylation the formation of several isomers should be encountered. Often the regioselectivity can be estimated or even predicted; for example, in hexopyranose acceptor substrates the affinity of hydroxyl groups is roughly 6-OH≫4-OH∼3-OH>2-OH resembling also the chemical reactivity of a nucleophilic OH-group.

There are reports that indicate that a change of configuration at the anomeric center should influence the regioselectivity [20,21], however, other subtle changes may be similarly effective. Quite often the regioselectivity is comparable for transglycosylation reactions and hydrolysis of the corresponding interglycosidic linkage. In many cases this could be demonstrated unequivocally by rational selection of the appropriate glycosidase as will be outlined below.

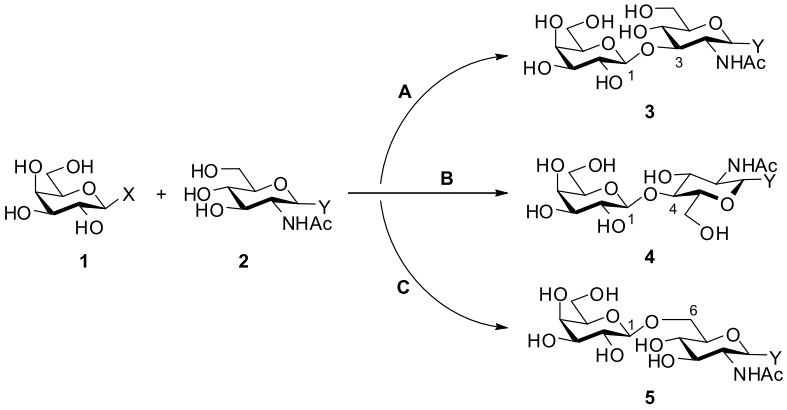
The stereoselectivity of glycosidases is always complete concerning the anomerity at the donor substrate site. In addition, many studies were performed to employ the chiral discrimination of these enzymes (as frequently used for lipases, esterases, etc.) regarding stereogenic centers in the aglyconic site. Both in the cleavage mode and also in transglycosylation such effects could be demonstrated [13,22,23,24], but will not be in the focus of this discussion.

2.1 β-Galactosidases

The availability and their easy handling has made β-galactosidases a preferred choice in preparative enzymatic glycoside synthesis often in positive competition to classical synthetic protocols. Employing as donor substrates galactopyranosyl derivatives **1** with various β-positioned leaving groups X and a suitable acceptor substrate, for example, the *N*-acetyl-glucopyranosyl component **2** with different anomeric functions Y, was shown to lead to the appropriate transglycosylation products **3**–**5**. Dependent on the β-galactosidase β,1–3-(**3**),β,1–4-(**4**, *N*-acetyllactosamine) or β,1–6-linked β-galactopyranosyl-*N*-acetylglucopyranosyl components (**5**) could be obtained. Advantageous for the formation of **3** were β-galactosidases from *Streptococcus* sp. or bovine testes [25], **4** could be favorably obtained employing β-galactosidase from *Bacillus circulans*, and **5** could be made using β-galactosidases from *Aspergillus oryzae*, *Penicillium multicolor*, *Kluyveromyces lactis*, and *Escherichia coli* [21] (Scheme 2).

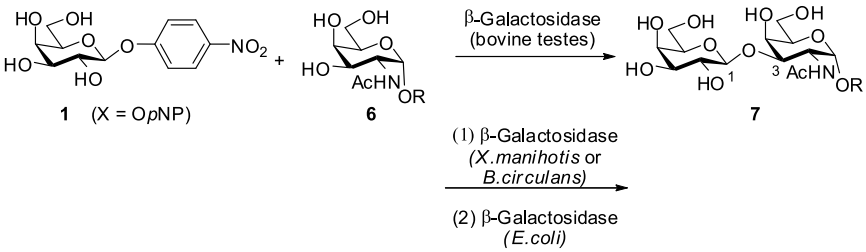
2.1.1 β,1–3-Galactosidases

En route to modified sialylated Thomsen–Friedenreich antigen components the core disaccharide motif Galβ1–3GalNAc needed to be easily accessible and thus β,1–3-galactosidase from bovine testes was employed [26,27]. A considerable number of GalNAc α-glycosides **6** could be transgalactosylated with **1** (X = *p*NP) to give the desired disaccharide derivatives **7** in 15–40% yield [25,28,29]. In more recent studies modifications could be employed to furnish yields up to 70% [30,31]. If β-galactosidases of *Xantomonas manihotis* (Xm) [32] or



Scheme 2

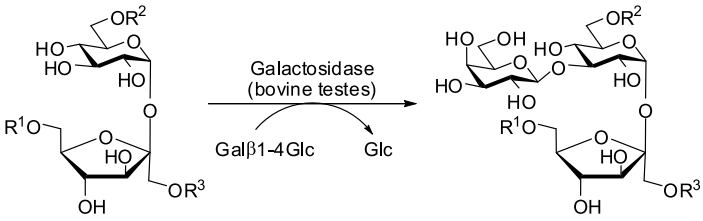
recombinant *Bacillus circulans* (bgaC-gene, Bc) [33] were employed for the same reaction in addition to **7** the corresponding regioisomeric Galβ1–6GalNAc derivative was formed [ratio 10 : 1 (with Xm) and 3 : 2 (with Bc)]. By in situ hydrolytic cleavage of the latter with β-galactosidase (*E. coli*) [26] the overall yields of **7** were 15–20% (with Xm) and about 30% (with Bc), respectively [30,31] (Scheme 3).



Scheme 3

Further acceptor modifications in these galactosylations were studied. Thus, 2-deoxy-galactose and *N*-acetylgalactose structures were recognized by all these galactosidases and gave the corresponding disaccharides in 20–70% yield [30]. With *N*-acetyl-chitobiose and -chitotriose β,1–3- and β,1–4-galactosylation (ratio 1 : 1) were reported [34]. Treatment of food saccharides such as sucrose (**8**), lactose, isomalt, isomaltulose, isomelezitose (**10**), raffinose (**12**), and 1-kestose (**14**) with lactose as donor substrate and β-galactosidase (bovine testes) furnished the terminally β,1–3-galactosylated-tri- to tetrasaccharides (**9**, **11**, **13**, and **15**) in 10–15% yield. The reactions could also be run on the 30–40 gram-scale and gave products in about 20% yield [30] (Scheme 4).

Modifications of donor substrates were also of interest. Generally, aromatic β-galactopyranosides were employed with good yields and easy UV-survey of the reaction due to the phenolic leaving group. Among them *o*- and *p*-nitrophenyl galactopyranosides represent standard donors, and also methylumbelliferyl as well as resorufinyl β-D-galactopyranosides were



8 $R^1 = R^2 = R^3 = H$: Sucrose **9**
10 $R^1 = \alpha$-Glc*p*, $R^2 = R^3 = H$: Isomelizitose **11**
12 $R^1 = R^3 = H$, $R^2 = \alpha$-Gal*p* : Raffinose **13**
14 $R^1 = R^2 = H$, $R^3 = \beta$-Fru*f* : 1-Kestose **15**

Scheme 4

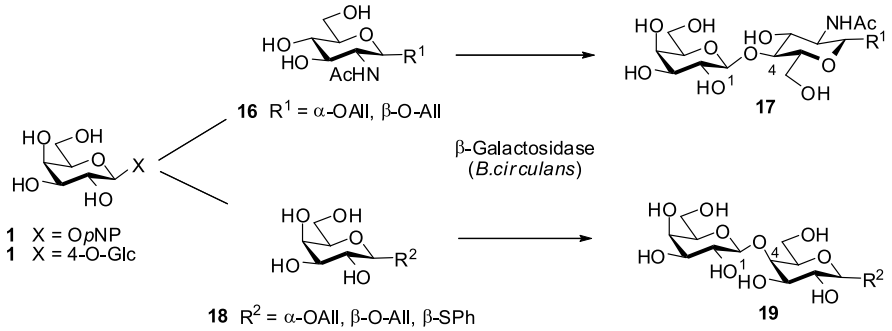
reported to be good donors. Often their solubility in aqueous–organic solvent mixtures causes problems. In a recent study [36] 16 aromatic and heteroaromatic β-galactopyranosides of type **1** were prepared. Their β,1–3-galactosylation potential for allyl *N*-acetyl-α-D-galactopyranoside was studied employing galactosidase from bovine testes to give the corresponding disaccharides in 10–75% yield. The solubility was also tackled in experiments studying various galactosides, however, *p*NP-β-D-Galp (**1**) was superior to lactose and was again much better than both lactulose and lactitol in donor substrate efficiency (about 30 : 10 : 2 : 1) [35].

2.1.2 β,1–4-Galactosidases

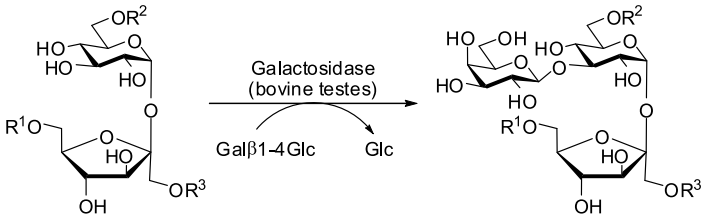
Early on, there were reports on the preparative use of β-galactosidase from *B. circulans* for the synthesis of β,1–4-galactosylated components [37]. This enzyme is commercially available under the name "Biolacta®", Lactase (EC 3.2.1.23) [38]. Biolacta is a quite stable enzyme ($t_{1/2}$ = 226 h at 22 °C and pH 7.0) that allows for an easy transfer reaction between pH 5.0–9.5 and up to 55 °C [38]. Its activity for transglycosylation in aqueous solvent mixtures was determined, and it showed considerable tolerance to several organic solvents. The activity in 30% aqueous organic solvent mixtures at pH 7.0 and 30 °C was measured after one hour and gave: DMSO (85%), acetone (70%), DMF (67%), dioxane (59%), and acetonitrile (12%) [39]. Studies toward the transglycosylation synthesis of *N*-acetyl-lactosamine and derivatives were of particular interest and performed by several groups [40,41,42]. As demonstrated [42] *p*NP-galactopyranoside (**1**, X = O*p*NP) was superior to lactose, and with α- as well as β-allyl *N*-acetyl-glucosaminide (**16**) the *N*-acetyllactosaminides could be obtained in up to 60% yield in acetonitrile/aqueous phosphate buffer. Correspondingly, with several α- and β-galactopyranosides (**18**) the appropriate Galβ1–4Gal disaccharides **19** could be synthesized in 50–70% yield (*Scheme 5*).

There was also interest in potential novel food additives that could be formed by β,1–4-galactosylation of natural oligosaccharides [43]. Compared to the results with β,1–3-galactosylation [35, see above] sucrose (**8**) and 1-kestose (**14**) were cleanly β,1–4 galactosylated at the 4 position of the incorporated glucopyranose residue to give the trisaccharide ("lactosucrose") **20** and the tetrasaccharide **21** in 20–30% yield (*Scheme 6*).

In contrast, both isomelezitose (**10**) and raffinose (**12**) gave rise to the formation of two isomeric tetrasaccharides. For **10** both available glucopyranose units were β,1–4-galactosylated



Scheme 5



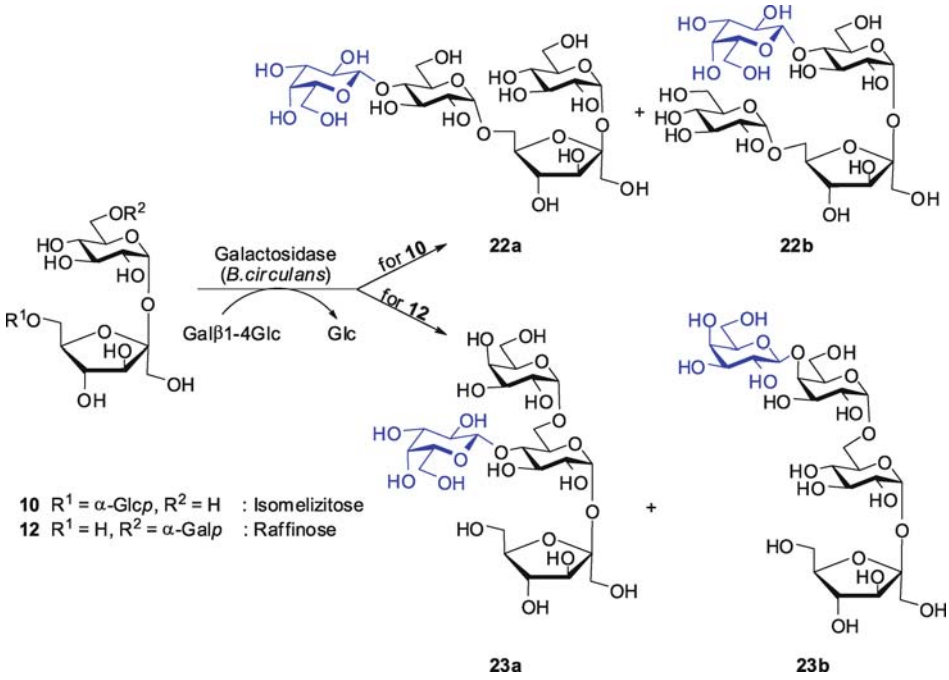
8 $R^1 = R^2 = R^3 = H$: Sucrose — **9**
10 $R^1 = \alpha$-Glc*p*, $R^2 = R^3 = H$: Isomelizitose — **11**
12 $R^1 = R^3 = H$, $R^2 = \alpha$-Gal*p* : Raffinose — **13**
14 $R^1 = R^2 = H$, $R^3 = \beta$-Fru*f* : 1-Kestose — **15**

Scheme 6

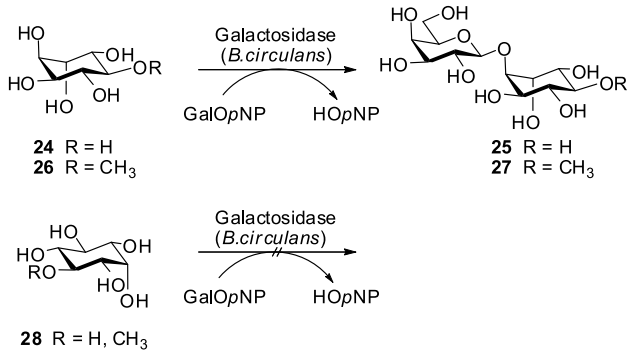
to give **22a** and **22b** (ratio 1 : 1, 21%), and for raffinose the β,1–4-galactosylation of the glucose residue gave **23a**, and that of the terminal galactopyranose ring led to **23b** (ratio 4 : 3, 38%) (*Scheme 7*).

Modifications of gluco-, galacto-, *N*-acetyl-gluco-, and *N*-acetyl-galacto-configured acceptor structures have been extensively studied. Generally, in the *gluco*-, equatorial-, and in the *galacto*-components the axial 4-hydroxy groups are galactosylated. With regard to the anomeric function, there were apparently no restrictions because alkyl- and aryl(thio)glycosides as well as *N*-acyl-glycosylamines could be employed [39,40,41,44,45,46,47,48]. Interestingly, galactosylation of inisitols could be also effected. Treatment of 1D-chiroinositol (**24**) as well as of 1D-pinitol (**26**) under standard conditions led to the galactosylated species **25** and **27**, respectively, in 45% yield. Apparently, these acceptors structures seem to resemble galactopyranose acceptors. In contrast, the isomers 1L-chiroinositol (**28**, R = H) and 1L-quebrachitol (**29**, R = CH_3) which rather resemble mannopyranose structures were not galactosylated [49] (*Scheme 8*).

Some recent work checked donor modifications in *B. circulans*-catalyzed galactosylations of allyl α-*N*-acetyl-hexosaminides (**30**) [50,51]. Whereas the 6-azido or 6-mesyloxy-galactopyranosides (**29** with R = CH_2N_3 or CH_2OMs) were not recognized, the 6-amino component

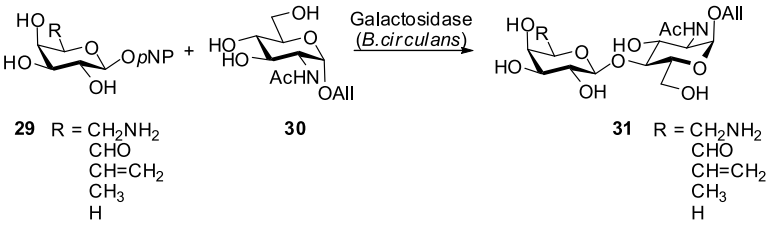


Scheme 7



Scheme 8

(**29**, R = CH_2NH_2) could be reacted to give the modified LacNAc glycoside structure **31** (R = CH_2NH_2). Reaction of the aldehyde **29** (R = CHO) to give **31** (R = CHO, 70% yield) could be confirmed [52] but surprisingly even the 6,7-olefin structure **29** (R = $CH{=}CH_2$) could be readily glycosylated to give the disaccharide **31** (R = $CH{=}CH_2$) in 30% yield. Also the 6-deoxy-galactopyranoside **29** (R = CH_3) acted as a quite good donor substrate to give the disaccharide **31** (R = CH_3) in 76% yield. In contrast, the configurationally corresponding L-arabino derivative **29** (R = H) showed only a transfer yield of 13% **31** (R = H) (*Scheme 9*).



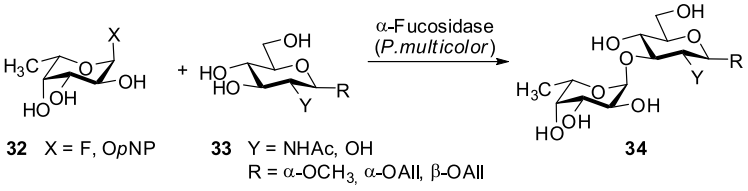
Scheme 9

2.1.3 β,1–6-Galactosidases

β-Galactosidases of *Aspergillus oryzae*, *Penicillium multicolor*, *Kluyveromyces lactis*, and *Escherichia coli* were reported to form β,1–6-linkages [21]. Employment of β-galactosidase of *Xantomonas manihotis* and recombinant *Bacillus circulans* (bga C gene) also gave considerable amounts of β,1–6-linked disaccharides [30].

2.2 α-Fucosidases

Fucosylated oligosaccharide structures associated with an array of biological recognition events represent important synthetic targets. α-L-Fucoside-fucohydrolases (EC 3.2.1.51) were isolated from prokaryotic and eukaryotic sources, namely bacteria and moulds, mollusks, plants, and mammals with different substrate specificities. Whereas many α-fucosidases are useful in oligosaccharide analysis only some of them could be used for synthetic purposes. Transglycosylation activities were observed by α-fucosidases from porcine liver [53,54], bovine kidney [55] and bovine testes [56], from *Aspergillus niger* [57], *Corynebacterium* sp., Ampullaria, and *Penicillium multicolor* [58,59], and from *Alcaligenes* sp. [54]. Recently, the enzyme from *P. multicolor* could be used to synthesize a disaccharide portion of Lewis x [60] employing either α-L-fucopyranosyl fluoride (**32**, X = F) or fucopyranoside (**32**, X = O*p*NP) as donor substrates. With α- and β-methyl or allyl glycosides of both *N*-acetyl-glucosamine (**33**, Y = NHAc) and glucose (**33**, Y = OH) the corresponding α,1–3-linked derivatives **34** (Y = NHAC, OH, R = α-OCH_3, α-*O*-All, β-OAll) were obtained in 10–40% yield (*Scheme 10*).



Scheme 10

2.3 β-Mannosidases

The conventional synthesis of 1,2-*cis*-configured glycosides is considerably demanding, and thus alternative approaches to β-mannopyranosides would be of interest. Whereas there are several reports on β-mannosidases (EC 3.2.1.25) with regard to their biochemistry and purification [e. g., 61], only β-mannosidase from guinea pig liver was used in transglycosylation to give disaccharides in low yields [62]. Later studies with a β-mannosidase from snail viscera and *p*NP β-D-Man*p* as donor substrates also gave minor yields of β,1–3- and β,1–4-linked mannobiosides [63,64]. Noteworthy may be the advantageous application of the complex of glycosidases from *Aspergillus oryzae* which contains a number of glycosidases (β-galactosidase, β-glucosidase, β-mannosidase, β-*N*-acetyl-glucosaminidase, β-*N*-acetyl-galactosaminidase) [65,66] for the preparation of β-mannosylated derivative [63,64].

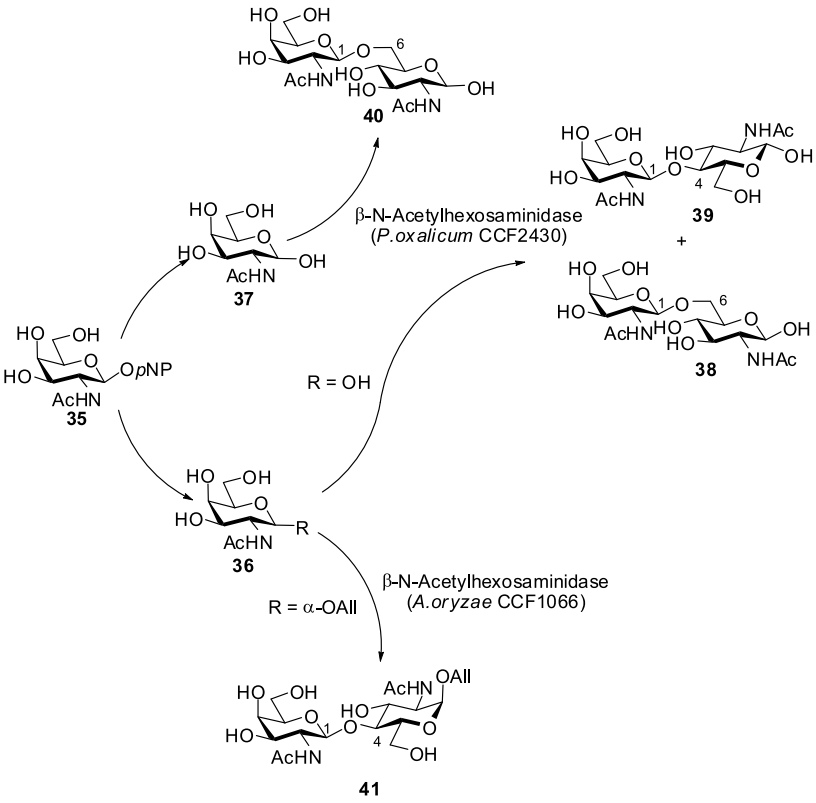
2.4 β-*N*-Acetylhexosaminidases

These enzymes (EC 3.2.1.52) show affinity for both β-GlcNAc and β-GalNAc moieties and thus both are hydrolyzed. Early reports proved them to transfer GlcNAc [67,68,69,70] and GalNAc [69,70] units to glucose-terminated oligosaccharides giving β,1–4- as well as β,1–6-linkages. Also, the transfer of GalNAc onto a β-GalNAc unit afforded GalNAcβ1–6-GalNAc [71]. Employing a β-*N*-acetylhexosaminidase from *P. oxalicum* CCF 2430 and *p*NPGalNAc (**35**) on *N*-acetyl-glucosamine (**36**, R = OH) gave both the regioisomeric glycosylation products GalNAcβ1–4GlcNAc (**38**) and GalNAcβ1–6GlcNAc (**39**) in 27% and 19% yield, respectively. With *N*-acetylgalactosamine (**37**), however, the same reaction gave exclusively the GalNAcβ1–6GalNAc disaccharide **40** in 87% yield. Further experiments used the same donor **35** with α-allyl *N*-acetyl-glucosaminide (**36**, R = α-OAll) and β-*N*-acetylhexosaminidase from *A. oryzae* OCF 1066 which gave only the disaccharide GalNAcβ1–4GlcNAcαOAll (**41**) in 78% yield [72] (*Scheme 11*).

In addition to transglycosylation sometimes selective degradation of oligosaccharides may be the method of choice to arrive at complex target structures. A recent convincing example started with an asparagine-linked biantennary complex-type sialylundecasaccharide **42** isolated from egg yolk. Mild acid hydrolysis gave a mixture of Asn-oligosaccharides which were FMOC-protected at Asn to allow for HPLC purification. Successive treatment with β-galactosidase and β-*N*-acetylhexosaminidase gave a collection of *N*-linked glycans such as the penta- (**45**) and the isomeric octasaccharides **43** and **44** which could be further used in solid-phase glycopeptides synthesis [73] (*Scheme 12*).

2.5 α-Sialidases

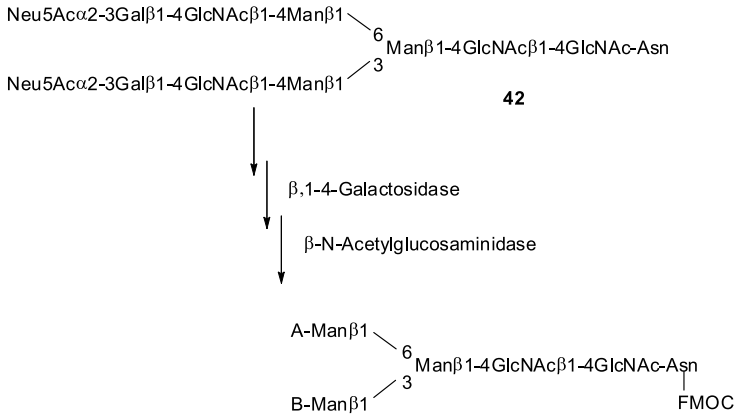
Neuraminic acid (Neu5Ac) is the major constituent in a variety of glycoconjugates such as glycoproteins, gangliosides, and oligosaccharides occurring in higher animals, viruses, bacteria, protozoa, and pathogenic fungi. The crucial role of sialylation in regulation of cellular and molecular recognition in biological systems associated with the variation in the glycosylation pattern of cell surface glycoconjugates led to particular interest in the synthesis of (mostly terminally) sialylated components. In addition to many studies using the difficult



Scheme 11

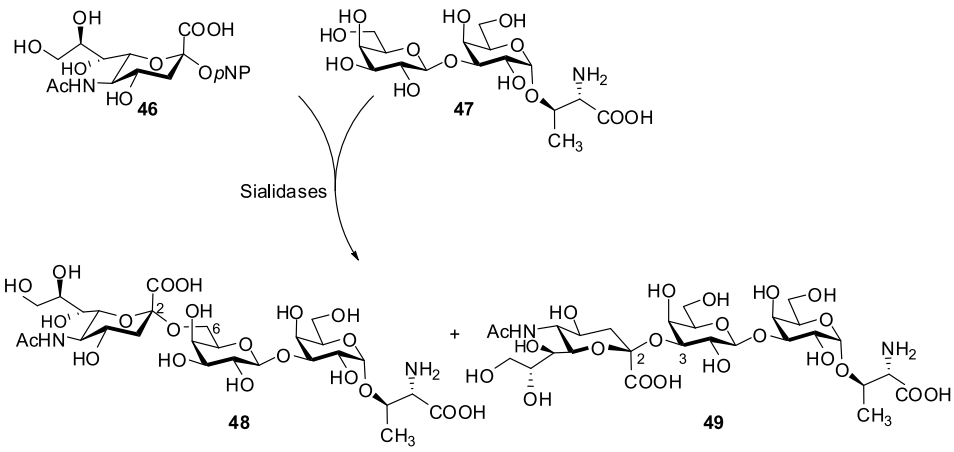
to access and handle sialyltransferases (cf. e. g. [25] for a combined hydrolase-transferase approach) there are some reports employing sialidases [74,75,76]. A comprehensive study revealed the synthetic potential of various sialidases in the preparation of sialylated Thomsen–Friedenreich (T) antigen structures [77]. Treatment of Neu5Ac*p*NP (**46**) with the α-threonin disaccharide glycoside **47** and several sialidases gave both the regioisomeric α,2–6- (**48**) and the α,2–3-linked trisaccharide derivatives (**49**) in varying amounts. Application of sialidases from *Vibrio cholerea* and *Chlostridium perfringens* gave virtually exclusively (> 99%) the α,2–6-amino acid trisaccharide **48**. On the other hand sialidase from *Salmonella typhimurium* (86–99%) and Newcastle disease virus sialidase (> 99%) showed almost exclusively the formation of the α,2–3- sialylated structure **49**. Generally, the yields ranged from about 10–25% which required subsequent selective hydrolytic glycosidase cleavages of the non-desired side product and separations (*Scheme 13*).

An extension of this method could be shown by employing MPEG polymer-supported acceptor glycoside structures. Again these could be sialylated with **46** in solution-phase giving about 15–25% yield. These sialylated polymer-bond components did not undergo glycosidic cleavage, however, the excess donor substrates did on subsequent treatment with β-galactosidase



43 A = Neu5Acα2-3Galβ1-4GlcNAcβ1-4, B = 0
44 B = Neu5Acα2-3Galβ1-4GlcNAcβ1-4, A = 0
45 A = B = 0

Scheme 12



Scheme 13

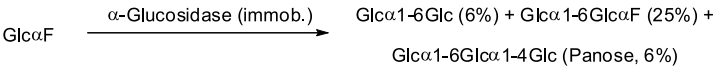
(*E. coli*) and *β-N*-acetylhexosaminidase (*A. oryzae*), and thus a facile separation and purification approach could be established [78].

2.6 Further Glycosidases

Biochemical studies constantly report and characterize novel glycosidases, and with time their fortunate stereochemistry and their easy access become attractive for chemoenzymat-

ic work. Well established are α-mannosylations employing α-mannosidase from jack bean. With *p*-nitrophenyl α-D-mannopyranoside ergot alkaloids could be successfully α-mannosylated [79,80]. The recently reported α-mannosidase from *Candida albicans* [81] still awaits synthetic application. Also α-L-rhamnosidases could be used in rhamnosylation of deoxy mannojirimycin [82]. A Galactosylation and change of the regioselectivity of the original enzyme was reported [83].

Early on [84,85] it could be demonstrated that glycosyl fluorides were excellent donor substrates for glycosidases [86]. In an example immobilized α-glucosidase (EC 3.2.1.20) could be used for self transfer of α-glucopyranosyl fluoride to give isomaltose (6%), α-isomaltosyl fluoride (25%), and the trisaccharide panose (6%) [87] (● *Scheme 14*).



● Scheme 14

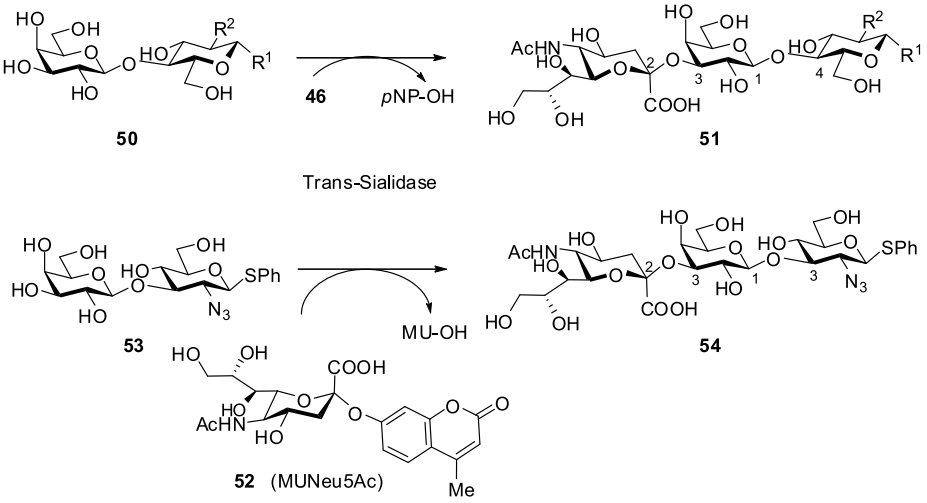
Besides glycosidases recognizing hexopyranose structures recently there were reports on the chemoenzymatic use of an α-L-arabinofuranosidase which could be also used to synthesize β-D-galacto-furanosides [88].

A quite interesting area with considerable application potential opens up with enzymes obtained from extremophiles [89]. First reports on properties and functions described a novel α-glucosidase from the acidophilic archaeon *Ferroplasma acidiphilium* strain Y [90] and two β-glucosidases from a thermo-tolerant yeast *Pichia etchellsii* [91].

There are a number of valuable reports on the use of further glycosidases in chemoenzymatic studies which due to spatial restrictions had to be omitted from this review.

3 *trans*-Sialidase

The pathogen responsible for Chagas disease is *Trypanosoma cruzi*. This upon infection expresses a *trans*-sialidase (TSia) that catalyzes transfer of terminal Neu5Ac of the host's sialoglycoconjugates onto its own surface mucins and thus masks its own epitopes to prevent degradation by the host's natural killer cells [92,93,94]. On the basis of its primary structure TSia belongs to the glycosidase family 33 which also includes multiple bacterial sialidases [95]; kinetic and crystallographic studies of the closely related TSia from *T. brucei* provided information on the binding pocket and the mechanism [96,97,98]. The easy handling of TSia and the excellent stereoselectivity in transferring Neu5Ac from various donors in α,2–3-linkages to galactose-terminated oligosaccharides was recognized and synthetically used [99,100,101,102,103]. Some recent extensions of this approach allowed us to synthesize an array of sialylated structures associated with T-antigen determinant components as well as Neu5Ac-terminated oligosaccharide motifs for binding of the myelin-associated glycoprotein (MAG) [104,105,106]. Treatment of various lactose and *N*-acetyllactosamine structures **50** (R^1 = SPh, R^2 = N_3; R^1 = R^2 = OH; R^1 = OAll, R^2 = OH; R^1 = OMe, R^2 = NHAc) with *p*-nitrophenyl sialylate **46** and *trans*-sialidase gave the corresponding trisaccharides **51** in 30–60% yield. Correspondingly, the β,1–3-linked isolactosamine precursor **53** could be reacted with muraminyl sialylate **52** to give **54** in 30% yield [107] (● *Scheme 15*).

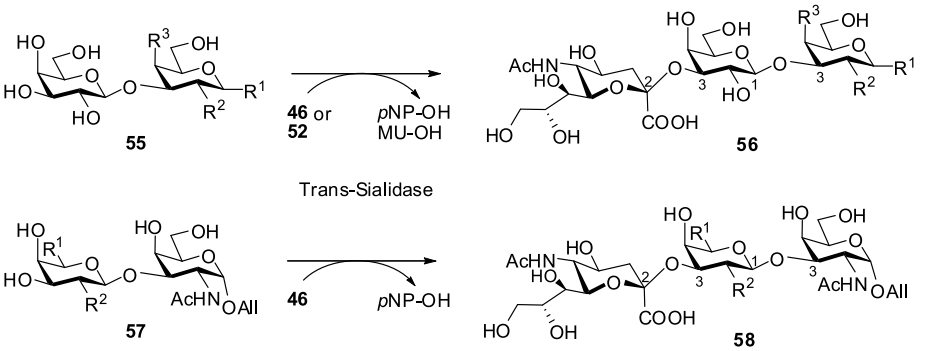


50 51 52 (MUNeu5Ac) 53 54

Scheme 15

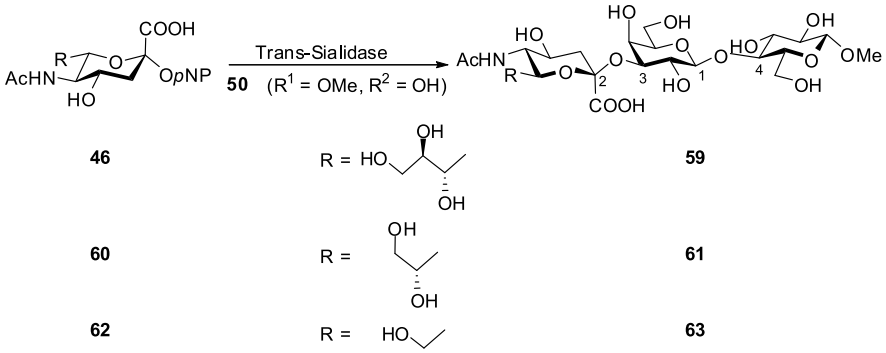
Further transformations were done with β1–3-galactobiose structure **55**. Here Galβ1–3-GalNAcαThr and Galβ1–3GalNAcαSer (**55**, $R^1 = \alpha$Thr or αSer, $R^2 = NHAc$, $R^3 = OH$) could be used as acceptor structures and also 2-deoxy ($R^2 = H$) and 2-azido ($R^2 = N_3$) were α,2–3-sialylated with TSia and **46** or **52** as donors to give the modified trisaccharides **56** in 30–50% yield [105,106]. In the LacNAc derivatives **57** the terminal galactose unit which directly accepts the sialyl residue, could be modified. 2-Deoxy-galactose-(**57**, $R^1 = H$, $R^2 = CH_2OH$) and also the D-fucose-(**57**, $R^1 = OH$, $R^2 = CH_3$) and the L-arabinose-(**57**, $R^1 = OH$, $R^2 = H$) terminated disaccharides were nicely sialylated to give the corresponding derivatives **58** (30–50% yield) [106] (*Scheme 16*).

In recent studies modified donor substrates were obtained from *p*NPNeu5Ac **46** by selective oxidative cleavage. Dependant on the ratio of compound **46**:oxidant the octose (**60**,



55 56 57 58

Scheme 16

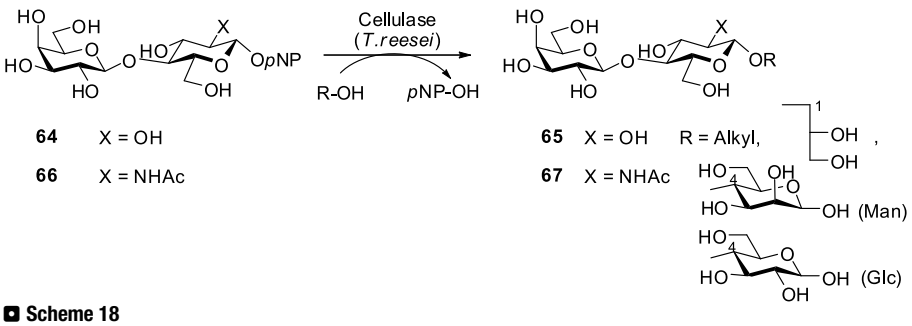


Scheme 17

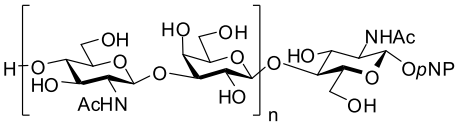
*p*NPOct5Ac) or the heptose (**62**, *p*NPHep5Ac) donor substrates were obtained in high yields. Surprisingly, their transfer onto methyl β-lactoside (**50**) under TSia catalysis did not show any problems; in fact, the yields of the octose (**61**) or heptose analog of neuraminic acid carrying trisaccharides (**63**) were identical to the original derivative **59**, that is 55–65% yield. Apparently, *trans*-sialidase will easily recognize these modified structures. In contrast, STD-NMR studies showed that only the original trisaccharide **59** is a well-binding substrate for MAG and the analogous structures **61** and **63** are not [107] (*Scheme 17*).

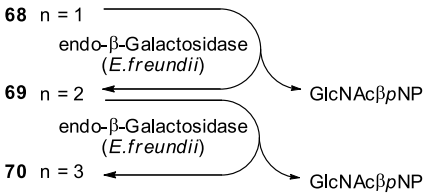
4 Endo-Glycosidases

In oligosaccharides or glycoconjugates internal glycosidic linkages are cleaved by endo-glycosidases, and under kinetic control their transglycosylation activity is favored. Some of them could be employed in preparative approaches [108,109,110], and recently the commercially available cellulas from *Trichoderma reesei* was shown to transfer entire lactose or lactosamine units employing the corresponding *p*-nitrophenyl β-glycosides **64** or **65**. With aliphatic alcohols as well as lower di- and triols yields of **65** and **67** in the range of 1–20% were obtained. However, with mannose or glucose LacNAcβ*p*NP (**66**) gave the corresponding β1–4-linked trisaccharides **67** in 13 and 9% yield, respectively [111] (*Scheme 18*).

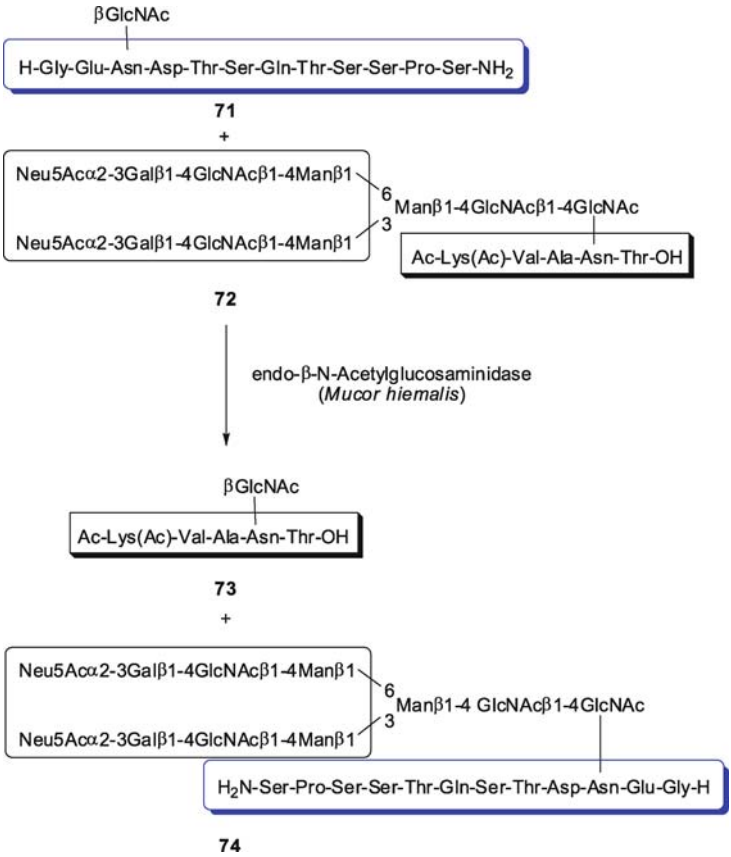


Scheme 18





Scheme 19



Scheme 20

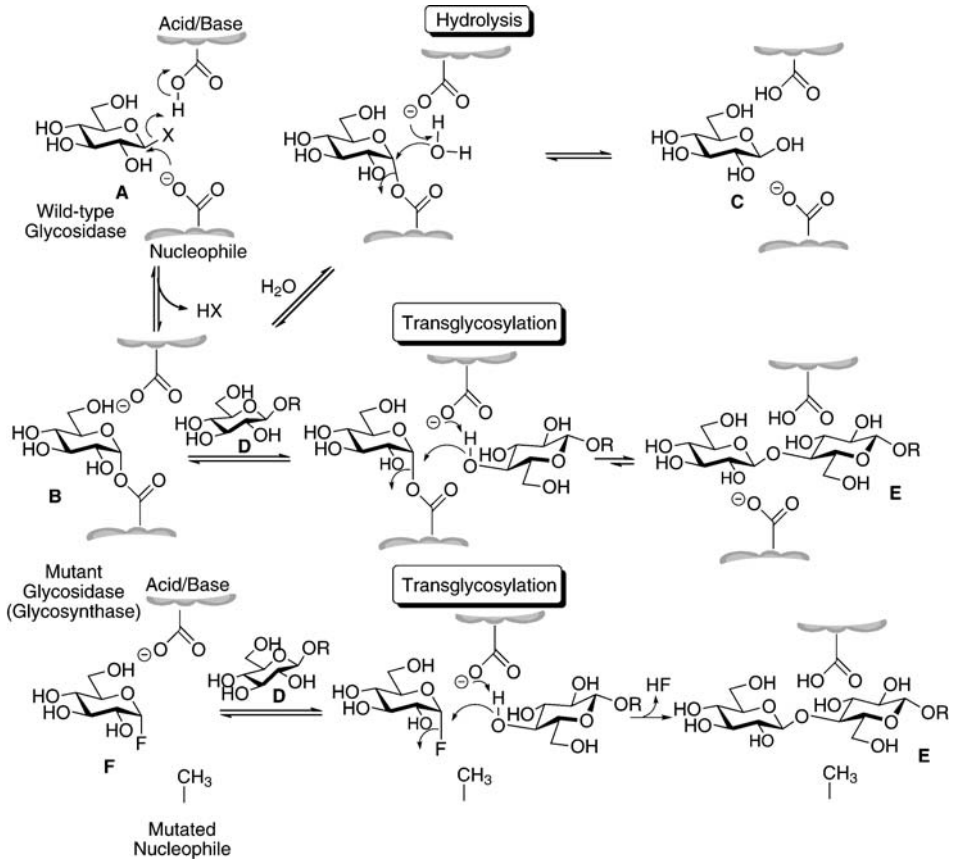
An interesting self-transglycosylation could be observed with endo β-galactosidase from *E. freudii* employing the *p*-nitrophenyl β-glycoside of the trisaccharide **68** (n = 1) in high concentration. Chain elongations to give regio- and stereoselectively the pentasaccharide **69** (n = 2) and also the heptasaccharide **70** (n = 3) in about 10% were observed with GlcNAcβ*p*NP as the leaving group [112,113] (● *Scheme 19*).

Endo-β-*N*-acetylglucosaminidase from *Mucor hiemalis* (endo-M) shows transglycosylation activity with a number of different substrates. It could be used to transfer high-mannose, hybrid-type, as well as complex-type oligosaccharides and thus glycopeptides became available [114]. Employing this enzyme for the transfer of a complex-type oligosaccharide chain in **72** from the GlcNAc-CD52 (**71**) gave the GlcNAc peptide **73** and the complex-type CD52 target structure **74** [115] (● *Scheme 20*).

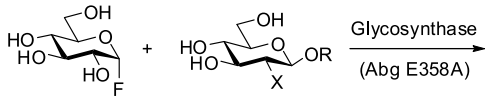
5 Glycosynthases

A quantum leap in using glycosidases for preparative oligosaccharide synthesis came along with the introduction of glycosynthases by Withers et al. [116]. As elaborated for retaining glycosidases the problem of product hydrolysis in transglycosylation could be solved by bioengineering of suitable mutants. Thus, the active-site catalytic nucleophile was removed and replaced by a non-nucleophilic amino acid side chain. As depicted in ● *Scheme 21* hydrolysis and transglycosylation for a wild-type retaining β-glycosidase (in this depiction a retaining β-glucosidase) led to hydrolysis and transglycosylation products. It is generally accepted that the donor substrate **A** reacts with the nucleophilic amino acid to give the glycosyl-enzyme intermediate **B**. This on treatment with water involving a proton transfer to the acid–base amino acid and subsequent release of the nucleophilic amino acid yields the hydrolysis product **C**. In transglycosylation the acceptor substrate **D** is transformed correspondingly to give the disaccharide product **E**. Mutant glycosidases fold apparently in the same way as the wild-type enzymes, however, devoid of a nucleophilic site they cannot form the intermediate **B**, and thus hydrolysis does not occur. However, if a donor substrate with opposite anomeric configuration is offered like in the glycosyl fluoride **F** with a small leaving group the mutant glycosidase, the glycosynthase, will recognize it as an analogue of the covalent glycosyl-enzyme intermediate **B**. Thus, employing the same transglycosylation mechanism as for the wild-type enzyme the acceptor sugar is transferred to give again the disaccharide **E**. Since these mutant enzymes cannot display hydrolytic activity the yields are generally excellent (● *Scheme 21*).

Within recent years a considerable number of glycosynthases have been described both from exo- as well as from endo-glycosidase and the subject was also reviewed several times [114,117,118,119,120]. In the original work a well-known β-glucosidase Abg from *Agrobacterium* sp. was studied. Mutation of the nucleophilic amino acid E358 from glutamate to alanine could be achieved. Condensation of α-glucosyl fluoride (**75**) with several acceptor glucopyranosides such as the *o*-nitrophenyl β-glucopyranoside **76** or the 2,4-dinitrophenyl β-2-deoxy-2-fluoro-glucopyranoside (**77**) gave transglycosylation yields of di- to tetrasaccharides **78a**–**c** and **79a**–**c**, respectively, in 76–84% overall yield [116] (● *Scheme 22*). Another glycosynthase could be employed by the same group [121] and transgalactosylation was effected with galactopyranosyl fluoride (**80**) to the PGA-resin-bound glucopentapeptide **81** to give cleanly the corresponding resin-bound lactosylpentapeptide **82**. In this study

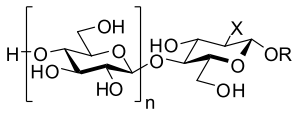


Scheme 21



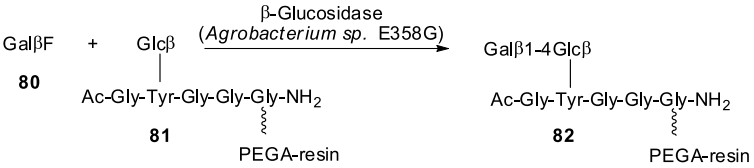
75

76 X = OH, R = *o*NP
77 X = F, R = *op*N_2P



78a-c n = 1-3, X = OH, R = *o*NP
79a-c n = 1-3, X = F, R = *op*N_2P

Scheme 22



Scheme 23

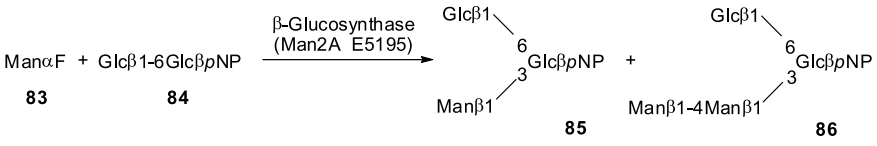
the use of solid phase PEGA resin bound acceptor structures in glycopeptide synthesis with glycosynthases could be demonstrated (*Scheme 23*).

As mentioned above β-mannosylation is an often realized structural feature in natural glycoconjugates and oligosaccharides, yet the chemical formation of β-mannopyranosides remains a challenging issue. Consequently, β-mannosyl-transferases were successfully employed, but too demanding for preparative purposes. In contrast, β-mannosylations using β-mannosidases could be performed only in poor yields. Now the present approach demonstrated convincingly the mutation of a retaining β-mannosidase from *C. fini* (Man2A) [122]. This mannosynthase catalyzed the β-mannosylation of several acceptor structures. Treatment of β-mannopyranosyl fluoride (**83**) with the *p*-nitrophenyl β-glycoside of gentiobiose (Glcβ1–6GlcβpNP) catalyzed by this glycosynthase afforded a mixture of tri- (**85**) and tetrasaccharides (**86**) with mono- as well as dimannosylation in 34% **85**) and 49% (**86**) yield, respectively. Whereas the first β-mannosylation in both products goes to the 3-position of the gentiobioside, the second mannosylation occurs at the site of the *manno*-unit and results in a Manβ1–4 Man substructure as in **86** (*Scheme 24*).

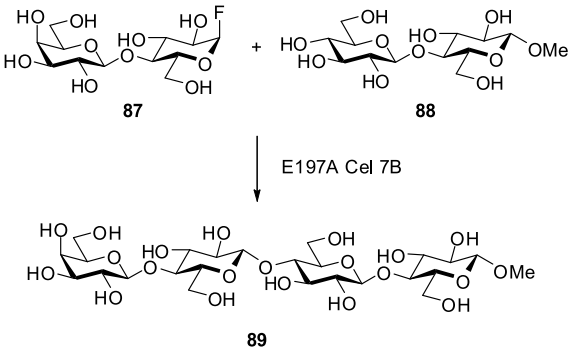
Glycosynthases could also be obtained from retaining endo-glycosidases. The first one was the endo-acting 1,3–1,4-β-glucanase of *Bacillus licheniformis*, in which the nucleophilic amino acid E134 was substituted with alanine. This enzyme could accept laminaribiosyl fluoride (Glcβ1–3GlcαF) as the donor substrate and glycosylate glucosides, cellobiosides as well as laminaribiosides in excellent yields [123].

The endo-glycosynthase obtained from the E197A mutant of the retaining cellulose Cel7B of *Humicola insolens* could be employed to transfer α-cellobiosyl fluoride (**87**) to methyl β-cellobioside (**88**) to afford the cellotetraoside **89** in virtually quantitative yield [124] (*Scheme 25*).

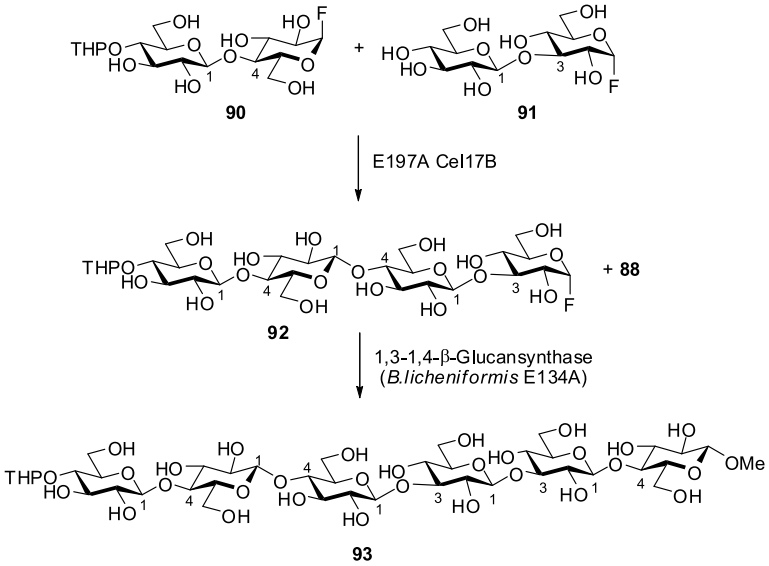
Another quite impressive example was reported by collaboration of the same Spanish-French scientists. In a tandem-type glycosylation using the aforementioned endo-glycosynthases, the α-cellobiosyl fluoride with a 4′-tetrahydropyran protection **90** was transglycosylated to the acceptor disaccharide α-laminaribiosyl fluoride **91** employing the mutant cellulose to give the tetrasaccharide **92**. Its further use as a donor structure with the mutant 1,3–1,4-β-glucanase and the acceptor cellobioside **88** gave the β,1–4- and β,1–3-linked hexasaccharide glucan **93** stere-



Scheme 24



Scheme 25



Scheme 26

oselectively and in excellent yield [125] (*Scheme 26*). An excellent resource for structural data of glycohydrolases as well as glycosytransferases and other carbohydrate active enzymes can be found in the database of Carbohydrate-Active enZymes [CAZy].

References

1. Nilsson KGI (1988) Trends Biotechnol 6:256
2. Monsan P, Paul F, Remaud M, Lopez A (1989) Food Biotechnol 3:11
3. Nilsson KGI (1996) In: Khan SH, O'Neill RA (eds) Modern Methods in Carbohydrate Synthesis. Harwood, Amsterdam, chap 21
4. Wong CH, Halcomb RL, Ichikawa Y, Kajimoto T (1995) Angew Chem Int Ed Engl 34:412, 521
5. Wong CH, Whitesides GM (1994) In: Baldwin JE, Magnus PD (eds) Tetrahedron Org Chem Ser, vol. 12. Pergamon Press, Oxford

6. Shoda SI (2001) In: Fraser-Reid BO, Tatsuta K, Thiem J (eds) Glycoscience. Springer, Berlin Heidelberg New York, p 1465
7. Vocadlo DJ, Withers SG (2000) In: Ernst B, Hart GW, Sinay P (eds) Carbohydrates in Chemistry and Biology. Wiley-VCH, Weinheim, 2: 723
8. Ajisaka K (2001) Trends Glycosc Glycotechnol 13:305
9. Ajisaka K, Yamamoto Y (2002) Trends Glycosc Glycotechnol 14:1
10. Thiem J (1995) FEMS Microbiol Rev 16:193
11. Kren V, Thiem J (1997) Chem Soc Rev 26:463
12. Wong CH (2002) In: Drauz K, Waldmann H (eds) Enzyme Catalysis in Organic Synthesis. Wiley-VCH, Weinheim, 2:609
13. Kornfeld R, Kornfeld S (1985) Annu Rev Biochem 54:631
14. Crout DGH, MacManus DA, Critchley P (1990) J Chem Soc, Perkin Trans 1:1865
15. Ooi Y, Hashimoto T, Mitsuo N, Satoh T (1985) Chem Pharm Bull 33:1808
16. Attal S, Bay S, Cantacuzene D (1993) Tetrahedron 48:9251
17. Kren V (1997) Top Curr Chem 186:45
18. Rajnochová E, Dvoráková J, Hunková Z, Kren V (1997) Biotechnol Lett 19:869
19. Okahata Y, Mori T (1996) J Chem Soc, Perkin Trans 1:2861
20. Nilsson KGI (1987) Carbohydr Res 167:95
21. Nilsson KGI (1988) Carbohydr Res 180:53
22. Boos W, Lehmann J, Wallenfels K (1968) Carbohydr Res 7:381
23. Björkling FB, Gotfredsen SE (1988) Tetrahedron 44:2957
24. Werschkun B, König WA, Kren V, Thiem J (1995) J Chem Soc Perkin Trans 1:2459
25. Kren V, Thiem J (1995) Angew Chem Int Ed Engl 34:893
26. Hedbys L, Johansson E, Morbach K, Larsson PO (1989) Carbohydr Res 186:217
27. Johansson E, Hedbys L, Larsson PO (1991) Enzyme Microb Technol 13:781
28. Gambert U, Thiem J (1997) Carbohydr Res 299:85
29. Gambert U, Gonzales Lio R, Farkas E, Thiem J, Verez Bencomo V, Lipták A (1997) Bioorg Med Chem 5:1285
30. Kröger L, Thiem J (2005) J Carbohydr Chem 24:717
31. Kröger L, Scudlo A, Thiem J (2006) Adv Synth Catal 348:1217
32. Fujimoto H (1997) J Carbohydr Chem 16:967
33. Fujimoto H, Miyasato M, Ito Y, Sasaki T, Ajisaka K (1998) Glyconj J 15:155
34. Gambert U, Conradt HS, Nimtz M, Thiem J (2000) J Carbohydr Chem 19:621
35. Schröder S, Schmidt U, Thiem J, Kowalczyk J, Kunz M, Vogel M (2004) Tetrahedron 60:2601
36. Kröger L, Thiem J (2007) Carbohydr Res 342:467
37. Sakai K, Katsumi H, Ohi H, Usui T, Ishido Y (1992) J Carbohydr Chem 11:553
38. Daiwa Kasei KK 7–12, Uehonmachi 5-chome Tennoji-ku Osaka 543
39. Priya K, Loganathan D (1999) Tetrahedron 55:1119
40. Vic G, Hastings JJ, Howarth OW, Crout DHG (1996) Tetrahedron: Asymm 7:709
41. Takayama S, Shimazaki M, Qiao L, Wong CH (1996) Bioorg Med Chem Lett 6:1123
42. Farkas E, Thiem J (1999) Eur J Org Chem 3073
43. Farkas E, Schmidt U, Thiem J, Kowalczyk J, Kunz M, Vogel M (2003) Synthesis 699
44. Usui T, Morimoto S, Hayakawa Y, Kawaguchi M, Murata T, Matahira Y, Nishida Y (1996) Carbohydr Res 285:29
45. Usui T, Kubota S, Ohi H (1993) Carbohydr Res 244:315
46. Herrmann GF, Ichikawa Y, Wandrey C, Gaeta FCA, Paulson JC, Wong CH (1993) Tetrahedron Lett 34:3091
47. Fang J, Xie W, Li J, Wang PG (1998) Tetrahedron Lett 39:919
48. Komba S, Ito Y (2002) Can J Chem 80:1174
49. Hart J, Falshaw A, Farkas E, Thiem J (2001) Synlett 329
50. Weingarten S, Thiem J (2004) Org Biomol Chem 2:961
51. Neumann J, Weingarten S, Thiem J (2007) Eur J Org Chem 1130
52. Kimura T, Takayama S, Huang H, Wong CH (1996) Angew Chem Int Ed Engl 35:2348
53. Korf U, Thiem J (1992) Kontakte (Darmstadt) 1:3
54. Murata T, Morimoto S, Zeng X, Watanabe S, Usui T (1999) Carbohydr Res 320:192
55. Nilsson KGI, Eliasson U, Larsson-Lorek U (1995) Biotechnol Lett 17:717
56. Nilsson KGI, Pan H, Larsson-Lorek U (1997) J Carbohydr Chem 16:459

57. Vetere A, Galateo C, Paoletti S (1997) Biochem Biophys Res Commun 234:358
58. Ajisaka K, Shikaraba M (1992) Carbohydr Res 224:291
59. Ayisaka K, Fujimoto H, Miyasato M (1998) Carbohydr Res 309:125
60. Farkas E, Thiem J, Ajisaka K (2000) Carbohydr Res 328:293
61. McCleary BV (1982) Carbohydr Res 101:75
62. Kyosaka S, Murata S, Tsuda Y, Tanaka M (1986) Chem Pharm Bull 34:5140
63. Taubken N, Sauerbrei B, Thiem J (1993) J Carbohydr Chem 12:651
64. Taubken N, Thiem J (1998) Glycoconj J 15:757
65. Crout DHG, Singh S, Swoboda BEP, Critchley P, Gibson WT (1992) J Chem Soc, Chem Commun 1992
66. Kren V, Scigelova M, Prikrylova V, Havlicek V, Sedmera P (1994) Biocatalysis 10:181
67. Kubisch J, Weignerova L, Kötter S, Lindhorst TK, Sedmera P, Kren V (1999) J Carbohydr Chem 18:975
68. Rajnochova E, Dvorakova J, Hunkova Z, Kren V (1997) Biotechnol Lett 19:869
69. Singh S, Packwood J, Samuel CJ, Critchley P, Crout DHG (1995) Carbohydr Res 279: 293
70. Singh S, Scigelova M, Critchley P, Crout DHG (1998) Carbohydr Res 305:363
71. Singh S, Scigelova M, Vic G, Crout DHG (1996) J Chem Soc, Perkin Trans 1:1921
72. Weignerova L, Vavruskova P, Pisvejcova A, Thiem J, Kren V (2003) Carbohydr Res 338:1003
73. Kayihara Y, Suzuki Y, Yamamoto N, Sasaki K, Sakakibara T, Juneja LR (2004) Chem Eur J 10:971
74. Thiem J, Sauerbrei B (1991) Angew Chem Int Ed Engl 30:1503
75. Ajisaka K, Fujimoto H, Isomura M (1994) Carbohydr Res 259:103
76. Makimura Y, Ishida H, Kondo A, Hasegawa A, Kiso M (1998) J Carbohydr Chem 17:975
77. Schmidt D, Sauerbrei B, Thiem J (2000) J Org Chem 65:8518
78. Schmidt D, Thiem J (2000) Chem Comm 1919
79. Scigelova M, Kren V, Nilsson KGI (1994) Biotechnol Lett 16:683
80. Kren V (1997) Top Curr Chem 186:45
81. Mora-Montes HM, Lopez-Romero E, Zinker S, Ponce-Noyola P, Flores-Carreon A (2004) Glycobiol 14:593
82. Bannwell MG, Ma X, Asano N, Ikedra K, Lambert JN (2003) Org Biomol Chem 1:2035
83. Dion M, Nisole A, Spangenberg P, Andre C, Glottin-Fleury A, Mattes R, Tellier C, Rabiller C (2001) Glycoconj J 18:215
84. Barnett JEG, Jarvis WTS, Munday KA (1967) Biochem J 105:669
85. Hehre EJ, Genghof DS (1971) Archiv Biochem Biophys 142:382
86. Williams SJ, Withers SG (2000) Carbohydr Res 327:27
87. Schlingmann M, Keller R, Wiesner M, Treder W, Thiem J (1990) Eur Pat Appl EP 0298438, Chem Abstr 112:75339g
88. Remond C, Plantier-Royon R, Aubry N, O'Donohue MJ (2005) Carbohydr Res 340: 637
89. Vorgias CE, Antranikian G (2000) Glycomicrobiol 313
90. Ferrer M, Golyshina OV, Plou FJ, Timmis KN, Golyshin PN (2005) Biochem J 391:269
91. Wallecha A, Mishra S (2003) Biochim Biophys Acta 1646:74
92. Schenkman S, Jiang MS, Hart GW, Nussenzweig V (1991) Cell 65:1117
93. Pontes de Carvalho LC, Tomlinson S, Vandekerckhove F, Bienen EJ, Clarkson AB, Hart GW, Nussenzweig V (1993) J Exp Med 177:465
94. Schenkman S, Eichinger D, Pereira ME, Nussenzweig V (1994) Annu Rev Microbiol 48:499
95. Henrissat B (1998) Biochem Soc Tans 26:153
96. Buschiazzo A, Tavares GA, Campetella O, Spinelli S, Cremona ML, Paris G, Amaya MF, Frasch AC, Alzari PM (2000) Embo J 19:16
97. Buschiazzo A, Cremona ML, Paris G, Amaya MF, Frasch AC, Alzari PM (2002) Mol Cell 10:752
98. Watts AG, Damager I, Amaya MF, Buschiazzo A, Alzari PM, Frasch AC, Withers SG (2003) J Am Chem Soc 125:7532
99. Sabesan S, Paulson JC (1986) J Am Chem Soc 108:2068
100. Vandekerckhove F, Schenkman S, Pontes de Carvalho L, Tomlinson M, Kiso M, Yoshida M, Hasegawa A, Nussenzweig V (1992) Glycobiol 2:541
101. Turnbull WB, Harrison JA, Kartha KPR, Schenkman S, Field RA (2002) Tetrahedron 58:3207
102. Agusti R, Mendoza VM, Gallo-Rodriguez C, de Lederkremer RM (2005) Tetrahedron Assym 16:541

103. Mendoza VM, Agusti R, Gallo-Rodriguez C, de Lederkremer RM (2006) Carbohydr Res 341:1488
104. Neubacher B, Thiem J (2005) Glycoconj J 22:203
105. Neubacher B, Schmidt D, Ziegelmüller P, Thiem J (2005) Org Biomol Chem 3:1551
106. Kröger L, Scudlo A, Thiem J (2006) Adv Synth Catal 348:1217
107. Neubacher B, Scheid S, Kelm S, Frasch AC, Meyer B, Thiem J (2006) ChemBioChem 7: 896
108. Yamamoto K, Fujimori K, Hameda K, Mizuno M, Inazu T, Kumagai H (1997) Carbohydr Res 305:415
109. Ashida H, Yamamoto K, Kunagai H (2001) Carbohydr Res 330:487
110. Akaike E, Tsutsumida M, Osumi K, Fujita M, Yamanoi T, Yamamoto K, Fujita K (2004) Carbohydr Res 339:719
111. Totani K, Yasutake N, Ohi H, Murata T, Usui T (2001) Arch Biochem Biophys 385:70
112. Murata T, Hattori T, Amarume S, Koichi A, Usui T (2003) Eur J Biochem 270:3709
113. Murata T, Honda H, Hattori T, Usui T (2005) Biochim Biophys Acta 1722:60
114. Thayer DA, Wong CH (2007) Top Curr Chem 267:37
115. Li HG, Singh S, Zeng Y, Song HJ, Wang LX (2005) Bioorg Med Chem Lett 15:895
116. Mackenzie LF, Wang Q, Warren RAJ, Withers SG (1998) J Am Chem Soc 120:5583
117. Jakeman DL, Withers SG (2002) Trends Glycosci Glycotechnol 14:13
118. Williams SJ, Withers SG (2002) Aust J Chem 55:3
119. Perugino G, Trincone A, Rossi M, Moracci M (2004) Trends Biotechnol 22:31
120. Perugino G, Cobucci-Ponzano B, Rossi M, Moracci M (2005) Adv Synth Catal 347:941
121. Tolberg JF, Petersen L, Jensen KJ, Mayer C, Jakeman DL, Warren RAJ, Withers SG (2002) J Org Chem 67:4143
122. Nashiru O, Zechel DL, Stoll D, Mohammadzadeh T, Warren RAJ, Withers SG (2001) Angew Chem Int Ed 40:147
123. Malet C, Planas A (1998) FEBS Lett 440:208
124. Fort S, Boyer V, Greffe L, Davies GL, Moroz O, Christiansen L, Schulein M, Cottaz S, Driguez H (2000) J Am Chem Soc 122:5429
125. Faijes M, Fairwether JK, Driguez H, Planas A (2001) Chem Eur J 7:4651

Part 6
Complex Polysaccharides

6.1 Polysaccharides: Occurrence, Significance, and Properties

James N. BeMiller
Department of Food Science, Whistler Center for Carbohydrate Research, Purdue University, West Lafayette, IN 47907–2009, USA
bemiller@purdue.edu

Abstract

Polysaccharides are present in all living organisms where they carry out one or more of their diverse functions. While there is no specific category or definition of a complex polysaccharide, most are structurally complex. Polysaccharides contain 1–5 different monosaccharide (sugar) units. The different sugar units may have different anomeric configurations and/or be joined by different glycosidic linkages. Polysaccharides may be linear or branched. Branches may be short saccharide units on a linear backbone or the molecule may have a branch-on-

In: *Glycoscience*. Fraser-Reid B, Tatsuta K, Thiem J (eds)
Chapter-DOI 10-1007/978-3-540-30429-6_34:

branch structure; in either case, the branches may be isolated or clustered. Polysaccharides may contain non-carbohydrate groups. Esters or cyclic acetal groups, when present, can be removed by appropriate treatments. All polysaccharides are polydisperse, i. e., are present in a range of molecular weights rather than having a single molecular weight. Most are polymolecular, i. e., differ in fine structure from molecule to molecule. So most polysaccharides can be said to be structurally complex. They may be attached to protein molecules or to other polysaccharide molecules. They are solvated by water. Most dissolve in aqueous systems, especially if they are alkaline. Polysaccharides can be depolymerized by acids and heat, specific enzymes, and high pH systems following oxidation. Their hydroxyl groups can be esterified (acylated), etherified (alkylated), and oxidized. Amino groups can be acylated (and deacylated). Carboxyl groups can be converted into esters, amides, and amines. Structural modification makes the molecules even more complex and polymolecular and, perhaps, polydisperse.

Keywords

Cell wall; Chemical reactions; Energy; Exopolysaccharide; Glycan; Glycosaminoglycan; Physical properties; Polysaccharide; Proteoglycan; Protein-polysaccharide; Structural analysis

Abbreviations

db	dry-weight basis
DMF	dimethylformamide
HPAEC-PAD	high-performance anion-exchange chromatography and pulsed amperometric detection
TEMPO	2,2,6,6-tetramethyl-1-piperidinyloxy

1 Introduction

Polysaccharides [1], as the name implies, are polymers of saccharide units, i. e., polymers whose monomer units are simple aldose and/or ketose sugars (monosaccharides). The average number of monosaccharide units in polysaccharides varies from about 20 (although it is usually at least 100) to perhaps 10 million. The number of different monosaccharide units found in polysaccharides is not known with any certainty; more than 100 have been identified. Most are hexoses, but pentoses are not uncommon; eight- and nine-carbon-atom sugars are less common, but found. The monosaccharide units are in either five-membered (furanosyl) or six-membered (pyranosyl) ring forms, most often the latter. These units are joined together in a head-to-tail fashion by glycosidic linkages (*Fig. 1*).

Polysaccharides may be linear or branched. There are numerous branched structures, including structures with only a few, very long branches; linear structures with short branches regularly spaced, irregularly spaced, or in clusters; and branch-on-branch structures with branches clustered or positioned to produce bush-like structures with or without decoration with short branches [2,3]. Each polysaccharide has one, and only one, reducing end, which is the end terminating in a hemiacetal (or carbonyl) group (*Fig. 1*). Each branch generates an additional non-reducing end, which consists of a glycosyl unit which is attached to another through

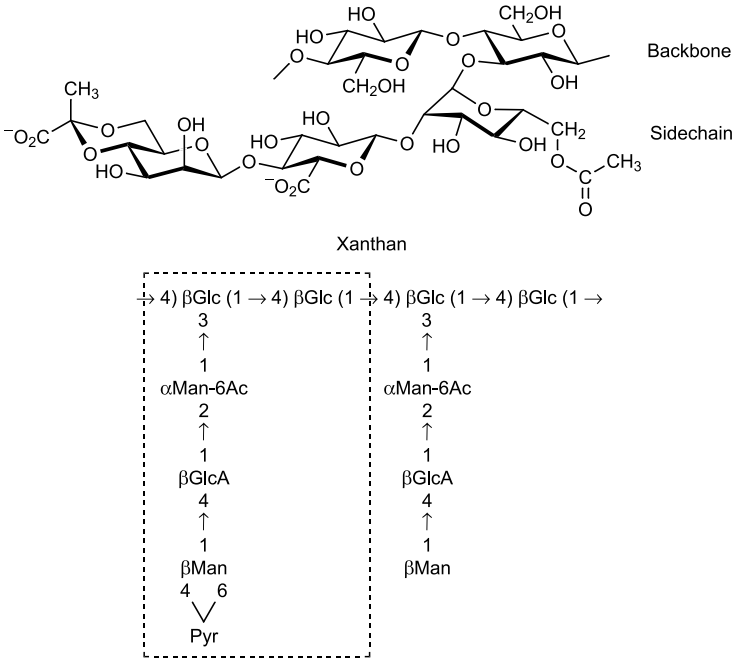


Figure 1

Structure of the pentasaccharide repeating unit of xanthan, a commercial exopolysaccharide produced by *Xanthomonas campestris* (Chap. 6.4). The main chain has the structure of cellulose, i. e., it has a repeating unit of cellobiose. The trisaccharide side chain has, at its non-reducing end, a β-D-mannopyranosyl unit (β-Man*p*) with pyruvic acid attached as a 4,6-cyclic acetal. (About one-half of these terminal mannosyl units in the polysaccharide have these 4,6-*O*-pyruvyl groups.) This non-reducing end unit is linked (1→4) to a β-D-glucuronopyranosyl unit (β-Glc*p*A), which is linked (1→2) to a 6-*O*-acetyl-α-D-mannopyranosyl unit (α-Man*p*6Ac), which in turn is linked (1→3) to a β-D-glucopyranosyl unit of the main chain. The terminal glucosyl unit on the far right of the main chain as the structure is written will have a terminal hydroxyl group, making the structure a hemiacetal. As a result, the ring can easily and rapidly open to an acyclic aldehydo form. This is the reducing end of the polysaccharide. The opposite end on the far left terminates with addition of a hydrogen atom. This is the non-reducing end of the main chain. Inside the box of the shorthand designation (bottom) is a pyruvylated pentasaccharide building block unit which is identical to the upper structure; outside the box is a non-pyruvylated repeating unit. [Reprinted by permission from Whistler RL, BeMiller JN (1997) Carbohydrate chemistry for food scientists. American Association of Cereal Chemists, St. Paul, MN]

its anomeric hydroxyl group but which has no glycosyl unit attached to it, making it a chain end (the terminal unit of the trisaccharide side-chain in *Fig. 1*). Therefore, a polysaccharide molecule may have many non-reducing ends. In terms of mass, linear polysaccharides are the most abundant, because they are structural components of higher plants and marine algae. However, there are many more branched polysaccharides than linear ones. (See Chap. 6.2, Chap. 6.4, and Chap. 6.5 for additional polysaccharides and additional information about polysaccharides not presented in this chapter.)

Polysaccharides may contain, in addition to monosaccharide units, ester, ether, and/or cyclic acetal moieties (*Fig. 2*).

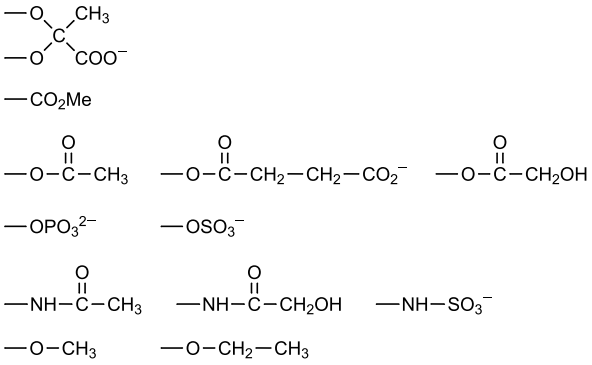


Figure 2
Some on-carbohydrate functional groups that may be part of the structure of a polysaccharide

Ester groups include acetate, glycolate, succinate, sulfate, and phosphate groups on the polysaccharide's hydroxyl groups. Methyl and ethyl ether groups may be present. Pyruvic acid may be present as a cyclic acetal. In those polysaccharides containing amino sugars, the amino groups are usually not free, but present as amido groups, with the acid moiety contributed by acetic, glycolic, or sulfuric acid. In polysaccharides containing uronic acid units (sugar units with a carboxylic acid group in place of the hydroxymethyl group), the carboxylic acid functions may be present as methyl esters.

Glycan is the formal generic term for a polysaccharide. For example, a xylan is a polysaccharide made up primarily of D-xylopyranosyl units (it may contain minor amounts of other sugars and still be called a xylan); a β-glucan is constructed of β-D-glucopyranosyl units; an arabinoxylan has L-arabinose and D-xylose as its monomeric units; a glucuronoxylomannan consists of D-glucuronic acid, D-xylose, and D-mannose; etc. If the polysaccharide has a backbone structure, it is the latter part of the name, e. g., galactomannans have a main mannan chain to which are attached D-galactopyranosyl units; the same is true of arabinoxylans. In a few cases, the anomeric configuration is used with the name. An example is the family of β-glucans, whose name indicates that they are made up of β-D-glucopyranosyl units (although so is cellulose) and distinguishes them from α-glycans (amylose, amylopectin, glycogen, and others). Polysaccharides/glycans that were named before there were systematic names do not follow either of these rules; examples are cellulose, amylose, pectin, glycogen, hyaluronic acid, heparin, and alginic acid. The term glycan, often misused, is restricted to polymers of simple sugars, i. e., to chains of more than 20 saccharide/ glycosyl units.

Polysaccharides may be attached to proteins, in which case they are known as protein-polysaccharides, when they originate from plants, or proteoglycans, when they originate from animals. Polysaccharides may also contain lipid components. Lipopolysaccharides and other mycobacterial constituents are covered in Chap. 6.5 and Chap. 7.

Many bacteria produce polysaccharides and either incorporate them in cell walls, incorporate them in capsules, or excrete them into the growth medium. When they are components of a capsule (capsular polysaccharides) or dissolve in the extracellular medium, they are known as exopolysaccharides.

Great structural diversity is found in the polysaccharides of both plants and microorganisms. Because their synthesis does not involve a template molecule (Chap. 6.4), polysaccharides are polydisperse, i. e., molecules of a specific polysaccharide from a single source are present in a range of molecular weights. In addition, the degree of polydispersity, the average molecular weight, and the range of molecular weights in a polysaccharide preparation vary from source to source. Most polysaccharides are also polymolecular, i. e., their fine structures vary from molecule to molecule. With the exception of cellulose and a few other plant polysaccharides, only bacterial polysaccharides have repeating-unit structures. Structures of other polysaccharides, and even of bacterial polysaccharides with regard to non-carbohydrate components, can vary between taxa and with growth conditions of the plant or microorganism and even between tissues of the same plant. As a result of the attributes described above and the fact that the structures of polysaccharides, whether of plant, algal, or microbial origin, vary with variations in the growth conditions of the source organism, the structures of most polysaccharides can be described as "complex". However, there is no specific category or definition of "complex carbohydrates", and a nutritionist would interpret the term differently than would a carbohydrate/polysaccharide chemist.

Polysaccharides are present in most living organisms. In fact, polysaccharides comprise about 70% of the dry weight of the total biomass. (Probably, at least 90% of all carbohydrate in nature occurs in polysaccharides.) They serve a variety of functions, not all of which are known. They are most abundant in higher plants:

1. as structural components of primary and secondary cell walls and in the middle lamella,
2. as reserve food materials in leaves, seeds, stems, roots, tubers, rhizomes, and other tissues,
3. as exudates of unknown function, and
4. as extractable material of unknown function.

They are cell-wall constituents, non-cell-wall constituents, and storage materials in algae. In microorganisms, they may be cellular or extracellular constituents. The polysaccharide chitin is a structural component in the exoskeletons of crustaceans and insects. The greatest amount of polysaccharides is found in plant cell walls (higher plants, algae, and fungi).

There is no ideal system of polysaccharide classification [3]. The best system should be that based on chemical structure. However, because of their polymolecularity, which limits descriptions to statistical structures in many cases, and the great variety of structures, classifying polysaccharides in this way has limitations. Combinations of the following categories are used [3]:

1. Branching:
    - unbranched,
    - linear with short branches or side units
        - regularly spaced,
        - irregularly spaced, or
        - clustered,
    - highly branched (branch-on-branch) structures;

2. Different kinds of monomer units:
   - one (homoglycan),
   - two (diheteroglycan),
   - three (triheteroglycan),
   - four (tetraheteroglycan),
   - five (pentaheteroglycan);
3. Charge:
   - neutral,
   - anionic (acidic),
   - cationic.

Specific-sugars and linkage types may then be used within each of these general groups. There may also be variability in any non-carbohydrate constituents. Polysaccharides used industrially are most often classified by source. Polysaccharides may also be categorized by function, the major two being structural and energy storage. However, especially in plants, it is not always clear whether a polysaccharide has a structural or a reserve role or both and, in both plants and animals, their functions are not always clearly and completely understood.

2 Animal Polysaccharides

2.1 Glycosaminoglycans

As their name implies, glycosaminoglycans are polysaccharides that contain amino sugars (specifically 2-amino-2-deoxy sugar units). Glycosaminoglycans are non-reserve polysaccharides that are usually covalently bound to proteins to form the proteoglycans of mammalian tissues [4]. Proteoglycans have different overall structures, but generally consist of a large number of polysaccharide chains attached to a polypeptide backbone to form structures that, in some cases, may resemble the brush part of a bottle brush (*Fig. 3*). About 90% of these structures are carbohydrate. There are considerable species-to-species and tissue-to-tissue differences in structures of the overall proteoglycan and the individual components. Most, but not all, proteoglycans are found in extracellular matrices of connective and certain other tissues, where they form a macromolecular framework.
Proteoglycans as a group contain three types of linkages attaching the reducing end of the polysaccharide chain to the core protein:

1. the GlcNAc-Asn (2-acetamido-2-deoxy-D-glucopyranosylasparagine) glycosylamine linkage as part of the core pentasaccharide unit found in many glycoproteins (Chap. 8) attaches keratan sulfate I to the core protein;
2. the GlcNAc-Ser(Thr) (2-acetamido-2-deoxy-D-glucopyranosylserine (or threonine) glycosidic linkage attaches keratan sulfate II to the core protein;
3. the Xyl-Ser (D-xylopyranosylserine) glycosidic linkage attaches chondroitin sulfates, dermatan sulfate, heparan sulfate, and heparin to the core protein.

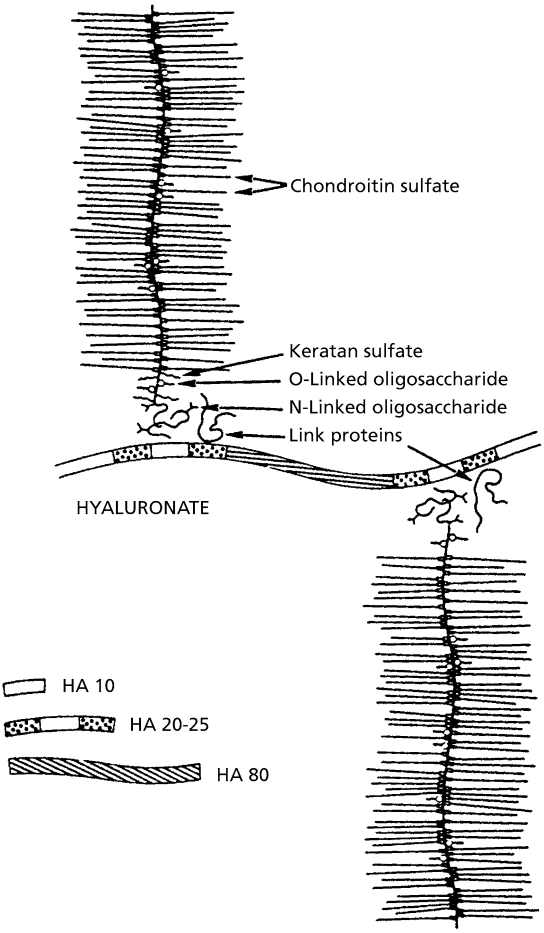


Figure 3
Proteoglycan aggregate of cartilage, consisting of a hyaluronic acid (hyaluronate, HA) chain to which proteoglycan molecules are bound non-covalently. [Reprinted by permission from reference [4, Fig. 15]

With these polysaccharides, the linkage is
$\rightarrow 4)\beta$Glc*p*A$(1\rightarrow 3)\beta$Gal*p*$(1\rightarrow 3)\beta$Gal*p*$(1\rightarrow 4)\beta$Xyl*p*$(1\rightarrow$Ser [4].
The proteoglycan molecules bind non-covalently to a decasaccharide segment of the hyaluronic acid (hyaluronate, HA) via a specific binding region, which is at one end of the protein core. Next to the HA-binding region is a segment that contains both *N*- and *O*-linked oligosaccharides and keratan sulfate. Finally, a more extended part of the core is heavily substituted with chondroitin sulfate chains plus scattered keratan sulfate molecules and oligosaccharides. The binding of proteoglycans to hyaluronate is stabilized by link proteins, which bind non-covalently, both to the hyaluronate chain and to the globular portion of the proteoglycan core. There

are ca. 40 hyaluronate disaccharide repeating units between each proteoglycan unit. The size of the aggregate depends on the size of the hyaluronate chain. Because the latter may comprise as many as 5000 disaccharide repeating units and the proteoglycan molecules are about 100 monosaccharide units apart, one hyaluronate chain may accommodate up to 100 proteoglycan chains [4].

Chondroitin sulfates have disaccharide repeating units of D-glucuronic acid and 2-acetamido-2-deoxy-D-galactose (*N*-acetylgalactosamine), viz, →4)βGlc*p*A(1→3)βGal*p*NAc(1→. Chondroitin 4-sulfate has a sulfate half-ester group on O4 of most GalNAc units; chondroitin 6-sulfate has a sulfate half-ester group on O6 of most GalNAc units. A single chondroitin sulfate molecule, however, can consist of both types of disaccharide units, and certain cells produce chondroitin 4,6-disulfate in which GalNAc units are doubly esterified. Linkage to the core protein has been previously described. Proteoglycans containing chondroitin sulfates are present in a wide variety of mammalian tissues. As much as 10% of the wet weight of cartilage and intervertebral disc tissue may be chondroitin sulfate. Chondroitin sulfates usually comprise at least 75% of the polysaccharides of cartilage; the proteoglycans of non-cartilaginous tissues contain higher proportions of dermatan sulfate. Chondroitin sulfate occurs as part of the proteoglycan stored in mast cell granules [4].

Structurally, dermatan sulfate is similar to the chondroitin sulfates, the difference being that, in dermatan sulfate, some of the D-glucuronic acid units are replaced by L-iduronic acid units, L-iduronic acid being the C5 epimer of D-glucuronic acid and always the primary uronic acid in dermatan sulfate. As with other polysaccharides of this family, dermatan sulfate is composed of various sulfated disaccharide units, the predominant ones in the case of dermatan sulfate being →4)αL-Ido*p*A(1→3)βGal*p*NAc(4-SO_3^-)-(1→, →4)βGlc*p*A(1→3)βGal*p*NAc(4-SO_3^-)-(1→, and →4)βGlc*p*A(1→3)βGal*p*NAc(6-SO_3^-)-(1→, with lesser amounts of →4)αL-Ido*p*A(2-SO_3^-)-(1→3)βGal*p*NAc(4-SO_3^-)-(1→ and →4)αL-Ido*p*A(2-SO_3^-)-(1→3)βGal*p*NAc(1→. Proteoglycans with dermatan sulfate chains are generally present, in concentrations of a few percent of the organic material, in fibrous connective tissue, in the cornea of some species, and in certain fibrous cartilages [4].

Keratan sulfate, which has a relatively low degree of sulfation, is composed of repeating units of *N*-acetyllactosamine, →3)βGal*p*(1→4)βGlc*p*NAc(1→. Sulfate half-ester groups are present on some of both monosaccharide units. As previously stated, keratan sulfate I and keratan sulfate II differ in the type of linkage to the core protein. Keratan sulfate I occurs in cornea, erythroid cell lines, and embryonic liver and lung tissue. Keratan sulfate II, which always occurs together with chondroitin sulfate in proteoglycans, is present in skeletal tissues such as bone, cartilage, and nucleus pulposus. In cartilage, a few keratan sulfate chains are clustered at one end of the core protein (*Fig. 3*); the remainder of the protein core is heavily substituted with chondroitin sulfate with some keratan sulfate chains scattered among them [4].

Heparan sulfate is the generic name for a family of glycosaminoglycans. The disaccharide unit of heparan sulfate contains either βGlc*p*A or αL-Ido*p*A linked to either αGlc*p*NAc or αGlc*p*NSO_3^-, all glycosidic linkages being (1→4). Not only do individual members of the family differ in chain structure, they differ in amounts of sulfate groups, which may be present on N2 and/or O6 of a 2-amino-2-deoxy-α-D-glucopyranosyl unit (α-D-glucosaminyl unit) and/or the O2 position of an α-L-idopyranosyluronic unit. Heparan sulfate proteoglycans are cell-surface components of many cell types [4].

Heparin has the same general structure as heparan sulfate. Heparin proteoglycan is found exclusively in granules of mast cells [4]. Heparin itself, which is obtained by degradation of the native proteoglycan, is a potent anticoagulant. Chondroitin 4,6-disulfate is also stored in mast cell granules.

Hyaluronic acid/hyaluronate is the only glycosaminoglycan that does not occur naturally as a proteoglycan. Hyaluronic acid has a disaccharide repeating unit of $\rightarrow 4)\beta$Glc*p*A$(1\rightarrow 3)\beta$Glc*p*NAc$(1\rightarrow$. It is widely distributed in mammalian tissues, viz, in synovial fluid, vitreous humor, and a variety of connective tissues and is synthesized by many bacteria. In cartilage, very large hyaluronic acid molecules form the backbones of aggregates of multiple proteoglycan molecules (*Fig. 3*), as each cartilage proteoglycan molecule contains a hyaluronate-binding region [4]. Free hyaluronic acid is found in synovial fluid and vitreous humor and in tissues with a high content of water such as skin.

As with other biopolymers, structures of the glycosaminoglycans, the proteoglycans of which they are a part, and the complexes of hyaluronic acid plus proteoglycans (*Fig. 3*) determine their physicochemical properties, which are very much related to their biological roles/functions. X-ray fiber diffraction and molecular modeling has shown that glycosaminoglycans can adopt several alternative chain conformations, but in general, the nature of the glycosidic linkages and their charged nature leads them to be in the form of extended helices [5]. When bound to a core protein molecule, the negative electrostatic charges also cause the extended chains to repel each other, leading to their ability to occupy very large hydrated domains. Because a proteoglycan of cartilage may have as many as 10,000 negative charges, the tissue has a great capacity to resist compressive forces that would force the like charges closer together. Proteoglycans also serve as a network for holding cations.

2.2 Chitin

Another structural polysaccharide of animals is chitin, which is also found in the cell walls of certain fungi [6]. Chitin is related to cellulose in that the structures are identical with replacement of the hydroxyl group at C2 of the β-D-glucopyranosyl units of cellulose with an acetamido group, i. e., chitin is a polymer of $(1\rightarrow 4)$-linked 2-acetamido-2-deoxy-β-D-glucopyranosyl (*N*-acetyl-β-D-glucosaminyl) units, although chitin is usually incompletely *N*-acetylated. Technically, chitin is a glycosaminoglycan because it contains an aminosugar as a monomer unit, but the term glycosaminoglycan is almost always restricted to those acidic/anionic mammalian polysaccharides discussed above, with the insoluble, somewhat cationic chitin being considered separately.

Like cellulose (*Sect. 3.1*), chitin imparts mechanical strength. Both chitin and cellulose form fibrils. Chitin fibrils are structural elements in the exoskeletons of insects and crustaceans. Fungal cell walls, which often contain 80–90% (dry-weight basis, db) polysaccharide, have been categorized into eight types by polysaccharide composition:

- cellulose-glycogen,
- cellulose-β-glucan,
- cellulose-chitin,
- chitosan-chitin,
- chitin-β-glucan (the most common),

- mannan-β-glucan,
- mannan-chitin,
- galactosaminoglycan-chitin.

The primary cell-wall component of many species is chitin. At least in fungal cell walls and probably elsewhere, chitin occurs as part of a proteoglycan.

2.3 Glycogen

Glycogen is a storage form of carbohydrate in many organisms. It is ubiquitous in mammals, especially in muscle and liver tissue, fish, mollusks, insects, other animals, bacteria, and fungi

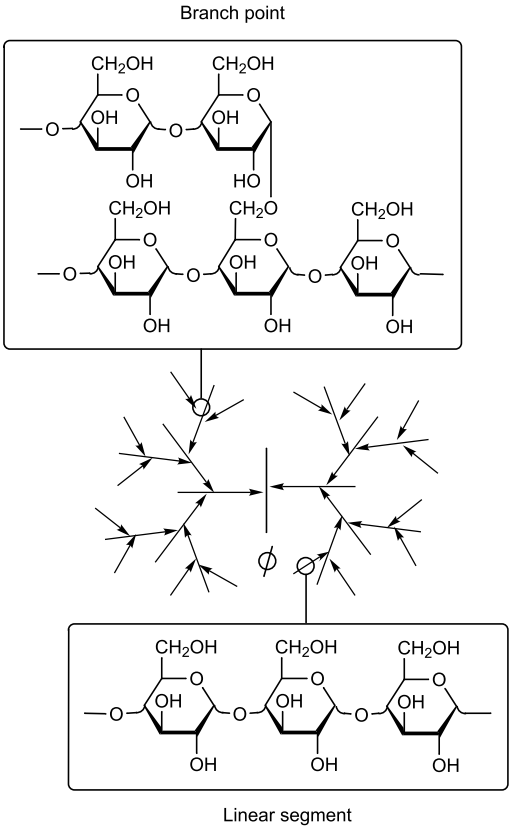


Figure 4
A schematic representation of a glycogen molecule with all but the innermost four tiers removed (see text), showing the linking of the α-D-glucopyranosyl units (in the Haworth representation) in the linear chains and at the branch points (arrowheads) and the overall arrangement of the chains according to Whelan and coworkers [8]. Φ represents the reducing end where the polysaccharide is attached to the protein primer. [Reprinted by permission from BeMiller JN (1986) Glycogen In: Mark-Bikales-Overberger-Menges: Encyclopedia of Polymer Science and Engineering, 2nd edn, vol 6. Wiley, New York]

and exists in different forms and amounts in different organisms. A related polysaccharide, phytoglycogen, which is also related to starch amylopectin, is found in some plants. The structure of the glycogen molecule is that of a tiered, branch-on-branch, bush-like structure, in which chains of (1→4)-linked α-D-glucopyranosyl units are joined by α-D-(1→6) linkages (*Fig. 4*) [7,8].

Because of its function to provide an energy source for muscle and to maintain blood sugar levels (liver), animal glycogen is always in a dynamic state. Its molecular weight, therefore, may vary from less than 1×10^6 to perhaps 2×10^9 (>6200 to at least 10^7 D-glucosyl units) depending upon the tissue of origin, the hormonal and nutritional state of the animal, and the presence or absence of disease. Even temperature, time of day, and season can affect the amount and nature of glycogen in some organisms. In any organism, at any time, the glycogen in a given tissue will be quite polydisperse, i. e., will vary considerably with regards to molecular weight/size. The highly tiered, branch-on-branch structure with approximately 50% of the D-glucosyl units in external chains gives a glycogen molecule up to tens of thousands non-reducing ends from which D-glucose 1-phosphate can be released quickly and rapidly by the action of phosphorylase when required. Glycogen in mammalian tissue occurs as single spherical (β) particles and as aggregates of as many as 100 β particles (α particles) [7].

Glycogen synthesis begins on glycogenin, a self-glucosylating protein that remains as a covalently bound part of the glycogen molecule [9,10,11]. Glycogen from certain sources contain small amounts of D-glucosamine (2-amino-2-deoxy-D-glucose) [12] and phosphate mono- and diester groups [13,14].

3 Higher Plant Cell-Wall Polysaccharides

Differences in compositions of the cell walls of higher plants occur within phyla, classes, families, and genera of higher plants, with location within a given plant (because different cells have different functions and exist in different environments), and with stage of development [15]. Nevertheless, some general features can be described. Polysaccharides are the primary constituents of plant cell walls.

3.1 Cellulose

Cellulose is the principal component of the cell walls of higher plants and, therefore, the most abundant organic compound on Earth [16,17,18,19]. Cellulose is also present in brown, red, and green algae and in certain fungi and slime molds and excreted extracellularly by certain bacteria. Even certain animals (truncates and protozoans) synthesize cellulose. Hence, cellulose is synthesized by both plants and animals, but almost exclusively by plants, and by both prokaryotes and eukaryotes (primarily by eukaryotes). The amount of cellulose in a plant varies greatly from species to species. Wood, which is about half cellulose on a dry-weight basis, has the highest percentage of cellulose, except for the fibrous seed hairs of cotton, which are ~90% cellulose. In wood, the majority of cellulose fibrils are found in thickened secondary cell walls where they provide mechanical strength and stability.

Cellulose differs from most other polysaccharides in being a homoglycan with a single type of linkage and in occurring naturally as a partially crystalline, high-molecular-weight poly-

mer with a preponderance of microcrystalline domains. In its native state, cellulose usually exists in the form of strong fibers. Chemically, all celluloses are the same, i. e., all are polymers of β-D-glucopyranosyl units linked (1→4), but these (1,4)-β-glucans differ in physical organization into fibrillar structures from source to source.

3.2 Other Plant Cell-Wall Polysaccharides

In both primary and secondary cell walls, cellulose is mixed with other polysaccharides (◉ *Table 1*) and lignin [20,21,22,23,24,25,26]. Historically, the other polysaccharides have been called hemicelluloses. The term hemicellulose indicates a polysaccharide closely associated with cellulose in cell walls. The relationship is physical, not structural. These polysaccharides are almost always heteroglycans; β-glucans and some polysaccharides that are composed almost exclusively of D-xylose are among the exceptions. Structures of hemicelluloses vary from linear to highly branched and bushlike.

Hemicelluloses have traditionally been defined by extraction procedures, i. e., hemicelluloses are those polysaccharides extracted by alkaline solutions from plant tissues after removal of low-molecular-weight substances with hot aqueous alcohol, removal of waxes and other lipid-soluble substances, delignification, and removal of pectin with an aqueous solution of a calcium ion chelator. The most abundant hemicelluloses in the primary cell walls of dicotyledons are xyloglucans. Xyloglucans have a backbone chain whose chemical structure is identical to that of cellulose. About 75% of the β-D-glucopyranosyl units in that chain are substituted

Table 1
Some cell-wall polysaccharides[a,b]

Higher land plants	*Marine algae*	*Fungi and yeasts*
Cellulose	Algins	Cellulose
Hemicelluloses	Cellulose	Chitin
Arabinoxylans	Galactans	β-Glucans
Galactoglucomannans	Agars	α-Glucans
β-Glucans	Carrageenans	Mannans
Glucomannans	Furcellarans	
Mannans	β-Glucans	
Xylans	L-Fucans	
Xyloglucans	Mannans	
Pectic polysaccharides	Xylans	
Arabinans		
Arabinogalactans		
Galactans		
Galacturonans		
Rhamnogalacturonans		

[a]Other than those of bacteria; [b]Many of these polysaccharides may contain monosaccharide units other than those indicated in the name. For example, xylans often contain uronic acid units, and the polysaccharides named L-fucans may contain, in addition to the principal sugar (L-fucose), D-galactose, D-glucuronic acid, D-mannose, and D-xylose

at O6 with β-D-xylopyranosyl units or oligosaccharide chains terminated at the reducing end (the end attached to the main chain) with a β-D-xylopyranosyl unit. In monocotyledons, the most abundant polysaccharides are branched glucuronoarabinoxylans. These polysaccharides have main chains of (1→4)-linked β-D-xylopyranosyl units, which is the structure of cellulose with the terminal hydroxymethyl group (C6) removed. Attached to the main chains at O2 and/or O3 are units of L-arabinose, D-galactose, or D-glucuronic acid. Both xyloglucans and glucuronoarabinoxylans are cross-linked via substituents attached via ester or ether bonds and may be linked to other cell wall components, viz., lignin, in the same way.

The predominant hemicelluloses in secondary cell walls of both monocotyledons and dicotyledons are xylans. These polysaccharides have a xylan main chain that is decorated with 4-*O*-methyl-D-glucopyranosyluronic acid, D-glucopyranosyluronic acid, L-arabinofuranosyl, and/or L-arabinopyranosyl units. These related polysaccharides are usually acetylated to various degrees. Gymnosperms contain glucomannans, also acetylated, as the principal secondary cell-wall hemicellulose [26].

Pectic substances, another family of polysaccharides (*Table 1*), are located in middle lamellae and primary cell walls. Their structures can be complex, with more than one type of polysaccharide chain in a single molecule [21]. For example, arabinogalactan chains are attached to poly(α-D-galactopyranosyluronic acid) chains in native pectins (rhamnogalacturonans). (Arabinogalactans also commonly occur in protein-polysaccharide structures.) The primary structural units of rhamnogalacturonans are α-D-galacturonopyranosyluronic acid units. Ten other monosaccharide units have been found in rhamnogalacturonans, including the rare sugars D-apiose, L-aceric acid, and 3-deoxy-D-*lyxo*-2-heptulosaric acid and a sugar usually associated with bacterial polysaccharides, 3-deoxy-D-*manno*-2-octulosonic acid (Kdo) [26]. Pectic substances comprise 1–4% of woody tissue. Biotechnology is being used to modify plant pectins.

The polysaccharides found in cell walls are specific to specific cell types. Most flowering plants have type-I cell walls. These walls have a cellulose-xyloglucan interlocking framework (~50% of the wall mass) embedded in a matrix of pectic polysaccharides [(~30%; rhamnogalacturonan I plus lesser amounts of partially methyl-esterified poly(D-galacturonic acids) (galacturonans)]. Grasses have type-II cell walls. The characteristic components of these walls are cellulose, glucuronoarabinoxylans, galacturonans, and mixed-linkage β-glucans/(1,3:1,4-β-glucans) [27].

Multicellular algae must be tough and flexible. Holding the plant upright is not a requirement as it is with land plants, which are too brittle for aquatic existence. Structural polysaccharides of algae vary greatly between phyla. Included as cell-wall components of algae are alginates, sulfated galactans (carrageenans, agars, etc.), glucans, fructans, xylans, mannans, cellulose, pectic substances, and other polysaccharides [28].

3.3 Plant Cell-Wall Polysaccharides that Have a Reserve Function

Plant cell walls are the source of the greatest quantity of polysaccharides. Some cell-wall-associated polysaccharides may serve storage functions in addition to or rather than structural functions. Such polysaccharides include the β-glucans [29,30], mannans [31], galactans [28], and arabinogalactans [23].

β-Glucans occur in cereal seeds from which they can be extracted [29,30]. Some are (1→3)-linked β-glucans. Barley endosperm cell walls contain ~75% of this type, some of which may be covalently bound in some way. (1→3)-Linked β-glucans are also found in many other monocots, especially in oat and other Graminae. Another group of β-glucans in the grasses, known as mixed-linkage β-glucans, contains polysaccharides with both (1→3) and (1→4) linkages.

4 Energy-Reserve Polysaccharides of Plants

Some polysaccharides provide a reserve energy supply for tissues and organisms (● *Table 2*). One polysaccharide that serves in this role, glycogen, is discussed in ● *Sect. 2.3*. Starch is the principal carbohydrate energy-storage substance of higher plants [32,33,34] and, after cellulose, the second most abundant carbohydrate end-product of photosynthesis. Starch is not only a reserve substance of many higher plants, it is an energy source for animals that feed on them. All higher plants produce starch sometime during their lifetime. Starch is found in leaves, where it serves as a transient D-glucose storage material, and in seeds (especially those of cereal grains), fruits, roots, rhizomes, stems, tubers, and trunks for long-term storage. Starches provide at least 70% of human caloric intake on a worldwide basis.

Starch is unique among carbohydrates in that it occurs in discrete particles called granules. Starch granules are relatively dense and insoluble. Most starch granules contain two polymers: an essentially linear polysaccharide called amylose and a highly branched polysaccharide called amylopectin. Each starch from each plant source and tissue is unique in terms of granule morphology, composition, polysaccharide structures, and behavioral characteristics. Several mutant variants of starch are known; two of them, waxy maize and high-amylose (amylose extender) maize, are major commercial products.

Other energy-storage polysaccharides include inulin and other fructans in roots, tubers, stems, and algae [35]; galactomannans in legume seeds [36, ● Chap. 6.4]; mannans [31]; glucomannans [31]; starch-like polysaccharides (floridean starch), fructans, and β-glucans of algae [37]; and α- and β-glucans of fungi [37].

It is not always clear whether cell-wall polysaccharides have a structural or reserve function because some, for example, the β-glucans of cereal grains, may have dual roles. However, it is likely that the β-glucans and other hemicelluloses in cell walls of tissues other than those in

Table 2
Some energy-storage polysaccharides

Higher land plants	*Marine algae*	*Fresh-water algae*	*Fungi and yeasts*	*Animals*
Fructans	Fructans	α-Glucans[a]	α-Glucans[a]	Glycogen
Galactans	α-Glucans[a]	β-Glucans	β-Glucans	
Galactomannans	β-Glucans			
Glucomannans	Xylans			
Starches				
Xyloglucans				

[a]Starch- and glycogen-like polymers

seeds function as structural components. During germination, seed components, including in some cases β-glucans, are used for energy and carbon sources for synthesis.
(1→3)-Linked β-glucans are cell-wall constituents of bacteria, blue-green and other algae, fungi, yeasts, and lichens [30]. Some of this class of polysaccharides in cell walls of some bacteria and some fungi are (1→3), (1→6)-β-glucans which have a (1→3)-linked main chain to which are attached β-D-glucopyranosyl and, in some cases, gentiobiosyl units via (1→6) linkages. Other β-glucans, other than cellulose, contain other linkages, e. g., all (1→6), all (1→2), and other mixed linkages [30].

5 Bacterial Polysaccharides

Bacterial cell walls contain a structural material called peptidoglycan or murein. This substance is composed of glycosaminoglycan chains cross-linked with peptides to form a three-dimensional network [38,39,40]. The glycosaminoglycan chains are composed of a disaccharide repeating unit in which one member is a 2-acetamido-2-deoxy-β-D-glucopyranosyl unit and the other is usually an *N*-acetylmuramic acid (a 2-acetamido-2-deoxy-β-D-glucopyranosyl unit etherified at O3 with D-lactic acid) unit. (Modifications of *N*-acetylmuramic acid are rare, but do occur.) Therefore, the glycan has the structure of chitin with alternating glycosyl units possessing a lactic acid ether group. Bacterial cell walls may also contain non-polysaccharide carbohydrate polymers (teichoic acids and teichuronic acids). In these polymers, alditols, monosaccharides, or oligosaccharides are joined by phosphodiester linkages. Teichoic and teichuronic acids are linked covalently to the peptidoglycan in the cell wall via a phosphodiester bond.
Some carbohydrate-containing polymers are components of capsules or are secreted into the culture medium. As stated by Kenne and Lindberg [41], structural variation in such bacterial polysaccharides is almost unlimited and unusual sugars are not uncommon. The overwhelming majority of bacterial polysaccharides are heteroglycans composed of oligosaccharide repeating units. A majority is also antigenic and responsible for the specific immunological properties of one type, one species, or one group of bacteria. Bacterial cell-wall polysaccharides, i. e., the lipopolysaccharides and other mycobacterial cell-wall polysaccharides, are described in Chap. 6.5 and Chap. 7.
Bacterial exopolysaccharides (extracellular polysaccharides) include those that are made from sucrose, viz., dextrans and fructans. Dextrans are branched α-glucans containing (1→3) and/or (1→6) and occasionally (1→2) linkages. Fructans contain β-D-fructofuranosyl units linked (2→6) or (2→1).

6 Fungal and Yeast Cell-Wall Polysaccharides

In addition to chitin (*Sect. 2.2*), yeast cell walls may contain a (1→3)-β-glucan, a (1→6)-β-glucan, various mannans, or a galactan. Other cell-wall components of certain fungi are rhamnomannans, glucomannans, galactomannans, xylomannans, glucuronoxylomannans, and other polysaccharides [42].

7 Physical Properties

Because the structures of polysaccharides vary widely, their physical properties vary. This section mentions in a cursory way two common properties: (1) that of being water-soluble or hydratable and (2) that of forming intermolecular associations. Neither is a universal property of polysaccharides. Polysaccharides that are unbranched (or only very slightly branched) and unsubstituted homopolymers with a single type of linkage (generally 1→3 or 1→4) are insoluble in water at temperatures up to 100 °C because of their ability to form intermolecular hydrogen bonds over much of their lengths; examples are cellulose, mannan, and crystalline amylose. Some polysaccharides that, while slightly branched or substituted, but which have significant chain segments that are unbranched may be soluble only in hot water or high-pH solutions; examples are locust bean (carob) gum, some xylans, and certain arabinoxylans (Chap. 6.4). Modification or derivatization of only a few hydroxyl groups on an insoluble polysaccharide appears to be sufficient to diminish intermolecular hydrogen bonding and provide water solubility. Polysaccharide molecules or bundles of molecules that can associate with other polymer molecules or bundles of molecules over only a short portion of their lengths will form three-dimensional networks and gel the aqueous system in which they are present (Chap. 6.4). In some cases, these associations involve hydrogen bonding between hydroxyl groups of portions of adjacent chains. Some polysaccharides will form gels by themselves; others require the presence of other substances to form a gel. Many commercial water-soluble polysaccharides will gel aqueous systems (Chap. 6.4). The physiological function of certain natural "complex polysaccharides" as described in *Sect. 2.1* requires them to form a gel matrix. Polysaccharides cannot be dissolved under conditions in which they would form a gel.

8 Structural Analysis

There are generally six aspects of determining the structure of a polysaccharide: (1) determination of the monosaccharide composition, (2) determination of linkages to the various monosaccharide constituents, (3) determination of the sequence of monosaccharide units, (4) determination of anomeric configurations, (5) determination of the presence and location of any substituent (ester, ether, cyclic acetal) groups, (6) determination of average molecular weight and degree of polydispersity. Determination of the monosaccharide composition begins with acid-catalyzed hydrolysis under conditions determined experimentally to give both maximum depolymerization and minimum destruction of the sugars [43]. Released monosaccharides are then determined both qualitatively and quantitatively by HPLC [44], usually by using high-performance anion-exchange chromatography and pulsed amperometric detection (HPAEC-PAD).

Linkages of most polysaccharides have been determined by methylation analysis, which can provide information about the linkage site, ring size, and the nature of the monosaccharide unit [45]. Methylation analysis determines the position(s) of linkage(s) to each monomer unit. It does not give information about the sequence of the units or their anomeric configuration(s). To provide information about the monosaccharide sequence(s), the polysaccharide may be partially depolymerized using enzyme- or acid-catalyzed hydrolysis or methanolysis to gener-

ate oligosaccharides, whose structures are then determined. One approach to determining the structure (other than anomeric configurations) of an oligosaccharide is via use of MALDI-TOF mass spectroscopy [46]. A technique that provides both linkage analysis and anomeric configurations of the monomer units is that known as reductive cleavage analysis [46]. Oxidative techniques can also provide structural information [46].

Combinations of methylation, various degradative techniques, and various ^{1}H and ^{13}C NMR spectroscopic techniques are used to characterize polysaccharides, except for molecular weight and degree of polydisperity. Ten selected examples are given in references [47,48,49,50,51,52,53,54,55,56]. Average molecular weights and molecular weight ranges are determined by high- or medium-angle laser light scattering (MALLS) analysis of the eluate. Mass spectrometric techniques are commonly used to sequence structures of the oligosaccharide portions of glycolipids and glycoproteins [57,58]. Mass spectrometry can be combined with other techniques in the structural analysis of polysaccharides [46,59,60].

Mishnick and co-workers have developed several techniques for determining the positions of substitutions on monosaccharide units and distribution of substituents along derivatized polysaccharide chains, primarily cellulose derivatives. Examples are found in references [61,62,63,64,65,66,67,68,69,70]. An important method involves perdeuteromethylation followed by partial acid-catalyzed hydrolysis, then reductive amination with propylamine, and finally permethylation to yield completely *O*- and *N*-alkylated charged oligosaccharides that are analyzed by MALDI-TOF-MS [68,70]. Statistical methods are then applied.

8.1 Molecular Conformations

Molecular conformations [71,72] and associations can be determined by CP-MAS-^{13}C-NMR and two-dimensional NMR methods [73], but the primary means of determining molecular conformations has been X-ray fiber diffraction analysis coupled with molecular modeling as practiced by Chandraskeran and co-workers. Examples are found in references [74,75,76,77,78,79,80,81,82,83,84]. The technique has also been used to study polysaccharide-polysaccharide interactions [85]. Recently, powder diffraction data has been used to determine a molecular conformation [86]. Modeling alone has been used to predict possible conformations [87].

9 Reactions of Polysaccharides

Although polysaccharide molecules may contain a variety of functional groups, including amino groups (especially in bacterial polysaccharides), carboxyl groups (in polysaccharides containing uronic acids and succinate half ester and pyruvyl cyclic acetal groups), and various organic and inorganic ester and amide and ether groups (*Sect. 1*), the primary reactions of polysaccharides involve hydroxyl groups. Reactions resulting in cleavage of glycosidic bonds are also important. Selected examples of the former and of non-hydroxyl group reactions are presented below.

9.1 Esterifications

Esterifications (acylations) with organic acids can be effected in two ways. One is to use a highly reactive reagent such as the acylpyridinium ion (e. g., acetic anhydride or acetyl chloride in pyridine), often with a tertiary amine for neutralization of the generated acid [88], or with an acid anhydride in an aqueous alkaline system. Commercially, organic esters are made in an aqueous alkaline medium using acid anhydrides. Starch is acetylated with acetic anhydride to a low-DS for the purpose of making a stabilized product [89] (Transesterification using vinyl acetate is used in countries in which acetic anhydride cannot be possessed [89].) Reaction with a difunctional reagent (acetic adipic mixed anhydride) is used to strengthen starch granules via cross-linking [90].

Practiced with cellulose is acylation to the extent that the molecules are soluble in organic solvents. The product is often thermoplastic. These products include cellulose triacetate, cellulose 2,3-diacetate (acetate rayon), cellulose acetate butyrate, cellulose acetate phthalate, and cellulose acetate propionate.

Inorganic esters, viz., the sulfate half ester and phosphate esters, are also produced. Sulfate half esters can be made using sulfur trioxide and a tertiary amine [91,92,93]. An example of the latter is sulfation by SO_3-trimethylamine in DMF. Sulfation of dermatan sulfate using this method resulted in initial substitution on the O6 of the 2-acetamido-2-deoxy-β-D-galactopyranosyl 4-sulfate units to produce the 4,6-disulfate. When this initial step was about 50% complete, sulfation occurred about equally on O2 and O3 of the α-L-idopyranosyluronic acid units, producing a product with β-GalNAc 4,6-disulfate, αLIdoA 2-SO_4, αLIdoA 3-SO_4, and αL-IdoA 2,3-disulfate units [94]. There are other regioselective approaches for polysaccharide functionalization by sulfation [95].

The reverse of sulfation is desulfation, which can be effected under mild enzymic conditions [96]. Similarly, enzymic depolymerization, closely guarded by trade secrets, is routinely used in the production of heparin-derived drugs.

Sulfate groups are also easily removed by aqueous acid with S–O bond scission so that the configuration of the carbohydrate is fully retained. The general order for desulfation is O6 > O2 > O2 [97].

Cyclic sulfates can be prepared from cyclic but not acyclic carbohydrate units by reaction with sulfuryl chloride in pyridine [98] or phenyl chlorosulfate and sodium hydride [99]. The reactivity of cyclic sulfates is quite similar to that of epoxides. Thus, they can be used to form fluoro or azido derivatives of polysaccharides by highly regiospecific approaches, usually, but not always, leading to *trans*-diaxial products from which the sulfate can be easily removed by mild, acid-catalyzed hydrolysis [100,101].

Both monostarch and distarch phosphate esters are commercial starch derivatives [90,102].

9.2 Etherifications

Etherifications (alkylations) are carried out for two reasons: structural analysis (*Sect. 8.1*) and to change the nature (properties) of the polysaccharide (Chap. 6.4). The basic reaction used is the classical Williamson ether synthesis, i. e., some of the hydroxyl groups on the polysaccharide are converted into the alkoxy form, then these molecules are reacted with

molecules with leaving groups, i. e., epoxides or alkyl or aryl halides [103,104,105]. In addition to the common water-soluble derivatives, called gums, described in Chap. 6.4, sulfoethyl, sulfopropyl, diethylaminoethyl, and other ether derivatives are made commercially as chromatography media.

Fundamentally, any reaction of a low-MW alcohol to produce an ester or ether can be done with a polysaccharide. Many derivatives have been made, including those made with difunctional reagents for the purpose of cross-linking molecules [101]. Only a few have been commercialized. Highly substituted ethylcellulose, a thermoplastic substance with organic solvent solubility, is made commercially.

9.3 *N*-Acylation of Aminopolysaccharides

Selective *N*-acylation of a number of aminopolysaccharides and glycosamino-glycuronans depends on the basicity of the nitrogen atom present in the polysaccharide chain [106] and they can be efficiently acylated using anhydrides in organic media [107]. Various free aminochitosan fractions, for example, can be quantitatively acylated at ambient temperature within a short period of time. The high efficiency of the above reaction can be attributed to the increased basicity of the amino function during the reaction. *N*-Alkylations are effected in a similar manner.

9.4 Oxidations

Partial oxidations [104,105] of polysaccharides is commonly done, both during structural analysis and to modify their properties. Oxidation introduces both carbonyl and carboxylate functions at different positions, and especially in alkaline systems, can result in chain cleavage, i. e., depolymerization [108]. Oxidation using sodium hypochlorite by itself [109] or in combination with sodium bromide [110] is practiced in the starch industry, both to introduce specific properties into the product and for depolymerization. This oxidation is non-specific.

Vicinal diol functions are oxidized with a high degree of specificity using periodic acid. Dialdehyde polysaccharides are formed [111,112].

Selective oxidation of primary hydroxyl groups can be achieved with 2,2,6,6-tetramethyl-1-piperidinyloxy (TEMPO) and hypobromite or hypochlorite at pH 10.5–11. A 98% selectivity was observed for starch, and selectivity of over 90% for inulin [113].

Another selective method for primary hydroxyl group oxidation is the oxidation by dinitrogen tetraoxide in the gas phase or dissolved in carbon tetrachloride [114]. When this reaction is applied to cellulose, formation of a high yield of D-glucuronic acid units incorporated into the cellulose chain results [115]. Selective oxidation of the primary hydroxyl group was increased to 88% by using phosphoric acid and sodium nitrite to generate the dinitrogen tetraoxide.

Other selective oxidation methods include the combination of dimethyl sulfoxide/acetic anhydride [116] and oxygen (using Adams catalyst) [117]. The specific combination of chromium trioxide/perchloric acid [118] was extremely effective for the selective oxidation of the primary hydroxyl groups of chitosan, producing a polymer made up entirely of 2-amino-2-deoxy-D-glucuronic acid units. An interesting mechanism explaining the formation of this product was proposed. The mechanism is based on the quaternization of the primary amino group

prior to oxidation and formation of a bulky ammonium perchlorate group, thus providing the required protection of the hydroxyl group at C3.
Selective oxidation of amylose at C3 has been achieved with dimethyl sulfoxide/phosphorus pentaoxide [119]. In contrast, the oxidation of unprotected cellulose with dimethyl sulfoxide/acetic anhydride [120] or dicyclohexylcarbodiimide/dimethyl sulfoxide [121] produced a product containing modified β-D-glucopyranosyl units – some with a carbonyl function at C2, some with a carbonyl function at C3, and some with carbonyl functions at both C2 and C3 [122].

9.5 Reactivity of Carboxyl and Carbonyl Groups of Polysaccharides

Carboxylic acid-containing polysaccharides can be easily functionalized by reactions such as esterification, amination, or amidation. Industrial esterification of polyuronides, including alginic acid (Chap. 6.4), in aqueous medium can be easily accomplished by reaction with ethylene, propylene, butylene, and pentylene oxides in 50–80% yield [123].
Amination can be achieved by reductive amination of carboxyl groups using sodium cyanoborohydride [124]. By this method, 2-amino-2-deoxy functionalized amylose and cellulose have been produced. Reductive amination can also be applied to the synthesis of other amino derivatives of various oxidized polysaccharides such us xanthan [125] or alginic acid and galactomannans [126].

References

1. Aspinall GO (ed) (1982, 1983, 1985) The polysaccharides, vol 1, vol 2, vol 3. Academic Press, New York
2. BeMiller JN (1999) Macromol Symp 140:1
3. BeMiller JN (2001) Classification, structure, and chemistry of polysaccharides in foods. In: Cho SS, Dreher M (eds) Handbook of dietary fiber and functional foods. Marcel Dekker, New York, chap 31
4. Fransson L-A (1985) Mammalian glycosaminoglycans. In: Aspinall GO (ed) The polysaccharides, vol 3. Academic Press, Orlando, Florida, chap 5
5. Rees DA, Morris ER, Thom D, Madden JK (1982) Shapes and interactions of carbohydrate chains. In: Aspinal GO (ed) The polysaccharides, vol 1. Academic Press, Orlando, FL, chap 5
6. Muzzarelli RAA (1985) Chitin. In: Aspinall GO (ed) The polysaccharides, vol 3. Academic Press, Orlando, FL, chap 6
7. Geddes R (1985) Glycogen: a structural viewpoint. In: Aspinall GO (ed) The polysaccharides, vol 3. Academic Press, Orlando, FL, chap 4
8. Gunja-Smith Z, Marshall JJ, Mercier C, Smith EE, Whelan WJ (1970) FEBS Lett 12:101
9. Whelan WJ (1986) BioEssays 5:136
10. Smythe C, Cohen P (1991) Eur J Biochem 200:625
11. Alonso MD, Lomako J, Lomako WM, Whelan WJ (1995) FASEB J 9:1126
12. Kirkman BR, Whelan WJ, Bailey JM (1989) BioFactors 2:123
13. Lomako J, Lomako WM, Whelan WJ, Marchase RB (1993) FEBS Lett 329:263
14. Lomako J, Lomako WM, Kirkman BR, Whelan WJ (1994) BioFactors 4:167
15. Bacic A, Harris PJ, Stone BA (1988) Structure and function of plant cell walls. In: Preiss J (ed) The biochemistry of plants, vol 14. Academic Press, New York, chap 8
16. Brown RM Jr (ed) (1982) Cellulose and other natural polymer systems: biogenesis, structure, and degradation. Plenum Press, New York
17. Tarchevsky IA, Marchenko GN (1991) Cellulose: biosynthesis and structure (translated from Russian). Springer, Berlin, Heidelberg, New York

18. Klemm D, Philipp B, Heinze T, Heinze U, Wagenknecht W (eds) (1998) Comprehensive cellulose chemistry, vols 1 and 2. Wiley-VCH, Weinheim
19. Atalla RH (1999) Celluloses. In: Pinto BM (ed) Carbohydrates and their derivatives including tannins, cellulose, and related lignins, vol 3, Comprehensive natural products chemistry. Elsevier, Amsterdam, chap 3.16
20. Whistler RL, Richards EL (1970) Hemicelluloses. In: Pigman W, Horton D (eds) The carbohydrates, vol IIA. Academic Press, New York, chap 37
21. Aspinall GO (1980) Chemistry of cell wall polysaccharides. In: Preiss J (ed) The biochemistry of plants, vol 3. Academic Press, New York, chap 12
22. McNeil M, Darvill AG, Fry SC, Albersheim P (1984) Ann Rev Biochem 53:625
23. Brinson K, Dey PM (1985) Polysaccharides containing xylose, arabinose, and galactose in higher plants. In: Dey PM, Dixon RA (eds) Biochemistry of storage polysaccharides in green plants. Academic Press, London, chap 10
24. Shimizu K (1991) Chemistry of hemicelluloses. In: Hon DN-S, Shiraishi N (eds) Wood and cellulosic chemistry. Marcel Dekker, New York, chap 5
25. Albersheim P, Darvill AG, O'Neill MA, Schols HA, Voragen AGJ (1996) An hypothesis: the same six polysaccharides are components of the primary cell walls of higher plants. In: Visser J, Voragen AGJ (eds) Pectins and pectinases. Elsevier, Amsterdam, pp 47–55
26. Mohnen D (1999) Biosynthesis of pectins and galactomannans. In: Pinto BM (ed) Carbohydrates and their derivatives including tannins, cellulose, and related lignins, vol 3, Comprehensive natural products chemistry. Elsevier, Amsterdam, chap 3.15
27. Carpita NC, Gibeaut DM (1993) Plant J 3:1
28. Percival E, McDowell RH (1985) Algal polysaccharides. In: Dey PM, Dixon RA (eds) Biochemistry of storage polysaccharides in green plants. Academic Press, London, chap 9
29. Dixon RA (1985) β-(1→3)-Linked glucans from higher plants. In: Dey PM, Dixon RA (eds) Biochemistry of storage carbohydrates in green plants. Academic Press, London, chap 6
30. Stone BA, Clarke AE (1992) Chemistry and biology of (1→3)-β-glucans. LaTrobe University Press, Victoria, Australia
31. Bewley JD, Reid JSG (1985) Mannans and glucomannans. In: Dey PM, Dixon RA (eds) Biochemistry of storage polysaccharides of green plants. Academic Press, London, chap 8
32. Manners DJ (1985) Starch. In: Dey PM, Dixon RA (eds) Biochemistry of storage carbohydrates in green plants. Academic Press, London, chap 4
33. Jenner CF (1988) Storage of starch. In: Preiss J (ed) The biochemistry of plants, vol 14. Academic Press, New York, chap 20
34. Whistler RL, BeMiller JN (eds) (1984) Starch: chemistry and technology. Academic Press, Orlando, FL
35. Pontis HG, del Campillo E (1985) Fructans. In: Dey PM, Dixon RA (eds) Biochemistry of storage carbohydrates in green plants. Academic Press, London, chap 5
36. Reid JSG (1985) Galactomannans. In: Dey PM, Dixon RA (eds) Biochemistry of storage carbohydrates in green plants. Academic Press, London, chap 7
37. Manners DJ, Sturgeon, RJ (1988) Reserve carbohydrates of algae, fungi, and lichens. In: Preiss J (ed) The biochemistry of plants, vol 14. Academic Press, New York, chap 12
38. Rogers HJ, Perkins HR, Ward JB (1980) Microbial cell walls and membranes. Chapman & Hall, London
39. Hammond SM, Lambert PA, Rycroft A (1983) The bacterial cell surface. Croom Helm, Beckenham, UK
40. Bugg TDH (1999) Bacterial peptidoglycan biosynthesis and its inhibition. In: Pinto BM (ed) Carbohydrates and their derivatives including tannins, cellulose, and related lignins, vol 3, Comprehensive natural products chemistry. Elsevier, Amsterdam, chap 3.10
41. Keene L, Lindberg B (1983) Bacterial polysaccharides. In: Aspinall GO (ed) The polysaccharides, vol 2. Academic Press, New York, chap 5
42. Gorin PAJ, Barreto-Bergter E (1983) The chemistry of polysaccharides of fungi and lichens. In: Aspinall GP (ed) The polysaccharides, vol 2. Academic Press, New York, chap 6
43. Biermann CJ (1989) Hydrolysis and other cleavage of glycosidic linkages. In: Biermann CJ, McGinnis GD (eds) Analysis of carbohydrates by GLC and MS. CRC Press, Boca Raton, FL, p 27
44. McGinnis GD, Laver ML, Biermann CJ (1989) High-performance liquid chromotography (HPLC) of carbohydrates. In: Biermann CJ,

McGinnis GD (eds) Analysis of carbohydrates by GLC and MS. CRC Press, Boca Raton, FL, p 19
45. Carpita NC, Shea EM (1989) Linkage structure of carbohydrates by gas chromotography-mass spectrometry (GC-MS) of partially methylated alditol acetates. In: Biermann CJ, McGinnis GD (eds) Analysis of carbohydrates by GLC and MS. CRC Press, Boca Raton, FL, p 157
46. Cui SW (2005) Structural analysis of polysaccharides. In: Cui SW (ed) Food carbohydrates. CRC Press, Boca Raton, FL, p 105
47. Bilan MI, Grachev AA, Shashkov AS, Nifantiev NE, Usov AI (2006) Carbohydr Res 341:238
48. Ahrazem O, Prieto A, Gimenez-Abian MI, Leal JA, Jiminez-Barbero J, Bernable M (2006) Carbohydr Res 341:246
49. Omarsdottir S, Petersen BO, Barsett H, Paulsen BS, Duus JO, Olafsdottir ES (2006) Carbohydr Polym 63:54
50. Urai M, Anazai H, Ogihara J, Iwabuchi N, Harayama S, Sunairi M, Nakajama M (2006) Carbohydr Res 341:766
51. Mondal S, Chakraborty I, Rout D, Islam SS (2006) Carbohydr Res 341:878
52. Jansson P-E, Stenutz R, Widmalm G (2006) Carbohydr Res 341:1003
53. Perepelov AV, Zablotni A, Shashkov AS, Knirel YA, Sidorczyk Z (2006) Carbohydr Res 341:1969
54. Snyder DS, Gibson D, Heiss C, Kay W, Azadi P (2006) Carbohydr Res 341:2388
55. Bilan MI, Vinogradova EV, Shashkov AS, Usov AI (2007) Carbohydr Res 342:586
56. Shashkov AS, Kocharova NA, Zatonsky GV, Blaszczyk A, Knirel YA, Rozalski A (2007) Carbohydr Res 342:653
57. Dell A (2002) Aust J Chem 55:27
58. Haslam SM, Khoo K-H, Dell A (2003) Sequencing of oligosaccharides and glycoproteins. In: Wong C-H (ed) Carbohydrate-based drug discovery, vol 2. Wiley VCH, Weinheim, p 461
59. Wilson JC, Hitchen PG, Martin F, Peak IR, Collins PM, Morris HR, Dell A, Grice ID (2005) Carbohydr Res 340:765
60. Forsberg LS, Reuhs BL (1997) J Bacteriol 179:5366
61. Wilke O, Mischnick P (1997) Starch/Stärke 49:453
62. Gohdes M, Mischnick P (1998) Carbohydr Res 309:109
63. Mischnick P (1998) Starch/Stärke 50:33
64. Mischnick P (2000) Angew Chem Int Ed 39:1222
65. Kern H, Choi SC, Wenz G, Heinrich J, Ehrhardt L, Mischnik P, Garidel P, Blume A (2000) Carbohydr Res 326:67
66. Mischnik P (2001) Cellulose 8:245
67. Adden R, Mischnik P (2005) Int J Mass Spectr 242:63
68. Adden R, Niedner W, Müller R, Mischnik P (2006) Anal Chem 78:1146
69. Adden R, Müller R, Brinkmalm G, Ehrler R, Mischnik P (2006) Macromol Biosci 6:435
70. Adden R, Müller R, Mischnick P (2006) Cellulose 13:459
71. Wang Q, Cui SW (2005) Understanding the conformations of polysaccharides. In: Cui SW (ed) Food carbohydrates. Taylor & Francis, Boca Raton, FL, p 219
72. Rao VSR, Balaji PV, Chandrasekaran R (1998) Conformation of carbohydrates. Harwood Academic Publ., Amsterdam, Chap 7 (p 223), Chap 8 (p 255), Chap 9 (p 277).
73. Stipanovic AJ, Giammatteo PJ (1989) Advan Chem Ser 223:73
74. Chandrasekaran R, Puigjaner LC, Joyce KL, Arnott S (1988) Carbohydr Res 181:23
75. Chandrasekaran R, Radha A, Thailambal VG (1992) Carbohydr Res 224:1
76. Chandrasekaran R (1997) Advan Carbohydr Chem Biochem 52:311
77. Chandrasekaran R (1998) Advan Food Nutr Res 42:131
78. Chandrasekaran R, Bian W, Okuyama K (1998) Carbohydr Res 312:219
79. Chandrasekaran R (1999) Macromol Symp 140:17
80. Bian W, Chandrasekaran R, Rinaudo M (2002) Carbohydr Res 337:45
81. Bian W, Chandrasekaran R, Ogawa K (2002) Carbohydr Res 337:305
82. Chandrasekaran R, Janaswamy S, Bian W (2002) Polym Preprints 43:193
83. Janaswamy S, Chandrasekaran R (2002) Carbohydr Res 337:523
84. Janaswamy S, Chandrasekaran R (2005) Carbohydr Polym 60:499
85. Chandrasekaran R, Janaswamy S, Morris VJ (2003) Carbohydr Res 338:2889
86. Janaswamy S, Chandrasekaran R (2005) Carbohydr Res 340:835
87. Chandrasekaran R, Janaswamy S (2002) Carbohydr Res 337:2211

88. Sun R, Fang JM, Tomkinson J (2000) J Agr Food Chem 48:1247
89. Jarowenko W (1986) Acetylated starch and miscellaneous organic esters. In: Wurzburg OB (ed) Modified starches: properties and uses. CRC Press, Boca Raton, FL, p 55
90. Wurzburg OB (1986) Cross-linked starches. In: Wurzburg OB (ed) Modified starches: properties and uses. CRC Press, Boca Raton, FL, p 41
91. Doctor WM, Lewis D, Coleman M (1991) Thrombosis Res 64:413
92. Osawa Z, Morota T, Hatanaka K, Akaike T (1993) Carbohydr Polym 21:283
93. Yamamoto I, Takayama K, Honma K (1991) Carbohydr Polym 14:53
94. Ludwig-Baxter KG, Liu Z, Perlin AS (1991) Carbohydr Res 214:245
95. Falshaw R, Furneaux RH, Slim GC (1999) Carbohydrate sulphates. In: Finch P (ed) Carbohydrates: structures, syntheses and dynamics. Kluwer Academic Publishers, Dordrecht, p 107
96. Linhardt RJ (1992) Chemical and enzymatic methods for the depolymerization and modification of heparin. In: Ogura H (ed) Carbohydrates—synthetic methods and application in medicinal chemistry. Kodansha/VCH Publishers, Tokyo, p 387
97. Roy AB, Turner J (1983) Carbohydr Res 124:338
98. Turvey JR (1965) Adv Carbohydr Chem Biochem 20:183
99. Abdel-Malik MM, Perlin AS (1989) Carbohydr Res 190:39
100. van der Klein PAM, Filemon W, Veeneman GH (1992) J Carbohydr Chem 11:837
101. Lohray BB (1992) Synthesis 1035
102. Solarek DB (1986) Phosphorylated starches and miscellaneous inorganic esters. In: Wurzburg OB (ed) Modified starches: properties and uses. CRC Press, Boca Raton, FL, p 97
103. Felcht UH (1985) In: Kennedy JF, Phillips GO, Wedlock DJ, Williams DJ (eds) Cellulose and its derivatives. Ellis Horwood, Chichester, p 273
104. Greenway TM (1994) In: Gilbert RD (ed) Cellulose polymers, blends and composites. Hanser, Munich, p 173
105. Phillip BJMS (1993) Pure Appl Chem A30:703
106. Hirano S, Kondo Y (1982) J Chem Soc Jpn 1622
107. Kurita K, Ichikawa H, Ishizeki S (1982) Macromol Chem 183:1161
108. BeMiller JN (2007) Whistler and BeMiller's Carbohydrate Chemistry for Food Scientists, 2nd edn. AACC International, St. Paul, MN, Chap 4
109. Santacesaria E, Trulli F, Brussani GF (1994) Carbohydr Polym 23:35
110. Besemer AC, van Bekkum H (1994) Starch 46:95
111. Fujimoto M, Fukami K, Tsuji K, Nagase T (1980) US Patent 4 186 024
112. Canteley M, Hough L, Pittet AO (1963) J Chem Soc 2527
113. de Noy AEJ, Besmer AC, van Bekkum H (1994) Rec Trav Chim Pays-Bas 113:165
114. Hoffman J, Larm O, Larsson K, Scholander G (1985) Acta Chem Scand Ser B 39:513
115. Allen TC, Cuculo JA (1973) J Poly Sci Macromol Rev 7:189
116. Snyder SL, Vigo TL, Welch CM, (1974) Carbohydr Res 34:91
117. Miyah H, Misaki A, Toru M (1973) Carbohydr Res 31:277
118. Horton D, Just E (1973) Carbohydr Res 29:73
119. Braun PJ, French D, Robyt JF (1985) Carbohydr Res 141:265
120. Defaye J, Driguez H, Gadelle A (1976) Appl Polym Symp 28:955
121. Horton D, Usui T (1978) In: Schweiger RG (ed) Carbohydrate sulfates. Am Chem Soc Symp Ser 77:95
122. Bredereck K (1967) Tetrahedron Lett 695
123. Steiner AB, McNeely WH (1950) US Patent 2 494 912
124. Yalpani M, Hall LD, Defaye J, Gadelle A (1984) Can J Chem 62:260
125. Yalpani M (1987) US Patent 4 683 298
126. Yalpani M, Hall LD (1982) J Polym Sci Polym Chem Edn 20:3399

6.2 Starch: Structure, Properties, Chemistry, and Enzymology

John F. Robyt
Laboratory of Carbohydrate Chemistry and Enzymology,
Department of Biochemistry, Biophysics, and Molecular Biology,
Iowa State University, Ames, IA 50011, USA
jrobyt@iastate.edu

In: *Glycoscience*. Fraser-Reid B, Tatsuta K, Thiem J (eds)
Chapter-DOI 10-1007/978-3-540-30429-6_35:

Abstract

Starch is a very important and widely distributed natural product, occurring in the leaves of green plants, seeds, fruits, stems, roots, and tubers. It serves as the chemical storage form of the energy of the sun and is the primary source of energy for the organisms on the Earth. Starch is composed of two kinds of polysaccharides, amylose and amylopectin, exclusively composed of D-glucose residues with α-(1→4) linkages in a linear amylose and α-(1→4) linkages and ~5% α-(1→6) branch linkages in amylopectin, both combined in a water-insoluble granule that is partially crystalline and whose size, shape, and morphology are dependent on its biological source. The properties, isolation, fractionation, enzymatic degradation, biosynthesis, chemical modification, and specific methods of analysis of starch are presented.

Keywords

Amylose; Amylopectin; Granules; Crystallinity; Hydrogen bonding; Hydrophobic bonding; Isolation; Fractionation; Gelatinization; Solubilization

Abbreviations

ADPGlc	adenosine diphosphate glucose
α-Glc-1-P	α-D-glucopyranosyl-1-phosphate
avg. d.p.	average degree of polymerization
CGTase	cyclomaltodextrin glucanyltransferase
d.s.	degree of substitution
DMSO	dimethyl sulfoxide
EDTA	ethylene diamine tetraacetic acid
FACE	fluorescence-assisted capillary electrophoresis
P_i	inorganic phosphate
PPO	2,5-diphenyloxazole
POPOP	1,4-bis-2-(5-phenyloxazoly)benzene
TEMPO	2,2,6,6-tetramethyl-1-piperidine oxoammonium ion
UDP-Glc	uridine diphosphate–glucose

1 Introduction, History, Development, and Uses

Starch is an abundant, naturally occurring polysaccharide, rivaling cellulose in the amount found on the Earth. It is found in the leaves of all green plants and in the seeds, fruits, stems,

roots, and tubers of most plants. Starch results as an end-product of photosynthesis and serves as the chemical storage form of the energy of the sun on the Earth. A high percentage of the energy available to non-photosynthetic organisms comes from starch, which is found in the principal food crops of the world: wheat, potatoes, rice, maize, barley, rye, beans, peas, sorghum, tapioca (or cassava), sweet potatoes, avocados, arrowroot, taro, bananas, mango, pineapple, sago (palm starch) and so forth.

It is estimated that 60–70% of the caloric intake by humans comes from starch. As such, starch has been of great importance in the evolution of organisms and especially for humans, where it also has had a role in the evolution of culture. Besides being used as an essential food, wheat starch was used to give body and the ability to hold ink to papyrus, a thin bark that was the earliest material used for writing (~4000 B.C.), especially in Egypt. When paper was developed in China, around 100 A.D., starch was also used to give body to the paper and to hold ink on the paper and it continues today to be used to size paper. The Romans (~100 B.C.) used starch to whiten cloth and to powder hair, and about 300 A.D., it was widely used to stiffen cloth and was often mixed with dyes to color the cloth. Colored starches (especially yellow and red) were also used as cosmetics.

Starch is one of three major biorenewable materials found on the Earth, sucrose and cellulose being the other two. Of the three, starch is today the most important because of its relative abundance and its relative ease of isolation in a highly pure form, which is relatively easily solubilized and enzymatically hydrolyzed to glucose and/or different maltodextrin products, or chemically modified. It, therefore, is a biorenewable, natural product that finds many industrial uses, such as the formation of glucose syrups, high-fructose syrups, maltodextrins with a wide range of average molecular weights, cyclomaltodextrins, D-glucitol (D-sorbitol) and the formation of ethanol, acetic acid, D-lactic acid, and other organic compounds by fermentation.

2 Isolation of Starch Granules

Because of the water-insolubility of starch granules, they are relatively easy to isolate from their plant sources. The source, for example seeds from maize, wheat, barley, rice, beans, and so forth are first steeped in water for 10–15 h at 50 °C. Steeping softens the outer parts of the seeds so the starch inside can be more easily obtained. The steep water often contains 1.63 g of sodium hydrogen sulfite per liter to remove protein. In some cases, the sulfite is omitted to retain the protein and give a starch as it more closely exists in the plant. The steeped material is then ground, milled, or blended to give a white insoluble suspension of starch that is filtered one time through six layers of cheese cloth and a second time through eight layers, to remove fibrous non-starch substances. The starch is allowed to settle and the water is poured off and fresh water is added and the starch is washed two or three times in this manner. The starch is then centrifuged or filtered and allowed to air dry at 20 °C. It is then lightly pounded into a fine powder.

If the starch is to be isolated from a tuber, stem or rhizome, the plant materials are washed and peeled and then cut into 1-cm pieces, placed in water, and blended. The resulting starch suspension is then obtained in a similar manner, as described above, beginning with filtering through the cheese cloth.

3 Chemical Composition of Starch

Most starches are composed of two kinds of polysaccharides, a linear α-(1→4) linked glucan, called amylose, and an α-(1→4) linked glucan with 4.2 to 5.9% α-(1→6) branch linkages, called amylopectin. See *Fig. 1* for the chemical structure of a segment of amylose and a segment of amylopectin and *Table 1* for the percent branching in amylopectin for several different starches.

The ratio of amylose to amylopectin also varies, depending on the source of the starch; it ranges from 17 to 70% amylose and a corresponding 83 to 30% amylopectin. See *Table 2* for a list of different starches and their composition of amylose and amylopectin. Usually amylopectin is present as the major component, with a common ratio of 1:3 or ~25% amylose and ~75% amylopectin. There are mutants, however, such as, waxy barley, waxy maize, waxy rice, waxy potato, and waxy sorghum that are 100% amylopectin and there are the high-amylose starches,

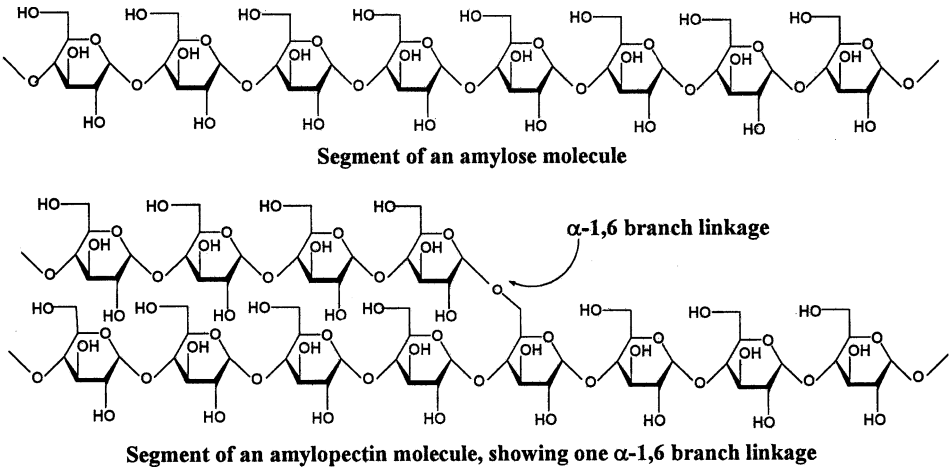


Figure 1
The structures of segments of amylose and amylopectin

Table 1
Percent branching of amylopectin from different starches

Amylopectins	% Branching[a]
Maize	4.2
Potato	4.5
Wheat	4.8
Barley	5.0
Oats	5.2
Waxy maize	5.9

[a]The avg. % branching was calculated from the avg. chain length determined by the enzymatic method [177]

Table 2
Composition of starch granules from ten plant sources

Starch	Amylose[a] (%)	Amylopectin[a] (%)	Lipid[b] (%)	Protein[b] (%)	Phosphate[b] (%)
Maize	25	75	0.80	0.35	0.090
Waxy maize	0	100	0.20	0.25	0.024
Amylomaize-5	53	47	0.70	0.30	0.090
Amylomaize-7	70	30	0.75	0.30	0.060
Potato	22	78	0.01	0.10	0.210
Wheat	23	77	0.90	0.40	0.180
Rice	19	81	0.59	0.30	0.090
Tapioca	17	83	0.02	0.10	0.009
Banana	20	80	0.48	0.32	0.060
Shoti	30	70	0.01	0.20	0.630

[a]The percent amylose and amylopectin were determined by precipitating the amylose with 1-butanol and then precipitating the amylopectin with two volumes of ethanol; the butanol-1 was removed and both were dehydrated by trituration with acetone several times and ethanol, followed by drying under vacuum at 40 °C, and weighed.
[b]From Ref. [178]

such as, amylomaize-5, with 53% amylose and 47% amylopectin, amylomaize-7, with 70% amylose and 30% amylopectin, and wrinkled pea with 66% amylose and 34% amylopectin.
The molecular weights of the two polysaccharides also vary, depending on the source of the starch. Amylose is much lower than amylopectin and both polysaccharides exist as a distribution of molecular weights, so that the determined molecular weight is an average. Amylomaize-7 amylose has an average degree of polymerization (avg. d.p.) of ~400 glucose residues or a number average molecular weight of ~64,800 Da; potato amylose has an avg. d.p. of ~1,000 glucose residues or a number average molecular weight of ~162,000 Da; and wheat amylose has an avg. d.p., of ~4,000 glucose residues or a number average molecular weight of 648,000 Da. Amylopectin is a much larger molecule, having number average molecular weights of 1×10^6 to 5×10^8 Da. The polydispersity of the sizes of the amylose component is lower than the amylopectin component.
It has been determined that for a starch granule of 15 μm diameter and a mass of 2.65×10^{-9} g, containing 25% amylose, with an avg. d.p. of 1,000 and 75% amylopectin, with an avg. d.p. of 100,000, there would be 2.5×10^9 molecules of amylose and 7.4×10^7 molecules of amylopectin in a granule [1].
Two minor carbohydrate components have been found in starch granules. The first is a water-soluble amylose that is slightly branched ~0.5 to ~2.0% [2,3,4]. The second is a very highly branched componentfound in 12 varieties of starch granules that were isolated and examined for their structure and properties and most probably occur in all starches. They ranged in amounts from 0.51% (w/w) in amylomaize-7 starch to 8.4% (w/w) in rice starch. They had relatively high molecular weights, 2.47 kDa for amylomaize-7 starch to 5.75 kDa for waxy maize starch, and a high degree of α-(1→6) branching, ranging from 15.6% for rice starch to 41.4% for shoti starch [5].
In addition to these major carbohydrate components, starch granules also contain minor non-carbohydrate components: lipids from 0.01 to 0.80% (w/w); proteins, from 0.10 to 0.40% (w/w); and phosphate 0.09 to 0.63% (w/w). See *Table 2* for the compositions of ten types

of starch granules. The lipids are primarily found in the cereal starches: maize, waxy maize, amylomaize-5, amylomaize-7, rice, wheat, barley, rye, and so forth; and the phosphate in these starches are part of the phospholipids. The phosphate in the tuber starches, canna, potato, shoti and others, is attached to the primary alcohol of the glucose residues in amylopectin. At least some of the protein, if not all of it, is the enzymes, starch synthases and starch branching enzymes that were responsible for the biosynthesis of the starch in the granule (see *Sect. 8*). In addition to the composition mentioned above, starch granules at equilibrium with their environment (20 °C) and humidity (40–50%) will absorb water and contain 10–15% (w/w) water of hydration. When starch granules are suspended in water, they swell to a limited extent and adsorb water to ~30% (w/w).

4 Occurrence of Starch as Water-Insoluble Granules

All starches are stored in plants as water-insoluble particles or granules in the chloroplasts in leaves and in amyloplasts in other plant tissues. The granules are relatively dense particles of compact molecules, with semi-crystalline properties. They were first observed in the early eighteenth century by van Leeuwenhoek, using his newly invented microscope. In the middle of the nineteenth century, Nägeli [6] described differences in the size and shapes of starch granules from various plant sources. In 1913, Reichert [7] published a book containing hundreds of photomicrographs of starch granules from widely different plant sources. In 1994, Jane et al. [8] published a comparative anthology of scanning electron micrographs of starch granules from 54 plant sources, showing widely different sizes of large (100 μm) to very small (0.2 μm) granules and different morphologies, such as smooth, oval, and spherical shapes, round pancake shapes with varying thicknesses, flat thin plates or lenticular-shaped disks, and irregularly shaped polygons whose edges have varying degrees of sharpness. Canna bulb starch granules are very large (60 × 100 μm) and potato starch granules are also large (20 × 75 μm). Many of the starches from the cereal grains, such as maize, waxy maize, oats, and sorghum are smaller (15–25 μm), with the exception of rice starch granules, which are quite small (3–5 μm). The cereal starches have irregular polygonal shapes, with a number of faces that have relatively sharp edges. Starches from beans and peas have smooth oval granules, with 10–45 μm diameters. They often are accompanied by an indentation in the center or at the end of the granule. There are some starches that are extremely small, such as the granules from Chinese taro (1–4 μm), amaranth (0.5–2 μm), and parsnip (1–3 μm); they generally have polygonal shapes, with sharp edges, similar to the cereal starches, but much smaller. Most leaf starches are very small, biconvex granules of ~1 μm in diameter. *Figure 2* illustrates these differences by pictures of scanning electron micrographs of nine starch granules with varying sizes and morphologies.

4.1 Structure of the Starch Granules

Starch granules also have varying degrees of crystallinity, as shown by X-ray powder pattern diffractions. Three distinct X-ray patterns have been observed, A-, B-, and *C*-patterns [9,10]. The A-patterns are characteristic of cereal grain starches, such as maize, waxy

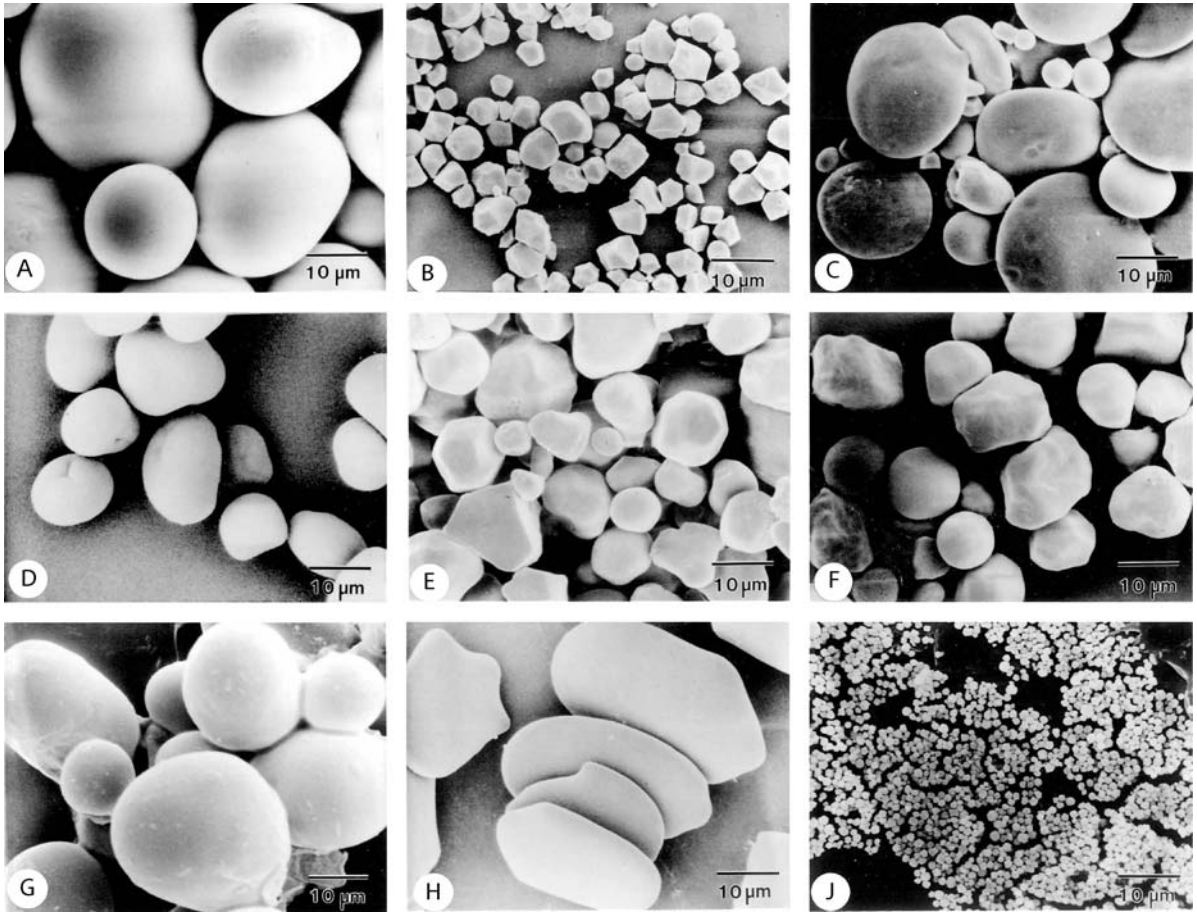


Figure 2
Scanning electron micrographs of starch granules from nine different botanical sources: A, potato; B, rice; C, wheat; D, mung bean; E, maize; F, waxy maize; G, tapioca; H, shoti; J, leaf starch. Magnification 1500X for each starch

maize, wheat, and rice. The B-patterns are characteristic of tuber, fruit, and stem starches, such as canna, potato, sago, banana starches and the mutant maize starches, amylomaize-5 and amylomaize-7 [9,10]. *C*-type patterns found in roots, beans, and peas are intermediate between A- and B-types [10,11]. Tapioca starch gives a typical *C*-type X-ray pattern. It has now been shown that *C*-type starch granules have a central core with a B-type structure, surrounded by an A-type structure [12].

When B- and *C*-type native starches are heat-moisture treated, 30 °C for B-type and ~50 °C for *C*-type, they are converted into A-type X-ray patterns [13,14,15,16]. Conversion of A-type starches into other crystalline forms only occurs when the original structure is completely destroyed and it is then allowed to recrystallize [17].

The interpretation of the X-ray diffraction patterns of starch granules and the significance of the A- and B-types of granules are problematic. There are differences in the ultrastructures of starch granules from different plant sources [18]. This was reinforced by a study of the kinetics and products of the reaction of glucoamylase with seven native starch granules in which three distinct categories were obtained: waxy maize starch was the most susceptible; barley, maize, and tapioca starches were intermediate; and amylomaize-7, shoti, and potato starches were quite resistant [19].

A general model for the crystalline parts of the starch granule has developed as a spherocrystalline assembly of disk-shaped amylopectin molecules [20] that are radially directed with the non-reducing-ends of the chains directed toward the surface of the granule [21,22]. The amylose chains in the cereal starches exist as lipid complexes that are left-handed extended helices, arranged tangentially in the granule. For non-complexed amylose chains in the cereal starches and in other starches, the amylose chains exist as both single and double helices [23,24]. In the granule, the chains of the amylopectin molecules can be considered as two-dimensional molecules that stack on top of each other, along with amylose chains, to make a three-dimensional micelle, giving radially oriented crystallites 9 nm in size. These units of both crystalline and amorphous phases form a network, extending throughout the granule in repeating units that are a common structural feature for all starches [25]. The amylose and amylopectin molecules form hydrogen bonds and hydrophobic bonds, both intermolecularly and intramolecularly, to varying degrees that hold the molecules together, giving a water-insoluble granule. The arrangement and quantity of these bonds in the granules gives rise to the differences in granule properties (degrees of crystallinity, types of X-ray diffraction patterns, gelatinization temperature, and susceptibility to enzyme hydrolysis [19,26,27]) found in starches from different plant sources. The α-(1→6) branch linkages in amylopectin are believed to occur in the amorphous regions of the granule. The branch linkages of amylopectin are clustered, with several chains of 15–20 glucose residues from the cluster, intertwined into double helices, contributing to the crystallinity [17,21].

4.2 Growth Rings in Starch Granules

Growth rings are readily seen in the large tuber starch granules, such as potato or canna starches, by optical microscopy. Growth rings are particularly apparent by scanning electron microscopy of granules that have been treated with acid or α-amylases [9]. ❯ *Figure 3* is a scanning electron micrograph of a potato starch granule that has been treated with *Bacillus*

Figure 3
Scanning electron micrograph of potato starch granule treated with *Bacillus amyloliquefaciens* α-amylase, showing growth rings of the crystalline α-amylase resistant starch in the granule

amyloliquefaciens α-amylase. During treatment with α-amylase, the amorphous layers in the granule are hydrolyzed and the crystalline layers remain intact. These so-called growth rings most likely originate during different phases of the biosynthesis of starch granules in which the deposition of crystalline layers of starch alternates with an amorphous layer of starch. These layers most probably are produced by fluctuations in the rate and/or mode of starch deposition during biosynthesis. Potato and canna starch granules show the growth rings the best when they are fully hydrated. When the starch granules are dried, the growth rings disappear [28].

4.3 Gelatinization of Starch Granules and the Solubilization of the Starch

The first step in the utilization of starch often involves the disruption of the granule, although by no means is this always the case (see *Sect. 9* on the chemical modification of starch granules). Disruption of starch granules can be accomplished by heating a suspension of starch granules in water, either by boiling the water (100 °C) or by autoclaving the suspension under pressure at 121 °C.

Starch granules swell when heated in water and have a "gelatinization temperature" that is characteristic of their source. When this temperature is reached, the swollen granules loose their birefringence, as evidenced by the loss of the "polarization cross" that can be observed with a polarizing microscope. The granules, however, still have a certain amount of structural integrity, but as the temperature is increased further, they continue to swell and eventually burst, releasing the amylose and amylopectin molecules, along with the minor components,

Table 3
Water solubility of seven native starches by autoclaving 220 mg in 10 mL at 121 °C for 20 min

Starches	Solubility (mg/mL)
Waxy maize	18.7
Tapioca	17.4
Potato	12.4
Maize	12.4
Rice	7.9
Amylomaize-7	5.5
Wheat	5.2

From Ref. [32]

into the aqueous solution. The heating process promotes increased motion of the molecules in the granule, eventually breaking the hydrogen and hydrophobic bonds between the molecules in the crystalline regions of the granule, which become hydrated and are released into the water. Cold, chemical methods of solubilization of starch can also be used, such as alkali (1 M NaOH) and dimethyl sulfoxide (DMSO) that gelatinize starch granules at low temperatures (20 °C). Gelatinization and the consequent solubilization of starch granules in DMSO occurs by a very different path from gelatinization in boiling water or autoclaving. Starch granules do not swell in DMSO as they do in boiling water [29,30], instead the granules slowly dissolve in anhydrous DMSO by fragmentation of the interior of the granules into smaller and smaller pieces. The addition of water to DMSO, however, greatly increases the rate of gelatinization and alters the pathway, which is different from gelatinization in anhydrous DMSO. Potato and canna starches retain their shape and birefringence during dissolution in aqueous DMSO and fragmentation does not occur. Instead the starch is peeled from the surface of the granules, with large pieces of starch gel adhering to the smooth surfaces of the ungelatinized portions of the granule [31]. When water is added, the DMSO breaks hydrogen bonds of the starch molecules at the surface of the granule and the water hydrates and gelatinizes the starch molecules at the surface, peeling the starch from the granules in layers [32].

Another method of chemically gelatinizing starch granules is by the use of relatively high concentrations of salt solutions to produce swelling and dissolution [33]; 4-M $CaCl_2$ has been used to gelatinize potato starch granules [34] and 13-M LiCl to gelatinize the starch from the surface of maize starch granules [35]. A method has been developed in which the surfaces of eight starch granules were peeled under controlled conditions, using aqueous-DMSO to give slow, gradual peeling from the surface of the granules in 8–10 h by selecting the amount of water from 5 to 15% (v/v) and the temperature from 0–15 °C [36].

There is evidence from two sources that the gelatinization process does not completely disrupt the hydrogen and hydrophobic bonds between the starch molecules, even though they are dissolved in the water. The first is the significant and greatly different activities of a single kind of α-amylase hydrolyzing solubilized starch from different plant sources [26]. The second is the differences that are observed in the water-solubility of starches from different plant sources and from acid/alcohol modified starches [32]. The maximum water solubilities for seven common starches are given in *Table 3*.

5 Fractionation of Starch into Amylose and Amylopectin

Early methods in 1940 for fractionating starch involved the leaching of an amylose component from swollen starch granules with hot water [37,38,39,40]. Starch granules (1–2% w/v) were successively treated with 70 °C water and the linear amylose component was obtained. This extraction, however, was found to be incomplete [41] and only the low molecular weight amylose was extracted. A few years later, the amylose and amylopectin components were completely separated by first solubilizing the starch granules in boiling water and selectively precipitating the amylose from the solution by 1-butanol [41,42]. This 1-butanol complex was crystalline, giving a V-type X-ray pattern [43]. After the removal of the 1-butanol-amylose precipitate by centrifugation, the amylopectin component was precipitated from the supernatant with two volumes of anhydrous ethanol. The 1-butanol can be removed from the amylose and the two components can be dehydrated by being treated several times with anhydrous acetone and once with anhydrous ethanol, followed by drying under vacuum at 40 °C to give dry powders. It is emphasized that the lipids complexed in the granules of the cereal starches must first be removed by defatting with methanol, ethanol, or 2-methoxyethanol (methyl cellosolve) extraction before solubilizing the granule and removing the amylose by adding an organic complexing agent [44], otherwise the removal of the amylose in these starches is incomplete.

Another organic compound that forms a complex with amylose is thymol, which also has been used to separate amylose from amylopectin [45]. Thymol is added to a starch solution as a 10% (w/v) solution in ethanol. Besides 1-butanol and thymol, many other organic compounds have been found to form complexes with amylose and give precipitates [46]. A list of organic compounds that form complexes with amylose and a list of those that do not form a complex and precipitate amylose are given in *Table 4*.

6 Properties of Amylose and Amylopectin

It has been known for almost 200 years that starch gives a deep blue color when a solution of potassium iodide and iodine is added [47]. More than a century later it was suggested that the complex consisted of a helical polysaccharide, with triiodide in the center of the helix [48]. Using flow dichroism, it was demonstrated that the triiodide was stacked in a linear structure, as required for the helical model [49]. Another study of the optical properties of crystals of the amylose-triiodide complex showed it to be consistent with a helical structure [50] and X-ray diffraction showed the triiodide complex gave the dimensions of a unit-cell of a helix with six glucose residues per turn [51]. This confirmed a helical structure for the amylose complex with triiodide that predated the helical models proposed by Pauling for polypeptides [52] and the double helical model for DNA by Watson and Crick [53] by 10 years.

The amylose-organic compound complexes give a particular kind of X-ray pattern that is called a "V-pattern" [54,55,56]. Amylose complexed with 1-butanol or 1-hexanol forms helices with six glucose residues per turn of the helix; branched chains or halogenated alkanes, such as tert-butyl alcohol (2-methyl-2-propanol) and tetrachloroethane form helices with seven glucose residues per turn of the helix [46,57,58,59], and bulkier molecules, such as 1-naphthol and thymol form helices with eight glucose residues per turn of the helix [60]. The length of the

Table 4
Hydrophobic compounds forming complexes with amylose and those that do not form complexes

Compounds forming complexes	
1,1,2,2-tetrachloroethane	bromocyclopentae
cycloheptane	1,2-dichloropropane
hexachloroethane	chloroform
(1.2.2)-bicycloheptene	fluorobenzene
1,1,2,2-tetrabromomethane	1,1,2-trichloroethane
cyclohexane	cyclopentane
carbontetrachloride	methylcyclohexane
cyclohexene	methylcyclopentane
cycloheptene	2,3-dibromobutane
cycloheptadiene	2,3-dimethylbutane
1,1,1-trichloroethane	thymol

Compounds that do not form complexes	
tetrachloroethylene	carbon disulfide
trichloroethylene	n-heptane
chlorobenzene	mineral oil
bromobenzene	1-bromopentane
iodobenzene	1-bromobutane
p-cymene	*p*-dibromobenzene
toluene	*p*-chlorobromobenzene
m-xylene	1,3,5-triethylbenzene
p-xylene	1-methylnaphthalene
mesitylene	

From Ref. [46]

helical chain that is folded back and forth, however, remains at 10 nm for all of the complexing agents.

α-Amylases hydrolyze the amorphous, folding areas on the surfaces of the lamella of packed helices, with the formation of resistant amylodextrin helical complexes of different degrees of polymerization (d.p.) values that are dependent on the number of glucose residues per turn of the helix: 1-butanol complex (six glucose residues per turn) gave d.p. 75 ± 4, tert-butyl alcohol and 1,1,2,2-tetrachloroethane (seven glucose residues per turn) gave d.p. 90 ± 3, and 1-naphthol (eight glucose residues per turn) gave d.p. 123 ± 2 [61]. See *Fig. 4* for the formation of these amylodextrin complexes by α-amylase hydrolysis of the surface of the folded amylose-organic complex.

The reason that amylose forms a complex with the organic molecules is that it is a relatively long, linear molecule, whose structural features inside the helix are hydrophobic, allowing hydrophobic bonding with the organic molecules, whereas amylopectin chains are much short-

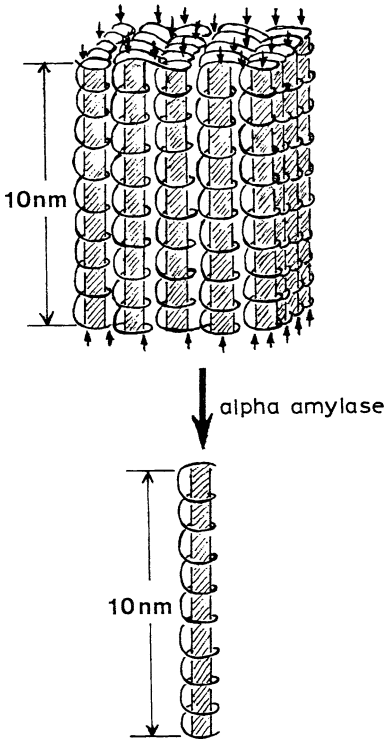


Figure 4
Alpha-amylase hydrolysis of the organic molecule (1-butanol, tetrachloroethane, or 1-naphthol) amylose complexes of the lamella-crystalline material in which the α-amylase hydrolysis takes place at the tops and bottoms (*arrows*) of the complexes, giving a resistant amylose-complex fragment of approximately the same size with avg. d.p. values of 108, 122, and 169, respectively, for the three organic complexes. From ref. [61], by permission of the authors and the publisher, Elsevier Press

er because of branching and do not form a sufficient length of helical complex with organic compounds that can give the lamella crystalline complexes, thus, allowing amylose to be separated from amylopectin.

It is the amylose component of starch that gives the blue color when KI/I_2 solution is added. To study the iodine-iodide color of amyloses of different d.p. values, maltodextrin-amylose molecules, with various avg. d.p. values from 6 to 568 were prepared by Bailey and Whelan [62], using phosphorylase, α-D-glucopyranosyl-1-phosphate, and maltohexaose. The colors of the various sized maltodextrins (1 mg) were observed when 10:1 (w/w) KI/I_2 solution was added. The first color to be observed was faint red for avg. d.p. 12; a red-purple color was observed for avg. d.p. 31; a purple color was observed for avg. d.p. 40; and a blue color was observed for avg. d.p. 45. The increase in the "blue value" was linear as a function of avg. d.p. up to $\sim$ avg. d.p. 60; the absorbance at 645 nm then slowly increased and reached a maximum at avg. d.p. of 400. The intensity of the iodine/iodide color in the low molecular weight range was dependent on the concentration of the iodine. When the concentration of the iodine was increased 10-fold, the intensity was increased 50% [62].

Amylose in a water solution is thought to behave as a random coil that has some helical character [63]. The random coiled amylose moves randomly in the aqueous solution with helices constantly being formed and deformed, giving a dynamic structure. An aqueous solution of amylose, 2 mg mL^{-1} or more is unstable and comes out of solution (retrogrades), as an insoluble precipitate, at a rate that is dependent on the concentration. At a concentration of 1 mg mL^{-1}, however, this same amylose does not retrograde and stays in solution indefinitely.

Acid hydrolysis of retrograded amylose (avg. d.p. 1,000) gives a resistant fragment with an avg. d.p. of 32 [61]. Hydrolysis of the retrograded amylose by human salivary α-amylase and porcine pancreatic α-amylase also gives a resistant fragment of avg. d.p. of 43, and hydrolysis by *Bacillus amyloliquefaciens* α-amylase gives a resistant fragment of avg. d.p. of 50 [61]. A structure for retrograded amylose was proposed in which crystalline double helical chains of 32 glucose residues of ~10 nm lengths are interspersed with amorphous regions to give resistant double helical structures. When the retrograded amylose is acid hydrolyzed for prolonged periods, most of the α-(1→4) glycosidic linkages in the amorphous regions are hydrolyzed, leaving the double helical, crystalline regions of 32 glucose residues. When the retrograded amylose, however, is hydrolyzed by α-amylases, they cannot hydrolyze the α-(1→4) glycosidic linkages close to the crystalline regions because of the numbers of their substrate-subsites, leaving stubs of glucose residues on the ends of the retrograded amylose chains, giving larger fragments than acid hydrolysis. Porcine pancreatic α-amylase has five subsites and human salivary α-amylase has six subsites, giving a chain with ~10–11 more glucose residues, 32 + 11, or 43 glucose residues, and *B. amyloliquefaciens* α-amylase has nine subsites, giving a chain with 18 more glucose residues than acid hydrolysis, nine glucose residues on each end, giving a total of 32 + 18 or 50 glucose residues [61].

Amylopectin acts very differently in aqueous solution. Even though it is 3- to 5-orders of magnitude larger than amylose, it is much more water soluble. This is primarily due to the presence of ~5% α-(1→6) branch linkages in the amylopectin molecule. The branch linkages of 50 for every 1,000 linkages give an average chain length of 20 glucose residues per chain. Because of the branching, the amylopectin chains are, thus, relatively short and, therefore, cannot line up in solution to give intermolecular double helices of sufficient length and they, hence, do not give a significant mixture of crystalline and amorphous regions in solution that retrograde. This then is the reason why amylopectin is much more soluble and stable in an aqueous solution than is amylose.

The amylopectin molecules are not compact or spherical as are glycogen molecules, which have 10–12% α-(1→6) branch linkages, randomly distributed. Amylopectin has relatively high viscosity, indicating an elongated structure. Electron microscopy, X-ray diffraction, chemical, and enzymological studies have shown that amylopectin has its branch linkages clustered at ~70 Å intervals [64]. Amylopectin molecules are ~100–150 Å in width and ~2,000–4,000 Å in length. Further, there is a regular alternation of elongated α-(1→4) linked chains and clustered α-(1→6) branch linkages that gives the 70 Å periodicity of the clustered branch linkages [64]. The amylopectin molecules can be considered to be two-dimensional, with the third dimension coming from stacking of the molecules through the formation of intermolecular hydrogen and hydrophobic bonds in the water-insoluble granule and also partially when in an aqueous solution [26,32]. See *Fig. 5* for a schematic drawing of amylopectin and its clustered branches and its branched chains.

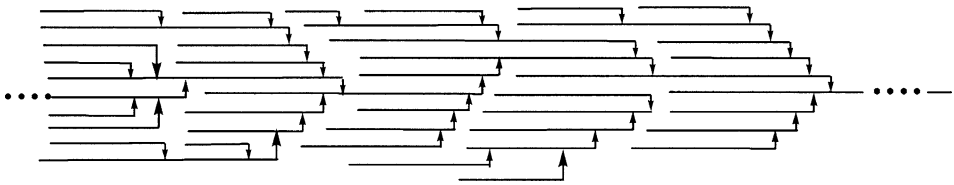


Figure 5
Schematic picture (looking from the top down) on a segment of amylopectin, showing the clustering of the α-(1→6) branch linkages in a repeating sequence

As was mentioned previously, the crystalline regions in the starch granules are primarily made up of amylopectin molecules, but with the amorphous regions being made up of the clustered branch linkages of amylopectin. The amylose molecule has one reducing-end and one non-reducing-end, and the amylopectin molecule also has one reducing-end, but many non-reducing-ends due to the branching.

Retrograded and dehydrated amylose is very water insoluble. It can, however, be put into an aqueous solution by first dissolving ~100 mg in 1 mL of DMSO by stirring and if necessary gently warming to 40–50 °C. After the amylose is completely dissolved in the DMSO, a thick clear gel-like solution is obtained. This gel can then be very slowly diluted with water, while stirring, to give the desired concentration of amylose in water. Amylopectin is readily dissolved in boiling water to give solutions of 2 to 5% (w/v).

7 Enzymatic Degradation of Starch

Several different kinds of enzymes degrade starch to give different kinds of products.

7.1 Endo-Acting Alpha-Amylase Hydrolysis of Solubilized Starch

The major category of enzymes that hydrolyze the α-(1→4) linkages of starch is the α-amylases. These enzymes are ubiquitous and are produced by bacteria, fungi, plants, and animals. In mammals, there are two specific sources, the salivary glands that secrete the enzyme into the mouth and the pancreas that secretes the enzyme into the small intestine. The α-amylases are endo-acting enzymes that attack the interior parts of the polymeric starch chains, producing a rapid drop in viscosity and a rapid change in the iodine-iodide color; because of these properties, they are sometimes called liquefying enzymes. When α-amylases encounter a starch chain and hydrolyze an α-(1→4) linkage, they also produce low molecular weight maltodextrins by a process called multiple attack on one of the two chains that were initially cleaved [65,66]. A common misconception about the action of α-amylases is that they act randomly, but this is not true. The major maltodextrin products that are formed are of specific sizes and depend on the particular biological source of the α-amylase and they, therefore, do not act randomly.

Human salivary and porcine pancreatic α-amylases primarily produce maltose, maltotriose, maltotetraose [67,68], and several kinds of α-(1→6) branched maltodextrins, ranging in size from four to seven glucose residues [69,70]. An even more common misconception about the

action of α-amylases is that glucose is the major product. Glucose, in fact, is only a very minor product, being produced by a slow secondary hydrolysis of the primary maltodextrin products. For example, salivary and pancreatic α-amylases slowly hydrolyzes maltotetraose to two molecules of maltose, the primary reaction; they also more slowly hydrolyze maltotetraose to give glucose and maltotriose and they hydrolyze maltotriose even more slowly to glucose and maltose; and they do not hydrolyze maltose at all [67,68].

A common bacterial α-amylase is excreted by *Bacillus subtilis* (whose name later was changed to *B. amyloliquefaciens*). It primarily produces maltotriose, maltohexaose, and maltoheptaose as the major products [71]. It also only produces glucose from a slow secondary hydrolysis of the primary maltodextrin products; it very slowly hydrolyzes maltohexaose to glucose and maltopentaose; it slowly hydrolyzes maltoheptaose to maltose and maltopentaose, and it slowly hydrolyzes maltoheptaose to maltohexaose and glucose; maltopentaose and smaller maltodextrins are not hydrolyzed [71]. Another α-amylase that is very similar to *B. amyloliquefaciens* α-amylase is the plant enzyme, barley malt α-amylase [72]. The bacterial α-amylase from *B. licheniformis*gives high yields (33%) of maltopentaose as a primary product [73]. It has a high temperature optimum of 70 °C and a broad pH optimum from 5 to 8, with significant activity in the alkaline pH range of 9–10 [74].

7.2 Exo-Acting Beta-Amylase and Glucoamylase Hydrolysis of Solubilized Starch

Starch exo-acting enzymes act exclusively at the non-reducing-ends of starch chains to form maltodextrin or glucose products. A primary exo-acting amylase is beta-amylase that is almost exclusively found in plants. The major β-amylases that have been studied are from sweet potato, barley, and soy beans [75]. β-Amylases catalyze the hydrolysis of the penultimate glycosidic bond at the non-reducing-ends of starch chains, forming β-maltose. Amylose is completely hydrolyzed; hydrolysis of amylopectin, however, only gives 55% β-maltose because the action of the enzyme is blocked by the α-(1→6) branch linkages that β-amylase cannot hydrolyze. This limitation leaves ~45% of a high molecular weight limit dextrin that contains all of the branch linkages of the original amylopectin molecules at the non-reducing-ends of the limit dextrins [76]. The first β-amylase from a source other than plants was reported from *B. polymyxa* [77]. It acted identically to the plant β-amylases. An α-amylase was reported from *Streptomyces* sp. that is an endo-acting enzyme, but produced 75–80% α-maltose [78].

Glucoamylase is another exo-acting enzyme that hydrolyzes the first glycosidic linkage at the non-reducing-ends of starch chains to give β-D-glucose. Many fungi, such as *Aspergillus niger*, *A. awamori*, *Rhizopus delemar*, and *R. niveus* elaborate glucoamylases [79,80,81,82,83]. In contrast to other types of amylases, glucoamylase can hydrolyze both α-(1→4) and α-(1→6) glycosidic linkages, although at different rates, with the α-(1→4) linkages being hydrolyzed ~600-times faster than the α-(1→6) linkages. But because of the hydrolysis of the α-(1→6) branch linkages, glucoamylases can completely hydrolyze all of the starch to β-D-glucose.

Another exo-acting starch-degrading enzyme is phosphorylase. It is produced by many plants and may be a primary enzyme involved in starch degradation in plants, although many plants

also have α- and β-amylases. Phosphorylase, however, is not a hydrolase and catalyzes the reaction between the α-(1→4) linkage of the terminal glucose residue at the non-reducing-ends of starch chains and inorganic phosphate to give α-D-glucose-1-phosphate [84]. Phosphorylase, like β-amylase, is unable to by-pass α-(1→6) branch linkages. It does not even remove all of the glucose residues close to the branch linkages and stops four glucose residues from the branch point on the two chains. Another enzyme, starch debranching enzyme, is required to hydrolyze the α-(1→6) branch linkages for further cleavage of the α-(1→4) linkages by phosphorylase.

Pseudomonas stutzeri produces an enzyme that forms maltotetraose from the non-reducing-ends of starch chains and, like β-amylase, cannot bypass the α-(1→6) branch linkages of amylopectin [85]. It gives 42% (w/w) α-maltotetraose and 58% limit dextrin. A similar exo-acting amylase is elaborated by *Aerobacter aerogenes* (*syn*: *Klebsiella pneumoniae* subsp. *aerogenes*; *Enterobacter aerogenes*) and exclusively forms maltohexaose from starch [86,87].

Another exo-acting starch enzyme forms cyclomaltodextrins. The enzyme is not a hydrolase, but a glucanyltransferase and is called cyclomaltodextrin glucanyltransferase (CGTase). The C4–OH group of the non-reducing glucose residue of starch chains is made to attack C1 of the sixth, seventh, and eighth α-(1→4) glycosidic bonds from the non-reducing-end to form cyclic, non-reducing α-(1→4) linked maltodextrins, containing six, seven, and eight glucose residues, respectively. Several different enzymes of this type are elaborated by different species of bacteria. The first CGTase to be isolated was from *Bacillus macerans*and it primarily forms cyclomaltohexaose, with smaller amounts of cyclomaltoheptaose and cyclomaltooctaose [88]. CGTases are also elaborated by *B. megaterium* [89] and *B. circulans* [90] that primarily form cyclomaltoheptaose, and a *Brevibacterium* sp. CGTase that primarily forms cyclomaltooctaose [91].

Isoamylases that exclusively hydrolyze the α-(1→6) branch linkages of starch were first recognized in plants and first isolated from broad beans [92]. A bacterial isoamylase was obtained from the culturing of *Pseudomonas amyloderamosa* [93,94,95,96] and has found wide use in studying the structure of amylopectin and related polysaccharides [97].

A new exo-acting enzyme, a starch lyase, has been discovered in red seaweeds and fungi. It removes glucose residues at the non-reducing-ends of starch chains and converts them into 1,5-anhydro-D–fructopyranose units. This product has been found to be an antioxidant and a new and versatile chiral building block for the formation of fine chemicals [98,99].

7.3 Reaction of Enzymes with Starch Granules

Starch granules generally have been thought to be resistant to hydrolysis by amylases [100], although as early as 1879 [101], they were reported to be digested by amylases. Ten percent hydrolysis by α-amylases of starch granules were reported [102]. In a study of the hydrolysis of starch granules by α-amylases from four sources (malt, fungi, bacteria, and pancreas), it was found that there were differences, with pancreatic α-amylase being the most effective, followed by malt, bacterial, and fungal α-amylases [103]. The activity of *B. subtilis* var. *amyloliquefaciens* varied according to the type of starch [104], with waxy maize being the most susceptible and high amylose starch (amylomaize-7 starch) the least susceptible.

The hydrolysis of starch granules by *R. niveus* glucoamylase also varied with the type of starch [105]. Reactions of glucoamylase on wheat and corn (maize) starch granules were followed visually by scanning electron microscopy and chemically by the determination of the amount of glucose released [106]. The scanning electron microscopy showed that wheat starch granules were attacked along the equatorial groove of the pancake-shaped granules and that extensive reaction with maize starch granules produced a Swiss-cheese appearance with many deep holes into the granules [106].

A systematic study of the kinetics of glucoamylase was made using seven kinds of starch granules, and it was found that they divided into three groups: waxy maize starch, which was the most susceptible, being converted with 200 units into 98% glucose in 32 h; barley, maize, and tapioca starches were intermediate, being converted into 82, 77, and 75% glucose, respectively, in 32 h of reaction; and the third group of amylomaize-7, shoti, and potato starches, which were the least susceptible, being converted into 21, 15, and 13%, respectively, in 32 h [107].

Scanning electron microscopy showed that when the starch granules in the first two categories were converted into ~50% glucose, there were numerous deep holes formed into the granules (see *Fig. 6*). Higher degrees of conversion showed that the remaining granules were hollow shells, with the interior parts hydrolyzed. It was concluded that glucoamylase enters the gran-

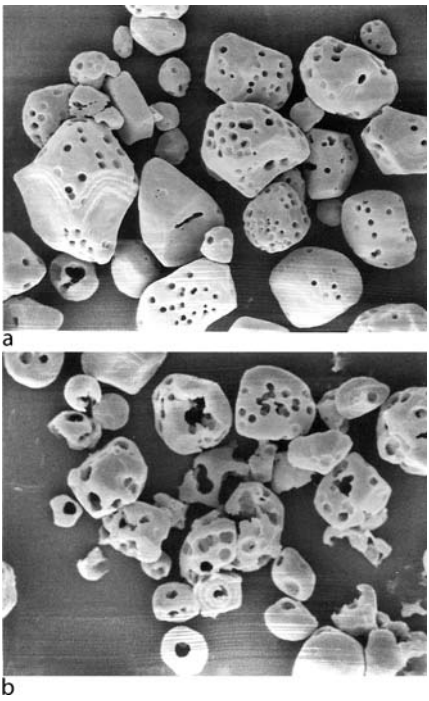


Figure 6
Scanning electron micrographs of maize starch granules that have been reacted with *Aspergillus niger* glucoamylase. A, 50 mg mL^{-1} starch granules, 20 IU mL^{-1} glucoamylase at 37 °C for 32 h; B, 50 mg mL^{-1} starch granules, 100 IU mL^{-1} glucoamylase at 37 °C for 32 h

ule and catalyzed reactions inside the granule [106,107]. The granules of the least susceptible starches in the third group did not show much change in the granule morphology, except for some minor pitting of the surface.

When the starches in the third group from potato, shoti, and amylomaize-7 were heated in water to their onset gelatinization temperatures of 60, 70, and 90 °C, respectively, for 0.5, 1, 2, and 4 h, the granules increased in size ~2–3 fold and birefringence was lost. Glucoamylase hydrolysis of potato and amylomaize-7 gelatinized starches was significantly increased after heating, from 13 to 86.6% for potato starch and from 21 to 45% for amylomaize-7 starch. The glucoamylase hydrolysis of shoti starch, however, only increased slightly from 14.7 to 21.1%, after being heated for 4 h. The results after 4 h of heating of shoti starch granules were only slightly higher than the results after 0.5 h [108].

Two of the three starches that were quite resistant to glucoamylase hydrolysis, amylomaize-7 and shoti starches, gave the highest percent reaction with the α-(1→6) debranching enzyme, *Ps. amyloderamosa* isoamylase of 11.9 and 11.5%, respectively [109]. Waxy maize, maize, and barley starches were intermediate in their susceptibility to isoamylase hydrolysis, giving 7.8, 7.3, and 6.2%, respectively.

Heating at the gelatinization temperatures of the starch granules for 1 h significantly increased the reaction with isoamylase: potato starch increased from 3.7 to 56%; amylomaize-7 starch increased from 11.9 to 36%; waxy maize starch was increased from 7.8 to 36%; and shoti starch was increased from 11.6 to 22%.

When gelatinized waxy maize starch was hydrolyzed 36% with isoamylase, only 16% of the hydrolyzed chains remained inside the granule. When 40% (v/v) ethanol was added to the waxy maize reaction digest, however, 44% of the isoamylase hydrolyzed chains remained inside the granule, and when 80% (v/v) ethanol was added, 100% of the hydrolyzed chains remained inside the granule [109]. With an 80% ethanol concentration, 100% of the isoamylase hydrolyzed chains also remained inside the granules for potato and shoti starches, and 98% of the isoamylase hydrolyzed chains remained inside the granules for amylomaize-7 starch for a total percent isoamylase hydrolysis of 41, 10, and 13%, respectively [109].

An in situ method has also been developed to retain 100% of the glucose inside the granule by reacting it with glucoamylase in a minimal amount of water, 50 mg of starch in 1.0 mL of buffer, containing glucoamylase [110]. This was called the "solid-granule" reaction system. The buffer, with glucoamylase, completely enters the granules and reaction takes place inside the granules. Although eventually the concentration of glucose produced in the granules is relatively high, only a very low amount (< 0.01%) of reversion products (maltose, isomaltose, nigerose) were formed. Using a sealed vessel, it was possible to convert waxy maize starch granules into 51% glucose, all inside the granules. The amount of starch conversion was proportional to the number of non-reducing ends. It was possible to control the percent conversion by heating the reaction vessel at different times for 20 min at 110–120 °C to inactivate the glucoamylase and thereby control the amount of glucose produced to give 10, 20, 30, 40, and 50% hydrolysis. The granules with 20% or more glucose have a sweet taste [110].

A similar reaction was conducted with cyclomaltodextrin glucanyltransferase (CGTase) reacting with waxy maize starch granules in a minimal aqueous environment. The yield was disappointing, giving only 1.3% (w/w) cyclomaltodextrins inside the granules. This could be increased to 3.4%, with 100% retention of cyclomaltodextrin inside the granule by adding a combination of CGTase and isoamylase to the solid-granule reaction system [111].

In general, the rates of reaction of enzymes with native starch granules is 3- to 4-orders of magnitude slower than they are with solubilized starch. The reaction of starch granules with enzymes is one method of modifying and altering the structure of starch granules. See *Sect. 9* for chemical methods of modifying starch granules. It also is important to recognize that a single enzyme, reacting with starch granules from different plant sources or with solubilized starches from different plant sources will have significantly different activities on the different starches [26]. Furthermore, this activity can be greatly increased and the enzyme stabilized, by the addition of 0.04% (w/v) polyethylene glycol [26], as a specific-sized polyethylene glycol will give the maximum enzyme activity for a specific enzyme, however, the activities will still remain significantly different on the different starches [112].

8 Biosynthesis of Starch

Starch is biosynthesized in plant organelles. In leaf or transitory starch, the organelles are chloroplasts and in storage starch in seeds, tubers, stems, and roots, the organelles are amyloplasts. The process of starch biosynthesis is divided, for convenience of discussion, into three steps: (1) initiation, (2) elongation, and (3) termination, and for amylopectin, a fourth step, branching.

Starch chains were the first polysaccharides to have a mechanism proposed for their biosynthesis. This was in 1940 and was based on the studies of Hanes on the action of potato phosphorylase [113]. Hanes proposed that potato phosphorylase transferred glucose from α-D-glucopyranosyl-1-phosphate (α-Glc-1-P) to the non-reducing ends of the starch chains and maltodextrins to give elongation of the starch chains. For starch biosynthesis by phosphorylase, it was recognized that the synthesis required a primer molecule that was either amylose, amylopectin, or a maltodextrin with a minimum of three glucose residues [114,115]. Phosphorylase is an enzyme that catalyzes a reaction that is reversible. The reaction can be either degradative or synthetic, depending on the starting substrates. It is degradative, if it is inorganic phosphate (P_i) or synthetic if it is α-Glc-1-P, as shown in the following reaction:

$$\underset{\text{starchchain}}{P_i + G-G-G-G-G-\cdots} \xrightarrow[\text{synthetic}]{\text{degradative}} \underset{\alpha\text{-Glc-1-P}}{G-P} + \underset{\text{degraded starchchain (putative primer)}}{G-G-G-G-\cdots} \qquad (1)$$

The synthetic reaction of the putative primer chain, however, only will occur when the ratio of P_i to α-Glc-1-P is less than the equilibrium value, which is 10.8 at pH 5, 6.7 at pH 6, and 3.1 at pH 7.0 [116]. These ratios are not obtained in vivo, as the concentration of P_i is 20- to 40-fold higher than the concentration of α-Glc-1-P [117,118] so that the in vivo conditions greatly favor degradation by phosphorylase, rather than synthesis.

Because a degraded starch chain is the product of the degradation reaction, the synthetic phosphorylase catalyzed reaction would require a starch chain or a maltodextrin chain to be a substrate in the synthetic reaction, as shown above. This was the origin for the primer required reaction for starch biosynthesis from the non-reducing-end of the primer, and it has pretty much been retained for 60 years [119,120,121].

In 1960, some 20 years after the phosphorylase experiments, it was found that UDP-Glc and ADPGlc were the high-energy glucosyl donors for the biosynthesis of starch chains, rather than α-Glc-1-P, and that active starch synthesizing enzymes, starch synthase and starch branching enzyme were entrapped inside the starch granules [122,123]. When ADP-[^{14}C]Glc was incubated with starch granules, ^{14}C-labeled glucose was incorporated into the starch in the granules. When this labeled starch was reacted with the exo-acting beta-amylase, ^{14}C-labeled maltose was obtained and it was assumed that the glucose from ADPGlc was, therefore, being added to the non-reducing-ends of starch primers [122,123,124]. This experiment has been widely considered as proof that starch is biosynthesized by the addition of glucose from ADPGlc to the non-reducing-ends of starch primer chains. This conclusion, however, is not necessarily correct in that if starch chains were synthesized from the reducing-end de novo, independent of a primer, the synthesized starch chains would have every glucose residue labeled when the starch granules were reacted with ADP-[^{14}C]Glc and reaction of this chain with beta-amylase would also give ^{14}C-labeled maltose.

Because of these problems and the lack of definitive experiments supporting the mechanism of starch chain elongation, Mukerjea and Robyt carried out a number of experiments to test the non-reducing-end, primer mechanism [125,126]. The first series of experiments was the pulsing of starch granules from eight different plant sources with ADP-[^{14}C]Glc and the chasing of the pulsed starch granules with non-labeled ADPGlc [125]. The pulsed starch granules gave starch chains that on reduction with $NaBH_4$ gave ^{14}C-labeled D-glucitol. The chased starch gave a significant decrease of ^{14}C-labeled D-glucitol. These experiments indicated that the starch chains were being elongated by the addition of glucose to the reducing-end and not to the non-reducing end, because if the glucose was being added to the non-reducing-end of a primer, it would be impossible to obtain any ^{14}C-labeled D-glucitol and further it would also be impossible to chase the label from the D-glucitol. In addition, it was found that both labeled-amylose and labeled-amylopectin components were synthesized and that a significant amount of starch was synthesized: maize starch had an avg. labeled d.p. of 827, waxy maize 436, taro 462, rice 467, wheat 476, and potato 524. There was also evidence that covalent intermediates were formed with the starch synthase during synthesis [125].

A two-site insertion mechanism for starch biosynthesis from the reducing-end was proposed, as shown in *Fig. 7*. The initiation step occurs with the glycosylation of the two sites, with the two nucleophilic X-groups attacking C1 of the glucose residues of ADPGlc. Polymerization proceeds by the C4–OH group of one of the covalently linked glucose residues attacking the C1 group of the other, displacing the X-group, which attacks C1 of another ADPGlc. The C4-OH of this glucose residue then attacks the C1 of the maltosyl unit, which is transferred to the glucose residue, freeing the other X-group, which makes an attack onto the glucose residue of another ADPGlc. This process continues, going back and forth between the two X-groups, giving the polymerization of the starch chain by addition of glucose from ADPGlc to the reducing-end of the growing chain. The starch chain is extruded from the active site of starch synthase. The elongation of the chain is eventually terminated by hydrolysis of the chain from the active site.

In the second set of experiments by Mukerjea and Robyt, the effect of adding the putative primers to the reaction of ADP-[^{14}C]Glc was studied with three kinds of starch granules from maize, wheat, and rice [126]. The reactions were examined in the absence and presence of increasing concentrations of maltose (G2), maltotriose (G3), and maltodextrin (d.p. 12). All

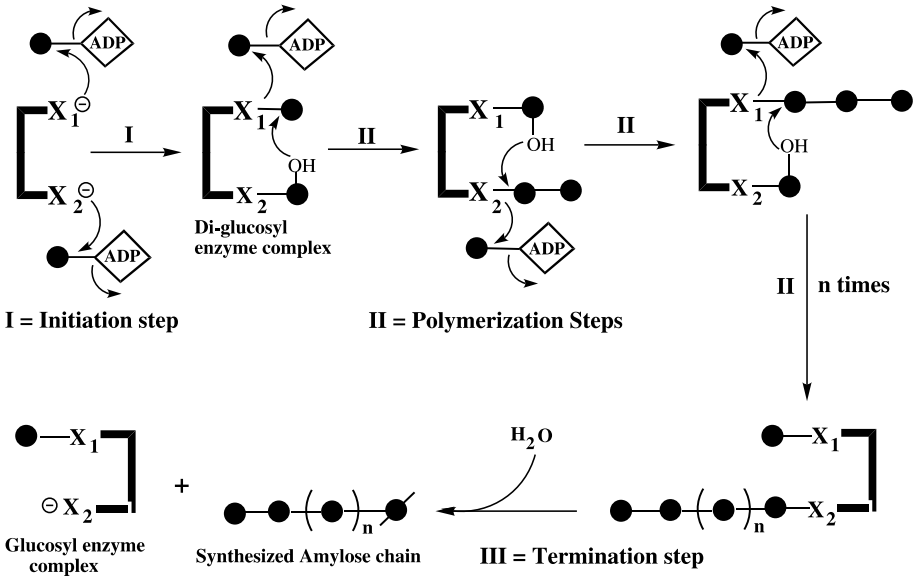


Figure 7

Proposed mechanism for the elongation of starch chains by starch synthase, showing the initiation step (I) and the polymerization steps (II) in which D-glucopyranose residues (*circles*) are added to the active site groups (X^-) from ADPGlc and inserted between X and the reducing-end of a growing glucan chain by a transglycosylation reaction to give a polymer of α-(1→4) linked D-glucopyranose residues. The polymerization is terminated (III) when water enters the active-site and hydrolyzes the X-acetal linkage of a reducing-end glucose residue attached to the starch chain. The mechanism gives a primer free elongation of the starch chain from the reducing-end. The *dark circles* are glucose residues and the *dark circle with a line through it* is a free hemi-acetal, reducing-end glucose residue

of the added putative primers were found to inhibit starch synthesis in increasing amounts, as the concentrations of the maltodextrins were increased, rather than stimulating synthesis, as would be expected for primers [126].

The putative primers did undergo a reaction. The major product in the presence of G2 was G3, with exponentially decreasing amounts of G4–G12; in the presence of G3, the major product was G4, with exponentially decreasing amounts of G5–G12. Further, the addition of glucose was to the non-reducing end of the putative primers. It was concluded that the putative primers were undergoing an alternate reaction in which they were acceptors that released glucose from the active site of starch synthase, with the glucose being added to C4–OH of their non-reducing-ends. This was confirmed by adding the acceptors to pulsed starch granules and observing the release of ^{14}C-label from the granules in the absence of ADPGlc [126]. This study showed that primers are not involved in starch biosynthesis and in fact their presence inhibited starch biosynthesis.

Leloir and co-workers [122,123,124] and Frydman and Cardini [127] also found that the incubation of starch granules with maltodextrins and UDPGlc or ADPGlc exclusively transferred glucose to the non-reducing-ends of maltose, maltotriose, and maltotetraose primarily to give

the next higher homologues, but they concluded that this was due to the primer non-reducing-end mechanism and did not recognize that starch synthesis was inhibited. This transfer of glucose to the non-reducing-end of carbohydrates is a well-known reaction in the synthesis of dextran by dextransucrase when low molecular carbohydrates are added to the reaction digest, glucose is transferred to C6–OH of the non-reducing-ends of the carbohydrates [128,129]. The carbohydrates are acceptors that inhibit dextran synthesis [129] just as the putative primers in starch biosynthesis accepted glucose and inhibited starch biosynthesis [126].

The third set of experiments by Mukerjea and Robyt involved two reactions with three varieties of starch granules from maize, wheat, and rice. In reaction I, the granules were reacted with 1-mM ADP-[^{14}C]Glc and in reaction II, a portion of the granules from reaction I was reacted with 1-mM nonlabeled ADPGlc [130]. The starch granules from the two reactions were solubilized and reacted with the exo-acting enzymes, glucoamylase and beta-amylase to an extent of 50% or less of the ^{14}C-label. The amounts of ^{14}C-labeled products from the two enzymes, respectively, were nearly equal for reaction I and reaction II. If the addition of glucose had been to the non-reducing-ends of primers, reaction II would not have given any labeled products. These results confirm that the elongation of the starch chain is the addition of glucose to the reducing-end by a de novo reaction and not by the addition of glucose to the non-reducing-end of a primer [130].

It must be emphasized that the mechanism of starch biosynthesis does not require a preformed starch chain primer or a maltodextrin chain primer and the starch chains begin de novo by forming covalent di-glucosyl intermediates from ADPGlc with starch synthase that proceed to synthesize amylose chains by adding glucose to the reducing-end of a growing chain, essentially by a transferase reaction of the enzyme covalently linked starch chain to the enzyme covalently linked glucose (see ❯ *Fig. 7*).

To produce amylopectin, it is necessary to have a starch branching enzyme that is capable of introducing the α-(1→6) branch linkages into the linear amylose molecules synthesized by starch synthase. Such an enzyme (Q-enzyme) was first identified in potatoes [131] and then in broad beans [132]. The potato Q-enzyme has been isolated and purified and the properties determined [133]. Q-enzyme is a glucanyltransferase and cleaves an α-(1→4) linkage of an amylose chain, and then transfers the chain to the C6–OH of another amylose-chain, rather than being transferred and attached to the residual part of the cleaved chain [134]. It was shown that water is not involved in these branching reactions [133].

There remain a number of important questions as to how branching enzymes convert the synthesized amylose chains into amylopectin and how the branch linkages appear in clusters in amylopectin, with a repeat distance of 70 Å. Another question is how starch granules are initiated and formed and how the percentages of amylose and amylopectin molecules are in a constant ratio that is specific for the particular plant source of the starch (see ❯ *Table 2*).

A logical assumption that can be surmised is that the enzymes that are found entrapped in the starch granules have most likely been responsible for the synthesis of the starch molecules in the granules. It has been hypothesized that starch synthase might exist both as a single enzyme and as a multi-enzyme complex of several starch synthases and branching enzymes. In the former case, amylose would be produced, and in the latter case, the multi-enzyme complex could produce amylopectin [135,136].

9 Chemical Modification of Starch

The chemical and physical properties of starch can be significantly altered by chemical modifications.

9.1 Acid Modifications of Starch Granules

One of the first reported modifications of starch granules was the treatment with aqueous acid in which there was hydrolysis of some or all of the glycosidic linkages. Obviously the molecular size of the polysaccharide molecules is decreased by acid hydrolysis. In 1874, Nägeli reported the treatment of native starch granules with 15% (w/v) sulfuric acid for one month at ~20 °C. An acid-resistant material resulted that was insoluble, but when isolated it could be solubilized in hot water [137]. This resistant material was called Nägeli Amylodextrin. It has an avg. d.p. of 24–30 and was fractionated by gel permeation chromatography to give three fractions [138]. Fraction I had the highest molecular weight and consisted of a mixture of branched maltodextrins with two or more branches per molecule. Fraction II, with a d.p. of ~ 25, primarily had a single branch linkage, with the branch located close to the reducing end [138]. Fraction III was linear with an avg. d.p. of 13 [138]. With continued acid hydrolysis, Fraction I disappears and Fractions II and III are better resolved on Sephadex G-50.

A soluble starch of much higher molecular weight than the Nägeli Amylodextrin was prepared by treating a suspension of starch granules with 7.5% (w/v) HCl for 1 week at 22–24 °C, or 3 days at 40 °C, followed by washing the granules free of acid [139]. The method is called the Lintner Procedure and is usually performed on potato starch granules. The Lintner Procedure is the method still used today to produce commercial soluble starch. The product is not completely water soluble and is polydisperse, with a relatively high reducing-value. The amylose component of Lintner soluble starch has an avg. d.p. of 340 and the amylopectin component has an avg. d.p. of 603, with a wide distribution of d.p. values [140].

Another method was reported in 1919 for obtaining soluble starch by refluxing the starch granules in 95% ethanol, containing 0.2–1.6% (w/v) HCl for 6–15 min [141]. The product is called Small's soluble starch. It is soluble (10 mg ml^{-1}) in hot water and has a low reducing value due to the formation of ethyl-glycosides at the hydrolytic points.

In 1987, Ma and Robyt reported the modification of starch granules by treating them with 0.36% (w/v) HCl in anhydrous methanol, ethanol, 2-propanol, and 1-butanol at 65 °C for 1 h. The starch granule was maintained, but the molecular sizes of the starch molecules progressively decreased in the different alcohols, as the number of carbons in the alcohols increased [140]. The molecular size distribution was greatly narrowed for these acid-hydrolyzed starches and the amylose component completely disappeared in the starch granules modified in 2-propanol and 1-butanol. The starch granules modified in 1-butanol also became quite soft. It was also found that acid hydrolysis was occurring inside the granules where there was 10–15% (w/w) water and the amount of acid also increased inside the granules as the size of the alcohols increased [140].

Fox and Robyt reported the kinetics of the acid-alcohol modifications by determining the avg. d.p. vs. time of reaction of the starch granules with 0.36 and 6.0% (w/v) HCl at 25 °C [142]. The kinetics showed that the avg. d.p. values dropped rapidly in the first 10 h and then leveled

off, becoming constant at ~30 h, d.p. values decreasing as the size of the alcohols increased and as the concentration of the acid increased. The avg. d.p. values varied from a high of 1717 to a low of 19, depending on the kind of alcohol and the acid concentration. The products represented new kinds of acid-hydrolyzed limit dextrins, with a relatively narrow distribution and widely different avg. d.p. values, and different proportions of amylose and amylopectin.

The acid hydrolysis of starch granules in the four alcohols was also studied as a variation of temperature from 5 to 65 °C, acid concentration from 0.36 to 5.0% (w/v), and in the concentration of the starch granules [143]. The avg. d.p. values of the limit dextrins were dependent on all three parameters: first, the acid concentration and the temperature, of which combinations can be obtained to give limit dextrins with avg. d.p. values ranging from 1800 to 12 for the starch granules. In addition, the concentration of the starch granules was inversely proportional to the avg. d.p. values. This latter result confirmed that the mechanism of the hydrolysis of the glycosidic bond was taking place by the water inside the granules, because as the concentration of the starch granules increased, the concentration of the acid decreased inside the granules, due to the fact that there were more granules.

The acid hydrolysis of potato starch granules was studied using mixtures of two alcohols, methanol and 2-propanol from 90:10 to 10:90 in 10% (v/v) intervals [144]. The avg. d.p. of the different modified starches leveled off in 48–72 h of reaction at 20 °C. The results gave a family of 11 curves, starting with 100% methanol and ending with 100% 2-propanol, all with different limiting avg. d.p. values. As the concentration of 2-propanol was increased and methanol decreased, the plots of avg. d.p. of the limit dextrins vs. the ratios of the two alcohols resembled titration curves, with different plateaus, in which the d.p. of the limit dextrin changed very slowly, or not at all. These plots were characteristic and dependent on the types of alcohols, the volume ratios of the two alcohols in the mixture, and the kind of starch [144]. It was proposed that the different ratios of the two alcohols produced the exposure of different glycosidic linkages to acid hydrolysis by converting different crystalline regions in the granule into amorphous regions that became susceptible to acid hydrolysis.

Another type of acid-modified starch is "thin-boiling" starches. Thin-boiling starches are produced by mild hydrolysis with 0.1–0.2% (w/v) HCl at 45–50 °C [145]. The treatment weakens the forces, holding the starch molecules together in the granule to give a starch that is more easily solubilized and has decreased viscosity. In the production of thin-boiling starches, the viscosity is monitored and the process is stopped when a particular desired viscosity is attained. The thin-boiling starches are widely used in the food industries to improve the appearance (clarity), stability, and texture of the food products in soups, sauces, dressings, bakery products, dairy products, and confectioneries.

9.2 Chemical Modification of the Starch Molecules in the Granules

The majority of the D-glucopyranose residues in starch granules have three hydroxyl groups, one each at C-2, C-3, and C-6, which can be modified. The amount of modification is called the "degree of substitution" (d.s.). If all three of the hydroxyl groups on all of the glucose residues are modified, the d.s. = 3; if two are modified, the d.s. = 2; and if only one is modified, the d.s. = 1; if one hydroxyl group out of every ten glucose residues is modified, the d.s. = 0.1, and so forth.

Three general kinds of modifications can be obtained: oxidation, esterification, and etherification to enhance hydrophilic character, to introduce hydrophobic character, to introduce positive or negative charges, to cross-link the starch chains, and to introduce color to the starch. The purposes for the modifications is varied, for example, to lower the gelatinization temperature, to decrease or prevent retrogradation, that is to decrease or prevent precipitation from solution, to enhance gelatinization, to increase the hydration of the starch molecules, to impart increased thickening and gelling properties, to increase binding of other molecules to starch, and to impart film-forming properties.

The function of modified starch is dependent on a number of things: the kind of starch (plant source), the amylose to amylopectin ratio, the nature of the substituents that are added, the degree of substitution, whether it is premodified by acid hydrolysis, and whether the starch is granular or gelatinized.

9.2.1 Oxidation of Starch

Alkaline hypochlorite is the most common method of oxidizing starch granules. The reaction is the oxidation of the secondary alcohols to ketones and the primary alcohol to a carboxyl group [146,147,148]. Alkaline hypochlorite oxidation is sometimes used to produce thin-boiling starches instead of acid hydrolysis.

Primary alcohol groups can be exclusively oxidized to aldehyde groups with pyridinium dichromate [149,150] and to carboxyl groups with the 2,2,6,6-tetramethyl-1-piperidine oxoammonium ion (TEMPO) [151]. The aldehydes can then be reduced to primary alcohols by reaction with NaB^3H_4 [150,152], giving radiolabeled ^{3}H-starch and the carboxyl group can be inverted by the action of *Azotobacter vinlandii* poly-β-D-mannuronic acid C-5-epimerase to give L-iduronic acid [153].

Vicinal alcohol groups, the 2,3-glycol structure of starch, can be oxidized by periodic acid or periodate to cleave the carbon–carbon bond and oxidize the two carbons to produce two aldehyde groups at C-2 and C-3 [154,155]. The product is called "dialdehyde starch." There are, however, very few actual aldehyde groups present. The main products are hydrated aldehyde groups and intra- and inter-molecular hemiacetals [154,156,157]. The amount of oxidation can be controlled by the amount of periodate used. The resulting "aldehydes" can react with alcohols, amines, hydrazines, and hydrazides to give additional modification products. Chlorous acid will oxidize the "aldehyde groups" to polycarboxylic acid polymers [154].

9.2.2 Formation of Ester Derivatives

Common starch esters include the acetates. High d.s. starch acetates have been formed by using acetic anhydride with either sodium acetate or pyridine catalysts at 90–100 °C [158,159,160]. The major use of starch acetates have been in the study of the structures of amylose and amylopectin after acid hydrolysis. Acetylation of granular starch in aqueous suspension by acetic anhydride at pH 10–11 is used to produce low d.s. starch acetates that are primarily used for the stabilization of their viscosity and for their water-soluble clarity. The acetylation decreases the hydrophilic character and increases the hydrophobic character of starch. The major uses of starch acetates is in the paper industry for surface sizing to give improved print quality, uniform porosity, surface strength, and resistance to various solvents [161].

The solubility of starch acetates depends on the d.s. and the d.p. Starch acetates of 10–15% are soluble in water at 50–100 °C and are insoluble in organic solvents. Starch acetates of 40% or more are soluble in aromatic hydrocarbons and halogenated aliphatic hydrocarbons.

Starch esters formed by reaction of acid chlorides give a wider range in the kinds of esters that can be synthesized. In particular, reaction of oxalyl dichloride gives an ester with a free carboxylate group; reaction with succinyl dichloride also gives an ester with a free carboxylate group, but in addition can give intra- and inter-molecular cross-linking of the starch chains; adipyl dichloride gives the same kind of reactions, but with greater potential for intermolecular cross-linking due to its larger size.

Starch granules have been cross-linked with phosphate by the reaction of an aqueous alkaline (pH 8–12) suspension reaction with phosphorus oxychloride [161]. Trimetaphosphate has also been used to produce phosphate cross-linkages. These reactions are primarily with the C-6–OH groups and only a minimal number with the C-3–OH groups [162]. Starch phospho esters can be obtained by phosphorylation with sodium triphosphate at pH of 8.5 [163].

The esterification of starch with p-toleunesulfonyl (tosyl) chloride in pyridine gives partial 6-*O*-tosyl starch [164,165]. This ester can be replaced by a number of nucleophilic groups, such as iodide to give 6-deoxy-6-iodo starch that in turn can be reduced with $NaBH_4$ to give partial 6-deoxy-starch or azide groups that on reduction give partial 6-amino-6-deoxy starch [166,167].

9.2.3 Formation of Ether Derivatives

Ethers are formed with the alcohol groups of starch when alkyl halides or epoxides are added to an alkaline suspension of starch granules. The starch ethers that are produced are frequently water-soluble. Methyl or ethyl ethers can be formed by the reaction of starch at pH 10 with methyl or ethyl chlorides at 70–120 °C to give d.s. values of 1.2–2.3 [168]. A common ether is the formation of 6-*O*-carboxymethyl starch by the reaction of 2-chloroacetic acid with alkaline starch [169]. Carboxymethyl starch is a water-soluble starch that is used as an emulsifier and thickening agent and provides stability against heat damage and amylase degradation.

Hydroxy ethers of starch are prepared by reaction of alkaline starch with epoxides, such as ethylene oxide and propylene oxide. The reaction is usually run on a 40–45% (w/v) suspension of starch granules in water under strong alkaline conditions (0.1–0.2-M NaOH) at 40–50 °C, under nitrogen. The hydroxy ethyl starches are primarily used as binders for pigmented coatings and as surface sizing agents in paper manufacturing. Hydroxypropyl starches are of importance in food applications, where they are used as an edible, water-soluble film coating [161]. The formation of hydroxyalkyl derivatives increases their water solubility and decreases or prevents retrogradation of the starch chains. Ethylene oxide and propylene oxide do undergo “chaining reactions” in which they keep adding to the newly generated hydroxyl groups to give a chain of hydroxy alkyl ethers. This chaining reaction can be reduced or eliminated by conducting the reaction in the presence of methyl or ethyl chloride, which cap the newly formed hydroxyl groups, preventing further additions.

Starches can be chemically dyed by reacting a procion dye with 2,4,6-trichloro-triazine or 2,4-dichloro-triazine, followed by reaction with alkaline starch [170,171]. A chloro group on the triazine ring of the dye derivative is replaced by a starch hydroxyl group to form a stable

ether linkage between the dye and the starch. Various colors of chlorotriazine dyes are available: Cibacron Blue, Procion Brilliant Blue, Procion Brilliant Red, Procion Brown, Procion Green, and Procion Yellow.

10 Analytical Methods in Starch Chemistry

10.1 Iodine/Iodide Colors

The formation of a blue-purple color by the interaction of starch and its components with iodine/iodide is a very characteristic and unique property of starch. The blue color that is observed is due to the amylose component and has a maximum absorbance at 645 nm. The other major component, amylopectin, gives a red-purple or maroon color, with a maximum absorbance at 490 nm. The ability of starch and its major components to give this iodine/iodide color provides a method for specifically detecting very low amounts of starch from as little as 1 μg mL^{-1}. This property of starch is unique in that no other polysaccharide reacts with iodine/iodide to give a color and it hence provides a specific qualitative method for detecting starch and its components. It also provides a very sensitive method for detecting the cleavage of the α-(1→4) glycosidic linkages produced by enzymatic and chemical degradations, as the color rapidly changes as the molecular size of the starch decreases [62].

The method: the iodine reagent to obtain the blue color of starch contains 0.2% (w/v) iodine in a 2.0% (w/v) potassium iodide solution [172]. Usually a solution with 10-times this concentration is prepared and then diluted 1→10 immediately before use. The analysis is carried out by the addition of 25 μL of the diluted iodine/iodide reagent to 50 μL of each sample in a microtiter plate well and the absorbance is measured at 600–620 nm, using a microtiter plate colorimeter. The starch triiodide blue color is most stable under acidic conditions (pH 1–2) and will not form under alkaline conditions (pH 8 and higher).

It should be emphasized that the reaction to give the "blue color" is the reaction of starch with triiodide and higher complexes and not with iodine. The quantitative use of the starch-triiodide color is problematic. As indicated, its most useful application is the qualitative identification of starch. An example is the identification of starch fractions from gel-permeation chromatography [140,142] and the detection of the cleavage of the α-(1→4) glycosidic linkages of starch by enzymatic or chemical degradation reactions, which gives a decrease and eventually a change in the color [65].

10.2 The Reducing Value

The reducing value, as the name implies, is the quantitative or qualitative measurement of the specific aldehyde or hemiacetal ends of the amylose and amylopectin molecules or the ends of their degradation products. There are several methods that might be used, such as the alkaline Somoygi–Nelson method [173] or the ferricyanide/ferrocyanide/cyanide method [174]. However, for amylose and amylopectin, there is only one reducing-end group per molecule, most of the reducing-value methods are not sensitive enough to determine the reducing-end of the starch macromolecules. In the last 15 years, however, a very sensitive, micro-method has been developed that can quantitatively measure the reducing-ends of the

starch macromolecules when the starch is in concentrations of 1–9 mg mL^{-1} [175]. The method can also be used to measure the reducing value of smaller oligosaccharides, as well. It involves the copper-bicinchoninate reagent as the oxidizing reagent that oxidizes the aldehyde and itself becomes reduced to give a purple color whose absorbance can be measured at 560 nm [175].

Two reagents are involved: Solution A consists of 97.1 mg of disodium 2,2′-bicinchoninate dissolved in 45 mL of water containing 3.2 g of sodium carbonate monohydrate and 1.2 g of sodium bicarbonate, with the final volume made to 50 mL with water. Solution B consists of 62 mg of copper sulfate pentahydrate and 63 mg of L-serine dissolved in 45 mL of water, with a final volume of 50 mL. The working reagent is prepared daily by mixing equal volumes of Solution A and B. Reducing carbohydrate standards consist of 1–20 μg mL^{-1}.

The reducing value is determined by adding 100 μL of the reducing carbohydrate to 100 μL of the working reagent to the wells of a 96-sample microtiter plate. The plate is covered with Saran wrap or some other method to seal the plate and prevent evaporation of the sample and reagent when the plate is heated in a water bath at 80 °C for 35 min. The plate is then cooled for 15 min and the absorbance is measured at 560 nm. Triplicate analyses are usually performed for each sample and for the maltose standards, along with a water blank [175].

It is important to use maltose or an equivalent disaccharide, such as isomaltose, cellobiose, nigerose, and so forth as standards, as glucose gives over-oxidation for all of the reducing-value methods and cannot be used, unless glucose is the only carbohydrate reducing compound that is being measured. The substituted di- and higher-oligosaccharides do not give over-oxidation for most reducing value reagents, with the exception of the 3,5-dinitrosalicylic acid method [176].

10.3 Determination of the Total Amount of Glucose-Carbohydrate in a Sample

The total amount of starch in a solution of pure starch can be measured, using the phenol-sulfuric acid micro method [175]. The procedure consists of adding 25 μL of sample (containing 10–200 μg of gluco-carbohydrate mL^{-1}) and 25 μL of 5% (w/v) phenol to the wells of a microtiter plate. Appropriate standards of glucose or maltose (10–200 μg mL^{-1}) are used, along with a water blank. This should be done in triplicate. The plate is then placed into a bed of ice and 125 μL of concentrated sulfuric acid is added to each well, containing sample and phenol. The plate is again mixed at slow speed on a vortex mixer for 30 sec and then incubated in a water bath at 80 °C for 30 min and cooled, and the absorbance is measured at 492 nm.

10.4 Specific Determination of Starch When in the Presence of Other Carbohydrates, Using the Glucoamylase/Glucose Oxidase Method

The solution containing starch (~1 g or less) in 100 mL of pyridinium acetate (pH 5.2) is hydrolyzed by 20 IU of glucoamylase (where IU equals the amount of enzyme that will produce 1 μmole of glucose at pH 5.2 and 37 °C in 1 min). The reaction is allowed to proceed at 37 °C for ~5 h. These conditions should specifically hydrolyze 1 g of starch completely to

glucose. The glucose is then specifically and quantitatively determined by the micro glucose oxidase method [175].

For the glucose oxidase determination of glucose, there are two solutions: Solution A consists of 7.5 mg of glucose oxidase and 0.75 mg of peroxidase (Sigma Chemical Co., St. Louis, MO, USA) which is dissolved in 12.5 mL of tris-phosphate-glycerol buffer (18.1 g of tris and 28.3 g of sodium dihydrogen phosphate dihydrate dissolved in 100 mL of water plus 100 mL of glycerol, diluted to 250 mL with water). Solution B consists of 50 mg of o-dianisidine dissolved in 250 mL of the buffer described for Solution A above. The working reagent consists of equal volumes of Solutions A and B, prepared fresh daily. Solution A is stable for 1 week stored at 4 °C.

The analysis: 50 µL of sample containing 10–100 µg mL^{-1} is added to 200 µL of the working glucose oxidase reagent in a micro titer plate. Each sample is analyzed in triplicate and a blank is prepared with water. The plate is briefly mixed for 30 sec on a vortex plate mixer; 50 µL of concentrated HCl is added to each sample well and the absorbance is measured at 490–500 nm [175].

10.5 Determination of the Average Degree of Polymerization of Starch Fractions

The number of glucose residues in a starch component or fraction is termed the "degree of polymerization" or d.p. The determination of the avg. d.p. is based on the counting of the number of molecules in a given weight of starch. A method of counting the number of molecules in a given sample is the reducing value as described in *Sect. 10.2*, using copper bicinchonate that "counts" the aldehyde/hemiacetal groups. The weight of the starch fraction can be obtained by using either the phenol-sulfuric acid method (*Sect. 10.3*) or the glucoamylase/glucose oxidase method (*Sect. 10.4*). The avg. d.p. is calculated by the following formula:

$$\text{avg. d.p.} = \frac{\text{total weight of starch (µg ofglucose)}}{\text{reducing value of starch (as µg of maltose)}} \times 1.9 \qquad (2)$$

The factor 1.9 is the quotient of the (molecular weight of maltose) / (molecular weight of glucose) or 342 / 180.

10.6 Determination of the Degree of Branching of Starch Fractions

A branched starch fraction whose degree of branching or percent of branching is to be determined, first has the avg. d.p. determined (*Sect. 10.5*). Then a sample of the branched compound is dissolved in 20-mM acetate buffer (pH 4.8). Then 40 mIU of isoamylase per 50 µg (where 1 IU = 1 µmole of α-(1→6) branch linkage is hydrolyzed per min by the enzyme). The isoamylase must be free of α-amylase and glucosidase activities and can be obtained from Megazyme International [Wicklow, Ireland]. After reaction with isoamylase, the avg. d.p. is determined as per *Sect. 10.5* The percent branching is then computed from the following:

$$\%\ \text{branching} = \frac{\text{avg. d.p. after debranching withisoamylase}}{\text{avg. d.p. before debranching}} \times 100 \qquad (3)$$

The average chain length of the branched chains in the starch fraction can then be calculated by the following:

$$\text{avg. chain length} = \frac{1}{\%\ \text{branching}} \times 100 \tag{4}$$

If a sample has 5% branching, the avg. chain length would be $(1/5) \times 100 = 20$ glucose residues. The individual maltodextrin chains can be separated and further analyzed by fluorescence-assisted capillary electrophoresis (FACE), (see Sect. 15.3 in Chap. 1.2).

10.7 Measurement of Amylase Activity on Starch and Related Substrates

Amylases, both α- and β-, hydrolyze the α-(1→4) glycosidic linkages of starch to give reducing maltodextrin products that can be measured by using the micro copper-bicinchoninate reducing value method (*Sect. 10.2*). Starch (1.5 mL of 0.5% w/v) is prepared in 25 mM of a buffer with the optimum pH of the α- or β-amylase, containing 1.0 mM for α-amylases; this is incubated for 10 min at the optimum temperature of the enzyme. The reaction is started by the addition of 25–50 µL of the enzyme to the starch solution. Aliquots of 150 µL are taken every 5 min for 30 min and added to 300 µL of 50 mM NaOH, which gives a pH of 12.5 and stops the reaction. The reducing value is measured, using the copper-bicinchoninate method in which 100 µL of the stopped reaction is added to 100 µL of the working reagent in triplicate in a 96-well microtiter plate. Similar amounts (100 µL) of maltose standards from 2.0–20.0 µg are added to 100 µL of the working reagent in triplicate. The plate is covered with Saran wrap to prevent evaporation and the plate is heated at 80 °C for 35 min, cooled, and the absorbance measured at 560 nm. Dilutions of the enzyme in the reaction mixture and the stopping reagent are used to determine the amount of reducing maltose produced by the α-amylase and this is determined for each time point. The average of the triplicate measurements are obtained and a plot of µg of maltose produced versus the time of the reaction is prepared and the slope of the linear part of the line determined, which gives the µg of maltose produced per min. This is converted into µmoles of α-(1→4) bonds hydrolyzed per min by dividing by the molecular weight of maltose (342), which gives the number of International Units of α- or β-amylase [175].

10.8 Measurement of the Activity of Starch Synthase

Starch synthase catalyzes the synthesis of amylose. It can be measured in starch granules, where the enzyme is entrapped or it can be measured in a solution obtained in a plant extract.

10.8.1 Measurement of Starch Synthase Activity in Starch Granules

Starch granules (100 mg) are suspended in 1.0 mL of 0.1-mM EDTA and 4-mM glycine buffer (pH 8.4). The reaction is initiated by adding 20-mM (0.2 µCi) ADP-[^{14}C]Glc and allowed to react for 1 h. The reaction is terminated by centrifuging and washing the granules 5-times with 1 mL of water to obtain background radioactivity. This removes the unreacted ADP-[^{14}C]Glc from the granules and terminates the reaction. The granules are then treated 3-times with 1 mL of anhydrous acetone to remove water. Small amounts of residual acetone are removed

from the granules by pulling a vacuum for 1 min. The acetone-dried starch is weighed and the amount of ^{14}C-glucose incorporated into the starch is determined by heterogeneous liquid scintillation counting by putting the granules in 10 mL of toluene scintillation cocktail [5.0 g PPO and 0.1 g POPOP in 1.0 L of toluene] and counting in a liquid scintillation spectrometer. The sample is counted for 10 min or 10,000 cpm, whichever comes first. The activity of the enzyme is nmoles of glucose incorporated into starch per h per 100 mg of starch granules [125].

10.8.2 Measurement of Soluble Starch Synthase Activity in a Plant Extract

Plant extract (25 µL) is added to 75 µL of 0.1-mM EDTA and 4-mM glycine buffer (pH 8.4), containing 20-mM (0.05 µCi) ADP-[^{14}C]Glc. The digest is incubated at 37 °C for 30min. An aliquot (25 µL) is then added to (1.5 × 1.5 cm) Whatman 3-mm paper and placed into 150 mL of methanol and stirred for 10 min. The washing is repeated two more times and the paper is then dried under a heat lamp. A blank is prepared with the 20-mM (0.05 µCi) ADP-[^{14}C]Glc by adding 25 µL of water instead of plant extract and then putting 25 µL onto the Whatman 3-mm paper and washing the paper 3-times with 150 mL of methanol. The dried papers are added to 10 mL of toluene liquid scintillation cocktail and counted by heterogeneous liquid scintillation counting. The blank value is subtracted from the sample value to obtain the amount of ^{14}C-glucose incorporated into starch.

10.9 Measurement of the Activity of Starch Branching Enzyme

The formation of amylopectin occurs through the action of starch branching enzyme on maltodextrins and/or amylose to form the α-(1→6) branch linkage. Amylose or maltodextrins can be used as the substrates, but the amylose must be at a concentration of 1 mg/mL or less, as higher concentrations leads to the precipitation (retrogradation) of amylose from solution.
The assay described here uses amylose. Amylose (10 mg) is dissolved in 1.0-mL dimethylsulfoxide by stirring on a hot plate and gently warming to 40–50 °C. After the amylose has been solubilized, the dimethylsulfoxide-amylose solution is slowly diluted with 0.1-mM EDTA and 4-mM glycine buffer (pH 8.4) to 10.0 mL. The solution containing starch branching enzyme (100 µL) is added to 200 µL of the amylose solution and incubated at 37 °C for 30 min. The reaction digest is heated in a boiling water bath for 5 min, cooled, and 200 µL added to 200 µL isoamylase (250 mIU) and incubated 30 min at 37 °C. The isoamylase reaction digest (100 µL) is added in triplicate to 100 µL of the working reagent of the copper-bicinchoninate and the reducing value determined, using maltose as a standard (see *Sect. 10.2* for the reagent and the procedure). The activity of the starch branching enzyme is the number of α-(1→6) bonds formed, which in this assay would be the µmoles of maltose determined $\text{min}^{-1} \times 1.9$.

References

1. French D (1984) Organization of Starch Granules. In: Whistler RJ, BeMiller JN, Paschall EF (eds) Starch: Chemistry and Technology, 2nd edn. Academic Press, New York, p 188
2. Hizukuri S, Takeda Y, Yesuda M (1981) Carbohydr Res 94:205
3. Takeda Y, Shirasaka K, Hizukuri S (1984) Carbohydr Res 132:83

4. Takeda Y, Hizukuri S, Takeda C, Suzuki A (1987) Carbohydr Res 165:139
5. Mukerjea Ru, Robyt JF (2003) Carbohydr Res 338:1811
6. Näglei CW (1858) Die Stärkekörner, Pflanzenphysiologischer Untersuchungen. Zurich, p 624
7. Reichert ET (1913) Basis for the Classification of Plants and Animals, Publ No 173. Carnegie Institution, Washington, DC
8. Jane J-l, Kasemsuwan T, Leas S, Zobel HF, Robyt JF (1994) Starch/Stärke 46:121
9. French D (1984) Organization of Starch Granules. In: Whistler RJ, BeMiller JN, Paschall EF (eds) Starch: Chemistry and Technology, 2nd edn. Academic Press, New York, pp 200–210.
10. Taylor WN, Zobel HF, White M, Senti FR (1961) J Phys Chem 65:1816
11. Sarko A, Wu HCH (1978) Starch/Stärke 30:73
12. Bogracheva TY, Morris VJ, Ring SC, Hedley C (1998) Biopolymers 45:323
13. Zobel HF (1988) Starch/Stärke 40:1
14. Fukui T, Nikuni Z (1969) Agric Biol Chem 33:460
15. Sair L (1967) Cereal Chem 44:8
16. Lebail P, Bizot H, Buléon A (1993) Carbohydr Polym 21:99
17. Zobel HF, Stephen AM (1995) Starch: Structure, Analysis, and Applications. In: Stephen AM (ed) Food Polysaccharides and Their Applications. Marcel Dekker, New York, p 35
18. Katz FR, Furcsik SL, Tenbarge FL, Hamber RJ, Friedman RB (1993) Carbohydr Polym 21:133
19. Kimura A, Robyt JF (1995) Carbohydr Res 277:87
20. Lelievre J (1992) Food Sci Technol Today 6:234
21. Zobel HF (1988) Starch 40:44
22. French D (1984) Organization of Starch Granules. In: Whistler RJ, BeMiller JN, Paschall EF (eds) Starch: Chemistry and Technology, 2nd edn. Academic Press, New York, p 184
23. Gidley MJ, Bociek SM (1985) J Am Chem Soc 107:7040
24. Gidley MJ, Bulpin PV (1987) Carbohydr Res 161:291
25. Jenkins PJ, Cameron ER, Donald AM (1993) Starch/Stärke 45:417
26. Mukerjea Ru, Slocum G, Mukerjea Ro, Robyt JF (2006) Carbohydr Res 341:2049
27. Slaughter SL, Ellis PR, Butterworth PJ (2001) Biochim Biophys Acta 1525:29
28. Hanssen E, Dodt E, Niemann EG (1953) Kolloid-Z 130:19
29. Killion PJ, Foster JF (1960) J Polym Sci 46:65
30. Leach HW, Schoch TJ (1962) Cereal Chem 39:318
31. Hall DM, Sayre JG (1971) Textile Res 41:404
32. Mukerjea Ru, Slocum G, Robyt JF (2007) Carbohydr Res 342:103
33. Gough BM, Pybus JN (1973) Stärke 25:123
34. Jane J-l, Shen JJ (1993) Carbohydr Res 247:279
35. Pan DD, Jane J-l (2000) Biomacromolecules 1:126
36. Mukerjea Ro, Mukerjea Ru, Robyt JF (2006) Carbohydr Res 341:757
37. Meyer KH, Brentano W, Bernfeld P (1940) Helv Chim Acta 23:845
38. Meyer KH, Bernfeld P, Wolff E (1940) Helv Chim Acta 23:854
39. Meyer KH (1940) Naturwissenschaften 28:397
40. Meyer KH (1942) Adv Coll Sci 1:143
41. Schoch TJ (1942) J Am Chem Soc 64:2957
42. Schoch TJ (1945) Adv Carbohydr Chem 1:258
43. Bear RS (1942) J Am Chem Soc 64:1388
44. Schoch TJ (1964) Methods Carbohydr Chem 4:56
45. Cowie JMG, Greenwood CT (1957) J Chem Soc 2862
46. French D, Pulley AO, Whelan WJ (1963) Die Stärke 15:349
47. Colin HH, de Claubry G (1814) Ann Phys 48:297
48. Hanes CS (1937) New Phytologist 36:198
49. Rundle RE, Baldwin RR (1943) J Am Chem Soc 65:554
50. Rundle RE, French D (1943) J Am Chem Soc 65:558
51. Rundle RE, French D (1943) J Am Chem Soc 65:1707
52. Pauling L, Corey RB, Branson HR (1951) Proc Natl Acad Sci US 37(242):2105
53. Watson JD, Crick FHC (1953) Nature 171:738
54. Rundle RE, Edwards FC (1943) J Am Chem Soc 65:2200
55. Bear RS (1944) J Am Chem Soc 66:2122
56. Mikus FF, Hixon RM, Rundle RE (1946) J Am Chem Soc 68:1115
57. Zaslow B (1963) Biopolymers 1:165
58. Simpson TD, Dintzis FR, Taylor NW (1972) Biopolymers 11:2591
59. Yamashita Y, Hirai N (1966) J Polym Sci, Part A-2 4:166
60. Yamashita Y, Monobe K (1971) J Polym Sci, Part A-2, 9:1471
61. Jane J-l, Robyt JF (1984) Carbohydr Res 132:105

62. Bailey JM, Whelan WJ (1961) J Biol Chem 236:969
63. Banks W, Greenwood CT (1968) Carbohydr Res 7:349
64. Yamaguchi M, Kainuma K, French D (1979) J Ultrastructure Res 69:249
65. Robyt JF, French D (1967) Arch Biochem Biophys 122:8
66. Robyt JF, French D (1970) Arch Biochem Biophys 138:622
67. Dube SK, Nordin P (1962) Arch Biochem Biophys 99:105
68. Robyt JF, French D (1970) J Biol Chem 245:3917
69. Kainuma K, French D (1969) FEBS Letters 5:257
70. Kainuma K, French D (1970) FEBS Letters 6:182
71. Robyt JF, French D (1963) Arch Biochem Biophys 100:451
72. MacGregor EA, MacGregor AW (1968) Carbohydr Res 142:223
73. Saito N (1973) Arch Biochem Biophys 155:290
74. Morgan FJ, Priest FR (1981) J Appl Bacteriol 50:107
75. Robyt JF (1984) Enzymes in the Hydrolysis and Synthesis of Starch. In: Whistler RJ, BeMiller JN, Paschall EF (eds) Starch: Chemistry and Technology, 2nd edn. Academic Press, New York, pp 102–103
76. Robyt JF, Whelan WJ (1968) General Aspects of Amylase Action, Chap. 13 α-Amylases, Chap. 14 β-Amylases. In: Radley JA (ed) Starch and Its Derivatives, 4th edn. Chapman and Hall, London, pp 477–497
77. Robyt JF, French D (1964) Arch Biochem Biophys 104:338
78. Hidaka H, Adachi K, Yoshida K, Nirva K (1978) Denpun Kagaku 25:148
79. Pazur JH, Ando T (1959) J Biol Chem 234:1966
80. Ueda S (1956) Bull Agric Chem Soc Jpn 20:148
81. Savel'ev AN, Sergev VR, Firsov LM (1982) Biochem (U.S.S.R.) 47:330
82. Hiromi K (1970) Biochem Biophys Res Commun 40:1
83. Tsujisaka Y, Fukumoto J, Yamoto T (1958) Nature 181:94
84. Stetten Jr D, Stetten MR (1960) Physiol Revs 40:513
85. Robyt JF, Ackerman RJ (1971) Arch Biochem Biophys 145:105
86. Kainuma K, Kobayashi S, Ito T, Suzuki S (1972) FEBS Letters 26:281
87. Kainuma K, Wako K, Kobayashi S, Nogami A, Suzuki S (1975) Biochem Biophys Acta 410:333
88. French D (1975) Adv Carbohydr Chem 12:189
89. Kitahata S, Tsuyama N, Okada S (1974) Agric Biol Chem 38:387
90. Blackwood AD, Bucke C (2000) Enzyme Microbiol Technol 27:704
91. Mori S, Hirose S, Takaichi O, Kitahata S (1994) Biosci Biotech Biochem 58:1968
92. Hobson PN, Whelan WJ, Peat S (1951) J Chem Soc 1451
93. Harada T, Yokobayashi K, Misaki A (1968) Appl Microbiol 16:1493
94. Yokobayashi K, Misaki A, Harada T (1969) Agric Biol Chem 33:625
95. Yokobayashi K, Misaki A, Harada T (1970) Biochim Biophys Acta 212:458
96. Harada T, Misaki A, Akai H, Yokobayashi K, Sugimoto K (1972) Biochim Biophys Acta 286:497
97. Kainuma K, Kobayashi S, Harada T (1978) Carbohydr Res 61:345
98. Andersen SM, Lundt I, Marcussen J, Yu S (2002) Carbohydr Res 337:873
99. Lee SS, Yu S, Withers SG (2003) Biochemistry 42:13081
100. Manners DJ (1971) Biochem J 123:1P
101. Brown HT, Heron J (1879) J Chem Soc 34:596
102. Stamberg OE, Bailey CH (1939) Cereal Chem 16:319
103. Sandstedt RM, Gates RL (1954) Food Res 19:190
104. Leach HW, Schoch TJ (1961) Cereal Chem 38:34
105. Rasper V, Perry G, Duitschaever CL (1974) Can Inst Food Sci Technol J 7:166
106. Smith JS, Lineback DR (1976) Stärke 28:243
107. Kimura A, Robyt JF (1995) Carbohydr Res 277:87
108. Kimura A, Robyt JF (1996) Carbohydr Res 288:233
109. Kimura A, Robyt JF (1996) Carbohydr Res 287:255
110. Kim Y-k, Robyt JF (1999) Carbohydr Res 318:129
111. Kim Y-k, Robyt JF (2000) Carbohydr Res 328:509
112. Yoon S-H, Robyt JF (2005) Enzyme Microb Technol 37:556
113. Hanes CS (1940) Proc Roy Soc B 129:174
114. Green DE, Stumpf PK (1942) J Biol Chem 142:355

115. French D, Wild GM (1953) J Am Chem Soc 75:4490
116. Trevelyan WE, Mann PFE, Harrison JS (1952) Arch Biochem Biophys 39:419
117. Ewart MH, Siminovitch D, Briggs DR (1954) Plant Physiol 29:407
118. Liu TT, Shannon JC (1981) Plant Physiol 67:525
119. Bocca SN, Rothschild A, Tandecarz JS (1997) Plant Physiol Biochem 35:205
120. Ball SG, van de Wal HBJM, Visser RGF (1998) Trends Plant Sci 3:1360
121. Tomlinson K, Denyer K (2003) Adv Bot Res 40:1
122. de Fekete MAG, Leloir LF, Cardini CE (1960) Nature 187:918
123. Recondo E, Leloir LF (1961) Biochem Biophys Res Commun 6:85
124. Leloir LF, de Fekete MAR, Cardini CE (1961) J Biol Chem 236:636
125. Mukerjea Ru, Yu L, Robyt JF (2002) Carbohydr Res 337:1015
126. Mukerjea Ru, Robyt JF (2005) Carbohydr Res 340:245
127. Frydman RB, Cardini CE (1964) Biochem Biophys Res Commun 17:406
128. Robyt JF, Walseth TF (1978) Carbohydr Res 61:433
129. Robyt JF, Eklund SH (1983) Carbohydr Res 121:279
130. Mukerjea Ru, Robyt JF (2005) Carbohydr Res 340:2206
131. Haworth WN, Peat S, Bourne EJ (1944) Nature 154:236
132. Hobson PN, Whelan WJ, Peat S (1950) J Chem Soc 3566
133. Drummond GS, Smith EE, Whelan WJ (1972) Eur J Biochem 26:168; (1970) FEBS Lett 9:136
134. Whelan WJ (1971) Biochem J 122:609
135. Schiefer S, Lee EYC, Whelan WJ (1973) FEBS Lett 30:129
136. Boyer CD, Preiss J (1978) Carbohydr Res 61:321
137. Nägeli W (1874) Justus Liebigs Ann Chem 173:218
138. Watanabe T, French D (1980) Carbohydr Res 84:115
139. Lintner CJ (1886) J Prakt Chem 34:378
140. Ma W-P, Robyt JF (1987) Carbohydr Res 166:283
141. Small JC (1919) J Am Chem Soc 41:113
142. Fox JD, Robyt JF (1992) Carbohydr Res 227:163
143. Robyt JF, Choe J-Y, Hahn RS, Fuchs EB (1996) Carbohydr Res 281:203
144. Robyt JF, Choe J-Y, Fox JD, Hahn RS, Fuchs EB (1996) Carbohydr Res 283:141
145. Arbuckle AW, Greenwood CT (1958) J Chem Soc 2226
146. Whistler RL, Linke EG, Kazeniac S (1956) J Am Chem Soc 78:4704
147. Hullinger CH, Whistler RL (1951) Cereal Chem 28:153
148. Whistler RL, Schweiger R (1957) J Am Chem Soc 79:6460
149. Corey EJ, Schmidt G (1979) Tetrahed Lett 5:399
150. Westlake RJ, Hill RD (1982) Carbohydr Res 104:334
151. Chang PS, Robyt JF (1996) J Carbohydr Chem 15:819
152. Lee S-B, Robyt JF (2001) Carbohydr Res 336:47
153. Chang PS, Mukerjea Ru, Fulton DB, Robyt JF (2000) Carbohydr Res 329:913
154. Mehltretter CL (1963) Stärke 15:313
155. Mehltretter CL (1966) Stärke 18:208
156. Mehltretter CL (1964) Methods Carbohydr Chem 4:316
157. Guthrie RD (1961) Adv Carbohydr Chem 16:105
158. Wurzurg OB (1964) Methods Carbohydr Chem 4:286
159. Friese H, Smith FA (1928) Ber 61B:1975
160. Radley JA (1968) In: Radley JA (ed) Starch and Its Derivatives, 4th edn. Chapman and Hall, London, p 357
161. Rutenberg MW, Solarek D (1984) Starch Derivatives: Production and Uses. In: Whistler RJ, BeMiller JN, Paschall EF (eds) Starch: Chemistry and Technology, 2nd edn. Academic Press, New York, pp 339–344
162. Gramera RE, Heerema J, Parrish FW (1966) Cereal Chem 43:104
163. Paschall EF (1964) Methods Carbohydr Chem 4:294
164. Whistler RL, Hirase S (1961) J Org Chem 26:4600
165. Bines BJ, Whelan WJ (1960) Chem Ind (London) 997
166. Weill CE, Kaminsky M, Guerrera J (1979) Carbohydr Res 73:337
167. Braun PJ, French D, Robyt JF (1985) Carbohydr Res 143:107
168. Coon I, St J Manley R (1963) Methods Carbohydr Chem 3:271
169. Stojanovic Z, Jeremic K, Jovanovic SM (2000) Starch/Stärke 52:413

170. Pavlenko AF, Ovodov YS (1970) J Chromatog 52:165
171. Marshall JJ (1970) Anal Biochem 37:466
172. Bourne EJ, Haworth WN, Macey A, Peat S (1948) J Chem Soc 924
173. Nelson N (1944) J Biol Chem 153:357
174. Robyt JF, Ackerman RJ, Keng JG (1972) Anal Biochem 45:517
175. Fox JD, Robyt JF (1991) Anal Biochem 195:93
176. Robyt JF, Whelan WJ (1972) Anal Biochem 45:510
177. Whistler RL, Daniel JR (1984) Molecular Structure of Starch. In: Whistler RJ, BeMiller JN, Paschall EF (eds) Starch: Chemistry and Technology, 2nd edn. Academic Press, New York, p 173
178. Gracza R (1965) Minor Constituents of Starch. In: Whistler RJ, Paschall EF, BeMiller JN, Roberts HJ (eds) Starch: Chemistry and Technology, vol 1. Academic Press, New York, pp 105–131

6.3 Cellulose and Associated Heteropolysaccharides

Wolfgang G. Glasser
Dept. Wood Science and Forest Products, Virginia Tech,
Blacksburg, VA 24061, USA
wglasser@vt.edu

Abstract

Polysaccharides constitute a class of polymers in which the elemental composition (i. e., O/C ratio) has been reduced from that of carbon dioxide to that of carbon monoxide plus hydrogen. The resulting structural diversity is employed by plants to support the life-sustaining functions of energy storage, structural support, and moisture management. Whereas cellulose is the plant's premiere structural support molecule, the associated heteropolysaccharides ("hemicelluloses") serve to provide for moisture sorption, fracture toughness, and storage energy. This variety of function is achieved by molecular manipulations resulting in strict regularity and uniformity on one hand, and careful tailoring of irregularity capable of preventing crystallization on the other. Synthetic structure manipulation strategies (i. e., derivatizations) employ

In: *Glycoscience*. Fraser-Reid B, Tatsuta K, Thiem J (eds)
Chapter-DOI 10-1007/978-3-540-30429-6_36:

chemical modification reactions qualified to maneuver between the different functions of plant-derived polysaccharides.

Keywords

Cellulose; Homopolysaccharides; Hemicelluloses; Heteropolysaccharides; Crystallinity; Cell wall architecture; Molecular interactions; Cellulose derivatives; Cellulose esters; Cellulose ethers

Abbreviations

AGU	anhydroglucose unit
CA	cellulose acetate
CAB	cellulose acetate butyrate
CAP	cellulose acetate propionate
CMC	carboxymethyl cellulose
CMCAB	carboxymethyl cellulose acetate butyrate
CTA	cellulose triacetate
DMSO	dimethyl sulfoxide
DP	degree of polymerization
DS	degree of substitution
GPC	gel permeation chromatography
HAMC	hydroxyethyl methyl cellulose
HEC	hydroxyethyl cellulose
HPC	hydroxypropyl cellulose
LCCE	long-chain cellulose ester
LODP	level off DP
MCC	microcrystalline cellulose
MHS	Mark–Houwink–Sakurada
MS	molar substitutions
NMMO	*N*-methylmorpholine *N*-oxide
TBAF	tetrabutylammonium fluoride
THF	tetrahydrofuran

1 Introduction

Carbohydrates, substances having the general composition of C_x $(H_2O)_x$, are central to planet Earth's survival and sustainability. Carbohydrates result from the solar energy-driven reduction of carbon dioxide by photosynthesis. Photosynthesis generates both carbohydrate matter (solids) and oxygen (gas). The reduction of carbon dioxide to carbohydrates and oxygen (plus water) requires 112 kcal/mol of CO_2 [1].

Carbohydrates include a large variety of different molecules and materials, different in composition, size, and function. The chemistry of C_x $(H_2O)_x$-molecules (i. e., carbohydrates) is extremely complex in terms of its molecular and stereo-chemical diversity. This complexity has been the subject of numerous texts [2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25].

The process of photosynthesis is central to the planet's survival. Without photosynthesis and without solar energy-collecting plant life, planet Earth would soon turn into a planet with an environment not unlike that of Mars, a planet which is described as having had water on its surface in the past. Water and plant life are central to Earth's ecology; and carbohydrates are central to plants.

The life-sustaining function of carbohydrates necessitates that they meet three general requirements: (A) carbohydrates must serve as **stored energy**, materials that are capable of readily releasing energy by degradation; (B) carbohydrates must serve for **structural support**, materials that provide mechanical strength to individual cells as well as plant tissues; and (C) carbohydrates must provide plants with the ability to manage **moisture**, materials that control the water household over a wide range of physical properties and physiological needs, materials capable of immobilizing water and controlling water's rheological properties. The purposes of **energy storage**, **structural support**, and **gel formation** are all central to life as we recognize it; and they are all addressed by carbohydrates.

For carbohydrates to meet these requirements, diversity is needed on both the molecular and the size-level. Only large carbohydrate molecules, **polysaccharides**, can provide the wide spectrum of storage, structural, and gel-forming abilities required by nature. Meeting these requirements has made it necessary for plants to produce polysaccharides that can be classified as **linear**, **branched**, and **crosslinked** polymers, as well as **homo-** and **heteropolymers** in accordance with terminology in common use in the polymer community (◉ *Fig. 1*) [26]. Nature has found need to adopt all different kinds of macromolecular architectures in pursuit of the three different functions of carbohydrates.

Cellulose ranks first among the strength-building polysaccharides in abundance. It is most prevalent in wood fibers where it helps trees gain the strength necessary to lift crowns (i. e., solar energy-collecting devices) above ground level. However, cellulose is not exclusive to trees or even plants in general. A very pure form of cellulose is present in the seed hairs (*trichomes*) of the cotton and several other plants; in soft and hard bast fibers from the stems

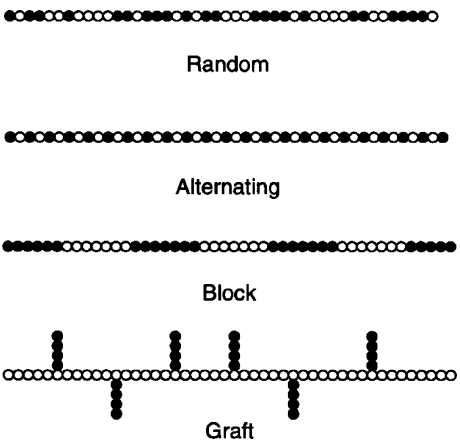


Figure 1
Common structures of heteropolymers with two types of repeat units

of such plants as hemp, ramie, flax, sisal, banana, pineapple, etc.; and cellulose is also produced by a variety of bacteria (esp. *Acetobacter xylinium*), algae (such as *Valonia ventricosa*), and even marine organisms (*Tunicatae*). The algal cellulose from *V. ventricosa* is of interest because of its large and highly oriented crystallites; and bacterial cellulose has become a commercial product for material, nutritional, and medicinal applications (such as *Cellulon* by Weyerhaeuser; *PrimaCel* by NutraSweet Kelco; and *Biofilm*, a skin-substitute) [12,13,27,28]. Whereas these different sources of cellulose differ in a wide range of characteristics and properties, they all consist of one form of a linear homopolysaccharide whose structural characteristics have been studied for the better part of almost two centuries. Most of this research has focused on wood-based cellulose generated by industrial delignification (pulping) processes. It is estimated that the total amount of cellulose grown and degraded annually on Earth is about $5–10 \times 10^{10}$ t [12]. Other estimates put the annual plant synthesis rate at $10^{11}–10^{12}$ t [18]. This compares to an annual pulp production (for paper and board) of about 200×10^{6} t worldwide [29].

2 Formation and Structure

2.1 Chemical Structure

Cellulose is a linear homopolymer of β-(1 → 4)-linked D-anhydroglucopyranose monomeric repeat units (AGU). Since every second glucose unit is sterically rotated by about 180 °C with respect to its neighbors, it can be considered a syndiotactic homopolymer of glucose or an isotactic homopolymer of cellobiose units (● *Fig. 2*) [18]. Arguments can be presented in favor of either viewpoint, cellulose as polymer of monomeric glucose or dimeric cellobiose units.

Each monomeric repeat unit possesses three hydroxyl (OH–) groups, in positions C-2, C-3, and C-6 of the pyranose unit. The homopolymer is asymmetric having two distinct end-units a **nonreducing** and a **reducing** end. The nonreducing end features an additional OH-group in position C-4, and the reducing end has carbonyl (reducing) functionality.

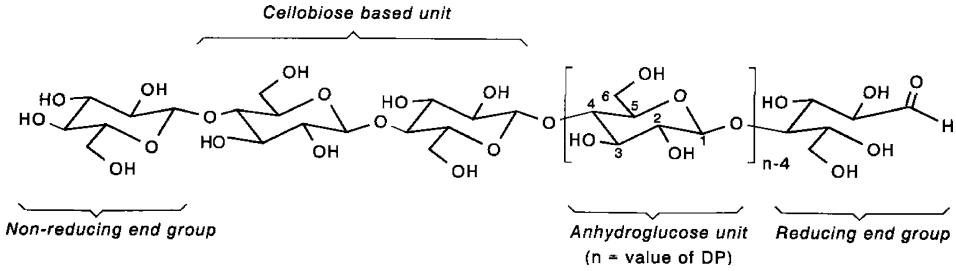


Figure 2
Chemical structure of cellulose indicating two different endgroups and repeat units in two optional interpretations, cellobiose, or anhydroglucopyranose. The rotation of every second anhydroglucose unit creates a syndiotactic linear poly[anhydroglucose]-molecule, called *cellulose* [18]

Table 1
Average degrees of polymerization (DP) of celluloses of different origin

Type of cellulose	DP	Molecular weight (g mol^{-1})
Cotton	7,000–14,000	1,125,000–2,250,000
Flax	8,000	1,300,000
Bacterial cellulose	2,700	500,000
Wood pulp	2,500–3,300	400,000–600,000
α-Cellulose	800–1100	130,000–180,000
Regenerated cellulose	250–500	40,000–80,000
Microcrystalline cellulose	100–200	15,000–30,000

Source: H. Krässig et al., 2004 [13].

The size of cellulose molecules is consistently expressed as **degree of polymerization** (DP), which expresses size in terms of the number of AGUs per individual chain. In natural state, cellulose is thought to be a homopolysaccharide with relative size-uniformity (**monodisperse**) consisting of somewhere around 14×10^3 glucose units (*Table 1*) [30,31,32]. This size, however, is limited to cellulose in intact plant tissues because its isolation always results in molecular degradation. Commercially available celluloses range in size from DP of 100 to 12,000 (for cotton) in relation to their method of isolation (*Table 1*) [13].
The asymmetric nature of cellulose molecules, with a distinctive carbonyl group at one end, is sometimes used for the determination of molecular mass by **end group determination**. More common methods of molecular weight determination, however, employ viscosimetry. Whereas the highly hydroxylated nature of cellulose serves as an almost ideal qualification for water solubility, its strict molecular regularity, linearity, and uniformity also qualify it for crystallinity. Water solubility is limited to DP < 6, and typical polymer properties arise already at DP 30 [14]. Sincecrystallization is spontaneous and exothermic, preventing it requires the introduction of irregularity, branching and/or diverse functionalization. This principal rule of polymer chemistry is adopted by nature when nonstructural, gel-forming polysaccharides are needed. The transition from structural to gel-forming and nutritional (storage energy) purposes is accomplished by nature strictly on the basis of variation in chemical structure [33].

2.2 Physical Structure

All molecules regardless of physical state (gas, liquid, or solid) wish to undergo Brownian motion. This desire results in such observations as gas pressure varying with temperature; solution behaviors (and viscosity) varying in relation to temperature and solvent parameters; and solids changing properties in relation to glass transition temperatures (T_g) and melting points (T_m). Resistance to motion originates from intra- and intermolecular interactions, and this relates to rotational freedom and unfreedom of polymer chains. A widely used illustration of the rotational freedom phenomenon is the *energy contour map* (*Fig. 3*) [3,16]. Energy contour maps are constructed on the basis of dynamic molecular modeling experiments. The models determine energy parameters involved in the rotation of two neighboring repeat units around the common interunit bond. When two adjacent anhydroglucose units are rotated with

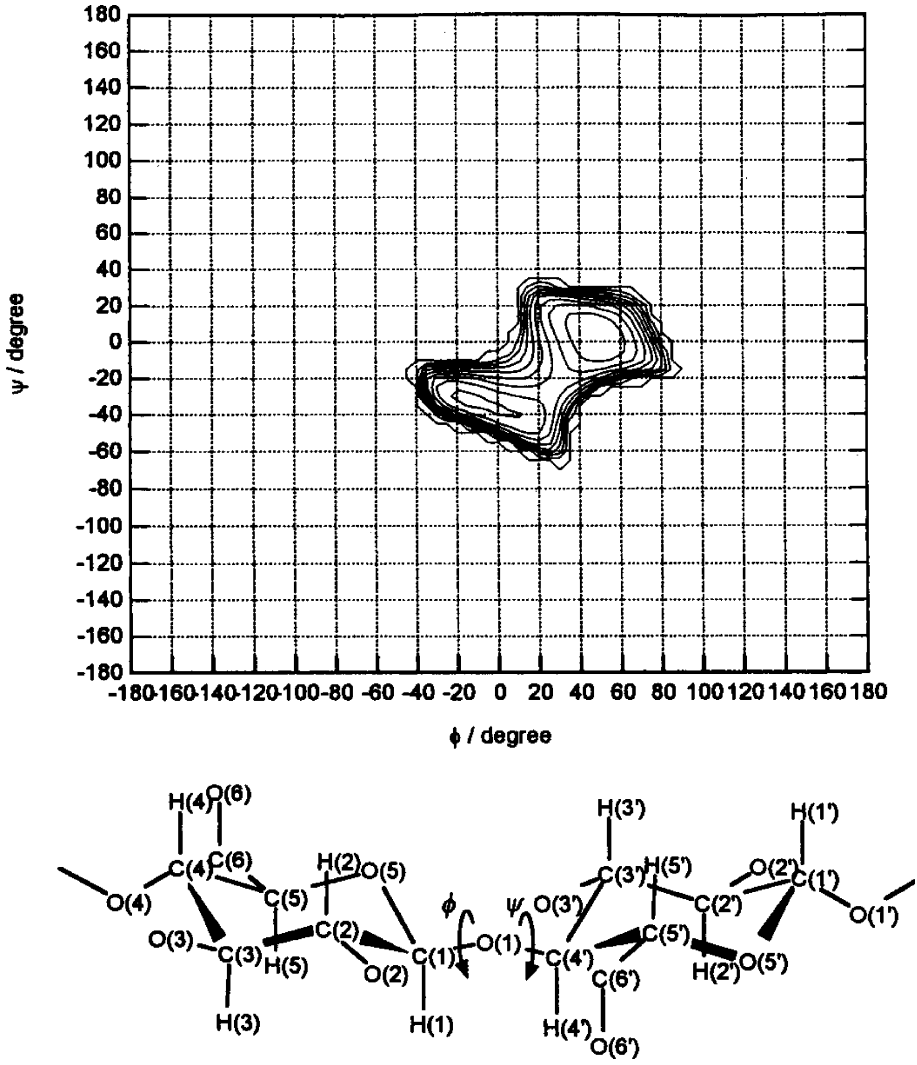


Figure 3

The *2-D energy contour map* illustrating the potential energy barriers to rotation around two dihedral angles, Φ and Ψ, of cellobiose. (Adopted from Kajiwara and Miyamoto [34])

respect to each other around the connecting glycosidic ether bond, certain conformations meet positions of higher and lower potential energy. The most stable conformations are characterized by a low energy state. When rotation is plotted as bond angles vs. stored energy, a three-dimensional energy profile is created that identifies positions of high and low energy states in a manner similar to a topographical map (*Fig. 3*).

With cellulose, the region of angular orientation that produces the lowest point on the energy contour map is narrow and well defined. Any distortion out of the most relaxed position involves significant energy input. Several positions of relaxed state are identified on the ener-

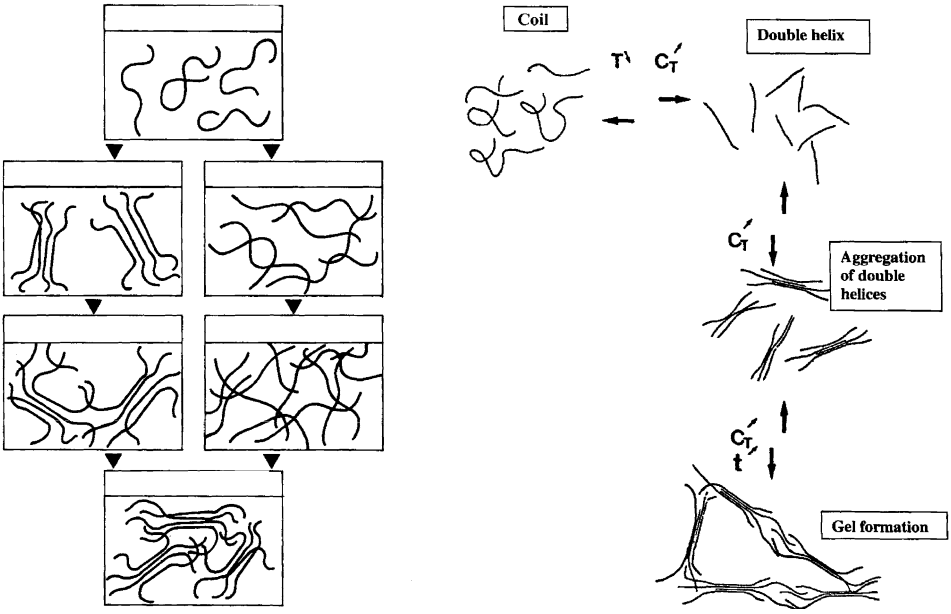


Figure 4
Organization of cellulose (left) and κ-carrageenan molecules (right) during desolvation. The figure illustrates the possible paths to the formation of solid structures: early formation of organized structures in solution (left side) leading to liquid crystalline structures of rigid-type molecules, and desolvation without organization (right) leading to a largely amorphous structure (left, according to Schurz, [35]). And gel formation by the aggregation of helices (right, according to Rinaudo, [33])

gy contour map for cellotriose (◉ *Fig. 3*). The consequence of this contour is the identification of several preferred conformations (in solution) and several lesser stable ones. The more pronounced an energy gradient is with regard to intermolecular rotation, the more rigid is a molecule in solution. The steeper the energy gradient, the stiffer is the molecule and the faster will it organize into a conformation of order when desolvated (◉ *Fig. 4*). This perspective suggests that order can be gained already in solution, and this phenomenon is known as *liquid crystallinity*. Cellulose exhibits distinct signs of liquid crystallinity [12].

In the case of cellulose, intramolecular order produces a helix with a two fold screw axis [4]. The particular helical arrangement of single cellulose molecules represents an extremely extended conformation, i. e., a helix with high pitch (◉ *Fig. 5*), that serves cellulose well for building stiffness and tensile strength in solid state. Hydrogen bonds play a pivotal role in achieving this helical organization.

In solid state, cellulose exists in mostly crystalline state. This state has been the subject of extensive studies for at least 100 years, and several important aspects need to be pointed out at the onset. Crystallinity is neither *uniform* (options exist) nor *static* (crystallinity can be lost as well as gained in relation to molecular mobility), nor is it *permanent* (conditions have been identified under which transitions of order take place). Cellulose crystallinity manifests itself through the existence of distinctive X-ray diffraction patterns. These patterns allow the determination of the overall dimensions of *unit cells* which are spatial units that represent the

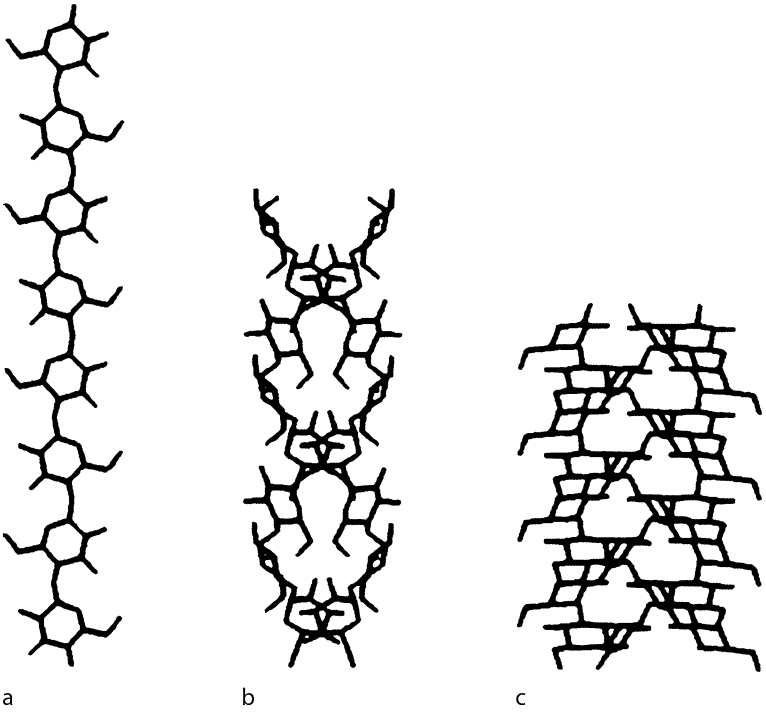


Figure 5
Three potential helical structures of polysaccharides. (a) Single helix with twofold screw axis (cellulose); (b) double helix (amylose); and (c) triple helix (β, 1–3 glucan) (according to Rees [4]). The particular structure depends largely on the constraints imposed by hydrogen-bonds and rotational freedoms around the glycosidic intermonomer bonds (i. e., *Fig. 3*, energy barriers to rotation)

smallest part of the crystal that can reproduce the entire crystal by translations of its contents along its edges [36]. Unit cell dimensions reflect the molecular arrangement of cellulose chains in the crystal. Cellulose crystallinity has most often been illustrated by use of the Meyer Misch monoclinic unit cell (*Fig. 6*), which has variably been revised since its first publication in 1938 [37]. The monoclinic unit cell is based on the understanding that hydrogen (H)-bonds exist intramolecularly between the H-atom of the OH group in position C-3 (i. e., O(3)H) and the oxygen-atom of the C-5 position (O5) of the adjacent unit in the same cellulose chain. Additional H-bonds exist between the O(6) of one glucose residue and the O(2)H of the adjacent unit. Since this H-bonding pattern results in every glucose unit of a cellulose chain being rotated by 180 °C with respect to its neighbor, a stable 2-fold screw axis is formed. The formation of intermolecular H-bonds between adjacent cellulose chains (in the O–B plane of *Fig. 7*) involves the position O(3) on one, and O(6)H on the adjacent unit. This arrangement results in the formation of a flat sheet. In crystallites, sheets are staggered on top of each other whereby cellulose chains are shifted from their respective neighbors. There appear to be no H-bonds between layers of sheets, but sheets are bonded to each other via Van der Waals bonds (hydrophobic interactions) [40].

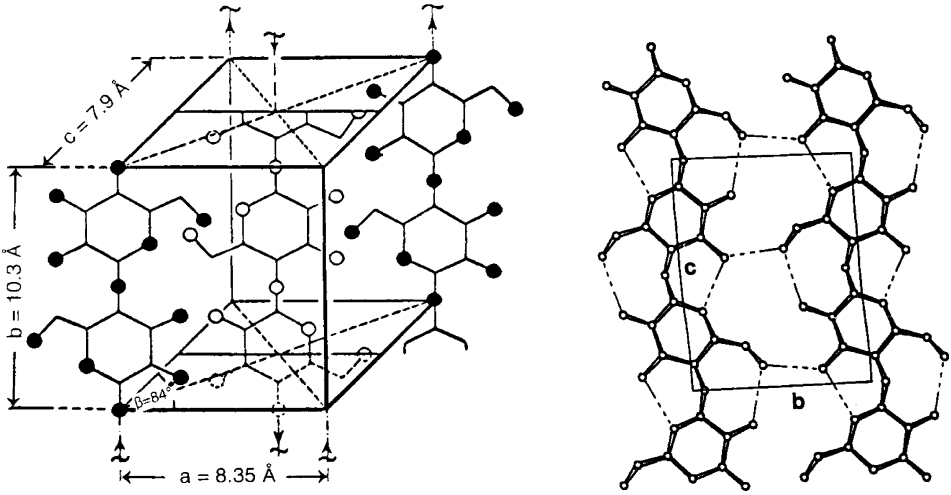


Figure 6
Monoclinic unit cell of *cellulose I* according to the model by Meyer and Misch [37], which shows the antiparallel orientation of adjacent chains (left); and the hydrogen-bonding network of two adjacent cellulose chains forming a sheet-like structure according to Gardner and Blackwell [38]

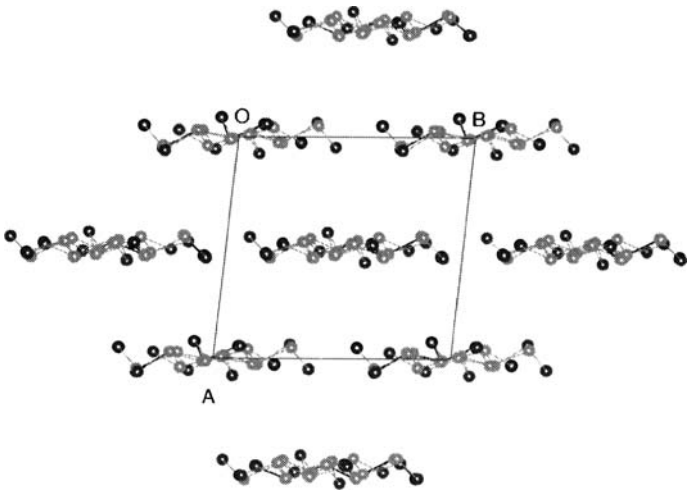


Figure 7
Segment of a crystal structure of cellulose I_β illustrating the relationship between unit cell and crystal structure. Hydrogen bonds (not shown) are in the direction of O to B creating sheets of aligned chains. (Adopted from Horii and Wada [39])

Cellulose crystallinity is not uniform. A simple experiment of immersing cellulose in cold concentrated alkali, a process used for enhancing the dye-absorbing quality of cotton fabrics called *mercerization*, was found already in the 1930s on the basis of X-ray diffraction to produce a cellulosic *allomorph* with different unit cell dimensions [13,14]. This was given the

Table 2
Unit cell dimensions for different cellulose polymorphs (in nm)

	a	b	c	β (degree)
Cellulose I	0.821	1.030	0.790	83.3
Cellulose II	0.802	1.036	0.903	62.8
Cellulose III	0.774	1.030	0.990	58.0
Cellulose IV	0.812	1.030	0.799	90.0

For unit cell definitions, see *Fig. 6*.

designation *cellulose II* in contrast to the native cellulose state, which was designated *cellulose I*. While cellulose II has the same types of hydrogen bonds, it is apparent that it is composed of antiparallel chains whereas cellulose I possesses parallel chains. Since new hydrogen bonds between the corner and center chains are formed during the transformation from I to II, cellulose II is the more stable crystalline configuration compared to cellulose I. Virtually all regenerated cellulose (regenerated from solution) exists in cellulose II-form [15].

In addition to the recognition of two significantly different cellulose allomorphs, cellulose I and II, several other crystalline arrangements have been identified (*Table 2* [15]). These evolve from treatment with strongly hydrogen-bonding, swelling agents, such as liquid ammonia (*cellulose III*) and treatment with glycerol at high temperatures (260 °C), which produces *cellulose IV*.

Whereas cellulose I had always been considered to consist of a uniform monoclinic crystal structure, evidence surfaced in the 1950s that cast doubt on this hypothesis. On the basis of mostly ^{13}C-NMR evidence, it was established in the 1980's that cellulose I exists in two separate and distinct subclasses, designated *cellulose* I_α and I_β. The more stable I_β-form (by 8.7 kJ mol^{-1} for cellobiose) predominates in plant celluloses, whereas I_α is more prevalent in some bacterial celluloses (*Valonia* sp.). Its unit cell is described as a 1-chain triclinic arrangement that has been reviewed recently [40]. Whereas crystal transformation from cellulose I to II involves treatment with highly concentrated solutions of alkali or regeneration from solution, crystal transformation from Cellulose I_α to I_β was found to be possible by annealing at high temperatures with saturated steam; annealing at high temperature in alkaline solution; or regeneration of cellulose from cellulose triacetate or from cellulose III crystal state [40]. The ease of transition of cellulose I_α to I_β, and of cellulose I to II, illustrates how the crystal structure of cellulose is dynamic and in some cases transient. Examples of these transitions have been described elsewhere [12,13,13,14,15,16].

Unit cells are the smallest repeatable spatial units within crystallites. A crystallite contains multiple unit cells in parallel arrangement giving rise to crystalline strands or *elementary fibrils* (*Fig. 8*). These fibrils are quoted with cross-sectional dimensions of 2.5 to 4 and 12–14 nm in length. In cotton they can have dimensions of 10 to 40 × 50–60 nm. Elementary fibrils aggregate to form microfibrils.

The formation of crystalline microfibrils has been elucidated using the bacterium *Acetobacter xylinum* as cellulose producing model. It appears that individual cellulose chains are "extruded" at multiple cellulose-synthesizing sites located in the cytoplasmatic membrane of the organism. Cellulose synthesis produces 12 to 16 cellulose chains into the surrounding medium

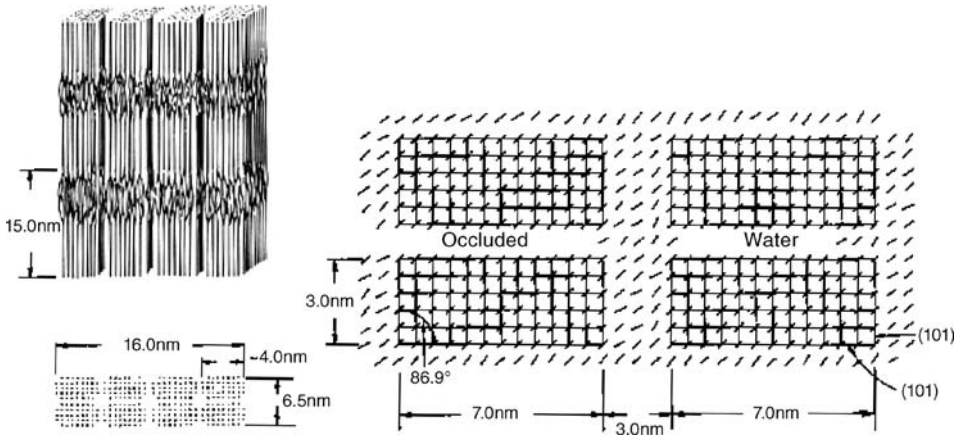


Figure 8
The architecture of elementary and microfibrils of cellulose illustrating the mixture of crystalline and amorphous (*paracrystalline*) phases. (After Kreaessig et al. [13])

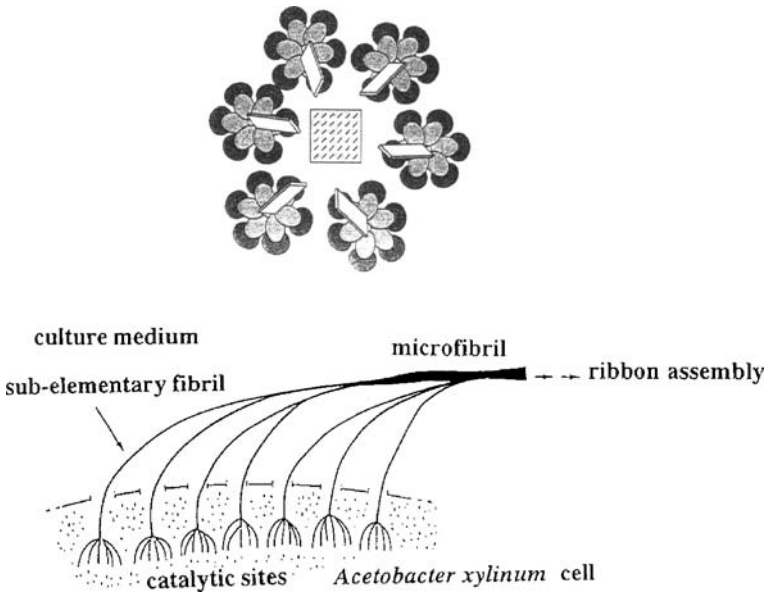


Figure 9
Schematic model of the cellulose synthase complex: crystalline cellulose I emerges from the *rosette terminal complex*, which is embedded in the plasma membrane. View from the top (above) and the side (below). (Adopted from Perez and Mazeau [42] and from Horii [40])

where they form sub elementary fibrils with cross sectional dimensions of about 1.5 nm. These fibrils aggregate with each other to form microfibrils that combine to form ribbon assemblies with diameters of 40 to 60 nm. The molecular biology and biosynthesis of cellulose formation has recently been reviewed [41]. A schematic model is given in ❯ *Fig. 9*.

2.3 Properties

The transition from crystalline to melt state, which is normal for crystalline polymers, is not observed with cellulose under normal conditions. It appears that the secondary bonds giving rise to the crystalline state are too strong and too numerous to be broken by a rise in temperature. Thermal degradation (beginning at ca. 180 °C) precedes melting under atmospheric pressure conditions. Nevertheless, a plastic deformation interpreted as melting has recently been reported for cellulose fibers exposed to laser radiation in a highly confined (pressurized) space [43]. The fracture surface of a thermoplastically deformed cellulose disc is shown in *Fig. 10*.

Parallel observations are made for solution behavior where interaction with conventional solvents fails to result in dissolution. Solvents with extremely high hydrogen bonding capability, or the reduction of H-bonds by derivatization, are required to allow crystalline cellulose to dissolve or to melt.

Cellulose is soluble in several *derivatizing* and *nonderivatizing* solvents [20,44,45,46]. These, in general, are solvents that modify OH-groups covalently or that form strong secondary bonds

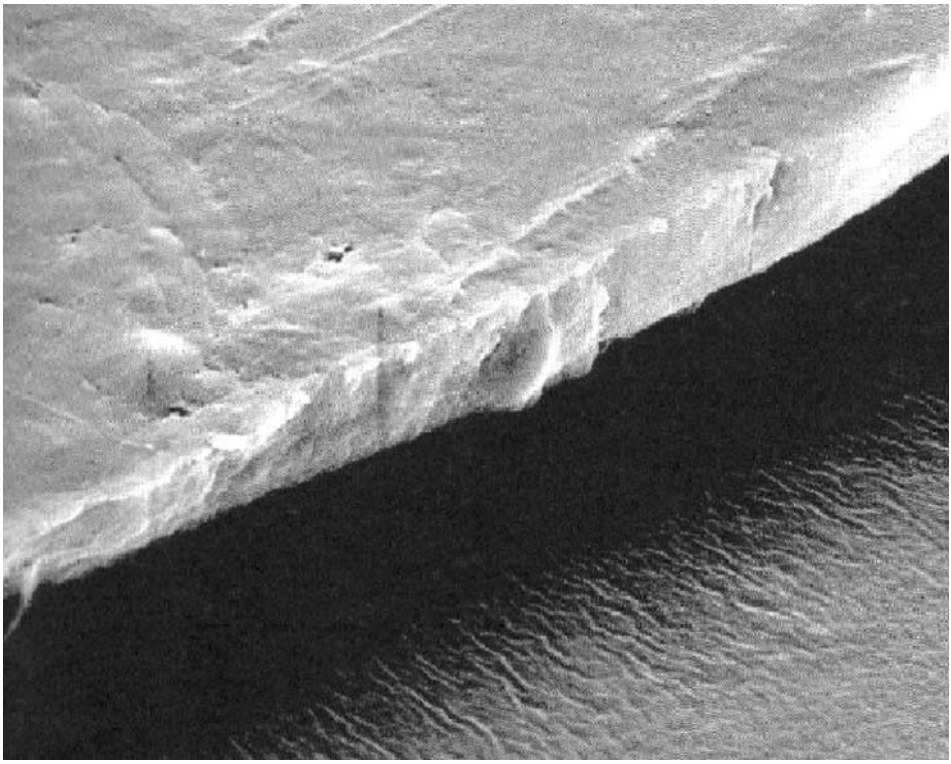


Figure 10
Thermoplastically deformed cellulose using laser radiation on a compressed disc of cotton fibers. (Schroeter [43])

Table 3
List of common aqueous cellulose solvents and their major applications

Solvent designation	Composition (of solvent or cellulose derivative)	Application[a]
Cuoxam (Schweizer's reagent)	Copper (II) tetra ammonium hydroxide	Viscosity molecular weight
Cuen	Copper (II) ethylenediamine hydroxide	Viscosity molecular weight
Cadoxen	Cadmium ethylene-diamine hydroxide	Viscosity molecular weight
Acetone or ethyl- or butyl acetate	Cellulose trinitrate	Viscosity molecular weight and molecular weight distribution
Acetone, THF, or dioxane	Cellulose tricarbanilate	Molecular weight distribution

[a]Viscosity molecular weight is calculated using the Mark–Houwink–Sakurada (MHS) equation of $[\eta] = K_m M_v^a$, where η is intrinsic viscosity, M_v is viscosity molecular weight, a is the MHS exponential factor (material and system-specific, between 0.9 and 1.0 for metal complex-based solvents), and K_m is a constant. Molecular weight distributions result from GPC experiments.

with OH-groups, respectively. A list of the most common aqueous solvent systems is given in *Table 3*.

Among the best-known nonderivatizing solvent systems is a combination between copper, alkali, and ammonia termed **Schweizer's reagent**. Solutions of cuprammonium hydroxide have been used for both analytical and industrial cellulose dissolution. Regenerated fibers with silk-like appearance and dialysis membrane have been (and partially continue to be) industrial products on the basis of cellulose dissolution in cuprammonium hydroxide. The success of this solvent is based on the ability of copper and ammonia to complex with the glycol functionality of cellulose as shown in *Fig. 11*. Because of the potential side reactions (oxidation and crosslinking, "Norman" compound formation), alternatives to both ammonia as well as copper have been developed. "Cuen" and "cadoxen" are related formulations based on the use of ethylene diamine and cadmium, respectively. The various combinations of alkali, ammonia,

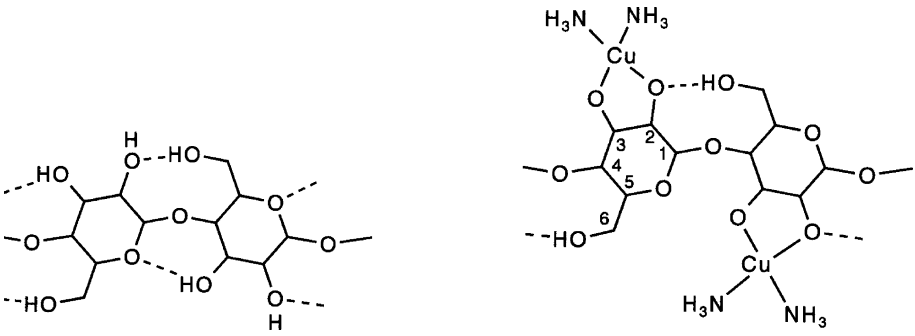


Figure 11
Complexation principle of cellulose with derivatizing solvent molecules illustrated for the case of cuprammonium hydroxide. Solvent molecules replace the existing hydrogen bonds with solvating Cu-complexes. (After Burchardt et al. [47])

and copper or cadmium continue to form the basis for numerous molecular weight determination methods for cellulose. The commerce of cellulose-rich (chemical) pulp relies heavily on molecular weight determinations using any one of these combinations (*Table 3*).
Advances of the past three decades, however, have produced alternatives to the heavy metal-based cellulose solvents. Prominent among them are dimethylacetamide/LiCl (DMAc/LiCl); N,MMN-O hydrate (NMMO); tetrabutylammonium fluoride/DMSO (TBAF/DMSO); and potassium thiocyanate/DMSO [48]. While many of these solvents have gained significant popularity among laboratory chemists, only the amine oxide solvent, N,MMNO, has achieved industrial practicality. This will be discussed in the section on regenerated cellulose fibers.
Several other important physical properties of cellulose are given in *Table 4*.

Table 4
Important physical properties of cellulose

Density[a]	1.61–1.62 g cm^{-3}
Heat of combustion[b]	17.5 Jg^{-1}
Heat of crystallization[c]	18.7–21.8 kJ mol^{-1} AGU
Specific internal surface[d]	10–200 m^2 g^{-1}
Refractive index[e]	
Parallel to fiber axis	1.62
Perpendicular to fiber axis	1.54
Breaking stress (MPa)	
- Wood[f]	500
- Ramie[a]	900
- Flax[f]	600–1100
- Sisal[f]	490–760
- Regenerated fiber (Fortisan®)[a]	2000
(Cordenka 700)[g]	825
(Lyocell)[h]	1400
Modulus (GPa)	
- Wood[f]	45
- Ramie	
- Flax[f]	45–100
- Sisal[f]	19–32
- Cellulose I fibers[k]	137–143
- Tunicate fibers[l]	143
- Regenerated fiber (Fortisan)[a]	
(Cordenka 700)[g]	19.5-27
(Lyocell)[h]	36
Degree of crystallinity[i]	
Cotton	83
Wood pulp[j]	46–56
Rayon	67

Sources: [a]French et al. [12]; [b, c, d, e, i]Kraessig et al. [13]; [f]Oksman, Selin [49]; [g]Ganster et al. [50] [h]Seavey et al. [51]; [j]Kadla, Dai [52]; [k]Sturcova et al. [53], and Sakurada et al. [54]; [l]Eichhorn [55]

2.4 Sources

Cellulose is the primary component of the cell walls of all major plants (Table 5). As cellulose content varies in relation to its source, so does the method of isolation.

Whereas cotton represents the purest form of cellulose, wood contributes the vast majority of it. Depending on end use, paper, board or chemical grade, wood conversion to some type of cellulose-rich fiber amounts to about 200×10^6 t/a a worldwide [29]. This compares to about 15×10^6 t/a cotton [13]. Only a minor amount, ca. 7×10^6 t/a, of high-purity cellulose is used for chemical purposes, mainly regenerated fibers (viscose rayon, lyocell) and derivatives (esters, ethers). Smaller amounts are also used for hydrocolloids [microcrystalline cellulose (MCC)].

Qualification of different cellulose sources for the various end use applications is determined on the basis of purity, molecular size, and *α-cellulose content*. α-cellulose refers to the portion of cellulose insoluble in 18% aqueous sodium hydroxide. Whereas the content of noncellulosic polysaccharides has proven to be a hindrance to the clarity of cellulose esters (determined as *haze* in otherwise clear films), α-cellulose content is important for the spinnability of cellulose solutions into regenerated fibers, and for viscosity characteristics of cellulose ethers. Molecular weights play an important role in various cellulose ethers.

The isolation of cellulose from wood may employ any one of several process options [13,15,16]. Chemical wood pulp may be derived by delignification using acidic solutions of sulfur

Table 5
Approximate cellulose content of different resources

Source		Cellulose content (%)	Content of heteropoly-saccharides (%)	Lignin content (%)	Other components (%)
Cotton[a]		92.7	5.7	0	1.0
Bast fibers[b]	Jute	64.4	12.2	11.8	1.6
	Flax	62.1	18.5	2.0	5.4
	Hemp	67.0	16.9	3.3	2.8
	Ramie	68.8	15.0	0.6	5.8
	Abaca	63.2	20.1	5.1	1.6
Wood[c]	Spruce (*Picea abies*)	43	27.1	29	0.4
	Birch (*Betula papyrifera)*	43	37.9	19	0.5
	Poplar (*Populus tremuloides*)	49	29.6	21	0.4
Paper pulp[d]	Softwood kraft	82			
	Hardwood kraft	72			
Dissolving pulp[e]	Sulfite	89–93			
	Kraft	93–98			

[a, b, c, d, e]Source: Kraessig et al. [13]
[c]Extractive-free composition
Source: Fengel and Wegener [15]

dioxide (*sulfite pulp*) or alkaline solutions most commonly containing sodium sulfide (*kraft* or *sulfate pulp*). An option based on the use of organic solvents (ethanol/water, *organosolv pulp*) has been practiced in Canada in the past. All process types may be applied to softwoods and hardwoods.

Commercial trade in chemical celluloses is assisted by the definition of standard methods of analysis. Extensive collections of quality-defining protocols have been assembled by ISO, the International Standards Organization; ASTM, the American Society for Testing Materials; TAPPI, the Technical Association of the Pulp and Paper Industry (of USA); DIN (German Industrial Norms); and many others.

3 Heteropolysaccharides ("Hemicelluloses")

3.1 Occurrence and Structure

The pure homopolysaccharide cellulose exists in nature at best only in the seed hairs of some plants (such as cotton). Whereas cellulose-producing viable cells always produce a pure, uncontaminated form of polymer, most cellulose generated exists as part of a multiphase composite material (wood and other lignocellulosic plant materials). Most cellulosic natural materials contain pure cellulose only on the nano-level where it interfaces with a variety of other molecules that contribute properties complementary to the pure homopolysaccharide. Whereas cellulose is well qualified to add mechanical strength to biological structures, it is limited by its high degree of crystallinity, which results in brittleness, low failure strain, and low ability to interact with other molecules. Yet the materials of nature typically have needs that transcend structural reinforcement. Control over such material properties as ductility, viscoelasticity, impact resistance (at all levels of temperature and humidity), biological degradability and digestibility, water retention, etc., all requires a level of variation in chemical structure that cannot be achieved with a highly ordered, crystalline homopolysaccharide alone. Since virtually all linear homopolymers crystallize, both biological organisms as well as synthetic organic chemists have learned to limit crystallinity by limiting order via structural variability. In plants, interfering with, and reducing, molecular regularity has been achieved by the introduction of several types of *different repeat units* into the polysaccharide chains, so as to create copolymers consisting of different sugar units; by introducing *branches* and *side groups*; and by introducing *derivatizing substituents* that interfere with crystallization and hydrogen bond formation. Nature uses all three different strategies for tailoring physical properties to the structural needs of a living plant.

Hemicelluloses represent a class of noncellulosic polysaccharides that is associated with cellulose in plant cell walls [56]. The term hemicelluloses was first used by Schulze in 1887 in the belief that *(hemi- or half)-celluloses* were perhaps components that were on the way of becoming cellulose; a term meant to distinguish this group of noncellulosic polysaccharides from that which makes up the cell wall structure. It has long been recognized that the term is unfortunate and misleading, and that *polyoses* or *heteropolysaccharides* are better descriptors [57]. However, "hemicelluloses" is an often-used designation for the noncellulosic heteropolysaccharide components in plants. The latter term shall be used in this text.

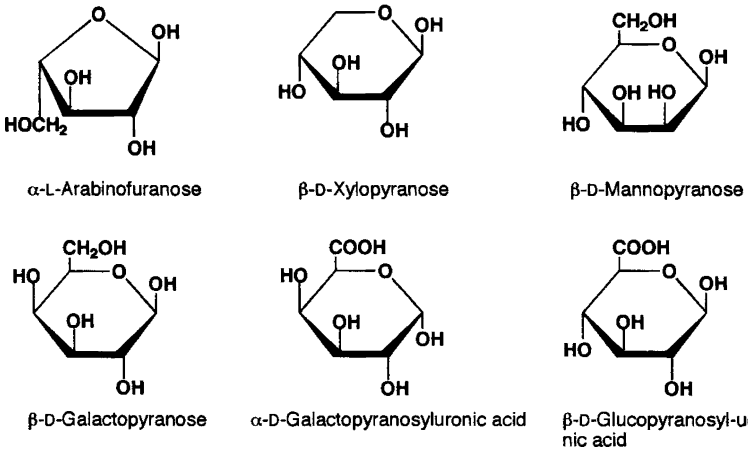


Figure 12
Most important monosaccharide units of plant heteropolysaccharides associated with cellulose (excl. β-D-glucopyranose)

It is the heteropolysaccharides of plants that bestow cellulosic composites with the ability to absorb impact, the ability to absorb moisture, and the ability to create pores in the form of free volume in amorphous (disordered or *para-crystalline*) materials [58,59]. Modification by reducing molecular regularity has the additional benefit of creating a transition from a focus on mechanical (structural) functions to an emphasis on energy storage and gel formation. Reduction in order translates into ease of hydrolysis, enzyme accessibility, rate of nutrient release for decay organisms, water absorption and swelling. Reduction in order is achieved by the introduction of monosaccharide units, and of bond types, which differ from those of cellulose. The principal monosaccharides involved in the heteropolysaccharides of plants are shown in *Fig. 12*.

Heteropolysaccharides are classified into five primary classes, xylans, glucomannans, arabinans, galactans, and glucans (*Table 6*) [60]. With few exceptions, heteropolysaccharides consist of linear homo- or copolymers with variable degrees of branching (usually by single monosaccharidic branches) and with occasional (3–13 wt.%) replacement of OH groups by *O*-acetyl groups.

Whereas cellulose contributes 40 to 60% by weight to plant mass, the balance is divided between primarily lignin and heteropolysaccharides (see *Table 5*). This division varies with plant source. The most prominent group of heteropolysaccharides in hardwoods and annual plants are the *xylans* (*Fig. 13*). Xylans have a backbone that consists of anhydro-D-xylopyranose units that are linked in β, $1 \rightarrow 4$ manner [60,61]. This gives xylan backbones the same structure as cellulose with the exception of the C-6 methylol group on the pyranose units. Pure and unbranched xylan therefore also crystallizes [63]. However, in plant tissues, xylans are heavily branched with 4-*O*-methyl-D-glucuronic acid and/or L-arabinose units. Many xylans also have a high degree of acetylation. It is known that low degrees of acetylation promote solubility in water.

Table 6
Classification of noncellulosic heteropolysaccharides

Polysaccharide type	Monosaccharide units present	Linkage types	Monosaccharide present in branches	Heteropolysaccharide designation
Xylan	D-Xyl *p*	β , 1 → 4	L-Ara *f*, 4–O-Methyl-D-Glu A *p*	Arabinoxylan, Glucuronoxylan Arabinoglucurono xylan
Glucomannan	D-Man *p* D-Glu *p*	β , 1 → 4	D-Gal *p*	Glucomannan, Galactoglucomannan
Arabinan	L-Ara *f*	α , 1 → 5	L-Ara *f*, D-Gal *p*	Arabinan
Galactan	D-Gal *p*	β , 1 → 4 and/or β, 1 → 3	L-Ara *f*, D-Gal *p* D-Glu A *p*	Arabinogalactan (2 types)
Glucan	D-Glu *p*	β , 1 → 4 or β, 1 → 3	D-Xyl *p*, D-Gal *p*, D-Ara *f* None	Xyloglucans β-Glucans

Source: Tenkanen [61]

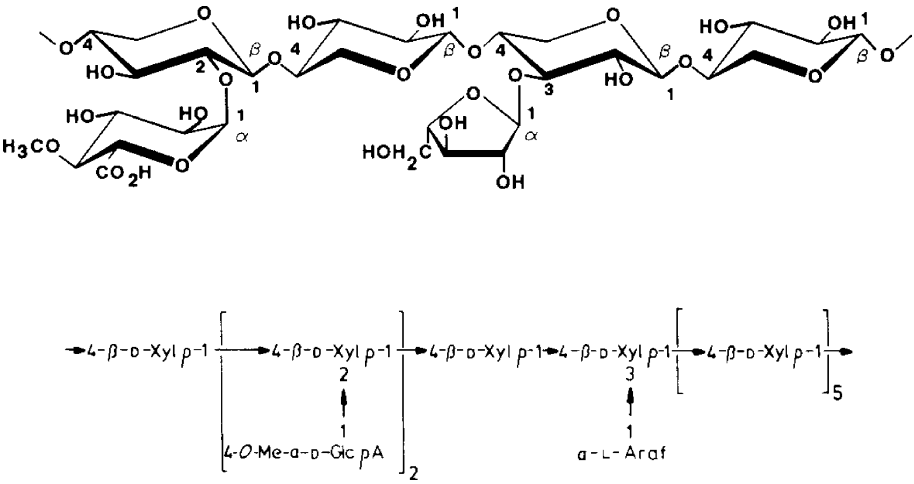


Figure 13
Schematic structure of arabinoglucuronoxylan (without acetyl groups). The formula below gives the abbreviation with some quantitative detail. (Taken from Sjostrom [16])

The second most abundant group of heteropolysaccharides are the *glucomannans* [60,61]. Glucomannans consist of a linear copolymer chain made up of glucose and mannose units in different ratios (although mostly favoring mannose). Glucomannans are also β,1 → 4-linked, and they occasionally carry single galactose side chains. Glucomannans are typical of soft-

wood species. Arabinans, galactans and glucans are present in various plant tissues to lesser extents.

3.2 Isolation

Heteropolysaccharides are not present in pure form although they often represent 20–30% of cell wall mass. Their chemical structures and some of their physical properties resemble those of the natural gums, although they are typically of modest molecular size (DP in the range of 50 to 200). They are also closely associated with pectins [58]. In woody plant tissues, heteropolysaccharides are typically connected to lignin via a variety of covalent bonds. This combination of two polymer classes with widely different chemical and physical properties, and with a molecular architecture resembling that of block copolymers, is responsible for the insolubility of both polymer components. In contrast to cellulose heteropolysaccharides are easier to dissolve in alkali or DMSO. Heteropolysaccharides are usually isolated from lignified plant tissues by pretreatment procedures that aim at removal of extractives, pectins, and lignin so as to produce a *holocellulose* that represents the combination of cellulose and heteropolysaccharides [62]. Holocellulose is subject to fractionation by (mostly) alkali extraction. The subject of heteropolysaccharide isolation from holocellulose has recently experienced a renaissance since the heteropolysaccharide fraction represents a significant mass fraction of plants [60,63]. Heteropolysaccharides as well as lignin both resist dissolution without depolymerization. Only the hydrolysis of covalent bonds existing between and within lignin and heteropolysaccharides will result in effective delignification or carbohydrate removal. This is why most heteropolysaccharide isolation protocols begin with the preparation of holocellulose. This involves various degrees of lignin extraction and/or lignin degradation (with strong, lignin-specific oxidants, alkali, and/or organic solvents) followed by heteropolysaccharide extraction using usually aqueous alkali.

Heteropolysaccharides are, in general, highly desirable components of paper-grade pulps while they are not tolerated in *dissolving-grade* pulps used for the production of regenerated cellulose products and cellulose derivatives. A vast body of literature exists on the effects of hemicelluloses on the papermaking process.

Although abundant in nature, and although several opportunities have been identified [64], heteropolysaccharides have largely escaped industrial exploitation. Exceptions are the conversion of pentoses into furfural (for temperature-resistant foundry resins) by acid hydrolysis; the fermentation of pentoses into single cell protein (*Torula yeast*); the catalytic reduction of xylose to xylitol, a dietary sweetener; and the use of larch (*Larix* sp.)-specific arabinogalactan extracts as dispersants in printing inks.

4 Microcrystalline Cellulose, Nanocrystals and Cellulose Whiskers

Being composed of linear chains of a highly crystalline polysaccharide, linked by β, $1 \rightarrow 4$, D-anhydroglucopyranose bonds that are hydrolyzed with relative ease, cellulose has long been the subject of partial acid hydrolysis to mixtures of hydrocolloid particles [65]. Hydrolysis of the disordered (paracrystalline) regions of cellulose liberates highly crystalline particles with dimensions in the 50 to 100 μm range. When dissolved in a suitable cellulose solvent, these par-

Table 7
Applications of microcrystalline celluloses

Application	Role of MCC
Suspensions	Thickening agent, stabilizer
Emulsions	Emulsifier, stabilizer
Ceramics	Green strength and glaze hardening
Food	Dietary filler reduces caloric value
Cosmetics	Moisture retention
Pharmaceuticals	Excipient, lubricant, building agent

Source: Battista [68]

ticles represent materials with molecular weights corresponding to DPs in the range between 50 and 400 depending on cellulose source. Since these molecular weights are determined by the hydrolysis of the amorphous regions existing between crystalline areas, they received the term *level off DP* (LODP).

The discovery that cellulose is capable of forming a stable colloidal suspension of microcrystallites if this hydrolysis is carried out with a nonpeptizing acid (such as HCl) followed by mechanical disintegration led to the introduction of a commercial *microcrystalline cellulose* product under the trade name AVICEL (FMC Corp., Philadelphia, PA) [66,67,68]. The successful production of MCC requires hydrolysis and sufficient mechanical disruption of a $>5\%$ aqueous suspension to allow the liberation of monocrystals in stable gel-form. Since its introduction in 1961 [66], MCC has been recognized for its unique rheological properties, which distinguish it as a gel with stable viscosity at ambient as well as elevated temperatures. Its smooth texture produces a favorable mouthfeel that, in turn, has helped MCC capture markets in the food industry. Its viscosity properties have helped MCC to find entrance into such applications ranging from ceramics to food to cosmetics and to pharmaceuticals (*Table 7*). In comparison to gels produced using hydrophilic macromolecules, such as natural gums, MCC does not form rubbery gels upon exposure to the air. Following dialysis and spray drying, MCC forms flowable powders of micron-sized particles that are widely accepted in tableting operations because of their excellent compression properties. Annual production figures of 550×10^3 t/a by the FMC Corp. have been reported in the literature [68].

Continued hydrolysis of MCC with sulfuric acid was found to lead to the separation of nanosized crystals, variably referred to as whiskers, monocrystals, nanocrystals, cellulose crystallites, etc., which represent the crystalline phase of cellulose [69,70,71]. Depending on source, these crystallites have dimensions of 5 to 15 nm in diameter and 200 to 1000 nm in length. Their aspect ratio (i. e., length divided by diameter) may vary from nearly 1 to >100 [71]. It is apparent that MCC consists of agglomerations of whiskers (*Fig. 14*). The retention of sulfate groups on whisker surfaces helps in the formation of stable dispersions in water. Whiskers are reported to have mechanical moduli of about 130 GPa, and this corresponds closely to values calculated for cellulose crystallites [71]. Cellulose whiskers have attracted significant attention as reinforcing agents in composites.

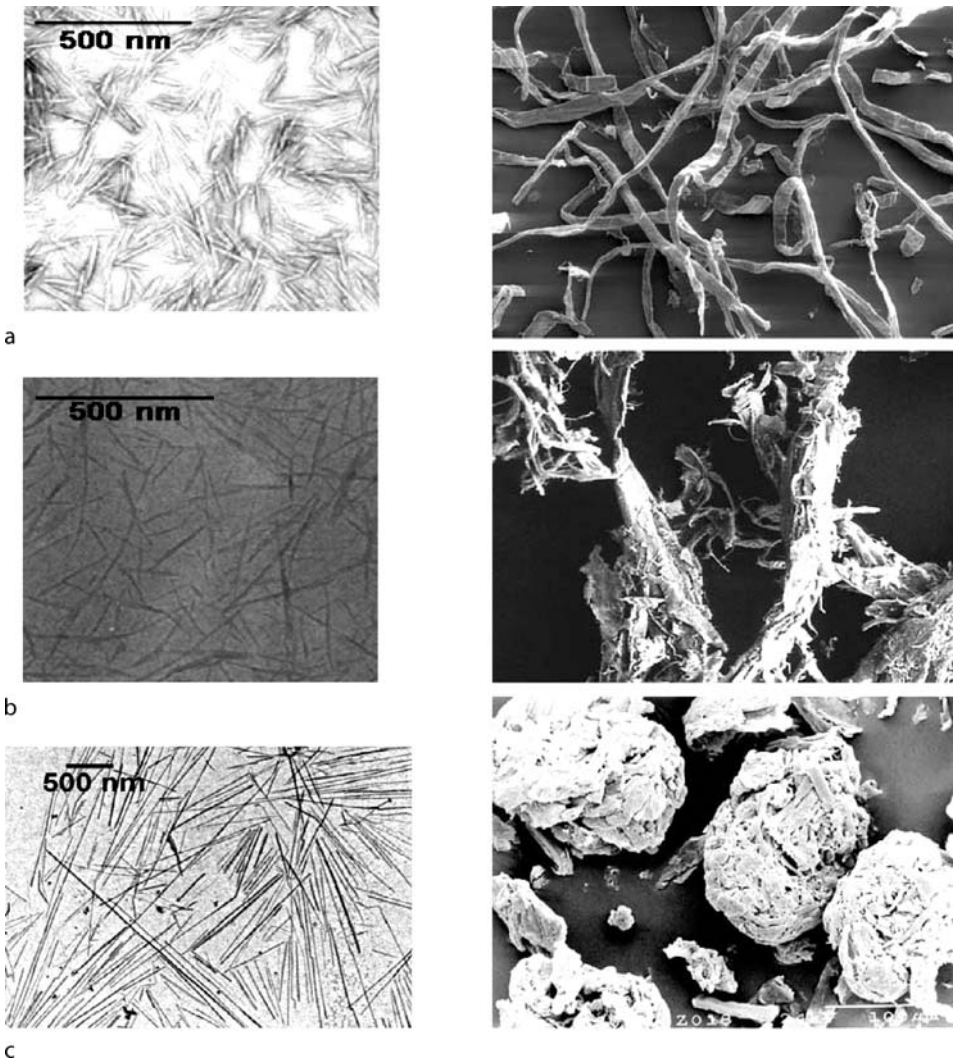


Figure 14
Transmission (left) and scanning (right) electron micrographs of nanocrystals (from cotton (a), sugar-beet pulp (b), and tunicin (c)), left; and of wood (chemical pulp) fibers (*top*), hydrolyzed pulp fibers (*center*), and microcrystalline cellulose (*bottom*), right. (From Dufresne et al. [71] and Kraessig et al. [13])

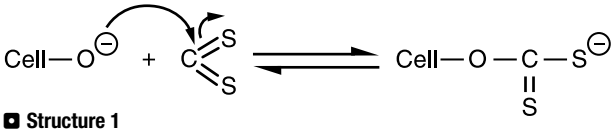
5 Cellulose Regenerates

Cellulose is soluble in a wide range of solvents from which it can be regenerated in unadulterated (crystalline) form [13]. This is accomplished in either fiber (especially staple fiber) form, in filament form or in any other form (sheets or films, sponges, etc.). Cellulose solvents are divided into *derivatizing* and *nonderivatizing* solvent systems [73]. The principle of deriva-

tizing solvents involves the formation of transient cellulose derivatives that have substituent groups that are covalently attached and that can easily be removed during the regeneration step. Nonderivatizing solvents, by contrast, are those solvents that are capable of dissolving cellulose on the basis of intermolecular interaction alone. However, the delineation is less precise when it comes to solvent systems that form strong secondary interactions. Three types of cellulose regeneration technologies are currently in industrial practice, and these are the *viscose rayon process* (derivatizing), the *lyocell process* (nonderivatizing) and the *cuprammonium hydroxide process* (nonderivatizing). Other cellulose fibers of commercial significance are cellulose acetate (CA with DS circa. 2.5) and cellulose triacetate (CTA with DS nearly 3.0) fibers and filaments. However, these do not qualify as cellulose regenerates as they retain their substituent groups.

The oldest and most widely practiced cellulose regeneration technology of the derivatizing solvent-type is the ***viscose rayon*** process. It is based, in part, on the discovery in 1857 by Cross, Bevin, and Beadle of the dissolution of sulfidized cellulose in alkali. Viscose fibers are by far the most important cellulose regenerates, amounting to an annual fiber production of 2.5×10^6 t worldwide [13,74].

The viscose rayon process (◉ *Fig. 15*) follows a process flow diagram that distinguishes several steps. Cellulose (usually in sheet form) is alkalized with 20% aqueous alkali and *ripened* to a reduced DP of approximately 300. This may be accomplished in continuous or discontinuous operations. The ripened alkalized cellulose is then subjected to a chemical modification with carbon disulfide (CS_2) in a *xanthation* step. Cellulose xanthate requires a DS of 0.5–0.6, amounting to the use of 150–400 kg of CS_2 per ton of fiber. Xanthation occurs on all three available OH-groups, but with predominance of C-2 and C-6. Following reaction with CS_2, the cellulose xanthate is dissolved in alkali to yield a sticky, honey-colored solution with 7–12% cellulose xanthate and 5–8% alkali. This solution is filtered and deaerated before it is spun (through spinnerets) into a precipitation bath containing sulfuric acid and zinc salts. The chemistry of the complex process is illustrated by the following equation.



Structure 1

Complexity, environmental difficulties (especially with CS_2 regeneration) and overall slow processing conditions (compared to melt-spun synthetic fibers) catalyzed the search for alternative cellulose regeneration processes.

The ***lyocell process*** is based on the regeneration of cellulose from a true (underivatized) solvated molecular state. The solvent, N, methyl morpholine-*N*-oxide (NMMO)-hydrate, dissolves cellulose on the basis of its strong hydrogen bonding capacity. NMMO is a substance with a melting point of 80 °C that can be fully recovered in the process. American Enka developed the process in 1976, and it is presently commercially operated by Accordis (formerly Courtaulds) with operations in Mobile, Alabama, USA, and Grimsby, United Kingdom (70×10^3 t/a) and in Heiligenkreuz, Austria (Lenzing, AG). Most operations generate staple fibers for textiles that are marketed under the trade name *Tencel*®; no filaments are produced as yet [13,74].

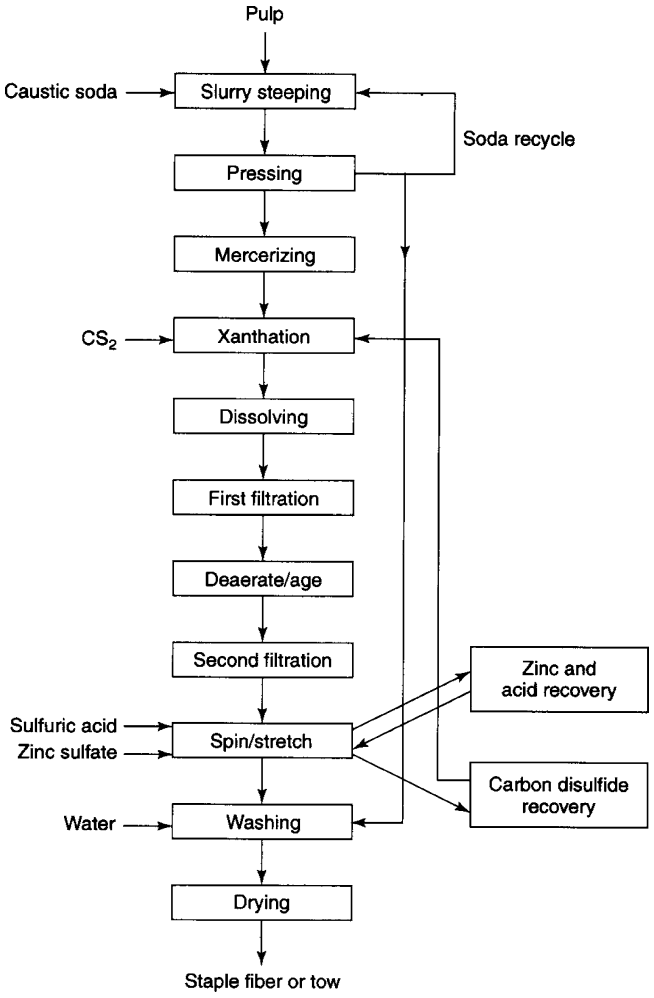


Figure 15
Process diagram of the *viscose rayon* process (Adopted from Woodings [74])

The process of cellulose regeneration in the form of lyocell fibers is significantly simpler than that of the viscose rayon process. It is illustrated in *Fig. 16*. A solution containing 14% cellulose, 10% water, and 76% NMMO plus stabilizers is extruded at a temperature slightly above 100 °C into an aqueous NMMO-bath from which cellulose is precipitated [74,75]. The extrusion has been described as a melt-spinning process that has recently made it possible to manufacture cellulosic self-bonded meltblown nonwovens as well [76].

The third type of cellulose regenerate with commercial significance operates on the basis of ***cuprammonium hydroxide*** (Schweizer's reagent). The process had significance for fibers in the first half of the 20th century, but it started to decline in the 1960s due to insurmountable obstacles in the recovery of copper ions. The main producer of cuprammonium fibers

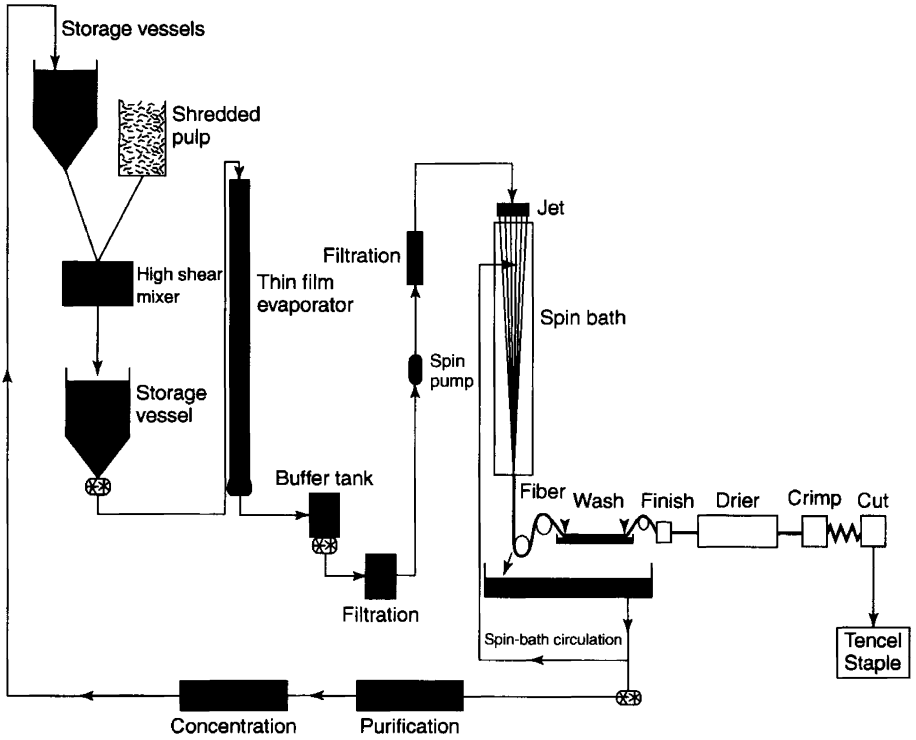


Figure 16
Regenerating cellulose in the form of lyocell fibers from a solution in NMMO according to the *Tencel* Process by Accordis (From Woodings [74])

(Bemberg AG of Germany) shifted the production to that of hemodialysis membranes in the form of flat sheets and hollow fibers. The membrane products are known under the name *Cuprophan* and are produced by *Membrana*, a division of AKZO Nobel Fibres [13,75].
The spinning of *cuprammonium* hydroxide fibers involved solutions of 4–11% cellulose, 4–6% Cu, and 6–10% NH_3. Solution viscosities were 200 Pa × s at 20 °C.
Cuprammonium fibers were popular for their superior fiber feel and their silk-like appearance, which are based on the extreme fineness of the filaments.
Alternatives to alkaline solutions of carbon disulfide are combinations of formaldehyde and DMSO; various alkyl silanes, and urea and alkali. The latter is currently being promoted for industrial applications. It is known under the term *carbamate process*. The carbamate process has advantages over the viscose rayon process while being able to use the same process technology.

6 Cellulose Derivatives

Cellulose derivatization has the purpose of (a) *reshaping* cellulose from its original cotton or wood fiber form; (b) *modifying* its chemical and/or physical properties (such as solution,

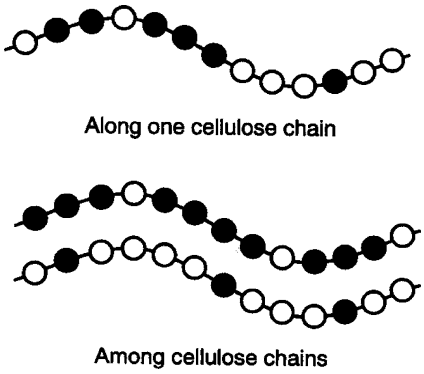


Figure 17
Schematic illustration of the possible distribution of substituents in cellulose derivatives with variation along single chains or among adjacent chains. The resulting *block-like* architecture of the former and the corresponding *over-* and *under*-derivatization of the latter both potentially restrict thermal and solubility properties of the derivatives

thermal, or physiological behaviors); or (c) making cellulose amenable to *analysis* by a specific analytical methodology, such as molecular weight or chemical structure determination.

Since the cellulose molecule is virtually monofunctional in terms of aliphatic hydroxyl groups, derivatization reactions target predominantly *substitutions* of OH-groups. In addition to substitution reactions, *oxidations*, at the reducing end group, at the C-6 methylol group, or at the glycolic group (C–C cleavage between C-2 and C-3 with formation of a dialdehyde), also play a role. A third type of reaction involves *hydrolytic depolymerization*.

The important class of *derivatization by substitution* is measured in terms of the degree of substitution, DS, which may range between 0 and a maximum of 3.

The introduction of substituents is rarely uniform in practice. In most cases, modification reactions occur preferentially in the most accessible regions of the molecule or at the most reactive OH-groups (◉ *Fig. 17*). The three OH groups are not entirely equivalent, as they differ in accessibility within the crystalline organization as well as in their relative reactivity. The reaction sequence usually progresses from C-6 (most accessible and reactive), to C-2 and C-3. In addition to variations in the pattern of substitution within an AGU repeat unit, differences are encountered along the chain (◉ *Fig. 17*). Since most reactions with cellulose begin as heterogeneous reactions, diffusion and accessibility limitations play an important role in the progression of the reaction. The mixture of crystalline and disordered regions within cellulose produces a more rapid derivatization within the amorphous phase, and this results in *blockiness* whereby regions of high DS neighbor regions of unsubstitution. The resulting irregularities tend to influence thermal properties and solution behaviors. Blockiness is prevented either by conducting derivatization reactions in homogeneous phase (i. e., in solution state); or by completing derivatizations to a high DS followed by substituent removal under homogeneous conditions.

Only reactions in homogeneous phase, in dilute solution, can be expected to minimize variations within and between cellulose derivative molecules.

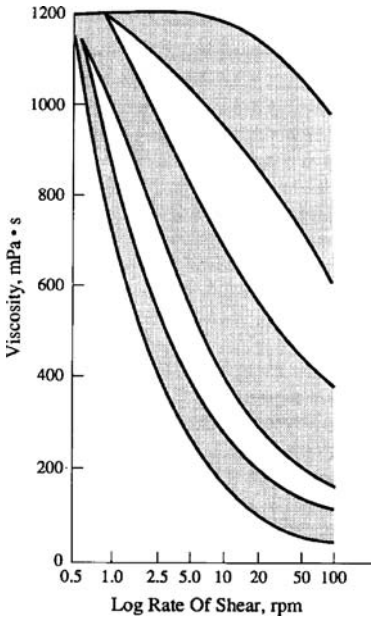


Figure 18
Relationship between viscosity and shear rate. The reduction in viscosity with rising shear rate describes the degree of *shear thinning*, an important material characteristic of water-soluble polysaccharides. The mouthfeel of foods becomes less slimy with increasing shear thinning behavior. (Adopted from Szczesniak and Farkas [77])

The chemical structure of cellulose ethers is typically described on the basis of total product hydrolysis and analysis of the resulting monosaccharide derivative units. The yield of mono-, di- and tri-substituted glucose units, and the distribution of substituents between the hydroxyl groups in positions C-2, C-3, and C-6, represents distinguishing molecular features. The determination of degrees of substitution (DS) and molar substitutions (MS), representing the number of substituent equivalents per anhydroglucose unit, widely dictate rheological and other properties. Additional analytical methods describing the properties of cellulose ethers involve viscosity measurements. Cellulose ethers may be both thixotropic and pseudoplastic (*shear thinning*) as is illustrated in *Fig. 18*.

Since cellulose and cellulose derivatives have found wide acceptance in the cosmetic, pharmaceutical and medical industries, much work is currently focusing on highly selective modification of cellulose ethers (as well as cellulose esters) for highly targeted controlled release characteristics. Interaction with water, and interaction with a variety of co-substrates under conditions varying in temperature, ionic strength, pH, etc., has opened avenues for the use of cellulose ethers in drug delivery. This field has recently been reviewed [78,79].

The two primary categories of cellulose derivatives, *cellulose esters* and *cellulose ethers*, which dominate the markets for cellulose derivatives (*Table 8*), have somewhat different raw material requirements. Whereas both types of modification technologies prefer to start with chemical cellulose of high α-cellulose content, ethers give preference to molecular weight considerations while esters are sensitive to impurities. This is a consequence of the predom-

Table 8
Estimated cellulosic materials markets excl. paper pulp and textiles (in tons/a worldwide)

Chemical pulp		4.4×10^6
Microcrystalline cellulose[1]		$0.5–1.0 \times 10^6$
Regenerated cellulose products	Viscose rayon	2.5×10^6
	Lyocell	82×10^3
	Cuprammonium hydroxide	Insignificant
Cellulose esters	Nitrates	$150–200 \times 10^3$
	Phosphates	Insignificant
	Sulfates	Insignificant
	Acetate and mixed acetates	850×10^3
Cellulose ethers	Carboxymethyl	230×10^3
	Methyl and hydroxyalkylmethyl	120×10^3
	Hydroxyethyl	60×10^3
	Hydroxypropyl	10×10^3

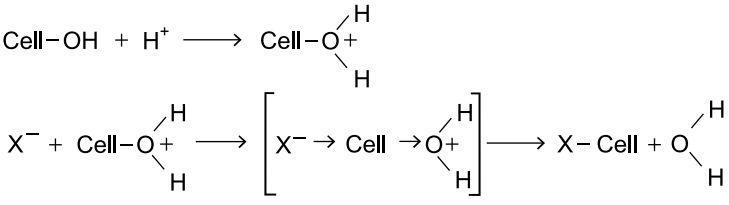
[1] Estimated based on the assumption that FMC has between 50 and 100% of world market.

inant application of both types: Whereas ethers find their greatest use in connection with the modification of aqueous solution viscosities (*rheology modifiers*), esters are often used for structural purposes where visual appearance and clarity matter most.

6.1 Cellulose Esters

Cellulose esters are used primarily in structural applications, such as films, fibers, coatings, etc. Since they are processed in solution or in melt state, their resistance to flow (i. e., melt viscosity) represents a distinct handicap. Commercial cellulose esters are all generated by heterogeneous modification reaction [80,81,82,83,84]. However, reactions in homogeneous phase (with various solvents) are increasing in number and variety, and they have recently been reviewed by Heinze et al. [82]. Cellulose esters are usually classified as *inorganic* and *organic* esters.

In general, all esters are based on the stepwise formation of oxonium ions followed by their nucleophilic substitution with an acid residue and the elimination of water (*Structure 2*). Because this substitution reaction reverses in the presence of water, water removal by binding



Structure 2

or other means is required. Esterification proceeds in heterogeneous phase, and derivatization does not necessarily destroy the fiber structure.

6.1.1 Inorganic Esters

Inorganic cellulose esters are formed with nitric, sulfuric, and phosphoric acid [13,80,85]. Of greatest commercial significance are the *cellulose nitrates* (sometimes erroneously designated *nitrocellulose*), which have first been described in 1832 by H. Braconot. Cellulose nitrates are formed by treating cotton (or chemical cellulose) with mixtures of nitric and sulfuric acid. Their ready inflammability has given them the name *gun cotton*. When A. Nobel discovered that mixtures of nitroglycerine and cellulose nitrate form a stable form of an explosive mixture known as *blasting gelatin* in 1875, this discovery, and the development of smokeless gun powder, changed human history forever.

Cellulose nitrates were used as base films in early photography, where they were plasticized with camphor and other substances. The frequent fires experienced with early motion pictures, which were produced using cellulose nitrate as basis, are still responsible for the stringent fire code regulations in today's movie theaters.

Cellulose nitrate continues to be used in fine lacquers and automotive coatings. The nitrates are easy to produce, and simple mixtures with other polymers and plasticizers are readily formulated. Their use depends on the DS, which is determined via nitrogen content. The target-DS is adjusted via the water content maintained in the reaction mixture (› *Fig. 19*). Cellulose nitrates of low DS (10.9 to 11.7% nitrogen) provide for solubility in a range of organic solvents, and derivatives with high DS are used for gun cotton.

Cellulose nitrates are often used for the determination of molecular weights and weight distributions of cellulose on analytical scale.

Cellulose sulfates are appealing water soluble polymers even at low DS. Cellulose sulfation, however, is complicated by serious chain degradative side reactions. Cellulose sulfation has been carried out using chlorosulfonic acid, sulfuric acid, or sulfur trioxide. At low degrees of substitution (DS 0.2 to 0.3), cellulose sulfates are used in oil well drilling, foodstuffs, cosmetics, and pharmaceuticals. The preparation of cellulose sulfates for symplex capsules can

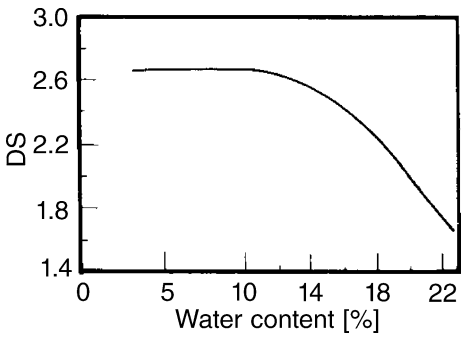


Figure 19
Relationship between DS and water content during cellulose nitration with HNO_3 and H_2SO_3. (From Klemm et al. [18])

be accomplished using cellulose acetate (DS 2.5) in DMSO-solution with SO_3-pyridine or SO_3-DMF. The sulfate substituents survive the ensuing deacetylation with NaOH. Interest in medical and pharmaceutical applications of cellulose sulfates is increasing lately [79,81].
Cellulose phosphate can be prepared by converting cellulose with molten H_3PO_4 and urea at 120 °C. The derivative with DS 0.3 to 0.4 has a high water retention value and is used in the textile industry for flame retardancy, as thickeners, and for membranes targeted for hemodialysis [79,86].

6.1.2 Organic Esters

Cellulose can be esterified with almost any organic acid [81,82,84]. Whereas this is possible in principle with both alkyl chlorides and carboxylic anhydrides, commercial practice has focused on the anhydride option. Although many esters have been described in the literature, industrially manufactured organic esters are prepared only with aliphatic fatty acids with between 2 and 4 carbon atoms in length (i. e., acetates to butyrates, CA to CB). Exceptions are some mixed esters with phthalic acid, which are used for enteric coatings in pharmaceutical applications; and a novel carboxymethyl cellulose acetate butyrate (CMCAB), which is used in water-borne coatings applications [81,84].
Cellulose esters have been known for 150 years [83,85]. However, their utility was not fully appreciated until after a significant industrial breakthrough (credited to G. W. Miles) revealed the well-soluble cellulose diacetate in 1904. Whereas the initial applications of cellulose diacetate focused on replacing the highly flammable cellulose nitrates used in airplane coatings during World War I, there was increasing demand for cellulose acetates throughout the first half of the 20th century. Principal applications for cellulose acetates were in the textile fiber business; in cigarette filter tows; in plastics; and in films and coatings. Applications for cellulose esters in specialty markets, such as LCD flat screens and controlled release agents, continue to be developed today [81,83].
A recent monograph on cellulose acetates lists more than ten acetate manufacturers worldwide (*Table 9*) [83]. The industrial process distinguishes several process steps (*Fig. 20*).

Table 9
List of cellulose acetate manufacturers

Eastman Chemical Company, USA
Rhodia Acetow, Germany
Celanese Acetate LLC, USA
Acordis, United Kingdom
Daicel Chemical Industries, Ltd., Japan
Acetati SPA, Italy
Voridian Chemical Company, USA
Inacsa, Spain
Mitsubishi Rayon Co., Ltd., Japan
Novceta SpA, Italy
SK Chemicals, S. Korea

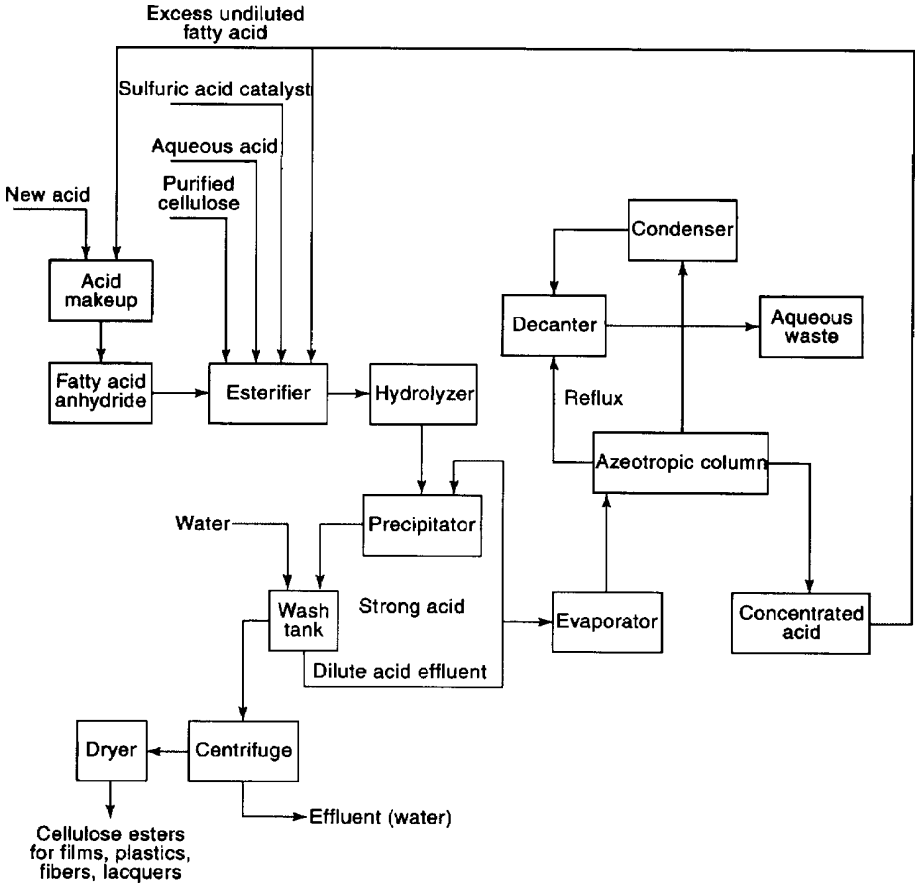


Figure 20
Flow diagram of cellulose ester (acetates, CA, and mixed acetates with propionic or butyric acids, CAP or CAB) production. Note the introduction of water via dilute acid serves the purpose of reducing DS and raising solubility in organic solvents. (Adopted from Edgar [87])

Pretreatment/activation involves swelling mostly chemical pulp fibers with glacial acetic acid and 4–7% moisture (on cellulose basis). Activation ensures uniform acetylation.
Esterification involves the treatment of preactivated cellulose with a mixture of acetic acid and acetic anhydride at a consistency of about 10%. The acetic anhydride content of the mixture is approximately 25%, and this amounts to a 10–40% excess over stoichiometric requirement. The reaction proceeds at 50 °C. ***Hydrolysis*** refers to a process step in which the DS of the esterification product is reduced from its level of > 2.9 to a level between 2.3 to 2.7 by the addition of 5 to 10% water. The temperature during hydrolysis is maintained at 40 to 80 °C. The consequence of hydrolysis is a uniform release of acetyl groups from the cellulose backbone so as to assure uniform solubility. ***Precipitation*** involves either the addition of the acetylation mixture into water or dilute acetic acid to produce a *flake* product, or the addition of water

Table 10
Solubility characteristics of commercial cellulose acetates in relation to DS (expressed in the form of acetic acid content). (Adapted from Balser et al. as quoted in Rustemeyer [83])

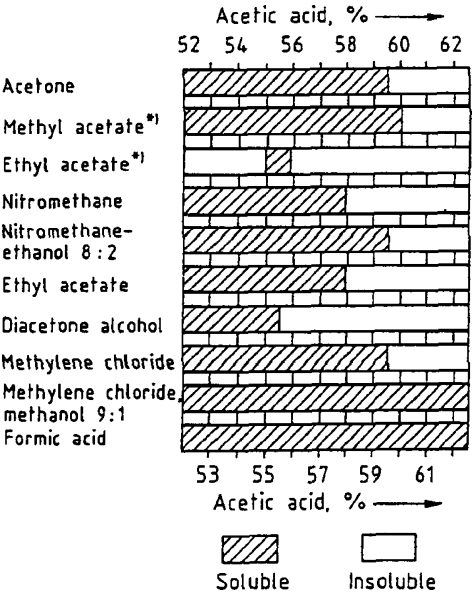


Table 11
Cellulose acetate solubility

DS	Acetyl content (%)	Soluble in
2.8–3.0	43–44.8	Methylene chloride
2.2–2.7	37–42	Acetone
1.2–1.8	24–32	Methoxyethanol
0.6–0.9	14–19	Water

into an ester mixture under stirring to generate a powderous product. Critical performance measures for cellulose acetate and mixed acetates involve solution behavior (including water dispersibility), thermal properties, and controlled release characteristics.

Cellulose ester solubility varies widely. Cellulose acetates with uniform substituent distribution are soluble in several organic solvents and these range from cellosolves (ethylene glycol ethers) to acetone and methyl chloride (*Table 10* and *Table 11*). Whereas the cellulose triacetate with >60% acetic acid content is barely soluble in chloroform or methylene chloride/methanol, a cellulose acetate with a DS of 0.6 to 0.9 (14 to 19% acetyl content) is fully soluble in water (*Table 11*). Important thermal properties depend largely on type and pattern of substitution (*Table 12*). Whereas cellulose acetates are not melt processible without the addition of large amounts of plasticizer, cellulose mixed acetates with propionyl or butyryl groups, CAP and CAB, have lower melting points depending on the degree of fatty acid substi-

Table 12
Cellulose acetate and mixed acetate properties (From Edgar et al. [88]; Rustemeyer [83])

	CTA	CA	CAP	CAB
Acetyl content (%)	44.8	37–42	0.6–2.5	2.0–29.5
Propionyl content (%)	–	–	42.5–46	–
Butyryl content (%)	–	–	–	17.0–52.0
Hydroxyl content (%)			1.8–5.0	1.0–1.8
Melting range (°C)			188–210	130–240
T_g (°C)		230[1]	142–149	96–161

[1] T_g (K) = 523–20.3 DS

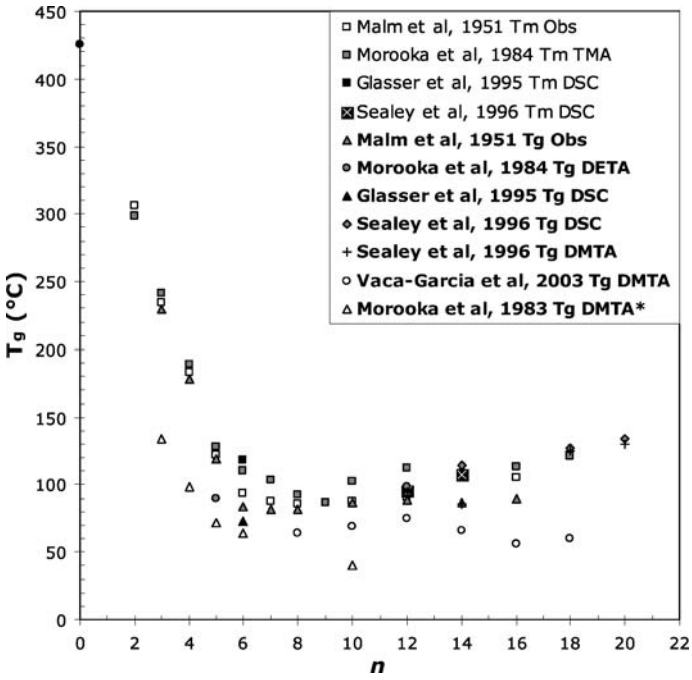


Figure 21
Relationship between apparent glass transition temperature (referred to frequently by different designations) and size of alkanoate substituents in C-atoms (*n*) of cellulose triesters. Data obtained from Malm et al. [90], Morooka et al. [91,92], Glasser et al. [93], Sealey et al. [94], Edgar et al. [88], Vaca-Garcia et al. [89]

tution. This has recently been evaluated for cellulose triesters including fatty and waxy esters (*Fig. 21*) [88,89].

Extensive research into the thermal properties of long-chain cellulose esters (LCCE) has produced an understanding of the melt- and flow properties of fully as well as partially substituted derivatives. It is apparent that T_g declines quickly to between 50 and 150 C as acyl chain length exceeds that of butyrate (*Fig. 21*).

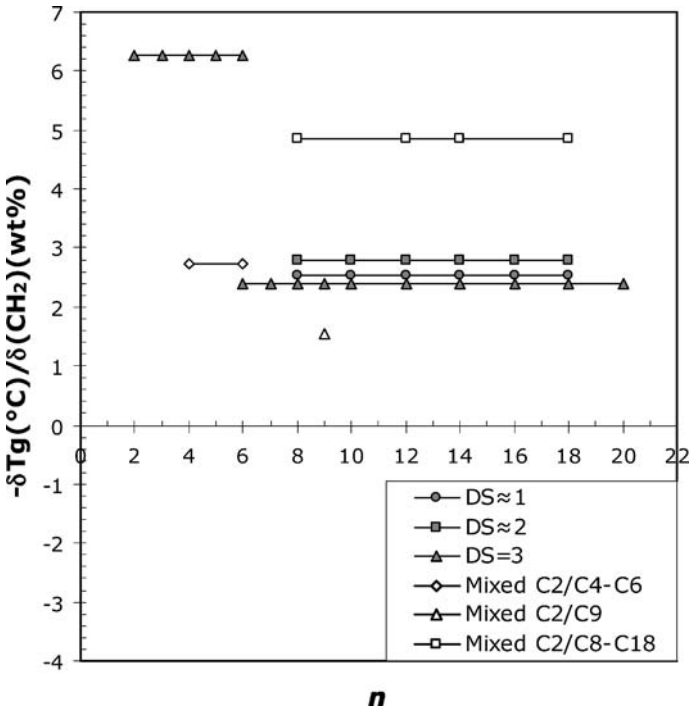


Figure 22
Relationship between the average reduction in apparent glass transition temperature per weight percent methylene, i. e., $\partial T_g/\partial(CH_2)$ (wt.%)-parameter, and alkanoate substituent size (n) for the different data sets of Fig. 21. (Corrected for liquid crystalline-phase content of cellulose esters with $n > 6$). (Adopted from work by Vaca-Garcia et al. [89])

A comparison of the extensive data sets available for LCCEs has recently produced the recognition that the thermal characteristics of cellulose esters are the result of an internal plasticization by methylene groups, for the most part (*Fig. 22*). Whereas the highest plasticization effect is observed for C2 to C6 triesters, most LCCEs with substituents greater than hexanoate experience a plasticization effect of ca. 2.5° C/CH_2(wt.%). Plasticization with methylene groups is reduced by the presence of OH-groups, and by the formation of liquid crystalline phases [89].
Controlled release applications for a variety of cellulose esters have recently been summarized by Marchessault et al. [86] as well as by Edger [79].

6.2 Cellulose Ethers

The formation of cellulose ethers aims, in general, at permanent and irrevocable disruption of crystallinity, and at controlling viscosity properties of aqueous solutions (i. e., *rheology modification*). Since "stiffening water" and gel formation are primary targets for many natural polysaccharides (in gums as well as body fluids), cellulose ethers aim to achieve solutions with

high viscosity at low solids content. This is why cellulose ether manufacturers prefer cotton over wood pulp, which has a lower degree of polymerization.

$$\text{Cell–}\overline{\underline{O}}\text{H} + \text{Na}^{+}\text{OH}^{-} \rightarrow \text{Cell–}\overline{\underline{O}}|^{-}\ \text{Na}^{+} + \text{H}_2\text{O}$$

$$\text{Cell–}\overline{\underline{O}}|^{-}\ \text{Na}^{+} + \overset{\delta+}{\text{CH}_3}\text{–}\overset{\delta-}{\text{Cl}} \rightarrow \text{Cell–}\overline{\underline{O}}\text{–CH}_3 + \text{Na}^{+}\text{Cl}^{-}$$

Structure 3

Cellulose ethers are generally formed by *alkalizing* high purity cellulose fibers in suspension, under conditions involving high alkalinity and low water content [95]. Industrial cellulose (*Structure 3*) ether reactions all involve heterogeneous conditions in organic solvent suspension. Low temperatures help preserve molecular weight. Much recent work has focused on the differences in properties between homo- vs. heterogeneously prepared cellulose ethers [46]. Homogeneous reactions, in general, produce derivatives with a more uniform distribution of substituents along the backbone of cellulose, and they allow a wider range of degree of modification. However, homogenous cellulose modification chemistry is not practiced commercially at present.

Cellulose ethers are formed from alkalized cellulose in (a) a typical Williamson synthesis using alkyl halides; (b) using ring opening reactions with oxiranes; or (c) by reaction with vinyl compounds. Cellulose ethers with industrial significance include various alkyl ethers (ca. 100,000 tons per year), hydroxyalkyl ethers (ca. 50,000 tons per year), and carboxymethylethers (ca. 300,000 tons per year) (see *Table 8*). These products are used as thickeners and dispersants in agriculture and food products, as well as in ceramics, cements, textiles, pharmaceuticals, cosmetics, paints, and detergents (*Table 13*). Four major cellulose ether categories are distinguished, and these include carboxymethyl cellulose (CMC); methylcellulose and hydroxyethyl methyl cellulose (HAMC); hydroxyethyl cellulose, also with varying degrees of hydrophobic substitution; and hydroxypropyl cellulose (HPC) (*Table 13*).

Table 13
Major applications for cellulose ethers

Application category	CMC	HAMC	HEC	HPC
Textile chemicals	✓		✓	
Paper chemicals	✓		✓	
Detergents	✓			
Cosmetics and pharma chemicals	✓	✓	✓	✓
Food	✓	✓		✓
Construction and building (plasters, etc.)		✓	✓	
Oil well chemicals	✓		✓	
Latex paints		✓	✓	
Ceramics				✓
Adhesives (wallpaper, tiles, etc.)		✓	✓	✓
Emulsion polymerization		✓		✓
Films				✓

Table 14
List of cellulose ether manufacturers

CP Kelko
Noviant
Dow Chemicals Comp.
Wolff Cellulosics
Shin-Etsu Chemical
Clariant AG
Hercules

Cellulose ethers swell or are colloidally soluble. They all raise viscosity and have specific rheology profiles (see *Fig. 18*). Many cellulose ethers are surface-active, and they reduce the surface activity of water by about 20 to 25% depending on type. Cellulose ethers are generally compatible with other hydrocolloids and many other substances. The field has recently been reviewed by Majewicz et al. [95]. The primary manufacturers of industrial cellulose ethers are listed in *Table 14*.

There are three distinguishing characteristics of cellulose ethers, and these relate to the behavior during dissolution; the structure and rheological properties of the solutions; and the ability to interact and blend with other solids. These behaviors depend on a variety of molecular features including type of substituents, number of substituents, distribution of substituents, and molecular mass. Most cellulose ethers are *pseudoplastic* or *shear thinning*, meaning that they exhibit reduced viscosity under conditions of applied shear (*Fig. 18*). Many of the cellulose ethers also coagulate with temperature rising: they are thermally reversible gels and flocculate

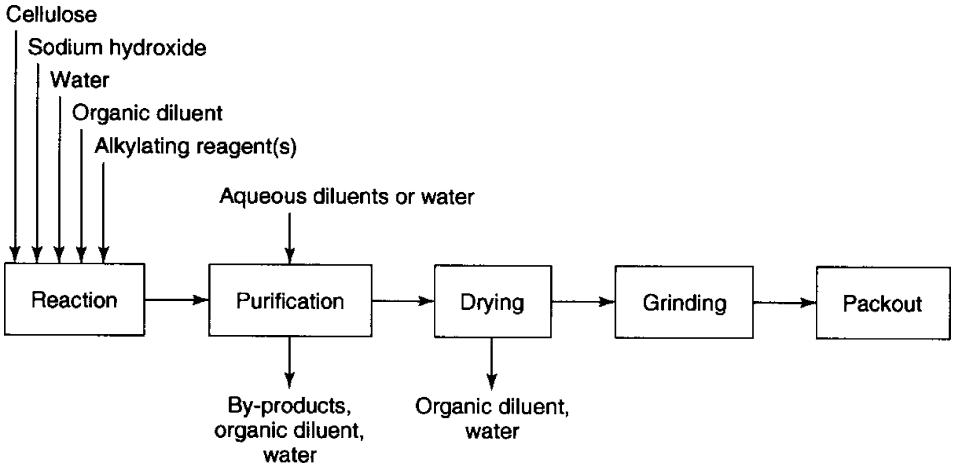


Figure 23
Process flow diagram for the manufacture of cellulose ethers. (Adapted from Majewicz et al.,95 2003)

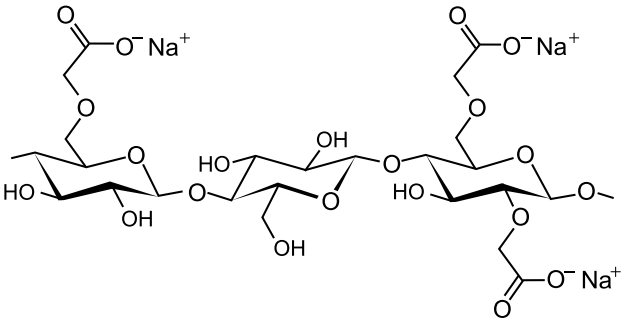
when heated. When derivatized with substituents that are capable of secondary substitution (i. e., hydroxyalkyl ethers and related derivatives), ethers may be formed in which the number of substituent moles/AGU exceeds the DS. This is expressed as number of derivatizing moles / AGU (i. e., *molar substitution* MS) in distinction of degree of substitution, DS. If MS exceeds DS, side chains or branches are formed on the cellulose backbone. Such branches have a significant influence on solubility and rheology.

The production of cellulose ethers proceeds in two major steps, *activation* by pretreatment with alkali followed by *reaction*.

The *activation* treatment involves steeping cellulose (cotton or pulp fibers in sheet form) with cold 15 to 20% aqueous alkali for several hours. This *alkali cellulose* is subsequently reacted with reagents that consume alkali (Williamson synthesis); or alternatively in reactions in which alkali serves as catalyst (alkoxylation). Reactions are carried out between room temperature and 110 °C. Following reaction, cellulose ethers in solid fiber form (by virtue of being suspended in a small polar solvent such as isopropanol) are washed with aqueous alcohol, dried, and powdered by granulation (◉ *Fig. 23*).

Specific products include the following:

Carboxylmethyl cellulose (◉ *Scheme 1*) is produced in many grades by many manufacturers owing to its various end use applications as well as simple reaction protocol. CMC is used in detergents, paper coatings, drilling mud additives in oil well drilling, cosmetics and pharmaceuticals (as fat free ointments for example) and in food products (soups, ketchups, etc.).



◘ Scheme 1
Structural scheme of carboxymethyl cellulose

Methyl cellulose and *hydroxyalky methyl cellulose* manufacture involves in part Williamson synthesis with methyl chloride and alkoxylation. The various ethers are used primarily in building materials and in industrial applications (◉ *Table 13*). Higher grades with greater purities consist of modified vegetable gums, and they find uses in controlled release applications in pharmaceuticals, and in food products and cosmetics where they serve as emulsifiers and texture agents.

Hydroxyethyl cellulose (◉ *Fig. 24*) is a nonflocculating, nonionic cellulose ether that has side chain extensions expressed in MS/DS-ratios of about 1.5 to 3.5. HECs are used in latex paints owing to their excellent interaction ability with solids.

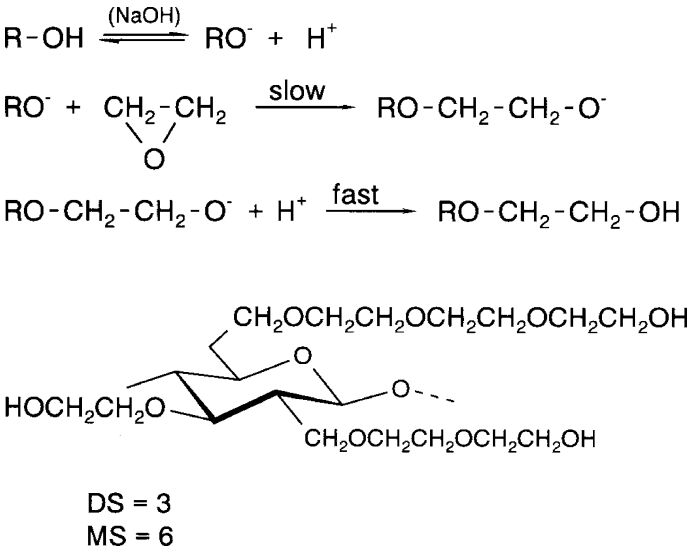
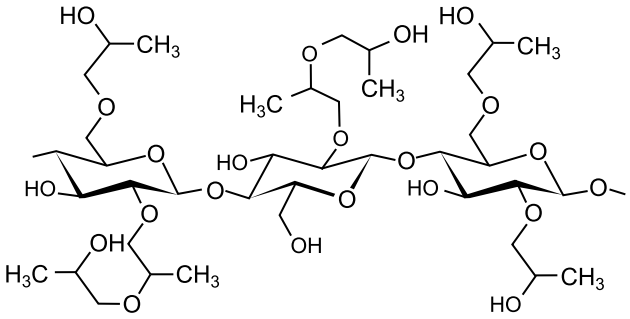


Figure 24
Reaction and chemical structure of hydroxyethyl cellulose with a (hypothetical) DS of 3 and an MS of 6 (for illustration of DS and MS, only). Commercial HEC products have an MS between 1.8 and 3.5. (From Klemm et al., [18])



Scheme 2
Structural scheme of hydroxypropyl cellulose

Hydroxypropyl cellulose (Scheme 2) may either be swellable in water or may dissolve, and it is thermoplastic. HPC is used in the pharmaceutical industry and in cosmetics. It coagulates at temperatures above around 40 to 50 °C. It forms liquid crystalline solutions.

7 Conclusions

Plant polysaccharides are highly energy-efficient polymers serving nature in three vital capacities: energy storage, structural support, and water management. Plants control functional char-

acteristics via molecular structures. Whereas cellulose is a highly regular, organized molecule giving rise to a variety of crystalline formations, heteropolysaccharides ("hemicelluloses") gain their energy-release and gel-formation (i. e., rheological) function by careful control of irregularity. The abundant availability of cellulose has invited industrial efforts to modify functions via chemical derivatization. Cellulose derivatives are available that are thermoplastically processible as well as water-soluble. Numerous chemical modification strategies are available for tailoring mechanical, thermal, and solubility characteristics to desired end-uses.

References

1. Buchanan B, Gruissem W, Jones R (eds) (2000) Biochemistry and Molecular Biology of Plants. American Society of Plant Physiologists
2. Robyt JF (1998) Essentials of Carbohydrate Chemistry. Springer, Berlin Heidelberg New York
3. Paul Finch (ed) (1999) Carbohydrates-Structure, Synthesis, Dynamics. Kluwer Academic Publishers, Dordrecht, 334 pp
4. Rees DA (1977) Polysaccharide Shapes. Chapman and Hall, London
5. Collins P, Ferrier R (1995) Monosaccharides – Chemistry and Role in Natural Products. Wiley, New York
6. Stick RV (2001) Carbohydrates – The Sweet Molecules of Life. Academic Press, New York
7. Stoddart JF (1971) Stereochemistry of Carbohydrates. Wiley – Interscience, New York
8. Stephen Hanessian (ed) (1997) Preparative Carbohydrate Chemistry. Marcel Dekker, New York
9. Tombs MP, Harding SE (1998) An Introduction to Polysaccharide Biotechnology. Taylor and Frances, London
10. MacGregor EA, Greenwood CT (1980) Polymers in Nature. Wiley, Chichester
11. Kaplan DL (ed) (1998) Biopolymers from Renewable Resources. Springer, Berlin Heidelberg New York, 420 pp
12. French AD, Bertoniere N, Brown RM, Chanzy H, Gray D, Hattori K, Glasser W (2003) Cellulose. Enc Polym Sci Tech, vol 5, p 477
13. Kraessig H, Schurz J, Steadman RG, Schliefer K, Albrecht W, Mohring M, Schlosser H (2004) Cellulose. Ullmans Enc Ind Chem
14. Dumitriu S (ed) (2005) Polysaccharides – Structural diversity and functional versatility, 2nd edn. Marcel Dekker, New York, 1204 pp
15. Fengel D, Wegener G (1984) Wood – Chemistry, ultrastructure, reactions. Walter de Gruyter, Berlin, 613 pp
16. Sjostrom E (1993) Wood chemistry – Fundamentals and applications, 2nd edn. Academic Press, New York, pp 293
17. Hon NS, Shiraishi N (eds) (2001) Wood and cellosic chemistry, 2nd edn. Marcel Dekker, New York, pp 914
18. Klemm D, Philipp B, Heinze T, Heinze U, Wagenknecht W (1998) Comprehensive Cellulose Chemistry, vol. 1 and 2. Wiley -VCH, Weinheim, pp 260 and 389
19. Heinze T (edn) (2005) Polysaccharides I – Structure, characterization and use. Springer, Berlin Heidelberg New York, 281 pp
20. Nevell TP, Zeronian SH (eds) (1987) Cellulose chemistry and its applications. Ellis Horwood, 552 pp
21. Whistler RL, Bemiller JN (eds) (1993) Industrial gums – polysaccharides and their derivatives, 3rd edn. Academic Press, San Diego, CA, 807 pp
22. Rinaudo M, Auzely R, Mazeau K (2003) Polysaccharides. Enc Polymer Sci Technol, vol 11, pp 200–261
23. Rinaudo M (2003) Polysaccharides. Kirk Othmer Enc Chem Technol, pp 1–42
24. Kamide K (2005) Cellulose and cellulose derivatives. Elseviar B.V., Amsterdam, pp 630
25. Gilbert RD (ed) (1994) Cellulosic polymers, blends and composites. Hanser Publishers, Cincinatti, OH, pp 244
26. Hiemenz PC (1984) Polymer Chemistry. Marcel Dekker, New York, 738 pp
27. Stephens RS, Westland JA, Neogi AN US Patent #4,960,763
28. Kent RA, Stephens RS, JA Westland (1991) Food Technol 45:108
29. Whiteman A, Brown C, Bull G (1999) Forest product market developments. FAO Corporate Document Repository

30. Marx-Figini M, Schulz GV (1966) Biochem Biophys Acta 112:1
31. Marx-Figini M (1983) J Appl Polym Sci: Appl Polym Symp 37:157–164
32. Marx-Figini M (1969) J Polym Sci Part C 23:57
33. Rinaudo M (2005) In: Dumitriu S (ed) Polysaccharides – Structural diversity and functional versatility, 2nd edn. Marcel Dekker, New York, pp 237–251
34. Kajiwara K, Miyamoto T (2005) In: Dumitriu S (ed) Polysaccharides – Structural diversity and functional versatility. Marcel Dekker, New York, pp 1–40
35. Schurz J (1977) Cell Chem Technol 11:30–28
36. Yui T, Ogawa K (2005) In: Dumitriu S (ed) Polysaccharides – Structural diversity and functional versatility, 2nd edn. Marcel Dekker, New York, pp 99–122
37. Meyer KH, Misch L (1937) Helv Chim Acta 20:232
38. Gardner KH, Blackwell J (1974) Biopolymers 13:1975
39. Horii F, Wada M (2005) Cellulose 12(5), 479–484.
40. Horii F (2001) In: Hon NS, Shiraishi N (eds) Wood and cellulosic chemistry, 2nd ed. Marcel Dekker, New York, pp 83–107
41. Brown RM, Saxena I (2004) Molecular Biology, Biosynthesis, and Structure of Cellulose. Special Issue of Cellulose 11(3/4):273–485
42. Perez S, Mazeau K (2005) Conformations, structures and morphologies of celluloses. In: Dumitriu S (ed) Polysaccharides – Structural diversity and functional versatility, 2nd edn. Marcel Dekker, New York, pp 41–68
43. Schroeter J, Felix F (2005) Cellulose 12:159–165
44. Johnson DC (1987) In: Nevell TP, Zeronian SH (eds) Cellulose chemistry and its applications. Ellis Horwood, pp 181–201
45. Heinze TJ, Glasser WG (eds) (1998) Cellulose derivatives – Modification, characterization, and nanostructures. ACS Symp Ser #688, 361 pp
46. Liebert T, Heinze T (1998) In: Heinze TJ, Glasser WG (eds) Cellulose derivatives – Modification, characterization, and nanostructures. ACS Symp Ser #688, pp 61–72
47. Burchardt W, Habermann N, Kluefers P, Seger B, Wilhelm U (1994) Angew Chem 106:936–939
48. Heinze T, Liebert T, Koschella A (2006) Esterification of polysaccharides. Springer, Berlin Heidelberg New York, 232 pp
49. Oksman K, JF Selin (2004) In: Wallenberger FT, Weston N (eds) Natural Fibers, Plastics and Composites. Kluwer Academic Publishers, Dordrecht, pp 149–165
50. Ganster J, Fink HP, Pinnow M (2006) Composites A37:1796–1804
51. Seavey K, Ghosh I, Davis R, Glasser WG (2001) Cellulose 8:149–159
52. Kadla JF, Dai Q (2003) Kirk-Othmer Encyc Chem Technol, vol 21. Wiley, New York, pp 1–47
53. Sturcova A, Davies GR, Eichhorn SJ (2005) Biomacromolecules 6:1055
54. Sakurada I, Nukushina Y, Ito T (1962) J Polym Sci 57:651
55. Eichhorn S (2006) J ACS Symp Ser 938:63–77
56. Timell TE (1965) Adv Carboh Chem 20:410–482
57. Thompson NS 2000 Hemicelluloses. Kirk Othmer Enc Chem Technol
58. Whistler RL (1993) In: Whistler RL, Bemiller JN (eds) Industrial gums – polysaccharides and their derivatives, 3rd edn. Academic Press, San Diego, pp 295–308
59. Gatenholm P, Tenkanen M (eds) (2004) Hemicelluloses: Science and technology. ACS Symp Ser 864:388
60. Tenkanen M (2004) In: Gatenholm P, Tenkanen M (eds) Hemicelluloses: Science and technology. ACS Symp Ser 864:292–311
61. Ebringerova A, Hromadkova Z, Heinze Th (2005) In: Heinze T (ed) Polysaccharides I – Structure, Characterization and Use. Springer, Berlin Heidelberg New York, pp 4–67
62. Glasser WG, Kaar WE, Jain RK, Sealey JE (2000) Cellulose 7:299–317
63. Marchessault RH (2004) Isolation and properties of xylan: Rediscovery and renewable resource. ACS Symp Ser 864:158–166
64. Soderquist-Lindblad M, Albertsson AC (2005) In: Dumitriu S (ed) Polysaccharides – Structural diversity and functional versatility, 2nd edn. Marcel Dekker, New York, pp 491–508
65. Nishinari K (ed) (2000) Hydrocolloids, Pt. 1. Elsevier, Amsterdam, 467 pp
66. Battista OA, Smith PA (1961) US Patent #2,978,446
67. Battista OA (1975) Microcrystal polymer science. McGraw-Hill, New York
68. Battista OA (1985) Cellulose, microcrystalline. In: Enc Polymer Sci Eng, vol 3, pp 86–90
69. Oksman K, Sain M (2006) Cellulose nanocomposites – processing, characterization and properties. ACS Symp Ser 938:256

70. Bondeson D, Kvien I, Oksman K (2006) Strategies for preparation of cellulose whiskers from MCC as reinforcement in nanocomposites. ACS Symp Ser 938:10–25
71. Samir ASA, Alloin F, Dufresne A (2005) Biomacromolecules 6:612–626
72. Mathew AP, Oksman K, Sain M (2005) J Appl Polym Sci 97:2014–2025
73. Heinze T, Glasser WG (1998). In: Heinze TJ, Glasser WG (eds) Cellulose derivatives – Modification, characterization, and nanostructures. ACS Symp Ser 688:2–18
74. Woodings C (2003) Cellulose fibers, regenerated. Enc Polym Sci Tech, vol 5, pp 532–569
75. Schuster KC, Rohrer C, Eichinger D, Schmidtbauer J, Aldred P, Firgo H (2004) In: Wallenberger FT, Weston N (eds) Natural fibers, plastics and composites. Kluwer Academic Publishers, Dordrecht, pp 123–146
76. Anonymous (2006) Cellulose melt blown nonwoven technology ready for commercialization. Technical Textiles 3/2006, E 129
77. Szczesniak AS, Farkas EH (1962) J Food Sci 27:381
78. El-Nokaly MA, Soini HA (eds) (1999) Polysaccharide applications – cosmetics and pharmaceuticals. ACS Symp Ser 737:347
79. Edgar K (ed) (2006) Polysaccharides in drug delivery. Special Issue of Cellulose 14(1):1–83
80. Shelton MC (2003) Cellulose Esters, Inorganic. Enc Polym Sci Tech, vol 9. Wiley, New York, pp 113–129
81. Edgar K (2003) Organic cellulose esters. Enc Polym Sci Tech, vol 9, p 129
82. Heinze T, Liebert T, Koschella A (2006) Esterification of polysaccharides. Springer, Berlin Heidelberg New York, 232 pp
83. Rustemeyer P (ed) (2004) Cellulose acetates – properties and applications. Wiley-VCH, Weinheim
84. Edgar KJ, Buchanan CM, Debenham JS, Rundquist PA, Seiler BD, Shelton MC, Tindall D (2001) Advances in cellulose ester performance and application. Progr Polym Sci 26:1605–1688
85. Balser K, Hoppe C, Eicher T, Wandel M, Astheimer HJ, Steinmaier H, Allen JM (2004) Cellulose Esters. Ullmanns Enc Ind Chem
86. Marchessault RH, Ravenelle F, Zhu XX (eds) (2006) Polysaccharides for drug delivery and pharmaceutical applications. ACS Symp Ser 934:365
87. Gedon S, Fengi R (2000) Cellulose esters-organic, Kirk-othmar End.Chem.Technol.
88. Edgar K, Pecorini TJ, Glasser WG (1998) In: Heinze TJ, Glasser WG (eds) Cellulose derivatives. ACS Symp Ser 688:38–60
89. Vaca-Garcia C, Gozzelino G, Glasser WG, Borredon ME (2003) J Polym Sci, Pt. B: Polym Phys 41:281
90. Malm CJ, Mench JW, Kendall DL, Hiatt GD (1951) Ind Eng Chem 43:684 and ibid 688
91. Morooka T, Norimoto M, Yamada T, Shiraishi N (1983) Wood Res 69:61
92. Morooka T, Norimoto M, Yamada T (1984) J Appl Polym Sci 29:3981
93. Glasser WG, Samaranayake G, Dumay M, Dave V (1995) J Polym Sci, Part B: Polym Phys 33:2045
94. Sealey JE, Samaranayake G, Todd JG, Glasser WG (1996) J Polym Sci, Part B: Polym Phys 34:1613
95. Majewicz TG, Erazo-Majewicz PE, Podlas TJ (2003) Cellulose ethers. Enc Polym Sci Tech, vol 5, pp 507–532

6.4 Gums and Related Polysaccharides

James N. BeMiller
Department of Food Science, Whistler Center for Carbohydrate Research, Purdue University, West Lafayette, IN 47909–2009, USA
bemiller@purdue.edu

Abstract

In the context of carbohydrates, gums are usually considered to be non-starch, water-soluble polysaccharides with commercial importance. When used as ingredients in processed foods, they may be called hydrocolloids. Gums are used because of the functionalities they impart to whatever system or product into which they are incorporated. As with other polymers, their chemical structures, together with the nature of the aqueous environment surrounding the molecules (pH, types and concentrations of salts or other solutes, temperature, shear, etc.) determines the shapes of the molecules; the chemical nature and shapes of the molecules determines the gum's physiochemical properties, and their physicochemical properties determines their functionalities. All gums have one similar property, i. e., the ability to thicken water and aqueous systems, but they may impart different rheological properties to the systems they thicken. Certain gums provide certain functionalities better than do other gums.

In: *Glycoscience*. Fraser-Reid B, Tatsuta K, Thiem J (eds)
Chapter-DOI 10-1007/978-3-540-30429-6_37:

Keywords

Gums; Hydrocolloids

Abbreviations

CMC	carboxymethylcellulose
Gal	D-galactosyl units
β**Glc*p***	β-D-glucopyranosyl units
β**Glc*p*A**	β-D-glucuronopyranosyl units
α**LGul*p*A**	α-L-guluronopyranosyl units
HEC	hydroxyethylcellulose
HM	high-methoxyl
HPMC	hydroxypropylmethylcellulose
LBG	locust bean (carob) gum
LM	low-methoxyl
Man	D-mannosyl units
β**Man*p***	β-D-mannopyranosyl units
β**Man*p*A**	β-D-mannuronopyranosyl units
MC	methylcellulose
MW	molecular weight
NSP	non-starch polysaccharides
PES	processed Euchema seaweed
PGA	propylene glycol alginate
PNG	Philippine natural grade
α**LRha*p***	α-L-rhamnopyranosyl units

1 Introduction

There is no chemical category of gums (or mucilages, a term once used to describe substances that produce slimy or tacky aqueous dispersions, which are properties manifested by many types of water-soluble plant, animal, and microbial polysaccharides, including starches and modified starches). The term gum is an umbrella term common to several different fields with different definitions in each field. As applied to water-soluble substances, it refers to water-soluble, non-starch polysaccharides (NSP), both those that are natural and those that are made from natural NSP by structural modification. In general, they have the ability to produce highly viscous aqueous solutions at low concentrations, but there are exceptions to this principal. Gum arabic and "low viscosity grades" of certain gums are the most notable exceptions; they require rather high concentrations to produce highly viscous aqueous solutions. The starch polysaccharides (amylose and amylopectin), although they and their derivatives are water-soluble, are not considered to be gums, but rather starches are put in a separate category. Those water-soluble polysaccharides (other than those from starches) that are used industrially are often called industrial gums. A few years ago, some in the food industry developed an aversion to referring to food ingredients as gums and started calling food gums hydrocolloids, even though many of them have gum in their name, so names such as guar gum appear on the ingredient

labels of food products. The term hydrocolloid is appropriate because many gums, at least those of higher molecular weights, do not form true solutions. Rather, they form dispersions of hydrated polymer molecules whose particle sizes are in the colloid range. The hydrocolloid category may also include proteins used as food ingredients, especially gelatin, which like the carbohydrate polymer hydrocolloids, is used as a stabilizer in foods. The terms gum and hydrocolloid generally imply commercial importance, as there is nothing else that distinguishes them from many other water-soluble polysaccharides; but again, this is not always true. There are water-soluble polysaccharides that have gum in their name, e. g., corn fiber gum, but are not produced and sold as commodities. There are others that are not produced as isolated substances but are useful as part of the natural matrix in which they occur; examples are the β-glucan known as oat gum and psyllium seed gum. Still, others, such as the arabinoxylans of wheat flour, have a considerable impact on the system (dough) in which the flour is used, but the material (in this case the flour) is used because it contains starch and gluten, not because it contains water-soluble NSP. Yet others, like the aloe polysaccharides, while commercially useful, are probably not thought of as gums. This chapter presents a brief synopsis of those gums/hydrocolloids used in food and/or non-food applications. Additional information can be found elsewhere. An elementary presentation can be found in reference [1]. More in depth presentations are available in references [2,3,4,5,6,7,8,9,10,11,12,13].

2 Classification

It is usually helpful for discussions of the properties of substances, to classify them. However, there is no especially useful way to classify gums [14]. They are usually classified according to source, but that is not useful, even as an indication of relative cost, except that those from bacterial sources, i. e., produced by fermentation processes, are more expensive on a per unit weight basis. But even that can be misleading because it is not the cost per unit weight, but the cost of the amount required to give the desired functionality that should be compared; for at least in some cases, gums can be interchanged. Gum sources are algae, higher plants, microorgan-

Table 1
Sources of common gums

Source		Gums
Marine algae		Agars, alginates, carrageenans, furcellaran
Higher plants	Extracts	Larch arabinogalactan, pectins
	Seeds	Guar gum, locust bean (carob) gum, tara gum
	Tubers and roots	Inulin, konjac glucomannans
	Exudates	Gum arabics
Microorganisms		Curdlan, dextrans, gellans, xanthans
Cellulosics (by chemical modification)		Carboxymethylhydroxyethylcelluloses, cetylhydroxyethylcelluloses, ethylhydroxyethylcelluloses, hydroxyethylcelluloses, hydroxyethylmethylcelluloses, hydroxypropylcelluloses, hydroxypropylmethylcelluloses, methylcelluloses

isms, and derivatives of cellulose (◉ *Table 1*). (There is one synthetic water-soluble polymer, polydextrose, made by reacting D-glucose and sorbitol in the presence of citric acid, that has the properties, of a low-viscosity gum, but which is not usually considered to be a gum.) There have been attempts to classify food gums as to whether they are used as thickeners and stabilizers or to form gels, but there is overlap in these properties. In fact, no classification system is effective because of the multiple properties of most gums. It would probably be most desirable to classify gums based on their structures (◉ *Table 2*). But here again, there are difficulties with a classification system based on structure for two reasons: (a) in one sense, most gums have similar structures; (b) in another sense, in most cases the structure of a given gum is sufficiently unique that it needs to be placed in a category by itself. Most of the commercially important hydrocolloids/industrial gums are linear polymer molecules with short side groups that either occur naturally or are introduced by chemical derivatization. When they occur naturally, they may be either saccharide units or functional groups (most often an organic or inorganic ester group). Ester and ether derivatives can be made chemically. Whether the side group is a saccharide unit or an ester or ether group, such groups are called "side chains", even though the longest is a trisaccharide unit. Exceptions to this general structure of linear polymers with short side chains are gum arabic, which has a highly branched, branch-on-branch structure, and curdlan, which is an unsubstituted linear molecule. Gellan and konjac mannan may be deacylated to give mostly unsubstituted linear molecules. Most, if not all, linear polysaccharides exist in some sort of helical shape.

All polysaccharide gum preparations are polydisperse, i. e., they contain populations of molecules with a range of molecular weights. All algal and higher plant polysaccharides are polymolecular, i. e., the fine structures of the molecules vary from molecule to molecule. (There are a few exceptions, most notably cellulose in this context. Cellulose molecules are composed of a single type of monomer unit joined by a single type of linkage (◉ Chap. 6.3), so there is no opportunity for variation in chemical structure.) Bacterial polysaccharides/gums (curdlan, gellans, xanthans) have constant repeating-unit structures, but they can have some degree of polymolecularity due to differences in attached ester (gellans, xanthans) and/or cyclical acetal (xanthans) groups. Bacterial polysaccharide preparations can also differ from each other depending on the strain of the organism and the culture conditions (medium, temperature, etc.) used to produce them. Therefore, only statistical structures can be determined and presented for the gums discussed here. Details of structures can be found elsewhere. This chapter emphasizes the characteristics of the principal polysaccharides called gums that place them in that category of substances, but the presentation is cursory.

◘ Table 2
General structures of some common gums

Overall structure	Gums
Linear, unbranched molecules	Alginates, carrageenans, curdlan, inulin, pectins
Linear molecules with short branches (side chains)	Gellans, guar gum, konjac, glucomannans, larch arabinogalactan, locust bean (carob) gum, tara gum, xanthans, all cellulosics
Branch-on-branch (bush-like) molecules	Gum arabics

3 General Properties

This section outlines the general properties of gums. Each gum has unique properties; those will be discussed under the discussion of that particular gum.

Gums are available as dry powders. Because polysaccharides possess a strong affinity for water and hydrate readily, gum particles take up water, swell, and usually undergo partial or complete dissolution when they are added to an aqueous system. As alluded to, a common property of gums is their ability to gel and/or thicken aqueous systems, but the rheological properties of the gels and/or viscous solutions made with gums varies with the gum used, its viscosity grade, its concentration, other components of the system, and temperature. The systems into which they are incorporated become viscoelastic. Another general property of gums, therefore, is their ability to change the flow properties of liquid products and the deformation properties of semisolid products, even though they are usually used in low concentrations (<0.5%), indicating their great ability to produce viscosity and/or to form gels [15,16,17,18,19,20,21]. The increase in viscosity with an increase in concentration of a gum is not linear, except at maybe very low concentrations. Rather viscosity increases geometrically with increases in gum concentration.

A common property of gums that form gels is their ability to form junction zones, either by themselves or in combination with another water-soluble polymer (polysaccharide or protein). A junction zone is an association of two or more polymer molecules over a portion of their lengths. The molecules can be held together in a junction zone by hydrogen bonding, ionic cross-bridges, or hydrophobic associations. When gum molecules form intermolecular associations with other polymer molecules over more than one portion of their lengths, a three-dimensional network entrapping water and whatever is dissolved in it, i. e., a gel, is formed.

In order to thicken or gel aqueous systems, gums must be soluble, which is why that is part of their definition. They are water-soluble because they are polyhydroxy substances. As a result, another common property is their ability to bind and hold water (to different degrees). Yet another common property, and a property of all polysaccharides, is their susceptibility to acid- and enzyme-catalyzed hydrolysis (to varying degrees). The extent of hydrolysis of their glycosidic linkages, i. e., the extent of depolymerization, in an acidic environment is a function of the structure of the polysaccharide (its monomer units and the types of linkages), acid concentration (pH), temperature, and time at that temperature and pH. Gums, like other polysaccharides, are also subject to enzyme-catalyzed hydrolysis. This process is controlled by the specificity of the enzyme, pH, temperature, time, and the presence of any required co-factors, such as metal ions. Required enzymes usually come from bacterial or fungal sources, but some may be of plant origin. They are generally specific for a gum or, in a few cases, the polysaccharides in a family of closely related gums such as the galactomannans.

Not to be forgotten aspects of gums are the facts that they are abundant and obtained from a variety of replenishable sources. They are obtained from marine algae (seaweeds), higher plants (as extracts and exudates), microorganisms (via fermentation), and from cellulose (from forest products) (❯ *Table 1*). They are amenable to chemical and biochemical modification via reactions of their hydroxyl groups and hydrolysis of their glycosidic linkages. They have a variety of structures (❯ *Table 2*). They are non-toxic and biodegradable. As mentioned, the terms gum and hydrocolloid usually, but not always, imply commercial importance. Some applications of gums are given in ❯ *Table 3*.

Table 3
Some applications of gums

Product category in which gums are used	Examples
Construction Materials	In tape joint compounds; in gypsum spray plaster; in ceramic tile adhesives, grouts, and mortars; in cement mortars, stuccos, and plasters; in concrete; as wallpaper adhesive; as a ceramic processing aid; as a thickener for latex paint
Consumer products	As absorbents, as binders, in detergents to encapsulate
Food products	As thickeners, to provide specific rheology, to form gels, to inhibit syneresis, to stabilize emulsions, to stabilize suspensions, to emulsify, to bind/hold water, to form films and coatings, to inhibit crystallization, to improve texture/mouthfeel, as processing aids, as whipping agents, to stabilize foams, as extrusion aids, as binders, to stabilize proteins, to encapsulate, to flocculate
Health care, medical, and pharmaceutical applications	In controlled-release products; as tablet and capsule binders, tablet and capsule excipients, wound healing materials, and blood volume expanders; and for constipation and diarrhea treatment
Mining operations	As flotation agents and suppressants
Personal care and cosmetic products	In hand lotions (as emollients/demulcents), in toothpaste, in cosmetics, in shampoo and hair conditioning formulations
Petroleum production	In drilling muds, cementing fluids, hydraulic fracturing fluids, workover fluids, completion fluids, pipeline cleaning fluids, and enhanced recovery fluids
Textiles	In printing pastes, as a warp size
Other	In suspension polymerization, as binders for fertilizers and pesticides, in tire sealants, in de-icing fluids

4 Specific Gums

4.1 Xanthans

Xanthan [22,23,24,25,26,27,28], known commercially as xanthan gum, is a high-molecular-weight bacterial polysaccharide produced in fermentation vats by the microorganism *Xanthomonas campestris*. It is a non-gelling polysaccharide that imparts unique rheological properties to its solutions. It is a linear polymer with a backbone structure identical to that of cellulose (*Sect. 4.3*). On every other D-glucopyranosyl unit of its main-chain backbone is a trisaccharide unit containing, from the reducing end outward, an α-D-mannopyranosyl 6-acetate unit, a β-D-glucuronopyranosyl unit, and a β-D-mannopyranosyl unit at the non-reducing end. Some of the non-reducing end units are terminated with a 4,6-*O*-pyruvyl cyclic acetal group. Xanthan molecules have the shape of a five-fold helix, rather than the two-fold helical shape of cellulose. They often are in the form of double helices and bundles of double helices. Xanthan undergoes a conformational transition as the properties of the solvent are altered. An ordered helical conformation is stabilized by high ionic strength and/or low temperature. Determined average MW values for xanthan range from 3×10^5 to 7.5×10^6 or greater, an indication of the ability of xanthan molecules to aggregate.

Xanthan molecules behave as rather rigid, linear molecules, and the stiffness of its molecules gives it unique properties. The actual properties of xanthan preparations can vary with the strain of the organism used and the conditions of fermentation, so not all xanthan products are identical. In addition, xanthan preparations are available in the following types: different particle sizes, different viscosity grades, transparent (clear-solution) types (devoid of the usual inactivated bacterial cells), salt/brine-tolerant types, easily dispersible types, rapidly hydrating types, delayed-hydrating types, smooth-flow types (reduced pseudoplasticity), types with greater suspending power, dust-free types (granular mesh), and cellulase-free types.

Xanthan forms the most pseudoplastic (instantaneous, reversible shear thinning) solutions of all the gums. This property is due to the stiffness of its molecules and/or intermolecular associations of two or more molecules. In plots of viscosity vs. concentration, there is a Newtonian (non-pseudoplastic) plateau at very low shear rates, which at least, makes its solutions appear to have a yield value (a yield value being the force required to initiate flow). As a result, xanthan is an excellent stabilizer for suspensions and emulsions.

The viscosities of solutions of xanthan, unlike those of all other water-soluble polymers, do not decrease as the temperature is raised. They are unaffected by temperature between 0 °C and 100 °C. And even though it is an anionic polymer, its solution viscosity is not affected by pH or most salts, i. e., although ionic, it behaves as if it were a neutral gum. Xanthan interacts synergistically with guar gum and locust bean gum (both galactomannans) (*Sect. 4.2*). An increase in solution viscosity results from its interaction with guar gum. Heat-reversible gelation results from its interaction with locust bean gum.

4.2 Galactomannans

Commercial gums that are in the galactomannan family of polysacharrides are guar gum, locust bean (carob) gum, and tara gum [29,30,31,32,33,34]. All are linear polymers with single-unit branches. The backbone chain of all three gums is composed of (1→4)-linked β-D-manopyranosyl units, a structure identical to that of cellulose, with the exception of the configuration of C2 of the monomeric units. A mannan of this type is highly crystalline and even more insoluble than is cellulose, but the branch units on the galactomannans give the gums water solubility. The branch units are α-D-galactopyranosyl units attached to *O*-6 positions of main-chain β-D-mannopyranosyl units. The differences between the different galactomannans are in the percentages and distributions of the branch units. The polysaccharide of guar gum (guaran) has a Gal : Man ratio of ~0.49–0.67 : 1 and a somewhat random distribution of the side units. The galactomannan of locust bean gum (LBG) has fewer branch units (Gal : Man = ~0.19–0.35 : 1) and a clustered arrangement of the branch units. Locust bean gum has stretches of ~80 unsubstituted D-mannopyranosyl units alternating with sections of ~50 units in which most of the main-chain units have an α-D-galactopyranosyl side-chain unit glycosidically attached to their *O*-6 positions. Because of the difference in structures, guar gum and LBG have somewhat different physical properties, even though both are galactomannans and are composed of long, somewhat rigid chains that provide high solution viscosity. (Reported average MW values for guaran are $1–2 \times 10^6$ and for the polysaccharide of LBG $3–4 \times 10^5$.) Because guaran has its galactosyl units fairly evenly placed along the chain, there are few locations on the chains that are suitable for intermolecular associations. However,

molecules of LBG, with their stretches of unsubstituted backbone units, can hydrogen bond with each other. This structural difference makes locust bean gum less soluble than is guar gum. Guar gum is rather soluble in cold water, but heating of dispersions of LBG in water to a temperature of at least 80 °C is required for its complete dissolution.

The unsubstituted backbone stretches of LBG can also hydrogen bond to xanthan (◉ *Sect. 4.1*) and κ-Carrageenan (◉ *Sect. 4.8*) helices, producing a three-dimensional network and forming rigid gels. In the case of xanthan and LBG, neither forms gels by itself, but their combination can result in rigid gels. Rigid gels can also be formed by the LBG-κ-carrageenan combination. (κ-Carrageenan itself is a gel-forming gum.)

Tara gum is used in much less quantities than are the other two galactomannans. Its structure is intermediate between those of the other two galactomannans with a Gal : Man ratio of ~0.33 : 1 (closer to that of LBG than to that of guar gum). Fenugreek galactomannan has a Gal : Man ratio of ~1 : 1. There are also other galactomannans, such as those of species of *Cassia*, *Crotalaria*, *Indigofera*, and *Sesbania* (all of the Leguminosae family), some of which find commercial use, primarily in India.

In no case is one of these galactomannans available and used as a pure polysaccharide. Rather they are the ground endosperm of the seeds of the plant from which they are obtained. As a result, the commercial gums also contain fiber, protein, ash, and other constituents. As already indicated, the polysaccharide present in guar gum is called guaran. Chemical names have not been given to the other galactomannans.

Guar gum is derivatized to modify its properties and functionalities. Available derivatives include the hydroxypropyl ether, the carboxymethyl ether, a cationic ether, a mixed hydroxypropyl cationic ether, the phosphate ester, and a mixed phosphate ester and cationic ether (amphoteric product).

4.3 Derivatives of Cellulose

Cellulose (◉ Chap. 6.1 and ◉ Chap. 6.3) is a high-molecular-weight, linear, insoluble polymer of repeating β-D-glucopyranosyl units in the $^4C_1$ conformation joined by (1→4) glycosidic linkages. The (1→4) linkages joining the β-D-glucopyranosyl units of cellulose molecules are equatorial → equatorial linkages, which gives the molecules a flat, ribbon-like structure in which each glucopyranosyl unit in the chain is turned upside down as compared to the units preceding and following it. This flat, linear structure enables cellulose molecules to associate with each other via hydrogen bonding over extended regions, forming polycrystalline, fibrous bundles. Crystalline regions are separated by amorphous regions. Cellulose is insoluble because, in order for it to dissolve, most of its multitude of hydrogen bonds would have to be released at once. Cellulose can, however, through derivatization, be converted into water-soluble gums. All cellulose derivatives that are gums are linear polymer molecules with chemically introduced ether groups as side units. Commercial cellulose derivatives are only partially etherified.

To make water-soluble derivatives of cellulose, purified wood pulp or cotton linters are treated with a solution of sodium hydroxide (18%) to produce the alkoxy form of cellulose, called alkali cellulose. In alkali cellulose, some of the hydroxyl groups exist as $-O^-$ and some as –OH. The alkali cellulose is nucleophilic and is reacted with a reagent containing a leaving

group to produce an ether linkage [35,36,37,38,39,40,41,42,43,44]. Detailed structural analyses (relative derivatization of O2, O3, and O6 and average distribution of substituent ether groups along the polymer chain) have been carried out for some cellulose derivatives.

4.3.1 Neutral Cellulose Derivatives

The primary neutral, water-soluble cellulose derivatives are hydroxyethylcelluloses (HEC), methylcelluloses (MC), and hydroxypropylmethylcelluloses (HPMC). To introduce hydroxyethyl ether groups (–O–CH_2–CH_2–OH), alkali cellulose is reacted with ethylene oxide. To introduce hydroxypropyl ether groups (–O–CH_2–CH_2OH–CH_3), alkali cellulose is reacted with propylene oxide. To introduce methyl ether groups (–O–CH_3), alkali cellulose is reacted with methyl chloride. Other water-soluble, neutral gums made commercially from cellulose are cetylhydroxyethylcelluloses, ethylhydroxyethylcelluloses, hydroxyethylmethylcelluloses, and hydroxypropylcelluloses. All cellulose derivatives are families of products. Different products within each family can be made by varying the amount of derivatization (with each group if products are doubly substituted), the order of reagent addition if two reagents are used, the regularity of distribution of substituent groups, average molecular weights of products (giving different viscosity grades), and the physical nature of the products. All products in these families of products are water soluble to different degrees. They are soluble because the ether group protrusions along the polymer chains prevent the intermolecular associations characteristic of native cellulose. While a few added ether groups spread along the chains enhance water solubility, they also decrease chain hydration by replacing water-binding hydroxyl groups with less polar ether groups, hydroxyethylcelluloses being the least affected. In products like the methylcelluloses and hydroxypropylmethylcelluloses, the ether groups restrict hydration of the chains to the point that they are on the borderline of water solubility. When aqueous solutions of these products are heated, the water molecules hydrating the polymer dissociate from the chain and hydration is decreased sufficiently that intermolecular associations increase (probably via van der Waals interactions) and gelation occurs. Lowering the temperature of the gel allows the molecules to rehydrate and redissolve, so the thermogelation is reversible. Because of non-polar groups, MC, HPMC, hydroxypropylcelluloses, cetylhydroxyethylcelluloses, ethylhydroxyethylcelluloses, and hydroxyethylmethylcelluloses are somewhat surface active and absorb at air–water and oil–water interfaces.

4.3.2 Anionic Cellulose Derivatives

The only anionic group added to cellulose is the carboxymethyl (–O–CH_2–CO_2^-) ether group. To introduce the carboxymethyl group, alkali cellulose is reacted with chloroacetic acid. (The carboxymethyl ether is also the ether of glycolic acid.) Also prepared commercially is carboxymethylhydroxyethylcellulose.

The ionic charge on the molecules makes them hydrate and dissolve rapidly. Electrostatic repulsion of the carboxylate groups causes the molecules of carboxymethylcellulose (CMC) and its hydroxyethyl ether derivative to repel each other and to be somewhat extended. As a result, solutions of CMC tend to be both highly viscous and stable.

Different products can be made by manipulating the same variables listed for the neutral cellulose derivatives.

4.4 Alginates

Alginates (algins) are polysaccharides extracted from brown marine algae [45,46,47,48,49,50]. They are salts (most often the sodium salt) of poly(uronic acids). As a result, they have a carboxylate group on every monomer unit of their structures. They are composed of two monomer units: β-D-mannuronopyranosyl units (βMan*p*A) and α-L-guluronopyranosyl units (αLGul*p*A), both linked (1→4). The αLGul*p*A : βMan*p*A ratio in alginates from different seaweed sources ranges from ~0.5 : 1.0 to ~2.5 : 1.0.

The monomer units in alginates are arranged so as to give the molecules a block copolymer structure. In an alginate polymer, there are stretches of only βMan*p*A units, stretches of only αLGul*p*A units, and stretches where the two monomer units alternate. Segments containing only βMan*p*A units, because the units are in the $^4C_1$ conformation and joined in (1→4) equatorial-equatorial linkages, are flat and ribbon-like. Segments containing only αLGul*p*A units, because the units are in the $^1C_4$ conformation and joined by (1→4) axial-axial linkages, have a pleated (corrugated) structure. Different percentages of the different block segments cause alginates from different seaweed sources (harvested in different regions of the world) to have different properties.

Solutions of sodium alginates are highly viscous. The calcium salt of alginates is insoluble. Insolubility results from interactions between calcium ions and the poly(αLGulA) regions of molecules. Calcium ions bind together stretches of poly(αLGulA) segments to form what is known as a junction zone. The result is a three-dimensional polymer network and a gel. The strength of the gel depends on the content of poly(αLGulA) regions in the alginate used and the concentration of calcium ions. Alginates with a greater content of poly(αLGulA) regions produce gels of higher strength.

The free carboxylic acid form of the polymer is alginic acid. Reaction of alginic acid with propylene oxide produces the propylene glycol ester. Commercial propylene glycol alginates are partial esters (50–85% of the carboxyl groups esterified). Solutions of propylene glycol alginates (PGA) are much less sensitive to low pH and polyvalent cations, including calcium ions, than are solutions of non-esterified alginates, because the esterified carboxyl groups cannot ionize. Also, the propylene glycol group introduces a protrusion that prevents the close association of chains. Therefore, PGA solutions are stable. The somewhat hydrophobic propylene glycol groups also give the molecule mild interfacial activity, i. e., foaming, emulsifying, and emulsion-stabilizing properties. Propylene glycol alginate is used when stability to acid, non-reactivity with calcium ions, or its surface activity is desired.

4.5 Pectins

Commercial pectins are partially esterified (methyl ester) polymers composed primarily of α-D-galacturonopyranosyl units linked (1→4) [51,52,53,54,55,56,57,58,59]. The native molecules present in the cell walls and intercellular layers of all land plants (Chap. 6.1), from which commercial pectins are obtained, are more complex molecules that are converted into the partially methyl esterified poly(α-D-galactopyranosyluronic acids) during extraction with acid. Commercial pectin is obtained from citrus peel and apple pomace. Generally, the highest quality pectin is obtained from lemon and lime peel.

The compositions and properties of pectins vary with source, the processes used during preparation, and any subsequent treatments. During extraction with mild acid, some hydrolytic depolymerization and hydrolysis of methyl ester groups occurs. The term pectin denotes a family of compounds. The term pectin is usually used in a generic sense to designate those water-soluble poly(galacturonoglycan) preparations of varying methyl ester contents and degrees of neutralization that are capable of forming gels. (Pectins produce little viscosity and, while they find some use as milk protein stabilizers, are not used as thickeners.) In all natural pectins, some of the carboxyl groups are in the methyl ester form. Depending on the manufacturing conditions, the remaining free carboxylic acid groups may be partly or fully neutralized, i. e., partly or fully present as sodium, potassium, or ammonium carboxylate groups. Typically, they are present in the sodium salt form.

By definition, preparations in which more than 50% of the carboxyl groups are in the methyl ester form ($-COOCH_3$) are classified as high-methoxyl (HM) pectins; the remainder of the carboxyl groups will be present as a mixture of free acid (–COOH) and salt (e. g., $-COO^-Na^+$) forms. Preparations in which less than 50% of the carboxyl groups are in the methyl ester form are called low-methoxyl (LM) pectins. Treatment of a pectin preparation with ammonia (often dissolved in methanol) converts some of the methyl ester groups into carboxamide ($-CONH_2$) groups (15–25%). In the process, a low-methoxyl (LM) pectin (by definition) is formed. These products are known as amidated LM pectins.

The principal and key feature of all pectin molecules is a linear chain of (1→4)-linked α-D-galactopyranosyluronic acid units. Neutral sugars, primarily L-rhamnosyl units, are also present. In citrus and apple pectins, the α-L-rhamnopyranosyl units are inserted into the polysaccharide chain at rather regular intervals. The inserted rhamnopyranosyl units may provide the necessary irregularities in the structure required to limit the size of the junction zones and effect gelation (as opposed to precipitation/complete insolubility). At least some pectins contain covalently attached, highly branched arabinogalactan chains and/or short side chains composed of D-xylosyl units. The presence of side chains may also be a factor that limits the extent of chain association. Pectins from some sources, such as sunflower heads, have acetate ester groups at O2 of the main-chain monomer units. These acetyl units hinder intermolecular associations and gel formation.

Junction zones are, therefore, formed between regular, unsubstituted pectin chains when the negative charges on the carboxylate groups are removed (addition of acid), hydration of the molecules is reduced (by addition of a cosolute, almost always sugar, to a solution of HM-pectin), and/or when polymer chains are bridged by calcium ions.

HM pectin solutions gel when sufficient acid and sugar is present. As the pH of a pectin solution is lowered, the highly hydrated and charged carboxylate groups are converted into uncharged, only slightly hydrated, carboxylic acid groups. As a result of losing some of their charge and hydration, the polymer molecules can now associate over a portion of their length, forming junction zones and a network of polymer chains that entraps the aqueous solution of solute molecules. Junction zone formation is assisted by the presence of a high concentration (~65%, at least 55%) of soluble solids (most often sugar), which competes with the pectin molecules for the water molecules and reduces hydration of the chains, allowing them to interact with one another.

LM pectin solutions gel only in the presence of divalent cations (only calcium ion is used in food applications, and almost all pectin is used as a food ingredient) which crosslink the poly-

mer chains. Increasing the concentration of divalent cations increases the gelling temperature and gel strength.

4.6 Gellans

Gellan [26,28,60,61,62,63], known commercially as gellan gum, is an extracellular, anionic polysaccharide produced by the bacterium *Sphingomonas elodea*, making it a fermentation gum. The gellan molecule is linear and is composed of β-D-glucopyranosyl, β-D-glucuronopyranosyl, and α-L-rhamnopyranosyl units in a molar ratio of 2:1:1. Native gellan (also called high-acyl gellan) contains two ester groups, an acetyl group at O6 and a glyceryl group at O2, on one of the two (the same) β-D-glucuronopyranosyl units. On average, there is one glycerate ester group per tetrasaccharide repeat unit and one acetate ester group for every two repeat units.

Some gellan is de-esterified by treatment with alkali. Removal of the acyl groups has a dramatic effect on the properties of the gel produced. The de-esterified form is known as low-acyl gellan. Its tetrasaccharide repeat unit structure is $\rightarrow$4)-αLRha*p*-(1$\rightarrow$3)- βGlc*p*-(1$\rightarrow$4)-βGlc*p*A-(1$\rightarrow$4)- βGlc*p*-(1$\rightarrow$, so one out of four monomer units is a uronic acid unit. Both high- and low-acyl gellans have average molecular weights of $1–2 \times 10^6$.

Three basic forms of gum are available: high-acyl (native), low-acyl clarified, and low-acyl unclarified. Blending of high- and low-acyl types results in products with intermediate properties.

Gellan can gel aqueous systems. It forms gels with both monovalent and divalent cations, divalent cations (Ca^{2+}) being about 10 times more effective. Gels can be formed with as little as 0.05% gum (99.95% water). Gelation is often effected by cooling a hot solution containing the required cation. Shearing during cooling of a hot gellan solution prevents the normal gelation mechanism from occurring and produces a smooth, homogeneous, thixotropic fluid (a pourable gel) that stabilizes emulsions and suspensions very effectively because of its high yield value. Gentle agitation of a weak gellan gel will also disrupt the gel structure and turn the gel into a smooth, pourable, thixotropic fluid with excellent emulsion and suspension stabilizing properties.

The low-acyl types of gellan form firm, brittle, non-elastic gels (with textures similar to those of gels made with agar and κ-carrageenan) (➤ *Sect. 4.8*). The high-acyl (native) type forms soft, elastic, non-brittle gels (with textures similar to those made with mixtures of xanthan and locust bean gum) (➤ *Sect. 4.1* and ➤ *Sect. 4.2*). A range of intermediate gel textures can be achieved by mixing the two basic types of gellan.

4.7 Gum Arabic

Gum arabic, also known as gum acacia [64,65,66], is another gum that is ionic because of the presence of uronic acid units. The structure of the most important polysaccharide of gum arabic is that of a highly branched, branch-on-branch (bushlike), anionic polymer that is attached to a polypeptide. The structure is not only complex, but variable. What is harvested as, and called, gum arabic is a heterogeneous material, but generally consists of two primary fractions. One, which accounts for about 70% of the gum, is composed of polysaccharide molecules with

little or no protein. The other fraction contains molecules of higher molecular weight that have protein as an integral part of their structures and is the functional component of gum arabic. The protein-polysaccharide fraction is itself heterogeneous with respect to protein content. The polysaccharide structures are covalently attached to the protein component via linkage to hydroxyproline and, perhaps, serine units, the two predominant amino acids in the polypeptide. The overall protein content is about 2 wt%, but fractions may contain as much as 25% protein. Some glycoprotein is also present.

Detailed analyses have been carried out on a number of gum arabic preparations. These analyses include composition, i. e., ratios of the various fractions, and structural analysis of the polysaccharide component. Both those polysaccharide components attached to protein molecules and those that are not are acidic arabinogalactans with the following approximate composition: D-galactose, 44%; L-arabinose, 24%; D-glucuronic acid, 14.5%; L-rhamnose, 13%; 4-*O*-methyl-D-glucuronic acid, 1.5%. They contain chains of (1→3)-linked β-D-galactopyranosyl units having two- to four-unit side chains consisting of (1→3)-linked β-D-galactopyranosyl units joined by (1→6)-linkages. Both the main chain and the numerous side chains have attached α-L-arabinofuranosyl, α-L-rhamnopyranosyl, β-D-glucuronopyranosyl, and 4-*O*-methyl-β-D-glucuronopyranosyl units. The uronic units occur most often as non-reducing end-units.

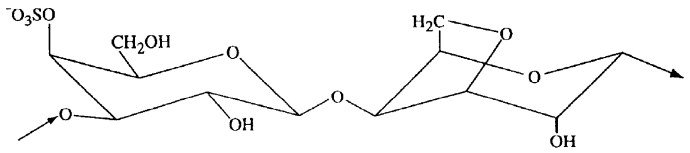
The unique structure of gum arabic, i. e., the anionic, highly branched, very compact structure of its polysaccharide component and its multiple attachments to a hydrophobic polypeptide, give it unique properties. Among its unique properties is the fact that even though it has a very high molecular weight, high concentrations in water produce only low viscosity; 40% concentrations are about as viscous as a thin sugar syrup, and solutions of 50% concentration can be made. So gum arabic is used for its properties other than thickening aqueous systems. It is used for its interfacial activity and for its compatibility with high concentrations of soluble solids (sugar). It does form gels at concentrations greater than 50%.

Gum arabic is an exudate of acacia trees, of which there are many species distributed over tropical and subtropical regions. The most important growing area for species that give the best gum is in Sudan. Purified, spray-dried forms of gum arabic are commonly used.

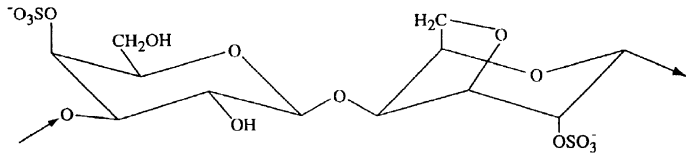
There are other exudate gums that were gums of commerce in ancient times and used up to about World War II that are used in only very small amounts today. They are gum karaya, gum ghatti, and gum tragacanth [64,65,66,67].

4.8 Carrageenans and Related Gums

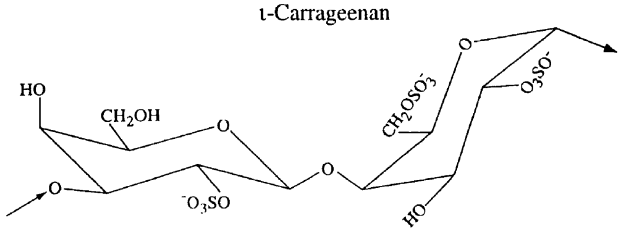
There is one large family of algal polysaccharides (the carrageenans and related polysaccharides [68,69,70,71,72,73]) that are acidic (anionic) because they contain sulfate half-ester groups. The term carrageenan denotes a group of sulfated galactans extracted from red marine algae (seaweeds) with dilute alkaline solutions. The sodium salt of a carrageenan is normally produced. Carrageenan polysaccharide molecules are linear chains of D-galactopyranosyl units joined by alternating (1→3)-α-D- and (1→4)-β-D-glycosidic linkages, with most sugar units having one or two sulfate half-ester groups esterified to O2 and/or O6. Their sulfate contents range from 15 to 40%. Units often contain a 3,6-anhydro ring. The principal structures are termed kappa (κ), iota (ι), and lambda (λ) carrageenans (❯ *Fig. 1*). Carrageenans, as



κ-Carrageenan



ι-Carrageenan



λ-Carrageenan

Figure 1
The predominant repeating unit structures in the three basic types of carrageenan

extracted, are mixtures of non-homogeneous polysaccharides. The disaccharide units shown in Figure 1 represent the predominate building block of each type, but are not repeated throughout the entire chain. Commercial carrageenan products have average molecular weights in the range 10^5–10^6.

Carrageenan products, of which a large number are usually available from a single supplier for different specific applications, contain different proportions of the three main behavioral types (kappa, iota, and lambda) produced by starting with mixtures of red seaweed species. Other substances, such as potassium ions and sugar (for standardization), may be added to the obtained powder to produce the commercial product.

Carrageenan products dissolve in water to form highly viscous solutions. The viscosity is quite stable over a wide range of pH values because the sulfate half-ester groups are always ionized, even under strongly acidic conditions, giving the molecules a negative charge. However, carrageenans undergo rather facile depolymerization in hot acidic solutions.

Segments of molecules of κ- and ι-type carrageenans exist as double helices of parallel chains. In the presence of potassium or calcium ions, thermoreversible gels form upon cooling a hot solution containing double-helical segments. Gelation can occur in water at concentrations as low as 0.5%. When κ-type carrageenan solutions are cooled in the presence of potassium ions,

a stiff, brittle gel results. Potassium and calcium ions together produce a high gel strength. Gels made with κ-type carrageenans are the strongest of the carrageenan gels. These gels tend to synerese as junction zones within the structure grow in length.

Iota-type carrageenans are a little more soluble than are kappa-types, but again, only the sodium salt form is soluble in cold water. Iota types gel best with calcium ions. The resulting gel is soft and resilient, has good freeze-thaw stability, and does not synerese, presumably because ι-type carrageenans are more hydrophilic and form fewer junction zones than do κ-type carrageenans.

During cooling of hot solutions of κ- or ι-type carrageenans, gelation occurs because the linear molecules are unable to form continuous double helices due to the presence of structural irregularities. The linear helical portions then associate to form a three-dimensional network in the presence of the appropriate cation. All salts of lambda-type carrageenans are soluble and non-gelling.

Under conditions in which double-helical segments are present, carrageenan molecules, particularly those of the kappa type, form junction zones with the unsubstituted segments of LBG (*Sect. 4.2*) to produce rigid brittle gels. This gelation occurs at a concentration one-third that needed to form a pure κ-type carrageenan gel. The synergistic effect between a κ-type carrageenan and LBG produces gels with greater elasticity and gel strength, and with less syneresis than occurs with gels made with potassium kappa-carrageenate alone. This combination is often used in food products.

A useful property of carrageenans is their reactivity with proteins, particularly those of milk [74]. Kappa-type carrageenans complex with kappa-casein micelles of milk, forming a weak, thixotropic, pourable gel. The thickening effect of κ-type carrageenans in milk is 5–10 times greater than it is in water.

Carrageenans are most often used because of their ability to form gels. They will form gels with both milk and water. Mixtures of carrageenan types are used to provide a wide range of products that are standardized with various amounts of sucrose, glucose (dextrose), buffer salts, or gelling aids, such as potassium chloride. The available commercial products form a variety of gels: gels that are clear or turbid, rigid or elastic, tough or tender, heat-stable or thermally reversible, and do or do not undergo syneresis. Carrageenan gels do not require refrigeration because they do not melt at room temperature. They are freeze-thaw stable.

Also prepared and used is an alkali-modified seaweed flour that has been called processed Euchema seaweed (PES), Philippine natural grade (PNG) carrageenan, semi-refined carrageenan, and alternatively refined carrageenan [71]; but is now often called carrageenan like the extracted carrageenan. To prepare semi-refined carrageenan, a red seaweed (usually a *Euchema* sp.) is treated with a solution of potassium hydroxide. Because the potassium salts of the kappa types of carrageenans found in these seaweeds are insoluble, the carrageenan molecules are not solubilized and not extracted out. Primarily low-molecular-weight soluble components are removed from the seaweed fronds during this treatment. The remaining seaweed is dried and ground to a powder. Semi-refined carrageenan is, therefore, a composite material that contains not only the molecules of carrageenan that would be extracted with dilute sodium hydroxide, but also other cell-wall materials.

Carrageenans are used in food products and toothpaste. The two other gums, agar [75,76,77,78,79] and furcellaran (also called Danish agar) [80], are also extracted from red seaweeds. Agar, whose structure is somewhat different than those of the carrageenans and furcellaran, is pri-

marily used to make media for growing microorganisms, although some is used in foods. Small amounts of furcellaran are used in processed milk products.

4.9 Cationic Gums

There are no natural cationic gums. The one that comes closest is chitosan [81,82,83,84,85,86]. Chitosan is largely deacetylated chitin. Chitin (Chap. 6.1) is a linear polymer of (1→4)-linked 2-acetamido-2-deoxy-β-D-glucopyranosyl (*N*-acetyl-β-D-glucosaminyl) units, i. e., a structure like that of cellulose, but with an acetamido (–NH–CO–CH_3) group on C2 of the monomer units in place of the hydroxyl group of cellulose. Chitosan, which is mostly deacetylated, has more amino (–NH_2) groups than acetamido groups along the polymer molecule. Chitosan, which is primarily obtained from crab and lobster shells using a hot, concentrated solution of sodium hydroxide, is soluble only in acidic solutions which convert the amino group to the ammonium ion (–NH_3^+) form. Carboxymethylchitosan is amphoteric. Chitosan can also be obtained from the mycelia of certain fungi. Little chitosan is used commercially.
One cationic gum (cationic guar gum) is produced by etherfication with a cationic reagent, viz., 1-chloro-2-hydroxy-3-trimethylammoniumpropane. In an alkaline solution, the reagent is converted to 3-trimethylammoniumpropylene oxide, which reacts with the guar gum.

4.10 Fermentation Gums other than Xanthans and Gellans

Curdlan is a bacterial polysaccharide made by *Agrobacterium biovar* [87,88,89]. It is a linear (1→3)-β-glucan (MW ~73,000) that forms a triple helix. Curdlan is insoluble in cold water. When aqueous dispersions of curdlan are heated, two types of gels form. First curdlan dissolves. When the solution reaches 55–66 °C, then is cooled, a reversible gel forms. The gel melts when held at about 60 °C. When the thermoreversible gel is heated to a temperature above 80 °C, an irreversible gel forms. Heating to higher temperatures results in stronger irreversible gels. Transition temperatures are a function of concentration.
Dextrans [90,91] are a group of bacterial polysaccharides composed of α-D-glucopyranosyl units. Their linkages are mainly (1→6); they also contain (1→3) linkages as branch points. Native dextrans are usually of very high molecular weight (at least 10^7), but lower molecular weight types are commercially available. They are not used as thickeners or gel formers. They are used as blood volume expanders, to make beads (via crosslinking) for size-exclusion chromatography, to complex iron ions for animal nutrition supplements, in the preparation of X-ray film, and in the Bayer process for aluminum refining.
A polysaccharide with the commercial name Diutan is highly pseudoplastic and proposed for use as a rheology modifier in concretes, grouts, and mortars. Diutan has the same backbone structure as gellan, but has a disaccharide side unit [28].
Another polysaccharide also has the same backbone structure as gellan, but has a single monosaccharide side on the backbone [28]. It is known commercially as welan. Welan was originally proposed as a thickener for oil well drilling fluids. It is now used in de-icing fluids, as a pigment suspending agent for concrete, and as a tire sealant.
Pullulan [92] is produced in cultures of *Aureobasidium* sp., *A. pullulans* being primarily used. Its structure is that of α-maltotriosyl units linked (1→6); a few (~6.6%) α-maltotetraosyl units

are present. Depending on the bacterial strains used and the conditions or times of fermentation, molecular weights of from 10^3 to more than 10^6 can be obtained. Pullulan dissolves in water, forming stable, low-viscosity, non-gelling solutions. It is useful for preparing water-soluble films.

Succinoglycan [89] produced by a strain of *Alcaligenes faecalis* contains D-glucopyranosyl and D-galactopyranosyl units in a 7 : 1 molar ratio, succinic acid ester groups, and pyruvyl cylic acetal groups at the end of a tetrasaccharide side unit that is found on every fourth unit of the main chain. The β-D-glucopyranosyl units of the main chain are joined by (1→4) and (1→3) linkages in a 3 : 1 ratio, making it a mixed-linkage β-glucan. The side units are linked (1→6) to the main chain.

4.11 Other Plant Gums

There are other water-soluble polysaccharides that have some economic importance, but are generally not employed as isolated substances. Many are beneficial because of the physiological effects they produce upon consumption rather than the physical properties they impart to aqueous systems. Some are presented here in alphabetical order.

Aloe polysaccharides [93] are obtained from the fleshy leaves of *Aloe* sp. Crude preparations of Aloe polysaccharides are especially used in cosmetic and pharmaceutical preparations to be applied to the skin, but have other uses as well.

Cereal brans contain a polysaccharide of a class called beta-glucans [94,95]. (Cellulose is a β-glucan, but unlike cellulose, the polysaccharides called β-glucans contain entirely or partly (1→3) linkages.) β-Glucan is the major non-starch polysaccharide in oat and barley brans. Oat bran [96] is the principal source. Oat β-glucan is a linear chain of β-glucopyranosyl units, about 70% of which are linked (1→4) and about 30% (1→3). Such (1→4, 1→3)-β-glucans are known as mixed-linkage β-glucans. They can be extracted with hot water, which may occur during cooking, eg., in the preparation of oatmeal. Oat and barley bran are added to certain food products like breads. The value of oat β-glucan, which has been called oat gum, especially lies in its effectiveness in reducing serum cholesterol and providing the other health benefits of soluble dietary fiber, but the health benefits imparted by soluble dietary fiber polysaccharides (reducing postprandial serum glucose levels and the accompanying insulin response and protecting against hypertension, stroke, cardiovascular disease, and type-2 diabetes) vary in degree from gum to gum. A negative side to β-glucans is that they cause problems in the brewing industry. They increase wort viscosity and decrease the rate of wort filtration, beer clarity, and stability.

Inulin [97,98,99,100] is the only commercial, water-soluble polysaccharide containing D-fructose, a ketose. It is a linear molecule of β-D-fructofuranosyl units linked (2→1). It is much smaller than other gums with chain lengths of only 15–22 units and is much more easily depolymerized under acidic conditions because its monomer units are in the furanosyl ring form. It is obtained mostly from roots of the chicory plant. Because of its small molecular size, hot solutions are relatively non-viscous even at concentrations of 50%. When hot solutions of > 25% concentration are cooled, a particulate gel forms. Inulin is used for its health benefits, viz, as a prebiotic.

Konjac glucomannan (also known by other names) [97,101,102] is a mostly linear, but slightly branched polymer containing β-D-mannopyranosyl units and β-D-glucopyranosyl units, with all inter-unit linkages being (1→4). The βMan*p* : βGlc*p* ratio varies from ∼1.5 : 1 to ∼2.0 : 1, depending on the source. The two units are randomly distributed within the polymer chains. Its molecular weight is $\sim 10^6$. Preparations contain 0.05–0.17 acetyl (acetate ester) groups per monomer unit. Konjac glucomannan molecules are rather acid stable (as compared to other polysaccharides). It forms strong, elastic, thermoreversible gels when used in combination with κ-carrageenan (*Sect. 4.8*) or xanthan (*Sect. 4.1*). Heat-stable gels are formed from deacetylated products. The gum is obtained from tubers of *Amorphallus konjac* (known by various local names) grown primarily in mountainous or hilly regions of Southeast Asia. Konjac glucomannan finds little use outside of Asia.

A low-viscosity gum, an arabinogalactan, is extracted from wood chips of *Latrix* sp. This substance, known both as larch gum and Larch arabinogalactan [66], consists of a backbone chain of (1→3)-linked β-D-galactopyranosyl units, each of which bears a side-chain substituent attached to O6. Most of the side chains are β-galactobiosyl (disaccharide) units. Other side units are β-L-arabinopyranosyl-(1→3)-α-L-arabinofuranosyl units or simply α-L-arabinofuranosyl units. Its molecular weight ranges from ∼16,000 to ∼100,000. Its solubility is high, and solutions with concentrations as high as 60% are fluid. Larch arabinogalactan is marketed as a prebiotic and a stimulator of the immune system.

Cereal flours contain polysaccharides that are classified as hemicelluloses (Chap. 6.1), but which, when they come from cereal flour, are called pentosans. Pentosans are mostly arabinoxylans and are divided into two classes: water soluble and water insoluble. Their structure is that of a main chain of (1→4)-linked β-D-xylopyranosyl units to which are attached short side units, mostly monosaccharide units and mostly L-arabinosyl units. Two side units may be on the same backbone unit, i. e., a backbone unit may be doubly branched. The water-soluble molecules have more side units than do the water-insoluble molecules. The molecules also contain ferulic acid ester groups. Of primary interest are wheat flour arabinoxylans because of their water-absorbing and water-binding capacities. Especially do arabinoxylan molecules crosslinked by oxidative coupling of the ferulic acid ester groups have high water-holding capacity. Even though they are present in very small amounts, arabinoxylans significantly affect distribution of water between bread dough constituents, the rheological properties of both the dough and the finished product, gas retention in the baked product, and texture of the baked product.

Arabinogalactans are also polysaccharide components of wheat flour. Arabinogalactans in wheat flour are attached to protein molecules and thus are part of the structures of arabinogalactan-protein molecules. Like the arabinoxylans, there are variations in both amounts and structures of water-extractable arabinogalactan-protein molecules in wheat flours.

Psyllium gum (gum psyllium) can be extracted from psyllium seed husks (hulls) [93]. The husks themselves are incorporated into products sold to prevent constipation and to a few food products as a source of dietary fiber. The gum is extracted when water or milk is added to these products.

Many other polysaccharides have been investigated as gums, but have not reached the status of commercial viability. Among these are polysaccharides from chia seeds [93], corn hull (corn fiber), flaxseed [93], okra [93], and tamarind seeds [93] (used in India).

Yet other water-soluble polysaccharides are used commercially but are not referred to as gums. They include the chondroiton sulfates (animal) (Chap. 6.1), hyaluronic acid (animal and bacterial) (Chap. 6.1), and polydextrose (synthetic).

References

1. BeMiller JN (2007) Carbohydrate chemistry for food scientists, 2nd edn. AACC International, St. Paul
2. Cui SW (2005) (ed) Food carbohydrates. CRC Press, Boca Raton, FL
3. Dumitriu S (2005) (ed) Polysaccharides, 2nd edn. CRC Press, Boca Raton, FL
4. Eliasson AC (2006) (ed) Carbohydrates in foods, 2nd edn. Taylor and Francis, Boca Raton, FL
5. Imeson A (1997) (ed) Thickening and gelling agents for food, 2nd edn. Blackie Academic and Industrial, London
6. Kennedy JF, Phillips GO, Wedlock DJ, Williams PA (1985) (eds) Cellulose and its derivatives. Ellis Horwood, Chichester, UK
7. Phillips GO, Williams PA (2000) (eds) Handbook of hydrocolloids. CRC Press, Boca Raton, FL
8. Stephen AM (1995) (ed) Food polysaccharides and their applications. Marcel Dekker, New York
9. Stephen AM, Phillips GO, Williams PA (eds) (2006) Food polysaccharides and their applications, 2nd edn. CRC/Taylor & Francis, Boca Raton, FL
10. Tomasik P (2004) (ed) Chemical and functional properties of food saccharides. CRC Press, Boca Raton, FL
11. Vandamme EJ, DaBaets S, Steinbüchel A (2002) (eds) Biopolymers, vol 5 Polysaccharides I. Wiley VCH, Weinheim
12. Vandamme EJ, Dabaets S, Steinbüchel A (2002) (eds) Biopolymers, vol 6 Polysaccharides II. Wiley VCH, Weinheim
13. Whistler RL, BeMiller JN (1992) (eds) Industrial gums, 3rd edn. Academic Press, San Diego
14. BeMiller JN (2001) Classification, structure, and chemistry of polysaccharides in foods. In: Cho SS, Dreher ML (eds) Handbook of dietary fiber. Marcel Dekker, New York, p 603
15. BeMiller JN Polysaccharides: Properties. In: Reference 1, Chapter 5
16. Wang Q, Cui S (2005) Understanding the physical properties of food polysaccharides. In: Reference 2, p 161
17. Doublier JL, Cuvelier G (2006) Gums and hydrocolloids: functional aspects. In: Reference 4, p 233
18. Williams PA, Phillips GO (2000) Introduction to food hydrocolloids. In: Reference 7, p 1
19. Clark AH (2000) Biopolymer gelation—the structure-property relationship. In: Williams PA, Phillips GO (eds) Gums and stabilisers for the food industry—10. The Royal Society of Chemistry, Cambridge, p 91
20. Clark AH (1996) Properties of biopolymer gels. In: Finch CA (ed) Industrial water soluble polymers. The Royal Society of Chemistry, Cambridge, p 106
21. Lapasin R, Pricl S (1995) Rheology of industrial polysaccharides. Chapman & Hall, New York
22. BeMiller JN Xanthan. In: Reference 1, Chapter 10
23. Stokke BT, Christensen BE, Smidsrød O (2005) Macromolecular properties of xanthan. In: Reference 3, p 433
24. Urlacher B, Noble O (1997) Xanthan gum. In: Reference 5, p 284
25. Sworn G (2000) Xanthan gum. In: Reference 7, p 103
26. Morris VJ (2006) Bacterial polysaccharides. In: Reference: 9, p 413
27. Born K, Langendorff V, Boulenguer P (2002) Xanthan. In: Reference 10, p 259
28. Kang KS, Pettitt DJ (1993) Xanthan, Gellan, Welan, and Rhamsan. In: Reference 13, p 341
29. Guar, locust bean, and tara gums. In: Reference 1, Chapter 8
30. Fox JE (1997) Seed gums. In: Reference 5, p 262
31. Wielinga WC (2000) Galactomannans. In: Reference 7, p 137
32. Gidley MJ, Reid JSG (2006) Galactomannans and other cell wall storage poysaccharides in seeds. In: Reference 9, p 181
33. Dierckx S, Dewettinck K (2001) Seed gums. In: Reference 12, p 321
34. Maier H, Anderson M, Karl C, Magnuson K, Whistler RL (1993) Guar, locust bean, tara and fenugreek gums. In: Reference 13, p 181

35. Cellulose and cellulosics. In: Reference 1, Chapter 7
36. Zecher D, Gerrish T (1997) Cellulose derivatives. In: Reference 5, p 60
37. Nicholson MD, Merritt FM (1985) Cellulose ethers. In: Reference: Reference 6, p 363
38. Reveley A (1985) A review of cellulose derivatives and their industrial applications. In: Reference 6, p 212
39. Felcht UH (1985) Cellulose ethers—synthesis, application and analytical aspects. In: Reference 6, p 273
40. Murray JCF (2000) Cellulosics. In: Reference 7, p 219
41. Coffey DG, Bell DA, Henderson A (2006) Cellulose and cellulose derivatives. In: Reference 9, p 147
42. Grover JA (1993) Methylcellulose and its derivatives. In: Reference 13, p 475
43. Desmaris AJ, Went RF (1993) Hydroxyalkyl and ethyl ethers of cellulose. In: Reference 113, p 505
44. Fedderson RL, Thorp SN (1993) Sodium carboxymethylcellulose. In: Reference 13, p 537
45. Algins/Alginates. In: Reference 1, Chapter 13
46. Onsøyen E (1997) Alginates. In: Reference 5, p 22
47. Draget KI (2000) Alginates. In: Reference 7, p 379
48. Draget KI, Moe ST, Skjåk-Braek G, Smidsrød O (2006) Alginates. In: Reference 9, p 289
49. Draget KI, Smidsrød O, Skjåk-Braek G (2002) Alginate from algae. In: Reference 12, p 215
50. Clare K (1993) Algin. In: Reference 13, p 105
51. Pectins. In: Reference 1, Chapter 14
52. Rolin C, Nielsen BU, Glahn PE (1998) Pectin. In: Reference 3, p 377
53. May CD (1997) Pectins. In: Reference 5, p 230
54. May CD (2000) Pectins. In: Reference 7, p 169
55. Voragen AGJ, Pilnik W, Thibault JF, Axelos MAV, Renard GMGC (1995) Pectins. In: Reference 8, p 287
56. Lopes de Silva JA, Rao MA (2006) Pectins: structure, functionality, and uses. In: Reference 9, p 353
57. MacDougall AJ, Ring SJ (2004) Pectin polysaccharides. In: Reference 10, p 181
58. Ralet MC, Bonnin E, Thibault JF (2002) Pectins. In: Reference 12, p 345
59. Rolin C (1993) Pectin. In: Reference 13, p 257
60. Gellans, curdlan, dextrans, and levans. In: Reference 1, Chapter 11
61. Gibson W, Sanderson GR (1997) Gellan gum. In: Reference 5, p 119
62. Sworn G (2000) Gellan gum. In: Reference 7, p 117
63. Giavasis I, Harvey LM, McNeill B (2000) Crit Rev Biotechnol 20:177
64. Gum arabic and other exudate gums. In: Reference 1, Chapter 15
65. Wareing MV (1997) Exudate gums. In: Reference 5, p 86
66. Williams PA, Phillips GO, Stephen AM, Churms SC (2006) Gums and mucilages. In: Reference 9, p 455
67. Weiping W (2000) Tragacanth and karaya. In: Reference 7, p 231
68. Carrageenans. In: Reference 1, Chapter 12
69. Thomas WR (1997) Carrageenan. In: Reference 5, p 45
70. Imeson A (2000) Carrageenan. In: Reference 7, p 87
71. Brixler HJ, Johndro KD (2000) Phillipine natural grade or semi-refined carrageenan. In: Reference 8, p 425
72. Piculell L (2006) Gelling carrageenans. In: Reference 9, p 239
73. Therkelsen GH (1993) Carrageenan. In: Reference 13, p 145
74. Hansen PMT (1993) Food hydrocolloids in the dairy industry. In: Nishinari K, Doi E (eds) Food hydrocolloids. Plenum Press, New York, p 211
75. Matsuhashi T (1998) Agar. In: Reference 3, p 335
76. Armisén R (1997) Agar. In: Reference 5, p 1
77. Armisén R, Galatas F (2000) Agar. In: Reference 8, p 21
78. Stanley NF (2006) Agars. In: Reference 9, p 217
79. Selby HH, Whistler RL (1993) Agar. In: Reference 13, p 87
80. Bjerre-Petersen E, Christiansen J, Hemmingsen P (1973) Furcellaran. In: Whistler RL, BeMiller JN (eds) Industrial gums. Academic Press, New York, p 123
81. Einbu A, Varum KM (2004) Structure-property relationships in chitosan. In: Reference 10, p 217
82. Kubota N, Kikuchi Y (1998) Macromolecular complexes of chitosan. In: Reference 3, p 595
83. Terbojevich M, Muzzarelli RAA (2000) Chitosan. In: Reference 7, p 367
84. Vårum KM, Smidsrød O (2006) Chitosans. In: Reference 9, p 497

85. Peter MG (2002) Chitin and chitosan in fungi. In: Reference 12, p 123
86. Peter MG (2002) Chitin and chitosan from animal sources. In: Reference 12, p 481
87. Nishinari K, Zhang H (2000) Curdlan. In: Reference 7, p 269
88. Lee IY (2002) Curdlan. In: Reference 11, p 135
89. Harada T, Terasaki M, Harada A (1993) Curdlan. In: Reference 13, p 427
90. Leathers TD (2002) Dextran. In: Reference 11, p 299
91. de Belder AN (1993) Dextran. In: Reference 13, p 399
92. Tsujiska Y, Mitsuhashi M (1993) Pullulan. In: Reference 13, p 447
93. BeMiller JN, Whistler RL, Barkalow DG, Chen CC (1993) Aloe, chia, flaxseed, okra, psyllium seed, quince seed, and tamarind gums. In: Reference 13, p 227
94. Stone BA, Clarke AE (1992) Chemistry and biology of (1→3)-β-glucans. LaTrobe University Press, Victoria
95. Morgan K (2000) Cereal β-glucans. In: Reference 7, p 287
96. Wood PJ (ed) (1993) Oat bran. American Association of Cereal Chemists, St. Paul
97. Inulin and konjac glucomannan. In: Reference 1, Chapter 9
98. Sensus Operations CV (2000) Frutafit® -inulin. In: Reference 7, p 397
99. Praznik W, Ciéslik E, Huber A (2004) Fructans: occurrence and application in food. In: Reference 10, p 197
100. Franck A, De Leenheer L (2002) Inulin. In: Reference 12, p 439
101. Thomas WR (1997) Konjac gum. In: Reference 5, p 169
102. Takigami S (2000) Konjac mannan. In: Reference 7, p 413

6.5 Bacterial Cell Wall Components

Cynthia Ginsberg[1,2], *Stephanie Brown*[1,2], *Suzanne Walker**[1,2]
[1] Department of Microbiology and Molecular Genetics, Harvard Medical School, Boston, MA 2115, USA
[2] Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA 02138, USA
suzanne_walker@hms.harvard.edu

In: *Glycoscience*. Fraser-Reid B, Tatsuta K, Thiem J (eds)
Chapter-DOI 10-1007/978-3-540-30429-6_38:

Abstract

Bacterial cells are surrounded by a variety of cell-surface structures that allow them to thrive in extreme environments. Components of the cell envelope and extracellular matrix are responsible for providing the cells with structural support, mediating intercellular communication, allowing the cells to move or to adhere to surfaces, protecting the cells from attack by antibiotics or the immune system, and facilitating the uptake of nutrients. Some of the most important cell wall components are polysaccharide structures. This review discusses the occurrence, structure, function, and biosynthesis of the most prevalent bacterial cell surface polysaccharides: peptidoglycan, lipopolysaccharide, arabinogalactan, and lipoarabinomannan, and capsular and extracellular polysaccharides. The roles of these polysaccharides in medicine, both as drug targets and as therapeutic agents, are also described.

Keywords

Cell wall; Peptidoglycan; Lipopolysaccharide; Arabinogalactan; Lipoarabinomannan; Capsular polysaccharides; Extracellular polysaccharides

Abbreviations

ACP	acyl carrier protein
AG	arabinogalactan
AraLAM	uncapped LAM
CPS	capsular polysaccharide
EPS	exocellular polysaccharide
GFP	green fluorescence protein
HMW	high molecular weight
IM	inner membrane
LAM	lipoarabinomannan
LM	lipomannan
LMW	low molecular weight
LPS	lipopolysaccharide
mAGP	mycolyl-arabinogalactan-peptidoglycan complex
ManLAM	mannose-capped LAM
MGT	monofunctional glycosyltransferase
OM	outer membrane
OMP	outer membrane protein
PBP	penicillin-binding protein
PGT	peptidoglycan glycosyltransferase
PI	phospho-inositol
PILAM	inositol phosphate-capped LAM
PIM	phosphatidylinositol mannoside
PMB	polymyxin B
TPase	transpeptidase
Dab	2,4-diaminobutyric acid
DAP	2,6-diaminopimelic acid
***m*DAP**	*meso*-2,6-diaminobutyric acid
HyDAP	*meso*-2,6-diamino-3-hydroxypimelic acid
Hyg	*threo-P*-hydroxyglutamic acid
HyLys	hydroxylysine
Orn	ornithine
f	furanose
p	pyranose
P	phosphate
PP	pyrophosphate
PEtN	2-aminoethyl phosphate
Ara	arabinose
Col	L-colitose
Fru	fructose
Fuc	fucose
FucNAc	*N*-acetylfucosamine
Gal	D-galactose
GalA	galacturonic acid

GalN	2-amino-2-deoxy-galactose
GalNAc	*N*-acetylgalactosamine
GalNAcA	*N*-acetylgalacturonic acid
Glc	D-glucose
GlcA	glucuronic acid
GlcN	2-amino-2-deoxy-glucose
GlcNAc	*N*-acetylglucosamine
Hep	L-glycero-D-*manno*-heptose
Kdo	2-keto-3-deoxyoctonate
Man	mannose
ManNAc	*N*-acetylmannosamine
MurNAc	*N*-acetylmuramic acid
NAG	*N*-acetyl-D-glucosamine
NAM	*N*-acetyl-D-muramic acid
Neu5Ac	5-*N*-acetyl-D-neuraminic acid
Qui	quinovose
QuiNAc	*N*-acetylquinovosamine
Rha	rhamnose
Rib	ribulose

1 Introduction

Bacteria have evolved over the last three billon years to occupy such diverse niches as υ deep sea vents and the GI tracts of multicellular organisms as different as nematodes and humans. Some, such as the mycobacteria that cause tuberculosis and leprosy, have even evolved to live inside other cells. Accordingly, different bacterial species produce wildly different cell-surface structures adapted to life in these diverse environs. Bacterial cell surface architecture is complex, and the assorted cell surface structures perform a host of important functions. They play structural roles, providing the cells with mechanical support and controlling cell shape. They form permeability barriers, controlling entry of molecules into the cells and protecting the bacteria from antibiotics. They also mediate interactions between cells, facilitating intercellular adhesion and invasion of host tissue, and allowing bacteria to evade the immune system. Several of the most important cell surface components contain saccharide structures, and these are the subjects of this review.

In most bacteria, the cytoplasmic membrane is surrounded by a cell wall containing layers of a crosslinked carbohydrate polymer called peptidoglycan [1]. Peptidoglycan is the major structural component of the cell wall and is essential for protecting bacteria from osmotic lysis. Disruption of the peptidoglycan layers causes the cell to burst. The structure and biosynthesis of peptidoglycan are remarkably conserved across bacterial species, making it a very attractive antibiotic target [2]. In fact, peptidoglycan biosynthesis is the target of several clinically important antibiotics, including the penicillins and vancomycin.

Other cell-surface structures, although less well conserved than peptidoglycan, typify broad groups of bacteria. For example, most bacteria can be divided into two groups that possess different general surface architectures, termed “Gram-positive” and “Gram-negative” (❯ *Fig. 1*).

Figure 1
Bacterial cell envelope architecture

This classification is based on a staining technique developed in 1884 by a Danish doctor, Hans Christian Gram. Gram observed that some bacteria (Gram-positive) were able to retain crystal violet dye, while others (Gram-negative) were not. This differential behavior hints at one major difference in the two types of cell envelopes: while both types of bacteria have a hydrophilic cell wall made of peptidoglycan, the thickness of the peptidoglycan layers range from less than 10 nM in Gram-negative cells to up to 80 nM in Gram-positive organisms [3]. The crystal violet dye diffuses through the thin peptidoglycan layers of Gram-negative bacteria and can be washed out, but is retained in the thicker matrix of Gram-positive organisms [4]. Perhaps to compensate for their thinner, and thus more fragile, peptidoglycan matrix, Gram-negative bacteria possess an outer membrane that blocks the passage of many different types of molecules into the periplasmic space that contains the peptidoglycan layers [5]. *E. coli* and many other Gram-negative bacteria are able to thrive in the human intestine because their outer membranes exclude bile salts and other toxic molecules. The outer membrane also acts as a barrier to many antibiotics, so some classes of drugs that are very effective against Gram-positive infections are virtually useless against Gram-negative pathogens. In addition to *E. coli*, some of the better-known Gram-negative pathogens include *Vibrio cholerae*, *Neisseria gonorrhoeae*, *Neisseria meningitidis*, *Pseudomonas aeruginosa*, and *Hemophilus influenzae*.

The outer membrane is an unusual structure consisting of an asymmetric lipid bilayer, with an inner leaflet containing phospholipids and an outer leaflet composed of lipopolysaccharide (LPS) [6]. LPS is a complex glucosamine-based glycolipid composed of three parts: the membrane anchor lipid A, a short core oligosaccharide and a highly variable *O*-antigen polysaccharide [5]. Lipid A, also referred to as "endotoxin," stimulates the innate immune response, promoting the release of pro-inflammatory cytokines and thereby producing some of the dramatic symptoms seen with systemic Gram-negative infections, including fever and sepsis [7]. Gram-negative *O*-antigen usually produces a strong antibody response, which can confer protection against subsequent infection [8]. For some Gram-negative pathogens, vaccine strategies employing synthetic polysaccharides mimicking fragments of *O*-antigen have been investigated for protective effects [9]; however, since most pathogenic species synthesize an impressive array of *O*-antigen structures (*E. coli*, for example, synthesizes more than 150 different types) [10], the immune protection conferred by infection or vaccination with one type of *O*-antigen may not confer protection against bacteria of the same species possessing different *O*-antigen structures. Other distinguishing features of the Gram-negative outer membrane include lipoproteins and integral membrane proteins (outer membrane proteins or OMPs) that are embedded in the outer membrane [11,12]. Some OMPs function as channels that allow the cell to take in nutrients and excrete waste products. Others are enzymes, or serve as adhesins. In addition to LPS, the outer membrane can also include other glycolipids such as enterobacterial common antigen, a structure specific to the Enterobacteriaceae including *E. coli* [13,14]. The outer membrane is also the attachment site for surface organelles, such as pili, which are involved in adhesion and allow bacteria to colonize cells or environmental surfaces [15].

The cell walls of Gram-positive bacteria, including the pathogens *S. aureus*, *Bacillus anthracis*, *Streptococcus pneumoniae*, *and Enterococcus faecalis* contain thick layers of peptidoglycan. The peptidoglycan layers serve as both a protective barrier and as a scaffold for the attachment of secondary cell wall polymers and surface proteins. Surface proteins include hydrolytic enzymes involved in peptidoglycan turnover, as well as structures such as pili

and adhesin/invasin proteins that help bacteria colonize surfaces and invade host tissue [16]. Important secondary cell wall polymers include the polyanionic wall teichoic acids (WTAs) and teichuronic acids, which are covalently linked to the peptidoglycan layers, and lipoteichoic acids (LTAs), which are anchored in the cytoplasmic membrane and extend through the cell wall into extracellular space [17,18]. Both WTAs and LTAs are involved in virulence in pathogenic organisms. LTAs, like lipid A, stimulate innate immunity; WTAs have been shown to mediate processes such as adhesion, biofilm formation, and antibiotic resistance [18]. Lipoteichoic acids also serve as the attachment site for a number of the surface proteins, including enzymes that hydrolyze peptidoglycan.

A subset of Gram-positive bacteria possess a cell envelope that differs markedly from other Gram-positive bacteria. These microorganisms, collectively classified as Corynebacterineae or mycolata, include notorious human pathogens such as *Mycobacterium tuberculosis*, *Corynebacterium diphtheriae*, and *Mycobacterium leprae*. Other Corynebacteria are industrially important for the production of amino acids [19]. The Corynebacterineae have a multilayered cell wall comprised of peptidoglycan covalently attached through a disaccharide linker to an arabinogalactan polysaccharide, which is tethered at its other end to long-chain beta-hydroxy fatty acids called mycolic acids [20,21]. The mycolic acids are perpendicular to the plane of the membrane, forming a thick lipid barrier functionally similar to the Gram-negative outer membrane.

Mycobacterial infections are particularly difficult to treat because most antibiotics cannot penetrate the densely mycolylated cell walls [22]. Cell wall biosynthesis inhibitors are usually administered in combination with drugs that have intracellular targets, so that the latter can get through the cell wall. For example, the standard regimen for TB treatment involves two months of the transcription inhibitor rifampicin delivered with the mycolic acid biosynthesis inhibitors pyrazinamide and isoniazid, followed by another six months of isoniazid and rifampicin [23].

In addition to the cell envelope structures described above, many Gram-negative and Gram-positive bacteria produce polysaccharide structures called capsular antigens. These include capsular polysaccharides, which are attached to the outer membrane of Gram-negative bacteria or the peptidoglycan layers of Gram-positive cells to form an outer protective layer called the ‘capsule’, and extracellular polysaccharides, which are excreted into the extracellular space [24]. The capsule can perform a variety of roles. For example, some pathogenic bacterial strains construct capsules that mimic mammalian cell surfaces to avoid detection by the immune system [25]. Many capsular polysaccharides, however are highly immunogenic, and have been used to produce vaccines against a number of bacteria including *Neisseria meningitidis* and *Streptococcus pneumoniae* [26]. Extracellular polysaccharides are an important component of biofilms, multicellular cooperative bacterial communities. Biofilms are more resistant to antibiotics and immune attack than free-swimming bacteria, and biofilm formation is implicated in the establishment of persistent infections by several common human pathogens [27].

This review will address the structure and biosynthesis of important microbial cell wall polysaccharides, as well as chemical syntheses of these structures, and their roles in human diseases and disease treatments. The document is divided into sections discussing bacterial peptidoglycan, the Gram-negative outer membrane, corynebacterial cell wall components, and capsular antigens.

2 Peptidoglycan

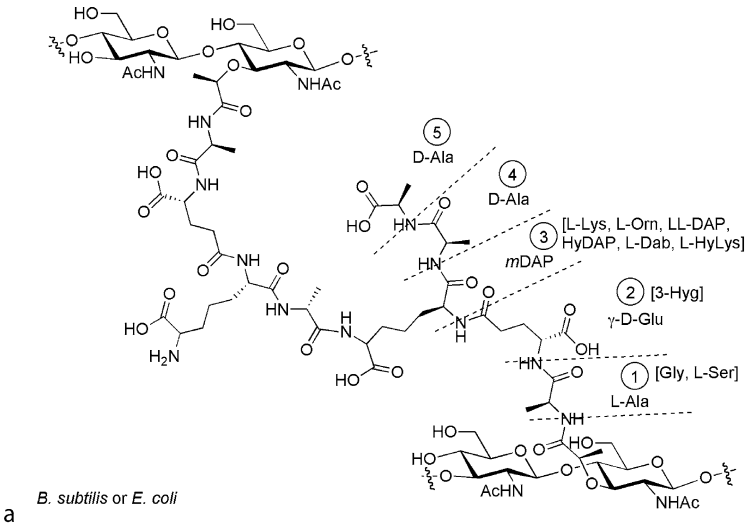
2.1 The Peptidoglycan Monomer Unit

2.1.1 Structure of the Peptidoglycan Monomer Unit

Peptidoglycan is constructed from monomers containing an *N*-acetylglucosamine-β(1,4)-*N*-acetylmuramic acid (NAG-NAM) disaccharide. The C3 position of *N*-acetyl-muramic acid (MurNAc) is modified with a peptide sidechain. The chemical composition of peptidoglycan polysaccharide was first described in the early 1960s using a method involving enzymatic degradation of cell wall extracts and subsequent chromatographic separation [28,29]. Shortly thereafter, Weidel and coworkers and Ghuysen and coworkers published methods for determining the amino acid sequence of the peptide side chain [30,31]. Both groups used enzymatic degradation of cell wall extracts to isolate peptide fragments for chemical structure determination.

Since the execution of these early experiments, the structure of peptidoglycan from many different bacteria has been determined. The structure of the NAG-NAM disaccharide is highly conserved among most bacteria. A notable variation occurs among mycobacteria, where the muramic acid is predominantly *N*-glycolylated rather than *N*-acetylated [32]. 1,6-anhydro MurNAc is also found at the end of the peptidoglycan strands in *E. coli* and other Gram-negative bacteria. Other possible modifications include acetylation or phosphorylation of the muramyl 6′ hydroxyl group, and occasional absence of an *N*-acetyl group or peptide chain.

In contrast with the disaccharide, the peptide chain shows a great deal of interspecies variation [33]. The general peptide chain composition and variations are shown in ❯ *Fig. 2*. The peptide is connected to the lactyl group of muramic acid through its N terminus, and consists of alternating D and L amino acids. D amino acids are not metabolized in eukaryotes, and their only known uses are in peptidoglycan and some prokaryotic secondary metabolites. The first and second amino acids in the chain are usually L-Ala and D-Glu, respectively. The most variation occurs at position 3, which usually contains a diamino acid involved in crossbridge formation as shown in ❯ *Fig. 2*. Meso-diaminopimelic acid (*m*DAP) occurs most frequently, and is found in most Gram-negative bacteria, but L-Lys is also commonly found, and occurs for example, in *S. aureus*. Position 4 is usually a D-Ala, connected by a crossbridge to another peptide chain. If no crossbridge is formed, this residue may either be stripped off (resulting in a tripeptide) or remain substituted by another D-Ala (resulting in a pentapeptide); mature peptidoglycan in most organisms contains primarily a mixture of tri- and tetrapeptides. The composition of the crossbridges is also species-specific [33]. The peptide chains are connected either directly as seen in *E. coli* and *B. subtilis*, or through an interpeptide bridge such as the pentaglycine bridge seen in *S. aureus*, (L-Ala)$_2$ seen in *Enterococcus faecalis* or L-Ser-L-Ala seen in *S. pneumoniae*. Representative structures are shown in ❯ *Fig. 2*. Other variations occur, including cross-links between two *m*DAP residues. The crossbridge variations can confer resistance to beta-lactam antibiotics [34,35,36]. The polypeptide chain variations have been reviewed extensively by Schleifer and Kandler [33].



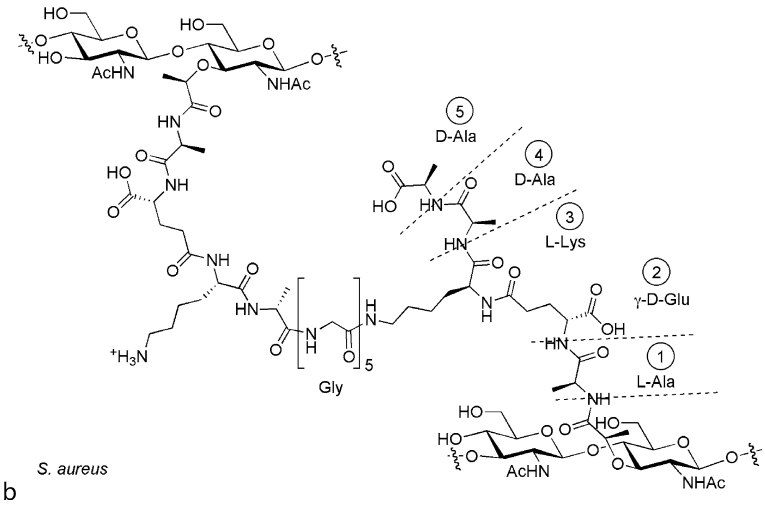


Figure 2
Structure of peptidoglycan. a) A segment of crosslinked peptidoglycan, showing the common variations in peptide chain composition and the direct crosslinking seen in *E. coli* and *B. subtilis*. b) A fragment of *S. aureus* peptidoglycan, showing the pentaglycine bridge

2.1.2 Biosynthesis of the Peptidoglycan Monomer Unit

Peptidoglycan monomers are assembled on a carrier lipid anchored in the cytoplasmic membrane, and then flipped across the membrane for incorporation into peptidoglycan. The intracellular synthesis of the peptidoglycan momomers is well understood (*Fig. 3*) [37]. The peptidoglycan monomers are made from the common cellular building block UDP-GlcNAc

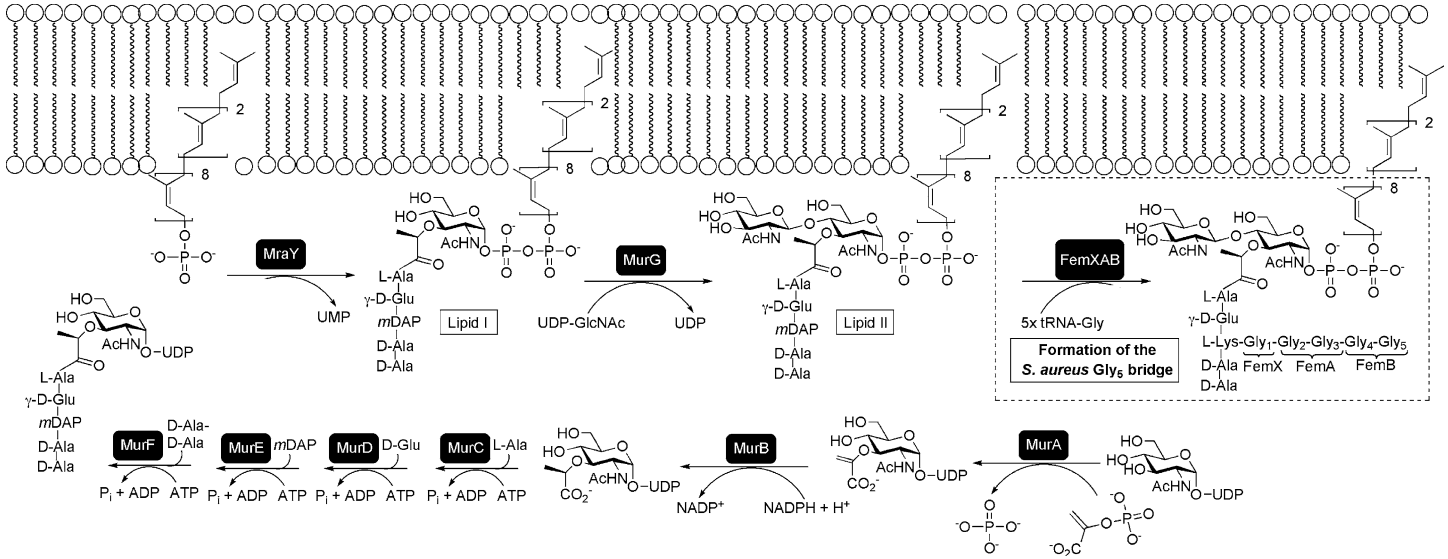


Figure 3
The intracellular steps of peptidoglycan biosynthesis

by the enzymes MurA-G and MraY. Briefly, MurA transfers an enolpyruvyl group to the 3-hydroxyl of UDP-GlcNAc. MurB then reduces the enolpyruvyl group to a lactyl group. The amino acid ligases MurC, D, and E then sequentially catalyze the addition of L-Alanine, D-glutamic acid, and *m*DAP or L-Lysine to the lactyl chain. MurF catalyzes addition of the terminal D-Ala-D-Ala dipeptide, completing the peptide chain. The MurC, D, E and F ligases all catalyze peptide bond formation in an ATP-dependent fashion. MraY transfers phospho-MurNAc-pentapeptide from UDP-MurNAc-pentapeptide to an undecaprenyl phosphate carrier lipid anchored in the cytoplasmic membrane, forming the first membrane linked intermediate (Lipid I). Finally, MurG catalyzes the transfer of GlcNAc from UDP-GlcNAc to C4 of MurNAc, forming GlcNAc-MurNAc-pp-undecaprenyl (Lipid II). Though most bacteria utilize undecaprenyl phosphate as a carrier lipid, the Corynebacterineae instead use decaprenyl phosphate (see ❯ *Section 4*). Lipid I and II intermediates from *Mycobacterium smegmatis* were recently isolated and analyzed, and found to contain decaprenyl rather than undecaprenyl carrier lipid chains [38].

The Lipid I and II building blocks may be further elaborated by many other enzymes that modify the sugars or amino acid chains. Branched peptides are added to the Lipid I and II peptide chains either by enzymes that act in an ATP-dependent fashion similar to the MurC-F ligases [39], or by enzymes that add amino acid residues from aminoacyl tRNA intermediates, such as the *S. aureus* enzymes FemA, FemB and FemX , which form the pentaglycine bridge (❯ *Fig. 3*) [36], and the *S. pneumoniae* enzymes FemM and FemN, which form an L-Ser-L-Ala or L-Ala-L-Ala dipeptide bridge [34,35]. Lipid II is also the substrate for the sortase enzymes that catalyze the attachment of surface proteins for incorporation into peptidoglycan [40].

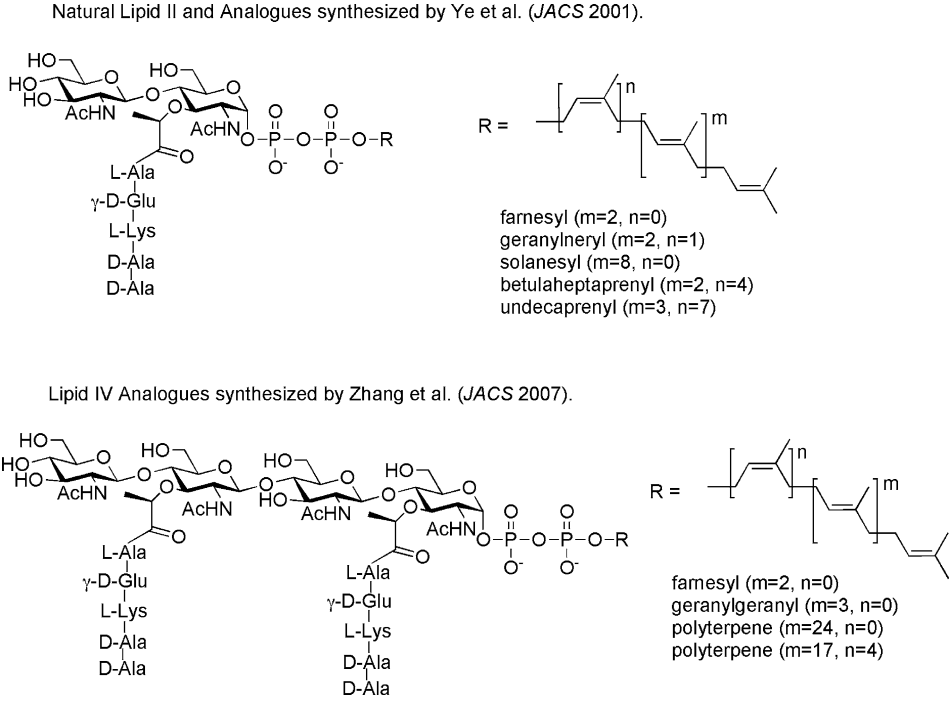
By 2001, mechanistic characterization had been completed for MurA-G from *E. coli*, and crystal structures had been obtained for all of these enzymes except MurC [41,42,43,44,45,46,47,48]. For a detailed discussion of the work completed on this pathway prior to 2000, see the review by van Heijenoort [37]. Since then, *E. coli* MurC and MurG have been crystallized [47,48,49], and homologues of the Mur enzymes from other bacteria have been purified and characterized as well. MraY has also now been purified and characterized [50], a significant accomplishment because this enzyme is an integral membrane protein with many hydrophobic segments, making it particularly hard to purify in significant quantities. However, there is still no crystal structure for this enzyme.

2.1.3 Chemical Synthesis of the Peptidoglycan Monomer Unit

Interest in the synthesis of Lipids I and II was motivated by difficulties in isolating these compounds from natural sources. These intermediates are present in only small amounts in bacterial cells, and the 55-carbon undecaprenyl membrane anchor of the natural compounds aggregates extensively, limiting their aqueous solubility [51,52]. The limited availability of Lipids I and II impeded characterization of MurG and downstream enzymes in the peptidoglycan biosynthetic pathway that use these lipid-linked compounds as substrates. Some characterization of these enzymes was accomplished using assays that measured incorporation of radioactivity from labeled precursors onto uncharacterized membrane acceptors [53], but these assays were not sufficient for detailed mechanistic investigations of individual enzymes. Guan et al. have recently reported a procedure for extracting natural Lipid II from *E. coli* in quantities sufficient for mass spectrometry [54], but extraxtion of larger quantities remains impracticable.

Lipid I and analogues thereof have been prepared using both chemical and enzymatic approaches. In 1997, Auger et al. reported a semi-synthetic approach to a Lipid I analogue containing an unnatural lipid chain that involved enzymatic degradation of UDP-MurNAc-pentapeptide to the corresponding anomeric phosphate followed by coupling to a heptaprenyl lipid phosphate [55]. Almost simultaneously, Men et al. reported the total chemical synthesis of a Lipid I analogue containing a citronellyl carbon chain (◉ *Fig. 4*) [56]. Both the heptaprenyl and citronellyl lipids were found to serve as substrates for *E. coli* MurG, despite substantial changes to the length and structure of the lipid chain. In 2001, two chemical syntheses of undecaprenyl Lipid I were reported [57,58]. Ye et al. synthesized undecaprenyl Lipid I and a range of analogues containing different lipids, and these compounds were tested as substrates for MurG [58]. Substrate analogs containing shorter lipids were shown to be better substrates than the 55-carbon lipid for in vitro studies of MurG, probably as a result of limited solubility of the C_{55} compound [59]. Substrate analogs with short lipid chains were subsequently used to develop several activity assays for MurG, which facilitated the mechanistic and structural characterization of this enzyme [48,60,61].

Syntheses of undecaprenyl Lipid II have also been developed. Ye et. al. took a chemo-enzymatic approach, converting chemically synthesized Lipid I to Lipid II using purified *E. coli* MurG [58]. Since MurG utilizes substrates containing a variety of lipid chains, this route has been used successfully to prepare a range of Lipid II analogues. Two fully chemical routes



◘ Figure 4
Synthetic Lipid II and Lipid IV

to the synthesis of Lipid II have also been developed [62,63,64]. In principle, the chemical approach allows the preparation of wider range of substrate analogues than the chemo-enzymatic route, but extensive protecting group manipulations are required to make the GlcNAc-MurNAc glycosidic linkage. In the chemical literature, the structures of undecaprenyl Lipids I and II are commonly drawn incorrectly, showing 7 *cis* and 3 *trans* isoprene units (and a terminal isoprene). The natural undecaprenyl chain is biosynthesized from farnesyl pyrophosphate, and contains instead 8 *cis* and 2 *trans* isoprene units (see *Fig. 3*) [65,66].

An addition enzymatic synthesis of Lipid II has been reported by Breukink et al., in which the UDP-GlcNAc, UDP-MurNAc pentapeptide, and prenyl phosphate substrates are incubated with bacterial membrane preparations containing MraY and MurG [67]. This method has been used to prepare natural Lipid II, as well as analogues containing polyprenyl chains of varying length. The development of multiple routes to Lipids I and II has made these compounds readily available, facilitating the characterization of the peptidoglycan biosynthetic enzymes as well as screens for inhibitors.

2.2 The Peptidoglycan Matrix

2.2.1 Structure of the Peptidoglycan Matrix

Bacterial peptidoglycan has a mesh-like structure consisting of linear glycan chains connected by peptide crossbridges. As mentioned in *Sect. 2.1.1*, the structure of the glycan is remarkably conserved between bacterial species, and consists of alternating *N*-acetylglucosamine and *N*-acetylmuramic acid residues (or in mycobacteria, *N*-glycolylmuramic acid) connected through $\beta(1,4)$ glycosidic linkages [33]. In Gram-negative bacteria, the glycan chains terminate in 1,6-anhydro MurNAc [68]. The peptidoglycan of a given bacterium contains a distribution of chain lengths, with both long and short chains. Glycan chain length distributions are typically measured by isolating bacterial sacculi, using an amidase to break the crosslinks between the glycan strands, and then evaluating the length distribution by separating the strands using C18 chromatography [69]. Using this method, *E. coli* glycan strands were found to have an average length of 21 and a maximum length of 80 disaccharide units, with ~70% of the chains averaging 8.9 disaccharide units, and 30% averaging 45.1 units [69,70]. This method yielded slightly shorter chain lengths for several *S. aureus* strains, with 85–95.5% of the glycans having a calculated average length of 6 disaccharide units, and the remaining 4.5–15% having lengths longer than 23–26 disaccharides [71]. The average chain lengths for a number of bacteria have also been determined using methods that involve reduction of the reducing termini with $NaBH_4$ followed by hydrolysis of the glycans using acid or muramidase, separation of the muropeptides by chromatography, and quantification of the reducing ends. This method cannot be used for *E. coli* because the reducing termini are capped with 1,6-anhydro-MurNAc, but similar methods for chain length analysis have been used where the non-reducing termini are derivatized enzymatically with D-galactose, or the amount of 1,6-anhydro-MurNAc is determined directly. These methods have been used to analyze glycan chains for a handful of bacterial species, and the chain lengths have been found to depend on the bacterial species as well as the growth conditions. Chains from different species were found to average between 10 and 65 disaccharide units, and for a given organism different lengths were calculated when it was grown as rods or spheres [31]. For *E. coli*, different average chain

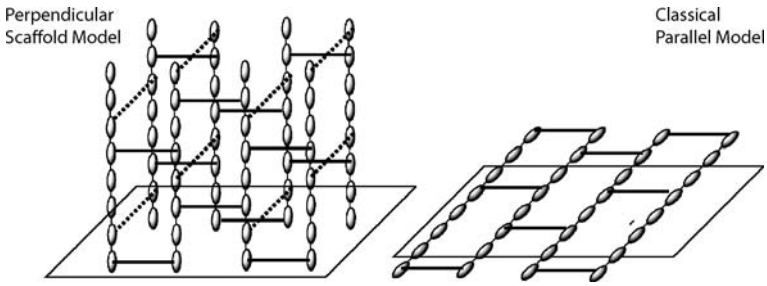


Figure 5
Models for the 3-dimensional arrangement of peptidoglycan

lengths (between ~18 and 33 disaccharide units) were obtained in different media, and in exponential and stationary phase growth [70]. Despite the existence of several different methods for chain length determination, the actual lengths are still subject to much debate. Different analytical methods can yield wildly different estimates of chain length. For example, longer average glycan chain lengths were estimated for *E. coli* strains using the galactose derivatization method (35–50 disaccharides) [72] than by 1,6-anhydro-MurNAc quantification (18–33 disaccharides) [70], and slightly longer average chain lengths (12–16 disaccharide units) were estimated for *S. aureus* using derivitization and digestion than by C18 chromatography of intact chains [31]. For *Bacillus* strains, average chain lengths of 45–140 disaccharides have been determined, with the length estimate varying two-fold depending on the method used for calculation [73]. Differences in measured chain lengths may also be a function of how the sacculi are prepared, since lytic transglycosylases may degrade the glycan chains during sample processing. In addition, the average lengths of glycan chains may vary as a function of their position within the matrix. These factors complicate interpretation of data concerning average chain lengths and variations thereof between organisms.

The degree of crosslinking also varies between bacterial species. In Gram-negative organisms the percentage of crosslinked peptides has been estimated as 40–50% [68,74], while in Gram-positives it is generally higher. For example, in Staphylococci including *S. aureus*, 80–90% of the peptides are crosslinked [75,76]. Gram-positive cells also contain considerably more peptidoglycan than Gram-negatives. The Gram-positive cell wall contains a thick layer of peptidoglycan that accounts for 90% of the weight of the cell wall, and 50% of the weight of the cell. Electron micrographs show the peptidoglycan to be 20–80 nm thick [3]. In contrast, Gram-negative bacteria contain only a thin layer of peptidoglycan surrounding the cytoplasmic membrane, and only 15–20% of the cell wall is peptidoglycan. The thickness of *E. coli* peptidoglycan has been determined using several techniques, including cryotransmission electron microscopy [77], small-angle neutron scattering [78], and atomic force microscopy [79]. All three techniques indicate a maximum thickness of 6 or 7 nm, and neutron scattering suggests a non-uniform peptidoglycan layer, where 75 to 80% of the surface is 2.5 nm thick and the remaining 20 to 25% is about 7 nm thick [78].

Despite the large amount of data available on the physical characteristics and chemical composition of peptidoglycan, the three-dimensional structure of the peptidoglycan matrix is unknown. Computer simulations have been used to produce structural models using experi-

mental data on the conformations of glycan chains, the lengths of the glycan chains, and the degree of crosslinking [70]. The cell wall has classically been pictured with the glycan chains arranged parallel to the membrane. Different models have predicted that the glycan chains are either regularly arranged or randomly coiled, depending on the parameters and constraints used in model construction [70,80]. An alternate "scaffold model", recently proposed by Dmitriev et al., contends that the glycan strands are oriented perpendicular to the membrane [81,82]. These two models are shown in ❯ *Fig. 5*. Dmitriev et al. use computer models of *E. coli* and *S. aureus* peptidoglycan to argue that perpendicular glycan strands would yield a cell wall with elastic properties and porosity more consistent with the observed properties of peptidoglycan. They also contend that the scaffold model more easily explains the synthesis of new peptidoglycan.

Arguments have subsequently been launched both for and against the scaffold model. Notable contributors to this debate include Meroueh et al. [83], who synthesized a tetrasaccharide fragment of peptidoglycan (corresponding to two monomer units), and obtained a solution NMR structure of the fragment. Structural constraints obtained from this model were then used to build a computational 3D model of peptidoglycan in accordance with the scaffold model. Meroueh et al. argue that their model demonstrates the plausibility of the scaffold model because the sizes of the pores in their peptodoglycan model are in accordance with the pore sizes of 50 to 500 Å seen in atomic force microscopic imaging of *S. aureus*.

Arguments against the scaffold model have been reviewed by Vollmer and Holtje [70]. Perhaps the most compelling arguments involve the estimated lengths of the glycan strands. In constructing their models, Dmitriev et al. assume glycan chains that average 8–12 disaccharides in length, much shorter than most of the measurements reported for *E. coli*. These longer glycan chains isolated from *E. coli* would far exceed the estimated width of the *E. coli* peptidoglycan layer, making vertical chain orientation unlikely.

2.2.2 Synthesis of the Peptidoglycan Matrix

Lipid II monomers are assembled into peptidoglycan on the outer face of the cytoplasmic membrane as shown in ❯ *Fig. 6*. The transport of the Lipid II monomers from the inner face to the outer face of the membrane is presumed to involve a transporter protein, but neither the gene nor the protein involved has been identified. The lipid II units are incorporated into peptidoglycan by peptidoglycan glycosyltransferases (PGTs) that form $\beta(1–4)$ glycosidic linkages between the disaccharide units, and transpeptidase (TPase) enzymes that catalyze formation of the peptide cross-bridges [84]. These reactions are believed to occur sequentially, with glycan chain formation preceding crosslinking [85,86]. Bifunctional enzymes that catalyze both transglycosylation and transpeptidation are thought to be responsible for the majority of peptidoglycan biosynthesis [87]. These enzymes have been dubbed "penicillin binding proteins" (PBPs) because they were first identified by their ability to bind radiolabeled penicillin. Bacteria typically contain several high molecular weight (HMW) PBPs, some of which ("Class A") have active PGT domains, and some of which ("Class B") contain only vestigial PGT domains in addition to the active TPase domains [88]. There are also several low molecular weight (LMW) PBPs that have only endopeptidase or carboxypeptidase activities. Some bacterial strains also contain monofunctional peptidoglycan glycosyltransferases (MGTs), which lack TPase domains [89,90,91].

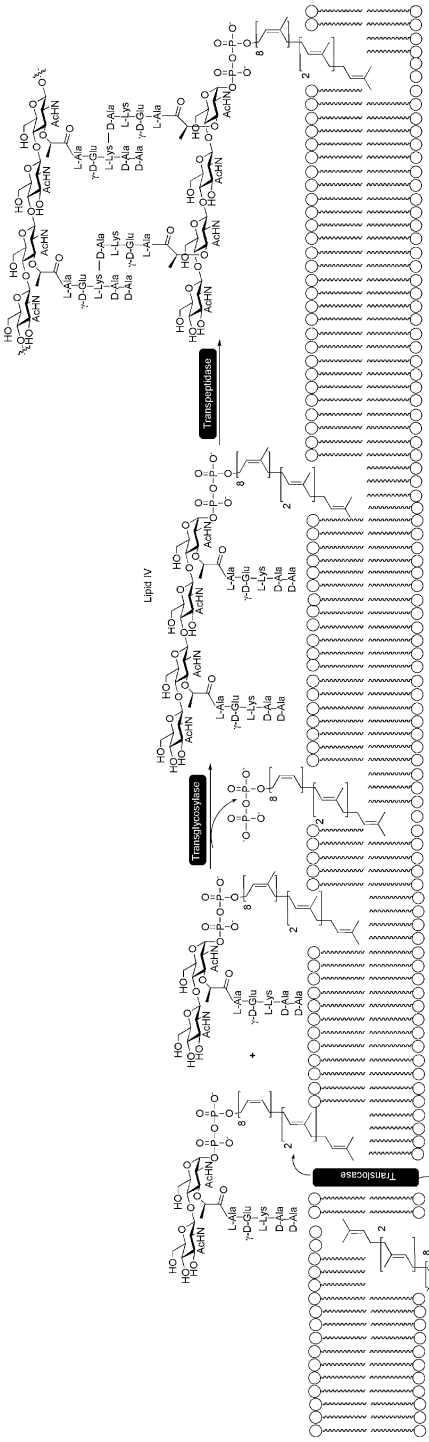


Figure 6
Extracellular steps of peptidoglycan biosynthesis

Most bacteria contain many PBPs, and it is thought that different enzymes play different roles in cell growth and division [33]. For example, *E. coli* contains twelve different PBPs and one monofunctional glycosyltransferase (MGT) [92,93,94,95]. From knock-out studies, two of the HMW PBPs have been shown to play important roles in cell elongation and cell division [94]. *E. coli* also contains four LMW PBPs that are thought to be important for cell shape [96]. Most LMW PBPs have carboxypeptidase rather than transpeptidase activity. Rather than crosslights forming, they moderate the amount of cross-linking found in the cell wall by cleaving the peptide chains, usually at the D-Ala-D-Ala terminal peptide bond. Other bacteria contain different numbers of these enzymes. For example, *S. aureus* contains four different PBPs, as well as two MGTs [97]. Different numbers and types of PBPs and MGTs may be required to maintain the cell sizes and shapes of different bacterial species.

The process of cell wall assembly has proven difficult to unravel. The roles of individual enzymes are largely unknown. Many of the PBPs have been knocked-out in an attempt to determine their specific functions, but these studies have yielded limited information, presumably because the enzymes have redundant functions so that when one enzyme is knocked out, another one can functionally compensate. *E. coli* cells, for example, remain viable even after the inactivation of at least 8 of the 12 PBPs present in this organism [94]. Moreover, it is not clear how cell wall synthesis is coordinated with the processes of cell growth and cell division. To grow and divide, bacteria must break specific bonds in existing peptidoglycan while simultaneously inserting new material so that the integrity of the cell wall is not compromised. This requires coordinating the action of the enzymes that make peptidoglycan with the muramidases, glucosaminidases, amidases, endopeptidases, and carboxypeptidases that specifically hydrolyze the bonds in existing peptidoglycan [92].

2.2.3 The Coordination of Peptidoglycan Biosynthesis

Researchers have used a number of different experimental approaches in an attempt to understand the spatial and temporal organization of cell wall synthesis. In general, experiments have been designed to either identify enzyme complexes, track the cellular localization of individual enzymes or label newly synthesized peptidoglycan. This work has been recently reviewed by Scheffers and Pinho [98].

Several techniques have established the association of enzymes in peptidoglycan assembly and degradation. In vivo crosslinking experiments in bacteria including *E. coli*, *B. subtilis*, and *H. influenzae* have shown that the PBPs are organized into several multienzyme complexes [99,100,101]. Data from co-immunoprecipitation studies support this conclusion [102], and affinity chromatography has additionally shown that the PBPs interact with peptidoglycan hydrolases [103,104,105].

The localization of different PBPs and peptidoglycan hydrolases in several bacterial species has been determined using immunofluorescence and, more recently, using GFP-fusions [106, 107,108,109,110]. From these studies, it appears that in coccoid bacteria cell wall synthesis occurs mainly at the septum, while in rod-shaped bacteria different PBPs are localized to sites of cell division and peripheral cell wall synthesis. Furthermore, the localization of the enzymes seems to be dependent on both interactions with other enzymes and with substrate [98]. These visualization techniques have also helped explain the apparent functional redundancy of the PBPs. For example, in *S. pneumoniae*, different PBPs are specifically localized to sites of

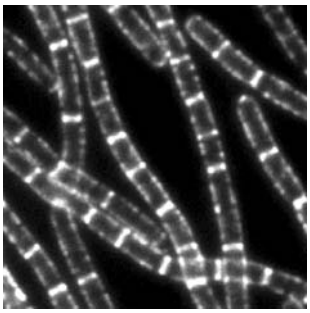


Figure 7
***B. subtilis* cells stained with fluorescent ramoplanin**

septal and peripheral peptidoglycan synthesis. However, if one of the PBPs is knocked out, a different PBP from the same class will compensate by taking over the normal localization of the missing enzyme [111].

A complimentary approach to determining localization of the peptidoglycan machinery is to determine the location of newly synthesized peptidoglycan. Early experiments sought to do this with pulse-chase experiments, to watch time-dependent incorporation of labeled precursors [112,113,114]. More recently, fluorescence microscopy has been used to image peptidoglycan synthesis in Gram-positive bacteria by using fluorescent analogs of certain antibiotics to label extracellular peptidoglycan precursors. This method was first demonstrated by Daniel and Errington, who used fluorescent vancomycin to suggest that peptidoglycan is synthesized in a helical pattern along the walls of *B. subtilis* cells [115]. Vancomycin recognizes the terminal D-Ala-D-Ala in the peptapeptide chain of Lipid II and un-crosslinked peptidoglycan (see ❯ *Sect. 2.3*). Tiyanont et al. subsequently used a fluorescent analog of the antibiotic ramoplanin to assess the roles of specific cytoskeletal proteins in directing the helical pattern of peptidoglycan synthesis (❯ *Fig. 7*)[116].

2.2.4 Kinetic and Structural Studies of the PBP Enzymes

Structural and mechanistic investigations of the PBPs face the same obstacles described in ❯ *Sect. 2.1.2* and ❯ *Sect. 2.1.3* for MurG and MraY. The PBPs are anchored to the membrane by a hydrophobic transmembrane domain, making them difficult to solubilize in active form, and the initial substrate of these enzymes is the membrane-anchored Lipid II produced by MurG. The LMW PBPs and the class B HMW PBPs have proven much easier to study than the class A HMW bifunctional enzymes. Soluble, active constructs of LMW and class B enzymes have been obtained in a straightforward manner by removal of the transmembrane domains. Moreover, many of these enzymes were found to accept structurally simple substrates, facilitating kinetic analysis of hydrolysis. For example, kinetic studies of the class B enzyme *Streptococcus pneumoniae* PBP2x employed a soluble form of the enzyme with a deleted transmembrane domain, and the thioester substrate *N*-benzoyl-D-Ala-thioacetic acid [117]. Mechanistic studies have now been completed for several LMW and class B PBPs, and a number of crystal structures have been solved for these enzymes, including co-complexes with beta-lactam antibiotics [117,118,119,120,121,122,123,124,125].

Kinetic and structural studies on class A PBPs were slower to emerge due to difficulties in obtaining soluble, active constructs. The first demonstration of PGT activity was reported by the Strominger laboratory in the 1960s [126], but no biochemical characterization of any PGT domains appeared for decades. Kinetic studies of the PGT domains were made possible by the development of preparative routes to Lipid II and analogs, discussed in Sect. 2.1.3. Studies with PBP1b, the major PGT in *E. coli* have shown that this enzyme has more stringent requirements than MurG for the lipid substrate, but compounds with shorter lipid chains still make better substrates for in vitro assays of enzyme activity than the natural undecaprenyl Lipid II substrate [58]. Kinetic characterization has now been reported for several PGTs [127,128,129,130,131,132].

The first crystal structures for the transpeptidase domains of class A PBPs were published in 2005 and 2006 [133,134,135]. Until very recently, there were no crystal structures for any PGT domains. However, in 2007 Yuan et al. reported the 2.1 Å crystal structure of the isolated PGT domain of *Aquifex aeolicus* PBP1A [136]. Simultaneously, Lovering et al. reported a 2.9 Å structures of *S. aureus* PBP2 containing both the PGT and TP domains. A 2.8 Å complex of the *S. aureus* enzyme bound to moenomycin was also reported [137]. Based on these crystal structures and additional biochemical information, each of these groups independently proposed a mechanism for glycan polymerization where lipid II acts as the glycosyl acceptor, and the growing polymer chain is the glycosyl donor. These structures have contributed significantly to the development of a mechanistic picture for peptidoglycan assembly.

2.2.5 Chemical Synthesis of Peptidoglycan Polymer

Access to discrete sections of polymeric peptidoglycan would facilitate studies of both the structure and enzymology of peptidoglycan assembly. As discussed in Sect. 2.2.1, the three-dimensional structure of peptidoglycan is not known. Structural studies have been hindered by difficulties obtaining pure samples for structural studies. As a step toward solving this problem, Meroueh et al. synthesized a tetrasaccharide fragment of peptidoglycan (corresponding to two monomer units), and obtained a solution NMR structure for this compound [83].

Discrete, polymeric peptidoglycan fragments are also needed for studying the PGTs. In the first coupling cycle, two molecules of Lipid II are condensed to form a tetrasaccharide, "Lipid IV", which is elongated to form polymer in subsequent coupling cycles. Access to polymeric substrates like Lipid IV is thus necessary for mechanistic studies of the PGTs and for characterization of inhibitors. Toward this end, Zhang et al. recently reported the synthesis of heptaprenyl Lipid IV (Fig. 4) [138]. Reaction of this substrate with several PGTs was demonstrated.

There have also been some studies addressing the assembly of three-dimensional peptidoglycan. The peptidoglycan cross-bridges of *Staphylococcus aureus*, *Enterococcus faecalis*, and *Enterococcus faecium* consist of the sequences Gly_5, L-Ala_2, and D-Asx, respectively. In one study, the *fmhB*, *femA*, and *femB* genes of *S. aureus* were expressed in *E. faecalis*, leading to the formation of Lipid II with mosaic side chains [139]. It was found that the *E. faecalis* PBPs were able use these chains for peptidoglycan cross-linking.

2.3 Medicine

Peptidoglycan biosynthesis is an important antibiotic target, and there is continuing interest in the development of new antibacterial drugs that target this pathway [2]. Several features make peptidoglycan synthesis a particularly attractive drug target. Peptidoglycan is essential in all bacteria, its structure and biosynthesis are highly conserved, and it contains structures like D-amino acids and *N*-acetylmuramic acid that are not produced by eukaryotes. Drugs that target peptidoglycan biosynthesis therefore have the potential to work against a wide range of bacteria while having few side effects in humans. Many existing classes of antibiotics, including the clinically important penicillins and glycopeptides, target peptidoglycan biosynthesis.

Most of the steps in the peptidoglycan biosynthetic pathway are targeted by one or more natural product inhibitors. Natural products targeting the intracellular steps of peptidoglycan have been reviewed by Bugg and Walsh (see *Fig. 9*) [140]. A few of these are in clinical use, including the broad-spectrum antibiotics D-cycloserine and phosphomycin (*Fig. 8*). D-cycloserine is a cyclic analog of D-alanine that targets alanine racemase and D-Ala-D-Ala ligase. Cycloserine is often administered with other drugs for treatment of TB. Phosphomycin, which is commonly used for treatment of urinary tract infections, inhibits MurA by binding irreversibly to enzyme. Bacitracin, which is used as a topical antibiotic, binds to the undecaprenyl carrier lipid on the outside of the cell membrane, preventing it from being recycled.

Several classes of natural products inhibit MraY, and though none of these are used commercially as antibiotics, there is much interest in developing them for clinical use. Research efforts toward this end, including structure-activity studies and mechanistic investigations, have been recently reviewed [141,142,143]. Several families of nucleoside antibiotics targeting MraY show promise for clinical development, including the caprazamicins, which showed activity against a mouse TB model [144], and the mureidomycins, which protected a mouse model from *P. aeruginosa* infection (*Fig. 8*) [145].

The PBPs, which are located outside the cytoplasmic membrane, may make especially attractive antibiotic targets. This extracellular location is advantageous because potential antibiotics aren't limited to structural types that can penetrate the cell membrane, and there are fewer mechanisms by which resistance can develop to compounds that inhibit extracellular targets. Furthermore, there are no human homologues of these enzymes, decreasing the likelihood of inhibitor toxicity.

PBP inhibitors can be grouped into two classes: substrate binders, which recognize structural elements in Lipid II or nascent peptidoglycan, and enzyme binders, which inhibit transglycosylation or transpeptidation by binding directly to the enzymes. Most of the inhibitors that target the PGTs are substrate binders. These include the glycopeptides vancomycin and teicoplanin, which bind to the terminal D-Ala-D-Ala in the pentapeptide chain of Lipid II or un-crosslinked peptidoglycan [146], lantibiotics such as nisin and mersacidin, which are believed to recognize the diphospho-sugar portion of Lipid II [67,147,148], and the lipoglycodepsipeptide antibiotics ramoplanin and enduracidin, which also recognize the diphospho-sugar portion of Lipid II (*Fig. 9*) [149,150].

Substrate binders are attractive as drugs because resistance to them tends to develop relatively slowly, and many hit more than one cellular target. For example, vancomycin acts by binding to D-Ala-D-Ala, blocking both the transglycosylation and transpeptidation reactions. Certain analogs of vancomycin are also thought to inhibit transglycosylation by binding to

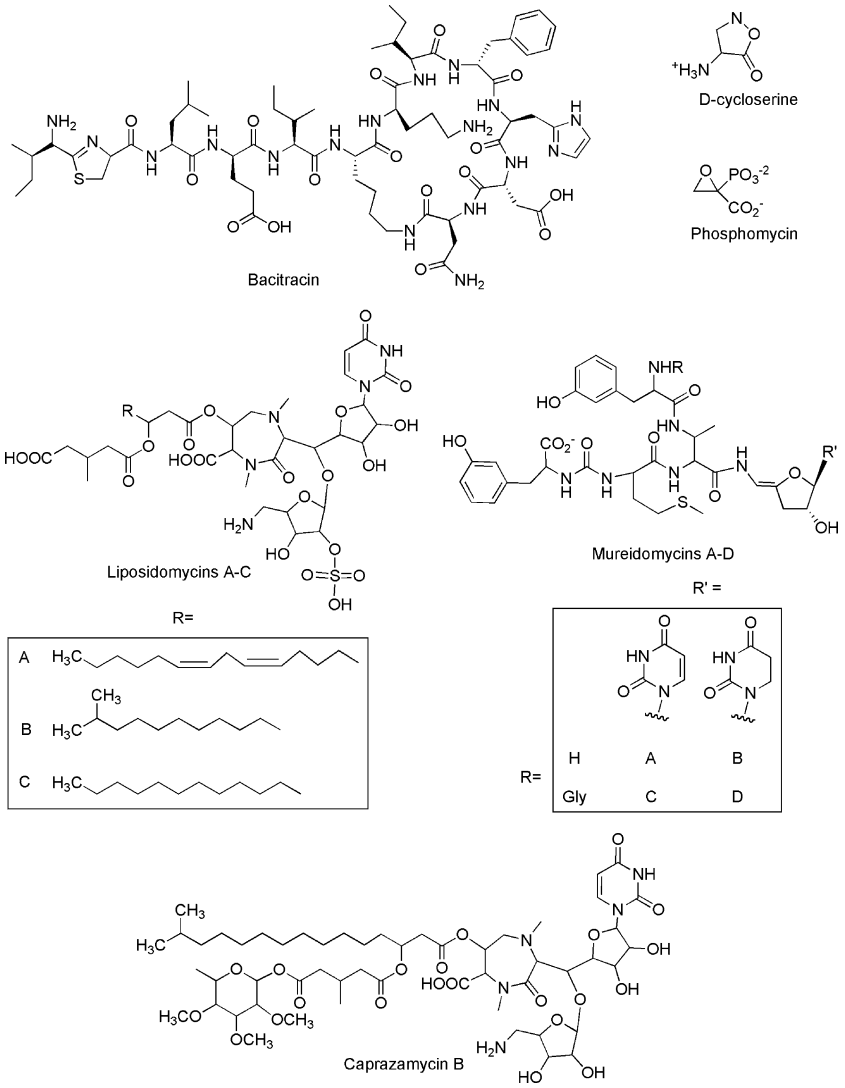


Figure 8
Antibiotics targeting the intracellular steps of peptidoglycan biosynthesis

the enzyme in addition to the substrate [151,152]. Nisin also uses multiple mechanisms: it is thought to block transglycosylation by binding to Lipid II, and to form pores in the bacterial membrane [148,153]. Some of these substrate binders are used clinically; vancomycin and teicoplanin have been in clinical use for many years, and ramoplanin is in clinical trials. Some lantibiotics are also being investigated for possible clinical development [148,149]. Substrate binders do have a few drawbacks: they tend to be large and polar, so they are generally only effective against Gram-positive bacteria because they are unable to penetrate the Gram-

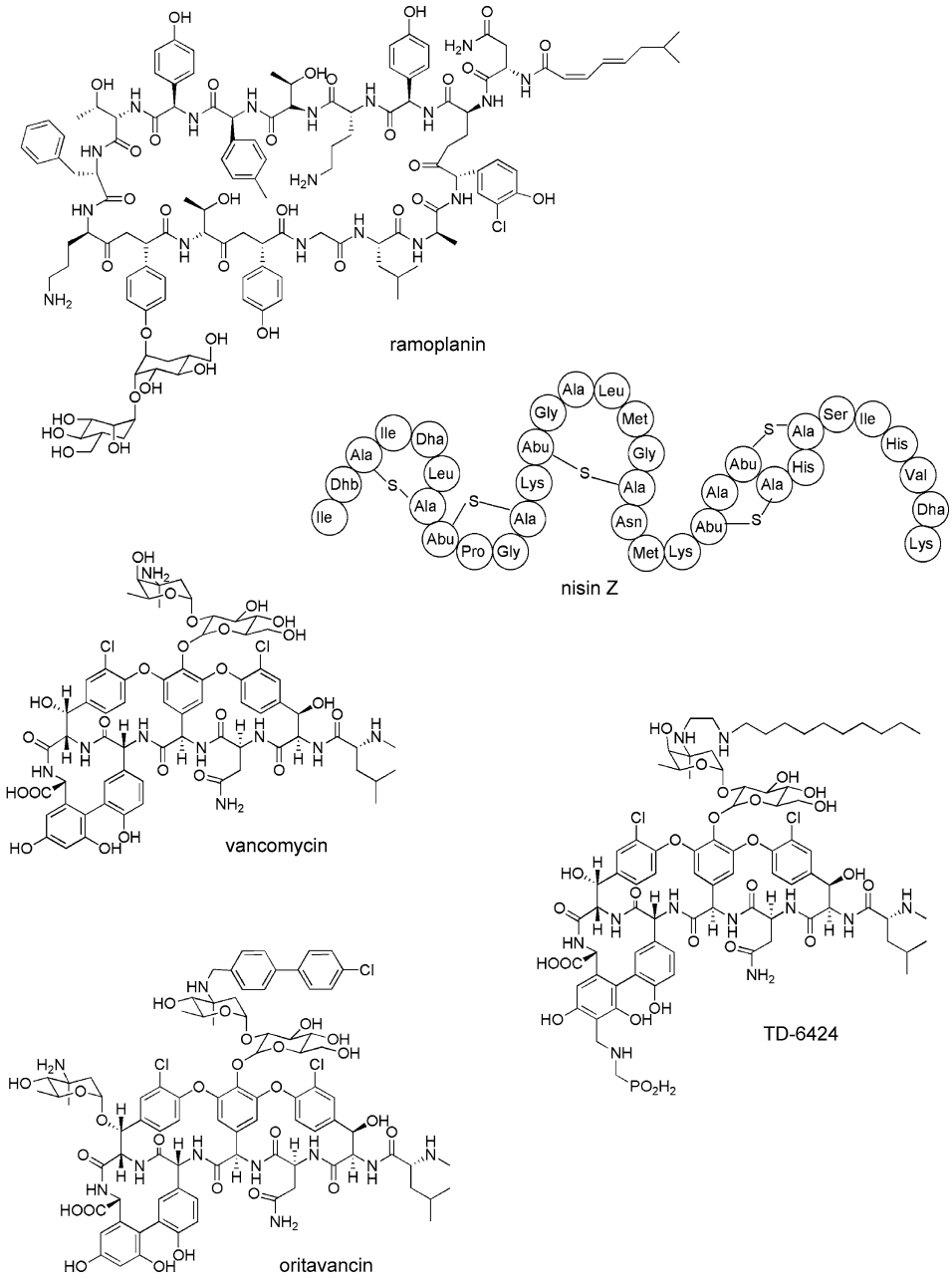


Figure 9
Substrate-binding antibiotics targeting transglycosylation and transpeptidation

negative outer membrane (see *Section 3*). They also require parenteral administration for systemic use because they are poorly absorbed when administered orally. The syntheses and modes of action of PGT inhibitors, including teicoplanin, ramoplanin, and nisin, have been reviewed by Welzel [154]. Synthetic routes to vancomycin have been reviewed by Nicolaou and Snyder [155].

There is only one family of antibiotics known to bind directly to the PGTs, the moenomycins [156]. The most prominent of these is moenomycin A, shown in *Fig. 10*. Moenomycin A is extremely potent, with minimum inhibitory concentrations (MICs) in the range of 0.01 to 0.1 μg mL^{-1} [152], but it is not currently used in humans because of undesirable properties that include a long half-life, toxicity, and poor oral bioavailability. Studies of the moenomycins are aimed at unraveling structure-function relationships in order to design derivatives with more favorable properties. Much work in this area has been carried out by Welzel and coworkers, who have identified features of the molecule that are important for bioactivity and for target interactions. They have suggested that moenomycin is a substrate mimic, though because methods to dissect the peptidoglycan glycosyltransferase mechanism are still lacking, the actual mode of inhibition is unknown. The total synthesis of moenomycin was recently completed, and synthetic methods have been developed to access moenomycin derivatives [157,158]. The gene cluster for moenomycin has also recently been identified, and a subset of these genes have been cloned and heterologously expressed to produce a moenomycin precursor [159]. With the availability of chemical and enzymatic routes to moenomycin derivatives, as well as crystal structures of the PGTs, it should be possible to dissect the mechanism of moenomycin and to develop new antibiotics that target the PGT enzymes.

The transpeptidase enzymes are targeted by the most widely used of all antimicrobial classes, the β-lactams (*Fig. 11*). β-lactams inhibit by binding irreversibly to the enzyme active site. Many different structural classes exist, including the penicillins, cephalosporins, cephamycins, carbapenems, monobactams, and monocarbams [160]. These drugs are widely used because they are highly effective and safe, and many derivatives are available. The original β-lactams were effective only against Gram-positive pathogens, but many years of development have produced drugs that have activity against Gram-negative strains, as well as some broad-spectrum antibiotics. There are two major resistance mechanisms against these drugs: production of β-lactamases that hydrolyze the drugs and production of altered PBPs that do not bind the drugs [161,162]. The latter type of resistance is seen in methicillin resistant *S. aureus* (MRSA). MRSA infections have become a major problem in hospitals, prompting great concern. Historically, the former type of resistance has been circumvented by administering β-lactams in combination with β-lactamase inhibitors. New β-lactams and β-lactamase inhibitors are continuously designed as new β-lactamases continue to emerge [163]. There are not yet any β-lactams in clinical use that can overcome MRSA-type resistance, but a cephalosporin with high affinity for the β-lactam resistant PBP expressed in MRSA, PBP2a, is entering phase III trials (ceftobiprole, *Fig. 11*) [164].

Efforts to design design new inhibitors of the PGTs have focused on making derivatives of known inhibitors. Boger and co-workers have developed a synthetic route to the ramoplanin aglycone, and have made a set of analogues with which to study the mechanism of action of this molecule [165,166,167]. Hydrophobic substituents appended to glycopeptides have been found to increase activity against resistant microorganisms, and there is some evidence that these compounds, in addition to binding peptidoglycan precursors, can also interact with

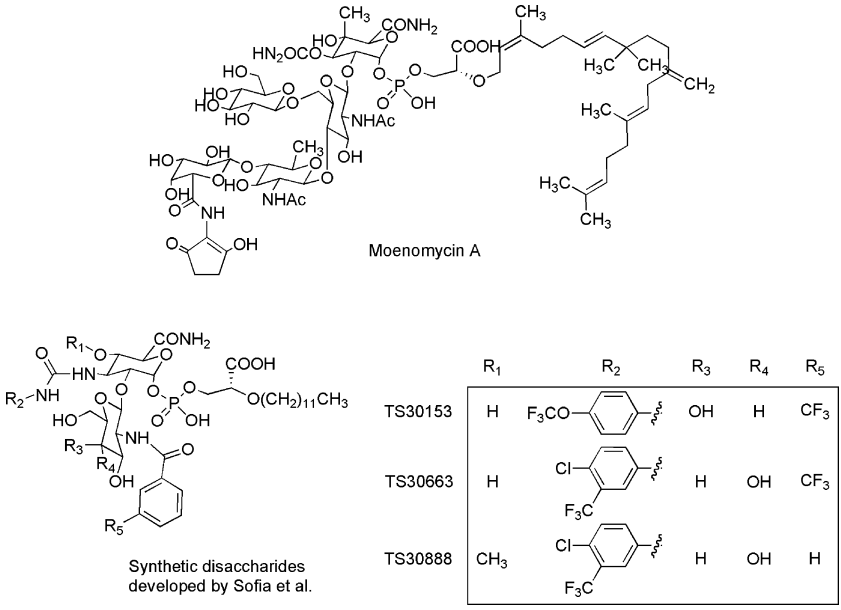


Figure 10
Moenomycin and bioactive disaccharide analogs developed by Sofia et al. [174]

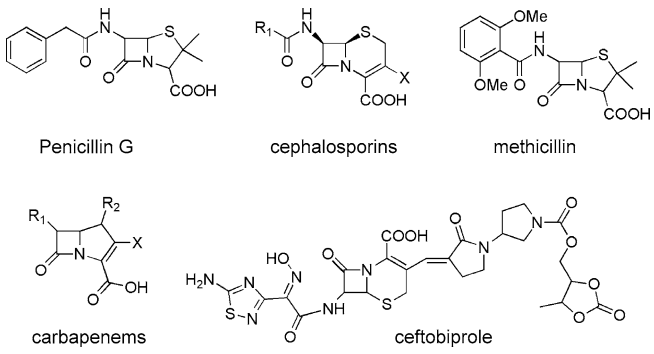


Figure 11
Beta-lactams. Ceftobiprole shows promising activity against MRSA, and is in phase III clinical trials [369]

the PGTs [146]. Four glycopeptide derivatives with hydrophobic substituents are currently in clinical trials: dalbavancin, oritavancin, TD-6424 and telavancin (*Fig. 9*) [146,168]. Several groups are also exploring the use of glycopeptide biosynthetic enzymes for the production of novel derivatives [168,169,170,171,172].

In addition to the aforementioned studies of substrate binding antibiotics, there have been efforts to design inhibitors that bind to the enzyme. These inhibitors have been patterned on either moenomycin or Lipid II, which are among the few compounds known to bind directly to PGTs. For example, Sofia et al. developed a combinatorial library approach to

synthesize disaccharide analogues of moenomycin [173]. 1300 analogs were prepared and tested, and some compounds with activity approaching that of moenomycin were identified (◉ *Fig. 10*) [174,175]. These compounds exhibited interesting differences in in vivo activity against different microorganisms, suggesting structural and functional differences in the PGTs expressed within and among bacteria. It is possible that the various inhibitors targeted different subsets of PGTs. This finding would not be unprecedented: β-lactam antibiotics are known to display differential activities against different PBPs. If moenomycin analogues are similarly able to discriminate amongst PGTs, they could be useful tools to dissect the functions of the different enzymes that participate in cell wall assembly. Two other recent papers describing modest PGT inhibitors based on moenomycin and Lipid II analogs have been reported [176,177].

2.4 High-Throughput Screens

There are ongoing efforts to develop screens for identification of peptidoglycan biosynthesis inhibitors. Many pathway enzymes have no known inhibitors, and identification of small molecules that target these enzymes would provide a starting point for drug design. There is also interest in identifying new classes of inhibitors that hit validated targets, such as non-β-lactamase inhibitors of the transpeptidases. Over the years, many different screens have been used for inhibitor identification. The history of peptidoglycan drug discovery has been recently reviewed by Silver [2].

Most of the in vitro screens used for inhibitor identification have been multi-target screens where several pathway enzymes were incubated together with starting materials, and inhibition of any of the enzymes is measured as a decrease in product formation. These types of assays are efficient because they allow simultaneous screening of multiple targets. Inhibitors of MurA and MurC have successfully been identified in this manner [178,179].

One potential drawback to multi-target screens is that they are unable to distinguish between different modes of inhibition, so are equally likely to identify substrate binders and enzyme binders. The latter are specifically needed to help unravel enzymatic mechanism. Affinity screens can be used to identify compounds that bind to the enzyme. This type of screen has been successfully used to identify inhibitors of MurF that were subsequently used to obtain a crystal structure to aid in drug design [180,181]. A slightly different approach is a displacement screen, which selects for inhibitors that compete with the substrate for binding. This method has been used successfully to identify small-molecule inhibitors of MurG [182,183].

3 The Gram-Negative Outer Membrane

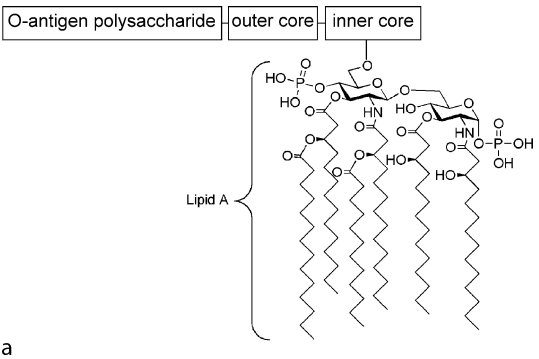
3.1 The Outer Membrane as a Permeability Barrier

The outer membrane (OM) is an asymmetric lipid bilayer, with an inner leaflet composed of phospholipids and an outer leaflet composed of lipopolysaccharide (◉ *Fig. 1*) [6]. The phospholipid composition of the inner leaflet of the OM of *E. coli* is similar to that of the inner membrane: predominantly phosphatidylethanolamine, with smaller amounts of phosphatidyl-

glycerol and cardiolipin [184]. In addition to LPS, the outer membrane also contains lipoproteins and integral OM β-barrel proteins (OMPs). *E. coli* has at least 90 lipoproteins, which protrude into the periplasm and are anchored to the inner leaflet of the outer membrane by their lipid tails [12]. The OMPs adopt a barrel-shape using antiparallel β-strands to span the OM. Many of the OMPs serve as channels that allow essential hydrophilic nutrients to diffuse across the OM [11]. There are also OMPs that function as enzymes. In addition to the components mentioned above, the OM also contains other types of lipid-linked polysaccharides, such as enterobacterial common antigen [13,14]. The OM is also the attachment site for surface organelles, such as pili, which play an important role in pathogenesis [15]. The OM serves as a permeability barrier that prevents toxic compounds, including detergents and antibiotics, from entering the cell [5]. The existence of this barrier enables *E. coli* and other *Enterobacteriaceae* to live in the human intestine, where there are high concentrations of bile salts. It also makes Gram-negative bacteria resistant to many antibiotics [185].

3.1.1 Structure of LPS

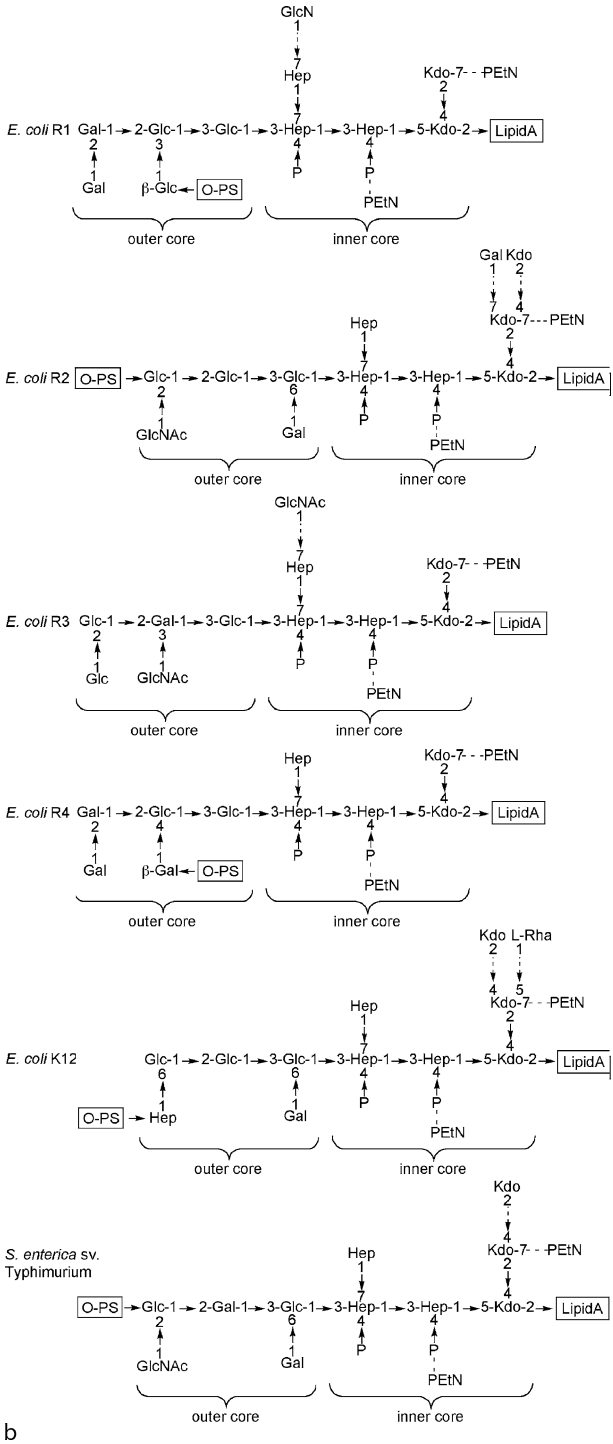
The structure of LPS is can be divided into three parts: the membrane anchored disaccharide lipid A, a short core oligosaccharide bearing multiple phosphoryl groups, and the structurally diverse *O*-antigen polysaccharide [7]. LPS has been best studied in the *Enterobacteriaceae*. Lipid A is the most highly conserved part of LPS. Lipid A typically consists of a β-(1′,6) linked glucosamine disaccharide that is phosphorylated at the 1 and 4′ positions of the disaccharide. The lipid A of *E. coli* and *S. enterica* is acylated by four (*R*)-3-hydroxymyristic acids at the 2, 3, 2′, and 3′ positions, and the fatty acids on the non-reducing glucosamine are further acylated on their hydroxyl groups (*Fig. 12a*). Lipid A is important for the structural integrity of the outer membrane, and a 30% reduction in lipid A renders *E. coli* hypersensitive to antibiotics. The acyl chains attached to lipid A are all unsaturated, allowing them to pack tightly together, and resulting in the very low fluidity (and thus low permeability) of the outer membrane. Lipid A is also responsible for inducing Gram-negative septic shock, and so is often referred to as ‘endotoxin’ [7]. The basic lipid A structure described above may be modified



a

Figure 12

a) The general structure of LPS. b) The five known core types in *E. coli*, and the *Salmonella* core structure seen most commonly in clinical isolates. *Dotted lines* indicate non-stoichiometric structural modifications


GlcN
Hep
Kdo-7- - -PEtN
E. coli R1 Gal-1→2-Glc-1→3-Glc-1→3-Hep-1→3-Hep-1→5-Kdo-2→LipidA
Gal
β-Glc
O-PS
P
PEtN
outer core
inner core
Gal Kdo
Kdo-7---PEtN
E. coli R2 O-PS→Glc-1→2-Glc-1→3-Glc-1→3-Hep-1→3-Hep-1→5-Kdo-2→LipidA
GlcNAc
Gal
outer core
inner core
GlcNAc
E. coli R3 Glc-1→2-Gal-1→3-Glc-1→3-Hep-1→3-Hep-1→5-Kdo-2→LipidA
Glc
GlcNAc
outer core
inner core
E. coli R4 Gal-1→2-Glc-1→4-Glc-1→3-Hep-1→3-Hep-1→5-Kdo-2→LipidA
Gal
β-Gal
O-PS
outer core
inner core
Kdo L-Rha
Kdo-7- - -PEtN
E. coli K12 Glc-1→2-Glc-1→3-Glc-1→3-Hep-1→3-Hep-1→5-Kdo-2→LipidA
O-PS→Hep
Gal
outer core
inner core
Kdo
S. enterica sv. Typhimurium
O-PS→Glc-1→2-Gal-1→3-Glc-1→3-Hep-1→3-Hep-1→5-Kdo-2→LipidA
GlcNAc
Gal
outer core
inner core

b

in response to environmental signals. For example, the charge may be altered by capping the lipid A phosphates with aminoarabinose or 2-aminoethylphosphate, which provides protection against cationic peptides.

The structures of the core oligosaccharides of many different bacteria have been determined [7]. The inner core, which is linked to lipid A, is more highly conserved than the outer core. The inner core is made up of 3-deoxy-D-*manno*-octulosonic acid (Kdo) and, in most bacteria, L-glycero-D-*manno*-heptose (Hep). The minimal structure of LPS required for the survival of *E. coli* is commonly thought to consist of the lipid A and attached Kdo residues [186,187]. However, Meredith et al. recently reported isolation of a viable mutant of *E. coli* devoid of Kdo, with an outer membrane composed of the tetra-acylated lipid A precursor [188]. There are five known core structures for *E. coli* (● *Fig. 12b*), which differ in the outer core. These structures are composed predominantly of α-glucose and α-galactose. The R1 core type is seen most commonly in clinical isolates. One core type predominates among the *Salmonella* (● *Fig. 12b*). The sugars of the inner core are commonly phosphorylated [189]. In *E. coli* and *S. enterica* these phosphoryl substituents are known to be critical for outer membrane stability, and it is thought that adjacent LPS molecules are linked together through divalent cations bound to the phosphoryl groups.

O-antigens are polysaccharides composed of repeating oligosaccharide units. The structures of the repeating units exhibit remarkable diversity in the type, linkage, and substitution of component sugars. More than 170 different *O*-antigens have been identified in *E. coli* [10], and approximately 60 different *O*-antigens are characterized for *S. enterica* [190]. The *O*-antigen structures of *E. coli* have been compiled into a searchable database, available on the web at http://www.casper.organ.su.se/ECODAB/ [191]. A few representative structures are shown in ● *Fig. 13*. The *O*-antigen polymers are usually between one and forty repeating units in length. *O*-antigens are known to play important roles in pathogenesis, by helping the bacteria evade immune attacks from the host and colonize host tissue [192]. *Helicobacter pylori*, for example, evades immune detection by synthesizing *O*-antigens that mimic the Lewis blood group antigens of the host [192]. Mucosal pathogens such as *Neisseria*, *Haemophilus*, and *Bordatella* have LPS lacking *O*-antigen chains. Instead, they synthesize shorter, more highly branched LPS structures referred to as lipooligosaccharides (LOS) (● *Fig. 13*). For more information about LOS structures, see the review by Preston et al. [193]. Pathogenic *Neisseria* and *Haemophilus* species synthesize LOS that mimic human glycolipids and glycosphingolipids. These structures have been shown to be important for both evading the immune system and for adhering to and invading human cells [194,195]. *Campylobacter jejuni* LOS mimic human gangliosides, and it is suspected that development of antibodies against these LOS in individuals infected with *C. jejuni* may lead to GBS, an autoimmune disease that attacks nerve cells [196].

3.1.2 The Outer Membrane as a Permeability Barrier

The OM serves an important barrier function, protecting Gram-negative bacteria against hydrophobic dyes and detergents, as well as hydrophobic antibiotics such as the macrolides, novobiocins, rifamycins, actinomycin D, and fusidic acid that are effective against Gram-positives [185]. The OM is much less permeable than the phospholipid bilayer of the IM because the LPS fatty acyl chains are tightly packed, resulting in low fluidity [5]. The tight

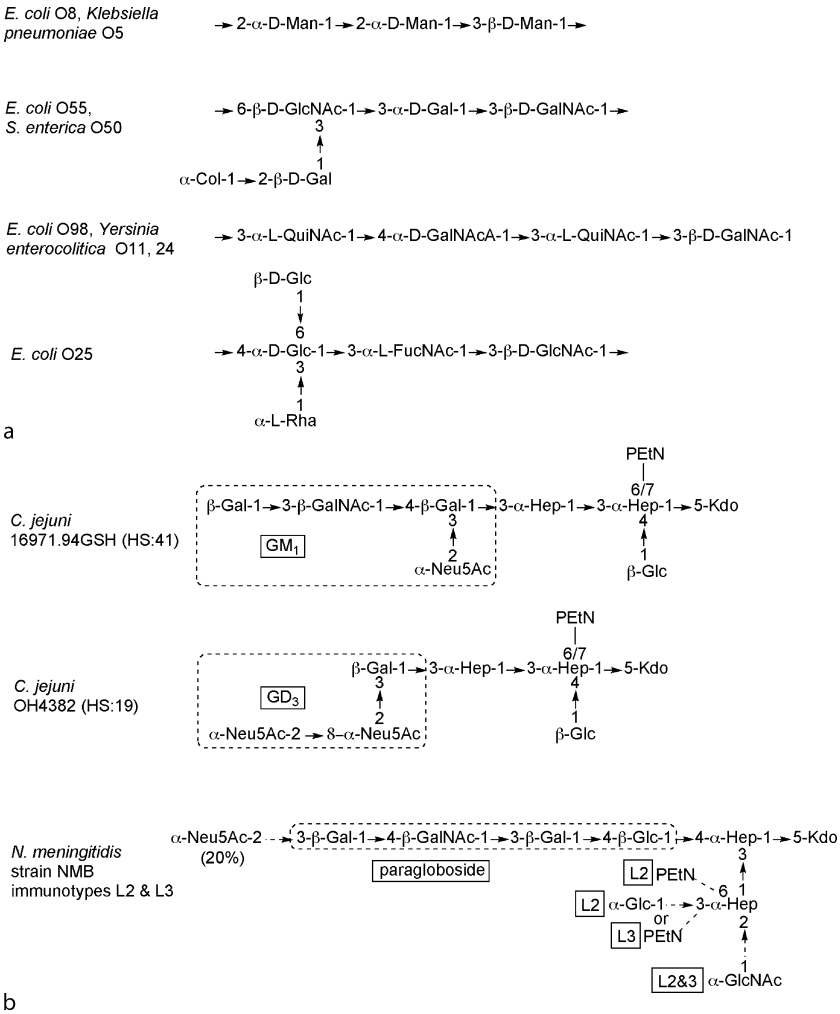


Figure 13

a) Structures of *E. coli O*-antigens. These structures are identical or nearly identical to the *O*-antigens from other bacteria, where noted. b) Lipooligosaccharide structures from *C. jejuni* and *N. meningitidis*. Structures corresponding to the human GM_1 and GD_3 gangliosides and paragloboside antigen are noted

packing is a result of several factors. There is a strong hydrophobic interaction between LPS molecules because LPS has a higher number of fatty acyl chains per molecule than normal glycerophospholipids, and the acyl chains are all saturated. LPS also contains a large number of H-bond donor and acceptor groups both in the sugar and in the lipid part of the molecule, and extensive H-bonds between LPS molecules stabilize the packing of the LPS leaflet and prevent the partitioning of hydrophobic molecules into the outer membrane.

In addition, divalent cations are recruited to neutralize the negatively charged groups in the inner core and lipid A regions, preventing electrostatic repulsion between LPS molecules. Chelators such as EDTA, which remove divalent cations from the cell surface, can cause

a "deep rough" phenotype characterized by hypersensitivity to detergents and antibiotics, and leakage of periplasmic proteins into the culture medium [197]. *E. coli* cells treated with EDTA release about half of their LPS into the medium [198]. Polycations, such as polymyxin B, also disrupt the OM (See ❯ *Sect. 3.4.1*) [199]. Polymyxin B binds to lipid A, primarily through electrostatic interactions.

3.2 Biosynthesis

The components of the outer membrane are synthesized inside the cell, and then transported across the inner membrane and periplasm to the outer membrane for assembly. The assembly process must be coordinated such that the structural integrity of the outer membrane is preserved. The intracellular biosynthesis of cell envelope components is well understood and is the subject of several reviews [7,200].

3.2.1 LPS Biosynthesis

The biosynthesis of lipid A was elucidated largely by work done by the Raetz group. Lipid A-Kdo_2 is synthesized from UDP-GlcNAc by a series of nine enzymes, as shown in ❯ *Fig. 14* [7,201]. Most of these enzymes have been characterized to some degree, and there are crystal structures for several. Inhibition of any of the first seven steps in lipid A biosynthesis results on cell death. Addition of the Kdo sugars necessarily precedes addition of the acyloxy groups in *E. coli* and *S. enterica*, but this not necessary for other Gram-negatives, such as *P. aeruginosa* [7,201]. The unusual Kdo sugar is added from a CMP-Kdo donor, for which the biosynthetic pathway has been determined [186]. The acyloxy groups attached by the final two acyltransferases are not essential, but are important for the biological effects exerted by endotoxin. LpxL and LpxM mutants constructed in pathogens including *E. coli*, *H. influenzae*, and *Salmonella* exhibit decreased virulence and provoke less of an inflammatory response [201]. Such mutants are being explored for use in vaccines [202,203].

After assembly of lipid A-Kdo, a series of glycosyltransferases attaches the core sugars from nucleotide-sugar donors [7]. For *E. coli*, all of the glycosyl and phosphoryl transferase genes involved in the synthesis of the R1 core have been characterized [204,205]. Homologs of the heptosyltransferases responsible for transferring the conserved inner core Hep residues from ADP-Hep have been identified in many bacteria, and most of the enzymes that form the unusual ADP-Hep donor are now known [7,206]. *E. coli* and *S. enterica* strains missing one or both core Hep residues exhibit a "deep rough" phenotype, characterized by hypersensitivity to hydrophobic detergents and antibiotics, and release of periplasmic enzymes from the cell [5,185]. Phosphorylation of the Hep residues was likewise shown to be important for antibiotic and detergent resistance. An *S. enterica* mutant lacking both phosphates was completely non-virulent, and in *P. aeruginosa* this mutant was non-viable [207,208].

The completed lipid A-core oligosaccharide structure is transported across the plasma membrane by the ATP-dependent transporter MsbA. In vitro experiments with *E. coli* MsbA suggest that the ATPase recognizes only fully hexaacylated Kdo_2-lipid A, but the transporter appears to be slightly less selective in vivo [209]. Crystal structures have been reported for MsbA homologues from *E. coli* and *Vibrio cholera* [210,211,212]. However, these structures

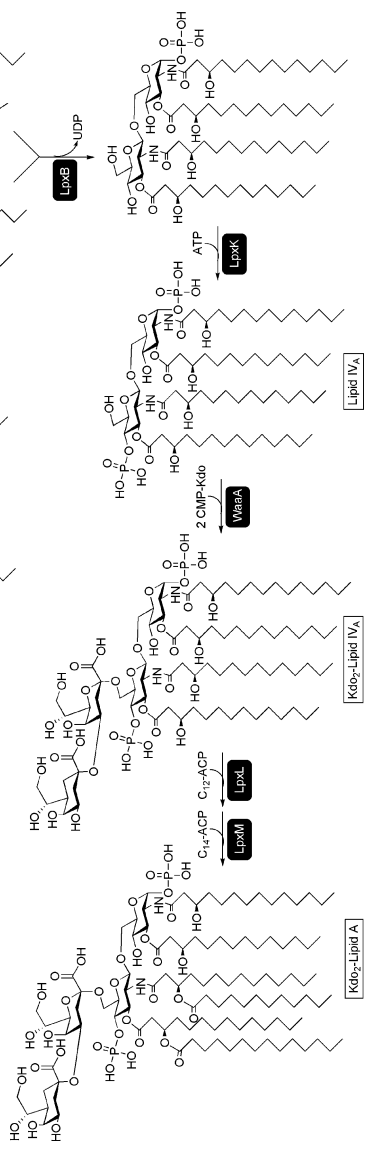


Figure 14
Biosynthesis of Kdo_2-lipidA

were incorrect due to an error in data processing, and were later retracted [213]. MsbA is expected to be structurally homologous to other ABC transporters for which structures exist, including the bacterial multidrug transporter Sav1866 [214].

O-antigens are synthesized on undecaprenyl phosphate carrier lipids anchored in the plasma membrane. There are three different mechanisms for chain assembly [7]. In the Wzy-dependent pathway, individual *O*-antigen repeat units are synthesized on carrier lipids, then flipped across the membrane to the periplasm by Wzx, and polymerized by Wzy. Chain length is regulated by Wzz. The Wxy pathway is also used in capsular polysaccharide biosynthesis (see ❯ *Sect. 5.2.1*). In the ABC transporter-dependent pathway, the entire *O*-antigen chain is synthesized on the cytoplasmic face of the membrane and then flipped to the periplasm by an ABC transporter which comprises Wzm and Wzt. There is also a third, synthase-dependent pathway, where *O*-antigen synthesis occurs concurrently with export [215].

The first step in chain synthesis for each of these pathways involves attachment of an initial sugar residue to the undecaprenyl membrane anchor. This step is relatively highly conserved; all initiating enzymes thus far identified are homologues of either *E. coli* WecA, which adds GlcNAc-1-P to the anchor [216,217], or *S. enterica* WbaP, which attaches Gal-1-P to the carrier lipid [218]. The chain is then elongated by a series of glycosyltransferases utilizing nucleotide-sugar donors [219], and exported by one of the mechanisms described above. The completed *O*-antigen is attached to the lipid A–core at the periplasmic face of the membrane by WaaL ligase. The WaaL proteins from different organisms have low sequence homology, but can functionally compensate for each other and are expected to have similar structures [220,221,222].

As mentioned in ❯ *Sect. 3.1.1*, the lipid A and the core oligosaccharides can undergo structural modifications in response to environmental signals [7]. Modifications of lipid A usually entail removal or modification of the phosphates and fatty acyl chains (❯ *Fig. 15*). Modifications occur at both early and late stages of OM biosynthesis. In *E. coli* grown at low tem-

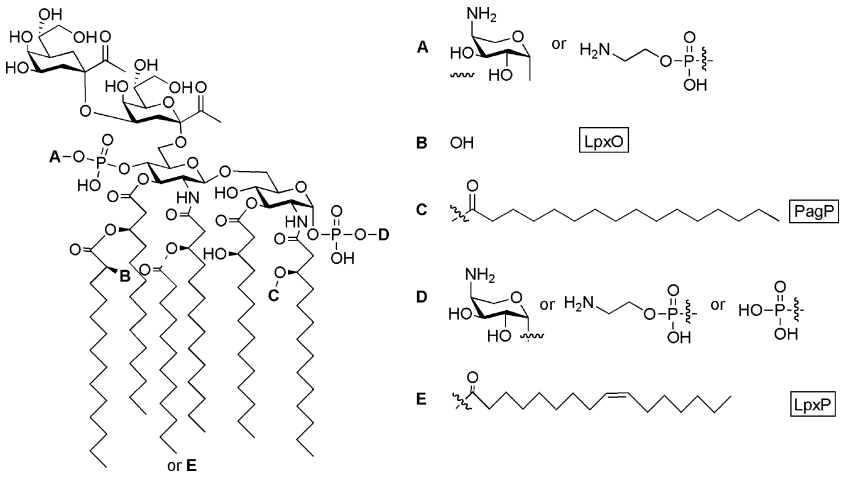


Figure 15
Possible modifications of lipid A in *E. coli* and *S. enterica*

peratures, the biosynthesis of lipid A uses the alternate acyltransferase LpxP, which incorporates an unsaturated palmitoleoyl group into the lipid A chains in place of the saturated acyl group added by LpxL. The unsaturated chains maintain membrane fluidity at low temperatures. A common phosphate modification is the addition of 4-amino-4-deoxy-L-arabinose to lipid A, which confers increased resistance to antimicrobial peptides by altering the charge on lipid A. This occurs at the outer surface of the inner membrane. Lipid A can also be modified after incorporation into the OM. For example, lipid A can be heptaacylated by transfer of palmitate from glycerophospholipids to lipid A by the OM enzyme PagP [223,224]. The extra acyl chain may be needed to increase resistance against the host's immune system during certain stages of infection. The genes and enzymes responsible for many such modifications are known, as are the regulatory pathways that control these modifications [7,201]. These are all potential targets for therapeutic intervention.

3.2.2 LPS Export and Assembly

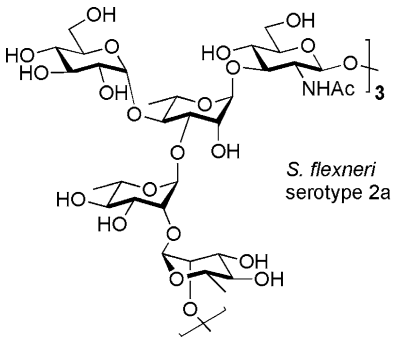
The transport of these components and their incorporation into the outer membrane are subjects of ongoing research. After the *O*-polysaccharide is ligated to the lipid A-core at the periplasmic side of the inner membrane [225], LPS is somehow targeted across the periplasm to the OM. How this occurs is still not well understood [226]. "Zones of adhesion" or "Bayer's Patches" have been observed where the IM is thought to contact the OM [227], and it has been proposed that the OM components reach the OM via these membrane contact sites, without traveling through the periplasm [228]. Consistent with this model, EM shows newly synthesized LPS at regions of membrane adhesion [229].

Recent publications have shed some light on how the LPS components are assembled. Wu et al. recently identified a complex of proteins in *E. coli* that is responsible for assembly of lipopolysaccharide [230]. Two genes, *lptA* and *lptB*, were also recently proposed to be involved in the transport of LPS from the IM to the OM of *E. coli* [231].

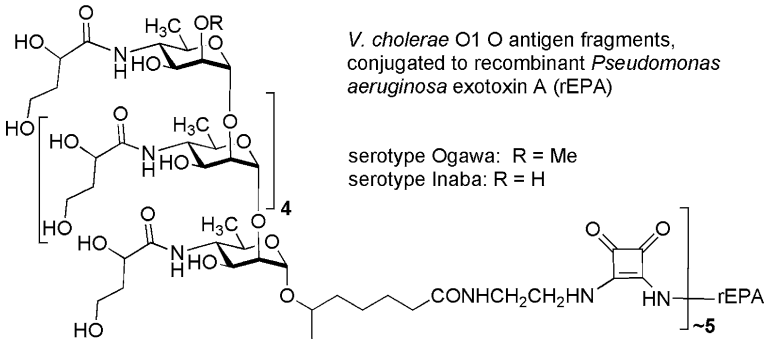
3.3 Chemical Synthesis

Lipid A was first synthesized by Imoto et al. in the mid 1980s [232,233], and the synthetic compound was found to possess all endotoxic activity of LPS. Many derivatives of Lipid A were subsequently synthesized in order to tease apart the importance of different structural elements for endotoxic activity. These investigations have been recently reviewed [234,235]. The tetraacylated lipid A precursor (lipid IV_A, ❯ *Fig. 14*) was found to compete with lipid A for binding to macrophages, but to have endotoxic activity 10^7 less than that of lipid A [235], with the result that it acts as a lipid A antagonist in human cells. In contrast, synthetic Re LPS (Kdo_2-lipid A, ❯ *Fig. 14*) was found to have enhanced immunostimulatory ability compared with lipid A.

Oligosaccharide mimics of the *O*-antigens of many bacterial pathogens, including *E. coli* [236], *Helicobacter pylori* [237], *Shigella flexneri* [238,239] and *Vibrio cholerae* [240,241] have been synthesized and evaluated for use in antibacterial vaccines (❯ *Fig. 16*).



a



b

Figure 16
Synthetic *O*-antigen mimics. a) Vaccination with the *Shigella flexneri* pentasaccharide corresponding to three repeating units induced a strong antibody response in mice [239]. b) Hexasaccharide fragments of *V. cholera O*-antigen conjugated to the rEPA carrier protein used in studies toward cholera vaccine development [240]

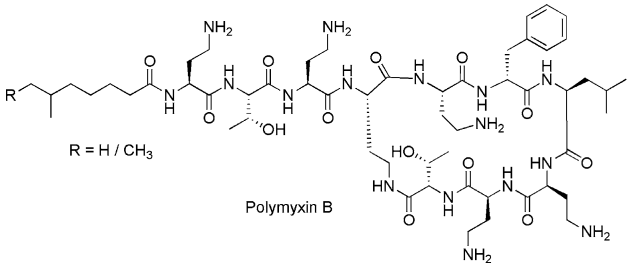
3.4 Medicine

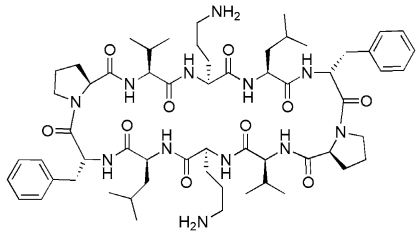
One of the major problems associated with Gram-negative infections is septic shock, which results from an overly vigorous immune response launched as a reaction to LPS recognition. During bacterial death or division, endotoxins are released into the blood stream, where the lipid A moiety is recognized by a circulating protein called the lipopolysaccharide binding protein (LBP) [242]. The LPS-LPB complex binds to the CD14 receptor on the surface of myeloid cells, which subsequently forms an activated complex with the soluble receptor MD-2 and the transmembrane signaling receptor Toll-like receptor-4 (TLR4) [243]. Activation of the receptor cluster results in an inflammatory cascade and the production of reactive oxygen species, which may eventually lead to multiple organ failure. Septic shock can thus result in death even if the bacterial infection is successfully controlled. There is therefore much interest in designing therapeutics that would prevent septic shock by neutralizing LPS.

3.4.1 Drugs that Neutralize LPS

Attempts to design therapeutics for LPS neutralization have largely focused on designing analogs of the naturally occurring molecules that bind LPS. Many cationic peptides with antimicrobial activity also neutralize LPS. These peptides penetrate the Gram-negative cell wall by binding to the cation binding sites on LPS or neutralizing the charge over a patch of the OM, resulting in defects in the LPS. The structures of several of these peptides are shown in ❯ *Fig. 17*. These include Polymyxin B, magainin, cecropin, the bactenecins, gramicidin S, and the defensins. Efforts toward peptide-based endotoxin neutralization have been recently reviewed [244].

The gold standard for anti-endotoxin peptides is Polymyxin B (PMB), a small cyclic lipopeptide (❯ *Fig. 17*). It neutralizes LPS very efficiently, and has been shown to employ a two-step binding process where it first attaches to the LPS by electrostatic interactions and then inserts its acyl chain into the lipid layer. PMB is used only topically because it is neurotoxic. However, there have been many attempts to design analogs with lower toxicity [245,246,247]. Synthetic anti-endotoxin peptides (SAEP) based on Polymyxin B have been engineered to have low toxicity while retaining the in vitro and in vivo detoxification ability of PMB. Furthermore,





cecropin	KWKLFKKIEKVGQNIRDGIIKAGPAVAVVGQATQIAK
BPI (86-99)	KISGKWKAQKRFLK (smallest BPI -derived peptide found to bind LPS)
magainin	GIGKFLHSAGKFGKAFVGEIMKS
Bac7(1-35)	RRIRPRPPRLPRPRPRPLPFPRPGPRPIPRPLPFP

Figure 17
Peptide antibiotics that neutralize LPS

a vaccine against *Neisseria meningitidis* LPS which uses purified LPS from pathogenic strains detoxified by complexation to SAEP has entered Phase I clinical trials.
Some bactenecin derivatives have also shown good activity in vitro and in vivo. In particular, the proline-rich synthetic bactenecin Bac7(1–35) proved to be as effective as PMB at reducing plasma endotoxin levels and decreasing mortality in endotoxin-challenged rats [248]. Some success has also been achieved with derivatives of cecropin, which binds lipid A, and magainin, which inserts itself into the OM as an alpha-helix, disrupting the arrangement of fatty acid chains. Hybrids of cecropin and the potent anti-microbial peptide mellitin have been shown to bind LPS and inhibit cytokine production as well as PMB, and novel analogs of magainin with improved amphiphilicity have exhibited affinities for LPS close to that of PMB [244].
Other synthetic LPS binding peptides have been based on fragments of larger proteins that bind LPS, such as serum amyloid P component (SAP), human neutrophil bactericidal/permeability-increasing protein (BPI), Limulus anti-LPS factor (LALF), lactoferrin, and histones [244]. BPI is an extracellular endotoxin receptor that neutralizes LPS in vivo by binding to it and forming large aggregates. Several synthetic peptides derived from BPI have shown good microbicidal and LPS neutralizing activity, and a recombinant protein derived from BPI provided protection meningococcal sepsis in a Phase III clinical trial [249]. Peptides derived from SAP and lactoferrin have also shown good in vitro and in vivo activity.

3.4.2 Antibacterial Vaccines

A common method used in vaccine development is to link polysaccharide antigens to immunogenic peptide "adjuvants" such as tetanus or diphtheria toxoid. Vaccination with such glycoconjugates results in the activation of T-cell dependent immunological memory. LPS-based glycoconjugate vaccines are currently in various stages of development for several bacteria, including *E. coli*, *Francisella tularensis*, *Pseudomonas aeruginosa*, *Salmonella*, *Shigella*, and *Vibrio cholerae* [26].
There is still no LPS-based vaccine in clinical use, but a *Shigella* vaccine based on detoxified LPS from *Shigella sonnei* was used in a trial by the Israeli Army several years ago, with the result that the infection rate was reduced by 75% in the vaccinated group [250]. A vaccine against *E. coli* using detoxified *E. coli* J5 lipopolysaccharide complexed to an group B meningococcal outer membrane protein (OMP) has also fared well in Phase I clinical trials [251]. Several synthetic derivatives of LPS components have also shown promise as vaccine candidates, including a synthetic triacylated lipid A derivative [252] and oligosaccharide mimics of an *O*-antigen from *Shigella flexneri* [239].

3.4.3 LPS in Cancer Treatment

LPS has well-documented antitumor effects, and cancer treatments employing LPS have been in development for quite some time [253]. The first cancer treatment investigations using LPS were clinical trials performed by Coley in 1898 [254]. Lipid A was shown to be responsible for the antitumor activity of LPS, and trials in animal models of cancer have generally shown that administration of lipid A increases survival and slows tumor growth. The mechanisms by which lipid A exerts its antitumor effects have largely been elucidated [253], though an in-depth discussion is beyond the scope of this review. Briefly, lipid A activates signaling cascades

in the host, leading to the production of cytokines that are either cytotoxic or stimulate the immune system. LPS and lipid A have been tested in Phase I and II clinical trials in cancer patients, and both were found to cause disease stabilization and cytokine production [253]. Lipid A has also been mixed with tumor cell homogenates or cell-surface antigens to create vaccines targeted to particular tumors. In these vaccine preparations, the lipid A acts as an adjuvant to promote the development of an immune response to the tumor antigens. Phase I and II trials with adjuvant vaccines look promising, and antitumor effects are correlated with an antibody response or with cell-mediated cytotoxicity [253].

Thus, the potent immunostimulatory ability of LPS results not only in its severe toxicity, but also in its potential utility as a therapeutic agent. The interactions between LPS and the immune system are complex, and there is continued research aimed at unraveling the signaling pathways responsible for the inflammatory and anti-tumor properties of LPS.

4 Cell Wall Components in the Corynebacterineae

The Corynebacterineae, or mycolata, are a subset of actinomycetes including the genera *Corynebacterium*, *Dietzia*, *Rhodococcus*, *Gordonia*, *Mycobacterium*, *Nocardia*, *Skermania*, *Tsukamurella*, *Turicella and Williamsia*. The best-known members of this group are the notorious pathogens *Mycobacterium tuberculosis*, *Mycobacterium leprae*, and *Corynebacterium diphtheriae*, the causative agents of tuberculosis, leprosy and diphtheria, respectively. Though they are Gram-positive, lacking an outer membrane, they possess a multi-layered cell wall containing an outer lipid bilayer that is functionally similar to the Gram-negative outer membrane. This bilayer structure is very dense in some species, creating an almost impenetrable barrier. For pathogens such as *Mycobacterium tuberculosis*, and *Mycobacterium leprae*, the result is a cell wall that is impermeable to many antibiotics, making infections caused by these organisms particularly difficult to treat [22]. Many other mycolata, including species belonging to the *Corynebacterium*, *Mycobacterium*, and *Nocardia* genera, cause infections primarily in immunocompromised individuals but are equally resistant to drug treatment.

In the previous edition of this book, Todd Lowary exhaustively reviewed what was known to date about the structure, function and biosynthesis of the mycobacterial cell wall polysaccharides, as well as chemical syntheses of substrates and inhibitors for the biosynthetic enzymes [21]. Cell wall biosynthesis remains better studied in mycobacteria than in the other mycolata because of the obvious medical significance of these bacteria. The purpose of this section is therefore to highlight more recent developments in the study of the mycobacterial cell wall, and to describe some of the variation in cell wall structure seen among different genera within the mycolata.

4.1 Structure of the Mycobacterial Cell Wall Polysaccharides

The basic structure of the mycobacterial cell wall is shown in ❯ *Fig. 1* (❯ *Sect. 1*). The major component of the cell wall is the mycolyl-arabinogalactan-peptidoglycan (mAGP) complex [21]. This consists of an inner peptidoglycan layer attached through a linker unit to an arabinogalactan polysaccharide, which is in turn esterified to mycolic acids. The mycolic acids are branched hydroxy fatty acids that form the inner leaflet of the lipid bilayer. The outer leaflet

is made up of free glycolipids and phospholipids, which also fill any gaps in the inner leaflet. These are the subject of several reviews, and will not be discussed here [20,255]. In addition to mAGP, the mycobacterial cell wall contains several other distinctive polysaccharide components, including lipoarabinomannan (LAM) and lipomannan (LM), glycolipids that are anchored in the cytoplasmic membrane and extend through the cell wall. LAM is the major antigenic component of the mycobacterial cell wall.

4.1.1 Arabinogalactan

The structure of mycobacterial AG has been elucidated by fragmentation using either enzymatic digestion or acid hydrolysis, followed by mass spectrometry (*Fig. 18*) [256]. The arabinogalactan is covalently bound to peptidoglycan through a phosphodisaccharide linker, α-L-Rha*p*-(1→3)-D-GlcNAc-P. This linker unit is highly conserved among all the mycolata. The linear galactan chain, which contains approximately 30 β-D-galactofuranose (Gal*f*) residues with alternating β-(1→5)- and β-(1→6)-linkages, is attached to the linker through C-4 of rhamnose. Approximately three arabinofuranose (Ara*f*) chains are attached to each Gal*f* chain through α(1→5) linkages to C5 of the β(1→6) linked Gal*f* residues. The arabinan chain is composed of α(1→5) linked D-Ara*f* residues with (1→3) linked α-D-Ara*f* branches. The non-reducing terminus of each arabinan is decorated with a [β-D-Ara*f*-(1→2)-α-D-Ara*f*]$_{2-}$3,5-α-D-Ara*f*-(1→5)-α-D-Ara*f* motif. Mycolic acids are anchored to both the terminal β-Ara*f* and the penultimate 2-α-Ara*f*.

The non-reducing terminal regions of the arabinan chains have been described for a number of bacteria. Because of the methods used for digestion, however, the structure of the remainder of the AG complex, including the location of branch points and the locations and types of modifications, remains largely speculative. Recently, a more detailed structural model of mycobacterial arabinan was obtained by digestion with an arabinase able to release large oligoarabinosyl fragments from AG, followed by advanced tandem mass spectrometry [256]. This study confirmed the presence of the Ara_{22} structural motif shown in *Fig. 18a*, and identified a galactosamine substituent on the C2 position of a portion of the internal 3,5-branched Ara residue of the AG of *M. tuberculosis.*

The arabinogalactan structures of other mycolata have been studied in less detail than that of the mycobacteria, but differences are apparent in both the saccharide linkages and in the composition of the non-reducing terminus. The arabinogalactan of *C. diphtheriae* contains roughly the same linkages as that of *M. tuberculosis*, but lacks 3,5-linked Ara*f* residues, suggesting an unbranched arabinan [257]. The AG of *C. diphtheriae* AG also contains a significant amount of mannose, which may form a mannose cap on the arabinan domain. Rhodococcal galactan differs from that of mycobacteria in that it contains (1→3) and (1→2) linkages in addition to (1→5) and (1→6) linkages [258]. The AG chains can also contain a variety of terminal motifs, including a branched tri-arabionsyl unit and an unbranched (1→5) linked linear terminus, and termini capped with mannose residues. Rhodococcal AG may also contain fewer than three arabinans per galactan.

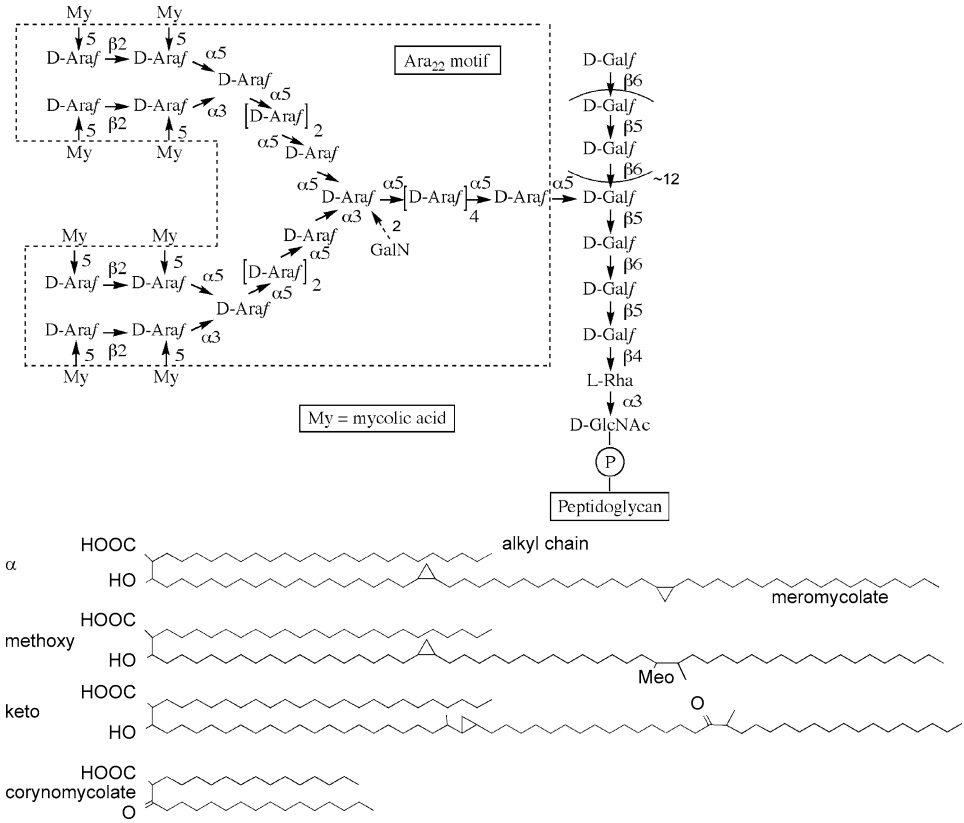


Figure 18
a) Structure of arabinogalactan from *M. tuberculosis*. b) Structures of mycolic acids from *M. tuberculosis*, and a corynomycolic acid

4.1.2 Mycolic Acids

Mycolic acids are α-alkyl branched, β-hydroxy long-chain fatty acids. The alkyl chain is considerably shorter than the main meromycolate chain. The mycolic acids of different bacterial species differ in both the length and structure of the long meromycolate [259]. The mycolic acids from mycobacteria contain 60–90 carbons, and are longer than those from the other mycolata. *Mycobacterium tuberculosis* contains three structural classes of mycolic acids, as shown in ❯ *Fig. 18b* [259]. The α-mycolic acid is the most abundant, with only 10–15% each of the methoxy- and keto-mycolic acids. The cyclopropyl groups of α-mycolic acid are in the *cis, cis* configuration, with variations in the length of the terminal alkyl group and the number of methylene groups between the cyclopropane rings and the carboxyl group. Methoxy- and keto-mycolic acids can contain either *cis*- or *trans*-cyclopropane rings. Corynebacterial corynomycolic acids (❯ *Fig. 18b*) are simple and short, containing only 22–38 carbons [260]. Rhodococcal mycolic acids contain 30–54 carbons, and like their mycobacterial counterparts

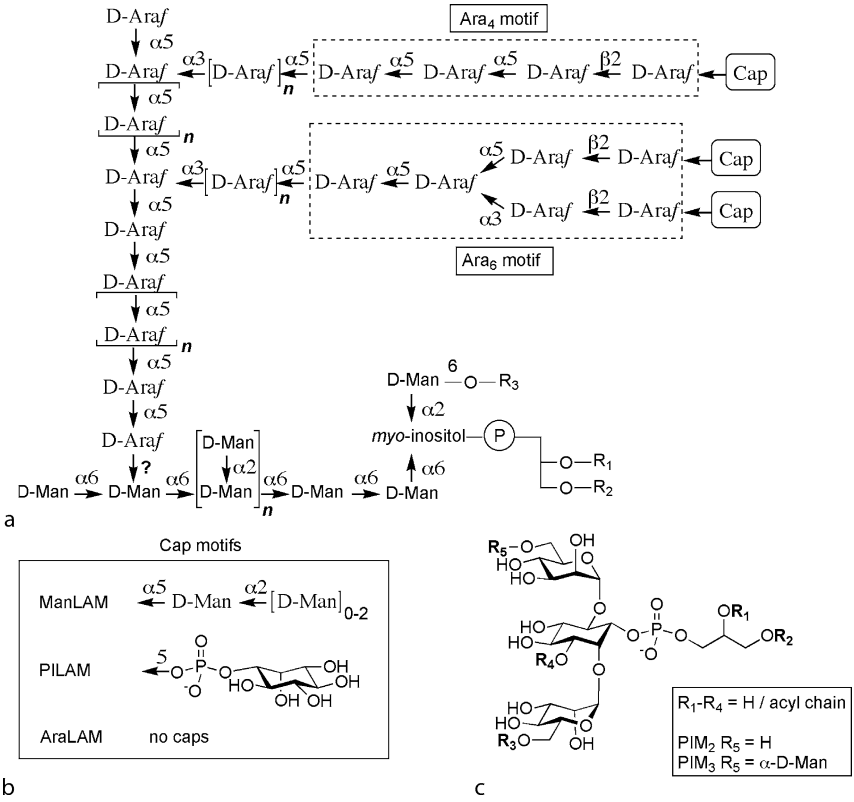


Figure 19

a) Structure of mycobacterial LAM, showing the Ara_4 and Ara_6 terminal arabinan motifs. b) LAM caps. c) Structure of the MPI anchor. The four potential sites of fatty acid attachment are shown

contain a short alkyl chain and a longer meromycolate chain. The rhodococcal meromycolate is unsaturated with up to four double bonds [258,261].

4.1.3 Lipoarabinomannan and Related Lipoglycans

The mycobacterial cell wall contains the essential lipoglycan structures lipoarabinomannan, lipomannan, and phosphatidyl-*myo*-inositol-mannosides (PIMs) [262]. These molecules are attached to the cytoplasmic membrane and extend through the cell wall into the extracellular environment (*Fig. 1*). PIMs, LMs, and LAMs are attached to the cytoplasmic membrane by a conserved anchor, which consists of an sn-glycero-3-phospho-(1-D-*myo*-inositol), with a single α-D-Man*p* linked at C2 of the *myo*-inositol (*Fig. 19c*). The mannan core is attached to C6 of the *myo*-inositol unit. There are four potential sites of fatty acid attachment on the anchor: both free hydroxyls of the glycerol unit, C6 of the Man*p* linked to C2 of *myo*-inositol and C3 of the *myo*-inositol. The fatty acids are most commonly palmitic and tuberculostearic (10-methyl-octadecanoic) acids, and less frequently, stearic acid. Other fatty acids are seen in

some mycobacteria. There is a lot of variation in the number, identity and location of the fatty acids. Different mycobacteria possess different characteristic fatty acid acylation patterns, and these have been determined for a number of species using ^{31}P NMR [262].

In PIMs, 1–6 mannoses are attached to the anchor. In LMs and LAMs, this chain is extended into a 20–25 residue mannan containing a linear chain of α(1,6)-linked Man*p* units, with single Man*p* units attached at C-2 in many species, including *M. tuberculosis* and *M. leprae* (● *Fig. 19a*) [263,264]. In LAMs, the mannan is linked to arabinan chains similar to those in AG, containing a linear $\alpha(1\rightarrow5)$-linked Ara*f* substituted at C3 of some residues with arabinan branches. The arabinan branches contain terminal linear tetra-Ara*f* and branched hexa-Ara*f* motifs (● *Fig. 19a*) [262]. Each LAM contains about 50–80 Ara*f* residues, though the number and length of arabinan chains attached to each mannan is uncertain [263]. The LAMs of many pathogenic species including *M. tuberculosis*, *M. leprae*, *M. avium*, and *M. kansasii* are capped with capped with one, two or three αMan*p* units (● *Fig. 19b*). The mannose capped LAMs (ManLAMs) are important for several aspects of infection (see ● *Sect. 4.4.3*). LAMs of many non-pathogenic strains, in contrast, are capped with inositol phosphate, and these PILAMs are highly inflammatory [264].

LAM-like lipoglycans have been identified in several non-mycobacterial strains, including members of the *Rhodococcus*, *Corynebacterium*, and *Gordonia* genera. These molecules all possess α(1,6)-Man*p* backbones, attached either to a few arabinose residues, or to longer arabinan chains [262]. The immunogenic properties of these LAMs are not well studied, but the LAM from the equine pathogen *Rhodococcus equi* was shown to induce a strong pro-inflammatory response.

4.2 Biosynthesis of Cell Wall Polysaccharides in Mycolata

In the past few years, much has been learned about the biosynthesis of cell wall components in mycolata. Credit for this recent surge of information belongs largely to bacterial genome sequencing projects. Before 2001, only one complete genome sequence, from a *M. tuberculosis* strain, was available for any of the mycolata. Since that time, complete genome sequences have been reported for several additional mycobacterial species including *M. leprae* and several corynebacterial species including *C. diphtheriae*, as well as species belonging to the *Nocardia* and *Rhodococcus* genera.

This section will summarize what is known about the biosynthesis of arabinogalactan and lipoarabinomannan. For a more comprehensive discussion of how glycan assembly is coordinated, and the glycosyltransferases involved, the reader is referred to a recent review by Berg et al. [265].

4.2.1 Biosynthesis of Donor Sugars

The AG and LAM polysaccharides contain unusual sugars, notably D-Ara*f*, D-Gal*f* and Rha. The biosynthetic pathways for the sugar donors thus hold potential as targets for anti-tuberculosis drugs. The glycosylation reactions in the early steps of AG and LAM biosynthesis utilize NDP-sugar donors, while those in the later stages use sugars linked to decaprenyl-phosphate. For most of the nucleotide-sugar donors in *M. tuberculosis*, the biosynthetic pathways have

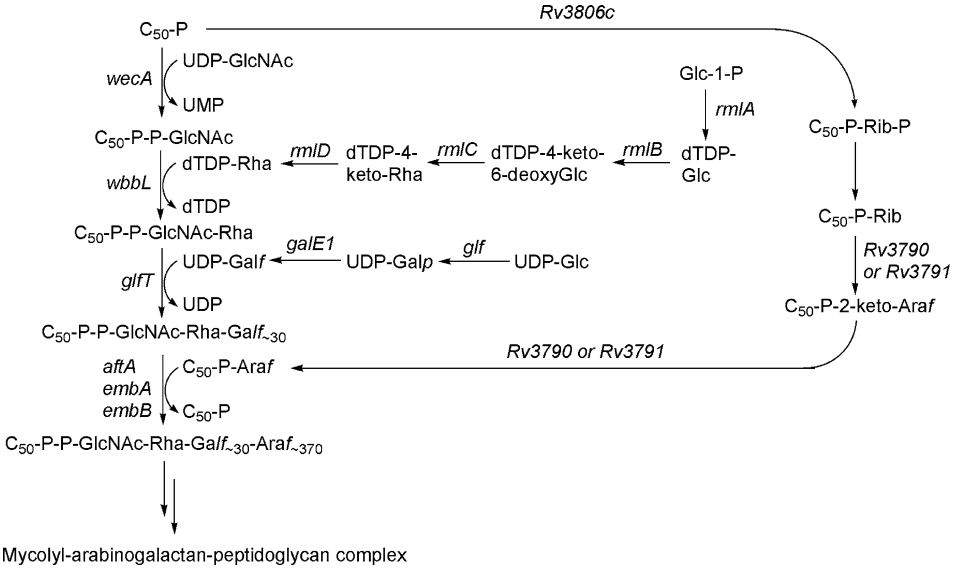


Figure 20
Biosynthesis of arabinogalactan. Biosynthetic genes are indicated, where known

been characterized or the genes involved have been annotated based on homology to other bacteria (*Fig. 20*) [265]. The Rml enzymes responsible for synthesis of dTDP-Rha have been identified and expressed in *E. coli*. The biosynthetic pathways for the polyprenyl-based sugar donors have also been largely characterized [265]. The mannosyltransferase involved in the synthesis of β-D-mannosyl-1-monophosphoryl-decaprenol (C_{50}-P-Man) was identified by homology with eukaryotic dolichol monophosphomannose (DPM) synthase (*Fig. 21*), and subsequently characterized [266]. β-D-arabinofuranosyl-1-monophosphoryl-decaprenol (C_{50}-P-Ara*f*) is the only known donor of Ara*f* in mycobacteria. C_{50}-P-Ara*f* is formed in an unusual process by epimerization of decaprenylphosphoryl ribose (*Fig. 20*), and the enzymes responsible for this process have been purified [267].

4.2.2 Biosynthesis of Arabinogalactan

Information about the biosynthesis of arabinogalactan was originally obtained in the laboratory of Patrick Brennan, in a series of studies where mycobacterial membrane preparations were incubated with radiolabeled sugar donors [32,265]. The biosynthesis was found to proceed in a stepwise fashion as shown in *Fig. 20*, with UDP-GlcNAc, dTDP-Rha, UDP-Gal*p* and C_{50}-P-Ara*f* sequentially incorporated into polymeric glycolipid product. The entire AG polymer is assembled on a decaprenyl membrane anchor. The enzyme responsible for initial attachment of GlcNAc to decaprenyl-P has not been characterized, though a candidate gene has been identified by homology to the gene for WecA, a GlcNAc phosphotransferase involved in lps enterobacterial common antigen synthesis in *E. coli* (see *Sect. 3.2.1*). The linkage unit is completed by transfer of Rha from dTDP-Rha by the enzyme WbbL [268]. Synthesis of the galactan is thought to involve at least two glycosyltransferases. The enzyme Rv3782 was

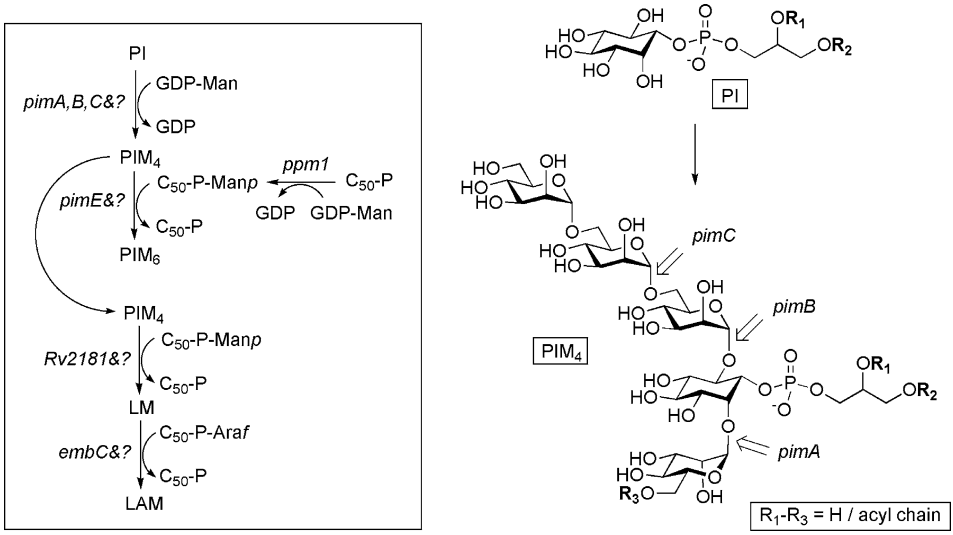


Figure 21
Biosynthesis of PIMs, LMs, and LAMs

shown to initiate galactan synthesis by formation of the Gal*f*-β (1→4)-Rha linkage to the rhamnose [269]. Synthesis of the main galactan chain is catalyzed by the bifunctional enzyme GlfT, which forms both the (1→5) and (1→6) glycosidic linkages [270].
Synthesis of the arabinan domain of AG is less well characterized, but several of the arabinosyltransferases involved have been identified and characterized to different degrees. AftA, which has been purified, was shown in an in vitro assay to add the first Ara*f* residue to the galactan domain of AG [271]. Two proteins, EmbA and EmbB were shown to be to be necessary for formation of the branched Ara_6 motif, as absence of either protein either in vivo or in vitro resulted in a linear glycan (lacking branching at the 3 position of the 3,5-linked Ara*f*) [272]. However, the exact roles of these proteins in glycan formation are unknown.

4.2.3 Biosynthesis of the Mycolic Acids

Mycolic acids are synthesized inside the cell, and then transferred to trehalose-6-phosphate to yield trehalose monomycolate (TMM), which is thought to be moved outside the cell through an ABC transporter. The mycolic acids are then attached to the arabinogalactan complex by a mycolyltransferase. The intracellular synthesis of the mycolic acids in *M. tuberculosis* is fairly well understood, and the mycolyltransferase responsible for peptidoglycan attachment has been identified. The intermediate processing steps, however, remain speculative. Synthesis of the mycolic acids has been recently reviewed by Takayama et al. [273], and will not be discussed in detail here.

4.2.4 Biosynthesis of LAM

LAM, LM, and the PIMs are thought to share a biosynthetic pathway, where PIM_4 is the last common intermediate, and can be extended to either PIM_6 or LM/ LAM (*Fig. 21*) [274].

The first step in the pathway is the synthesis of PI from CDP-diacylglycerol and *myo*-inositol by PgsA (Rv2612c) [275]. *Myo*-inositol is uncommon in prokaryotes, but inositol phosphates and inositol-containing phosphoglycans are important in eukaryotic signaling pathways. The mycobacterial enzymes responsible for *myo*-inositol biosynthesis have been characterized [276,277].

PI is elaborated by a series of mannosyltransferases to make $PIMs_{1-6}$ (*Fig. 21*). Inhibition studies with amphomycin suggest that the first three enzymes that synthesize PIM_1 through PIM_3 use GDP-Man as the glycosyl donor, while those that make $PIMs_{4-6}$ use C_{50}-P-Man [265]. The genes *pimA*, *pimB*, and *pimC* were found encode mannosyltransferases that sequentially attach the first three mannose residues to form $Pims_{1-3}$. While *pimA* is essential, *pimB* and *pimC* are neither essential nor required for PIM/ LM/ LAM synthesis, suggesting the presence of other mannosyltransferases with PimB and PimC activities [265]. Understanding of the biosynthetic pathway after PIM_3 is somewhat vague, but three C_{50}-P-Man-dependent mannosyltransferase genes have been partially characterized: *pimE*, necessary for conversion of PIM_4 to PIM_5 or PIM_6 [278], Rv2181, which attaches α(1,2)-linked Man*P* to the LM backbone [274], and Rv1635c, which attaches the first mannose residue in the cap of ManLAM [279]. An arabinosyltransferase, *embC*, necessary for the synthesis of the LAM arabinan chains, has also been identified [280]. For a more detailed description of LAM biosynthesis, see the review by Nigou et al. [262].

4.3 Chemical Syntheses

4.3.1 Synthesis of Small Oligosaccharides

The enzymes involved in AG and LAM biosynthesis are potential targets for novel therapeutic agents, and there is a lot of interest in developing tools with which to study them. Small polysaccharide analogs of AG and LAM have been explored as potential substrates and inhibitors with which to study the enzymes, and some of these compounds have shown antimicrobial activity. Several groups have demonstrated the utility of small synthetic arabinofuranosyl oligosaccharides as substrates for mycobacterial arabinosyltransferases [281,282,283], and arabinosyltransferase assays using synthetic substrates have been used to probe the functions of several arabinosyltransferases [284]. Synthetic disaccharides have likewise been employed to probe the substrate specificity of mannosyltransferases (*Fig. 22d*) [285] and galactofuranosyltransferase [270].

Disaccharides have also been synthesized as potential cell wall biosynthesis inhibitors, including analogs of the arabinogalactan linkage disaccharide (*Fig. 22e*) [286], and small lipid disaccharides containing arabinose and galactose [287]. Some of these have exhibited moderate in vitro activity against TB.

4.3.2 Synthesis of Larger AG and LAM Fragments

The synthesis of AG or LAM arabinan is complicated by the presence of β-linked arabinofuranosyl moieties. The beta-linkage is difficult to construct because the substituents at C1 and C2 are *cis*. Sterics favor formation of the *trans*-glycoside, and neighboring group participa-

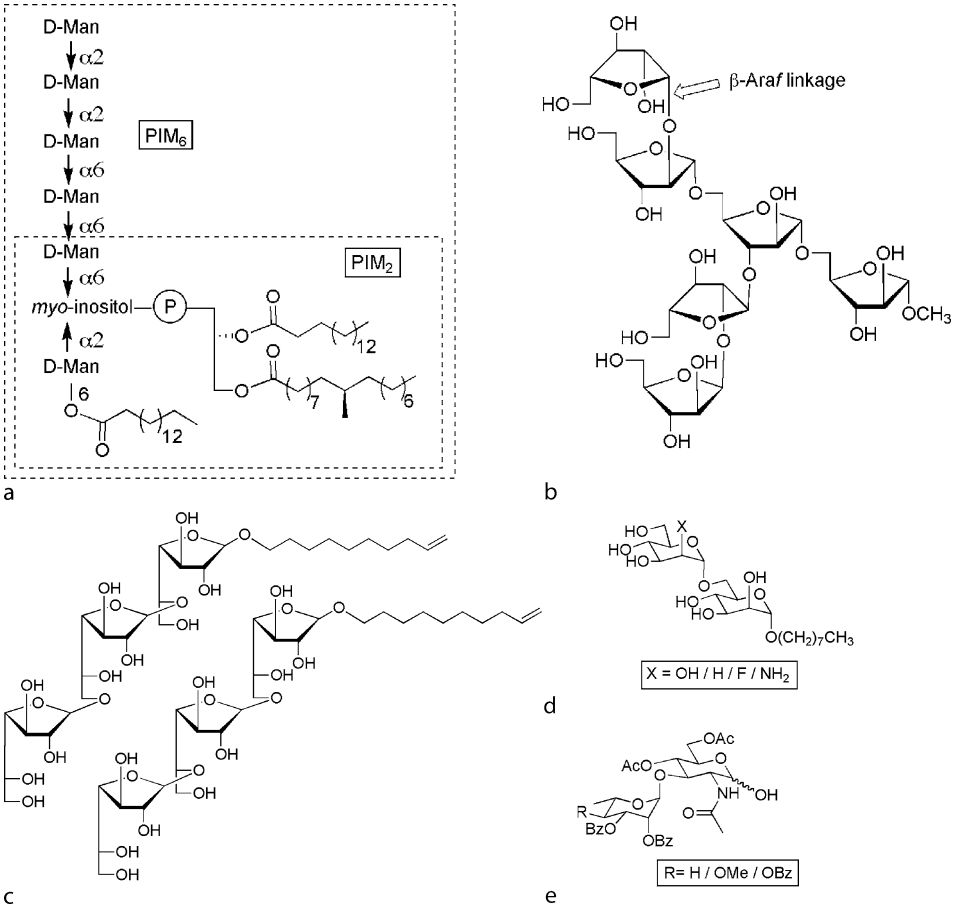


Figure 22
Synthetic disaccharide and oligosaccharide fragments of AG and LAM discussed in the text

tion by a C2 protecting group cannot be used to direct cis-glycosidic bond formation. Much work on arabinan synthesis has been carried out by the Lowary group, and they have explored the use of 2,3-anhydrosugar thioglycosides and sulfoxides for stereoselective construction of β-arabinofuranosyl linkages [288]. Using an epoxide protecting group at C2 and C3, they are able to achieve stereoselective formation of the β-glycosidic bond. Subsequent nucleophilic opening of the epoxide ring gives the arabinofuranoside as the major product. They have used this methodology to construct the arabinofuranosyl hexasaccharide motif (*Fig. 22b*), which is found in AG and LAM.

The synthesis of oligosaccharides containing Gal*f* residues has also been explored. An assay employing synthetic dec-9-enyl glycolipid acceptors was used to elucidate the function of glfT, the bifunctional UDP-galactofuranosyltransferase in galactan biosynthesis [289], and the synthesis of two dec-9-enyl trisaccharide fragments of galactan has been reported (*Fig. 22c*) [290].

Synthetic routes to PIM analogs have also been developed, and the Seeberger group recently reported the first total synthesis of PIM_2 and PIM_6 (*Fig. 22a*) [291]. Shortly thereafter, Ainge et al. reported the first synthesis of PIM_4, and demonstrated that synthetic PIM_2 and PIM_4 were immunogenic, enhancing interferon-gamma production in a transgenic mouse model [292].

4.4 Medicine

4.4.1 Pathogenic Mycolata

Among corynebacterineae, the most notorious human pathogens are *Mycobacterium tuberculosis*, *Corynebacterium diphtheriae* and *Mycobacterium leprae*. Leprosy is no longer considered a major public health problem. Leprosy has been brought under control thanks to free-multi-drug therapy provided by the WHO since 1995, and in 2006 only 200,000 cases were reported worldwide [293]. Diphtheria is now rare in the United States, thanks to rigorous vaccination. However, it continues to be a problem in other parts of the world. Notably, there was a large diphtheria epidemic in the former Soviet Union in the early 1990s, with upwards of 10,000 cases reported [294]. Tuberculosis has proven harder to control than leprosy and diphtheria, and remains a global concern. According to the WHO, approximately a third of the world's population is infected with TB, and more than 8 million people worldwide develop active TB each year [295]. In 2005, the CDC reported 14,097 cases of active TB in the United States [296], and an additional 10 to 15 million people in the US are estimated to have latent TB.

In addition to the three discussed above, there are a number of other pathogens within the mycolata. Infections by non-TB mycobacteria including *M. avium-intracellulare*, *M. marinum* *M. ulcerans*, and *M. kansasii* are rarely seen in healthy individuals, but are common sources of skin infections and lung disease in AIDS patients [297]. In addition, there are increasing reports of human infections caused by multidrug-resistant non-diphtheria Corynebacteria, especially in immunocompromised patients [298,299].

4.4.2 Current TB Treatments

The drug regimens currently used for TB infection require at least 6 months of treatment with multiple drugs, typically a combination of rifampicin, isoniazid, ethambutol, and pyrazinamide [23]. Rifampicin inhibits transcription, while these other drugs target the synthesis of mycobacterial cell wall structures (see *Sect. 4.4.4* and *Fig. 23*). These treatment regimens are harsh, and rifampicin, isoniazid, and pyrazinamide can all cause hepatitis [300]. Treatment is also complicated by the prevalence of TB among HIV-positive individuals. Rifampicin and other rifamycins accelerate the metabolism of many of the protease inhibitors and reverse transcriptase inhibitors used in HIV treatment, leading to sub-therapeutic levels of the HIV drugs [23]. TB is becoming increasingly resistant to the drugs currently used in treatment. In 2005, the CDC reported that 7.3 percent of tuberculosis cases in the US were resistant to isoniazid, and 1.0 percent were resistant to both isoniazid and rifampin [296]. According to the WHO, between 1999 and 2003 10.2% of new TB cases showed resistance to at least one of the

first-line TB drugs, and 4.3% of all TB cases worldwide were multidrug-resistant (MDR-TB), showing resistance to at least isoniazid and rifampicin [295]. Unfortunately, no new anti-TB medications with novel mechanisms of actions have been developed since the introduction of rifampicin in 1966.

There is also a vaccine for TB, the BCG vaccine, made from live attenuated *M. bovis* [301]. The WHO recommends vaccination of infants in parts of the world where TB is common. However, BCG does not protect adults very well against TB, and people who have received the vaccine may show a false positive reaction on the TB skin test. Because of these limitations, BCG is rarely used in the US [302].

4.4.3 Roles of Polysaccharide Structures in Infection and Disease Progression

M. tuberculosis bacteria persist in the harsh environment of the host's alveolar macrophages. These bacteria are able to survive by downregulating the host's immune response. Specifically, *M. tuberculosis* prevents activation of the infected macrophages. Macrophage activation would lead to production of pro-inflammatory cytokines, such as IL-12, IL-18 and tumor necrosis factor α (TNF-α), and to production of interferon-γ producing T cells, allowing the host to combat the infection [303]. Both the LAMs and the mycolic acids of *M. tuberculosis* are able to modulate the host's immune response.

ManLAM and PILAM seem to have opposite effects on the immune response [262,264]. ManLAM was shown to inhibit macrophage activation, the production of the pro-inflammatory cytokines IL-12 and TNF-α and, *M. tuberculosis* -induced macrophage apoptosis. PILAM, in contrast, induces the release of pro-inflammatory cytokines. By suppressing the immune response, ManLAM thus promotes survival of pathogenic mycobacteria, while PILAM promotes killing of non-pathogenic mycobacteria. ManLAM and PILAM were also shown to interact with different cell-surface receptors on phagocytic cells [262,264].

Mycolic acid structure is also important for virulence. Deletion of the proximal cyclopropane ring of the α-mycolate leads to significantly decreased growth in a mouse model of TB infection, as does deletion of the methoxy- and keto-mycolates. Deletion of keto-mycolates was also shown to result in restricted growth in macrophages [273,304].

4.4.4 LAMs and mAG as Drug Targets

Several currently used TB drugs target the biosynthesis of the unique mycobacterial cell wall structures (*Fig. 23*). Ethambutol is an arabinosyltransferase inhibitor [305], and mycolic acid synthesis is targeted by several commonly used antituberculosis drugs, including isoniazid, ethionamide, isoxyl, thiolactomycin, and triclosan. In addition, pyrazinamide was shown to inhibit the synthesis of mycolic acid precursors [306].

There have been efforts to target new inhibitors to most of the steps in AG and LAM biosynthesis, including synthesis of the decaprenylphosphoarabinose and UDP-Gal*f* donor sugars, and the construction of the arabinan and galactan chains. The approaches taken to design inhibitors of LAM and AG biosynthesis were reviewed by Lowary in 2003 [307]. Thus far, the most successful novel inhibitors have been directed at the proven drug targets: the arabinosyltransferases and the mycolic acids. Efforts to synthesize analogs of the diamine antibiotic ethambutol as arabinosyltransferase inhibitors produced SQ109 (*Fig. 23*), which was found

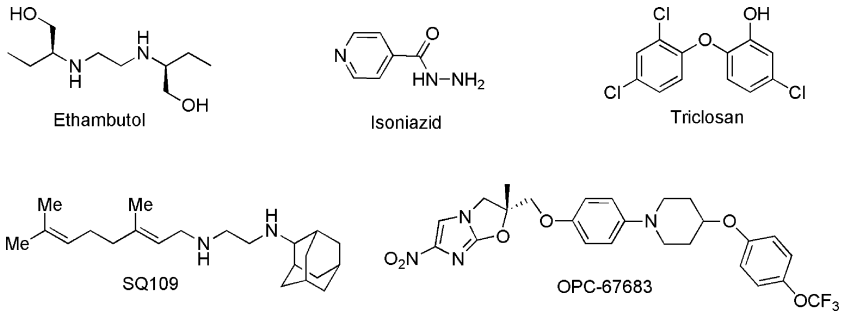


Figure 23
TB drugs that inhibit the arabinosyltransferases or mycolic acid synthesis

to be more effective than ethambutol in treating a mouse model of TB [308]. In addition, screens have been used to identify new structural classes of inhibitors. One promising compound identified in this manner is the nitro-dihydro-imidazooxazole OPC-67683 (*Fig. 23*), recently reported to be a novel highly active inhibitor of mycolic acid biosynthesis [309]. This compound was found to be highly potent against TB, including MDR-TB, with an MIC range of 0.006–0.024 μg mL^{-1}. In combination therapy with rifampicin and pyrazinamide, this compound showed improved efficacy in a mouse model of TB when compared to the standard regimen of rifampicin, isoniazid, ethambutol, and pyrazinamide. Furthermore, the drug was reported to have no effect on liver function.

Thus, though mycobacterial infections are difficult to treat under even the best of circumstances using currently available drugs, novel antibiotics targeting the biosynthesis of the unique mycobacterial cell wall structures show promise for clinical development. Further characterization of the enzymes involved in the biosynthesis of cell wall structures will facilitate the development of such antibiotics. Furthermore, studies of the cell wall biosynthetic pathways will likely uncover new potential targets for drug design efforts.

5 Capsular Antigens

Capsular antigens comprise both capsular polysaccharides, which are attached to the bacterial cell surface, and exocellular polysaccharides, which are excreted into the bacterial environment. These two types of polysaccharides have similar structures and properties [24]. Capsular polysaccharides (CPS) are the major constituents of the extracellular capsule layers that encase many bacteria. The capsules serve as an extra layer of defense and protect the enclosed cells from dehydration. Exocellular polysaccharides form a slime that surrounds the producing cells. They are a crucial component of biofilms, which contribute substantially to the severity and persistence of bacterial infections (see *Sect. 5.4.3*) [310]. Capsular antigens are produced by both Gram-negative and Gram-positive bacteria. Most capsular antigens are highly immunogenic, and vaccination with them produces high antibody titer [311]. Antibodies interact with immunodeterminants (epitopes) of capsular antigens, and bacteria are commonly classified into serotypes possessing different immunodeterminant groups. The capsules of pathogenic

organisms are frequently used for vaccine development [26]. This section will describe the structure and synthesis of capsular polysaccharides from some common pathogens, and discuss their use in vaccines.

5.1 Structure

5.1.1 Capsular Antigen Structure

The structures of capsular antigens are very diverse. The polysaccharides are composed of repeating units, which can be either homo- or heteropolymers containing either branched or unbranched chains (*Fig. 24*) [24]. The chains contain a range of sugars, and can also be substituted with both organic and inorganic molecules. For example, the capsules of pneumococci are built from glucose, galactose, rhamnose, and glucuronic acid and various glycosamines (in particular, *N*-acetyl derivatives of D-glucosamine, D-galactosamine, D-mannosamine, D-fucosamine and 2,4-diamino-2,4,6-trideoxy-D-galactose). In addition, some pneumococcal capsular antigens show structural similarities to teichoic acids, containing residues of ribitol phosphate and glycerol phosphate, and many contain pyruvate residues. Several groups of streptococcal polysaccharides contain hyaluronic acid, while the chains of another group contain terminal sialic acid residues. The *E. coli* capsular antigens can be divided into four groups that share general structural features and biosynthesis [312]. Groups 1 and 4 are acidic polysaccharides resembling *O*-antigens. Group 2 and 3 capsules are found in *E. coli* isolates that cause extraintestinal infections. The structures are diverse, and some contain phosphate residues in their backbone structures, while others resemble vertebrate glycoconjugates. See *Fig. 24*. *S. pneumoniae* and *E. coli* each possess more than 80 different CPS. For an overview of capsular antigen structures, the reader is referred to a recent review by Ovodov [24].

5.1.2 Capsule Structure

Capsular polysaccharides are covalently linked to the cell surface, forming a highly hydrated protective barrier around the cell. In Gram-negative bacteria, CPS are thought to be attached to either phospholipid or lipid A (*Fig. 25*). The *E. coli* capsular antigens of groups 1 and 4 are both found in two forms: either linked to a Lipid A core (K_{LPS}), or as a high molecular weight capsular antigen with an unknown mode of attachment, while groups 2 and 3 are attached to the OM through a diacylglycerolphosphate anchor [312]. The CPS of many Gram-positive pathogens, including *Staphylococcus aureus* [9], *Streptococcus pneumoniae* [313], and group B *Streptococcus* [314], are covalently linked to peptidoglycan (*Fig. 25*). The mode of attachment is not well-documented, but in group B *Streptococcus* the CPS is attached through a phospodiester bind to a linker oligosaccharide, which is in turn attached to the GlcNAc residues in peptidoglycan (*Fig. 25*) [314]. In some Gram-positives, including one serotype of *S. pneumoniae*, CPS may be membrane-associated rather than covalently attached [315].

5.1.3 Biofilm Structure

A biofilm is a community of microorganisms that settles on a surface and is covered by an exopolymer matrix. Cells in biofilms are slow-growing, and exhibit clear phenotypic differences from free-living planktonic cells. Biofilm structure is both complex and dynamic, and

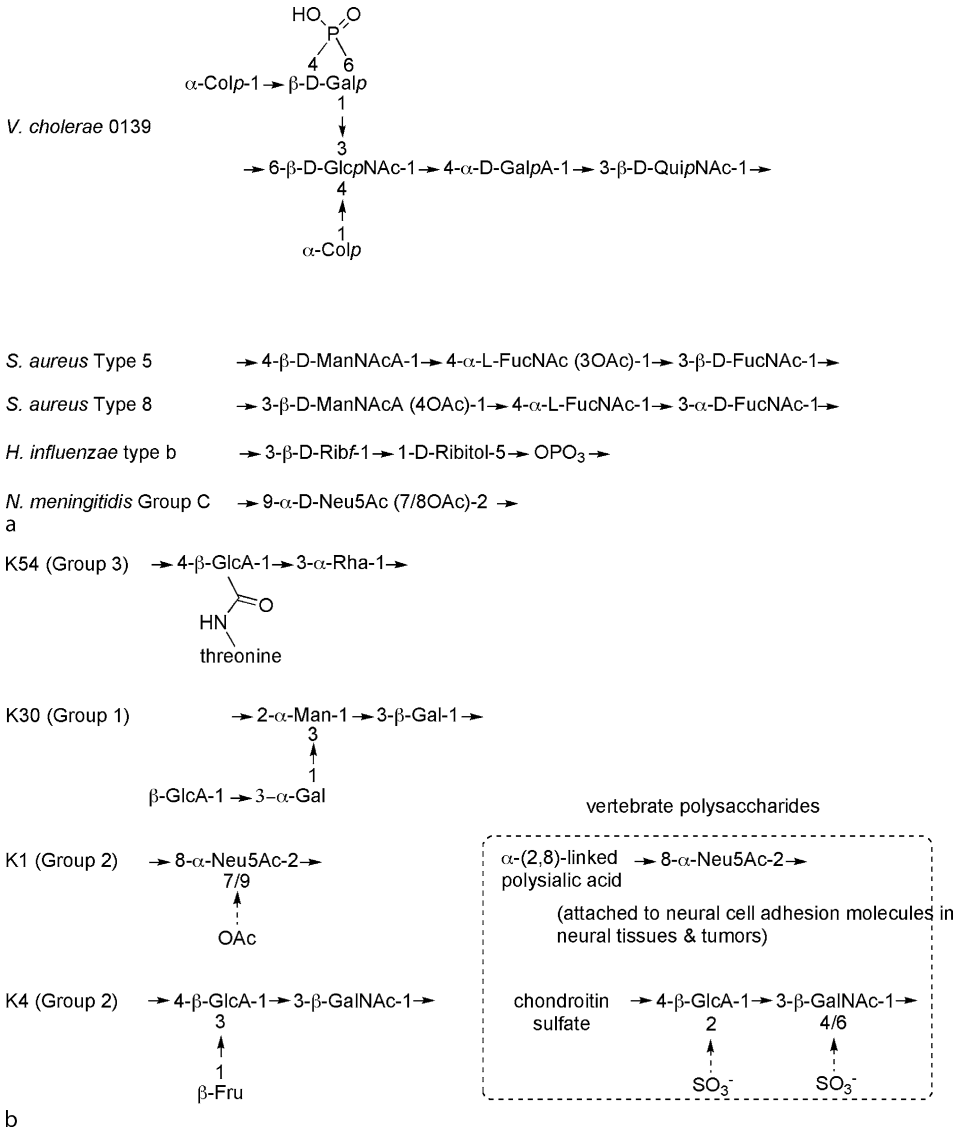


Figure 24
Capsular antigen structures. a) Some CPS currently used in vaccines. *S. aureus* type 5 and 8 are the CPs in StaphVAX [355]. b) Some *E. coli* capsular antigens. For structures that mimic vertebrate glycoconjugates, the corresponding human structures are shown

the composition of biofilms varies greatly depending on the bacterial species as well as on the growth conditions. Biofilm composition and architecture has been reviewed by Sutherland [316]. The cells are encased in a webbing made of exocellular polysaccharides (EPS), and production of EPS is essential for construction of the biofilm matrix [317]. Different bac-

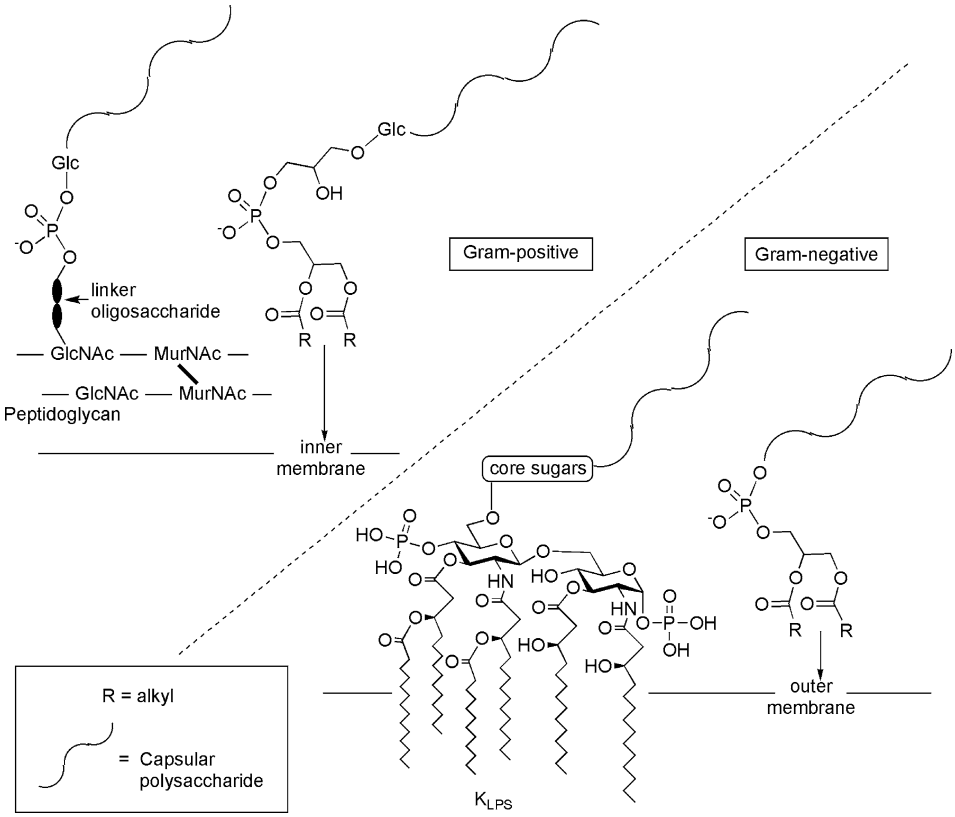


Figure 25
Attachment of capsular polysaccharides to the cell surface. Gram-negative CPS are attached to the outer membrane while Gram-positive CPS are attached to peptidoglycan or anchored in the inner membrane

teria produce a variety of EPS structures, but cellulose and β(1,6)-linked *N*-acetylglucosamine are the most common biofilm components of many bacteria [318]. Cellulose, for example, is a crucial component of biofilms for *Salmonella typhimurium*, *Escherichia coli*, and many other bacteria. β(1,6)-linked *N*-acetylglucosamine (PNAG, also called PIA) is the main component of biofilms produced by most *Staphylococcus epidermidis* and *Staphylococcus aureus* strains, and is thought to be produced by other bacteria including *E. coli* and *Yersinia pestis*. *E. coli* biofilms also commonly contain the extracellular polysaccharide colonic acid.

In addition to extracellular polysaccharides, biofilms contain proteins, DNA and RNA, as well as peptidoglycan, lipids, phospholipids and other cell components. Some of these materials are secreted by the bacteria, while others are products from cell lysis, or environmental material [316,319]. Biofilms are highly hydrated, and up to 97% of the biofilm can be made of water [320]. Confocal scanning laser microscopy has shown that cells within biofilms are aggregated into microcolonies separated by channels that allow the passage of nutrients and waste products [316].

5.2 Biosynthesis

Each polysaccharide structure is synthesized by a distinct set of enzymes, and the biosyntheses of the individual structures will not be discussed here. The biosynthetic gene clusters for many different capsular and extracellular polysaccharides are known, and many of the enzymes involved have been characterized. For a general review of the genetics and biochemistry of CPS production, the reader is referred to a review by Roberts [321]. There are also many excellent reviews discussing capsule formation in individual bacteria, including *H. influenzae* [322], *S. Aureus* [323], *E. coli* [312], and pathogenic streptococci [324,325].

5.2.1 Conserved Features of Capsular Antigen Biosynthesis

Despite the diversity of capsular antigen structures, the genes involved in CPS assembly and export in a number of different bacteria share a high degree of homology, suggesting that the export and assembly processes are conserved. Two different pathways have been described for the assembly and export of CPS in *E. coli*: the Wzy-dependent pathway and the ABC transporter-dependent pathway (*Fig. 26*). Other bacteria are expected to utilize similar pathways based on the presence of homologous gene clusters [321,326,327]. The Wzy-dependent pathway, for example, has been studied in Gram-positive bacteria including streptococci [328], staphylococci and lactococci [329]. The *E. coli* pathways were recently reviewed by Whitfield [312].

E. coli capsule groups 1 and 4 are exported and assembled via the Wzy-dependent pathway, which is also used for LPS biosynthesis (see *Sect. 3.2.1*) [312]. Individual CPS repeat units are assembled on an undeacprenyl phosphate carrier lipid by the sequential action of several glycosyltransferases. The lipid-linked repeat units are then thought to be flipped across the plasma membrane by Wzx, and polymerized by Wzy. New units are added to the reducing end of the growing polymer. At this point, polymerized products are either transferred from the carrier lipid to lipid A-core acceptor by WaaL ligase (forming K_{LPS}, *Fig. 25*), or incorporated into capsules in a process that depends on Wzc, a tyrosine autokinase located in the plasma membrane, Wzb, a tyrosine phosphatase, and Wza, an outer membrane lipoprotein [312].

Based on EM data and the crystal structure of Wza, the Whitfield and Naismith groups have recently proposed a model in which Wza and Wzc form a complex for CPS export that spans the periplasm [330]. Their model is significant because it provides a mechanism by which CPS may move from the periplasm through the outer membrane. The export processes for macromolecules such as LPS, CPS and OM proteins, and coordination of export and biosynthesis are poorly understood.

E. coli capsule groups 2 and 3 are exported and assembled in a completely different fashion than groups 1 and 4 [14,312]. The CPS is fully assembled in the cytoplasm by addition of new units to the non-reducing terminus by processive glycosyltransferases. The polymer is attached to the cytoplasmic membrane by a diacylglycerolphosphate or diacylglycerolphosphate-Kdo anchor (*Fig. 25*). After polymerization, the chain is exported by an ABC transporter (KpsM and KpsT). The CPS is then thought to be shuttled from the IM to the OM by KpsE and exported by the OM protein KpsD, though the roles of these proteins are speculative. See the recent review by Whitfield for further discussion [312].

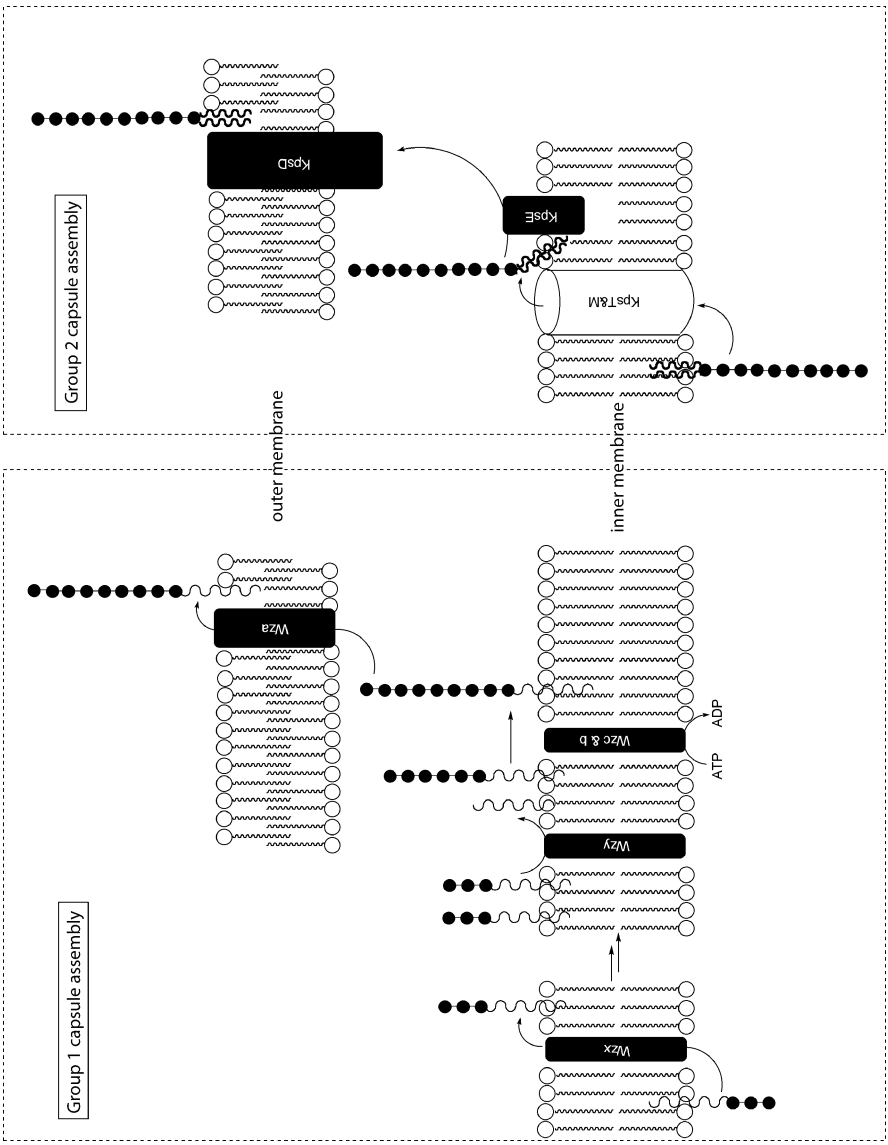


Figure 26
Capsular polysaccharide assembly and export in *E. coli*

5.2.2 Biofilm Assembly

In response to environmental signals, bacteria are able to switch between a free-living planktonic state and one in which they function as part of a multicellular biofilm community. The processes involved in biofilm formation are very complex and not yet well understood. Genetic and molecular approaches to studying biofilm development, as well as the genes and regulatory circuits known to be important for the formation and maturation of biofilms and the return of biofilm microorganisms to planktonic growth have been recently reviewed [331].
Microscopy techniques including confocal laser scanning microscopy and low-temperature scanning electron microscopy have also been used to help unravel the process of biofilm formation. For example, Moscato et al. used microscopy to observe early steps in biofilm formation by *Streptococcus pneumoniae* grown on polystyrene or glass surfaces [332]. Using *S. pneumoniae* mutants deficient in the synthesis of certain cell envelope components, they were then able to determine which of these components were important for biofilm formation.

5.3 Chemical Synthesis

Synthetic investigations of capsular polysaccharides have largely focused on the synthesis of fragments for use in vaccines. Routes have been developed, for example to CPS fragments from pathogenic strains of *N. meningitidis*, *V. cholerae*, *H. influenzae*, *S. pneumoniae*, and group B *Streptococcus* for use in vaccine construction and immunological studies (*Fig. 27*) [333,334,335,336,337,338,339,340].
In addition, a synthetic route was recently reported for β-(1,6)-linked glucosamine oligosaccharides containing up to 11 glucosamine residues, corresponding to fragments of the extracellular polysaccharide poly-*N*-acetylglucosamine, important in biofilm formation in many bacteria (*Fig. 27*) [341].
Lactic acid bacteria are used to produce fermented milk products, and the exopolysaccharides produced by the bacteria influence the texture of the resulting products. More importantly, these exopolysaccharides are thought to have several health benefits. There is evidence they lower cholesterol, modulate the immune system, help prevent colon cancer, and fight ulcers [342,343]. There is thus interest in establishing structure-function relationships for these structures, as well as in metabolic engineering of lactic acid bacteria to produce capsular polysaccharides with the desired properties [342,343].

5.4 Medicine

5.4.1 Roles of Capsular Antigens in Bacterial Infections

Capsular and extracellular polysaccharides are involved in several aspects of cellular behavior that are tied to bacterial survival and virulence [321]. The capsule layer provides a physical barrier that prevents the bacteria from drying out, aiding in survival outside a host. CPS are also involved in colonization and biofilm formation. In some bacteria CPS promote adherence to surfaces, aiding colonization and biofilm formation, while CPS in other bacteria inhibit adhesion and biofilm formation [344].

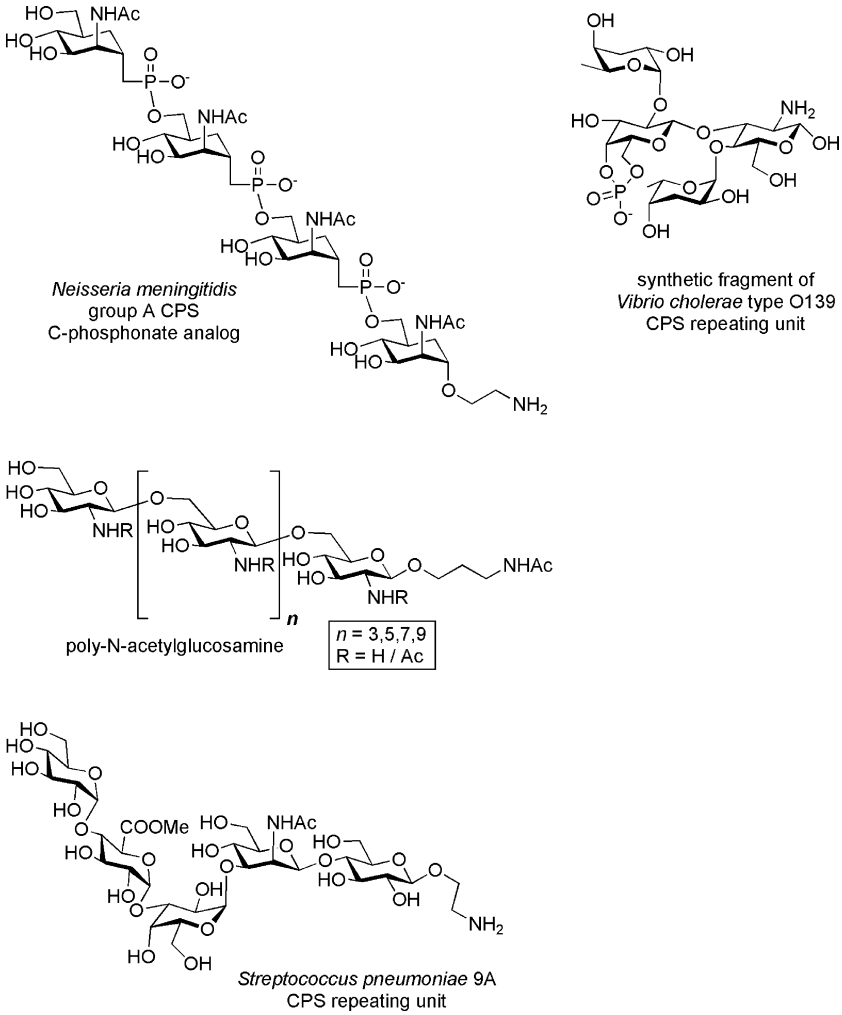


Figure 27
Synthetic capsular polysaccharide structures

Importantly, CPS also help bacteria evade the host's immune system. Following infection by a bacterial pathogen, the host's first lines of defense involve the innate immune system. Highly conserved bacterial structures including peptidoglycan, LPS and lipoteichoic acids are recognized by receptors on macrophages or by soluble pattern recognition molecules, leading to activation of the complement pathway, phagocytosis of the invading bacterial cells, and activation of an inflammatory cascade [345]. Bacterial capsules mask underlying bacterial cell surface structures, preventing recognition by the innate immune system and activation of the complement pathway. Furthermore, bacteria including *E. coli*, *N. meningitidis*, and *Group B Streptococcus* suppress activation of the innate immune system by making capsules containing *N*-acetylneuraminic acid (sialic acid). Vertebrates decorate their cell surfaces with sialic acid

as a mechanism to avoid accidentally killing their own healthy cells. The sialic acid binds complement regulatory proteins, and helps prevent complement activation on the surface of host cells. Bacterial capsules containing sialic acid are thought dampen an innate immune response by several mechanisms, including suppression of complement activation [346,347,348].
In both *S. pneumoniae* and *N. meningitidis*, the thickness of the capsule has been shown to vary at different points in infection. In *Neisseria meninigitidis* decreased capsule production enhances tissue invasion, while increased capsule production is essential for survival in systemic infections [349]. Likewise, studies in pneumococci have suggested that the capsule prevents bacterial adhesion to epithelial cells, as well as to endothelial cells [350,351,352]. Bacteria producing less capsular polysaccharide more efficiently colonize mucosal surfaces, while those producing more capsule are more virulent in systemic infections [350,353].

5.4.2 Anti-Bacterial Vaccines Based on CPS Structures

Most capsular polysaccharides strongly stimulate antibody production, and antibodies raised against particular capsular epitopes have been shown to prevent infection by bacteria possessing these capsules. Capsular polysaccharides are thus commonly used to develop vaccines against both Gram-negative and Gram-positive pathogens. Early vaccines contained just polysaccharides, while more recent formulations are based on glycoconjugates, where the polysaccharide antigens are bound to protein carriers [26]. Glycoconjugate vaccines offer several advantages over polysaccharide vaccines.
Polysaccharides are generally thymus-independent antigens, meaning that they stimulate antibody production by B-cells without the aid of helper T-cells [345]. This type of immune response does not result in the formation of memory B-cells or T-cells. Vaccination with polysaccharides thus provides protection for only a limited amount of time. Moreover, polysaccharide vaccines elicit a poor antibody response in young children and older adults [354]. In glycoconjugate vaccines, the protein carrier stimulates a T cell response. Immunological memory is thus established, and re-vaccination is not usually required (unlike with polysaccharide vaccines, which generally need to be re-administered every 5 years). The antibodies produced with T cell-mediated immunity are also of higher avidity. At present, several CPS vaccines are commonly used, and there are many more in development. For a thorough discussion, the reader is directed to a recent review by Jones [26]. Polysaccharide vaccines are currently used against *Salmonella typhi*, *Neisseria meningitidis* and *Streptococcus pneumoniae*. Glycoconjugate vaccines are licensed for use against *H. influenzae*, *N. meningitidis* and *S. pneumoniae*.
CPS vaccines against other pathogenic bacteria, including enterococcal and streptococcal strains, are in various stages of development and clinical trials. Notably, a vaccine against *S. aureus* called StaphVAX has been tested in Phase 3 clinical trials. StaphVAX is made from the two types of capsular polysaccharides most commonly found in clinical isolates (*Fig. 25*), each bound to recombinant *P. aeruginosa* exotoxin A (as a protein carrier). The vaccine was found to provide some protection from *S. aureus* infection, but only up to 40 weeks after injection [355].
An alternative to polysaccharide or glycoconjugate vaccines involves using peptides that mimic CPS antigens. Methods for obtaining peptide mimics include selection from phage-display libraries and use of anti-idiotype antibodies [354]. Anti-idiotype antibodies are antibodies that mimic the antigen to which they are raised, and are generated during a normal antibody-medi-

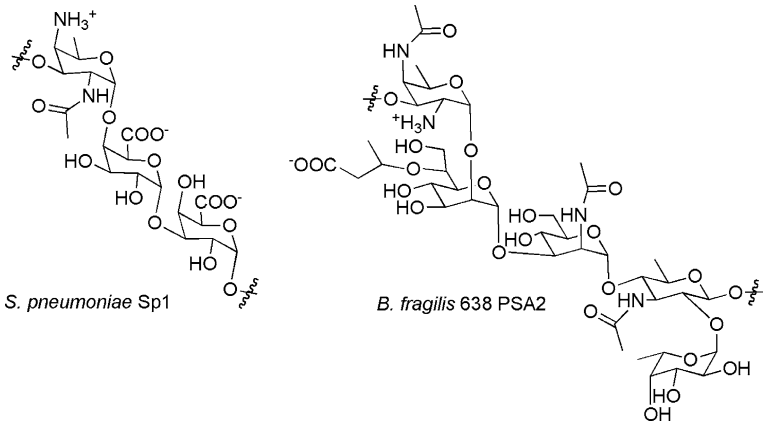


Figure 28
Zwitterionic capsular polysaccharides

ated immune response. Vaccines based on peptide mimics have the advantage of being able to stimulate a T-cell response, but the peptides are often degraded rapidly in vivo. The potential of this vaccine strategy has been demonstrated for bacteria including *Neisseria meningitidis*. Mice immunized with peptides mimicking *N. meningitidis* CPS were subsequently protected against infection by this pathogen [356,357].

While most polysaccharide antigens act in a T-cell independent fashion, zwitterionic polysaccharides have been shown to induce T-cell mediated immunity [358]. Several bacteria produce zwitterionic CPS, which contain both free amine and free carboxyl groups in the repeating unit. These include CPS types 5 and 8 of *Staphylococcus aureus*, PSA from *B. fragilis* and the type 1 *S. pneumoniae* polysaccharide (*Fig. 28*). Vaccination with zwitterionic polysaccharides have been shown to produce T-cell mediated immunity in mice [359], and vaccines based on zwitterionic CPS may be useful for combating several common bacterial pathogens.

5.4.3 Roles of Extracellular Polysaccharides in Bacterial Infections

Extracellular polysaccharides are a crucial component of biofilms. Biofilm formation is increasingly being recognized as an important factor in the transmission and persistence of bacterial infections. According to the CDC, more than 60% of bacterial infections involve biofilms, and these infections are problematic because biofilms are resistant to antibiotic therapies and host immune attack. The basis for the antibiotic resistance of biofilms is not completely understood, but is thought to involve a combination of factors [360]. Within the biofilm sub-populations of cells are living in a variety of different microenvironments, growing at different rates, and expressing different genes, so that some subset of cells is likely to show decreased susceptibility to any given antibiotic. In addition, the biofilm is highly charged, retarding the passage of some charged compounds and antibiotics [361,362].

Biofilms play key roles in several chronic human infections including infectious kidney stones, bacterial endocarditis, and cystic fibrosis lung infections, and the roles of biofilm formation in these diseases have been reviewed [27]. Biofilms also serve as environmental reservoirs for

pathogens. Biofilms can form on a variety of surfaces, and patients have obtained bacterial infections from biofilms persisting, for example, in hospital water lines, dialysis equipment and endoscopes [363]. In 1999, the CDC estimated that up to 270 000 infections may be transmitted by flexible endoscopes each year [364].

Patients with implanted medical devices, including replacement joints, cardiac pacemakers, urinary catheters and dialysis catheters, also commonly develop infections as a result of biofilm formation. A variety of Gram-positive and Gram-negative bacteria have been isolated from colonized devices, and biofilms often contain multiple bacterial species [363]. Strategies being investigated to prevent such infections involve development of infection resistant materials for use in implants [365], and covering medical devices with antimicrobial coatings [366]. Biofilm formation requires the coordinated action of many individual bacteria, and the importance of bacterial quorum sensing in biofilm formation has been demonstrated. For example, knocking out genes involved in quorum sensing in *Pseudomonas aeruginosa* results in reduced virulence [367]. Quorum sensing inhibitors are therefore being explored to prevent biofilm formation [368]. Greater understanding of the bacterial processes involved in biofilm formation will provide insight into better strategies for preventing biofilm-associated infections.

References

1. Koch AL (2006) J Mol Microbiol Biotechnol 11:115
2. Silver LL (2006) Biochem Pharmacol 71:996
3. Cabeen MT, Jacobs-Wagner C (2005) Nat Rev Microbiol 3:601
4. Beveridge TJ, Davies JA (1983) J Bacteriol 156:846
5. Nikaido H (2003) Microbiol Mol Biol Rev 67:593
6. Kamio Y, Nikaido H (1976) Biochemistry 15:2561
7. Raetz CR, Whitfield C (2002) Annu Rev Biochem 71:635
8. Reeves P (1995) Trends Microbiol 3:381
9. Jones C (2005) Carbohydr Res 340:1097
10. Hull S (1997) In: Sussman M (ed) Escherichia coli: Mechanisms of Virulence. Cambridge University Press, Cambridge, UK, p 145
11. Schulz GE (2002) Biochim Biophys Acta 1565:308
12. Tokuda H, Matsuyama S (2004) Biochim Biophys Acta 1694:IN1
13. Kuhn HM, Meier-Dieter U, Mayer H (1988) FEMS Microbiol Rev 4:195
14. Whitfield C, Roberts IS (1999) Mol Microbiol 31:1307
15. Pizarro-Cerda J, Cossart P (2006) Cell 124:715
16. Navarre WW, Schneewind O (1999) Microbiol Mol Biol Rev:174
17. Ward JB (1981) Microbiol Rev 45:211
18. Neuhaus FC, Baddiley J (2003) Microbiol Mol Biol Rev 67:686
19. Hermann T (2003) J Biotechnol 104:155
20. Brennan PJ (2003) Tuberculosis (Edinb) 83:91
21. Lowary TL (2001) In: Fraser-Reid B, Tatsuta K, Thiem J (eds) Glycoscience: Chemistry and Chemical Biology, vol 3. Springer-Verlag, Berlin Heidelberg New York, p 2005
22. Lambert PA (2002) J Appl Microbiol 92:Suppl:46S
23. American Thoracic Society, CDC, and Infectious Diseases Society of America (2003) MMWR 52:1
24. Ovodov YS (2006) Biochemistry (Mosc) 71:937
25. Comstock LE, Kasper DL (2006) Cell 126:847
26. Jones C (2005) An Acad Bras Cienc 77:293
27. Parsek MR, Singh PK (2003) Annu Rev Microbiol 57:677
28. Primosigh J, Pelzer H, Maass D, Weidel W (1961) Biochim Biophys Acta 46:68
29. Glauner B (1988) Anal Biochem 172:451
30. Weidel W, Pelzer H (1964) Adv Enzymol Relat Areas Mol Biol 26:193
31. Ghuysen JM (1968) Bacteriol Rev 32:425
32. Crick DC, Mahapatra S, Brennan PJ (2001) Glycobiology 11:107R
33. Schleifer KH, Kandler O (1972) Bacteriol Rev 36:407
34. Fiser A, Filipe SR, Tomasz A (2003) Trends Microbiol 11:547

35. Filipe SR, Severina E, Tomasz A (2001) Microb Drug Resist 7:303
36. Rohrer S, Berger-Bachi B (2003) Antimicrob Agents Chemother 47:837
37. van Heijenoort J (2001) Nat Prod Rep 18:503
38. Mahapatra S, Yagi T, Belisle JT, Espinosa BJ, Hill PJ, McNeil MR, Brennan PJ, Crick DC (2005) J Bacteriol 187:2747
39. Staudenbauer W, Strominger JL (1972) J Biol Chem 247:5095
40. Perry AM, Ton-That H, Mazmanian SK, Schneewind O (2002) J Biol Chem 277:16241
41. Benson TE, Walsh CT, Hogle JM (1996) Structure 4:47
42. Skarzynski T, Mistry A, Wonacott A, Hutchinson SE, Kelly VA, Duncan K (1996) Structure 4:1465
43. Schonbrunn E, Sack S, Eschenburg S, Perrakis A, Krekel F, Amrhein N, Mandelkow E (1996) Structure 4:1065
44. Bertrand JA, Auger G, Fanchon E, Martin L, Blanot D, van Heijenoort J, Dideberg O (1997) Embo J 16:3416
45. Gordon E, Flouret B, Chantalat L, van Heijenoort J, Mengin-Lecreulx D, Dideberg O (2001) J Biol Chem 276:10999
46. Yan Y, Munshi S, Leiting B, Anderson MS, Chrzas J, Chen Z (2000) J Mol Biol 304:435
47. Hu Y, Chen L, Ha S, Gross B, Falcone B, Walker D, Mokhtarzadeh M, Walker S (2003) Proc Natl Acad Sci USA 100:845
48. Ha S, Walker D, Shi Y, Walker S (2000) Protein Sci 9:1045
49. Deva T, Baker EN, Squire CJ, Smith CA (2006) Acta Crystallogr D Biol Crystallogr 62:1466
50. Bouhss A, Crouvoisier M, Blanot D, Mengin-Lecreulx D (2004) J Biol Chem 279:29974
51. van Heijenoort Y, Gomez M, Derrien M, Ayala J, van Heijenoort J (1992) J Bacteriol 174:3549
52. Pless DD, Neuhaus FC (1973) J Biol Chem 248:1568
53. Mengin-Lecreulx D, Texier L, Rousseau M, van Heijenoort J (1991) J Bacteriol 173:4625
54. Guan Z, Breazeale SD, Raetz CR (2005) Anal Biochem 345:336
55. Auger G, Crouvoisier M, Caroff M, Heijenoort Jv, Blanot D (1997) Lett Peptide Sci 4
56. Men H, Park P, Ge M, Walker S (1998) J Am Chem Soc 120:2484
57. VanNieuwenhze MS, Mauldin SC, Zia-Ebrahimi M, Aikins JA, Blaszczak LC (2001) J Am Chem Soc 123:6983
58. Ye X-Y, Lo M-C, Brunner L, Walker D, Kahne D, Walker S (2001) J Am Chem Soc 123:3155
59. Chen L, Men H, Ha S, Ye XY, Brunner L, Hu Y, Walker S (2002) Biochemistry 41:6824
60. Ha S, Chang E, Lo M-C, Men H, Park P, Ge M, Walker S (1999) J Am Chem Soc 121:8415
61. Liu H, Ritter TK, Sadamoto R, Sears PS, Wu M, CH CHW (2003) Chembiochem 4:603
62. VanNieuwenhze MS, Mauldin SC, Zia-Ebrahimi M, Winger BE, Hornback WJ, Saha SL, Aikins JA, Blaszczak LC (2002) J Am Chem Soc 124:3656
63. VanNieuwenhze MS (2004) In: Harmata M (ed) Strategies and Tactics in Organic Synthesis, vol 4. Elsevier, Amsterdam, p 293
64. Schwartz B, Markwalder JA, Wang Y (2001) J Am Chem Soc 123:11638
65. Allen CM, Keenan MV, Sack J (1976) Arch Biochem Biophys 175:236
66. Rogers HJ, Perkins HR, Ward JB (1980) Microbial Cell Walls and Membranes. Chapman & Hall, London
67. Breukink E, van Heusden HE, Vollmerhaus PJ, Swiezewska E, Brunner L, Walker S, Heck AJ, de Kruijff B (2003) J Biol Chem 278:19898
68. Glauner B, Holtje JV, Schwarz U (1988) J Biol Chem 263:10088
69. Harz H, Burgdorf K, Holtje JV (1990) Anal Biochem 190:120
70. Vollmer W, Holtje JV (2004) J Bacteriol 186:5978
71. Boneca IG, Huang ZH, Gage DA, Tomasz A (2000) J Biol Chem 275:9910
72. Schindler M, Mirelman D, Schwarz U (1976) Eur J Biochem 71:131
73. Ward JB (1973) Biochem J 133:395
74. Quintela JC, Caparros M, de Pedro MA (1995) FEMS Microbiol Lett 125:95
75. Snowden MA, Perkins HR, Wyke AW, Hayes MV, Ward JB (1989) J Gen Microbiol 135:3015
76. Gally D, Archibald AR (1993) J Gen Microbiol 139:1907
77. Matias VR, Al-Amoudi A, Dubochet J, Beveridge TJ (2003) J Bacteriol 185:6112
78. Labischinski H, Goodell EW, Goodell A, Hochberg ML (1991) J Bacteriol 173:751
79. Yao X, Jericho M, Pink D, Beveridge T (1999) J Bacteriol 181:6865
80. Koch AL (1998) Res Microbiol 149:689
81. Dmitriev BA, Toukach FV, Holst O, Rietschel ET, Ehlers S (2004) J Bacteriol 186:7141

82. Dmitriev BA, Toukach FV, Schaper KJ, Holst O, Rietschel ET, Ehlers S (2003) J Bacteriol 185: 3458
83. Meroueh SO, Bencze KZ, Hesek D, Lee M, Fisher JF, Stemmler TL, Mobashery S (2006) Proc Natl Acad Sci USA 103:4404
84. Edelman A, Bowler L, Broome-Smith JK, Spratt BG (1987) Mol Microbiol 1:101
85. Tipper DJ, Strominger JL (1968) J Biol Chem 243:3169
86. Van Heijenoort Y, Derrien M, Van Heijenoort J (1978) FEBS Lett 89:141
87. Nakagawa J, Tamaki S, Matsuhashi M (1979) Agric Biol Chem 43:1379
88. Ghuysen JM (1991) Annu Rev Microbiol 45:37
89. van Heijenoort J (2001) Glycobiology 11:25R
90. Di Berardino M, Dijkstra A, Stuber D, Keck W, Gubler M (1996) FEBS Lett 392:184
91. Hara H, Suzuki H (1984) FEBS Lett 168:155
92. Holtje JV (1998) Microbiol Mol Biol Rev 62:181
93. Nanninga N (1998) Microbiol Mol Biol Rev 62:110
94. Denome SA, Elf PK, Henderson TA, Nelson DE, Young KD (1999) J Bacteriol 181:3981
95. Wang QM, Peery RB, Johnson RB, Alborn WE, Yeh WK, Skatrud PL (2001) J Bacteriol 183:4779
96. Popham DL, Young KD (2003) Curr Opin Microbiol 6:594
97. Pucci MJ, Dougherty TJ (2002) J Bacteriol 184:588
98. Scheffers DJ, Pinho MG (2005) Microbiol Mol Biol Rev 69:585
99. Bhardwaj S, Day RA (1997) In: Marshak D (ed) Techniques in protein chemistry, vol 8. Academic Press, New York
100. Alaedini A, Day RA (1999) Biochem Biophys Res Commun 264:191
101. Simon MJ, Day RA (2000) Anal Lett 33:861
102. Figge RM, Divakaruni AV, Gober JW (2004) Mol Microbiol 51:1321
103. von Rechenberg M, Ursinus A, Holtje JV (1996) Microb Drug Resist 2:155
104. Vollmer W, von Rechenberg M, Holtje JV (1999) J Biol Chem 274:6726
105. Romeis T, Holtje JV (1994) J Biol Chem 269:21603
106. Wang L, Khattar MK, Donachie WD, Lutkenhaus J (1998) J Bacteriol 180:2810
107. Weiss DS, Pogliano K, Carson M, Guzman LM, Fraipont C, Nguyen-Disteche M, Losick R, Beckwith J (1997) Mol Microbiol 25:671
108. Daniel RA, Harry EJ, Errington J (2000) Mol Microbiol 35:299
109. Pedersen LB, Angert ER, Setlow P (1999) J Bacteriol 181:3201
110. Lewis PJ (2004) Mol Microbiol 54:1135
111. Morlot C, Zapun A, Dideberg O, Vernet T (2003) Mol Microbiol 50:845
112. de Pedro MA, Quintela JC, Holtje JV, Schwarz H (1997) J Bacteriol 179:2823
113. De Pedro MA, Schwarz H, Koch AL (2003) Microbiology 149:1753
114. Pooley HM, Schlaeppi JM, Karamata D (1978) Nature 274:264
115. Daniel RA, Errington J (2003) Cell 113:767
116. Tiyanont K, Doan T, Lazarus MB, Fang X, Rudner DZ, Walker S (2006) Proc Natl Acad Sci USA 103:11033
117. Thomas B, Wang Y, Stein RL (2001) Biochemistry 40:15811
118. Dessen A, Mouz N, Gordon E, Hopkins J, Dideberg O (2001) J Biol Chem 276:45106
119. Kelly JA, Kuzin AP (1995) J Mol Biol 254:223
120. Kelly JA, Kuzin AP, Charlier P, Fonze E (1998) Cell Mol Life Sci 54:353
121. Lim D, Strynadka NC (2002) Nat Struct Biol 9:870
122. Morlot C, Pernot L, Le Gouellec A, Di Guilmi AM, Vernet T, Dideberg O, Dessen A (2005) J Biol Chem 280:15984
123. Silvaggi NR, Josephine HR, Kuzin AP, Nagarajan R, Pratt RF, Kelly JA (2005) J Mol Biol 345:521
124. Macheboeuf P, Contreras-Martel C, Job V, Dideberg O, Dessen A (2006) FEMS Microbiol Rev 30:673
125. Pernot L, Chesnel L, Le Gouellec A, Croize J, Vernet T, Dideberg O, Dessen A (2004) J Biol Chem 279:16463
126. Anderson JS, Meadow PM, Haskin MA, Strominger JL (1966) Arch Biochem Biophys 116:487
127. Barrett D, Leimkuhler C, Chen L, Walker D, Kahne D, Walker S (2005) J Bacteriol 187:2215
128. Schwartz B, Markwalder JA, Seitz SP, Wang Y, Stein RL (2002) Biochemistry 41:12552
129. Terrak M, Nguyen-Disteche M (2006) J Bacteriol 188:2528
130. Zawadzka-Skomial J, Markiewicz Z, Nguyen-Disteche M, Devreese B, Frere JM, Terrak M (2006) J Bacteriol 188:1875
131. Di Guilmi AM, Dessen A, Dideberg O, Vernet T (2003) J Bacteriol 185:4418

132. Barrett DS, Chen L, Litterman NK, Walker S (2004) Biochemistry 43:12375
133. Contreras-Martel C, Job V, Di Guilmi AM, Vernet T, Dideberg O, Dessen A (2006) J Mol Biol 355:684
134. Lovering AL, De Castro L, Lim D, Strynadka NC (2006) Protein Sci 15:1701
135. Macheboeuf P, Di Guilmi AM, Job V, Vernet T, Dideberg O, Dessen A (2005) Proc Natl Acad Sci USA 102:577
136. Yuan Y, Barrett D, Zhang Y, Kahne D, Sliz P, Walker S (2007) Proc Natl Acad Sci USA
137. Lovering AL, de Castro LH, Lim D, Strynadka NC (2007) Science 315:1402
138. Zhang Y, Fechter EJ, Wang TS, Barrett D, Walker S, Kahne DE (2007) J Am Chem Soc
139. Arbeloa A, Hugonnet JE, Sentilhes AC, Josseaume N, Dubost L, Monsempes C, Blanot D, Brouard JP, Arthur M (2004) J Biol Chem 279:41546
140. Bugg TD, Walsh CT (1992) Nat Prod Rep 9:199
141. Bugg TD, Lloyd AJ, Roper DI (2006) Infect Disord Drug Targets 6:85
142. Kimura K, Bugg TD (2003) Nat Prod Rep 20:252
143. Dini C (2005) Curr Top Med Chem 5:1221
144. Igarashi M, Nakagawa N, Hattori S, N. Doi TN, Yamazaki T, Miyake T, Ishizuka M, Naganawa H, Shomura T, Omoto S, Yano I, Hamada M, Takeuchi T (2002) 42nd Interscience Conference on Antimicrobial Agents and Chemotherapy, San Diego, California, p 232
145. Isono F, Katayama T, Inukai M, Haneishi T (1989) J Antibiot (Tokyo) 42:674
146. Kahne D, Leimkuhler C, Lu W, Walsh C (2005) Chem Rev 105:425
147. Hsu ST, Breukink E, Tischenko E, Lutters MA, de Kruijff B, Kaptein R, Bonvin AM, van Nuland NA (2004) Nat Struct Mol Biol 11:963
148. Chatterjee C, Paul M, Xie L, van der Donk WA (2005) Chem Rev 105:633
149. Walker S, Chen L, Hu Y, Rew Y, Shin D, Boger DL (2005) Chem Rev 105:449
150. Fang X, Tiyanont K, Zhang Y, Wanner J, Boger D, Walker S (2006) Mol Biosyst 2:69
151. Leimkuhler C, Chen L, Barrett D, Panzone G, Sun B, Falcone B, Oberthur M, Donadio S, Walker S, Kahne D (2005) J Am Chem Soc 127:3250
152. Chen L, Walker D, Sun B, Hu Y, Walker S, Kahne D (2003) Proc Natl Acad Sci USA 100:5658
153. Hasper HE, Kramer NE, Smith JL, Hillman JD, Zachariah C, Kuipers OP, deKruijff B, Breukink E (2006) Science 313:1636
154. Welzel P (2005) Chem Rev 105:4610
155. Nicolaou KC, Snyder SA (2003) Wiley-VCH, Weinheim, Germany, p 239
156. Welzel P (1993) In: Krohn K, Kirst H, Maag H (eds) Antibiotics and Antiviral Compounds—Chemical Synthesis and Modification. VCH, New York, p 373
157. Taylor JG, Li X, Oberthur M, Zhu W, Kahne DE (2006) J Am Chem Soc 128:15084
158. Adachi M, Zhang Y, Leimkuhler C, Sun B, LaTour JV, Kahne DE (2006) J Am Chem Soc 128:14012
159. Ostash B, Saghatelian A, Walker S (2007) Chem Biol 14:257
160. Essack SY (2001) Pharm Res 18:1391
161. Fisher JF, Meroueh SO, Mobashery S (2005) Chem Rev 105:395
162. Koch AL (2000) Crit Rev Microbiol 26:205
163. Koch AL (2003) Clin Microbiol Rev 16:673
164. Livermore DM (2006) Clin Microbiol Infect 12 Suppl 2:11
165. Shin D, Rew Y, Boger DL (2004) Proc Natl Acad Sci USA 101:11977
166. Rew Y, Shin D, Hwang I, Boger DL (2004) J Am Chem Soc 126:1041
167. Jiang W, Wanner J, Lee RJ, Bounaud PY, Boger DL (2003) J Am Chem Soc 125:1877
168. Higgins DL, Chang R, Debabov DV, Leung J, Wu T, Krause KM, Sandvik E, Hubbard JM, Kaniga K, Schmidt DE, Jr., Gao Q, Cass RT, Karr DE, Benton BM, Humphrey PP (2005) Antimicrob Agents Chemother 49:1127
169. Bister B, Bischoff D, Nicholson GJ, Valdebenito M, Schneider K, Winkelmann G, Hantke K, Sussmuth RD (2004) Biometals 17:471
170. Kruger RG, Lu W, Oberthur M, Tao J, Kahne D, Walsh CT (2005) Chem Biol 12:131
171. Losey HC, Jiang J, Biggins JB, Oberthur M, Ye XY, Dong SD, Kahne D, Thorson JS, Walsh CT (2002) Chem Biol 9:1305
172. Lu W, Oberthur M, Leimkuhler C, Tao J, Kahne D, Walsh CT (2004) Proc Natl Acad Sci USA 101:4390
173. Sofia MJ, Allanson N, Hatzenbuhler NT, Jain R, Kakarla R, Kogan N, Liang R, Liu D, Silva DJ, Wang H, Gange D, Anderson J, Chen A, Chi F, Dulina R, Huang B, Kamau M, Wang C, Baizman E, Branstrom A, Bristol N, Goldman R, Han K, Longley C, Axelrod HR, et al. (1999) J Med Chem 42:3193

174. Baizman ER, Branstrom AA, Longley CB, Allanson N, Sofia MJ, Gange D, Goldman RC (2000) Microbiology 146 Pt 12:3129
175. Goldman RC, Baizman ER, Branstrom AA, Longley CB (2000) Bioorg Med Chem Lett 10:2251
176. Garneau S, Qiao L, Chen L, Walker S, Vederas JC (2004) Bioorg Med Chem 12:6473
177. Ruhl T, Daghish M, Buchynskyy A, Barche K, Volke D, Stembera K, Kempin U, Knoll D, Hennig L, Findeisen M, Oehme R, Giesa S, Ayala J, Welzel P (2003) Bioorg Med Chem 11:2965
178. El Zoeiby A, Beaumont M, Dubuc E, Sanschagrin F, Voyer N, Levesque RC (2003) Bioorg Med Chem 11:1583
179. Barbosa MD, Ross HO, Hillman MC, Meade RP, Kurilla MG, Pompliano DL (2002) Anal Biochem 306:17
180. Comess KM, Schurdak ME, Voorbach MJ, Coen M, Trumbull JD, Yang H, Gao L, Tang H, Cheng X, Lerner CG, McCall JO, Burns DJ, Beutel BA (2006) J Biomol Screen 11:743
181. Longenecker KL, Stamper GF, Hajduk PJ, Fry EH, Jakob CG, Harlan JE, Edalji R, Bartley DM, Walter KA, Solomon LR, Holzman TF, Gu YG, Lerner CG, Beutel BA, Stoll VS (2005) Protein Sci 14:3039
182. Helm JS, Hu Y, Chen L, Gross B, Walker S (2003) J Am Chem Soc 125:11168
183. Hu Y, Helm JS, Chen L, Ginsberg C, Gross B, Kraybill B, Tiyanont K, Fang X, Wu T, Walker S (2004) Chem Biol 11:703
184. Osborn MJ, Gander JE, Parisi E, Carson J (1972) J Biol Chem 247:3962
185. Nikaido H, Vaara M (1985) Microbiol Rev 49:1
186. Raetz CR (1990) Annu Rev Biochem 59:129
187. Brabetz W, Muller-Loennies S, Holst O, Brade H (1997) Eur J Biochem 247:716
188. Meredith TC, Aggarwal P, Mamat U, Lindner B, Woodard RW (2006) ACS Chem Biol 1:33
189. Holst O (1999) In: Brade H, Opal SM, Vogel SN, Morrison DC (eds) Endotoxin in Health and Disease. Marcel Dekker Inc., New York, p 115
190. Reeves P (1993) Trends Genet 9:17
191. Stenutz R, Weintraub A, Widmalm G (2006) FEMS Microbiol Rev 30:382
192. Lerouge I, Vanderleyden J (2002) FEMS Microbiol Rev 26:17
193. Preston A, Mandrell RE, Gibson BW, Apicella MA (1996) Crit Rev Microbiol 22:139
194. Harvey HA, Swords WE, Apicella MA (2001) J Autoimmun 16:257
195. Mandrell RE, Apicella MA (1993) Immunobiology 187:382
196. Moran AP, Prendergast MM (2001) J Autoimmun 16:241
197. Leive L (1974) Ann N Y Acad Sci 235:109
198. Leive L (1965) Biochem Biophys Res Commun 21:290
199. Vaara M, Vaara T (1983) Antimicrob Agents Chemother 24:114
200. Cronan JE (2003) Annu Rev Microbiol 57:203
201. Trent MS (2004) Biochem Cell Biol 82:71
202. Fisseha M, Chen P, Brandt B, Kijek T, Moran E, Zollinger W (2005) Infect Immun 73:4070
203. Anisimov AP, Shaikhutdinova RZ, Pan'kina LN, Feodorova VA, Savostina EP, Bystrova OV, Lindner B, Mokrievich AN, Bakhteeva IV, Titareva GM, Dentovskaya SV, Kocharova NA, Senchenkova SN, Holst O, Devdariani ZL, Popov YA, Pier GB, Knirel YA (2007) J Med Microbiol 56:443
204. Heinrichs DE, Yethon JA, Amor PA, Whitfield C (1998) J Biol Chem 273:29497
205. Yethon JA, Heinrichs DE, Monteiro MA, Perry MB, Whitfield C (1998) J Biol Chem 273:26310
206. Kneidinger B, Marolda C, Graninger M, Zamyatina A, McArthur F, Kosma P, Valvano MA, Messner P (2002) J Bacteriol 184:363
207. Walsh AG, Matewish MJ, Burrows LL, Monteiro MA, Perry MB, Lam JS (2000) Mol Microbiol 35:718
208. Yethon JA, Gunn JS, Ernst RK, Miller SI, Laroche L, Malo D, Whitfield C (2000) Infect Immun 68:4485
209. Doerrler WT, Raetz CR (2002) J Biol Chem 277:36697
210. Chang G (2003) J Mol Biol 330:419
211. Chang G, Roth CB (2001) Science 293:1793
212. Reyes CL, Chang G (2005) Science 308:1028
213. Chang G, Roth CB, Reyes CL, Pornillos O, Chen YJ, Chen AP (2006) Science 314:1875
214. Dawson RJ, Locher KP (2006) Nature 443:180
215. Keenleyside WJ, Whitfield C (1996) J Biol Chem 271:28581
216. Lehrer J, Vigeant KA, Tatar LD, Valvano MA (2007) J Bacteriol 189:2618
217. Meier-Dieter U, Barr K, Starman R, Hatch L, Rick PD (1992) J Biol Chem 267:746
218. Wang L, Liu D, Reeves PR (1996) J Bacteriol 178:2598
219. Samuel G, Reeves P (2003) Carbohydr Res 338:2503
220. Abeyrathne PD, Daniels C, Poon KK, Matewish MJ, Lam JS (2005) J Bacteriol 187:3002

221. Heinrichs DE, Monteiro MA, Perry MB, Whitfield C (1998) J Biol Chem 273:8849
222. Schild S, Lamprecht AK, Reidl J (2005) J Biol Chem 280:25936
223. Bishop RE, Gibbons HS, Guina T, Trent MS, Miller SI, Raetz CR (2000) Embo J 19:5071
224. Hwang PM, Choy WY, Lo EI, Chen L, Forman-Kay JD, Raetz CR, Prive GG, Bishop RE, Kay LE (2002) Proc Natl Acad Sci USA 99:13560
225. Osborn MJ, Gander JE, Parisi E (1972) J Biol Chem 247:3973
226. Doerrler WT (2006) Mol Microbiol 60:542
227. Bayer ME (1968) J Gen Microbiol 53:395
228. Bayer ME (1991) J Struct Biol 107:268
229. Muhlradt PF, Menzel J, Golecki JR, Speth V (1973) Eur J Biochem 35:471
230. Wu T, McCandlish AC, Gronenberg LS, Chng SS, Silhavy TJ, Kahne D (2006) Proc Natl Acad Sci USA 103:11754
231. Sperandeo P, Cescutti R, Villa R, Di Benedetto C, Candia D, Deho G, Polissi A (2007) J Bacteriol 189:244
232. Imoto M, Yoshimura H, Sakaguchi N, Kusumoto S, Shiba T (1985) Tetrahedron Letters 26:1545
233. Imoto M, Yoshimura H, Yamamoto M, Shimamoto T, Kusumoto S, Shiba T (1987) Bulletin of the Chemical Society of Japan 60:2197
234. Kusumoto S, Fukase K (2007) Chem Rec 6:333
235. Rietschel E, Kirikae T, Schade F, Mamat U, Schmidt G, Loppnow H, Ulmer A, Zahringer U, Seydel U, Padova FD (1994) FASEB J 8:217
236. Agnihotri G, Misra AK (2006) Carbohydr Res 341:2420
237. Kwon YT, Lee YJ, Lee K, Kim KS (2004) Org Lett 6:3901
238. Belot F, Wright K, Costachel C, Phalipon A, Mulard LA (2004) J Org Chem 69:1060
239. Phalipon A, Costachel C, Grandjean C, Thuizat A, Guerreiro C, Tanguy M, Nato F, Vulliez-Le Normand B, Belot F, Wright K, Marcel-Peyre V, Sansonetti PJ, Mulard LA (2006) J Immunol 176:1686
240. Wade TK, Saksena R, Shiloach J, Kováč P, Wade WF (2006) FEMS Immunology & Medical Microbiology 48:237
241. Zhang J, Kovac P (1999) Carbohydr Res 321:157
242. Ulevitch RJ, Tobias PS (1995) Annu Rev Immunol 13:437
243. Triantafilou M, Triantafilou K (2005) J Endotoxin Res 11:5
244. Jerala R, Porro M (2004) Curr Top Med Chem 4:1173
245. Tsubery H, Ofek I, Cohen S, Eisenstein M, Fridkin M (2002) Mol Pharmacol 62:1036
246. Thomas CJ, Surolia A (1999) FEBS Lett 445:420
247. Thomas CJ, Surolia N, Surolia A (1999) J Biol Chem 274:29624
248. Ghiselli R, Giacometti A, Cirioni O, Circo R, Mocchegiani F, Skerlavaj B, D'Amato G, Scalise G, Zanetti M, Saba V (2003) Shock 19:577
249. Levy O (2002) Expert Opin Investig Drugs 11:159
250. Cohen D, Ashkenazi S, Green MS, Gdalevich M, Robin G, Slepon R, Yavzori M, Orr N, Block C, Ashkenazi I, Shemer J, Taylor DN, Hale TL, Sadoff JC, Pavliakova D, Schneerson R, Robbins JB (1997) Lancet 349:155
251. Cross AS, Opal SM, Palardy JE, Drabick JJ, Warren HS, Huber C, Cook P, Bhattacharjee AK (2003) Vaccine 21:4576
252. Savoy F, Nicolle DM, Rivier D, Chiavaroli C, Ryffel B, Quesniaux VF (2006) Immunobiology 211:767
253. Reisser D, Pance A, Jeannin JF (2002) Bioessays 24:284
254. Coley W (1898) J Am Med Assoc 31:589
255. Asselineau J, Laneelle G (1998) Front Biosci 3:e164
256. Lee A, Wu SW, Scherman MS, Torrelles JB, Chatterjee D, McNeil MR, Khoo KH (2006) Biochemistry 45:15817
257. Puech V, Chami M, Lemassu A, Laneelle MA, Schiffler B, Gounon P, Bayan N, Benz R, Daffe M (2001) Microbiology 147:1365
258. Sutcliffe IC (1998) Antonie Van Leeuwenhoek 74:49
259. Barry CE, 3rd, Lee RE, Mdluli K, Sampson AE, Schroeder BG, Slayden RA, Yuan Y (1998) Prog Lipid Res 37:143
260. Dover LG, Cerdeno-Tarraga AM, Pallen MJ, Parkhill J, Besra GS (2004) FEMS Microbiol Rev 28:225
261. Nishiuchi Y, Baba T, Hotta HH, Yano I (1999) J Microbiol Methods 37:111
262. Nigou J, Gilleron M, Puzo G (2003) Biochimie 85:153
263. Khoo KH, Douglas E, Azadi P, Inamine JM, Besra GS, Mikusova K, Brennan PJ, Chatterjee D (1996) J Biol Chem 271:28682
264. Briken V, Porcelli SA, Besra GS, Kremer L (2004) Mol Microbiol 53:391
265. Berg S, Kaur D, Jackson M, Brennan PJ (2007) Glycobiology

266. Gurcha SS, Baulard AR, Kremer L, Locht C, Moody DB, Muhlecker W, Costello CE, Crick DC, Brennan PJ, Besra GS (2002) Biochem J 365:441
267. Mikusova K, Huang H, Yagi T, Holsters M, Vereecke D, D'Haeze W, Scherman MS, Brennan PJ, McNeil MR, Crick DC (2005) J Bacteriol 187:8020
268. Mills JA, Motichka K, Jucker M, Wu HP, Uhlik BC, Stern RJ, Scherman MS, Vissa VD, Pan F, Kundu M, Ma YF, McNeil M (2004) J Biol Chem 279:43540
269. Mikusova K, Belanova M, Kordulakova J, Honda K, McNeil MR, Mahapatra S, Crick DC, Brennan PJ (2006) J Bacteriol 188:6592
270. Rose NL, Completo GC, Lin SJ, McNeil M, Palcic MM, Lowary TL (2006) J Am Chem Soc 128:6721
271. Alderwick LJ, Seidel M, Sahm H, Besra GS, Eggeling L (2006) J Biol Chem 281:15653
272. Escuyer VE, Lety MA, Torrelles JB, Khoo KH, Tang JB, Rithner CD, Frehel C, McNeil MR, Brennan PJ, Chatterjee D (2001) J Biol Chem 276:48854
273. Takayama K, Wang C, Besra GS (2005) Clin Microbiol Rev 18:81
274. Kaur D, Berg S, Dinadayala P, Gicquel B, Chatterjee D, McNeil MR, Vissa VD, Crick DC, Jackson M, Brennan PJ (2006) Proc Natl Acad Sci USA 103:13664
275. Jackson M, Crick DC, Brennan PJ (2000) J Biol Chem 275:30092
276. Norman RA, McAlister MS, Murray-Rust J, Movahedzadeh F, Stoker NG, McDonald NQ (2002) Structure 10:393
277. Nigou J, Dover LG, Besra GS (2002) Biochemistry 41:4392
278. Morita YS, Sena CB, Waller RF, Kurokawa K, Sernee MF, Nakatani F, Haites RE, Billman-Jacobe H, McConville MJ, Maeda Y, Kinoshita T (2006) J Biol Chem 281:25143
279. Dinadayala P, Kaur D, Berg S, Amin AG, Vissa VD, Chatterjee D, Brennan PJ, Crick DC (2006) J Biol Chem 281:20027
280. Berg S, Starbuck J, Torrelles JB, Vissa VD, Crick DC, Chatterjee D, Brennan PJ (2005) J Biol Chem 280:5651
281. Ayers JD, Lowary TL, Morehouse CB, Besra GS (1998) Bioorg Med Chem Lett 8:437
282. Lee RE, Brennan PJ, Besra GS (1997) Glycobiology 7:1121
283. D'Souza FW, Ayers JD, McCarren PR, Lowary TL (2000) J Am Chem Soc 122:1251
284. Khasnobis S, Zhang J, Angala SK, Amin AG, McNeil MR, Crick DC, Chatterjee D (2006) Chem Biol 13:787
285. Subramaniam V, Gurcha SS, Besra GS, Lowary TL (2005) Bioorg Med Chem 13:1083
286. Wen X, Crick DC, Brennan PJ, Hultin PG (2003) Bioorg Med Chem 11:3579
287. Pathak AK, Pathak V, Suling WJ, Gurcha SS, Morehouse CB, Besra GS, Maddry JA, Reynolds RC (2002) Bioorg Med Chem 10:923
288. Lowary TL (2003) Curr Opin Chem Biol 7:749
289. Kremer L, Dover LG, Morehouse C, Hitchin P, Everett M, Morris HR, Dell A, Brennan PJ, McNeil MR, Flaherty C, Duncan K, Besra GS (2001) J Biol Chem 276:26430
290. Gandolfi-Donadio L, Gallo-Rodriguez C, de Lederkremer RM (2003) J Org Chem 68:6928
291. Liu X, Stocker BL, Seeberger PH (2006) J Am Chem Soc 128:3638
292. Ainge GD, Parlane NA, Denis M, Hayman CM, Larsen DS, Painter GF (2006) Bioorg Med Chem 14:7615
293. World Health Organization (2007) Leprosy elimination website http://www.who.int/lep/en/ Accessed 11 Nov 2007
294. Vitek CR (2006) Curr Top Microbiol Immunol 304:71
295. Aziz MA, Wright A, Laszlo A, De Muynck A, Portaels F, Van Deun A, Wells C, Nunn P, Blanc L, Raviglione M (2006) Lancet 368:2142
296. US Department of Health and Human Services, CDC (September 2006)
297. Wagner D, Young LS (2004) Infection 32:257
298. Otsuka Y, Kawamura Y, Koyama T, Iihara H, Ohkusu K, Ezaki T (2005) J Clin Microbiol 43:3713
299. Fernandez-Natal I, Guerra J, Alcoba M, Cachon F, Soriano F (2001) Eur J Clin Microbiol Infect Dis 20:514
300. Forget EJ, Menzies D (2006) Expert Opin Drug Saf 5:231
301. Skeiky YA, Sadoff JC (2006) Nat Rev Microbiol 4:469
302. Advisory Council for the Elimination of Tuberculosis and the Advisory Committee on Immunization Practices (1996) MMWR Recomm Rep 45 RR-4:1
303. Russell DG (2007) Nat Rev Microbiol 5:39
304. Riley LW (2006) J Clin Invest 116:1475
305. Lee RE, Protopopova M, Crooks E, Slayden RA, Terrot M, Barry CE, 3rd (2003) J Comb Chem 5:172

306. Schroeder EK, de Souza N, Santos DS, Blanchard JS, Basso LA (2002) Curr Pharm Biotechnol 3:197
307. Lowary TL (2003) Mini Rev Med Chem 3:689
308. Nikonenko BV, Protopopova M, Samala R, Einck L, Nacy CA (2007) Antimicrob Agents Chemother
309. Matsumoto M, Hashizume H, Tomishige T, Kawasaki M, Tsubouchi H, Sasaki H, Shimokawa Y, Komatsu M (2006) PLoS Med 3:e466
310. Lasa I (2006) Int Microbiol 9:21
311. Ovodov YS (2006) Biochemistry (Mosc) 71:955
312. Whitfield C (2006) Annu Rev Biochem 75:39
313. Sorensen UB, Henrichsen J, Chen HC, Szu SC (1990) Microb Pathog 8:325
314. Deng L, Kasper DL, Krick TP, Wessels MR (2000) J Biol Chem 275:7497
315. Bender MH, Cartee RT, Yother J (2003) J Bacteriol 185:6057
316. Sutherland IW (2001) Trends Microbiol 9:222
317. Sutherland IW (1999) In: Wingender J, Neu TR, Flemming HC (eds) Microbial Extracellular Polymeric Substances. Springer, Berlin Heidelberg New York, p 73–92
318. Lasa I (2006) Int Microbiol 9:21
319. Branda SS, Vik S, Friedman L, Kolter R (2005) Trends Microbiol 13:20
320. Zhang X, Bishop PL, Kupferle MJ (1998) Water Sci Technol 37:345
321. Roberts IS (1996) Annu Rev Microbiol 50:285
322. Kroll JS (1992) J Infect Dis 165 Suppl 1:S93
323. O'Riordan K, Lee JC (2004) Clin Microbiol Rev 17:218
324. Llull D, Lopez R, Garcia E (2001) Curr Mol Med 1:475
325. Garcia E, Llull D, Munoz R, Mollerach M, Lopez R (2000) Res Microbiol 151:429
326. Whitfield C, Paiment A (2003) Carbohydr Res 338:2491
327. Silver RP, Prior K, Nsahlai C, Wright LF (2001) Res Microbiol 152:357
328. Xayarath B, Yother J (2007) J Bacteriol 189:3369
329. Kleerebezem M, Kranenburg Rv, Tuinier R, Boels IC, Zoon P, Looijesteijn E, Hugenholtz J, Vos WMd (1999) Antonie Van Leeuwenhoek 76:357
330. Collins RF, Beis K, Dong C, Botting CH, McDonnell C, Ford RC, Clarke BR, Whitfield C, Naismith JH (2007) Proc Natl Acad Sci USA 104:2390
331. O'Toole G, Kaplan HB, Kolter R (2000) Annu Rev Microbiol 54:49
332. Moscoso M, Garcia E, Lopez R (2006) J Bacteriol 188:7785
333. Alpe M, Oscarson S (2003) Carbohydr Res 338:2605
334. Hansson J, Garegg PJ, Oscarson S (2001) J Org Chem 66:6234
335. Slattegard R, Teodorovic P, Kinfe HH, Ravenscroft N, Gammon DW, Oscarson S (2005) Org Biomol Chem 3:3782
336. Teodorovic P, Slattegard R, Oscarson S (2006) Org Biomol Chem 4:4485
337. Turek D, Sundgren A, Lahmann M, Oscarson S (2006) Org Biomol Chem 4:1236
338. Demchenko AV, Boons GJ (2001) J Org Chem 66:2547
339. Berkin A, Coxon B, Pozsgay V (2002) Chemistry 8:4424
340. Pozsgay V, Brisson JR, Jennings HJ (1990) Carbohydr Res 205:133
341. Gening ML, Tsvetkov YE, Pier GB, Nifantiev NE (2007) Carbohydr Res 342:567
342. Welman AD, Maddox IS (2003) Trends Biotechnol 21:269
343. Jolly L, Vincent SJ, Duboc P, Neeser JR (2002) Antonie Van Leeuwenhoek 82:367
344. Joseph LA, Wright AC (2004) J Bacteriol 186:889
345. Janeway C, Travers P, Walport M, Shlomchik M (2004) Immunobiology: The Immune System in Health and Disease—Sixth Edition. Taylor & Francis, London
346. Carlin AF, Lewis AL, Varki A, Nizet V (2007) J Bacteriol 189:1231
347. Jarvis GA, Vedros NA (1987) Infect Immun 55:174
348. Marques MB, Kasper DL, Pangburn MK, Wessels MR (1992) Infect Immun 60:3986
349. Hammerschmidt S, Muller A, Sillmann H, Muhlenhoff M, Borrow R, Fox A, van Putten J, Zollinger WD, Gerardy-Schahn R, Frosch M (1996) Mol Microbiol 20:1211
350. Weiser JN (2000) In: Fischetti V, Novick RP, Ferretti JJ, Portnoy DA, Rood JI (eds) Gram-positive pathogens. ASM Press, Washington, D.C., p 225
351. Adamou JE, Wizemann TM, Barren P, Langermann S (1998) Infect Immun 66:820
352. Talbot UM, Paton AW, Paton JC (1996) Infect Immun 64:3772

353. Hammerschmidt S, Wolff S, Hocke A, Rosseau S, Muller E, Rohde M (2005) Infect Immun 73:4653
354. Weintraub A (2003) Carbohydr Res 338: 2539
355. Shinefield H, Black S, Fattom A, Horwith G, Rasgon S, Ordonez J, Yeoh H, Law D, Robbins JB, Schneerson R, Muenz L, Fuller S, Johnson J, Fireman B, Alcorn H, Naso R (2002) N Engl J Med 346:491
356. Prinz DM, Smithson SL, Westerink MA (2004) J Immunol Methods 285:1
357. Passo CL, Romeo A, Pernice I, Donato P, Midiri A, Mancuso G, Arigo M, Biondo C, Galbo R, Papasergi S, Felici F, Teti G, Beninati C (2007) J Immunol 178:4417
358. Cobb BA, Kasper DL (2005) Cellular Microbiology 7:1398
359. Malley R, Srivastava A, Lipsitch M, Thompson CM, Watkins C, Tzianabos A, Anderson PW (2006) Infect Immun 74:2187
360. Mah TF, O'Toole GA (2001) Trends Microbiol 9:34
361. Jefferson KK, Goldmann DA, Pier GB (2005) Antimicrob Agents Chemother 49:2467
362. Fux CA, Costerton JW, Stewart PS, Stoodley P (2005) Trends Microbiol 13:34
363. Lindsay D, von Holy A (2006) J Hosp Infect 64:313
364. FDA and CDC (1999) FDA and CDC public health advisory: Infections from endoscopes inadequately reprocessed by an automated endoscope reprocessing system
365. Arciola CR, Alvi FI, An YH, Campoccia D, Montanaro L (2005) Int J Artif Organs 28:1119
366. von Eiff C, Kohnen W, Becker K, Jansen B (2005) Int J Artif Organs 28:1146
367. Smith RS, Iglewski BH (2003) Curr Opin Microbiol 6:56
368. Musk DJ, Jr., Hergenrother PJ (2006) Curr Med Chem 13:2163
369. Noel GJ (2007) Clin Microbiol Infect Suppl 2:25

Part 7
Glycolipids

7.1 Glycolipids: Occurrence, Significance, and Properties

Otto Holst
Structural Biochemistry, Research Center Borstel, 23485 Borstel, Germany
oholst@fz-borstel.de

Abstract

This chapter focuses on the occurrence and the physicochemical properties of glycolipids in Nature. Owing to space limitations, the presented overview must be incomplete, and, thus, mainly publications of the past 15 years are included. However, all review articles cited herein inform the interested reader about earlier work. Although lipopolysaccharides (LPS), lipoarabinomannan (LAM), lipomannan, lipoglycans, and lipoteichoic acids are not understood as glycolipids per definition, their occurrence and properties are also described in this chapter. GPI-anchored lipids is a main topic of ❯ Chap. 7.4.

Keywords

Glycolipids; Diacylglycerol ethers; Diacylglyceryl esters; Sulfated glycolipids; Trehalose-containing lipo-oligosaccharides; Glycopeptidolipids; Phenolic glycolipids; Lipopolysaccharides; Lipoglycans; Lipoteichoic acids

Abbreviations

BHT bacteriohopanetetrol
CAC critical aggregate concentration

In: *Glycoscience*. Fraser-Reid B, Tatsuta K, Thiem J (eds)
Chapter-DOI 10-1007/978-3-540-30429-6_39:

DTBP 2,6-di-*tert*-butylpyridine
FTIR Fourier-transformed infrared spectroscopy
GPL glycopeptidolipids
HIV human immunodeficiency virus
LAM lipoarabinomannan
LPS lipopolysaccharides
LTA lipoteichoic acids

1 Introduction

Glycolipids are membrane components and occur in all kingdoms of organisms, i. e., bacteria, plants, and animals including man. They are glycoconjugates and comprise a very heterogeneous group of molecules which possess one or more monosaccharide units linked by a glycosyl linkage to a hydrophobic moiety such as acylglycerol, ceramide, or prenyl phosphate [1]. In general, glycolipids are important factors of membrane stabilization and cell surface rigidization. They also play an important role in a broad variety of biological processes, e. g., in photosynthetic electron transport in plants, in cell-cell communication, in receptor modulation, or in signal transduction.

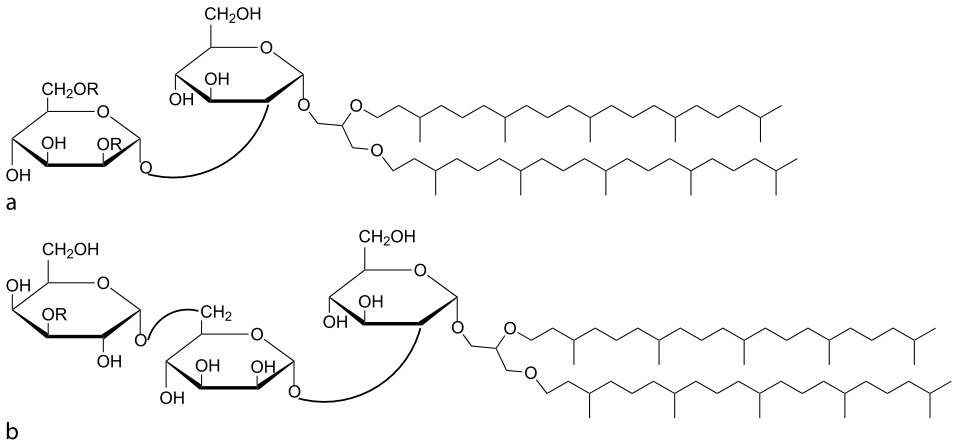
In general, glycolipids are amphiphilic molecules comprising a hydrophilic headgroup and apolar hydrophobic hydrocarbon (acyl) chains. Thus, glycolipids tend to aggregate to clusters in aqueous environments, i. e., to lipid bilayers or membranes. Possible supramolecular arrangements are micellar, lamellar, hexagonal, non-lamellar cubic, and inverted-hexagonal structures [2]. Which of these structures is furnished, is influenced by the primary chemical structure of a molecule and by extrinsic factors like temperature, pH, and the concentration of bivalent ions. Elevated temperature leads to the transition of the gel (β) to the liquid crystalline (α) phase of the acyl chains which occurs at the transition temperature T_c. A major difference between glycolipids and phospholipids results from the strong influence of hydrogen bonds on the phase behavior of the former, which leads to a rather high T_c (providing structural integrity in glycolipid-containing membranes) and to metastable polymorphism.

In cells, a particular composition and the physicochemical properties of the lipid matrix at given ambient conditions represent the basis for a normal functioning [3]. Any disturbance of the composition due, e. g., to different net charges or acylation patterns may result in changes of membrane fluidity/permeability, in phase separation, or in disruption of the lamellar membrane architecture. However, cells are potentially able to overcome such changes in lipid matrix composition by altering it ("homovicious adaptation") [4]. If an adaptation of this type does not take place, severe dysfunction of a cell may occur.

Various means and methodologies have been used to investigate the physicochemical properties of synthetic, semisynthetic, and natural glycolipids, e. g., NMR spectroscopy, X-ray diffraction, electron microscopy, or Fourier-transformed infrared spectroscopy (FTIR). In the following paragraphs, the occurence and physicochemical properties of different biological glycolipid classes and of lipopolysaccharides, lipoteichoic acids, and mycobacterial mycolates are briefly discussed.

2 Prokaryotic Glycolipids

Membrane glycolipids of the family Halobacteriaceae (taxon Archaea, formerly Archaebacteria) represent derivatives of *sn*-2,3-di-*O*-alkylglycerol (where alkyl is in most cases phytanyl, ● *Fig. 1*) [5,6] or its dimer. In quite a few cases, these glycolipids are sulfated in their sugar portion [6]. There is evidence that sulfated glycolipids of this type contribute to the control of the ion permeability of the cytoplasmic membrane. Also, these lipids represent useful taxonomic markers of halophilic bacteria. Recently, a novel glycolipid has been isolated from the purple membranes of *Halobacterium salinarum*, with the structure 3-HSO_3-β-Gal*p*-(1→6)-α-Man*p*-(1→2)-α-Glc*p*-(1→1)-[*sn*-2,3-di-*O*-phytanylglycerol]-6-[phospho-*sn*-2,3-di-*O*-phytanylglycerol] [7]. Later, it could be shown that an increase of this glycolipid is induced by osmotic shock [8].



● Figure 1
Sulfoglycodialkylglycerolipids from the halobacteria, (a) *Natralba asiatica*, and (b) *Halobacterium salinarum* and *Hb. saccharovorum*. *R* = SO_3^-

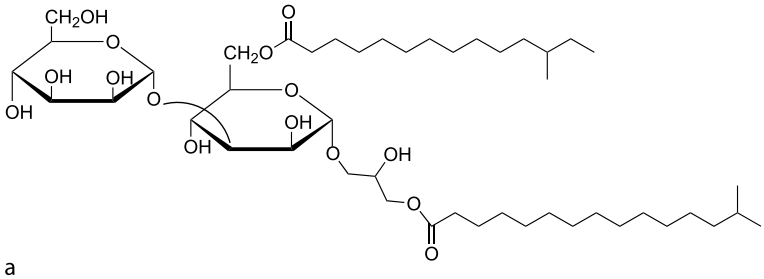
In cyanobacteria, as in photosynthetic bacteria, glycolipids are constituents of the thylakoid membranes. However, they occur also in the heterocyst [9] which represents a specialized cell for nitrogen reduction, and it is believed that they contribute to minimize the content of oxygen in the cell interior. Of these glycolipids, two groups of glycosides have been described, of which in the assumed predominant one α-Glc*p* (● *Fig. 2*) or α-Gal*p* is linked glycosidically to the primary hydroxy group of linear polyhydroxyalkanes or polyhydroxyketones.

In addition to other diacylglycerol-containing glycolipids described earlier, several acylated sulfoquinovosyl-[6,10] and digalacto-[10] diacylglycerol glycolipids have been isolated from different cyanobacteria and were found to be inhibitors of the reverse transcriptase of human immunodeficiency virus (HIV) type 1 [10].

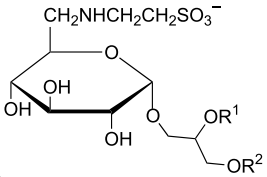
Sulfoquinovosyldiacylglycerol and digalactodiacylglycerol are prominent glycolipids of plants [6,9] where they occur in the plastids; however, members of both lipid types were also identified as major membrane components of bacteroids of *Bradyrhizobium japonicum* [11,12].

Figure 2
Structures of α-D-glucopyranosides and an α-1-*O*-acyl-D-glucopyranose from heterocysts of cyanobacteria

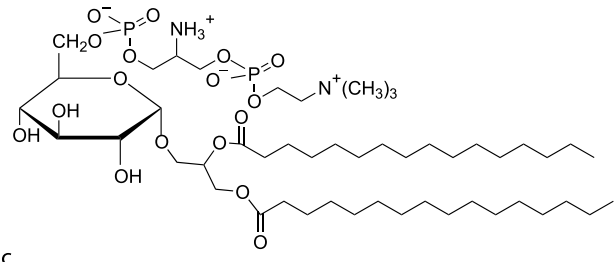
Bacteroids are the nitrogen-fixing form of these and other bacterial species that live symbiotically in leguminoses. Sulfoquinovosyldiacylglycerol lipids have also been identified in phototrophic and a few other bacterial species [6,13]. Different mono- to pentaglyco(di)acylglycerol lipids have been found in membranes of other bacteria possessing a broad variety of sugars and decorations (e. g., methyl or acetyl groups). Another glycolipid, i. e., α-D-Man*p*-(1→3)-6-*O*-acyl-α-D-Man*p*-(1→3)-*O*-acylglycerol (Fig. 3) has been identified in the Gram-positive, coryneform bacterium *Arthrobacter atrocyaneus* [14], and a similar one (different fatty acids) in *Rothia mucilaginosa* [15]. These lipids are unusual with regard to the monoacyl substitution (at *sn*-1) of glycerol and a second fatty acid that is linked to O6 of the first Man residue. Dimannosyl lipids of Gram-positive bacteria have been known for many years; however, it has mostly not unequivocally been shown whether both of their acyl residues are linked to the glycerol moiety. With regard to sugar acylation, α-D-Gal*p*-(1→6)-β-D-Gal*p*-(1→6)-β-D-Gal*p*NAcyl-(1→2)-α-D-Glc*p*-(1→1)-diacylglycerol has been identified in the thermophilic bacterium *Meiothermus taiwanensis* [16] and novel 6-*O*-acyl-α-D-Glc*p*-(1→3)-α-D-Glc*p*-(1→1)-2-*O*-acyl-3-*O*-alkylglycerol [17]. A novel phosphoglycolipid that was alkylated at its glycerol residue was found in thermophilc *Thermus* and *Meiothermus* species [18]. Other 1,2-diacyl-3-α-D-glycopyranosyl-*sn*-glycerols were found in the Gram-negative budding bacterium *Hyphomonas jannaschina* [19], containing as novel glycolipid 1,2-diacyl-3-α-D-glucuronopyranosyl-*sn*-glycerol taurinamide (Fig. 3) [20], *Streptococcus* sp. [21] and *Spirochaeta aurantia* [22], the latter comprising an oligosaccharide of 13 sugars. A phosphocholine-containing amino-



a



b



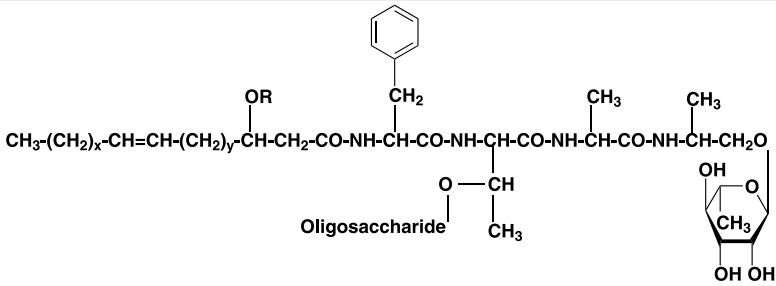
c

Figure 3

Structures of (a) a novel glycolipid from *Arthrobacter atrocyaneus*, (b) 1,2-diacyl-α-D-glucopyranosyl-*sn*-glycerol taurinamide from *Hyphomonas jannaschina*, and (c) a novel phosphocholine-containing glycolipid from *Mycoplasma fermentans*. R^1, predominantly *cis*-16:1ω7 and *cis*-18:1ω7; R^2, predominantly *n*-16:0, *n*-18:0, and *n*-19:0

glycoglycerolipid was isolated from *Mycoplasma fermentans* and its structure characterized (Fig. 3) [23,24]. As a very unusual bacterial glycolipid, cholesteryl 6-*O*-acyl-β-D-galactopyranoside has been isolated from *Borrelia burgdorferi* [25,26]. This molecule developed host immunity during Lyme disease and may be used in diagnosis and vaccination.

Mycobacteria possess three major classes of glycolipids, i. e., trehalose-containing lipooligosaccharides, glyceropeptidolipids (GPL), and phenolic glycolipids [27,28]. Other important cell wall lipoglycans are the lipoarabinomannan (LAM) [29], the lipomannan, and the mycolyl arabinogalactan which is covalently linked to the peptidoglycan [30]. GPL (Fig. 4) are unique to Mycobacteria and are the basis for serotyping in the *M. avium–M. intracellulare* complex. These lipids are also found in other mycobacterial species, e. g., *M. smegmatis* and *M. chelonae*; however, they are not found in species that contain phenolic glycolipids. All



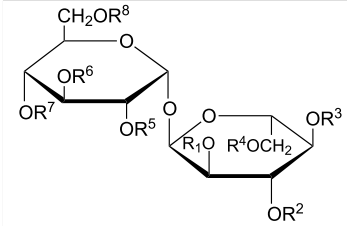
Serovar	Oligosaccharide
2	4-*O*-Ac-2,3-di-*O*-Me-α-L-Fuc*p*-(1→3)-α-L-Rha*p*-(1→2)-6-deoxy-L-Tal*p*
9	4-*O*-Ac-2,3-di-*O*-Me-α-L-Fuc*p*-(1→4)-β-D-GlcpA-(1→4)-2,3-di-*O*-Me-α-L-Fuc*p*-(1→3)-α-L-Rha*p*-(1→2)-6-deoxy-L-Tal*p*
20	2-*O*-Me-α-D-Rha*p*-(1→3)-2-*O*-Me-α-L-Fuc*p*-(1→3)-α-L-Rha*p*-(1→2)-6-deoxy-L-Tal*p*
26	2,4-di-*O*-Me-α-L-Fuc*p*-(1→4)-β-D-Glc*p*A-(1→4)-2-*O*-Me-α-L-Fuc*p*-(1→3)-α-L-Rha*p*-(1→2)-6-deoxy-L-Tal*p*

Figure 4
Examples of GPL from *Mycobacterium avium* serovars. $x = 13–15$, $y = 7–11$, $R = H$ or CH_3

GPL investigated to date possess the disaccharide α-L-Rha*p*-(1→2)-α-L-6-deoxy-Tal*p* that is *O*-glycosidically linked to Thr of the acylated tetrapeptide core. The disaccharide is part of an oligosaccharide, the composition of which is responsible for serospecificity.

α,α-Trehalose-containing lipooligosaccharides (*Fig. 5*) have been found in a variety of mycobacteria, e. g., *M. smegmatis* and *M. kansasii*. These lipids comprise oligosaccharides linked to trehalose which in turn may be di-, tri-, or even tetraacylated at various positions. Trehalose-containing lipids represent another class of mycobacterial glycolipid antigens, also found in other *Actinomycetales* like *Rhodococcus* [31]. Phenolic glycolipids (*Fig. 6*) are found, e. g., in *M. leprae*, in some strains of *M. tuberculosis*, and in *M. bovis*. They consist of a similar aglycone and a variable oligosaccharide that is glycosidically linked to the phenol. Each oligosaccharide represents a specific antigen.

Lipopolysaccharides (LPS, endotoxin) are specific compounds of the Gram-negative cell wall where they occur in the outer leaflet of the outer membrane [32]. They consist of a lipid moiety {the lipid A [33,34] which in general comprises an acylated β-(1→6)-linked amino sugar disaccharide that may be phosphorylated at positions O1 and O4′} which is substituted at O6′ by a polysaccharide that consists of the core region, a heterooligosaccharide that is linked to lipid A [35,36], and the *O*-specific polysaccharide [37,38] that substitutes the core region. LPS of this type are called *S*-form LPS, in distinction to LPS comprising only lipid A and the core region (*R*-form LPS). Both LPS forms can be found in wild-type bacteria. A most general LPS structure is depicted in *Fig. 7*. LPS represent the main surface antigens of Gram-negative bacteria and play an important role in severely infected hosts, i. e., by leading to various pathophysiological reactions such as fever, hypotension, and multi-organ failure.



Species	Substitution pattern
M. fortuitum bv *fortuitum*	R_1, R_2, R_4, and R_5= acylgroups, R_3, R_6, and R_7 = H, R_8 = β-D-Glc*p*
M. tuberculosis Canetti	R_1, R_2, R_4, and R_2–R_4 = acylgroups (proportion = 2:3), R_3, R_5, and R_7 = H, R_6 = 4-acylamino-4,6-dideoxy-Gal*p*-(1→4)-2-*O*-Me-α-L-Fuc*p*-(1→3)-β-D-Glc*p*-(1→3)-2-*O*-Me-α-L-Rhap-(1→3)-2-*O*-Me-α-L-Rha*p*-(1→3)-β-D-Glc*p*-(1→3)-4-*O*-Me-α-L-Rha*p*, R_8 = Me
M. butyricum	R_2, R_3, and R_5 = acylgroups, R_1, R_4, R_6, and R_7 = H, R_8 = 4,6-*O*-(methyl 1-carboxyethylidene)-3-*O*-Me-β-D-Glc*p*-(1→3)-4,6-*O*-(methyl 1-carboxyethylidene)-β-D-Glc*p*-(1→4)-β-D-Glc*p*

Figure 5
Examples of α,α-trehalose-containing lipids from mycobacteria

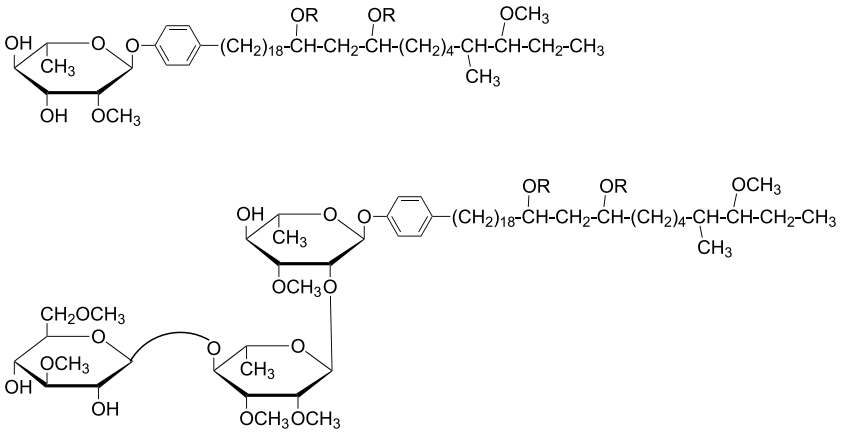


Figure 6
Phenolic glycolipids from *Mycobacterium bovis* (top) and *M. leprae* (bottom)

The characteristic components of the cytoplasmic membrane of Gram-positive bacteria are the amphiphilic macromolecules lipoglycan and lipoteichoic acid [39,40] (*Fig. 8*). Lipoteichoic acids possess in their saccharide chain alditol phosphates as characteristic components, whereas lipoglycans are linear or branched polysaccharides linked to diacylglycerol. These molecules may be further substituted by phosphoglycerol residues. Since lipoteichoic acids and lipoglycans are not found in the same bacterium they are believed to replace each other

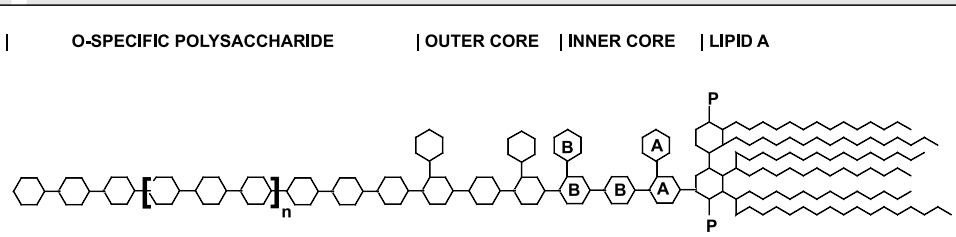


Figure 7
Schematic structure of bacterial *S*-form lipopolysaccharide (LPS) as, in particular, occurring in *Enterobacteriaceae*. *R*-Form LPS consists only of the core region and lipid A. A: 3-desoxy-D-*manno*-oct-2-ulopyranosonic acid (Kdo); B: heptose

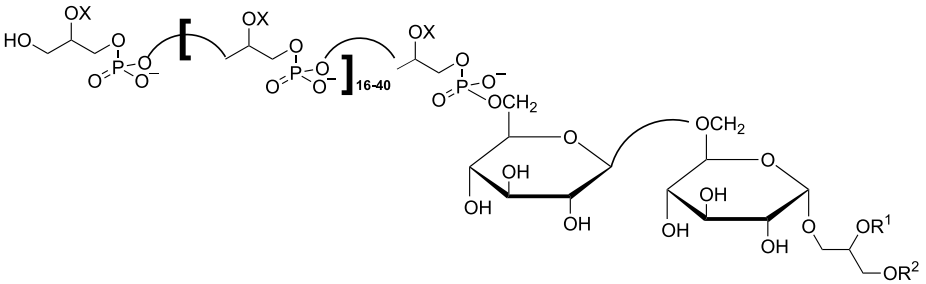


Figure 8
Structure of the most widespread lipoteichoic acid, as found e. g. in baciili, enterococci, listeria, and staphylococci. *x* = H, D-Ala, or glycosyl residues

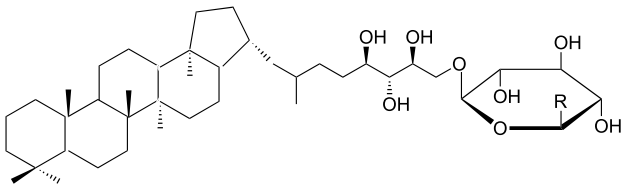


Figure 9
Examples of hopanoid glycosides. *R* = CH_2OH: from *Rhodospirillum acidophila*; *R* = CH_2NH_2: from the cyanobacterium *Synechocystis*; *R* = COOH: from *Rhodospirillum rubrum*

functionally. It should be noted that a putative precursor of lipoteichoic acid biosynthesis has been isolated from *Bacillus subtilis* [41].

Hopanoids are pentacyclic triterpenoids that make up a class of essential membrane lipids [42,43,44]. They are believed to represent the most abundant natural products on earth. They occur widespread in Gram-negative and Gram-positive bacteria, where they play an important role in membrane stability and rigidity. In particular, hopanoids are thought to function as surrogates of cholesterol. Some hopanoids represent glycolipids, some examples of which are depicted in ❯ *Fig. 9*.

3 Plant Glycolipids

Plant glycolipids can be grouped in esters of glucose and sucrose, steryl glycosides, glycosphingolipids, and glycosyl diacylglycerols [9], the last of which occur most abundantly.

Because of their occurrence in all forms of plastids, galactosyl- and sulfoquinovosyldiacylglycerols are the major lipid components in plants (● *Fig. 10*). 1,2-Di-*O*-acyl-3-*O*-β-D-galactopyranosyl-*sn*-glycerol is the precursor of higher galactosyl diacylglycerols formed by sequential addition of α-D-Gal*p* residues to O6 of the preceding Gal*p*. These mono- to tetragalactosyl lipids represent the predominant glycosyl diacylglycerols in plants. However, Glc*p* is also found in such plant lipids, e. g., 1,2-di-*O*-acyl-3-*O*-β-D-glucopyranosyl-*sn*-glycerol in rice bran. Other lipids derive from galactosyl diacylglycerols by addition or removal of fatty acids (● *Fig. 10*), including the acylation at O6 of galactosyl residues. A galactolipid containing (15*R*)-hydroxy-(9*Z*),(12*Z*)-octadecadienoic acid (● *Fig. 10*) has been isolated from oat seeds [45].

Plant glycosphingolipids are *N*-acylated sphinganine derivatives (● *Fig. 11*). It is thought that these lipids play a role in signaling events in plants, analogous to the function of sphingosine derivatives in animals (see below). Two types of glycosphingolipids are *differentiated* in plants, i. e., neutral cerebrosides and phytoglycolipids which are inositol-containing derivatives of ceramide 1-phosphate (● *Fig. 12*). There exist β-D-Man*p* and β-D-Glc*p* monoglycerol ceramides, and it is mainly the latter derivative that is used for chain elongation of up to four sugar residues (● *Fig. 11*). A galactosylceramide which is rather unusual in plants was found in the red alga *Corrallina pilulifera* [46]. However, oligo-α-mannose-type glycans have also been identified, i. e. in *Candida albicans* [47] and *Aspergillus nidulans* [48]. In addition, such glycans have also been found in some basidiomycetes, however, most investigated species contained β-Gal*p* and α-Fuc*p* residues linked to *O*-6 of the Man in the general structure α-D-Man*p*-2Ins1-[PO_4]-Cer [49]. The most abundant amide-linked fatty acids are C_{14}–C_{26} (*R*)-2-hydroxy fatty acids, and in a few cases 2,3-dihydroxy fatty acids have been identified. In plants, a variety of sterol derivatives are also found, e. g., steryl glycosides which possess up to five linearly linked glycosyl residues, and others, like the saponins, that may possess branched carbohydrate moieties. In the former derivatives, the glycosyl residues are linked to the OH group at C3 of ring A (compare with ● *Fig. 13*). A huge variety of glycosyl residues has been identified in steryl glycosides, including β-D-Glc*p*, β-D-Glc*p*A, β-D-Gal*p*, α-D-Rib*f*, and others. Furthermore, there exist acylated derivatives, which usually carry a fatty acid (e. g., C_{16} or C_{18}) at C6 of Glcp; however, other acyl derivatives have also been identified occasionally. Steryl biosides that originate from steryl glycosides via further glycosylation are also frequently observed. Their acylated derivatives, however, occur only in rare cases.

Esters of glucose and sucrose have been identified as viscous exudates in *Solanaceae* species in amounts of up to 25% of the plant dry mass. In the first group, di- to penta-*O*-acyl-α- and -β-D-glucopyranoses occur, containing acetyl and (branched) acyl chains of up to C_{12} without any general acylation pattern. In sucrose esters, generally the Glc residue is higher, even fully, acylated than the fructose residue. A novel oxo-phytodienoic acid-containing galactolipid (● *Fig. 14*) has been identified in *Arabidopsis thaliana* [50], as well as a novel class of oxylipins, namely *sn*1-*O*-(12-oxophytodienol)-*sn*2-*O*-(hexadecatrienoyl)-monogalactosyl diglyceride [51].

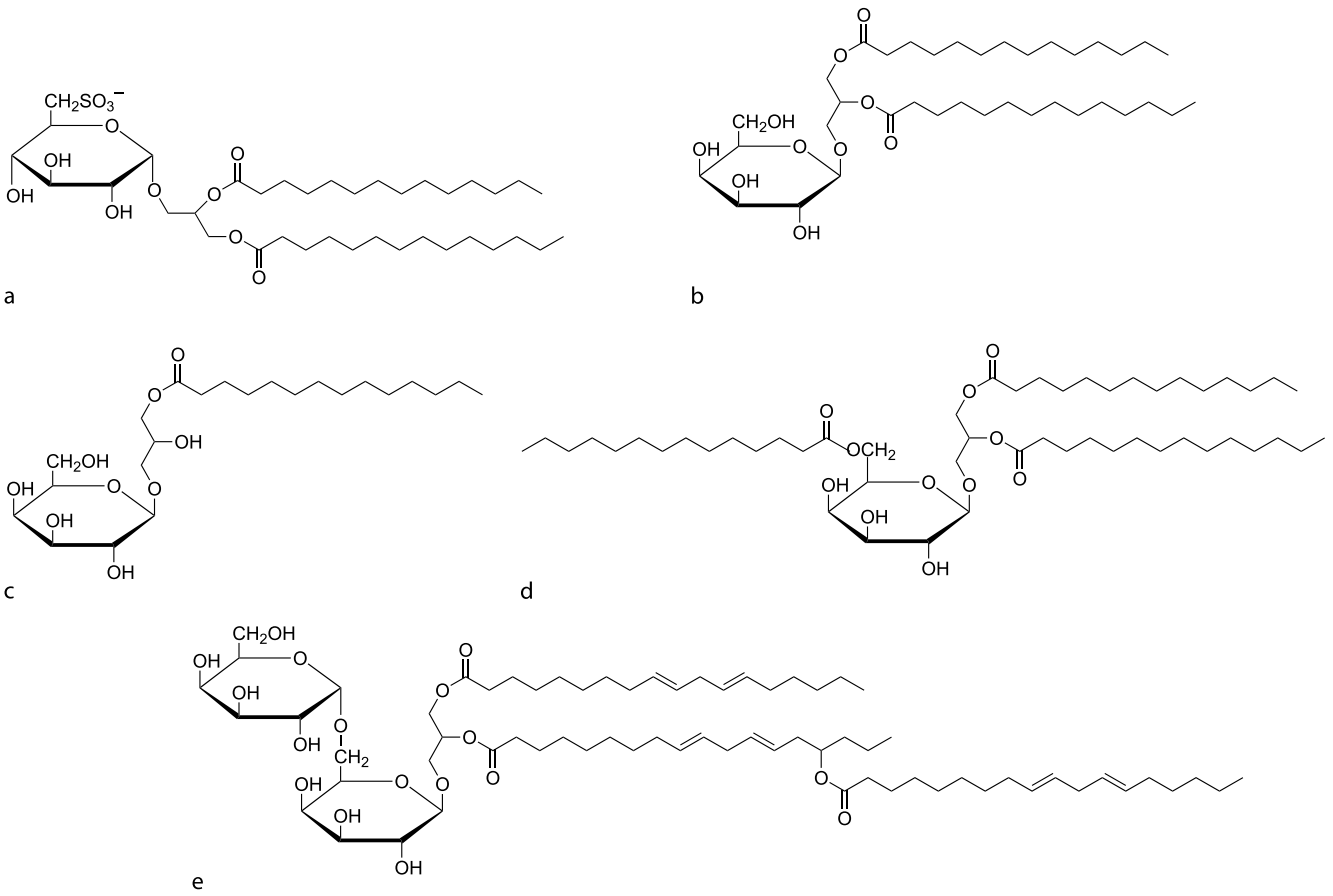


Figure 10
(a–d) Structures of glycosyl acylglycerols from plants, (e) the structure of a novel galactolipid from oat seeds

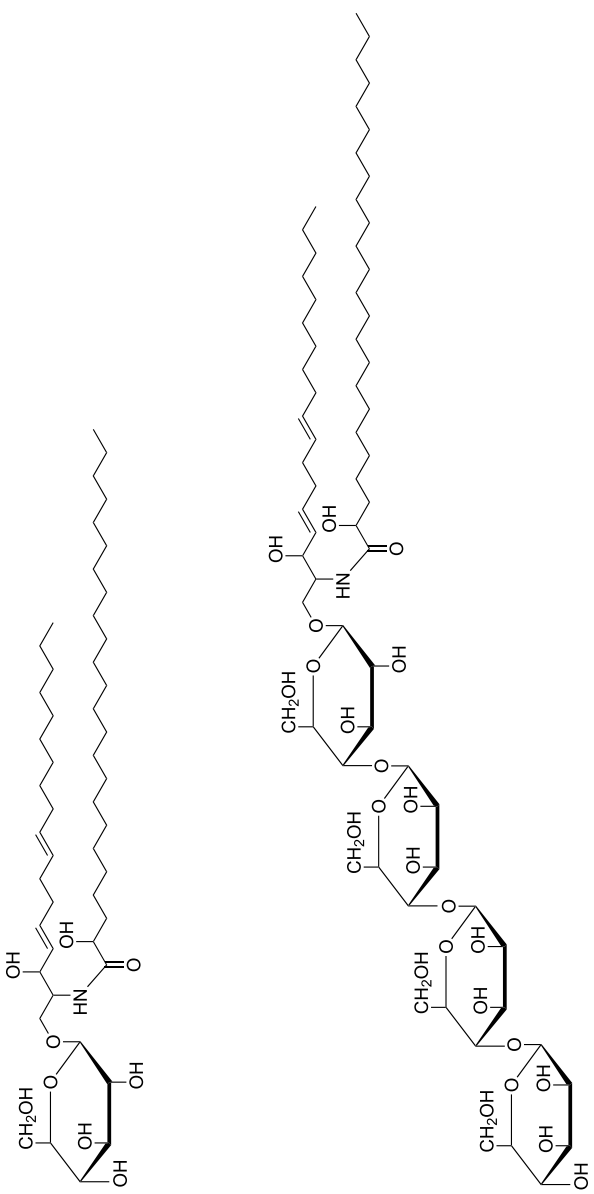


Figure 11
Structures of neutral cerebrosides from plants

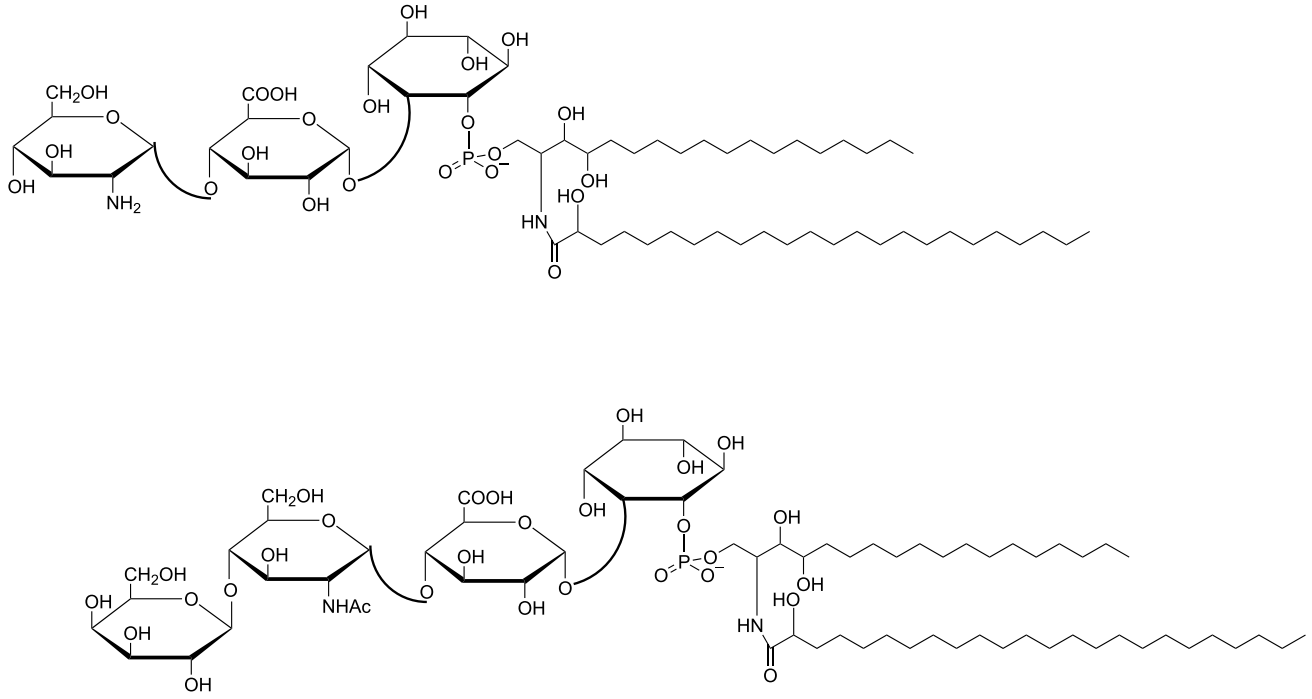


Figure 12
Phytoglycolipid structures

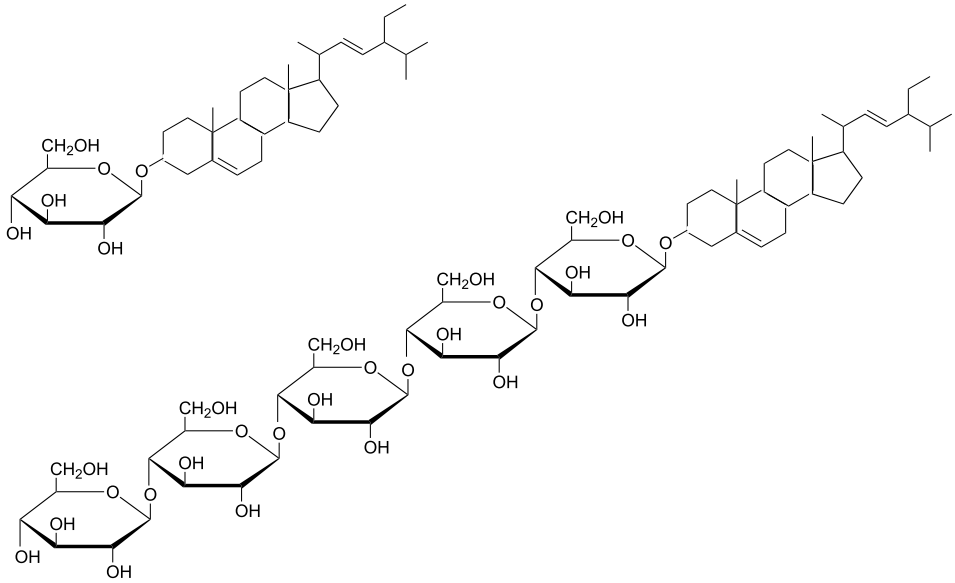


Figure 13
Structures of steryl glycosides from plants

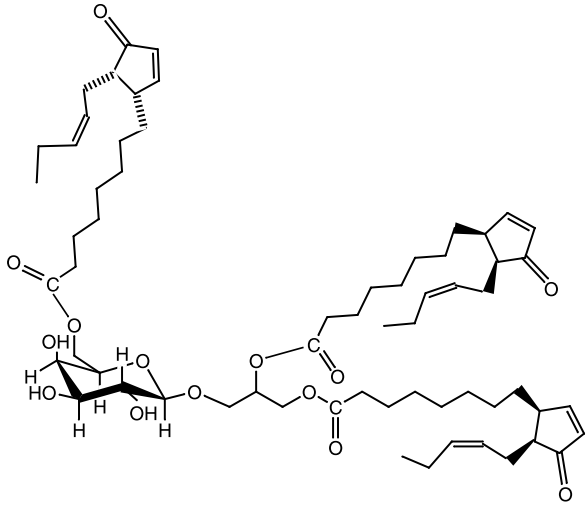


Figure 14
The structure of the oxophytodienoic acid-containing galactolipid from *Arabidopsis thaliana*

4 Animal Glycolipids

In animal tissues, glycolipids are important constituents of the plasma membrane. They play a significant role in cell-cell communications, cellular differentiation, and proliferation, and also in oncogenesis, as indicated by alterations in their compositions and in glycosyltransferase activities during differentiation and growth of cells.

The most prominent lipid anchor of animal glycolipids is ceramide (*Fig. 15*) which consists of sphingosine that is substituted at its amino group by fatty acids of various chain lengths. Sugars are linked glycosidically to ceramide at the primary hydroxy group of sphingosine. Lipids of this type are called glycosphingolipids. The glycosphingolipids together with sphingomyelin and cholesterol are clustered in microdomains that serve in membrane-linked processes [52,53,54], e. g., signal transduction and internalization of bacteria and viruses.

Glycosphingolipids [55] can be divided in the two groups: neutral and acidic glycosphingolipids. Both comprise mono- or oligosaccharidic sugar moieties. Negative charges in the latter group can be provided by sialic acids, uronic acids, sulfates, phosphates, and phosphonates. Of the neutral glycosphingolipids, galactosylceramide (*Fig. 13*) is the most prominent member in the mammalian brain. This molecule and its sulfated form, i. e., 3-*O*-sulfogalactosylceramide (*Fig. 15*) belong to the dominant and essential glycolipids of the myelin layer of axons [56,57] where they are present in the outer leaflet of the myelin membrane [56], together with cholesterol and sphingomyelin. Less polar glycolipids than GalCer which co-appear with it during myelinogenesis were purified from rat brain. Two structures were identified, i. e. 3-*O*-acetyl-sphingosine-GalCer with non-hydroxy and hydroxy fatty *N*-acylation, respectively [58]. Other varieties, namely plakosides and agelasphins, were isolated from sponges and comprise a bis-hydroxylated ceramide moiety [59].

Higher glycosylated glycosphingolipids are found in all tissues. They are grouped in series, e. g., 'ganglio', 'globo', or 'lacto' in order to distinguish them with regard to structural and functional relationships (*Fig. 15*, *Table 1*). Other higher glycosylated glycosphingolipids were identified in *Schistosoma mansonii* eggs [60], carrying α-Fuc-(1→3)-GalNAc, β-GalNAc-(1→4)-[α-Fuc-(1→3)-GlcNAc], and β-Gal-(1→4)-[α-Fuc-(1→3)-GlcNAc] termini. From the tapeworm *Diphyllobothrium hottai*, ten glycosphingolipids were isolated from adults and four from plerocercoids, comprising mono—tetrasaccharide structures [61]. β-Gal-(1→4)-[α-Fuc-(1→3)]β-Glc-(1→3)-β-Gal-(1→Cer was found only in adults, whereas β-Gal-(1→4)-[α-Fuc-(1→3)]β-Glc-(1→3)-[β-Gal-(1→6)-]β-Gal-(1→Cer was identified in both, adults and plerocercoids. A novel class of fucosylated glycosphingolipids containing polyunsaturated very long chain fatty acids was identified to be essential for spermatogenesis and male fertility in mice [62].

Of the acidic glycosphingolipids, gangliosides are characterized by the presence of one or more sialic acid residues which are linked to O3 or O6 of Gal or GalNAc of the core oligosaccharides [63]. Polysialogangliosides were identified in skate brain [64]. Gangliosides are present in the outer leaflet of the plasma membrane but have also been identified, e. g., in cytosol membranes of some endocrine cells [65]. A broad variety of gangliosides was identified as tumor- and blood-group antigens. Furthermore, gangliosides may also act in autoimmune disorders as autoantigens, e. g., in the Guillain–Barre syndrome and in autoimmune diabetes [65], and there is good evidence that various gangliosides are able to modify the activities of functionally different receptor molecules [66]. In the nervous system, the composition of gangliosides is dif-

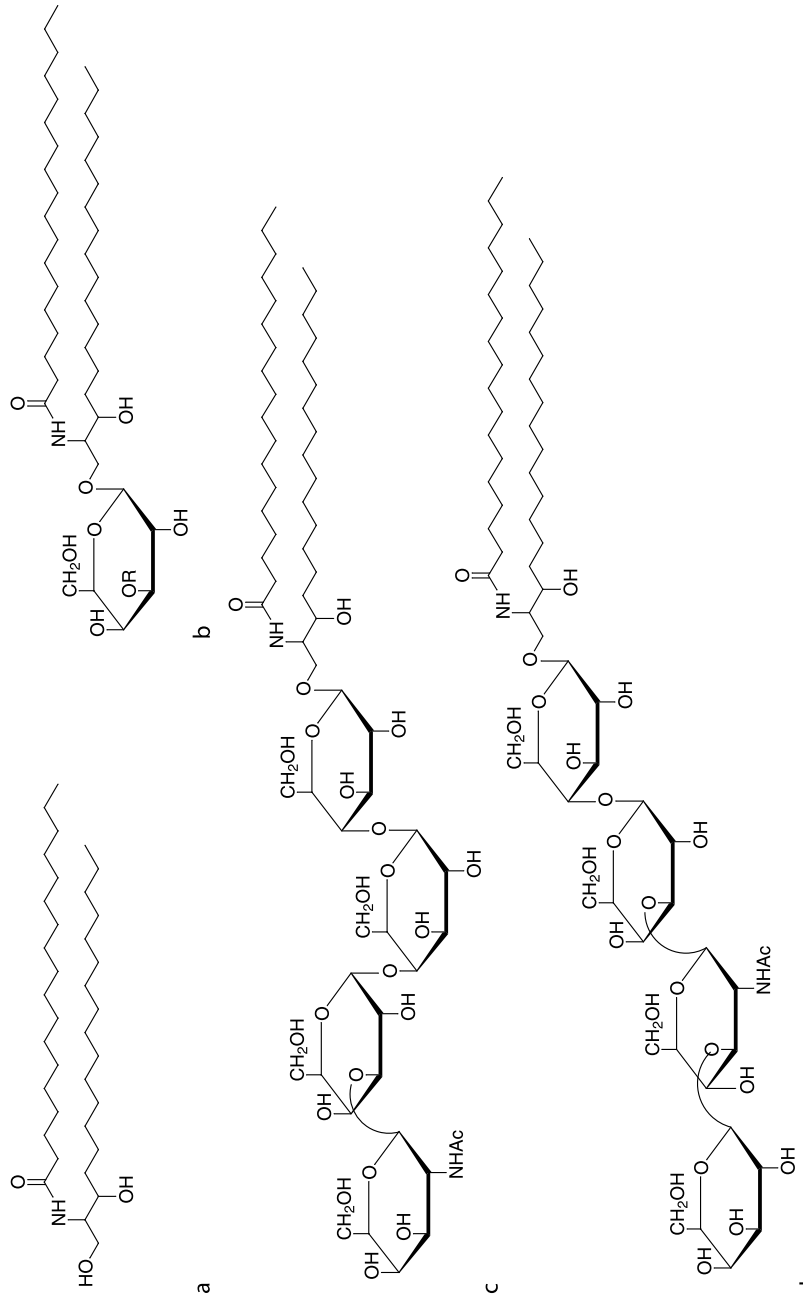


Figure 15
(a) Ceramide, (b) galactosyl ceramide (R = H) and 3-*O*-sulfogalactosylceramide ($R = SO_3^-$), (c) globotetraosylceramide, (d) lactotetraosylceramide

Table 1
Root names and structures of oligosaccharide chains of neutral glycosphingolipids

Root	Structure
ganglio	β-D-Gal*p*-(1→3)-β-D-Gal*p*NAc-(1→4)-β-D-Gal*p*-(1→4)-D-Glc*p*-
neolacto	β-D-Gal*p*-(1→4)-β-D-Glc*p*NAc-(1→3)-β-D-Gal*p*-(1→4)-D-Glc*p*-
isoglobo	β-D-Gal*p*NAc-(1→3)-α-D-Gal*p*-(1→3)-β-D-Gal*p*-(1→4)-D-Glc*p*-
mollu	β-D-Glc*p*NAc-(1→2)-α-D-Man*p*-(1→3)-β-D-Man*p*-(1→4)-D-Glc*p*-
arthro	β-D-Gal*p*NAc-(1→4)-β-D-Glc*p*NAc-(1→3)-β-D-Man*p*-(1→4)-D-Glc*p*-

ferent in the peripheral nervous system (abundance of the neolacto series, compare *Table 1*) and the central nervous system (abundance of the globo series) [67]. Recently, novel hybrid gangliosides comprising isolacto-, neolacto- and ganglio-core were isolated from the Pacific salmon, *Oncorhynchus keta* [68].

Another group of the acidic glycosphingolipids are the sulfoglycosphingolipids (formerly called 'sulfatides'). This field is comprehensively reviewed in [6]. Most abundant in mammalian tissues is 3-*O*-sulfogalactosylceramide (*Fig. 15*) which occurs, e. g., in human kidney, rat lung, and in the human central and peripheral nervous system. Also, lactosyl II^3-sulfate ceramide is found in mammalian tissues, e. g., in human kidney and liver. Higher glycosylated molecules are present, e. g., in human kidney, and in the rat and human central and peripheral nervous systems.

Glycosphingolipids are also present in invertebrates. Insects contain arthrosides (see *Table 1*) that are structurally similar to gangliosides but contain Man instead of Gal and GlcUA instead of sialic acids [69]. In Drosophila, various zwitterionic and acidic glycosphingolipids were found [70]. Zwitterionic glycosphingolipids have also been identified in worms like *Ascaris suum* [71] or in the Annelida (for references see [71]). Interestingly, a glycosphingolipid with the structure α-D-Gal*p*-(1→2)-Ino-(1→P-Cer was identified in *Ascaris suum* [72] which represents the first 'phytoglycolipid' (see above) that was identified in animals. A highly fucosylated glycosphingolipid with a GlcNAc-(1→3)-GalNAc-(1→4)-Glc-(1→Cer core was identified in eggs of *Schistosoma mansonii* [73]. Also, a broad variety of marine organisms like sponges, starfish, and mollusks possess neutral and phosphorus-containing glycosphingolipids and gangliosides, and sulfoglycosphingolipids, e. g., 3-*O*-sulfogalactosylceramide. This field is excellently reviewed in [46].

Sponges contain also other interesting glycolipids, as was shown in a case study with *Plakortis simplex* [59]. In particular, polyisoprenoid glycolipids like plakopolyprenoside and plaxyloside possessing a C-35 linear polyisoprenoid alcohol as aglycon and a β-Xyl*p*-composed sugar chain were identified, and simplexides, containing a very long chain secondary alcohol (34–37 C-atoms) as lipid moiety, which is hydroxylated in the middle of the chain. This hydroxy group is glycosylated by a Glc*p*-(1→4)-Gal*p* disaccharide.

In protozoa, e. g., *Leishmania* and *Trypanosoma*, glycolipids have been identified that possess a GPI anchor (see Chap. 7.4) [74]. Characteristic for *Leishmania* is lipophosphoglycan [74,75,76,77] which covers the entire surface of the promastigote. The molecule can be divided in four regions, i. e., lipid anchor, core region, repeating units, and cap, the last of which vary in different species.

5 Properties

5.1 Lipopolysaccharides

Like other amphiphilic molecules, LPS aggregate and build up clusters in aqueous solutions. This process occurs only above a critical aggregate concentration (CAC) which has been established for a few lipids, but not, however, for LPS owing to extreme experimental difficulties. Estimations based on comparisons to values of other lipids led to the assumption of a CAC of $< 10^{-10}$ M for hexaacyl-lipid A (*Fig. 16*). For the lipid A precursor IVa (*Fig. 16*) which represents a tetraacyl-lipid A, a CAC of $< 10^{-7}$ has been reported [78].

Since the conformation of a particular LPS molecule depends on the phase state of its hydrocarbon moiety, it is variable and, thus, may exist in a highly ordered gel state or in a less ordered liquid-crystalline state. These states and the changes from one to the other are dependent on the phase-transition temperature and other parameters, like water and ion concentration [79]. The state of order of LPS with a longer saccharide moiety shows a stronger dependence on water and ions than that of LPS with shorter saccharide chains. The addition of cations, e. g., Mg^{2+}, results in the aggregation of LPS due to intermolecular bridging.

Synthetic lipid A analogues were used in FTIR investigations to study possible intermolecular conformations of neighboring molecules [79]. There is good evidence that the bisphosphorylated β-(1→6)-linked GlcN disaccharide backbone of lipid A is inclined (20–40° relative to the membrane normal), thus, the phosphate linked to O1 is reaching to the outside and that at O4′ is buried in the membrane. This model is supported by data from transition temperature measurements and calorimetric experiments. However, it should be noted that other authors reported the reverse conformation with the O4′ phosphate on the surface and the O1 phosphate in the membrane [80,81].

The determination of the relation between the conformation and the supramolecular structure of a given LPS molecule is possible using a simple geometric model described by $S = a_h / a_o$

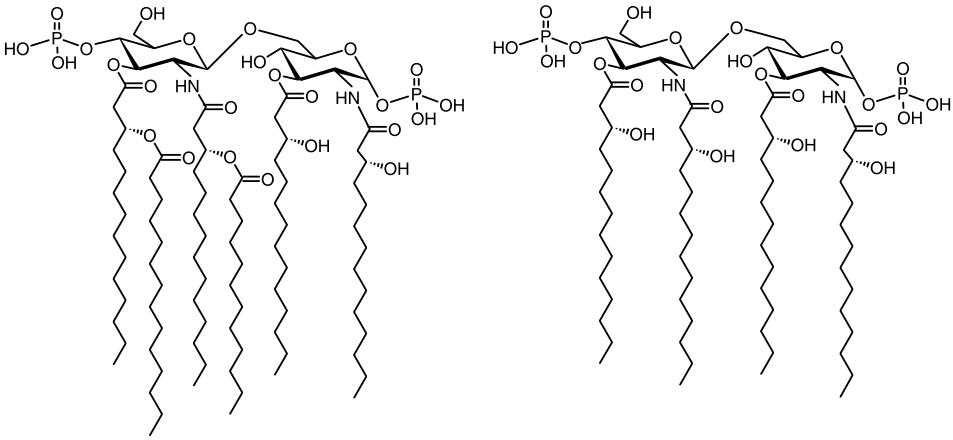


Figure 16
Chemical structures of hexaacyl-lipid A from *Escherichia coli* (left) and precursor IVA from *Escherichia coli* (right)

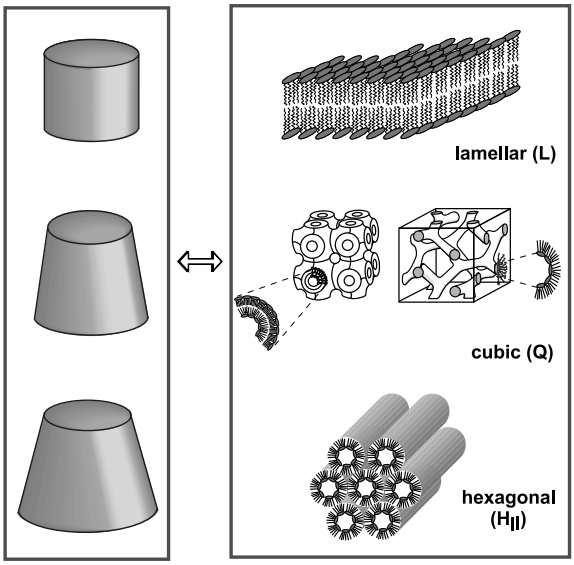


Figure 17
Molecular conformations and supramolecular structures of endotoxins

(*S*, final structure; a_h, effective cross-sectional areas of the hydrophilic region; a_o, effective cross-sectional areas of the hydrophobic region) [82,83]. The value for a_h is strongly influenced by *gauche*-conformers in the acyl chain: their presence (liquid-crystalline state) provokes an increase of a_h, their absence (gel state) a decrease. The presence of ions and variations in the hydration state influence the value for a_o. Furthermore, headgroup charges and/or asymmetric distribution of acyl chains leading to the presence or absence of an axis of the molecule determine the supramolecular structure, the most relevant of which are shown in *Fig. 17*. They are the lamellar, cubic, and hexagonal structures [84]. The phase behavior of LPS and lipid A was shown to depend on the length of the sugar moiety that is present in the molecule [78,83,85,86,87].

It has been well established that the expression of the full biological activity of the lipid A molecule is due to a particular chemical structure, i. e., a bisphosphorylated β-(1→6)-linked Glc*N*-disaccharide that is acylated by 6 fatty acids of chain lengths between C_{10} and C_{16} [88]. Any change in this chemical structure leads to changes of the physical conformation and of the biological properties. Thus, it is a unique lipid A conformation that is responsible for biological activity, a concept that could be verified by comparison of chemical structures of different lipid A and their respective biological activities [89,90,91,92,93,94,95,96,97].

5.2 Lipoteichoic Acids

In distinction to LPS, the critical micellar concentration of various lipoteichoic acids (LTA) has been measured and determined to be in the range of $2.8 \cdot 10^{-7}$ M to $5 \cdot 10^{-6}$ M [98,99], indicating a significantly lower hydrophobicity than for glycerophospholipids which is due to the

large hydrophilic carbohydrate chains [100]. Independent of the different structures of their hydrophilic regions, LTA possess similar spherical micellar structures in aqueous solutions, even over a wide range of temperature [101,102]. It has been determined that LTA of *Streptococcus pneumoniae* furnishes an average micelle of about 150 molecules with a diameter of 22 nm, in which an outer layer of 8.5 nm thickness is built up by the hydrophilic part of the LTA, indicating a highly coiled conformation of this region [102]. Since the effective cross-sectional area of the lipid moieties of LTA is smaller than that of the hydrophilic carbohydrate chains, LTA possess a conical supramolecular structure which does not allow the molecule to form membrane layers. Thus, LTA can only be inserted into a layer of the lamellar phase lipids, and it has been shown that stable membranes formed from membrane lipids and LTA can be obtained [103].

5.3 Mycolates

The structure of the cell envelope of mycobacteria [30,104] is highly complex. Outside the cytoplasma membrane the so-called cell wall skeleton is present, consisting of the three covalently attached polymers, peptidoglycan, arabinogalactan, and mycolic acids. Also present in this skeleton are lipoarabinomannan and lipomannan, and a variety of (glyco)lipids. Mycolic acids are long-chain α-alkyl-β-hydroxyfatty acids that often comprise more than 80 C-atoms and possess functional groups such as *cis*- or *trans*-cyclopropanes, α-methyl methyl ethers, or α-methyl ketones. They are ester-linked to terminal arabinose residues of the arabinogalactan and extend perpendicular to this and the peptidoglycan. Since other lipids (possibly glycolipids) intercalate into the layer of mycolic acids, some kind of bilayer is formed. The mycolic acids furnish the inner leaflet of this bilayer and represent an effective barrier to the penetration of antibiotics and chemotherapeutic agents.

Evidence for this mycolate leaflet has been obtained, among others, by using X-ray diffraction on purified mycobacterial envelopes which gave a strong reflection at 4.2 Å and a rather diffuse reflection at 4.5 Å that were assigned to highly and less ordered fatty acyl chain regions, respectively [105]. Also, the perpendicular orientation of the acyl chains was demonstrated. Such purified mycobacterial cell walls have also been investigated using differential scanning calorimetry in order to examine phase transitions [106]. One important observation was that the lipids of the cell wall possess a melting point above the highest temperature at which the bacteria are viable, and that this is significantly higher than that of other biological membranes. This confirmed the presence of a compact and highly ordered mycolate layer in the cell wall. Further investigations employing electron-spin resonance experiments using spin-labeled fatty acids showed that these enter less ordered regions (more fluid), possibly occupied by the less ordered lipids that build up the outer leaflet of the bilayer and intercalate into the mycolate leaflet. Taken together, evidence is provided that an outer permeability barrier is present in the mycobacterial cell wall, which is constructed as an asymmetric membrane comprising a highly ordered inner (mycolates) and a less ordered outer (lipids) leaflet.

5.4 Hopanoids

Hopanoids are thought to represent surrogates of cholesterol in bacteria, acting as efficient membrane stabilizers. This is supported by the findings that both sterols and hopanoids are

(i) amphiphilic molecules, (ii) possess a rigid, hydrophobic polycyclic skeleton and a polar headgroup, and (iii) possess close conformational similarity, as shown by molecular modeling of cholesterol and bacteriohopanetetrol (BHT) [43,44].
Other proof was obtained from experiments with artificial membranes in which BHT or its glycoside isolated from *Bacillus acidocaldarius* led to a condensation of C16:0 and abolished the clear phase transition between the gel-like and the liquid-crystalline phases. Furthermore, it was found that 40 mol-% of BHT completely suppresses this phase transition in phospholipid bilayers.

5.5 Glycoglycerolipids

Glycoglycerolipids represent one of the major glycolipid classes, and they play important functional roles in furnishing membranes and the maintenance of membrane fluidity as well as in specific lipid-protein interactions [107,108,109] and cell surface recognition. In the latter functions (as well in those of glycosphingolipids), the relation between structure and action is determined by the sequence and conformation of the saccharide chains and, however, influenced by the structures of their lipid moieties and the lipid environment of the whole molecule. Conformational analyses of glucosyldiacylglycerol (Glc-DAG, and glucosyldialkylglycerol) using molecular mechanics indicated that in the membrane only a reduced number of possible conformations appear [110]. Obviously, the saccharide conformation of GlcDAG is in lipid-water systems in the hexagonal phase, owing to a close headgroup packing which is in clear difference to the lamellar arrangement of glycosylceramides.
Model membranes have been used to investigate the influence on membrane structure and function of both polar headgroups and hydrophobic regions of glycolipids, by application of thermodynamic, structural, and spectroscopic methods. With regard to the hydrophobic moiety, the stability of various phases of the lipids was studied with a special focus on length and degree of saturation of the alkyl chains [111]. Investigations using calorimetry showed that size and stereochemistry of the carbohydrate moiety, the anomeric configuration of the glycosidic linkage of the sugar to glycerol, the charge of the lipid moiety, and the stereochemistry of glycerol are essential for transition characteristics. With regard to the hydrophobic moiety, the temperatures for the transitions from lamellar gel to the lamellar liquid crystalline phase ($L_\beta \rightarrow L_\alpha$) and from the lamellar crystalline (L_c) phase to L_α of α-Glc containing DAG increase with the presence of longer acyl chains. On the other hand, the high-temperature transition L_α to non-lamellar phase decreases with increasing chain lengths, e. g., for a di-C18:0 glycerol a direct transition from L_α to the inverted hexagonal liquid crystalline phase is observed [112,113,114,115]. A change from the α- to the β-Glc anomer has some impact on the phase properties. Between the L_c and L_β phases, an additional 'L_c-like' phase is observed. The transition $L_\beta \rightarrow L_\alpha$ is also faster than that for α-glucosides, and the final L_c phase is less stable. Thus, a direct $L_c H_{II}$ transition is not present in the case of β-glucodiacylglycerols, but, instead, two-step $L_c \rightarrow L_\beta \rightarrow H_{II}$ or three-step $L_c \rightarrow L_\beta \rightarrow L_\alpha \rightarrow H_{II}$ transitions have been observed for C20:0/C20:0 and C18:0/C18:0 β-glucodiacylglycerols, respectively. Thermodynamic and NMR data indicate that β-glucodiacylglycerols tend to form inverted non-bilayer phases. Between β-gluco- and β-galactodiacylglycerols, the only difference is the higher stability of the L_c phase of the latter. Although only few data exist in the literature, it seems that mannodiacylglycerols possess a rather simple phase transition behavior, compared to that of

gluco- and galactodiacylglycerols. Glucuronoglycerolipids show a different phase transition behavior, owing to the negative charge of the carboxyl group at C6 of the sugar moiety. As expected, the charge destabilizes the rather ordered gel-phase at low temperature, resulting in a decreased phase transition temperature. At pH values above the pK of the carboxyl group, the lamellar gel phase undergoes a reversible structural change between 40 and 50 °C which is below the melting point at about 69 °C and indicates a disordered bilayer stacking. In di- and trisaccharide diacylglycerols, the introduction of a second or third sugar residue disables these lipids to form non-bilayer structures. Here, the L_α phase represents the final high-temperature phase. It should be noted that a computerized database (LIPDAT) is available which provides the reader with published data on phase behavior of synthetic, semisynthetic, and natural polar lipids [111].

Differential scanning calorimetry, differential scanning densimetry, and small and wide angle X-ray scattering have been used to investigate the polymerization of β-gluco-, β-galacto-, and β-maltodiacylglycerols with varying acyl chain lengths [116]. The aim of the study was to elucidate (i) whether the structural changes occur synchronous with the enthalpy and volume changes and (ii) whether the cooperativity of the transitions is predominantly determined by the geometry of the phase change.

The results show that enthalpy and changes of volume and structure occur in synchrony with the transition temperature. All phase changes may be described under equilibrium conditions as two-state processes. The value for the size of the cooperative unit is the same, independent of the method by which it was determined. The lowest cooperative unit was found during the L_cL_β and L_cL_β transitions (about 37), whereas $L_\beta L_\alpha$ possessed a medium (about 120) and $L_\alpha H_{II}$ a high (several hundreds) number.

The physicochemical properties of a major glycolipid from *Mycoplasma fermentans* [24], namely 6′-*O*-(3″-phosphocholine-2″-amino-1″-phospho-1′,′3″-propanediol)-α-D-glucopyranosyl-(1′→3′)-1,2-diacyl-*sn*-glycerol (MfGl-II), have been characterized [117,118]. Interestingly, the properties of this molecule resembled those of LPS/lipid A, i. e., the $\beta \rightarrow \alpha$ gel-to-liquid crystalline phase transition behaviour of the hydroacarbon chains had a T_c of 30 °C, and a unilamellar/cubic supramolecular structure was identified. At high water contents, MfGl-II formed a L_α phase and isotropic aggregates (vesicles).

5.6 Glycosphingolipids

The LIPDAT data base contains also thermodynamic data that have been published for synthetic, semisynthetic, and natural glycosphingolipids [119]. Of the latter, sphingomyelins, gangliosides, and cerebrosides have been investigated. Natural sphingomyelins undergo mesophase transitions between 20 and 40 °C. On the basis of these and other data and the packing mode of long-chain sphingomyelins, it is thought that these lipids are important for membrane transport. Of the cerebrosides, mainly the monoglycosides have been investigated, some of which possess low-temperature gel phase metastability. A majority of molecules displays unusually high lamellar order to disorder temperatures between 50 and 70 °C, indicating their role in the structural integrity of membranes. The large difference in the gel-to-liquid phase transition temperatures [120] of monoglycosylcerebrosides and phospholipids indicate a phase-separation-induced clustering of the former in membranes, an observation that is indeed of high biological significance (see below). FTIR was used to investigate the transfer of metastable

gluco- and galactocerebrosides to their stable forms [121], and the data led to the conclusion that the presence of an α-OH group in the acyl chain results in a less tightly packed structure, owing to reduced intermolecular hydrogen bonding. These findings were in accord with the reduced enthalpy of the gel-to-liquid crystalline phase transition. Other investigations using bovine brain cerebrosides [122] indicated the presence of two gel states, one of which is in the metastable form. The shift from the metastable to the stable form showed again the formation of hydrogen bonds. The spectra of the metastable and the anhydrous forms were very similar which confirmed the observed identical X-ray diffraction patterns for both forms [124]. In gangliosides, sialic acid residues introduce negative charges to the polar headgroup. The general involvement of gangliosides in the modification of membrane functions is obviously based on their tendency to aggregate. Higher oligosaccharide-containing gangliosides form micelles (critical micellar concentration 10^{-8}–10^{-5} M).

As mentioned above, the different physical properties of sphingolipids and phospholipids in membranes results in clustering of the former, i. e., in microdomain formation [125,126]. This process is furthermore influenced by the differences in the length of the acyl chains which range from 20 to 26 carbon atoms in glycosphingolipids and from 16 to 22 carbon atoms in phospholipids, and by the degree of saturation which occurs more often in the latter class. A model of microdomains (or rafts, in the case where they are associated with GPI-anchored proteins) proposes densely packed glycosphingolipids and cholesterol molecules. Rafts seem to be important in membrane traffic and signal transduction events.

Glycosphingolipids are thought to play important roles in cell recognition and adhesion processes which involve the interaction between the carbohydrate moieties. Thus, saccharide headgroups should possess preferred conformations. In accord with this and as with glycosyldiacylglycerols, the number of possible conformations of glycosphingolipids seems to be reduced by hindrance of the membrane environment, as was measured for glucosylceramide [110]. The orientation of the saccharide chain is furthermore dependent on intrinsic preferences of the saccharide-ceramide linkage. Recently, the adhesion between lactosylceramide-coated bilayers was measured, and it was found that the non-specific van der Waals attraction between the membranes is not increased by the presence of the carbohydrate moiety and, thus, is weak [126]. Therefore, in this case the measured forces can neither be strong enough to overcome intercellular electrostatic repulsion nor be responsible to arrest cells. However, it should be noted that only few such investigations have been performed; thus, this field awaits further clarification.

Acknowledgments

I am indebted to Prof. Dr. U. Seydel (Research Center Borstel) for help with *Fig. 17*.

References

1. IUPAC-IUB Joint Commission on Biochemical Nomenclature (JCBN) (1998) Eur J Biochem 257:293
2. Luzzati V (1968) X-ray diffraction studies of lipid-water systems. In: Chapman D (ed) Biological membranes, physical fact and function. Academic Press, London, p 71

3. Shinitzky M (1984) Membrane fluidity and cellular function. In: Shinitzky M (ed) Physiology of membrane fluidity, vol. 1. CRC Press, Boca Raton, p 1
4. Cossins AR, Sinensky M (1984) Adaptation of membranes to temperature, pressure, and exogenous lipids. In: Shinitzki M (ed) Physiology of membrane fluidity, vol. 2. CRC Press, Boca Raton, p 1
5. Kates M (1996) J Microbiol Meth 25:113
6. Ishizuka I (1997) Progr Lipid Res 36:245
7. Corcelli A, Colella M, Mascolo G, Fanizzi FP, Kates M (2000) Biochemistry 39:3318
8. Lobasso S, Lopalco P, Lattanzio VMT, Corcelli A (2003) J Lipid Res 44:2120
9. Heinz E (1996) Plant glycolipids: structure, isolation and analysis. In: Christie WW (ed) Lipid methodology–three. The Oily Press, Dundee, UK, p 211
10. Reshef V, Mizradri E, Maretzki T, Silberstein C, Loya S, Hizi A, Carmeli S (1997) J Nat Prod 60:1251
11. Cedergren RA, Hollingsworth RI (1994) J Lipid Res 35:1452
12. Tang Y, Hollingsworth RI (1997) Glycobiology 7:935
13. Kates M (1990) Glyco-, phosphoglyco-, and sulfoglycoglycerolipids of bacteria. In: Kates M (ed) Handbook of lipid research. 6. Glycolipids, phosphoglycolipids, and sulfoglycolipids. Plenum Press, New York
14. Niepel T, Meyer H, Wray V, Abraham WR (1997) Tetrahedon 53:3593
15. Pasciak M, Holst O, Lindner B, Mierzchala M, Grzegorzewicz A, Mordarska H, Gamian A (2004) Biochim Biophys Acta 1675:54
16. Yang FL, Lu CP, Chen CS, Chen MY, Hsiao HL, Su Y, Tsay SS, Zou W, Wu SH (2004) Eur J Biochem 271:4545
17. Pasciak M, Holst O, Lindner B, Mordarska H, Gamian A (2003) J Biol Chem 278:3948
18. Yang YL, Yang FL, Jao SC, Chen MY, Tsay SS, Zou W, Wu SH (2006) J Lipid Res 47:1823
19. Batrakov SG, Nikitin DJ, Pitryuk IA (1996) Biochim Biophys Acta 1303:39
20. Batrakov SG, Nikitin DJ, Pitryuk IA (1996) Biochim Biophys Acta 1302:167
21. Roethlisberger P, Iida-Tanaka N, Hollemeyer K, Heinzle E, Ishizuka I, Fischer W (2000) Eur J Biochem 267: 5520
22. Vinogradov E, Paul CJ, Li J, Zhou Y, Lyle EA, Taping RI, Kropinski AM, Perry MB (2004) Eur J Biochem 271:4685
23. Matsuda K, Ishizuka I, Kasami T, Handa S, Yamamoto N, Taki T (1997) Biochim Biophys Acta 1349:1
24. Zähringer U, Wagner F, Rietschel ET, Ben-Menachem G, Deutsch J, Rottem S (1997) J Biol Chem 272:26262
25. Ben-Menachem G, Kubler-Kielb J, Coxon B, Yergey A, Schneerson R (2003) Proc Natl Acad Sci USA 100:7913
26. Schröder NWJ, Schombel U, Heine H, Göbel UB, Zähringer U, Schumann RR (2003) J Biol Chem 278:33645
27. Aspinall GO, Chatterjee D, Brennan PJ (1995) Adv Carbohydr Chem Biochem 51:169
28. Daffé M, Draper P (1998) Adv Microbial Physiol 39:131
29. Chatterjee D, Khoo KH (1998) Glycobiology 8:113
30. Brennan PJ, Nikaido H (1995) Ann Rev Biochem 64:29
31. Ueda S, Fujiwara N, Naka T, Sakaguchi I, Ozeki Y, Yano I, Kasama T, Kobayashi K (2001) Microb Pathogen 30:91
32. Raetz CRH, Whitfield C (2002) Ann Rev Biochem 71:635
33. Zähringer U, Lindner B, Rietschel ET (1994) Adv Carbohydr Chem Biochem 50:211
34. Zähringer U, Lindner B, Rietschel ET (1999) Chemical structure of lipid A. In: Brade H, Morrison DC, Opal S, Vogel S (eds) Endotoxin in health and disease. M Dekker Inc., New York, p 93
35. Holst O (2000) Chemical structure of the core region of lipopolysaccharides. In: Brade H, Morrison DC, Opal S, Vogel S (eds) Endotoxin in health and disease. M Dekker Inc., New York, p 115
36. Holst O (2002) Trends Glycosci Glycotech 14:87
37. Knirel YA, Kochetkov NK (1994) Biochemistry (Moscow) 59:1325
38. Jansson PE (1999) The chemistry of the *O*-polysaccharide chains in bacterial lipopolysaccharides. In: Morrison DC, Brade H, Opal S, Vogel S (eds) Endotoxin in health and disease. M Dekker Inc., New York, p 155
39. Fischer W (1994) Lipoteichoic acids and lipoglycans. In: Ghuysen JM, Hakenbeck R (eds) Bacterial cell wall. Elsevier Science, Amsterdam, p 199
40. Neuhaus FC; Baddiley J (2003) Microbiol Mol Biol Rev 67:686
41. Li Y, Gray GR (1996) Biochemistry 35:16299

42. Ourisson G, Albrecht P (1992) Acc Chem Res 25:398
43. Ourisson G, Rohmer M (1992) Acc Chem Res 25:403
44. Sahm H, Rohmer M, Bringer-Mayer S, Sprenger GA, Welle R (1993) Adv Microbial Physiol 35:247
45. Hamberg M, Liepinsh E, Otting G, Griffiths W (1998) Lipids 33:355
46. Fattorusso E, Mangoni A (1997) Fortschr Chem Org Naturstoffe 72:215
47. Trinel PA, Maes E, Zanetta JP, Delplace F, Coddeville B, Jouault T, Strecker G, Poulain D (2002) J Biol Chem 277:37260
48. Bennion B, Park C, Fuller M, Lindsey R, Momany M, Jennemann R, Levery SB (2003) J Lipid Res 44:2073
49. Jennemann R, Geyer R, Sandhoff R, Gschwind RM, Levery SB, Gröne HJ, Wiegandt H (2001) Eur J Biochem 268:1190
50. Andersson MX, Hamberg M, Kourtchenko O, Brunnström A, McPhail KL, Gerwick WH, Göbel C, Feussner I, Ellerström M (2006) J Biol Chem 281:31528
51. Stelmach BA, Müller A, Hennig P, Gebhardt S, Schubert-Zsilavecz M, Weiler EW (2001) J Biol Chem 276:12832
52. Harder T, Simons K (1997) Curr Op Cell Biol 9:534
53. Hakomori SI, Yamamura S, Honda K (1998) Ann NY Acad Sci 845:1
54. Hooper NM (1998) Curr Biol 8:R114
55. Hakomori SI (2000) Glycoconj J 17:627
56. Stoffel W, Bosio A (1997) Curr Op Neurobiol 7:654
57. Coetzee T, Suzuki K, Popko B (1998) Trends Neurosci 21:126
58. Dasgupta S, Levery SB, Hogan EL (2002) J Lipid Res 43:751
59. Constantino V, Fattorusso E, Menna M, Taglialatela-Scafati, O (2004) Curr Med Chem 11:1671
60. Wuhrer M, Kantelhardt SR, Dennis RD, Doenhoff MJ, Lochnit G, Geyer R (2002) Eur J Biochem 269:481
61. Iriko H, Nakamura K, Kojima H, Iida-Tanaka N, Kasama T, Kawakami Y, Ishizuka I, Uchda A, Murata Y, Tamai Y (2002) Eur J Biochem 269:3549
62. Sandhoff R, Geyer R, Jennemann R, Paret C, Kiss E, Yamashita T, Gorgas K, Sijmonsma TP, Iwamori M, Finaz C, Proia RL, Wiegandt H, Gröne HJ (2005) J Biol Chem 280:27310
63. Wiegandt H (1995) Behav Brain Res 66:85
64. Nakamura K, Kojima H, Suzuki M, Suzuki A, Tamai Y (2000) Eur J Biochem 267:5198
65. Misasi R, Dionisi S, Farrilla L, Carabba B, Lenti L, Di Mario U, Dotta F (1997) Diabetes – Metabolism Rev 13:163
66. Yates AJ, Rampersaud A (1998) Ann NY Acad Sci 845:57
67. Ogawa-Goto K, Abe T (1998) Neurochem Rev 23:305
68. Niimura Y (2006) Carbohydr Res 341:2669
69. Wiegandt H (1992) Biochim Biophys Acta 1123:117
70. Seppo A, Moreland M, Schweingruber H, Tiemeyer M (2000) Eur J Biochem 267:3549
71. Lochnit G, Dennis RD, Ulmer AJ, Geyer R (1998) J Biol Chem 273:466
72. Sugita M, Mizunoma T, Aoki K, Dulaney JT, Inagaki F, Suzuki M, Suzuki A, Ichikawa S, Ohta S, Kurimoto A (1996) Biochim Biophys Acta 1302:185
73. Khoo KH, Chatterjee D, Caulfield JP, Morris HR, Dell A (1997) Glycobiology 7:653
74. McConville MJ (1995) Aust N Z J Med 25:768
75. McNeely TB, Doyle PS (1996) Arch Biochem Biophys 334:1
76. Opat A, Ng K, Currie G, Handman E, Bacic A (1996) Glycobiology 6:387
77. Mengeling BJ, Beverly SM, Turco SJ (1997) Glycobiology 7:873
78. Hofer M, Hampton RY, Raetz CRH, Yu H (1991) Chem Phys Lipids 59:167
79. Brandenburg K, Seydel U (1998) Chem Phys Lipids 96:23
80. Labischinski H, Barnickel G, Bradaczek H, Naumann D, Rietschel ET, Giesbrecht P (1985) J Bacteriol 162:9
81. Naumann D, Schultz C, Sabisch A, Kastowsky M, Labischinski H (1989) J Mol Struct 214:213
82. Israelachvili JN (1991) Intermolecular and surface forces. Academic Press, London
83. Israelachvili JN, Marcelja S, Horm RG (1980) Q Rev Biophys 13:121
84. Brandenburg K, Seydel U, Schromm AB, Loppnow H, Koch MHJ, Rietschel ET (1996) J Endotox Res 3:173
85. Brandenburg K, Seydel U (1984) Biochim Biophys Acta 775:225
86. Brandenburg K, Seydel U (1990) Eur J Biochem 191:229
87. Seydel U, Brandenburg K (1992) Supramolecular structure of lipopolysaccharides and lipid A.

In: Morrison DC, Ryan JL (eds) Bacterial endotoxic lipopolysaccharides, vol 1, molecular biochemistry and cellular biology. CRC Press, Boca Raton, p 225
88. Rietschel ET, Brade H, Holst O, Brade L, Müller-Loennies S, Mamat U, Zähringer U, Beckmann F, Seydel U, Brandenburg K, Ulmer AJ, Mattern T, Heine H, Schletter J, Loppnow H, Schönbeck U, Flad HD, Hauschildt S, Schade UF, DiPadova F, Kusumoto S, Schumann RR (1996) Curr Top Microbiol Immunol 216:39
89. Brandenburg K, Mayer H, Koch MHJ, Weckesser J, Rietschel ET, Seydel U (1993) Eur J Biochem 218:555
90. Brandenburg K, Schromm AB, Koch MHJ, Seydel U (1995) Progr Clin Biol Res 392:167
91. Seydel U, Brandenburg K, Rietschel ET (1994) Progr Clin Biol Res 388:17
92. Mayer H, Krauss JH, Yokota A, Weckesser J (1990) Natural variants of lipid A. In: Friedman H, Klein TW, Nakano M, Nowotny A (eds) Endotoxin. Plenum Press, New York, p 45
93. Takayama K, Qureshi N, Ribi E, Cantrell JL (1984) Rev Infect Dis 6:439
94. Alving CR, Verma JN, Rao M, Krzych U, Anselem S, Green SM, Wassef NM (1992) Res Immunol 143:197
95. Kiener PA, Marck F, Rodgers G, Lin PF, Warr G, Desidero J (1988) J Immunol 141:870
96. Loppnow H, Libby P, Freudenberg M, Kraus JH, Weckesser J, Mayer H (1990) Infect Immun 58:3743
97. Moran AP (1995) FEMS Immunol Med Microbiol 11:121
98. Wicken AJ, Evans JD, Knox KW (1986) J Bacteriol 166:72
99. Courtney HS, Simpson WA, Beachy EH (1986) Infect Immun 51:414
100. Fischer W (1990) Handb Lipid Res 6:123
101. Labischinski H, Naumann D, Fischer W (1991) Eur J Biochem 202:1269
102. Fischer W, Markwitz S, Labischinski H (1997) Eur J Biochem 244:913
103. Gutberlet T, Markwitz S, Labischinski H, Bradaczek H (1991) Macromol Chem Macromol Symp 46:283
104. Chatterjee D (1997) Curr Op Chem Biol 1:579
105. Nikaido H, Kim SH, Rosenberg EY (1993) Mol Microbiol 8:1025
106. Liu J, Rosenberg EY, Nikaido H (1995) Proc Natl Acad Sci USA 92:11254
107. Clausen H, Hakomori SI (1989) Vox Sang 56:1
108. Karlsson KA (1989) Annu Rev Biochem 58:309
109. Lingwood CA (1992) Curr Opin Struct Biol 2:693
110. Nyholm PG, Pascher I (1993) Biochemistry 32:1225
111. Koynova R, Caffrey M (1994) Chem Phys Lipids 69:181
112. Mannock D, Lewis R, McElhaney R (1990) Biochemistry 29:7790
113. Mannock D, Lewis R, McElhaney R (1990) Chem Phys Lipids 55:309
114. Lewis R, Mannock D, McElhaney R (1990) Zentralbl Bacteriol Suppl 20:643
115. Mannock DA, Lewis RNAH, McElhaney RN, Harper PE, Turner DC, Gruner SM (2001) Eur Biophys J 30:537
116. Köberl M, Schöppe A, Hinz HJ, Rapp G (1998) Chem Phys Lipids 95:59
117. Ben-Menachem G, Byström T, Rechnitzer H, Rottem S, Rilfors L, Lindblom G (2001) Eur J Biochem 268:3694
118. Brandenburg K, Wagner F, Müller M, Heine H, Andrä J, Koch MHJ, Zähringer U, Seydel U (2003) Eur J Biochem 270:3271
119. Koynova R, Caffrey M (1995) Biochim Biophys Acta 1255:213
120. Boggs JM (1987) Biochim Biophys Acta 906:353
121. Lee DC, Miller IR, Chapman D (1986) Biochim Biophys Acta 859:266
122. Jackson M, Johnston DS, Chapman D (1988) Biochim Biophys Acta 944:497
123. Ruocco MJ, Atkinson D, Small DM, Skarjune RP, Oldfield E, Shipley GG (1981) Biochemistry 20:5957
124. Harder T, Simons K (1997) Curr Opin Cell Biol 9:534
125. Stoffel W, Bosio A (1997) Curr Opin Neurobiol 7:654
126. Yu ZW, Calvert TL, Leckband D (1998) Biochemistry 37:1540

7.2 Synthesis of Glycolipids

Masahiro Wakao, *Yasuo Suda*
Department of Nanostructure and Advanced Materials,
Graduate School of Science and Engineering, Kagoshima University,
890-0065 Kagoshima, Japan
wakao@eng.kagoshima-u.ac.jp, ysuda@eng.kagoshima-u.ac.jp

Abstract

Glycolipids, composed of hydrophilic carbohydrate and hydrophobic aliphatic residues, have a wide variety of biological activity. Natural glycolipids are known to be very complex and heterogeneous, and sometimes contain lipophilic contaminants, which cause confusion in the understanding of their original functions in biological process. To investigate their function at the molecular level, structurally defined glycolipids are necessary. A chemical synthetic approach may be the only method to overcome this issue. So far, a lot of synthetic efforts have been devoted in this field. This chapter summarizes syntheses of natural and related glycolipids, sphingoglycolipids, glyceroglycolipids, lipoteichoic acid, mycoloyl arabinan, lipopolysaccharide, glycophosphatidylinositol, etc.

In: *Glycoscience*. Fraser-Reid B, Tatsuta K, Thiem J (eds)
Chapter-DOI 10-1007/978-3-540-30429-6_40:

Keywords

Sphingoglycolipid; Glycosphingolipid; Ceramide; Sphingosine; Dihydrosphingosine; Phytoshingosine; Sialic acid; Glycosylation; Sialylation; DSLe[a]; Cerebroside; Plakoside

Abbreviations

Alloc allyloxycarbonyl
Azmb (2-azidomethyl)benzoyl
BINAP 2,2′-bis(diphenylphosphino)-1,1′-binaphthyl
Bt butyryl
BCG bacillus Calmette–Guérin
CAc chloroacetyl
Cp cyclopentadienyl
DTBMP 2,6-di-*tert*-butyl-4-methylpyridine
EDC *N*-ethyl-*N*′-(3-dimethylaminopropyl)carbodiimide
GPI glycosylphosphatidylinositol
HOAt 1-hydroxy-7-azabenzotriazole
Kdo 3-deoxy-D-*manno*-2-octuronic acid
LAM lipoarabinomannan
Lev levulinoyl
LPS lipopolysaccharide
LTA lipoteichoic acid
MA mycolic acid
MBn *p*-methylbenzyl
MP methoxyphenyl
NAP naphthyl
Neu5Ac sialic acid
NIS *N*-iodosuccinimide
Nos nosyl (4-nitrobenzenesulfonyl)
PAB *p*-(pivaloylamino)benzyl
PCC pyridinium chlorochromate
PIM phosphatidylinositol mannoside
PMT *p*-methoxytrityl
PyBOP (benzotriazol-1-yloxy)tripyrrolidinophosphonium hexafluorophosphate
SE 2-trimethylsilylethyl
TBHP *tert*-butyl hydroperoxide
TBSA tuberculostearic acid
TEMPO 2,2,6,6-tetramethyl-1-piperidinyloxy, free radical
TES triethylsilyl
TFAc trifluoroacetyl
TLR Toll-like receptor
Troc 2,2,2-trichloroethoxycarbonyl

1 Introduction

To date, more than 300 natural glycolipids have been isolated. Structural features of glycolipids are composed of covalently bound mono- or oligosaccharides and lipophilic parts such as alkyl or acyl long chains, which are sometimes modified by additional functional groups such as phosphates, sulfates, glycerols, etc. They work as constituents of the biomembrane, and play significant roles in many biological events such as cell proliferation, cell differentiation, cell recognition, cell adhesion, etc. In addition, some bacterial glycolipids like lipopolysaccharide (LPS) exhibit immunostimulating activity against vertebrates. Recently, structurally unique glycolipids have been isolated and identified from various organisms such as marine sponges, insects, and also in Chinese traditional medicine. These glycolipids exhibit unique bioactivity including anti-tumor and immunostimulating activity.

Evaluation of structure–function relationships and structural assignment of new complex glycolipids have been performed on many glycolipids. Chemical synthesis, which can provide structurally defined molecules, is very important not only for the accurate structural assignment, but also for the precise investigation of their biological functions, since natural glycolipids are usually obtained as inseparable complex mixtures owing to the heterogeneity in the lipid moieties, and may contain lipophilic contaminants.

In this chapter, we reviewed recent works on chemical syntheses of sphingoglycolipids, glyceroglycolipids, bacterial glycolipids such as mycoloyl arabinan partial structure, lipoteichoic acid (LTA), Re-type lipopolysaccharides, glycophosphatidylinositols, and other glycolipids with unique structures.

2 Synthesis of Sphingoglycolipids

Sphingoglycolipids are components of the cell membrane and play a significant role in various biological processes [1,2]. Sphingoglycolipids are composed of at least one monosaccharide residue and long-chain aliphatic components known as ceramides (Cer). A typical structure is illustrated in ❯ *Fig. 1*. They have diverse structural variations on both the ceramide and saccharide residues, and they are classified into many types according to carbohydrate sequence. The representative families are summarized in ❯ *Table 1*. In this section, recent syntheses

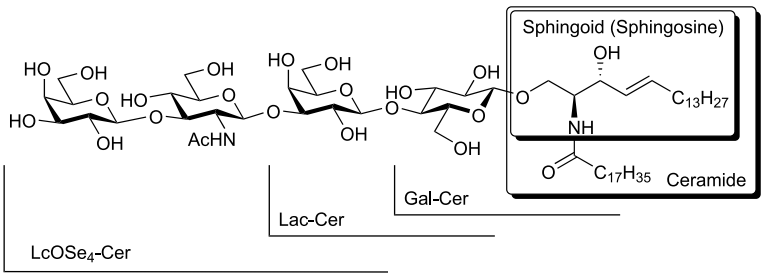


Figure 1
Chemical structure of $LcOSe_4$

Table 1
The structure of the representative sphingoglycolipids

Series	Abbreviation	Structure
Lacto	($LcOSe_4$)	Gal(β1-3)GlcNAc(β1-3)Gal(β1-4)Glc(β1-1)Cer
Neolacto	($LcnOSe_4$)	Gal(β1-4)GlcNAc(β1-3)Gal(β1-4)Glc(β1-1)Cer
Globo	($GbOSe_4$)	GalNAc(β1-3)Gal(α1-4)Gal(β1-4)Glc(β1-1)Cer
Paraglobo		Gal(β1-4)GalNAc(β1-3)Gal(β1-4)Glc(β1-1)Cer
Isoglobo	($GbiOSe_4$)	GalNAc(β1-3)Gal(α1-3)Gal(β1-4)Glc(β1-1)Cer
Ganglio	($GgOSe_4$)	Gal(β1-3)GalNAc(β1-4)Gal(β1-4)Glc(β1-1)Cer
Muco	($MucOSe_4$)	Gal(β1-3)Gal(β1-3)Gal(β1-4)Glc(β1-1)Cer
Gala	($GalOSe_2$)	Gal(α1-4)Gal(β1-1)Cer
Sulfatide		3-*O*-Sulfo-Gal(β1-1)Cer

regarding sphingosine, gangliosides noting recent progress in sialylation, and cerebroside are described.

2.1 Synthesis of Ceramide

Ceramide is usually composed of an *N*-linked acyl chain and sphingoid. Sphingoid consists of a polar head group and a lipophilic chain. Dihydrosphingosine, sphingosine, and phytosphingosine are known as major sphingoids (see *Fig. 2*). Among them, sphingosine is the most common sphingoid, and a number of synthetic methods have been developed for the syntheses of sphingosine and its protected equivalents, which have been summarized in several reviews [3,4,5,6,7,8,9,10,11,12]. Most of these methods are based on a chiral pool strategy utilizing amino acid (L-serine [13]), carbohydrate (D-glucose [14], D-glucal [15], D-glucosamine [16], D-galactose [17], D-galactal [18], D-xylose [17,19], and D-arabinose [19]), and other chiral sources such as glycidol [20], D-glyceraldehyde [21], D-threose [22], D-tartaric acid [23,24], D-mannitol [25], and so on. An asymmetric strategy via enantioselective aldol reaction [26,27] including aldol reactions with chiral auxiliaries [28], epoxidation [29], alkenylation [30], or *N*-sulfinylaziridination [31] is also effective for the synthesis.
Recently, methods with efficient, simple, versatile, and rapidly accessible routes have been required because sphingosine and its derivatives are potentially useful as a pharmacore. Divergent synthesis of the polar-head group of the sphingosine has been reported [32,33,34]. Hudlicky et al. achieved the synthesis of four stereoisomers (**2**, **3**, **4**, and **5**) of a sphingo-

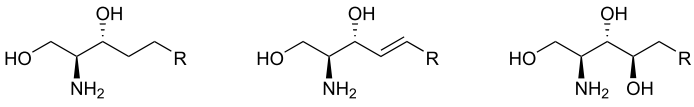
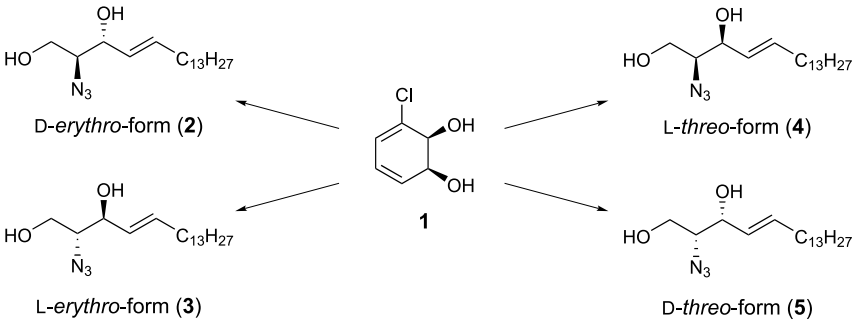
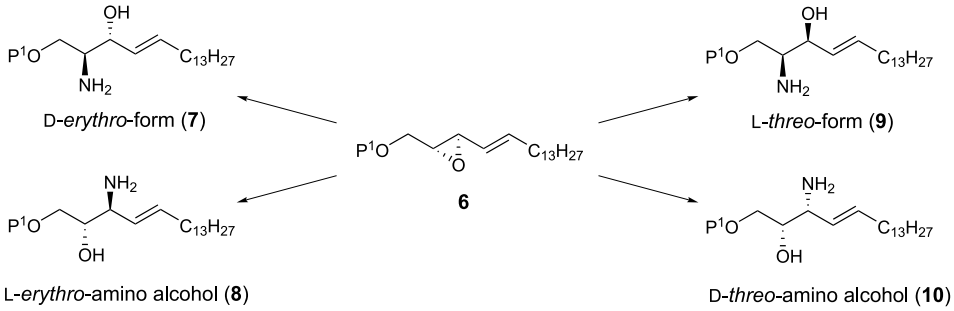


Figure 2
Chemical structure of sphingoids (R = alkyl chain, normally -$C_{13}H_{27}$)



Scheme 1



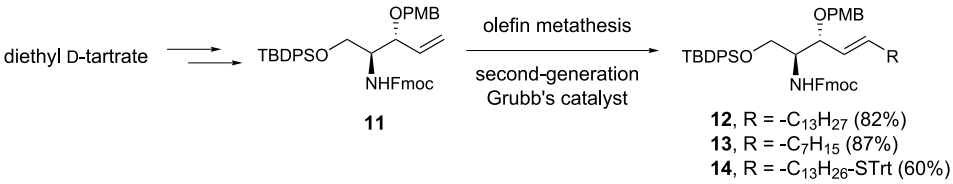
Scheme 2

sine derivative using a chemoenzymatically prepared chiral diol **1** as a starting material [32] (*Scheme 1*). Somfai et al. also successfully synthesized four stereo- and regioisomers (**7**, **8**, **9**, **10**) from chiral epoxide **6** [33] (*Scheme 2*). This method may be more efficient than the former because it can be accessible to the syntheses of the corresponding enantiomers by using *ent*-**6**.

Construction of a lipophilic chain towards divergent synthesis of sphingosine has been performed by cross-metathesis olefination [35,36,37,38]. This strategy is attractive in terms of high *E*-selectivity with good yield in the product formation and the functionality tolerance under metathesis conditions. Basu et al. synthesized sphingosine core building block **11** with an orthogonal protecting group, which was derived from diethyl D-tartrate in several steps. Olefin metathesis of the building block **11** with various alkenes in the presence of Grubbs generation II catalyst [39] gave the corresponding alkenes **12**, **13**, and **14** in good yields [36] (*Scheme 3*).

For synthesis of phytosphingosine [40] and dihydrosphingosine [7], many methods have been reported. These utilized a chiral pool strategy with amino acid or carbohydrate similar to that of sphingosine.

Labeled ceramide derivatives are also important for investigation of its biological behavior. Synthesis of radio-labeled [41,42,43,44,45,46] and fluorescence-labeled [47,48,49,50,51,52,53,54] compounds has been reported.



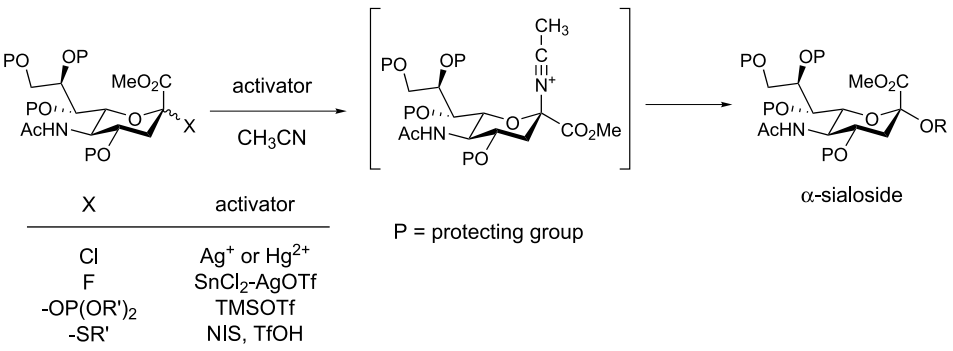
Scheme 3

2.2 Synthesis of Gangliosides

Gangliosides are generic terms of glycosphingosines containing sialic acid (neuramic acid) and are recognized to possess very important biological functions. Although there are many efforts for synthesizing gangliosides reported in the literature [12], glycosylation with sialic acid (sialylation) is, even now, one of the most important remaining issues for the total synthesis of gangliosides or other compounds containing sialic acid. Efficient methods for sialylation are required. This section deals with recent progress on sialylation and a total synthesis of ganglioside, DSLe^a (disialyl Lewis A).

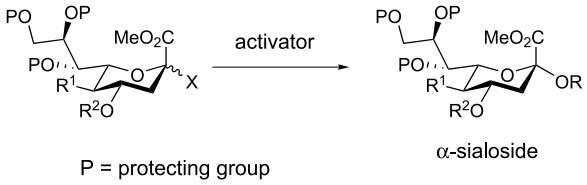
2.2.1 Glycosylation with Sialic Acid

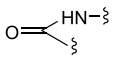
Sialic acid is regularly located at the non-reducing terminal of glycoconjugates through α-glycosidic bonds. Much effort has been devoted to the synthesis of sialoside [55]. Generally, an α-sialoside is obtained by kinetically controlled α-selective glycosylation. In most cases, the reactions are carried out in a solvent with α-directing effect such as CH_3CN or EtCN [56,57]. The representative glycosyl donors are shown in *Scheme 4*. A glycosyl halide is activated by mercury salt [58]. When secondary hydroxy groups in the acceptor were used for sialylation, the yields were modest. Thioglycosides [59,60,61,62,63,64] and glycosyl phosphates [65,66] afforded α-sialoside in moderate yields with high selectivity. However, when sterically hindered secondary hydroxy groups, such as the 8-OH group of neuramic acid or 3-OH group of galactose, are used in the acceptor, the yields and selectivities are both decreased due to an



X	activator
Cl	Ag^+ or Hg^{2+}
F	$SnCl_2$-AgOTf
-OP(OR')$_2$	TMSOTf
-SR'	NIS, TfOH

Scheme 4

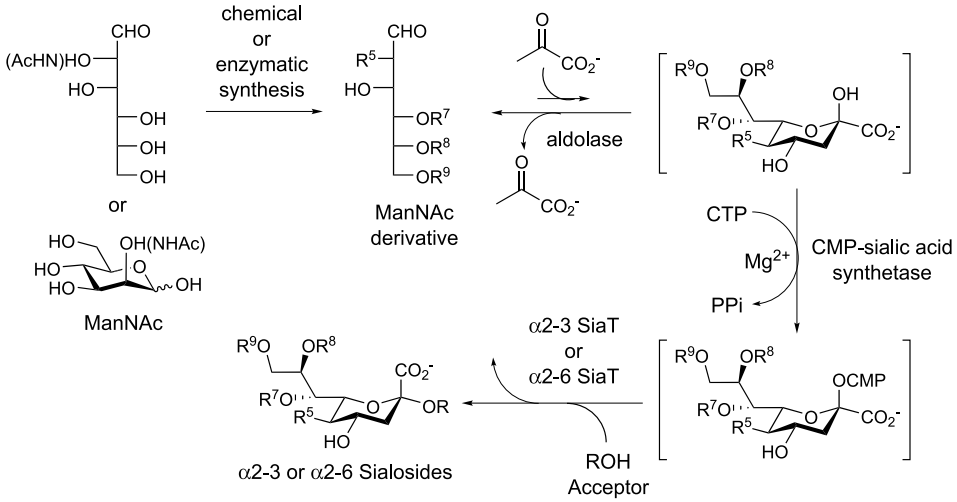


X	R^1	R^2
-SR	Ac_2N-	Ac
-SR	TFAcHN-	Ac
-SR	N_3-	Ac
-SR	TrocHN-	Ac
-SR		
-OC(=NPh)CF_3	AcHN-	Ac
-OC(=NPh)CF_3	phthN-	Ac

Scheme 5

undesired elimination in the sialyl donor. Thus, sialylation with 100% α-glycosidic linkage and high yield still remains a significant challenge.

Recently, it was found that the *N*-protecting group of the sialyl donor at the 5-position affected the yield and selectivity of the obtained sialoside (*Scheme 5*). Boons et al. first reported the *N*,*N*-diacetylated sialyl donor in 1998 [67,68]. The *N*,*N*-diacetylated sialyl donor is more reactive than the *N*-acetylated one, giving α-sialoside in good yield. After that, various sialyl donors were prepared, i. e. *N*-trifluoroacetyl [69], *N*-trichloroacetyl [70], azido [71], *N*-2,2,2-



R^5 = NHAc, OH, NHGc; or their derivatives
R^7 = H or Ac
R^8 = H, Ac, SO_3H, or Me
R^9 = H, Ac, or lactyl
R = molecules containing Gal(NAc) as a terminal or penultimate residue

Scheme 6

trichloroethoxy carbonyl [70,72], *N*-fluorenylmethoxy carbonyl [70], *N*,*N*-phthaloyl [73], and 5-*N*,4-*O* carbonyl [74] derivatives. Although the reaction mechanisms of these sialylations is still unclear, it seems that a delocalization of electroproperty and a fixation of dipole moment on the nitrogen atom is important for enhancing reactivity of the sialyl donor. In addition, Yu et al. also reported a *N*-phenyl trifluoroimidate donor [75], which showed a higher reactivity than those previously reported. Thus, α-sialylation is constantly being improved.

A chemoenzymatic sialylation is also an efficient method for obtaining sialosides. Wong et al. reported a one-pot two-enzyme system, in which sialyltransferases and CMP-NeuAc recycling enzymes were used [76,77]. The resultant sialoside was obtained in good yield. Recently, a one-pot three-enzyme system was reported by Chen et al. [78,79] (*Scheme 6*). Mannosamine was used as a starting material. This system afforded the corresponding sialosides in high yield, and could be applied to a variety of protected mannosamine derivatives.

2.2.2 Total Synthesis of DSLe^a

A total synthesis of ganglioside DSLe^a, α(2-3)/α(2-6)-disialyl Lewis A, which was isolated from human colonic adenocarcinoma [80], was recently reported by Kiso's group as illustrated in *Fig. 3* and *Scheme 7* [81,82]. DSLe^a is characterized by a sterically hindered tri-antennary structure at *O*-3, *O*-4, and *O*-6 of glucosamine residue. The crucial points of the synthesis

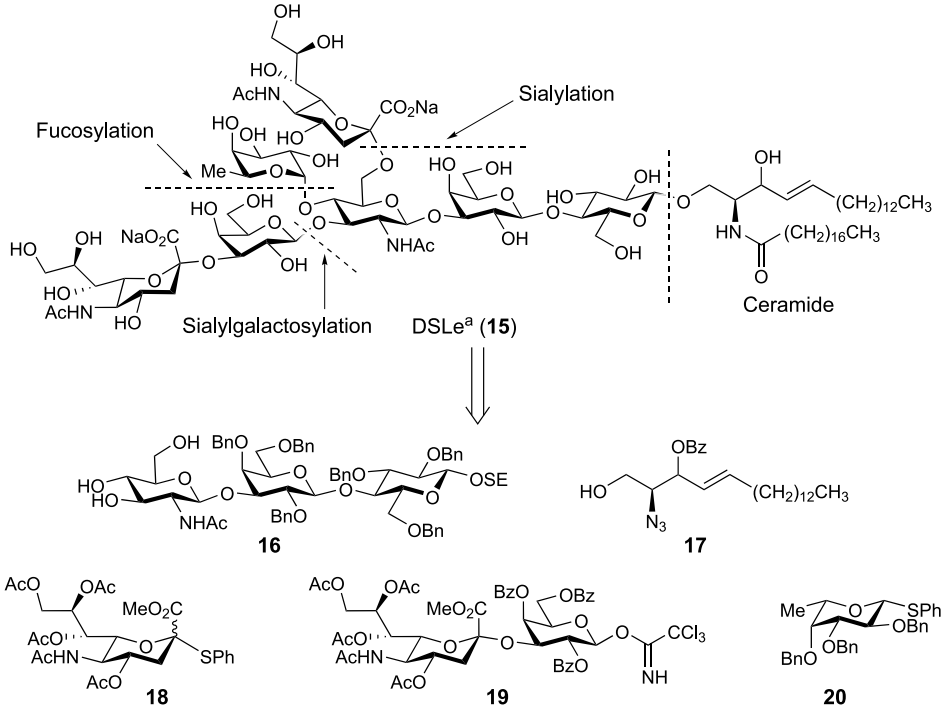
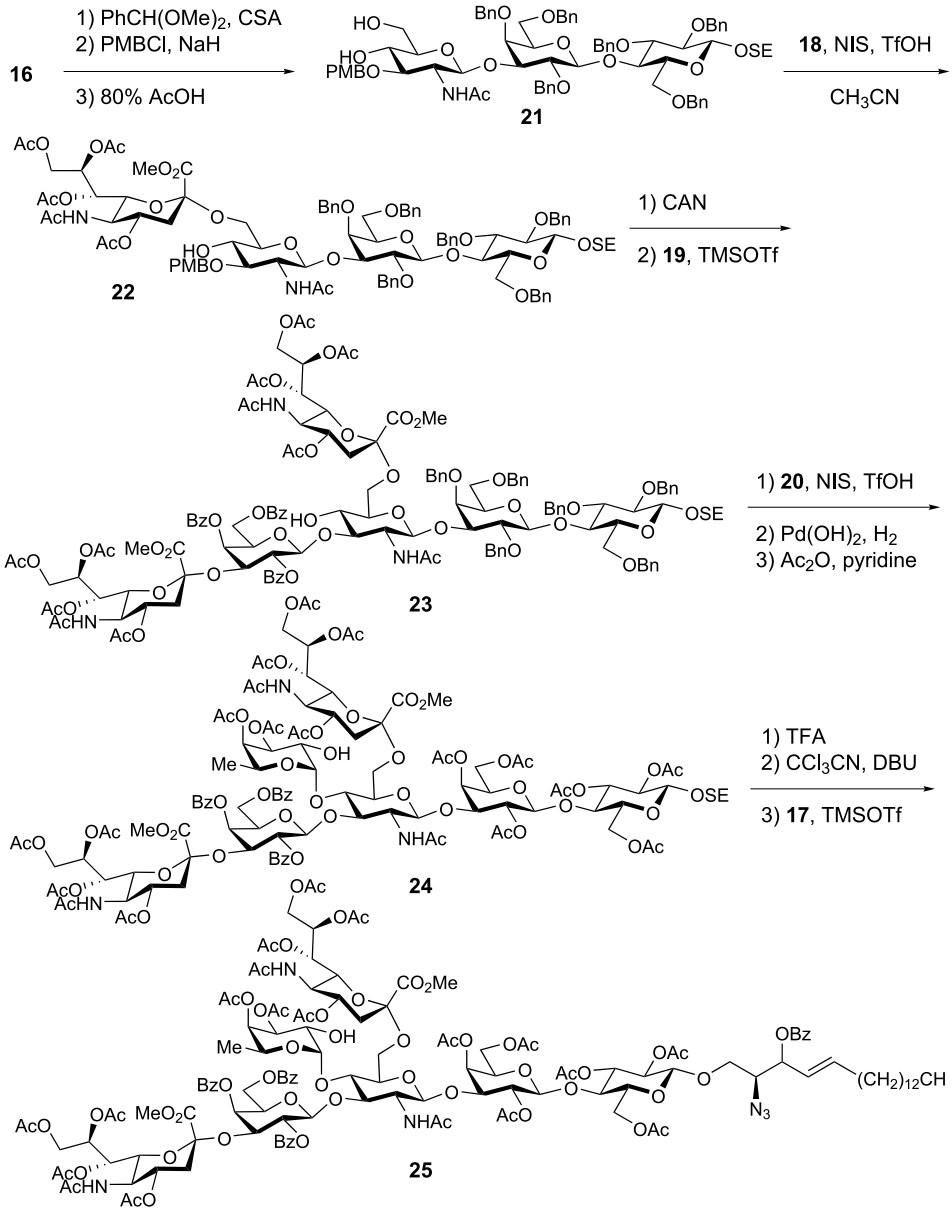


Figure 3
Retrosynthetic analysis of DSLe^a (15)


16
1) PhCH(OMe)2, CSA
2) PMBCl, NaH
3) 80% AcOH
21
18, NIS, TfOH
CH3CN
22
1) CAN
2) 19, TMSOTf
23
1) 20, NIS, TfOH
2) Pd(OH)2, H2
3) Ac2O, pyridine
24
1) TFA
2) CCl3CN, DBU
3) 17, TMSOTf
25
1) H2S
2) octadecanoic acid, EDC•HCl
3) NaOMe, MeOH, then H2O
DSLea (15)

Scheme 7

are regioselective α-sialylation at the 6-position, regioselective β-sialylgalactosylation at the 3-position, and α-fucosylation at the 4-position of the glucosamine moiety. According to the retrosynthetic analysis: DSLe^a is divided into the five known building blocks **16**, **17**, **18**, **19**, and **20** (*Fig. 3*). The synthetic pathway is shown in *Scheme 7*.

Trisaccharide **16** was converted to 3-*O*-*p*-methoybenzyl form **21** [83] in three steps. Sialylation of **21** with thioglycoside **18** [84] under the established conditions [61,62,63,64,85], with *N*-iodosuccinimide (NIS) and trifluoromethanesulfonic acid (TfOH) as promoters at −35 °C in acetonitrile, gave predominantly α-sialyl tetrasaccharide **22** (74%) with β-sialoside (16%). After removal of the PMB group, glycosylation of tetrasaccharide with sialylgalactosyl donor **19** [86] preferentially afforded β(1-3)-linked hexasaccharide **23** (53%) with β(1-4)-linked hexasaccharide (13%). Fucosylation of the remaining hydroxy group at the 4-position in GlcNAc was performed only when an excess amount of fucosyl donor **20** [87] (15 mol equiv.) was used, affording the desired heptasaccharide in 75% yield. Removal of the benzyl group by hydrogenolysis, followed by full acetylation of the resultant hydroxy group, furnished heptasaccharide **24**. Heptasaccharide **24** was then led to sphingoglycoside **25** via removal of the SE group, trichloro-imidation, and glycosylation with azidosphingosine **17**. Target compound DSLe^a **15** was synthesized by reduction of the azide group, *N*-acylation with octadecanoic acid, *O*-deacylation, and saponification of methyl ester.

2.3 Synthesis of Cerebrosides

Cerebroside is the generic term for monoglycosylceramides, which have been studied for over a century. Recently, there has been renewed focus on cerebrosides because of their interesting biological activities, such as immunostimulation and suppression. They are also recognized as synthetically attractive targets with regard to a wide variety of functional groups in the lipophilic part. A few cerebrosides are illustrated in *Fig. 4*. Typhoniside A is a component of traditional Chinese medicine [88]. Both OCH [89] and KRN7000 [90,91] possessing α-Gal structure exhibited immunostimulating activity. In this section, the total synthesis of plakoside A, which has a highly complex structure in the lipophilic part, is described.

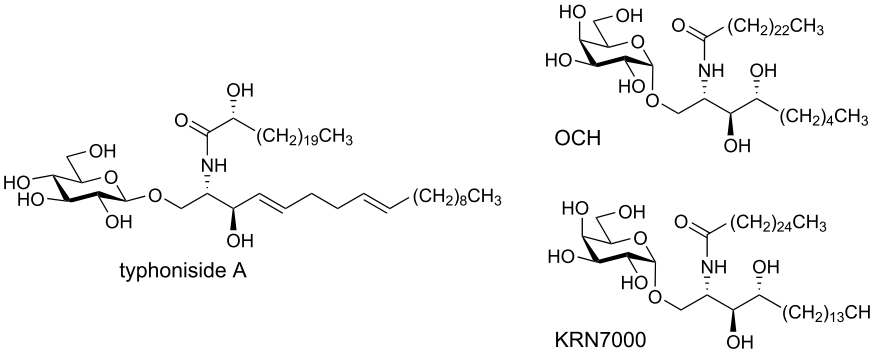


Figure 4
Chemical structures of cerebrosides

2.3.1 Total Synthesis of Plakoside A

Plakoside A and B are naturally occurring glycosphingosines, which were isolated from the marine sponge *Plakortis simplex* in 1997 by Fattorusso's group [92] (*Fig. 5*). They are composed of a prenylated D-galactose and cyclopropane-containing ceramide, and possess strong immunosuppressive activity without any cytotoxicity. Total synthesis of Plakoside A has so far been accomplished by two groups, Nicolaou's group in 2000 [93] and Mori's group in 2001 [94]. Herein, we will focus on Nicolaou's method.

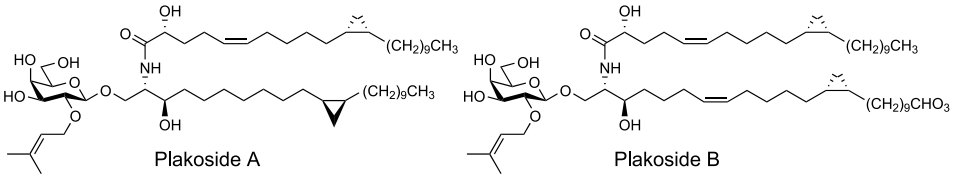
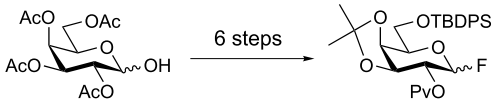


Figure 5
Structures of plakosides A and B

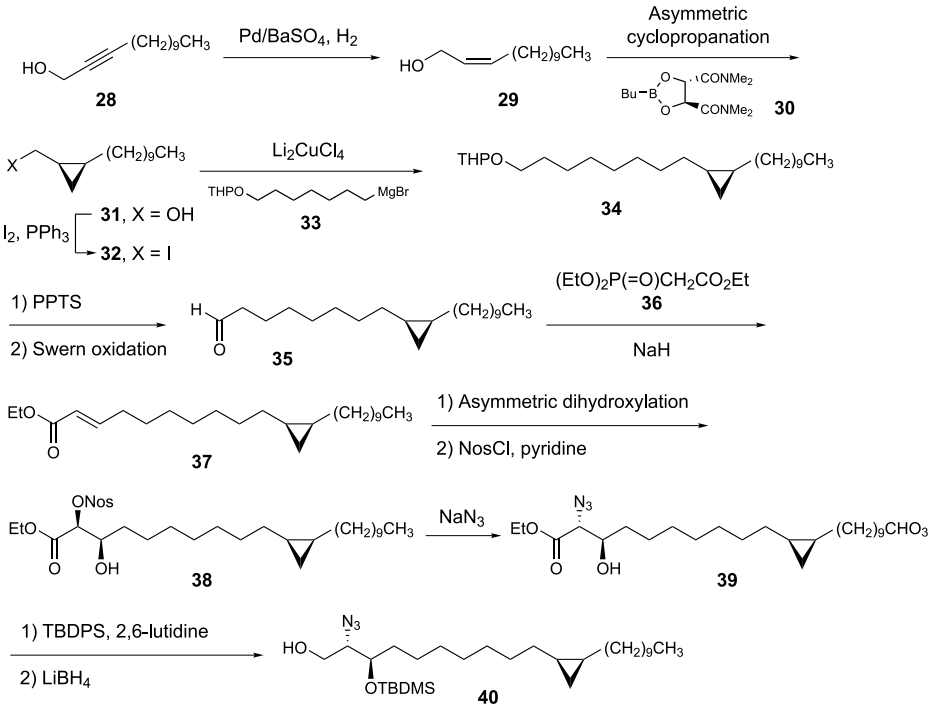
Plakoside A was synthesized by connecting three building blocks, galactose, the α-hydroxy acyl chain, and the sphingosine parts. Introduction of stereogenic centers of the cyclopropane ring and hydroxy group in the acyl side chain was carried out by Charette cyclopropanation [95,96,97,98] and Sharpless asymmetric dihydroxylation [99,100], respectively. Galactosyl fluoride **27** was synthesized from commercially available galactosyl pentaacetate **26** in six steps (*Scheme 8*).



26 **27**

Scheme 8

Azide alcohol sphingosine **40** was prepared by an asymmetric cyclopropanation and dihydroxylation as key reactions (*Scheme 9*). Commercially available tridec-2-yn-1-ol **28** was selectively hydrogenated by Pd on $BaSO_4$ under an H_2 atmosphere to give (*Z*)-alcohol **29**. Charette asymmetric cyclopropanation using bis(iodomethyl)zinc•DME complex and (*R*,*R*)-dioxaborolane **30** afforded **31** in 96% yield with 84% ee. Alcohol **31** was converted to the corresponding iodide **32** by exposure to iodine and triphenylphosphine. Elongation of the aliphatic chain was performed by the cross coupling with iodide **32** and Grignard reagent **33** in the presence of a catalytic amount of Li_2CuCl_4. Acidic cleavage of THP ether of **34**, followed by Swern oxidation of the resulting hydroxy group gave aldehyde **35** in good yields. Horner–Emmons olefination of aldehyde **35** with phosphate **36** predominantly furnished the (*E*)-α,β-unsaturated ester **37** (*E*/*Z* > 95:5). Asymmetric dihydroxylation of **38** with AD mix-β in the presence of $CH_3SO_2NH_2$ afforded the corresponding diol in 96% yield, then the resulting α-hydroxy group was selectively nosylated with nosyl chloride to give nosylate **38** in 90% yield. S_N2 reaction of **38** on the α-position with NaN_3 afforded α-azido-β-hydroxy

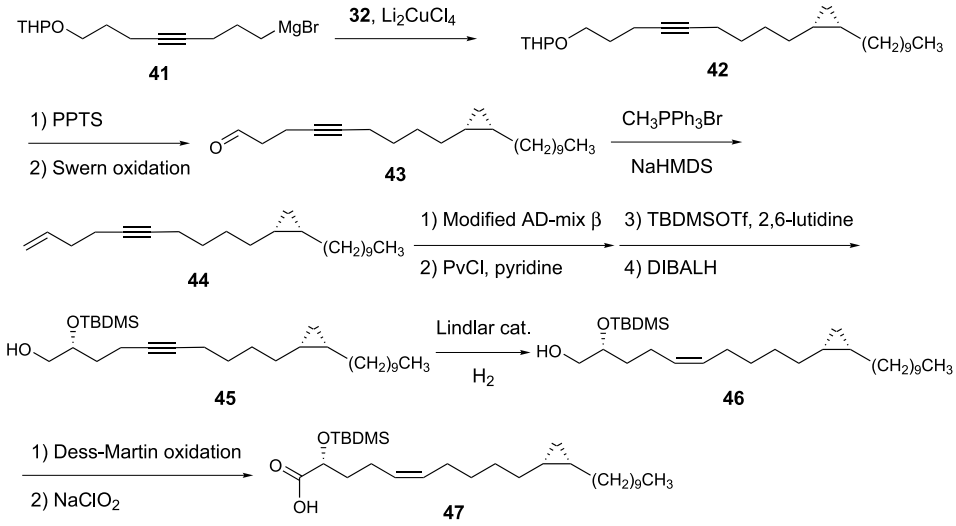


Scheme 9

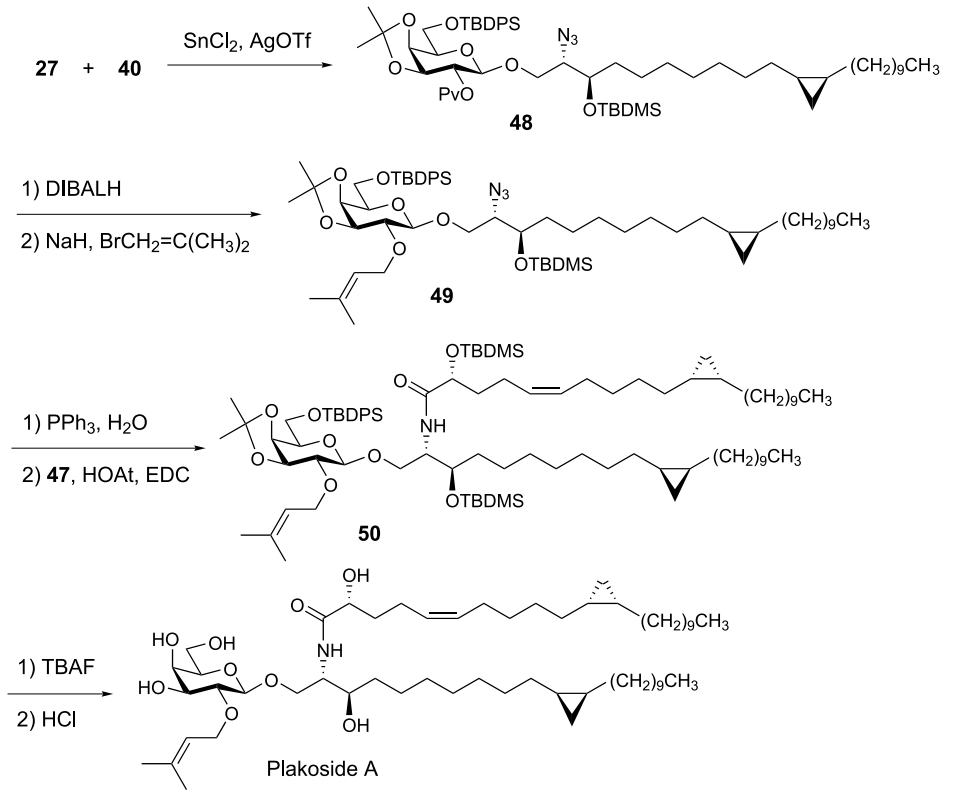
ester **39**. Silylation of **39** with *tert*-butyldimethylsilyl trifluoromethanesulfonate (TBDMSOTf) and 2,6-lutidine, and subsequent reduction of the ester moiety with lithium borohydride furnished azidosphingosine **40**.

Acyl side chain **47** was prepared as a similar strategy for the synthesis of **40** (see *Scheme 10*). Cross coupling of Grignard reagent **41** and iodide **32** in the presence of catalytic Li_2CuCl_4 gave THP ether **42** in 84% yield. THP ether of **42** was cleaved under acidic conditions, and then the resulting hydroxy group was converted to aldehyde **43** by Swern oxidation. Wittig reaction of aldehyde **43** with CH_3PPh_3Br afforded the C_1 homolog **44**. Asymmetric dihydroxylation of **44** was then performed with AD mix-β and $CH_3SO_2NH_2$ to furnish the corresponding diol, in which the primary hydroxy group was temporally protected with a pivaloyl group and the secondary one with a *tert*-butyldimethylsilyl group. After reductive cleavage of pivaloate, alkyne **45** was subjected to Lindlar hydrogenation conditions selectively and quantitatively to give (*Z*)-olefin **46**. A stepwise oxidation procedure via Dess–Martin oxidation and $NaClO_2$ oxidation afforded carboxylic acid **47**.

The three building blocks **27**, **40**, and **47** were connected as shown in *Scheme 11*. First glycosylation of azidosphingosine **40** with **27** was performed stereoselectively in the presence of $SnCl_2$ and AgOTf to give β-glycoside **48**. After removal of the pivaloyl group at the 2-position in the galactose moiety, the resulting hydroxy group was condensed with prenyl bromide in the presence of NaH to afford compound **49**. The azide group of **49** was reduced to an amino



Scheme 10



Scheme 11

group by the Staudinger reaction, which involved coupling with carboxylic acid **47** in the presence of HOAt and EDC to give compound **50**. Final deprotection was performed by sequential treatment of TBAF and aqueous HCl to afford plakoside A.

3 Synthesis of Glyceroglycolipids

Glyceroglycolipids are important constituents of cell membrane. These lipids are composed of the carbohydrate polar head group and glycerolipid [101] (*Fig. 6*). The carbohydrate moiety consists of one to eight sugar residues. Glycerolipids typically include diacylglycerol, acylalkylglycerol, or dialkylglycerol. Major structural components of cell membrane are glycosyl diacylglycerol with mono-, di-, and trisaccharides. Synthetic studies at the early stage were summarized by Gigg et al. [102].

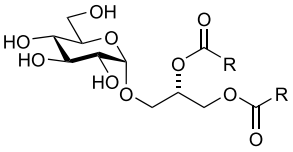


Figure 6
Chemical structure of typical glycosyl diacylglycerol

Glycerolipids from archaebacteria have very unique chemical structures. The structural feature is a tetraether-type macrocyclic component composed of two glycerols and two long saturated isoprenoid alkyl chains possessing one or two polar headgroups derived from the D-galactofuranosyl moiety and/or phosphate at the terminal ends of the lipid core [101] (*Fig. 7*). Challenges for the synthesis of macrocyclic glycolipids were (1) macrocyclic ring formation, (2) stereoselective preparation of a long isoprenoid alkyl chain, and (3) regioselective introduction of sugar and/or phosphate moieties. Macrocyclic ring formation was achieved by several groups using McMurrry [103,104,105], Glaser [106,107], or olefin metathesis [108] coupling reactions in the last decade. Elegant synthesis of the lipid moiety was achieved by Kakinuma and co-workers, in which stereoselective preparation of a long isoprenoid alkyl chain and construction of a macrocyclic framework was performed.

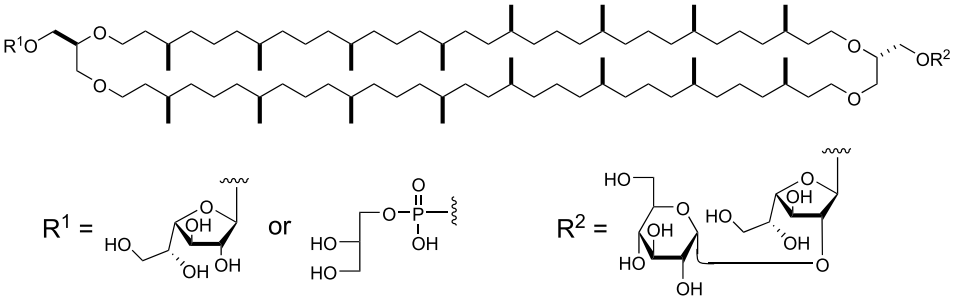


Figure 7
The chemical structure of glyceroglycolipid found in methanogenic Archaea

4 Synthesis of Lipoteichoic Acid

Lipoteichoic acid is a component of the Gram-positive bacterial cell surface. The typical chemical structure of *Staphylococcus aureus* LTA is shown in ❯ *Fig. 8* [109,110,111], which is composed of a glycolipid moiety and a hydrophilic poly(glycerol phosphate) moiety. The former moiety shows bacterial species-specific structural variation. The glycerol units in the later moiety are partially substituted with (oligo)-D-alanyl (~70%) and/or (oligo)-*N*-acetyl-α-D-glucosaminyl (~15%) residues.

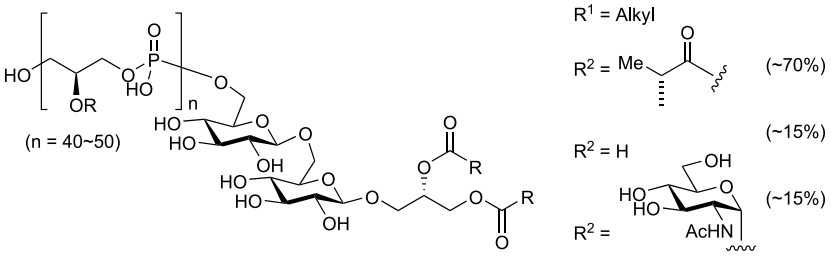
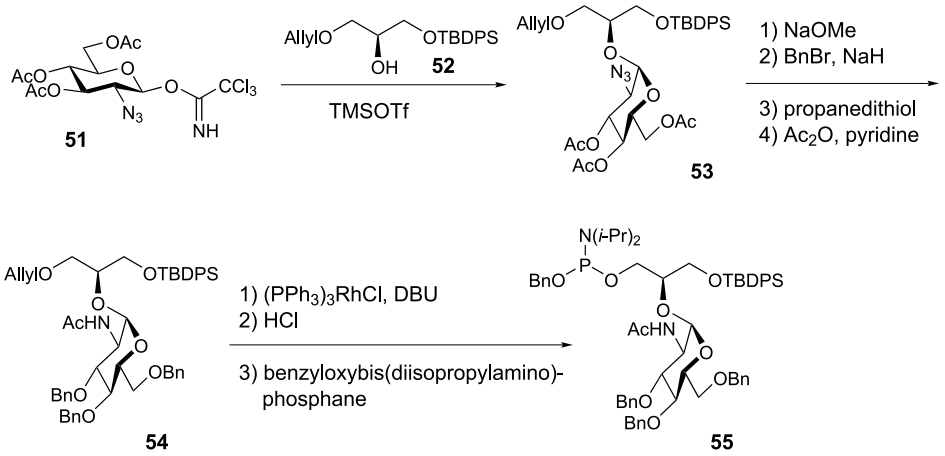


Figure 8
General structure of LTA from *Staphylococcus aureus*

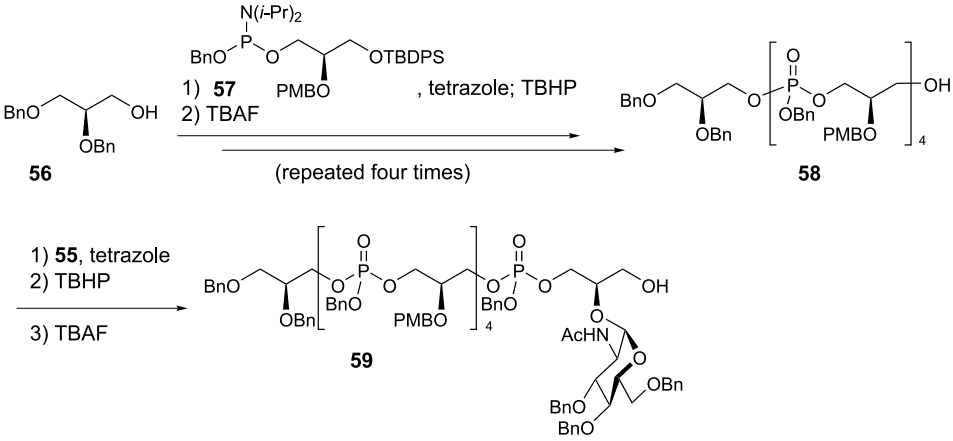
The first synthetic study regarding LTA was reported by van Boeckel's group in 1985 [112]. The syntheses of fundamental structures of *Staphylococcus pyrogenes* and *Enterococcus hirae* LTA were then reported by Kusumoto's group in 1992 [113] and 1994 [114], respectively. Recently, *Staphylococcus aureus* LTA analogue **64** was reported by Schmidt's group [115,116]. This analogue showed a similar biological activity in terms of cytokine induction from human blood leukocytes. This section deals with the synthetic work on *Staphylococcus aureus* LTA.

Target compound **64** consists of one *N*-acetyl-α-D-glucosaminyl glycerol phosphate residue, one glycerol phosphate residue, one glycolipid residue, and five D-alanyl glycerol phosphate residues in regard to the relative abundance of *N*-acetyl-α-D-glucosamine to the D-alanine residue in the natural *S. aureus* LTA. Four intermediates **55**, **56**, **57**, and **62** were prepared as key components of the *N*-acetyl-α-D-glucosaminyl glycerol phosphate, glycerol phosphate, glycolipid, and D-alanyl glycerol phosphate residues, respectively. Incorporation of alanine residues was performed at the final stage of the synthesis because of lability of alanyl ester.

N-acetyl-α-D-glucosaminyl glycerol phosphate residue **55** was prepared as shown in ❯ *Scheme 12*. Glycosylation of glycerol derivative **52** with the known azidoglycosyl donor under trimethylsilyl trifluoromethanesulfonate (TMSOTf) proceeded smoothly to give the desired α-glycoside **53** in 75% yield. After displacement of the acetyl group to the benzyl group, reduction of the azide group with propanedithiol and subsequent acetylation afforded compound **54**. The allyl group on the glycerol chain was then removed by stepwise treatment of rhodium catalyst and hydrogen chloride, and the resulting hydroxy group was phosphited to furnish phosphite **55**.



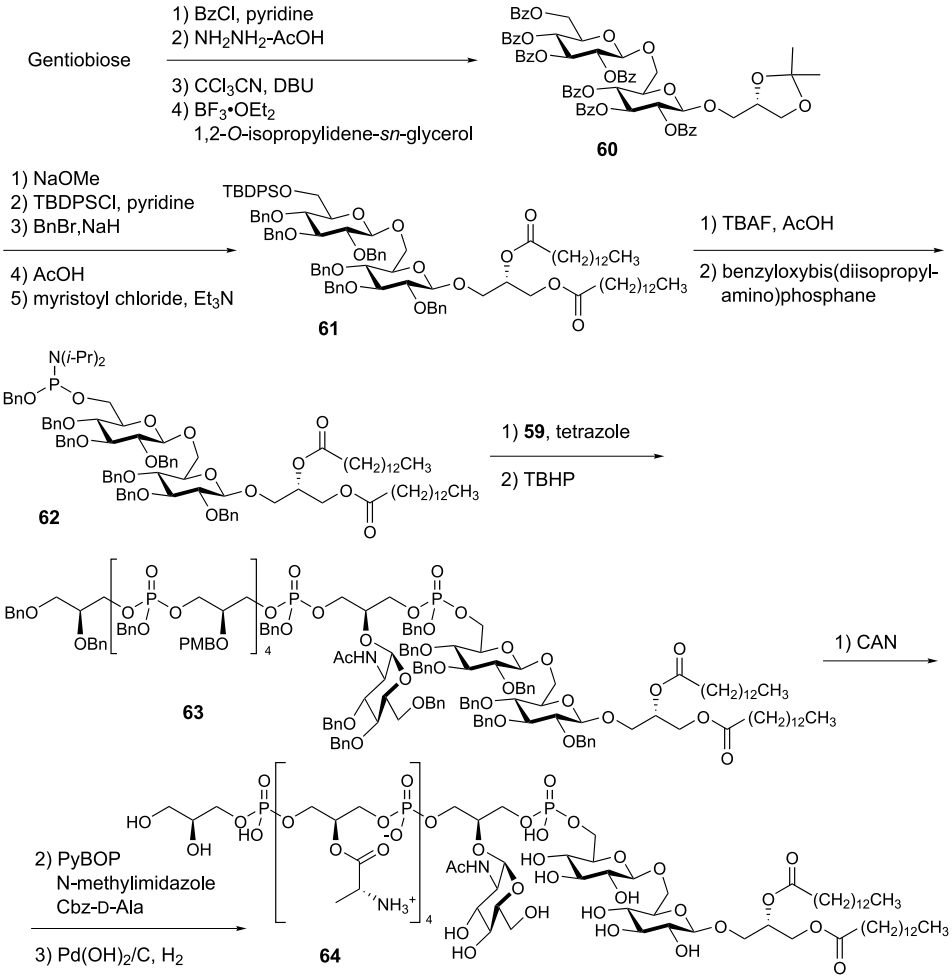
Scheme 12



Scheme 13

Hexaphosphoglycerol **59** was prepared as shown in *Scheme 13*. Condensation of **56** with **57** in the presence of tetrazole, addition of *tert*-butyl hydroperoxide, and then treatment with TBAF was performed. This reaction was repeated four times to give tri-phosphoester intermediate **58**, which was successively reacted with **55**. Oxidation and desilylation of the resultant complex furnished the desired intermediate **59**.

Synthesis of glycolipid **62** was initiated from commercially available gentiobiose (*Scheme 14*). Gentiobiose was fully benzoylated under standard conditions and treated with hydradine-acetic acid. The resulting 1-hydroxy group was reacted with trichloroacetonitrile under DBU, which was then condensed with 1,2-*O*-isopropylidene-*sn*-glycerol in the presence of a catalytic amount of boron trifluoride ether complex to give glycoside **61**. After removal of the benzoyl group, the primary hydroxy group was selectively reacted with TBDP-



Scheme 14

SCl and the remaining hydroxyl group was benzylated. Acidic cleavage of the isopropylidene group and then condensation of the resulting hydroxy group with myristoryl chloride furnished compound **61**. Desilylation of **61** with TBAF followed by the same phosphitylation gave the desired glycolipid **62**. Reaction of **62** with **59** in the presence of tetrazole and then addition of *tert*-butyl hydroperoxide (TBHP) afforded compound **63** in 75% yield. Treatment of **63** with ceric(IV) ammonium nitrate (CAN) resulted in four glycerol hydroxy groups, which were then coupled with benzyloxycarbonyl (Cbz) protected D-alanine residue in the presence of benzotriazol-1-yloxy)tripyrrolidinophosphonium hexafluorophosphate (PyBOP) to afford the fully protected LTA analogue. Target molecule **64** was generated by hydrogenolysis of the resultant complex with Pearlman's catalyst.

5 Synthesis of Mycoloyl Arabinan

Mycobacterium bovis Bacillus Calmette–Guérin (BCG) was established for the vaccine against tuberculosis. The BCG cell wall skeleton (BCG-CWS) contains a very complex and polymeric glycoconjugate, composed of mycolated arabinogalactan linked to peptidoglycan (PG) through linker disaccharide (β-Rhap-(1-3)-α-GlcNAc) (*Fig. 9*) [117]. BCG-CWS is related to the innate immunity [118,119]. It was suggested that it activates macrophages via Toll-like receptors [120,121,122]. However, it is not obvious which structures are responsible to the activation due to the inherent complexity of BCG-CWS. Recently, synthesis of a non-reducing terminal structure of BCG-CWS consisting of pentaarabinofuranose and mycolate was achieved by Ito's group [123], and the activity was evaluated extensively. This section deals with their work.

The non-reducing terminal structure of BCG-CWS contains both α- and β-arabinofuranoses. To construct the structure [124,125,126,127,128], the methodology for both α- and β-con-

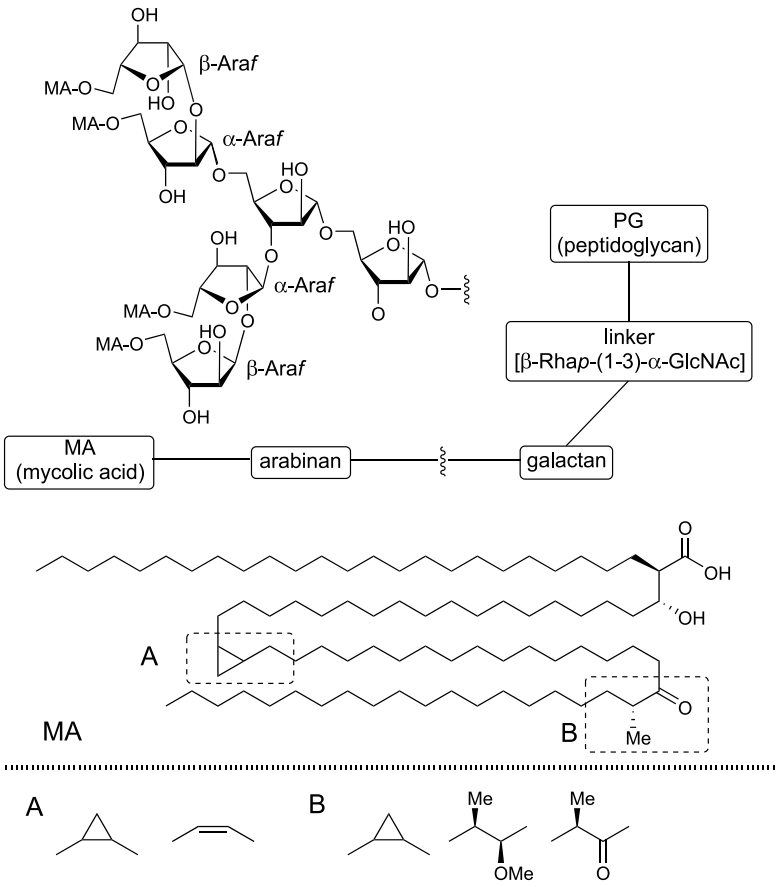
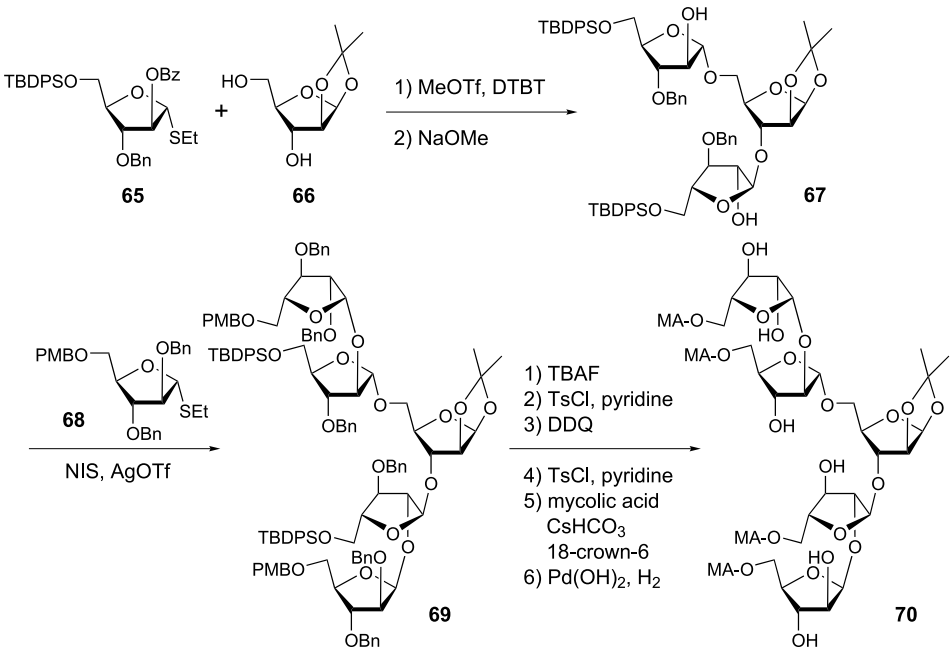


Figure 9
BCG-CWS structure and synthetic target, terminal pentaarabino mycolate



Scheme 15

figurated arabinofuranoside was required. However, the basic method for the stereoselective synthesis of the furanosidic linkage had not matured, compared with that of the pyranosidic linkage. In addition, the construction of β-configurated arabinofuranoside was a challenging issue because of the 1,2-*cis* orientation [129].

Synthesis of mycolated pentafuranoside **70** was performed as shown in *Scheme 15*. Glycosylation of **66** [130] with thioglycoside **65** [131] under conditions using methyl trifluoromethanesulfonate and 2,6-di-*tert*-butyl-4-methylpyridine (DTBMP) and subsequent deacetylation gave trisaccharide **67**. The resulting hydroxy groups in **67** were then glycosylated with thioglycoside **68**, which was the optimized substrate for α-linked glycoside, to give the desired pentaarabinoside **69** in 57% isolated yield, together with 35% yield of a stereoisomeric mixture. The mycolate moieties were efficiently introduced via S_N2-type alkylative esterification with tosylate as the eliminating group. After deprotection of TBDPS and PMB groups and subsequent tosylation, the resulting tosylated pentaarabinoside was treated with mycolic acid (MA) in the presence of cesium hydrogencarbonate and 18-crown-6 ether. Finally, the mycolated pentaarabinoside **70** was prepared by removing the benzyl protecting groups with $Pd(OH)_2$ and H_2.

6 Synthesis of Re-Type Lipopolysaccharide (Re-LPS)

LPS is located in the outer membrane of Gram-negative bacteria and composed of *O*-antigen polysaccharide, core polysaccharide, and lipid A. It was found that lipid A is the key compo-

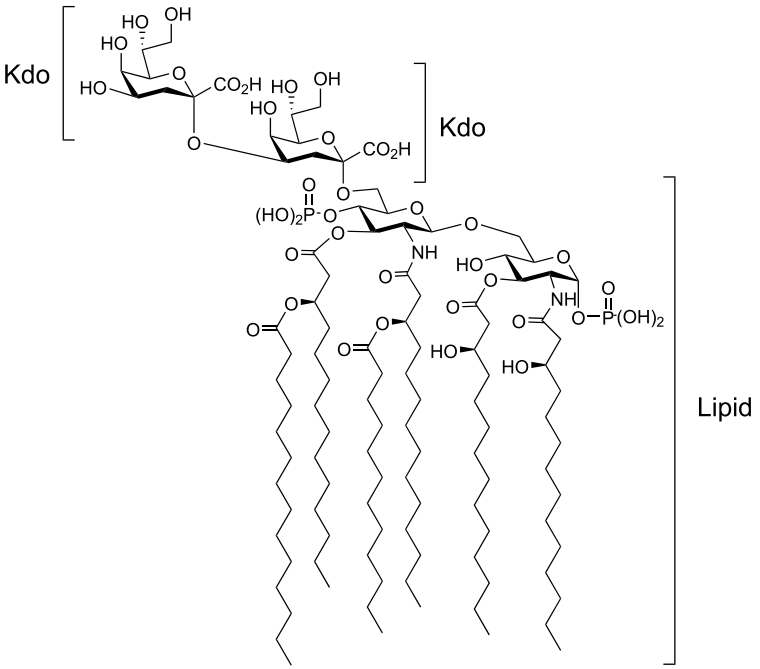
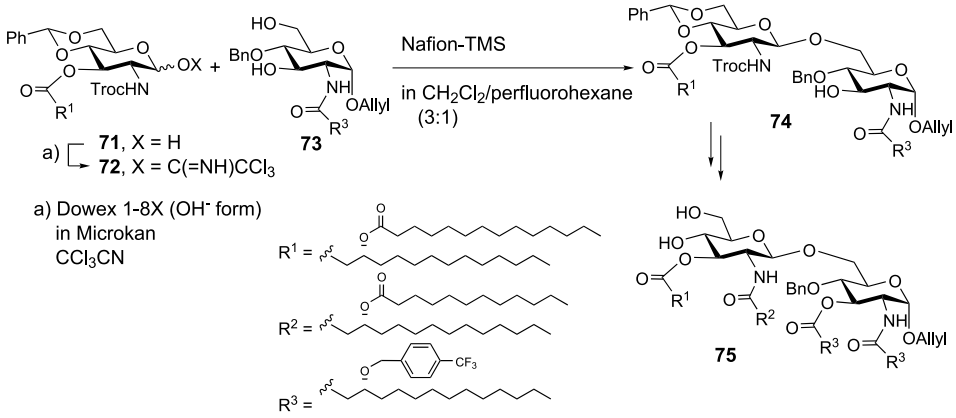


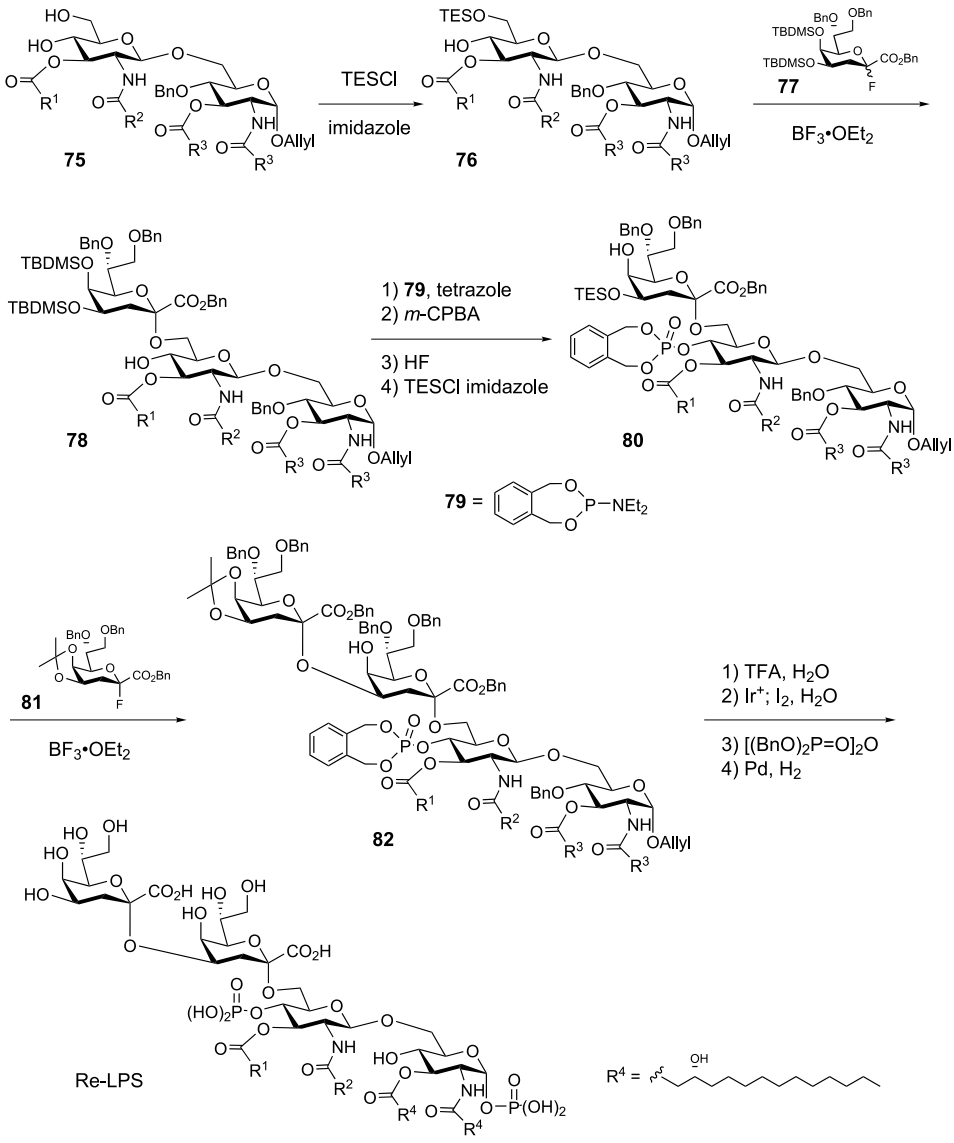
Figure 10
Structure of Re-type LPS

nent for the endotoxic activity of LPS [132]. Among LPS, Re-type LPS from the *Escherichia coli* Re mutant strain is known as the smallest LPS. The structure of Re-LPS is composed of lipid A and two units of 3-deoxy-D-*manno*-2-octuronic acid (Kdo) (*Fig. 10*) [133]. The structure contains highly acid-labile glycosyl phosphate and base-labile ester functional



Scheme 16

groups. For these reasons, although partial syntheses of Re-LPS had been reported [134,135, 136], the first total synthesis of Re-LPS was achieved by Kusumoto et al. in 2001 [137]. An acylated disaccharide **75** was efficiently prepared using polymer-supported reagents. Formation of glycosyl imidate **72** was carried out with Dowex 1-8X (OH^- form) as a polymer-supported base, and glycosylation of acceptor **73** with the imidate **72** was carried out using



Scheme 17

Nafion-TMS as a Lewis acid [138,139,140]. The work-up in both reactions was very easy compared with conventional methods using inorganic or organic bases because of the removal of resin from the reaction mixture (*Scheme 16*). The desired β(1-6)-glycoside **74** was thus prepared in 73% yield. Disaccharide **75** was then synthesized via 3-*O*-acylation, removal of the Troc group, *N*-acylation, and debenzylidenation.

Synthesis of Re-type LPS was then started with disaccharide **75** (*Scheme 17*). The key step in the synthesis was an efficient and stereoselective glycosylation with Kdo residue. The glycosylation was accomplished using TBDMS-protected fluoride **77** as a donor and TES ether **76** as an acceptor in the presence of BF_3Et_2O [141], affording the desired α(2-6)-glycoside in 89% yield with complete regio- and stereoselectivity. The remaining hydroxy group was phosphorylated, and subjected to desilylation of TBDMS and subsequent selective silylation to afford TES trisaccharide **80**. A second glycosylation proceeded successfully with isopropylidene-protected fluoride **81** to furnish the desired α(2-4)-glycoside **82** in 75% yield with complete regio- and stereoselectivity. In contrast, using TBDMS-protected fluoride **77** gave a low yield of **82** (26%), probably due to the steric repulsion between both substrates. After deprotection of isopropylidene and allyl groups, the anomeric hydroxy group was selectively phosphorylated. Finally, the protected precursor **82** was then subjected to hydrogenation with Pd black to afford Re-LPS. Pure Re-LPS was obtained by liquid-liquid partition column chromatography using a two-layered solvent system (*n*-BuOH/THF/H_2O/MeOH = 16/7/1/20) on Sephadex LH-20 gel.

7 Synthesis of Glycosylphosphatidylinositol (GPI)

Glycosylphosphatidylinositols are naturally occurring glycophospholipids on a cell surface. Full structural assignment was reported by Ferguson et al. in 1988 [142,143]. The core structure was composed of a carbohydrate residue with an ethanolamine side chain, inositol residue, and a phosphoglycerolipid residue (*Fig. 11*). The carbohydrate residue, ethanolamine side chain or lipid moiety is species specific [144,145,146,147,148], and depends on species and

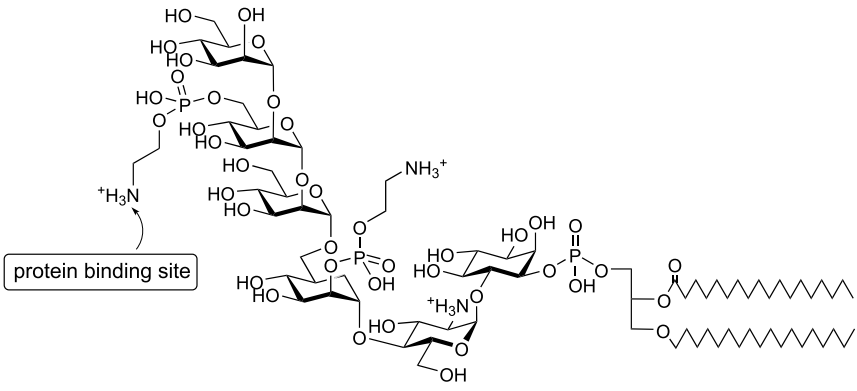
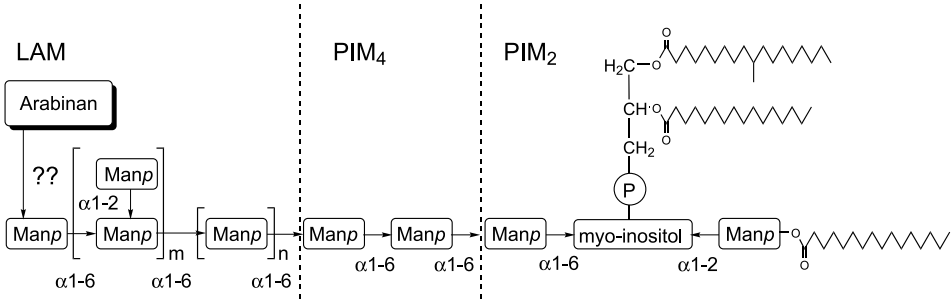


Figure 11
Typical structure of GPI anchor of rat brain Thy-1

on cell-type. Total synthesis of GPI anchors was started in the early 1990s and successfully carried out for a ceramide containing GPI anchors of yeast [149,150] and for acylglycerol containing GPI anchors of *Trypanosoma brucei* [151,152], of rat brain Thy-1 [153,154], of *Plasmodium falciparum* [155], and of *Mycobacterium tuberculosis* [156,157,158,159,160,161,162,163,164]. This section deals with recent work by Seeberger et al. on the total synthesis of phosphatidylinositol mannoside (PIM).

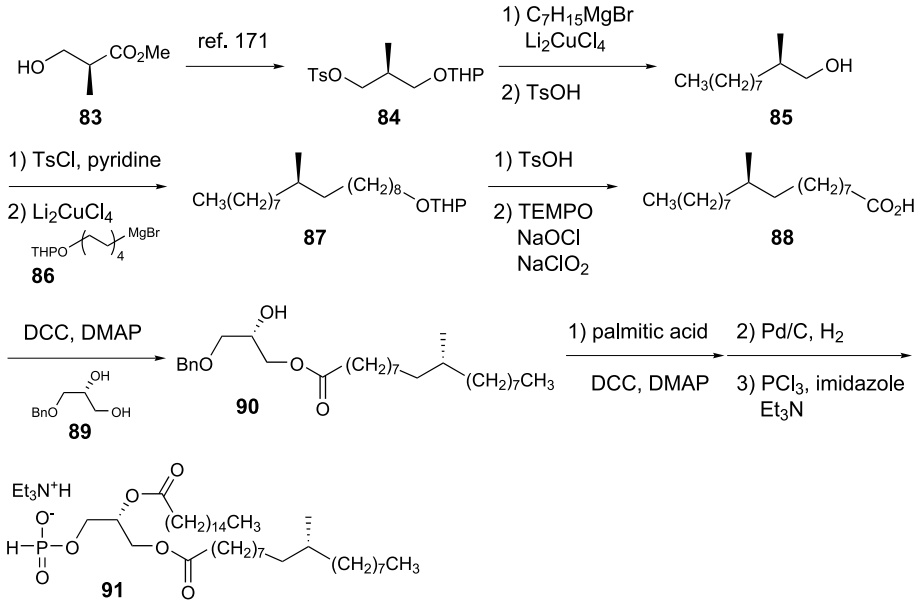
PIM is included in the fraction of lipoarabinomannan (LAM), which is the major component of the mycobacterial envelope accompanied with mycoloyl arabinogalactan-peptidoglycan complex. The structures of PIMs and LAMs are shown in ❯ *Fig. 12* [165,166]. PIMs are terminal components of LAMs, which consist of diacylglycerphosphorylinositol residue, acyl mannosyl residue, and oligomannoside residue. The diacylglycerol moiety contains the carboxylic acid named tuberculostearic acid (TBSA) [167,168]. Although LAMs have been shown to exert profound physiological effects and emerge to modulate the host immune response [169,170], PIMs, particularly PIM_2 and PIM_6, that are predominantly found in the cell wall, show an immune response and act as agonists of Toll-like receptor 2 (TLR2) [165,166].



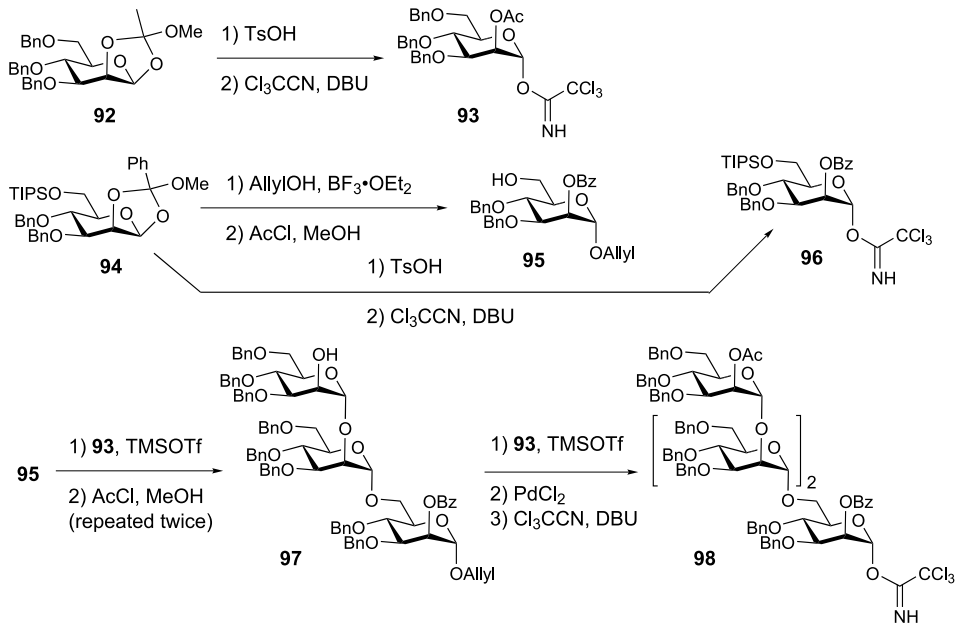
❯ Figure 12
Structure of LAM (typical lipid composition shown)

The glycerophospholipid moiety was prepared as shown in ❯ *Scheme 18*. The synthesis was started with (*S*)-3-hydroxy-2-methylpropionate **83**, a commercially available chiral building block, which was converted to tosylate **84** according to the literature [171]. Cross-coupling of tosylate **84** and heptylmagnesium bromide in the presence of a catalytic amount of Li_2CuCl_4 followed by treatment with toluenesulfonic acid gave alcohol **85**. Alcohol **85** was then tosylated and was subjected to the second Cu-catalyzed alkylation. After removal of the THP group in **87**, the resultant complex was converted to the desired tuberculostearic acid **88** by TEMPO-catalyzed oxidation. Regioselective acylation of glycerol **89** with acid **88** using DCC and DMAP gave the monoacylated compound **90** in 75% yield. The remaining hydroxy group was acylated with palmitic acid using DCC, and the benzyl protecting group was removed by hydrogenolysis. H-Phosphonate **91** was obtained by the treatment with phosphoryl trichloride and imidazole.

Syntheses of mannosyl and oligomannosyl moieties are shown in ❯ *Scheme 19*. Mannosyl imidates **93** and **96** were prepared from the appropriate orthoester **92** and **94** via hydrolysis with TsOH and then treatment with CCl_3CN and DBU, respectively. Allyl mannoside **95** was also prepared from orthoester **94** via glycosylation with allyl alcohol and then desilylation of



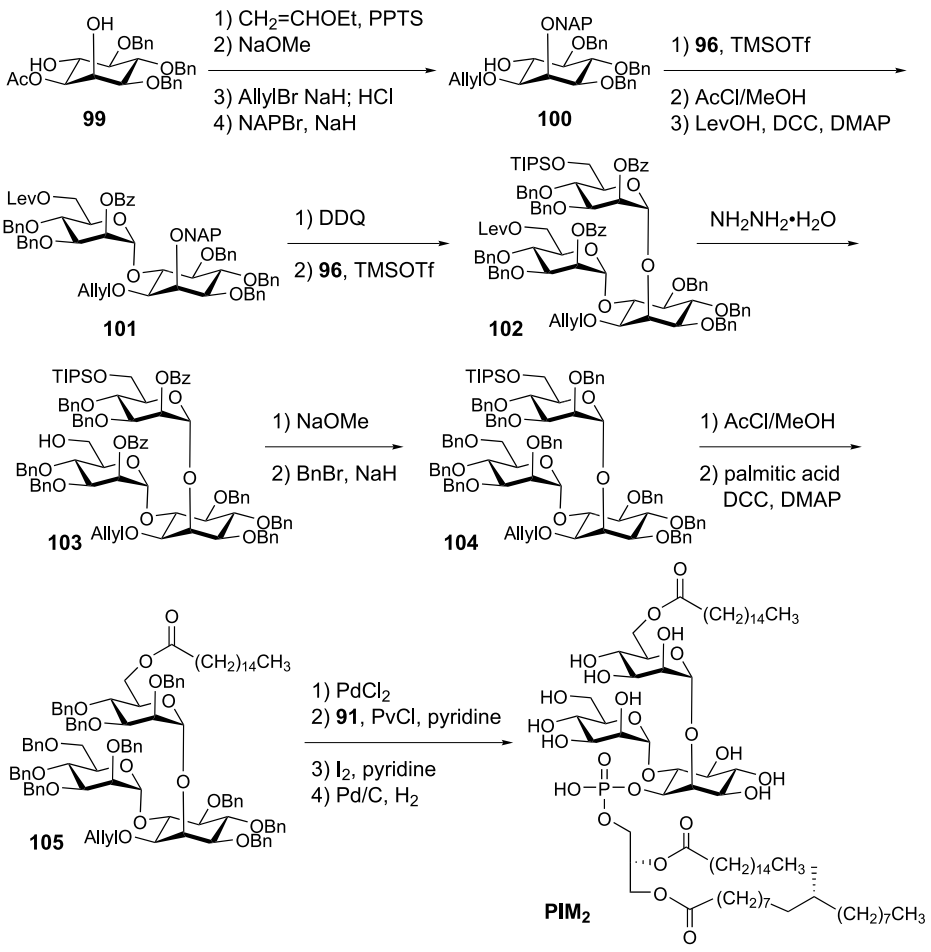
Scheme 18



Scheme 19

the TIPS group. Trimannoside **97** was derived from imidate **93** and allyl mannoside **95** via sequential glycosylation and deprotection of the acetyl group. Tetramannosyl imidate **98** was prepared from **97** by the further glycosylation with imidate **93**, removal of the anomeric allyl group, and treatment with CCl_3CN and DBU.

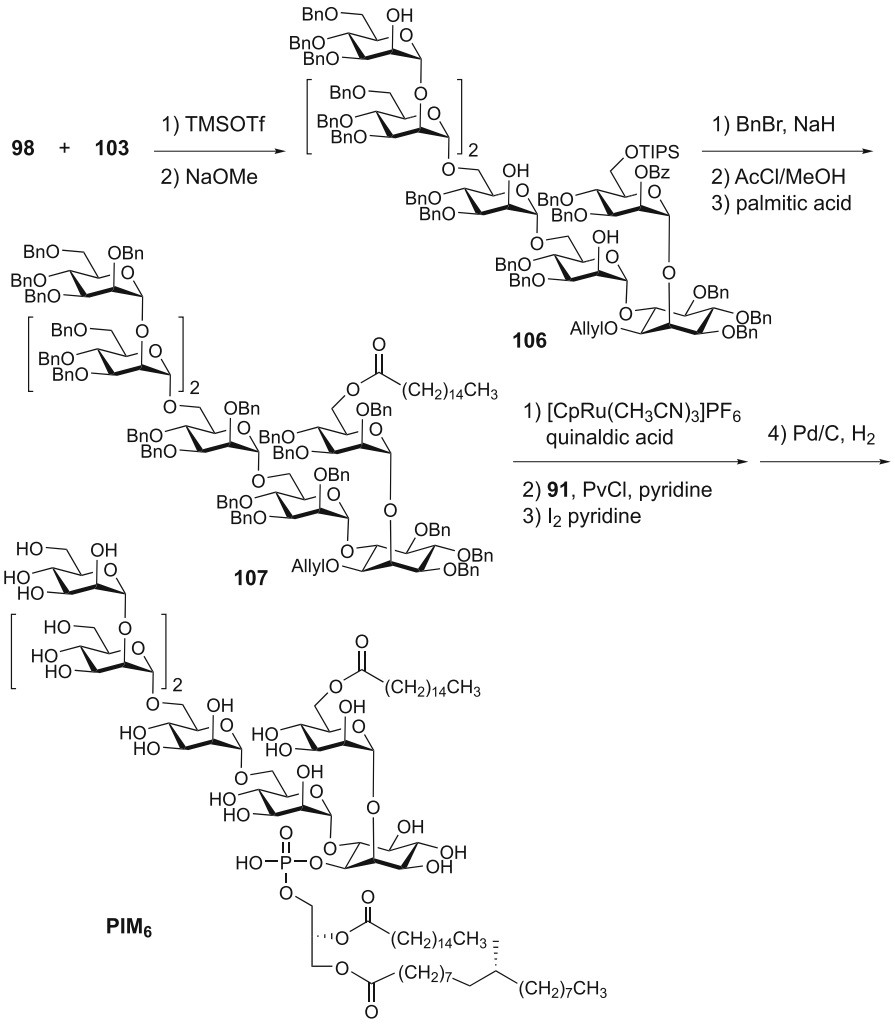
Synthesis of PIM_2 is illustrated in *Scheme 20*. The enantio pure inositol derivative **99** prepared from methyl α-D-glucoside [172] was led to the compound **100** via temporal masking with an ethoxyethyl group at the 2- and 6-positions, deacetylation and subsequent allylation at the 1-position, and introduction of a naphthyl (NAP) group at the 2-position. After glycosylation of **100** with imidate **96**, the triisopropylsilyl (TIPS) group was converted to the levulinoyl (Lev) group. The NAP group of mannosyl inositol **101** was removed by treatment with DDQ, and the resulting hydroxy group was glycosylated with imidate **96** again. Removal of the Lev group gave a common intermediate **103** for the syntheses of PIM_2 and PIM_6. After removal of



Scheme 20

benzoyl groups, the resulting hydroxy groups were benzylated. Desilylation and subsequent acylation with palmitic acid afforded palmitate **105**. The desired PIM_2 was obtained via deallylation, condensation with H-phosphonate **91**, oxidation of phosphonate to phosphate, and hydrogenolysis. At this stage, removal of the allyl group in inositol was found to be a problematic reaction because of a simultaneously occurring Wacker-type oxidation. In addition, an iridium complex was also ineffective in this reaction. This problem was solved using a cationic Ru^{II} complex [173] as described in the synthesis of PIM_6.

PIM_6 was synthesized as shown in ❯ *Scheme 21*. Glycosylation of dimannosyl inositol **103** with tetramannosyl imidate **98** was performed using TMSOTf, followed by deacetylation to



Scheme 21

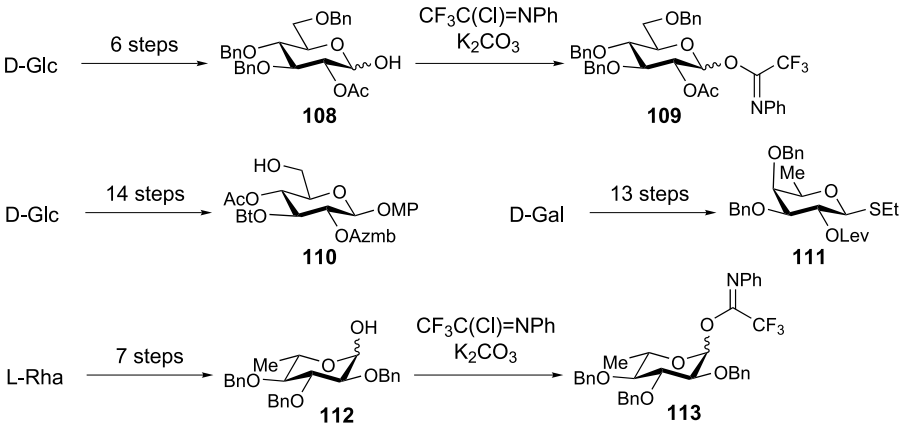
give hexamannosyl inositol **106**. After benzylation of hydroxy groups, the TIPS group was removed. The resulting hydroxy group was acylated using palmitic acid and DCC. Removal of the allyl group in compound **107** was then evaluated. Among many explorations, it was found that a stoichiometric cationic $CpRu^{II}$ was effective for the deallylation, while the use of cationic Pd^{II} or Ir^{I} was ineffective. After removal of the allyl group, the resulting hydroxy was phosphorylated with H-phosphonate **91**. Finally, PIM_6 was accomplished by hydrogenolysis of the obtained phosphate.

8 Synthesis of Other Glycolipids

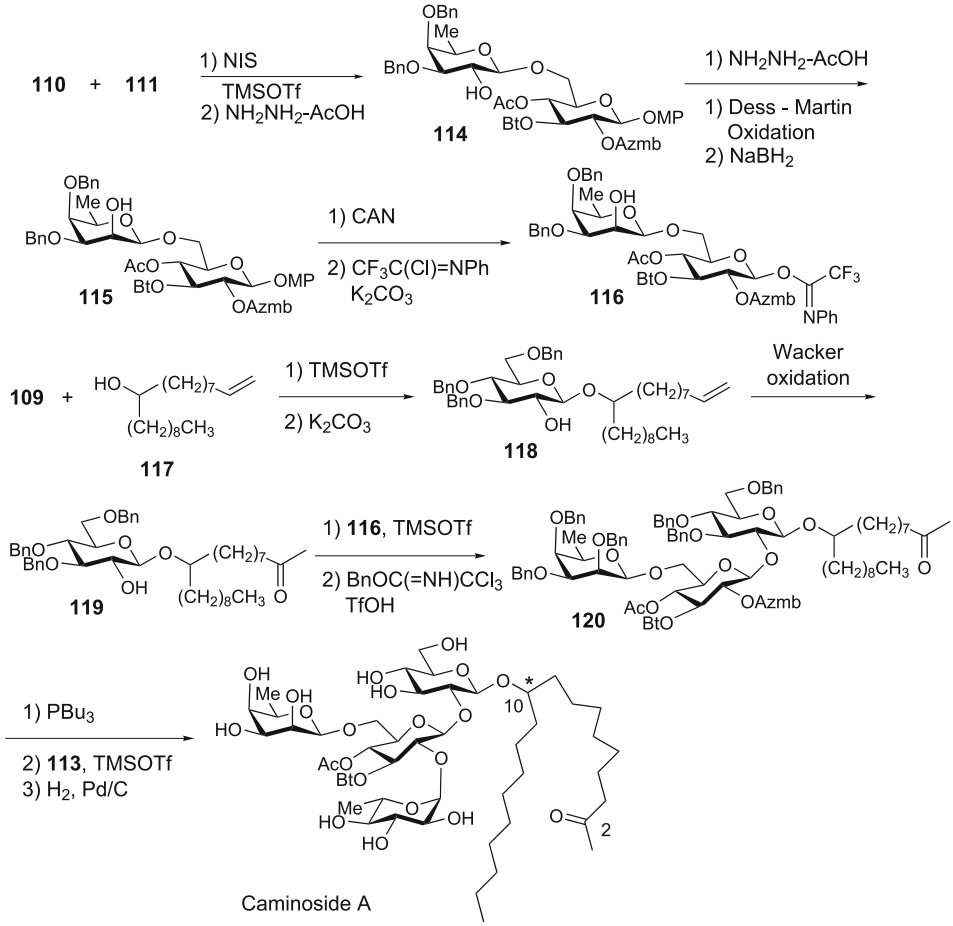
8.1 Total Synthesis of Caminoside A

Caminoside A is a novel antimicrobial glycolipid containing a tetrasaccharide part, which was isolated from marine sponge *Caminus sphaeroconia*, and was reported to have an inhibitory activity against a bacterial type III secretory system [174]. The tetrasaccharide part is composed of two D-glucose, 6-deoxy-D-talose, and L-quinovose with a 1,2-*cis*-β-mannopyranose-type linkage of 6-deoxy-D-talose. That may be the most difficult linkage to construct stereoselectively. Other structural features are (i) the center glucose residue is fully substituted, (ii) 6-deoxy-D-talose and L-quinovose are rare in nature, and (iii) methyl ketone lipid aglycon, in which the stereochemistry at the 10-position was recently determined [175], is unprecedented in marine sponge metabolites.

The first total synthesis of Caminoside A was recently achieved by Yu's group [176]. In their synthetic strategy, an indirect approach for the construction of the 1,2-*cis*-β-mannopyranose-type linkage of 6-deoxy-D-talose was performed: β-selective glycosylation with 1-thiofucopyranoside **111** introduced 2-*O*-Lev as a directing group followed by an inversion of the 2-OH configuration affords the desired 6-deoxy-1,2-*cis*-β-talopyranosidic linkage. Monosaccharide building blocks **109**, **110**, **111**, and **113** were prepared (*Scheme 22*).



Scheme 22



Scheme 23

As shown in *Scheme 23*, glycosylation of **110** with **111** using NIS/TMSOTf [177,178] afforded the β-disaccharide **114** in 76% yield. After selective removal of the 2-*O*-Lev group in the presence of Ac, Bt, and Azmb, the resulting equatorial 2-OH group was converted to the axial one via oxidation, followed by selective reduction. No equatorial isomer was obtained in the reduction step. The oxidative cleavage of the anomeric MP group followed by the treatment with trifluoroacetimidoyl chloride gave the disaccharide trifluoroimidate **116**. The axial 2-OH group was not reacted due to the steric hindrance. Glycosylation of 1-nonadecen-10-ol (**117**) with 2-*O*-Ac glucosyl donor **109** was carried out using TMSOTf [179] and subsequent removal of the 2-*O*-Ac group provided the desired glycoside **118**. A terminal C=C double bond in the lipid moiety was converted into a methyl ketone functional group by Wacker oxidation [180,181] at this stage. Glycosylation of **119** with the disaccharide trifluoroacetimide **116** was performed under conditions similar to those for **118**. The remaining free axial

2-OH group of the talose moiety was then protected with a benzyl group by the reaction with benzyl trichloroimidate in the presence of a catalytic amount of TfOH, to prevent an unexpected glycosylation at the next step. After selective removal of the 2-*O*-Azmb group of **120** with tributylphosphine, the resulting 2-OH group was glycosylated with trifluoroimidate **113** under conditions similar to those for **118** and **120**. The obtained fully protected caminoside A derivative was then deprotected by hydrogenolysis in the presence of Pd/C to accomplish the synthesis of caminoside A.

8.2 Total Synthesis of Woodrosin I

Resin glycosides obtained from plants belonging to the morning glory family (*Convolvulaceae*) are composed of complex oligosaccharides and (11*S*)-hydroxyhexadecanoic acid (jalapinolic acid) as a common aglycon in this series. This aglycon lipid moiety frequently forms a macrolactone ring spanning two or more sugar units of the backbone [182,183,184,185,186,187,188,189,190,191,192]. This intricate structure stimulated the interest of synthetic chemists, and many works have been devoted to it [193,194,195,196,197,198,199,200,201]. Among them, three independent approaches for tricolorin A [196,197,198,199,200,201] and for the congener tricolorin G [200,201] are focused on here. Their structures are illustrated in ❯ *Fig. 13*. The first synthetic study of tricolorin A was reported by Heathcock'group in 1997 [196,197], where Yonemitsu protocol lactonization was utilized as a key macrocyclization.

Fürstner's group reported syntheses of tricolorin A and G [200,201] based on macrocyclization by ring-closing metathesis (RCM), affording macrocycles in high yields. In this section, total synthesis of woodrosin I [202], which may possess the most intricate structure in this series, reported by Fürstner's group is described [203,204,205].

Woodrosin I is composed of four glucose, one rhamnose, and lipid moieties. The lipid consists of a 27-membered macrolide ring spanning four glucose units. The convergent synthesis of woodrosin I was performed via intermediates **134** and **135**, and the macrocyclization was achieved by RCM using Grubbs carbene ruthenium complex or phenylindenylidene ruthe-

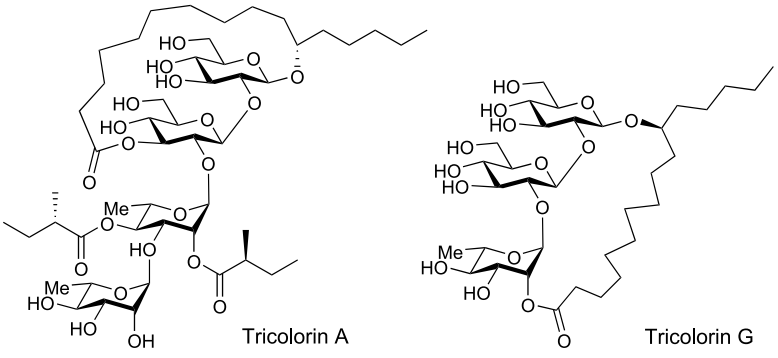
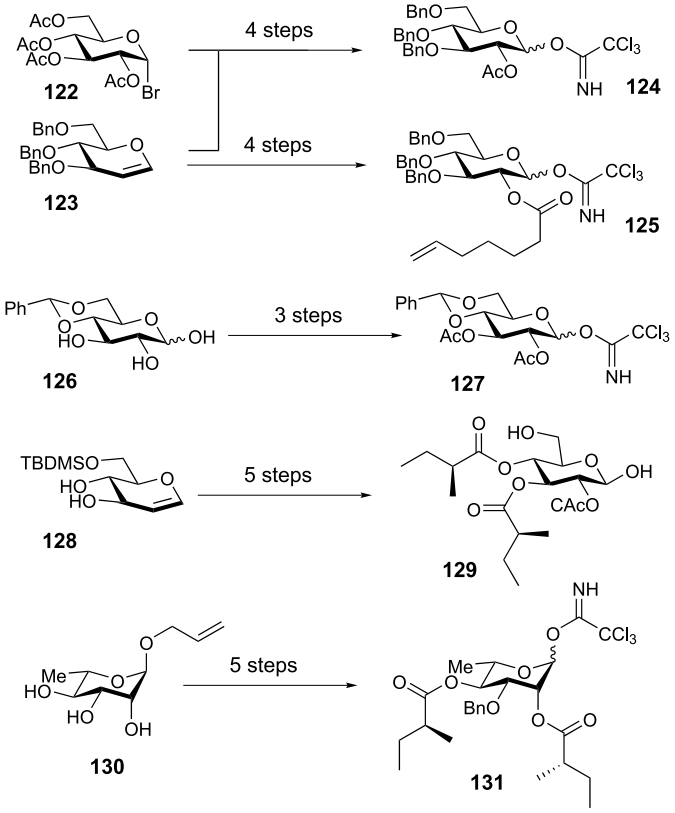


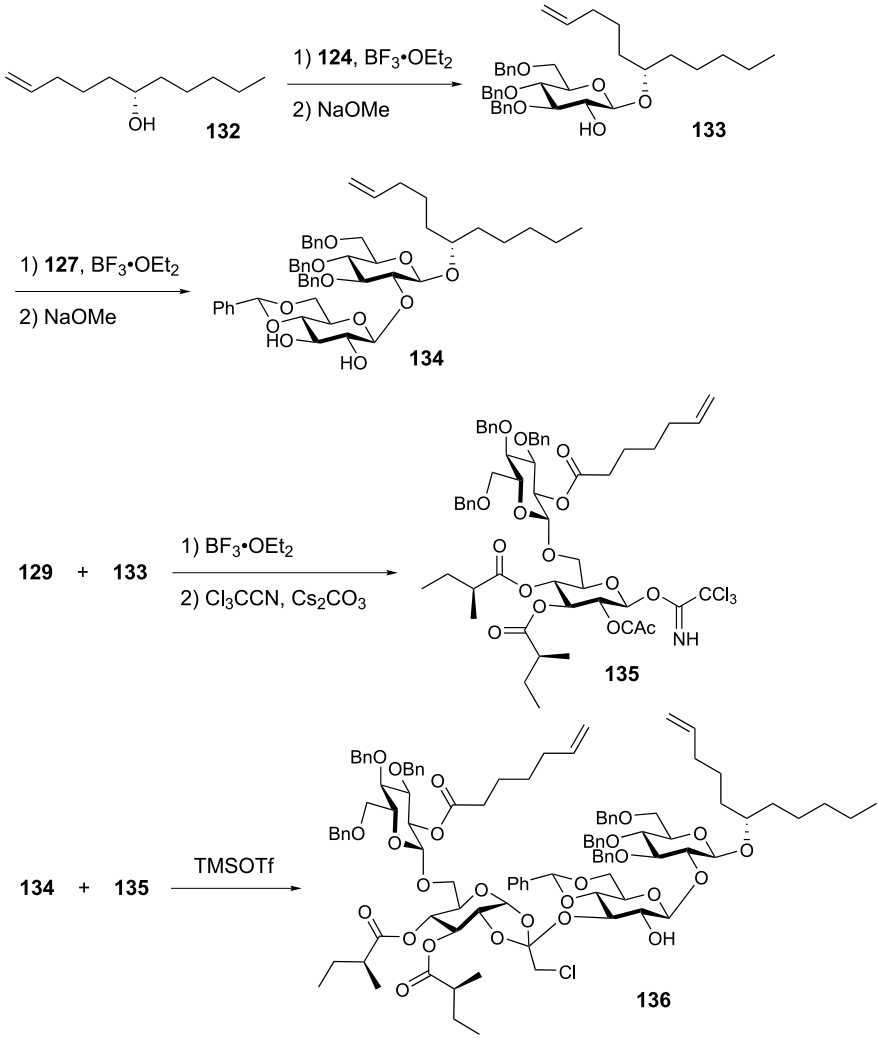
Figure 13
Chemical structures of tricolorin A and G



Scheme 24

nium complex. Glucose and rhamnose building blocks **124**, **125**, **127**, **129**, and **131** were derived from appropriate monosaccharide components as shown in *Scheme 24*. The reducing terminal glucosyl donor **124** was prepared from glycosyl bromide **122** or commercially available benzyl glucal **123** in four steps. Non-reducing terminal glucosyl donor **125** was also prepared from benzyl glucal **123**. A second glucosyl donor **127** was derived from 4,6-benzylidene acetal **126** [206] in three steps. A third glucose moiety **129** was prepared from TBDMS-protected glucal **128** [207] in five steps. Rhamnosyl donor **131** was derived from allyl L-rhamnoside **130** [208].

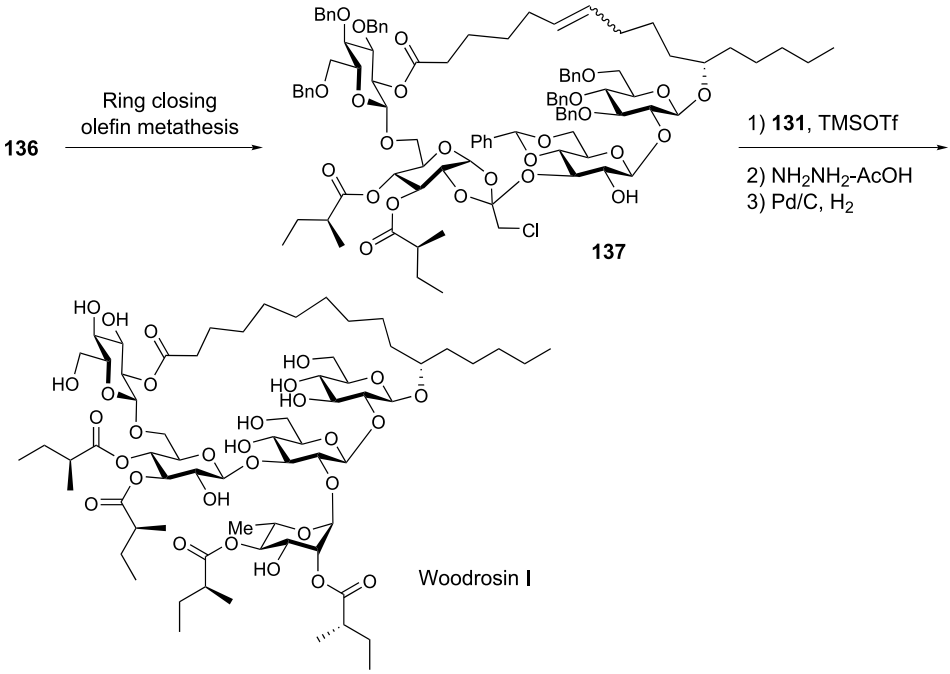
As shown in *Scheme 25*, the glycosylation of alcohol **132**, which was prepared by an enantioselective condensation of 5-hexanal and dipentylzinc in the presence of $Ti[O(i\text{-}Pr)]_4$ and chiral bis-trifluoromethanesulfonamide [209], with glucosyl donor **124** followed by treatment of NaOMe gave glycoside **133**. The resultant 2-OH group was then glycosylated with 4,6-benzylidene donor **127** under similar conditions to those described above. Removal of the acetyl group at the 2′- and 3′-positions was then carried out to afford right-hand building block **134**. Glycosylation of acceptor **133** with donor **129** proceeded regioselectively at the 6-position of the acceptor to form the corresponding glycoside, which was then treated with trichloroace-



Scheme 25

tonitrile and cesium carbonate to furnish left-hand building block **135**. Then, the glycosylation of diol **134** with **135** in the presence of TMSOTf was performed, where the coupling reaction proceeded regioselectively at the more reactive 3′-position without unexpected formation of orthoester **136**, although the neighboring effect of the chloroacetyl group may engage β-glycoside in the presence of a strong Lewis acid catalyst. In addition, rhamnosylation to the 2′-position of orthoester **131** did not occur because of steric hindrance exerted by the orthoester moiety.

The formation of the macrocycle (*Scheme 26*), which had been a serious problem, was achieved by ring-closing olefin metathesis (RCM). RCM of orthoester **136** was performed



Scheme 26

with Grubbs carbene ruthenium complex [210,211] or phenylindenylidene ruthenium complex [212] as a catalyst to afford the corresponding macrocycle **137** in high yield with an *E*:*Z* mixture (*E*:*Z* = 9:1). The rearrangement to β-glycoside and the introduction of the rhamnose moiety to the 2′-position of **137** were effectively accomplished by an "inverse glycosylation procedure" developed by Schmidt [213], where glycosyl donor **131** was added slowly in a premixed solution of acceptor **137** and TMSOTf. Final deprotection was performed by treatment with hydrazinium acetate and subsequent hydrogenolysis with Pd/C to furnish the desired woodrosin I.

8.3 Total Synthesis of Cycloviracin B_1

Cycloviracin B_1 (see *Fig. 14*) [214,215], isolated from the soil microorganisms *Kibdelosporangium albatum* (R761-7) [216], is a macrocyclic glycolipid showing antiviral activity [214,215]. This compound is composed of five glucose and two acyl units with different acyl chain lengths. Two of the glucose units form the macrocyclic dilactone (macrodiolide). The constitution of the cycloviracin was elucidated by an extensive NMR study. However, several structural details, such as the absolute stereochemistry of the six chiral centers on the acyl units, were solved by the total synthesis of Fürstner's group [217,218].

The synthesis was performed according to a two-directional synthesis strategy [219] since the macrodiolide core part in the target compound has C_2-symmetry. The synthetic pathway is

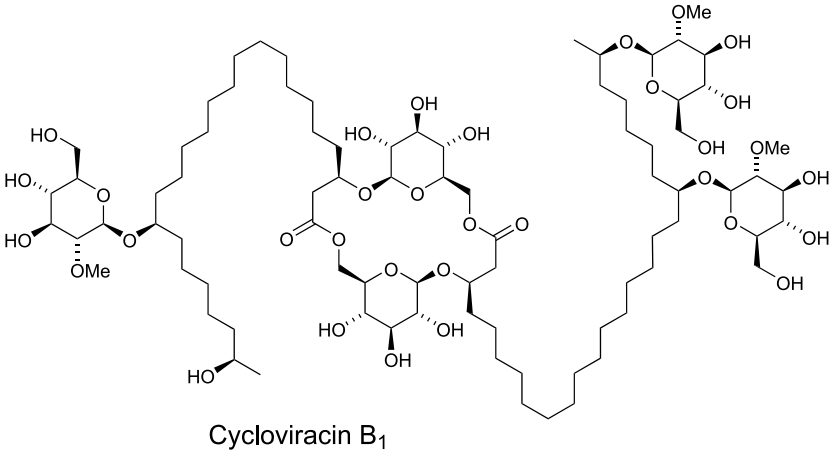
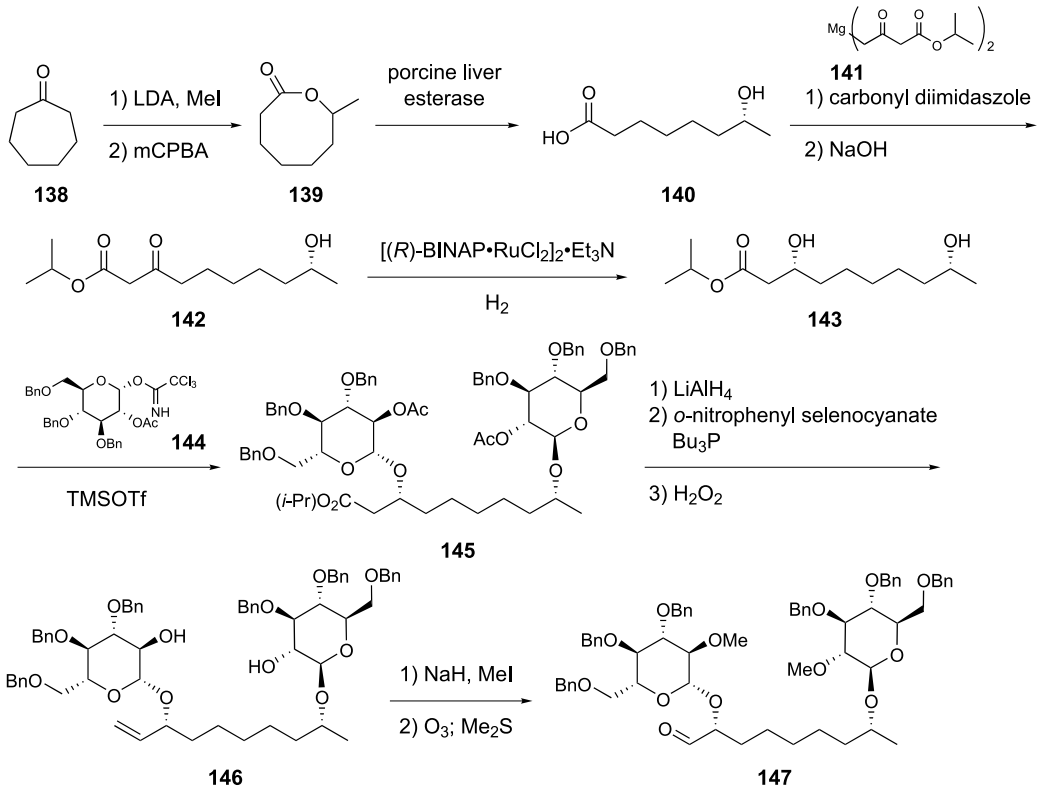


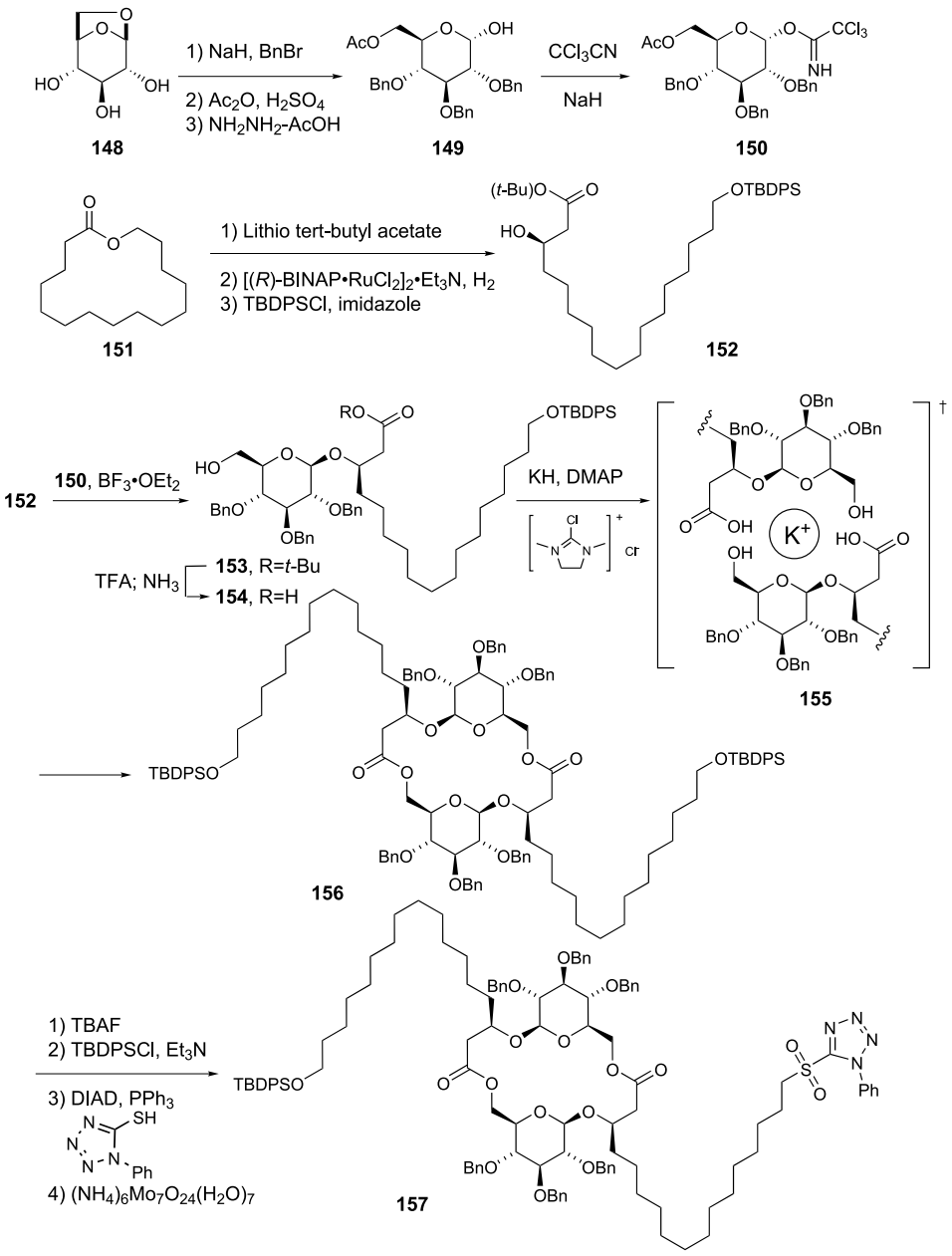
Figure 14
Structure of cycloviracin B_1

shown in *Scheme 27*, *Scheme 28* and *Scheme 29*. α-Methylation of cycloheptanone **138** and subsequent Baeyer–Villiger oxidation gave the desired lactone **139**, which was then subjected to enzymatic hydrolysis using an esterase from porcine liver, resulting in carboxylic acid **140** in a high enantiopurity (95% ee). Compound **140** was activated with carbonyl diimidazole and reacted with magnesium carboxylate **141** [220,221] to afford β-keto ester **142**. Stereoselective hydrogenation of the ketone at the β-position of **142** was carried out in the presence of catalytic [(*R*)-BINAPRuCl$_2$]$_2$•Et$_3$N [222,223] to afford diol **143** in quantitative yield with a high diastereoselectivity (de = 99%). Glycosylation of diol **143** with glycosyl imidate donor **144** [224] in the presence of TMSOTf gave glycoside **145** with a complete β-selectivity. Subsequent reduction of ester groups in **145** with $LiAlH_4$ gave the corresponding triol, whose primary alcohol then reacted with *o*-nitrophenylselenocyanate in the presence of Bu_3P. The resulting selenide was oxidized with aqueous H_2O_2, and was converted to the corresponding alkene **146** with spontaneous elimination of selenide [225,226]. Methylation of the remaining secondary hydroxy groups followed by ozonolysis of the alkene moiety in **146** furnished aldehyde **147**, which represents a fully functional right-side component.

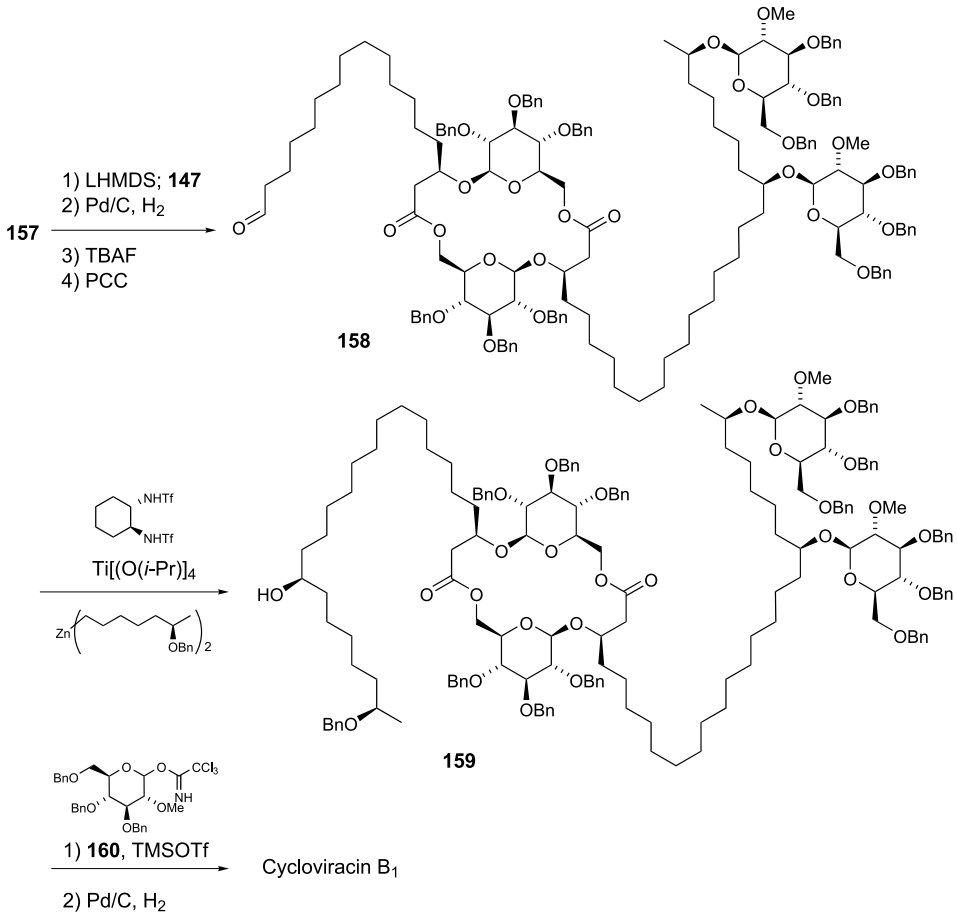
Macrodiolide core **156** was synthesized as shown in *Scheme 28*. Commercially available 1,6-anhydoglucose **148** was benzylated, and then subjected to acetolysis with Ac_2O and H_2SO_4. Selective deacetylation with hydrazine acetate gave acetal **149**. Glycosyl donor **150** was prepared by the treatment with trichloroacetonitrile and a catalytic amount of NaH. Claisen condensation of pentadecanolide **151** with lithio *tert*-butyl acetate gave the corresponding β-keto ester, which was subsequently hydrogenated with [(*R*)-BINAPRuCl$_2$]$_2$•Et$_3$N as described above. Silylation of the resultant primary hydroxy group with TBDPSCl afforded β-hydroxy ester **152**. Glycosylation of **152** with glycosyl imidate **150** was performed in the presence of BF_3•OEt using the β-directing effect of acetonitrile to furnish the corresponding glycoside **153** (68%, α:β = 1:5). Subsequent treatment of **153** with TFA gave carboxylic acid **154**. Macrodiolide core **156** was constructed by a template-directed cyclodimerization. Namely, cyclization precursor **154** was reacted with 2-chloro-1,3-dimethylimidazolium chlo-


1) LDA, MeI
2) mCPBA
porcine liver esterase
1) carbonyl diimidaszole
2) NaOH
138
139
140
141
[(R)-BINAP•RuCl2]2•Et3N
H2
142
143
144
TMSOTf
1) LiAlH4
2) o-nitrophenyl selenocyanate
Bu3P
3) H2O2
145
1) NaH, MeI
2) O3; Me2S
146
147

Scheme 27


1) NaH, BnBr
2) Ac2O, H2SO4
3) NH2NH2-AcOH
CCl3CN
NaH
148
149
150
1) Lithio tert-butyl acetate
2) [(R)-BINAP•RuCl2]2•Et3N, H2
3) TBDPSCl, imidazole
151
152
150, BF3•OEt2
KH, DMAP
TFA; NH3
153, R=t-Bu
154, R=H
155
156
1) TBAF
2) TBDPSCl, Et3N
3) DIAD, PPh3
4) (NH4)6Mo7O24(H2O)7
157

Scheme 28



Scheme 29

ride as a dehydrating agent in the presence of a K^+ cation to furnish macrodiolide **156** in 71% isolated yield. In this reaction, it was suggested that the K^+ cation worked as a promotor for preorganization of the cyclization precursor **154** and resulted in chelate intermediate **155**. Desymmetrization of terminal functional groups on lateral chains was accomplished by desilylation and monosilylation of hydroxy groups. Mitsunobu reaction of the remaining terminal hydroxy group and subsequent oxidation gave sulfone compound **157**.

Introduction of the right segment **147** into **157** was performed by Julia–Kocienski olefination [227], and the resultant alkene was hydrogenated. After removal of the TBDPS group, the resulting hydroxy group was oxidized by pyridinium chlorochromate (PCC) to afford aldehyde **158**. Left lateral chains of cycloviracin was constructed by condensation of organozinc in the presence of Ti[O(*i*-Pr)]$_4$ and (*S*,*S*)-ditriflate. Compound **159** was then glycosylated with **160**, which was subjected to final deprotection by hydrogenolysis, affording the desired cycloviracin B_1.

References

1. Bell RM, Hannun YA, Merrill AH (1993) Advances in lipid research: Sphingolipids and their metabolism. Academic Press, Orland, FL, 25:26
2. Kanfer JN, Hakomori S (1983) Sphingolipid biochemistry. In: Handbook of lipid research. Plenum Press, New York
3. Devant RM (1992) Kontakte (Darmstadt) 11
4. Byun HS, Bittman R (1993) Chemical preparation of sphingosine and sphingolipids: a review of enantioselective synthesis. In: Cevc G (ed) Phospholipids handbook. Dekker, New York, p 97
5. Vanlerberghe G (1996) Oil, Crops Grass, Lipids 3:365
6. Schmidt RR (1996) Spec Publ R Soc Chem 180:93
7. Koskinen PM, Koskinen AMP (1998) Synthesis 1075
8. Jung KH, Schmidt RR (1999) In: Gunstone FD (ed) Lipid synthesis and manufacture. Sheffield Academic, Sheffield, UK, p 208
9. Curfman C, Liotta D (1999) Methods Enzymol 311:391
10. Koskinen PM, Koskinen AMP (1999) Methods Enzymol 311:458
11. Vanker YD, Schmidt RR (2000) Chem Soc Rev 29:201
12. Liao J, Tao J, Lin G, Liu D (2005) Tetrahedron 61:4715
13. Boutin RH, Rapoport H (1986) J Org Chem 51:5320
14. Koike K, Numata M, Sugimoto M, Nakamura Y, Ogawa T (1986) Carbohydr Res 158:113
15. Bettelli E, Chinzari P, D'Andrea P, Passacantilli P, Piancatelli G, Topai A (1996) Korean J Med Chem 6:339
16. Murakami T, Minamikawa H, Hato M (1994) Tetrahedron Lett 35:745
17. Zimmermann P, Schmidt RR (1988) Liebigs Ann Chem 663
18. Hirata N, Yamagiwa Y, Kamikawa T (1991) J Chem Soc Perkin Trans 1 2279
19. Yadav JS, Vidyanand D, Rajagopal D (1993) Tetrahedron Lett 34:1191
20. Katsumura S, Yamamoto N, Fukuda E, Iwama S (1995) Chem Lett 393
21. Mulzer J, Brand C (1986) Tetrahedron 42:5961
22. Schmidt RR, Maier T (1988) Carbohydr Res 174:169
23. Somfai P, Olsson R (1993) Tetrahedron 49:6645
24. Metz K, Honda M, Komori T (1993) Liebigs Ann Chem 55
25. Yamanoi T, Akiyama T, Ishida E, Abe H, Amemiya M, Inazu T (1989) Chem Lett 335
26. Ito Y, Sawamura M, Hayashi T (1988) Tetrahedron Lett 29:239
27. Kobayashi S, Hayashi T, Kawasuji T (1994) Tetrahedron Lett 35:9573
28. Groth U, Schoelkopf U, Tiller T (1991) Tetrahedron 47:2835
29. Umemura T, Mori K (1987) Agric Biol Chem 51:1973
30. Soai K, Takahashi K (1994) J Chem Soc Perkin Trans 1 1257
31. Davis FA, Reddy GV (1996) Tetrahedron Lett 37:4349
32. Nugent TC, Hudlicky T (1998) J Org Chem 63:510
33. Olofsson B, Somfai P (2003) J Org Chem 68:2514
34. Lee JM, Lim HS, Chung SK (2002) Tetrahedron Asymmetry 13:343
35. Torssell S, Somfai P (2004) Org Biomol Chem 2:1643
36. Rai AN, Basu A (2004) Org Lett 6:2861
37. Rai AN, Basu A (2005) J Org Chem 70:8228
38. Chaudhari VD, Kumar KSA, Dhavale DD (2005) Org Lett 7:5805
39. Chatterjee AK, Choi TL, Sander DP, Grubbs RH (2003) J Am Chem Soc 125:11360
40. Howell AR, Ndakala AJ (2002) Curr Org Chem 6:365
41. Mitsutaka S, Kita K, Okino N, Ito M (1997) Anal Biochem 247:52
42. Toyokuni T, Nisar M, Dean B, Hakomori S (1991) J Labelled Compd Radiopharm 29:567
43. Wild R, Schmidt RR (1994) Tetrahedron Asymmetry 5:2195
44. Anand JK, Sadozai KK, Hakomori S (1996) Lipids 31:995
45. Yatomi Y, Yamamura S, Ruan F, Kume S, Ozaki Y, Igarashi Y (1997) FEBS Lett 417:341
46. Bielawska A, Hannum YA, Szulc Z (2000) Methods Enzymol 311:480
47. Imbs AB, Molotkovskii YG, Bergel'son LD (1986) Bioorg Khim 12:527
48. Pagano RE, Martin OC (1988) Biochemistry 27:4439

49. Mikhalev II, Molotkovskii YG, Bergel'son LD (1990) Bioorg Khim 16:685
50. Marchesini S, Demasi L, Cestone P, Preti A, Agmon V, Dagan A, Navon R, Gatt S (1994) Chem Phys Lipids 72:143
51. Gege C, Oscarson S, Schmidt RR (2001) Tetrahedron Lett 42:377
52. Zhou XT, Forestier C, Goff RD, Li C, Teyton L Bendelac A, Savage PB (2002) Org Lett 4:1267
53. Vo-Hoang Y, Micouin L, Ronet C, Gachelin G, Bonin M (2003) Chem Bio Chem 4:27
54. Schwarzmann G, Wendeler M, Sandhoff K (2005) Glycobiology 15:1302
55. Boons BJ, Demchenko AV (2000) Chem Rev 100:4539
56. Kanie O, Kiso M, Hasegawa A (1988) J Carbohydr Chem 7:501
57. Murase T, Ishida H, Kiso M, Hasegawa A (1988) Carbohydr Res 184:C1
58. Ogura H, Furuhata K, Ito M, Shitori Y (1986) Carbohydr Res 158:37
59. Hasegawa A, Ohki H, Nagahama T, Ishida H, Kiso M (1991) Carbohydr Res 212:277
60. Hasegawa A, Ogawa M, Kojima Y, Kiso M (1992) J Carbohydr Chem 11:333
61. Ishida H, Ishida H, Kiso M, Hasegawa A (1994) J Carbohydr Chem 13:655
62. Hotta K, Komba S, Ishida H, Kiso M, Hasegawa A (1994) J Carbohydr Chem 13:665
63. Ishida H, Ishida H, Kiso M, Hasegawa A (1994) Carbohydr Res 260:C1
64. Hasegawa A, Nagahama T, Ohki H, Hotta K, Ishida H, Kiso M (1991) J Carbohydr Chem 10:493
65. Martin TJ, Schmidt RR (1992) Tetrahedron Lett 33:6123
66. Kondo H, Aoki S, Ichikawa Y, Halcomb RL, Ritzen H, Wong CH (1994) J Org Chem 59:864
67. Demchenko AV, Boons GJ (1998) Tetrahedron Lett 39:3065
68. Demchenko AV, Boons GJ (1999) Chem Eur J 5:1278
69. Meo CD, Demchenko AV, Boons GJ (1999) J Org Chem 66:5490
70. Adachi M, Tanaka H, Takahashi T (2004) Synlett 609
71. Yu CS, Niikura K, Lin CC, Wong CH (2001) Angew Chem Int Ed 40:2900
72. Ando H, Koike Y, Ishida H, Kiso M (2003) Tetrahedron Lett 44:6883
73. Tanaka K, Goi T, Fukase K (2005) Synlett 2958
74. Tanaka H, Nishiura Y, Takahashi T (2006) J Am Chem Soc 128:7124
75. Cai S, Yu B (2003) Org Lett 5:3827
76. Ichikawa Y, Lin YC, Dumas DP, Shen Gj, Garcia-Junceda E, Williams MA, Bayer R, Ketcham C, Walker LE, Paulson JC, Wong CH (1992) J Am Chem Soc 114:9283
77. Herrmann GF, Ichikawa Y, Wandrey C, Gaeta FCA, Paulson JC, Wong CH (1993) Tetrahedron Lett 34:3091
78. Yu H, Chokhawala H, Karpel R, Yu H, Wu B, Zhang J, Zhang Y, Jia Q, Chen X (2005) J Am Chem Soc 127:17618
79. Yu H, Huang S, Chokhawala H, Sun M, Zheng H, Chen X (2006) Angew Chem Int Ed 45:3938
80. Nudelman E, Fukushi Y, Levery SB, Higuchi T, Hakomori S (1986) J Biol Chem 261:5487
81. Ando T, Ishida H, Kiso M (2001) J Carbohydr Chem 20:425
82. Ando T, Ishida H, Kiso M (2003) Carbohydr Res 338:503
83. Kameyama A, Ishida H, Kiso M, Hasegawa A (1989) Carbohydr Res 193:C1
84. Marra A, Sinaÿ P (1989) Carbohydr Res 187:35
85. Yoshida M, Uchimura A, Kiso M, Hasegawa A (1993) Glycoconjugate J 10:3
86. Hasegawa A, Suzuki N, Ishida H, Kiso M (1996) J Carbohydr Chem 15:623
87. Komba S, Ishida H, Kiso M, Hasegawa A (1996) Bioorg Med Chem 4:1833
88. Chen X, Wu YL, Chen D (2002) Tetrahedron Lett 43:3529
89. Miyamoto K, Miyake S, Yamamura T (2001) Nature 413:531
90. Natori N, Koezuka Y, Higa T (1993) Tetrahedron Lett 34:5591
91. Morita M, Motoki K, Akimoto K, Natori T, Sakai T, Sawa E, Yamaji K, Koezuka Y, Kobayashi E, Fukushima H (1995) J Med Chem 38:2176
92. Costantino V, Fattorusso E, Mangoni A, Rosa MD, Ianaro J (1997) J Am Chem Soc 119:12465
93. Nicolaou KC, Li J, Zenke G (2000) Helv Chim Acta 83:1977
94. Seki M, Kayo A, Mori K (2001) Tetrahedron Lett 42:2357
95. Charette AB, Juteau H, Lebel H, Molinaro C (1998) J Am Chem Soc 120:11943
96. Charette AB, Prescott S, Brochu C (1995) J Org Chem 60:1081
97. Charette AB, Marcoux JF (1996) J Am Chem Soc 118:4539
98. Charette AB, Lebel H (1999) Org Synth 76:86
99. Kolb HC, VanNieuwenhze MS, Sharpless KB (1994) Chem Rev 94:2483

100. Sharpless KB, Amberg W, Bennani YL, Crispino GA, Hartung J, Jeong KS, Kwong HL, Morikawa K, Wang ZM, Yu D, Zhang ZL (1992) J Org Chem 57:2768
101. Sprott GD (1992) J Bioenerg Biomembr 24:555
102. Gigg J, Gigg R (1990) Synthesis of glycoglycerolipids. In: Kates M (ed) Handbook of lipid research, vol 6. Plenum, New York, chap 7
103. Eguchi T, Arakawa K, Terachi T, Kakinuma K (1997) J Org Chem 62:1924
104. Eguchi T, Kano H, Arakawa K, Kakinuma K (1997) Bull Chem Soc Jpn 70:2545
105. Eguchi T, Ibaragi K, Kakinuma K (1998) J Org Chem 63:2689
106. Menger FM, Chen XY, Brocchini S, Hopkins HP, Hamilton D (1993) J Am Chem Soc 115:6600
107. Menger FM, Chen XY (1996) Tetrahedron Lett 37:323
108. Arakawa K, Eguchi T, Kakinuma K (1998) J Org Chem 63:4741
109. Fischer W (1994) Med Microbiol Immunol 183:61
110. Morath S, Geyer A, Hartung T (2001) J Exp Med 193:393
111. Morath S, Stadelmaier A, Geyer A, Schmidt RR, Hartung T (2002) J Exp Med 195:1635
112. van Boeckel CAA, van Boom JH (1985) Tetrahedron 41:4567
113. Fukase K, Matsumoto T, Ito N, Yoshimura T, Kotani, S, Kusumoto S (1992) Bull Chem Soc Jpn 65: 2643
114. Fukase K, Yoshimura T, Kotani, S, Kusumoto S (1994) Bull Chem Soc Jpn 67:473
115. Stadelmaier A, Morath S, Hartung T, Schmidt RR (2003) Angew Chem Int Ed 42:916
116. Stadelmaier A, Figueroa-Perez I, Deininger S, von Aulock S, Hartung T, Schmidt RR (2006) Bioorg Med Chem 14:6239
117. Besra GS, Khoo KH, McNeil M, Dell A, Morris R, Brennan PJ (1995) Biochemistry 34:4257
118. Alexandroff AB, Jackson AM, O'Donnell MA, James K (1999) Lancet 353:1689
119. Begum NA, Ishii K, Kurita-Takeuchi M, Tanabe M, Kobayashi M, Moriwaki Y, Matsumoto M, Fumumori Y, Azuma I, Toyoshima K, Seya T (2004) Infect Immun 72:937
120. Tsuji S, Matsumoto M, Takeuchi O, Akira S, Azuma I, Hayashi A, Toyoshima K, Seya T (2000) Infect Immun 68:6883
121. Underhill DM, Ozinsky A, Smith KD, Aderem A (1999) Proc Natl Acad Sci USA 96:14459
122. Krieg AM (2003) Nat Med 9:831
123. Ishiwata A, Akao H, Ito Y, Sunagawa M, Kusunose N, Kashiwazaki Y (2006) Bioorg Med Chem 14:3049
124. D'Souza FW, Lowary TL (2000) Org Lett 2:1493
125. Yin H, D'Souza FW, Lowary TL (2002) J Org Chem 67:892
126. Mereyala HB, Hotha S, Gurjar MK (1998) Chem Commun 685
127. Wang H, Ning J (2003) J Org Chem 68:2521
128. D'Souza FW, Ayers JD, McCarren PR, Lowary TL (2000) J Am Chem Soc 122:1251
129. Demchenko AV (2003) Synlett 1225
130. Martin OR, Rao SP, El-Shenawy HA, Kurz KG, Cutler AB (1988) J Org Chem 53:3287
131. Callam CS, Gadikota RR, Krein DM, Lowary TL (2003) J Am Chem Soc 125:13112
132. Brade H, Opal SM, Vogel SN, Morrison DC (eds) (1999) Endotoxin in health and disease. Marcel Dekker, New York
133. Zähringer U, Lindner B, Seydel U, Rietschel ET, Naoki H, Unger FM, Imoto M, Kusumoto S, Shiba T (1985) Tetrahedron Lett 26:6321
134. Imoto M, Kusunose N, Kusumoto S, Shiba T (1988) Tetrahedron Lett 29:2227
135. Kusumoto S, Kusunose N, Kamikawa T, Shiba T (1988) Tetrahedron Lett 29:6325
136. Fukase K, Kurosawa M, Kusumoto S (1994) J Endotoxin Res 1:149
137. Yoshizaki H, Fukuda N, Sato K, Oikawa M, Fukase K, Suda Y, Kusumoto S (2001) Angew Chem Int Ed 40:1475
138. Demuth M, Mikhail G, George MV (1981) Helv Chim Acta 64:2759
139. Olah GA, Iyer PS, Prakash GKS (1986) Synthesis 513
140. Haunert F, Bolli MH, Hinzen B, Ley SV (1998) J Chem Soc Perkin Trans 1:2235
141. Nicolaou KC, Chucholowski A, Dolle RE, Randall JL (1984) J Chem Soc Chem Commun 1:2235
142. Ferguson MAJ, Homans SW, Dwek RA, Rademacher TW (1988) Science 239:753
143. Homans SW, Ferguson MAJ, Dwek RA, Rademacher TW, Anand R, Williams, AF (1988) Nature 333:269
144. Ferguson MAJ, Williams AF (1988) Annu Rev Biochem 57:285
145. Ferguson MAJ (1991) Curr Opin Struct Biol 1:522
146. Thomas R, Dwek RA, Rademacher TW (1990) Biochemistry 29:5413
147. Cross GAM (1990) Annu Rev Cell Biol 6:1

148. Field MC (1992) Glycoconjugate J 9:155
149. Mayer TG, Kratzer B, Schmidt RR (1994) Angew Chem Int Ed Engl 33:2177
150. Mayer TG, Kratzer B, Schmidt RR (1999) Eur J Org Chem 1153
151. Murakata C, Ogawa T (1992) Carbohydr Res 234:75
152. Murakata C, Ogawa T (1992) Carbohydr Res 235:95
153. Campbell AS, Fraser-Reid B (1995) J Am Chem Soc 117:10387
154. Tailler D, Ferriéres V, Pekari K, Schmidt RR (1999) Tetrahedron Lett 40:679
155. Liu XY, Kwon YU, Seeberger PH (2005) J Am Chem Soc 127:5004
156. Elie CJJ, Dreef CE, Verduyn R, Vandermarel GA, van Boom JH (1989) Tetrahedron 45:3477
157. Elie CJJ, Verduyn R, Dreef CE, Brounts DM, Vandermarel GA, van Boom JH (1990) Tetrahedron 46:8243
158. Elie CJJ, Verduyn R, Dreef CE, Vandermarel GA, van Boom JH (1992) J Carbohydr Chem 11:715
159. Watanabe Y, Yamamoto T, Ozaki S (1996) J Org Chem 61:14
160. Watanabe Y, Yamamoto T, Okazaki T (1997) Tetrahedron 53:903
161. Stadelmaier A, Schmidt RR (2003) Carbohydr Res 338:2557
162. Jayaprakash KN, Lu J, Fraser-Reid B (2004) Bioorg Med Chem Lett 14:3815
163. Stadelmaier A, Biskup MB, Schmidt RR (2004) Eur J Org Chem 3292
164. Liu XY, Bridget LS, Stocker Seeberger PH (2006) J Am Chem Soc 128:3638
165. Gilleron M, Ronet C, Mempel M, Monsarrat B, Gachelin G, Puzo G (2001) J Biol Chem 276:34896
166. Gilleron M, Quesniaux VFJ, Puzo G, (2003) J Biol Chem 278:29880
167. Anderson RJ, Chargaff E (1929) J Biol Chem 85:77
168. Spielman MA (1934) J Biol Chem 106:87
169. Chatterjee D, Khoo KH (1998) Glycobiology 8:113
170. Nigou J, Gilleron M, Rojas M, Garcia LF, Thurnher M, Puzo G (2002) Microb Infect 4:945
171. Mori K, Takikawa H (1991) Libigs Ann Chem 497
172. Jia ZJ, Olsson L, Fraser-Reid B (1998) J Chem Soc Perkin Trans 1 631
173. Tanaka S, Saburi H, Ishibashi Y, Kitamura M (2004) Org Lett 6:1873
174. Linington RG, Robertson M, Gauthier A, Finlay BB, van Soest R, Andersen RJ (2002) Org Lett 4:4089
175. MacMillan JB, Linington RG, Andersen RJ, Molinski TF (2004) Angew Chem Int Ed 43:5946
176. Sun J, Han X, Yu B (2005) Synlett 433
177. Konradsson P, Udodong UE, Fraser-Raid B (1990) Tetrahedron Lett 31:4313
178. Hunt DK, Seeberger PH (2002) Org Lett 4:2751
179. Yu B, Tao H (2001) Tetrahedron Lett 42:2405
180. Smith AB, Cho YS, Friestad GK (1998) Tetrahedron Lett 39:8765
181. Nicolaou KC, Gray D, Tea J (2001) Angew Chem Int Ed 40:3675
182. Pereda-Miranda R, Mata R, Anaya AL, Wickramaratne DBM, Pezzuto JM, Kinghorn AD (1993) J Nat Prod 56:571
183. Noda N, Yoda S, Kawasaki T, Miyahara K (1992) Chem Pharm Bull 40:3163
184. Noda N, Ono M, Miyahara K, Kawasaki T, Okabe M (1987) Tetrahedron 43:3889
185. Ono M, Honda F, Karahashi A, Kawasaki T, Miyahara K (1997) Chem Pharm Bull 45: 1955
186. Noda N, Takahashi N, Kawasaki T, Miyahara K, Yang CR (1994) Phytochemistry 36:365
187. Ono M, Kawasaki T, Miyahara K (1989) Chem Pharm Bull 37:3209
188. Noda N, Tsuji K, Miyahara K, Yang CR (1994) Chem Pharm Bull 42:2011
189. Kitagawa I, Baek NI, Ohashi K, Sakagami M, Yoshikawa M, Shibuya H (1989) Chem Pharm Bull 37:1131
190. Ono M, Kuwabata K, Kawasaki T, Miyahara K (1992) Chem Pharm Bull 40:2674
191. Gaspar EMM (1999) Tetrahedron Lett 40:6861
192. Enriquez RG, Leon I, Perez F, Walls F, Carpenter KA, Puzzuoli FV, Reynolds WF (1992) Can J Chem 70:1000
193. Jiang ZH, Geyer A, Schmidt RR (1995) Angew Chem Int Ed Engl 34:2520
194. Furukawa J, Kobayashi S, Nomizu M, Nishi N, Sakairi N (2000) Tetrahedron Lett 41: 3453
195. Fürstner A, Radkowski K, Grabowski J, Wirtz C, Mynott R (2000) J Org Chem 65:8758
196. Larson DP, Heathcock CH (1996) J Org Chem 61:5208
197. Larson DP, Heathcock CH (1997) J Org Chem 62:8406
198. Lu SF, O'yang Q, Guo ZW, Yu B, Hui YZ (1997) Angew Chem Int Ed Engl 36:2344

199. Lu SF, O'yang Q, Guo ZW, Yu B, Hui YZ (1997) J Org Chem 62:8400
200. Fürstner A, Müller T (1999) J Am Chem Soc 121:7814
201. Fürstner A, Müller T (1998) J Org Chem 63:424
202. Ono M, Nakagawa K, Kawasaki T, Miyashita K (1993) Chem Pharm Bull 41:1925
203. Fürstner A, Jeanjean F, Razon P (2002) Angew Chem Int Ed 41:2097
204. Fürstner A, Jeanjean F, Razon P, Wirtz C, Mynott R (2003) Chem Eur J 9:307
205. Fürstner A, Jeanjean F, Razon P, Wirtz C, Mynott R (2003) Chem Eur J 9:320
206. Barili PL, Berti G, Catelani G, Cini C, D'Andrea F, Mastrorilli E (1995) Carbohydr Res 278:43
207. Nicolaou KC, Mitchell HJ, Jain NF, Bando T, Hughes R, Winssinger N, Natarajan S, Koumbis AE (1999) Chem Eur J 5:2648
208. Zhang J, Mao J, Chen H, Cai M (1994) Tetrahedron Asymmetry 5:2283
209. Takahashi H, Kawakita T, Ohno M, Yoshioka M, Kobayashi S (1992) Tetrahedron 48: 5691
210. Schwab P, France MB, Ziller JW, Grubbs RH (1995) Angew Chem Int Ed Engl 34:2039
211. Schwab P, Grubbs RH, Ziller JW (1996) J Am Chem Soc 118:100
212. Fürstner A, Guth O, Düffels A, Seidel G, Liebl M, Gabor B, Mynott R (2001) Chem Eur J 7:4811
213. Schmidt RR, Toepfer A (1991) Tetrahedron Lett 32:3353
214. Tsunakawa M, Komiyama N, Tenmyo O, Tomita K, Kawano K, Kotake C, Konishi M, Oki T (1992) J Antibiot 45:1467
215. Tsunakawa M, Kotake C, Yamasaki T, Moriyama T, Konishi M, Oki T (1992) J Antibiot 45:1472
216. Tomita K, Hoshino Y, Miyaki T (1993) Int J Syst Bacteriol 43:297
217. Fürstner A, Mlynarski J, Albert M (2002) J Am Chem Soc 124:10274
218. Fürstner A, Albert M, Mlynarski J, Matheu M, DeClercq E (2003) J Am Chem Soc 125:13132
219. Poss CS, Schreiber SL (1994) Acc Chem Res 27:9
220. Brooks DW, Lu LD, Masamune S (1979) Angew Chem Int Ed Engl 18:72
221. Page PCB, Moore JPG, Mansfield I, McKenzie MJ, Bowler WB, Gallagher JA (2001) Tetrahedron 57:1837
222. Noyori R (1994) Asymmetric Catalysis in Organic Synthesis. Wiley, New York
223. Ohkuma T, Noyori R (1999) Comprehensive Asymmetric Catalysis. In: Jacobsen EN, Pfaltz A, Yamamoto H (eds) Springer, Berlin Heidelberg New York, Vol. 1, p 199
224. Fürstner A, Konetzki I (1998) Tetrahedron Lett 39:5721
225. Sharpless KB, Young MW (1975) J Org Chem 40:947
226. Grieco PA, Gilman S, Nishizawa M (1976) J Org Chem 41:1485
227. Blakemore PR, Cole WJ, Kocienski PJ, Morley A (1998) Synlett 26

7.3 Gangliosides in the Nervous System: Biosynthesis and Degradation

*Robert K. Yu**, *Toshio Ariga*, *Makoto Yanagisawa*, *Guichao Zeng*
Institute of Molecular Medicine and Genetics, Institute of Neuroscience, Medical College of Georgia, Augusta, GA 30912–2697, USA
ryu@mcg.edu

Abstract

Gangliosides, abundant in the nervous system, are known to play crucial modulatory roles in cellular recognition, interaction, adhesion, and signal transduction, particularly during early developmental stages. The expression of gangliosides in the nervous system is developmentally regulated and is closely related to the differentiation state of the cell. Ganglioside biosynthesis occurs in intracellular organelles, from which gangliosides are transported to the plasma membrane. During brain development, the ganglioside composition of the nervous system undergoes remarkable changes and is strictly regulated by the activities of glycosyltransferases, which can occur at different levels of control, including glycosyltransferase gene transcription and posttranslational modification. Genes for glycosyltransferase involved in ganglioside biosynthesis have been cloned and classified into families of glycosyltransferases based on their amino acid sequence similarities. The donor and acceptor substrate specificities are determined by enzymatic analysis of the glycosyltransferase gene products. Cell-type specific regulation of these genes has also been studied. Gangliosides are degraded by lysosomal exoglycosidases. The action of these enzymes occurs frequently in cooperation with activator proteins. Several human diseases are caused by defects of degradative enzymes, resulting in massive accumulation of certain glycolipids, including gangliosides in the lysosomal compartment and other organelles in the brain and visceral organs. Some of the rep-

In: *Glycoscience*. Fraser-Reid B, Tatsuta K, Thiem J (eds)
Chapter-DOI 10-1007/978-3-540-30429-6_41:

resentative lysosomal storage diseases (LSDs) caused by the accumulation of lipids in late endosomes and lysosomes will be discussed.

Keywords

Glycosphingolipid; Ganglioside; Biosynthesis; Catabolism; Nervous system; Endoplasmic reticulum; Golgi; Glycosyltransferase; Glycosidase; Lipid-storage disorder

Abbreviations

GSL	glycosphingolipid
GEM	GSL-enriched microdomain
ER	endoplasmic reticulum
S1P	sphingosine-1-phosphate
GT	glycosyltransferase
ST	sialyltransferase
Gal-T	galactosyltransferase
GalNAc-T	*N*-acetylgalactosaminyltransferase
UDP-	uridine-diphospho-
SAP	sphingolipid activator protein
Hex	hexosaminidase
LSD	lysosomal storage disease
CNS	central peripheral nervous system
PNS	peripheral nervous system

Nomenclature for gangliosides and related glycosphingolipids based on that of Svennerholm [1] in accordance with the IUPAC recommendations [2].

1 Introduction

Gangliosides, sialic acid-containing glycosphingolipids (GSLs), are diverse and highly complex molecules located primarily on plasma membranes and are particularly abundant in the central and peripheral nervous systems (CNS and PNS). Structural diversity of gangliosides is a hallmark of these molecules. About 200 gangliosides differing in carbohydrate components are known today [3,4]. The number of ganglioside molecular species increases manyfold if heterogeneity of the lipophilic components is considered. Gangliosides are known to have various biological functions depending on their ceramide and carbohydrate structures. Many studies have indicated that they play a variety of functions in cellular growth and differentiation, modulation of signal transduction, and immune reactions, as well as in several aspects of cancer metastasis and malignant transformation [5,6,7]. Gangliosides are also involved in the pathogenesis of certain immune-mediated neurological diseases because of their highly antigenic carbohydrate moieties that facilitate the production of antibodies that react with individual GSLs [3,8,9,10,11].

In the plasma membrane, GSLs are believed to segregate into microdomains that are also enriched in glycosylphosphatidylinositol-anchored proteins, sphingomyelin, and cholesterol [12]. Because of their specific localization on the plasma membrane, especially in

specialized microdomains or lipid raft structures, they are also involved in intercellular signaling [13,14,15,16,17] and in trafficking and/or sorting [18,19]. Certain biological functions of GSLs are better understood based on clustered units of GSLs, termed lipid rafts or GSL-enriched microdomains (GEMs) [17]. Although the exact functions of these clusters have not yet been fully elucidated, some of the functional roles of GSLs will be discussed based on this concept.

Ganglioside biosynthesis is catalyzed by specific glycosyltransferases (GTs) that transfer specific sugar residues from activated donor substrates, usually a sugar nucleotide, to a proper lipid or an oligosaccharide receptor. GTs are typically grouped into different families based on the type of sugar they transfer (e. g., galactosyltransferases, sialyltransferases, and so on). Over the past decade, many GT genes, including those involved in the ganglioside biosynthetic pathway (❯ *Fig. 1* and ❯ *Fig. 2*), have been cloned and classified into GT families by amino-acid-sequence similarities [20,21,22] (available at http://cazy.org/fam/acc_GT.html and http://glycob.oupjournals.org). These GTs differ in their donor- and acceptor-substrate specificities and other parameters, such as their microenvironment, local pH, and subcellular local-

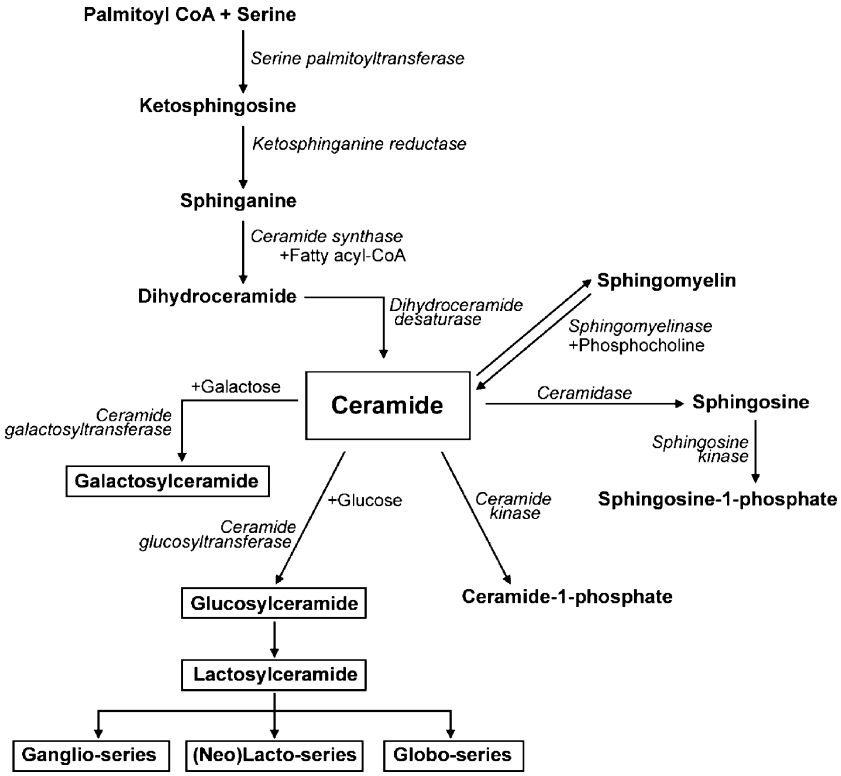


Figure 1
Synthesis of glycosphingolipids. *De novo* synthesis of GSLs occurs at the cytosolic membranes of endoplasmic reticulum

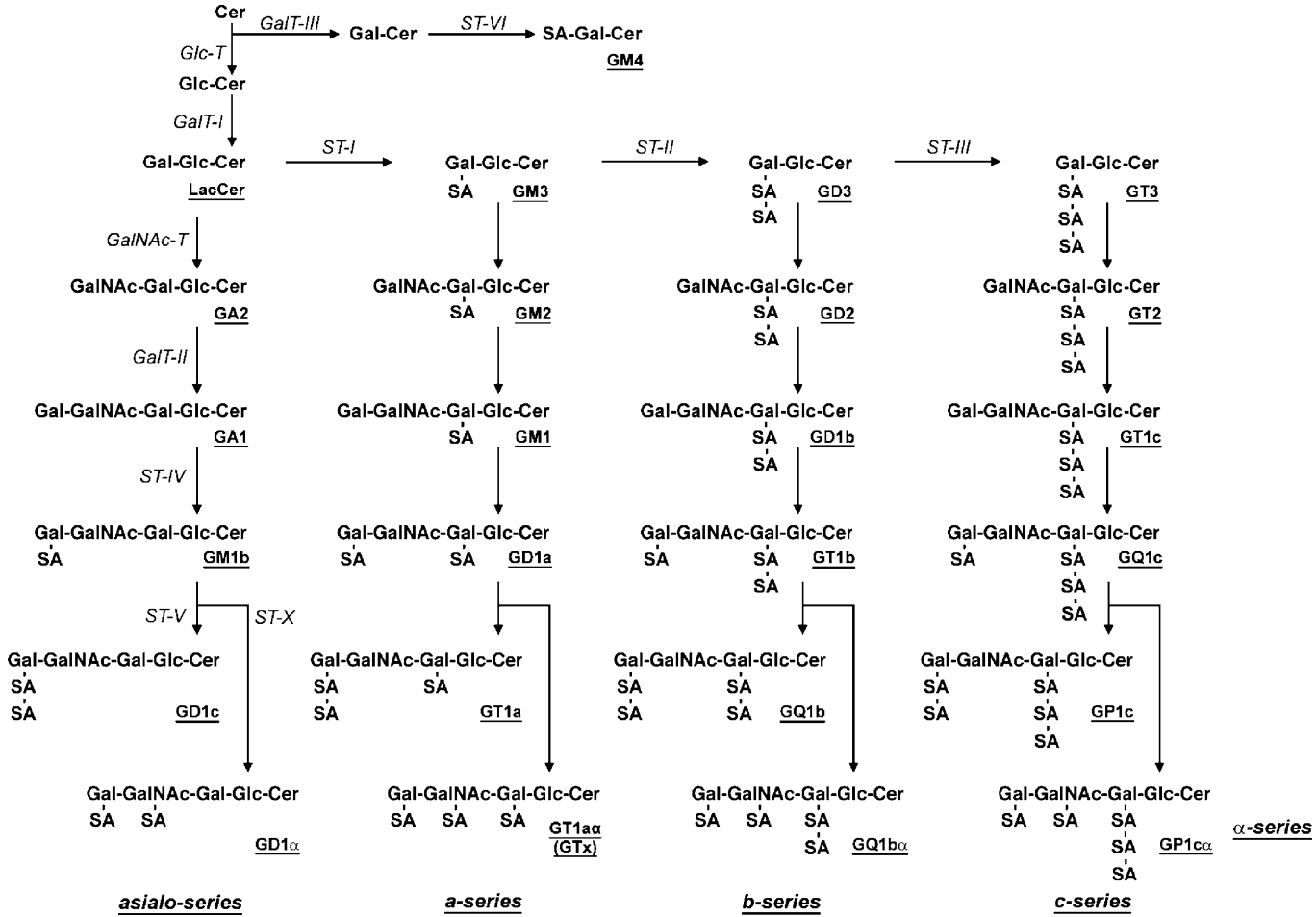

Cer
GalT-III
Gal-Cer
ST-VI
SA-Gal-Cer
GM4
Glc-T
Glc-Cer
GalT-I
Gal-Glc-Cer
LacCer
ST-I
Gal-Glc-Cer
SA
GM3
ST-II
Gal-Glc-Cer
SA
SA
GD3
ST-III
Gal-Glc-Cer
SA
SA
SA
GT3
GalNAc-T
GalNAc-Gal-Glc-Cer
GA2
GalNAc-Gal-Glc-Cer
SA
GM2
GalNAc-Gal-Glc-Cer
SA
SA
GD2
GalNAc-Gal-Glc-Cer
SA
SA
SA
GT2
GalT-II
Gal-GalNAc-Gal-Glc-Cer
GA1
Gal-GalNAc-Gal-Glc-Cer
SA
GM1
Gal-GalNAc-Gal-Glc-Cer
SA
SA
GD1b
Gal-GalNAc-Gal-Glc-Cer
SA
SA
SA
GT1c
ST-IV
Gal-GalNAc-Gal-Glc-Cer
SA
GM1b
Gal-GalNAc-Gal-Glc-Cer
SA
SA
GD1a
Gal-GalNAc-Gal-Glc-Cer
SA
SA
SA
GT1b
Gal-GalNAc-Gal-Glc-Cer
SA
SA
SA
SA
GQ1c
ST-V
ST-X
Gal-GalNAc-Gal-Glc-Cer
SA
SA
GD1c
Gal-GalNAc-Gal-Glc-Cer
SA
SA
SA
GT1a
Gal-GalNAc-Gal-Glc-Cer
SA
SA
SA
GQ1b
Gal-GalNAc-Gal-Glc-Cer
SA
SA
SA
SA
GP1c
Gal-GalNAc-Gal-Glc-Cer
SA
SA
GD1α
Gal-GalNAc-Gal-Glc-Cer
SA
SA
SA
GT1aα
(GTx)
Gal-GalNAc-Gal-Glc-Cer
SA
SA
SA
SA
GQ1bα
Gal-GalNAc-Gal-Glc-Cer
SA
SA
SA
SA
SA
GP1cα
α-series
asialo-series
a-series
b-series
c-series

Figure 2
Synthesis of gangliosides. Cer, ceramide; SA, sialic acid; GalNAc-T, *N*-acetylgalactosaminyltransferase I or GM2/GD2/GT2-synthase; GalT-I, galactosyltransferase I; GalT-II, galactosyltransferase II or GM1-synthase; GalT-III, galactosyltransferase III; GlcT, glucosyltransferase; ST-I, sialyltransferase I or GM3-synthase; ST-II, sialyltransferase II or GD3-synthase; ST-III, sialyltransferase III or GT3-synthase; ST-IV, sialyltransferase IV or GD1a-synthase; ST-V, sialyltransferase V or GT1a-synthase; ST-VI, sialyltransferase VI or GM4-synthase; ST-X, sialyltransferase x of α 2-6 sialosyl linkage to *N*-acetylgalactosaminyltransferase ("α-" series gangliosides)

ization [7,20,22,23]. This sequence-based classification is assumed to integrate both structural and mechanistic features within each family [20,21,22]. In vitro enzymatic analysis, however, indicates that multiple GTs can synthesize occasionally one linkage, and that one GT may be responsible for synthesis of multiple products [23]. Tissue-specific expression of these genes has been demonstrated by hybridization using cDNA fragments. Following cloning of the ganglioside synthase genes, the promoters of the key GT genes in the ganglioside biosynthetic pathway have been analyzed [7,23]. All promoters of these GT genes are TATA-less and do not possess any known core promoter element, and little is known about how genes that lack any of the conventional core promoter elements are regulated. Although these genes are characterized as housekeeping genes, recent studies on the promoters of these GT genes indicate that their expression is subject to complex developmental and cell type-specific regulation.
Here, we will focus on the current topics of biosynthesis of GSLs, including gangliosides in the ER and Golgi apparatus. We will review the cloned GT genes involved in ganglioside biosynthesis and their transcriptional regulation. We will also describe degradation or catabolism of gangliosides and, in particular, aspects of ganglioside metabolism in lipid-storage diseases that may be involved in GSL-trafficking via detergent insoluble GEMs, lipid rafts, or caveolae.

2 Biosynthesis of Glycosphingolipids

Glycosphingolipids (GSLs) are primarily synthesized in the endoplasmic reticulum (ER) and Golgi apparatus [24,25,26] and are internalized by endocytosis [27]. The known biosynthetic pathways of GSLs and gangliosides are depicted in *Fig. 1* and *Fig. 2*, respectively.
The synthesis of GSLs commences with the synthesis of ceramides in ER membranes. At least three enzymatic reactions are known to take place in this compartment: serine palmitoyltransferase, which catalyzes the condensation of palmitoyl coenzyme A and serine for the formation of 3-ketosphinganine; (dihydro)ceramide synthase, which converts sphingoid bases as well as fatty acid coenzyme A to produce dihydroceramide; and dihydroceramide desaturase, which coverts dihydroceramides to ceramides [24,25,28]. Ceramide is the common precursor of all GSLs and sphingomyelin. Ceramide, sphingosine, and a metabolite of sphingosine, sphingosine-1-phosphate (S1P), not only serve as important building blocks of complex sphingolipids but also have emerged as a new class of lipid biomodulators of various cellular functions. These metabolites are known to function as intracellular second messengers and as ligands in extracellular space. S1P especially has numerous functions as an important extracellular mediator that binds to cell-surface S1P receptors. Recent studies have also shown that sphingolipid-metabolizing enzymes function in intracellular organelles and in extracellular space, including the outer leaflet of the plasma membrane [29].

To continue the biosynthesis of complex sphingolipids, ceramide is transferred to the Golgi apparatus with the aid of a transfer protein, CERT [30]. The transferred ceramide is converted to glucosylceramide (GlcCer), which is catalyzed by GlcCer synthase, which transfers a glucose residue from uridine-diphospho (UDP)-glucose to the 1-position of ceramide. The reaction occurs at the cytosolic leaflet of the Golgi apparatus [31]. GlcCer is then translocated to the luminal face of the Golgi membranes, presumably by multidrug transporters [32,33]. Ceramide can also be galactosylated by the action of ceramide galactosyltransferase to form galactosylceramide (GalCer), which also serves as the precursor of sulfatides and GM4 ganglioside [34]. GalCer, sulfatides, and GM4, once formed, are then transferred to the luminal Golgi [27]. Lactosylceramide (LacCer), the common precursor for many GSLs, such as lacto(neo)-series, globo-series (see *Fig. 1*), and ganglio-series GSLs, is formed by the addition of a galactose moiety from UDP-galactose to GlcCer. The reaction is catalyzed by galactosyltransferase-I that has been purified and cloned from rat brain [28,35].

From LacCer, two main families of complex GSLs are derived: the neutral and acidic GSLs, including gangliosides (*Fig. 2*). These GSLs are synthesized by stepwise addition of monosaccharide residues onto LacCer. Subsequent glycosylation reactions are catalyzed by glycosyltransferases in the Golgi apparatus [7,36]. In the case of gangliosides, the limited specificity of some of the transferases [7,37] gives rise to complex GSL patterns on the cell surface within a combinatorial biosynthetic pathway [24]. The biosynthesis of these GSLs is strictly regulated by the activities of glycosyltransferases [7].

With the exception of GM4, a major ganglioside of the myelin membrane [38], gangliosides are all derived from LacCer. The addition of the first sialic acid residue converts LacCer to GM3, the precursor of most complex brain gangliosides. The reaction is catalyzed by either sialyltransferase I (ST-I) or GM3-synthase. Further addition of sialic acid residues generates GD3 (catalyzed by ST-II or GD3-synthase) and GT3 (catalyzed by ST-III or GT3-synthase). ST-I, ST-II, and ST-III are all distinct enzyme entities arising from different genes. GM3, GD3, and GT3 represent the entry substrates for biosynthesis of complex-type gangliosides in the a-, b-, and c-series pathways, respectively (*Fig. 2*). These enzymes include *N*-acetylgalactosaminyltransferase I (GalNAcT-I) or GM2/GD2-synthase, galactosyltransferase II (GalT-II) or GM1/GD1b-synthase, ST-IV or GD1a/GT1b-synthase, and ST-V or GT1a/GQ1b-synthase. The development of a parallel c-pathway was achieved by Yu and Ando [39] and Ando and Yu [40] after structural characterization of the c-series gangliosides, including GT3, GT2, GT1c, and GQ1c.

In order to produce asialo-series gangliosides, LacCer is converted to GA1 by two enzymatic reactions catalyzed by GalNAcT-I and GalT-II. ST-IV and ST-V are responsible for sialylation to GM1b from GA1 and to GD1c from GM1b, respectively. The complex gangliosides with sialic acid moieties α2,6-glycosidically linked to *N*-acetylgalactosamine residues, such as GD1α, GT1α, GQ1bα, and GP1cα, have been added later to the biosynthetic scheme [28,41,42,43]. In adult human tissues, gangliosides from the asialo- and c-series are found only in trace amounts [28] (*Fig. 2*). So far, it should be noted that (a) ST-I, ST-IV, and ST-V catalyze the formation of α2,3 sialyl linkage to galactose; (b) ST-II and ST-III are involved in the formation of α2,8 sialyl linkage to sialic acid; and (c) sialyl-transferase (ST-X, possibly ST-V) involving the formation of α2,6 sialyl linkage to *N*-acetylgalactosamine has not been characterized [27] although Irie et al. [44], using a rat liver Golgi preparation, reported GT1aα and GQ1bα are synthesized from GD1a and GD1b, respectively. Further synthesis

of the complex hexosamine-containing gangliosides is unique in that several identical GTs may participate in catalyzing the addition of various sugar residues to different ganglioside acceptor substrates [7].

Unlike lysosomal lipid-storage diseases (see below), very few metabolic diseases are known to be related to mutations of glycosyltransferases in ganglioside synthesis. The one new finding has been from that of Simpson et al. [45], who showed that a nonsense mutation of ST-I (GM3-synthase) is the cause of human autosomal recessive infantile-onset symptomatic epilepsy syndrome. Because ST-I is a key enzyme in the synthesis of all complex gangliosides, this report clearly indicates that the precise expression of complex gangliosides has critical biological functions in human nervous system development.

At this time, the precise intracellular localization of the enzymes that catalyze biosynthesis of gangliosides has not yet been achieved. It is assumed that the sequence of events in GSL biosynthesis resembles that of glycoprotein biosynthesis, i. e., early reactions occur in early subcellular organelles (endoplasmic reticulum and *cis*/medial Golgi) of the pathway, and late reactions occur in late stations (trans Golgi/trans Golgi network) [46]. GSLs are also known to recycle via the Golgi apparatus, especially in endosomes [47,48].

3 Glycosyltransferase Genes Involved in Ganglioside Biosynthesis and their Transcriptional Regulation

Glycosyltransferases (GTs; EC 2.4.x.y) constitute a large family of enzymes, which have been classified into families by amino acid sequence similarities [21]. By the end of 2006, the database at http://cazy.org/fam/acc_GT.html comprises more than 12,000 known and putative GT sequences that have been divided into 87 families. Here, we review the cloned GTs that are involved in the ganglioside biosynthetic pathway (*Fig. 2*), with respect to their classification, donor, and acceptor substrate specificities, as well as the linkages these enzymes catalyze. We also discuss the regulatory mechanisms of these genes, emphasizing their cell type-specific regulation.

3.1 Genes of GTs for Ganglioside Synthesis

GTs involved in ganglioside biosynthetic pathways (*Fig. 2*) include at least six sialyltransferases (STs, EC 2.4.99.y), three galactosyltransferases (GalT, EC 2.4.1.x), one galactosaminyltransferase (GalNAcT, EC 2.4.1.92), and one glucosyltransferase (GlcT, EC 2.4.1.80). All these GTs have been cloned except STIII, which remains unknown (see below). All cloned STs share four peptide-conserved motifs: L (large), S (small) [49,50], motif III [51], and VS (very small) [52,53] and are classified into three ST families (ST3Gal, ST8Sia, and ST6GalNAc) based on the linkage in which the sialic acid is transferred and the acceptor saccharide-specificity. The ST3Gal family consists of ST3Gal I, II, III, IV, V, and VI, which fall into two main branches (subfamilies) by phylogenetic analysis: ST3Gal I/II and ST3Gal III/IV/V/VI [54]. The enzymes in the ST3Gal family transfer Neu5Ac residues in the α2,3-linkage to terminal galactose residues found in glycoproteins and GSLs, but only ST3Gal I and V can use a GSL as an acceptor substrate, and thus are defined as ganglioside synthases. The cDNA for ST3Gal I (EC 2.4.99.4) was cloned from mouse [55] and

human [56]. The transiently expressed product of the ST3Gal I cDNA in host cells catalyzes the formation of GM1b, GD1a, and GT1b [55], indicating that this enzyme is an STIV (GM1b/GD1a/GT1b-synthase). The cDNA for human, mouse, and rat ST3Gal V (EC 2.4.99.9) has been cloned [53,57,58]. Assays for the expressed enzyme activity showed that ST3Gal V exclusively uses lactosylceramide (Galβ1-4Glcβ1-Cer) as an acceptor substrate to catalyze the synthesis of the ganglioside GM3, and thus is an STI (GM3-synthase) [57,58]. It should be noted that ST3Gal V (GM3-synthase) is the only ganglioside synthase to use Galβ1-4Glcβ1-1′Cer (LacCer) to synthesize GM3.

Two other ganglioside synthases, STII and STV, are classified into the ST8Sia I and III of the ST8ia family. This family consists of six members (ST8Sia I, II, III, IV, V and VI) in two main branches in the phylogenetic tree (ST8Sia I/V/VI and ST8Sia II/III/IV) [54]. The ST8Sia I gene has been cloned from human [59,60,61], mouse [62], and rat [63] and show high sequence similarity among the species. Analysis of these cDNA products showed that the cloned ST8Sia I catalyzes the formation of GD3 and thus is an STII (GD3 synthase). However, it was reported that this enzyme could utilize gangliosides GM1b, GD1a, and GT1b as exogenous substrates, in addition to GM3, indicating that this enzyme catalyzes the formation not only of GD3 but also of GD1c, GT1a, and GQ1b in vitro [64]. In addition, a cloned human ST8Sia I can also efficiently synthesize GT3 in vitro using GD3 as an acceptor [65]. The latter results are likely due to an artifact as our study showed that over-expression of exogenous GD3-synthase cDNA in cells resulted in both GD3 and GT3 [66] and that many cell lines expressing high levels of GD3 do not necessarily synthesize GT3. ST8Sia V of this family also catalyzes the synthesis of several gangliosides. The cDNAs for mouse [67] and human [68] ST8Sia V have been cloned. Enzymatic assays show ST8Sia V to be involved in the synthesis of gangliosides GD1c, GT1a, GQ1b [69] and that ST8SiaV is considered an STV. However, the ST8Sia I (GD3-synthase) may also synthesize GD1c, GT1a, and GQ1b [67]. Therefore, which (or both) of the ST8Sia I and V enzymes is responsible for the synthesis of GD1c/GT1a/GQ1b remains unclear. In vitro assays also indicate that GT3 is an additional product of ST8Sia V [67] but this result also may be an in vitro artifact as discussed in the case of ST8Sia I. Besides, the cDNA for mouse [69], rat [70], and human [71]. ST8Sia III has been cloned, and the common product was initially suggested to be GT3. However, further analysis of the expressed products of the ST8Sia III cDNA showed extremely high catalytic activity in transferring sialic acid residue through an α2,8-linkage to glycoproteins, and this enzyme is much more specific to *N*-linked oligosaccharides of glycoproteins than of GSLs. Therefore, whether a unique cDNA coding only for STIII (GT3-synthase) exists is still an open question.

ST6GalNAc catalyzes the transfer of Neu5Ac residues in an α2,6 linkage to GalNAc residues found in *O*-glycosylproteins or GSLs. A number of the ST6GalNAc genes have been cloned [72,73,74,75,76,77,78] and classified into ST6GalNAc I, II, III, IV, V, and VI. These six members fall by phylogenetic analysis into two main subfamilies, ST6GalNAc I/II and ST6GalNAc III/IV/V/VI [54]. Interestingly, these two subfamilies are in agreement with their substrate specificities: ST6GalNAc I and II of the first subfamily exhibit broad substrate specificity and catalyze the transfer of Neu5Ac onto Galβ1-3GalNAc peptides [75,76,77], whereas ST6GalNAc III/IV/V/VI of the second subfamily exhibits more restricted substrate specificity, only utilizing sialylated acceptor substrates (Neu5Acα2-3Galβ1–3GalNAc-R) found either in glycoproteins or GSLs [72,73,74,78]. As far as can be determined by in vitro assays of the expressed recombinant ST6GalNAc, only two members of the second

subfamily, ST6GalNAc III [72,73] and V [74], can act on the synthesis of GSLs GM1b and GD1α. Therefore, at least four expressed recombinant sialyltransferases, ST8Sia V, ST3Gal I, ST6GalNAc III, and ST6GalNAc V, catalyze the formation of GM1b in vitro. It is unclear, however, whether GM1b is actually synthesized from GlcCer by the four GTs (aside-series; ❯ *Fig. 2*) under physiological conditions.

Three galactosyltransferases (GalTI, II, and III) in the ganglioside biosynthetic pathways (❯ *Fig. 2*) have been cloned, and all exhibit a conserved putative *N*-linked glycosylation site. GalTI (UDP-galactose:glucosylceramide β1,4-galactosyltransferase) catalyzes the synthesis of LacCer by transferring galactose from UDP-Gal to GlcCer, generating a β1,4-glycosidic linkage. The cDNA for the enzyme catalyzing this reaction has been cloned from rat [79] and human [80]. The deduced amino-acid sequence of the rat GalTI shows 39% homology with mouse β1,4-galactosyltransferase (EC 2.4.1.38), which catalyzes the transfer of Gal to β1,4-GlcNAc in glycoproteins. The GalTI prefers GlcCer as a substrate for GSL synthesis, however, indicating that this cloned GalTI is different from the glycopeptide β1,4-Gal T (EC 2.4.1.38) [79]. GalTII (UDP-galactose: GD2 β1,3-galactosyltransferase) catalyzes the synthesis of GD1b, GM1, or GA1 by transferring a galactose moiety from UDP-Gal to β-GalNAc in the receptor substrates, generating a β1,3-glycosidic linkage. The cDNA for GalTII has been cloned from rat [81], human [82], and mouse brains [83]. The GalTII gene has no significant homology with other galactosyltransferase genes and consists of a single exon, which is different from other GT genes. In addition, GalTII, as well as GalTIII, contains a very hydrophobic putative stem region differing significantly from other animal GTs [82]. GalTIII (ceramide UDP-galactosyltransferase, EC 2.4.1.45) catalyzes the synthesis of GalCer, the precursor of ganglioside GM4, in embryonic chicken brain and in mouse brain. The GalTIII cDNA was first cloned from rat [84], and then from mouse [85,86] and then from human [87,88]. Analysis of the deduced amino-acid sequences showed common structural features of GalTIII in three species, such as the putative *N*-glycosylation sites, the *C*-terminal hydrophobic part of the protein, the terminal signal peptide, and the KKVK endoplasmic reticulum retention signal [85].

GalNAcT (β1,4 *N*-acetylgalactosaminyltransferase) catalyzes the transfer of *N*-acetylgalactosamine onto the receptor substrates to form a β1,4 linkage. The cDNA for GalNAcT has been cloned from human [89], rat [90], and mouse [91,92]. The deduced amino-acid sequences of the cDNA indicate that GalNAcT is a typical type II membrane protein, similar to other GTs, and consists of a short *N*-terminal residue, a transmembrane region, and a long *C*-terminal residue, including the catalytic domain. GalNAcT shows no significant homology with other GalNAc transferases but shows 39% identity of the amino-acid sequence with another β1,4-*N*-acetylgalactosaminyltransferase, which is responsible for biosynthesis of the blood group Sda/Cad antigen [93,94]. In vitro assays of the expressed recombinant GalNAcT showed that it catalyzes the formation not only of GM2 and GD2 from GM3 and GD3, respectively, but also of GA2 from LacCer, and thus GalNAcT is designated as a GA2/GM2/GD2/GT2-synthase. This enzyme lacks activity to transfer GalNAc to glycoproteins; another purified GalNAcT, however, can transfer GalNAc to both GSL and oligosaccharides [95]. It is unclear whether these two GalNAcTs are the same or different enzyme entities.

GlcT (UDP-glucose:*N*-acylsphingosine D-glucosyltransferase) transfers glucose from UDP-Glc to ceramide in brain tissues, which is the first step in biosynthesis of the major gangliosides except GM4 (❯ *Fig. 1*). The cDNA of GlcT has been cloned from human [96], mouse [97], and

rat [98]. The amino-acid sequence shows that GlcT has a type III structure: a short *N*-terminal segment to the exoplasmic side, a single transmembrane segment, and a long cytosolic tail [99]. Moreover, the catalytic domain of GlcT is in the cytosolic face of the Golgi membrane [96,99]. These two features of GlcT are different from most of the features of the cloned GTs. This enzyme is involved only in GSL synthesis, as suggested by the fact that no β-glucosyl residues have been found in either glycoproteins or the lipid precursors.

In summary, STI and GlcT are highly specific in both donor and receptor substrates, which exclusively catalyze the reactions for synthesis of the gangliosides GM3 and GlcCer, respectively. The other GTs discussed above show broad substrate specificities and are responsible for the synthesis of more than one ganglioside or even both GSLs and glycoproteins. This sequence-based classification is assumed to integrate both structural and mechanistic features within each family [21]. However, prediction of the donor and acceptor specificity of a GT based on sequence homology can be problematic because there are many examples of closely related sequences, such as the sequences of blood group A and B transferases with only four different amino acids [100], having different catalytic activity. Determination of the X-ray crystal structures of these transferases will provide a structural basis accounting for the specificities as well as for the catalytic reactions. More than 100 crystal structures have been described for proteins corresponding to 23 different GTs from prokaryotes and eukaryotes [101]. The 3-dimensional GT database, containing structural information for 17 distinct GT families, is available at http://www.cermav.cnrs.fr/glyco3d. Unfortunately, due to difficulties with expression level, purification, and crystallization, no crystalline structure for ganglioside synthases have so far been determined.

3.2 Transcriptional Regulation of the Genes of GTs for Ganglioside Synthesis

Cells can switch between expressing simple and complex gangliosides or among different series of gangliosides during brain development and cell differentiation, indicating that expression of the GTs for ganglioside synthesis is under strict regulation. Interestingly, all of the promoters of ganglioside synthase genes analyzed so far are TATA-less, lacking a CCAAT box but containing G*C*-rich boxes, characteristic of the housekeeping genes [7]. The term "housekeeping" means that there is little or no strict regulation. This lack of regulation may no longer be considered accurate because prevalence of the TATA box has in the past been overestimated in the eukaryotic genome; the majority ($\sim$76%) of eukaryotic genes, in fact, are TATA-less [102]. Very little is known about TATA-less promoter activation. The GT genes for ganglioside synthases are TATA-less genes under strict regulation in a cell type-specific manner.

The promoter of ST8Sia I (GD3-synthase, STII), which has attracted much more attention than have other promoters of the ganglioside synthase genes, has been cloned from rat [103], mouse [104] and human [105]. The promoter of all three species is TATA-less and does not possess known core- promoter elements, including the initiator (INR), the downstream promoter (DPE), the TFIIB recognition elements (BRE), and the "motif ten element" (MTE). Similar to other TATA-less promoters, it has a high GC content and is enriched in Sp1-binding sites. The proximal promoter region is defined within 500 or 700 bp upstream of the ATG

codon in rat and mouse promoters or human promoter, respectively [103,104,105]. The rat and mouse promoters have a single transcription start site [103,104], and the human promoter possesses multiple initiation sites [105]. Although there are no common binding sequences among the species, all the promoters have several Sp1-binding sites in the proximal promoter region, and deletion of the Sp1 sites results in a dramatic loss of the proximal promoter activities of the mouse [104] and rat gene (Zeng and Yu, unpublished data), suggesting that Sp1 plays a significant role in the transcriptional regulation of the gene. Our recent data from the DNA pull-down experiments show that at least seven transcription factors, Sp1, AP1, AP2, GATA, Ets-1, NFκB, and MRE, bound to their consensus sites on the proximal promoter fragment of the rat ST8Sia I gene (Zeng and Yu, unpublished data). In vivo binding of these factors to the promoter is cell-type specific as determined by ChIP assays in the rat F-11 that express a high level of ST8Sia I and in the rat PC12 cells that express little ST8Sia I (Zeng and Yu, unpublished data). The results show that Sp1 and NFκB bind to the promoter in both cells, while AP2 or Ets-1 binds to the promoter only in F-11 or PC12 cells, respectively. No binding of AP1 and GATA has been found in vivo in either of the cells. Results from deletion of the corresponding binding sites have confirmed the involvement of these transcription factors in regulation of the proximal ST8Sia I promoter in a cell-type specific manner (Zeng and Yu, unpublished data). We have also identified an unknown mutation close to the 3′-end of the rat promoter that decreases proximal promoter activity in F-11 cells and increases it in PC12 cells (Zeng and Yu, unpublished data). ST8Sia I is highly expressed in melanoma and in many types of brain tumors, although ST8Sia I is almost undetectable in melanocytes and adult brain; thus it is clinically significant to understand the mechanism regulating ST8Sia I gene expression in a cell-type-specific manner. Other features of this promoter include a negative control region and a GT/CG repeat sequence in the distal regulatory region [103,104,105]. These unique GT/CG repeats present a structure of Z-DNA and are located in the negative control regions. Whether the GT/CG repeat sequence contributes to the suppressive activity in the transcriptional regulation of the ST8Sia I gene remains to be elucidated.

ST3Gal V (GM3-synthase, STI) catalyzes the synthesis of GM3, a substrate for synthesis of all gangliosides and possibly also of GM4 (❯ *Fig. 2*). Regulation of ST3Gal V expression is critical for initiation of ganglioside synthesis. There are four isoforms of ST3Gal V mRNA in human fetal brain, differing in the 5′-untranslated region, but only one of these isoforms can be detected in adult brain [106], suggesting that expression of the ST3Gal V gene is developmentally regulated, and that alterative splicing or transcriptional initiation of the ST3Gal V RNA occurs in a tissue-specific manner [58,107]. Analysis of the promoters of the human [108,109] and mouse [110] ST3Gal V genes found that the mouse ST3Gal V gene uses multiple transcriptional initiation sites, including several downstream sites [110], and the human ST3Gal V gene uses at least two initiation sites [111], consistent with the fact that there are multiple isoforms of transcripts from the ST3Gal V genes [106]. Minimum promoter regions of 177 bp and 254 bp lie upstream of the initiation site of the human and mouse promoters, respectively, and display the highest level of promoter activity [108,109,110], although longer promoter fragments from the proximal regions to 1,600 bp do not significantly contribute to transcriptional activity [108,109,110]. The proximal region contains no TATA or CCAAT boxes but has multiple Sp1 and AP2 sites. It has recently been reported that a juxtaposed Sp1/AP2-binding site and one of the six AP2-binding sites in the proximal region are essential in maintaining a high lev-

el of promoter activity [110]. Chromatin immunoprecipitation confirmed binding of the transcription factors Sp1 and AP2 to the two sites inside cells [110]. DNA pull-down experiments revealed that 7 factors (CREB, PPAR, Pax-5, Smad3/4, Stat6, MEF-1, and MEF-2) bound to both human and mouse promoter fragments [109,110]. Among them, the CREB-binding site has been further demonstrated to be critical in transcription of the human ST3Gal V gene in HL-60 cells when activated by PMA (phorbol 12-myristate 13-acetate) [111]. These results indicate that multiple factors, including Sp1, AP2, and CREB, contribute to activation of ST3Gal V gene transcription and that expression of human and mouse ST3Gal V genes may be under a similar regulatory mechanism. Future studies may focus on how these factors and regulatory elements function coordinately in vivo in development-dependent or cell-type specific expression of the gene.

GalNAcT is a key enzyme controlling the expression of more complex gangliosides (*Fig. 2*), and the GalNAcT gene is highly expressed in tumor cell lines such as in neuroblastoma and melanoma [112,113], and also in mouse brain at the late stage of development [62,90]. Studies on the promoter and genomic organization of the human GalNAcT gene show that it has three transcription start sites and three alternative exons (exon 1a, exon 1b, and exon 1c). Three promoter regions (P1, P2, and P3) in GalNAcT have been defined for individual transcription start sites, respectively and one of them is located in the first intron [114]. There are consensus binding sites for transcription factors EGR-1, HNF-5, Sp1, and PEA3 in the 5′-flanking region of exon 1a, three binding sites for Sp1 and one for AP2 and PEA3 in the 5′-flanking region of exon 1b, and one site for AP2 and E2F and S1 HS sites in the 5′-flanking region of exon 1c. The activities of promoters P1 and P2 are strongly enhanced by a sequence (enhancer) located in exon 1. Furukawa et al. [114] also proposed a suppressive sequence (silencer) of 42 bp located upstream of exon 1c, which suppressed the transcriptional activity of the P3 promoter in a melanoma cell line in which no promoter activity had been detected. These results indicate that alternative promoter regions or the enhancer/silencer may regulate cell type-specific expression of the GalNAcT gene.

The three GTs above are the key enzymes that regulate synthesis of the more complex gangliosides: “a”, “b”, and “c” series. The transcriptional regulation of these GT genes is itself complex and tissue-/cell-specific. Other promoters for ganglioside synthase genes, including promoters of the GlcT, GalTII, and GalTIII genes, have also been cloned and analyzed [7]. They are all TATA-less and contain no CCAAT box, but they do have G*C*-rich boxes. The proximal promoter (578 bp) of the mouse GlcT gene contains four Sp1-binding sites located between −558 and −439 bp upstream of the ATG codon [115]. Deletion of a 41-bp fragment (from −538 to −578) dramatically decreases the activity of the proximal promoter. This 41-bp fragment contains one Sp1-binding site, indicating an essential role of this site in maintaining a high level of proximal promoter activity [115]. Since GlcT catalyzes the formation of glucosylceramide, which serves as a core structure for more than 300 species of GSLs, it is unclear whether regulation of the GlcT gene pertains to ganglioside biosynthesis.

Analysis of the promoter of the mouse GalTII gene reveals several features different from other ganglioside synthase promoters [116]. The proximal 550-bp fragment that shows the highest level of promoter activity contains no Sp1-binding site, but does show a number of other consensus sequences. The GalTII gene has a single exon and a single transcription start site. A 350-bp fragment beyond the proximal region strongly suppresses promoter activity. As many as 27 transcription factors have been characterized as binding to their consensus

sites in the promoter region in vitro [116]. Those data characterizing the transcription factors may indicate a quite different mechanism for regulation of GalTII gene expression. GalTII is responsible for synthesis of GA1, GM1, GD1b, and GT1c. GM1 is known to play important roles in the development and functions of the neural system [117]. How GalTII gene expression is regulated, particularly during neuronal differentiation, remains to be elucidated.

The promoter of the human GalTIII gene has been cloned and analyzed [118,119]. The gene is transcribed from a single initiation site 322 bp upstream of the ATG codon. The *cis*-acting regulatory elements were found in two small regions at −292/−256 and −747/−688. The region at −292/−256 contains putative binding sites for transcription factors Ets and Sp1, and the region at −747/−688 exhibits binding sites for ERE, NF-1, TGGCA-BP, and CRE [118]. A third *cis*-acting region, distally located at −1325/−1083, consists of multiple binding sites for TCF-1, TGGCA-BP, NF-IL6, CF1, bHLH, NF-1, GATA, and Q-IRE. A negative regulatory domain localized in a far region at −1594/−1326 was also identified [118]. Further analysis of the two small regulatory regions demonstrated that the functional elements are a Sp1-binding site at −292/−256 and a CRE-binding site at −747/−688 [119]. The cell type-specific activities of those promoters have been observed between the human oligodendroglioma (HOG) and human neuroblastoma (LAN-5) cell lines, and the factor CREB may account for the different levels of GalTIII promoter activity in those lines [119]. GalTIII is a key enzyme in the biosynthesis of Galcer, the most abundant GSL in the myelin sheath. Galcer serves as the precursor for the biosynthesis of GM4, the simplest ganglioside present in the myelin of the central nervous system [120]. Whether cell-type specific regulation of the GalTIII gene controls GM4 synthesis in cells is a subject for further investigation.

4 Ganglioside Catabolism and Gangliosidoses

4.1 Ganglioside Degradative Pathways

The constitutive degradation of GSLs occurs primarily in the acidic compartments, the endosomes and the lysosomes [121]. Here, water-soluble glycosidases sequentially cleave off the terminal carbohydrate residues from GSLs (*Fig. 3*). Ganglioside catabolism consists of the sequential removal of individual sugar residues, starting from the non-reducing terminal unit by (exo)glycohydrolases, which are soluble enzymes [27]. The catabolism of gangliosides requires the so-called sphingolipid activator proteins (SAPs) [122] or saposins [123], which are derived from a protein, SAP-precursor or prosaposin, by proteolytic processing in the ER [124,125].

Lysosomal sialidase is required for the degradation of complex gangliosides for the formation of GM1 ganglioside [126]. GM1 is degraded to GM2 by the action of a β-galactosidase in the presence of either the GM2-activator protein (AP) or SAP-B [127]. GM2 is degraded to GM3 by hexosaminidase (HEX) A. Hydrolysis of GM2 requires a normal GM2 ganglioside-GM2 activator-Hex A complex. Hex A consists of a dimer and has the structure α-β. The α subunit is encoded by the *HEXA* gene at band 15q23-q24 [122]. GM3 is degraded to LacCer and sialic acid by SAP-B [126]. The galactose residue in the GSLs is removed by either galactosylceramide-β-galactosidase or GM1-β-galatosidase to form glucosylceramide in the presence of either SAP-B or -C [28,128]. The membrane-bound lysosomal hydrolase, glu-

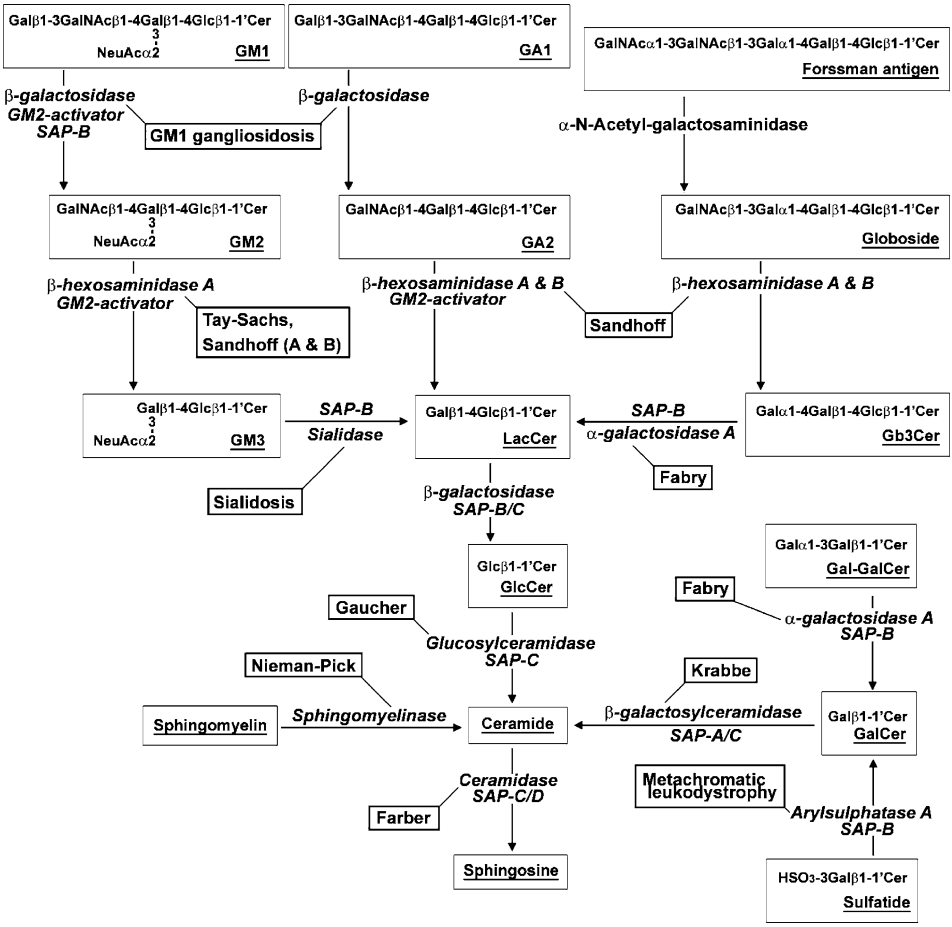


Figure 3

Degradation of GSLs in the lysosomes. Figure includes enzymes of individual lysosomal storage diseases (LSDs). Sphingolipid activator proteins (SAPs) are required in this process

cocerebrosidase (glucosylceramide-β-glucosidase), participates in the degradation of glucosylceramide to yield glucose and ceramide [129]. The end product, ceramide, is eventually degraded into a long-chain base residue and a fatty acid moiety by ceramidase [130]. The lysosomal hydrolysis of ceramide is catalyzed by acid ceramidase and requires sphingolipid activator proteins (SAP-D) as a co-factor in vivo [131].

4.2 Glycosphingolipid Lysosomal Storage Diseases

The importance of gangliosides in human health is evident in neurodegenerative diseases associated with defects in the degradation of gangliosides. It is well known that serious neurological deficits develop in humans with aberrant GSLs or ganglioside metabolism. Most

Table 1
Shingolipisosis

Disease	Defected enzyme	Accumulated substance
GM1 gangliosidosis	β-galactosidase	GM1, oligosaccharides & glycoproteins having terminal β-galactose
GM2 gangliosidosis		
B variant (Tay–Sachs disease)	β-hexosaminidase A	GM2
O variant (Sandhoff disease)	β-hexosaminidase A & B	GM2, GA2, oligosaccharides & glycoproteins having terminal β-*N*-acetylglucosamine
Fabry disease	α-galactosidase	Globotriaosylceramide
Metachromatic leukodystrophy	Arylsulfatase A	Sulfatide
Krabbe disease (Globoid-cell leukodystrophy)	Galactosylceramidase	Galactosylceramide & psycosine
Gaucher disease	Glucosylceramidase	Glucosylceramide
Nieman–Pick disease (types A & B)	Sphingomyelinase	Sphingomyelin
Farber disease	Ceramidase	Ceramide

of the human lysosomal-storage diseases are associated with GSL catabolism, the so-called degradation disorders or lipid-storage disorders (LSDs or lipidoses) [33]. LSDs are caused by deficiencies in lysosomal enzymes and abnormalities in lysosomal functions within the lysosomes [132]. LSDs are generally classified by accumulated substrates, which include sphingolipidoses, glycoproteinoses, mucolipidoses, mucopolysaccharidoses, and others. Glycosphingolipidoses are a group of inherited metabolic disorders in which harmful amounts of GSLs accumulate, resulting in an increase in the size and number of lysosomes and endosomes. This excessive storage of GSLs can cause permanent cellular and tissue damage, particularly in the brain, peripheral nervous system, liver, spleen and bone marrow (*Fig. 3* and *Table 1*). In addition, defects in co-factors, such as SAPs, that are involved not only in the lysosomal degradation of GSLs, but also in targeting of the transport systems in the degradation process can also lead to LSDs [33,121]. LSDs are clinically heterogeneous with respect to the age of onset, progression of symptoms, and the particular organs involved. Varying levels of residual enzyme activity, associated with different defective alleles that cause the respective diseases, are responsible in part for this clinical heterogeneity [133]. These events can possibly occur with the accumulation of lipids in lipid rafts in late endosomes/lysosomes [134,135]. In this section, we will focus on gangliosidoses that are caused by inherited defects within the ganglioside degradative pathway.

GM1 gangliosidosis is an autosomal recessive disorder caused by a deficiency of lysosomal β-galactosidase, resulting in accumulation of GM1 in the brain and keratosulfate-like mucopolysaccharide in the visceral organs. GM1 gangliosidosis is a progressive neurological disease with three clinical subtypes, infantile (type 1)-, juvenile (type 2)- and adult-onset variant (type 3) [136,137,138]. In type 1, motor and mental retardation are most rapidly progressive, which resembles features of some neurolipidosis (i. e., neurodegeneration, progressive organomegaly, dysostosis multiplex, facial coarsening, macular cherry-red spots) Patients with type 1 gangliosidosis are usually associated with mucopolysaccharidosis-like dysmorphic

features (e. g., visceromegaly, dysostosis multiplex, and coarsened facial features) that typically develop between birth and 6 months. Death usually occurs during the second year of life due to infection (usually pneumonia, resulting from recurrent aspiration) and cardiopulmonary failure [139]. Type 2 disease develops later (age 1–2) with onset of progressive psychomotor retardation. Some visceromegaly and milder skeletal disease develop compared to the infantile form. Death usually occurs before the second decade of life. Type 3 disease shows a late onset (in the teens) and typically presents during childhood or adolescence as a slowly progressive dementia with Parkinsonian features and extrapyramidal disease, particularly dystonia. Chronic forms of type 3 disease are variable and reported in fewer cases [140] than the other two types. Age at death may vary greatly, from 20 to 50. Several β-galactosidase-deficient animal models have been reported, i. e., in mice [141,142,143], cat [144,145], sheep [146], and dog [141,148,149]. These animal models appear to mimic closely the pathological, biochemical, and clinical abnormalities of human disease [141].

In type 1 disease, the deposition of GM1 in tissues is significantly increased, with less deposition in type 2. In type 3, GM1 deposits appear primarily in the basal ganglia. In some cases, no increase of GM1 is found in plasma or in cerebrospinal fluid; the metabolism of GM1 in the cultured skin fibroblasts of individuals with type 3 is almost normal although the residual activity of GM1 β-galactosidase activity is only 10% of normal [138]. The storage of substances in the peripheral tissues decreases from type 1 to type 3. Kobayashi et al. [150] have demonstrated the accumulation of lyso-GM1 in late-infantile type and adult type GM1 gangliosidosis. Lyso-asialo-GM1 is also expressed in late-infantile type, but not in adult type.

GM2-gangliosidosis is a family of three autosomal recessive, lysosomal storage disorders, characterized by the intralysosomal accumulation of GM2 ganglioside and related GSLs, primarily in the central and peripheral neural tissues [33,151,152]. GM2 gangliosidosis is caused by mutations in at least 1 of 3 recessive genes: *HEXA, HEXB*, and *GM2A* [153]. Recently, Lemieux et al. [153] have reported that the crystal structure of Hex A reveals an α-β heterodimer, with each subunit having a functionally active site. Only the α-subunit active site can hydrolyze GM2 gangliosides because of its flexible loop structure, removed post-translationally from β, and the presence of αAsn423 and αArg424. The loop structure is involved in binding the GM2 activator protein, while αArg424 is critical for binding the carboxylate group of the *N*-acetylneuraminic acid (NeuAc) residue of GM2. The β-subunit lacks these key residues and has βAsp452 and βLeu453 in its place; the β-subunit therefore cleaves efficiently only neutral substrates. Deficiencies of either the α- or the β-subunit of the heterodimeric β-hexosaminidase A (Hex A) protein or of the small monomeric GM2 activator protein (a substrate-specific co-factor for Hex A) lead to the three devastating phenotypic neurodegeneration-associated disorders: Tay–Sachs disease, Sandhoff disease, and the AB variant form [153].

Tay–Sachs disease is the B-variant of GM2 gangliosidosis due to α-chain deficiency and to the subsequent deficiency of hexosaminidases A and S, but with normal hexosaminidase B. Depending on the residual enzyme activity of β-hexosaminidase, the onset of symptoms may occur anywhere from late infancy to adulthood and are usually subclassified into infantile (type 1)-, juvenile (type 2)-, chronic-, and adult-onset forms [33]. In type 1, the most common disease with a carrier frequency of 1 in 27 among Ashkenazi Jews [154], patients are normal at birth but then show symptoms, such as mild motor weakness, between 3 and 6 months, resulting in hypotonia, poor head control, decreasing attentiveness, and visual symptoms (cherry red

spot). About 10 months later, infants with type 1 Tay–Sachs have severe motor, mental, and visual disabilities, usually leading to death between age 2 and 4. In type 2, the first onset of motor symptoms occurs between 2 and 6 years. Juveniles exhibit abnormalities such as loss of speech and vision, increasing spasticity, seizures. and progressive dementia. Death often occurs due to infections. In the chronic form of Tay–Sachs disease, onset occurs between 2 and 5 years, and patients can live to about age 40. In the adult form, the symptoms are heterogeneous. Individuals show neurological disorders, such as spinal muscular atrophy and psychoses, but their intelligence and visual functions are not affected.

Sandhoff disease is the 0-variant of GM2 gangliosidosis due to a β-chain deficiency, resulting in deficient activity of β-hexosaminidases A and B but with normal β-hexosaminidase S, which causes accumulation of GM2 and asialo-GM2 in the brain and globoside in the visceral organs [33]. Sandhoff disease is classified into three types: infantile (type 1)-, juvenile (type 2)-, and adult (type 3)-forms [33]. In type 1, organomegaly and slight bone deformations may occur in addition to the clinical and pathological manifestations of classical Tay–Sachs disease. In the type 2, symptoms such as slurred speech, cerebellar ataxia, and psychomotor retardation appear at age 3 to 10. In addition, juveniles with type 2 have increasing spasticity and gradual deterioration of mental function although visual function is normal. In type 3, the onset of symptoms is delayed to later in adult life. The clinical manifestations in this chronic variant are similar to those of Tay–Sachs disease variants of corresponding ages. The AB-variant of GM2-gangliosidosis is caused by deficiency of the GM2 activator protein, but β-hexosaminidase A, B, and S activities are normal, resulting in the accumulation of GSLs, e. g., GM2 and asialo-GM2 (GA2) [33]. In this AB variant, the clinical picture resembles that of classical Tay–Sachs disease with a delayed appearance of symptoms. In its histopathological aspects, Sandoff disease in these patients is characterized by swollen neurons with massive accumulations of stored lipids in membranous cytoplasmic bodies (MCBs) throughout the central and peripheral nervous systems. Signs and symptoms begin in infancy. Infants appear normal until age 3 to 6 months. Children with the AB variant can live only into early childhood. Several β-hexosaminidase-deficient animal models have been reported, i. e., mice [155,156,157,158], cat [159,160], sheep [159], and dog [159,161,162]. Lyso-GM2 was expressed in the brain of Tay–Sachs patients [163]. Kobayashi et al. [150] have also demonstrated the accumulation of lyso-GM2 and lyso-GA2 in the brain of individuals with Sandhoff disease and lyso-GM2 in the brain of those with Tay–Sachs disease. Li et al. [164] reported the presence of an unusual ganglioside, taurine-conjugated GM2, as minor components in the brain of individuals with Tay–Sachs disease. Itoh et al. [165] reported the accumulation of GalNAc-GM1b in the brain of those with Tay–Sachs disease. Further study revealed that GalNAc-GD1a was expressed in the brain of individuals with the same disease [166].

With respect to gangliosidoses, a unique form of gangliosidosis in emu (*Dromaius novaehollandiae*) has been reported [167,168]. Ultrastructural examination of the affected neurons reveals numerous MCBs similar in appearance to those seen in mammalian gangliosidoses. Neurological examination reveals hypermetric gait, persistent head tremor, and mild ataxia. Histopathologic and electron microscopic abnormalities are found, which include swollen, pale neurons were present in the cerebrum, pons, medulla, cerebellum, spinal cord, spinal ganglia, autonomic ganglia, myenteric plexus, and the ganglion cell layer of the retina. Analysis of brain gangliosides of the affected 7-month-old emu revealed 14- and 25-fold increases of GM1 and GM3 gangliosides, respectively, compared with control emus.

5 Concluding Remarks

Early research on gangliosides focused on understanding the genetic basis of ganglioside storage diseases, which led to major advances in the analysis of ganglioside structure, tissue and cellular/subcellular distributions. In the 1960s, extensive research was extended to the metabolism of these GSLs, which formed the basis of developing the various biosynthetic and catabolic pathways that have now become known. With the advent of contemporary molecular and cellular biology, remarkable success has been made in understanding the genetic and regulatory mechanisms of ganglioside metabolism during development. Current efforts have focused on the functional roles of gangliosides during development, and the concept of lipid rafts, where many of the membrane signaling events involving gangliosides are launched, has revolutionized our studies on their biological functions. Understanding the genetic basis of the various ganglioside storage disorders has also provided rational strategies for their diagnosis and treatment with enzyme replacement or gene therapy. Future studies will most likely continue to focus on areas related to their biological functions, both in normal development and in disease.

Acknowledgement

Studies from our own laboratories presented in this review have been supported by a grant from USPHS NIH (NS11853) to RKY.

References

1. Svennerholm L (1963) Chromatographic separation of human brain gangliosides. J Neurochem 10:613–623
2. The nomenclature of lipids (Recommendations 1976) IUPAC-IUB Commission on Biochemical Nomenclature (1978) Biochem J 171:21–35
3. Schwarz A, Futerman AH (1996) The localization of gangliosides in neurons of the central nervous system: the use of anti-ganglioside antibodies. Biochim Biophys Acta 1286:247–267
4. Yu RK, Yanagisawa M, Ariga T (2007) Glycosphingolipid structures. In: Kamerling J, Voragen AGJ, Lee Y, Boons G-J, Suzuki A and Taniguchi N (eds) Comprehensive Glycoscience (3). Elsevier, Amsterdam, pp 73–122
5. Ledeen RW (1989) Biosynthesis, metabolism and biological effects of gangliosides. In: Margolis RU and Margolis RK (eds) Neurobiology of glycoconjugates. Plenum Press, New York, pp 43–83
6. Yu RK, Saito M (1989) Structure and localization of gangliosides. In: Margolis RU and Margolis RK (eds) Neurobiology of glycoconjugates. Plenum Press, New York, pp 1–42
7. Yu RK, Bieberich E, Xia T, Zeng G (2004) Regulation of ganglioside biosynthesis in the nervous system. J Lipid Res 45:783–793
8. Yu RK, Ariga T (1998) The role of glycosphingolipids in neurological disorders. Mechanisms of immune action. Ann NY Acad Sci 845:285–306
9. Fredman P, Lekman A (1997) Glycosphingolipids as potential diagnostic markers and/or antigens in neurological disorders. Neurochem Res 22:1071–1083
10. Ariga T, Yu RK (2005) Antiglycolipid antibodies in Guillain-Barre syndrome and related diseases: review of clinical features and antibody specificities. J Neurosci Res 80:1–17
11. Yu RK, Usuki S, and Ariga T (2006) Ganglioside molecular mimicry and its pathological roles in Guillain-Barre syndrome and related diseases. Infect Immun 74:6517–6527
12. Rajendran L, Simons K (2005) Lipid rafts and membrane dynamics. J Cell Sci 118:1099–1102

13. Hakomori S (1990) Bifunctional role of glycosphingolipids. Modulators for transmembrane signaling and mediators for cellular interactions. J Biol Chem 265:18713–18716
14. Hakomori S, Igarashi Y (1995) Functional role of glycosphingolipids in cell recognition and signaling. J Biochem 188:1091–1103
15. Mutoh T, Tokuda A, Miyadai T, Hamaguchi M, Fujiki N (1995) Ganglioside GM1 binds to the Trk protein and regulates receptor function. Proc Natl Acad Sci USA 92:5087–5091
16. Yates AJ, Rampersaud A (1998) Sphingolipids as receptor modulators. An overview. Ann N Y Acad Sci. 845:57–71
17. Hakomori S-I (2000) Cell adhesion/recognition and signal transduction through glycosphingolipid microdomain. Glycoconj J 17: 143–151
18. van der Bijl P, Lopes-Cardozo M, van Meer G (1996) Sorting of newly synthesized galactosphingolipids to the two surface domains of epithelial cells. J Cell Biol 132:813–821
19. Hakomori S-I (2004) Carbohydrate-to-carbohydrate interaction, through glycosynapse, as a basis of cell recognition and membrane organization. Glycoconj J 21:125–137
20. Kapitonov D, Yu RK (1999) Conserved domains of glycosyltransferases. Glycobiology 9:961–978
21. Coutinho PM, Deleury E, Davies GJ, Henrissat B (2003) An evolving hierarchical family classification for glycosyltransferases. J Mol Biol 328:307–317
22. Harduin-Lepers A, Mollicone R, Delannoy P, Oriol R (2005) The animal sialyltransferases and sialyltransferase-related genes: a phylogenetic approach. Glycobiology 15:805–817
23. Zeng G, Yu RK (2007) Cloning and transcriptional regulation of genes responsible for synthesis of gangliosides. Curr Drug Targets (in press)
24. Kolter T, Proia RL, Sandhoff K (2002) Combinatorial ganglioside biosynthesis. J Biol Chem 277:25859–25862
25. Merrill AH Jr. (2002) De Novo sphingolipid biosynthesis: a necessary, but dangerous, pathway. J Biol Chem 277:25843–25846
26. Bieberich E, Silva J, Wang G, Krishnamurthy K, Condie BG (2004) Selective apoptosis of pluripotent mouse and human stem cells by novel ceramide analogues prevents teratoma formation and enriches for neural precursors in ES cell-derived neural transplants. J Cell Biol 167:723–734
27. Tettamanti G (2004) Ganglioside/glycosphingolipid turnover: new concepts. Glycoconj J 20:301–317
28. Sandhoff K, Kolter T (2003) Biosynthesis and degradation of mammalian glycosphingolipids. Philos Trans R Soc Lond B Biol Sci 358:847–861
29. Tani M, Ito M, Igarashi Y (2007) Ceramide/sphingosine/sphingosine 1-phosphate metabolism on the cell surface and in the extracellular spaces. Cell Signal 19:229–237
30. Hanada K, Kumagai K, Yasuda S, Miura Y, Kawano M, Fukasawa M, and Nishijima M (2003) Molecular machinery for non-vesicular trafficking of ceramide. Nature 426:803–809
31. Jeckel D, Karrenbauer A, Burger KN, van Meer G, Wieland F (1992) Glucosylceramide is synthesized at the cytosolic surface of various Golgi subfractions. J Cell Biol 117:259–267
32. Eckford PD, Sharom FJ (2005) The reconstituted P-glycoprotein multidrug transporter is a flippase for glucosylceramide and other simple glycosphingolipids. Biochem J 389:517–526
33. Kolter T, Sandhoff K (2006) Sphingolipid metabolism diseases. Biochim Biophys Acta 1758:2057–2079
34. Basu S, Schultz A, Basu M, Roseman S (1971) Enzymatic synthesis of galactocerebroside by a galactosyltransferase from embryonic chicken brain. J Biol Chem 246:4272–4279
35. Nomura T, Takizawa M, Aoki J, Arai H, Inoue K, Wakisaka E, Yoshizuka N, Imokawa G, Dohmae N, Takio K, Hattori M, Matsuo N (1998) Purification, cDNA cloning, and expression of UDP-Gal: glucosylceramide β1,4-galactosyltransferase from rat brain. J Biol Chem 273:13570–13577
36. Giraudo CG, and Maccioni HJ (2003) Ganglioside glycosyltransferases organize in distinct multienzyme complex in CHO-K1 cells. J Biol Chem 278:40262–40271
37. Pohlentz G, Klein D, Schwarzmann G, Schmitz D, Sandhoff K (1988) Both GA2, GM2 and GD2 syntheses and GM1b, GD1a and GT1b synthases are single enzymes in Golgi vesicles from rat liver. Proc Natl Aca Sci USA 85:7044–7048
38. Ledeen RW, Yu RK, Eng LF (1973) Gangliosides of human myelin: sialosylgalactosylceramide (G7) as a major component. J Neurochem 21:829–839
39. Yu RK, Ando S (1980) Structures of new complex gangliosides of fish brain. Adv Exp Med Biol 125:33–45

40. Ando S, Yu RK (1979) Isolation and characterization of two isomers of brain tetrasialogangliosides. J Biol Chem 254:12224–12229
41. Nakamura K, Inagaki F, Tamai Y (1988) A novel ganglioside in dogfish brain. Occurrence of a trisialoganglioside with a sialic acid linked to N-acetylgalactosamine. J Biol Chem 263:9896–9900
42. Nakamura K, Kojima H, Suzuki M, Suzuki A, Tamai Y (2000) Novel polysialogangliosides of skate brain structural determination of tetra, penta and hexasialogangliosides with a NeuAcGalNAc linkage. Eur J Biochem 267:5198–5208
43. Hidari K, Kawashima I, Tai T, Inagaki F, Nagai Y, Sanai Y (1994) In vivo synthesis of disialoganglioside (GD1α) from asialo-GM1 using sialyltransferses in rat liver Golgi vesicles. Eur J Biochem 221:603–609
44. Irie F, Hidari K, Tai T, Li Y-T, Seyama Y Hirabayashi Y (1994) Biosynthetic pathway for a new series of gangliosides, GT1α and GQ1α. FEBS Lett 351:291–294
45. Simpson MA, Cross H, Proukakis C, Priestman DA, Neville DC, Reinkensmeier G, Wang H, Wiznitzer M, Gurtz K, Verganelaki A, Pryde A, Patton MA, Dwek RA, Butters TD, Platt FM, Crosby AH (2004) Infantile-onset symptomatic epilepsy syndrome caused by a homozygous loss-of-function mutation of GM3 synthase. Nat Genet 36:1225–1229
46. Lannert H, Gorgas K, Meissner I, Wieland FT, Jeckel D (1998) Functional organization of the Golgi apparatus in glycosphingolipid biosynthesis. Lactosylceramide and subsequent glycosphingolipids are formed in the lumen of the late Golgi. J Biol Chem 273:2939–2946
47. Schapiro FB, Lingwood C, Furuya W, Grinstein S (1998) pH-independent retrograde targeting of glycolipids to the Golgi complex. Am J Physiol 274:C319–C332
48. Kok JW, Hoekstra K, Eskelinen S, Hoekstra D (1992) Recycling pathways of glucosylceramide in BHK cells: distinct involvement of early and late endosomes. J Cell Sci 103:1139–1152
49. Drickamer K (1993) A conserved disulphide bond in sialyltransferases. Glycobiology 3:2–3
50. Livingston BD, Paulson JC (1993) Polymerase chain reaction cloning of a developmentally regulated member of the sialyltransferase gene family. J Biol Chem 268:11504–11507
51. Jeanneau C, Chazalet V, Auge C, Soumpasis DM, Harduin-Lepers A, Delannoy P, Imberty A, Breton C (2004) Structure-function analysis of the human sialyltransferase ST3Gal I: role of n-glycosylation and a novel conserved sialylmotif. J Biol Chem 279:13461–13468
52. Geremia RA, Harduin-Lepers A, Delannoy P (1997) Identification of two novel conserved amino acid residues in eukaryotic sialyltransferases: implications for their mechanism of action. Glycobiology 7:v–vii
53. Kapitonov D, Bieberich E, Yu RK (1999) Combinatorial PCR approach to homology-based cloning: cloning and expression of mouse and human GM3-synthase. Glycoconj J 16:337–350
54. Harduin-Lepers A, Mollicone R, Delannoy P, Oriol R (2005) The animal sialyltransferases and sialyltransferase-related genes: a phylogenetic approach. Glycobiology 15:805–817
55. Lee YC, Kurosawa N, Hamamoto T, Nakaoka T, Tsuji S (1993) Molecular cloning and expression of Galβ1,3GalNAc α2,3-sialyltransferase from mouse brain. Eur J Biochem 216:377–385
56. Kitagawa H, Paulson JC (1994) Differential expression of five sialyltransferase genes in human tissues. J Biol Chem 269:17872–17878
57. Ishii A, Ohta M, Watanabe Y, Matsuda K, Ishiyama K, Sakoe K, Nakamura M, Inokuchi J, Sanai Y, Saito M (1998) Expression cloning and functional characterization of human cDNA for ganglioside GM3 synthase. J Biol Chem 273:31652–31655
58. Kono M, Takashima S, Liu H, Inoue M, Kojima N, Lee YC, Hamamoto T, Tsuji S (1998) Molecular cloning and functional expression of a fifth-type α2,3-sialyltransferase (mST3Gal V: GM3 synthase). Biochem Biophys Res Commun 253:170–175
59. Nara K, Watanabe Y, Maruyama K, Kasahara K, Nagai Y, Sanai Y (1994) Expression cloning of a CMP-NeuAc:NeuAc α2-3Gal β1-4Glc β1-1′Cer α2,8-sialyltransferase (GD3 synthase) from human melanoma cells. Proc Natl Acad Sci USA 91:7952–7956
60. Haraguchi M, Yamashiro S, Yamamoto A, Furukawa K, Takamiya K, Lloyd KO, Shiku H, Furukawa K (1994) Isolation of GD3 synthase gene by expression cloning of GM3 α2,8-sialyltransferase cDNA using anti-GD2 monoclonal antibody. Proc Natl Acad Sci USA 91:10455–10459
61. Sasaki K, Kurata K, Kojima N, Kurosawa N, Ohta S, Hanai N, Tsuji S, Nishi T (1994) Expression cloning of a GM3-specific α-2,8-sialyltransferase (GD3 synthase). J Biol Chem 269:15950–15956

62. Yamamoto A, Haraguchi M, Yamashiro S, Fukumoto S, Furukawa K, Takamiya K, Atsuta M, Shiku H, Furukawa K (1996) Heterogeneity in the expression pattern of two ganglioside synthase genes during mouse brain development. J Neurochem 66:26–34
63. Zeng G, Gao L, Ariga T, Yu RK (1996) Molecular cloning of cDNA for rat brain GD3-synthase. Biochem Biophys Res Commun 226:319–323
64. Nara K, Watanabe Y, Kawashima I, Tai T, Nagai Y, Sanai Y (1996) Acceptor substrate specificity of a cloned GD3 synthase that catalyzes the biosynthesis of both GD3 and GD1c/GT1a/GQ1b. Eur J Biochem 238:647–652
65. Nakayama J, Fukuda MN, Hirabayashi Y, Kanamori A, Sasaki K, Nishi T, Fukuda M (1996) Expression cloning of a human GT3 synthase. GD3 AND GT3 are synthesized by a single enzyme. J Biol Chem 271:3684–3691
66. Yanagisawa M, Liour SS, Yu RK (2004) Involvement of gangliosides in proliferation of immortalized neural progenitor cells. J Neurochem 91:804–812
67. Kono M, Yoshida Y, Kojima N, Tsuji S (1996) Molecular cloning and expression of a fifth type of α2,8-sialyltransferase (ST8Sia V) Its substrate specificity is similar to that of SAT-V/III, which synthesize GD1c, GT1a, GQ1b and GT3. J Biol Chem 271:29366–29371
68. Kim YJ, Kim KS, Do S, Kim CH, Kim SK, Lee YC (1997) Molecular cloning and expression of human α2,8-sialyltransferase (hST8Sia V). Biochem Biophys Res Commun 235:327–330
69. Yoshida Y, Kojima N, Kurosawa N, Hamamoto T, Tsuji S (1995) Molecular cloning of Sia α2,3Gal β1,4GlcNAc α2,8-sialyltransferase from mouse brain. J Biol Chem 270:14628–14633
70. Zeng G, Gao L, Yu RK (1997) Cloning of the cDNA coding for rat brain CMP-NeuAc:GD3 α2-8 sialyltransferase. Gene 187:131–134
71. Lee YC, Kim YJ, Lee KY, Kim KS, Kim BU, Kim HN, Kim CH, Do SI (1998) Cloning and expression of cDNA for a human Sia α2,3Gal β1,4GlcNA:α2,8-sialyltransferase (hST8Sia III). Arch Biochem Biophys 360:41–46
72. Sjoberg ER, Kitagawa H, Glushka J, van Halbeek H, Paulson JC (1996) Molecular cloning of a developmentally regulated N-acetylgalactosamine α2,6-sialyltransferase specific for sialylated glycoconjugates. J Biol Chem 271:7450–7459
73. Lee YC, Kaufmann M, Kitazume-Kawaguchi S, Kono M, Takashima S, Kurosawa N, Liu H, Pircher H, Tsuji S (1999) Molecular cloning and functional expression of two members of mouse NeuAcα2,3Galβ1,3GalNAc GalNAcα2,6-sialyltransferase family, ST6GalNAc III and IV. J Biol Chem 274:11958–11967
74. Okajima T, Fukumoto S, Ito H, Kiso M, Hirabayashi Y, Urano T, Furukawa K (1999) Molecular cloning of brain-specific GD1α synthase (ST6GalNAc V) containing CAG/Glutamine repeats. J Biol Chem 274:30557–30562
75. Kurosawa N, Kojima N, Inoue M, Hamamoto T, Tsuji S (1994) Cloning and expression of Gal β1,3GalNAc-sp ecific GalNAc α2,6-sialyltransferase. J Biol Chem 269: 19048–19053
76. Kurosawa N, Inoue M, Yoshida Y, Tsuji S (1996) Molecular cloning and genomic analysis of mouse Galβ1, 3GalNAc-specific GalNAc α2,6-sialyltransferase. J Biol Chem 271:15109–15116
77. Samyn-Petit B, Krzewinski-Recchi MA, Steelant WF, Delannoy P, Harduin-Lepers A (2000) Molecular cloning and functional expression of human ST6GalNAc II Molecular expression in various human cultured cells. Biochim Biophys Acta 1474:201–211
78. Ikehara Y, Shimizu N, Kono M, Nishihara S, Nakanishi H, Kitamura T, Narimatsu H, Tsuji S, Tatematsu M (1999) A novel glycosyltransferase with a polyglutamine repeat; a new candidate for GD1α synthase (ST6GalNAc V)(1). FEBS Lett 463:92–96
79. Nomura T, Takizawa M, Aoki J, Arai H, Inoue K, Wakisaka E, Yoshizuka N, Imokawa G, Dohmae N, Takio K, Hattori M, Matsuo N (1998) Purification, cDNA cloning, and expression of UDP-Gal: glucosylceramide β1,4-galactosyltransferase from rat brain. J Biol Chem 273:13570–13577
80. Takizawa M, Nomura T, Wakisaka E, Yoshizuka N, Aoki J, Arai H, Inoue K, Hattori M, Matsuo N (1999) cDNA cloning and expression of human lactosylceramide synthase. Biochim Biophys Acta 1438:301–304
81. Miyazaki H, Fukumoto S, Okada M, Hasegawa T, Furukawa K (1997) Expression cloning of rat cDNA encoding UDP-galactose:GD2 β1,3-galactosyltransferase that determines the expression of GD1b/GM1/GA1. J Biol Chem 272:24794–24799

82. Amado M, Almeida R, Carneiro F, Levery SB, Holmes EH, Nomoto M, Hollingsworth MA, Hassan H, Schwientek T, Nielsen PA, Bennett EP, Clausen H (1998) A family of human β3-galactosyltransferases Characterization of four members of a UDP-galactose:β-N-acetyl-glucosamine/β-N-acetyl-galactosamine β-1,3-galactosyltransferase family. J Biol Chem 273:12770–12778
83. Daniotti JL, Martina JA, Zurita AR, Maccioni HJ (1999) Mouse β1,3-galactosyltransferase (GA1/GM1/GD1b synthase): protein characterization, tissue expression, and developmental regulation in neural retina. J Neurosci Res 58:318–327
84. Schulte S, Stoffel W (1993) Ceramide UDP-galactosyltransferase from myelinating rat brain: purification, cloning, and expression. Proc Natl Acad Sci USA 90:10265–10269
85. Bosio A, Binczek E, Stoffel W (1996) Molecular cloning and characterization of the mouse CGT gene encoding UDP-galactose ceramide-galactosyltransferase (cerebroside synthetase). Genomics 35:223–226
86. Coetzee T, Li X, Fujita N, Marcus J, Suzuki K, Francke U, Popko B (1996) Molecular cloning, chromosomal mapping, and characterization of the mouse UDP-galactose:ceramide galactosyltransferase gene. Genomics 35:215–222
87. Kapitonov D, Yu RK (1997) Cloning, characterization, and expression of human ceramide galactosyltransferase cDNA. Biochem Biophys Res Commun 232:449–453
88. Bosio A, Binczek E, Le Beau MM, Fernald AA, Stoffel W (1996) The human gene CGT encoding the UDP-galactose ceramide galactosyl transferase (cerebroside synthase): cloning, characterization, and assignment to human chromosome 4, band q26. Genomics 34: 69–75
89. Nagata Y, Yamashiro S, Yodoi J, Lloyd KO, Shiku H, Furukawa K (1992) Expression cloning of β1,4 N-acetylgalactosaminyltransferase cDNAs that determine the expression of GM2 and GD2 gangliosides. J Biol Chem 267:12082–12089
90. Hidari JK, Ichikawa S, Furukawa K, Yamasaki M, Hirabayashi Y (1994) β1-4N-acetylgalactosaminyltransferase can synthesize both asialoglycosphingolipid GM2 and glycosphingolipid GM2 in vitro and in vivo: isolation and characterization of a β1-4N-acetylgalactosaminyltransferase cDNA clone from rat ascites hepatoma cell line AH7974F. Biochem J 303:957–965
91. Takamiya K, Yamamoto A, Yamashiro S, Furukawa K, Haraguchi M, Okada M, Ikeda T, Shiku H, Furukawa K (1995) T cell receptor-mediated stimulation of mouse thymocytes induces up-regulation of the GM2/GD2 synthase gene. FEBS Lett 358:79–83
92. Sango K, Johnson ON, Kozak CA, Proia RL (1995) β1,4-N-Acetylgalactosaminyltransferase involved in ganglioside synthesis: cDNA sequence, expression, and chromosome mapping of the mouse gene. Genomics 27:362–365
93. Lowe AM, Lambert PA, Smith AW (1995) Cloning of an Enterococcus faecalis endocarditis antigen: homology with adhesins from some oral streptococci. Infect Immun 63:703–706
94. Lo Presti L, Cabuy E, Chiricolo M, Dall'Olio F (2003) Molecular cloning of the human β1,4 N-acetylgalactosaminyltransferase responsible for the biosynthesis of the Sd(a) histo-blood group antigen: the sequence predicts a very long cytoplasmic domain. J Biochem 134:675–682
95. Yanagisawa K, Taniguchi N, Makita A (1987) Purification and properties of GM2 synthase, UDP-N-acetylgalactosamine: GM3 β-N-acetyl-galactosaminyltransferase from rat liver. Biochim Biophys Acta 919:213–220
96. Ichikawa S, Sakiyama H, Suzuki G, Hidari KI, Hirabayashi Y (1996) Expression cloning of a cDNA for human ceramide glucosyltransferase that catalyzes the first glycosylation step of glycosphingolipid synthesis. Proc Natl Acad Sci USA 93:4638–4643
97. Ichikawa S, Ozawa K, Hirabayashi Y (1998) Molecular cloning and expression of mouse ceramide glucosyltransferase. Biochem Mol Biol Int 44:1193–1202
98. Wu K, Marks DL, Watanabe R, Paul P, Rajan N, Pagano RE (1999) Histidine-193 of rat glucosylceramide synthase resides in a UDP-glucose- and inhibitor (D-threo-1-phenyl-2-decanoylamino-3-morpholinopropan-1-ol)-binding region: a biochemical and mutational study. Biochem J 341:395–400
99. Spiess M (1995) Heads or tails—what determines the orientation of proteins in the membrane. FEBS Lett 369:76–79
100. Yamamoto F, Clausen H, White T, Marken J, Hakomori S (1990) Molecular genetic basis of the histo-blood group ABO system. Nature 345:229–233

101. Breton C, Snajdrova L, Jeanneau C, Koca J, Imberty A (2006) Structures and mechanisms of glycosyltransferases. Glycobiology 16:29R–37R
102. Yang C, Bolotin E, Jiang T, Sladek FM, Martinez E (2007) Prevalence of the initiator over the TATA box in human and yeast genes and identification of DNA motifs enriched in human TATA-less core promoters. Gene 389: 52–65
103. Zeng G, Gao L, Yu RK (1998) Isolation and functional analysis of the promoter of the rat CMP-NeuAc:GM3 α2,8 sialyltransferase gene 1. Biochim Biophys Acta 1397:126–130
104. Takashima S, Kono M, Kurosawa N, Yoshida Y, Tachida Y, Inoue M, Kanematsu T, Tsuji S (2000) Genomic organization and transcriptional regulation of the mouse GD3 synthase gene (ST8Sia I): comparison of genomic organization of the mouse sialyltransferase genes. J Biochem 128:1033–1043
105. Furukawa K, Horie M, Okutomi K, Sugano S, Furukawa K (2003) Isolation and functional analysis of the melanoma specific promoter region of human GD3 synthase gene. Biochim Biophys Acta 1627:71–78
106. Kim KW, Kim SW, Min KS, Kim CH, Lee YC (2001) Genomic structure of human GM3 synthase gene (hST3Gal V) and identification of mRNA isoforms in the 5′-untranslated region. Gene 273:163–171
107. Fukumoto S, Miyazaki H, Goto G, Urano T, Furukawa K, Furukawa K (1999) Expression cloning of mouse cDNA of CMP-NeuAc: Lactosylceramide α2,3-sialyltransferase, an enzyme that initiates the synthesis of gangliosides. J Biol Chem 274:9271–9276
108. Kim SW, Lee SH, Kim KS, Kim CH, Choo YK, Lee YC (2002) Isolation and characterization of the promoter region of the human GM3 synthase gene. Biochim Biophys Acta 1578:84–89
109. Zeng G, Gao L, Xia T, Tencomnao T, Yu RK (2003) Characterization of the 5′-flanking fragment of the human GM3-synthase gene. Biochim Biophys Acta 1625:30–35
110. Xia T, Zeng G, Gao L, Yu RK (2005) Sp1 and AP2 enhance promoter activity of the mouse GM3-synthase gene. Gene 351:109–118
111. Chung TW, Choi HJ, Lee YC, Kim CH (2005) Molecular mechanism for transcriptional activation of ganglioside GM3 synthase and its function in differentiation of HL-60 cells. Glycobiology 15:233–244
112. Cheung NK, Saarinen UM, Neely JE, Landmeier B, Donovan D, Coccia PF (1985) Monoclonal antibodies to a glycolipid antigen on human neuroblastoma cells. Cancer Res 45:2642–2649
113. Yamashiro S, Ruan S, Furukawa K, Tai T, Lloyd KO, Shiku H, Furukawa K (1993) Genetic and enzymatic basis for the differential expression of GM2 and GD2 gangliosides in human cancer cell lines. Cancer Res 53:5395–5400
114. Furukawa K, Soejima H, Niikawa N, Shiku H (1996) Genomic organization and chromosomal assignment of the human β1, 4-N-acetylgalactosaminyltransferase gene Identification of multiple transcription units. J Biol Chem 271:20836–20844
115. Ichikawa S, Ozawa K, Hirabayashi Y (1998) Molecular cloning and characterization of the mouse ceramide glucosyltransferase gene. Biochem Biophys Res Commun 253:707–711
116. Xia T, Gao L, Yu RK, Zeng G (2003) Characterization of the promoter and the transcription factors for the mouse UDP-Gal:βGlcNAc β1,3-galactosyltransferase gene. Gene 309:117–123
117. Ledeen RW, Wu G, Lu ZH, Kozireski-Chuback D, Fang Y (1998) The role of GM1 and other gangliosides in neuronal differentiation Overview and new finding. Ann N Y Acad Sci 845:161–175
118. Tencomnao T, Yu RK, Kapitonov D (2001) Characterization of the human UDP-galactose:ceramide galactosyltransferase gene promoter. Biochim Biophys Acta 1517:416–423
119. Tencomnao T, Kapitonov D, Bieberich E, Yu RK (2004) Transcriptional regulation of the human UDP-galactose:ceramide galactosyltransferase (hCGT) gene expression: functional role of GC-box and CRE. Glycoconj J 20:339–351
120. Yu RK, Lee SH (1976) In vitro biosynthesis of sialosylgalactosylceramide (G7) by mouse brain microsomes. J Biol Chem 251:198–203
121. Sandhoff K, Kolter T (1997) Biochemistry of glycosphingolipid degradation. Clin Chim Acta 266:51–61
122. Kytzia HJ, Sandhoff K (1985) Evidence for two different active sites on human hexosaminidase A: interaction of GM2 activator protein with hexosaminidase. J Biol Chem 260:7568–7572
123. Morimoto S, Martin BM, Kishimoto Y, O'Brien JS (1988) Saposin D: a sphingomyelinase activator. Biochim Biophys Res Commun 156:403–410

124. Furst W, Machleidt W, Sandhoff K (1988) The precursor of sulfatide activator protein is processed to three different proteins. Biol Chem Hoppe-Seyler 369:317–328
125. O'Brien JS, Kretz KA, Dewji N, Wenger DA, Esch F, Fluharty AL (1988) Coding of two sphingolipid activator proteins (SAP1 and SAP2) by same genetic locus. Science 241:1098–1101
126. Fingerhut R, Van der Horst GT, Verheijen FW, Conzelmann E (1992) Degradation of gangliosides by the lysosomal sialidase requires an activator protein. Eur J Biochemi 208:623–629
127. Wilkening G, Linke T, Uhlhorn-Dierks G, Sandhoff K (2000) Degradation of membrane-bound ganglioside GM1. J Biol Chem 275:35814–35819
128. Zschoche A, Furst W, Schwarzmann G, Sanhoff K (1994) Hydrolysis of lactosylceramide by human galacrtosylceramidase and GM1-β-galactosidase in a detergent-free system and its stimulation by sphingolipid activator protein, sap-B and sup-C. Activator proteins stimulate lactosylceramide hydrolysis. Eur J Biochem 222:83–90
129. Brady RO, Gal AE, Kanfer JN, Brady RM (1965) The metabolism of glucocerebrosides. 3. Purification and properties of a glucosyl- and galactosylceramide-cleaving enzyme from rat intestinal tissue. J Biol Chem 240:3766–3770
130. Huwiler A, Kolter T, Pfeilschifter J, Sandhoff K (2000) Physiology and pathophysiology of sphingolipid metabolism and signaling. Biochim Biophys Acta 1485:63–99
131. Linke T, Wilkening G, Sadeghlar F, Mozcall H, Bernardo K, Schuchman E, Sandhoff K (2001) Interfacial regulation of acid ceramidase activity. J Biol Chem 276:5760–5768
132. Jeyakumar M, Dwek RA, Butters TD, Platt FM (2005) Storage solutions: treating lysosomal disorders of the brain. Nat Rev Neurosci 6:713–725
133. Gieselmann V (2005) What can cell biology tell us about heterogeneity in lysosomal storage diseases? Acta Paediatr Suppl 94:80–86
134. Simons K, Gruenberg J (2000) Jamming the endosomal system: lipid rafts and lysosomal storage diseases. Trends Cell Biol 10:459–462
135. Walkley SU (2004) Secondary accumulation of gangliosides in lysosomal storage disorders. Cell Dev Biol 15:433–444
136. Beratis NG, Varvarigou-Frimas A, Beratis S, Sklower SL (1989) Angiokeratoma corporis diffusum in GM1 gangliosidosis, type 1. Clin Genet 36:59–64
137. Gascon GG, Ozand PT, Erwin RE (1992) GM1 gangliosidosis type 2 in two siblings. J Child Neurol 7:S41–50
138. Inui K, Namba R, Ihara Y, Nobukuni K, Tanike M, Midorikawa M, Tsukamoto H, Okada S (1990) A case of chronic GM1 gangliosidosis presenting as dystonia: a clinical and biochemical studies. J Neurol 237:491–493
139. Suzuki Y, Oshima A (1993) A β-galactosidase gene mutation identified in both Morquio B disease and infantile GM1 gangliosidosis. Hum Gent 91:407
140. Goldman JE, Katz D, Rapin I, Purpura DP, Suzuki K (1981) Chronic GM1 gangliosidosis presenting as dystonia: I. Clinical and pathological features. Ann Neurol 9:465–475
141. Hahn CN, del Pilar Martin M, Schroder M, Vanier MT, Hara Y, Suzuki K, Suzuki K, d'Azzo A (1997) Generalized CNS disease and massive GM1-ganglioside accumulation in mice defective in lysosomal acid β-galactosidase. Hum Mol Genet 6:205–211
142. Takaura N, Yagi T, Maeda M, Nanba E, Oshima A, Suzuki Y, Yamano T, Tanaka A (2003) Attenuation of ganglioside GM1 accumulation in the brain of GM1 gangliosidosis mice by neonatal intravenous gene transfer. Gene Ther 10:1487–1493
143. Kasperzyk JL, d'Azzo A, Platt FM, Alroy J, Seyfried TN (2005) Substrate reduction reduces gangliosides in postnatal cerebrum-brainstem and cerebellum in GM1 gangliosidosis mice. J Lipid Res 46:744–751
144. De Maria R, Divari S, Bo S, Sonnio S, Lotti D, Capucchio MT, Castagnaro M (1998) β-galactosidase deficiency in Korat cat: a new form of feline GM1-gangliosidosis. Acta Neuropathol 96:307–314
145. Cox NR, Ewald SJ, Morrison NE, Gentry AS, Schuler M, Baker HJ (1998) Thymic alterations in feline GM1 gangliosidosis. Vet Immunol Immunopathol 63:335–353
146. Murnane RD, Wright RW, Ahern-Rindell AJ, Prieur DJ (1991) Prenatal lesions in an ovine fetus with GM1 gangliosidosis. Am J Med Genet 39:106–111
147. Saunders GK, Wood PA, Myers RK, Shell LG, Carithers R (1988) GM1 gangliosidosis in Portuguese water dogs: pathogenic and biochemical findings. Vet Pathol 25:265–269
148. Alroy J, Orgad U, DeGasperi R, Richard R, Warren CD, Knowles K, Thalhammer JG, Raghavan

SS (1992) Canine GM1-gangliosidosis. A clinical, morphologic, hisochemical and biochemical comparison of two different models. Am J Pathol 140:675–689
149. Yamato O, Ochiai K, Masuoka Y, Hayashida E, Tajima M, Omae S, Iijima M, Umemura T, Maeda Y (2000) GM1 gangliosidosis in shiba dogs. Vet Rec 146:493–496
150. Kobayashi T, Goto I, Okada S, Orii T, Ohno K, Nakano T (1992) Accumulation of lysosphingolipids in tissues from patients with GM1 and GM2 gangliosidoses. J Neurochem 59:1452–1458
151. Gravel RA, Clarke JTR, Kaback MM, Mahuran D, Sandhoff K, Suzuki K (1995) GM2 gangliosidoses. In: Scriver CR (ed) The Metabolic and Molecular Basis of Inherited Diseases. McGraw-Hill, New York, pp 2839–2879
152. Mahuran DJ (1999) Biochemical consequences of mutations causing the GM2 gangliosidoses. Biochim Biophys Acta 1455:105–138
153. Lemieux MJ, Mark BL, Cherney MM, Withers SG, Mahuran DJ, James MN (2006) Crystallographic structure of human β-hexosaminidase A: Interpretation of Tay–Sachs mutations and loss of GM2 ganglioside hydrolysis. J Mol Biol 359:913–929
154. Gravel RA, Kaback MM, Proia RL, Sandhoff K, Suzuki K, Suzuki K (2001) The GM2 gangliosdoses. In: Scriver CR, Beaudet AL, Valle D, and Sly WS (eds) The Metabolic and Molecular Bases of Inherited Diseases. McGraw-Hill, New York, pp 3827–3876
155. Sango K, Yamanaka S, Hoffmann A, Okuda Y, Grinberg A, Westphal H, McDonald MP, Crawley JN, Sandhoff K, Suzuki K, Proia RL (1995) Mouse models of Tay–Sachs and Sandhoff diseases differ in neurologic phenotype and ganglioside metabolism. Nat Genet 11:170–176
156. Sango K, Takano M, Ajiki K, Tokashiki A, Arai N, Kawano H, Horie H, Yamanaka S (2005) Impaired neurite outgrowth in the retina of a murine model of Sandhoff disease. Invest Ophthalmol Vis Sci 46:3420–3425
157. Phaneuf D, Wakamatsu N, Huang JQ, Borowski A, Peterson AC, Fortunato SR, Ritter G, Igdoura SA, Morales CR, Benoit G, Akerman BR, Leclerc D, Hanai N, Marth JD, Trasler JM, Gravel RA (1996) Dramatically different phenotypes in mouse models of human Tay–Sachs and Sandhoff disease. Hum Mol Genet 5:1–14
158. Andersson U, Smith D, Jeyakumar M, Butters TD, Borja MC, Dwek RA, Platt FM (2004) Improved outcome of N-butyldeoxygalactonojirimycin-mediated substrate reduction theraphy in a mouse model of Sandhoff disease. Neurobiol Dis 16:506–515
159. Walkley SU, Zervas M, Wiseman S (2000) Gangliosides as modulators of dendritogenesis in normal and storage disease-affected pyramidal neurons. Cereb Cortex 10:1028–1037
160. Martin DR, Cox NR, Morrison NE, Kennamer DM, Peck SL, Dodson AN, Gentry AS, Griffin B, Rolsma MD, Baker HJ (2005) Mutation of the GM2 activator protein in a feline model of GM2 gangliosidosis. Acta Neuropathol 110:443–450
161. Singer HS, Cork LC (1989) Canine GM2 gangliosidosis: morphological and biochemical analysis. Vet Pathol 26:114–120
162. Yamato O, Satoh H, Matsuki N, Ono K, Yamasaki M, Maede Y (2004) Laboratory diagnosis of canine GM2-gangliosidosis using blood and cerebrospinal fluid. J Vet Diagn Invest 16:39–44
163. Neuenhofer S, Conzelmann E, Schwarzmann G, Egge H, Sandhoff K (1986) Occurrence of lysoganglioside lyso-GM2 (II3-neu5Ac-gangliotriaosylsphingosine) in GM2 gangliosidosis brain. Biol Chem Hoppe Seyler 367:241–244
164. Li YT, Maskos K, Chou CW, Cole RB, Li SC (2003) Presence of an unusual GM2 derivative, taurine-conjugated GM2, in Tay–Sachs brain. J Biol Chem 278:35286–35291
165. Itoh T, Li YT, Li SC, Yu RK (1981) Isolation and characterization of a novel monosialosylpentahexosyl ceramide from Tay–Sachs brain. J Biol Chem 256:165–169
166. Yu RK, Itoh T, Yohe H, Macala LJ (1983) Characterization of some minor gangliosides in Tay–Sachs brains. Brain Res 275:47–52
167. Bermudez AJ, Johnson GC, Vanier MT, Schroder M, Suzuki K, Stogsdill PL, Johnson GS, O'Brien D, Moore CP, Fry WW (1995) Gangliosidosis in emu (Dromaius novaehollandiae). Avian Dis 39:292–303
168. Freischutz B, Tokuda A, Ariga T, Bermudez AJ, Yu RK (1997) Unusual gangliosidosis in emu (Dromaius novaehollandiae). J Neurochem 68:2070–2078

7.4 Chemical Synthesis of Glycosylphosphatidylinositol (GPI) Anchors and GPI-Linked Structures

Zhongwu Guo, *Lee Bishop*
Department of Chemistry, Wayne State University,
5101 Cass Avenue, Detroit, MI 48202, USA
zwguo@chem.wayne.edu

Abstract

Glycosylphosphatidylinositols (GPIs) are a class of natural glycolipids that are ubiquitously expressed by eukaryotic cells. One of the fundamental functions of GPIs is to anchor proteins, glycoproteins, and other molecules onto cell surfaces. The structures of GPIs are diverse and heterogeneous. To obtain homogeneous and well-defined GPIs for various studies, the chemical synthesis of GPIs has become a hot topic. This chapter summarizes the recent progress in the chemical synthesis of GPIs as well as related structures, such as the GPI-linked peptides and glycopeptides.

In: *Glycoscience*. Fraser-Reid B, Tatsuta K, Thiem J (eds)
Chapter-DOI 10-1007/978-3-540-30429-6_42:

Keywords

Glycosylphosphatidylinositol; GPI; Glycolipid; Glycoconjugate; Carbohydrate; Inositol; Phosphate; Lipid

Abbreviations

Ac	acetyl
AChE	human acetylcholinesterase
acyl-CoA	acyl coenzyme A
AG	acyl glycerol
AIBN	azoisobutyronitrile
alk-AG	alkyl acyl glycerol
All	allyl
Bn	benzyl
Boc	*t*-butyloxycarbonyl
***t*-Bu**	*t*-butyl
Bz	benzoyl
CAM	camphanic
camph	camphylidene
Cbz	benzyloxycarbonyl
ClAc	2-chloroacetyl
***m*-CPBA**	*m*-chloroperbenzoic acid
DCC	*N*,*N*′-dicyclohexylcarbodiimide
Dol-P-Man	dolicholphosphomannose
EDC	1-[3-(dimethylamino)propyl]-3-ethylcarbodiimide hydrochloride
FBP	folate binding protein
Fmoc	9-fluorenylmethoxycarbonyl
Gal	galactose
GIPL	glycoinositol phospholipids
Glc	glucose
GlcNAc	*N*-acetylglucosamine
$GlcNH_2$	glucosamine
GPI	glycosylphosphatidylinositol
HOBt	*N*-hydroxybenzotriazole
KLH	keyhole limpet haemocyanin
Lev	levulinoyl
LHMDS	lithium hexamethyldisilazide
LPG	lipophosphoglycan
Man	mannose
MBn	4-methoxybenzyl
Me	methyl
Mnt	menthyloxycarbonyl
MPh	4-methoxyphenyl
NeuNAc	*N*-acetyl neuraminic acid (sialic acid)
NIS	*N*-iodosuccinimide

PA	phenoxyacetyl
PARP	procyclic acidic repetitive protein
Pent	pentenyl
Ph	phenyl
Phth	phthalyl
PI	phosphatidylinositol
PI-PLC	PI-specific phospholipase C
P-OEtNH$_2$	phosphoethanolamine
***i*-Pr**	*iso*-propyl
Prp	prion protein
PsA	prespore-specific antigen
TBDMS	*t*-butyldimethylsilyl
TBDPS	*t*-butyldiphenylsilyl
TClAc	trichloroacetyl
TFA	trifluoroacetic acid
TIPDS	1,1,3,3-tetraisopropyldisiloxyl
TIPS	triisopropylsilyl
TPAP	tetrapropylammonium perruthenate
Trt	trityl (triphenylmethyl)
VSG	variant surface glycoprotein

1 Introduction

Glycosylphosphatidylinositol (GPI) anchors are a class of natural glycolipids which share a common core structure: Manα(1→4)GlcNH$_2\alpha$(1→6)-*myo*-Inositol-1-PO$_4$-lipid (**1**, ❯ *Fig. 1*). Presently, there are two types of GPIs identified. One type of GPIs, such as free glycoinositol phospholipids (GIPL) and the GPIs of lipophosphoglycans (LPGs) [1,2], anchor non-protein extracellular glycoconjugates onto cell surfaces. These GPIs are observed in protozoal parasites, and the common structure that they share is the basic GPI core mentioned above. Another type of GPIs are membrane protein and glycoprotein anchors ubiquitously expressed by mammals and by lower eukaryotes as well [3]. All GPI anchors of the second type have the conserved structure Manα(1→2)Manα(1→6)Manα(1→4)GlcNH$_2\alpha$(1→6)-*myo*-Inositol-1-PO$_4$-lipid (**2**, ❯ *Fig. 1*). The *C*-termini of GPI-anchored proteins and glycoproteins are coupled to the 6-*O*-position of the non-reducing end mannose of GPIs through a phosphoethanolamine group (P-OEtNH$_2$).

As depicted in ❯ *Fig. 1*, the lipid chains of GPIs can insert into cell membranes, enabling the anchoring of proteins and glycoproteins onto cell surfaces. Consequently, GPIs, as well as GPI-anchored proteins and glycoproteins, play an important role in a broad range of biological and pathological processes, from cell recognition, binding, and signal transduction to extracellular enzymatic reactions and microorganism invasion [3,4,5,6,7,8].

The first complete structural characterization of a GPI, achieved by Ferguson and coworkers in 1988, was of the *T. brucei* variant surface glycoprotein (VSG) anchor [9]. This study was assisted by the fact that the African trypanosome system exhibits abundant VSG (ca. 10^7 copies per cell) [10], which also facilitated the development of the first cell-free

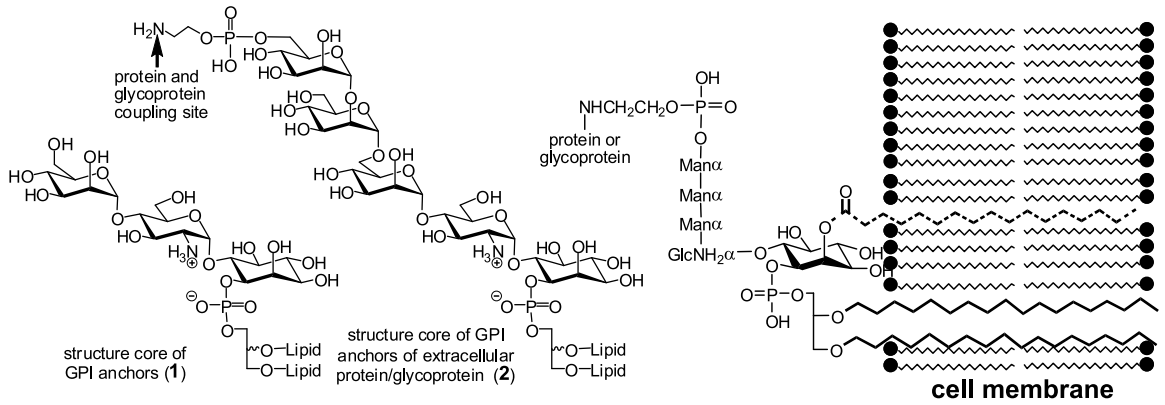


Figure 1
The core structures of glycosylphosphatidylinositol (GPI) anchors and the GPI anchoring of proteins and glycoproteins onto cell surfaces

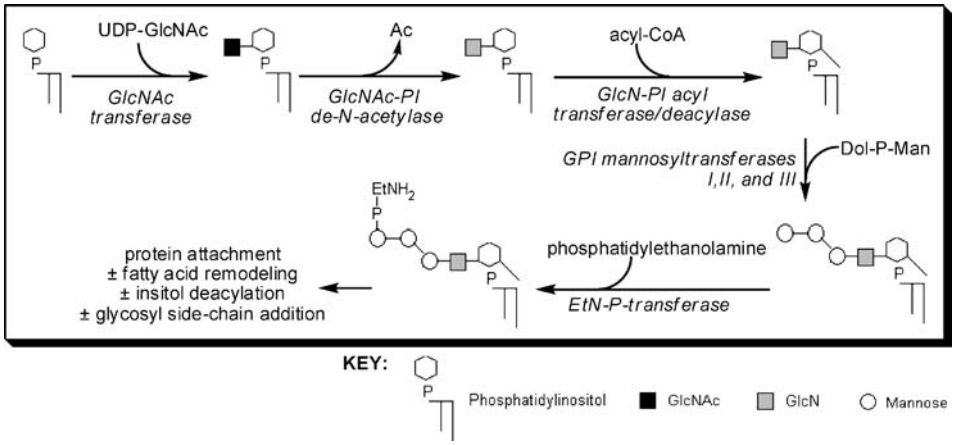


Figure 2

Biosynthesis of GPIs in mammalian cells. After a GlcNAc residue is added to PI followed by deacetylation, an acyl group is added to the inositol 2-*O*-position. Then, three mannose residues are added sequentially to form the GPI glycan core. The final step of GPI biosynthesis is the addition of a P-OEtNH_2 function to the 6-*O*-position of Man III. The GPI is added to proteins or glycoproteins before further modifications. Some mammalian biosynthetic intermediates are found to contain other phosphoethanolamine residues linked to either mannose I or III

system by Masterson and coworkers [11] that was employed to examine the biosynthesis of GPIs and GPI-linked proteins/glycoproteins. The conserved glycan core of GPI anchors suggests a generally conserved biosynthetic pathway among different species, which has been depicted in *Fig. 2*.

The biosynthesis of GPI anchors begins with the transfer of GlcNAc to phosphatidylinositol (PI) from UDP-GlcNAc to form GlcNAc-PI, which is followed by de-*N*-acetylation to form GlcNH_2-PI. The third step in the biosynthetic pathway is the reversible acylation of the 2-*O*-position of the inositol ring. This acylation is dependent on the presence of acyl coenzyme A (acyl-CoA), and has been shown to be stimulated in cell-free systems by its mere presence [12]. Inositol acylation confers resistance of the final GPI anchor product to cleavage catalyzed by PI-specific phospholipase C (PI-PLC) [13,14]. Most GPI anchors undergo cleavage of this acyl group prior to the addition to protein; however, some GPI anchors, such as the procyclic acidic repetitive protein (PARP) of procyclic *T. brucei*, the human CD52 antigen, and the human acetylcholinesterase (AChE), retain the inositol acyl group [13,15,16,17,18,19,20,21,22,23]. Thereafter, the biosynthetic pathway is continued by sequential addition of three mannose residues from the donor dolicholphosphomannose (Dol-P-Man) catalyzed by GPI mannosyltransferases I, II, and III, respectively. The final biosynthetic step completing the GPI anchor core is addition of P-OEtNH_2 to the 6-*O*-position of the terminal mannose from donor phosphatidylethanolamine. Following this step, the original lipid anchors are replaced in a process called fatty acid remodeling [24,25]. The completed GPI core is then added to proteins/glycoproteins through what is thought to be a transamidation reaction. The addition of glycosyl side chains and other functional groups to the GPI anchor core occurs after the completed anchor core is added to proteins/glycoproteins.

Table 1
GPI anchor structures

Anchor	R6	R5	R4	R3	R2	R1	Inositol-Acyl	Lipid	Ref
T. brucei VSG	-	-	-	±Galα(1→2)Galα(1→6)-[Galα(1→2)]Galα(1→3)		-	-	di-AG	[29]
T. brucei PARP	-	[	$NeuNAc_5Gal_9GlcNAc_9$, positions unknown			]-	palmitoyl	*lyso*-AG	[16,30]
T. congolense VSG	-	-	Galβ(1→6)GlcNAcβ(1→4)		-	-	-	di-AG	[31]
T. cruzi 1G7	Manα(1→2)	-	-	-	-	-	-	alk-AG/ceramide	[32,33]
T. cruzi Tc85	-	-	-	-	-	-	palmitoyl	alk-AG	[18,34]
T. cruzi mucins	Manα(1→2)	-	[	± Hexose		]P-OEtNH_2	-	alk-AG/ceramide	[35,36]
Leishmania PSP	-	-	-	-	-	-	-	alk-AG	[37,38]
T. gondii gp23	-	-	±Glcα(1→4)GalNAcα(1→4)		-	-	-	di-AG	[39,40]
N. caninum	-	-	-	-	-	-	acyl	n.d.	[41]
P. falciparum	Manα(1→2)	-	-	-	-	-	±myristoyl	di-AG	[19]
P. C. chabaudi	Manα(1→2)	-	-	-	-	-	±acyl	di-AG	[42]
Paramecium	-	-	[	Manα-PO_4		]-	-	ceramide	[43]
S. cerevisiae	±Manα(1→2)[±Manα(1→3)]Manα(1→2)			-	-	-	-	ceramide/di-AG	[44]
D. discoideum PsA	±Manα(1→2)	[	1–2 P-OEtNH_2			]-	-	ceramide	[45]
Torpedo AChE	Glcα(1→2)	±P-OEtNH_2	±GalNAcβ(1→4)	-	P-OEtNH_2	-	-	di-AG	[46,47]
Rat Thy-1	Manα(1→2)	-	GalNAcβ(1→4)	-	P-OEtNH_2	-	-	alk-AG	[9]
Hamster Prp	±Manα(1→2)	[	±NeuNAc±Gal-GalNAc-**?**		]P-OEtNH_2	-	-	n.d.	[48]
Mouse NCAM	±Manα(1→2)	n.d.	±GalNAcβ(1→4)	-	n.d.	-	-	n.d.	[49]
Bovine 5′-NT	±Manα(1→2)	±P-OEtNH_2	[	±HexNAc-**?**/P-OEtNH_2		]-	-	n.d.	[50]
Porcine MDP	-	±P-OEtNH_2	±Galβ(1→3)GalNAcβ(1→4) ±NeuNAc±GalNAcβ(1→4)		P-OEtNH_2	-	-	di-AG	[51]

Table 1 (continued)

Anchor	R6	R5	R4	R3	R2	R1	Inositol-Acyl	Lipid	Ref
Human MDP	±Manα(1→2)	n.d.	±Galβ(1→3)GalNAcβ(1→4)		n.d.	-	-	n.d.	[51]
Human AChE	-	±P-OEtNH$_2$	[	P-OEtNH$_2$		] -	palmitoyl	alk-AG	[13,17,52]
Human PLAP	-	[		P-OEtNH$_2$			] -	alk-AG	[53]
Human CD52	±Manα(1→2)	±P-OEtNH$_2$	-	-	P-OEtNH$_2$	-	±palmitoyl	di-AG	[20]
Human CD59	±Manα(1→2)	-	±GalNAcβ(1→4)	-	P-OEtNH$_2$	-	-	n.d.	[54,55]
Human FBP	-	-	-	-	-	-	±acyl	di-AG	[21,56]

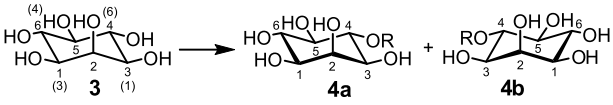
Notes: All membrane protein/glyocprotein GPI anchors contain the conserved structure shown above with various substituents (R^1-R^6) and lipids as noted. Brackets denote residues for which it has not been determined to which sugar the residue is linked. Question marks denote residues for which the position and orientation (α/β) of the linkage to the specified sugar have not been determined. A ± symbol was used where the associated residue is found only on a proportion of the structures. Other abbreviations include: n. d., complete structure not determined at a specified position; AG, acyl glycerol; alk-AG, alkyl acyl glycerol; NeuNAc, *N*-acetyl neuraminic acid (sialic acid); VSG, variant surface glycoprotein; PARP, procyclic acidic repetitive protein; PSP, promastigote surface protease; PsA, prespore-specific antigen; AChE, acetylcholinesterase; Prp, prion protein; FBP, folate binding protein

To date, more than 20 GPI anchors of extracellular proteins and glycoproteins have been characterized, all of which retain the general core structure described above. Some of these GPI anchors are depicted in *Table 1*. As shown, the structures of GPI anchors are highly diversified, and their structural diversity arises mainly from the variation in the side chains and other functional groups attached to the glycan core of GPIs, as well as the long lipid chains attached to the inositol residue. As a result, it is very difficult to obtain homogeneous GPI anchors and GPI-linked proteins and glycoproteins from nature, while the demand for these molecules from various studies has been ever growing. Consequently, the chemical synthesis of homogeneous and well-defined GPIs and related structures has become a very hot topic in carbohydrate research.

Thus far, a number of GPI anchors have been chemically synthesized [26,27], and there are many more reports about the chemical synthesis of related structures such as various inositol derivatives [28] and GPI fragments. For a short review like this, it is impossible to cover the literature completely. Therefore, this chapter will briefly summarize the recent developments in this field by especially focusing on the synthetic studies of intact GPI anchors, in their natural and unnatural forms, and GPI-linked glycoconjugates, as well as a few representative syntheses of some complex GPI core glycans by unique strategies such as solid-phase synthesis.

2 General Issues in GPI Synthesis

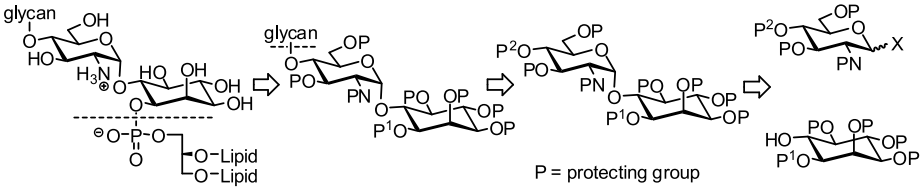
Consisting of carbohydrates, lipids, inositol, and often several phosphate moieties, the structure of GPI anchors is complex. Therefore, the chemical synthesis of GPIs denotes a significant challenge [27]. GPI anchors share some common structural features that deserve special consideration. First, all GPIs contain a *myo*-inositol residue **3** (*Scheme 1*) which, when free, is symmetric; thus, its positions 1 and 4 are chemically identical to positions 3 and 6, respectively. However, substitution at any of these sites will produce a racemic mixture, and commonly, *myo*-inositol in natural products, such as GPIs, are optically active. Consequently, the chemical synthesis of GPIs usually starts with the preparation of enantiomerically pure and 1,6-*O*- or 1,2,6-*O*-differentially protected *myo*-inositol derivatives, which itself poses an interesting problem in GPI synthesis [28]. Numerous strategies based on resolution of racemic mixtures or based on de novo synthesis [27,28] have been developed to deal with this problem, and some of these strategies will be discussed in the presentation of specific GPI anchor syntheses.



Scheme 1
The structure of *myo*-inositol and its optically active derivatives

Second, the glucosamine residue in the GPI core is linked to the inositol residue through an α-glycosidic linkage which is relatively difficult to achieve chemically. To facilitate the desired α-glycosylation during the coupling reaction between glucosamine and inositol, the amino group of a glucosaminosyl donor must be masked by a non-participating protecting group. For this purpose, most GPI syntheses employ derivatives of 2-azido-2-deoxy-D-glucose as donors,

because an azido group, which can be readily converted to a free amino group later on, may serve dually as a "latent amino group" and as a non-participating protecting group. Other non-participating protecting groups have also been used as well. Still, this glycosylation reaction often gives relatively low stereoselectivity or poor yield; therefore, it is usually realized at the initial stage of GPI synthesis (◉ *Scheme 2*).



◘ Scheme 2
Generally adopted retrosynthetic scheme for GPI syntheses

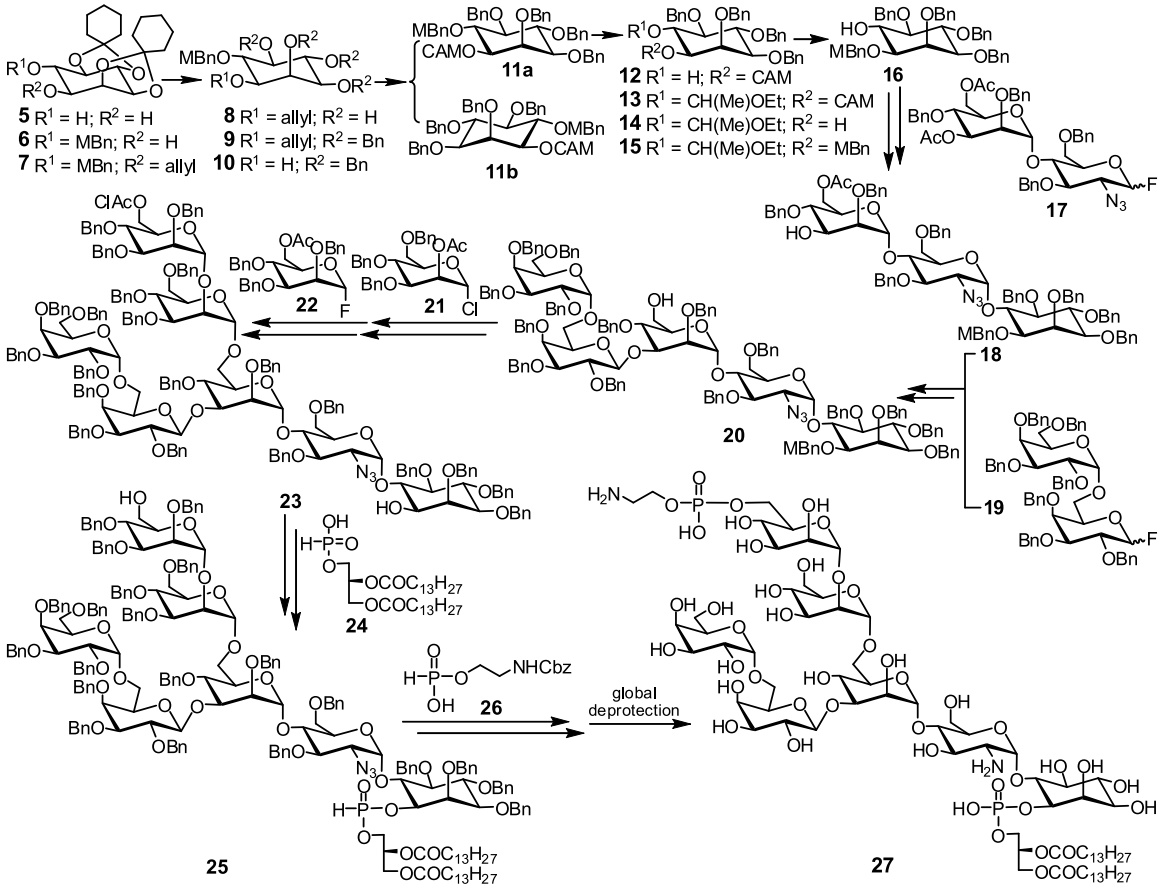
Third, including the phospholipid, all GPI anchors contain two or more phosphate moieties. Because of the concern of potential chemical incompatibility, the vast majority of reported GPI syntheses have adopted a strategy in which the phosphate functionalities are introduced at the final stage of the synthesis (◉ *Scheme 2*).

3 Chemical Synthesis of GPIs

3.1 Synthesis of the *T. brucei* GPI Anchor by Ogawa and Coworkers

Ogawa and coworkers [57,58,59,60,61] achieved the first total synthesis of a natural GPI anchor, i. e., the *T. brucei* GPI anchor, in 1991. They began their synthesis with the preparation of an optically active inositol derivative starting from the racemic diol **5** [62], as shown in ◉ *Scheme 3*. Diol **5** was treated first with dibutyltin oxide and then with 4-methoxybenzyl (MBn) chloride to give **6** (60%) which was readily converted to a tetraol **8** upon allylation and hydrolysis. After benzylation of the resulting tetraol, the allyl group was selectively removed to afford racemic **10**. To resolve the two enantiomers of **10**, it was transformed into diastereomeric (−)-ω-camphanates **11a** and **11b**, which were separable by column chromatography. After a series of manipulations of the protecting groups in **11a**, the desired alcohol **16** was obtained. In **16**, the 6-*O*-position was free and the 1-*O*-position was protected by an MBn group to differentiate them from other positions of inositol.

What was unique concerning this synthesis was that a disaccharide **17**, instead of a glucosamine derivative, was employed as a donor for the glycosylation of **16**. The reaction of glycosyl fluoride **17** and inositol derivative **16** was carried out in the presence of zirconocene dichloride and silver perchlorate in dry ether [63] to give the α- to β-pseudotrisaccharide with a 3.7 : 1 ratio. This was followed by deacetylation and regioselective acylation to give **18**. Subsequently, a galactobiose was introduced to **18** by the same glycosylation method to give the desired pentasaccharide (68%) along with 8% of its β anomer, which was followed by deacetylation. Further elongation of the resulting glycosyl acceptor **20** was realized by linear assembly of monosaccharides **21** and **22**. These glycosylation reactions were promoted by mercury(II)



Scheme 3
Ogawa's synthesis of the *T. brucei* GPI anchor

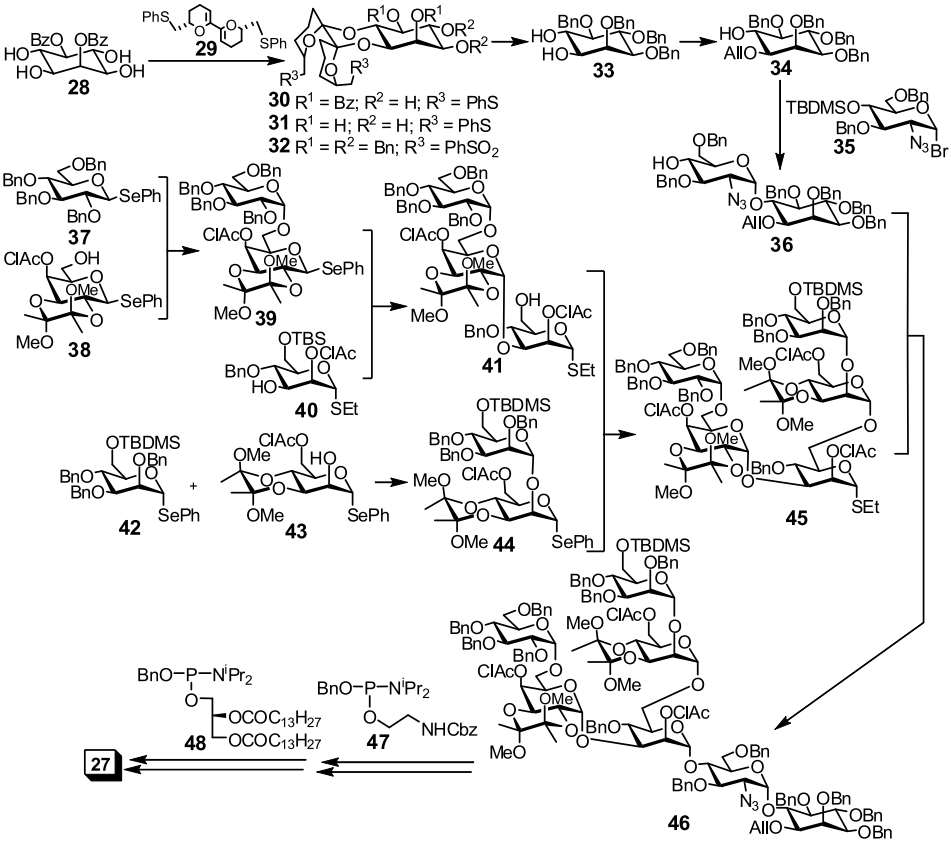
bromide/mercury(II) cyanide and zirconocene dichloride/silver perchlorate, respectively. Protecting group manipulation of the product eventually afforded **23** which was ready for introduction of the phosphate groups.

First, the phospholipid moiety was introduced to the 1-*O*-position of inositol by the phosphonate approach developed by Lindh and coworkers [64]. The reaction between **23** and **24** in the presence of pivaloyl chloride in pyridine gave the phosphitylation product in 64% yield. Next, the 2-chloroacetyl group (ClAc) on the non-reducing end mannose unit was selectively removed, and again the phosphonate approach was employed to attach a phosphityl group to the 6-*O*-position, which gave 40% yield of the desired product as a mixture of four diastereomers due to the presence of two chiral phosphorus atoms. It is worthy pointing out that the attempt to introduce a phosphodiester functionality to the non-reducing end mannose by the phosphorimidite method was unsuccessful. However, it was observed that, before phospholipidation of the inositol residue, phosphitylation of the non-reducing end mannose via the phosphorimidite methodology was successful, but removal of the MBn protection of the inositol group was a problem. Obviously, the functional groups on inositol could affect the reactivity of the non-reducing end mannose, and vice versa, even though these two residues seem well separated from each other. Finally, oxidation of the phosphitylation products with iodine (68%) followed by global deprotection by hydrogenolysis afforded the desired *T. brucei* GPI anchor **27**.

3.2 Synthesis of the *T. brucei* GPI Anchor by Ley and Coworkers

In their total synthesis of the *T. brucei* GPI anchor (◉ *Scheme 4*) [65,66], Ley and coworkers used a tetraol **28** [67] to prepare the asymmetric inositol derivatives. Its desymmetrization was realized by reacting with chiral bis(dihydropyran) **29** to offer a single diastereomer **30** (71%). Debenzoylation of **30** followed by benzylation and oxidation produced the sulfone **32**, of which the dispiroketal moiety was removed using lithium hexamethyldisilazide (LHMDS). Selective allylation of **33** by means of tin complex eventually gave the desired inositol derivative **34** (65%). Thereafter, **34** was glycosylated by bromide donor **35** using Lemieux's glycosylation protocol [68] to give the α-linked pseudodisaccharide with excellent selectivity, which was finally desilylated to yield the pseudodisaccharide acceptor **36**.

Further elongation of the carbohydrate chain of **36** was achieved in a highly convergent fashion (◉ *Scheme 4*). Therefore, after the properly protected monosaccharide units were obtained, they were assembled to form short oligosaccharide blocks, including a trisaccharide **41** and a disaccharide **44**. Then, **41** and **44** were coupled to afford a pentasaccharide **45**, which was finally linked to **36** to give the GPI backbone **46**. This convergent assembly was made possible largely owing to the successful use of selective glycosylation strategies based on "armed-disarmed" and semiorthogonal glycosyl donors. For example, the presence of a rigid cyclic 1,2-diketal group in **38** and **43** creates significant torsional strains in these structures, making them "disarmed" (less reactive); as a result, the armed glycosyl donors **37** and **42** could be selectively activated in the presence of **38** and **43**, even though all these substrates had the same phenylseleno group at their anomeric centers. On the other hand, because selenoglycosides are more reactive than thioglycosides, it was also possible to activate glycosyl donors **39** and **44** in the presence of **40** and **41**, respectively, to achieve their selective glycosylations. Meanwhile,

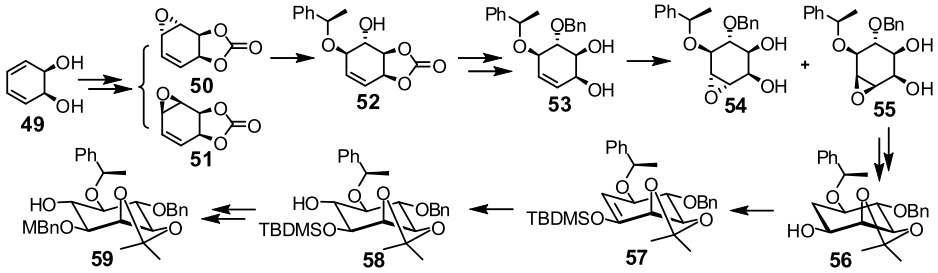


Scheme 4
Ley's synthesis of the *T. brucei* GPI anchor

thioglycoside **45** was activated by a stronger promoter to be directly used as a donor for glycosylation of the pseudodisaccharide **36**. By using selective activation for glycan assembly, transformations required to manipulate the anomeric centers in the traditional glycosylation methods were eliminated. Consequently, this synthesis could effectively save a few potentially necessary synthetic steps.

For installing the phosphate functionalities, the phosphorimidite method was employed. First, the *t*-butyldimethylsilyl (TBDMS) group on the non-reducing end mannose unit was selectively removed for the successful attachment of a P-OEtNH$_2$ group following tetrazole-promoted phosphitylation by phosphorimidite **47** and oxidation by *m*-chloroperbenzoic acid (*m*-CPBA). The phospholipid moiety was introduced to the inositol 1-*O*-postion by the same method using **48**. Finally, global deprotection was achieved in three steps to afford the *T. brucei* GPI anchor **27**.

Ley and coworkers [69,70,71] have also developed a de novo synthesis of optically pure derivatives of inositol starting from benzene (*Scheme 5*). Conversion of benzene to an optically

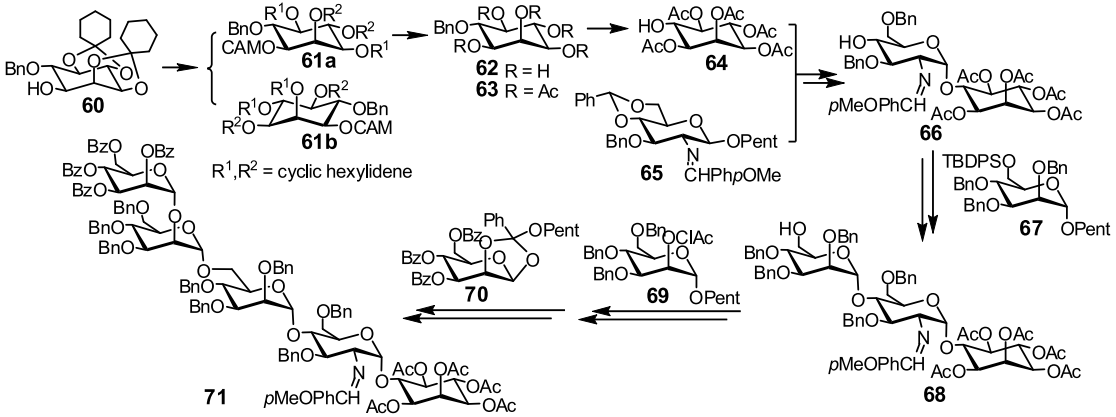


Scheme 5
Ley's alternative synthesis of the optically pure inositol derivative

active diol **49** by microbial oxidation was achieved with *P. putida*, and subsequent acylation gave a cyclic carbonate which was stereoselectively epoxidized by *m*-CPBA to form the α- and β-epoxides **50** and **51** in a 4.6 : 1 ratio. Epoxide ring opening of **50** by 1-phenyl ethanol with camphorsulfonic acid as catalyst proved to be regioselective, which was followed by benzylation and deacylation to produce **53**. Epoxidation of **53** using *m*-CPBA afforded **54** and **55** in a 1 : 9 ratio. Separation of the epoxides was more easily achievable after their conversion to acetonides, and then, the desired product was subjected to regioselective ring opening by lithium aluminum hydride to yield **56** (81%). Further oxidization with tetrapropylammonium perruthenate (TPAP) and subsequent treatment with TBDMS chloride gave a silyl enol ether **57**. Hydroboration of **57** turned out to be regio- and stereoselective to afford the differentially protected *myo*-inositol derivative **58**. An interesting feature of this reaction was the silyl migration from C1 to C6, occurring prior to *p*-methoxybenzylation of C-1. Further manipulation of protecting groups gave **60** that was used to construct a partial structure of the *T. brucei* GPI anchor, following a strategy similar to that discussed above.

3.3 Synthesis of the GPI Anchors of Rat Brain Thy-1 and *P. falciparum* by Fraser-Reid and Coworkers

In their synthesis of the glycan core of GPIs, Fraser-Reid and coworkers [72] developed a strategy to synthesize optically pure inositol derivatives (*Scheme 6*), which was similar to that of Ogawa and coworkers [57,58,59,60,61] in that both started from **2** and both used camphanic acid to resolve the enantiomers. However, after regioselective benzylation via phase transfer catalysis to get racemic **60**, Fraser-Reid and coworkers [73] immediately resolved it into **61a** and **61b** [74]. Following further manipulation of the protecting groups, an optically pure inositol derivative **64** was obtained, which only had the 6-*O*-position differentiated. Evidently, **64** was only suitable to synthesize the GPI glycan core but not the whole GPI, since regioselective 1-*O*-phospholipidation of the inositol residue was impossible. Compound **64** was subsequently glycosylated by **65**, and the product was subjected to reductive ring opening by sodium cyanoborohydride to produce pseudodisaccharide **66** in a good overall yield. In **65**, the glucosamine amino group was protected with the non-participating *p*-methoxybenzaldehyde to promote α-glycosylation. The carbohydrate chain was further elongated by sequential assem-



Scheme 6
Fraser-Reid's synthesis of the core glycan of GPI anchors

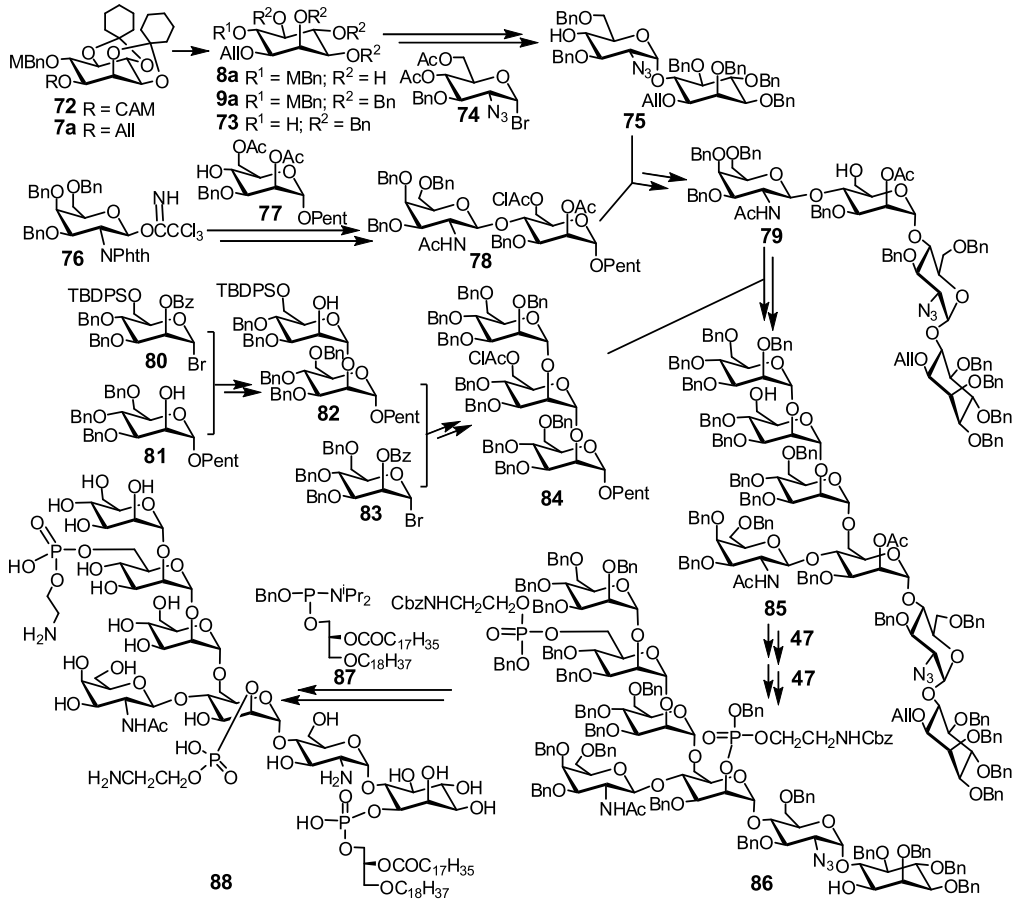
bly of monosaccharides **67**, **69**, and **70** to yield the GPI core glycan. In this synthesis, all glycosidic linkages were constructed by their own glycosylation method using *n*-pentenyl glycosides or *n*-pentenyl orthoesters as donors. It is interesting to point out that an attempt to prepare the same structure by direct coupling of a disaccharide unit to **68** was unsuccessful.

In their total synthesis of a fully phosphorylated GPI anchor of rat brain Thy-1, Fraser-Reid and coworkers [75,76,77,78] developed a process to prepare optically pure inositol derivatives with 1-*O*- and 6-*O*-positions differently protected (*Scheme 7*). Therefore, after **2** was *p*-methoxybenzylated to get **6** along with its regioisomer, the two enantiomers of **6** were resolved as camphanates to obtain the desired diastereomer **72**. Then, the camphanoic group in **72** was replaced with an allyl group, and the optically pure product **7a** was treated with acid to remove the ketals, followed by benzylation. The regioisomer of **6** was also treated by a similar procedure and converted to **9a**. Thereafter, MBn was removed under acidic conditions, and the resultant **73** was coupled to the azido sugar **74** using silver perchlorate as a promoter [79] to afford the α-product in a 63% yield along with 22% of the β-anomer. Further manipulation of the protecting groups eventually afforded the pseudodisaccharide **75** as a glycosyl acceptor.

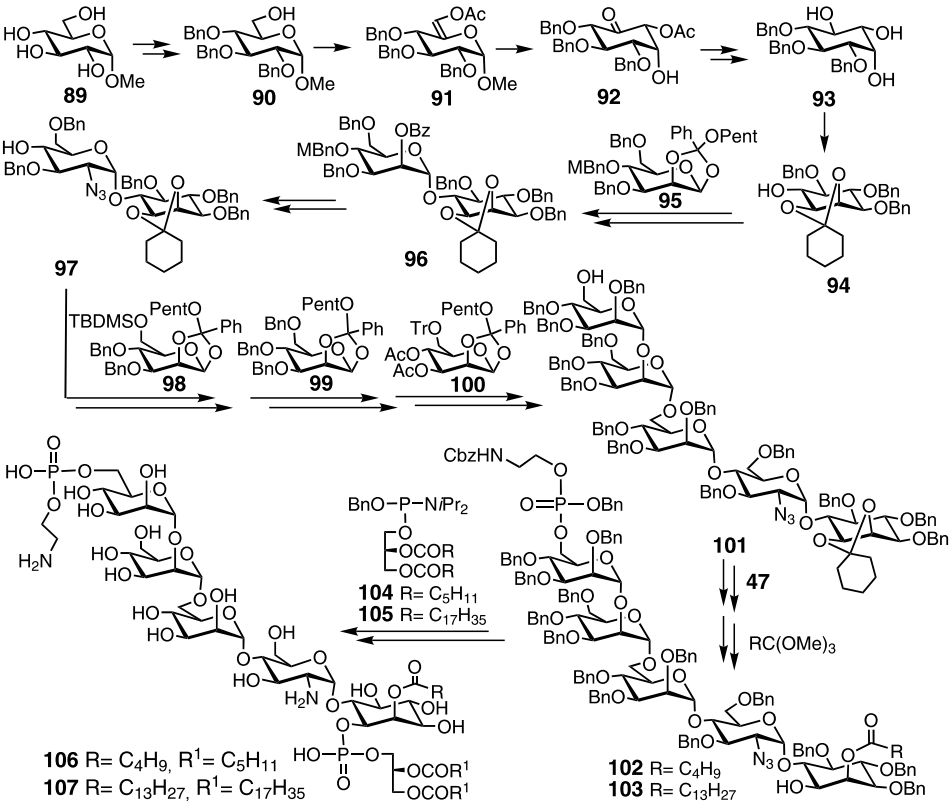
A highly convergent strategy was employed to assemble the desired GPI (*Scheme 7*). After **75** was available, a disaccharide **78** and a trisaccharide **84** were prepared from properly protected and activated monosaccharides via semiorthogonal glycosylations. Then, **78** was coupled to **75** first with *N*-iodosuccinimide (NIS) as a promoter, which was followed by coupling of the resultant tetrasaccharide **79** to **84** to yield the GPI glycan **85**. The dechloroacetylation steps were achieved with thiourea. Next, two phosphoethanolamine functionalities were separately introduced to Man-I and Man-III by the phosphorimidite method described above. The phospholipid moiety was introduced to **86** also by the phosphorimidite method to finally afford a mixture of all eight stereoisomers of the fully protected GPI. Global deprotection was performed by hydrogenolysis over palladium hydroxide to give the GPI anchor **88** of rat brain Thy-1.

Fraser-Reid and coworkers [80] also developed another method to prepare optically pure inositol derivatives using methyl α-D-glucopyranoside **89** as the starting material (*Scheme 8*). After **90** was obtained from **89** via several steps of protecting group manipulation, it was converted to an enol acetate **91**. Next, stereoselective Ferrier reaction of **91** afforded the cyclic hexanone derivative **92** (63%). Eventually, the inositol derivative **93** was obtained after chelation-mediated reduction and subsequent methanolysis to remove the acetyl group. Regioselective 1,2-*O*-protection of the triol **93** by a cyclohexylidene group was possible to afford the glycosyl acceptor **94** which was used to synthesize the *P. falciparum* GPI anchor [81].

In their synthesis of *P. falciparum* GPI anchor, Fraser-Reid and coworkers [81,82] utilized a unique strategy to achieve the pseudodisaccharide. As shown in *Scheme 8*, a mannose donor **95**, instead of a glucosamine derivative, was employed to glycosidate **94** to give a benzoylated pseudodisaccharide **96** in an excellent yield (99%). After the benzoyl group was replaced with a triflic group, Deshong's azide displacement method [83] was used to transform the mannose unit into an azido sugar, which served as a latent glucosamine unit, and give pseudodisaccharide **97**. Further elongation of the GPI glycan was realized by linear assembly of monosaccharides **98**, **99**, and **100**. After the GPI glycan was constructed, the trityl group on the nonreducing end mannose residue was removed to give **101**, to which a phosphoethanolamine group was introduced by the phosphorimidite method. This was followed by acidic removal of the cyclohexylidene group of the inositol moiety and regioselective 2-*O*-acylation of the



Scheme 7
Fraser-Reid's synthesis of the rat brain Thy-1 GPI anchor

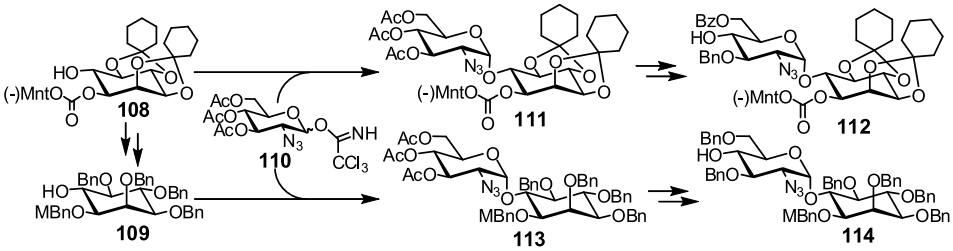


Scheme 8
Fraser-Reid's synthesis of the *P. falciparum* GPI anchor

resultant diol based on stereoelectronic ring opening of a cyclic orthoester developed by King and Allbutt [84]. Accordingly, treatment of the diol with trimethyl orthoesters produced the corresponding cyclic orthoesters that were treated with ytterbium(III) triflate to afford **102** and **103** as major products. Phospholipidation of the free inositol hydroxyl group was achieved through reaction with **104** or **105** followed by one pot oxidation using *m*-CPBA. Hydrogenolysis over Pearlman's catalyst finally gave the fully deprotected *P. falciparum* GPI anchors **106** and **107**. An important conclusion about 2-*O*-acylated derivatives of GPIs was that no sign of acyl migration was observed over a period of several months based on ^{1}H NMR studies.

3.4 Synthesis of the GPI Anchors of Rat Brain Thy-1, *T. gondii*, and *S. cerevisiae* by Schmidt and Coworkers

As shown in *Scheme 9*, two types of inositol derivatives were prepared and used by Schmidt and coworkers [85,86,87,88,89] in their synthesis of several GPI anchors. The synthesis of *S. cerevisiae* GPI started with optically pure **108** [86,90], while the syntheses of Thy-1 [87,88]

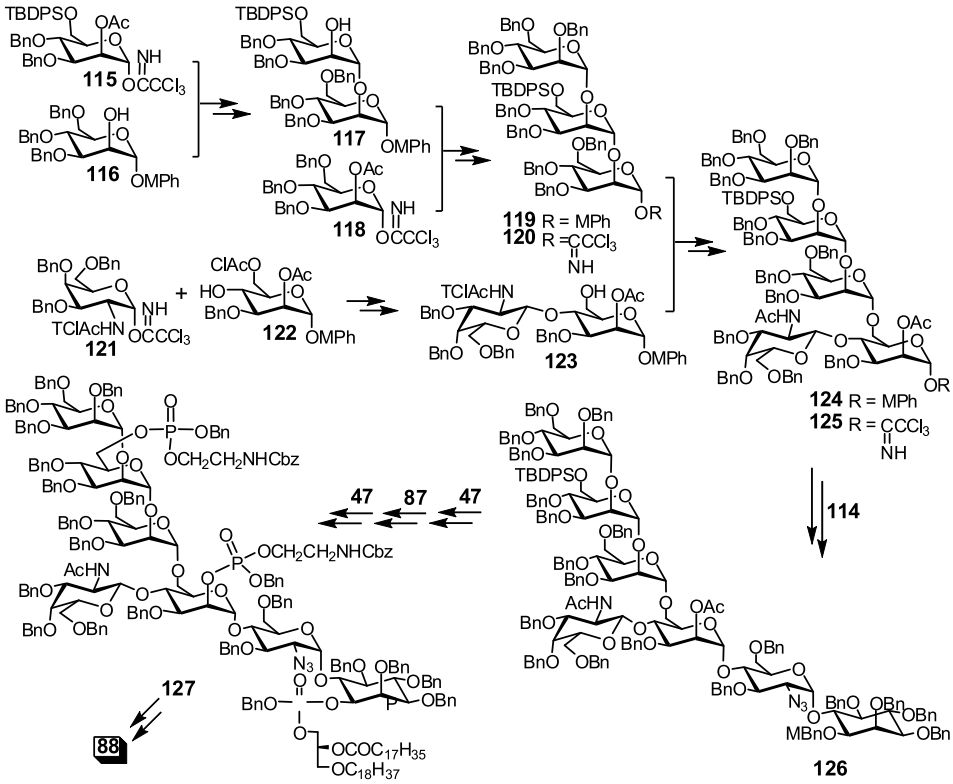


Scheme 9
Schmidt's synthesis of the pseudodisaccharides as glycosyl acceptors

and *T. gondii* [89] GPIs started with **109**. Treating **2** [62] with bis(tributyltin)oxide followed by (−)-menthyl chloroformate led to **108** and its diastereomer that were easily separable by crystallization in petroleum ether [91]. Compound **108** was then converted to **109** following a series of transformations including replacement of the menthyloxycarbonyl group with an MBn group, acid-catalyzed hydrolysis of the cyclohexylidene groups, benzylation of the exposed hydroxyl groups, and finally, selective 6-de-*O*-allylation. The trichloroimidate glycosylation method using **110** as a glycosyl donor was used to introduce the α-linked glucosamine unit to **108** and **109**. The reaction of **108** and **110** afforded **111** exclusively, which, after protecting group manipulation, was converted to **112** that was used to synthesize *S. cerevisiae* GPI [85,86]. On the other hand, the reaction of **109** and **110** gave **113** exclusively, which was converted to **114** following deacetylation, 4,6-*O*-benzylidenation, 3-*O*-benzylation, and then regioselective opening of the benzylidene ring. Pseudodisaccharide **114** was used to synthesize rat brain Thy-1 [87,88] and *T. gondii* [89] GPI anchors.

The heptaosyl core of rat brain Thy-1 GPI was constructed by a highly convergent approach (*Scheme 10*) [87], beginning with the preparation of the trimannose **120** and disaccharide **123** blocks. So, after monosaccharides **115**, **116** and **118** were obtained, they were assembled to form **119** by the trichloroimidate glycosylation method. Then, the reducing end of **119** was selectively deprotected by oxidative removal of the *p*-methoxyphenyl group followed by reaction with trichloroacetonitrile to form trichloroacetimidate **120** that was used as a glycosyl donor. Meanwhile, glycosylation of **121** by **122** and then selective removal of ClAc group by thiourea afforded the glycosyl acceptor **123**. Next, **120** and **123** were coupled to give α- and β-linked pentasaccharides in a 94% yield and 9 : 4 ratio. The *N*-trichloroacetyl group (TClAc) in the desired α-pentasaccharide was then changed to the acetyl group by treatment with tributyltin hydride and azoisobutyronitrile (AIBN) [92] to give **124** that was converted to trichloroacetimidate **125** as described above. The pentasaccharide was coupled to pseudodisaccharide **114** to give the glycan core **126**. Thereafter, the 2-*O*-, 1-*O*-, and 6-*O*-positions of Man-I, inositol, and Man-III were sequentially deprotected and phosphorylated using **47**, **87**, and **47**, respectively, to give the fully protected GPI anchor **134**. Finally, **134** was deprotected to yield the free GPI anchor **88** of rat brain Thy-1.

In another synthesis of the rat brain Thy-1 GPI anchor, Schmidt and coworkers [88] have adopted the linear strategy to construct the pentasaccharide building block. As shown in *Scheme 11*, after all monosaccharide units were obtained, they were assembled sequentially by the trichloroimidate glycosylation method to form pentasaccharide **130**. It was manipulated

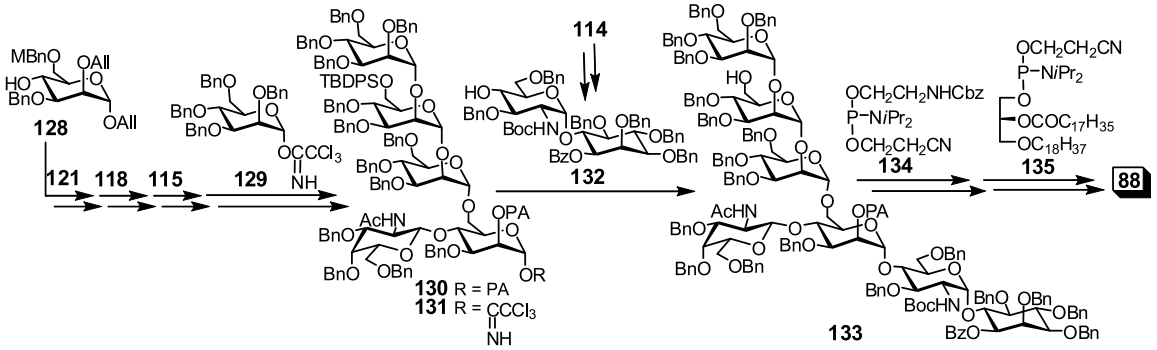


Scheme 10
Schmidt's synthesis of the rat brain Thy-1 GPI anchor

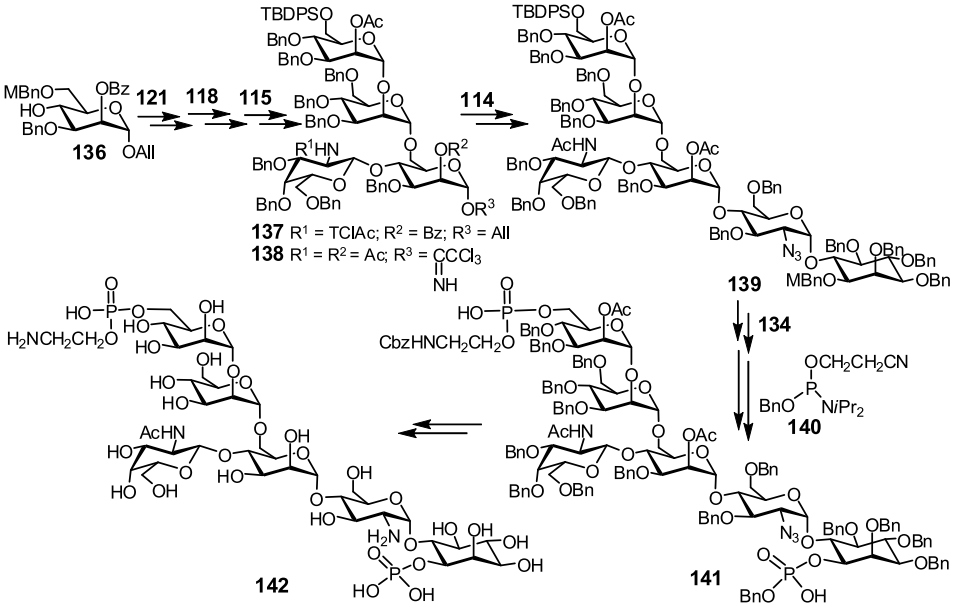
by the same procedure described above to give trichloroimidate donor **131** and then coupled to **132** that was derived from **114** to subsequently afford the core heptasaccharide **133**. The phosphate moieties were installed by the phosphorimidite method using **134** and **135** as substrates. In this synthesis, the cyanoethyl group was employed as a protecting group of phosphorimidites and phosphates. The cyanoethyl group could be removed concomitantly during the deprotection of the phenoxyacetyl and benzoyl groups. Global deprotection of the protected GPI anchor in two steps eventually produced the free GPI anchor **88** of rat brain Thy-1.

By similar strategies, Schmidt and coworkers have also prepared a partial structure of the GPI anchor of *T. gondii* (*Scheme 12*) [89]. First, the trichloroacetimidate **138** of a tetrasaccharide as the glycosyl donor was prepared and utilized to glycosidate pseudodisaccharide **114**. The resultant hexasaccharide **139** was then selectively deprotected and phosphorylated at Man III and inositol respectively to yield **141**. Quantitative deacylation followed by debenzylation over Pearlman's catalyst gave the deprotected GPI anchor **142** in a 92% yield.

Furthermore, Schmidt and coworkers [86,90] also achieved the *S. cerevisiae* GPI anchor (*Scheme 13*). After a tetramannose block **144** was assembled from monosaccharides **143**,



Scheme 11
Schmidt's alternative synthesis of the rat brain Thy-1 GPI anchor

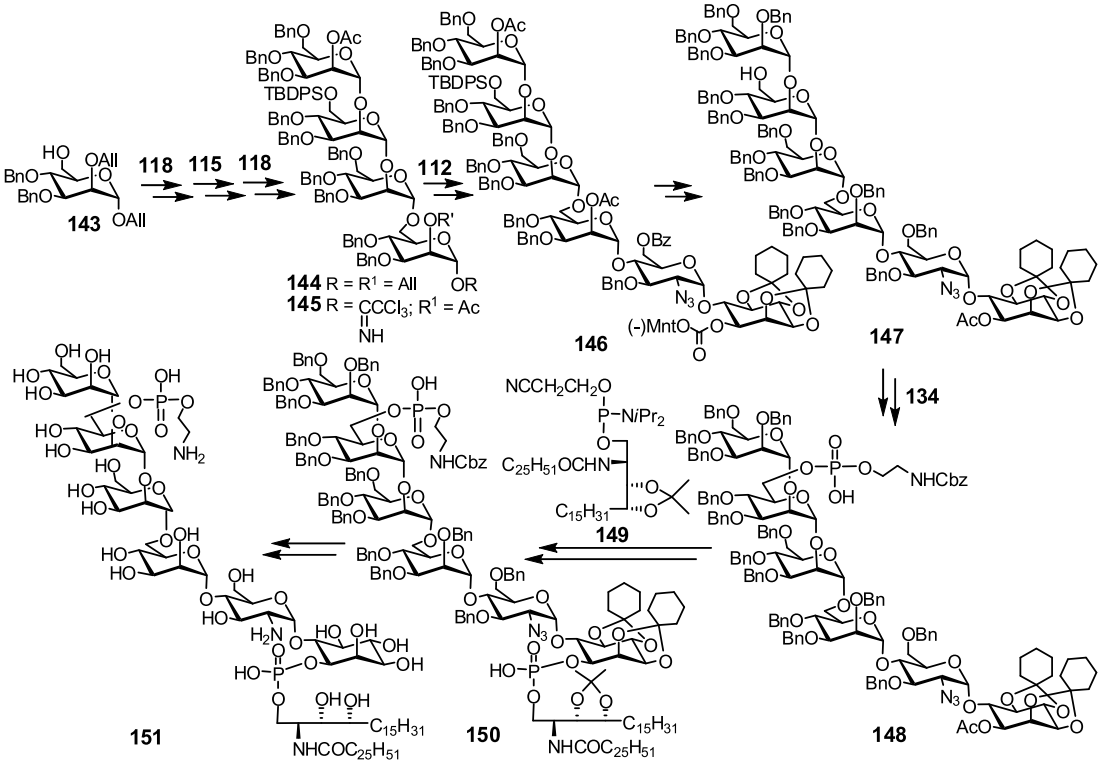


Scheme 12
Schmidt's synthesis of the *T. gondii* GPI anchor

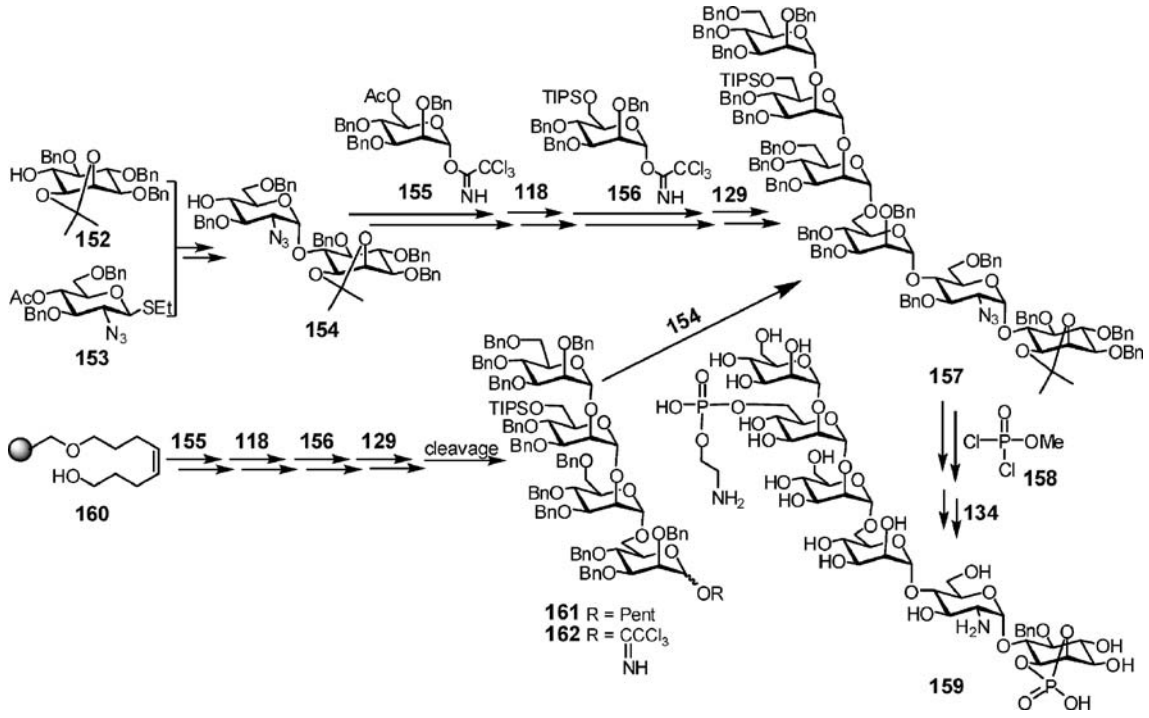
115, and **118** in a linear fashion, it was converted to the glycosyl donor **145** following a series of protecting group manipulation and trichloroacetimidation. Then, **145** was coupled to the pseudodisaccharide **112** to yield the hexasaccharide core **146** of *S. cerevisiae* GPI anchor. Further modification of protecting groups afforded **147** which was ready for installing the phosphate moieties. The phosphate attached to 6-*O*-position of Man III was introduced first using **134**, which was followed by phospholipidation using **149** to give the fully protected GPI anchor **150**. Finally, all protecting groups were removed to produce the free GPI anchor **151** of *S. cerevisiae*.

3.5 Synthesis of the GPI Anchor of *P. falciparum* by Seeberger and Coworkers

Seeberger and coworkers have achieved the non-lipidated GPI anchor of *P. falciparum* by both solution phase [93] and solid phase [94] synthesis (*Scheme 14*). Their syntheses started from the optically pure inositol derivative **152** which was prepared according to a method reported by Fraser-Reid and coworkers [80,81]. After **152** was coupled to an azido sugar by the thioglycoside glycosylation method to get pseudodisaccharide **154**, further elongation of its carbohydrate chain was realized by either linear assembly or coupling of a tetrasaccharide block prepared by solid-phase synthesis. In the first case, monosaccharide units were installed through a series of glycosylation reactions performed in solution to the give fully protected pseudohexasaccharide **157**. In the later case, a tetrasaccharide was first assembled on the Mer-



Scheme 13
Schmidt's synthesis of the *S. cerevisiae* GPI anchor

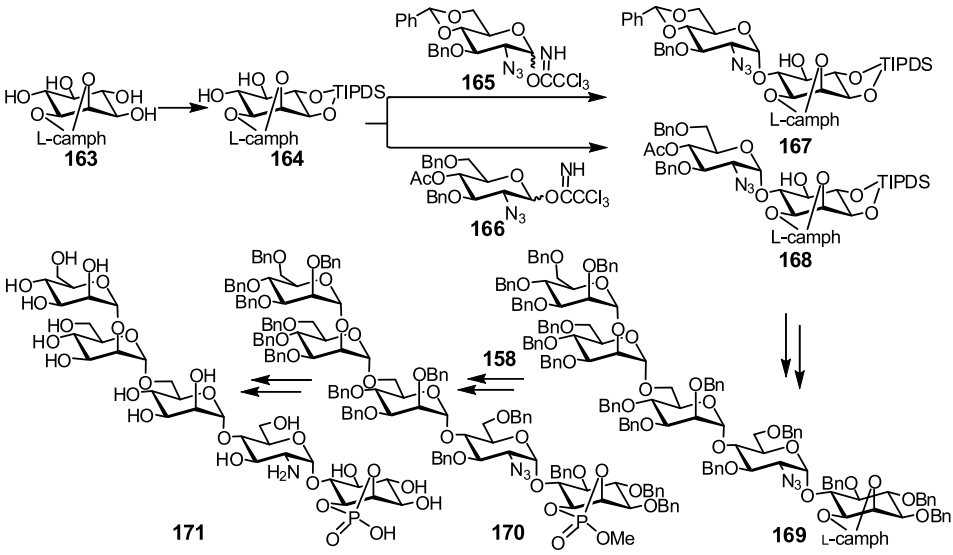


Scheme 14
Seeberger's synthesis of the *P. falciparum* GPI anchor

rifield resin with octenediol as a linker [95]. After the resin was treated with Grubb's catalyst in an ethylene atmosphere, the *n*-pentenyl glycoside **161** was obtained. The attempt to directly couple **161** to **154** failed; therefore, **161** was converted to trichloroacetimidate **162**. Glycosylation of **154** with **162** as a donor produced the desired GPI glycan core **157** in a 32% yield. Thereafter, the isopropylidene group on the inositol residue was removed, and to the exposed hydroxyl groups was introduced a cyclic phosphate using **158**. A phosphoethanolamine group was finally attached to the nonreducing end mannose, which was followed by global deprotection to afford the non-lipidated GPI anchor **159** of *P. falciparum*. The product was linked to a carrier protein to form a promising glycoconjugate vaccine [93].

3.6 Synthesis of the Core Glycan of GPI Anchors by Martin-Lomas and Coworkers

Martin-Lomas and coworkers [96] desymmetrized *myo*-inositol according to a procedure originally reported by Bruzik and Tsai [97]. Reaction of *myo*-inositol with L-camphor dimethyl ketal followed by resolution of the diastereomeric products gave optically pure ketal **163** in a 39% yield. Treatment of **163** with 1,3-dichloro-1,1,3,3-tetraisopropyldisiloxane ($TIPDSCl_2$) gave the diol **164** (93%), the 6-*O*-position of which could be regioselectively glycosylated to afford desired pseudodisaccharides (*Scheme 15*) [98]. Glycosylation of **164** with **165** as a donor offered good stereoselectivity (α/β 5.5 : 1) but poor regioselectivity, as 16% of the 5-*O*-glycosylation product was formed. The glycosylation of **164** with **166** as a donor gave the pseudodisaccharides in a 75% yield as a 3 : 2 mixture of α- and β-anomers. The latter synthetic route was more practical due to its higher yield and easy separation of the anomers. After iso-



Scheme 15
Martin-Lomas's solution-phase synthesis of the glycan core of GPI anchors

lation and then deacetylation of the desired product **168**, its carbohydrate chain was elongated to get **169**. Thereafter, the L-camphor ketal was selectively removed, which was followed by phosphorylation to form a cyclic phosphate. Global deprotection with concomitant reduction of the azido group eventually afforded the monophosphorylated GPI glycan **171**.

Martin-Lomas and coworkers [99] have also studied solid-phase synthesis of the glycan core of GPI anchors (*Scheme 16*). However, the pseudodisaccharide was prepared first in solution and then attached to the resin. This can avoid solid-phase formation of the problematic glycosidic linkage, so as to increase the efficiency of solid-phase synthesis and the product purification. The carbohydrate chain was elongated through sequential introduction of monosaccharide units **173**, **174** and **175** by the trichloroimidate glycosylation method. The 9-fluorenylmethoxycarbonyl (Fmoc) group which can be easily removed under moderate basic conditions was utilized as the temporary protection of the involved hydroxyl groups. The resultant glycan core **176** was finally released from the resin and isolated in a 20% overall yield.

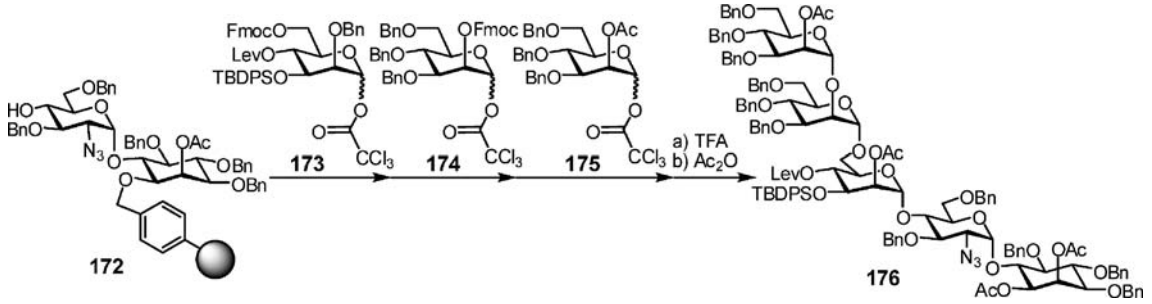
3.7 Synthesis of the *T. cruzi* GPI Anchor by Vishwakarma and Coworkers

An important feature of the synthesis of *T. cruzi* GPI reported by Vishwakarma and coworkers (*Scheme 17*) [100] was that the racemic mixture of an inositol derivative **177** was employed to prepare the desired pseudodisaccharide. After glycosylation with **110** as a glycosyl donor, a mixture of diastereomeric pseudodisaccharides **178** (1 : 1) was obtained. These diastereomers were modified and then separated by silica gel column chromatography to produce the desired isomer **114**. In this way, the process of resolving a racemic mixture was avoided. The carbohydrate chain elongation of **114** was achieved in a highly convergent fashion. Therefore, two disaccharides **181** and **183** were prepared first and then coupled to form the tetramannose segment **184**. Next, **184** was coupled to pseudodisaccharide **114** to yield the GPI glycan core **185** of *T. cruzi* GPI anchor. Thereafter, the 6-*O*-position of Man III was deprotected and phosphorylated, which was followed by deprotection of inositol 1-*O*-position and phospholipidation. Global deprotection and concomitant reduction of the azido group was realized by hydrogenation to eventually afford the *T. cruzi* GPI anchor containing saturated lipid chains.

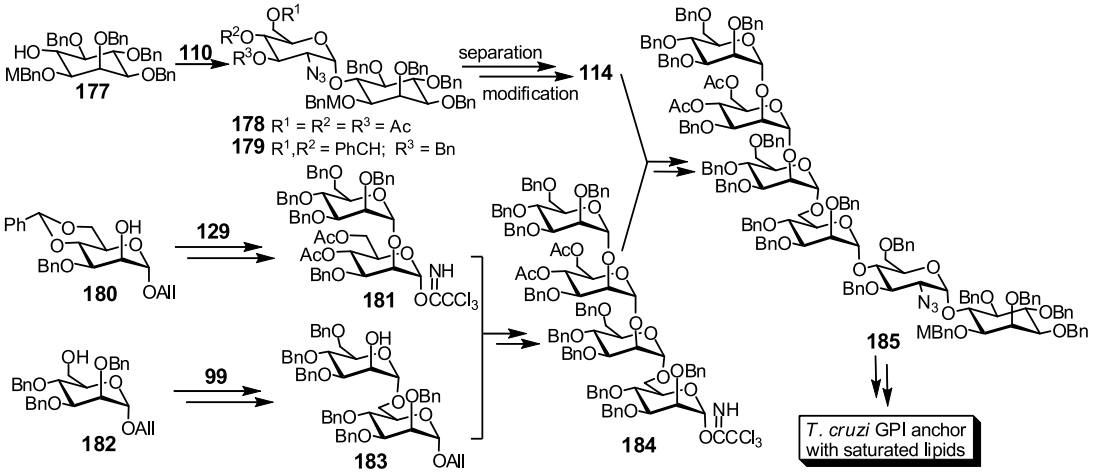
3.8 Synthesis of the *T. cruzi* GPI Anchor by Nikolaev and Coworkers

One of the most outstanding features of the GPI synthesis reported by Nikolaev and coworkers (*Scheme 18*) [101] was that the benzoyl group (Bz), instead of the benzyl group (Bn), was utilized as the permanent protecting group. Since Bz can be selectively removed under basic conditions at the final stage, this allowed the presence of unsaturated lipid chains in the synthetic GPI anchor, which was not easily achievable when Bn was employed as the permanent protecting group in most other syntheses.

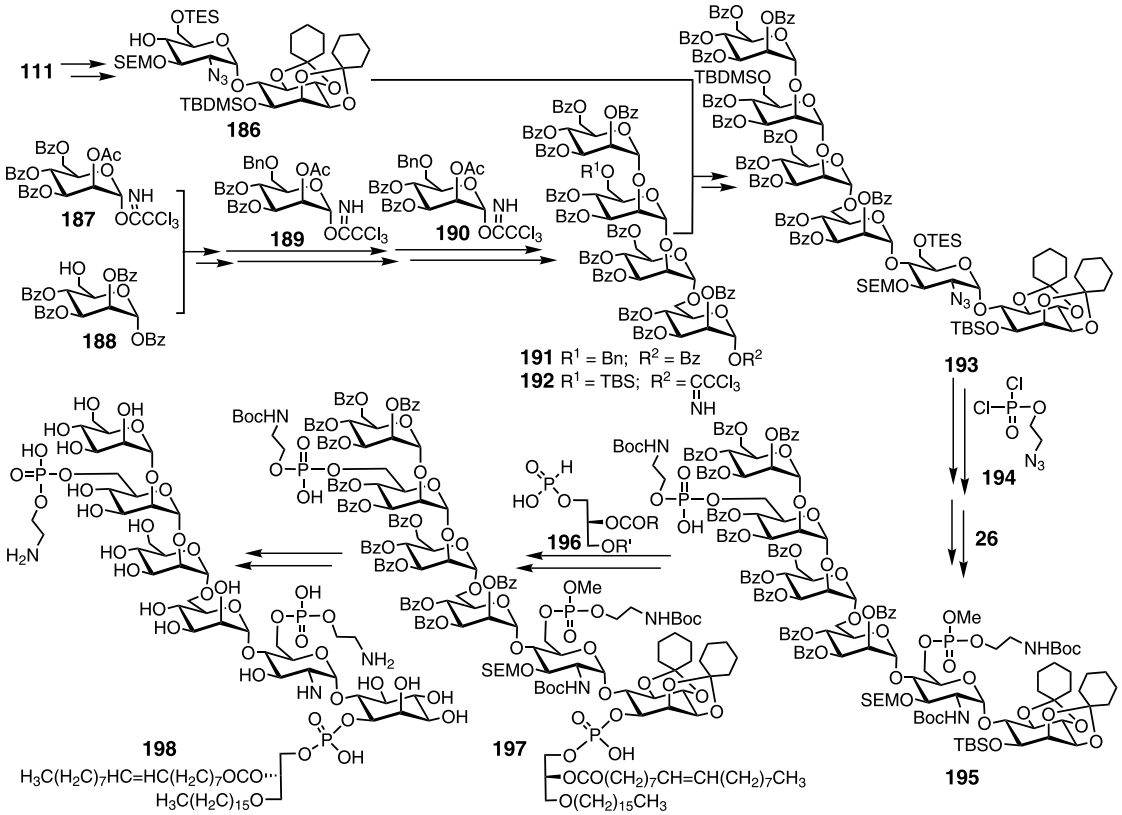
For pseudodisaccharide synthesis, Nikolaev and coworkers [101] adopted a strategy similar to that of Schmidt and coworkers [86,90], using the diastereomeric (−)-menthyl carbonates of inositol as the starting materials. However, instead of separating the diastereomeric mixture at the initial stage and using optically pure inositol derivative for glycosylation, both **108** and its diastereomer were glycosylated by **110**, and the diastereomeric pseudodisaccharides



Scheme 16
Martin-Lomas's solid-phase synthesis of the glycan core of GPI anchors



Scheme 17
Vishwakarma's synthesis of the *T. cruzi* GPI anchor



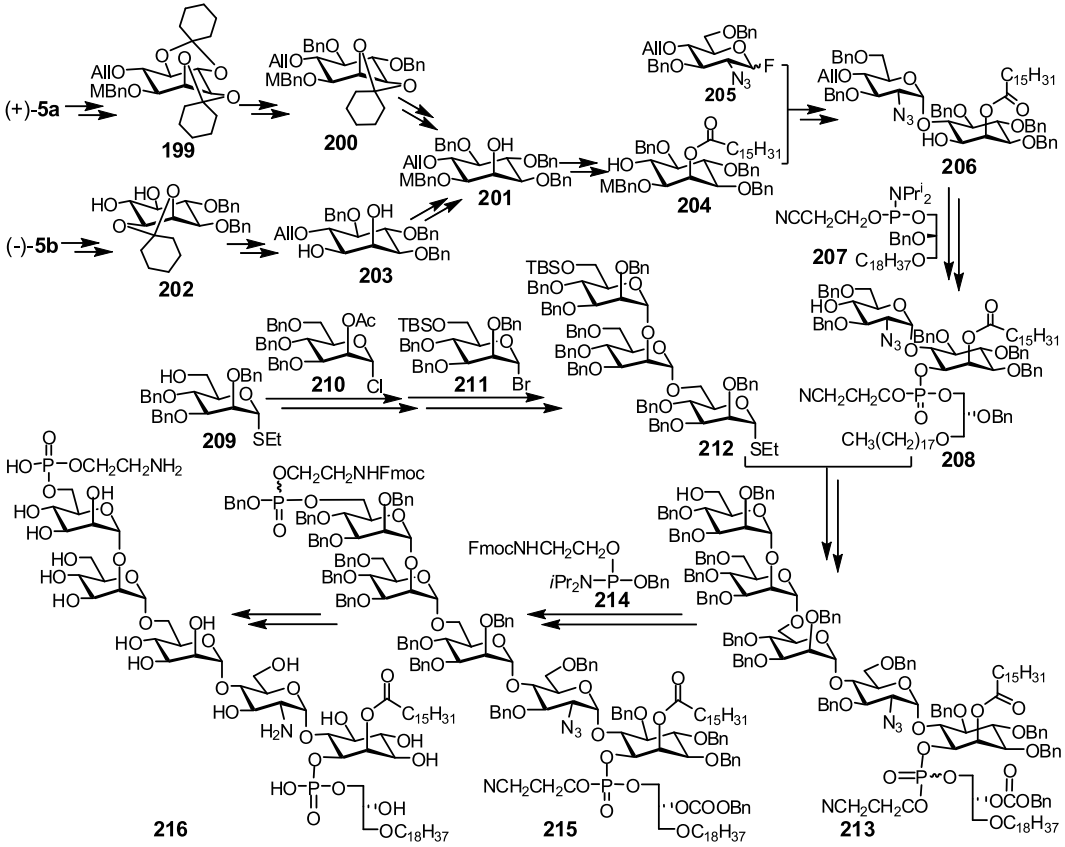
Scheme 18
Nikolaev's synthesis of the *T. cruzi* GPI anchor

were then separated to obtain the desired isomer **111**. After manipulation of the protecting groups, the product **186** was applied to carbohydrate chain elongation, which was achieved in a convergent fashion. Thus, a tetramannose segment **191** was first constructed through sequential assembly of monosaccharide units **187–190** and protecting group manipulation. Then, the tetrasaccharide was coupled to **186** to obtain the GPI glycan core **193** of *T. cruzi*, which was followed by sequential deprotection and phosphorylation of the glucosamine 6-*O*-position and Man III 6-*O*-position, respectively. Finally, the inositol 1-*O*-position was selectively deprotected and phospholipidated to produce the protected GPI anchor **197**. Global deprotection of **197** was achieved in 3 separate steps, namely, removal of the methyl phosphate by thiophenol, de-*O*-benzoylation using 0.05 M sodium methoxide in methanol, and then removal of the *t*-butyloxycarbonyl (Boc) groups by 10% trifluoroacetic acid (TFA) in water, to afford the free GPI anchor **198** of *T. cruzi* that contained an unsaturated lipid chain.

3.9 Synthesis of the Human Sperm CD52 GPI Anchor by Guo and Coworkers

As mentioned above, the GPI anchor of CD52 antigen retains the inositol 2-*O*-acyl group, which will increase the synthetic challenge, including the preparation of necessary inositol derivatives and the assembly of the target structure. Guo and coworkers [102,103] started their synthesis with the preparation and resolution of enantiomers **5a** and **5b** according to the reported chemoenzymatic procedure [104]. Both optically pure enantiomers **5a** and **5b** were used to synthesize the same inositol derivative **201**, in which the 1-*O*-, 2-*O*-, 6-*O*-positions were differentiated from each other and from other inositol positions as well (➤ *Scheme 19*). These syntheses were highlighted by several steps of regioselective deprotection reactions and tin-complex directed regioselective alkylation reactions. After **201** was obtained, a palmitoyl group was introduced to its 2-*O*-position, which was followed by deallylation to give **204** that was ready for installation of the carbohydrate chain. The glucosamine unit was introduced first with glycosyl fluoride **151** as a donor, affording a mixture that contained both α- and β-pseudodisaccharides, slightly in favor of the former (4 : 3). These isomers were easily separated after removal of the MBn group to obtain the desired pseudodisaccharide product **206**.

Before further elongation of the carbohydrate chain, Guo and coworkers [102,103] first installed the phospholipid moiety to the inositol 1-*O*-position, because they found that final stage installation of the phospholipid moiety, as reported for all other GPIs without the inositol 2-*O*-acyl group, caused problems. After the phospholipidated pseudodisaccharide **208** was obtained by reacting **206** with **207** followed by oxidation using *t*-butyl peroxide, the carbohydrate chain of **208** was elongated in a convergent fashion. Thus, monosaccharides **209**, **210**, and **211** were used to construct a trimannose **212**, which was directly coupled to **208** to yield the pseudopentasaccharide core **213**. Thereafter, a phosphoethanolamine moiety was introduced to the 6-*O*-posiiton of Man III to give the protected GPI anchor **215**. Global deprotection of **215** was accomplished in two steps, including base treatment to remove Fmoc and cyanoethyl groups and hydrogenolysis to remove Bn groups, to finally afford the GPI anchor **216** of CD52 antigen.



Scheme 19
Guo's synthesis of the CD52 GPI anchor

4 Chemical Synthesis of GPI-Linked Peptides/Proteins and Glycopeptides

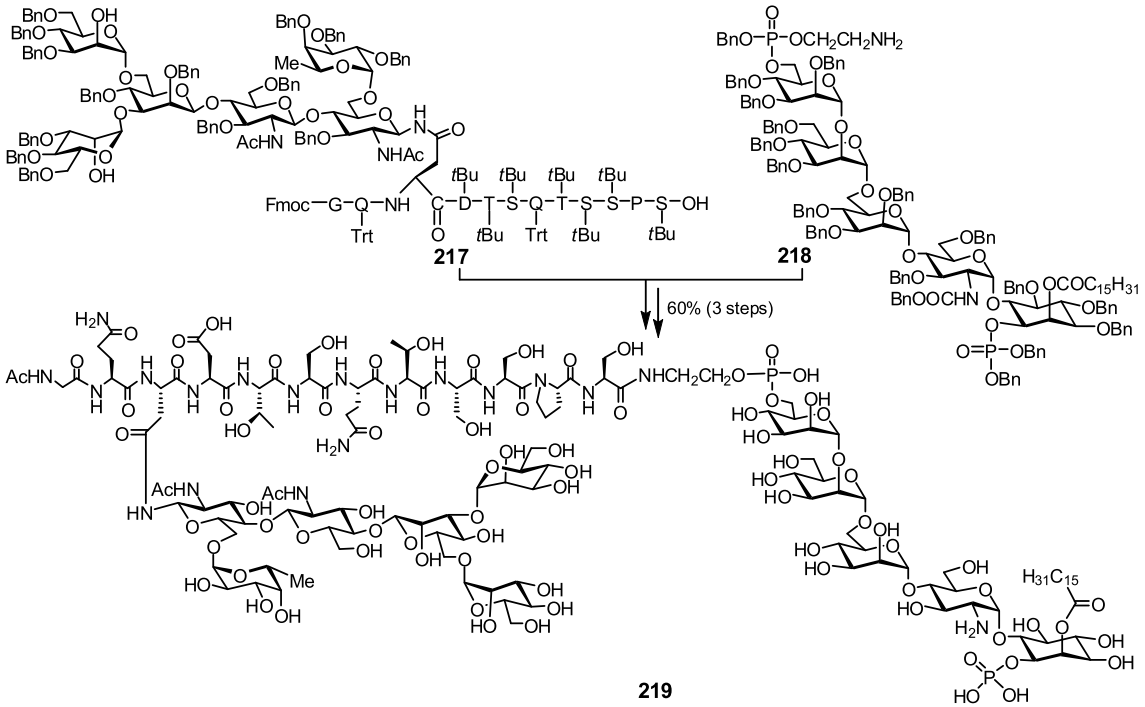
In the literature, there are very few reports concerning the chemical synthesis of GPI-linked structures. In the study of GPI-based malaria vaccines, Seeberger and coworkers [93] coupled their synthetic GPI **159** to a carrier protein to form an immunogenic glycoconjugate. For this purpose, the GPI was treated with 2-iminothiolane to introduce a thio group at the ethanolamine moiety and then conjugated to maleimide-activated ovalbumin or to keyhole limpet haemocyanin (KLH). This material was used to immunize mice for induction of protective immune responses against malaria. Evidently, the linkage between the GPI and the carrier protein was artificial.

Nakahara and coworkers [105] have recently studied the coupling between amino acids or peptides and GPIs through the native amide linkage using an unprotected dimannose containing a P-OEtNH_2 group as the GPI model. Both thioesters and pentafluorophenyl esters of amino acids were explored for this coupling reaction. The pentafluorophenyl esters gave complex results, as amino acids were attached to the free hydroxyl groups of the sugar units as well. The thioesters using silver nitrate as the promoter turned out to give a clean reaction with the desired product formed in a good yield (76%). However, this coupling method was significantly less effective for free peptides. Then, they proved that the native ligation method could couple free peptides to the GPI model in a good yield (78%). These results indicate that direct coupling of free peptides or proteins to free GPIs may not be easy, but the native ligation method is potentially useful for this problem.

Guo and coworkers have recently achieved the first chemical synthesis of a GPI-peptide [103] and a GPI-glycopeptide **219** [106] that contained the native amide linkage between the GPI and the peptide backbone. In these syntheses, they employed protected GPIs and peptides or glycopeptides for the coupling reaction with *N*-hydroxybenzotriazole (HOBt) and *N*,*N*′-dicyclohexylcarbodiimide (DCC) or 1-[3-(dimethylamino)propyl]-3-ethylcarbodiimide hydrochloride (EDC) as the condensation reagents. For example, the coupling reaction between **217** and **218** in the presence of EDC/HOBt (*Scheme 20*) was conducted in a homogeneous solution in dichloromethane and *N*-methylpyrrolidinone (2 : 1) to give an excellent yield (70%) of the desired conjugation product that was easily purified by column chromatography. Two-step global deprotection of this product, including catalytic hydrogenolysis to remove the benzyl groups and further treatment with 15% TFA in dichloromethane to remove the peptide side chain protections, eventually afforded the backbone structure of CD52 antigen **129**. The final product was then purified by reversed-phase HPLC.

5 Summary

Since the first total synthesis of a GPI anchor by Ogawa and coworkers in 1991 [59], several GPIs have been synthesized chemically. In this process, many generally useful strategies have been established for the synthesis of GPIs and related structures. However, these synthetic strategies are not always applicable to all GPIs. In fact, any additional functional group or even the change of the linking position of a functional group may necessitate completely different protection and/or synthetic strategies, so each GPI synthesis needs to be taken



Scheme 20
Guo's synthesis of a principal structure of CD52 antigen

as a unique challenge that requires specific attention. For example, most of the reported GPI syntheses were based on benzyl group-protected sugars, which are not compatible with GPIs that contain unsaturated lipid chains. Only lately have Nikolaev and coworkers [101] achieved some progress with regard to this issue by using benzoyl group, instead of benzyl group, as the universal and permanent protection of hydroxyl functionalities. Currently, an unsolved and important problem is the synthesis of natively-linked GPI-protein and glycoprotein conjugates. Although glycoconjugates with short peptides or glycopeptides attached to GPIs have been achieved by Guo and coworkers [103,106] through coupling protected peptides or glycopeptides to protected GPIs, the reactions between large proteins or glycoproteins and GPIs are much more complex. A hopeful solution to this problem may be the native ligation between proteins and GPIs, as suggested by Nakahara's work [105]. Nevertheless, because one of the most important functions of GPIs is to anchor proteins and glycoproteins onto cell surfaces and almost all GPIs in nature are coupled with proteins, glycoproteins, or other molecules, it is highly desirable to have a practical method to synthesize GPI-protein/glycoprotein conjugates. In conclusion, during the past two decades, the chemical synthesis of GPI anchors has gained significant progress, as proved by the number of GPIs prepared chemically. With these synthetic, homogeneous, and well-defined GPIs and GPI conjugates available, other related studies of GPIs will witness rapid growth. These studies will eventually help our understanding of the structures and activities of GPIs and GPI-anchored molecules, while presently the mechanism concerning how GPIs function is still largely unclear. Consequently, the synthetic studies of GPIs and GPI-linked glycoconjugates will have a major impact on many other areas of GPI research.

Acknowledgement

ZG thanks his devoted students and research associates who have made great contributions to our research programs. We appreciate the financial supports for our research from the US National Science Foundation (CHE-0407144) and National Institutes of Health (R01-CA95142).

References

1. Ilg T, Etges R, Overath P, McConville MJ, Thomas-Oates J, Thomas J, Homans SW, Ferguson MAJ (1992) J Biol Chem 267:6834
2. McConville MJ, Schnur LF, Jaffe C, Schneider P (1995) Biochem J 310:807
3. Ferguson MAJ (1999) J Cell Sci 112:2799
4. Field MC (1997) Glycobiology 7:161
5. Tiede A, Bastisch I, Schubert J, Orlean P, Schmidt RE (1999) Biol Chem 380:503
6. Treumann A, Guther MLS, Schneider P, Ferguson MAJ (1998) In: Hounsell EF (ed) Methods in molecular biology. Humana Press, Totawa, p 213
7. Stevens VL (1995) Biochem J 310:361
8. McConville MJ, Ferguson MAJ (1993) Biochem J 294:305
9. Homans SW, Ferguson MAJ, Dwek RA, Rademacher TW, Anand R, Williams AF (1988) Nature 333:269
10. Cross GAM (1975) Parasitol 71:393
11. Masterson WJ, Doering TL, Hart GW, Englund PT (1989) Cell 56:793
12. Stevens VL, Zhang H (1994) J Biol Chem 269:31397
13. Roberts WL, Santikarn S, Reinhold VN, Rosenberry TL (1988) J Biol Chem 263:18776

14. Ferguson MAJ (1992) Biochem J 284:297
15. Field MC, Menon AK, Cross GAM (1991) EMBO J 10:2731
16. Ferguson MAJ, Murray P, Rutherford H, McConville MJ (1993) Biochem J 291:51
17. Deeg MA, Humphrey DR, Yang SH, Ferguson TR, Reinhold VN, Rosenberry TL (1992) J Biol Chem 267:18573
18. Abuin G, Cuoto AS, de Lederkremer RM, Casal OL, Galli C, Colli W, Alves MJM (1996) Exp Parasitol 82:290
19. Gerold P, Schofield L, Blackman MJ, Holder AA, Schwarz RT (1996) Mol Biochem Parasitol 75:131
20. Treumann A, Lifely MR, Schneider P, Ferguson MAJ (1995) J Biol Chem 270:6088
21. Luhrs CA, Slomiany BL (1989) J Biol Chem 264:21446
22. Mayor S, Menon AK, Cross GAM (1990) J Biol Chem 265:6174
23. Walter EI, Roberts WL, Rosenberry TL, Ratnoff WD, Medof ME (1990) J Immunol 144:1030
24. Ralton JE, McConville MJ (1998) J Biol Chem 273:4245
25. Reggiori F, Canivenc-Gansel E, Conzelmann A (1997) EMBO J 16:3506
26. Gigg R, Gigg J (1997) In: Large DG, Warren CD (eds) Glycopeptides and related compounds: synthesis, analysis, and applications. Marcel Dekker, New York, p 327
27. Guo Z, Bishop L (2004) Euro J Org Chem 3585
28. Sureshan KM, Shashidhar MS, Praveen T, Das T (2003) Chem Rev 103:4477
29. Furgeson MAJ, Homans SW, Dwek RA, Rademacher TW (1988) Science 239:753
30. Field MC, Menon AK (1991) TIGG 3:107
31. Gerold P, Striepen B, Reitter B, Geyer H, Geyer R, Reinwald E, Risse HJ, Schwarz RT (1996) J Mol Biol 261:181
32. Guther MLS, Cardoso de Almeida ML, Yoshida N, Ferguson MAJ (1992) J Biol Chem 267:6820
33. Heise N, de Almeida MJC, Ferguson MAJ (1995) Mol Biochem Parasitol 70:71
34. Cuoto AS, de Lederkremer RM, Colli W, Alves MJM (1993) Eur J Biochem 217:597
35. Previato JO, Jones C, Xavier MT, Wait R, Travassos LR, Parodi AJ, Mendonca-Previato L (1995) J Biol Chem 270:7241
36. Serrano AA, Schenkman S, Yoshida N, Mehlert A, Richardson JM, Ferguson MAJ (1995) J Biol Chem 270:27244
37. Schneider P, Ferguson MAJ, McConville MJ, Mehlert A, Homans SW, Bordier C (1990) J Biol Chem 265:16955
38. McConville MJ, Collidge TAC, Ferguson MAJ, Schneider P (1993) J Biol Chem 268:15595
39. Tomavo S, Dubremetz JF, Schwarz RT (1993) Biol Cell 78:155
40. Striepen B, Zinecker CF, Damm JBL, Melgers PAT, Gerwig GJ, Koolen M, Vliegenthart JFG, Dubremetz JF, Schwarz RT (1997) J Mol Biol 266:797
41. Schares G, Zinecker CF, Schmidt J, Azzouz N, Conraths FJ, Gerold P, Schwarz RT (2000) Mol Biochem Parasitol 105:155
42. Gerold P, Vivas L, Ogun SA, Azzouz N, Brown KN, Holder AA, Schwarz RT (1997) Biochem J 328:905
43. Azzouz N, Striepen B, Gerold P, Capdeville Y, Schwarz RT (1995) EMBO J 14:4422
44. Fankhauser C, Homans SW, Thomas-Oates JE, McConville MJ, Desponds C, Conzelmann A, Ferguson MAJ (1993) J Biol Chem 268:26365
45. Haynes PA, Gooley AA, Ferguson MAJ, Redmond JW, Williams KL (1993) Eur J Biochem 216:729
46. Butikofer P, Kuypers FA, Shackleton C, Brodbeck U, Stieger S (1990) J Biol Chem 265:18983
47. Mehlert A, Varon L, Silman I, Homans SW, Ferguson MAJ (1993) Biochem J 296:473
48. Stahl N, Baldwin MA, Hecker R, Pan KM, Burlingame AL, Prusiner SB (1992) Biochem 31:5043
49. Mukasa R, Umeda M, Endo T, Kobata A, Inoue K (1995) Arch Biochem Biophys 318:182
50. Taguchi R, Hamakawa N, Harada-Nishida M, Fukui T, Nojima K, Ikezawa H (1994) Biochem 33:1017
51. Brewis IA, Ferguson MAJ, Mehlert A, Turner AJ, Hooper NM (1995) J Biol Chem 270:22946
52. Roberts WL, Myher JJ, Kuksis A, Low MG, Rosenberry TL (1988) J Biol Chem 263:18766
53. Redman CA, Thomas-Oates JE, Ogata S, Ikehara Y, Ferguson MAJ (1994) Biochem J 302:861
54. Nakano Y, Noda K, Endo T, Kobata A, Tomita M (1994) Arch Biochem Biophys 311:117
55. Sugita Y, Nakano Y, Oda E, Noda K, Tobe T, Miura NH, Tomita M (1993) J Biochem 114:473
56. Lee HC, Shoda R, Krall JA, Foster JD, Selhub J, Rosenberry TL (1992) Biochem 31:3236
57. Murakata C, Ogawa T (1990) Tetrahedron Lett 31:2439

58. Murakata C, Ogawa T (1992) Carbohydr Res 235:95
59. Murakata C, Ogawa T (1991) Tetrahedron Lett 32:671
60. Murakata C, Ogawa T (1991) Tetrahedron Lett 32:101
61. Murakata C, Ogawa T (1992) Carbohydr Res 234:75
62. Garegg PJ, Iversen T, Johansson R, Lindberg B (1984) Carbohydr Res 130:322
63. Matsumoto T, Maeta H, Keisuke S, Tsuchihashi G (1988) Tetrahedron Lett 29:3567
64. Lindh I, Stawinski J (1989) J Org Chem 54:1338
65. Baeschlin DK, Chaperon AR, Charbonneau V, Green LG, Ley SV, Lucking U, Walther E (1998) Angew Chem Int Ed Engl 37:3423
66. Baeschlin DK, Chaperon AR, Green LG, Hahn MG, Ince SJ, Ley SV (2000) Chem Eur J 6:172
67. Watanabe Y, Mitani M, Morita T, Ozaki S (1989) J Chem Soc, Chem Commun 482
68. Lemieux RU, Hendriks KB, Stick RV, James K (1975) J Am Chem Soc 97:4056
69. Boons GJ, Grice P, Leslie R, Ley SV, Yeung LL (1993) Tetrahedron Lett 34:8523
70. Ley SV, Parra M, Redgrave AJ, Sternfirle F (1990) Tetrahedron 46:4995
71. Ley SV, Yeung LL (1992) Synlett 997
72. Mootoo DR, Konradsson P, Fraser-Reid B (1989) J Am Chem Soc 111:8540
73. Garegg PJ, Iversen T, Oscarson S (1976) Carbohydr Res 50:c12
74. Vacca JP, de Solms JS, Huff JR (1987) J Am Chem Soc 109:3478
75. Udodong UE, Madsen R, Roberts C, Fraser-Reid B (1993) J Am Chem Soc 115:7886
76. Roberts C, Madsen R, Fraser-Reid B (1995) J Am Chem Soc 117:1546
77. Madsen R, Udodong UE, Roberts C, Mootoo DR, Konradsson P, Fraser-Reid B (1995) J Am Chem Soc 117:1554
78. Campbell AS, Fraser-Reid B (1995) J Am Chem Soc 117:10387
79. Kovac P, Edgar KJ (1992) J Org Chem 57:2455
80. Jia ZJ, Olsson L, Fraser-Reid B (1998) J Chem Soc, Perkin Trans 1 631
81. Lu J, Jayaprakash KN, Fraser-Reid B (2004) Tetrahedron Lett 45:879
82. Lu J, Jayaprakash KN, Schlueter U, Fraser-Reid B (2004) J Am Chem Soc 126:7540
83. Soli ED, Manoso AS, Patterson MC, DeShong P, Favor DA, Hirschmann R, Amos B. Smith I (1999) J Org Chem 64:3171
84. King JF, Allbutt AD (1970) Can J Chem 48:1754
85. Mayer TG, Schmidt RR (1999) Eur J Org Chem 1153
86. Mayer TG, Kratzer B, Schmidt RR (1994) Angew Chem Int Ed Engl 33:2177
87. Tailler D, Ferrieres V, Pekari K, Schmidt RR (1999) Tetrahedron Lett 40:679
88. Pekari K, Schmidt RR (2003) J Org Chem 68:1295
89. Pekari K, Tailler D, Weingart R, Schmidt RR (2001) J Org Chem 66:7432
90. Mayer TG, Schmidt RR (1999) Eur J Org Chem 1153
91. Mayer TG, Schmidt RR (1997) Liebigs Ann/Recueil 859
92. Blattner G, Beau JM, Jacquinet JC (1994) Carbohydr Res 260:189
93. Schofield L, Hewitt MC, Evans K, Siomos MA, Seeberger PH (2002) Nature 418:785
94. Hewitt MC, Snyder DA, Seeberger PH (2002) J Am Chem Soc 124:13434
95. Andrade RB, Plante OJ, Melean LG, Seeberger PH (1999) Org Lett 1:1811
96. Martin-Lomas M, Khiar N, Garcia S, Koessler JL, Nieto PM, Rademacher TW (2000) Chem Eur J 6:3608
97. Bruzik KS, Tsai MD (1992) J Am Chem Soc 114:6361
98. Watanabe Y, Yamamoto T, Ozaki S (1996) J Org Chem 61:14
99. Reichardt NC, Martin-Lomas M (2003) Angew Chem Int Ed 42:4674
100. Ali A, Gowda DC, Vishwakarma RA (2005) Chem Commun 519
101. Yashunsky DV, Borodkin VS, Ferguson MAJ, Nikolaev AV (2006) Angew Chem Int Ed 45: 468
102. Xue J, Guo Z (2002) Bioorg Med Chem Lett 12:2015
103. Xue J, Shao N, Guo Z (2003) J Org Chem 68:4020
104. Gou D, Liu Y, Chen C (1992) Carbohydr Res 234:51
105. Tanaka Y, Nakahara Y, Hojo H, Nakahara Y (2003) Tetrahedron 59:4059
106. Shao N, Xue J, Guo Z (2004) Angew Chem Int Ed 43:1569

Part 8
Glycoproteins

8.1 Glycoproteins: Occurrence and Significance

Valentin Wittmann
Fachbereich Chemie, Universität Konstanz, 78457 Konstanz, Germany
mail@valentin-wittmann.de

Abstract

Protein glycosylation is regarded as the most complex form of post-translational modification leading to a heterogeneous expression of glycoproteins as mixtures of glycoforms. This chapter describes the structure and occurrence of glycoproteins with respect to their glycan chains. Discussed are different carbohydrate–peptide linkages including GPI anchors, common structures of *N*- and *O*-glycans, and the structure of glycosaminoglycans contained in proteoglycans. Also covered are the bacterial cell wall polymer peptidoglycan and the glycopeptide

In: *Glycoscience*. Fraser-Reid B, Tatsuta K, Thiem J (eds)
Chapter-DOI 10-1007/978-3-540-30429-6_43:

antibiotics of the vancomycin group. Properties and functions of the glycans contained in glycoproteins are dealt with in the next chapter of this book.

Keywords

Glycoproteins; Glycopeptides; Glycoforms; Microheterogeneity; *N*-Glycans; *O*-Glycans; Antigenic determinants; GPI anchors; Mucins; Proteoglycans

Abbreviations

AGEs advanced glycation end products
CE capillary electrophoresis
DATDH 2,4-diacetamido-2,4,6-trideoxyhexose
ER endoplasmic reticulum
ERGIC ER-Golgi intermediate compartment
EGF epidermal growth factor
ECM extracellular matrix
GPI glycosyl-phosphatidylinositol
GPC gel permeation chromatography
HPAEC high pH anion exchange chromatography
HPLC high performance liquid chromatography
IGF-I insulin-like growth factor-I
MS mass spectrometry
***m*-DAP** *meso*-diaminopimelic acid
MRSA methicillin-resistant *Staphylococcus aureus*
OT oligosaccharyl transferase
PAGE polyacrylamide gel electrophoresis
PNGase peptide *N*-glycanase
RAAM reagent array analysis method
SBA soybean agglutinin
tPA tissue plasminogen activator
VSG variant surface glycoprotein

1 Introduction

This chapter covers natural glycoconjugates in which carbohydrates are covalently linked to amino acids which themselves are part of a peptide or protein. Three classes of biologically important compounds fulfill this criterion:

- glycoproteins including their substructures, the glycopeptides,
- the bacterial cell wall polymer peptidoglycan, and
- glycopeptide antibiotics.

Glycoproteins constitute the most diverse of these classes of glycoconjugates. Their substructures, the glycopeptides, occur as natural and artificial degradation products of glycoproteins during metabolism and structural analysis, respectively, or they are obtained synthetically in

order to probe their structure and/or function. Peptidoglycan is commonly not regarded as a glycopeptide. It forms a bag-like macromolecule (called the sacculus) wrapping bacteria and thereby contributing essentially to their mechanical stability. However, it is connected with glycopeptides not only by virtue of its composition but also because its biosynthesis is inhibited by (beside β-lactam antibiotics) glycopeptide antibiotics. The latter are secondary metabolites produced in actinomycetes and are dealt with at the end of this chapter.

2 Glycoproteins

Glycoproteins [1,2,3,4,5,6,7,8,9,10,11,12,13] consist of carbohydrate covalently linked with protein and are ubiquitous in all forms of life. Once thought to be restricted to eukaryotes, protein glycosylation is now being increasingly reported in prokaryotes [14,15,16,17]. The carbohydrate content of glycoproteins is variable from less than 1%, as in some collagens, to over 99%, as in glycogen. The carbohydrate may be in the form of a monosaccharide or disaccharide, but more frequently in the form of oligosaccharides and polysaccharides (up to hundreds of monosaccharides in size), or their derivatives (e. g. sulfo- or phospho-substituted), linear or branched, generally referred to as glycans. The number of glycans present in a glycoprotein differs greatly from a single to more than a hundred. The carbohydrate chains are attached to the polypeptide backbone by characteristic carbohydrate–peptide linkages.

Being the most complex co- and post-translational modification, glycosylation is found in most proteins including enzymes, antibodies, receptors, hormones, cytokines, and structural proteins. Glycoproteins occur inside cells, both in the cytoplasm and in subcellular organelles, in extracellular fluids as well as embedded in cell membranes. In the latter case the glycans are located extracellularly. Especially blood serum is a rich source of glycoproteins. Of the almost 100 proteins which have been identified in this fluid, almost all are glycosylated. Serum albumin is one of the rare exceptions in this respect although a genetic variant which is glycosylated has been discovered [18].

A striking feature of almost all glycoproteins is the polymorphism associated with their glycan moieties, a phenomenon known as microheterogeneity. This type of diversity derives from the fact that glycans are secondary gene products and is manifested in that individual molecules of a given glycoprotein carry different oligosaccharides at the same glycosylation site of the protein backbone. The resulting variants, referred to as glycoforms, were first observed in 1962 by Schmid et al. in α_1-acid glycoprotein from human serum by electrophoresis [19]. This glycoprotein contains five glycosylation sites modified with di-, tri-, and tetra-antennary glycans of the *N*-acetyllactosamine type [20]. In addition, the glycans can be fucosylated [21] and sialylated at different levels. Microheterogeneity is also observed in proteins with a single glycosylation site like chicken ovalbumin. Close to 20 different oligosaccharides have been identified at that site (Asn-293) [22]. One of the very rare cases in which microheterogeneity is absent is soybean agglutinin (SBA), a plant glycoprotein with a single uniform oligosaccharide per subunit, namely $Man_9(GlcNAc)_2$ [23]. This is mainly why SBA is the best source for a preparative isolation of this oligosaccharide. Up to now, many thousands of primary glycan structures have been characterized. There are several glycan databases available via the internet, for example, Glycosciences.de provided by the German Cancer Research Center (http://www.glycosciences.de) [24], KEGG Glycan, part of the Kyoto Encyclopedia of Genes and Genomes (http://www.genome.jp/kegg/glycan) [25], and the Glycan Database of the Consor-

tium for Functional Glycomics (http://www.functionalglycomics.org) [26]. These resources are increasingly linking to each other, and additional levels of integration are in development, including the EuroCarbDB initiative (http://www.eurocarbdb.org).

The observation of microheterogeneity gives rise to many interesting questions regarding the origin of this phenomenon and its biological relevance. Whereas in the early view the heterogeneity of the glycans was thought to be random and mainly a result of the lack of fidelity in their synthesis, today it appears that under constant physiological conditions the populations in a set of glycoforms are reproducible and highly regulated. Moreover, the populations change under certain conditions such as cell growth, cell differentiation, and disease, including malignant transformation, suggesting that the presence of different but defined glycoforms is indeed required for the normal functioning of an organism.

Today we know there is no single unifying function for the carbohydrates present in glycoproteins [27]. Perhaps their major function is to participate in numerous physiological and pathological molecular recognition events. In addition, they may modify the physical, chemical, and biological properties of the proteins they are attached to. It has been shown that the glycans alter charge and solubility of proteins and influence the conformation and dynamic properties of the polypeptide chain. Because of their large size, the oligosaccharides may cover functionally important areas of the proteins and thus regulate their interactions with other biomolecules or protect the protein from proteolytic degradation. Properties and functions of glycoproteins are covered in Chap. 8.2 of this book.

2.1 Monosaccharide Constituents

For a long time less than a dozen monosaccharides were considered to be the main constituents of glycoproteins. Beside these "classical" saccharides, however, refinement of the analytical methods brought up many new monosaccharides, originally considered as rare but which now appear to be more common than previously thought. *Table 1* [6,28,29,30] gives an overview of the classical and examples of rare monosaccharide constituents of glycoproteins.

2.2 Carbohydrate–Peptide Linkages

There are three major types of linkages between carbohydrates and proteins [1,2,3,4,5,6,7,8,9,30,31]:

- *N*-glycosidic, between the reducing end monosaccharide and the side chain amide group of asparagine (*N*-glycans),
- *O*-glycosidic, between the reducing end monosaccharide and the side chain hydroxy group of most commonly serine and threonine but also hydroxyproline, hydroxylysine, and tyrosine (*O*-glycans), and
- via ethanolamine phosphate, between the *C*-terminal amino acid of the protein and an oligosaccharide attached to phosphatidylinositol, generally known as the glycosyl-phosphatidylinositol (GPI) anchor.

Beside these, a number of uncommon linkages are found in nature including *C*-glycosides and carbohydrates linked via a phosphodiester bridge. The non-enzymatic condensation of

Table 1
Monosaccharides found in glycoproteins [6,28,29,30]

Monosaccharide[a]	Abbreviation	Comment
"*Classical*"		
Galactose	Gal	
Glucose	Glc	Mainly in collagens
Mannose	Man	
N-Acetylgalactosamine	GalNAc	
N-Acetylglucosamine	GlcNAc	
L-Fucose	Fuc	
L-Arabinofuranose	Ara*f*	In plant glycoproteins
Xylose	Xyl	In proteoglycans and plant glycoproteins
Glucuronic acid	GlcA	In proteoglycans
L-Iduronic acid	IdoA	In proteoglycans
N-Acetylneuraminic acid	Neu5Ac	Mainly in higher vertebrates and invertebrates
Rare		
2-Acetamido-4-amino-2,4,6-trideoxyglucose		*Clostridium symbiosum*
6-Deoxyaltrose		Salmonid fish eggs
3-Deoxy-D-*glycero*-*galacto*-nonulosonic acid	Kdn	Salmonid fish eggs
2,3-Diacetamido-2,3-dideoxymannuronic acid		*Bacillus stearothermophilus*
2-*O*-Methyl-L-fucose	Fuc2Me	Nematodes
Galactofuranose	Gal*f*	Bacteria, trypanosoma, fungi
3-*O*-Methylgalactose	Gal3Me	Snail
4-*O*-Methylgalactose	Gal4Me	Nematodes
6-*O*-Methylgalactose	Gal6Me	Algae
Galactose-3-sulfate	Gal3*S*	Thyroglobulin, mucins in cystic fibrosis
N-Acetylgalactosamine-4-sulfate	GalNAc4*S*	Pituitary hormones, Tamm–Horsfall glycoprotein, urokinase
3-*O*-Methylglucose	Glc3Me	*Methanothermus fervidus*
3-*O*-Methyl-*N*-acetylglucosamine	GlcNAc3Me	*Clostridium thermocellum*
N-Acetylglucosamine-6-sulfate	GlcNAc6*S*	Thyroglobulin
Gulose	Gul	Algae
3-*O*-Methylmannose	Man3Me	Snail
Mannose-4-sulfate	Man4*S*	Ovalbumin
Mannose-6-sulfate	Man6*S*	Ovalbumin, slime mold
Mannose-6-methylphosphate	Man6*P*Me	Slime mold
N-Acetylmannosamine	ManNAc	*Clostridium symbiosum*

Table 1 (continued)
Monosaccharides found in glycoproteins [6,28,29,30]

Monosaccharide[a]	Abbreviation	Comment
4,8-Anhydro-*N*-acetylneuraminic acid		Edible bird's nest
8-*O*-Methyl-9-*O*-acetyl-*N*-glycolylneuraminic acid	Neu5Gc8Me9Ac	Starfish
8-*O*-Methyl-7,9-di-*O*-acetyl-*N*-glycolylneuraminic acid	Neu5Gc7,9Ac$_2$8Me	Starfish
L-Rhamnose	Rha	Eubacteria
Pseudaminic acid (5,7-diacetamido-3,5,7,9-tetradeoxy-L-glycero-L-*manno*-nonulosonic acid)	Pse	*Campylobacter jejuni* flagellin, pili of *Pseudomonas aeruginosa*
2,4-Diacetamido-2,4,6-trideoxyhexose	DATDH	Pili of *Neisseria meningitidis* and *Neisseria gonorrhoeae*
N-Acetylfucosamine	FucNAc	Pili of *Pseudomonas aeruginosa*

[a] Unless otherwise stated, all monosaccharides are of D-configuration and the ring form is pyranose

reducing monosaccharides, such as glucose, with the side chain amino group of lysine leads via initial Schiff bases and subsequent rearrangements to the formation of so-called advanced glycation end products (AGEs). AGEs have been implicated in alterations of proteins during aging and long-term diabetes [32,33] and are not dealt with in this chapter.

2.2.1 *N*-Glycosides

Until recently, the only type of *N*-glycosidic linkage that had been characterized in glycoproteins regardless of their origin was the *N*-acetylglucosaminyl(β1–N)asparagine bond discovered by Neuberger et al. [34] and almost simultaneously by others [35,36]. During biosynthesis, the enzyme oligosaccharyl transferase (OT) transfers a triantennary tetradecasaccharide ($Glc_3Man_9GlcNAc_2$) from dolichol pyrophosphate to the amide nitrogen of an Asn side chain in the nascent polypeptide [37]. In the subsequent trimming process this oligosaccharide is modified by action of several glycosylhydrolases and glycosyltransferases resulting in a structural diversity of glycans sharing a common core pentasaccharide (*Fig. 1*). The primary peptide sequence requirements for OT-catalyzed glycosylation include a minimum Asn-Xaa-Ser/Thr (sometimes also Cys) tripeptide recognition motif where Xaa can be any of the 20 natural amino acids except proline. However, it is found that not all such consensus tripeptides (sequons) are glycosylated. A plausible explanation arises from mechanistic studies on the OT-catalyzed glycosylation of linear and constrained model peptides [38]. It could be demonstrated that only peptides which are able to adopt an Asx-turn conformation [39] are efficient substrates of OT leading to the proposed mechanism of oligosaccharyl transfer shown in *Fig. 2a* [40,41]. In this model, the apparent failure of the glycosylation machinery most likely results from conformational influence by the neighboring polypeptide sequence that may override the potential for the tripeptide acceptor sequence to adopt an Asx-turn and therefore limit its compatibility with the OT active site. However, although necessary, formation of an Asx-turn might not be sufficient [42]. An alternate mechanistic model which can also explain both specificity and enhanced amide nitrogen nucleophilicity was proposed by Bause

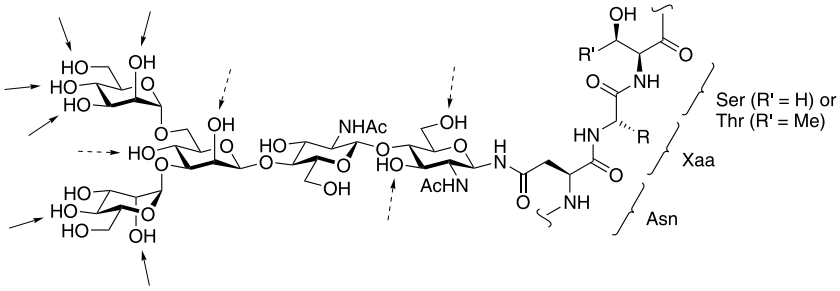
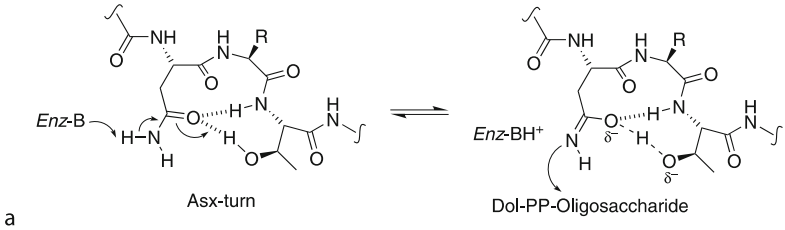


Figure 1
Core pentasaccharide Man(α1–6)[Man(α1–3)]Man(β1–4)GlcNAc(β1–4)GlcNAc of *N*-glycoproteins attached to the consensus sequence Asn-Xaa-Ser/Thr. *Solid arrows* indicate the points of attachment of the outer arm saccharides forming carbohydrate chains, called antennae. In addition, the inner-core may be substituted by several monosaccharides (*dashed arrows*)



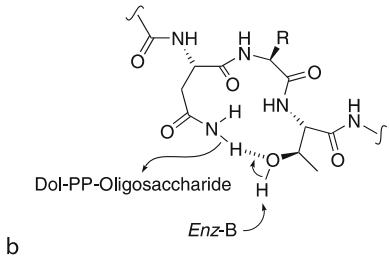


Figure 2
Proposed mechanisms of asparagine activation for oligosaccharyl transferase by (a) Imperiali et al. [40,41] and (b) Bause et al. [43,44,45]

et al. [43,44,45]. In this model the hydroxyamino acid acts as a hydrogen-bond acceptor rather than a donor (*Fig. 2b*).

During recent years, linkages between asparagine side chains and other carbohydrates such as α- [46] and β-glucose [47], β-*N*-acetylgalactosamine [48] and L-rhamnose[49] have been discovered mainly in bacterial glycoproteins [50]. However, the β-glucosyl–asparagine linkage has been found also in the mammalian protein laminin [51]. Whenever carefully studied, the saccharides are attached to the sequon Asn-Xaa-Ser/Thr except for the glycopeptide nephritogenoside in which glucose is α-linked to the amide side chain of the *N*-terminal tripeptide Asn-Pro-Leu [46]. Another example of an *N*-glycosidic bond is the linkage of β-glucose to

the guanidino group of arginine which has been reported to occur in amylogenin, a self-glucosylating protein from sweet corn [52].

2.2.2 *O*-Glycosides

In contrast to *N*-glycoproteins, biosynthesis of the glycan chains of *O*-glycoproteins is an entirely post-translational process and starts with addition of a single monosaccharide to a hydroxy amino acid of the protein backbone [53]. Therefore, a variety of carbohydrate–protein linkages are found in *O*-glycoproteins (● *Fig. 3*). The *N*-acetylgalactosaminyl (α1–O) serine/threonine linkage (also called Tn antigen) was first demonstrated in mucins and is widely distributed in nature in the so-called mucin-type glycoproteins. An increasingly important *O*-glycosidic bond is that between β-*N*-acetylglucosamine and the serine/threonine hydroxyl [54,55,56,57]. This type of attachment is present predominantly in intracellular glycoproteins (nuclear pore, chromatin proteins, transcription factors, and cytoplasmic inclusions) and indeed represents the first reported example of glycosylated proteins found outside of the secretory channels. Interestingly, no other sugars are connected to the *N*-acetylglucosamine residue. The *N*-acetylglucosaminyl (β1–O) serine/threonine modification shares many common traits with protein phosphorylation. Both are dynamic modifications processed by specific enzymes that modify serine/threonine residues and rapidly respond to extracellular stimuli.

In the epidermal growth factor (EGF) modules of several blood coagulation and fibrinolytic proteins, L-fucose was characterized to be α-linked to serine or threonine in the consensus sequence Cys-Xaa-Xaa-Gly-Gly-Ser/Thr-Cys [58]. Later on, the consensus sequence was broadened to Cys-$(Xaa)_{3-5}$-Ser/Thr-Cys to account for observations made with the proteins Notch, Serrate/Jagged, and Delta [59]. EGF modules also contain glucose β-*O*-linked to serine within the consensus sequence Cys-Xaa-Ser-Xaa-Pro-Cys, an example being the bovine blood clotting factor IX [60]. α-Galactose bound to serine is found in plant glycoproteins like extensin or potato lectin and in the glycoproteins from the cellulosome, an extracellular complex of cellulases, produced by cellulolytic bacteria [61]. When human insulin-like growth factor-I (IGF-I) was expressed in yeast, a new form of IGF-I was characterized in addition to IGF-I in which mannose α-glycosidically was linked to threonine [62]. The mannosyl Ser/Thr linkage [63] has also been reported in Ser/Thr-rich domains of the bovine peripheral nerve α-dystroglycan [64] and in rat brain proteoglycans [65,66]. Xylosyl (β1–O) serine is found in animal proteoglycans and, at present, considered to be confined to these glycoproteins. Pseudaminic acid (5,7-diacetamido-3,5,7,9-tetradeoxy-L-glycero-L-*manno*-nonulosinic acid) α-glycosidically linked to serine/threonine is an unusual bond that has been recently identified in *Campylobacter jejuni* flagellin [67,68]. It represents the first example of an acidic monosaccharide directly linked to protein. In the pili of *Neisseria meningitidis*, a 2,4-diacetamido-2,4,6-trideoxyhexose (DATDH) as part of the trisaccharide Gal(β1–4)Gal(α1–3)DATDH is linked to serine or threonine with unknown anomeric configuration [69]. In the pili of *Neisseria gonorrhoeae*, a disaccharide Gal(α1–3)DATDH has been reported to be *O*-glycosidically linked to serine [70] in contrast to an earlier publication in which Gal(α1–3)GlcNAc was assigned to be this disaccharide [71]. The *N*-acetylfucosaminyl (β1–O) serine/threonine linkage represents another recently described bacterial linkage that occurs in the pili of *Pseudomonas aeruginosa* [72].

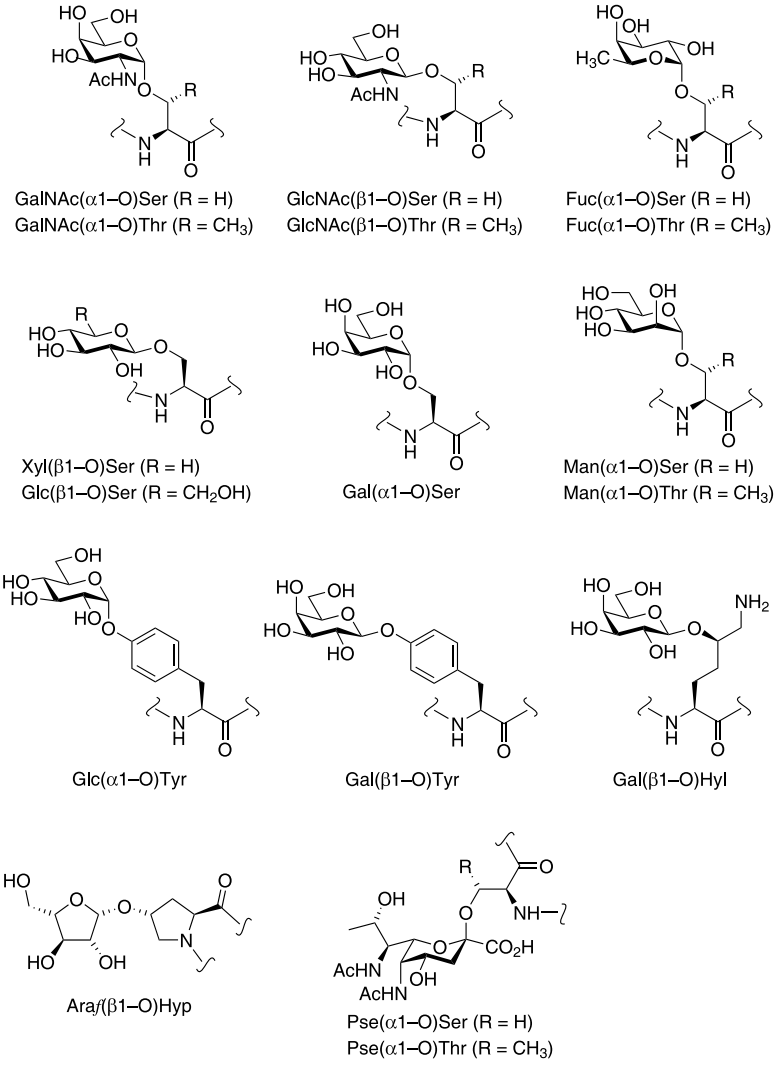


Figure 3
Carbohydrate–peptide linkages found in *O*-glycoproteins

O-Glycosylation is, however, not limited to serine and threonine. In glycogenin, the priming enzyme for glycogen synthesis, glucose is α-linked to the phenolic hydroxy group of a tyrosine residue [73,74] and galactosyl (β1–O) tyrosine has been identified in glycoproteins of the crystalline surface layers (*S*-layers) of eubacteria [75,76]. In the collagens, β-galactose is linked to hydroxylysine [77,78]. Hydroxyproline, finally, is glycosylated with α- [79] and β-L-arabinofuranose [80,81,82] in certain plant glycoproteins, with β-galactose in plants and eubacteria [83,84], and with *N*-acetylglucosamine in cytoplasmic glycoprotein of *Dictyostelium* [85].

2.2.3 GPI Anchors

Glycosyl-phosphatidylinositol anchors provide a means for anchoring proteins with a wide variety of structures and functions to the external surface of the plasma membrane of eukaryotic cells. The existence of lipid anchors for membrane proteins was initially demonstrated by the ability of phosphatidylinositol-specific phospholipase C (PI-PLC) to release alkaline phosphatase from various mammalian tissues into the medium [86,87]. Later, it was observed that fatty acids and ethanolamine were attached to the carboxy terminus of the rat Thy-1 antigen which provided direct evidence for a glycolipid anchor being covalently attached to protein [88]. In 1985 these and other results were combined to establish a common mode of membrane attachment via a GPI anchor covalently bound to the *C*-terminus through ethanolamine [89]. Today, we know well over 100 proteins which are GPI-anchored (reviewed in [90,91,92,93,94,95,96,97,98]) and the structures of several GPI anchors have been elucidated, the first being that on the *Trypanosoma brucei* variant surface glycoprotein (VSG) [99,100]. All characterized GPI anchors share a common core of ethanolamine-PO_4-6Man(α1–2)Man(α1–6)Man(α1–4)GlcN(α1–6)*myo*-Ino-1-PO_4 [99,100].

All characterized GPI anchors share a common core of ethanolamine-PO_4-6Man(α1–2)Man(α1–6)Man(α1–4)GlcN(α1–6)*myo*-Ino-1-PO_4-lipid. This may be further processed in a cell-type and protein-specific manner (*Fig. 4*). The rat brain Thy-1 anchor, for example, contains additional Man(α1–2) and GalNAc(β1–4) residues, whereas the tetrasaccharide backbone of the trypanosome VSG anchor is modified with a branched chain of α-linked galactoses. In human erythrocyte acetylcholinesterase [101] and decay accelerating factor [102] the first and second mannoses carry additional phosphoethanolamines. Attached to the phosphoinositol are lipids of varying chain length and saturation including diacylglyc-

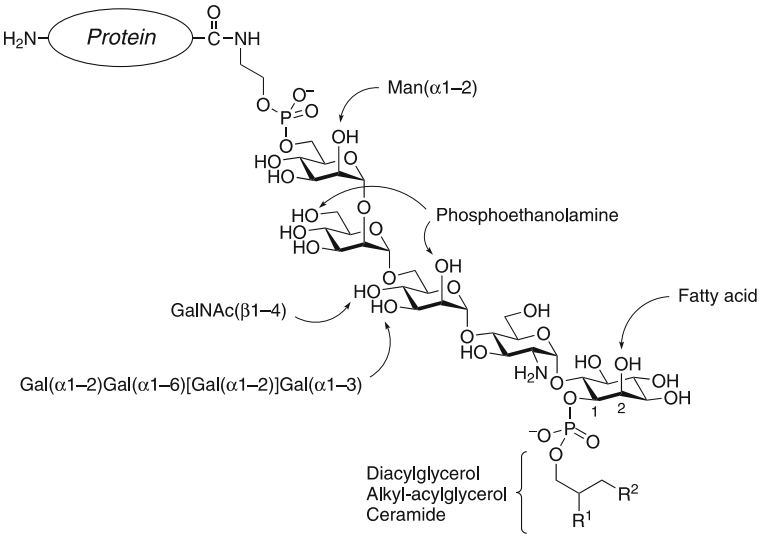


Figure 4
Minimal structure and some modifications of GPI anchors

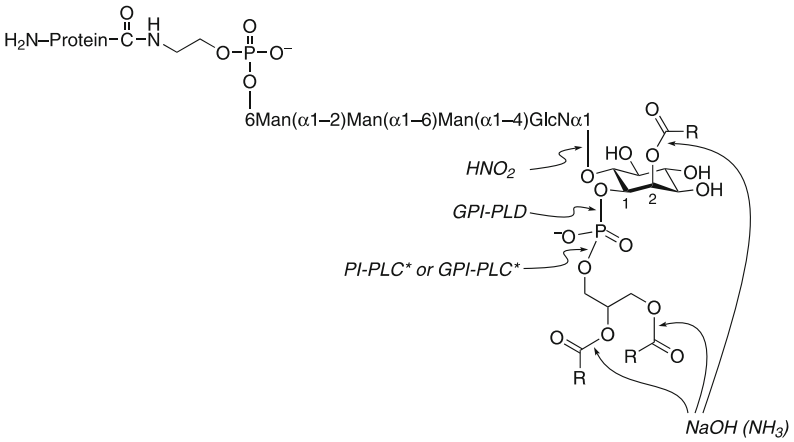


Figure 5
Enzymatic and chemical cleavage sites of GPI anchors most commonly used for the identification and structural analysis of GPI anchors. In addition, all phosphoric esters are cleavable by aqueous hydrofluoric acid. GPI-PLC*, GPI-specific phospholipase C; GPI-PLD, GPI-specific phospholipase D; PI-PLC*, phosphatidyinositol-specific phospholipase C. * Cleaves only if the inositol is unmodified

erol, alkyl-acylglycerol, stearoyl-*lyso*glycerol, or ceramide. In the VSG anchor, for example, dimyristyl glycerol has been identified. Furthermore, the inositol may be acylated with an additional fatty acid, most commonly palmitoyl, at position 2 of the inositol ring.

Biosynthesis of GPI anchors starts with the core structure assembly by sequential addition of UDP-GlcNAc (followed by *N*-deacetylation), dolichol-phosphate-mannose, and phosphoethanolamine to phosphatidylinositol and culminates in the en bloc transfer to protein shortly after the protein is synthesized. However, the biosynthetic pathways can differ strikingly between different organisms with respect to specific modifications and fatty acid remodeling occurring after completion of the core glycan. This also applies for the point when certain modifications are introduced, e. g. before or after the transfer of the GPI-moiety to the protein. GPI anchors can be cleaved at defined positions by an array of enzymatic and chemical methods, respectively (◉ *Fig. 5*). Thus, it becomes possible to identify GPI-anchored proteins and, moreover, analyze the structure and biosynthesis of GPI anchors [103].

The most obvious function of GPI anchors is to provide a stable, oriented attachment of proteins onto membranes, usually the extracellular surface of plasma membranes. However, there is much discussion over whether this highly conserved, multiple-step, complex mechanism for anchoring proteins has additional functions and it has been proposed [96,98,104,105] that GPI anchors:

- increase the lateral mobility of proteins,
- mediate the release or secretion of proteins,
- target proteins to apical surfaces,
- mediate endocytosis or protein turnover,
- play a role in protein sorting in the secretory and endocytic pathways, and
- participate in transmembrane signaling mechanisms.

In the meantime, several GPIs and GPI-anchored peptides have been synthesized as reviewed in [106].

2.2.4 Uncommon Linkages

In 1994 a previously unknown type of carbohydrate–protein linkage has been identified in human RNase 2 (identical to RNase U_s) with a mannose residue α-*C*-glycosidically attached to C2 of the indole ring of Trp-7 as a result of a post-translational modification (*Fig. 6*) [107,108,109]. In the meantime, *C*-mannosylation of tryptophan has been found in several proteins, including interleukin-12 [110] and properdin [111]. Recently, a method predicting the location of *C*-mannosylation sites in proteins was developed using a neural network approach [112]. Another uncommon linkage is the attachment of sugar to protein via a phosphodiester bridge [113] which had been first characterized in the lysosomal proteinase I of *Dictyostelium discoideum* in which α-*N*-acetylglucosamine-1-phosphate is bound to the side chain hydroxyl of serine [114,115]. Attachment of α-mannosyl phosphate to serine has been observed in several major proteins of *Leishmania* species [116].

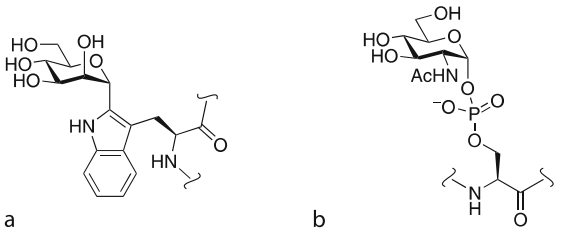


Figure 6
Uncommon carbohydrate–peptide linkages identified in (a) human RNase 2 and (b) proteinase I of *D. discoideum*

2.3 Methods of Glycoprotein-Glycans Analysis

Structural analysis of glycans contained in glycoproteins [117,118,119,120,121,122,123,124,125] is a challenging task due to the lack of a general oligosaccharide sequencing method like those available for peptides or oligonucleotides. The situation is further complicated by the phenomenon of microheterogeneity which makes only minute amounts of homogenous material available. Fractionation of individual glycoforms of a given glycoprotein can be achieved using capillary electrophoresis but is not generally applicable because the differences in physicochemical properties between neutral glycoforms are often relatively small. Therefore, protein glycosylation analysis is usually carried out after release of the glycans either from the whole glycoprotein or from glycopeptides obtained by proteolytic digestion. A general strategy consists of four steps:

1. Release of glycans from their conjugate polypeptide.
2. Labeling of released glycans.
3. Fractionation of glycan mixtures.
4. Sequencing of individual glycans.

2.3.1 Release of Glycans from Glycoprotein

In order to release glycans from a glycoprotein a general method is required that is independent of the protein to which the saccharides are attached. In this respect, chemical release is often preferred over the use of endoglycosidases or glycoamidases. Hydrazinolysis [126] has initially been described for release of *N*-glycans producing intact glycans with a free reducing terminus (*Fig. 7*) but later on it was shown that under controlled conditions the method is also suitable for *O*-glycans [127]. In the case of *O*-glycans, however, to some extent degradation of reducing-end monosaccharides (so-called peeling) remains a problem. As long as a reducing saccharide is not required at the cleavage point, *O*-glycans can be more mildly released in the form of oligosaccharide alditols by β-elimination and subsequent reduction upon treatment with alkaline borohydride [128]. This procedure, however, prevents subsequent labeling of the glycans by reductive amination (see below). To avoid this drawback, modified β-elimination procedures that deliver glycans with intact reducing ends have been developed [129,130].

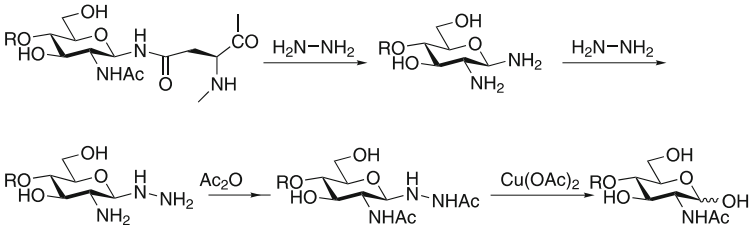


Figure 7
Hydrazinolytic release of *N*-linked glycans

Enzymatic release [131] of *N*-glycans is most frequently effected by the peptide *N*-glycanases (PNGases) F or A which, under denaturing conditions, generally cleave the Cγ–Nδ bond of glycosylated asparagines. Released are the intact *N*-linked glycans as glycosylamines, which are readily converted to regular glycans. Asparagine, in turn, is converted into aspartic acid resulting in a mass difference of 1 Dalton that may be used to assess the glycosylation site by mass spectrometry. In contrast to PNGase A, PNGase F is not able to cleave *N*-glycans containing an (α1–3)-linked fucose attached to the reducing-terminal GlcNAc residue. Alternatively, endoglycosidases can be applied which cleave between the first and second *N*-acetylglucosamine residue attached to asparagine. The various enzymes differ in their substrate specificities. Endo H, for example, cleave oligomannose-type and most hybrid-type glycans, whereas, Endo F2 and Endo F3 cleave certain complex-type chains [131]. Enzymatic release of *O*-glycans is much more difficult since most *O*-glycanases currently available are highly specific. Therefore, after PNGase treatment of *N*,*O*-glycoproteins, remaining *O*-linked glycans are often released by alkaline borohydride degradation.

2.3.2 Labeling and Fractionation of Glycans

Unless mass spectrometric analysis of the glycan pool is carried out, the released sugars need to be tagged to enable them to be detected during separation. Classically, the reducing end of the oligosaccharides had been reduced with alkaline sodium borotritide [132].

Today, radioactive labeling has been largely replaced by reductive amination with fluorescent compounds like 2-aminobenzamide [133], anthranilic acid [133], 8-aminonaphthalene-1,3,6-trisulfonic acid [134], 2-aminopyridine [135], 2-aminoacridone [136], or 1-aminopyrene-3,6,8-trisulfonate [137]. These fluorophores are compatible with a range of separation techniques including HPLC (normal phase, reversed phase, and weak anion exchange), high pH anion exchange chromatography (HPAEC), lectin-affinity chromatography, polyacrylamide gel electrophoresis (PAGE), capillary electrophoresis (CE), and gel permeation chromatography (GPC). Derivatization with ionizable functional groups has gained importance for high-resolution analysis of oligosaccharides by CE [138,139] and to enhance their otherwise low ionization efficiencies enabling sensitive detection of glycans by mass spectrometry even in the presence of peptides [140].

2.3.3 Sequencing of Glycans

For complete structural analysis of oligosaccharides, a combination of several physical, chemical, and biochemical techniques including NMR spectroscopy, mass spectrometry (MS), and chemical and enzymatic degradation is required. Mass spectrometry has developed a key technology for the analysis of glycans in glycomics [119,120,121,122,125]. The advantages of MS techniques are high sensitivity (detection limits of oligosaccharides in the low pico- to femtomolar range), high accuracy, and the possibility of being directly coupled to many separation methods. ESI-MS/MS, ESI-ion trap-MS^n, and MALDI-TOF/TOF-MS are currently the most frequently used methods. Structural information may be obtained from two types of fragmentation: cleavages of glycosidic bonds between monosaccharide units provide information on glycan sequence and branching, whereas cleavages within sugar rings (so-called cross-ring cleavages) provide extended information of branching and monosaccharide linkage positions. It is important, however, to keep in mind that mass spectrometry is not able to distinguish between different stereoisomeric sugars such as GlcNAc/GalNAc or Gal/Man/Glc all of which have the same mass. Therefore, knowledge of biosynthetic pathways of the organism which the glycans were derived from is used to reduce the number of possible glycans that are in agreement with measured data. Several algorithms for automated interpretation of MS data with integrated links to glycan databases have been developed and progress in this area has been reviewed [141].

Enzymatic analysis using highly specific exoglycosidases is a powerful means of determining the sequence of glycans especially if only very small amounts of material (picomoles or less) are available. This method is based on the determination of the susceptibility of a glycan to a series of exoglycosidases of defined specificity. In the case of the reagent array analysis method (RAAM) [142,143], the process has been automated. The glycan is divided into several aliquots and each aliquot is incubated with a defined mixture of exoglycosidases called a reagent array. Degradation occurs in each vial until a linkage is reached which is resistant to all exoglycosidases present in that mix. The remaining "stop point" fragments containing the labeled reducing end saccharides are characterized by size (GPC or MS). From these data, a computer program constructs the sequence of the glycan. Later, an improved approach was developed involving the simultaneous digestion of aliquots of a total pool of fluorescently labeled oligosaccharides with a series of multiple enzyme arrays [144]. For further details on the analysis of glycoproteins, the reader is referred to Chap. 8.5 of this book.

2.4 Structure of *N*-Glycans

Because of their biosynthesis, all *N*-glycoproteins share the common core pentasaccharide Man(α1–6)[Man(α1–3)]Man(β1–4)GlcNAc(β1–4)GlcNAc. The tremendous diversity of asparagine-linked oligosaccharides derives from attachment of several monosaccharides as well as up to five different carbohydrate chains, the so-called antennae, to the core. On the basis of the structure and the location of the glycans attached to the trimannosyl core, *N*-glycoproteins can be classified into three main groups (*Fig. 8*). These are:

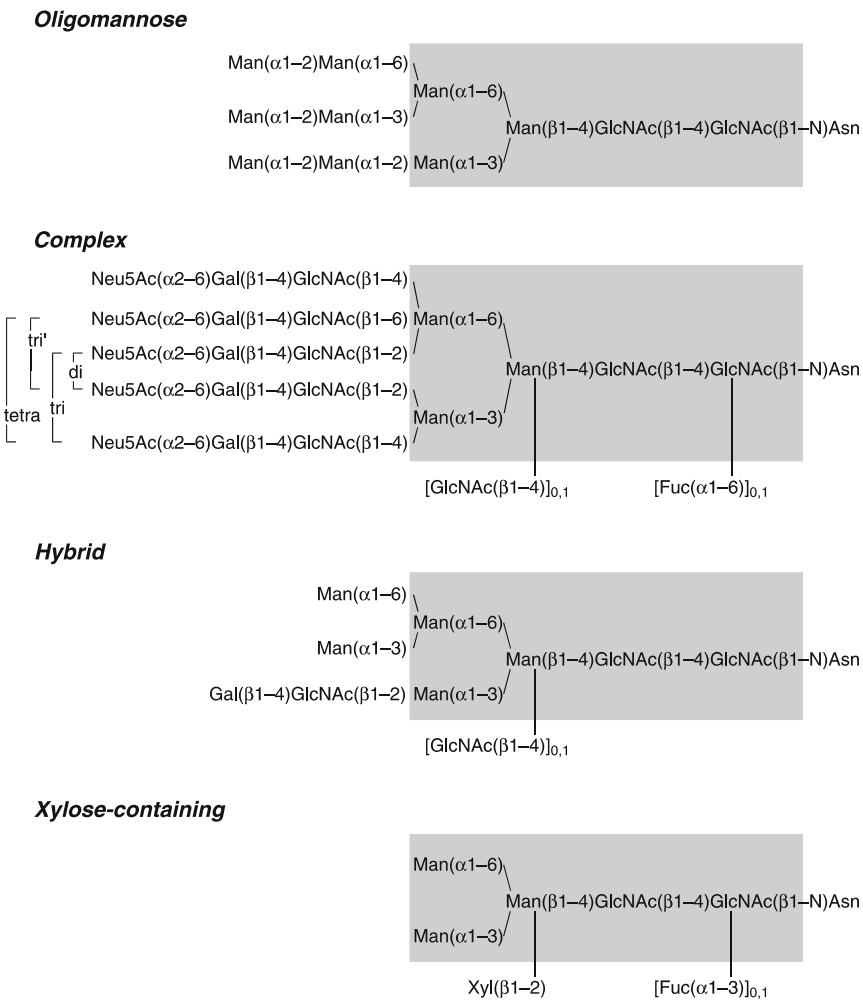


Figure 8
Examples of the main types of *N*-linked oligosaccharides of glycoproteins. The structure within the *gray box* represents the pentasaccharide core common to all *N*-glycans

- oligomannose (high mannose),
- complex, and
- hybrid type.

It has been suggested to add the xylose-containing type as a fourth group [28].
Oligomannose-type glycans usually contain two to six α-mannose residues bound to the core. The largest high mannose oligosaccharide thus contains nine mannose and two *N*-acetylglucosamine residues which were originally discovered in bovine thyroglobulin [145]. High molecular weight oligomannose chains with up to 100–200 mannose residues are, however, produced by yeasts.
Complex-type glycans contain no mannose residues other than those in the core. To the outer two α-mannose residues up to five units of the disaccharide Gal(β1–4)GlcNAc (*N*-acetyllactosamine) are β-linked. In the bi-, tri-, and tetraantennary glycans, these branches are attached to specific positions of the core as indicated in ❯ *Fig. 8*. The *N*-acetyllactosamine units are frequently substituted with sialic acid residues or additional repeats of *N*-acetyllactosamine, resulting in so-called poly-*N*-acetyllactosamines which are capped by sialic acids. Sialic acid is most commonly α-linked to the 3- or 6-position of galactose residues. Further diversification can be achieved by attachment of α1–6 linked fucose to the most inner core *N*-acetylglucosamine residue and of an *N*-acetylglucosamine residue linked β1–4 to the β-mannose of the core ("bisecting" GlcNAc). In addition, various monosaccharides can be found in the antennae.
Oligosaccharides of the hybrid typehave characteristic features of both complex-type and oligomannose-type glycans. One or two mannose residues are α-linked to the Man(α1–6) branch of the core pentasaccharide and one or two antennae like those found in complex-type oligosaccharides are β-linked to the Man(α1–3) branch of the core. This type of glycan results from partial processing of oligomannose-type glycans produced during biosynthesis and subsequent addition of sugars.
The fourth group of *N*-glycans is named xylose-containing and is characterized by a xylose residue linked β1–2 to the β-mannose of the core. Often, there is also fucose α1–3 linked to the innermost *N*-acetylglucosamine of the core.

2.4.1 Diversity of *N*-Glycans

As mentioned earlier, the structural diversity of glycans is provided by the variation of monosaccharides attached to the core, the degree of branching (antennary), and, particularly, the tremendous variation in the structure of the antennae. Some side chains commonly occurring in complex *N*-glycans are depicted in ❯ *Fig. 9*. These structures can be roughly classified into five groups. The first group is represented by *N*-acetyllactosamine which is the starting point for further modification but still being frequently found. Group two consists of side chains terminated by sialic acid, fucose, or galactose. Once these modifications have taken place, no further chain elongation is possible except for formation of polysialic acids. Following sialylation it is also still possible to transfer fucose or sialic acid to the *N*-acetylglucosamine residues, resulting in some of the structures of group three, e. g. Neu5Ac(α2–3)Gal(β1–4)[Fuc(α1–3)]GlcNAc (sialyl Lewis x), Neu5Ac(α2–3)Gal(β1–3)[Fuc(α1–4)]GlcNAc (sialyl Lewis a), or Neu5Ac(α2–3)Gal(β1–3)[Neu5Ac(α2–

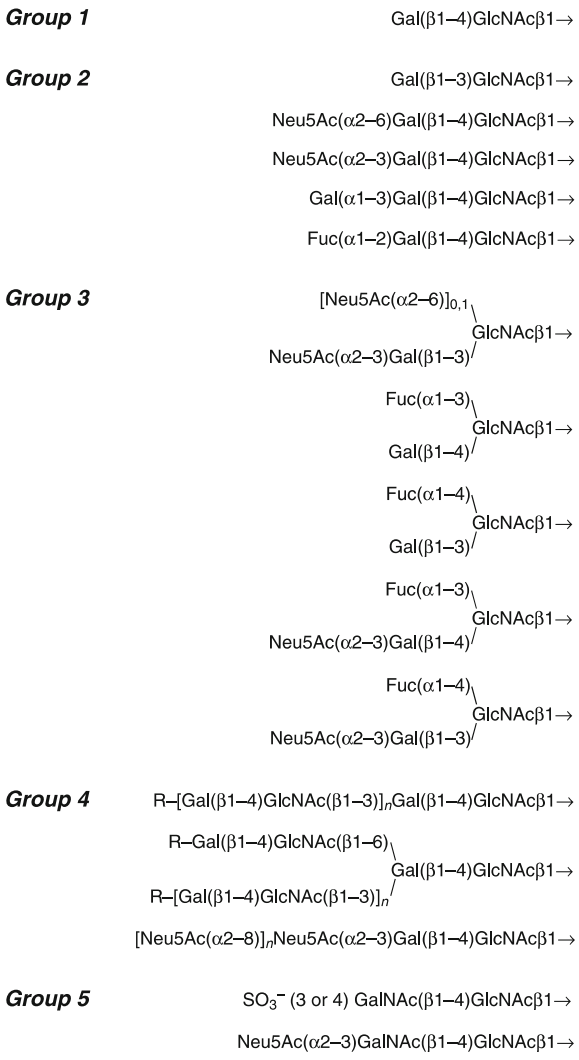


Figure 9
Commonly occurring structures of complex-type *N*-glycan side chains. The *arrows* indicate the point of attachment to the pentasaccharide core

6)]GlcNAc. On the other hand, fucosylation of a subterminal *N*-acetylglucosamine residue halts the chain elongation in nonsialylated, fucosylated form. Group four consists of long-chain structures containing either linear or branched poly-*N*-acetyllactosamine or α2–8 linked sialic acid repeatedly added to α2–3 linked sialic acid, forming polysialyl side chains. Group five side chains contain *N*-acetylgalactosamine instead of galactose. These structures may be further modified by sialylation or sulfation. There are many more variations in *N*-glycan chains identified to date and many more are expected to be found in future studies.

2.4.2 Cell-Type Specific Expression of *N*-Glycans

As described above, the enormous diversity of *N*-glycans is largely provided by variation of structures in the antennae. However, expression of specific carbohydrate epitopes, so-called antigenic determinants, is highly controlled and often cell-type specific. Critical findings revealed that poly-*N*-acetyllactosamines that are contained in many complex-type *N*-glycoproteins (but also in *O*-glycoproteins and glycolipids) provide a preferable backbone for many of the cell-type specific glycosylations, such as the ABH(O), I/i, and Lewis (Le^a, Le^x, and sialylated forms thereof) blood group and differentiation antigenic determinants [146]. Poly-*N*-acetyllactosamine repeats are not uniformly distributed among different antennae attached to the trimannosyl core. Because of the branch specificity of β1–3-*N*-acetylglucosaminyltransferase—the key enzyme for formation and elongation of poly-*N*-acetyllactosamine chains—poly-*N*-acetyllactosamine extensions are more common at the α1–6 linked mannose of the core. Tetraantennary glycans are particularly good acceptor substrates for poly-*N*-acetyllactosamine chain formation. Termination of chain elongation by sialylation leads to structures summarized in ❯ *Fig. 10*.

Glycosylation patterns of glycoproteins (and glycolipids) in human erythrocytes and granulocytes are well-established examples for cell-type specific expression of glycans. Some important carbohydrate antigenic determinants are listed in ❯ *Table 2*. Erythrocytes and granulocytes directly differentiate from the same precursor stem cells. In fetal erythrocytes, poly-*N*-acetyllactosamine is linear and expresses blood type i activity. In adult erythrocytes, poly-*N*-acetyllactosamines are branched at galactose, forming Gal(β1–4)GlcNAc(β1–3)[Gal(β1–4)GlcNAc(β1–6)]Gal(β1–4)GlcNAc, the blood group I antigen. The change from i to I specificity occurs gradually during the first year of life. Terminal galactose residues of the branched poly-*N*-acetyllactosamine chains of erythrocytes are then modified by α1–2 linked fucose, forming the H antigen. This reaction is catalyzed by an α1–2-fucosyltransferase. The H structure, Fuc(α1–2)Gal(β1–4/3)GlcNAc, is present in *O*-type blood group individuals. This structure now serves as the acceptor for α1–3-*N*-acetylgalactosaminyltransferase or α1–3-galactosyltransferase, forming A or B blood group antigens, respectively.

In granulocytes, no branching of poly-*N*-acetyllactosamines is observed. α1–3-Fucosyltransferase-catalyzed modification of the linear chains leads to the Lewis x antigen. If α2–3-sialyltransferase acts before fucosylation, sialyl Lewis x is formed. These carbohydrate structures are involved in a number of cell-cell interactions during the inflammatory response and

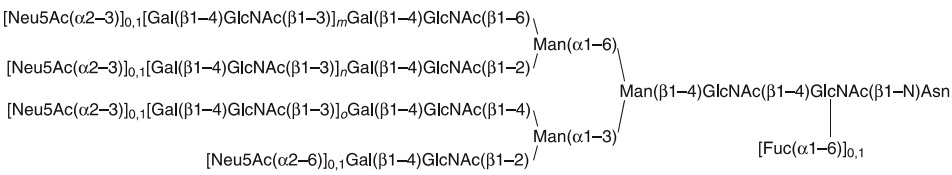


Figure 10

Representative structures of tetraantennary poly-*N*-acetyllactosamine containing *N*-glycans based on structures reported in various cells. The length of the *N*-acetyllactosamine repeats generally decreases in the order *m* > *n* > *o*. Sialic acids attached to poly-*N*-acetyllactosamine chains are usually α2–3 linked whereas α2–6 sialylation preferentially occurs at the side chain linked β1–2 to the Man(α1–3) branch of the core

Table 2
Antigenic determinants. LacNAc stands for Gal(β1–4)GlcNAc

Determinant structure	Name
Fuc(α1–2)Gal(β1–4/3)GlcNAc	Blood group H (O), type 1 or 2[a]
GalNAc(α1–3)[Fuc(α1–2)]Gal(β1–4/3)GlcNAc	Blood group A, type 1 or 2[a]
Gal(α1–3)[Fuc(α1–2)]Gal(β1–4/3)GlcNAc	Blood group B, type 1 or 2[a]
Gal(β1–3)[Fuc(α1–4)]GlcNAc	Blood group Lewis a (Le^a)
Neu5Ac(α2–3)Gal(β1–3)[Fuc(α1–4)]GlcNAc	Sialyl Lewis a (sLe^a)
Fuc(α1–2)Gal(β1–3)[Fuc(α1–4)]GlcNAc	Blood group Lewis b (Le^b)
Gal(β1–4)[Fuc(α1–3)]GlcNAc	Lewis x (Le^x)
Neu5Ac(α2–3)Gal(β1–4)[Fuc(α1–3)]GlcNAc	Sialyl Lewis x (sLe^x)
Fuc(α1–2)Gal(β1–4)[Fuc(α1–3)]GlcNAc	Lewis y (Le^y)
$[LacNAc(\beta1–3)]_n$	Blood group i
LacNAc(β1–3)[LacNAc(β1–6)]LacNAc	Blood group I

[a] Type 1 contains the sequence Gal(β1–3)GlcNAc, type 2 contains the sequence Gal(β1–4)GlcNAc

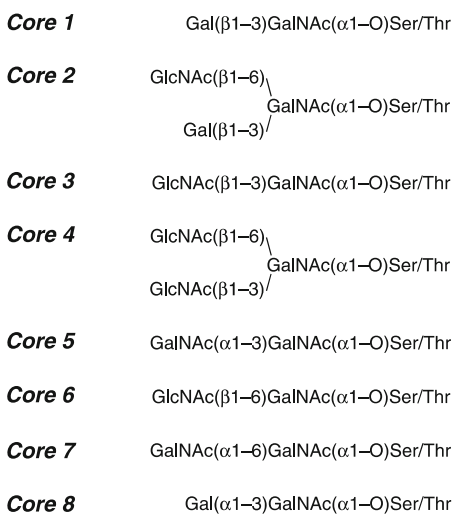
malignant transformations [147,148]. Production of Lewis x and sialyl Lewis x is restricted to myeloid cell lineage and only granulocytes and monocytes are enriched with them in blood cells. In contrast, expression of the ABH blood group antigens is restricted to erythroid cells, and no other blood cells contain ABH blood group antigens. This cell-type specific expression of glycans is caused by the presence of certain key glycosyltransferases that are essential for formation of the corresponding carbohydrate structures and of which expression is restricted to that particular cell type.

2.5 Structure of *O*-Glycans

The biosynthesis of *O*-glycans differs fundamentally from that of the *N*-glycans. *O*-Glycosylation does not begin with the transfer of an oligosaccharide from a dolichol precursor to the nascent peptide chain which is then further processed. Instead, *O*-glycosylation is mainly a post-translational and post-folding event and is initiated by glycosyltransferase-catalyzed addition of a single monosaccharide to a side chain hydroxyl of the protein. Depending on the carbohydrate–peptide linkage, various types of *O*-glycans can be distinguished. By definition, mucin-type *O*-glycoproteins are characterized by the GalNAc(α1–O)Ser/Thr bond [149,150]. This structural element, often referred to as the Tn antigen, was first identified in the mucins but is also found in other glycoproteins. Subsequent stepwise enzymatic glycosylation of the monoglycosylated polypeptide and further modification by sulfation and acetylation leads to a high degree of heterogeneity. Mucin-type *O*-glycan chains typically consist of one to 20 monosaccharide residues and they carry many of the blood group antigens and recognition signals required for intercellular and molecular interactions as well as cancer-associated and differentiation antigens.

2.5.1 Mucin-Type *O*-Glycans

Biosynthesis of mucin-type *O*-glycans starts with the transfer of *N*-acetylgalactosamine (GalNAc) from UDP-GalNAc to the side chain of serine or threonine catalyzed by an *N*-acetylgalactosaminyltransferase. More than 20 homologous *N*-acetylgalactosaminyltransferases are known to date [151]. These transferases are expressed in a tissue-specific manner and have overlapping but different substrate specificities. As a consequence, the site-specificity of mucin *O*-glycosylation is tissue specific and no simple consensus sequence can be found. However, based on statistical analysis of occupied *O*-glycosylation sites, some general rules can be deduced [9]. *O*-Glycosylation occurs preferentially at exposed regions of the protein surface like β-turns or regions with extended conformations. Such areas are usually of low hydrophobicity and, consequently, seldom contain amino acids like Trp, Leu, Ile, and Phe. Instead, Ser, Thr, and Pro are commonly found before and after occupied *O*-glycosylation sites. Also, no bulky amino acids are found near *O*-glycosylation sites, probably due to steric hindrance. However, in contrast to *N*-linked glycans which are usually well separated from each other, *O*-glycans frequently occur in clusters on short peptide sections consisting of repeating units of Ser, Thr, and Pro. It seems to be a general feature that Thr is glycosylated more efficiently than Ser. On the basis of statistical analyses, algorithms allowing the prediction of *O*-glycosylation sites have been developed [152]. The location of the subcellular compartment where *O*-glycosylation is initiated is still controversial and may depend on the type of *N*-acetylgalactosaminyltransferase. Several suggestions that have been made include subregions of the endoplasmic reticulum (ER), a proximal Golgi compartment, an intermediate ER-Golgi compartment (ERGIC), and the Golgi apparatus [153,154,155]. Subsequent stepwise enzymatic elongation of the monoglycosylated polypeptide then leads to several core structures, eight of which have been identified to date (◉ *Fig. 11*). However, under certain circumstances (e. g. in particular transformed cells), elongation does not occur and the Tn



◘ Figure 11
Core structures of *O*-glycans. Core 1 is also called T antigen

antigen is exposed which might be modified to sialyl Tn (sTn) (Neu5Ac(α2–6)GalNAc(α1–O)Ser/Thr) [156,157,158].

Depending on the cell type and its activation and differentiation status, the core structures can be further elongated or terminated by various glycosyltransferases resulting in a large number of *O*-glycans, of which some are depicted in ❯ *Fig. 12*. Many of these glycosyltransferases have been isolated and their genes have been cloned [159,160]. There are glycosyltransferases which specifically act on *O*-glycans (mainly those responsible for assembly of the core structures) and there are others with an activity restricted to *N*-glycans. However, a third group is able to assemble both *N*- and *O*-glycans. The antigenic determinants mentioned in ❯ *Sect. 2.4.2* are generated by such transferases, i. e. the structures listed in ❯ *Table 2* are found on both *N*- and *O*-glycans.

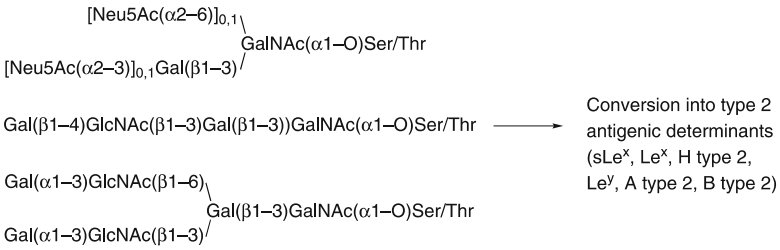
Core 1 and core 2 are the most common core structures in mucins and other secreted and cell surface glycoproteins. Core 1 (the T antigen) is usually not exposed but is monosialylated (Neu5Ac(α2–3)Gal(β1–3)GalNAc(α1–O)Ser/Thr and Neu5Ac(α2–6)[Gal(β1–3)]GalNAc(α1–O)Ser/Thr) or disialylated (Neu5Ac(α2–6)[Neu5Ac(α2–3)Gal(β1–3)]GalNAc(α1–O)Ser/Thr). Core 1 as well as other *O*-glycans is often elongated by *N*-acetyllactosamine, which is a substrate for conversion into the type 2 antigenic determinants sLe^x, Le^x, H type 2, Le^y, A type 2, and B type 2. Core 1 is, however, a poor acceptor for poly-*N*-acetyllactosamine formation. To efficiently achieve this type of extension, core 1 first has to be transformed into core 2. Then, poly-*N*-acetyllactosamine formation can take place at the new GlcNAc(β1–6) branch. Thus, the expression of core 2 *N*-acetylglucosaminyltransferase (core 2 GlcNAcT), the enzyme responsible for the conversion of core 1 into core 2, is rate limiting for the biosynthesis of poly-*N*-acetyllactosamine extensions. The fact that poly-*N*-acetyllactosamine extensions provide an excellent backbone for the expression of antigenic determinants involved in numerous recognition events highlights the biological importance of core 2 GlcNAcT.

So far, core 3 and 4 have been found only in mucins. Core 5 is contained in glycoproteins from several species and has been reported in human adenocarcinoma [161] and meconium [162,163]. It has been speculated that core 6 might be generated by a β-galactosidase-catalyzed degradation of core 2 [159]. Core 7 occurs in bovine submaxillary mucin [164]. All core structures may be unsubstituted or elongated and may carry terminal antigens. Elongation can result in type 1 chains (containing the sequence Gal(β1–3)GlcNAc) or type 2 chains (containing the sequence Gal(β1–4)GlcNAc, e. g. i antigen, poly-*N*-acetyllactosamines) and branched structures (I antigen). Neu5Ac(α2–3 and –6), Fuc(α1–2, –3, and –4), GalNAc(α1–3 and –6), GalNAc(β1–4), GlcNAc(α1–4), Gal(α1–3), and sulfated residues are typically found as terminal structures on *O*-glycans although sialic acids, sulfate, and Fuc(α1–3 or –4) may be also attached to internal carbohydrate residues. Many of the terminal structures are of the Lewis and ABH blood group system (❯ *Table 2*) and are also found on *N*-glycoproteins and glycolipids.

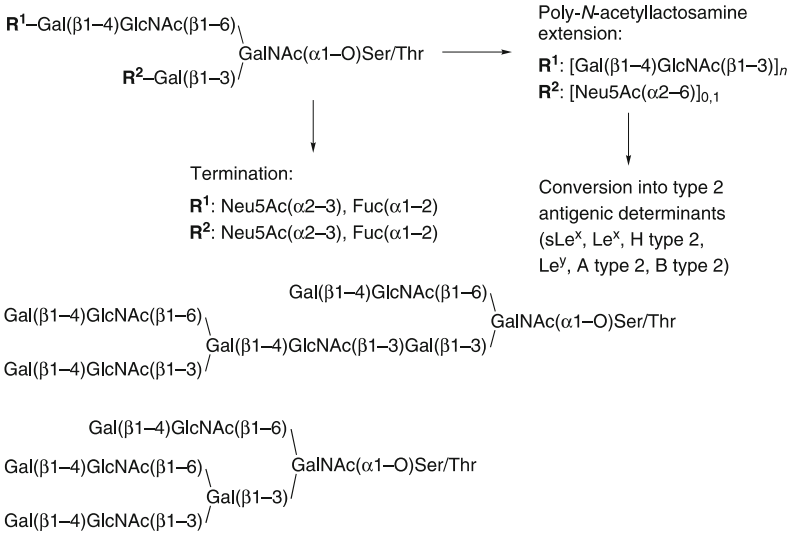
2.5.2 Non-Mucin *O*-Glycans

Beside the GalNAc(α1–O)Ser/Thr linkage of the mucin-type *O*-glycans a number of additional *O*-linkages have been identified. These include mainly fucose α-*O*-linked to Ser/Thr and glucose β-*O*-linked to Ser found in the epidermal growth factor domains of different proteins [58] and *N*-acetylglucosamine β-*O*-linked to Ser/Thr on cytoplasmic and nuclear proteins.

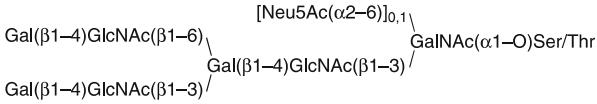
Structures generated from Core 1



Structures generated from Core 2



Structures generated from Core 3



Structures generated from Core 4

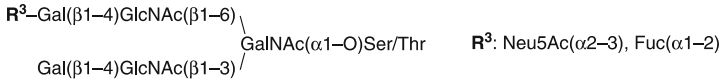
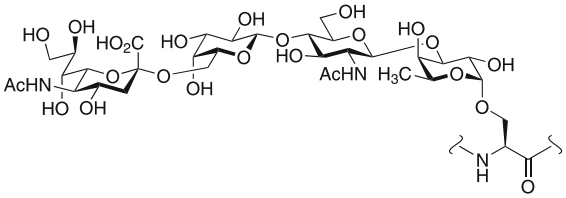
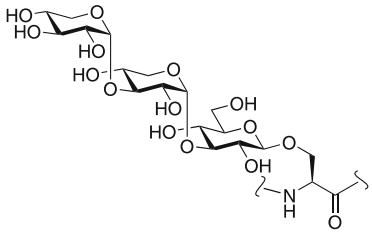


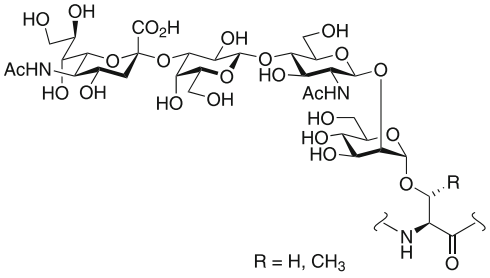
Figure 12
Structures of *O*-glycans commonly found in mucin-type *O*-glycoproteins



a



b



c

Figure 13
Some structures of non-mucin *O*-glycans (cf. text)

O-Linked fucose has been found on the EGF domains of urokinase, tissue plasminogen activator (tPA), factor VII, and factor XII, attached to the consensus sequence Cys-Xaa-Xaa-Gly-Gly-Ser/Thr-Cys. The fucosyltransferase responsible for the initiation of this type of glycosylation has been cloned [165]. On human (but not bovine) factor IX the tetrasaccharide Neu5Ac(α2–6)Gal(β1–4)GlcNAc(β1–3)Fuc(α1–O)Ser (*Fig. 13a*) was found [166,167]. In other EGF domains only Fuc(α1–O)Ser/Thr was detected.

On the same EGF modules glucose β-*O*-linked to the Ser of the conserved consensus sequence Cys-Xaa-Ser-Xaa-Pro-Cys occurs. In several proteins like the human and bovine blood clotting factors VII and IX, human and bovine protein Z, tPA, thrombospondin, and murine fetal antigen 1 (mFA1) Xyl(α1–3)Xyl(α1–3)Glc(β1–O)Ser (*Fig. 13b*) and substructures thereof (Xyl(α1–3)Glc(β1–O)Ser and Glc(β1–O)Ser) have been identified. The transferase activity responsible for addition of the second xylose has been detected [168].

The *N*-acetylglucosaminyl (β1–O) serine/threonine linkage is common on cytoplasmic and nuclear proteins [54,55,56,57]. No other sugars are connected to the GlcNAc residue. Formation of this type of modification is independent of the activity of the glycosylation machinery

in the endoplasmic reticulum and the Golgi apparatus. Addition of *O*-GlcNAc is catalyzed by a highly conserved *O*-GlcNAc transferase cloned independently by two groups [169,170]. The specificity of this enzyme resembles that of proline-directed kinases. *O*-GlcNAc is found in serine/threonine-rich sequences near Pro and Val residues. Often a negative charge is located nearby. Since the turnover of the *O*-GlcNAc glycan is faster than that of the peptide to which the sugar is attached, *O*-GlcNAc glycosylation is assumed to be reversible. Further evidence for this is the existence of a soluble *N*-acetyl-β-D-glucosaminidase [171]. It is likely that *O*-GlcNAc glycosylation plays a regulatory role that is analogous to protein phosphorylation. The mannosyl (α1–O) Ser/Thr linkage is uncommon in mammalian cells. In bovine peripheral nerve α-dystroglycan clusters of Neu5Ac(α2–3)Gal(β1–4)GlcNAc(β1–2)Man(α1–O)Ser/Thr (❯ *Fig. 13c*) have been observed in Ser/Thr-rich domains [64]. A similar glycan (Neu5Ac(α2–3)Gal(β1–4)GlcNAc(β1–3)Man(α1–O)Ser/Thr) has been found in rat brain proteoglycans [65,66].

2.6 Proteoglycans

Proteoglycans consist of glycosaminoglycan chains covalently attached to a protein backbone [172,173,174,175,176,177]. Therefore, they can be classified as glycoproteins. However, for historical reasons and because they differ substantially from other glycoproteins, they are usually treated as a separate class of compounds. Glycosaminoglycans are linear polysaccharides consisting of hexosamine residues like *N*-sulfonylglucosamine (GlcN*S*), *N*-acetylglucosamine (GlcNAc), or *N*-acetylgalactosamine (GalNAc) alternating with glucuronic acid (GlcA) or L-iduronic acid (IdoA) or galactose (Gal) residues. Different types of glycosaminoglycans are characterized by specific disaccharide repeat patterns. These oligosaccharides are substituted to varying degrees with sulfate linked to free amino and/or hydroxyl groups and range in size from ca. 15 disaccharide units to several hundred. The glycosaminoglycans contained in proteoglycans are chondroitin sulfate, dermatan sulfate, heparan sulfate, heparin, and keratan sulfate. Usually one type of glycosaminoglycan predominates but hybrid proteoglycans do exist. Another glycosaminoglycan (hyaluronan or hyaluronic acid) is not covalently attached to protein. A striking feature of all glycosaminoglycans is their negative charge caused by their sulfate and carboxyl groups which distinguishes these molecules from other glycoproteins. In addition to the glycosaminoglycan chains some proteoglycans contain *O*- and *N*-linked oligosaccharides similar or identical to those of various classes of glycoproteins. Proteoglycans are widely distributed in the animal and plant kingdom and they possess both structural and regulatory roles [177,178,179,180]. They are found in the extracellular matrix (ECM) and associated with specialized structures of the ECM like basement membranes. They also occur as part of plasma membranes where they function as receptors and co-receptors, and they are found intracellularly and intravesicularly in many hematopoietic cells. Proteoglycans have important physiological and homeostatic roles, e. g. during development, inflammation, and the immune response. Glycosaminoglycans such as heparan sulfate may encode information that is (cell) specific, spatially and temporally regulated, and instructive. Many diseases like chondrodystrophies [181], atherosclerosis [182], different types of cancer [183], or Alzheimer's disease [184] are related to abnormalities in the biosynthesis and processing of proteoglycans [172].

2.6.1 Carbohydrate–Peptide Linkage Region

Chondroitin sulfate, dermatan sulfate, heparan sulfate, and heparin are linked with their reducing end to the common core tetrasaccharide GlcA(β1–3)Gal(β1–3)Gal(β1–4)Xyl with the Xyl β-*O*-linked to a Ser residue (➤ *Fig. 14*), usually adjacent to a Gly in the protein backbone [172,173,174,177]. The GlcA residue of this tetrasaccharide is considered as part of the linkage region because the transferase involved in the addition of this residue to the Gal-Gal disaccharide is different from the enzyme that catalyzes the incorporation of GlcA into the rest of the glycosaminoglycan. The xylose may be phosphorylated and one or both galactose residues may be sulfated.

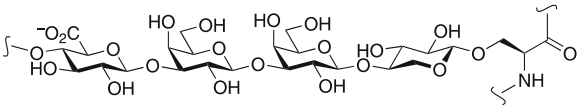


Figure 14
Structure of the linkage region GlcA(β1–3)Gal(β1–3)Gal(β1–4)Xyl(β1–O)Ser of proteoglycans

Skeletal keratan sulfate and corneal keratan sulfate are linked to core protein through *O*-linked and *N*-linked oligosaccharides, respectively, of the general types described in ➤ *Sect. 2.4* and ➤ *Sect. 2.5*.

2.6.2 Structure of Glycosaminoglycans

Hyaluronan (hyaluronic acid) is the simplest glycosaminoglycan. It consists of alternating GlcA and GlcNAc residues which are (β1–3) and (β1–4) linked, respectively (➤ *Fig. 15*) [172,173,174,177]. Hyaluronan is not sulfated and is not bound to protein, thus it is not a component of proteoglycans.

Chondroitin sulfate glycosaminoglycans are composed of the disaccharide repeat unit GlcA(β1–3)GalNAc(β1–4) with variable amounts of sulfation at the positions indicated in ➤ *Fig. 15*. Dermatan sulfate is formed from the precursor chondroitin sulfate by action of an appropriate uronosyl epimerase which converts GlcA into IdoA, giving rise to the disaccharide repeat unit IdoA(α1–3)GalNAc(β1–4) [185]. Thus, both glycosaminoglycans are found on the same protein core. However, there is still no unambiguous convention as to what the proportion of IdoA must be to name a glycosaminoglycan dermatan sulfate. Ordinarily, IdoA is only formed adjacent to 4-sulfated GalNAc and not adjacent to 6-sulfated or non-sulfated GalNAc.

Heparan sulfate glycosaminoglycans and heparin are more complex. They contain two main types of disaccharide residues, although the proportions differ greatly between heparan sulfate and heparin. Heparan sulfate consists of areas of non-sulfated GlcA(β1–4)GlcNAc(α1–4) and areas of sulfated IdoA(α1–4)GlcN*S*(α1–4) disaccharide repeat units with the latter being obtained from the former by action of modifying enzymes during biosynthesis. Heparin differs from heparan sulfate by its higher content of IdoA, *N*-sulfate, and *O*-sulfate with most of the disaccharide units containing 2-sulfated IdoA alternating with GlcN*S* which is partially 6-sulfated and occasionally 3-sulfated as well. On some heparan sulfates occasional 3-sulfation of

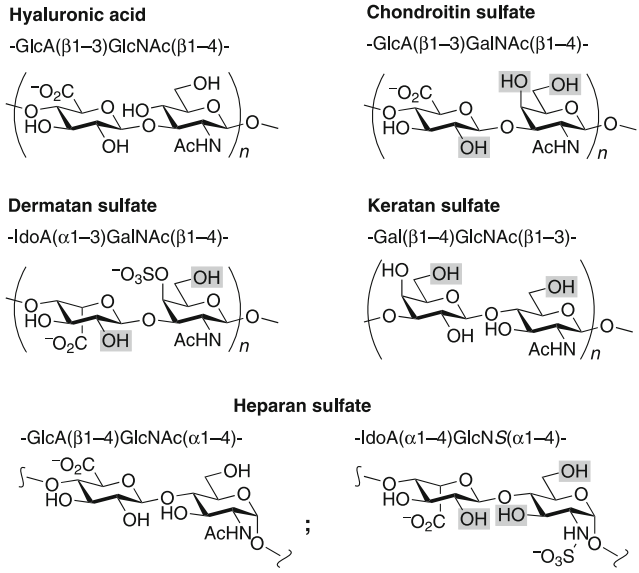
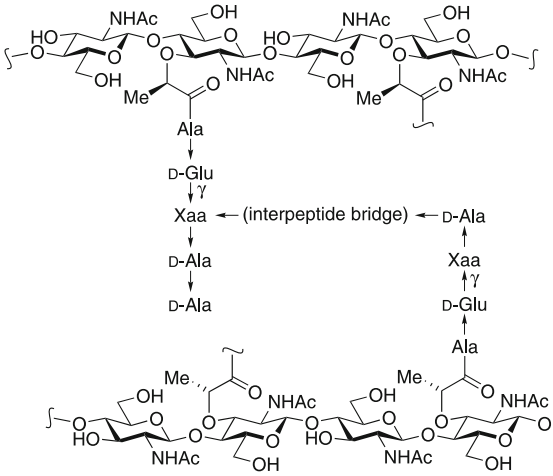


Figure 15
Disaccharide repeats of different glycosaminoglycans. Hydroxyl groups which may be modified by sulfation are marked gray

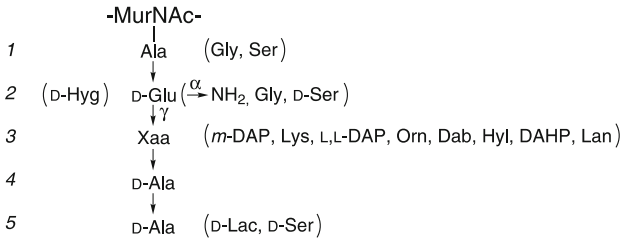
GlcN*S* is also found. Although heparan sulfate and heparin are structurally similar one should keep in mind that both are found on different core proteins.
Keratan sulfate glycosaminoglycans are sulfated poly-*N*-acetyllactosamines of the type found in *N*- and *O*-glycoproteins. The degree of sulfation which takes place at the primary hydroxyl groups is variable along the oligosaccharide chains with a higher degree at the non-reducing end.

3 Peptidoglycan

Peptidoglycan (synonym: murein) [186,187,188] is the essential cell wall polymer of most eubacteria, both Gram-positive and Gram-negative. Although it is not regarded as a glycopeptide it will be briefly treated in this chapter since it consists of carbohydrates covalently attached to peptides. Furthermore, knowledge of peptidoglycan structure is helpful for an understanding of the mode of action of the glycopeptide antibiotics of the vancomycin group. Peptidoglycan is composed of three structural features: a matrix of polysaccharide chains, the pentapeptide side chains, and the interstrand peptide cross-links, called interpeptide bridges (*Fig. 16a*). The glycan component of all peptidoglycans consists of alternating β1–4 linked GlcNAc and *N*-acetylmuramic acid (MurNAc) (a GlcNAc with a lactyl ether at O-3) residues. Average glycan-chain lengths of ca. 8 to 40 disaccharide units have been reported depending on the method used for determining them [188,189,190]. This structure is common to all bacterial peptidoglycans examined, with only minor variations. In *Staphylococcus aureus* for



a



b

Figure 16

(a) Part of the common primary structure of bacterial peptidoglycan. *Arrows* indicate the polarity of peptide chains pointing from the *N*-terminus towards the *C*-terminus. (b) Variations in the pentapeptide chain. Residues in *parentheses* may replace corresponding amino acids or, in the case of α-carboxylate modification of γ-D-Glu, may be missing. *Abbreviations*: Dab, 2,4-diaminobutyric acid; DAHP, 2,6-diamino-3-hydroxypimelic acid; DAP, diaminopimelic acid; Hyg, *threo*-3-hydroxyglutamic acid; Hyl, hydroxylysine; Lac, lactate; Lan, lanthionine

example, up to 50% of the MurNAc residues are 6-*O*-acetylated [191] and in *Micrococcus lysodeikticus*, up to 40% of the MurNAc residues are not attached to peptide chains [192].

Attached to the carboxylate of MurNAc is a pentapeptide with the common sequence Ala-γ-D-Glu-Xaa-D-Ala-D-Ala (non-specification of the amino acid configuration implies the natural L-form), where Xaa is *meso*-diaminopimelic acid (*m*-DAP) for Gram-negative bacteria and Lys for most Gram-positive bacteria, with some variations [193]. During peptidoglycan biosynthesis, the pentapeptide is cross-linked through amino acid Xaa to another peptide strand. This transpeptidation involves attack of the ε-amino group of the *m*-DAP or Lys residue (or amino terminus of the interpeptide bridge) onto the penultimate D-Ala of another chain, resulting in loss of the terminal D-Ala of the second peptide chain and peptide bond formation. As mentioned, variations in the pentapeptide side chains are possible and they are summarized in ❯ *Fig. 16b*. Variations in position 1 are rare and γ-D-Glu is found universally at position 2.

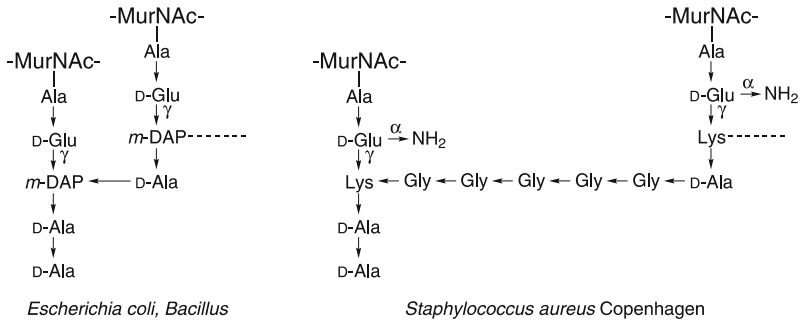


Figure 17
Peptide cross-links of *Escherichia coli*, *Bacillus*, and *Staphylococcus aureus* Copenhagen peptidoglycan

However, its α-carboxylate is often amidated or linked to an additional amino acid and in *Microbacterium lacticum* γ-D-Glu is replaced by *threo*-3-hydroxy-D-glutamic acid (3-Hyg). Most variation is found in position 3. If *m*-DAP (D,L-configuration) is incorporated, the L center is found in the pentapeptide chain and the D center in the side chain. Positions 4 and 5 (D-Ala-D-Ala) were for a long time thought to be invariant. However, the incidence of bacterial resistance to vancomycin-type glycopeptide antibiotics which recognize specifically the *N*-acyl-D-Ala-D-Ala terminus led to the discovery that these resistant strains contain altered substituents at position 5 as indicated in *Fig. 16b* [194,195,196].

Peptidoglycan strands are connected to varying degrees by a network of cross-links formed between the pentapeptide side chains. The structures of these cross-links which are quite heterogeneous have been reviewed [186,189,193] and this feature has been used for taxonomic classification of bacteria [193]. Here only two examples are presented (*Fig. 17*). In *Escherichia coli* and strains of *Bacillus* the simplest type of linkage is realized, i. e. the direct attachment of the ε-amino group of *m*-DAP to D-Ala at position 4 of the second peptide. If L-lysine is found in position 3 of the peptide chain, which is the case for most Gram-positive bacteria, usually 1–5 amino acid residues (L-configured or glycine) are incorporated between the ε-amino function of Lys and the D-Ala of the second peptide . As such, the interpeptide bridge of the well-studied *Staphylococcus aureus* Copenhagen strain consists of pentaglycine. However, other strains of *Staphylococcus* contain Ala, Ser, or Thr.

Currently, there is no generally accepted model for the three-dimensional structure of peptidoglycan. Two alternative models in which the glycan strands are arranged either parallel (classical model) or perpendicular ('scaffold' model) to the cytoplasmic membrane are discussed [188,190]. However, the models do not need to be mutually exclusive. It may also be that the cell wall exists as a mosaic of microdomains with different structures, or perhaps new cell wall is inserted in one form before being converted to another [197].

4 Glycopeptide Antibiotics

The glycopeptide antibiotics are glycosylated secondary metabolites of bacteria and fungi that are synthesized by non-ribosomal peptide synthetases. The term glycopeptide antibiot-

ic often solely refers to vancomycin and its analogues excluding other compounds, such as the bleomycins, ramoplanin, and the mannopeptimycins, that also belong to the group of glycopeptide antibiotics [198]. This section, however, will focus on the vancomycin group because it represents the biggest class of glycopeptide antibiotics.

4.1 Discovery and Medical Use

Vancomycin was discovered in the mid-1950s in a soil sample from the jungles of Borneo by coworkers of the American pharmaceutical company Eli Lilly [199]. It is produced as a secondary metabolite by the microorganism *Streptomyces orientalis* (reclassified as *Amycolatopsis orientalis*) [200] and displays antibiotic activity against the majority of Gram-positive bacteria. The discovery of vancomycin was followed a year later by the isolation of ristocetin [201]. These two antibiotics were recognized as belonging to a chemical class of antibiotics called glycopeptides [194,198,202,203,204]. Vancomycin was first used clinically in 1959. However, toxicity problems due to inadequacies during its purification precluded its widespread use in the early years and vancomycin was for the moment overshadowed by semisynthetic penicillins and later cephalosporins. However, with improved purification techniques and as the result of the emergence of multiple resistant and methicillin resistant *Staphylococcus aureus* (MRSA), vancomycin has become the drug of choice in the fight against these 'superbugs' [203]. Ristocetin appeared to be a promising antibacterial agent in the late 1950s but it was withdrawn following the death of several patients [205]. Avoparcin has shown growth-promoting activity in farm animals and has been marketed in Europe since 1976 as a feed additive. It is also used in the dairy industry for improvement of milk production of lactating dairy cows. Today two members of the group of the glycopeptide antibiotics are in clinical use—vancomycin and teicoplanin. Together with the aminoglycoside gentamycin, they are the antibiotics of last resort in our hospitals. However, in the meantime vancomycin resistant *S. aureus* strains are also being observed [206,207].

4.2 Structure and Classification

The first structural work on vancomycin were degradation studies carried out by Marshall [208] followed by extensive NMR examinations in the laboratory of Williams [209]. On the basis of these pioneering studies and on the X-ray analysis of the degradation product CDP-1 [210], Harris and Harris published the complete structure of vancomycin in 1982 [211]. The determination of the vancomycin structure then served as the basis for the structural characterization of hundreds of natural and semisynthetic glycopeptide antibiotics.

The glycopeptide antibiotics of the vancomycin group consist of a peptide backbone of 7 amino acids (see ❯ *Fig. 18* for selected examples). The side chains of amino acids 2 and 4 as well as 4 and 6 and in most cases also 5 and 7 are linked to each other, either via biaryl ether or direct biaryl bonds. The resulting structural elements give rise to the phenomenon of atropisomerism. Bound to this peptide core are 1–4 saccharides (mono- to tetrasaccharides). The vancomycin-type glycopeptide antibiotics can be subdivided in four or five groups based on chemical structure [198,204,212]. Group I, or the vancomycin type, has aliphatic amino acids at positions 1 and 3 while groups II–IV, illustrated by avoparcin, ristocetin, and

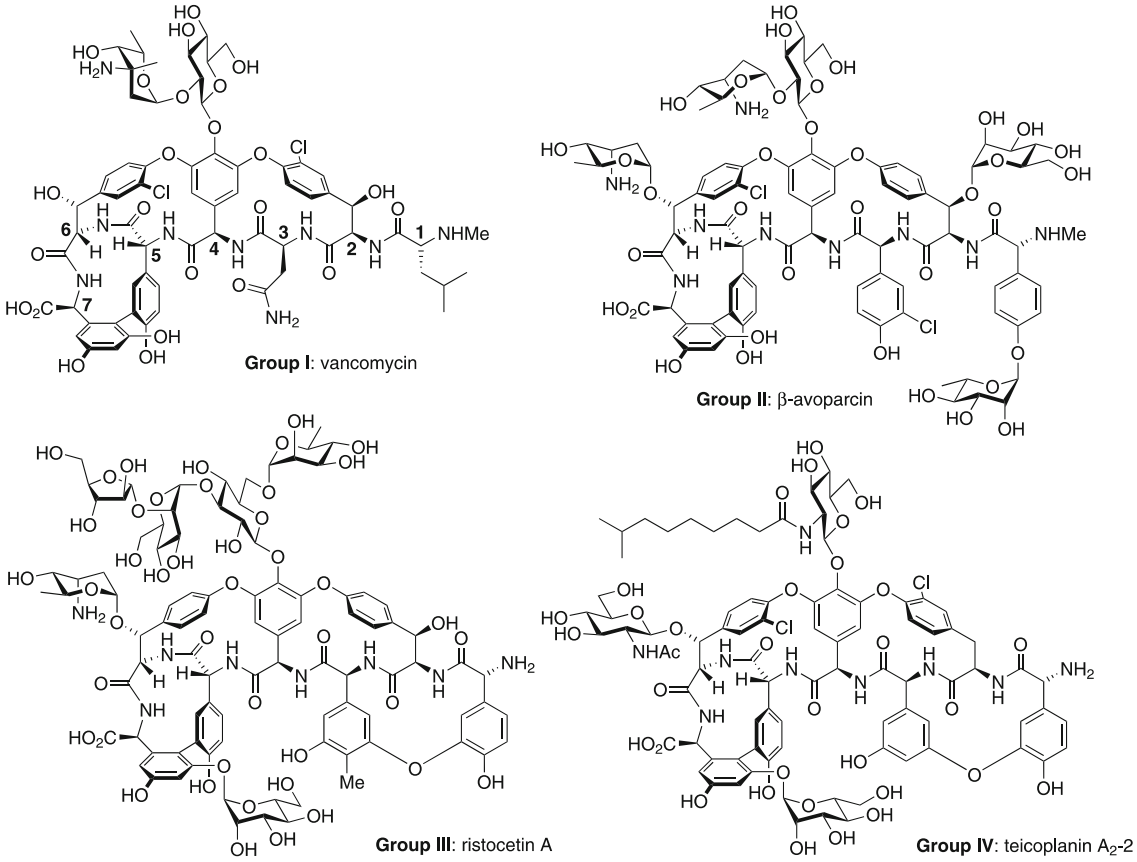


Figure 18
Selected structures of vancomycin-type glycopeptide antibiotics belonging to groups I–IV

teicoplanin, respectively, are characterized by aromatic residues at these positions. Members of group III and IV have an additional biaryl ether bridge between the side chains of amino acids 1 and 3. Type IV glycopeptide antibiotics could be considered a subgroup of III since the arrangement of the amino acids in the peptide core is the same. The antibiotics of this group have a fatty acid residue attached to an amino sugar. Finally, a fifth type of antibiotics can be defined containing a characteristic tryptophan residue at position 2. However, the members of this group identified so far (chloropeptin I, complestatin, and kistamycin A and B) are not glycosylated.

More than four centuries after its discovery and some 20 years after its structure elucidation, total syntheses of vancomycin [213,214,215,216,217] and its aglycon [218,219,220] have been accomplished by several research groups.

References

1. Fukuda M, Hindsgaul O (eds) (1994) Molecular Glycobiology. IRL Press, Oxford
2. Montreuil J, Vliegenthart JFG, Schachter H (eds) (1995) Glycoproteins. Elsevier, Amsterdam
3. Montreuil J, Vliegenthart JFG, Schachter H (eds) (1997) Glycoproteins II. Elsevier, Amsterdam
4. Gabius HJ, Gabius S (eds) (1997) Glycosciences: Status and Perspectives. Chapman & Hall, Weinheim
5. Varki A, Cummings R, Esko J, Freeze H, Hart G, Marth J (eds) (1999) Essentials of Glycobiology. Cold Spring Harbor Laboratory Press, Cold Spring Harbor
6. Lis H, Sharon N (1993) Eur J Biochem 218:1
7. Dwek RA (1996) Chem Rev 96:683
8. Rudd PM, Dwek RA (1997) Crit Rev Biochem Mol Biol 32:1
9. Van den Steen P, Rudd PM, Dwek RA, Opdenakker G (1998) Crit Rev Biochem Mol Biol 33:151
10. Taylor ME, Drickamer K (2006) Introduction to Glycobiology, 2 edn. Oxford University Press, Oxford
11. Lehle L, Strahl S, Tanner W (2006) Angew Chem 118:6956
12. Wittmann V (ed) (2007) Glycopeptides and Glycoproteins: Synthesis, Structure, and Application (Topics in Current Chemistry, vol. 267). Springer-Verlag, Berlin Heidelberg New York
13. Arnold JN, Wormald MR, Sim RB, Rudd PM, Dwek RA (2007) Annu Rev Immunol 25:21
14. Hitchen PG, Dell A (2006) Microbiology 152:1575
15. Szymanski CM, Wren BW (2005) Nat Rev Microbiol 3:225
16. Upreti RK, Kumar M, Shankar V (2003) Proteomics 3:363
17. Benz I, Schmidt MA (2002) Mol Microbiol 45:267
18. Brennan SO, Myles T, Peach RJ, Donaldson D, George PM (1990) Proc Natl Acad Sci USA 87:26
19. Schmid K, Binette JP, Kamiyama S, Pfister V, Takahashi S (1962) Biochemistry 1:959
20. Fournet B, Strecker G, Montreuil J, Vliegenthart JFG, Schmid K, Binette JP (1978) Biochemistry 17:5206
21. Van Halbeek H, Dorland L, Vliegenthart JFG, Montreuil J, Fournet B, Schmid K (1981) J Biol Chem 256:5588
22. Kato Y, Iwase H, Hotta K (1984) Anal Biochem 138:437
23. Ashford DA, Dwek RA, Rademacher TW, Lis H, Sharon N (1991) Carbohydr Res 213:215
24. Lütteke T, Bohne-Lang A, Loss A, Goetz T, Frank M, von der Lieth C-W (2006) Glycobiology 16:71R
25. Hashimoto K, Goto S, Kawano S, Aoki-Kinoshita KF, Ueda N, Hamajima M, Kawasaki T, Kanehisa M (2006) Glycobiology 16:63R
26. Raman R, Venkataraman M, Ramakrishnan S, Lang W, Raguram S, Sasisekharan R (2006) Glycobiology 16:82R
27. Varki A (1993) Glycobiology 3:97
28. Vliegenthart JFG, Montreuil J (1995) Primary structure of glycoprotein glycans. In: Montreuil J, Vliegenthart JFG, Schachter H (eds) Glycoproteins. Elsevier, Amsterdam, p 13

29. Sharon N, Lis H (1997) Glycoproteins: Structure and Function. In: Gabius H-J, Gabius S (eds) Glycosciences: Status and Perspectives. Chapman & Hall, Weinheim, p 133
30. Spiro RG (2002) Glycobiology 12:43R
31. Taylor CM (1998) Tetrahedron 54:11317
32. Cho SJ, Roman G, Yeboah F, Konishi Y (2007) Curr Med Chem 14:1653
33. Bierhaus A, Hofmann MA, Ziegler R, Nawroth PP (1998) Cardiovasc Res 37:586
34. Johansen PG, Marshall RD, Neuberger A (1961) Biochem J 78:518
35. Nuenke RH, Cunningham LW (1961) J Biol Chem 236:2451
36. Yamashina I, Makino M (1962) J Biochem 51:359
37. Helenius A, Aebi M (2004) Ann Rev Biochem 73:1019
38. Imperiali B, Shannon KL, Rickert KW (1992) J Am Chem Soc 114:7942
39. Abbadi A, McHarfi M, Aubry A, Premilat S, Boussard G, Marraud M (1991) J Am Chem Soc 113:2729
40. Imperiali B, Shannon KL, Unno M, Rickert KW (1992) J Am Chem Soc 114:7944
41. Imperiali B (1997) Acc Chem Res 30:452
42. Gibbs BS, Coward JK (1999) Bioorg Med Chem 7:441; see also corrigendum on page 2121
43. Bause E, Legler G (1981) Biochem J 195: 639
44. Bause E, Breuer W, Peters S (1995) Biochem J 312:979
45. Bause E, Wesemann M, Bartoschek A, Breuer W (1997) Biochem J 322:95
46. Shibata S, Takeda T, Natori Y (1988) J Biol Chem 263:12483
47. Wieland F, Heitzer R, Schaefer W (1983) Proc Natl Acad Sci USA 80:5470
48. Paul G, Lottspeich F, Wieland F (1986) J Biol Chem 261:1020
49. Messner P, Sleytr UB (1988) FEBS Lett 228:317
50. Review: Messner P (1997) Glycoconjugate J 14:3
51. Schreiner R, Schnabel E, Wieland F (1994) J Cell Biol 124:1071
52. Singh DG, Lomako J, Lomako WM, Whelan WJ, Meyer HE, Serwe M, Metzger JW (1995) FEBS Lett 376:61
53. Wopereis S, Lefeber DJ, Morava E, Wevers RA (2006) Clin Chem 52:574
54. Hart GW (1997) Annu Rev Biochem 66:315
55. Snow DM, Hart GW (1998) Int Rev Cytol 181:43
56. Zachara NE, Hart GW (2002) Chem Rev 102:431
57. Zachara NE, Hart GW (2004) Biochim Biophys Acta 1673:13
58. Harris RJ, Spellman MW (1993) Glycobiology 3:219
59. Shao L, Haltiwanger RS (2003) Cell Mol Life Sci 60:241
60. Hase S, Nishimura H, Kawabata S-I, Iwanaga S, Ikenaka T (1990) J Biol Chem 265:1858
61. Gerwig GJ, Kamerling JP, Vliegenthart JFG, Morag E, Lamed R, Bayer AE (1992) Eur J Biochem 205:799
62. Gellerfors P, Axelsson K, Helander A, Johansson S, Kenne L, Lindqvist S, Pavlu B, Skottner A, Fryklund L (1989) J Biol Chem 264:11444
63. Endo T (1999) Biochim Biophys Acta 1473: 237
64. Chiba A, Matsumura K, Yamada H, Inazu T, Shimizu T, Kusunoki S, Kanazawa I, Kobata A, Endo T (1997) J Biol Chem 272:2156
65. Krusius T, Finne J, Margolis RK, Margolis RU (1986) J Biol Chem 261:8237
66. Margolis RK, Margolis RU (1993) Experientia 49:429
67. Thibault P, Logan SM, Kelly JF, Brisson J-R, Ewing CP, Trust TJ, Guerry P (2001) J Biol Chem 276:34862
68. Young KT, Davis LM, DiRita VJ (2007) Nat Rev Microbiol 5:665
69. Stimson E, Virji M, Makepeace K, Dell A, Morris HR, Payne G, Saunders JR, Jennings MP, Barker S, Panico M, Blench I, Moxon ER (1995) Mol Microbiol 17:1201
70. Hegge FT, Hitchen PG, Aas FE, Kristiansen H, Lovold C, Egge-Jacobsen W, Panico M, Leong WY, Bull V, Virji M, Morris HR, Dell A, Koomey M (2004) Proc Natl Acad Sci USA 101:10798
71. Parge HE, Forest KT, Hickey MJ, Christensen DA, Getzoff ED, Tainer JA (1995) Nature 378:32
72. Castric P, Cassels FJ, Carlson RW (2001) J Biol Chem 276:26479
73. Alonso MD, Lomako J, Lomako WM, Whelan WJ (1995) FASEB J 9:1126
74. Lomako J, Lomako WM, Whelan WJ (2004) Biochim Biophys Acta 1673:45
75. Bock K, Schuster-Kolbe J, Altman E, Allmaier G, Stahl B, Christian R, Sleytr UB, Messner P (1994) J Biol Chem 269:7137
76. Messner P, Christian R, Neuninger C, Schulz G (1995) J Bacteriol 177:2188

77. Butler WT, Cunningham LW (1966) J Biol Chem 241:3882
78. Spiro RG (1967) J Biol Chem 242:4813
79. Yamagishi T, Matsuda K, Watanabe T (1976) Carbohydr Res 50:63
80. Lamport DTA (1967) Nature 216:1322
81. Lamport DTA, Miller DH (1971) Plant Physiol 48:454
82. Allen KA, Desai NN, Neuberger A, Creeth JM (1978) Biochem J 171:665
83. Strahm A, Amado R, Neukom H (1981) Phytochem 20:1061
84. Qi W, Fong C, Lamport DTA (1991) Plant Physiol 96:848
85. Teng-umnuay P, Morris HR, Dell A, Panico M, Paxton T, West CM (1998) J Biol Chem 273:18242
86. Ikezawa H, Yamanegi M, Taguchi R, Miyashita T, Ohyabu T (1976) Biochim Biophys Acta 450:154
87. Low MG, Finean JB (1977) Biochem J 167:281
88. Campbell DG, Gagnon J, Reid KBM, Williams AF (1981) Biochem J 195:15
89. Ferguson MAJ, Low MG, Cross GA (1985) J Biol Chem 260:14547
90. Ferguson MAJ, Williams AF (1988) Annu Rev Biochem 57:285
91. Cross GAM (1990) Annu Rev Immunol 8:83
92. Englund PT (1993) Annu Rev Biochem 62:121
93. McConville MJ, Ferguson MAJ (1993) Biochem J 294:305
94. Field MC, Menon AK (1993) Glycolipid anchoring of cell surface proteins. In: Schlesinger MJ (ed) Lipid Modifications of Proteins. CRC Press, Boca Raton, FL, p 83
95. Eckert V, Gerold P, Schwarz RT (1997) GPI-Anchors: Structure and Functions. In: Gabius H-J, Gabius S (eds) Glycosciences: Status and Perspectives. Chapman & Hall, Weinheim, p 223
96. Cole RN, Hart GW (1997) Glycosyl-phosphatidylinositol anchors: structure, biosynthesis and function. In: Montreuil J, Vliegenthart JFG, Schachter H (eds) Glycoproteins II. Elsevier, Amsterdam, p 69
97. Ikezawa H (2002) Biol Pharm Bull 25:409
98. Hoessli DC, Tauzin S, Nasir ud D, Borisch B (2007) Curr Org Chem 11:619
99. Ferguson MAJ, Homans SW, Dwek RA, Rademacher TW (1988) Science 239:753
100. Homans SW, Ferguson MAJ, Dwek RA, Rademacher TW, Anand R, Williams AF (1988) Nature 333:269
101. Deeg MA, Humphrey DR, Yang SH, Ferguson TR, Reinhold VN, Rosenberry TL (1992) J Biol Chem 267:18573
102. Walter BI, Roberts WL, Rosenberry TI, Ratnoff WD, Medof ME (1990) J Immunol 144:1030
103. Ferguson MAJ (1992) Chemical and enzymic analysis of glycosyl-phosphatidylinositol anchors. In: Hooper NM, Turner AJ (eds) Lipid Modification of Proteins: A Practical Approach. IRL Press, Oxford, p 191
104. Mayor S, Riezman H (2004) Nat Rev Mol Cell Biol 5:110
105. Baumann NA, Menon AK, Rancour DM (2000) Functions of Glycosyl Phosphatidylinositols. In: Ernst B, Hart GW, Sinaÿ P (eds) Carbohydrates in Chemistry and Biology, vol 4. Wiley-VCH, Weinheim, p 757
106. Guo Z, Bishop L (2004) Eur J Org Chem 3585-3596
107. Hofsteenge J, Mueller DR, de Beer T, Löffler A, Richter WJ, Vliegenthart JFG (1994) Biochemistry 33:13524
108. de Beer T, Vliegenthart JFG, Löffler A, Hofsteenge J (1995) Biochemistry 34:11785
109. Loeffler A, Doucey M-A, Jansson AM, Mueller DR, de Beer T, Hess D, Meldal M, Richter WJ, Vliegenthart JFG, Hofsteenge J (1996) Biochemistry 35:12005
110. Doucey M-A, Hess D, Blommers MJJ, Hofsteenge J (1999) Glycobiology 9:435
111. Hartmann S, Hofsteenge J (2000) J Biol Chem 275:28569
112. Julenius K (2007) Glycobiology 17:868
113. Haynes PA (1998) Glycobiology 8:1
114. Gustafson GL, Gander JE (1984) Methods Enzymol 107:172
115. Mehta DP, Ichikawa M, Salimath PV, Etchison JR, Haak R, Manzi A, Freeze HH (1996) J Biol Chem 271:10897
116. Guha-Niyogi A, Sullivan DR, Turco SJ (2001) Glycobiology 11:45R
117. Hounsell EF (1997) Methods of Glycoconjugate Analysis. In: Gabius H-J, Gabius S (eds) Glycosciences: Status and Perspectives. Chapman & Hall, Weinheim, p 15
118. Rudd PM, Dwek RA (1997) Curr Opin Biotechnol 8:488
119. Kameyama A (2006) Trends Glycosci Glycotechnol 18:323
120. Haslam SM, North SJ, Dell A (2006) Curr Opin Struct Biol 16:584
121. Geyer H, Geyer R (2006) Biochim Biophys Acta 1764:1853

122. Budnik BA, Lee RS, Steen JAJ (2006) Biochim Biophys Acta 1764:1870
123. Bardor M, Cabrera G, Rudd PM, Dwek RA, Cremata JA, Lerouge P (2006) Curr Opin Struct Biol 16:576
124. Mechref Y, Novotny MV (2002) Chem Rev 102:321
125. Zamfir AD, Bindila L, Lion N, Allen M, Girault HH, Peter-Katalinic J (2005) Electrophoresis 26:3650
126. Kobata A, Endo T (1993) In: Fukuda M, Kobata A (eds) Glycobiology: A Practical Approach. Oxford University Press, Oxford, p 79
127. Patel T, Bruce J, Merry A, Bigge C, Wormald M, Jaques A, Parekh R (1993) Biochemistry 32: 679
128. Piller F, Piller V (1993) In: Fukuda M, Kobata A (eds) Glycobiology: A Practical Approach. Oxford University Press, Oxford, p 291
129. Huang Y, Mechref Y, Novotny MV (2001) Anal Chem 73:6063
130. Huang Y, Konse T, Mechref Y, Novotny MV (2002) Rapid Commun Mass Spectrom 16:1199
131. Review: O'Neill RA (1996) J Chromatogr A 720:201
132. Yoshima H, Mizuochi T, Ishii M, Kobata A (1980) Cancer Res 40:4276
133. Bigge JC, Patel TP, Bruce JA, Goulding PN, Charles SM, Parekh RB (1995) Anal Biochem 230:229
134. Jackson P (1996) Mol Biotechnol 5:101
135. Honda S, Makino A, Suzuki S, Kakehi K (1990) Anal Biochem 191:228
136. Jackson P (1991) Anal Biochem 196:238
137. Guttman A, Chen F-TA, Evangelista A, Cooke N (1996) Anal Biochem 233:234
138. Chiesa C, O'Neill RA (1994) Electrophoresis 15:1132
139. Guttman A, Chen FT, Evangelista RA (1996) Electrophoresis 17:412
140. Shinohara Y, Furukawa J-I, Niikura K, Miura N, Nishimura S-I (2004) Anal Chem 76:6989
141. von der Lieth C-W, Lütteke T, Frank M (2006) Biochim Biophys Acta 1760:568
142. Edge CJ, Rademacher TW, Wormald MR, Parekh RB, Butters TD, Wing DR, Dwek RA (1992) Proc Natl Acad Sci USA 89:6338
143. Rudd PM, Guile GR, Kuster B, Harvey DJ, Opdenakker G, Dwek RA (1997) Nature 388:205
144. Rudd PM, Morgan BP, Wormald MR, Harvey DJ, van den Berg CW, Davis SJ, Ferguson MAJ, Dwek RA (1997) J Biol Chem 272:7229
145. Ito S, Yamashita K, Spiro RG, Kobata A (1977) J Biochem 81:1621
146. Watkins WM (1995) Molecular basis of antigenic specificity in the ABO, H and Lewis blood-group systems. In: Montreuil J, Vliegenthart JFG, Schachter H (eds) Glycoproteins. Elsevier, Amsterdam, p 313
147. Simanek EE, McGarvey GJ, Jablonowski JA, Wong C-H (1998) Chem Rev 98:833
148. Kannagi R (2002) Curr Opin Struct Biol 12:599
149. Hang HC, Bertozzi CR (2005) Bioorg Med Chem 13:5021
150. Hanisch F-G (2001) Biol Chem 382:143
151. Ten Hagen KG, Fritz TA, Tabak LA (2003) Glycobiology 13:1R
152. Julenius K, Molgaard A, Gupta R, Brunak S (2005) Glycobiology 15:153
153. Perez-Vilar J, Hidalgo J, Velasco A (1991) J Biol Chem 266:23967
154. Roth J, Wang Y, Eckhardt AE, Hill RL (1994) Proc Natl Acad Sci USA 91:8935
155. Roth J (1995) Compartmentation of glycoprotein biosynthesis. In: Montreuil J, Vliegenthart JFG, Schachter H (eds) Glycoproteins. Elsevier, Amsterdam, p 287
156. Yamashita Y, Chung YS, Horie R, Kannagi R, Sowa M (1995) J Natl Cancer Inst 87:441
157. Brockhausen I, Schutzbach J, Kuhns W (1998) Acta Anat 161:36
158. Brockhausen I, Yang J, Lehotay M, Ogata S, Itzkowitz S (2001) Biol Chem 382:219
159. Brockhausen I (1995) Biosynthesis of O-glycans of the N-acetylgalactosamine-α-Ser/Thr linkage type. In: Montreuil J, Vliegenthart JFG, Schachter H (eds) Glycoproteins. Elsevier, Amsterdam, p 201
160. Brockhausen I, Schachter H (1997) Glycosyltransferases Involved in *N*- and *O*-Glycan Biosynthesis. In: Gabius H-J, Gabius S (eds) Glycosciences: Status and Perspectives. Chapman & Hall, Weinheim, p 79
161. Kurosaka A, Nakajima H, Funakoshi J, Matsuyama M, Nagayo T, Yamashina I (1983) J Biol Chem 258:11594
162. Capon C, Leroy Y, Wieruszeski J-M, Ricart G, Strecker G, Montreuil J, Fournet B (1989) Eur J Biochem 182:139
163. Hounsell EF, Lawson AM, Feeney J, Gooi HC, Pickering NJ, Stoll MS, Lui SC, Feizi T (1985) Eur J Biochem 148:367
164. Chai W, Hounsell EF, Cashmore GC, Rosankiewicz JR, Bauer CJ, Feeney J, Feizi T, Lawson AM (1992) Eur J Biochem 203:257

165. Wang Y, Lee GF, Kelley RF, Spellman MW (1996) Glycobiology 6:837
166. Nishimura H, Takao T, Hase S, Shimonishi Y, Iwanaga S (1992) J Biol Chem 267:17520
167. Harris RJ, van Halbeek H, Glushka J, Basa LJ, Ling VT, Smith KJ, Spellman MW (1993) Biochemistry 32:6539
168. Minamida S, Aoki K, Natsuka S, Omichi K, Fukase K, Kusumoto S, Hase S (1996) J Biochem 120:1002
169. Kreppel LK, Blomberg MA, Hart GW (1997) J Biol Chem 272:9308
170. Lubas WA, Frank DW, Krause M, Hanover J (1997) J Biol Chem 272:9316
171. Dong DL, Hart GW (1994) J Biol Chem 269:19321
172. Wight TN (1999) Biosynthesis of Proteoglycans. In: Pinto BM (ed) Carbohydrates and their Derivatives Including Tannins, Cellulose, and Related Lignins (Comprehensive Natural Products Chemistry), vol 3. Elsevier, Amsterdam, p 161
173. Silbert JE, Bernfield M, Kokenyesi R (1997) Proteoglycans: a special class of glycoproteins. In: Montreuil J, Vliegenthart JFG, Schachter H (eds) Glycoproteins II. Elsevier, Amsterdam, p 1
174. Kresse H (1997) Proteoglycans - Structure and Functions. In: Gabius H-J, Gabius S (eds) Glycosciences: Status and Perspectives. Chapman & Hall, Weinheim, p 201
175. Mulloy B, Rider CC (2006) Biochem Soc Trans 34:409
176. Taylor KR, Gallo RL (2006) FASEB J 20:9
177. Bülow HE, Hobert O (2006) Annu Rev Cell Dev Biol 22:375
178. Wight TN, Kinsella MG, Qwarnstrom EE (1992) Curr Opin Cell Biol 4:793
179. Williams KJ, Fukui IV (1997) Curr Opin Lipidol 8:253
180. Bishop JR, Schuksz M, Esko JD (2007) Nature 446:1030
181. Superti-Furga A, Haestbacka J, Rossi A, Van der Harten JJ, Wilcox WR, Cohn DH, Rimoin DL, Steinmann B, Lander ES, Gitzelmann R (1996) Ann N Y Acad Sci 785:195
182. Wight TN (1996) In: Fuster V, Ross R, Topol EJ (eds) Atherosclerosis and Coronary Artery Disease. Raven Publishers, Philadelphia, PA, p 421
183. Iozzo RV (1988) Cancer Metastasis Rev 7:39
184. Snow AD, Wight TN (1989) Neurobiol Aging 10:481
185. It should be kept in mind that the anomeric configuration of the uronic acid remains the same. It is only the terminology that has changed due to the conversion of a D-sugar (GlcA) into an L-sugar (IdoA)
186. Bugg TDH (1999) Bacterial Peptidoglycan. Biosynthesis and its Inhibition. In: Pinto BM (ed) Carbohydrates and Their Derivatives Including Tannins, Cellulose, and Related Lignins (Comprehensive Natural Products Chemistry, vol 3). Elsevier, Amsterdam, p 241
187. Labischinski H, Maidhof N (1994) Bacterial peptidoglycan: overview and evolving concepts. In: Ghuysen J-M, Hakenbeck R (eds) Bacterial Cell Wall (New Compr Biochem, vol 27). Elsevier, Amsterdam, p 23
188. Vollmer W, Hoeltje J-V (2004) J Bacteriol 186:5978
189. Ghuysen JM (1968) Bacteriol Rev 32:425
190. Dmitriev B, Toukach F, Ehlers S (2005) Trends Microbiol 13:569
191. Tipper DJ, Strominger JL, Ensign JC (1967) Biochemistry 6:906
192. Leyh-Bouille M, Ghuysen J-M, Tipper DJ, Strominger JL (1966) Biochemistry 5:3079
193. Schleifer KH, Kandler O (1972) Bacteriol Rev 36:407
194. Williams DH, Bardsley B (1999) Angew Chem, Int Ed 38:1173
195. Walsh C (2000) Nature 406:775
196. Wright GD (2003) Curr Opin Chem Biol 7: 563
197. Young KD (2006) Trends Microbiol 14:155
198. Wolter F, Schoof S, Süssmuth RD (2007) Top Curr Chem 267:143
199. McCormick MH, Stark WM, Pittenger GE, Pittenger RC, McGuire JM (1956) Antibiot Annu 1955–1956:606
200. Lechevalier MP, Prauser H, Labeda DP, Ruan J-S (1986) Int J Syst Bacteriol 36:29
201. Philip JE, Schenck JP, Hargie MP (1957) Antibiot Annu 1955–1956:699
202. Nagarajan R (1994) Glycopeptide Antibiotics. Marcel Dekker, New York
203. Williams DH (1996) Nat Prod Rep 13:469
204. Nicolaou KC, Boddy CNC, Bräse S, Winssinger N (1999) Angew Chem Int Ed 38:2096
205. Howard MA, Firkin BG (1971) Thromb Diath Haemorrh 26:362
206. Hiramatsu K (1998) Drug Resist Updates 1:135
207. Levy SB (1998) Sci Am 3:46
208. Marshall FJ (1965) J Med Chem 8:18
209. Williams DH, Kalman JR (1977) J Am Chem Soc 99:2768

210. Sheldrick GM, Jones PG, Kennard O, Williams DH, Smith GA (1978) Nature 271:223
211. Harris CM, Harris TM (1982) J Am Chem Soc 104:4293
212. Lancini GC (1989) Fermentation and biosynthesis of glycopeptide antibiotics. In: Bushell ME, Graefe U (eds) Bioactive Metabolites from Microorganisms (Prog Ind Microbiol, vol 27). Elsevier, Amsterdam, p 283
213. Nicolaou KC, Mitchell HJ, Jain NF, Winssinger N, Hughes R, Bando T (1999) Angew Chem Int Ed 38:240
214. Nicolaou KC, Li H, Boddy CNC, Ramanjulu JM, Yue T-Y, Natarajan S, Chu X-J, Bräse S, Rübsam F (1999) Chem Eur J 5:2584
215. Nicolaou KC, Boddy CNC, Li H, Koumbis AE, Hughes R, Natarajan S, Jain NF, Ramanjulu JM, Bräse S, Solomon ME (1999) Chem Eur J 5:2602
216. Nicolaou KC, Koumbis AE, Takayanagi M, Natarajan S, Jain NF, Bando T, Li H, Hughes R (1999) Chem Eur J 5:2622
217. Nicolaou KC, Mitchell HJ, Jain NF, Bando T, Hughes R, Winssinger N, Natarajan S, Koumbis AE (1999) Chem Eur J 5:2648
218. Evans DA, Wood MR, Trotter BW, Richardson TI, Barrow JC, Katz JL (1998) Angew Chem Int Ed 37:2700
219. Boger DL, Miyazaki S, Kim SH, Wu JH, Loiseleur O, Castle SL (1999) J Am Chem Soc 121:3226
220. Boger DL, Miyazaki S, Kim SH, Wu JH, Castle SL, Loiseleur O, Jin Q (1999) J Am Chem Soc 121:10004

8.2 Glycoproteins: Properties

Valentin Wittmann
Fachbereich Chemie, Universität Konstanz, 78457 Konstanz, Germany
mail@valentin-wittmann.de

Abstract

This chapter focuses on the biological roles of the glycans contained in glycoproteins. Today we know there is no unifying function for the carbohydrates present in glycoproteins. They rather span the complete spectrum from being obviously unimportant to being crucial for the survival of an organism. In a crude scheme, their biological functions can be classified into two groups. On one hand, the carbohydrates can modify intrinsic properties of a protein by altering its size, charge, solubility, accessibility, structure, or dynamic properties. On the other hand, the glycans themselves may be specifically recognized by carbohydrate-binding proteins and thus participate in adhesion processes and signal transduction. A selection of some of these processes, that are well characterized, will be highlighted.

Keywords

Glycoproteins; Glycopeptides; Antifreeze glycoproteins; Protein stability; Protein activity; Glycopeptide conformation; Blood group antigens; Glycosyltransferases; Glycosidases; Protein folding

In: *Glycoscience*. Fraser-Reid B, Tatsuta K, Thiem J (eds)
Chapter-DOI 10-1007/978-3-540-30429-6_44:

Abbreviations

AFGPs antifreeze glycoproteins
CD circular dichroism
CHO Chinese hamster ovary
CRD carbohydrate recognition domain
DAF decay accelerating factor
EGF epidermal growth factor
ER endoplasmic reticulum
FDA Food and Drug Administration
FRET fluorescence resonance energy transfer
NOE nuclear Overhauser enhancement
tPA tissue plasminogen activator

1 Introduction

Whereas the last chapter provided an overview of the structures of glycoproteins and glycopeptides occurring in nature, this chapter focuses on the biological roles of their carbohydrate units. However, the frequently asked question "What is the function of glycosylation?" is actually as unreasonable as would be the question "What is the function of proteins?" Today we know there is no unifying function for the carbohydrates present in glycoproteins. They rather span the complete spectrum from being relatively unimportant to being crucial for the survival of an organism. Moreover, the same glycan may have different functions at different locations on a given protein, or in different cells or developmental stages of an organism. Thus, each glycoprotein has to be studied individually in order to unravel the roles of its glycans. The aim of this chapter is to indicate some general principles of protein glycosylation.

Diverse approaches are being employed in order to uncover the roles of carbohydrates contained in glycoconjugates [1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18]. They include the localization of specific glycans using lectins or antibodies, the modification of glycans by glycosyltransferases and glycosidases, and the use of inhibitors of glycan biosynthesis or processing. Natural or synthetic ligands can be used to identify specific carbohydrate receptors. The methods of molecular biology provide a powerful tool to study glycan function. Thus, it is possible to generate cell mutants with altered glycosyltransferase expressions. Alternatively, recombinant glycoproteins can be expressed in different cells with different glycosylation properties. Chinese hamster ovary (CHO) cells, for example, are not able to generate sialic acid α2–6 linkages and galactose α1–3 linkages and bacteria like *Escherichia coli* produce completely non-glycosylated proteins. Recently, the chemical synthesis of pure glycoforms of glycoproteins or glycoprotein mimetics has become feasible providing a powerful tool for structure-activity relationship studies [19]. Metabolic engineering is another technique developed by chemists that gives access to glycoproteins containing functional groups that are selectively addressable by chemical ligation reactions [20,21]. Carbohydrate chips are used to probe carbohydrate recognition by proteins and other binding partners in a high-throughput manner [14,22,23,24,25].

In a crude scheme, the biological functions of glycans can be classified into two groups. On one hand, the carbohydrates can modify intrinsic properties of a protein by altering its size, charge, solubility, accessibility, structure, or dynamic properties. On the other hand, the glycans themselves may be specifically recognized by carbohydrate-binding proteins and thus participate in adhesion processes and signal transduction.

2 Modulatory and Structural Roles of Glycans

In the following sections some modulatory and structural roles of glycans are discussed. However, it should be kept in mind that the chosen classification of the roles of glycans is somehow arbitrary since the individual effects often cannot be separated from each other. Thus, varied stability or biological function of a protein upon glycosylation is often a result of altered structural or dynamic parameters caused by the carbohydrates.

2.1 Modulation of Physicochemical Properties

The modification of physicochemical properties of proteins by attached carbohydrates is often observed, especially in glycoproteins with high carbohydrate content [1,7,12,26]. Sialylated or sulfated glycans, e. g., change the overall charge of a protein and increase its solubility. This is important for the highly glycosylated mucins (high sialic acid content) and the highly sulfated proteoglycans. Contained in mucous secretions of most epithelial cells, they can provide a gelation function due to their ability to retain water. Thus, they function as lubricants and protection for epithelial surfaces and mediate transport. Examples include synovia, buffering excessive gastric acid, and transport of chyme. In the extracellular matrix, proteoglycans provide elasticity and tensile strength. Furthermore, both mucins [27,28] and proteoglycans [29,30,31,32,33] act as adhesion molecules in numerous cell-cell, cell-matrix, and cell-microbe recognition events.

Antifreeze glycoproteins (AFGPs) [34,35,36,37,38] circulate in the blood of Antarctic fish and enable them to avoid freezing in their perpetually icy environment where the temperature is frequently as low as −1.9 °C. These mucin-type *O*-glycoproteins are composed of repeats of the glycotripeptide unit (Ala-Ala-[Gal(β1–3)GalNAc(α1–*O*)]Thr)$_n$ (*Fig. 1*). Eight distinct fractions of these proteins have been isolated (AFGP 1–8) with the number *n* of glycotripeptide

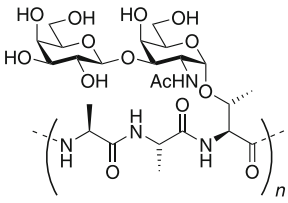


Figure 1
Structure of the glycotripeptide repeating unit (Ala-Ala-[Gal(β1–3)GalNAc(α1–O)]Thr)$_n$ contained in antifreeze glycoproteins

repeats ranging from 50 (molecular mass 33 kDa) to 4 (2.6 kDa). In smaller-sized AFGPs, proline replaces some of the alanine residues following threonine. Antifreeze glycoproteins function in a noncolligative manner by binding to and inhibiting the growth of ice crystals that enter the fish and maybe also by preventing the nucleation of ice crystals. This results in a freezing point depression without an appreciable change in the melting point. The difference between the melting and freezing temperatures is termed thermal hysteresis and is used as a measure for the magnitude of the antifreeze activity. In addition to thermal hysteresis, AFGPs effect alteration of the morphology of ice crystals into a hexagonal bipyramid. Studies revealed that chemical modification (acetylation, priodate oxidation, and complexation with borate but not oxidation of the two primary hydroxyls with galactose oxidase to give the bisaldehyde) or removal of the sugar residues results in a complete loss of antifreeze activity [35,38].
Recently, Nishimura et al. presented a detailed study on synthetic AFGPs composed of one to seven glycosylated Ala-Thr-Ala repeating units [39]. It was shown that a minimum of two

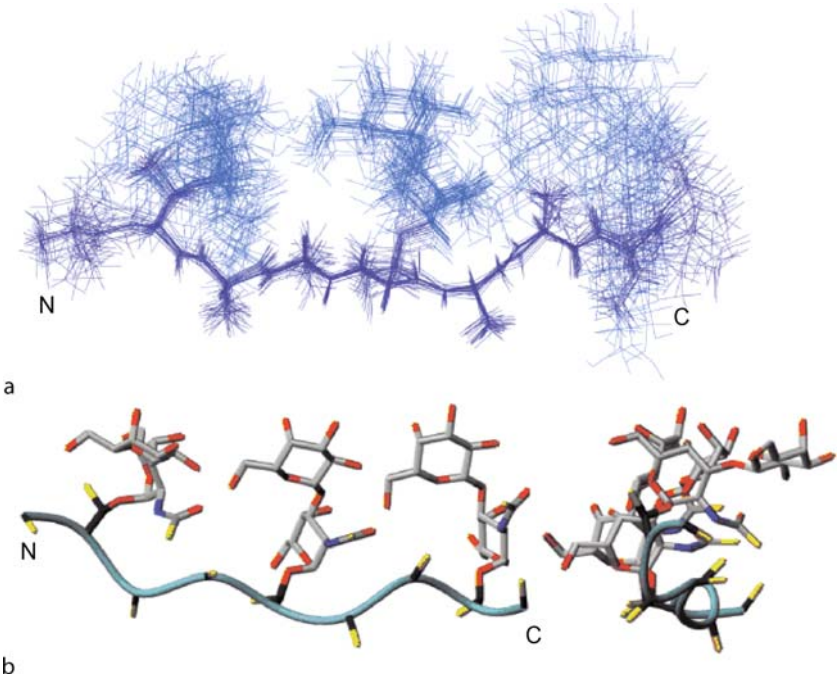


Figure 2
a Superposition of the 25 lowest energy structures of (Ala-[Gal(β1–3)GalNAc(α1–O)]Thr-Ala)$_3$ calculated from NMR-based constraints. The peptide backbone is navy, and the carbohydrate moieties are royal blue. b Structure closest to the average of the 25 best calculated models: yellow, methyl carbon; white, carbon in carbohydrate; gray, carbon in peptide side chains except for methyl carbon; blue, nitrogen; red, oxygen; the *N*- and *C*-terminal ends are identified. The peptide backbone folds into a left-handed helix similar to the polyproline type II helix. The disaccharides point to one side of the amphiphatic helix, thus constructing a hydrophilic face, while the Ala-CH_3 groups and acetyl methyl groups in the GalNAc residues are clustered into the hydrophobic face of the molecule. (Reprinted from [39] with permission from Willey-VCH)

repeating glycopeptide units is necessary to obtain compounds that give rise to thermal hysteresis. All compounds including the monomeric glycotripeptide, however, were capable to alter the morphology of ice crystals into hexagonal bipyramids. Varying the disaccharide part of oligomeric AFGP revealed three key motifs required for antifreeze activity, namely the *N*-acetyl group at the C-2 position of the reducing hexosamine, α-configuration of the *O*-glycosidic linkage between sugars and peptide chain, and the γ-methyl group of the threonyl residue. Strong evidence indicated that the peptide backbone conformation plays a critical role for antifreeze activity. Compounds lacking the mentioned key motifs either lost secondary structure or changed from a polyproline type II helix in the active compounds to a largely α-helical structure in the case of the glycosylated Ala-Ser-Ala repeating unit as determined by CD spectroscopy. Despite the absence of long- and medium-range NOE contacts, a surprisingly well-defined NMR-based structure could be calculated for the three repeating units containing glycopeptide (Ala-Ala-[Gal(β1–3)GalNAc(α1–O)]Thr)$_3$ (*Fig. 2*). The authors propose that the amphiphatic nature of the observed left-handed helix is crucial for antifreeze activity. However, even though the study by Nishimura et al. presents a consistent picture, the situation for naturally occurring AFGPs seems to be more complicated and structural studies on these molecules have a long history of controversial interpretations [38].

Genetic studies have shown that the AFGPs present in the two geographically and phylogenetically distinct Antarctic notothenioids and Arctic cod have evolved independently, in a rare example of convergent molecular evolution [38]. Recently, a novel antifreeze glycoprotein has been isolated from *Pleuragramma antarcticum* which was shown to be a proteoglycan [35]. Beside these antifreeze glycoproteins, several non-glycosylated antifreeze proteins are known [34,35,36,37,38].

2.2 Protective and Stabilizing Functions

There is little doubt that the "coating" of oligosaccharides on many glycoproteins can protect the peptide backbone from recognition by proteases and antibodies [1,7,12,26]. An example is the decay accelerating factor (DAF, CD55) which is proteolyzed upon removal of its *O*-linked sugars [40]. Decoration of the surfaces of most types of cells with different kinds of glycoconjugates gives rise to the so-called glycocalyx which can present a substantial physical barrier. Glycosylation can also alter the heat stability of proteins which has been shown for two different β1–3/4 glucanases from *Bacillus* species. Expression in *Saccharomyces cerevisiae* resulted in heavily glycosylated enzymes (carbohydrate content of about 45%) which were significantly more heat stable than their non-glycosylated counterparts expressed in *Escherichia coli* [41]. Such findings are of special interest for the industrial use of proteins. Other examples of the stabilizing function of carbohydrates [42,43,44] are ovine submaxillary mucin [45], an isolated tailpiece from human serum immunoglobulin M [46], RNase B [47,48,49], human CD2 [50], and the protease inhibitor PMP-C [51]. In these cases, the glycans had little overall effect on the conformation of the proteins but rather globally decreased the dynamic fluctuations of the glycoproteins, as revealed by NMR spectroscopy. The stabilizing function can be achieved even by a single carbohydrate unit [50,51].

2.3 Modulation of Biological Activity

There are many examples for the ability of carbohydrates to modulate the biological activity of functional proteins [1,7,12]. Bovine pancreas ribonuclease (RNase) for example occurs in unglycosylated (RNase A) and glycosylated (RNase B) forms, the latter being a set of nine different glycoforms of the high mannose type ($Man_{5-9}[GlcNAc]_2$) with respect to the only *N*-glycosylation site (Asn-34) of the protein [7]. Using double stranded RNA as substrate, it was shown that RNase A is more than three times as active as RNase B [49]. Furthermore, enzyme activities of several glycoforms of RNase, prepared by exomannosidase treatment of naturally occurring RNase B, have been determined and may be ranked in terms of decreasing activity as: RNase A > RNase ($Man_0[GlcNAc]_2$) > RNase ($Man_1[GlcNAc]_2$) > RNase ($Man_5[GlcNAc]_2$) > RNase B. These differences in activities were attributed to an overall increase in dynamic stability of the protein upon glycosylation and to steric hindrance between the oligosaccharides and the RNA substrate which is also supported by molecular modeling [47,48,49,52]. The dynamic properties of the protein were studied by NMR spectroscopic determination of proton/deuterium exchange rates of the various glycoforms. The presence of the glycan reduces solvent access to many regions of the peptide backbone both close to and remote from the glycosylation site as much as 30 Å away (*Fig. 3*), suggesting

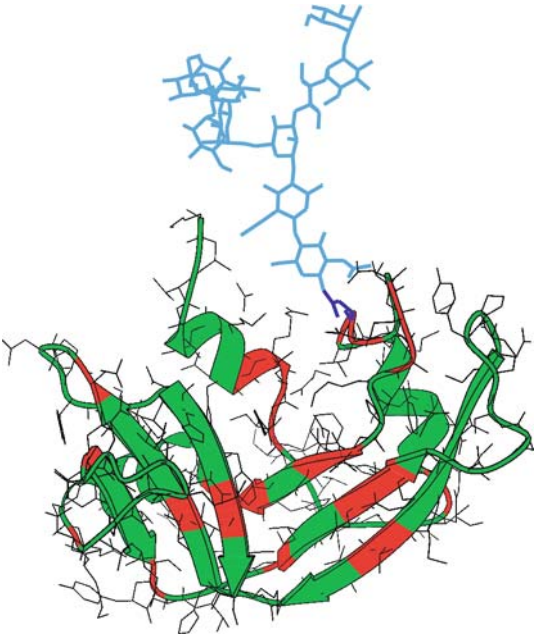


Figure 3
The presence of the glycan (light blue) on ribonuclease B reduces the amide proton/deuterium exchange rates compared with ribonuclease A for extensive regions of the peptide backbone (shown in red) both local to and remote from the glycosylation site [47]. The glycoprotein structure is based on the crystal structure of ribonuclease B [53] and one of the structures of $Man_9GlcNAc_2$ determined by NMR and molecular dynamics. (Reprinted from [44] with permission from Elsevier), [54]

that it reduces fluctuations of the backbone. This was confirmed by circular dichroism (CD) studies that indicated that the glycan has a small stabilizing effect on the peptide fold.

Tissue plasminogen activator (tPA) is another enzyme whose activity is affected by glycosylation [7]. tPA is a serine protease which converts plasminogen into plasmin which itself is a serine protease with fibrinolytic activity. There are two major classes of glycoforms of naturally occurring tPA. Type I contains three *N*-linked glycans at Asn-117, Asn-184, and Asn-448, whereas type II has only two, at Asn-117 and Asn-448. Variable occupancy of Asn-184 affects the population of glycoforms at Asn-448. Plasminogen is also a mixture of two major glycoforms containing one *O*-glycan and one *N*-glycan (type 1) or one *O*-glycan (type 2). For an efficient proteolytic activity of tPA, formation of a ternary complex with plasminogen and fibrin is required. Rate of formation and turnover of the complex is dependent on the glycosylation site occupancy of both tPA and plasminogen [55]. Thus, clot lysis occurs 2–3 times faster with type II tPA in combination with type 2 plasminogen versus participation of type I tPA and type 1 plasminogen. Both the *N*-glycan linked to Asn-184 in type I tPA and the *N*-glycan contained in type 1 plasminogen reduce the enzymatic activity of the serine proteases. In contrast to this tuning function of the *N*-glycans, it has been shown that the *O*-glycan contained in plasminogen is crucial for its proteolytic activity.

Glycosylation of a ligand can also mediate such an on-off or switching effect. Deglycosylation of the hormone human β chorionic gonadotropin (β-HCG) for example leads to a species which is still able to bind to its receptor with similar affinity but fails to stimulate adenylate cyclase [1]. In most cases, however, such effects of glycosylation are incomplete, i. e., the carbohydrates provide a means of tuning the primary function of the proteins. From numerous examples studied, it was concluded that the relatively large *N*-linked glycans generally downmodulate the activities of enzymes and signal molecules whereas *O*-glycosylation can result in both a downregulation and an upregulation [10,56,57].

2.4 Influence on Peptide Secondary Structure

Glycosylation can affect protein structure in several capacities [7,9,10,42,43,44,58,59]. Structural roles of glycans associated with protein stability and the regulation of protein function have already been mentioned in the preceding sections. Oligosaccharides attached to matrix molecules like collagens and proteoglycans are important in the physical maintenance of tissue structure, porosity, and integrity. *N*-Glycosylation is a co-translational process that is believed to play a major role in the initiation of correct folding of the nascent polypeptide chain in the rough endoplasmic reticulum [9,43,58,60,61]. Indeed, numerous examples exist in which removal of certain *N*-glycosylation sites by site-directed mutagenesis results in improper folding of glycoproteins [9,58,62,63,64,65,66]. The impact of *N*-glycosylation on the conformation of model peptides was nicely demonstrated by Imperiali et al. in a series of publications using time-resolved fluorescence resonance energy transfer (FRET) [67] and NMR techniques [68,69]. The examined peptides were derived from the A282-A288 sequence (Ile-Thr-Pro-Asn-Gly-Ser-Ile) of the hemagglutinin glycoprotein from influenza virus containing the critical Asn-A285 glycosylation site. This sequence represents a β-turn surface loop in the native protein which is of considerable interest as this motif is a common feature among the *final* structures of glycosylation sites in many glycoproteins.

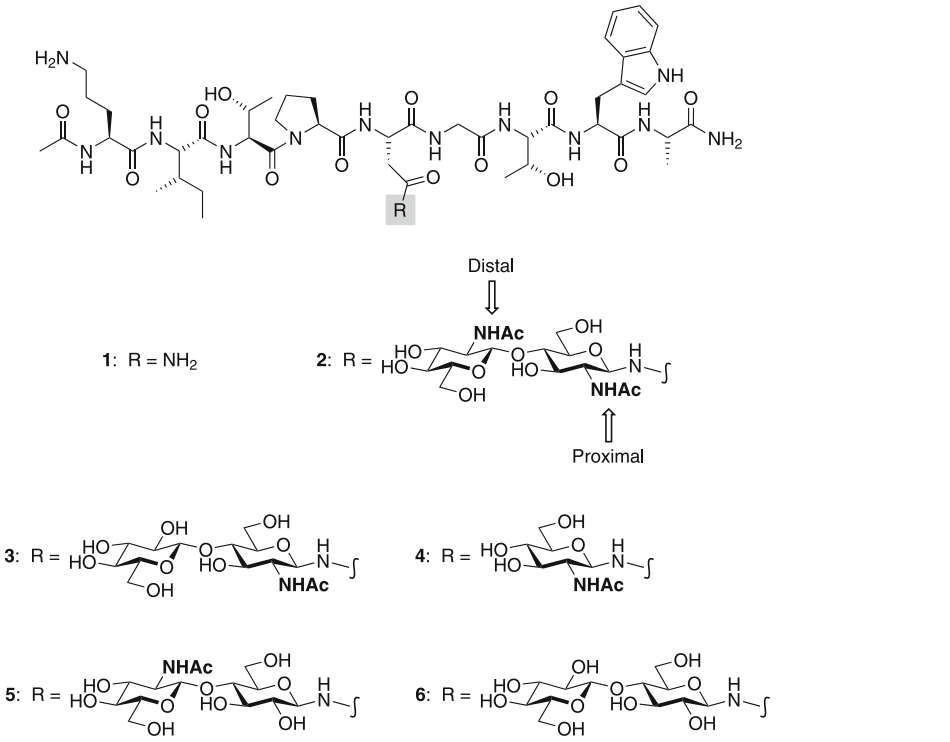


Figure 4
Peptide 1 and glycopeptides 2–6 synthesized in order to probe the influence of asparagine-linked glycosylation on peptide secondary structure [67,68,69]

In the FRET study [67], fluorescently labeled analogues of peptide **1** and glycopeptide **2** (*Fig. 4*) with a dansyl group at N_δ of Orn were examined. The study revealed that glycosylation of **1** with a chitobiosyl moiety (→ **2**) promoted the adoption of a more compact peptide secondary structure. Subsequent 2D ^{1}H NMR investigations in aqueous solution [68] supported this analysis and indicated that peptide **1** adopts an open and extended Asx-turn conformation prior to glycosylation whereas glycopeptide **2** exhibits a compact type I β-turn conformation (*Fig. 5*), quite similar to that observed in the final native protein structure. Both the FRET and NMR studies provide direct evidence that glycosylation of peptide **1** with chitobiose, a disaccharide representing the first two *N*-acetylglucosamine residues of the native tetradecasaccharide (cf. Chap. 8.1), induces a conformational switch in the peptide backbone. This observation is important with respect to the role glycosylation plays in the correct folding of glycoproteins. It was suggested that *N*-linked glycosylation may serve as a critical trigger to help the polypeptide chain to adopt a conformation that is populated in the native folded protein, but not in the nascent unmodified sequence.

Despite the strong influence of glycosylation on peptide conformation, no specific interactions between the chitobiose moiety and the peptide backbone were detected in the NMR analysis in aqueous solution. Therefore, it was proposed that the conformational change observed upon

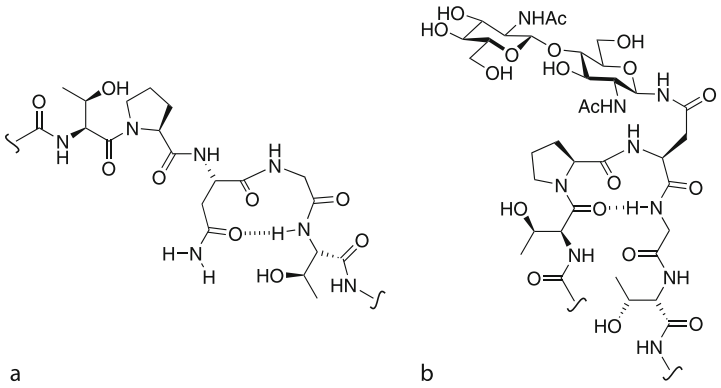


Figure 5
Glycosylation-induced conformational switching from an (a) Asx-turn to a (b) type I β-turn [68]

glycosylation results from either a steric effect in which the disaccharide alters the conformational space available to the peptide or from a modulation of the local water structure that influences the environment that the peptide experiences [68]. To address the important question if chitobiose may be replaced with other disaccharides, glycopeptides **3**–**6**, containing the saccharides Glc(β1–4)GlcNAc, GlcNAc, GlcNAc(β1–4)Glc, and Glc(β1–4)Glc, respectively, were examined (*Fig. 4*) [69]. In all cases, less well-ordered peptide conformations as compared to **2** were determined by NMR analysis. The study revealed that the *N*-acetyl group of the proximal sugar is critical for maintaining a β-turn conformation. Surprisingly, the *N*-acetyl group of the distal sugar also plays an important role in rigidifying both the saccharide and the peptide.

Recently, Imperiali et al. reported a synergistic experimental and computational study of a glycopeptide analogous to glycopeptide **2** but with the chitobiose moiety α-*N*-glycosidically linked to asparagine [70]. Thereby, the effect of the stereochemistry of the carbohydrate-peptide linkage on glycopeptide structure could be evaluated. It was shown that only the β-linked glycopeptide **2** adopted a type I β-turn whereas the α-anomer adopted an Asx-turn-like conformation comparable to that of the unmodified peptide **1**. In this regard, it is worth mentioning that statistical analyses of the conformation of *N*-glycosylation sites revealed that only about a quarter of glycosylated asparagines are located at a β-turn [71,72,73]. This could mean that the induction of a β-turn structure by *N*-glycosylation is of transient character and of major importance only during the early stages of the folding process [59]. A lectin-mediated mechanism by which *N*-glycans participate in quality control of protein folding will be discussed in *Sect. 3.3.2*.

An influence of *N*-glycosylation on peptide conformation has also been observed by Danishefsky et al. upon attachment of the non-natural trisaccharide Gal(β1–6)Gal(β1–6)GlcNAc β-*N*-glycosidically to the Asn side chain of the model peptide H-Ala-Leu-Asn-Leu-Thr-OH [74]. Whereas the unglycosylated peptide failed to manifest any appreciable secondary structure, the glycopeptide was assumed to exist in an equilibrium between an ordered and a random state. When carried out at –12 °C using a 90:10 mixture of H_2O/acetone-d_6, NMR analysis

revealed nuclear Overhauser enhancement (NOE) effect crosspeaks between the methyl of the sugar *N*-acetyl group and backbone amide protons.

In contrast to *N*-glycosylation, *O*-linked glycosylation is an entirely post-translational and post-folding event in mammalian cells. Nevertheless, *O*-Glycans may have major impact on the secondary, tertiary, and quaternary structure of fully folded proteins [10]. Otvos et al. showed that glycosylation distorts the α-helicity of an epitope on the rabies virus glycoprotein, and that *O*-glycosylation of threonine is more effective in perturbing the secondary structure of the peptide than *N*-glycosylation of asparagine [75]. In several NMR studies it has been shown that the peptide backbone of model peptides responds to *O*-glycosylation as evidenced by changes in sequential amide-amide NOE interactions [76,77,78,79,80]. The examples include the mucin-type GalNAc(α1–O)Thr [76,77,78] as well as the GlcNAc(β1–O)Thr linkages [79,80]. In all cases, the NMR data, which were supplemented in part by CD and fluorescence measurements [78,79], molecular modeling calculations [78,79,80], and chemical evidence [80], were indicative for a glycosylation-induced conformational change from a random structure to a turn-like structure. The conformational response is further modulated by whether the sugar component is a mono-, di-, or oligosaccharide [77,80].

In recent times, more detailed conformational analyses of *O*-glycopeptides have been published by several groups [81,82,83,84,85,86,87,88]. Very recently, Griesinger and Kunz and coworkers reported the effect of *O*-glycosylation with the α(2,6)-sialyl-T antigen on the conformational propensities of a 20-residue peptide representing the full length tandem repeat sequence of the human mucin MUC1 [89]. The peptide contained both the GVTSAP sequence, which is an effective substrate for GalNAc transferases, and the PDTRP fragment, a known epitope recognized by several anti-MUC1 monoclonal antibodies. NMR experiments were carried out in a H_2O/D_2O mixture (9 : 1) close to physiological conditions (25 °C, pH 6.5). Cluster analysis of the conformational ensemble yielded a rod-like structure with the GalNAc methyl group oriented towards the *C*-terminus, similar to results obtained by Danishefsky et al. [81].

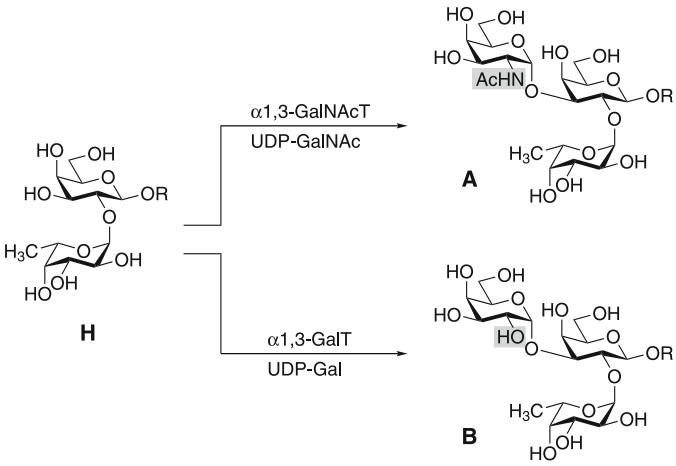
3 Involvement of Glycans in Recognition Events

Oligosaccharides have an enormous information-storing potential, being substantially higher than that of oligopeptides and oligonucleotides [90]. Thus, it is not surprising that the carbohydrates contained in glycoconjugates, beside their role in modulating intrinsic properties of glycoproteins, can act as recognition markers in numerous physiological and pathological processes [1,7,12,15,91,92,93,94,95,96,97,98,99,100]. Responsible for deciphering the encoded information are carbohydrate-binding proteins. They can be subdivided into antibodies, enzymes involved in sugar utilization and glycoconjugate turnover, and carbohydrate-binding proteins which are neither immunoglobulins nor enzymes. The latter have been referred to as lectins [101,102,103,104,105,106,107]. Carbohydrate-lectin interactions play a crucial role in many cellular recognition processes including clearance of glycoproteins from the circulatory system, control of intracellular traffic of glycoproteins, bacterial and viral adhesion to host cells, recruitment of leukocytes to inflammatory sites, cell interactions in the immune system, and tumor metastasis just to name a few. In the following sections some of these processes will be highlighted.

3.1 Carbohydrate Recognition by Blood Group Antibodies

The discovery of the ABO blood group system 100 years ago by Landsteiner et al. [108] was based on the observation that humans could be divided into different groups according to the presence or absence of serum constituents that would agglutinate red cells isolated from other humans. Although they were not aware of the underlying glycan basis, the work of Landsteiner and colleagues laid the basis for the safe transfusion of blood from one individual to another [109]. Today we know that the agglutinating serum constituents are antibodies and that their cognate antigens are oligosaccharides whose structures are genetically polymorphic [12,110,111,112,113]. The first discovery of these antigens on the surface of human erythrocytes led them to be classified as "blood-group antigens". However, they are also found in human secretions and mucosal tissues and it was suggested that they are more accurately defined as "histo-blood group antigens" [114].

The A, B, and H antigens (the latter being the one expressed by blood group O individuals) are formed by sequential action of distinct glycosyltransferases on different precursor saccharides, most often terminal residues of glycolipids and proteins. According to the nature of these peripheral disaccharide core structures on which the blood group antigens are synthesized, different types are distinguished. In type-1 structures, the antigen synthesis starts from Gal(β1–3)GlcNAc(β). Similarly, modification of Gal(β1–4)GlcNAc(β), Gal(β1–3)GalNAc(α), and Gal(β1–3)GalNAc(β) leads to type-2, type-3, and type-4 antigens, respectively. Fucosylation of these precursors by an α1–2 fucosyltransferase encoded by the *H* or the *Secretor* locus produces the blood group H determinant represented by the disaccharide Fuc(α1–2)Gal(β) (*Scheme 1*). Further glycosylation by an α1–3 *N*-acetylgalactosaminyltransferase (α1,3-GalNAcT) (corresponding to the *A* allele of the *ABO* locus) or an α1–3 galactosyltransferase (α1,3-GalT) (corresponding to the *B* allele of the *ABO* locus) then leads to GalNAc(α1–3)[Fuc(α1–2)]Gal(β) (blood group A determinant) and Gal(α1–3)[Fuc(α1–2)]Gal(β) (blood



Scheme 1
Biosynthesis of the blood group antigens A and B (minimal determinant structures)

Table 1
The ABH(O) system of antigens on erythrocytes and antibodies and glycosyltransferases in plasma

Geno-type	Phenotype	Blood group antigens on red cells (minimal determinant structure)	Antibodies in plasma	Glycosyltransferases in plasma
AA AO	A	**A**	anti-B	α 1,3-GalNAcT
BB BO	B	**B**	anti-A	α1,3-GalT
AB	AB	**A** and **B**	–	α1,3-GalNAcT, α1,3-GalT
OO	O	**H**	anti-A, anti-B	–

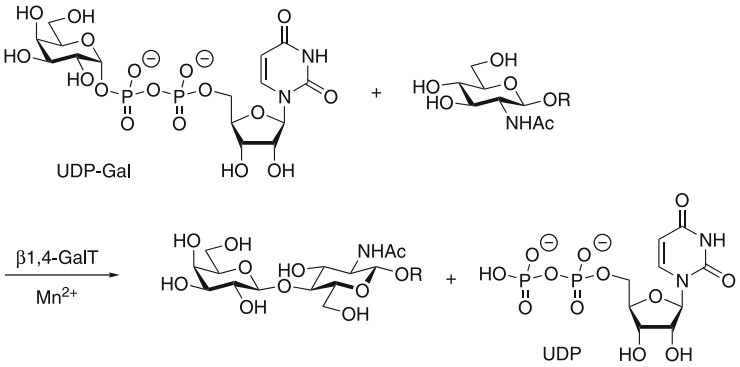
Table 2
Selected antigenic determinants (LacNAc = Gal(β1–4)GlcNAc)

Determinant structure	Name
Gal(β1–3)[Fuc(α1–4)]GlcNAc	Blood group Lewis a (Le^a)
Neu5Ac(α2–3)Gal(β1–3)[Fuc(α1–4)]GlcNAc	Sialyl Lewis a (sLe^a)
Fuc(α1–2)Gal(β1–3)[Fuc(α1–4)]GlcNAc	Blood group Lewis b (Le^b)
Gal(β1–4)[Fuc(α1–3)]GlcNAc	Lewis x (Le^x)
Neu5Ac(α2–3)Gal(β1–4)[Fuc(α1–3)]GlcNAc	Sialyl Lewis x (sLe^x)
Fuc(α1–2)Gal(β1–4)[Fuc(α1–3)]GlcNAc	Lewis y (Le^y)
[LacNAc(β1–3)]$_n$	Blood group i
LacNAc(β1–3)[LacNAc(β1–6)]LacNAc	Blood group I
GalNAc(α1–O)Ser/Thr	Tn
Gal(β1–3)GalNAc(α1–O)Ser/Thr	T
Fuc(α1–2)Gal(β1–3)GalNAc(β1–3)Gal(α1–4)Gal(β1–4)Glc	Globo H

group B determinant), respectively. The ABO classification is based on the presence or absence of the A and B antigens and the two antibodies anti-A and anti-B which always occur in the plasma when the corresponding antigen is missing (*Table 1*). With rare exceptions the H antigen is expressed on the cells of all blood group O individuals, but in persons belonging to phenotypes A, B, and AB there is complete or partial masking of H activity. Despite the vast accumulated knowledge of serology, chemistry, and genetics of the blood group structures, it is still not possible to assign a clearly defined physiological function to the *ABO* locus.
Beside the ABO blood group system a number of additional carbohydrate-based antigenic determinants are known (*Table 2*), which are, however, not only recognized by antibodies. Some of them are more highly expressed on tumor cells and therefore can be described as tumor-associated antigens [115]. Synthetic tumor-associated antigens such as Lewis y or Globo H have been used in the development of antitumor vaccines [116,117].

3.2 Carbohydrate-Modifying Enzymes

Of the many enzymes involved in carbohydrate metabolism, glycosyl transferases and glycosidases are of special interest for the biosynthesis and processing of glycoconjugates. The

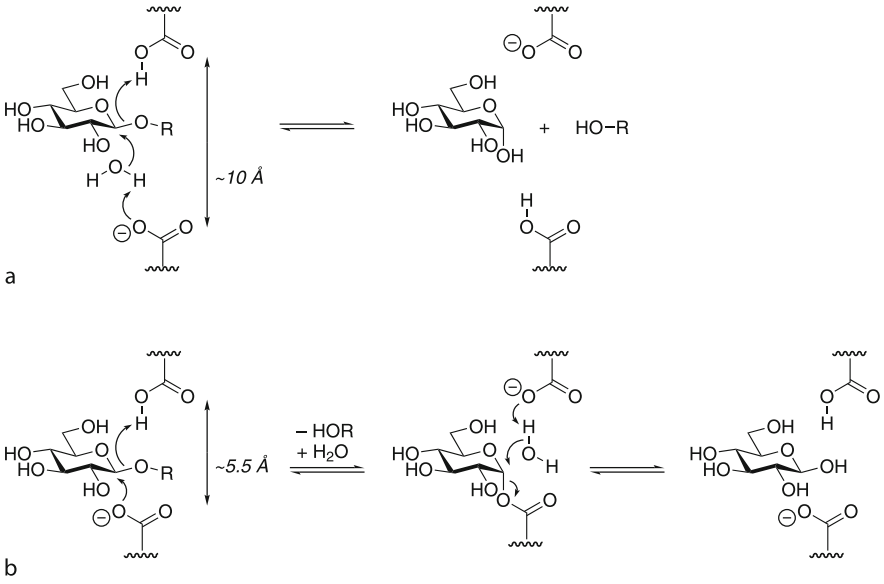


Scheme 2
Enzymatic galactosyl transfer from UDP-Gal to the 4-position of GlcNAc catalyzed by β1–4 galactosyltransferase (β1,4-GalT)

glycosyl transferases of the Leloir pathway [118] which are responsible for the synthesis of most glycoproteins and glycolipids in mammalian systems [119] are membrane-bound or membrane-associated enzymes. They utilize as glycosyl donors monosaccharides which are activated as glycosyl esters of nucleoside mono- and diphosphates (so-called sugar nucleotides) to transfer a glycosyl moiety to an acceptor [120,121,122]. The eight sugar nucleotides applied for the synthesis of most oligosaccharides are uridine 5′-diphosphoglucose (UDP-Glc), UDP-*N*-acetylglucosamine (UDP-GlcNAc), UDP-galactose (UDP-Gal), UDP-*N*-acetyl-galactosamine (UDP-GalNAc), UDP-glucuronic acid (UDP-GlcA), guanosine 5′-diphosphomannose (GDP-Man), GDP-fucose (GDP-Fuc), and cytidine 5′-monophospho *N*-acetylneuraminic acid (CMP-Neu5Ac). Non-Leloir transferases typically utilize glycosyl phosphates as donors. For each sugar nucleotide glycosyl donor, several Leloir transferases exist each of which acts on a different acceptor in a regio- and stereospecific manner. The transferases are generally considered to be specific for a given glycoside bond, however, some deviations from this picture of absolute specificity have been observed, both in the glycosyl donors and acceptors. β1–4 Galactosyltransferase for example catalyzes the transfer of galactose from UDP-Gal to the 4-position of β-linked GlcNAc residues to produce the Gal(β1–4)GlcNAc substructure (◉ *Scheme 2*). In the presence of lactalbumin, glucose is the preferred acceptor, resulting in the formation of lactose, Gal(β1–4)Glc [123].

Many glycosyltransferases have been employed in the in vitro synthesis of oligosaccharides [124,125]. Since the nucleoside mono- or diphosphates released during the reaction, e. g. UDP, usually cause product inhibition, in situ regeneration cycles have been developed in which they are transformed back to the sugar nucleotides [126]. Alternatively, the nucleoside mono- or diphosphates may be degraded by use of alkaline phosphatase [127].

Glycosidases (glycosyl hydrolases, EC 3.2.1) [128] catalyze the hydrolysis of glycosidic bonds. For the naturally occurring polysaccharides cellulose and starch typical rate constants up to 1000 s^{-1} are observed. Taking the half-lives for spontaneous hydrolysis of these materials (about 5 million years [129]) into account, these enzymes achieve rate enhancements of up to 10^{17}-fold, placing them amongst the most proficient of enzymes. In nature, glycosidases play important roles ranging from the degradation of polysaccharides as food sources



Scheme 3
Mechanisms of (a) inverting and (b) retaining glycosidases. Both mechanisms involve transition states with substantial oxocarbenium character present at the anomeric center

through to the trimming process of glycoproteins. The hydrolysis reaction can occur with either inversion or retention of anomeric configuration. Depending on the stereochemical outcome, two different mechanisms have been suggested 1953 by Koshland [130] and later on refined [131,132,133,134]. Both mechanisms involve a pair of carboxylic acids at the active site (*Scheme 3*). In inverting glycosidases, these residues are approximately 10 Å apart from each other on average. Reaction occurs by a single-displacement mechanism with one carboxylic acid acting as a general base activating a water molecule and the other as a general acid (*Scheme 3a*). In retaining enzymes, the carboxyl groups are only approximately 5.5 Å apart and the hydrolysis proceeds by a double-displacement mechanism involving a covalent glycosyl-enzyme intermediate (*Scheme 3b*). In both mechanisms, substantial oxocarbenium character has been shown to be present at the anomeric center in each transition state. The design of transition-state analogs mimicking the shape and charge of the oxocarbenium species has been extensively used to arrive at glycosidase inhibitors [135]. An example of a designed inhibitor which was approved by the Food and Drug Administration (FDA) in 2000 for the treatment of influenza virus A and B infections is the sialidase inhibitor zanamivir (Relenza^{TM}) [136]. The search for a drug with improved pharmacokinetic properties led to the zanamivir mimic oseltamivir that came on the market in 2000 under the name of Tamiflu^{TM} which is the phosphate salt of an ethyl ester prodrug [137].

In contrast to glycosyltransferases, glycosidases show broader, sometimes overlapping substrate specificities, especially with respect to the aglycon. Endoglucanases, for example, whilst typically considered to be cellulases, are also active to various degrees on xylan, xyloglucan, β-glucan, and various artificial substrates. Retaining glycosidases, run in the transglycosyla-

tion mode, have also been used for the construction of glycosidic bonds, albeit in moderate yields, the major problem being the thermodynamically favored hydrolysis of the reaction product [125,138]. Withers et al. could overcome this problem with a mutated β-glucosidase, wherein the nucleophilic carboxyl had been substituted with alanine [139]. The mutated glycosidase (a so-called glycosynthase [140,141,142]) folds correctly but lacks hydrolase activity since it cannot form the requisite α-glycosyl-enzyme intermediate. It possesses, however, high transglycosylation activity. Employing an activated α-glycosyl fluoride having the opposite anomeric configuration to that of the normal substrate, the ligation to a suitable acceptor sugar bound in the aglycon pocket is catalyzed. Once the oligosaccharide is formed, it cannot be hydrolyzed by the enzyme, thus allowing yields of over 90% to be achieved.

3.3 Carbohydrate-Lectin Interactions

3.3.1 Classification of Lectins

Lectins [102,103,104,105,143,144,145,146] are carbohydrate-binding proteins other than immunoglobulins without enzymatic activity towards the recognized sugars [101]. They are present throughout nature including the animal kingdom and the microbial world. Originally, they were classified according to their monosaccharide binding specificity. However, with the advent of molecular cloning a more consistent classification based on amino acid sequence homology and evolutionary relatedness emerged [147]. Whereas the biological function of plant lectins is still unclear, their use as biochemical tools has made an enormous contribution to the understanding of the structures and functions of carbohydrate structures in animal cells. Today, we have a much more detailed picture of the function of animal lectins than of plant lectins. ❯ *Table 3* shows selected examples of the 13 currently known families of animal lectins [12,102,105].

C-Type lectins were named after their requirement of calcium ions for recognition. They are all characterized by an extracellular carbohydrate recognition domain (CRD) and bind a diver-

Table 3
Selected families of animal lectins classified according to known sequence homologies[a] [12,102,105]

Lectin family	Defining structural motif	Carbohydrate ligand	Calcium dependence
C-type (includes selectins, collectins, endocytic lectins)	Conserved CRD	Variable	Yes (most)
Galectins (formerly *S*-type)	Conserved CRD	β-galactosides	No
P-type	Unique repeating motif	Mannose-6-phosphate on high-mannose-type *N*-glycans	Variable
I-type (includes Siglec family)	Immunoglobulin-like CRD	Variable (Siglecs: sialic acids)	No
Calnexin, calreticulin, calmegin	Homology with each other	Glucosylated high-mannose-type *N*-glycans in the ER	Yes
Hyaluronan-binding proteins	Homologue CRD	Hyaluronan chains	No

[a] Abbreviations: CRD, carbohydrate recognition domain; ER, endoplasmic reticulum

sity of sugars. However, not all calcium-requiring lectins are *C*-type lectins, as exemplified by calnexin and calcireticulin which recognize glucose residues on newly synthesized *N*-glycoproteins. Galectins, formerly named *S*-type lectins due to their dependency on free thiols for full activity, are soluble β-galactoside-specific lectins that combine preferentially with lactose and *N*-acetyllactosamine. P-Type lectins bind mannose-6-phosphate as their ligand and I-type lectins share a common immunoglobulin-like CRD. Another class of evolutionarily very ancient circulating soluble lectins are the pentraxins. They are characterized not so much by primary sequence homologies, but by a pentameric arrangement of their subunits and a probable role in the primary host immune response.

Multivalency appears to play an important role in lectin-mediated interactions [106,148,149,150,151], and many lectins are found to recognize individual carbohydrate epitopes only with low affinity. Indeed, most lectins are either intrinsically multivalent because of their defined multisubunit structure or by virtue of having multiple CRDs within a single polypeptide or they can become functionally multimeric by noncovalent association or by clustering on cell surfaces. High-avidity binding can result from multiple interactions of adequately presented low-affinity single sites, and this appears to be a common mechanism of modulating lectin function in vivo [144,152,153,154].

The following sections will focus on selected, well-characterized carbohydrate-lectin recognition processes.

3.3.2 Lectin Control of Protein Folding

N-Linked glycan moieties of glycoproteins which are incorporated co-translationally in the endoplasmic reticulum (ER) during biosynthesis are in many cases essential for proper folding of the protein. Their conformational influence on nascent or newly synthesized glycoproteins may be either direct by inducing local structure elements, such as β-turns, (cf. ❯ *Sect. 2.4*) or indirect by interaction with calnexin or calreticulin within the so-called calnexin-calreticulin-cycle (❯ *Fig. 6*) [92,96,155,156]. Membrane-bound calnexin (CNX) and soluble calreticulin (CRT) are homologous lectins found in the ER of nearly all eukaryotes. They bind to a monoglucosylated form of *N*-glycoproteins after two of the glucoses have been trimmed away by glucosidases I and II. Binding by calnexin or calreticulin retains the unfolded glycoprotein in the ER, permitting foldases to help folding into the correct 3D arrangement. Folding factor ERp57 is a thiol oxidoreductase homologue of protein disulfide isomerase that also binds to both calnexin and calreticulin and thereby is exposed to the substrate glycoprotein. If the glycoprotein contains cysteines, the formation of proper disulfide bonds is catalyzed through the formation of transient mixed disulfides with ERp57. Association of the glycoprotein with calnexin and calreticulin is dynamic and exposes the single terminal glucose residue to hydrolysis by glucosidase II. If the deglucosylated glycoprotein is incorrectly or incompletely folded, it is either reglucosylated by UDP-glucose glycoprotein glucosyltransferase, which has the remarkable function of a folding sensor, or degraded. Reglucosylation allows another cycle of binding by calnexin or calreticulin and interaction with ERp57. Once correctly folded, the protein is no longer recognized by the glucosyltransferase and, therefore, does not bind back any more to calnexin and/or calreticulin but can leave the ER. Exit of certain glycoproteins from the ER to the Golgi complex is assisted by another membrane-bound lectin, ERGIC-53, which binds to mannose residues.

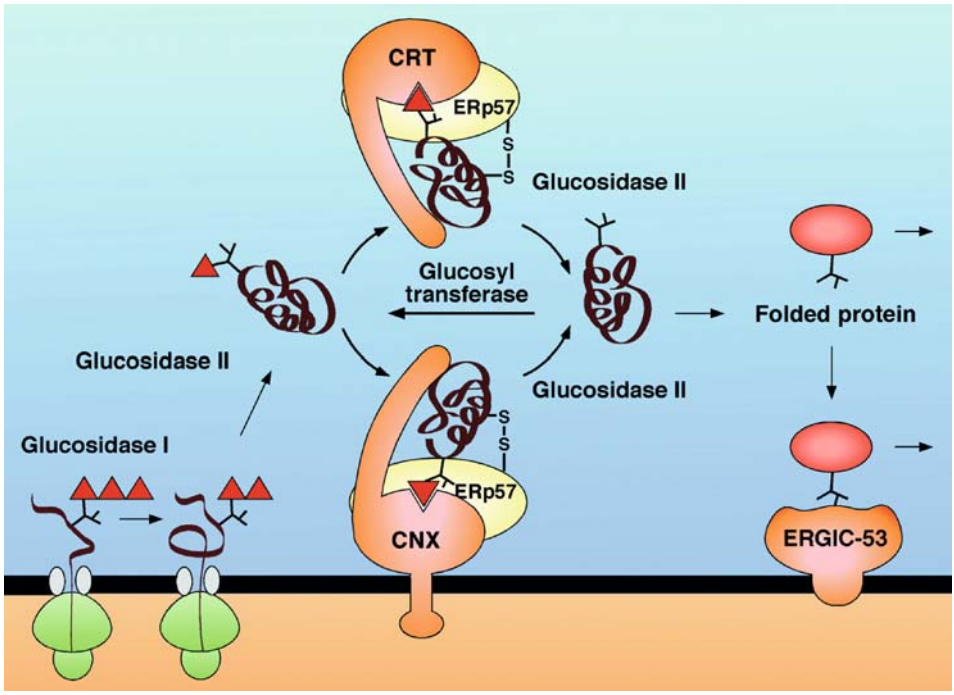


Figure 6
The calnexin-calreticulin-cycle (cf. text). Triangles represent glucose residues (Reprinted from [92] with permission from AAAS)

The calnexin-calreticulin-cycle promotes correct folding, inhibits aggregation of folding intermediates, and provides quality control by preventing incompletely folded glycoproteins from exiting to the Golgi complex. It seems to be essential in vivo. Transgenic mice devoid of calreticulin die on embryonic day 18 [157].

3.3.3 Clearance and Targeting of Glycoproteins

Some effects of glycosylation on the stability of proteins have already been mentioned in *Sect. 2.2* and they can presumably affect their half-life in single cells. In the intact organism, recognition of glycan structures by certain receptors can result in removal of the glycoconjugate or even a whole cell from the circulation. This has first been observed in the late 1960s by Ashwell and coworkers who serendipitously found that desialylation of glycoproteins resulted in significantly shorter serum half-lives [158,159]. This led to the characterization of the asialoglycoprotein receptor [160] (a hepatic lectin of the *C*-type) which has recently been crystallized [161] and not only recognizes desialylated, i. e. galactose terminated, *N*-linked oligosaccharides but also GalNAc structures as they occur on some *O*-linked glycans. These findings were traditionally interpreted as representing a physiological clearing mechanism for glycoproteins [160] which has, however, not yet proven beyond doubt [162]. Another exam-

ple of a receptor responsible for glycoprotein clearing is the GalNAc-4-sulfate receptor. Rapid removal of luteinizing hormone, for example, which contains this (*N*-linked) sugar is important to generate a circadian rhythm [163].

The mannose 6-phosphate (Man6*P*) receptors are the best understood examples of receptors responsible for intracellular trafficking of glycoproteins [164,165,166]. These P-type lectins mediate the routing of lysosomal hydrolases from the trans-Golgi network to their final destination in the lysosomes. Two such receptors are known, one cation-dependent and of low molecular weight (ca. 45 kDa) which has been crystallized [165], the other cation-independent and of high molecular weight (ca. 275 kDa). The targeting is mediated by recognition of Man6*P* residues on oligomannose-type *N*-glycans of lysosomal enzymes by the Man6*P* receptors. A defect in the synthesis of the Man6*P* markers, caused by a deficiency of GlcNAc-phosphotransferase (the first enzyme in the mannose phosphorylation pathway), results in I-cell disease (also called mucolipidosis II or MLII), an inherited lysosomal storage disease.

3.3.4 Leukocyte Trafficking

The leukocyte trafficking to inflammatory sites is a highly regulated multistep process, referred to as the inflammatory cascade [167]. The initial events, the tethering and rolling of leukocytes along the vascular endothelium, are mediated by the interaction of a family of adhesion molecules, termed selectins, and their carbohydrate containing ligands [168,169,170]. The cascade begins with the release of cytokines and other signaling molecules at the site of injury that stimulate the transient expression of E- and P-selectin on the endothelium surface. These *C*-type lectins bind to their ligands displayed on the circulating leukocytes and promote leukocyte adhesion to the stimulated endothelial cells. L-Selectin is constitutively expressed on leukocytes, and it recognizes its ligands on endothelial cells. The rolling leads to activation of integrins on the leukocytes that interact with their counter-receptors on endothelial cells (e. g. intercellular adhesion molecule-1, ICAM-1) and promote firm adhesion. This stronger interaction then allows emigration or extravasation of the leukocytes into the underlying damaged tissue. However, if too many leukocytes are recruited to the site of injury, normal cells can also be destroyed. This occurs in the condition known as septic shock, in chronic inflammatory diseases such as psoriasis and rheumatoid arthritis, and in the reperfusion tissue injury that occurs following a heart attack, stroke or organ transplant.

The selectins are membrane-bound proteins comprising five domains: a cytosolic tail that may play a role in signal transduction, a transmembrane domain, several complement regulatory domain repeats, an epidermal growth factor (EGF) domain, and an *N*-terminal, calcium-dependent *C*-type carbohydrate recognition domain (CRD). Both the EGF domain and the CRD are required for ligand binding, although the site of binding has been localized to the CRD. The tetrasaccharide sialyl Lewis x (sLe^x) Neu5Ac(α2–3)Gal(β1–4)[Fuc(α1–3)]GlcNAc (*Fig. 7*) is the minimum structure recognized by all three selectins. It is bound, however, only weakly with K_D values of 0.7, 3.9, and 7.8 mM for binding to E-, L-, and P-selectin, respectively [171]. Since the interaction of the selectins with their natural glycoprotein ligands which contain sLe^x and modifications thereof as terminating structures is much stronger, additional receptor-ligand contacts seem to be important for high-affinity binding. The natural ligand for L-selectin (GlyCAM-1) for example contains sLe^x sulfated at Gal-6, GlcNAc-6, or both positions and in the natural ligand for P-selectin, PSGL-1, the 19 amino acid *N*-terminus of the

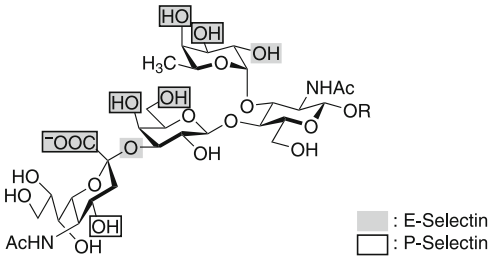


Figure 7
The tetrasaccharide Neu5Ac(α2–3)Gal(β1–4)[Fuc(α1–3)]GlcNAc (sialyl Lewis x), the minimum structure recognized the selectins. Indicated are functional groups involved in specific contacts to E-selectin (*gray boxes*) and P-selectin (*light boxes*) according to crystal structures [172]

protein which contains several sulfated tyrosine residues beside the *O*-linked, sLe^x-bearing glycan was found to be critical for binding. Recently, the crystal structures of human E- and P-selectin constructs containing the lectin and EGF (LE) domains co-complexed with sLe^x as well as the crystal structure of P-selectin LE in complex with the *N*-terminal domain of human PSGL-1 modified by both tyrosine sulfation and sLe^x have been reported [172]. This study offers a structural basis for affinity differences between sLe^x and E- and P-selectin, respectively and the high-affinity interaction between P-selectin and PSGL-1. In addition, it may well be that multivalent interactions between the selectins clustered on cell surfaces with multiple sLe^x residues presented on the highly glycosylated mucin-type counter-receptors contribute to high-affinity binding in vivo.

Inhibition of the selectin-ligand interactions is an attractive strategy for treating inflammation-related diseases [173,174,175]. As such, sLe^x has been intensively used as lead structure for development of anti-inflammatory drugs both in industry and academia [176,177,178].

References

1. Varki A (1993) Glycobiology 3:97
2. Lis H, Sharon N (1993) Eur J Biochem 218:1
3. Fukuda M, Hindsgaul O (eds) (1994) Molecular Glycobiology. IRL Press, Oxford
4. Karlsson K-A (1995) Curr Opin Struct Biol 5:622
5. Montreuil J, Vliegenthart JFG, Schachter H (eds) (1995) Glycoproteins. Elsevier, Amsterdam
6. Montreuil J, Vliegenthart JFG, Schachter H (eds) (1997) Glycoproteins II. Elsevier, Amsterdam
7. Dwek RA (1996) Chem Rev 96:683
8. Gabius H-J, Gabius S (eds) (1997) Glycosciences: Status and Perspectives. Chapman & Hall, Weinheim
9. Rudd PM, Dwek RA (1997) Crit Rev Biochem Mol Biol 32:1
10. Van den Steen P, Rudd PM, Dwek RA, Opdenakker G (1998) Crit Rev Biochem Mol Biol 33:151
11. Drickamer K, Taylor ME (1998) Trends Biochem Sci 23:321
12. Varki A, Cummings R, Esko J, Freeze H, Hart G, Marth J (eds) (1999) Essentials of Glycobiology. Cold Spring Harbor Laboratory Press, Cold Spring Harbor
13. Haines N, Irvine KD (2003) Nat Rev Mol Cell Biol 4:786
14. Paulson JC, Blixt O, Collins BE (2006) Nat Chem Biol 2:238
15. Lehle L, Strahl S, Tanner W (2006) Angew Chem Int Ed 45:6802

16. Taylor ME, Drickamer K (2006) Introduction to Glycobiology, 2nd ed. Oxford University Press, Oxford
17. Arnold JN, Wormald MR, Sim RB, Rudd PM, Dwek RA (2007) Annu Rev Immunol 25:21
18. Seifert GJ, Roberts K (2007) Annu Rev Plant Biol 58:137
19. Wittmann V (ed) (2007) Glycopeptides and Glycoproteins: Synthesis, Structure, and Application (Topics in Current Chemistry, vol. 267). Springer-Verlag, Berlin Heidelberg New York
20. Dube DH, Bertozzi CR (2003) Curr Opin Chem Biol 7:616
21. Prescher JA, Bertozzi CR (2005) Nat Chem Biol 1:13
22. Blixt O, Head S, Mondala T, Scanlan C, Huflejt ME, Alvarez R, Bryan MC, Fazio F, Calarese D, Stevens J, Razi N, Stevens DJ, Skehel JJ, van Die I, Burton DR, Wilson IA, Cummings R, Bovin N, Wong C-H, Paulson JC (2004) Proc Natl Acad Sci USA 101:17033
23. Feizi T, Chai W (2004) Nat Rev Mol Cell Biol 5:582
24. Shin I, Park S, Lee M-r (2005) Chem Eur J 11:2894
25. de Paz JL, Seeberger PH (2006) QSAR Comb Sci 25:1027
26. Sharon N, Lis H (1997) Glycoproteins: Structure and Function. In: Gabius H-J, Gabius S (eds) Glycosciences: Status and Perspectives. Chapman & Hall, Weinheim, p 133
27. Hanisch F-G (2001) Biol Chem 382:143
28. Hang HC, Bertozzi CR (2005) Bioorg Med Chem 13:5021
29. Mulloy B, Rider CC (2006) Biochem Soc Trans 34:409
30. Bülow HE, Hobert O (2006) Annu Rev Cell Dev Biol 22:375
31. Taylor KR, Gallo RL (2006) FASEB J 20:9
32. Imberty A, Lortat-Jacob H, Perez S (2007) Carbohydr Res 342:430
33. Bishop JR, Schuksz M, Esko JD (2007) Nature 446:1030
34. Yeh Y, Feeney RE (1996) Chem Rev 96: 601
35. Wöhrmann APA (1998) Soc Exp Biol Semin Ser 66:266
36. Ewart KV, Lin Q, Hew CL (1999) Cell Mol Life Sci 55:271
37. Fletcher GL, Hew CL, Davies PL (2001) Annu Rev Physiol 63:359
38. Harding MM, Anderberg PI, Haymet ADJ (2003) Eur J Biochem 270:1381
39. Tachibana Y, Fletcher GL, Fujitani N, Tsuda S, Monde K, Nishimura S-I (2004) Angew Chem Int Ed 43:856
40. Coyne KE, Hall SE, Thompson S, Arce MA, Kinoshita T, Fujita T, Anstee DJ, Rosse W, Lublin DM (1992) J Immunol 149: 2906
41. Olsen O, Thomsen KK (1991) J Gen Microbiol 137:579
42. Wyss DF, Wagner G (1996) Curr Opin Biotechnol 7:409
43. Imperiali B, O'Connor SE (1999) Curr Opin Chem Biol 3:643
44. Wormald MR, Dwek RA (1999) Structure 7:R155
45. Gerken TA, Butenhof KJ, Shogren R (1989) Biochemistry 28:5536
46. Wormald MR, Wooten EW, Bazzo R, Edge CJ, Feinstein A, Rademacher TW, Dwek RA (1991) Eur J Biochem 198:131
47. Joao HC, Scragg IG, Dwek RA (1992) FEBS Lett 307:343
48. Joao HC, Dwek RA (1993) Eur J Biochem 218:239
49. Rudd PM, Joao HC, Coghill E, Fiten P, Saunders MR, Opdenakker G, Dwek RA (1994) Biochemistry 33:17
50. Wyss DF, Choi JS, Li J, Knoppers MH, Willis KJ, Arulanandam ARN, Smolyar A, Reinherz EL, Wagner G (1995) Science 269:1273
51. Mer G, Hietter H, Lefèvre J-F (1996) Nat Struct Biol 3:45
52. Woods RJ, Edge CJ, Dwek RA (1994) Nat Struct Biol 1:499
53. Williams RL, Greene SM, McPherson A (1987) J Biol Chem 262:16020
54. Woods RJ, Pathiaseril A, Wormald MR, Edge CJ, Dwek RA (1998) Eur J Biochem 258:372
55. Mori K, Dwek RA, Downing AK, Opdenakker G, Rudd PM (1995) J Biol Chem 270:3261
56. Opdenakker G, Rudd PM, Ponting CP, Dwek RA (1993) FASEB J 7:1330
57. Opdenakker G, Rudd PM, Wormald M, Dwek RA, Van Damme J (1995) FASEB J 9:453
58. O'Connor SE, Imperiali B (1996) Chem Biol 3:803
59. Meyer B, Möller H (2007) Top Curr Chem 267:187
60. Imperiali B (1997) Acc Chem Res 30:452
61. Imperiali B, O'Connor SE, Hendrickson T, Kellenberger C (1999) Pure Appl Chem 71:777
62. Sodora DL, Cohen GH, Muggeridge MI, Eisenberg RJ (1991) J Virol 65:4424

63. Kern G, Kern D, Jaenicke R, Seckler R (1993) Protein Sci 2:1862
64. Duranti M, Gius C, Sessa F, Vecchio G (1995) Eur J Biochem 230:886
65. Feng W, Matzuk MM, Mountjoy K, Bedows E, Ruddon RW, Boime I (1995) J Biol Chem 270:11851
66. Lehmann S, Harris DA (1997) J Biol Chem 272:21479
67. Imperiali B, Rickert KW (1995) Proc Natl Acad Sci USA 92:97
68. O'Connor SE, Imperiali B (1997) J Am Chem Soc 119:2295
69. O'Connor SE, Imperiali B (1998) Chem Biol 5:427
70. Bosques CJ, Tschampel SM, Woods RJ, Imperiali B (2004) J Am Chem Soc 126:8421
71. Imberty A, Pérez S (1995) Protein Eng 8: 699
72. Petrescu A-J, Milac A-L, Petrescu SM, Dwek RA, Wormald MR (2004) Glycobiology 14: 103
73. Petrescu A-J, Wormald MR, Dwek RA (2006) Curr Opin Struct Biol 16:600
74. Live DH, Kumar RA, Beebe X, Danishefsky SJ (1996) Proc Natl Acad Sci USA 93:12759
75. Otvos L, Jr., Krivulka GR, Urge L, Szendrei GI, Nagy L, Xiang ZQ, Ertl HCJ (1995) Biochim Biophys Acta 1267:55
76. Andreotti AH, Kahne D (1993) J Am Chem Soc 115:3352
77. Liang R, Andreotti AH, Kahne D (1995) J Am Chem Soc 117:10395
78. Huang X, Smith MC, Berzofsky JA, Barchi JJ Jr (1996) FEBS Lett 393:280
79. Simanek EE, Huang DH, Pasternack L, Machajewski TD, Seitz O, Millar DS, Dyson HJ, Wong CH (1998) J Am Chem Soc 120:11567
80. Wu WG, Pasternack L, Huang DH, Koeller KM, Lin CC, Seitz O, Wong CH (1999) J Am Chem Soc 121:2409
81. Coltart DM, Royyuru AK, Williams LJ, Glunz PW, Sames D, Kuduk SD, Schwarz JB, Chen X-T, Danishefsky SJ, Live DH (2002) J Am Chem Soc 124:9833
82. Schuster O, Klich G, Sinnwell V, Kranz H, Paulsen H, Meyer B (1999) J Biomol NMR 14:33
83. Naganagowda GA, Gururaja TL, Satyanarayana J, Levine MJ (1999) J Pept Res 54:290
84. Kirnarsky L, Prakash O, Vogen SM, Nomoto M, Hollingsworth MA, Sherman S (2000) Biochemistry 39:12076
85. Grinstead JS, Koganty RR, Krantz MJ, Longenecker BM, Campbell AP (2002) Biochemistry 41:9946
86. Kirnarsky L, Suryanarayanan G, Prakash O, Paulsen H, Clausen H, Hanisch FG, Hollingsworth MA, Sherman S (2003) Glycobiology 13:929
87. von Mensdorff-Pouilly S, Kirnarsky L, Engelmann K, Baldus SE, Verheijen RH, Hollingsworth MA, Pisarev V, Sherman S, Hanisch FG (2005) Glycobiology 15:735
88. Möller H, Serttas N, Paulsen H, Burchell JM, Taylor-Papadimitriou J, Meyer B (2002) Eur J Biochem 269:1444
89. Dziadek S, Griesinger C, Kunz H, Reinscheid UM (2006) Chem Eur J 12:4981
90. Laine RA (1997) The Information-Storing Potential of the Sugar Code. In: Gabius H-J, Gabius S (eds) Glycosciences: Status and Perspectives. Chapman & Hall, Weinheim, p 1
91. Lee YC, Lee RT (1995) Acc Chem Res 28:321
92. Helenius A, Aebi M (2001) Science 291:2364
93. Rudd PM, Elliott T, Cresswell P, Wilson IA, Dwek RA (2001) Science 291:2370
94. Lowe JB, Marth JD (2003) Annu Rev Biochem 72:643
95. Gabius H-J, Siebert H-C, André S, Jiménez-Barbero J, Rüdiger H (2004) ChemBioChem 5:740
96. Helenius A, Aebi M (2004) Ann Rev Biochem 73:1019
97. Freeze HH, Aebi M (2005) Curr Opin Struct Biol 15:490
98. Dube DH, Bertozzi CR (2005) Nat Rev Drug Discovery 4:477
99. Freeze HH (2006) Nat Rev Genetics 7:537
100. Haltiwanger RS (2002) Curr Opin Struct Biol 12:593
101. Barondes SH (1988) Trends Biochem Sci 13:480
102. Gabius H-J (1997) Eur J Biochem 243:543
103. Lis H, Sharon N (1998) Chem Rev 98:637
104. Kilpatrick DC (2000) Handbook of Animl Lectins: Properties and Biomedical Applications. Wiley, Chichester
105. Gabius H-J (ed) (2002) Animal Lectins, Biochim Biophys Acta 1572:163
106. Dam TK, Brewer CF (2002) Chem Rev 102:387
107. Sharon N (2007) J Biol Chem 282:2753
108. Landsteiner K (1900) Zentralbl Bakt Parasitenkd Infektionskrankh Abt 1 27:357
109. Klein HG, Anstee DJ (2005) Mollison's Blood Transfusion in Clinical Medicine, 11th edn. Blackwell, Malden

110. Watkins WM (1995) Molecular basis of antigenic specificity in the ABO, H and Lewis blood-group systems. In: Montreuil J, Vliegenthart JFG, Schachter H (eds) Glycoproteins. Elsevier, Amsterdam, p 313
111. Daniels GL, Anstee DJ, Cartron JP, Dahr W, Issitt PD, Jorgensen J, Kornstad L, Levene C, Lomasfrancis C, Lubenko A, Mallory D, Moulds JJ, Okubo Y, Overbeeke M, Reid ME, Rouger P, Seidl S, Sistonen P, Wendel S, Woodfield G, Zelinski T (1995) Vox Sang 69:265
112. Stamatoyannopoulos G, Majerus PW, Perlmutter RM, Varmus H (2000) The Molecular Basis of Blood Diseases, 3rd ed. W. B. Saunders, Orlando, Florida
113. Morgan WTJ, Watkins WM (2000) Glycoconjugate J 17:501
114. Clausen H, Hakomori S-I (1989) Vox Sang 56:1
115. Hakomori S-I, Zhang Y (1997) Chem Biol 4:97
116. Keding SJ, Danishefsky SJ (2003) Synthetic Carbohydrate-Based Vaccines. In: Wong C-H (ed) Carbohydrate-Based Drug Discovery, vol 1. Wiley-VCH, Weinheim, p 381
117. Warren J, Geng X, Danishefsky S (2007) Top Curr Chem 267:109
118. Leloir LF (1971) Science 172:1299
119. Brockhausen I, Schachter H (1997) Glycosyltransferases Involved in *N*- and *O*-Glycan Biosynthesis. In: Gabius H-J, Gabius S (eds) Glycosciences: Status and Perspectives. Chapman & Hall, Weinheim, p 79
120. Kornfeld R, Kornfeld S (1985) Annu Rev Biochem 54:631
121. Ünligil UM, Rini JM (2000) Curr Opin Struct Biol 10:510
122. Hu Y, Walker S (2002) Chem Biol 9:1287
123. Ramakrishnan B, Qasba PK (2001) J Mol Biol 310:205
124. Koeller KM, Wong C-H (2000) Chem Rev 100:4465
125. Trincone A, Giordano A (2006) Curr Org Chem 10:1163
126. Wong CH, Haynie SL, Whitesides GM (1982) J Org Chem 47:5416
127. Unverzagt C, Kunz H, Paulson JC (1990) J Am Chem Soc 112:9308
128. Henrissat B, Davies G (1997) Curr Opin Struct Biol 7:637
129. Wolfenden R, Lu X, Young G (1998) J Am Chem Soc 120:6814
130. Koshland DE (1953) Biol Rev 28:416
131. Zechel DL, Withers SG (2001) Curr Opin Chem Biol 5:643
132. Zechel DL, Withers SG (2000) Acc Chem Res 33:11
133. Vasella A, Davies GJ, Bohm M (2002) Curr Opin Chem Biol 6:619
134. Heightman TD, Vasella AT (1999) Angew Chem 111:794
135. Lillelund VH, Jensen HH, Liang X, Bols M (2002) Chem Rev 102:515
136. von Itzstein M, Wu W-Y, Kok GB, Pegg MS, Dyason JC, Jin B, Van Phan T, Smythe ML, White HF, Oliver SW, Colman PM, Varghese JN, Ryan DM, Woods JM, Bethell RC, Hotham VJ, Cameron JM, Penn CR (1993) Nature 363:418
137. McClellan K, Perry CM (2001) Drugs 61:261
138. Crout DHG, Vic G (1998) Curr Opin Chem Biol 2:98
139. Mackenzie LF, Wang Q, Warren RAJ, Withers SG (1998) J Am Chem Soc 120:5583
140. Williams SJ, Withers SG (2002) Aust J Chem 55:3
141. Hancock SM, Vaughan MD, Withers SG (2006) Curr Opin Chem Biol 10:509
142. Perugino G, Cobucci-Ponzano B, Rossi M, Moracci M (2005) Adv Synth Catal 347:941
143. Drickamer K, Taylor ME (1993) Annu Rev Cell Biol 9:237
144. Rini JM (1995) Annu Rev Biophys Biomol Struct 24:551
145. Weis WI, Drickamer K (1996) Annu Rev Biochem 65:441
146. Vijayan M, Chandra N (1999) Curr Opin Struct Biol 9:707
147. Drickamer K (1988) J Biol Chem 263:9557
148. Drickamer K (1995) Nat Struct Biol 2:437
149. Lee RT, Lee YC (2000) Glycoconjugate J 17:543
150. Mammen M, Choi S-K, Whitesides GM (1998) Angew Chem Int Ed 37:2754
151. Lundquist JJ, Toone EJ (2002) Chem Rev 102:555
152. Rini JM, Lobsanov YD (1999) Curr Opin Struct Biol 9:578
153. Crocker PR, Feizi T (1996) Curr Opin Struct Biol 6:679
154. Collins BE, Paulson JC (2004) Curr Opin Chem Biol 8:617
155. Schrag JD, Procopio DO, Cygler M, Thomas DY, Bergeron JJM (2003) Trends Biochem Sci 28:49
156. Moremen KW, Molinari M (2006) Curr Opin Struct Biol 16:592
157. Mesaeli N, Nakamura K, Zvaritch E, Dickie P, Dziak E, Krause KH, Opas M, MacLennan DH, Michalak M (1999) J Cell Biol 144:857

158. Van den Hamer CJA, Morell AG, Scheinberg IH, Hickman J, Ashwell G (1970) J Biol Chem 245:4397
159. Morell AG, Gregoriadis G, Scheinberg IH, Hickman J, Ashwell G (1971) J Biol Chem 246:1461
160. Ashwell G, Harford J (1982) Annu Rev Biochem 51:531
161. Meier M, Bider MD, Malashkevich VN, Spiess M, Burkhard P (2000) J Mol Biol 300:857
162. Ashwell G (1994) The physiological role of the hepatic lectin for asialoglycoproteins: An evaluation of current hypotheses. In: Beuth J, Pulverer G (eds) Lectin Blocking: New Strategies for the Prevention and Therapy of Tumor Metastasis and Infectious Diseases (Zentralbl. Bakteriol., Suppl. 25). Gustav Fischer Verlag, Stuttgart, p 26
163. Hooper LV, Manzella SM, Baenziger JU (1997) The Biology of Sulfated Oligosaccharides. In: Gabius H-J, Gabius S (eds) Glycosciences: Status and Perspectives. Chapman & Hall, Weinheim, p 261
164. Kornfeld S (1990) Biochem Soc Trans 18: 367
165. Roberts DL, Weix DJ, Dahms NM, Kim J-JP (1998) Cell 93:639
166. Dahms NM, Hancock MK (2002) Biochim Biophys Acta 1572:317
167. Springer TA (1995) Annu Rev Physiol 57:827
168. Lasky LA (1995) Annu Rev Biochem 64:113
169. Kansas GS (1996) Blood 88:3259
170. McEver RP (2002) Curr Opin Cell Biol 14:581
171. Poppe L, Brown GS, Philo JS, Nikrad PV, Shah BH (1997) J Am Chem Soc 119:1727
172. Somers WS, Tang J, Shaw GD, Camphausen RT (2000) Cell 103:467
173. Newman W, Mirabelli CK (1998) Expert Opin Invest Drugs 7:19
174. Bendas G (1999) Pharm Ztg 144:3754
175. Wardlaw AJ (1999) Drugs Future 24:279
176. Simanek EE, McGarvey GJ, Jablonowski JA, Wong C-H (1998) Chem Rev 98:833
177. Kaila N, Thomas BE, IV (2002) Med Res Rev 22:566
178. Ernst B, Kolb HC, Schwardt O (2006) Carbohydrate Mimetics in Drug Discovery. In: Levy DE, Fügedi P (eds) The Organic Chemistry of Sugars. CRC Press, Boca Raton, FL, p 803

8.3 Biologically Relevant Glycopeptides: Synthesis and Applications

Clay S. Bennett[1], *Richard J. Payne*[2], *Kathryn M. Koeller*[1], *Chi-Huey Wong**[1]
[1] Chemistry Department, The Scripps Research Institute, La Jolla, CA 92037, USA
[2] School of Chemistry, The University of Sydney, Sydney, NSW, Australia
wong@scripps.edu

In: *Glycoscience*. Fraser-Reid B, Tatsuta K, Thiem J (eds)
Chapter-DOI 10-1007/978-3-540-30429-6_45:

Abstract

Over the past two decades interest in glycopeptides and glycoproteins has intensified, due in part to the development of new and efficient methods for the synthesis of these compounds. This includes a number of chemical and enzymatic techniques for incorporating glycosylation onto the peptide backbone as well as the introduction of powerful peptide ligation methods for the construction of glycoproteins. This review discusses these methods with a special emphasis on biologically relevant glycopeptides and glycoproteins. This includes the development of a number of antigens which hold promise as potential vaccines for HIV, cancer, or a host of other clinically important diseases. In addition the development of new antibiotics aimed at overcoming the problem of resistance will be discussed. Finally, chemical and enzymatic methods for the construction of glycopeptide mimetics will be described.

Keywords

Solid-phase; Ligation; Chemoenzymatic; Vaccine; Antigen; Conjugation; Antibiotic; Immune response; Inflammatory response; Conformation; Mimetic

1 Introduction

Glycosylation is a post- or co-translational modification of proteins that has extensive biological significance [1,2,3]. In general, protein function is influenced by the covalent attachment of an oligosaccharide to its surface. Altered protein glycosylation patterns often modify intercellular recognition processes, through modulation of protein–protein or protein–carbohydrate interactions [4]. Moreover, glycosylation increases peptide hydrophilicity, resulting in increased aqueous solubility and altered clearance properties compared with the naked peptide [5]. Glycosylation may also play a role in protein folding, as well as functioning as an inducer of protein secondary structure [6]. As such, the fields of glycobiology and carbohydrate chemistry offer a myriad of research opportunities in the biological arena.

At the structure-function level, however, the study of glycoproteins has been hampered by glycoprotein microheterogeneity [7]. Glycoproteins with identical peptide backbones intrinsically exist in numerous isoforms as a result of the attachment of diverse arrays of heterogeneous oligosaccharide chains. A single glycoprotein may present several different glycans, thereby making assessment of the contribution of a single glycoform to bioactivity difficult. To avoid this problem, many studies have taken an "all-or-none" approach, comparing function between glycosylated and fully de-glycosylated proteins [8,9]. Although research of this nature has provided valuable information, it does not allow exploration of the intricacies of specific glycan function.

Methods for the preparation of complex homogeneous glycopeptides have recently been developed [10,11,12,13,14]. Although research in this area is being actively pursued, it remains a relatively new field with a practical total synthesis of a natural homogeneous glycoprotein still to be achieved. The most comprehensive studies of specific glycoprotein function to date have been derived through the characterization of glycopeptide activity. Partial structures of relevant glycoproteins can usually be prepared synthetically in homogenous form, and in adequate quantity for biological research. Nevertheless, it is unclear whether structure-function relationships derived from synthetic glycopeptides are relevant to those of native glycoproteins. Recent efforts have therefore been largely directed toward the development of efficient methods for the preparation of homogeneous glycoproteins with well-defined carbohydrate structure. The development of new reliable chemical and enzymatic methods for glycopeptide synthesis has been instrumental in this field.

1.1 Naturally Occurring Glycopeptide Structure

The majority of all naturally occurring glycosidic linkages can be categorized into two groups. The first are the *N*-linked glycosides that are attached to the side chain amide of asparagine residues. The other, more diverse structures are *O*-linked glycosides, which are linked to the side chain hydroxyl groups of serine, threonine, or tyrosine. Less common linkages, which include *O*-glycosidic linkages to hydroxylysine or hydroxyproline, and *C*-glycosidic linkages to tyrosine, will not be discussed in the context of this chapter. A more detailed discussion of all known glycosyl-peptide linkages can be found in a number of detailed reviews [15,16,17].

1.1.1 *N*-Linked Glycans

Glycosylation through the side-chain amide of asparagine is referred to as *N*-linked glycosylation. This biosynthetic modification of proteins occurs co-translationally in the endoplasmic reticulum through the action of oligosaccharyltransferase (OT) [18]. The characteristic *N*-linked glycan contains a reducing terminal *N*-acetylglucosamine (GlcNAc) β-linked to asparagine (Fig. 1). In nature, this terminal GlcNAc is linked to an extensive network of other carbohydrates of the high mannose, hybrid, or complex type. OT is responsible for glycosylation of the peptide chain at the sequence Asn-X-Ser/Thr, referred to as the "Asx-turn motif" [19].

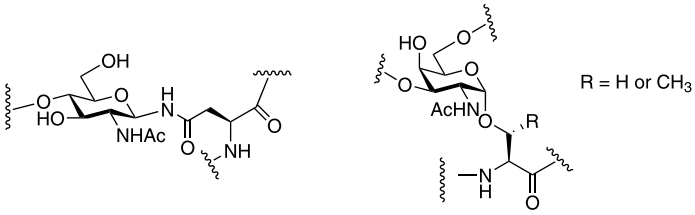


Figure 1
***N*-linked and *O*-linked glycosylation of peptides**

1.1.2 *O*-Linked Glycans

Glycosylation through the side chain hydroxyl of serine, threonine, or tyrosine is referred to as *O*-linked glycosylation. This glycosidic linkage is formed by transfer of the carbohydrate from a nucleotide sugar to the peptide chain, a reaction catalyzed by one of several glycosyltransferases [20]. The most commonly displayed *O*-linked glycan consists of a reducing terminal *N*-acetylgalactosamine (GalNAc) α-linked to serine or threonine (Fig. 1). Addition of subsequent carbohydrates to the GalNAc bridgehead is catalyzed by a number of tissue specific glycosyltransferases, hundreds of which have been characterized to date [3,21]. There are many other *O*-linked glycans observed in nature, examples include GlcNAc, fucose (Fuc), xylose (Xyl), mannose (Man), and glucose (Glc). In general no specific consensus sequence for *O*-glycosylation can be identified, however an important exception is the ser/thr/pro rich regions in the so called mucin domains. Mucins are expressed on the surface of a variety of epithelial cell types and contain numerous *O*-glycosylation sites [22]. These proteins represent an important class of tumor-associated antigens and, as such, are potential targets for tumor therapy.

1.2 Biologically Relevant Glycopeptide Target Structures

Synthetic glycopeptide target structures of interest are largely those associated with disease processes. For example, certain glycopeptide structures found on malignant cells, the so called tumor-associated glycopeptide antigens, have potential as vaccines for certain cancers [23,24]. In addition, the study of glycopeptide antigens mimicking glycoproteins on HIV has been

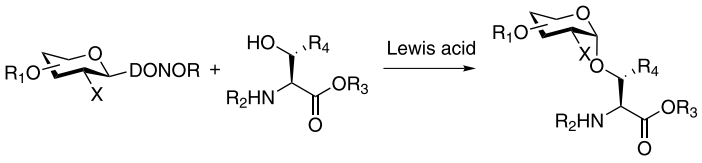
the subject of extensive efforts directed towards the development of HIV vaccines [25,26]. Furthermore, glycopeptides associated with intercellular adhesion events may be suitable for interference of the inflammatory pathway in reperfusion injury and asthma [27]. Certain glycopeptides are also potent antibacterial and antimicrobial agents [28]. Glycoprotein therapeutics have also found wide applicability for the effective treatment of anemia [29]. This chapter provides an overview of synthetic glycopeptide target structures of biological relevance along with synthetic and chemoenzymatic approaches aimed toward the synthesis of such targets. When available, results of biological studies are also included.

2 Synthetic Strategies

Synthetic approaches toward glycopeptide targets is a rapidly maturing field, and methodology has advanced to the point where most desired glycosylated amino acids are now routinely accessible, or even commercially available. Synthetic strategies for the construction of these amino acids have been reviewed in great detail previously and will not be discussed in detail here [15,16,30,31]. This section will highlight general synthetic and chemoenzymatic methods for the synthesis of glycopeptides available to the glycoscientist, including a discussion of methods for ligating glycopeptides to form more complex glycopeptide and glycoprotein structures.

2.1 Sequential Coupling of Glycosyl Amino Acids

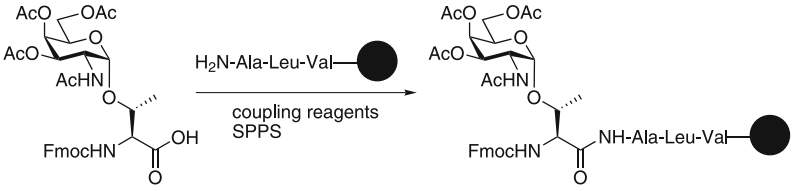
Numerous methods are available for the glycosylation of amino acids [32]. In the synthesis of *O*-linked constructs, the reaction of a suitable glycosyl donor with serine or threonine as the glycosyl acceptor is performed under conditions of Lewis acid catalysis (*Scheme 1*).



Scheme 1

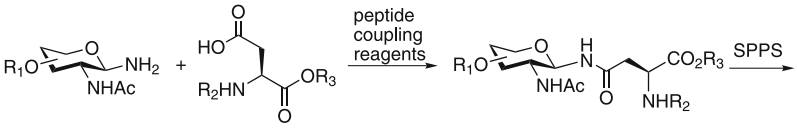
Depending on the carbohydrate protecting group pattern, α or β linkages can be formed specifically. Glycosyl donors employed can be any appropriate donor generally used for the synthesis of oligosaccharides. Suitably protected glycosyl amino acids can then be directly incorporated into solution-phase or solid-phase peptide synthesis (SPPS) [33] (*Scheme 2*). In one notable example, the Bock group has developed an efficient multiple column solid-phase synthetic method that allows facile construction of glycopeptide libraries. This procedure utilizes glycosyl amino acids activated as pentafluorophenyl esters [34]. More recently, Nishimura and coworkers demonstrated that microwave irradiation can facilitate coupling of hindered glycosyl amino acids [35].

In the synthesis of *N*-linked glycosyl amino acids, the coupling strategy is usually not a glycosylation at all, but merely amide bond formation [36]. In most common practice, glycosyl



Scheme 2

amines are synthesized and then attached to a protected aspartic acid through standard amide coupling strategies (*Scheme 3*). This type of *N*-linked building block can also be directly incorporated into SPPS to afford glycopeptides.

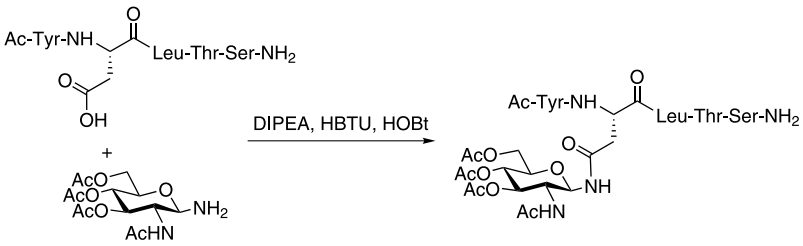


Scheme 3

2.2 Convergent Coupling of Glycans to Peptide Sequences

A convergent approach to glycopeptide synthesis was pioneered by Lansbury and coworkers [37,38]. This method allows the creation of *N*-linked glycopeptides by coupling of a glycosyl amine directly to a polypeptide chain containing an aspartic acid residue. The reaction conditions were optimized to minimize aspartamide side product formation (*Scheme 4*). Due to the nature of the coupling conditions, extensive protection of the peptide backbone and oligosaccharide are not necessary with this methodology. Danishefsky and co-workers have demonstrated the utility of this method through the synthesis of several glycopeptides [39].

Examples of convergent strategies for the synthesis of *O*-Glycopeptides have been reported, where the serine or threonine hydroxyl is used as a glycosyl acceptor [40]. These methods require extensive protection of the peptide and carbohydrate. Other pitfalls in this strategy



Scheme 4

include poor solubility of the peptide in organic solvents required for the glycosylation reaction, and lack of stereocontrol, thus leading to anomeric mixtures of the final product.

2.3 Chemical Glycopeptide Ligation Strategies

The methods described above permit facile access to structurally diverse, but relatively short glycopeptide structures (consisting of 10–50 amino acid residues). Above this limit, the linear nature of SPPS results in low yields of the target glycopeptides. For example, SPPS of a 50 amino acid peptide, averaging 95% yield per coupling step, would produce the desired oligomer in an overall yield of 7.7%. In addition, the accumulation of uncoupled sequences, side products and epimerization greatly increases the difficulty of purification of the final product. To overcome these limitations, the convergent ligation of fully unprotected glycopeptides has emerged as a viable strategy for the total synthesis of glycoproteins. Over the past decade many efficient chemoselective ligation methods have been developed, a summary of these is covered in the sections that follow. For a more detailed account of these methods, and their applications in the synthesis of glycopeptides and glycoproteins, the reader is referred to several recently published review articles [10,14,41,42].

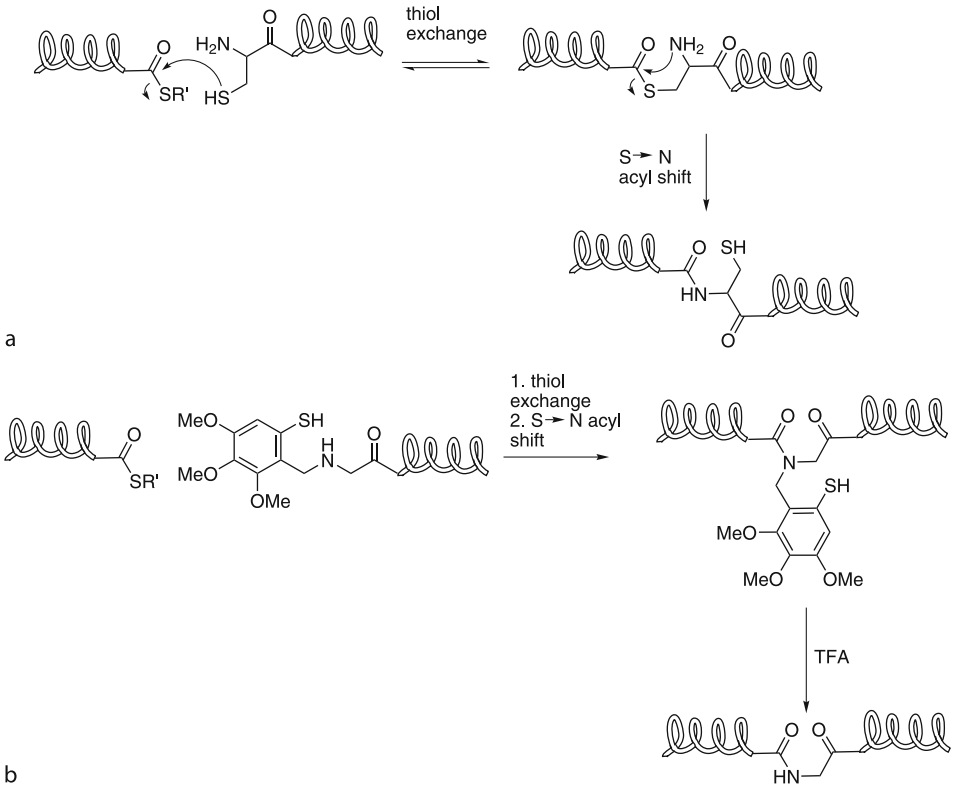
2.3.1 Native Chemical Ligation (NCL)

Native chemical ligation (NCL) is a chemoselective condensation reaction between a peptide thioester and a peptide bearing an *N*-terminal cysteine residue (❯ *Scheme 5a*). The first step of the ligation involves a rapid, reversible transthioesterification step. The resulting thioester intermediate then undergoes a spontaneous and irreversible S→N acyl transfer resulting in the formation of a new amide bond. The concept of NCL dates back to the 1950's with the pioneering work of Wieland et al. [43], however it was not until the 1990's that this method gained widespread attention as an efficient method for the ligation of peptides [44]. The scope of NCL has been extensively examined, and most amino acids undergo facile ligations, with the exception of those containing sterically encumbered side chains (Leu, Thr, Val, Ile, and Pro) [45]. The effectiveness of this technique is evident by its application in the synthesis of hundreds of different proteins [46]. The success of this method for the synthesis of peptides and proteins has inspired many laboratories to employ this technology for the synthesis of glycopeptides and glycoproteins. One key application of NCL was in a study towards the total synthesis of the glycopeptide antibiotic diptericin, which will be discussed in more detail in ❯ *Sect. 4.2* [47].

The obvious limitation of NCL is the requirement for a cysteine residue at the ligation junction. Cysteine has a relatively low abundance in proteins (ca. 1.7%) therefore there is a high probability that the target protein does not have cysteine at a synthetically useful position. This led to the development of a number of "cysteine-free ligation" strategies, namely auxiliary-assisted NCL, sugar-assisted ligation (SAL) and the traceless Staudinger ligation.

2.3.2 Cysteine-Free NCL

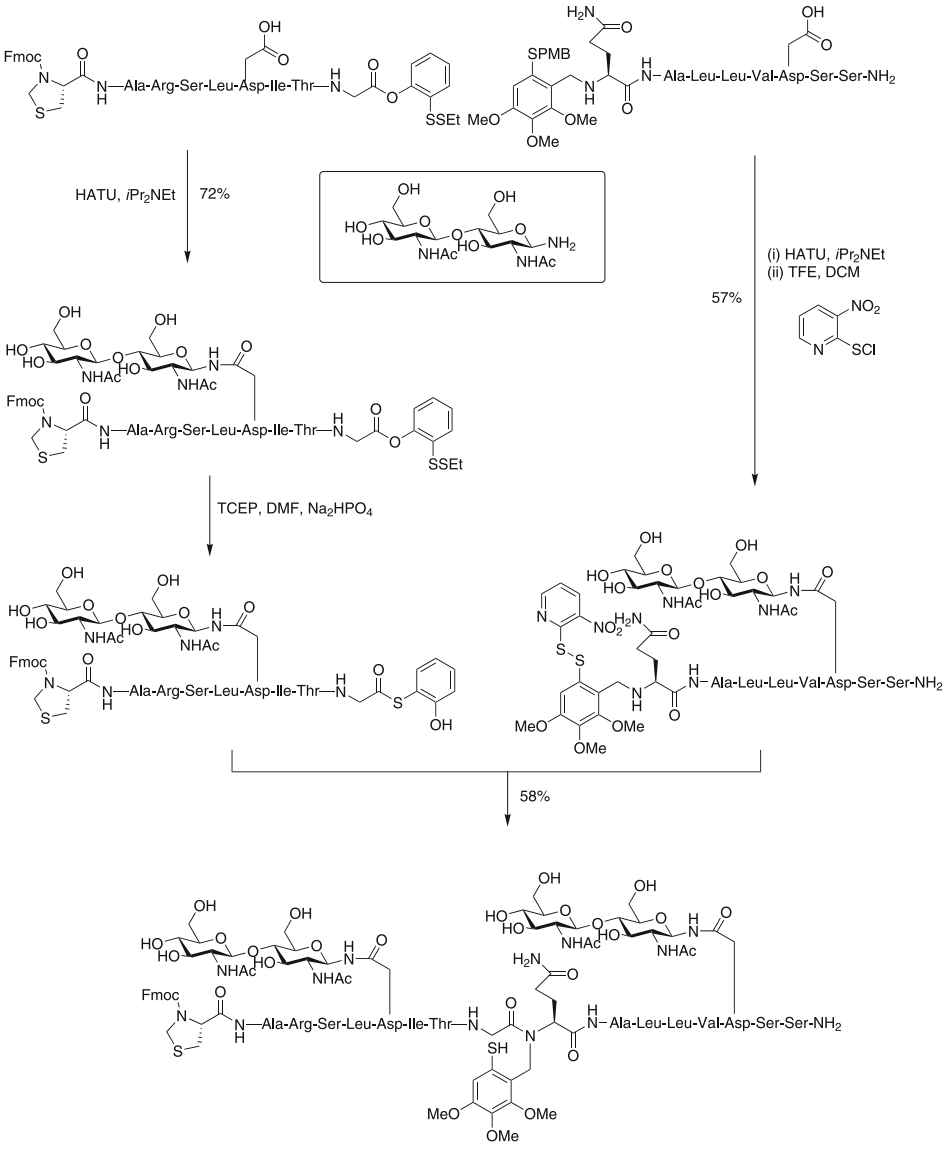
Initial studies directed at overcoming the requirement of an *N*-terminal cysteine in NCL led to the development of two thiol auxiliaries, 1-aryl-2-mercaptoethyl and 4,5,6-trimethoxy-2-mer-



Scheme 5

captobenzyl (Tmb) (*Scheme 5b*) [48,49]. These moieties are attached to the *N*-terminus of a synthetic peptide and can be readily removed following ligation by treatment with trifluoroacetic acid (TFA). The thiol of the auxiliary partakes in the chemical ligation in an identical fashion to the side chain of a cysteine residue. The application of this method to glycopeptide synthesis was first demonstrated by Macmillan and co-workers who used both auxiliaries to synthesize fragments of the *O*-linked glycoprotein GlyCAM-1 [50].

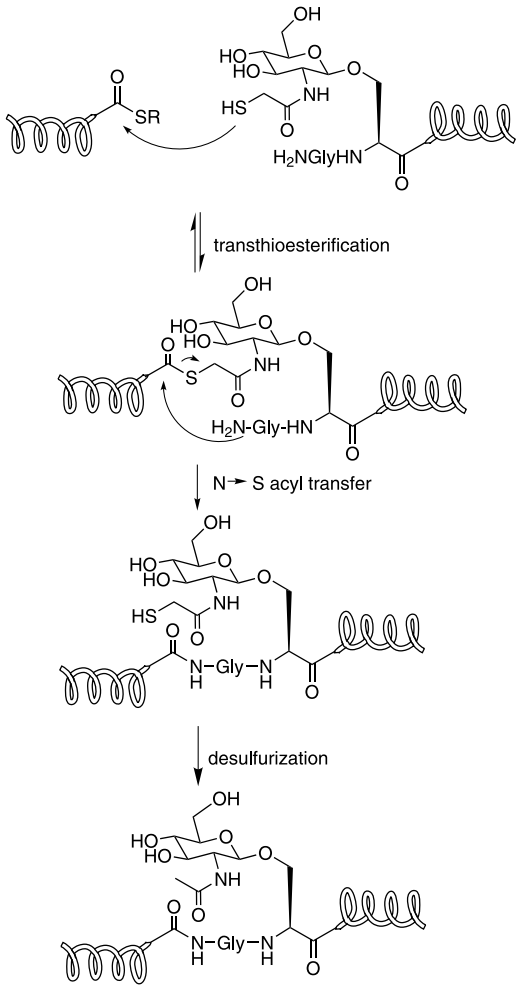
More recently, several elaborate examples of cysteine-free NCL were reported by Danishefsky and co-workers [51]. One example of this study is outlined in *Scheme 6*. The first step involved the synthesis of a masked glycopeptide thioester and a Tmb containing glycopeptide, using SPPS and the convergent Lansburn aspartylation for the introduction of glycans. The masked thioester was equipped with a *C*-terminal phenolic ester, which is converted to the desired thioester via an O→S acyl migration under the ligation conditions. The resulting thioester undergoes ligation with the Tmb containing glycopeptide (following in situ reduction of the disulfide bond) to afford a new peptide containing two glycosylation sites. This example is notable considering the established difficulty of ligations at a Gly-Gln junction. Following the ligation and removal of the auxiliary, the *N*-terminal cysteine can be unmasked to afford a handle for further ligations.



Scheme 6

2.3.3 Sugar Assisted Ligation (SAL)

Recently, another cysteine-free ligation strategy has been reported for the synthesis of glycopeptides. The method, dubbed sugar assisted ligation (SAL), utilizes a peptide thioester and a glycopeptide in which the carbohydrate is derivatized with a thiol auxiliary at the C-2 acetamide [52,53]. The proposed mechanism of this ligation, depicted in *Scheme 7*, involves

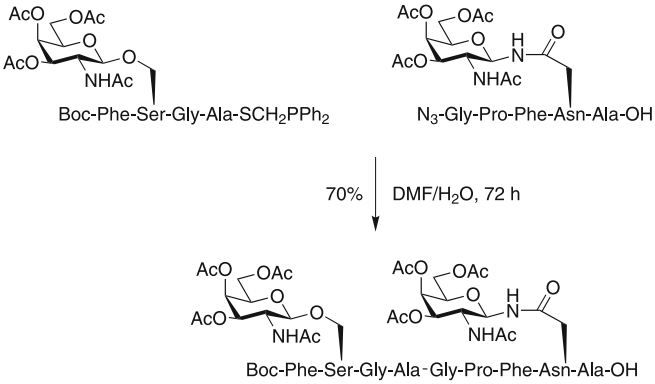


Scheme 7

an initial transthioesterification step, followed by an S→N acyl shift to afford the native peptide bond. The method has shown applicability to both *O*- and *N*-linked glycopeptides and tolerates a number of different amino acids at the ligation junction. Upon completion of the ligation, the thiol auxiliary can be efficiently cleaved under reducing conditions (H_2/Pd) to yield the native acetamide functionality at C-2.

2.3.4 Traceless Staudinger Ligation

A conceptually different ligation strategy that is currently being pursued in the synthesis of glycoproteins and glycopeptides is the "traceless Staudinger ligation." In this method an azide reacts with a phosphinothioester to form an iminophosphorane, which undergoes



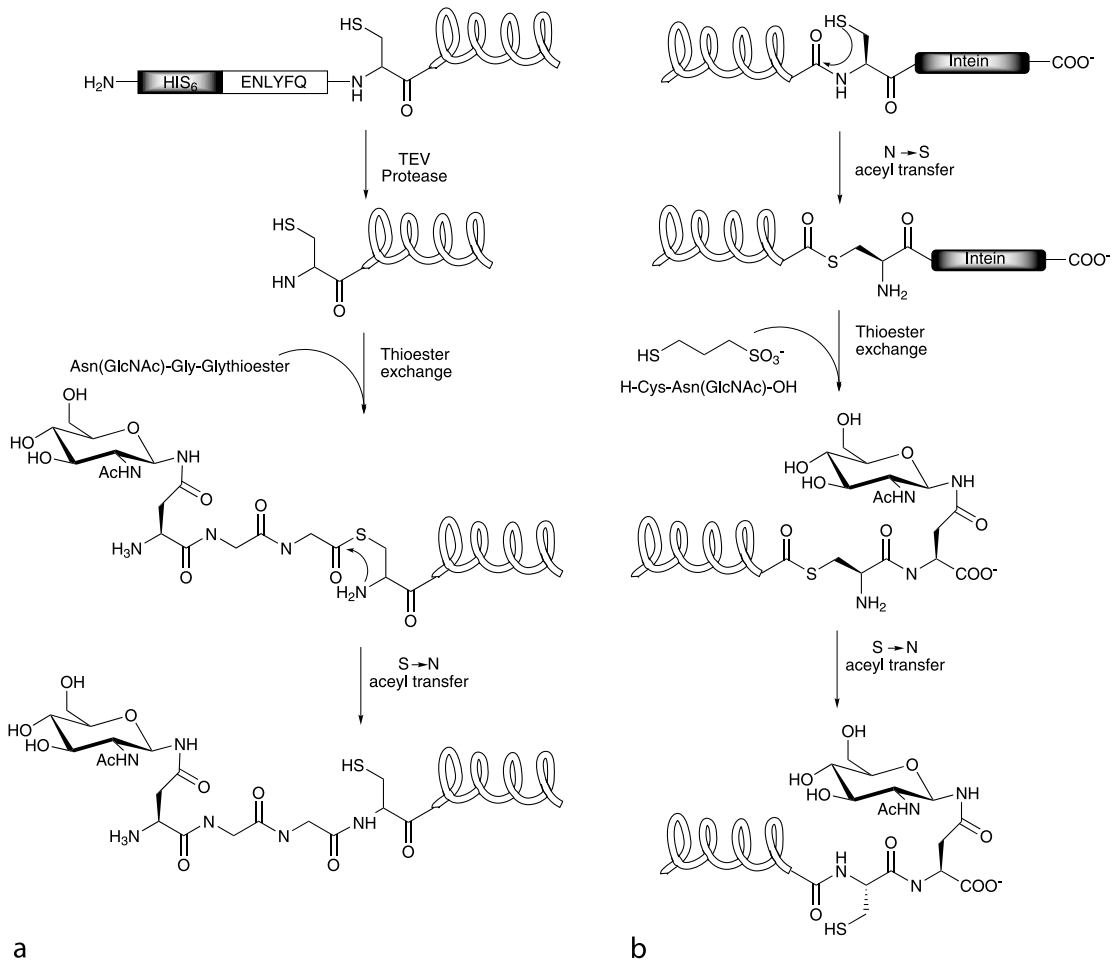
Scheme 8

an intramolecular S→N acyl transfer to generate an amidophosphonium salt intermediate (*Scheme 8*). Subsequent hydrolysis of the amidophosphonium salt yields the desired amide product [54]. Raines and coworkers recently reported that the 'traceless Staudinger ligation' proceeds in high yields (80–99% for dipeptide model systems) and is compatible with unprotected peptides [55,56]. More recently, this method has been extended to the synthesis of glycopeptides [57]. While the reaction requires sterically unencumbered amino acids at the ligation junction, both coupling partners were shown to tolerate glycosylation. The authors have also demonstrated that the *N*-terminal azide can be enzymatically coupled to a peptide using the engineered protease subtilisin BPN, opening up the possibility of extending this method to expressed protein ligation (EPL).

2.3.5 Expressed Protein Ligation (EPL)

The peptide ligation methods described above are powerful tools for the construction of complex glycopeptides, however for larger targets it often desirable to produce one of the two coupling partners recombinantly, a process known as expressed protein ligation (EPL). Methods have been developed for recombinant production of both *C*-terminal and *N*-terminal coupling partners. Production of the *C*-terminal coupling partner involves introduction of an *N*-terminal affinity tag followed by a protease recognition sequence and a cysteine residue. After the protein is isolated by affinity chromatography, it can be treated with an appropriate protease to afford a protein containing an *N*-terminal cysteine which can undergo NCL. Sequences which are susceptible to cleavage by both Factor Xa and tobacco etch virus protease (TEV protease) have been used in this process (*Scheme 9a*) [58,59].

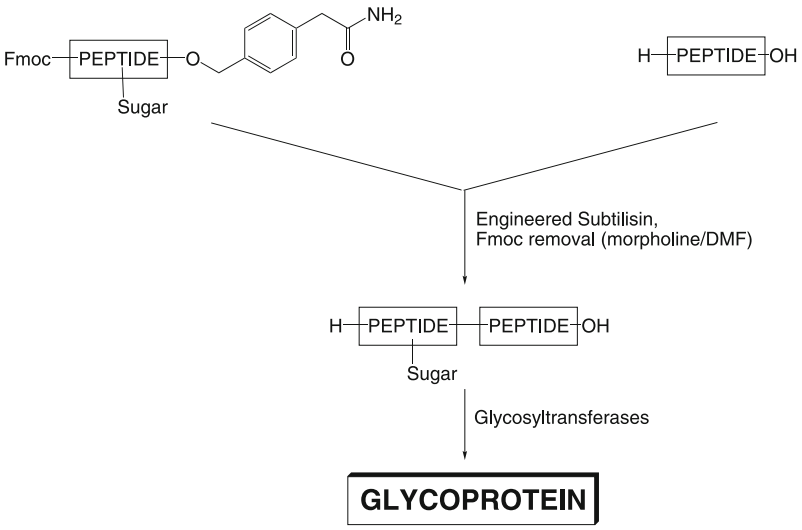
Production of an *N*-terminal coupling partner requires that the protein be expressed as a thioester. This can be achieved by taking advantage of naturally occurring self-splicing elements called inteins (*Scheme 9b*) [60,61]. Inteins catalyze their excision from a protein through a series of acyl-transfer reactions in which a cysteine thioester is a key intermediate. This intermediate can be isolated, and upon exposure to an appropriate ligation partner will undergo NCL to afford a new protein. It is important to note that the two approaches to EPL are orthogonal, a property which Macmillan and Bertozzi took advantage of in a recent synthesis of the glycoprotein GlyCAM-1 [62].


HS
H2N
HIS6
ENLYFQ
N
H
O
TEV
Protease
HS
HN
O
Asn(GlcNAc)-Gly-Glythioester
Thioester
exchange
OH
HO
HO
O
NH
AcHN
O
H3N
N
S
H2N
S→N
aceyl transfer
a
HS
O
N
H
Intein
COO-
N→S
aceyl transfer
S
NH2
HS
SO3-
H-Cys-Asn(GlcNAc)-OH
b

Scheme 9

2.4 Enzymatic Ligation and Glycosyltransferase Reactions

Enzymatic synthesis has become an invaluable tool in the assembly of glycopeptides. Proteases such as subtilisin can be utilized as catalysts for *O*-linked glycopeptide couplings (*Scheme 10*) [63]. Similarly, thiosubtilisin [64] has been utilized as a catalyst in the synthesis of both *O*-linked and *N*-linked glycopeptides. Protease-catalyzed peptide bond formation has the advantage of proceeding at neutral conditions that are compatible with acid- and base-labile glycosidic linkages.



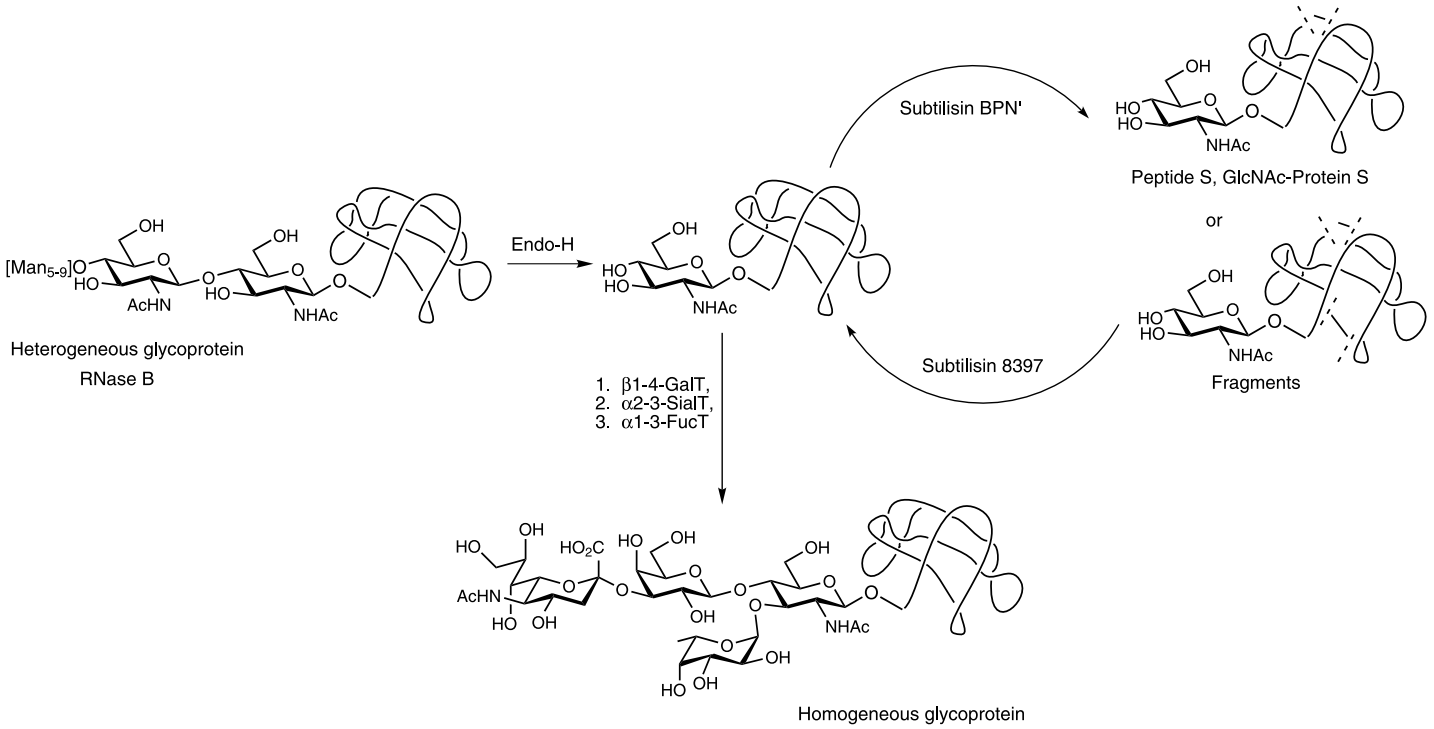
Scheme 10

Glycosyltransferases are also essential tools for glycopeptide synthesis [65,66]. Their use allows formation of specific linkages in a carbohydrate chain, and the enzymes often tolerate an array of functionality at the carbohydrate reducing terminus. Glycosyltransferase activity derived from whole cells has also been utilized in glycopeptide synthesis [67].

Glycosylhydrolases have also found utility in oligosaccharide synthesis. While these enzymes have the advantage of being less expensive and easier to handle than glycosyltransferases, they suffer from the drawbacks of poor regioselectivities and competitive hydrolysis of the desired product. Recently, Withers and coworkers reported a solution to this problem through the use of mutant glycosylhydrolases termed glycosynthases, which use glycosyl fluorides as donors [68]. To date, numerous glycosynthases have been created for a wide variety of sugars [69]. In addition to peptide and sugar synthesis, enzymes are also useful reagents for the removal of protecting groups [70,71] or cleavage of glycopeptides from solid supports under neutral conditions [72].

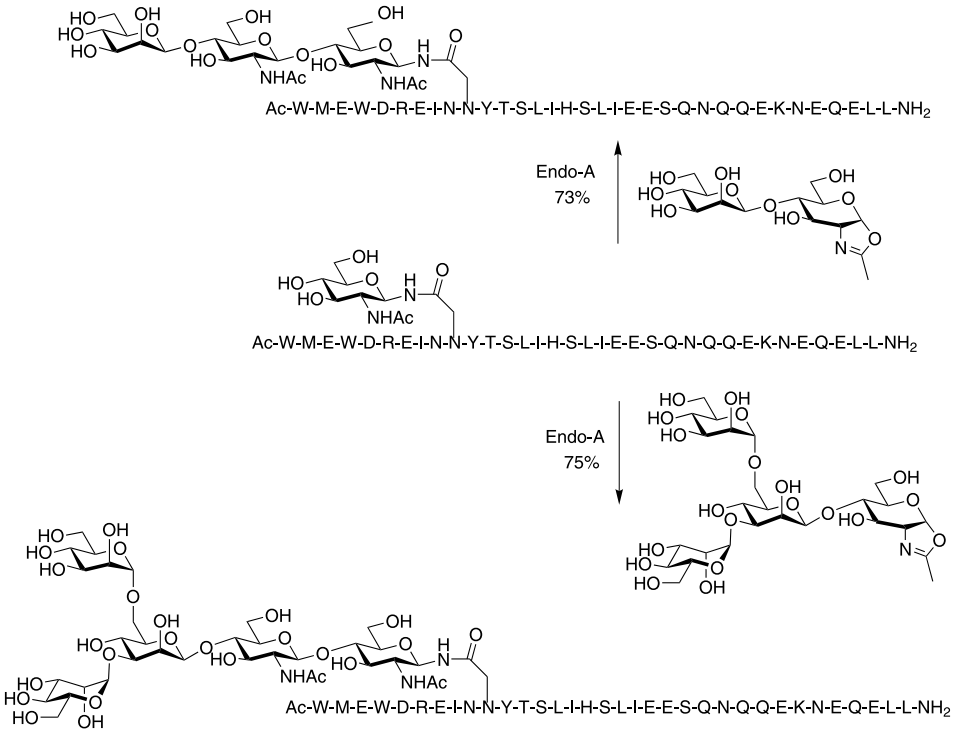
2.4.1 Endo-Glycosidase Activity in Synthesis

A special class of glycosylhydrolases, the endo-β-*N*-acetylglucosaminidases (Endo), have been utilized for the synthesis of homogeneous glycopeptides and glycoproteins. These


[Man5-9]O
HO
AcHN
OH
O
HO
NHAc
Heterogeneous glycoprotein
RNase B
Endo-H
HO
HO
NHAc
Subtilisin BPN'
Peptide S, GlcNAc-Protein S
or
Subtilisin 8397
Fragments
1. β1-4-GalT,
2. α2-3-SialT,
3. α1-3-FucT
HO2C
Homogeneous glycoprotein

Scheme 11

enzymes function by hydrolyzing the chitobiose core of *N*-linked glycans. The resulting product is a protein where each *N*-linked glycosylation site bears a single GlcNAc residue which can be used as a handle for enzymatic elaboration (Scheme 11) [12].
In addition, two of these enzymes, Endo A and Endo M, have been shown to catalyze the transfer of complex sugars to GlcNAc residues on glycopeptides and glycoproteins via transglycosylation [11,13]. While the yield of the reaction is low, it has been found that it can be increased through the addition of organic solvents to the reaction media [73]. More recently, Wang and coworkers have demonstrated that the enzymes catalyze the transfer of oligosaccharide oxazolines, putative transition-state mimetics for hydrolysis, in high yield in aqueous media (Scheme 12) [74]. Using this technology they have been able to synthesize numerous potential glycopeptide-based HIV vaccines (see Sect. 3.2).



Scheme 12

The above methods for glycopeptide synthesis have started to gain tremendous attention for applications in the preparation of various biologically relevant glycopeptides and glycoproteins. The role of these techniques in the areas of vaccine development, antibiotics, and novel therapeutics is discussed in the sections that follow.

3 Glycopeptide-Based Vaccines

Vaccination is the procedure whereby the immune system is induced to create antibodies against a foreign molecule involved in a specific disease process or viral infection. The strategy relies on the fact that if the body is "challenged" with the molecule again at a later date, the immune response to eradicate it will be enhanced and rapid. In effect, a vaccine educates the immune system to foreign entities that may be harmful to the body. In many disease states, the oligosaccharide chains presented on cell surface glycoproteins are altered [75], and therefore represent targets for therapeutic intervention. Since glycoproteins play major roles in cell-cell recognition, it can be expected that any alteration in carbohydrate structure will interfere with communication between cells [76]. This is thought to be the underlying cause of tumor cell metastasis and invasiveness into surrounding tissues. In addition, many pathogens have been shown to possess unique glycosylation patterns on their surface. The development of vaccines for certain cancers, HIV and other pathogens using glycopeptide antigens has been the subject of extensive research over the past decade.

3.1 Synthesis of Glycopeptide Cancer Antigens

Traditional cancer therapies based on surgery, irradiation and chemotherapy are accompanied by significant side effects. The first of these involves invasive removal of the tumor, while the latter two destroy diseased cells by capitalizing on differences between cancerous and healthy cells. It has been recognized for decades, that a phenotype of malignant cells is characterized by aberrantly glycosylated proteins [77,78], often arising from down regulation of certain cell-bound glycosyltransferases that normally function to elongate the saccharide chains. In the case of T_N and T antigens [79], the polysaccharide chains have been abbreviated to mono- and disaccharide structures, respectively. These truncated glycans are often observed on mucin-type proteins, thus affecting the conformation of the protein and allowing the tandem repeat region to become accessible to the immune system.

In addition to simplified glycan core structures, alterations in terminal glycan structures are also associated with malignancy. Certain glycosyltransferases, for example fucosyltransferases and sialyltransferases, tend to be over-expressed in tumor tissue, leading to the over-production of these terminal glycans. Examples of these glycan epitopes, commonly found on certain cancerous cells, include sialyl Lewis x (sLe^x), sialyl Tn (sTn), Globo H, Lewis y (Le^y) and polysialic acid (PSA) (*Fig. 2*) [80,81,82].

3.1.1 Carbohydrate-Based Vaccines

As the structures of cancer cell glycans differ from those that appear on their healthy counterparts, many research groups have focused on recruiting the immune system to target cancer cells on the basis of these altered glycosylation patterns. To date the synthesis of many tumor-associated glycans have been accomplished. Potential anti-cancer vaccines are then created by linking multiple copies of the glycan to immunogenic carrier proteins such as bovine serum albumin (BSA) and keyhole limpet haemocyanin (KLH) [83].

Several carbohydrate-based vaccines are currently undergoing clinical evaluation [83,84], however, vaccines based on immunization with oligosaccharides alone have been met with

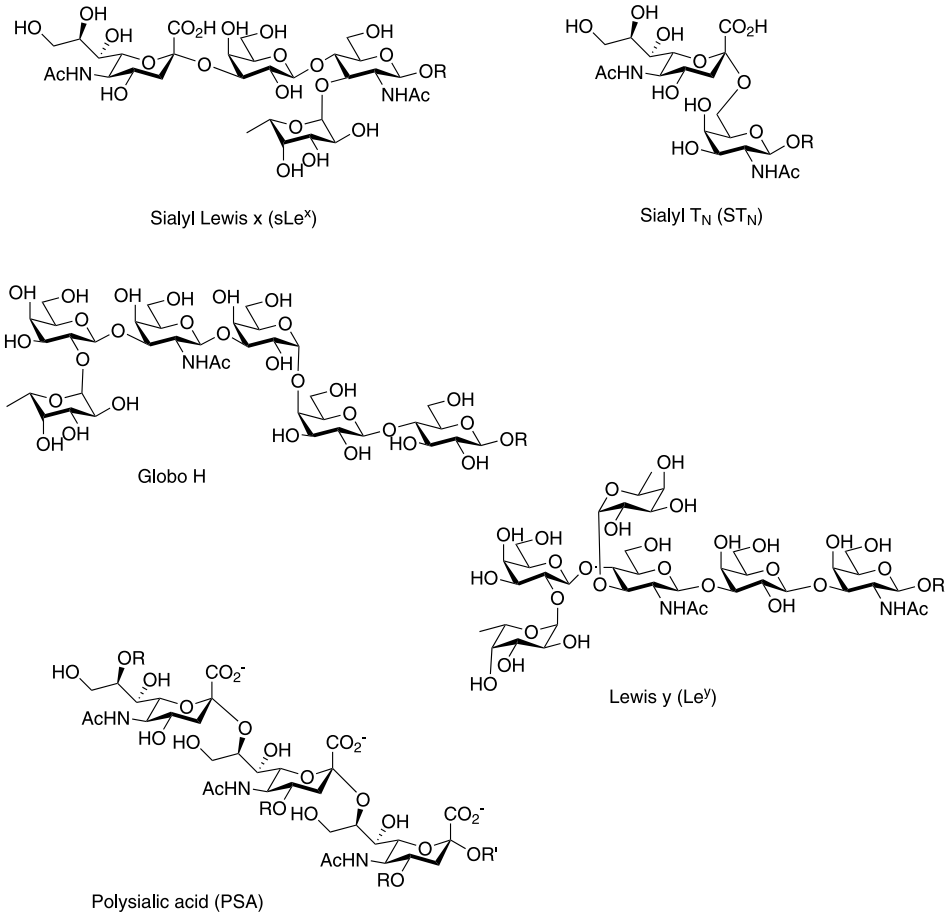


Figure 2
Glycans associated with cancer

limited success in most cases. Exceptions to this are the sialyl-T_N vaccine [85,86] developed by Biomira and the globo H vaccine developed by Danishefsky and coworkers, both of which are currently in clinical trials [87]. Experimental results suggest that the peptide chain may be as important as the carbohydrate portion of the antigen in terms of antibody recognition. The peptide may provide a scaffold for correct spatial orientation of the carbohydrate, perhaps by constraining it in a clustered array. Alternatively, the peptide may provide additional contacts for binding with the antibody.

3.1.2 Glycopeptide-Based Vaccines

Another common feature of tumors is the over-production of certain glycoproteins and glycolipids. In particular, epithelial tumors often overproduce mucin glycoproteins, which are

characterized by dense clustering of *O*-linked glycans on the protein surface. Mucins represent important diagnostic markers for cancer, and function as scaffolds for most of the cancer-associated glycan epitopes listed previously [88,89].

3.1.3 The T_N Antigen

The T_N antigen is a GalNAc α-linkage to serine or threonine on a mucin-like core peptide and is known to accumulate on the surface of tumor cells as a result of incomplete glycan synthesis (*Fig. 3*) [90]. Therefore, several groups have focused on, and reported syntheses of glycopeptides incorporating the T_N antigen. Most of these syntheses target multivalent α-T_N presentation, in order to mimic the proposed clustered array of sugars in the serine/threonine rich mucin domain.

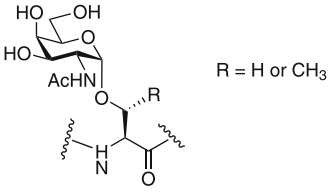


Figure 3
The T_N antigen

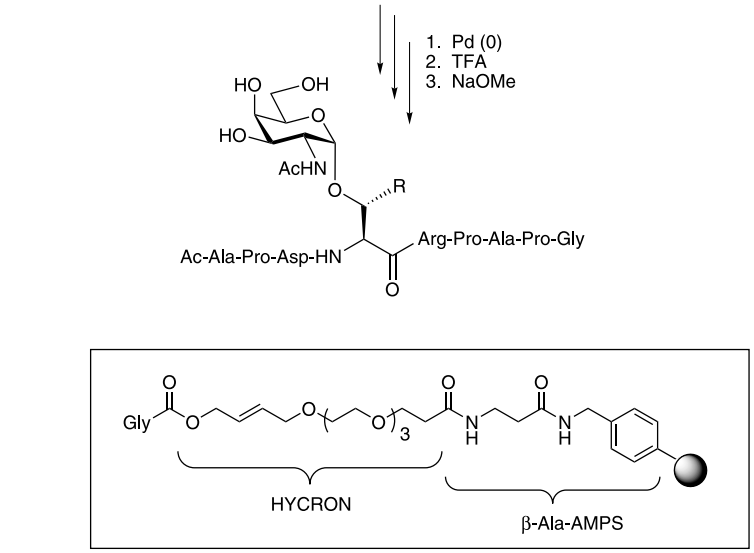
The Kunz laboratory has reported several formulations of the T_N antigen. In one study, tripeptides containing two sites of α-*O*-GalNAc glycosylation were synthesized by chemical methods, then conjugated to BSA for generation of antibodies [91]. In later work on the same system, lipase M was utilized for deprotection of the glycopeptide *C*-terminal carboxylate under neutral conditions [70]. Other instrumental work by this group has focused on the generation of novel solid support linkers, such as HYCRON. With this resin, cleavage of the glycopeptide from the solid-phase occurs under conditions that are not destructive to sensitive glycopeptide functionalities. Employing this methodology, several mucin glycopeptides have been synthesized. *Scheme 13* depicts a reported synthetic sequence of the MUC1 mucin displaying the T_N antigen [92]. Other glycopeptide sequences synthesized by HYCRON methodology include the homophilic recognition domain of Epithelial Cadherin-1 [93] and a partial sequence of MAdCAM-1 [94].

Solid-phase synthesis of another highly T_N glycosylated sequence of MUC1 was reported by Dupradeau et al. [95]. After conjugation to KLH, reactivities of the glycopeptide and unglycosylated control peptide against specific antibodies were assessed. In this case, the authors did not observe a difference between the reactivity of the glycosylated versus the unglycosylated peptides.

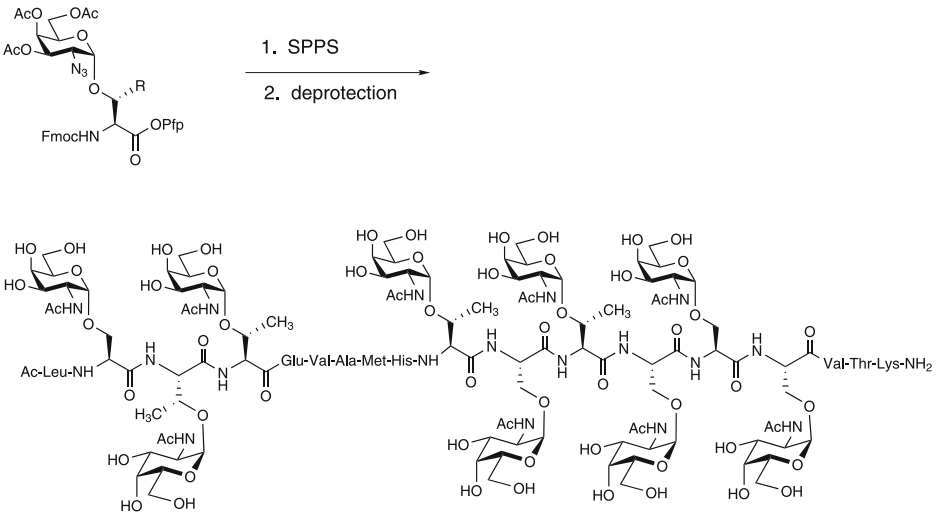
Klich et al. have also accomplished the synthesis of extensively glycosylated sequences of glycophorin A containing the T_N antigen [96] (*Scheme 14*).

Danishefsky and coworkers have prepared tripeptide T_N structures which incorporate handles for conjugation to immunogenic carriers [97] such as KLH, BSA or the lipopeptide Pam_3Cys, which was included as an immunostimulatory feature known to activate B cells and macrophages (*Scheme 15*) [98]. The synthetic strategy centered around the coupling of glycosyl amino acid building blocks in a sequential fashion.

Ac-Ala-Pro-Asp(OBu^t)-Thr(αAc$_3$GalNAc)-Arg(Mtr)-Pro-Ala-Pro-Gly-HYCRON-βAla-AMPS

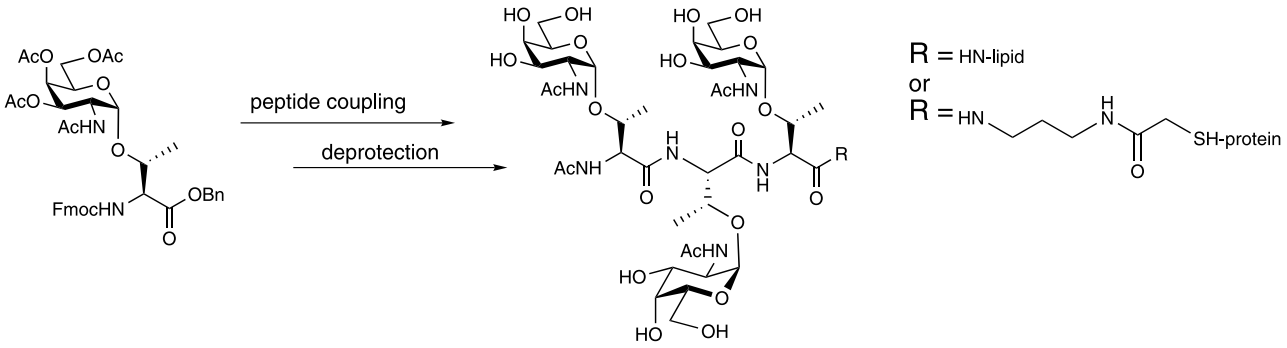


Scheme 13



Scheme 14

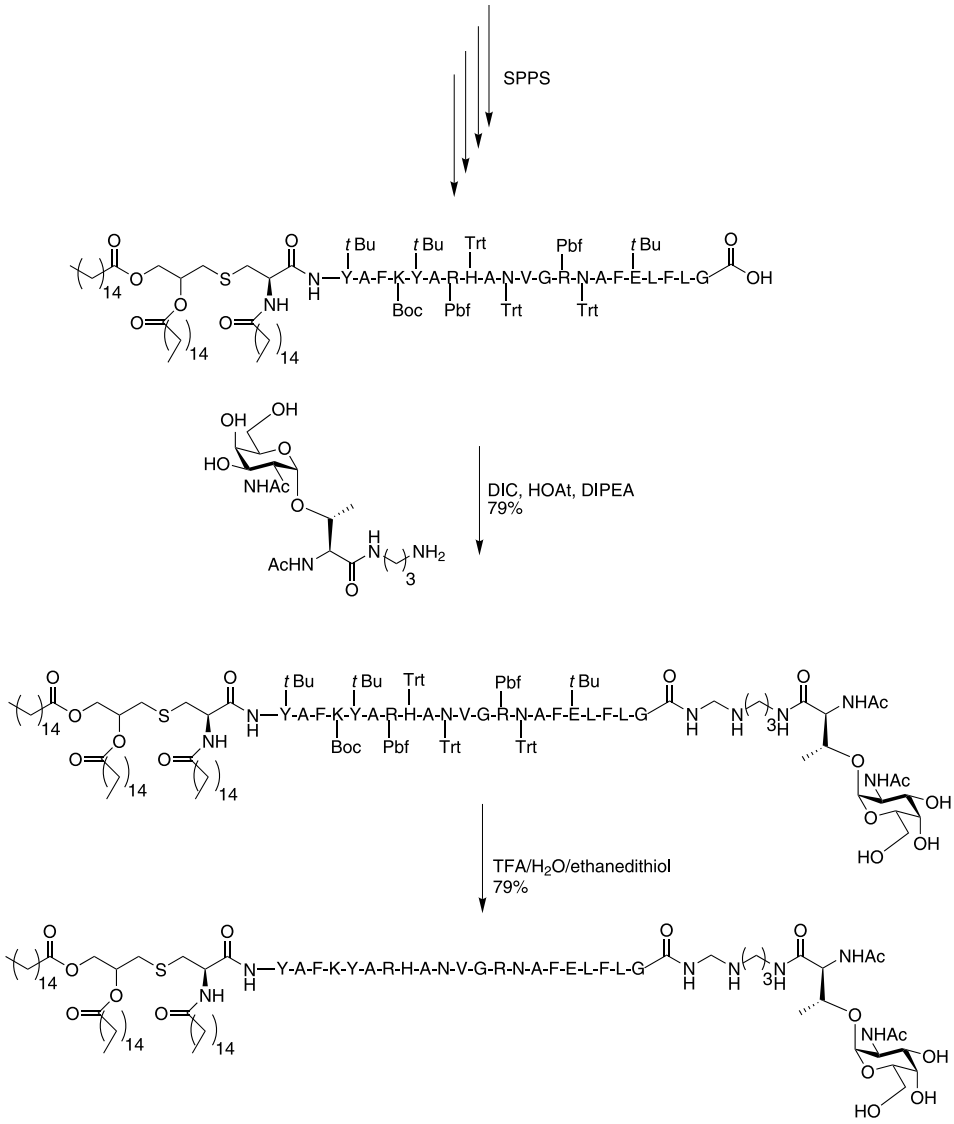
In this case, all conjugates were subject to immunological studies in mice. The KLH conjugate induced high antibody titers, while the lipopeptide conjugate failed to show a strong reaction [83]. This result emphasizes that more than one factor needs to be considered when attempting to generate an immune response. Furthermore, the method of hapten conjugation may be as important as the antigenic structure itself. Interestingly, the antibody response


AcO
OAc
AcO
AcHN
O
FmocHN
OBn
O
peptide coupling
deprotection
HO
OH
HO
O
AcHN
AcNH
H
N
O
N
H
R
AcHN
HO
O
HO
OH
R = HN-lipid
or
R = HN
H
N
SH-protein
O

Scheme 15

observed was polyclonal, whereby the antibodies generated recognized several different clustered antigen presentations.

Boons and coworkers have also reported the synthesis of a lipated T_N containing glycopeptide as a potential cancer antigen [99]. In this study the T_N antigen was linked to the 20-amino acid peptide YAF, a MHC class II restricted site for human T cells. It was envisioned that the YAF epitope would induce a T cell dependant immune response leading to the production of IgG antibodies [100]. The target was constructed by coupling the T_N antigen to the



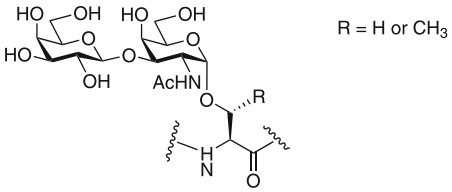
Scheme 16

lipopeptide (obtained through SPPS) through a polyamine linker (◉ *Scheme 16*). Following incorporation into liposomes, the antigen was able to successfully elicit IgM and IgG responses in mice.

To date, the T_N antigen has been studied to a greater level as a vaccine than the other antigenic glycopeptide cancer epitopes. The reason for this can be attributed to its simplicity and ease of synthesis. Recently, more complex antigens have been synthesized, using the approaches established on the T_N core structure as a foundation for which to build upon.

3.1.4 The T Antigen

The T antigenic structure (the "Thomsen-Friedenreich" or "TF" antigen) is derived from the precursor T_N antigen [90]. The disaccharide is also α-*O*-linked to serine or threonine in the peptide backbone (◉ *Fig. 4*).



◘ Figure 4
The T antigen

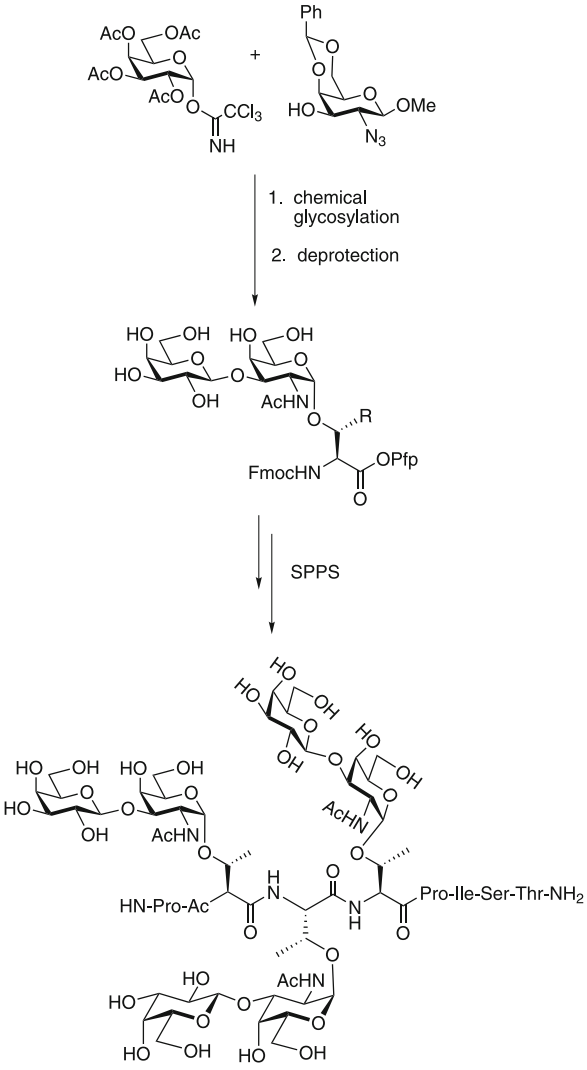
Syntheses of these structures have often been reported in conjunction with that of the T_N antigen [91,97,101]. Again, the clustered glycopeptide motif is often the form of interest. Notably, Paulsen et al. accomplished the construction of eight triply glycosylated mucin sequences, as well as the corresponding mono-glycosylated sequences in a single study [102] (◉ *Scheme 17*).

Kuduk et al. have also prepared clustered T antigens through glycal methodology that are appropriate for conjugation [97] (◉ *Scheme 18*). The approach centered on the synthesis of the Gal-β-1,3-GalNAc-α-Thr building block, which was subsequently coupled to a protected diamine handle for conjugation studies to appropriate carrier proteins (BSA or KLH). This cassette was then sequentially incorporated into a tripeptide T cluster. Immunological evaluation of these constructs is currently under investigation [83].

3.1.5 Sialyl-T_N and Sialyl-T Antigens

The next level of complexity, in the case of synthetic methods, is construction of the sialylated antigenic structures. The sialylated antigens contain sialic acid in either an α2-6 linkage to the GalNAc residue of the T_N or T antigens or in an α2-3 linkage to the terminal galactose of the T antigen [90] (◉ *Fig. 5*).

Nakahara et al. accomplished an early synthesis of a trimeric α2-6 sialyl-T_N glycopeptide. The resultant glycopeptide corresponds to a partial structure of glycophorin A [103] (◉ *Scheme 19*).

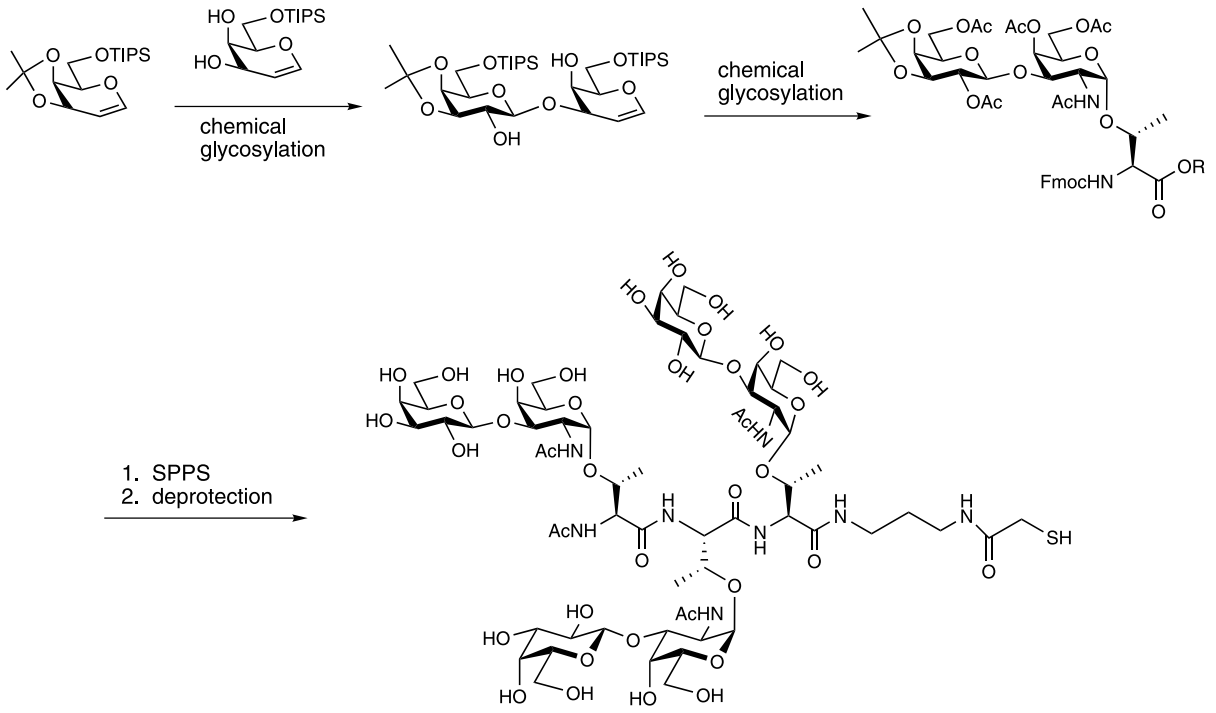


Scheme 17

Liebe and Kunz made further use of the HYCRON linker methodology in the construction of the α2-6 sialyl-T_N antigen attached to a portion of the mucin MUC1 repeat peptide [104] (*Scheme 20*).

Danishefsky's laboratory has recently reported syntheses of the clustered motif of both the α2-6-sialyl-T_N and α2-6 sialyl-T antigens (*Fig. 6*) using a similar procedure to that described for the clustered T antigen described above [105,106]. Conjugation and immunological evaluation of these constructs is currently in progress.

Only one of the aforementioned glycopeptide vaccine candidates has exhibited a proliferating effect on the peripheral blood lymphocytes, an essential property for a vaccine. This led to


chemical glycosylation
chemical glycosylation
1. SPPS
2. deprotection

Scheme 18

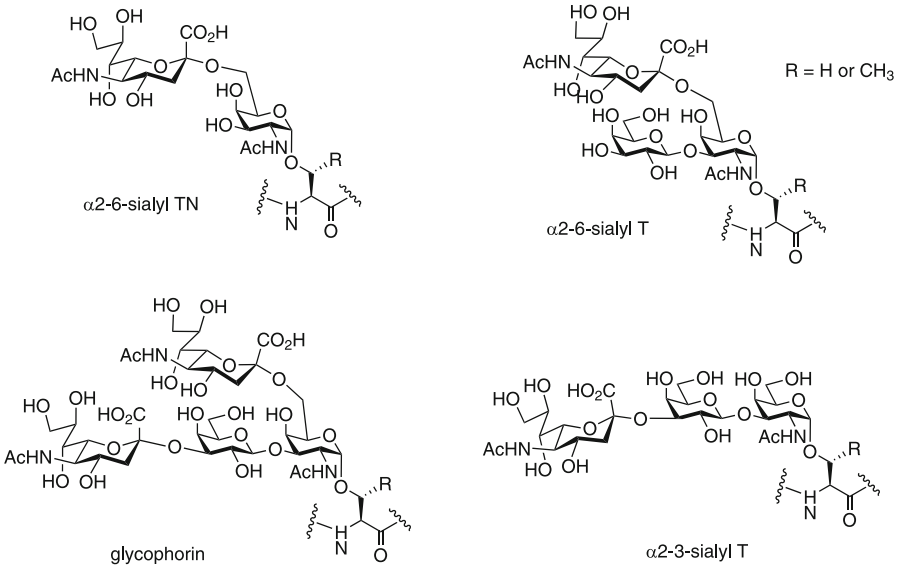
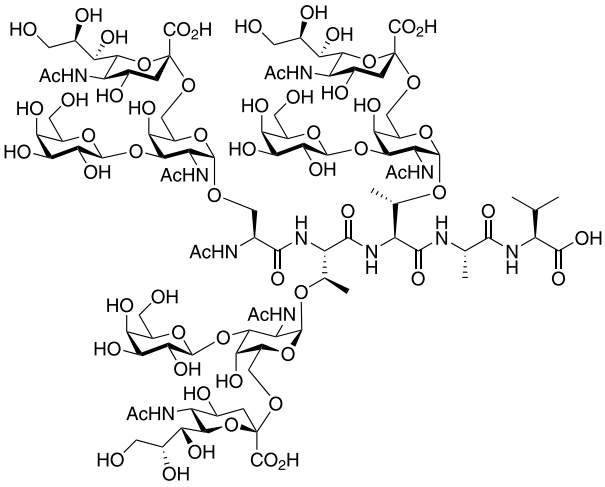


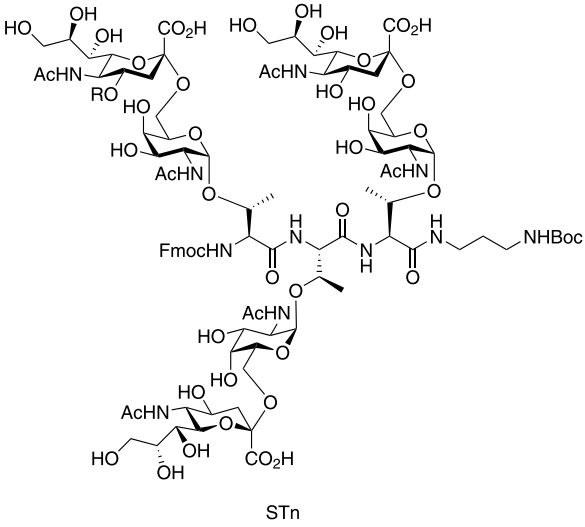
Figure 5
Sialic acid-containing cancer-associated antigens

the development of a novel concept for the incorporation of glycopeptides into antitumor vaccines, whereby the tumor associated MUC1 glycopeptide antigen was combined with a T-cell epitope of tetanus toxin as an immunostimulant [107]. Following the synthesis of the desired glycopeptide, using HYCRON linker strategy SPPS [92], the resin bound glycododecapeptide was coupled to the tetanus toxin peptide prior to cleavage from the resin (*Scheme 21*). The synthetic vaccine was able to induce proliferation of human blood lymphocytes after three stimulations with the conjugate. Importantly, proliferation was observed only in the presence of antigen presenting cells, indicating an antigen specific response. This conjugate was found to induce the proliferation of $CD3^+$ cells, including cytotoxic $CD8^+$ cells. It is important to note that it was necessary to incorporate all of the elements on the construct into the antigen to proliferate $CD8^+$ cells.

In addition to the production of $CD8^+$ cells, a sialyl-T_N containing glycopeptide has been used to elicit a T-helper cell mediated immune response. This approach involved coupling the sialyl-T_N glycododecapeptide to a T_H cell peptide epitope (*Scheme 22*) [108]. The concept was to utilize the T_H cell peptide epitope, residues 323–339 from ovalbumin (OVA), to initiate strong production of antibodies specific for the MUC1 epitope also present in the vaccine. It was envisaged that binding of the MUC1 epitope to surface Ig receptors of a B lymphocyte would result in endocytosis of the entire construct. In this approach the glycopeptide antigen was synthesized by SPPS using a 2-phenyl-2-trimethylsilyl-ethyl ester (PTMSEL) linker, which allows facile cleavage of the peptide from the resin under mild conditions [109]. The glycopeptide was then conjugated to the OVA sequence through a triethylene glycol linker. This conjugate was found to elicit an immune response in transgenic mice whose T cells expressed a receptor for the ovalbumin T cell epitope presented by MHC(III) molecules.



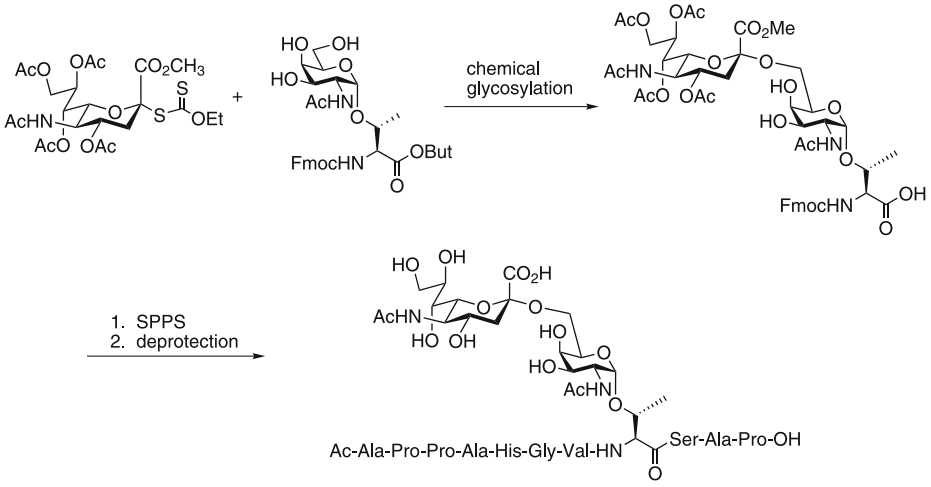
2,6-ST



STn

Scheme 19

Following three rounds of immunization one of the mice was found to exhibit significantly increased levels of anti-T_N antibodies. Importantly, significant increases in antibody concentration were observed between the first and second rounds of immunization, implying the establishment of an immunological memory (specific IgG response).



Scheme 20

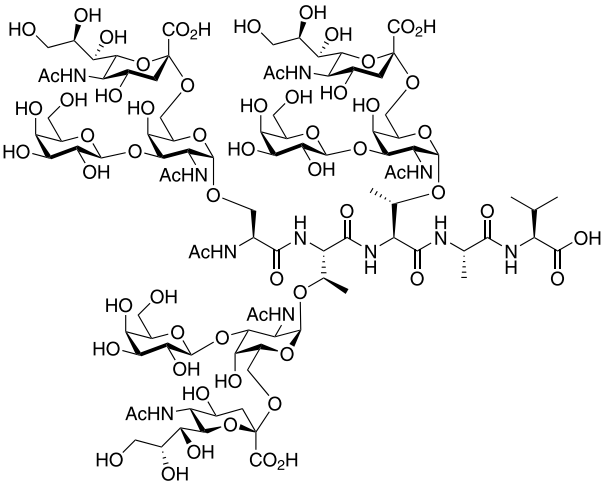
3.1.6 Lewis^y-Glycopeptides

With the successful synthetic studies on clustered T_N, T, sialyl-T_N and sialyl-T motifs in hand, Danishefsky and coworkers embarked on the synthesis of clustered glycopeptides containing the more complex Lewis^y (Le^y) determinant (*Scheme 23*) [110]. The synthesis was achieved using the cassette strategy in which a complex glycan donor was coupled to α-GalNAc-Ser to afford the protected Le^y hexasaccharide derivative. The glycosyl amino acid was subsequently submitted to iterative solution phase peptide couplings as executed in previous cluster syntheses to provide the trimeric Le^y cluster. Global deprotection, followed by coupling of the immunostimulatory Pam_3Cys moiety to the *C*-terminus of the peptide gave the desired antigenic cluster. The reactivities of the Pam_3Cys glycopeptide along with a Le^y-ceramide control compound against the anti-Le^y antibody 3S193 (elicited against tumor cells displaying the cell surface mucin motif) were determined using an ELISA assay. The compound displayed comparable reactivity to the control suggesting that the hexasaccharide was a reliable mimic of the tumor-associated Le^y mucin. Vaccinations of mice with the antigen resulted in a powerful IgM immune response, producing antibodies that bound strongly to the natural Le^y-mucin moiety.

3.1.7 Other Glycopeptide Cancer Antigens

Von dem Bruch and Kunz have prepared *N*-linked divalent Lewis x glycopeptides as potential tumor antigens (*Scheme 24*). Following chemical block synthesis, the glycohexapeptides were conjugated to BSA or KLH for immunological study [111].

Chen et al. have also reported the synthesis of a related cancer antigen, the F1α antigen [112]. Although the core peptidic-glycan linkage remains conserved, the pendant sugar is lactosamine, rather than sialic acid (*Fig. 7*).



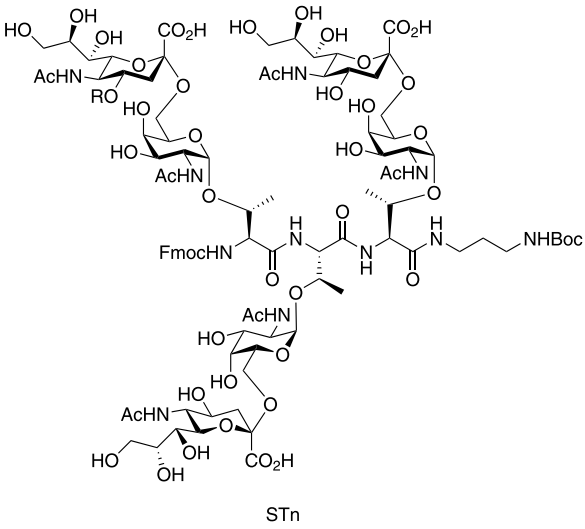
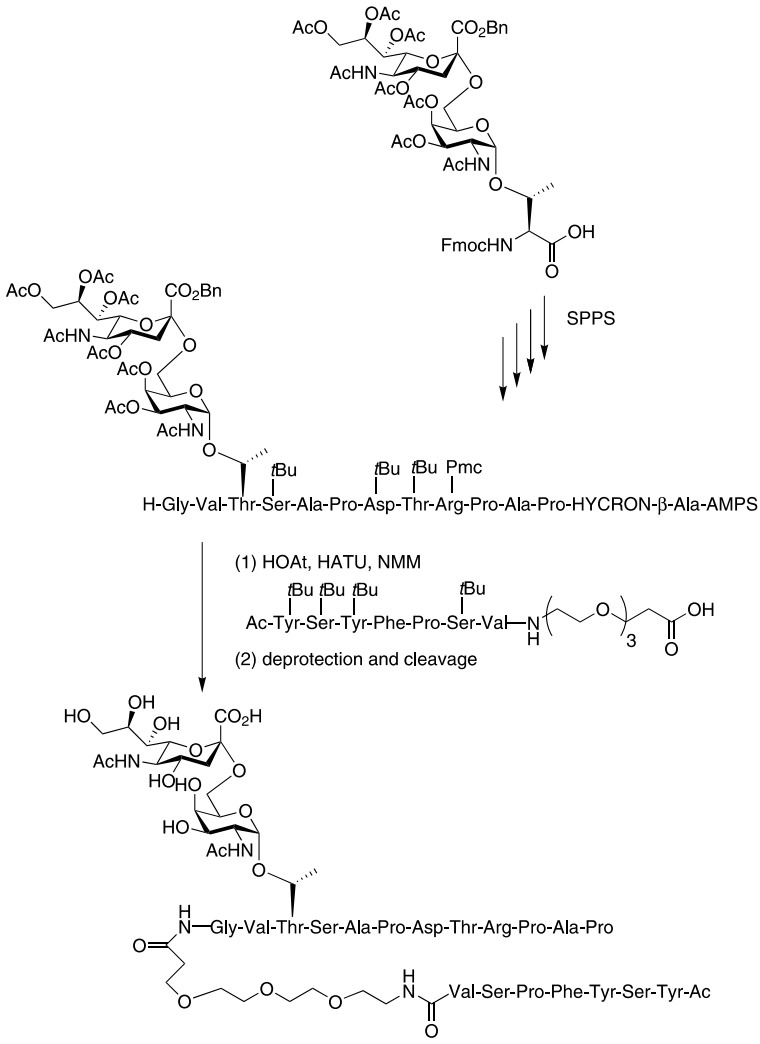


Figure 6
Clustered glycopeptide antigens as cancer vaccine candidates.

More recently Danishefsky and coworkers have evaluated multiantigenic glycopeptides as cancer vaccine candidates [84]. It was reasoned that vaccination strategies utilizing a range of tumor-associated antigens attached to one mucin-type peptide may trigger a more diverse immune response than those presenting only one antigenic glycan. An example of such a mucin-based construct displays T, Le^y, and Tn antigens (*Fig. 8*). Immunogenicity of the multivalent species (conjugated to KLH) was investigated and was shown to induce IgM

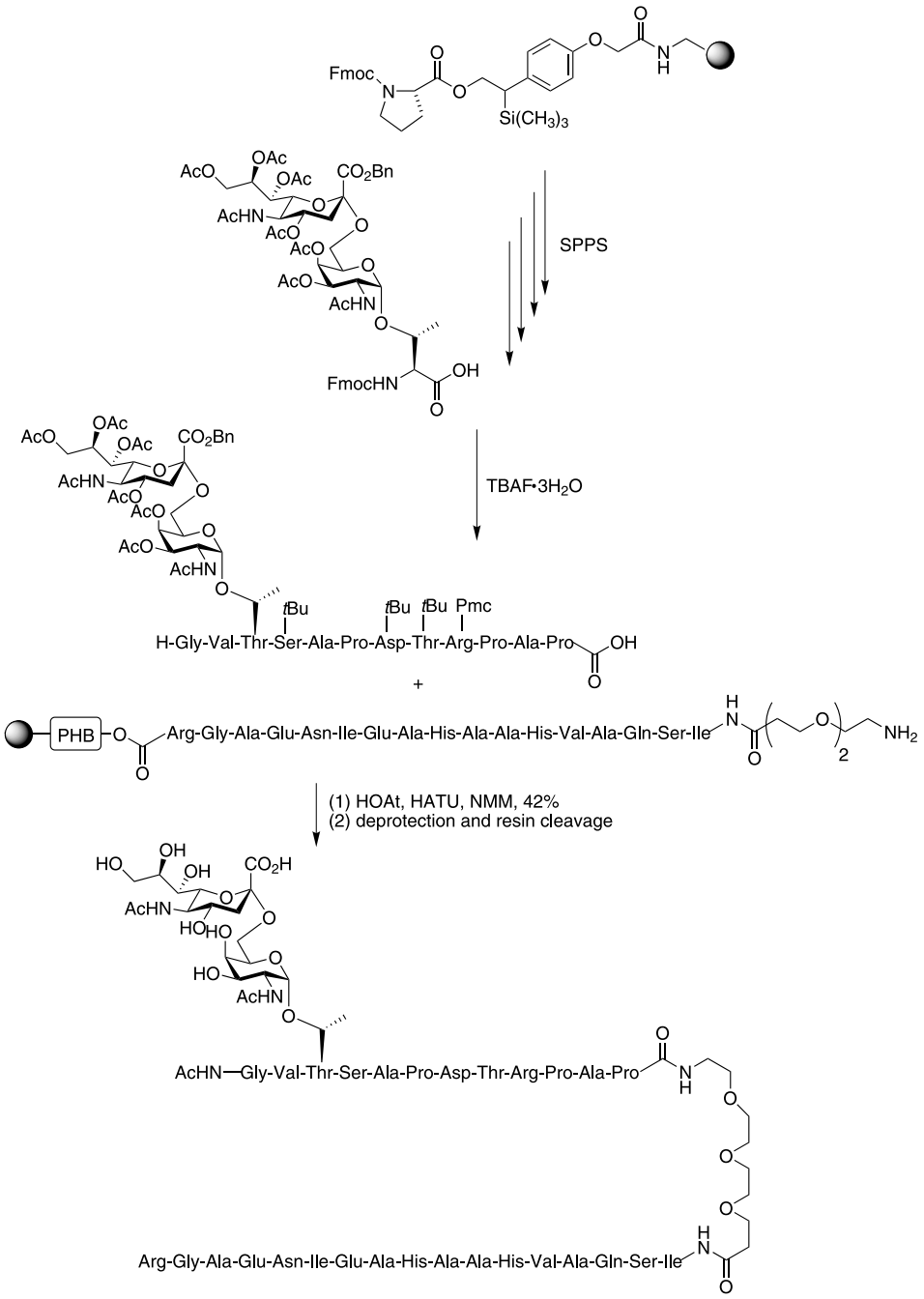


Scheme 21

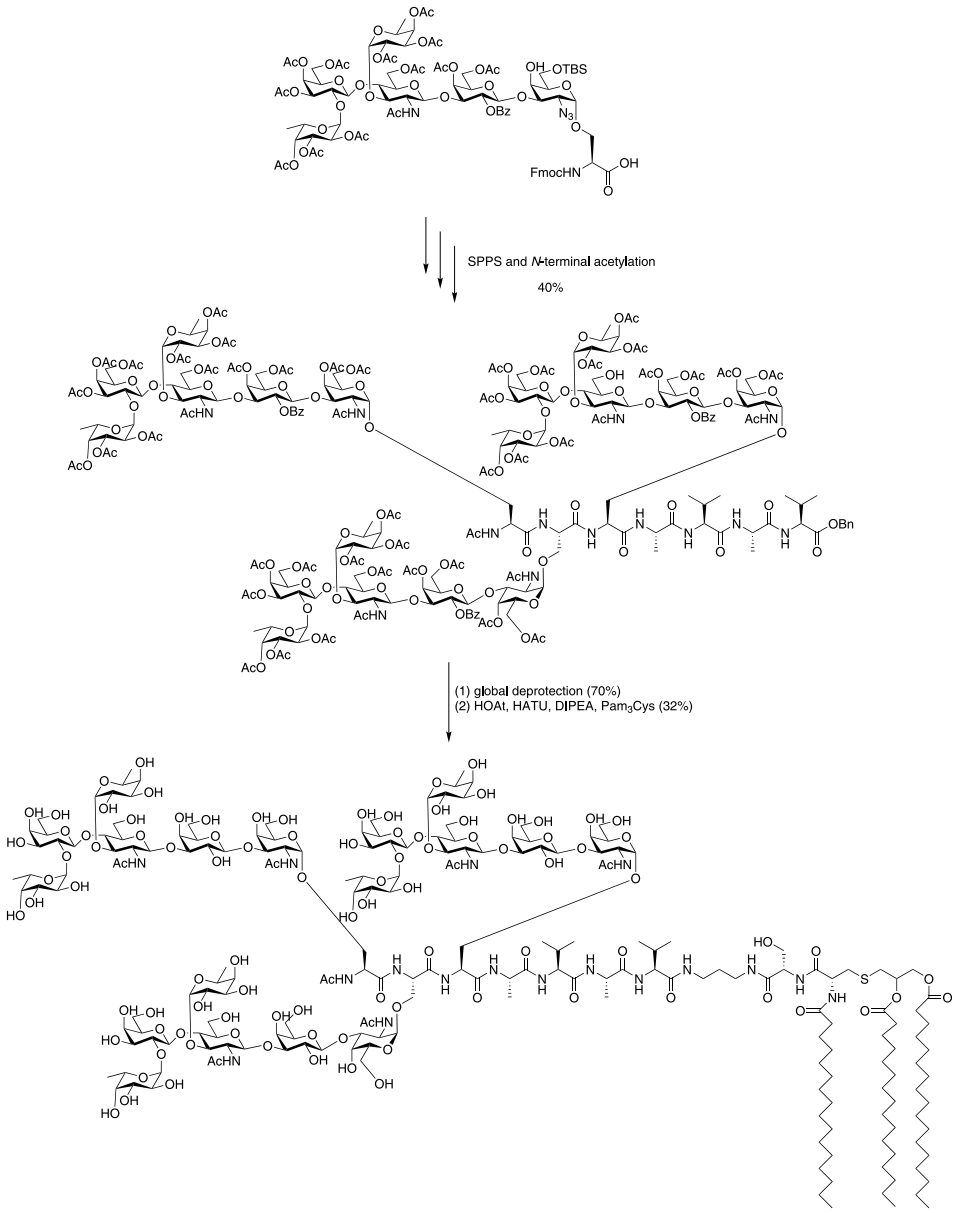
and IgG antibodies when injected into mice. Interestingly, the glycopeptide containing the T antigen displayed significantly less antigenic behavior than the trishomoserine construct containing the Globo-H epitope.

3.2 Synthesis of Glycopeptide HIV Antigens

Acquired immune deficiency syndrome (AIDS) is caused by infection with the human immunodeficiency virus (HIV) [113]. At present 40 million people are infected with the virus worldwide, with an estimated five million new infections in 2004 alone [114]. A vaccine that limits HIV-associated disease states or prevents HIV infection would be a valuable tool for stemming


Fmoc
Si(CH3)3
SPPS
FmocHN
TBAF•3H2O
H-Gly-Val-Thr-Ser-Ala-Pro-Asp-Thr-Arg-Pro-Ala-Pro
PHB
Arg-Gly-Ala-Glu-Asn-Ile-Glu-Ala-His-Ala-Ala-His-Val-Ala-Gln-Ser-Ile
(1) HOAt, HATU, NMM, 42%
(2) deprotection and resin cleavage
AcHN—Gly-Val-Thr-Ser-Ala-Pro-Asp-Thr-Arg-Pro-Ala-Pro
Arg-Gly-Ala-Glu-Asn-Ile-Glu-Ala-His-Ala-Ala-His-Val-Ala-Gln-Ser-Ile

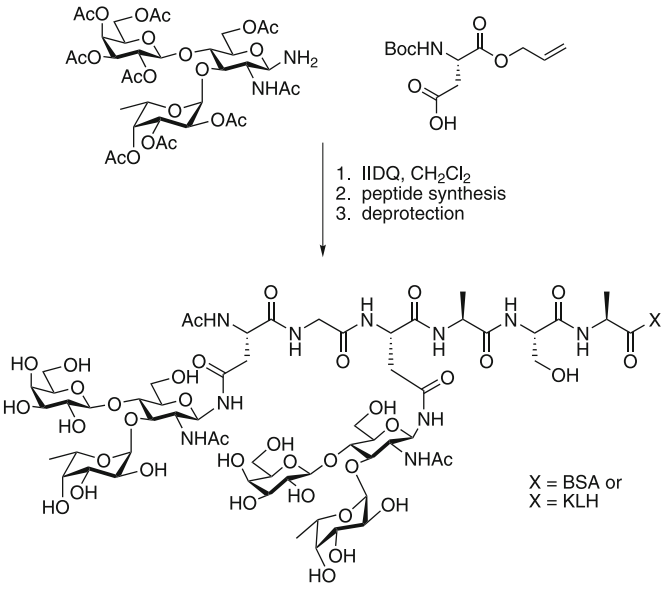
Scheme 22



Scheme 23

the HIV epidemic [25]. Currently, more than 30 HIV vaccine candidates are in various stages of clinical trials through-out the world, however, the problem is extremely complex and most have been met with limited success to date [115,116].

One of the major impediments of vaccine development is the extensive glycosylation of HIV envelope proteins, which shields them from antibody recognition. HIV-1 contains two enve-



Scheme 24

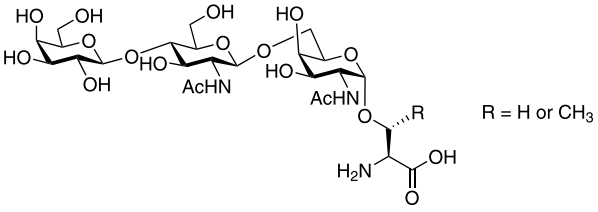


Figure 7
Synthetic F1α antigen

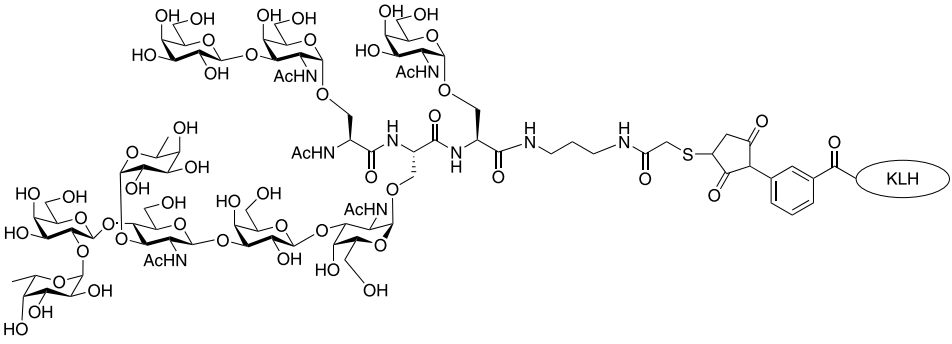
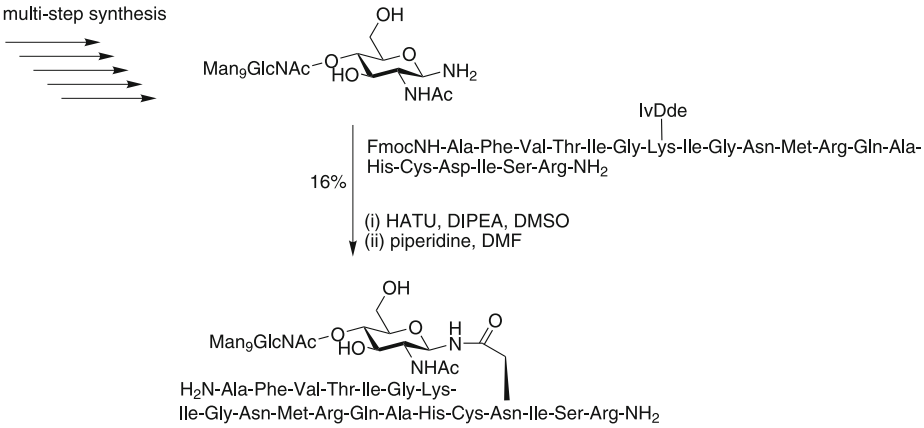


Figure 8
Glycopeptide conjugate displaying multiple carbohydrate antigens

lope glycoproteins, gp120 and gp41 that form trimeric complexes on the viral surface. The outer envelope glycoprotein gp120 typically contains 24 *N*-linked glycans [117,118,119], while the inner envelope glycoprotein gp41 contains four conserved *N*-glycans [120,121]. Two distinct oligosaccharide cluster domains are observed on the surface of gp120. One consists of high mannose-type *N*-linked glycans, while the other consists of complex type *N*-linked glycans [119]. Recently, the concept of using gp120 carbohydrates as antigens for eliciting broadly neutralizing immune responses gained recognition when the structure of the HIV neutralizing antibody 2G12 was solved [122]. A number of detailed studies by several research groups have shown that the antibody recognizes key features in the oligomannose epitope, in particular the mannose-α1,2-mannose sugar residues [122,123]. Importantly, studies with synthetic oligosaccharides showed that Man_9 glycans were necessary for optimal binding [124,125,126]. Recent X-ray crystal studies with these synthetic oligomannoses verified that the mannose-α1,2-mannose subunit is critical for recognition along with a sufficient oligomannose arm [127].

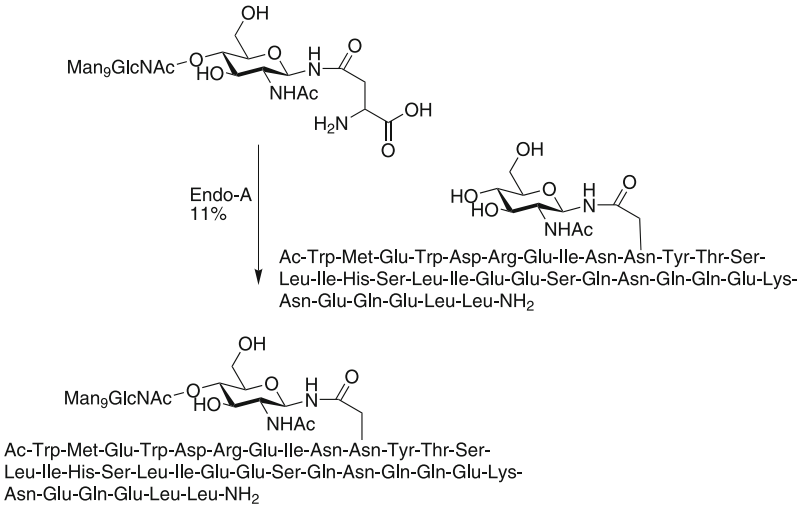
The wealth of information gained from the 2G12 binding studies has been applied to the development of novel synthetic glycopeptide vaccine candidates. Danishefsky and coworkers reported the chemical synthesis of HIV-1 gp120 glycopeptide fragments (amino acid residues 316–335), displaying either a hybrid-type or a high mannose-type *N*-linked glycans [128,129]. The preparation of the glycopeptides involved a multi-step synthesis of the fully elaborated 1-amino-oligosaccharides [130], followed by ligation to a selectively protected peptide displaying a free aspartic acid (corresponding to Asn-332 in the natural sequence). This was achieved using the block condensation approach described in ❯ *Sect. 2.2* (❯ *Scheme 25*).

Wang and co-workers have adopted a chemoenzymatic approach, using Endo A, to construct potentially immunogenic HIV-1 glycopeptides. This strategy involves enzymatic transglycosylation of a complex *N*-linked glycan to synthetic glycopeptides containing a single GlcNAc moiety to yield the desired glycopeptide. Several glycoheptapeptides (corresponding to amino acids 336 to 342) of HIV-1 gp120 possessing Man_9, Man_6 and Man_5 glycans were synthesized using this methodology [131].



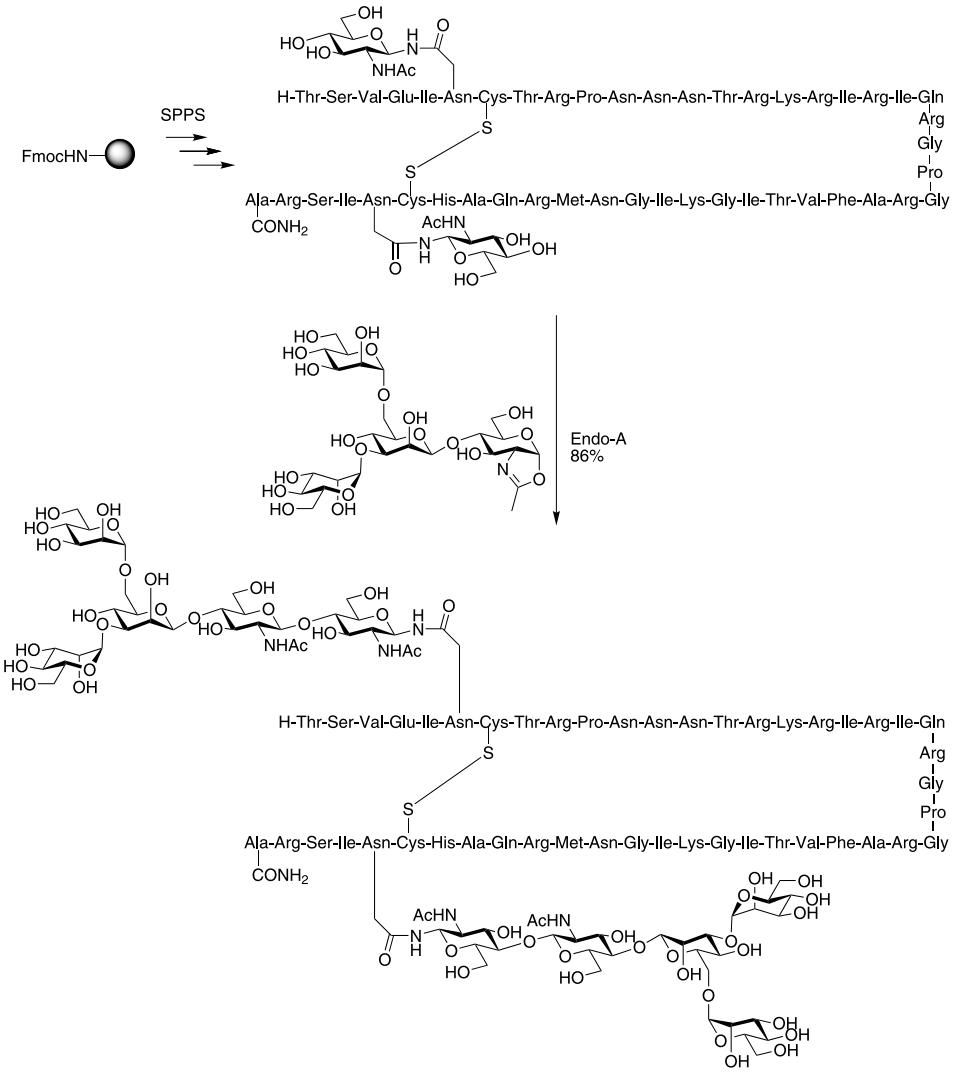
Scheme 25

This group also utilized endoglycosidases to synthesize homogeneous glycoforms of the peptide C34, derived from the *C*-terminal ectodomain of HIV-1 gp41 [132]. This sequence contains a conserved *N*-linked glycan at Asn-637 in the native gp41, and is responsible for the formation of α-helical bundles with the gp41 fragment N36, providing the driving force for the fusion of virus-host membranes [133]. Endo A was used to transfer $Man_9GlcNAc$ to a synthetic C34 peptide possessing a GlcNAc residue at Asn-637, providing the desired conjugate in 11% yield (▶ *Scheme 26*). The resulting $Man_9GlcNAc_2$-C34 conjugate was assessed for its inhibitory activities against HIV-1 using a cell-fusion assay. The glycopeptide exhibited an IC_{50} of 7.7 nM, less potent than the unglycosylated C34 peptide ($IC_{50} = 1.1$ nM). However, with respect to the development of an anti-HIV therapeutic, $Man_9GlcNAc_2$-C34 glycopeptide may be superior to C34 due to its increased solubility and resistance to proteolytic degradation. Conformational studies of the glycosylated C34 using circular dichroism (CD) analysis showed the C34 glycoprotein to have a similar solution conformation to the unglycosylated peptide. Glycosylation did, however, disrupt the ability of C34 to form six-helix bundles with the peptide N36. This led to the proposal that the less-compact six α-helix bundle may mimic some transition-state character of gp41 exposed during membrane fusion, providing novel conformational epitopes for the development of HIV vaccines.



Scheme 26

More recently, Wang and coworkers have exploited sugar oxazolines in the pursuit of HIV active glycopeptides [74]. Two C34 fragments containing core tri- and pentasaccharides were synthesized using sugar oxazoline donors in the Endo A catalyzed reactions. The new methodology significantly improved the yield of the target protein (73–75%). This method has also been applied a to 47-mer cyclic glycopeptide corresponding to the third variable (V3) domain of gp120 [134]. The V3 domain represents an important target for the development of HIV-1 vaccines and as such was once named "the principal neutralization determinant" [135]. This region of the native V3 domain contains two high-mannose type *N*-linked glycans at two con-



Scheme 27

served Asn residues (Asn-295 and Asn-332) which are recognized by the antibody 2G12 [123]. The synthesis of the V3 domain involved SPPS of the cyclic 47-mer V3 domain peptide containing two GlcNAc-Asn residues. Endo A was used to transfer disaccharide and tetrasaccharide oxazoline substrates to the GlcNAc residues to afford products containing tri- and pentasaccharide glycans respectively (*Scheme 27*). The effect of glycosylation on the conformations of the V3 domain was investigated by CD spectroscopy and attenuated total reflection Fourier transform infrared spectroscopy (ATR-FTIR). These studies again suggested that glycosylation could be exploited as a tool to induce favorable epitopes for an effective gly-

copeptide-based HIV vaccine. In addition, the glycopeptides were significantly more stable to protease digestion when compared to unglycosylated V3 peptide, suggesting increased in vivo lifetimes.

3.3 Other Glycopeptide-Based Vaccines

Reichel et al. have reported the construction of a lipoglycopeptide that may serve as a vaccine against bacterial meningitis [136]. This novel conjugate comprises several different recognition motifs. The carbohydrate portion mimics the inner core of meningococcal lipopolysaccharides (LPS, the "B-epitope"). The peptide chain serves as a T-epitope for immunological memory. Finally, the lipid portion functions in the activation of B-cells, as well as membrane anchoring (*Fig. 9*).

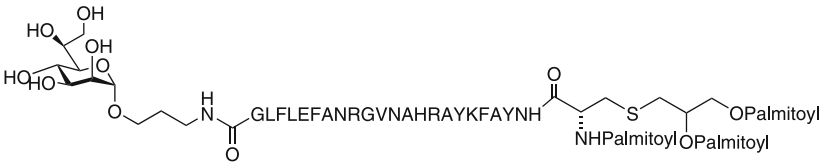


Figure 9
Lipoglycopeptide vaccine of bacterial meningitis

Antigenic haptens to *Shigella*, a genus of bacteria which causes dysentery in humans, have also been reported [137]. Glycoconjugates of this nature may function as a vaccine against dysentery. Numerous synthetic carbohydrates based on the repeating *O*-linked saccharides found in *Shigella* were conjugated to human serum albumin (HSA) carrier protein through a spacer arm for immunization purposes. These glycoconjugates were found to induce immune responses in mice [138]. More recently, potential *Shigella* vaccines were synthesized by attachment of oligosaccharide antigens to the non-natural T-helper peptide human leukocyte antigen (HLA) DR-binding epitope [139,140].

Helicobacter pylori which binds to blood group determinants on the gastric epithelium is responsible for stomach ulcers and gastro adenocarcinoma along with a host of other ailments [141,142]. The specific carbohydrate structures to which the bacterium adheres have been elucidated recently [143]. These antigens are also prime targets for the creation of a vaccine. Capsular polysaccharide conjugates serve as additional vaccine target structures [144]. This type of vaccine may aid in the eradication of pathogenic bacteria from the body through targeting of the carbohydrate coat presented on the bacterial cell surface.

Plasmodium falciparum is a single celled parasite, infections of which are responsible for malaria. The mortality and morbidity associated with this disease are largely the result of inflammatory cascades triggered by the glycosylphosphatidylinositol (GPI) expressed in large quantities on the pathogen's surface [145]. Recently, Seeberger and coworkers developed malarial vaccine candidates by conjugating multiple copies of synthetic GPI to BSA (*Fig. 10*) [146]. Preclinical evaluation of this synthetic vaccine was extremely promising, with vaccinated mice exhibiting a 60–75% survival rate when exposed to the parasite, as opposed to a 0–9% survival rate in untreated mice.

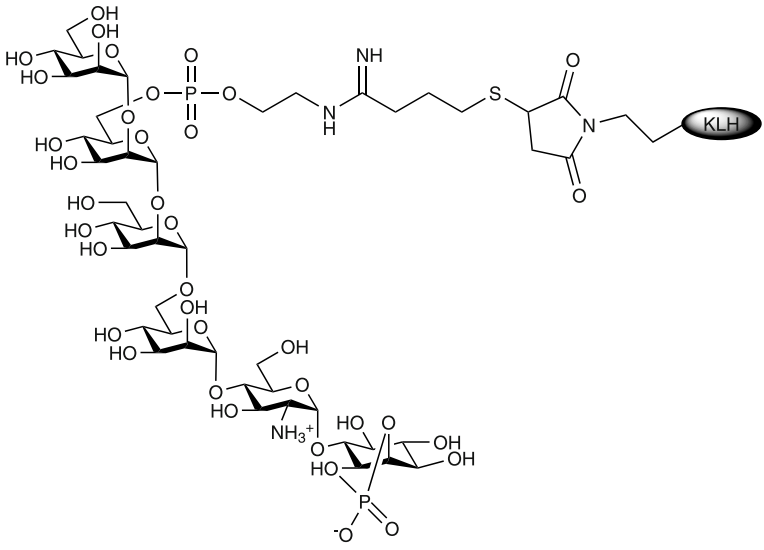


Figure 10
Synthetic GPI-KLH conjugate as a potential malarial vaccine

The unusual lipophosphoglycan (LPG) found on protozoa of the genus *Leishmania* is a promising target for the development of vaccines against these parasites. Recently, Hewitt, Seeberger and coworkers have described the synthesis of two potential Leishmaniasis vaccines based on conjugating the cap tetrasaccharide of the LPG to either KLH or Pam_3Cys [147]. More recently, this group has developed antigens formed by conjugating the cap tetrasaccharide to immunostimulating reconstituted influenza virosomes (IRIVs) through phospholipids or hemagglutinin [148]. Both of these later constructs were shown to induce production of IgG which recognize *Leishmania donovani*.

4 Glycopeptides as Antibiotics and Other Drug Targets

Glycopeptides have the potential to function as bactericidal agents, and as such, are currently used clinically in the treatment of microbial infections. Bacteria have the capability to evolve resistance to antibiotics [149], and as such, constant research into the development of new antibacterial agents is necessary.

Vancomycin is most often the antibiotic of last resort for the treatment of resistant bacterial strains, however, bacterial strains resistant to vancomycin are now emerging [19]. The health threat posed by these strains has led to intense research into both the mechanism by which resistance develops and the development of pharmaceutical antibacterial agents with novel modes of action.

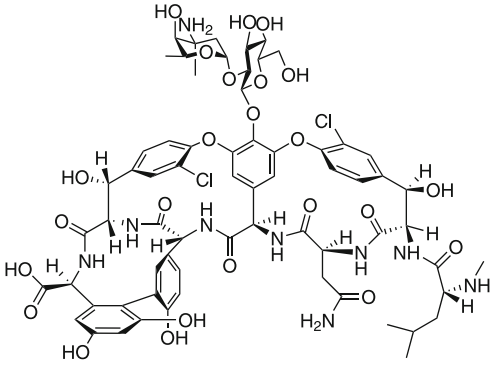
Glycopeptides are also being explored as drug candidates for other means. Addition of a carbohydrate to a peptide backbone often decreases proteolytic degradation and increases oral availability of peptide-based drugs. These are favorable qualities of pharmaceuticals that need to be optimized for agents to reach clinical trials.

4.1 Cyclic Glycopeptide Antibiotics

4.1.1 Vancomycin and Teicoplanin

Vancomycin and teicoplanin are complex cyclic glycopeptide antibiotics (*Fig. 11*). Although the structures of these compounds were elucidated decades ago, total syntheses of vancomycin aglycone [150,151,152,153,154,155], fully glycosylated vancomycin [156], and teicoplanin [157,158] have only recently been reported.

With the emergence of bacterial strains that are resistant to these glycopeptides, the first line of defense is often to synthetically modify the natural antibiotic. In the case of these glycopeptides, synthetic alteration of the aglycone structure is not practical, due to their complexity. However, the carbohydrate portion of the molecule [vancosamine(α1-2)glucose in the case of vancomycin] is readily accessible to modification. This strategy has led to new antibiotics that have potential for treatment of vancomycin-resistant strains.



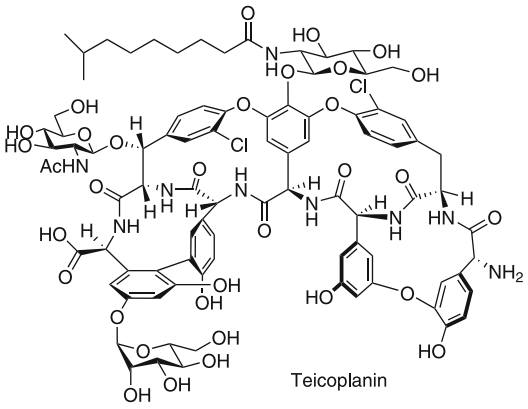


Figure 11
Structures of vancomycin and teicoplanin

It has been observed that the incorporation of lipophilic chains into vancomycin can increase the potency of the drug [159,160,161,162]. This is best exemplified by the observation that teicoplanin retains activity against some vancomycin-resistant bacterial strains. In one early study directed at improving the efficacy of these antibiotics, Ge et al. investigated the mechanism of action of a chloro-biphenyl derivative of vancomycin that showed activity against resistant bacteria [163]. The sulfoxide method of glycosylation was utilized to achieve the desired modification of the gluco-vancomycin precursor (*Scheme 28*).

The analog containing the biaryl moiety was found to act by a different mechanism than vancomycin itself. While vancomycin inhibits transpeptidase activity, the modified structure is thought to inhibit transglycosylase activity (*Scheme 29*). This finding reveals an alternative target for the development of antimicrobial agents. The glycosylation strategy described above provides facile access to numerous novel glycopeptide antibiotics based on the vancomycin core.

While chemical methods provide a powerful tool for the construction of vancomycin derivatives, they suffer from the need for extensive protection/deprotection chemistry. Over the past several years a significant amount of work has been devoted to identifying the glycosyltransferases (Gtfs) responsible for the synthesis of the vancomycin and teicoplanin glycosides [161]. The substrate specificities of these enzymes has been extensively studied and it was found that certain Gtfs were highly promiscuous, i. e. able to transfer a variety of non-natural UDP and TDP sugars to the glycopeptide aglycone [164,165,166]. Importantly, it has been shown that one of these enzymes (GtfE) is tolerant of azido modification of the glucose sugar, opening up the possibility for selective chemical modification of vancomycin derivatives via Huisgen 1,3-dipolar cycloaddition (*Scheme 30*) [167].

In addition to the glycosyltransferase enzymes, the substrate specificity of the teicoplanin acyltransferase (Atf), responsible for the installation of the fatty acid amide in teicoplanin, has been examined [168]. This enzyme appears to be tolerant to modifications in the amide chain length, carbohydrate structure and nature of the aglycone. This finding is of particular interest as it has been shown that installation of the lipid containing monosaccharide from teicoplanin onto vancomycin can restore antibiotic activity against several resistant strains [159].

Besides modification of the glycan, other strategies to overcome vancomycin resistance have been undertaken. Dimeric and trimeric vancomycin constructs bind the D-Ala-D-Ala peptide more tightly than vancomycin itself and have shown antibacterial activity against vancomycin resistant strains [169,170,171]. A potential inhibition pathway that acts through dimerization and membrane-anchoring is also being actively pursued [172,173].

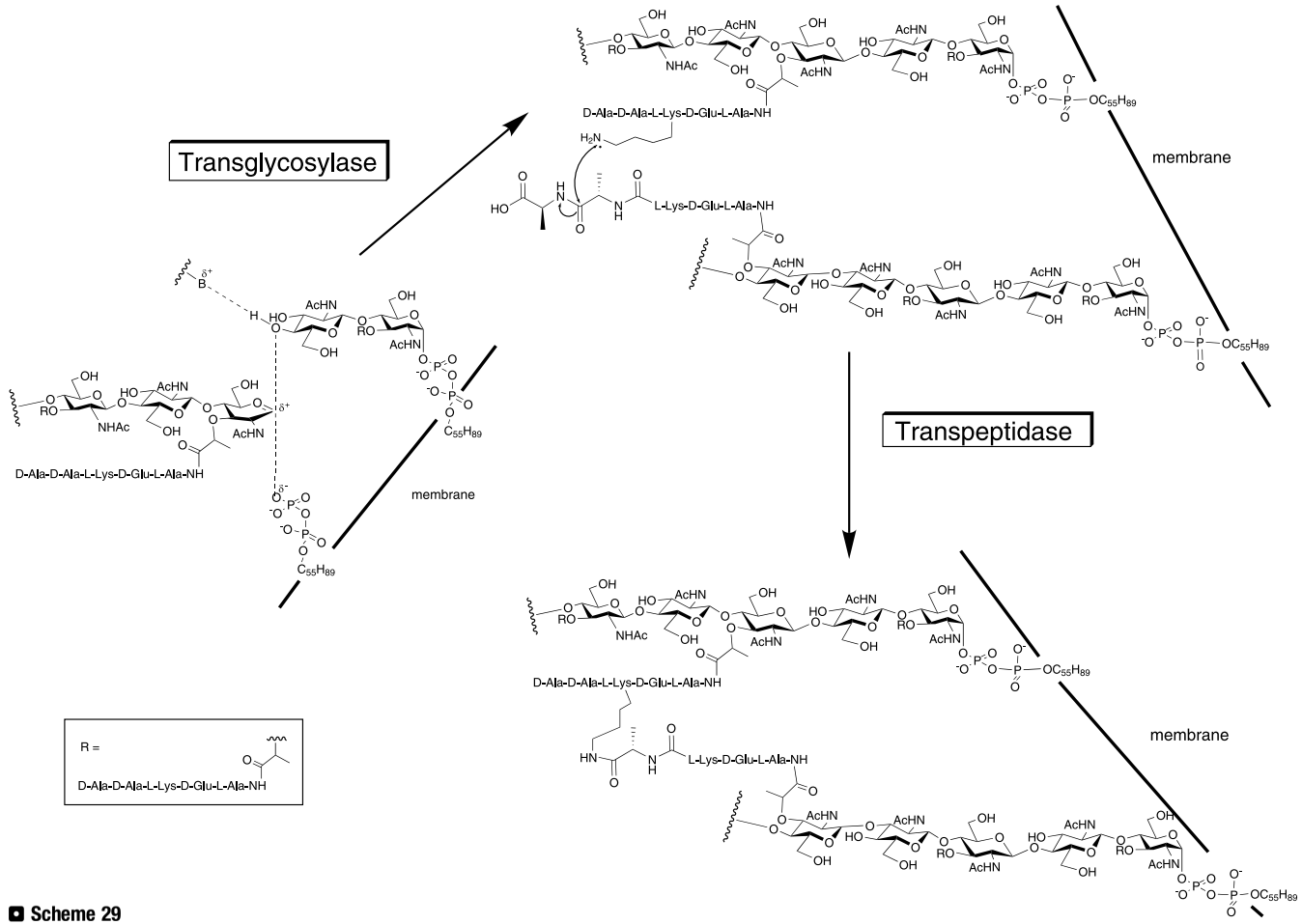
In addition to vancomycin and teicoplanin, several other fused cyclic glycopeptide derivatives have been studied [161,174,175]. In all cases examined, the glycan portion of the molecule is necessary for optimal activity.

4.1.2 Mannopeptimycins

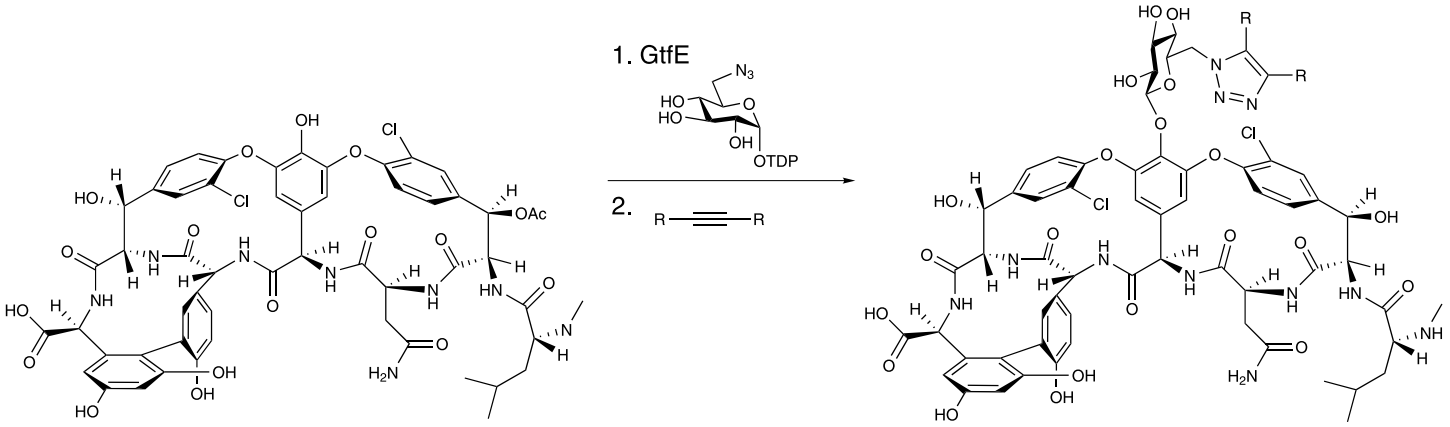
The mannopeptimycins are cyclic (*DL*)-glycohexapeptides which display potent antibacterial activity against gram positive bacteria including methicillin and vancomycin resistant enterococci (*Fig. 12*) [176]. This class of glycopeptides are thought to interfere with bacterial cell wall synthesis by binding to lipid II, although the exact mechanism of action remains unknown [177]. While there is limited SAR data on these molecules, it has been demonstrated

1. chemical glycosylation
2. protecting group manipulations
3. deprotection

Scheme 28


Transglycosylase
Transpeptidase
membrane
membrane
membrane
D-Ala-D-Ala-L-Lys-D-Glu-L-Ala-NH
L-Lys-D-Glu-L-Ala-NH
R =
D-Ala-D-Ala-L-Lys-D-Glu-L-Ala-NH

Scheme 29



Scheme 30

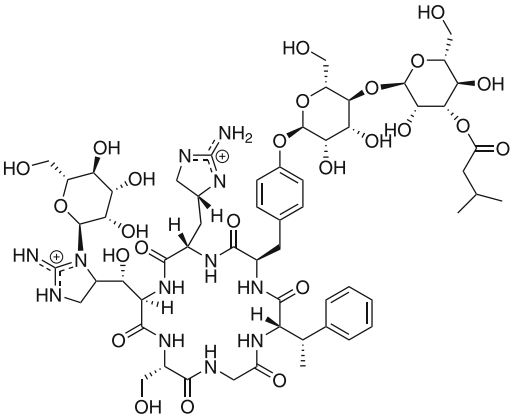


Figure 12
Structure of a mannopeptinycin

that the presence of an isovaleryl ester on the terminal mannose group is essential for retaining activity.

4.1.3 Ramoplanin

Ramoplanin is a lipoglycodepsipeptide isolated from a fermentation broth of Actinoplanes which showed potent antibiotic activity against a number of vancomycin resistant bacterial strains [178]. This compound is currently undergoing phase III clinical trials. While the mechanism of action of ramoplanin remains to be elucidated, it appears that the glycans are not necessary for biological activity and their function remains to be established.

4.2 Linear Glycopeptide Antibiotics

4.2.1 Drosocin

Several linear glycopeptides also show promise as antibacterial agents. Drosocin is a 19 amino acid peptide glycosylated with the disaccharide Gal(β1-3)GalNAc(α1-*O*)Thr (the T antigen) at position 11 (*Fig. 13*). The attached sugars are necessary for optimal biological activity. A recent study by Bulet et al. examined native drosocin, two truncated forms, and the

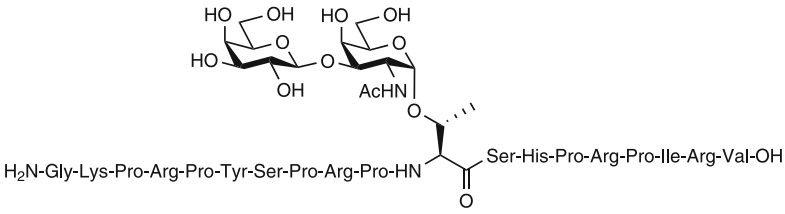


Figure 13
Structure of drosocin

unglycosylated peptide (synthesized by SPPS) in antibacterial assays against several bacterial strains [179]. The requirement of the sugar portion of the glycopeptide for full activity was confirmed by these studies.

4.2.2 Diptericin

The antimicrobial diptericin is an 82-amino acid glycopeptide containing two glycosylation sites. In nature the peptide exists as a mixture of glycoforms, the simplest of which possesses only two *O*-linked GalNAc residues [180]. In addition to two syntheses of this compound by SPPS [181,182], Bertozzi and coworkers have synthesized an analogue of the simplest glycoform using NCL (*Scheme 31*) [47]. As the peptide is devoid of cysteines it was necessary to introduce a Gly25-Cys mutation to facilitate NCL, as well as Asp29-Glu and Asp45-Glu mutations to facilitate the synthesis. The synthesis was initiated by SPPS of a 24-residue *N*-terminal thioester and a 58-residue *C*-terminal fragment containing an *N*-terminal cysteine. In order to circumvent problems associated with thioester hydrolysis, the authors used Ellman's "safety-catch" linker in the synthesis of the *N*-terminal fragment [183]. These fragments underwent smooth NCL to afford, after deprotection, synthetic diptericin which retained antimicrobial activity despite the amino acid substitutions. This report is of note as it was the first complex glycopeptide to be synthesized using NCL.

4.2.3 β-KDO

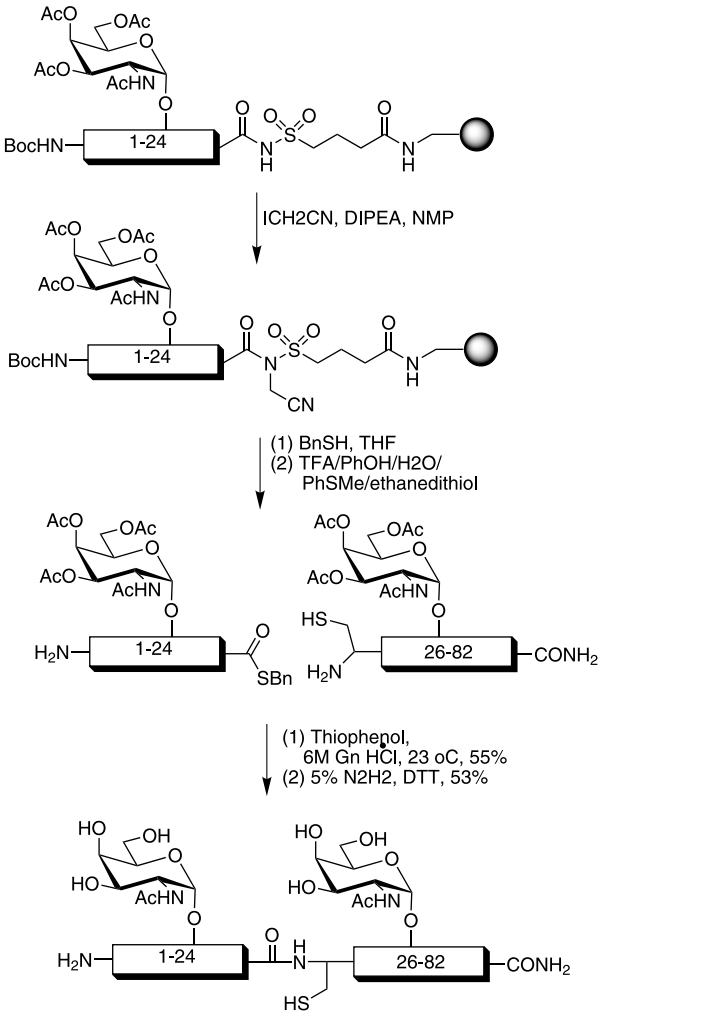
Another interesting antibiotic construct is a peptide-conjugated form of β-KDO. β-KDO is an inhibitor of the CMP-KDO synthetase that is involved in lipopolysaccharide (LPS) synthesis, a major constituent of the bacterial cell wall. Antibacterial activity is thus accomplished by interference with cell wall construction. β-KDO is incapable of membrane permeation, and therefore is not useful as an antibiotic in its native form. Hammond et al. have shown that conjugation of a β-KDO analog to certain dipeptides resolves this problem [184] (*Fig. 14*). Attachment of the β-KDO analog to a dipeptide allows the glycopeptide construct to permeate the bacterial membrane. Inside the cell, proteases hydrolyze the peptide and release the inhibitor, resulting ultimately in bacterial cell death.

4.3 Other Drug Targets

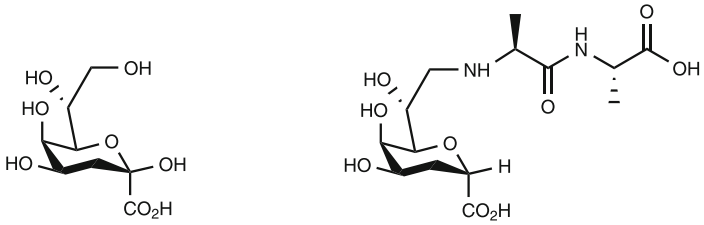
4.3.1 Erythropoietin (epo)

Erythropoietin is a 166 amino acid glycoprotein used to treat anemia arising from a wide range of clinical conditions. The glycoprotein exists as a complex mixture of glycoforms possessing three *N*-linked glycosylation sites and one *O*-linked site. It has been demonstrated that optimal in vivo activity is observed when the *N*-linked glycosylation sites possess tetrasialylated tetra-antennary glycans [29]. Recently, Danishefsky and coworkers have initiated a program directed at synthesizing a homogeneous glycoform of epo in order to gain better insight into the role of the carbohydrate moieties on biological activity. To date this group has synthesized two fragments, the 23–37 glycopeptide domain, and the 114–166 domain.

The synthesis of the 23–37 domain entailed coupling of a complex biantennary glycan to a masked peptide thioester, corresponding to residues 22–28 of epo, under conditions

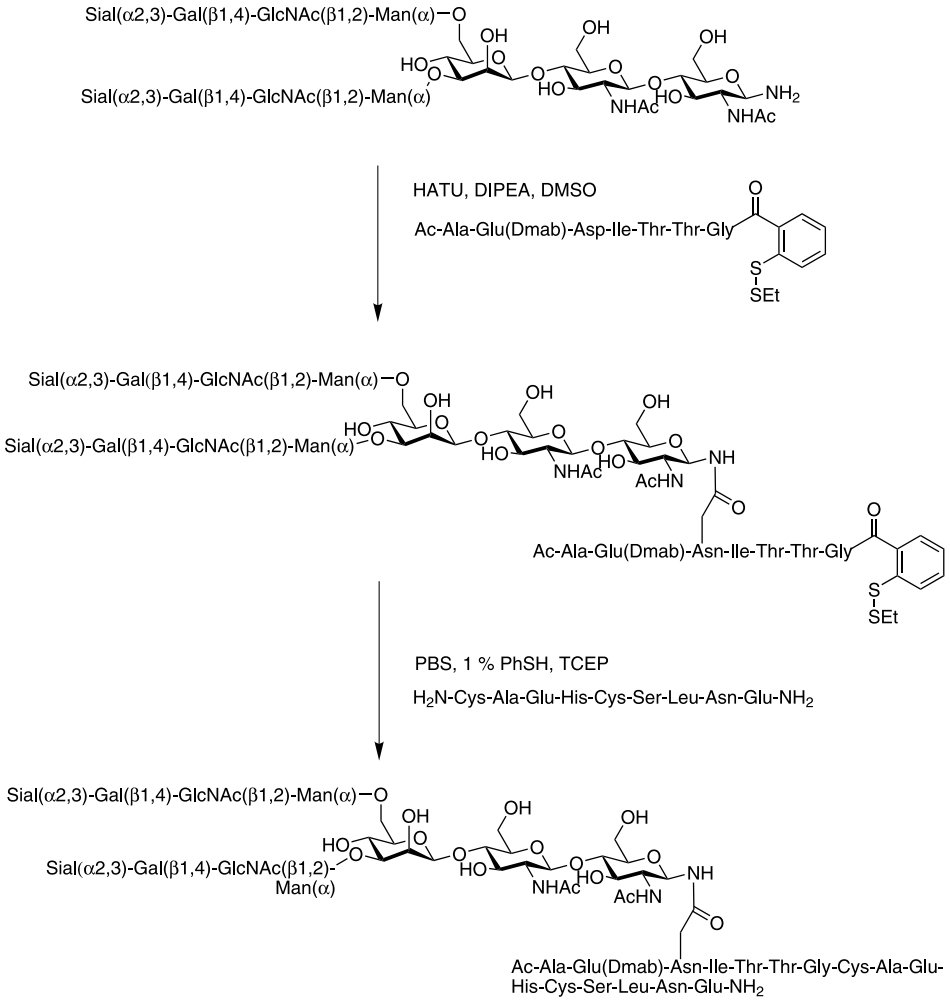


Scheme 31



β-KDO β-KDO di-Alanine analog

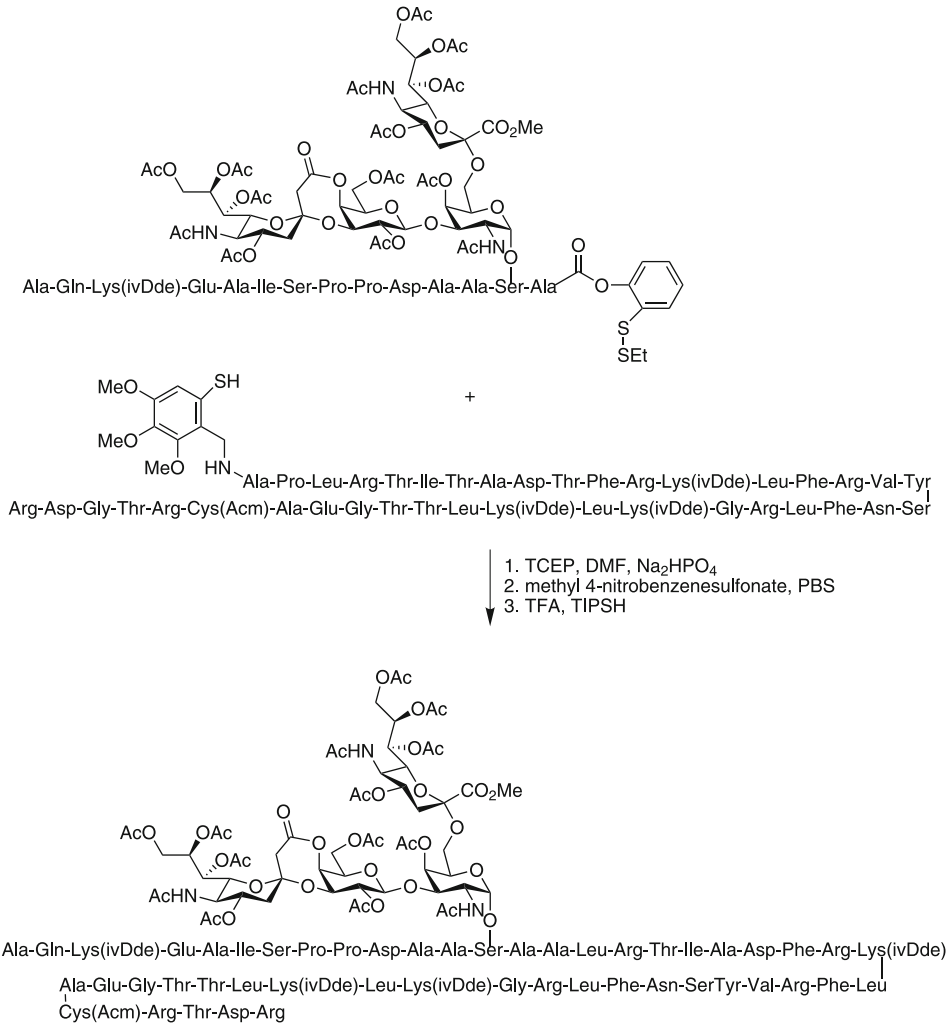
Figure 14
Structure of β-KDO and a glycopeptide variant



Scheme 32

described by Lansbury and coworkers (*Scheme 32*) [38,185]. The resulting glycopeptide was then exposed to reducing conditions where it underwent a reversible O→S acyl transfer to afford the corresponding thioester. This thioester then underwent in situ NCL with a peptide corresponding to residues 29–37 of epo to afford the desired product.

Synthesis of the 114–166 glycopeptide domain proved to be more challenging due to the presence of the glycophorin tetrasaccharide and the need to protect a number of cysteines in the molecule in order to carry out an auxiliary assisted cysteine-free NCL (*Scheme 33*) [186]. To this end, the glycophorin glycopeptide fragment (residues 114–127) was synthesized using the cassette approach and SPPS to afford a protected thioester. This fragment underwent smooth ligation with a Tmb containing acceptor corresponding to epo residues 128–166. Following



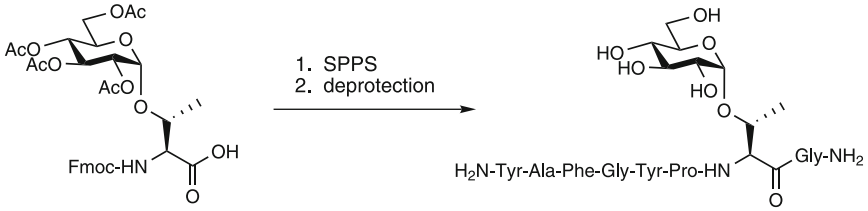
Scheme 33

ligation removal of the Tmb moiety proved to be problematic due to acid induced N→S acyl transfer reactions. It was therefore necessary to mask the aromatic thiol prior to auxiliary removal. It is of note that the sialic acid residue in the glycophorin was not affected by the acidic conditions used for auxiliary removal.

4.3.2 Opioid Receptor Agonists

Glycopeptide derivatives of the μ-and δ-opioid receptor agonists deltorphin and dermorphin have been examined for in vivo activity [187,188]. Peptide agonists themselves generally only reach the central nervous system in low quantities, mainly due to their inability to cross the

blood brain barrier. However, it was hypothesized that conjugation of known peptide agonists to glucose might facilitate uptake through the endothelial barrier by the glucose transporter GLUT-1 (◉ *Scheme 34*).



dermorphin glycosylated analogs

Scheme 34

In keeping with theory, it was found that the deltorphin and dermorphin analogs had excellent oral availability profiles, while still retaining selectivity for the appropriate opioid receptor. Polt and colleagues have examined the effect of glycosylation on enkephalin-based opioid peptides by synthesizing glycopeptides containing monosacccharides, disaccharides or trisaccharides [189]. In this study it was found that the glycopeptides presenting a disaccharide were the most active compounds, with antinociceptive potencies 100–200 times that of morphine. Cyclic glycopeptide enkephalin analogs have also been prepared by Polt et al. [190]. The constructs were prepared by solid-phase methods, and found to be active against the μ- and δ-opioid receptors.

4.3.3 Renin Inhibitors

Glycoconjugate inhibitors of the renin aspartyl protease have also been formulated for increased solubility and altered clearance. Known peptide inhibitors were conjugated to either GlcNAc or mannose, and then evaluated for renin inhibition [191] (◉ *Fig. 15*).

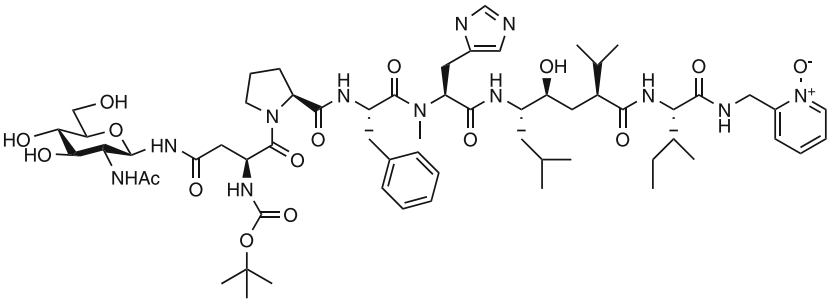


Figure 15
Glycopeptide-based renin inhibitor

The resulting compounds were found to have excellent oral activity as well as elevated serum concentrations due to slower clearance. Other studies of a similar nature have also been reported [192].

5 Other Biologically Relevant Glycopeptide Targets

Glycopeptides serve as target structures in the investigation of a variety of other biological events. For example, peptides presenting sialyl Lewis x (sLe^x) may act as inhibitors of excessive leukocyte extravasation in the inflammatory response [193].
Glycopeptide constructs may also interfere with protein-protein interactions that lead to platelet aggregation and transcriptional regulation. Moreover, studies to determine glycosyltransferase specificity in mucin biosynthesis have also been undertaken. Certain target structures can also be used to detect antibody responses in diseases such as multiple sclerosis. Overall, the specific biological roles of glycopeptides are continually pursued as novel glycoprotein functions are identified.

5.1 Anti-Inflammatory Agents

The inflammatory response is initiated by stimuli released from sites of tissue injury that results in the expression of selectins on the endothelial layer. These selectins (E(endothelial)-selectin and P(platelet)-selectin) function through recognition of oligosaccharides on the opposing leukocyte cell surface [194]. This interaction eventually weakly tethers the leukocyte to the endothelial layer, at which point integrin binding events lead to firm adhesion and extravasation of the leukocyte into the tissue. In certain disease processes, excessive leukocyte infiltration becomes deleterious to the body, and inhibitors of this process are desirable. Rheumatoid arthritis, asthma, organ transplant rejection, and reperfusion injury are just a few of the cases in which these events occur [27].

5.1.1 Polyvalent Selectin Ligands

Several groups have prepared glycopeptide constructs as inhibitors of specific selectins. Sprengard et al. have synthesized *N*-linked trivalent sialyl $Lewis^x$ (sLe^x) cyclic glycopeptides as inhibitors of E-selectin (*Fig. 16*) [195]. Cyclic heptapeptides were prepared, followed by peptide coupling to sLe^x formulated as the glycosyl amine. The trivalent constructs were found to be two to three times as effective as monovalent sLe^x in the inhibition of HL-60 cells binding to immobilized E-selectin.
Baisch and Ohrlein have synthesized unnatural sLe^x containing glycopeptides by chemoenzymatic methods [196]. Asparagine-linked GlcNAc was incorporated into a peptide chain, followed by glycosyltransferase mediated addition of galactose, sialic acid, and fucose. These enzymatic coupling reactions afforded divalent and trivalent sLe^x target structures as E-selectin ligands.
Bruning and Kiessling have prepared divalent glycopeptides bearing 3-sulfo-lactose as inhibitors of L(leukocyte)-selectin [197]. The aim of this study was to investigate the distance-dependence of carbohydrate attachment for optimal selectin inhibition.

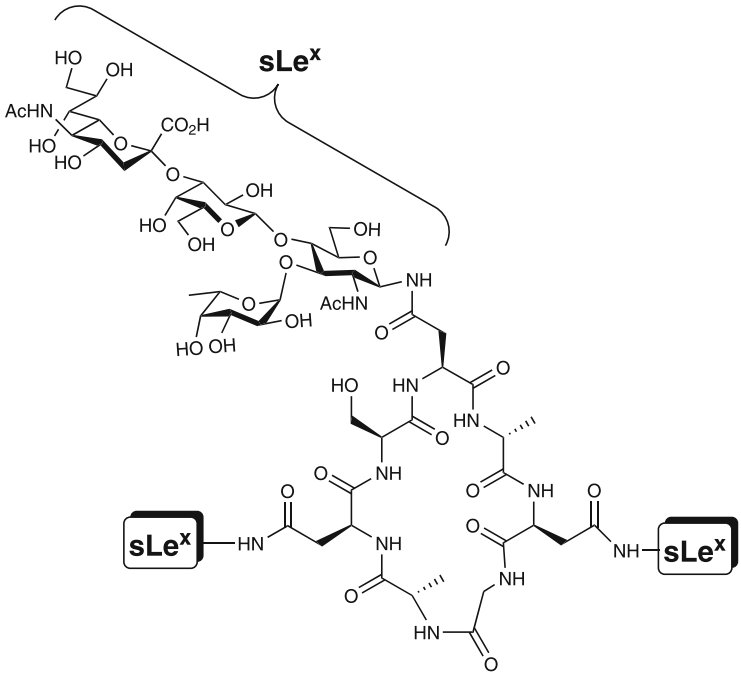


Figure 16
Cyclic trivalent sialyl Lewis X glycopeptide

5.1.2 Dual Function Ligands

Dual function inhibitors of the inflammatory response are also of interest. Sprengard et al. have synthesized an Arg-Gly-Asp-sialyl Lewis^x (RGD-sLe^x) glycopeptide as an "adhesion hybrid" (*Fig. 17*) [198]. The sLe^x portion of the molecule provides a selectin ligand, while the RGD portion targets recognition through an integrin binding motif. The *N*-linked glycopeptide construct was found to bind with higher affinity than most carbohydrate-based selectin ligands to immobilized P-selectin as an inhibitor of leukocyte adhesion.

Matsuda et al. have also reported the synthesis of a Sialyl Lewis^x-RGD conjugate [199]. In this example the authors placed a flexible linker between the two binding epitopes and examined their ability to bind to both integrin and P-selectin. Not only did this conjugate bind tightly to the selectin and integrin molecules, it was found to form stable crosslinks between the two molecules.

Another selectin target ligand whose carbohydrate structure has recently been characterized is P-Selectin Glycoprotein Ligand-1 (PSGL-1) (*Fig. 18*) [200].

The *N*-terminal portion of the dimeric protein contains both a sLe^x glycan and a tyrosine sulfate residue (*Fig. 19*). Both the sulfate and the glycan are required for recognition by P-selectin and L-selectin, whereas E-selectin only requires the presence of the glycan [201]. More detailed studies using a small library of PSGL-1 fragments showed that optimal binding to P-selectin required not only sialyl Lewis^x present on a core 2 mucin structure, but also the presence of three sulfated tyrosine residues [202].

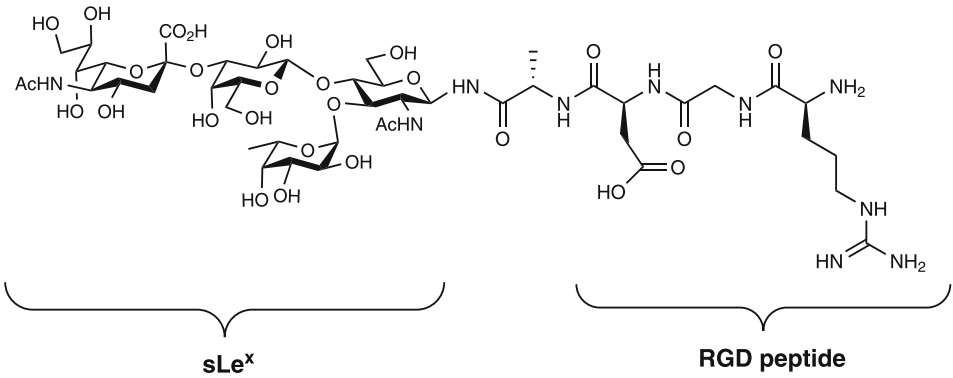
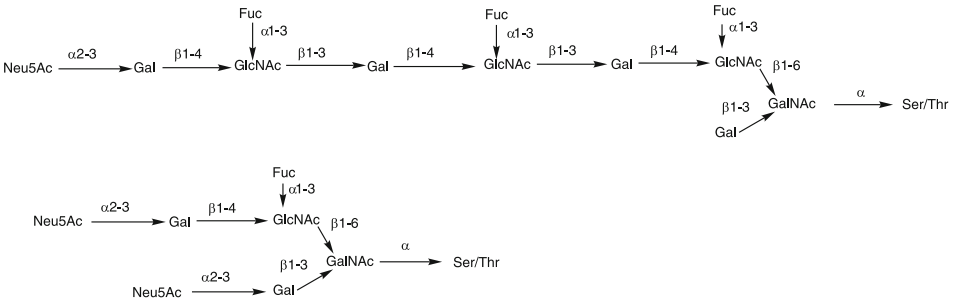
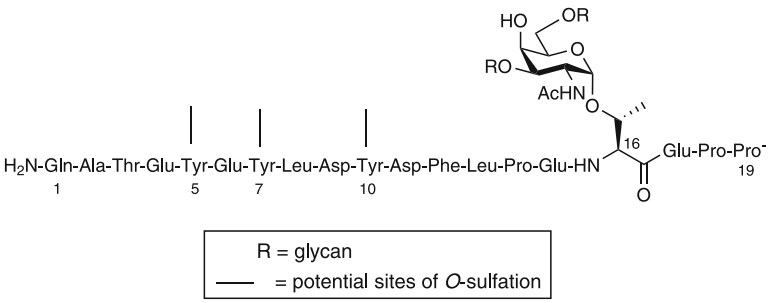


Figure 17
RGO glycopeptide bearing sialyl Lewis X



O-glycans from PSGL-1

Figure 18
***O*-glycans from PSGL-1**



PSGL-1 N-terminal structure

Figure 19
PSGL-1 *N*-terminal structure

5.1.3 Inhibitors of Mannose-6-Phosphate Receptors (MPRs)

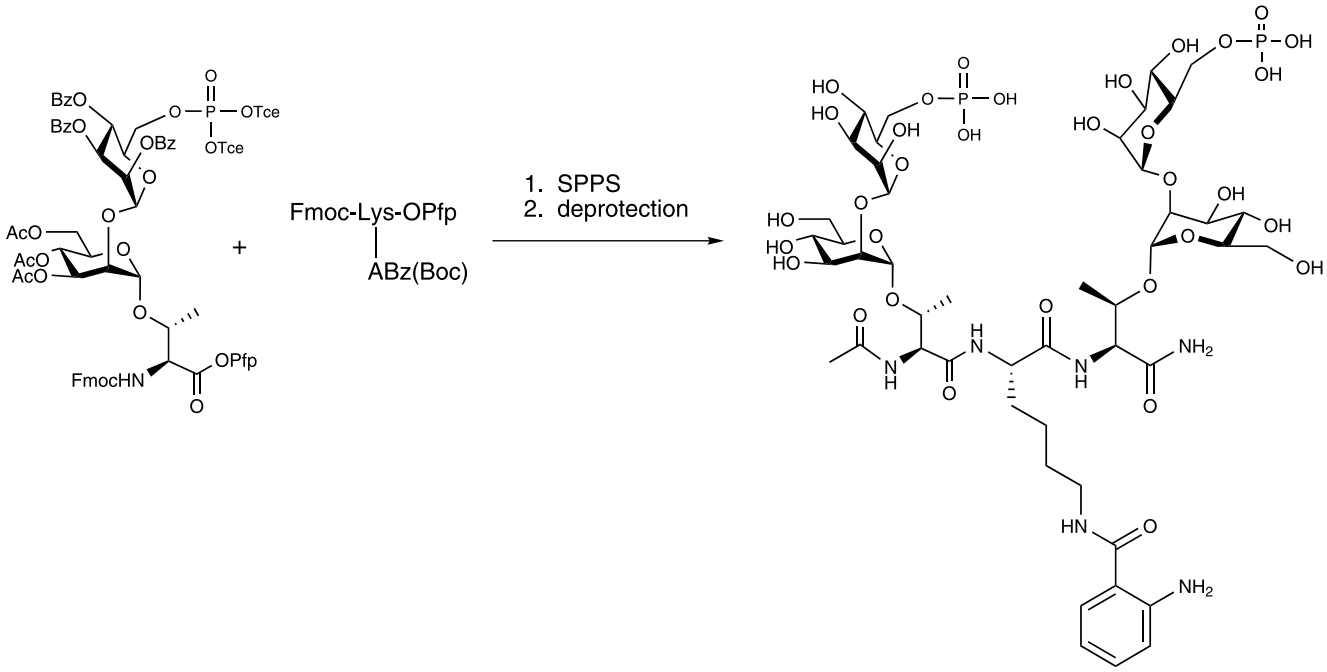
An alternative pathway for inhibition of the inflammatory response targets the mannose-6-phosphate receptors (MPRs). Mannose-6-phosphate inhibits the inflammatory process through binding an MPR, possibly interfering with the transport of lysosomal enzymes. Franzyk et al. have prepared divalent linear and constrained glycopeptides bearing phosphorylated dimannosyl residues (*Scheme 35*) [203]. The resulting compounds were assayed for inhibition of MPR binding to an immobilized phosphomannan core.
Optimal positioning of the glycosidic linkage between the mannose residues was assessed, as well as peptide spacer length between the glycosyl amino acids. In this case, the linear constructs showed more potent inhibition than the constrained cyclic analogs.

5.2 Other Glycopeptide-Based Inhibitors

RGD-containing glycopeptides have also been prepared as inhibitors of platelet aggregation. Vegad et al. have synthesized *O*-linked rhamnose, mannose, or glucose-containing linear and cyclic glycopeptides by solid-phase methodology [204]. In this case, addition of the sugar to the peptide resulted in compounds of equal or lesser potency in platelet aggregation assays.
An *N*-linked glycopeptide from vitamin K-dependent protein S was prepared by Holm et al. [205]. The chitobiose containing polypeptide corresponding to amino acid sequence 447–461 of protein S was synthesized by solid-phase protocols, utilizing the glycosyl amino acid. It was observed that the glycopeptide was a more efficient inhibitor of complex formation between compliment associated proteins C4BP and protein S than the unglycosylated peptide. These results suggest that the peptide sequence containing the potential *N*-glycosylation site (Asx-turn motif) may naturally present an *N*-glycan.

5.3 Studies of Glycosyltransferase Specificity

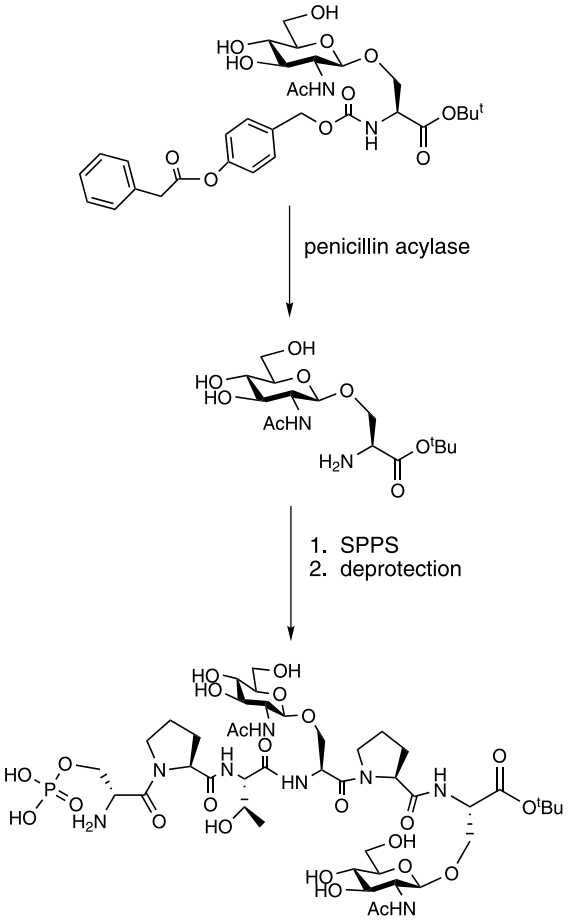
Glycopeptides corresponding to the core one through core seven structures present in mucin glycoproteins have been prepared. The aim of these studies was to synthesize various glycopeptides for the analysis of glycosyltransferase specificity in mucin biosynthesis. The Bock group has simultaneously prepared a library of glycopeptides by utilizing multiple column SPPS methodology [206,207]. Instrumental to the synthetic strategy was the use of glycosyl amino acids activated by pentafluorophenyl esters. The resulting library can be employed to screen specificity of core one through seven glycosyltransferases, as well as sialyltransferases and sulfotransferases.
Polypeptide *N*-α-acetylgalactosaminyltransferases (ppGalNAcTs) are responsible for transferring GalNAc to Ser/Thr on the backbone of mucin type proteins [22]. In order to probe the substrate preferences of this family of enzymes, a library of mucin type glycopeptides containing a variety of glycan clustering motifs was prepared [208]. This study demonstrated that different ppGalNAcTs preferred different densities of glycosylation along the peptide backbone for catalytic activity.


BzO
BzO
OBz
O
OTce
OTce
O
AcO
AcO
AcO
O
O
FmocHN
OPfp
O
+
Fmoc-Lys-OPfp
ABz(Boc)
1. SPPS
2. deprotection
HO
HO
HO
HO
HO
HO
OH
OH
OH
OH
OH
OH
OH
OH
NH2
NH2
HN
O
H
N

Scheme 35

5.4 Studies of Transcriptional Regulation

RNA Polymerase II (RNA Pol II) contains a recently identified mono-GlcNAc modification attached to the polypeptide chain [209]. This β-*O*-GlcNAc residue is thought to act in a regulatory fashion, potentially in a reciprocal role with phosphorylation [210]. In the cell, glycosylated RNA Pol II is transported from the cytosol into the cell nucleus. In the nucleus, the glycosides are cleaved and RNA Pol II is heavily phosphorylated. Only after this transformation occurs can the enzyme function in its transcriptional role. Pohl and Waldmann have synthesized a glycophosphopeptide that is a postulated intermediate in this regulatory process (*Scheme 36*) [211]. The glycophosphopeptide was synthesized utilizing enzymatic protecting group cleavage and block chemical synthesis.



Scheme 36

5.5 Detection of Antibody Response

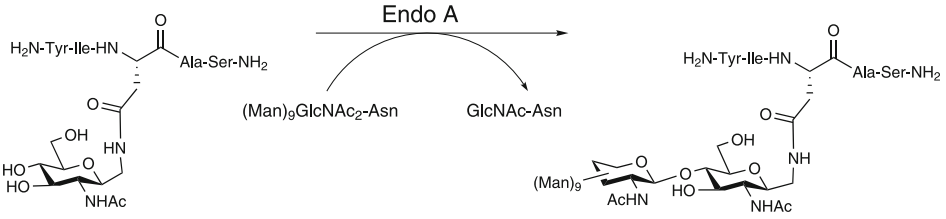
Mazzucco et al. have prepared glycopeptides of human myelin oligodendrocyte glycoprotein (hMOG) to detect antibody responses in multiple sclerosis (MS) [212]. SPPS afforded three glycopeptide constructs with varying glycosylation patterns. Antibody reactivity to the *N*-linked glycopeptide was found in the serum of patients affected by MS, but not in that of normal donors. Neither the unglycosylated peptide nor *O*-linked glycopeptides produced antibody responses in the serum of patients affected by MS. These results suggest that a specific glycopeptide epitope is required for detection of MS antibodies generated in affected patients.

6 Glycopeptide Mimetics

Mimetics of naturally occurring structures often find use as pharmaceuticals. In other cases, they may help clarify issues of biological relevance that are not accessible using natural products. Often times, designed mimetic structures are isosteric with the structure they are designed to resemble, but contain a non-hydrolyzable linkage.

6.1 *C*-Linked Glycopeptide Glycoamidase Inhibitor

Wang et al. have reported the synthesis of a *C*-linked high mannose containing glycopeptide that acts as a broad-spectrum inhibitor of glycoamidases [213]. The *C*-linked glycosylated asparagine residue was synthesized from the corresponding anomeric cyanide, and was subsequently incorporated into the peptide chain using standard peptide coupling methods. Endo A was then utilized to transfer $Man_9GlcNAc$ from $Man_9GlcNAc_2$-Asn to the synthetic peptide to give the target *C*-glycopeptide (*Scheme 37*). The natural *N*-linked glycopeptide was synthesized in an identical fashion, starting with the glycosyl amine.



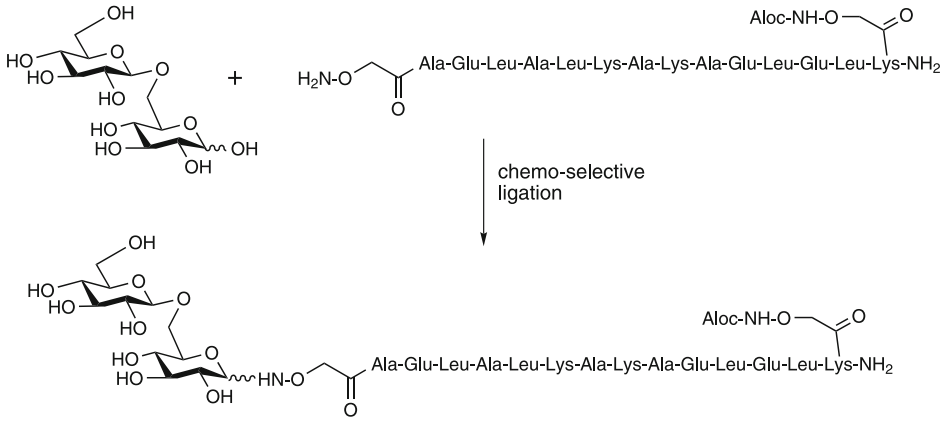
Scheme 37

The *C*-glycopeptide was assayed for inhibition against several glycoamidases of bacterial, plant, and animal origin. The assay results indicated that the *C*-glycopeptide was capable of inhibiting the entire range of enzymes tested. As such, mechanistic studies of further enzymatic activity are possible with constructs of a similar nature.

6.2 Chemoselective Ligation

Another potential avenue to the construction of glycopeptide mimetics is through "chemoselective" ligation of functional groups with matched reactivities. Cervigni et al. have report-

ed the ligation of carbohydrates to a peptide scaffold through the formation of an oxime bond [214]. The ingenuity of this strategy lies in the fact that both the carbohydrate and peptide can be ligated in largely unprotected forms. The ligation was found to be selective, independent of the peptide sequence (*Scheme 38*).



Scheme 38

A second approach relies on the traceless Staudinger ligation between an appropriately functionalized phosphine and an azide to produce a new amide bond [54]. This method has found extensive utility in the labeling of peptides and proteins [215]. Importantly, this method has also been used for the construction of native *N*-linked glycans through the reaction between a glycosyl azide and a phosphine containing amino acid [216,217].

The Huisgen [3+2] cycloaddition between azides and alkynes is another bioorthognal ligation reaction for incorporation of probes into protein and peptide scaffolds. Two variants of this reaction have been developed using either copper(I) [218] or strained cyclooctyne molecules [219] to promote the reaction. As with the Staudinger ligation this method has found extensive use in protein and peptide labeling studies.

In addition to using bioorthogonal reactions, several investigators have synthesized neoglycoproteins by harnessing the reactivity of naturally occurring cysteine residues [220]. In one recent example, it was shown that PhSeBr could promote the reaction between a 1-thiosugar and a cysteine containing protein to form a neoglycoconjugate [221]. It was found that regardless of which reaction partner was activated with PhSeBr, the reaction proceeds through a (phenylselenyl)sulfide containing protein.

6.3 Chemoselective Ligation and Enzymatic Methods

As an extension of this method, Rodriguez et al. combined SPPS, enzymatic functional group conversion, and chemoselective ligation to construct a mimetic of the glycopeptide antibiotic drosocin [222]. Initially, GalNAc was incorporated into the peptide sequence in a standard α-*O*-linkage to threonine. The primary 6-hydroxyl was then oxidized to the aldehyde utilizing galactose oxidase. Chemoselective ligation of the aldehyde with the β-glycosyl amine of

Scheme 39

GalNAc then provided the desired oxime linkage (*Scheme 39*).

The glycopeptide mimetic structure was found to have antibacterial activity that exceeded that of the unglycosylated drosocin construct. As such, the authors concluded that altered carbohydrate structure outside of the conserved glycosyl-peptide bond was a tolerated modification in the biological system.

These authors have described other functionality that can be utilized for chemoselective ligation to form glycopeptide mimetics. Glycopeptides with an inserted ketone functionality were conjugated to carbohydrates formulated as anomeric thiosemicarbazides or aminooxy groups [223]. The resulting "neoglycopeptides" were mimetics of selectin ligands or the *N*-linked glycoprotein motif.

More recently, using suppressor tRNA technology the Schultz and Wong labs were able to incorporate a ketone moiety into expressed proteins [224]. Using an orthogonal tRNA/tRNA synthetase pair from *Methanococcus jannaschii*, which was evolved to charge amber suppressing tRNA, the investigators were able to incorporate *p*-acetyl-L-phenylalanine site specifically into proteins in response to a stop codon (TAG). Using this methodology, a mutant of the Z domain of staphylococcal protein A containing an unnatural ketone moiety was expressed in *E. coli*. Following derivatization with 1-aminoxy-GlcNAc the authors were able to incorporate galactose and sialic acid into the protein backbone using glycosyltransferases. Alternatively,

it was demonstrated that the aminooxy trisaccharide could be assembled prior to conjugation to the protein without loss in efficiency. Subsequent work in this area has allowed the direct introduction of GlcNAc-β-serine [225] and GlcNAc-α-threonine [226] into proteins expressed in *E. coli.*

An even broader extension of this methodology has been applied to expression of chemoselective ligation sites on the cell surface. By exploiting the biosynthetic pathway that produces cell surface sialic acid, Mahal et al. were able to produce an unnaturally occurring ketone functionality on Jurkat cells [227]. Through the use of ligation procedures successfully demonstrated on isolated glycopeptides, the authors have an accessible route to "molecular landscaping" on the cell surface. This work has recently been extended to include the incorporation of azido moieties (for Staudinger ligations [228] and "click chemistry" [229,230]) and alkynes (for "click chemistry" [230]) into whole cell expression systems.

7 Concluding Remarks

Glycopeptide synthesis represents an essential tool for the study of the structure and function of glycoproteins. Despite enormous progress, a superior route for the preparation of adequate amounts of material for biological study remains elusive. The area of glycoprotein and glycopeptide research will continue to supply valuable information about biological systems and is expected to intensify in the future.

References

1. Varki A (1993) Glycobiology 3:97
2. Dwek RA (1996) Chem Rev 96:683
3. Sears P, Wong CH (1998) Cell Mol Life Sci 54:223
4. Crocker PR, Feizi T (1996) Curr Opin Struc Biol 6:679
5. Jenkins N, Parekh RB, James DC (1996) Nat Biotechnol 14:975
6. O'Connor SE, Imperiali B (1996) Chem Biol 3:803
7. Schachter H (1986) Biochem Cell Biol 64:163
8. Piesecki S, Alhadeff JA (1992) Biochim Biophys Acta Protein Struct Mol Enzymol 1119:194
9. Yamaguchi K, Akai K, Kawanishi G, Ueda M, Masuda S, Sasaki R (1991) J Biol Chem 266:20434
10. Brik A, Ficht S, Wong CH (2007) Curr Opin Chem Biol 10:683
11. Takegawa K, Tabuchi M, Yamaguchi S, Kondo A, Kato I, Iwahara S (1995) J Biol Chem 270:3094
12. Witte K, Sears P, Martin R, Wong CH (1997) J Am Chem Soc 119:2114
13. Haneda K, Inazu T, Mizuno M, Iguchi R, Yamamoto K, Kumagai H, Aimoto S, Suzuki H, Noda T (1998) Bioorg Med Chem Lett 8:1303
14. Liu L, Bennett CS, Wong CH (2006) Chem Commun: 21
15. Taylor CM (1998) Tetrahedron 54:11317
16. Arsequell G, Valencia G (1997) Tetrahedron Asymmetry 8:2839
17. Walsh CT, Garneau-Tsodikova S, Gatto GJ (2005) Angew Chem Int Ed 44:7342
18. Silberstein S, Gilmore R (1996) FASEB J 10:849
19. Imperiali B, Spencer JR, Struthers MD (1994) J Am Chem Soc 116:8424
20. Elhammer AP, Poorman RA, Brown E, Maggiora LL, Hoogerheide JG, Kezdy FJ (1993) J Biol Chem 268:10029
21. Bourne Y, Henrissat B (2001) Curr Opin Struc Biol 11:593
22. Hang HC, Bertozzi CR (2005) Bioorg Med Chem 13:5021
23. Ragupathi G, Koganty RR, Qiu DX, Lloyd KO, Livingston PO (1998) Glycoconjugate J 15:217
24. Becker T, Dziadek S, Wittrock S, Kunz H (2006) Curr Cancer Drug Targets 6:491

25. Spearman P (2006) Curr Pharm Design 12:1147
26. Wang LX (2006) Curr Opin Drug Disc 9:194
27. Carlos TM, Harlan JM (1994) Blood 84:2068
28. Nagarjan R (1993) J Antibiot 46:1181
29. Yuen CT, Storring PL, Tiplady RJ, Izquierdo M, Wait R, Gee CK, Gerson P, Lloyd P, Cremata JA (2003) Brit J Haematol 121:511
30. Kunz H (1987) Angew Chem Int Ed Engl 26:294
31. Garg HG, Vondembruch K, Kunz H (1994) Adv Carbohydr Chem Biochem 50:277
32. Kunz H (1997) In: Hanessian H (ed) Preparative Carbohydrate Chemistry. Marcel Dekker, Inc, New York, NY
33. Bielfeldt T, Peters S, Meldal M, Bock K, Paulsen H (1992) Angew Chem Int Ed Engl 31:857
34. Peters S, Bielfeldt T, Meldal M, Bock K, Paulsen H (1992) J Chem Soc Perkin Trans 1:1163
35. Matsushita T, Hinou H, Fumoto M, Kurogochi M, Fujitani N, Shimizu H, Nishimura SI (2006) J Org Chem 71:3051
36. Meinjohanns E, Meldal M, Paulsen H, Dwek RA, Bock K (1998) J Chem Soc Perkin Trans 1:549
37. Anisfeld ST, Lansbury PT (1990) J Org Chem 55:5560
38. Cohenanisfeld ST, Lansbury PT (1993) J Am Chem Soc 115:10531
39. Roberge JY, Beebe X, Danishefsky SJ (1998) J Am Chem Soc 120:3915
40. Schleyer A, Meldal M, Manat R, Paulsen H, Bock K (1997) Angew Chem Int Ed Engl 36:1976
41. Macmillan D (2006) Angew Chem Int Ed 45:7668
42. Pratt MR, Bertozzi CR (2005) Chem Soc Rev 34:58
43. Wieland T, Bokelmann E, Bauer L, Lang HU, Lau H (1953) Liebigs Ann Chem 583:129
44. Dawson PE, Muir TW, Clarklewis I, Kent SBH (1994) Science 266:776
45. Hackeng TM, Griffin JH, Dawson PE (1999) Proc Natl Acad Sci USA 96:10068
46. Dawson PE, Kent SBH (2000) Annual Review of Biochemistry 69:923
47. Shin Y, Winans KA, Backes BJ, Kent SBH, Ellman JA, Bertozzi CR (1999) J Am Chem Soc 121:11684
48. Botti P, Carrasco MR, Kent SBH (2001) Tetrahedron Lett 42:1831
49. Offer J, Boddy CNC, Dawson PE (2002) J Am Chem Soc 124:4642
50. Macmillan D, Anderson DW (2004) Org Lett 6:4659
51. Wu B, Chen JH, Warren JD, Chen G, Hua ZH, Danishefsky SJ (2006) Angew Chem Int Ed 45:4116
52. Brik A, Yang YY, Ficht S, Wong CH (2006) J Am Chem Soc 128:5626
53. Brik A, Ficht S, Yang YY, Bennett CS, Wong CH (2006) J Am Chem Soc 128:15026
54. Saxon E, Armstrong JI, Bertozzi CR (2000) Org Lett 2:2141
55. Nilsson BL, Soellner MB, Raines RT (2005) Annu Rev Biophys Biomol Struct 34:91
56. Soellner MB, Nilsson BL, Raines RT (2006) J Am Chem Soc 128:8820
57. Liu L, Hong ZY, Wong CH (2006) ChemBioChem 7:429
58. Erlanson DA, Chytil M, Verdine GL (1996) Chem Biol 3:981
59. Tolbert TJ, Franke D, Wong CH (2005) Bioorg Med Chem 13:909
60. Muir TW, Sondhi D, Cole PA (1998) Proc Natl Acad Sci USA 95:6705
61. Evans TC, Benner J, Xu MQ (1998) Protein Sci 7:2256
62. Macmillan D, Bertozzi CR (2004) Angew Chem Int Ed 43:1355
63. Witte K, Seitz O, Wong CH (1998) J Am Chem Soc 120:1979
64. Wong CH, Schuster M, Wang P, Sears P (1993) J Am Chem Soc 115:5893
65. Unverzagt C, Kunz H, Paulson JC (1990) J Am Chem Soc 112:9308
66. Schultz M, Kunz H (1993) Tetrahedron Asymmetry 4:1205
67. Herrmann GF, Wang P, Shen GJ, Wong CH (1994) Angew Chem Int Ed Engl 33:1241
68. Mackenzie LF, Wang QP, Warren RAJ, Withers SG (1998) J Am Chem Soc 120:5583
69. Perugino G, Cobucci-Ponzano B, Rossi M, Moracci M (2005) Adv Synth Catal 347: 941
70. Braun P W, H, Kunz, H (1993) Bioorg Med Chem 1:197
71. Kunz H, Kowalczyk D, Braun P, Braum G (1994) Angew Chem Int Ed Engl 33:336
72. Schuster M, Wang P, Paulson JC, Wong CH (1994) J Am Chem Soc 116:1135
73. Fan JQ, Takegawa K, Iwahara S, Kondo A, Kato I, Abeygunawardana C, Lee YC (1995) J Biol Chem 270:17723
74. Li B, Zeng Y, Hauser S, Song HJ, Wang LX (2005) J Am Chem Soc 127:9692

75. Kobata A (1996) In: Montreuil J, Vliegenthart JFG, Schachter H (ed) Glycoproteins and Disease. Elsevier, Amsterdam, p 183
76. Carlstedt I, Davies JR (1997) Biochem Soc Trans 25:214
77. Meezan E, Wu HC, Black PH, Robbins PW (1969) Biochemistry 8:2518
78. Kim YJ, Varki A (1997) Glycoconjugate J 14:569
79. Schachter H, Brockhausen I (1992) In: Allen HJ, Kesailus EC (ed) Glycoconjugates: Composition, Structure and Function. Dekker, New York, NY, p 263
80. Hakomori S, Zhang YM (1997) Chem Biol 4:97
81. Sell S (1990) Hum Pathol 21:1003
82. Taylorpapadimitriou J, Epenetos AA (1994) Trends Biotechnol 12:227
83. Danishefsky SJ, Allen JR (2000) Angew Chem Int Ed 39:836
84. Ragupathi G, Coltart DM, Williams LJ, Koide F, Kagan E, Allen J, Harris C, Glunz PW, Livingston PO, Danishefsky SJ (2002) Proc Natl Acad Sci USA 99:13699
85. Sandmaier BM, Oparin DV, Holmberg LA, Reddish MA, MacLean GD, Longenecker BM (1999) Journal of Immunotherapy 22:54
86. Koganty RR, Reddish MA, Longenecker BM (1996) Drug Discovery Today 1:190
87. Ragupathi G, Slovin SF, Adluri S, Sames D, Kim IJ, Kim HM, Spassova M, Bornmann WG, Lloyd KO, Scher HI, Livingston PO, Danishefsky SJ (1999) Angew Chem Int Ed 38:563
88. Hollingsworth MA, Swanson BJ (2004) Nat Rev Cancer 4:45
89. Dube DH, Bertozzi CR (2005) Nat Rev Drug Discov 4:477
90. Hakomori SI (1996) In: Montreuil J, Vliegenthart JFG, Schachter H (ed) Glycoproteins and Disease. Elsevier, Amsterdam, p 243
91. Kunz H, Birnbach S, Wernig P (1990) Carbohydr Res 202:207
92. Seitz O, Kunz H (1995) Angew Chem Int Ed Engl 34:803
93. Habermann J, Kunz H (1998) Tetrahedron Lett 39:4797
94. Seitz O, Wong CH (1997) J Am Chem Soc 119:8766
95. Dupradeau FY, Stroud MR, Boivin D, Li LS, Hakomori S, Singhal AK, Toyokuni T (1994) Bioorg Med Chem Lett 4:1813
96. Klich G, Paulsen H, Meyer B, Meldal M, Bock K (1997) Carbohydr Res 299:33
97. Kuduk SD, Schwarz JB, Chen XT, Glunz PW, Sames D, Ragupathi G, Livingston PO, Danishefsky SJ (1998) J Am Chem Soc 120:12474
98. Toyokuni Y, Hakomori SI, Singhal AK, (1994) Bioorg Med Chem 2:1119
99. Buskas T, Ingale S, Boons GJ (2005) Angew Chem Int Ed 44:5985
100. Wiertz E, Vangaansvandenbrink JAM, Gausepohl H, Prochnickachalufour A, Hoogerhout P, Poolman JT (1992) J Exp Med 176:79
101. Ciommer M, Kunz H (1991) Synlett: 593
102. Paulsen H, Peters S, Bielfeldt T, Meldal M, Bock K (1995) Carbohydr Res 268:17
103. Nakahara Y, Iijima H, Sibayama S, Ogawa T (1990) Tetrahedron Lett 31:6897
104. Liebe B, Kunz H (1997) Angew Chem Int Ed Engl 36:618
105. Schwarz JB, Kuduk SD, Chen XT, Sames D, Glunz PW, Danishefsky SJ (1999) J Am Chem Soc 121:2662
106. Sames D, Chen XT, Danishefsky SJ (1997) Nature 389:587
107. Keil S, Claus C, Dippold W, Kunz H (2001) Angew Chem Int Ed 40:366
108. Dziadek S, Hobel A, Schmitt E, Kunz H (2005) Angew Chem Int Ed 44:7630
109. Wagner M, Dziadek S, Kunz H (2003) Chem Eur J 9:6018
110. Glunz PW, Hintermann S, Schwarz JB, Kuduk SD, Chen XT, Williams LJ, Sames D, Danishefsky SJ, Kudryashov V, Lloyd KO (1999) J Am Chem Soc 121:10636
111. von dem Bruch K, Kunz H (1994) Angew Chem Int Ed Engl 33:101
112. Chen XT, Sames D, Danishefsky SJ (1998) J Am Chem Soc 120:7760
113. Gallo RC, Montagnier L (2003) N Engl J Med 349:2283
114. UNAIDS, AIDS Epidemic Update 2004, Geneva, 2004
115. Garber DA, Silvestri G, Feinberg MB (2004) Lancet Infect Dis 4:397
116. McMichael AJ, Hanke T (2003) Nat Med 9:874
117. Leonard CK, Spellman MW, Riddle L, Harris RJ, Thomas JN, Gregory TJ (1990) J Biol Chem 265:10373
118. Mizuochi T, Matthews TJ, Kato M, Hamako J, Titani K, Solomon J, Feizi T (1990) J Biol Chem 265:8519
119. Zhu XG, Borchers C, Bienstock RJ, Tomer KB (2000) Biochemistry 39:11194
120. Perrin C, Fenouillet E, Jones IM (1998) Virology 242:338

121. Lee WR, Yu XF, Syu WJ, Essex M, Lee TH (1992) J Virol 66:1799
122. Calarese DA, Scanlan CN, Zwick MB, Deechongkit S, Mimura Y, Kunert R, Zhu P, Wormald MR, Stanfield RL, Roux KH, Kelly JW, Rudd PM, Dwek RA, Katinger H, Burton DR, Wilson IA (2003) Science 300:2065
123. Scanlan CN, Pantophlet R, Wormald MR, Saphire EO, Stanfield R, Wilson IA, Katinger H, Dwek RA, Rudd PM, Burton DR (2002) J Virol 76:7306
124. Wang LX, Ni JH, Singh S, Li HG (2004) Chem Biol 11:127
125. Lee HK, Scanlan CN, Huang CY, Chang AY, Calarese DA, Dwek RA, Rudd PM, Burton DR, Wilson IA, Wong CH (2004) Angew Chem Int Ed 43:1000
126. Adams EW, Ratner DM, Bokesch HR, McMahon JB, O'Keefe BR, Seeberger PH (2004) Chem Biol 11:875
127. Calarese DA, Lee HK, Huang CY, Best MD, Astronomo RD, Stanfield RL, Katinger H, Burton DR, Wong CH, Wilson IA (2005) Proc Natl Acad Sci USA 102:13372
128. Geng XD, Dudkin VY, Mandal M, Danishefsky SJ (2004) Angew Chem Int Ed 43:2562
129. Mandal M, Dudkin VY, Geng XD, Danishefsky S (2004) Angew Chem Int Ed 43:2557
130. Dudkin VY, Orlova M, Geng XD, Mandal M, Olson WC, Danishefsky SJ (2004) J Am Chem Soc 126:9560
131. Singh S, Ni JH, Wang LX (2003) Bioorg Med Chem Lett 13:327
132. Wang LX, Song HJ, Liu SW, Lu H, Jiang SB, Ni JH, Li HG (2005) ChemBioChem 6:1068
133. Chan DC, Fass D, Berger JM, Kim PS (1997) Cell 89:263
134. Li HG, Li B, Song HJ, Breydo L, Baskakov IV, Wang LX (2005) J Org Chem 70:9990
135. Zolla-Pazner S (2004) Nat Rev Immunol 4:199
136. Reichel F, Ashton PR, Boons GJ (1997) Chem Commun: 2087
137. Pozsgay V (1998) Angew Chem Int Ed Engl 37:138
138. Pozsgay V, Coxon B, Glaudemans CPJ, Schneerson R, Robbins JB (2003) Synlett: 743
139. Wright K, Guerreiro C, Laurent I, Baleux F, Mulard LA (2004) Org Biomol Chem 2:1518
140. Belot F, Guerreiro C, Baleux F, Mulard LA (2005) Chem Eur J 11:1625
141. Graham DY (1991) J Gastroenterol Hepatol 6:105
142. Forman D, Coleman M, Debacker G, Eider J, Moller H, Damotta LC, Roy P, Abid L, Debacker G, Tjonneland A, Boeing H, Haubrich T, Wahrendorf J, Manousos O, Tulinius H, Ogmundsdottir H, Palli D, Cipriani F, Fukao A, Tsugane S, Miyajima Y, Zatonski W, Tyczynski J, Calheiros J, Zakelj MP, Potocnik M, Webb P, Knight T, Wilson A, Kaye S, Potter J (1993) Lancet 341:1359
143. Boren T, Falk P, Roth KA, Larson G, Normark S (1993) Science 262:1892
144. Sood RK, Fattom A, Pavliak V, Naso RB (1996) Drug Discovery Today 1:381
145. Berhe S, Schofield L, Schwarz RT, Gerold P (1999) Mol Biochem Parasitol 103:273
146. Schofield L, Hewitt MC, Evans K, Siomos MA, Seeberger PH (2002) Nature 418:785
147. Hewitt MC, Seeberger PH (2001) J Org Chem 66:4233
148. Liu XY, Siegrist S, Amacker M, Zurbriggen R, Pluschke G, Seeberger PH (2006) ACS Chem Biol 1:161
149. Malabarba A, Nicas TI, Ciabatti R (1997) Eur J Med Chem 32:459
150. Evans DA, Wood MR, Trotter BW, Richardson TI, Barrow JC, Katz JL (1998) Angew Chem Int Ed Engl 37:2700
151. Evans DA, Dinsmore CJ, Watson PS, Wood MR, Richardson TI, Trotter BW, Katz JL (1998) Angew Chem Int Ed Engl 37:2704
152. Nicolaou KC, Natarajan S, Li H, Jain NF, Hughes R, Solomon ME, Ramanjulu JM, Boddy CNC, Takayanagi M (1998) Angew Chem Int Ed Engl 37:2708
153. Nicolaou KC, Jain NF, Natarajan S, Hughes R, Solomon ME, Li H, Ramanjulu JM, Takayanagi M, Koumbis AE, Bando T (1998) Angew Chem Int Ed Engl 37:2714
154. Nicolaou KC, Takayanagi M, Jain NF, Natarajan S, Koumbis AE, Bando T, Ramanjulu JM (1998) Angew Chem Int Ed Engl 37:2717
155. Boger DL, Miyazaki S, Kim SH, Wu JH, Loiseleur O, Castle SL (1999) J Am Chem Soc 121:3226
156. Nicolaou KC, Mitchell HJ, Jain NF, Winssinger N, Hughes R, Bando T (1999) Angew Chem Int Ed 38:240
157. Boger DL, Kim SH, Miyazaki S, Strittmatter H, Weng JH, Mori Y, Rogel O, Castle SL, McAtee JJ (2000) J Am Chem Soc 122:7416
158. Evans DA, Katz JL, Peterson GS, Hintermann T (2001) J Am Chem Soc 123:12411

159. Dong SD, Oberthur M, Losey HC, Anderson JW, Eggert US, Peczuh MW, Walsh CT, Kahne D (2002) J Am Chem Soc 124:9064
160. Nagarajan R, Schabel AA, Occolowitz JL, Counter FT, Ott JL (1988) J Antibiot 41:1430
161. Kahne D, Leimkuhler C, Wei L, Walsh C (2005) Chem Rev 105:425
162. Nagarajan R, Schabel AA, Occolowitz JL, Counter FT, Ott JL, Feltyduckworth AM (1989) J Antibiot 42:63
163. Ge M, Chen Z, Onishi HR, Kohler J, Silver LL, Kerns R, Fukuzawa S, Thompson C, Kahne D (1999) Science 284:507
164. Losey HC, Jiang JQ, Biggins JB, Oberthur M, Ye XY, Dong SD, Kahne D, Thorson JS, Walsh CT (2002) Chem Biol 9:1305
165. Oberthur M, Leimkuhler C, Kruger RG, Lu W, Walsh CT, Kahne D (2005) J Am Chem Soc 127:10747
166. Fu X, Albermann C, Jiang JQ, Liao JC, Zhang CS, Thorson JS (2003) Nat Biotechnol 21:1467
167. Fu X, Albermann C, Zhang CS, Thorson JS (2005) Org Lett 7:1513
168. Kruger RG, Lu W, Oberthur M, Tao JH, Kahne D, Walsh CT (2005) Chem Biol 12:131
169. Sundram UN, Griffin JH, Nicas TI (1996) J Am Chem Soc 118:13107
170. Rao JH, Whitesides GM (1997) J Am Chem Soc 119:10286
171. Rao JH, Lahiri J, Isaacs L, Weis RM, Whitesides GM (1998) Science 280:708
172. Sharman GJ, Try AC, Dancer RJ, Cho YR, Staroske T, Bardsley B, Maguire AJ, Cooper MA, O'Brien DP, Williams DH (1997) J Am Chem Soc 119:12041
173. Cooper MA, Williams DH, Cho YR (1997) Chem Commun: 1625
174. Pavlov AY, Preobrazhenskaya MN, Malabarba A, Ciabatti R (1998) J Antibiot 51:525
175. Malabarba A, Ciabatti R, Scotti R, Goldstein BP, Ferrari P, Kurz M, Andreini BP, Denaro M (1995) J Antibiot 48:869
176. He HY, Williamson RT, Shen B, Graziani EI, Yang HY, Sakya SM, Petersen PJ, Carter GT (2002) J Am Chem Soc 124:9729
177. Ruzin A, Singh G, Severin A, Yang YJ, Dushin RG, Sutherland AG, Minnick A, Greenstein M, May MK, Shlaes DM, Bradford PA (2004) Antimicrob Agents Chemother 48:728
178. Walker S, Chen L, Hu YN, Rew Y, Shin DW, Boger DL (2005) Chem Rev 105:449
179. Bulet P, Urge L, Ohresser S, Hetru C, Otvos L (1996) Eur J Biochem 238:64
180. Bulet P, Hegy G, Lambert J, Vandorsselaer A, Hoffmann JA, Hetru C (1995) Biochemistry 34:7394
181. Winans KA, King DS, Rao VR, Bertozzi CR (1999) Biochemistry 38:11700
182. Cudic M, Bulet P, Hoffmann R, Craik DJ, Otvos L (1999) Eur J Biochem 266:549
183. Backes BJ, Virgilio AA, Ellman JA (1996) J Am Chem Soc 118:3055
184. Hammond SM, Claesson A, Jansson AM, Larsson LG, Pring BG, Town CM, Ekstrom B (1987) Nature 327:730
185. Wu B, Tan ZP, Chen G, Chen JH, Hua ZH, Wan Q, Ranganathan K, Danishefsky SJ (2006) Tetrahedron Lett 47:8009
186. Chen JH, Chen G, Wu B, Wan Q, Tan ZP, Hua ZH, Danishefsky SJ (2006) Tetrahedron Lett 47:8013
187. Negri L, Lattanzi R, Tabacco F, Orru L, Severini C, Scolaro B, Rocchi R (1999) J Med Chem 42:400
188. Tomatis R, Marastoni M, Balboni G, Guerrini R, Capasso A, Sorrentino L, Santagada V, Caliendo G, Lazarus LH, Salvadori S (1997) J Med Chem 40:2948
189. Elmagbari NO, Egleton RD, Palian MM, Lowery JJ, Schmid WR, Davis P, Navratilova E, Dhanasekaran M, Keyari CM, Yamamura HI, Porreca F, Hruby VJ, Polt R, Bilsky EJ (2004) J Pharmacol Exp Ther 311:290
190. Polt R, Szabo L, Treiberg J, Li YS, Hruby VJ (1992) J Am Chem Soc 114:10249
191. Fisher JF, Harrison AW, Bundy GL, Wilkinson KF, Rush BD, Ruwart MJ (1991) J Med Chem 34:3140
192. Laupichler L Sowa CE, Theim J (1994) Bioorg Med Chem 2:1281
193. Kansas GS (1996) Blood 88:3259
194. Rosen SD, Bertozzi CR (1994) Curr Opin Cell Biol 6:663
195. Sprengard U, Schudok M, Schmidt W, Kretzschmar G, Kunz H (1996) Angew Chem Int Ed Engl 35:321
196. Baisch G, Ohrlein R (1996) Angew Chem Int Ed Engl 35:1812
197. Bruning J, Kiessling LL (1996) Tetrahedron Lett 37:2907
198. Sprengard U, Kretzschmar G, Bartnik E, Huls C, Kunz H (1995) Angew Chem Int Ed Engl 34:990
199. Matsuda M, Nishimura SI, Nakajima F, Nishimura T (2001) J Med Chem 44:715
200. Wilkins PP, McEver RP, Cummings RD (1996) J Biol Chem 271:18732

201. Li FG, Wilkins PP, Crawley S, Weinstein J, Cummings RD, McEver RP (1996) J Biol Chem 271:3255
202. Leppanen A, Yago T, Otto VI, McEver RP, Cummings RD (2003) J Biol Chem 278:26391
203. Franzyk H, Christensen MK, Jorgensen RM, Meldal M, Cordes H, Mouritsen S, Bock K (1997) Bioorg Med Chem 5:21
204. Vegad H, Gray CJ, Somers PJ, Dutta AS (1997) J Chem Soc Perkin Trans 1:1429
205. Holm B, Linse S, Kihlberg J (1998) Tetrahedron 54:11995
206. Rio-Anneheim S, Paulsen H, Meldal M, Bock K (1995) J Chem Soc Perkin Trans 1:1071
207. Mathieux N, Paulsen H, Meldal M, Bock K (1997) J Chem Soc Perkin Trans 1:2359
208. Pratt MR, Hang HC, Ten Hagen KG, Rarick J, Gerken TA, Tabak LA, Bertozzi CR (2004) Chem Biol 11:1009
209. Kelly WG, Dahmus ME, Hart GW (1993) J Biol Chem 268:10416
210. Hart GW, Greis KD, Dong LYD, Blomberg MA, Chou TY, Jiang MS, Roquemore EP, Snow DM, Kreppel LK, Cole RC, Hayes BK (1995) Pure Appl Chem 67:1637
211. Pohl T, Waldmann H (1997) J Am Chem Soc 119:6702
212. Mazzucco S, Mata S, Vergelli M, Fioresi R, Nardi E, Mazzanti B, Chelli M, Lolli F, Ginanneschi D, Pinto F, Massacesi L, Papini AM (1999) Bioorg Med Chem Lett 9:167
213. Wang LX, Tang M, Suzuki T, Kitajima K, Inoue Y, Inoue S, Fan JQ, Lee YC (1997) J Am Chem Soc 119:11137
214. Cervigni SE, Dumy P, Mutter M (1996) Angew Chem Int Ed Engl 35:1230
215. Agard NJ, Baskin JM, Prescher JA, Lo A, Bertozzi CR (2006) ACS Chem Biol 1:644
216. He Y, Hinklin RJ, Chang JY, Kiessling LL (2004) Org Lett 6:4479
217. Bianchi A, Bernardi A (2006) J Org Chem 71:4565
218. Wang Q, Chan TR, Hilgraf R, Fokin VV, Sharpless KB, Finn MG (2003) J Am Chem Soc 125:3192
219. Agard NJ, Prescher JA, Bertozzi CR (2004) J Am Chem Soc 126:15046
220. Davis BG (2002) Chem Rev 102:579
221. Gamblin DP, Garnier P, van Kasteren S, Oldham NJ, Fairbanks AJ, Davis BG (2004) Angew Chem Int Ed 43:828
222. Rodriguez EC, Winans KA, King DS, Bertozzi CR (1997) J Am Chem Soc 119:9905
223. Rodriguez EC, Marcaurelle LA, Bertozzi CR (1998) J Org Chem 63:7134
224. Liu HT, Wang L, Brock A, Wong CH, Schultz PG (2003) J Am Chem Soc 125:1702
225. Zhang ZW, Gildersleeve J, Yang YY, Xu R, Loo JA, Uryu S, Wong CH, Schultz PG (2004) Science 303:371
226. Xu R, Hanson SR, Zhang ZW, Yang YY, Schultz PG, Wong CH (2004) J Am Chem Soc 126:15654
227. Mahal LK, Yarema KJ, Bertozzi CR (1997) Science 276:1125
228. Dube DH, Bertozzi CR (2003) Curr Opin Chem Biol 7:616
229. Rabuka D, Hubbard SC, Laughlin ST, Argade SP, Bertozzi CR (2006) J Am Chem Soc 128:12078
230. Sawa M, Hsu TL, Itoh T, Sugiyama M, Hanson SR, Vogt PK, Wong CH (2006) Proc Natl Acad Sci USA 103:1237

8.4 Glycosylation Engineering of Glycoproteins

Reiko Sadamoto[1], *Shin-Ichiro Nishimura*[1,2]
[1] Graduate School of Advanced Life Science, Hokkaido University, Sapporo 001–0021, Japan
[2] Drug-Seeds Discovery Research Laboratory, National Institute of Advanced Industrial Science and Technology (AIST), Sapporo 062–8517, Japan
reikosd@glyco.sci.hokudai.ac.jp, shin@glyco.sci.hokudai.ac.jp

Abstract

Naturally occurring glycosylation of glycoproteins varies in glycosylation site and in the number and structure of glycans. The engineering of well-defined glycoproteins is an important technology for the preparation of pharmaceutically relevant glycoproteins and in the study of the relationship between glycans and proteins on a structure-function level. In pharmaceutical applications of glycoproteins, the presence of terminal sialic acids on glycans is particularly important for the in vivo circulatory half life, since sialic acid-terminated glycans are not recognized by asialoglycoprotein receptors. Therefore, there have been a number of attempts to control or modify cellular metabolism toward the expression of glycoproteins with glycosylation profiles similar to that of human glycoproteins. In this chapter, recent methods for glycoprotein engineering in various cell culture systems (mammalian cells, plant, yeast, and *E. coli*) and advances in the chemical approach to glycoprotein formation are described.

In: *Glycoscience*. Fraser-Reid B, Tatsuta K, Thiem J (eds)
Chapter-DOI 10-1007/978-3-540-30429-6_46:

Keywords

Glycoprotein; Glycosylation; CHO cell; Yeast; Plant; *E. coli*; Chemical ligation

Abbreviations

CHO Chinese hamster ovary
NCL native chemical ligation
SPPS solid-phase peptide synthesis

1 Introduction

Naturally occurring glycoproteins vary in glycosylation site as well as in the number and structure of glycans within the molecule, and typical glycan structures often vary from species to species and even between cell types within the same species. However, though glycosylation is essential for stability and activity of many proteins, the precise role of this process has not yet been fully clarified. Nevertheless, the engineering of well-defined glycoproteins is an important technology with two main objectives; the preparation of pharmaceutically relevant glycoproteins and the study of the relationship between glycans and proteins on a structure-function level. There have been a multitude of studies focusing on protein glycosylation; however, in this chapter, we focus on recent advances in the methodology of glycoprotein engineering. Details of pharmaceutical applications and synthesis of biologically relevant glycoproteins can be found in other chapters.

2 Protein Expression in Cells

2.1 Mammalian Cells

The most commonly used method for formation of glycoproteins is through their expression as recombinant proteins in mammalian cell cultures [1]. However, the glycosylation process is affected by many factors, particularly the enzyme repertoire of the host cell and the availability of intracellular sugar-nucleotide donors. Therefore, one of the goals of recombinant glycoprotein formation is to achieve maximum and consistent sialylation of these recombinant glycoproteins. The presence of terminal sialic acids on the glycans is particularly important to the in vivo circulatory half life of glycoproteins, since sialic acid-terminated glycans are not recognized by asialoglycoprotein receptors. Toward this end, there have been a number of attempts to control or modify cellular metabolism to allow the expression of human glycosylation profiles in producer cell lines such as genetically engineered Chinese hamster ovary (CHO) cells.

2.1.1 Using Metabolic Control in Mammalian Cells

With the exception of sialic acid, which is activated as CMP-sialic acid in the nucleus, most nucleotide sugars are synthesized in the cytosol and are unable to penetrate the microsomal

membranes. Thus, nucleotide sugar transporter proteins are essential to the translocation of the nucleotide sugars from the cytosol into the Golgi lumen in which glycosylation proceeds as a series of enzyme-catalyzed reactions.

A major focus in studies on improving glycosylation efficiency is the manipulation of intracellular sugar pools through the addition of nucleotide sugar precursors. *N*-acetyl mannosamine (ManNAc), in particular, has been used to increase intracellular sialic acid pools, resulting in a significant increase in the sialylation of recombinant human interferon-gamma (IFN-γ) produced by CHO cells [2].

However, the augmentation of intracellular sialic acid pools by the introduction of ManNAc was found to be ineffective in increasing the overall level of sialylation of human tissue inhibitor of metalloproteinases 1 (TIMP-1) produced by CHO cells and mouse NS0 cells [3]. Therefore, the overexpression of nucleotide sugar transporters, particularly the CMP-sialic acid transporter (CMP-SAT), was investigated as a means of improving the sialylation process in CHO cells. Increasing cellular CMP-SAT levels by recombinant means results in increased transport of CMP-sialic acid into the Golgi, and subsequent increase in the CMP-sialic acid intra-lumenal pool affords increased sialylation of the proteins expressed [4].

2.1.2 Glycosyl Modification Combined with Metabolic Control by Unnatural Sugars

A metabolic oligosaccharide engineering technique that exploits the tolerance of unnatural substrates by carbohydrate biosynthetic pathways has opened new avenues to the chemical modification of glycoproteins. The introduction of unnatural substrates is limited by the range of substrates that an enzyme will accept. One solution to overcome this limitation is the use of simple enzyme-recognizable sugars possessing functional groups for participation in subsequent chemical reactions. The technique makes use of ketone-containing C2-carbon isosteres of the 2-*N*-acetamidosugars, *N*-acetylglucosamine (GlcNAc) and *N*-acetylgalactosamine (GalNAc) [5,6]. The 2-keto isostere of GalNAc could be incorporated into secreted glycoproteins as well as cell surface molecules allowing the ketone group to react selectively with aminooxy or hydrazide reagents. This technique enables the introduction of reactive groups for the further modification of secreted glycoproteins produced in large-scale recombinant expression systems.

Saxon et al. reported the use of mannosamine and glucosamine derivatives containing azide groups for conversion to cell-surface sialosides [7]. Here, Staudinger ligation [8] is used to label glycoconjugate-bound azidosugars. Cell-surface Staudinger ligation can also be utilized for the formation of hydrazones from metabolically introduced ketones. Both strategies afford an avenue to the selective modification of cell-surface glycans via the use of unnatural sugars [7,9].

2.1.3 Antisense Strategies

Antisense gene transfer technology has been used to limit enzymatic activity in the oligosaccharide biosynthesis pathway, allowing modification of glycans on glycoproteins [10]. Prati et al. use antisense technology to target the fucosyltransferase VI (FucT-VI) enzyme, which is essential to Sialyl Lewis X (sLex) synthesis, and thereby reduce the expression of sLex on the

cell surface [11]. The positive results from this method suggest that the targeting of glycosyltransferase by antisense technology can be applied to the production of novel glycoproteins in CHO cells.

2.2 Baculovirus-Insect Cell Expression Systems

Baculovirus-insect cell expression systems have been used to form glycosylated recombinant glycoproteins [12,13], though this system is incompatible with the formation of terminally sialylated *N*-glycans. The development of transgenic insect cell lines expressing mammalian genes has allowed the expression of "mammalianized" recombinant glycoproteins, since these cell lines are able to salvage sialic acids from extracellular sources, such as proteins found in fetal bovine serum, for glycoprotein sialylation. Transformation of SfSWT-1 cells with two additional mammalian genes encoding sialic acid synthase and CMP-sialic acid synthetase, led to a new insect cell line (SfSWT-3) that was capable of synthesizing its own sialic acid and CMP-sialic acid, and could sialylate recombinant glycoproteins when cultured in a serum-free growth medium supplemented with ManNAc [14]. This new cell line thus offers an improvement on conventional baculovirus-mediated recombinant glycoprotein production systems.

2.3 Yeast Expression Systems

A yeast expression system by itself cannot offer human type glycosylation. However, considering the advantages associated with the use of yeast expression systems, such as chemically defined culture media and shorter expression time, protein expression using yeast systems are expected to be applied to the production of therapeutic glycoproteins. The engineering of an artificial glycosylation pathway has allowed the replication of human *N*-glycosylation in yeast [15,17]. As a recent example for engineering of glycosylation in yeast, Bobrowicz et al. reported that the blocking of core oligosaccharide assembly succeeded in formation of humanized glycoproteins with terminal galactose moieties. This method allowed the synthesis of GlcNAc2Man3GlcNac2 *N*-glycans in yeast strain *Pichia pastoris* in high yield and with high homogeneity [18].

2.4 Plant Expression Systems

As well as insect cell lines, transgenic plant cell lines have also been produced to allow the *N*-glycosylation of therapeutic glycoproteins. The use of transgenic plants offer a number of advantages in the production of therapeutic glycoproteins in that they significantly reduce the potential for contamination with human pathogens, allow for easier genetic manipulation and lower production costs [19].

In order to compensate for the lack of β1,4-galactosylated and sialylated glycans, which play important roles in biological functions in animal cells, Palacpac et al. used mammalian glycosyltransferase to alter the glycosylation pathway of plant cells. The transformed cells containing the full-length human galactosyltransferase gene can be used to produce glycoproteins free from the immunogenic glycans common in plant cells. Using this approach, *N*-linked glycans with modified and defined sugar chain structures approximating those in mammalian

glycoproteins can be synthesized in plants [20]. Further to this approach, the introduction into tobacco suspension-cultured cells of genes encoding human CMP-*N*-acetylneuraminic acid synthetase and CMP-sialic acid transporter opened an avenue to a novel sialic biosynthetic pathway. Recombinant human proteins expressed in this plant cell line system maintain their biological activities, demonstrating that plant cells can be effectively utilized for the production of human glycoproteins [21].

3 Enzymatic Introduction of Glycans onto (Mutated) Proteins Expressed in *E. coli*

Unlike the abovementioned systems, systems based on *E. coli* do not permit the expression of glycoproteins without enzymatic modification. Instead, enzymatic approaches are employed to attach the targeted glycan to the proteins expressed in the cells. The great benefits of this enzymatic approach are (i) that proteins can be readily produced in *E. coli* on a far larger scale than in systems based on mammalian, insect, plant or yeast cells and (ii) easy application of enzymatic reactions in vitro.

Initially, the site-specific modification of chimeric proteins was performed using a guinea pig liver transglutaminase (G-TGase)-mediated procedure. However, more recently, various groups have exploited an alternative approach that utilizes microbial transglutaminase (M-TGase) as the catalyst [22]. An example of this methodology uses microbial transglutaminase (TGase) for the enzymatic preparation of *N*-linked neoglycopeptides from suitably modified oligosaccharides. *N*-allyl glycosidic derivatives are coupled with cysteamine to yield amino-terminated thioether spacers. Transglutaminase then accepts these spacers for the transamidation of the side-chain carboxamide group in the dipeptide Z-Gln-Gly [22].

Sato et al. [23] expanded the applicability of the TGase-mediated preparation of clinical protein conjugates by use of recombinant human interleukin 2 (rhIL-2) as a target protein and two synthetic alkylamine derivatives of poly(ethyleneglycol) (PEG12; MW 12 kDa) and galactose-terminated triantennary glycosides ((Gal)3) as modifiers.

The enzymatic approach using TGase also enabled the site-specific introduction of sialic acid-terminated glycans into proteins. To achieve this, a lactose derivative is introduced by TGase and then sialic acid is transferred to the lactose moiety using 2,6-sialyltransferase (α2,6-SiaT). Using this approach we reported the synthesis of cyclic glycopeptides, toward a tailored blocker of influenza virus hemagglutinin [24] A cyclic peptide, cyclo(Ser-Gly-Gly-Gln-Ser-His-Asp)3, provides an excellent scaffold for the synthesis of a cyclic glycopeptide carrying GM3 oligosaccharides with a potent inhibitory effect on the hemagglutination induced by the influenza virus.

In another example, a glycosylation site was first introduced into insulin by point mutation without loss of biological activity. Next, a lactose derivative was attached to the mutated site by TGase and, finally, sialic acid was transferred to the lactose moiety using 2,6-sialtransferase (α2,6-SiaT) [25] (*Fig. 1*). The results with mice clearly show that the introduction of a sialic acid residue is crucial in prolonging glucose-lowering activity in the blood. Using a similar approach, divalent- and trivalent-sialyloligosaccharide-displaying insulins were also created, with the trivalent sugar-displaying insulin showing further prolongation of the glucose-lowering activity [26].

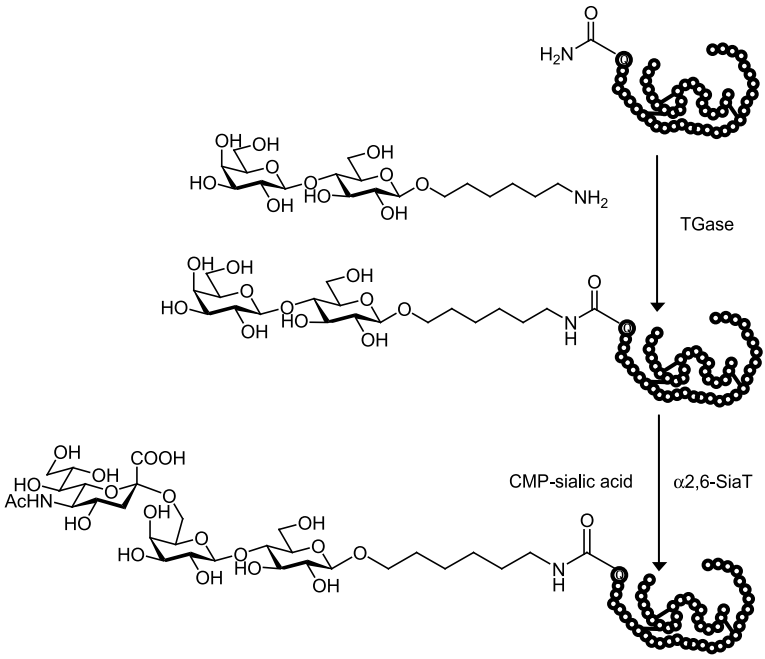


Figure 1
Introduction of sialyllactose into insulin using transglutaminase and sialyltransferase

4 Unnatural Amino Acid Incorporation into Proteins

The presence of a number of enzyme-recognizable amino acid residues, such as glutamine residues in the case of TGase, restricts the site-specific introduction of glycans. When the target protein has many glutamine residues, site specific modification can not be performed artificially. The introduction of an unnatural amino acid, however, allows the exact glycosylation site to be determined.

For the introduction of unnatural amino acids, several unnatural codons are used; for example, the stop (amber) codon and 4-base (or 5-base) codon [27,29]. An example of site-specific incorporation of an unnatural amino acid containing a keto group into a protein using the amber nonsense codon was reported by Liu et al. [30]. The introduced keto group can react specifically with an aminooxy saccharide derivative to give homogeneous glycoprotein mimetics.

We also demonstrated the direct and chemoselective coupling between unmodified simple sugars and *N*-methylaminooxy group displayed on the engineered streptavidin using the 4-base codon, which allowed for the combinatorial synthesis of novel glycoprotein mimetics (*Fig. 2*) [31].

In another study, proteins containing both mono- and diglycosylated amino acids, including glycosylated serine and tyrosine moieties, have been obtained by the suppression of nonsense codons in a cell-free expression system through the application of misacylated suppressor tRNAs activated with glycosylated serine and tyrosine derivatives (*Fig. 3*) [32].

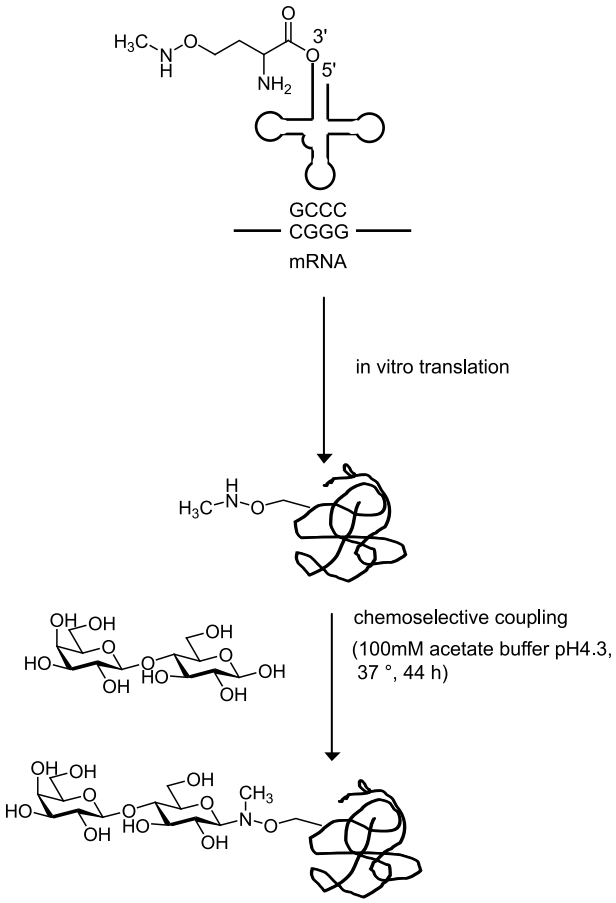


Figure 2
Chemoselective coupling of lactose and an oxylamino-functionalized amino acid residue

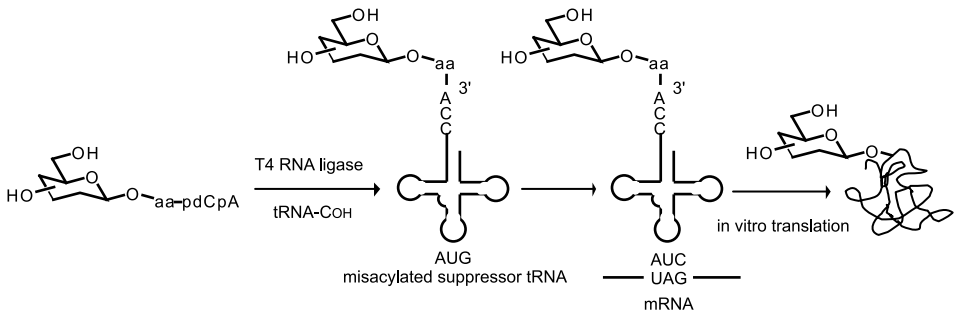


Figure 3
Cell-free expression of protein containing a glycosylated amino acid

5 Chemical Ligations and Synthesis

Whereas relatively short peptides can be readily obtained by solid-phase peptide synthesis (SPPS), native chemical ligation (NCL) [33] has now enabled the total chemical synthesis of proteins with molecular weights far in excess of those achievable by conventional stepwise SPPS. This method proceeds through the ligation of a C-terminal thioester and an *N*-terminal cysteine residue, respectively, by two unprotected peptides to give a protein with a native amide linkage at the ligation site. However, glycopeptides synthesis using NCL is problematic since the reaction conditions used in tert-butoxycarbonyl(Boc)-based SPPS is known to damage glycosidic linkages and those used in 9-fluorenylmethoxycarbonyl (Fmoc)-based SPPS to cleave the thioesters of *O*-linked glycans.

To overcome the problems associated with these reaction conditions, Backes et al. developed an alkanesulfonamide "safety-catch" linker for solid-phase synthesis [34]. In this method, acylation of a sulfonamide support affords a support-bound *N*-acylsulfonamide that is able to withstand the basic and strongly nucleophilic reaction conditions required for Fmoc-based SPPS. On completion of the solid-phase synthesis sequence, iodoacetonitrile treatment yields *N*-cyanomethyl derivatives that can be cleaved under mild nucleophilic reaction conditions to afford the target compounds. Coupling conditions of alkanesulfonamide resin with Boc- and Fmoc-amino acids were developed to minimize the racemization. Using this method, a short synthesis sequence, followed by iodoacetonitrile activation and nucleophilic displacement is able to form dipeptide products from a number of support-bound amino acids incorporating diverse side-chain functionalities.

Based on this method, Shin et al. reported the total synthesis of an 82-residue glycoprotein with two glycosylation sites through the ligation of a 24-residue glycopeptide-thioester, prepared using standard Fmoc chemistry protocols, with a 58-residue glycopeptide bearing an *N*-terminal cysteine residue (*Fig. 4*) [35].

The expressed chemical ligation method was developed for the total synthesis or recombinant formation of proteins displaying natural or unnatural post-translational modifications. This method is based on the construction of semi-synthetic glycoproteins, in which recombinantly derived peptide α-thioesters are covalently joined by NCL with synthetic glycopeptides. Macmillan et al. applied an extension of this technique to the construction of glycoproteins bearing mucin-like domains [36]. Macmillan et al. also reported a cysteine-free NCL method. The use of mild conditions for cleavage reaction and rapid synthesis in this method is useful especially for the preparation of glycoproteins [37].

Chemical synthesis of glycopeptides and glycoproteins based on NCL also allows the synthesis of glycopeptides from cysteine-free peptides. This method utilizes a peptide thioester and a glycopeptide in which a thiol handle is engineered into the sugar moiety at the C-2 position. Once the ligation reaction is complete, reduction of the thiol handle with H_2/metal to the acetamide moiety affords unmodified glycopeptides [38].

Dudkin et al. describe the synthesis of globally deprotected multi-branched glycans and these glycans are aminated prior to aspartylation with a hexapeptide, which is then extended using NCL to afford *N*-linked glycopeptides for use as possible antigens against prostate cancer (*Fig. 5*) [39].

Aziridine-incorporated peptides, capable of site- and stereoselective modification, have been formed using FmocAzyOH in Fmoc solid-phase peptide synthesis. This nonproteinogenic

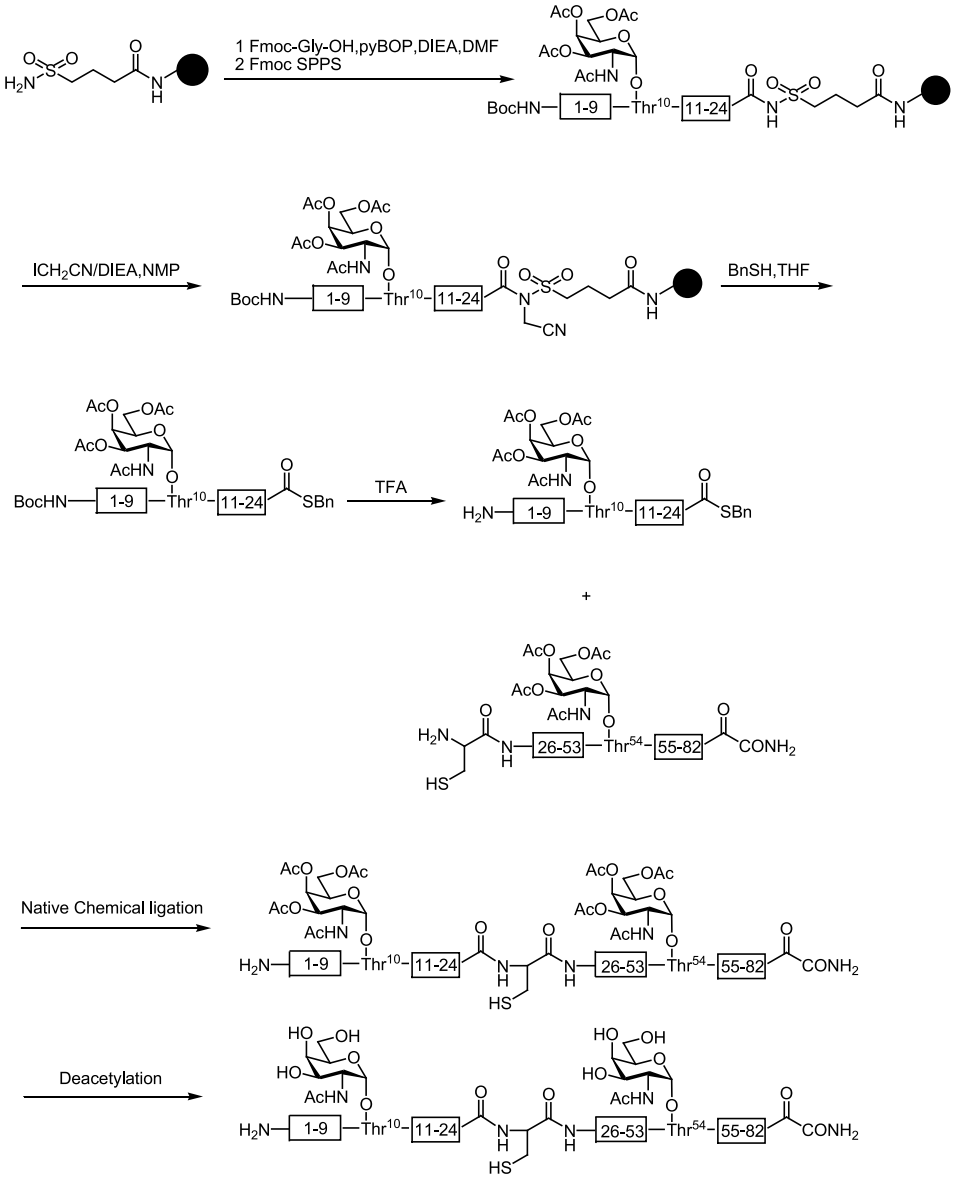


Figure 4
Chemical synthesis of a glycoprotein via native chemical ligation

amino acid possesses a unique electrophilic nature that enables site-selective conjugation with a variety of thiol nucleophiles, such as anomeric carbohydrate thiols, farnesyl thiol, as well as biochemical tags. The combination of this strategy with NCL offers a convergent and rapid avenue to the generation of complex thioglycoconjugates (*Fig. 6*) [40,41].

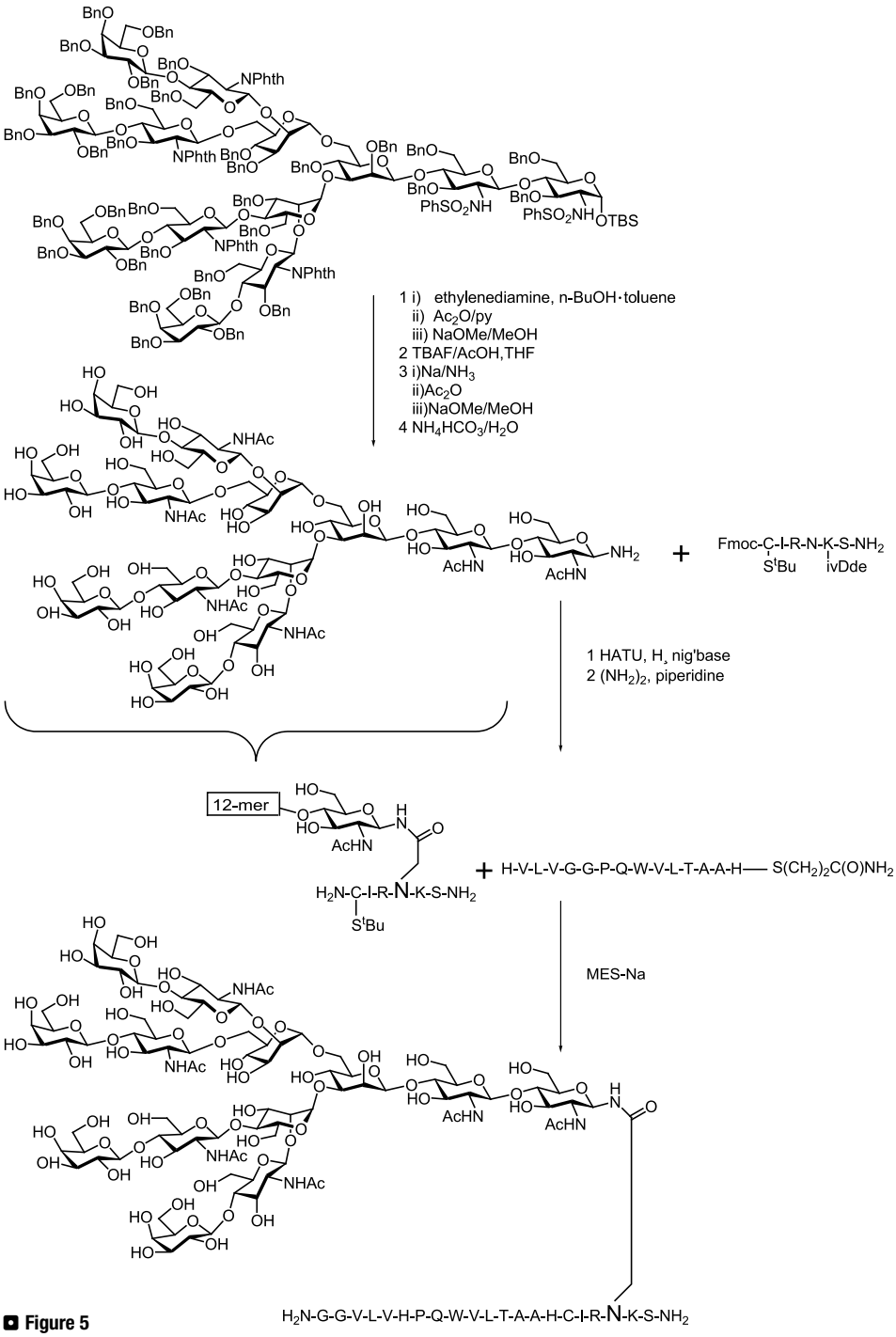


Figure 5
Synthesis of multi-carbohydrate-branched glycopeptides

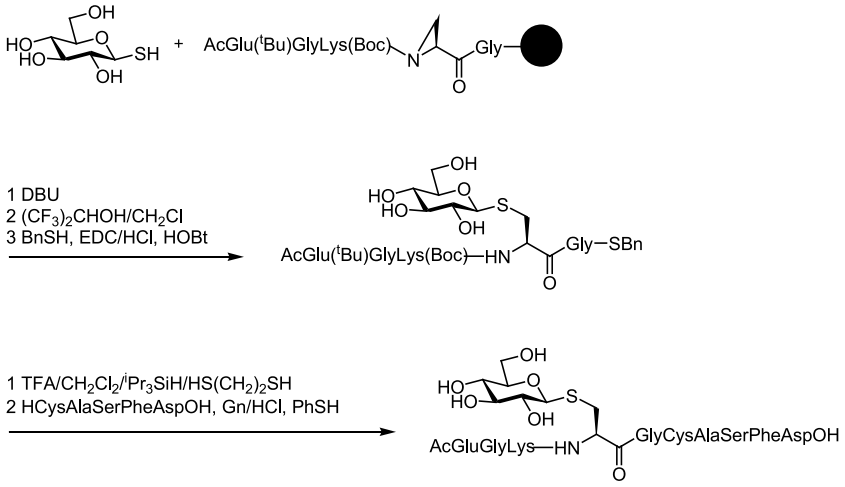


Figure 6
Synthesis of a thioglycoconjugate using aziridine-incorporated peptide

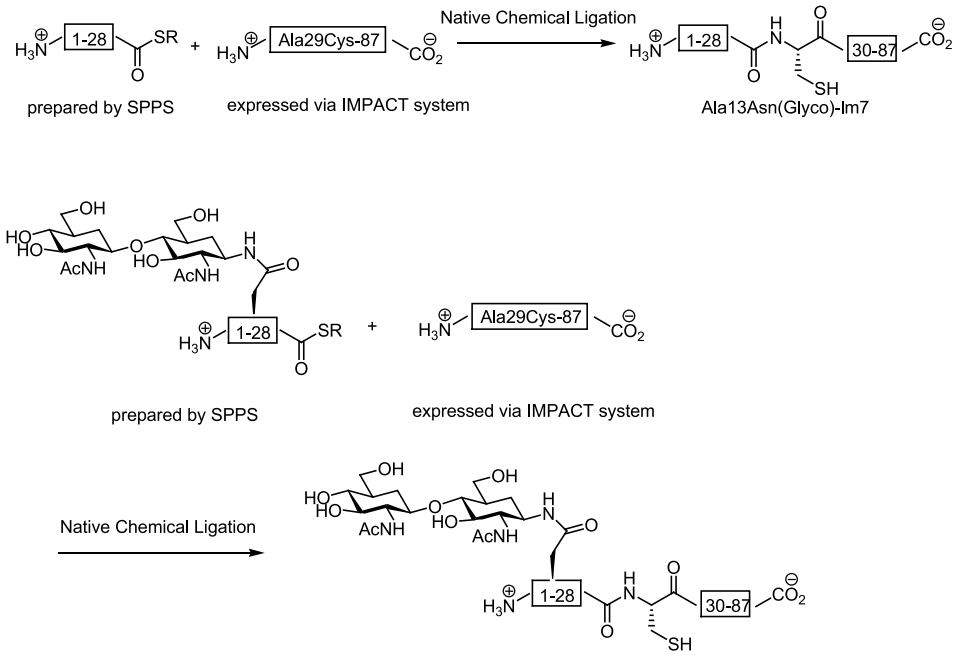


Figure 7
Glycosylated and unglycosylated analogs of Im7 prepared by native chemical ligation

Chemical synthesis allows the preparation of proteins with different glycosylation patterns aiding the study of the effects of glycosylation on protein folding. To examine the influence of glycosylation on protein folding pathways, Hackenberger et al. utilized NCL in the preparation of *N*-linked chitobiose glycoprotein analogues of the immunity protein Im7. The NCL method afforded both glycosylated and unglycosylated analogs, enabling further study of the influence of glycosylation on peptide structure (◉ *Fig. 7*) [42].

6 Summary

The use of glycosylation engineering in various expression systems has aided the move toward "human-type" glycosylation and efficient sialylation. The incorporation of unnatural amino acids into the protein back bone with site specificity has also enabled the precise control of glycosylation sites. Further, chemical modification of expressed proteins or peptides and/or enzymatic glycan elongation has been used to solve problems of length- and quantity limitation in the expression of glycoproteins.

References

1. Butler M (2006) Cytotechnology 50:57
2. Gu X, Wang DI (1998) Biotechnol Bioeng 58: 642
3. Baker KN, Rendall MH, Hills AE, Hoare M, Freedman RB, James DC (2001) Biotechnol Bioeng 73:188
4. Wong NSC, Yap MGS, Wang DIC (2006) Biotechnol Bioeng 93:1005
5. Hang HC, Bertozzi CR (2001) J Am Chem Soc 123:1242
6. Hang HC, Bertozzi CR (2001) Acc Chem Res 34:727
7. Saxon E, Bertozzi CR (2001) Annu Rev Cell Dev Biol 17:1
8. Kohn M, Breinbauer R (2004) Angew Chem Int Ed 43:3106
9. Saxon E, Luchansky SJ, Hang HC, Yu C, Lee SC, Bertozzi CR (2002) J Am Chem Soc 124:14893
10. Warner TG (1999) Glycobiology 9:841
11. Prati EGP, Scheidegger P, Sburlati AR, Bailey JE (1998) Biotechnol Bioeng 59:445
12. Marchal I, Jarvis DL, Cacan R, Verbert A (2001) Biol Chem 382:151
13. Harrison RL, Jarvis DL (2006) Adv Virus Res 68:159
14. Aumiller JJ, Hollister JR, Jarvis DL (2003) Glycobiology 13:497
15. Wildt S, Gerngross TU (2005) Nat Rev Microbiol 3:119
16. Gerngross TU (2004) Nat Biotechnol 22:1409
17. Wu JM, Lee CK, Hsu TA (2002) Cell Eng 3: 215
18. Bobrowicz P, Davidson RC, Li H, Potgieter TI, Nett JH, Hamilton SR, Stadheim TA, Miele RG, Bobrowicz B, Mitchell T, Rausch S, Renfer E, Wildt S (2004) Glycobiology 14: 757
19. Hellwig S, Sack M, Spiegel H, Drossard J, Fischer R (2003) Am Biotechnol Lab 21:50
20. Palacpac NQ, Yoshida S, Sakai H, Kimura Y, Fujiyama K, Yoshida T, Seki T (1999) Proc Natl Acad Sci USA 96:4692
21. Misaki R, Fujiyama K, Seki T (2006) Biochem Biophys Res Commun 339:1184
22. Ramos D, Rollin P, Klaffke W (2001) J Org Chem 66:2948
23. Sato H, Hayashi E, Yamada N, Yatagai M, Takahara Y (2001) Bioconjugate Chem 12:701
24. Ohta T, Miura N, Funitani N, Nakajima F, Niikura K, Sadamoto R, Guo CT, Suzuki T, Suzuki Y, Monde K, Nishimura SI (2003) Angew Chem Int Ed 42:5186
25. Sato M, Sadamoto R, Niikura K, Monde K, Kondo H, Nishimura SI (2004) Angewandte Chemie, Int Ed 43:1516
26. Sato M, Furuike T, Sadamoto R, Fujitani N, Nakahara T, Niikura K, Monde K, Kondo H, Nishimura S (2004) J Am Chem Soc 126:14013

27. Wang L, Xie J, Schultz PG (2006) Annu Rev Biophys Biomol Struct 35:225
28. Xie J, Schultz PG (2005) Curr Opin Chem Biol 9:548
29. Hohsaka T, Sisido M (2002) Curr Opin Chem Biol 6:809
30. Liu H, Wang L, Brock A, Wong CH, Schultz PG (2003) J Am Chem Soc 125:1702
31. Matsubara N, Oiwa K, Hohsaka T, Sadamoto R, Niikura K, Fukuhara N, Takimoto A, Kondo H, Nishimura SI (2005) Chem Eur 11:6974
32. Fahmi NE, Dedkova L, Wang B, Golovine S, Hecht SM (2007) J Am Chem Soc 129:3586
33. Dawson PE, Muir TW, Clark-Lewis I, Kent SB (1994) Science 266:776
34. Backes BJ, Ellman JA (1999) J Org Chem 64:2322
35. Shin Y, Winans KA, Backes BJ, Kent SBH, Ellman JA, Bertozzi CR (1999) J Am Chem Soc 121:11684
36. Macmillan D, Bertozzi CR (2000) Tetrahedron 56:9515
37. Macmillan D, Anderson DW (2004) Org Lett 6:4659
38. Brik A, Yang YY, Ficht S, Wong CH (2006) J Am Chem Soc 128:5626
39. Dudkin VY, Miller JS, Danishefsky SJ (2004) J Am Chem Soc 126:736
40. Galonic DP, van der Donk WA, Gin DY (2004) J Am Chem Soc 126:12712
41. Galonic DP, Ide ND, van der Donk WA, Gin DY (2005) J Am Chem Soc 127:7359
42. Hackenberger CPR, Friel CT, Radford SE, Imperiali B (2005) J Am Chem Soc 127:12882

8.5 Glycoprotein Analysis

Daryl Fernandes, *Daniel Spencer*
Ludger Ltd., Culham Science Centre, Abingdon, Oxfordshire OX14 3EB, UK
daryl.fernandes@ludger.com, daniel.spencer@ludger.com

Abstract

This chapter provides an overview of practical methods for glycosylation analysis of glycoprotein therapeutics. The topics include glycoprofiling methods for glycoforms, monosaccharides (neutral and *N*-acetylated species as well as sialic acids), oligosaccharides (chemical and enzymatic methods for glycan release, post-release purification, labeling and derivatization, different types of glycan HPLC and MS), and glycosylation site profiling.

Keywords

Glycoprofiling; Biopharmaceutical glycosylation

Abbreviations

2-AA	2-aminobenzoic acid
2-AB	2-aminobenzamide
AA-Ac	3-(acetylamino)-6-aminoacridine
AP	2-aminopyridine
APTS	8-aminopyrene-1,3,6-trisulfonate
CE	capillary electrophoresis
EPO	erythropoeitin
DMB	1,2-diamino-4,5-methylenedioxybenzene dihydrochloride

In: *Glycoscience*. Fraser-Reid B, Tatsuta K, Thiem J (eds)
Chapter-DOI 10-1007/978-3-540-30429-6_47:

EIR	electron interaction resin
GC	gas chromatography
GHP	glucose homopolymer
HILIC	hydrophilic interaction liquid chromatography
HPAE-PAD	high pH anion exchange with pulse amperometric detection
HPAE-FD	high pH anion exchange with fluorescence detection
HPLC	high performance liquid chromatography
mAbs	monoclonal antibodies
MS	mass spectrometry
NeuAc	*N*-acetyl-neuraminic acid
NeuGc	*N*-glycolyl-neuraminic acid
PAT	process analytical technology
SPE	solid phase extraction
QC	quality control

1 Introduction

This chapter focuses on procedures for glycosylation analysis of recombinant therapeutic glycoproteins such as monoclonal antibodies (mAbs) and erythropoietin (EPO). For these and many other biopharmaceuticals, glycosylation can profoundly influence drug safety and efficacy and so must be measured and tightly controlled during biomanufacturing. However, measurement is not always straightforward since current approved glycoprotein therapeutics are invariably highly heterogeneous populations of glycoforms, often bearing complex glycosylation and with multiple glycosylation sites. Most of these drugs have *N*-linked oligosaccharides and some have both *N*- and *O*-linked glycans. The regulations governing biopharmaceutical characterization are strict and glycoprofiling of therapeutics must be performed using well-characterized analytical procedures and be demonstrated to be accurate and reproducible both over time and between different laboratories. Accuracy in relative quantitation of the many glycan species found in each drug batch is particularly important so that biomanufacturers can reliably demonstrate comparability with reference batches. The procedures used for glycan characterization of biopharmaceuticals can be extended to analysis of other glycoproteins.

2 Glycoprofiling Scheme

Figure 1 shows a typical biopharmaceutical glycoprofiling scheme consisting of a number of analysis modules each of which measures a different glycosylation parameter of the glycoprotein drug sample. The data from these modules are used to build a forensic profile of the glycosylation [1].

With current technology, there is no single universal glycoanalysis technique–the complexity of biopharmaceutical glycosylation makes it imperative to use several orthogonal analysis methods.

The choice of analysis modules employed in a glycoprofiling scheme and the procedures used in each module is influenced by many factors including what types of glycans are possible in

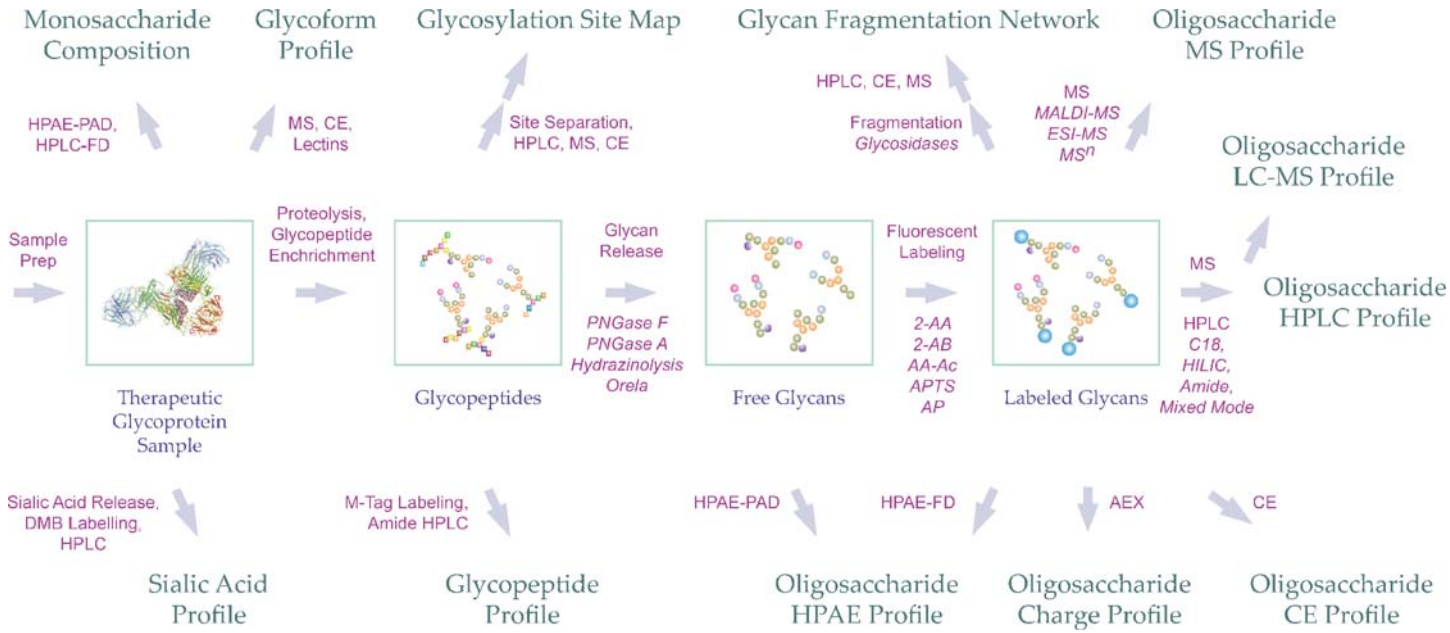


Figure 1
A typical biopharmaceutical glycoprofiling scheme

the glycoprotein expression system, the desired glycosylation pattern, aberrant glycosylation that could compromise drug safety or efficacy and the analytical instrumentation available.

3 Glycoform Profile

The glycoform profile is a map of the glycoprotein sub-populations that share the same peptide backbone but are decorated with different glycan structures. Detailed glycoform profiling remains a challenge but valuable low-resolution maps can be obtained by using a variety of methods. For some glycoproteins, capillary electrophoresis (CE) can give useful separation of the glycoforms [2] while Rosenfeld et al. [3] have employed lectin and antibody arrays to detect changes in glycoform distributions during biopharmaceutical production.
Both MALDI-MS and ESI-MS can be used to obtain mass distribution profiles of intact glycoproteins. The main experimental challenges here are (a) to get uniform and linear responses in MS by every glycoform and (b) to achieve sufficient accuracy and resolution of complex mixtures of such large molecules. The problem of overlapping of isobaric glycoforms is reduced by combining MS of intact glycoproteins with analysis of the glycopeptides or oligosaccharides. Glycoform profiling by MS is now routinely used by some biomanufacturers as part of a process analytical technology (PAT) system to monitor the quality of glycoprotein therapeutics during production.

4 Monosaccharide Profile

For biopharmaceuticals, quantitative monosaccharide profiling allows detection of residues indicative of undesired glycosylation (e. g. NeuGc) or the classes of glycans (for example, the presence of GalNAc indicates *O*-glycosylation). For technical reasons, neutral and *N*-acetylated residues are generally profiled separately from sialic acids.

4.1 Neutral and *N*-Acetylated Monosaccharides

The general form for neutral and *N*-acetylated monosaccharide profiling is: release of the monomer residues by acid hydrolysis, derivatization then analysis by HPLC, CE, or GC. In practice, reliable monosaccharide quantitation is an issue. In a study comparing monosaccharide profiles of glycoprotein standards from several laboratories, Townsend et al. [4] found that values of neutral and amino sugars determined using HPAE-PAD analyses varied by up to 50% between different labs. In our experience, the keys to obtaining accurate, reproducible profiles are (a) ensuring clean, stoichiometric release and derivatization, (b) using relevant, well-characterized quantitative internal and external standards and (c) choosing an analysis system that gives good resolution of all the monosaccharide species found in the analyte. The following method has proven to be suitable for use with most biopharmaceuticals:

1. Addition of 3-*O*-methyl-glucose as a quantitative internal standard.
2. Use of quantitative external standards matched to the analytes–e. g. fetuin or antibody for glycoprotein analyses and the equivalent glycan libraries for oligosaccharide analyses.

3. Release of monosaccharide residues by acid hydrolysis (typically using 4 mol dm^{-3} trifluoroacetic acid, 100 °C, 4 h). This will liberate hexoses and de-acetylated hexosamines.
4. Re-*N*-acetylation of the amino sugars using acetic anhydride in aqueous sodium bicarbonate buffer.
5. De-salting on a cation exchange resin (e. g. AG50WX12 H^+ form).
6. Derivatization by reductive amination with a small hydrophilic fluorophore or chromophore such as 2-AA (see below).
7. HPLC with fluorescence or UV detection either on a C18 column using a gradient of acetonitrile / sodium acetate (aq) or on HPAE-FD.

This method is not suitable for profiling of sialic acids, which are destroyed by the harsh acid hydrolysis and cannot be labeled by reductive amination, but does give reliable data for neutral monosaccharides.

4.2 Sialic Acids

The two main types of sialyl residues found in recombinant biopharmaceuticals are NeuAc and NeuGc. These usually occur as terminal structures attached to galactose residues at the non-reducing termini of both *N*- and *O*-linked glycans. Controlling the ratio of NeuAc to NeuGc is critical for biomanufacturers–NeuAc is the desired, normal human-type sialylation while NeuGc is found in non-human glycoproteins and is considered an undesired form of sialylation for injected therapeutics [5].

Our preferred method of sialic acid profiling is to use the DMB / HPLC system. The first step is release of the sialic acids from the glycoprotein by mild acid hydrolysis (typical conditions are incubation with 2 mol dm^{-3} acetic acid for 2 h at 80 °C) followed by fluorescent labeling using DMB and stabilization of the tagged conjugates by reduction with sodium dithionite. Analysis of the DMB labeled sialic acids is done by HPLC with a C18 column and fluorescence detection (λ_{ex} = 295 nm, λ_{em} = 352 nm).

With DMB analysis (as with all glycoprofiling) good calibration is essential for accurate, reliable results. For NeuAc:NeuGc determinations we recommend calibration using three standards:

a. A sialic acid reference panel containing a mixture of sialic acids found in human and animal tissue.
b. A quantitative NeuAc standard. Serial dilutions of this are used to obtain samples for a standard concentration- fluorescence response curve on the HPLC.
c. A qualitative NeuGc standard. This can either be used alone or can be added to aliquots of the NeuAc standard for calibration of retention on HPLC.

Alternatives to DMB analysis include HPLC of OPD (o-phenylenediamine) labeled sialic acids [6] as well as HPAE-PAD and hybrid variants of these procedures.

5 Oligosaccharide Profile

The general form for oligosaccharide profiling is: release of the glycans, derivatization (e. g. fluorescent labeling), then analysis by HPLC, CE, or MS.

5.1 Glycan Release

Glycan release can be achieved either enzymatically with endoglycosidases (e. g. PNGase F or PNGase A) or chemically. **PNGase F** is widely used for enzymatic release of *N*-glycans from therapeutic glycoproteins [7]. Sample preparation prior to PNGase F treatment typically involves disruption of the protein backbone with a chaotropic agent then reduction and alkylation to open disulfide bridges. The aim of this is to make the oligosaccharide-protein linkages more accessible to the enzyme. However, in some cases (e. g. with certain mAbs), better release kinetics and stoichiometry are obtained with the intact glycoprotein. One possible reason for this is steric hindrance of the glycan attachment sites after disruption of the protein. Various preparations of both natural and recombinant forms of PNGase F are available but commercial preparations do vary significantly in quality (particularly in their stability and levels of contamination). In our experience, the natural enzyme is generally more stable and works better in simple, clean formulations than the recombinant form. This may be related to alterations to the natural protein structure required for expression and isolation of the recombinant enzyme.
PNGase A is used for release of *N*-glycans from glycopeptides. It has a broader oligosaccharide substrate specificity than PNGase F and is more suitable for analysis of glycopeptides from plant and insect expression systems. PNGase A is particularly useful for analysis of tryptic digests of glycoproteins from proteomics gels [8].
Currently, there are no broad-specificity endoglycosidases for *O*-linked glycans. To liberate these one must turn to chemical release methods which include hydrazinolysis and beta-elimination.
Beta elimination is a well-established and widely used method for releasing *O*-glycans from glycoproteins. The process gives high yields, however, the glycans are reduced to alditols. These cannot be labeled by reductive amination and so are usually analyzed by techniques such as HPAE-PAD or MS which can give good structural information but generally give less reliable quantitation data than fluorescent HPLC [9].
Hydrazinolysis involves use of anhydrous hydrazine to break the sugar-peptide links and release free glycans in the reducing form. These are suitable for fluorescent labeling by reductive amination which makes hydrazinolysis the method of choice for a number of glycoprofiling laboratories. A typical protocol is incubation of the dried, salt-free glycoprotein with anhydrous hydrazine (85–100 °C, 6–16 h) under an inert, dry atmosphere, removal of excess hydrazine by evaporation, *N*-acetylation with acetic anhydride in sodium bicarbonate buffer, de-salting on a cation exchange resin in H^+ form then purification of the released oligosaccharides. Optimization of the temperature-time profile of the hydrazine incubation step allows tuning to release either *O*-glycans (under mild conditions), *N*-glycans (harsher conditions) or a mixture of both. Successful hydrazinolysis requires use of very pure hydrazine with low water content (<0.1%) to minimize 'peeling' side reactions (i. e. progressive removal of reducing terminus residues). However, when performed with good quality reagents, hydrazinolysis is capable of giving very reproducible glycan release with high yield and good stoichiometry [10].

5.2 Post-Release Purification

Whichever method is used for glycan detachment, the released oligosaccharides should be purified to remove non-carbohydrate material that could interfere with subsequent derivatization and analysis. Post-release methods described in the literature include precipitation of peptide material in cold ethanol, retention of peptides on hydrophobic resins or precipitation of the oligosaccharides on HILIC resins. The most effective post-release purification that we have seen for biopharmaceutical oligosaccharides has been use of SPE cartridges containing electron interaction resin (EIR). EIR is a versatile matrix that binds a wide range of biopharmaceutical glycans with high efficiency, allowing effective removal of contaminating salts, detergents, and protein. The washing regime for EIR cartridges are similar to those for C18 SPE purification of hydrophobic analytes with glycans being eluted in a mixture of volatile components. The benefits of glycan purification using EIR include higher efficiencies and better stoichiometry for fluorescent labeling and cleaner profiles and better signal to noise on HPLC and MS. Furthermore, microplate SPE versions of EIR are currently being developed for high-throughput analyses.

5.3 Labeling and Derivatization

Fluorophore or chromophore labeling of the purified oligosaccharides allows quantitative detection on HPLC and CE and can improve analysis of certain glycan species on MS. The most common labels are aromatic amines which are attached to the reducing termini by reductive amination. These include 2-AA and 2-AB which are small and hydrophilic, conjugate to the glycan efficiently and have good spectral properties. Furthermore, they have minimal effect on the specificities of glycosidases used in enzymatic glycan sequencing. A typical one-pot labeling reaction involves dissolution of purified, dried glycans with free reducing termini in a solution of 1 mol dm^{-3} of fluorescent label and 1 mol dm^{-3} sodium cyanoborohydride in 30% acetic acid and 70% dimethyl sulfoxide (v/v) followed by incubation at 65 °C for 3 h [11]. Post-reaction purification of the labeled glycans can be achieved using *S*-cartridges which contain a hydrophilic glycan binding membrane. The labeling mixture is applied to the membrane then air-dried. Excess label is removed by washing with 96% acetonitrile (aq) then the bound glycans are eluted with water. *S*-cartridges give very reproducible post-labeling cleanup but can be slow. However, faster SPE alternatives including those in microplate format for high throughput work are becoming available.

2-AA and 2-AB are routinely used for characterization of biopharmaceutical glycosylation at all stages of the drug development cycle, including glycoprofiling during production and final lot release tests for approved drugs. For most applications the two labels can be used interchangeably, giving similar profiles and reliable performance on amide HPLC (see below). The main differences are seen in anion exchange HPLC where 2-AA labeled glycans are retained longer than those tagged with 2-AB (2-AA bears a negative charge while 2-AB has no charge) and performance in mass spectrometry.

2-AP is another common glycan label and is especially popular with academic research groups in Japan. Takahashi and colleagues have developed a sophisticated system for multi-dimensional mapping of 2-AP labeled oligosaccharides on orthogonal chromatography phases (including amide, C18 and anion exchange HPLC) [12].

APTS is a label used for glycoprofiling by CE. The molecule bears three strong anionic groups which allows fast transit through the CE column towards the cathode [13].
AA-Ac is a versatile label which is gaining popularity for high throughput glycoprofiling and glycoproteomics. The molecule has an acridine ring structure that improves HPLC on hydrophobic phases, enhances analysis by MALDI–MS and electrospray MS and allows good separation on CE. Furthermore, the labeling reaction is faster than other labels (typically 1 h), post-labeling purification is fast and efficient and stoichiometry and efficiency seem excellent [14].

5.4 HPLC

HPLC is a key glycoprofiling tool that allows partial identification of species by their retention time, quantitation from peak areas, and separation of complex glycan mixtures into different fractions that can be taken for further analysis. Typically, a range of different HPLC columns with a variety of separation mechanisms is used to slice up the glycan population according to different physical properties such as molecule size and charge. The following describes the main HPLC profiles currently used for biopharmaceutical glycoprofiling.
High performance anion exchange (**HPAE**) chromatography relies on deprotonation of the sugar hydroxyl groups in a high pH buffer to yield anion species. These can then be separated on a high pH stable, anion exchange column. A typical washing regime is equilibration of the column in 150 mmol dm^{-3} sodium hydroxide and elution of glycans of increasingly high charge to mass ratio with a gradient of 0–250 mmol dm^{-3} sodium acetate (aq). Eluted glycans can be followed using pulsed amperometric detection (PAD) which is a specialized electrochemical detector. The system is capable of excellent resolution of glycan species, does not require derivatization of the sugars, and can be used for profiling both monosaccharides and large, highly charged oligosaccharides. In the past, for many labs, these factors have made **HPAE-PAD** the method of choice for biopharmaceutical glycoprofiling. However, reproducibility of quantitation is a concern–different PAD cells give different relative responses to standard sugars so HPAE-PAD profiles from different instruments are not always comparable. Fortunately, this problem can be overcome by labeling the glycans with a small, noncharged hydrophilic fluorophore such as 2-AB and switching to **HPAE-FD**. This combines the superb resolving power of HPAE chromatography with the reliability of 2-AB fluorescence quantitation [15]. HPAE-FD is now routinely used by some biomanufacturers for in-process checking and final lot release QC of biopharmaceutical glycosylation.
Glycan charge profiles can be determined by weak or strong anion exchange chromatography (**WAX** or **SAX**). At neutral pH, fluorophore labeled glycans, typically 2-AB label, will separate on a strong anion exchange column according to their charge state. Uncharged glycans will pass straight through the column while glycans bearing an anionic charge (e. g. sialylation or phosphorylation) will be retained and can be eluted with a salt gradient (typically 0–500 mmol dm^{-3} ammonium acetate, pH 4.5). This technique allows determination of the relative ratios of charged glycans in any given sample, but generally gives little information on the detailed structure of the glycans.
EIR HPLC allows reproducible separation of fluorophore labeled glycans with high resolution. The solvents and gradient regime are similar to those used for analysis of hydrophobic

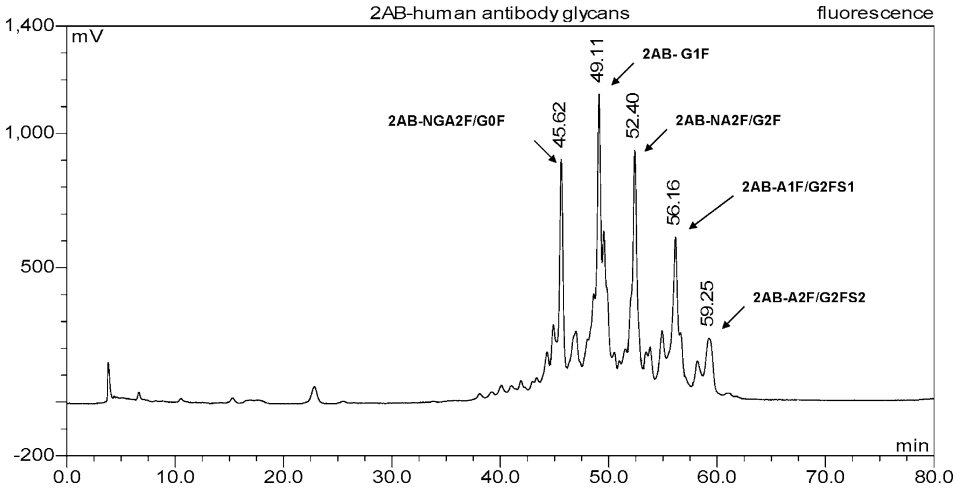


Figure 2
LudgerSepN1-amide HPLC of 2AB labeled human antibody glycans

analytes on C18 HPLC except that, generally, higher concentrations of the organic solvent–typically up to 60% (v/v) acetonitrile in 0.1% trifluoroacetic acid (aq)–are used to elute the glycans. EIR columns have excellent retention of a broad range of glycans and glycan derivatives including hydrophilic species that are not retained by C18 resins.

Hydrophilic interaction chromatography (**HILIC**) on **amide** columns (sometimes called 'normal phase' HPLC) is now widely used for detailed oligosaccharide profiling of biopharmaceuticals. The technique relies on interactions of the glycan hydroxyls with amide groups on the chromatography resin in a gradient of decreasing organic solvent. The separation roughly follows the size of the glycan, larger glycans having a greater number of hydroxyl groups. A typical elution regime is a gradient from 65–35% (v/v) acetonitrile in 250 mM ammonium formate, pH 4.4 (aq) over 75 minutes. *Figure 2* shows a typical amide HPLC profile of glycans from human serum antibody. To allow comparison of profiles between different HPLC runs, the retention times of glycans can expressed in terms of glucose units (GU) by comparison with an external standard of 2-AB labeled glucose homopolymer (GHP) ladder which contains chains from 1 to 23 glucose residues. Tables of glycan GU values are available for oligosaccharides found on monoclonal antibodies and other biopharmaceuticals [16]. Glycan amide HPLC is capable of giving excellent resolution and reproducibility and can be performed on oligosaccharides tagged with many types of fluorophore label including AA-Ac, 2-AB, 2-AA and 2-AP. Furthermore, the separation is sensitive to small changes in glycan structure allowing resolution of closely related and isobaric oligosaccharides and the solvents are volatile making it an excellent partner to detailed structural analysis by mass spectrometry.

5.5 MS

The details of glycan analysis by mass spectrometry are presented elsewhere in this book. Here, we concentrate on the compatibility of glycan fluorophore labeling and HPLC with MS.

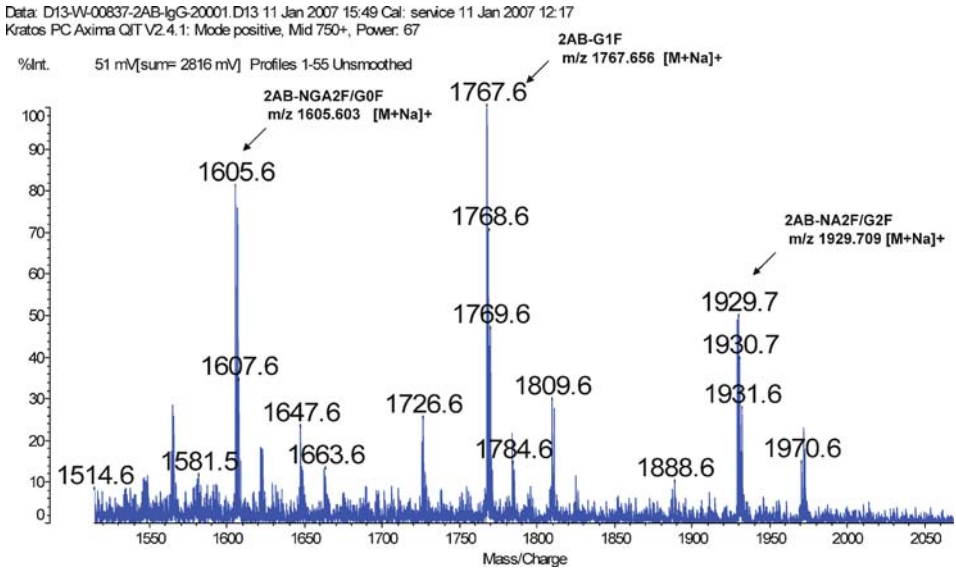


Figure 3
MS mass spectrum of 2AB labeled human antibody glycans. Axima-QIT, positive ion mode, DHB matrix

As well as aiding glycan detection, labeling with fluorophores such as 2-AB can significantly enhance ionization in a mass spectrometer. This is particularly useful when dealing with anionic oligosaccharides which can be particularly difficult to analyze by MS, suffering from severe signal attenuation by non-glycan contaminants such as salts and peptides. *Figure 3* shows a typical MS profile of 2-AB labeled glycans from human serum antibody.

Of the glycan fluorophores available, one of the most useful for HPLC and MS is AA-Ac. The structure of this label allows very efficient ionization of a broad range of glycans including highly sialylated species which often give poor signals on MS. This increases sensitivity and will typically allow femtomol amounts of glycan to be readily detected. LC/MS can be performed with AA-Ac labeled glycans run on either amide or EIR based chromatography before the electrospray mass spectrometer. Advantages of LC/MS over straight MS include the purification of glycans away from contaminants (e. g. salts and peptides) that interfere with ionization in the MS and separation of isobaric species that would overlap in the mass spectrum. Furthermore, the extra information given by the retention on HPLC can reduce the need for MS fragmentation, simplifying the analysis of routine samples.

AA-Ac labeled glycans also give excellent results in MALDI-MS and on current ion-trap instruments detailed MS^n studies can be performed with sub-picomole samples. The combination of this analysis with fractionation of complex samples on EIR or amide HPLC is particularly useful allowing reliable quantitation of the AA-Ac tagged glycans by HPLC and detailed structural characterization by MS^n fragmentation.

One of the biggest challenges in biopharmaceutical glycoprofiling is analysis of the negatively charged sialylated oligosaccharides. The main problems are non-ionization of the anionic glycans–especially in the presence of salts and peptides–and the lability of the sialyl residues

which can fall off as a result of in-source decay. These can be reduced by derivatization of the sialyl residues by methyl esterification or permethylation of the hydroxyls. The derivatized glycans will have altered characteristics on HPLC–which can be used to advantage–and the sialyl residues will be stabilized.

Successful glycoprofiling by MS requires optimization of MS conditions and calibration of the mass spectrometer. These must be done with well-characterized analytical standards. In the past these have been difficult to obtain but the range and availability of commercially available standards are improving. It is important to match the standards to the analytes–particularly in relation to the size of the glycans, degree of branching, degree of sialylation and other anionic groups, derivatization (particularly fluorophores or modifications to sialyl residues and hydroxyl groups). If possible, use pre-derivatized standards as well as glycans that you have derivatized in parallel to your own samples. Careful use of standards and instrument calibration will help you to get develop a well-characterized, robust mass spectrometry system for glycoprofiling.

In summary, the combination of HPLC and mass spectrometry is a very powerful technique, allowing detailed characterization of complex glycosylation with good relative quantitation.

6 Glycosylation Site Profile

With current technology detailed glycosylation site profiling is a lengthy process and not widely performed. However, it can yield valuable information, particularly during glycoprotein drug development and glycan characterization to support regulatory submissions or intellectual property rights. The general scheme is to isolate individual glycosylation sites as glycopeptides then perform a detailed glycan profile of each site. The challenges are to isolate pure sites, without contamination or selective loss of glycopeptide species, and to retain relative quantitation information between sites.

The simplest method for site isolation is to perform an exhaustive proteolytic digestion of the glycoprotein. Typically this is done using a very high-grade modified trypsin after reduction and alkylation. The sample should be monitored for completeness of digestion and aberrant digestion by contaminating proteases. If digestion yields glycopeptides with multiple glycosylation sites then one should use another complementary protease to affect cleavage between the connected sites. It is important to use proteases that give glycopeptides with sufficient size of peptides to allow good separation on the stage which is hydrophobic interaction chromatography (typically C18 or C8 HPLC with a wide pore resin using elution by aqueous acetonitrile with trifluoroacetic acid as an ion pair reagent). With care, this should give a peptide map with separation of glycosylation sites and co-elution of differently glycosylated variants of each site. Identification of glycosylation sites can be done by LC/MS using precursor ion scanning for masses indicative of glycosylation or collection of HPLC fractions then MS of each fraction. After this, each glycosylation site can be subjected to the full range of glycoprofiling analyses described above.

7 Summary

With current technology, quantitative glycosylation analysis of therapeutic glycoproteins can be very challenging. The glycosylation patterns can be extremely complex with species that

can be difficult to separate and refractory to simple analysis on mass spectrometry. However, with care, one can build a scheme for reliable quantitative glycoprofiling using a combination of glycoprofiling modules composed of selected sample preparation and derivation methods, HPLC and mass spectrometry. This chapter has overviewed some of the most useful glycoprofiling modules and how these tools can be integrated into a robust, flexible glycoprofiling system.

References

1. Fernandes DL (2004) Eur Biopharm Rev Winter (2004):92–97
2. Kamoda S, Kakehi K (2006) Electrophoresis 27:2495
3. Rosenfeld R, Bangio H, Gerwig GJ, Rosenberg R, Aloni R, Cohen Y, Amor Y, Plaschkes I, Kamerling JP, Maya RB (2007) J Biochem Biophys Methods 70:415–426
4. Townsend RR, Manzi A, Merkle RK, Rohde MF, Spellman M, Smith A, Carr SA (1997) Quantitative Monosaccharide Analysis: A Multi-Center Study. ABRF News, epublication
5. Fernandes DL (2006) Eur Biopharm Rev Spring (2006):100–104
6. Anumula KR (1995) Anal Biochem 230:24
7. O'Neill RA (1996) J Chromatogr A 720:201
8. Tarentino AL, Gomez CM, Plummer TH Jr (1985) Biochemistry 24:4665
9. Merry T, Astrautsova S (2003) Methods Mol Biol 213:27
10. Patel TP, Parekh RB (1994) Methods Enzymol 230:57
11. Bigge JC, Patel TP, Bruce JA, Goulding PN, Charles SM, Parekh RB (1995) Anal Biochem 230:229
12. Nakagawa H, Kawamura Y, Kato K, Shimada I, Arata Y, Takahashi N (1995) Anal Biochem 226:130
13. Chen FT, Evangelista RA (1998) Electrophoresis 19:2639
14. Charlwood J, Birrell H, Gribble A, Burdes V, Tolson D, Camilleri P (2000) Anal Chem 72:1453
15. Kotani N, Takasaki S (1998) Anal Biochem 264:66
16. Routier FH, Hounsell EF, Rudd PM, Takahashi N, Bond A, Hay FC, Alavi A, Axford JS, Jefferis R (1997) J Immunol Methods 213:113

Part 9
Glycomimetics

9.1 Azaglycomimetics: Natural Occurrence, Biological Activity, and Application

Naoki Asano
Faculty of Pharmaceutical Science, Hokuriku University,
920-1181 Kanazawa, Japan
n-asano@hokuriku-u.ac.jp

Abstract

A large number of alkaloids mimicking the structures of monosaccharides or oligosaccharides have been isolated from plants and microorganisms. The sugar mimicking alkaloids with a nitrogen in the ring are called azasugars or iminosugars. Naturally occurring azasugars are classified into five structural classes: polyhydroxylated piperidines, pyrrolidines, indolizidines, pyrrolizidines, and nortropanes. They are easily soluble in water because of their polyhydroxylated structures and inhibit glycosidases because of a structural resemblance to the sugar moiety of the natural substrate. Glycosidases are involved in a wide range of anabolic and catabolic processes, such as digestion, lysosomal catabolism of glycoconjugates, biosynthesis of glycoproteins, and the endoplasmic reticulum (ER) quality control and ER-associated degradation of glycoproteins. Hence, modifying or blocking these processes in vivo by inhibitors is of great interest from a therapeutic point of view. Azasugars are an important class of glycosi-

In: *Glycoscience*. Fraser-Reid B, Tatsuta K, Thiem J (eds)
Chapter-DOI 10-1007/978-3-540-30429-6_48:

dase inhibitors and are arousing great interest for instance as antidiabetics, antiobesity drugs, antivirals, and therapeutic agents for some genetic disorders. This review describes the recent studies on isolation, characterization, glycosidase inhibitory activity, and therapeutic application of azaglycomimetics.

Keywords

Azasugar; Azapyranose; Azafuranose; Bicyclic azasugar; Glycosidase inhibitor; Therapeutic application

Abbreviations

DAB	1,4-dideoxy-1,4-imino-D-arabinitol
DMDP	1,4-dideoxy-1,4-imino-D-mannitol
DNJ	1-deoxynojirimycin
ER	endoplasmic reticulum
ERAD	E*R*-associated degradation
FDA	US Food and Drug Administration
GSL	glycosphingolipid
HNJ	homonojirimycin
IRI	immunoreactive insulin
NIDDM	non-insulin-dependent diabetes mellitus
NJ	nojirimycin
STZ	streptozotocin

1 Introduction

Alkaloids mimicking the structures of monosaccharides are now believed to be widespread in plants and microorganisms [1,2]. Azasugars (or iminosugars) in which the ring oxygen has been replaced by nitrogen must be one of the most interesting discoveries in the field of natural products of recent years. Naturally occurring azasugars are classified into five structural classes: polyhydroxylated piperidines, pyrrolidines, indolizidines, pyrrolizidines, and nortropanes. They bind specifically to the active sites of glycosidases by mimicking the corresponding natural substrates. Glycosidases are involved in a wide range of important biological processes, such as intestinal digestion, post-translational processing of glycoproteins, quality control systems in the endoplasmic reticulum (ER) and ER-associated degradation (ERAD) mechanism, and the lysosomal catabolism of glycoconjugates. Inhibition of these glycosidases can have profound effects on carbohydrate catabolism in the intestines, and maturation, transport, and secretion of glycoproteins. The realization that glycosidase inhibitors have enormous therapeutic potential in many diseases such as diabetes, viral infection, and lysosomal storage disorders has led to increasing interest and demand for them [3,4]. This review describes the recent studies on isolation, characterization, glycosidase inhibitory activity, and therapeutic application of azaglycomimetics.

2 Natural Occurrences

2.1 Piperidines (Azapyranose)

In 1966, NJ (nojirimycin) (**1**) was discovered as the first natural glucose mimic, with a nitrogen atom in place of the ring oxygen [5]. NJ was originally described as an antibiotic produced by *Streptomyces roseochromogenes* R-468 and *S. lavendulae* SF-425 [5,6], and was shown to be a potent inhibitor of α- and β-glucosidases from various sources. Thereafter, *manno*-NJ (nojirimycin B) (**2**) [7] and *galacto*-NJ (galactostatin) (**3**) [8] as azasugars with the hydroxyl group at C-1 were also isolated from the fermentation broths of species of *Streptomyces*. *Manno*-NJ was co-produced with NJ by *S. lavendulae* SF-425, while *galacto*-NJ was isolated as a potent β-galactosidase inhibitor produced by *S. lydicus* PA-5726. However, because these iminosugars with the hydroxyl group at C-1 are fairly unstable, they are usually stored as bisulfite adducts; they may be reduced by catalytic hydrogenation with the platinum catalyst or by $NaBH_4$ to the stable 1-deoxy derivatives (see ❯ *Fig. 1* for structures of **1**–**24**).

DNJ (1-deoxynojirimycin) (**4**) was first prepared by the reduction of nojirimycin as described above [5] and chemically synthesized from L-sorbofuranose [9] but later isolated from the roots of mulberry trees and called moranoline [10]. DNJ is also produced by many strains of the genera *Bacillus* [11] and *Streptomyces* [12,13]. The first naturally occurring *N*-methyl derivative of DNJ was isolated from the leaves and roots of *Morus* spp. (mulberry trees) and, furthermore, the genus *Morus* has been shown to co-produce many kinds of glycosides of DNJ such as 2-*O*-, 3-*O*-, 4-*O*-α-D-glucosides, 2-*O*-, 3-*O*-, 4-*O*-, 6-*O*-β-D-glucosides, and 2-*O*-, 6-*O*-α-D-galactosides [14,15]. *Manno*-DNJ (1-deoxymannojirimycin) (**5**) was first isolated from the seeds of the legume *Lonchocarpus sericeus*, a native of the West Indies and tropical America [16] and later isolated from the neotropical liana, *Omphalea diandra* (Euphorbiaceae) [17], and the legume *Angylocalyx pynaertii* [18] growing in tropical African forests. *Manno*-DNJ was also isolated from the culture broth of *Streptomyces lavendulae* GC-148 [19], which had already been reported as a high-yielding strain of DNJ [13]. Very recently, our group has isolated *manno*-DNJ in high yields of 0.083% and 0.17% w/w from Thai medicinal plants "Thopthaep" and "Cha em thai" [20]. The biological origin of "Thopthaep" and "Cha em thai" were identified as *Connarus ferrugineus* (Connaraceae) and *Albizia myriophylla* Benth (Leguminosae), respectively. *C. ferrugineus* is used traditionally to treat scabies, as an ointment, and to treat stomach ache and constipation as an oral drug, while the roots and wood of *A. myriophylla* are used traditionally to relieve thirst and sore throats and to substitute for licorice owing to a sweet taste. Furthermore, 2-*O*-α-D-galactopyranosyl-*manno*-DNJ and 3-*O*-β-D-glucopyranosyl-*manno*-DNJ from *C. ferrugineus* and 2-*O*-β-D-glucopyranosyl-*manno*-DNJ and 4-*O*-β-D-glucopyranosyl-*manno*-DNJ from *A. myriophylla* have also been isolated [20].

Streptomyces subrutilus ATCC 27467 produces both DNJ and *manno*-DNJ in its culture broth, and *manno*-NJ has been suggested as an intermediate in the biosynthesis of *manno*-DNJ [21]. In this organism, it has also been suggested that epimerization of mannojirimycin can occur at C-2 to give NJ which is then dehydrated and reduced to DNJ. *Agrobacterium* spp. strain 19-1 has been shown to be able to epimerize DNJ to *manno*-DNJ [22]. Although the natural occurrence of *galacto*-DNJ (1-deoxygalactonojirimycin) (**6**) has not been reported to date,

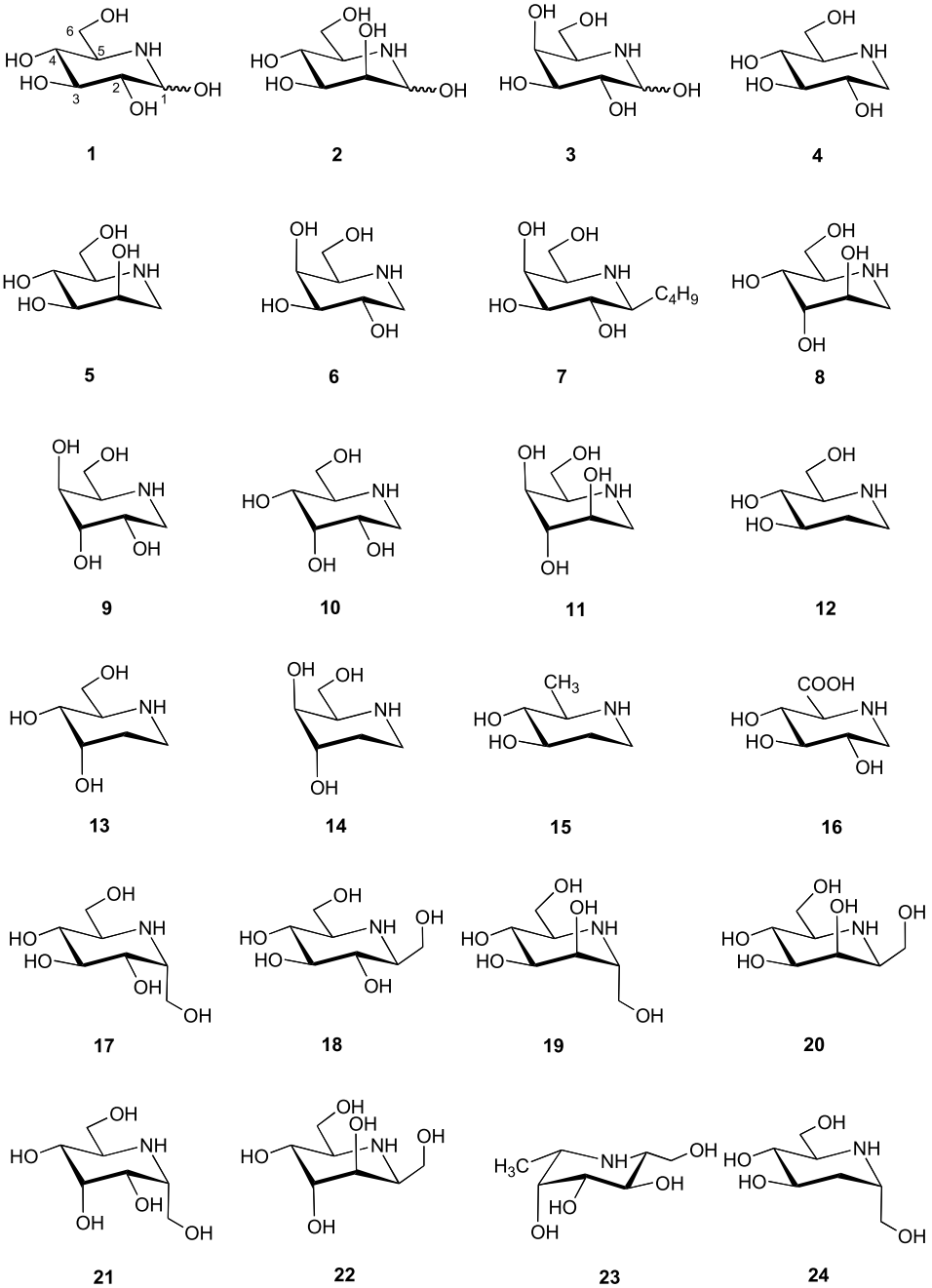


Figure 1
Structures of 1–24

β-1-*C*-butyl-*galacto*-DNJ (**7**) has been isolated from a commercially available Chinese crude drug "Sha-sheng," which is the roots of *Adenophora* spp. (Campanulaceae) [23].

Five naturally occurring 1-deoxyazasugars have been isolated to date. DNJ and *manno*-DNJ have been isolated from many plants and microorganisms as described above. *Altro*-DNJ (1-deoxyaltronojirimycin) (**8**) and *gulo*-DNJ (1-deoxygulonojirimycin) (**9**) have been discovered from *Scilla sibirica* (Hyacinthaceae) [24] and *Angylocalyx pynaetii* (Legminosae) [25]. The ^{1}H and ^{13}C NMR spectroscopic data of *allo*-DNJ (1-deoxyallonojirimycin) (**10**) isolated from "Thopthaep" [20] were in accord with those of synthetic *allo*-DNJ [26]. Recently, the L-enantiomers of DNJ, *manno*-DNJ, *allo*-DNJ, *altro*-DNJ, *galacto*-DNJ, *gulo*-DNJ, and *ido*-DNJ (**11**) were enantiospecifically synthesized [27]. The absolute configurations of natural DNJ, *manno*-DNJ, *allo*-DNJ, *altro*-DNJ, and *gulo*-DNJ were determined to be D-, D-, D-, D-, and L-enantiomers, respectively, from the value and sign of the optical rotation.

1,2-Dideoxynojirimycin, fagomine (**12**), was isolated from the seeds of Japanese Buckwheat (*Fagopyrum esculentum*) [28] and the Moreton Bay chestnut (black bean), *Castanospermum australe* (Leguminosae) [29], and it also occurs as the 4-*O*-β-D-glucoside in the seeds of the legume *Xanthocercis zambesiaca* [30]. The search for azasugars in the leaves and roots of *X. zambesiaca* led to the isolation of fagomine, 3-*epi*-fagomine (**13**), 3,4-di-*epi*-fagomine (**14**), and 3-*O*-β-D-glucopyranosylfagomine in addition to the 4-*O*-β-D-glucoside [31]. From the Chinese crude drug "Ti-koppi," the roots of *Lycium chinense* (Solanaceae), fagomine and 6-deoxyfagomine (**15**) have been isolated, together with many kinds of polyhydroxylated nortropane and tropane alkaloids [32]. 2*S*-Carboxy-3*R*,4*R*,5*S*-trihydroxypiperidine (**16**), which is a mimic of glucuronic acid, was discovered from the seeds of the legume *Baphia racemosa* as a novel specific inhibitor of human liver β-D-glucuronidase and α-L-iduronidase [33].

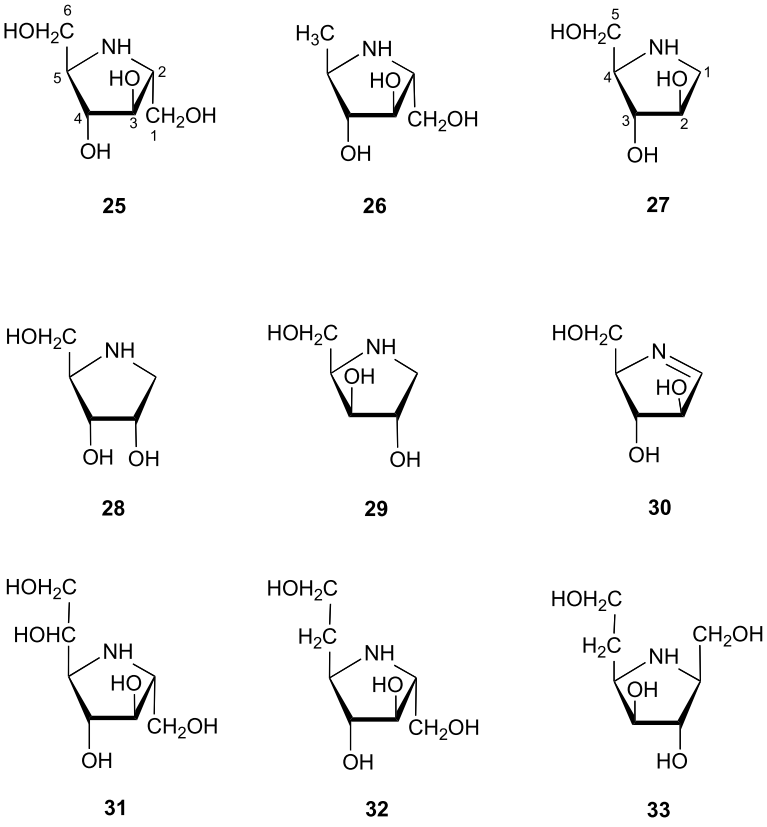
In 1988, α-HNJ (α-homonojirimycin) (**17**) was isolated from *O. diandra*, together with *manno*-DNJ [17], as the first report of a naturally occurring DNJ derivative with a carbon substituent at C-1. However, before the isolation of the natural product, its 7-*O*-β-D-glucoside (MDL 25637) had been designed as a potential drug for the treatment of diabetes mellitus [34,35]. α-HNJ has been detected in adults, pupae, and eggs of the neotropical moth, *Urania fulgens*, whose larvae feed on *O. diandra* [36]. α-HNJ has also been isolated from whole plants of *Aglaonema treubii* (Araceae) [37,38] and the bulbs of *Hyacinthus orientalis* (Hyacinthaceae) [39]. These plants also co-produce the isomers and glycosides of α-HNJ such as β-HNJ (β-homonojirimycin) (**18**), α-*manno*-HNJ (α-homomannojirimycin) (**19**), β-*manno*-HNJ (β-homomannojirimycin) (**20**), α-*allo*-HNJ (α-homoallonojirimycin) (**21**) [40], β-*altro*-HNJ (β-homoaltronojirimycin) (**22**) [38], and the 5-*O*-α-D-galactoside and 7-*O*-β-D-glucoside of α-HNJ. The isomers **21** and **22** and the 5-*O*-α-D-galactoside are new natural products, while the isomers **18**, **19**, and **20** have been chemically synthesized previously [41,42,43]. Amazingly, the 7-*O*-β-D-glucoside of α-HNJ (MDL 25637), which had been developed as an antidiabetic agent, was also isolated from these plants. The 7-deoxy derivative of β-*manno*-HNJ was isolated from the pods of *A. pynaetii* (Legminosae) [25] and the bulbs of *Scilla sibirica* (Hyacinthaceae) [24]. This compound was identical with β-L-fuco-HNJ (**23**) which had been chemically synthesized as a powerful inhibitor of α-L-fucosidase [44].

In the course of a search for α-glucosidase inhibitors, Kitaoka et al. found that such inhibitors are present in some Thai traditional crude drugs [45]. For example, α-HNJ occurs in "Non tai yak" at a level of 0.1% (dry weight). The "Non tai yak" sample is known to be *Stemona tuberosa* (Stemonaceae), which has been used in China and Japan for various medicinal pur-

poses. In particular, an extract from the freshly tuberous roots of *S. tuberosa* is used to treat respiratory disorders, including pulmonary tuberculosis and bronchitis, and also recommended as an insecticide [46,47,48]. In 2005, the re-examination of polyhydroxylated alkaloids in *S. tuberosa* led to the isolation of 13 alkaloids containing **4**, **5**, **17** (0.1% of dry weight), **18**, **19**, **20**, **21**, **22**, and α-5-deoxy-HNJ (**24**) as piperidine alkaloids [20].

2.2 Pyrrolidines (Azafuranose)

In 1976, DMDP (2*R*,5*R*-bis(dihydroxymethyl)-3*R*,4*R*-dihydroxypyrrolidine, 2,5-dideoxy-2,5-imino-D-mannitol) (**25**; see ● *Fig. 2* for structures of **25**–**33**), mimicking β-D-fructofuranose, was found in leaves of the legume *Derris elliptica* [49] and later shown to be present in many disparate species of plants and microorganisms [2], which would indicate that this is a common metabolite. The 6-deoxy derivative of DMDP (6-deoxy-DMDP, 2,5-imino-1,2,5-trideoxy-D-mannitol) (**26**) was isolated from the seeds of *Angylocalyx pynaertii* as



● **Figure 2**
Structures of 25–33

a weak β-mannosidase inhibitor [18]. Removal of one hydroxymethyl group from DMDP leads to DAB (1,4-dideoxy-1,4-imino-D-arabinitol) (**27**), which was first found in the fruits of *Angylocalyx boutiqueanus* [50] and subsequently in many disparate species of plants as well as DMDP [2]. DAB also occurred as the 2-*O*-β-D-glucoside in the leaves and roots of *Morus* spp. [14,15]. The 2-epimer of DAB, 1,4-dideoxy-1,4-imino-D-ribitol (**28**), has also been isolated from *Morus* spp. [15] and 1,4-dideoxy-1,4-imino-D-xylitol (**29**) from *A. pynaertii* [25]. The polyhydroxypyrroline nectrisine (FR-900483) (**30**) was isolated as an immunomodulator from the culture broth of the fungus *Nectria ludica* [51] and further shown to be a powerful inhibitor of yeast α-glucosidase [52]. 2,5-Dideoxy-2,5-imino-D-*glycero*-D-taloheptitol has been chemically synthesized as an iminofuranose [53]. In 1997, the first naturally occurring 2,5-dideoxy-2,5-iminoheptitol, homoDMDP (2,5-dideoxy-2,5-imino-DL-*glycero*-D-*manno*-heptitol) (**31**), was found in the leaves of bluebell (*Hyacinthoides non-scripta*) [54]. This compound was isolated later from the bulbs of Hyacinths (*Hyacinthus orientalis*) [39] and *Scilla campanulata* (Hyacinthaceae) [55] and also isolated as the 7-*O*-apioside and 7-*O*-β-D-xyloside in *H. non-scripta* and *S. campanulata* [54,55]. Hyacinths further co-produced 6-deoxy-homoDMDP (2,5-imino-2,5,6-trideoxy-D-*manno*-heptitol) (**32**) and 2,5-imino-2,5,6-trideoxy-D-*gulo*-heptitol (**33**) [39].

The structure of homoDMDP has been given as 2,5-dideoxy-2,5-imino-DL-*glycero*-D-*manno*-heptitol or its enantiomer from its NMR spectroscopic data [54]. However, the relative configuration at C-6 has not yet been determined since it cannot be determined from the NMR data. In 1999, in the course of the synthesis of a series of five-membered iminosugars, Takebayashi et al. reported the enantiospecific synthesis of (1′*S*,2*R*,3*R*,4*R*,5*R*)-3,4-dihydroxy-2-(1,2-hydroxyethyl)-5-hydroxymethyl)pyrrolidine [56]. The ^{1}H NMR and ^{13}C NMR spectroscopic data of homoDMDP reported in ref. 54 and 55 were superimposable with those of the synthetic compound in ref. 56, and the optical rotation {$[\alpha]_D$ +28.8° (c 3.75, H_2O)} of homoDMDP was also similar to that {$[\alpha]_D$ +25.6° (c 0.3, H_2O)} of the synthetic compound. Hence, we determined that both compounds are identical and the structure of homoDMDP is 2,5-dideoxy-2,5-imino-D-*glycero*-D-*manno*-heptitol. The structure of 6-deoxy-homoDMDP isolated from *H. orientalis* has been determined to be 2,5-dideoxy-2,5-imino-D-*manno*-heptitol or its enantiomer [39]. Fleet et al. have synthesized both enantiomers of 6-deoxy-homoDMDP (the detailed synthesis and their biological activities will be reported elsewhere). From comparison of the optical rotation values {natural product, $[\alpha]_D$ +59.2° (c 1.77, H_2O); D-enantiomer, $[\alpha]_D$ +62.2° (c 0.19, H_2O); L-enantiomer, $[\alpha]_D$ −89.1° (c 0.55, H_2O)}, the natural product was determined to be D-enantiomer.

2.3 Indolizidines

There are three groups of bicyclic polyhydroxyheterocycles, with two five-membered rings (pyrrolizidines) or one five-membered ring and one six-membered ring (indolizidines and nortropanes) fused together. These bicyclic alkaloids have a less obvious structural relationship to monosaccharides but in each case the configuration of hydroxyl substituents on the ring can be compared to those of sugars.

Certain poisonous plants often cause serious livestock losses. The Australian legumes, *Swainsona*, are known as “poison peas,” and sheep eating them develop a syndrome called “pea

struck" [57,58]. There is also the livestock poisoning by the closely related *Astragalus* and *Oxytropis* species, which are found throughout most of the world, and intoxication of livestock by certain of those species known as locoweeds in the western United States is called "locoism" [58,59]. The poisoning is characterized by cytoplasmic vacuolation of neuronal cells due to accumulation of mannose-rich oligosaccharides in lysosomes [60]. The trihydroxyindolizidine alkaloid swainsonine (**34**) occurs in these legumes and has been identified as a causative agent in locoism [59,61]. The toxicity of the other legume *Castanospermum australe* for livestock led to the isolation of the toxic principle castanospermine (**35**) [62] and these two alkaloids gave rise to a great impetus in research on *N*-containing sugars and their application (see ❯ *Fig. 3* for structures of **34**–**41**).

In 1981, castanospermine was first isolated from the immature seeds, with a yield of 0.057% [62]. X-ray crystallography showed that the stereogenic centers of the six-membered ring of castanospermine correspond to the *gluco* configuration [62], while 6-*epi*-castanospermine (**36**) isolated later from the seeds has the D-*manno* configuration in the piperidine ring [63]. *C. australe* co-produces 7-deoxy-6-*epi*-castanospermine (**37**) [64] and 6,7-di*epi*-castanospermine (**38**) [65]. Very recently, we have isolated a new castanospermine isomer 6,8-di*epi*-castanospermine (**39**) from the leaves and twigs, which contain castanospermine at the high level of 1% (unpublished data). In 2003, the first naturally occurring glycoside of **35**, castanospermine-8-*O*-β-D-glucoside, was isolated from the seeds [66]. Lentiginosine (**40**) and 2-*epi*-lentiginosine (**41**) have been isolated from the leaves of *Astragalus lentiginosus* and these two dihydroxyindolizidines are probably biosynthesized from 1-hydroxyindolizidine by hydroxylation at C-2 [67].

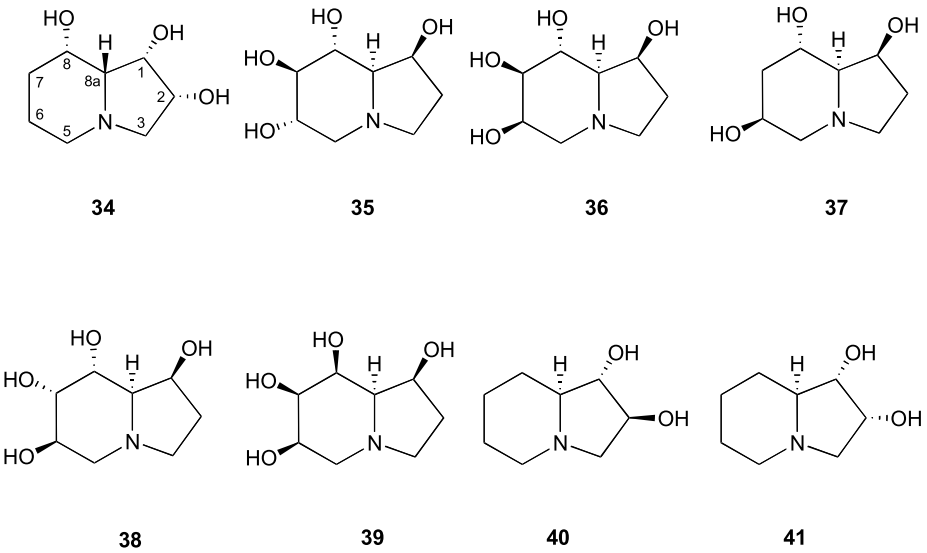


Figure 3
Structures of 34–41

2.4 Pyrrolizidines

In 1988, alexine (**42**), a polyhydroxylated pyrrolizidine alkaloid, was isolated from the pods of legume *Alexa leiopetala* [68] (see ❯ *Fig. 4* for structures of **42–48**). Although the broad class of pyrrolizidine alkaloids bear a carbon substituent at C-1 [69,70], alexine is the first example of a pyrrolizidine alkaloid with a carbon substituent at C-3. At about the same time, australine (**43**) was isolated from the seeds of *C. australe* and found to be 7a-*epi*-alexine from X-ray crystallographic analysis [71]. The isolation of 1-*epi*-australine (**44**), 3-*epi*-australine (**45**), and 7-*epi*-australine (**46**) from the same plant was later reported [72,73,74]. The structure of 1-*epi*-australine was firmly established by X-ray crystallographic analysis of the corresponding 1,7-isopropylidene derivative [73], and the absolute configurations of 3-*epi*-australine were also identified by X-ray crystal structure analysis [73]. Alkaloid **46** was tentatively assigned as 7-*epi*-australine based on the difference between its NMR parameters and those reported for australine [74]. The unambiguous synthesis of australine [75] and 7-*epi*-australine [76,77] and extensive NMR studies on the natural and synthetic isomers of australine [77] by Denmark et al. elucidated that the natural product reported as 7-*epi*-australine is really australine. This means that 7-*epi*-austaline has not yet been found as a natural product. Reinvestigation of the natural occurrence of 7-*epi*-australine in *C. australe* led to the isolation of new alkaloids, 2,3-di*epi*-australine (**47**), 2,3,7-tri*epi*-australine (**48**), and 2-*O*-β-D-glucopyranosyl-1-*epi*-australine [66].

Polyhydroxylated pyrrolizidine alkaloids with a hydroxymethyl substituent at C-3 have been thought to be of very restricted natural occurrence. The alexines and australines have been reported in only two small genera of the Leguminosae (*Castanospermum* and *Alexa*). However, a number of such pyrrolizidine alkaloids were found from a quite different family, Hyacinthaceae. In 1999, new polyhydroxylated pyrrolizidines different from

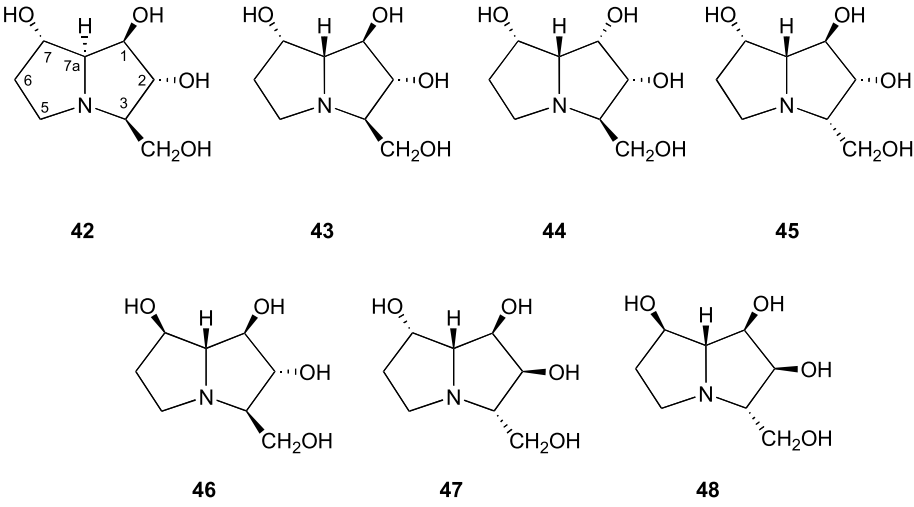


Figure 4
Structures of 42–48

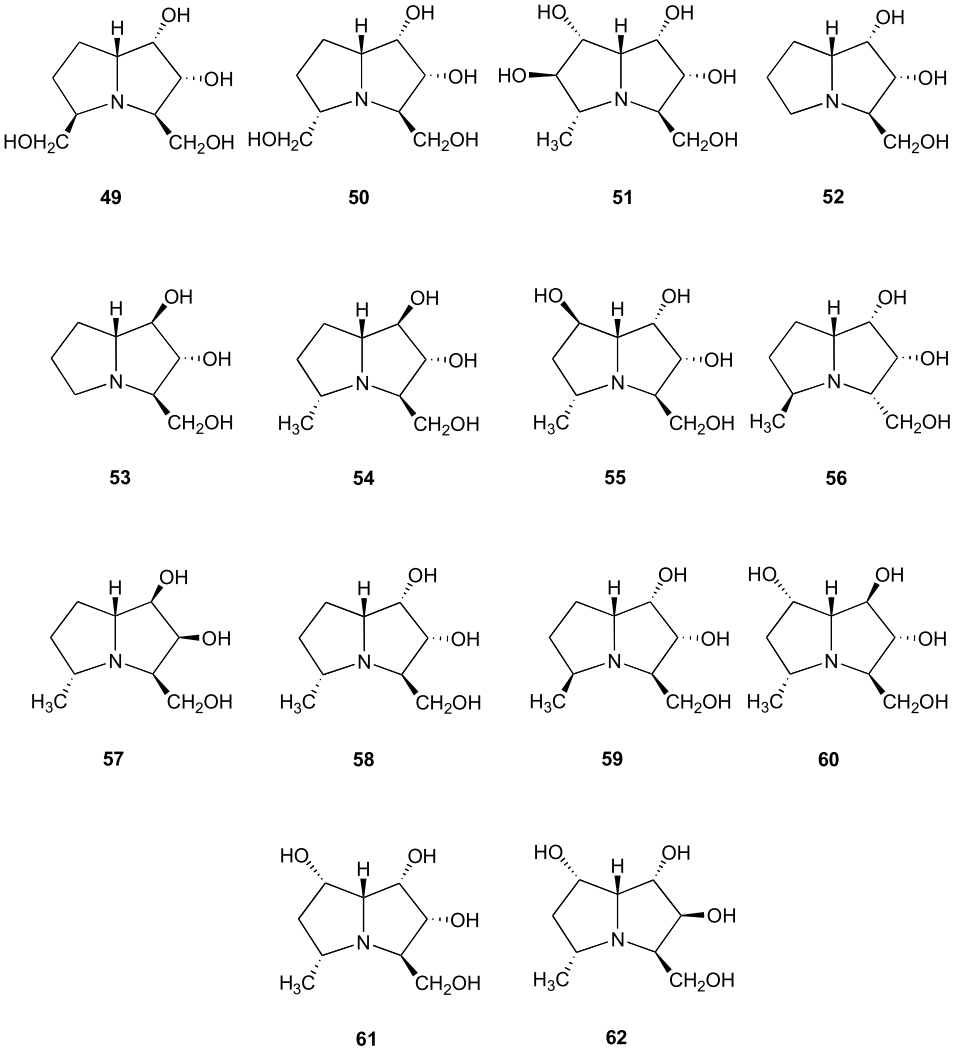
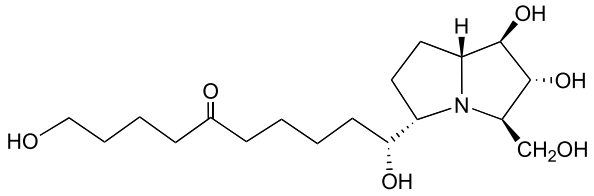


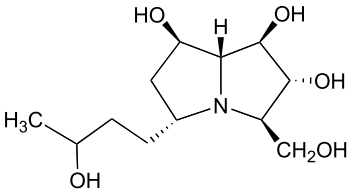
Figure 5
Structures of 49–62

alexines and australines were isolated from the Hyacinthaceae family plants, designated as hyacinthacines: hyacinthacines B_1 (**49**) and B_2 (**50**) from the immature fruits and stalks of *Hyacinthoides non-scripta* and hyacinthacine C_1 (**51**) from the bulbs of *Scilla campanulata* [55] (see *Fig. 5* for structures of **49–62**). Shortly after their isolation, four new hyacinthacines A_1 (**52**), A_2 (**53**), A_3 (**54**), and B_3 (**55**) were isolated from the bulbs of *Muscari armeniacum* in addition to **50** [78]. In 2001, Martin and co-workers reported the first synthesis of (+)-hyacinthacine A_2 from commercially available 2,3,5-tri-*O*-benzyl-D-arabinofuranose, and confirmed the absolute configuration of the natural (+)-hyacinthacine

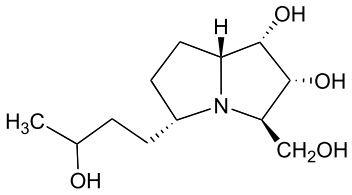
A_2 as (1*R*,2*R*,3*R*,7a*R*)-1,2-dihydroxy-3-hydroxymethylpyrrolizidine [79]. Subsequently, the natural (+)-hyacinthacine A_3 was enantiospecifically synthesized from an adequately protected DMDP and its absolute configuration was determined to be (1*R*,2*R*,3*R*,5*R*,7a*R*)-1,2-dihydroxy-3-hydroxymethyl-5-methylpyrrolizidine [80]. More recently, the absolute configuration of the natural (+)-hyacinthacine A_1 has been determined to be (1*S*,2*R*,3*R*,7a*R*)-1,2-dihydroxy-3-hydroxymethylpyrrolizidine from its total synthesis [81]. Many species of the genera *Muscari*and *Scilla*are very common as garden plants. The GC-MS analysis of the extract of commercially available bulbs of *S. sibirica* elucidated the existence of many kinds of polyhydroxylated alkaloids, five pyrrolidines and two pyrrolidine glycosides, six piperidines and one piperidine glycoside, and eight pyrrolizidines [82]. Surprisingly, seven pyrrolizidines other than the known alkaloid **55** were new hyacinthacines. They are hyacinthacines A_4 (**56**), A_5 (**57**), A_6 (**58**), A_7 (**59**), B_4 (**60**), B_5 (**61**), and B_6 (**62**).



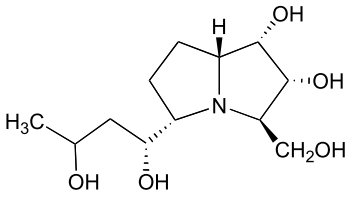
63



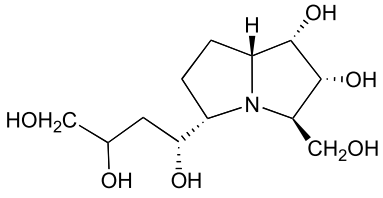
64



65



66



67

Figure 6
Structures of 63–67

In 1999, it was reported that *Broussonetia kajinoki* (Moraceae) produces a pyrrolizidine alkaloid with the C_{10} side chain at the C-5α position, which was designated as broussonetine N (**63**; see *Fig. 6* for structures of **63**–**67**), together with pyrrolidine alkaloids with a long (C_{13}) side chain [83]. Broussonetine N can be regarded as the α-5-*C*-(1,19-dihydroxy-6-oxodecyl)-hyacinthacine A_2. In 2004, four new hyacinthacine A_1 and 7-*epi*-australine derivatives with the hydroxybutyl side chain at the C-5α position were produced from the bulbs of *S. peruviana*, which also co-produces pyrrolidine alkaloids with a highly hydroxylated long side chain [84]. These four pyrrolizidine alkaloids were determined to be α-5-*C*-(3-hydroxybutyl)-7-*epi*-australine (**64**), α-5-*C*-(3-hydroxybutyl)hyacinthacine A_1 (**65**), α-5-*C*-(1,3-dihydroxybutyl)hyacinthacine A_1 (**66**), and α-5-*C*-(1,3,4-trihydroxybutyl)-hyacinthacine A_1 (**67**). The NOE experiments of pyrrolizidines **66** and **67** suggested that the absolute configurations of C-1′ are *R*.

2.5 Nortropanes

Until recently, four structural classes were encompassed as naturally occurring azasugar mimetics: polyhydroxylated piperidines, pyrrolidines, indolizidines, and pyrrolizidines. An entirely new class has now been added to this group with the discovery of the calystegines. Calystegines possess three structural features in common: a nortropane ring system; two to four hydroxyl groups varying in position and stereochemistry; and a novel aminoketal functionality, which generates a tertiary hydroxyl group at the bicyclic ring bridgehead. Calystegines were first discovered as secondary metabolites of plants and were implicated in the establishment and maintenance of specific plant-bacterium relationships [85]. Calystegines are abundant in the underground organs and root exudates of *Calystegia sepium*, *Convolvulus arvensis* (both Convolvulaceae), and *Atropa belladonna* (Solanaceae). The calystegines were subsequently extracted from root cultures of *C. sepium* transformed from *Agrobacterium rhizogenes*, grown in vitro and classified into two groups, calystegines A and B, on the basis of their relative mobility on paper electrophoresis [85]. The calystegine A group was separated by HPLC into four components (A_1, A_2, A_3, and A_4), of which the structure of the major component, calystegine A_3 (**68**; see *Fig. 7* for structures of **68**–**87**), was identified as 1α,2β,3α-trihydroxynortropane by a combination of high-resolution mass spectrometry and ^{1}H and ^{13}C NMR spectroscopy [86]. The calystegine B group was separated into two constituents, B_1 and B_2, and the structures of B_1 (**69**) and B_2 (**70**) were determined to be 1α,2β,3α,6α-tetrahydroxynortropane and 1α,2β,3α,4β-tetrahydroxynortropane, respectively, in a similar manner. A recent thorough examination of *M. bombycis* and *M. alba* (both of the Moraceae) resulted in the discovery of calystegines B_1, B_2, and C_1 (**71**) [14,15,87]. Calystegine C_1 was determined to be 1α,2β,3α,4β,6α-pentahydroxynortropane by a thorough NMR study. Since tropane alkaloids had not hitherto been isolated from the Moraceae, their occurrence therein is unique at the present time. A recent survey for the occurrence of calystegines in plants of Solanaceae and Convolvulaceae discovered that they are widely distributed in these families [1,2,88,89,90].

The occurrence in Solanaceae is documented for 12 genera, *Atropa*, *Brunfelsia*, *Datura*, *Duboisia*, *Hyoscyamus*, *Lycium*, *Mandragora*, *Nicandra*, *Physalis*, *Scopolia*, *Solanum*, and *Withania* [2,89]. Our group isolated and characterized calystegines A_3, A_5 (**72**), B_1, B_2, and

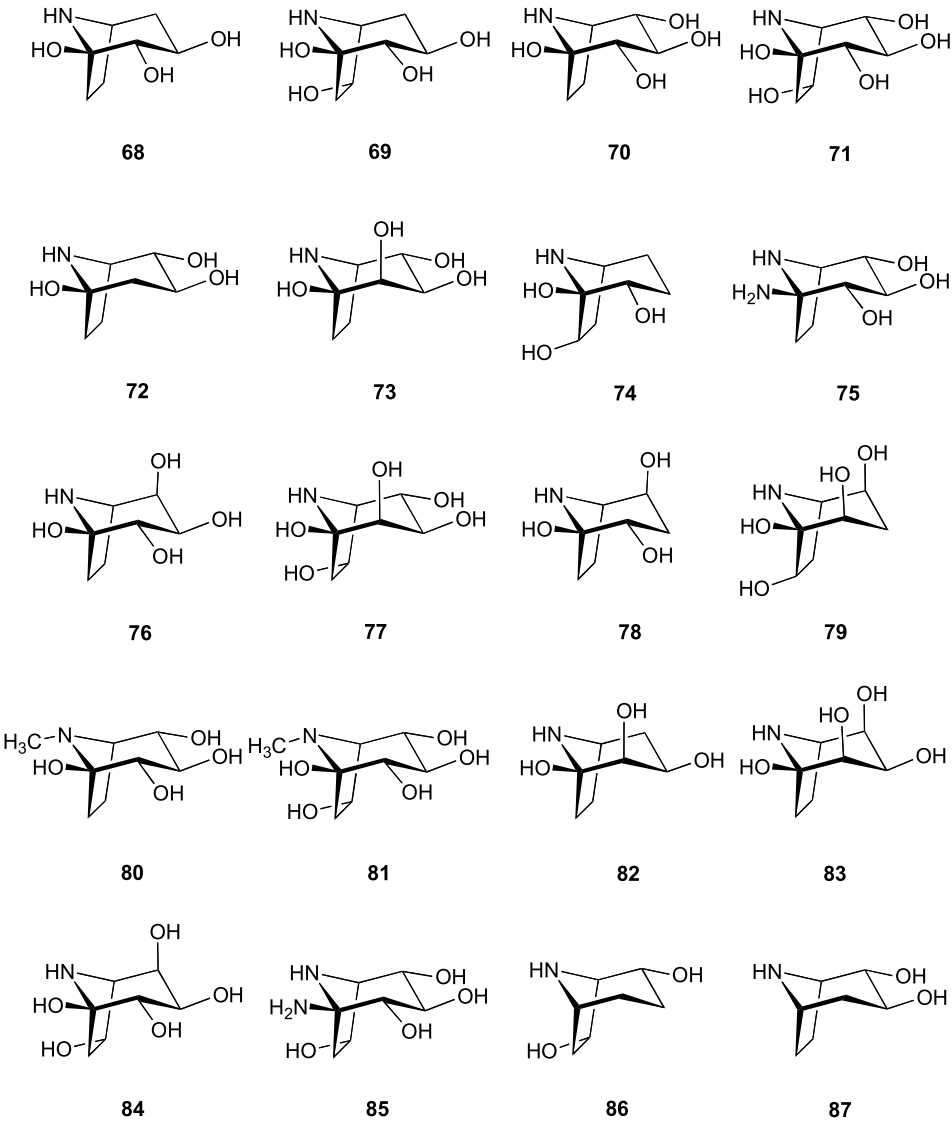


Figure 7
Structures of 68–87

B_3 (**73**) from the roots of *Physalis alkekengi* var. *francheti* [91], calystegines A_5, A_6 (**74**), B_1, B_2, B_3, and N_1(**75**) from the whole plant of *Hyoscyamus niger* [92], calystegines A_3, A_5, B_1, B_2, B_3, B_4 (**76**), and C_1 from the roots of *Scopolia japonica* [93], and calystegines B_1, B_2, B_4, C_1, and C_2 (**77**) from the leaves and twigs of *Duboisia leichhardtii* [94]. An examination of the roots of *Lycium chinense* led to the discovery of two new calystegines A_7 (**78**) and

B_5 (**79**), and two novel tropane alkaloids *N*-methylcalystegines B_2 (**80**) and C_1 (**81**), unlike the previously reported nortropane alkaloids [95]. Our recent works elucidated the presence of calystegine A_8 (**82**) in *H. niger*, and calystegine B_6 (**83**) in *S. japonica*, and calystegines C_3 (**84**) and N_2 (**85**) in *D. leichhardtii* (unpublished data). Calystegines N_1 and N_2 are assigned to an entirely new group of calystegines (the N series). The FABMS analysis of calystegines N_1 and N_2 gives odd-numbered $[M+H]^+$ ions of *m*/*z* 175 and 191 due to the replacement of an OH group by a NH_2 group relative to calystegines B_2 and C_1, respectively. The additional amino groups are located on C-1 in the parent alkaloids with the chemical shifts of the sole quaternary carbon at δ 78.3 (N_1) or 79.2 (N_2) in the ^{13}C NMR, in contrast to all other calystegines in which the hydroxyl-substituted quaternary carbon resonance occurs at an essentially invariant value of δ 93–94 ppm.

Besides free calystegines, calystegine B_1 occurs as the 3-*O*-β-D-glucoside in *Nicandra physalodes* fruits [96], and *Atropa belladonna* contains several glycosides including 3-*O*-β-D-glucopyranosylcalystegine B_1 and 4-*O*-α-D-galactopyranosylcalystegine B_2 [97]. The latter galactoside is also found in the mulberry fruits, as previously described [87]. Microbial β-transglucosylation of calystegine B_1 or B_2 using the whole cells of the yeast *Rhodotorula lactosa* gives 3-*O*-β-D-glucopyranosylcalystegine B_1or 4-*O*-β-D-glucopyranosylcalystegine B_2, respectively [98]. The glucose transfer to calystegine B_1 by commercially available rice α-glucosidase provides 3-*O*-α-D-glucopyranosylcalystegine B_1, but this enzyme does not transfer D-glucose to calystegine B_2 [98]. The lack of α-glucosyl transfer to calystegine B_2 could be due to the inhibition of rice α-glucosidase by calystegine B_2.

The enantioselective syntheses of (+)- and (−)-calystegines B_2 have determined that (+)-calystegine, (1*R*,2*S*,3*R*,4*S*,5*R*)-1,2,3,4-tetrahydroxynortropane, is the natural molecule [99,100], while the absolute configuration of the natural (−)-calystegine A_3 has been established as (1*R*,2*S*,3*R*,5*R*)-1,2,3-trihydroxynortropane by the syntheses of both enantiomers [101]. Later, natural (+)-calystegine B_3 and (−)-calystegine B_4 were prepared from D-galactose and D-mannose, respectively, for the first time and their absolute configuration confirmed [102,103].

Calystegines appear to be further widely distributed in the Convolvulaceae [88,90]. Eich and co-workers focused on the occurrence of seven calystegines (A_3, A_5, B_1, B_2, B_3, B_4, and C_1) identified by GC-MS analysis with authentic samples as references and analyzed the extracts of 65 Convolvulaceous species from predominantly tropical provenances belonging to all continents except Australia [88]. Consequently, they revealed their occurrence in 30 species belonging to 15 genera in this family. The qualitatively dominant alkaloid of the A series is calystegine A_3, detected in 43% of the 30 positive species followed by A_5 (20%). In the B series, calystegine B_2 exhibits the highest occurrence (86%) followed by B_1 (70%), B_3 (20%), and B_4 (13%). Calystegine C_1 is found in less than 1% of the species.

Convolvulaceous species produce polyhydroxylated alkaloids other than calystegines, as seen in ❯ *Table 1*. 2α,7β-Dihydroxynortropane (**86**) was isolated from seven species in Convolvulaceae, while only *Calystegia soldanella* contained both 2α,7β-dihydroxynortropane and 2α,3β-dihydroxynortropane (**87**) [90]. More interestingly, *Ipomoea carnea*produces the indolizidine alkaloids swainsonine (**34**) and 2-*epi*-lentiginosine (**41**) in addition to calystegines B_1, B_2, B_3, and C_1 [104]. *I. carnea* is a plant of tropical American origin but is now widely distributed in the tropical regions of the world. Chronic ingestion of *I. carnea* sometimes causes outbreak of natural poisoning in livestock. Swainsonine is also found in other species of the Convolvulaceae family. *Ipomoea* sp. Q6 aff. *calobra* (Weir Vine) grows in a small

Table 1
Distribution of polyhydroxylated nortropane and indolizidine alkaloids in the Convolvulaceae

Plant	Organ[a]	86	87	A_3 (68)	A_5 (72)	B_1 (69)	B_2 (70)	B_3 (73)	B_4 (76)	C_1 (71)	Sw[b] (34)	El[c] (41)
Calystegia japonica Choisy	C	•	–[d]	•	–	•	•	–	–	–	–	–
Calystegia sepium (L.) R. Br.	D	•	–	•	–	•	•	–	–	–	–	–
Calystegia soldanella (L.) Roem. et Schult.	A	•	•	•	–	•	•	•	–	–	–	–
Ipomoea batatas Lam. var. *edulis* Makino	B	•	–	–	–	•	•	–	–	–	–	–
Ipomoea carnea Jacq. (Japan)	B	•	–	–	–	•	•	–	–	–	•	–
Ipomoea carnea Jacq. (Brazil)	B	–	–	–	–	•	•	•	–	•	•	•
Ipomoea nil (L.) Roth	A	–	–	–	–	–	–	–	–	–	–	–
Ipomoea obscura Ker.	A	–	–	–	–	•	•	•	•	•	–	–
Ipomoea pes-caprae (L.) Sweet	B	–	–	–	–	–	•	–	–	–	–	–
Ipomoea vitifolia Vanprul	B	•	–	–	–	–	•	•	•	•	–	–
Quamoclit angulata Bojer.	B	•	–	–	–	•	•	–	–	–	–	–

[a]A: Whole parts. B: Aerial parts. C: Roots. D: Root cultures. [b]Swainsonine. [c]2-*epi*-Lentiginosine. [d]Not determined

area of southern Queensland in Australia and is reported to produce neurological disorders in livestock. The clinical symptoms are similar to those caused by swainsonine-containing legumes. Molyneux et al. detected swainsonine in the seeds and estimated the level as 0.058% by GC-MS and furthermore, the existence of swainsonine and calystegine B_2 in the seeds of *I. polpha* collected in the Northern Territory [105].

3 Chemical Biology/Biomedicine

Glycosidase inhibitors are currently of great interest as potential therapeutic agents because modifying or blocking biological processes using glycosidase inhibitors can significantly affect carbohydrate anabolism and catabolism. It has been increasingly realized that glycosidase inhibitors have enormous potentials in many diseases such as diabetes, viral infection, and lysosomal storage disorders [1,2,3,4].

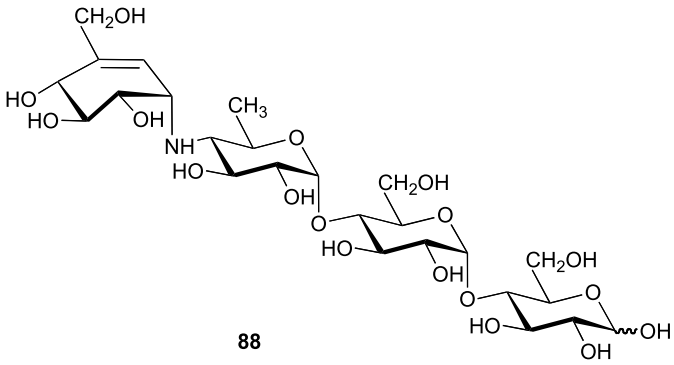
3.1 Glycosidase Inhibition

3.1.1 α-Glucosidase Inhibitors

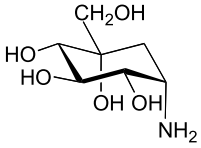
The intestinal oligo- and disaccharidases are fixed components of the cell membrane of the brush border region of the wall of the small intestine. These enzymes digest dietary carbohydrate to monosaccharides which are absorbed through the intestinal wall. They include sucrase, maltase, isomaltase, lactase, trehalase, and hetero-β-glucosidase. In the late 1970s, it was realized that inhibition of all or some of these activities by inhibitors could regulate the absorption of carbohydrate and these inhibitors could be used therapeutically in the oral treatment of the non-insulin-dependent diabetes mellitus (NIDDM or type 2 diabetes) [11]. Acarbose (**88**; see *Fig. 8* for structures of **88**–**92**) is a potent inhibitor of pig intestinal sucrase with an IC_{50} value of 0.5 µM and this was also effective in carbohydrate loading tests in rats and healthy volunteers, reducing postprandial blood glucose and increasing insulin secretion [106]. After intensive clinical development, acarbose (Glucobay) was first launched in Germany in 1990 and has since then been successfully marketed in Europe and Latin America. In 1996, it was introduced onto the market in the United States under the brand name Precose.

Valiolamine (**89**), an aminocyclitol produced by *Streptomyces hygroscopicus* var. *limoneus*, is a potent inhibitor of pig intestinal maltase and sucrase, with IC_{50} values of 2.2 and 0.049 µM, respectively [107]. Numerous *N*-substituted valiolamine derivatives were synthesized to enhance its α-glucosidase inhibitory activity in vitro and the very simple derivative voglibose (**90**), which is obtained by reductive amination of valiolamine with dihydroxyacetone, was selected as the potential oral antidiabetic agent [108]. Its IC_{50} values toward maltase and sucrase were 0.015 and 0.0046 µM, respectively. Voglibose (the brand name Basen) has been commercially available for the treatment of type 2 diabetes in Japan since 1994.

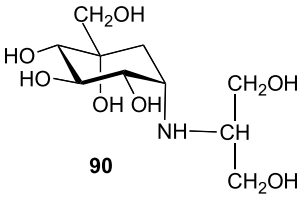
Mulberry leaves have traditionally been used to cure "Xiao-ke" (diabetes) in Chinese medicine. The strong inhibition of digestive α-glucosidases by DNJ attracted the interest of various research groups and a large number of *N*-substituted DNJ derivatives were prepared in the hope of increasing the in vivo activity. Miglitol (**91**) was identified as one of the most favorable candidates showing a desired glucosidase inhibitory profile [109]. Miglitol differs from acarbose in that it is almost completely absorbed from the intestinal tract, and may possess systemic effects in addition to the effects in the intestinal border [110,111]. In 1996, Glyset (miglitol) tablets were granted market clearance by the US Food and Drug Administration (FDA) and introduced onto the market in 1999 as a more effective second-generation α-glucosidase inhibitor with fewer gastrointestinal side effects. In 2006, it was introduced onto the market in Japan under the brand name Seibule.



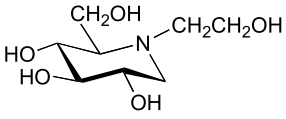
88



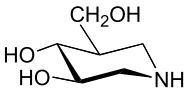
89



90



91



92

Figure 8
Structures of 88–92

α-Glucosidase inhibitors are especially suited for patients whose blood glucose levels are slightly above normal and can also benefit those who have high blood glucose right after they eat, a condition known as postprandial hyperglycemia. These drugs slow the rate at which carbohydrates are broken down into monosaccharides in the digestive tract and therefore lengthen the digestive process. Other antidiabetic agents such as sulfonylureas and biguanides sometimes are prescribed in combination with α-glucosidase inhibitors to help increase the effectiveness of this therapy. Protective effects of the α-glucosidase inhibitors have been reported for various diabetic complications. Interestingly, α-glucosidase inhibitors are also being studied as a possible treatment for heart disease, a common complication in diabetic patients. Although repetitive postprandial hyperglycemia increases ischemia/reperfusion injury, this effect can be prevented by treatment with α-glucosidase inhibitors [112].

3.1.2 Glycogen Phosphorylase Inhibitors

In type 2 diabetes, hepatic glucose production is increased [113]. A possible way to suppress hepatic glucose production and lower blood glucose in type 2 diabetes patients may be through inhibition of hepatic glycogen phosphorylase [114]. In enzyme assay, Fosgerau et al. reported

that DAB (**27**) is a potent inhibitor of hepatic glycogen phosphorylase [115]. Furthermore, in primary rat hepatocytes, DAB was shown to be the most potent inhibitor (IC_{50} 1 μM) of basal and glucagon-stimulated glycogenolysis ever reported [116]. Recently, Jakobsen et al. have reported that (3*R*,4*R*,5*R*)-5-hydroxymethylpiperidine-3,4-diol (isofagomine) (**92**) synthesized chemically is a potent inhibitor of hepatic glycogen phosphorylase, with an IC_{50} value of 0.7 μM, and furthermore, is able to prevent basal and glucagon-stimulated glycogen degradation in cultured hepatocytes with IC_{50} values of 2–3 μM [117]. However, its *N*-substitution always resulted in a loss of activity compared to the parent compound, and fagomine ((2*R*,3*R*,4*R*)-5-hydroxymethylpiperidine-3,4-diol) (**12**) was a weak inhibitor of this enzyme, with an IC_{50} value of 200 μM [117]. Glycogen phosphorylase inhibitors would be a beneficial target to attack in the development of new antihyperglycemic agents.

3.1.3 Herbal Medicines

Current scientific evidence demonstrates that morbidity and mortality of diabetes can be eliminated by aggressive treatment with diet, exercise, and new pharmacological approaches to achieve better control of blood glucose levels. In recent years, the possibility of preventing the onset of diabetes using dietary supplements and/or herbal medicines has attracted increasing attention. Mulberry leaves have been shown to have some antidiabetic properties. It was found that the mulberry leaf extract administered in a single dose of 200 mg/kg led to significant improvements in blood glucose levels in streptozotocin (STZ)-induced diabetic mice [118]. A study of 24 humans with type 2 diabetes found that patients treated with the mulberry agent had significant improvement in blood glucose control compared to a group treated with glibenclamide [119]. It is known that DNJ, fagomine, and DAB are contained in the mulberry leaves [14]. The evaluation for antihyperglycemic effects in STZ-induced diabetic mice has been carried out with fagomine [120]. Fagomine significantly reduced the blood glucose level 2 hours after intraperitoneal administration and its effect was sustained over 2 to 6 hours after administration. The effect of fagomine on immunoreactive insulin (IRI) release was investigated with the perfused pancreas of normal rats. The 8.3 mM glucose-induced IRI release was increased in the presence of fagomine in a concentration-dependent manner. The antihyperglycemic effects of the mulberry leaf extract would be a combination of α-glucosidase inhibition by DNJ and related compounds, the insulin releasing effect by fagomine, and glycogen phosphorylase inhibition by DAB.

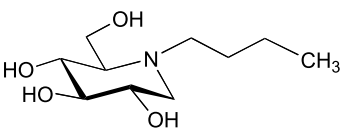
Thus, although traditional herbal medicines which have been used for a long time are candidates for diabetes prevention, it is very important to give scientific evidence for their antidiabetic effects. *Commelina communis*(Commelinaceae) has been used in traditional Chinese medicine as an antipyretic for noninfectious fever and to treat ascites, edema, and hordeolum [121], and is now very popular in Korea for the treatment of diabetes [122]. The MeOH extract of this plant shows a strong inhibitory activity against porcine intestinal α-glucosidases and contains DNJ (**4**), *manno*-DNJ (**5**), α-HNJ (**17**), 7-*O*-β-D-glucopyranosyl-α-HNJ, and DMDP (**25**) [123]. Alkaloids **4**, **17**, and 7-*O*-β-D-glucopyranosyl-α-HNJ are very potent inhibitors of digestive α-glucosidases, and DMDP as well as fagomine shows antihyperglycemic effects in STZ-induced diabetic mice [120]. These results support the pharmacological basis of this herb which has been used as a folklore medicine for the treatment of diabetes.

3.2 Molecular Therapy for Lysosomal Storage Disorders

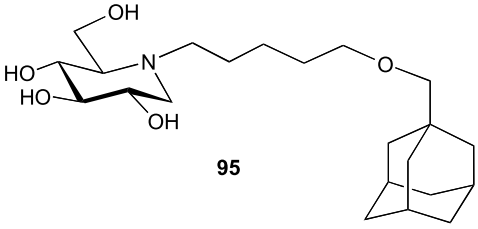
Recent experimental data show that some human genetic diseases are due to mutations in proteins that influence their folding and lead to retaining of mutant proteins in the ER and successive degradation [124,125]. Lysosomes are membrane-bound cytoplasmic organelles that serve as a major degradative compartment in eukaryotic cells. The degradative function of lysosomes is carried out by more than 50 acid-dependent hydrolases contained within the lumen [126]. The glycosphingolipid (GSL) storage diseases are genetic disorders in which a mutation of one of the GSL glycohydrolases blocks GSL degradation, leading to lysosomal accumulation of undegraded GSL [127]. Possible strategies for the treatment of these lysosomal storage diseases include enzyme replacement therapy, gene therapy, substrate deprivation, and bone marrow transplantation. The successful treatment for such diseases to date is the enzyme replacement therapy for patients with type 1 Gaucher disease and Fabry disease. However, this enzyme replacement therapy is useful only in the absence of neuropathology since enzymes do not cross the blood–brain barrier, and another problem with this therapy is the cost, which prevents many patients from obtaining this treatment. In recent years, remarkable progress has been made in developing a molecular therapy for the GSL storage disorders [3,4,128,129]. There are two novel approaches in this field. One is substrate reduction therapy and another is pharmacological (or chemical) chaperone therapy.

3.2.1 Substrate Reduction Therapy

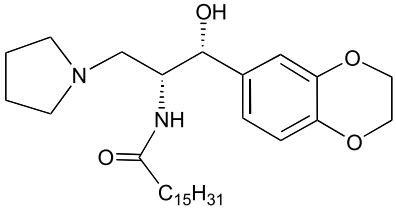
As long as the biosynthesis of substrate continues under the decrease of corresponding enzyme activity, the pathological accumulation of undegraded substrate in the lysosomes proceeds.



93



95



94

Figure 9
Structures of 93–95

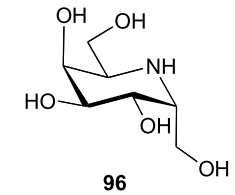
The aim of substrate reduction therapy is to reduce GSL substrate influx into the lysosomes by inhibitors of GSL synthesis. *N*-Butyl-DNJ (miglustat, Zavesca) (**93**; see *Fig. 9* for the structures of **93**–**95**) is an inhibitor of ceramide-specific glucosyltransferase [130]. Miglustat is the first orally active agent in the treatment of type 1 Gaucher disease. Gaucher disease is the most common lysosomal storage disorder caused by a deficiency of lysosomal β-glucosidase (known as β-glucocerebrosidase), resulting in the progressive accumulation of glucosylceramide. Type 1 Gaucher disease is non-neuronopathic and sometimes called the "adult" form. Ceredase in 1991 and its recombinant successor Cerezyme in 1994 were introduced as the enzyme replacement therapy of this type.

In 2001, Genzyme released preclinical data supporting Genz-78132 (the 3′ of D-threo-1-phenyl-2-palmitoilamino-3-pyrrolidino-1-propanol) (**94**) as the second-generation substrate reduction agent. The company has reported that Genz-78132 is 100–5000 times more potent in vitro for inhibition of cell surface ganglioside GM1, an indicator of glycosphingolipid synthesis, than the first-generation miglustat and *N*-(5-adamantane-1-yl-methoxypentyl)-DNJ (**95**). Furthermore, Genz-78132 is at least 20 times more potent in vivo than **95**, the most potent DNJ derivative [131]. Although there is no supporting clinical data, in preclinical studies Genz-78132 has a substantially greater therapeutic index than the first-generation inhibitors, which have shown limited efficacy and significant toxicity.

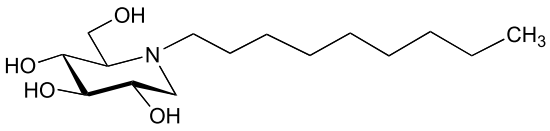
3.2.2 Pharmacological Chaperone Therapy

The concept of pharmacological chaperone therapy is that intracellular activities of misfolded mutant enzymes can be restored by administering competitive inhibitors that serve as pharmacological chaperones. These inhibitors appear to act as a template that stabilizes the native folding state in the ER by occupying the active site of the mutant enzyme, thus allowing its maturation and trafficking to the lysosome [128]. This concept was first introduced with Fabry disease [132]. Residual α-galactosidase A (α-Gal A) activity in lymphoblasts derived from Fabry patients and in tissues of R301Q α-Gal A transgenic mice was enhanced by treatment with *galacto*-DNJ (**6**), a competitive inhibitor of α-Gal A with a K_i value of 40 nM. Very recently, Yam et al. showed that *galacto*-DNJ induces trafficking of ER-retained R301Q α-Gal A to lysosomes of transgenic mouse fibroblasts and *galacto*-DNJ treatment results in efficient clearance of the substrate, globotriaosylceramide (Gb3) [133]. By testing a series of α-Gal A inhibitors for both in vitro inhibitory and chaperoning activities in lymphoblasts from Fabry patients, it was demonstrated that a potent inhibitor shows an effective chaperoning activity, whereas less potent inhibitors require higher concentrations to achieve the same effect [134]. *Galacto*-DNJ, α-*galacto*-HNJ (α-homogalactonojirimycin) (**96**; see *Fig. 10* for structures of **96**–**100**), **7**, and **21** are inhibitors of α-Gal A with IC_{50} values of 0.04, 0.21, 4.3, and 16 µM, respectively, and the respective addition at 100 µM to culture medium of Fabry lymphoblasts increases the intracellular α-Gal A activity by 14-, 5.2-, 2.4-, and 2.3-fold. Thus, potent and specific inhibitors of lysosomal glycosidases are expected to have therapeutic effects at lower concentrations.

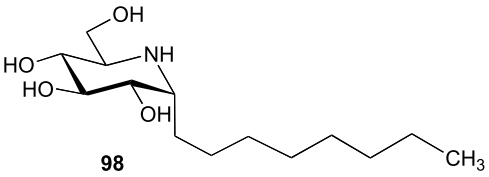
Sawker et al. reported that *N*-nonyl-DNJ (**97**) is a potent inhibitor of lysosomal β-glucosidase, with an IC_{50} value of 1 µM, and the addition of subinhibitory concentration (10 µM) of this compound to a fibroblast culture medium leads to a 2-fold increase in the mutant (N370S) enzyme activity [135]. Very recently, examination of a series of DNJ analogues on the resid-



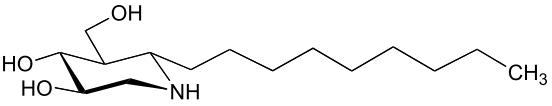
96



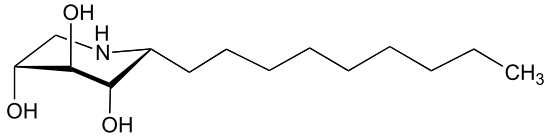
97



98



99



100

Figure 10
Structures of 96–100

ual activities of various lysosomal β-glucosidase variants has revealed that the nature of the alkyl moiety greatly influences their chaperoning activity: *N*-butyl-DNJ is inactive, the DNJ derivatives with *N*-nony and *N*-decyl chains are active, and *N*-dodecyl-DNJ is predominantly inhibitory [136]. However, it is also known that *N*-nonyl-DNJ is a potent inhibitor of ER

processing α-glucosidases like *N*-butyl-DNJ and hence has potential as anti-viral agents to inhibit folding and trafficking of viral envelope glycoproteins [137,138]. Inhibitors targeting a host function such as ER processing α-glucosidases must be carefully considered in terms of side-effects since they may inhibit folding, secretion, and trafficking of other glycoproteins in patient's cells or may inhibit directly lysosomal α-glucosidase after being taken up into cells. In fact, addition of *N*-nonyl-DNJ at 10 μM lowered the cellular lysosomal α-glucosidase activity by 50% throughout the assay period (10 days) in spite of the excellent chaperoning activity for the mutant β-glucosidase. The inhibition of lysosomal α-glucosidase as the side effect may induce storage of glycogen in the lysosomes, as observed in Pompe disease. On the other hand, α-1-*C*-octyl-DNJ (**98**), with a K_i value of 0.28 μM, showed a novel chaperoning activity for N370S Gaucher variants, minimizing the potential for undesirable side effects such as lysosomal α-glucosidase inhibition [139]. In addition, isofagomine [140], α-6-*C*-nonylisofagomine (**99**) [141], and α-1-*C*-nonyl-1,5-dideoxy-1,5-imino-D-xylitol (**100**) [142] are very potent inhibitors of lysosomal β-glucosidase and candidates as pharmacological chaperones for Gaucher disease.

A number of different pharmacological chaperones that correct folding states of mutant proteins and enable their trafficking to the proper locale would be excellent candidates for a new molecular therapy of human genetic disorders. It can be expected that orally administered pharmacological chaperones will become the standard treatment for broad genetic disorders with protein misfolding.

References

1. Asano N, Nash RJ, Molyneux RJ, Fleet GWJ (2000) Tetrahedron-Asymmetr 11:1
2. Watson AA, Fleet GWJ, Asano N, Molyneux RJ, Nash RJ (2001) Phytochemistry 56:265
3. Asano N (2003) Glycobiology 13:93R
4. Butters TD, Dwek RA, Platt FM (2005) Glycobiology 15:43R
5. Inoue S, Tsuruoka T, Niida T (1966) J Antibiot 19:288
6. Inoue S, Tsuruoka T, Ito T, Niida T (1968) Tetrahedron 24:2125
7. Niwa T, Tsuruoka T, Goi H, Kodama Y, Itoh J, Inoue S, Yamada Y, Niida T, Nobe M, Ogawa Y (1984) J Antibiot 37:1579
8. Miyake Y, Ebata M (1988) Agric Biol Chem 52:661
9. Paulsen H, Sangster I (1967) Chem Ber 100:802
10. Yagi M, Kouno T, Aoyagi Y, Murai H (1976) Nippon Nogei Kagaku Kaishi 50:571
11. Schmidt DD, Frommer W, Müller L, Truscheit E (1979) Naturwissenshaften 66:584
12. Murao S, Miyata S (1980) Agric Biol Chem 44:219
13. Ezure Y, Murao S, Miyazaki K, Kawamata M (1985) Agric Biol Chem 49:1119
14. Asano N, Tomioka E, Kizu H, Matsui K (1994) Carbohydr Res 253:235
15. Asano N, Oseki K, Tomioka E, Kizu H, Matsui K (1994) Carbohydr Res 259:243
16. Fellows LE, Bell EA, Lynn DG, Pikiewicz F, Miura I, Nakanishi K (1979) J Chem Soc Chem Commun 1979:977
17. Kite GC, Fellows LE, Fleet GWJ, Liu PS, Scofield AM, Smith NG (1988) Tetrahedron Lett 29:6483
18. Molyneux RJ, Pan YT, Tropea JE, Elbein AD, Lawyer CH, Hughes DJ, Fleet GWJ (1993) J Nat Prod 56:1356
19. Ezure Y, Ojima N, Konno K, Miyazaki K, Yamada N, Sugiyama M (1988) J Antibiot 41:1142
20. Asano N, Yamauchi T, Kagamifuchi K, Shimizu N, Takahashi S, Takatsuka H, Ikeda K, Kizu H, Chuakul W, Kettawan A, Okamoto T (2005) J Nat Prod 68:1238
21. Hardick DJ, Hutchinson DW, Trew SJ, Wellington EMS (1992) Tetrahedron 48:6285

22. Furumoto T, Asano N, Kameda Y, Matsui K (1989) J Antibiot 42:1302
23. Ikeda K, Takahashi M, Nishida M, Miyauchi M, Kizu H, Kameda Y, Arisawa M, Watson AA, Nash RJ, Fleet GWJ, Asano N (2000) Carbohydr Res 323:73
24. Yamashita T, Yasuda K, Kizu H, Kameda Y, Watson AA, Nash RJ, Fleet GWJ, Asano N (2002) J Nat Prod 65:1875
25. Asano N, Yasuda K, Kizu H, Kato A, Fan JQ, Nash RJ, Fleet GWJ, Molyneux RJ (2001) Eur J Biochem 268:35
26. Takahata H, Banba Y, Sasatani M, Nemoto H, Kato A, Adachi I (2004) Tetrahedron 60:8199
27. Kato A, Kato N, Kano E, Adachi I, Ikeda K, Yu L, Okamoto T, Banba Y, Ouchi H, Takahata H, Asano N (2005) J Med Chem 48:2036
28. Koyama M, Sakamura S (1974) Agric Biol Chem 38:1111
29. Molyneux RJ, Benson M, Wong RY, Tropea JH, Elbein AD (1988) J Nat Prod 51:1198
30. Evans SV, Hayman AR, Fellows LE, Shing TKM, Derome AE, Fleet GWJ (1985) Tetrahedron Lett 26:1465
31. Kato A, Asano N, Kizu H, Matsui K, Watson AA, Nash RJ (1997) J Nat Prod 60:312
32. Asano N, Kato A, Miyauchi M, Kizu H, Tomimori T, Matsui K, Nash RJ, Molyneux RJ (1997) Eur J Biochem 248:296
33. Cenci de Bello I, Dorling P, Fellows LE, Winchester B (1984) FEBS Lett 176:61
34. Liu PS (1987) J Org Chem 52:4717
35. Rhinehart BL, Robinson KM, Liu PS, Payne AJ, Wheatley ME, Wagner SR (1987) J Pharmacol Exp Ther 241:915
36. Kite GC, Horn JM, Romero JT, Fellows LE, Lees DC, Scofield AM, Smith NG Phytochemistry 29:103
37. Asano N, Nishida M, Kizu H, Matsui K, Watson AA, Nash RJ (1997) J Nat Prod 60:98
38. Asano N, Nishida M, Kato A, Kizu H, Matsui K, Shimada Y, Itoh T, Baba M, Watson AA, Nash RJ, Lilley PmdeQ, Watkin DJ, Fleet GWJ (1998) J Med Chem 41:2565
39. Asano N, Kato A, Miyauchi M, Kizu H, Kameda Y, Watson AA, Nash RJ, Fleet GWJ (1998) J Nat Prod 61:625
40. Martin OR, Compain P, Kizu H, Asano N (1999) Bioorg Med Chem Lett 9:3171
41. Holt KE, Leeper FJ, Handa S (1994) J Chem Soc Perkin Trans 1994:31
42. Martin OR, Saavedra OM (1995) Tetrahedron Lett 36:799
43. Bruce I, Fleet GWJ, Cenci di Bello I, Winchester B (1992) Tetrahedron 48:10191
44. Fleet GWJ, Namgoong SK, Barker C, Baines S, Jacob GS, Winchester B (2989) Tetrahedron Lett 30:4439
45. Kitaoka M, Ichikawa K, Sakurai Y, Matsushita Y, Iijima Y, Akiyama T, Boriboon M (1993) Annu Rep Sankyo Res Lab 45:99
46. Saralamp W, Chuakul W, Temsiririrkkul T, Clayton T (1996) (eds) Medicinal plants in Thailand, vol 1. Mahidol University, Bangkok
47. Terada M, Sano M, Ishii AI, Kino H, Fukushima S, Noro T (1982) Nippon Yakurigaku Zasshi 79:93
48. Sakata K, Aoki K, Chang CF, Sakurai A, Murakoshi J (1978) Agric Biol Chem 42: 457
49. Welter A, Jadot J, Dardenne G, Marlier M, Casimir J (1976) Phytochemistry 15:747
50. Nash RJ, Bell EA, Williams JM (1985) Phytochemistry 24:1620
51. Shibata T, Nakayama O, Tsurumi Y, Okuhara M, Terano H, Kohsaka M (1988) J Antibiot 41: 296
52. Kayakiri H, Takase S, Setoi H, Uchida I, Terano H, Hashimoto M (1988) Tetrahedron Lett 29:1725
53. Myerscough PM, Fairbanks AJ, Jones AH, Bruce I, Choi SS, Fleet GWJ (1992) Tetrahedron 48:10177
54. Watson AA, Nash RJ, Wormald MR, Harvey DJ, Dealler S, Lees E, Asano N, Kizu H, Kato A, Griffiths RC, Cairns AJ, Fleet GWJ (1997) Phytochemistry 46:255
55. Kato A, Adachi I, Miyauchi M, Ikeda K, Komae T, Kizu H, Kameda Y, Watson AA, Nash RJ, Wormald MR, Fleet GWJ, Asano N (1999) Carbohydr Res 316:95
56. Takebayashi M, Hiranuma S, Kanie Y, Kajimoto T, Kanie O, Wong CH (1999) J Org Chem 64:5280
57. James LF, Van Kampen KR, Hartley WJ (1970) Pathol Vet 7:116
58. Hartley WJ, Baker DC, James LF (1989) In: James LF, Elbein AD, Molyneux RJ, Warren CD (eds) Swainsonine and related glycosidase inhibitors. Iowa State University Press, Ames, p 50
59. Molyneux RJ, James LF (1982) Science 216:190
60. Dorling PR, Huxtable CR, Vogel P (1978) Neuropathol Appl Neurobiol 4:285
61. Colegate SM, Dorling PR, Huxtable CR (1979) Aus J Chem 32:2257

62. Hohenschutz LD, Bell EA, Jewss PJ, Leworthy DP, Pryce RJ, Arnold E, Clardy J (1981) Phytochemistry 20:811
63. Molyneux RJ, Roitman JN, Duuheim G, Szumio T, Elbein AD (1986) Arch Biochem Biophys 251:450
64. Molyneux RJ, Tropea JE, Elbein AD (1990) J Nat Prod 53:609
65. Molyneux RJ, Pan YT, Tropea JE, Benson M, Kaushal GP, Elbein AD (1991) Biochemistry 30:9981
66. Kato A, Kano E, Adachi I, Molyneux RJ, Watson AA, Nash RJ, Fleet GWJ, Wormald MR, Kizu H, Ikeda K, Asano N (2003) Tetrahedron-Asymmetr 14:325
67. Pastuszak I, Molyneux RJ, James LF, Elbein AD (1990) Biochemistry 29:1886
68. Nash RJ, Fellows LE, Dring JV, Fleet GWJ, Derome AE, Hamor TA, Scofield AM, Watkin DJ (1988) Tetrahedron Lett 29:2487
69. Wrobel JT (1985) In: Brossi A (ed) The alkaloids: chemistry and pharmacology, vol 26. Academic Press, New York, p 327
70. Robins DJ (1995) Cordell GA (ed) The alkaloids: chemistry and pharmacology, vol 46. Academic Press, New York, p 1
71. Molyneux RJ, Benson M, Wong RY, Tropea JE, Elbein AD (1988) J Nat Prod 51:1198
72. Harris CM, Harris TM, Molyneux RJ, Tropea JE, Elbein AD (1989) Tetrahedron Lett 30:5685
73. Nash RJ, Fellows LE, Dring JV, Fleet GWJ, Girdhar A, Ramsden NG, Peach JM, Hegarty MP, Scofield AM (1990) Phytochemistry 29:111
74. Nash RJ, Fellows LE, Plant AC, Fleet GWJ, Derome AE, Baird PD, Hegarty MP, Scofield AM (1988) Tetrahedron 44:5959
75. Denmark SE, Martinborough EA (1999) J Am Chem Soc 121:3046
76. Denmark SE, Herbert B (1998) J Am Chem Soc 120:7357
77. Denmark SE, Herbert B (2000) J Org Chem 65:2887
78. Asano N, Kuroi H, Ikeda K, Kizu H., Kameda Y, Kato A, Adachi I, Watson AA, Nash RJ, Fleet GWJ (2000) Tetrahedron-Asymmetr 11:1
79. Rambaud L, Compain P, Martin OR (2001) Tetrahedron-Asymmetr 12:1807
80. Izquierdo I, Plaza MT, Franco F (2002) Tetrahedron-Asymmetr 13:1581
81. Chabaud L, Landais Y, Renaud P (2005) Org Lett 7:2587
82. Yamashita T, Yasuda K, Kizu H, Kameda Y, Watson AA, Nash RJ Fleet GWJ, Asano N (2002) J Nat Prod 65:1875
83. Shibano M, Tsukamoto D, Kusano G (1999) Chem Pharm Bull 47:907
84. Asano N, Ikeda K, Kasahara M, Arai Y, Kizu H (2004) J Nat Prod 67:846
85. Tepfer D, Goldmann A, Pamboukdjian N, Maille M, Lepingle A, Chevalier D, Denarie J, Rosenberg C (1988) J Bacteriol 170:1153
86. Goldmann A, Milat ML, Ducrot PH, Lallemand JY, Maille M, Lepingle A, Charpin I, Tepfer D (1990) Phytochemistry 29:2125
87. Asano N, Yamashita T, Yasuda K, Ikeda K, Kizu H, Kameda Y, Kato A, Nash RJ, Lee HS, Ryu KS (2001) J Agric Food Chem 49:4208
88. Schimming T, Tofern B, Mann P, Richter A, Jenett-Siems K, Dräger B, Asano N, Cupta P, Correa MD, Eich E (1998) Phytochemistry 49:1989
89. Bekkouche K, Daali Y, Cherkaoui S, Veuthey JL, Christen P (2001) Phytochemistry 58:455
90. Asano N, Yokoyama K, Sakurai M, Ikeda K, Kizu H, Kato A, Arisawa M, Höke D, Dräger B, Watson AA, Nash RJ (2001) Phytochemistry 57:721
91. Asano N, Kato A, Oseki K, Kizu H, Matsui K (1995) Eur J Biochem 229:369
92. Asano N, Kato A, Yokoyama Y, Miyauchi M, Yamamoto M, Kizu H (1995) Carbohydr Res 284:169
93. Asano N, Kato, Kizu H, Matsui K, Watson AA, Nash RJ (1996) Carbohydr Res 293:195
94. Kato A, Asano, N, Kizu H, Matsui K, Suzuki S, Arisawa M (1997) Phytochemistry 45:425
95. Asano N, Kato A, Miyauchi M, Kizu H, Tomimori T, Matsui K, Nash RJ, Molyneux RJ (1997) Eur J Biochem 248:296
96. Griffiths RC, Watson AA, Kizu H, Asano N, Sharp HJ, Jones MG, Wormald MR, Fleet GWJ, Nash RJ (1996) Tetrahedron Lett 37:3207
97. Nash RJ, Watson AA, Winters AL, Fleet GWJ, Wormald MR, Dealer S, Lees E, Asano N, Molyneux RJ (1998) In: Garland T, Barr C (eds) Toxic plants and other natural toxicants. CAB International, Wallingford, p 276
98. Asano N, Kato A, Kizu H, Matsui K, Griffiths RC, Jones MG, Watson AA, Nash RJ (1997) Carbohydr Res 304:173
99. Duclos O, Mondange M, Depezay JC (1992) Tetrahedron Lett 33:8061
100. Boyer FD, Lallemand JY (1994) Tetrahedron 50:10443

101. Johnson CR, Bis SJ (1995) J Org Chem 60:615
102. Skaanderup PR, Madsen R (2001) Chem Comm 2001:1106
103. Skaanderup PR, Madsen R (2003) J Org Chem 68:2115
104. Haraguchi M, Gorniak SL, Ikeda K, Minami Y, Kato A, Watson AA, Nash RJ, Molyneux RJ (2003) J Agric Food Chem 51:4995
105. Molyneux RJ, McKenzie RA, O'Sullivan BM, Elbein AD (1995) J Nat Prod 58:878
106. Puls W, Keup U, Krause HP, Thomas G, Hoffmeister (1977) Naturwissenschaften 64:536
107. Kameda Y, Asano N, Yoshikawa M, Takeuchi M, Yamaguchi T, Matsui K, Horii S, Fukase H (1984) J Antibiot 37:1301
108. Horii S, Fukase H, Matsuo T, Kameda Y, Asano N, Matsui K (1986) J Med Chem 29:1038
109. Junge B, Matzke M, Stoltefuss J (1996) In: Kuhlmann J, Puls W (eds) Handbook of Experimental Pharmacology, vol 119. Springer, Berlin Heidelberg New York, p 411
110. Joubert PH, Foukaridis GN, Bopape ML (1987) Eur J Clin Pharmacol 31:723
111. Joubert PH, Venter HL, Foukaridis GN (1990) Br J Clin Pharmacol 30:391
112. Franz S, Calvillo L, Tillmanns J, Elbing I, Dienesch C, Bischoff H, Ertl G, Bauersachs J (2005) FASEB J 19:591
113. Defronzo RA, Bonadonna RC, Ferrannini E (1992) Diabetes Care 15:318
114. Martin JL, Veluraja K, Ross K, Johnson LN, Fleet GWJ, Ramsden NG, Bruce I, Orchard MG, Oikonomakos NG, Papageorgiou AC, Leonidas DD, Tsitoura HS (1991) Biochemistry 30:10101
115. Fosgerau K, Westergaard N, Quistorff B, Grunner N, Kristiansen M, Lundgren K (2000) Arch Biochem Biophys 380:274
116. Andersen B, Rassov A, Westergaard N, Lundgren K (1999) Biochem J 342:545
117. Jakobsen P, Lundbeck JM, Kristiansen M, Breinholt J, Demuth H, Pawlas J, Torres Candra MP, Andersen B, westergaard N, Lundgren K, Asano N (2001) Bioorg Med Chem 9:733
118. Chen F, Nakashima N, Kimura I, Kimura M (1995) Yakugaku Zasshi 115:476
119. Andallu B, Suryakantham V, Lakshmi Srikanthi B, Reddy GK (2001) Clin Chim Acta 314:47
120. Nojima H, kimura I, Chen F, Sugihara Y, Haruno M, Kato A, Asano N (1998) J Nat Prod 61:397
121. Huang KC (1993) The pharmacology of Chinese herbs. CRC Press, Boca Raton, Fl, p 296
122. Kim OK, Park, SY, Cho KH (1991) Kor J Pharmacogn 22:225
123. Kim HS, Kim YH, Hong YS, Paek NS, Lee HS, Kim TH, Kim KW, Lee JJ (1999) Planta Med 65:437
124. Bychova VE, Ptitsyn OB (1995) FEBS Lett 359:6
125. Welch WJ, Brown CR (1996) Cell Stress Chap 1:109
126. de Duve C (1963) In: de Reuck AVS, Cameron MP (eds) Lysosome. Churchill, London, p 1
127. Kornfeld S, Mellman I (1989) Annu Rev Cell Biol 5:483
128. Fan JQ (2003) TREND Pharmacol Sci 24:355
129. Cohen FE (2003) Nature 426:905
130. Platt FM, Neises GR, Dwek RA, Butters TD (1994) J Biol Chem 269:8362
131. Overkleeft HS, Renkema GH, Neele J, Vianello P, Hung IO, Strijland A, van der Burg AM, Koomen GP, Pandit UK, Aerts JMFG (1998) J Biol Chem 273:26522
132. Fan JQ, Ishii S, Asano N, Suzuki Y (1999) Nat Med 5:112
133. Yam GHF, Zuber C, Roth J (2005) FASEB J 19:12
134. Asano N, Ishii S, Kizu H, Ikeda K, Yasuda K, Kato A, Martin OR (2000) Eur J Biochem 267:4179
135. Sawker AR, Chen WC, Beautler E, Wong CH, Baich WE, Kelly JW (2002) Proc Natl Acad Sci USA 99:15428
136. Sawker AR, Adamsky-Werner SL, Chen WC, Wong CH, Beaytler E, Zimmer KP, Kelly JW (2005) Chem Biol 12:1235
137. Block TM, Lu X, Mehta AS, Blumberg BS, Tennant B, Ebling M, Korba B, Lansky DM, Jacob GS, Dwek RA (1998) Nat Med 4:610
138. Zitzmann N, Mehta AN, Carrouée S, Butters TD, Platt FM, McCauley J, Blumberg BS, Dwek RA (1999) Proc Natl Acad Sci USA 96:11878
139. Yu L, Ikeda K, Kato A, Adachi I, Godin G, Compain P, Martin OR, Asano N (2006) Bioorg Med Chem 14:7736
140. Chang HH, Asano N, Ishii S, Ichikawa Y, Fan JQ (2006) FEBS J 273:4082
141. Zhu X, Sheth KA, Li S, Chang HH, Fan JQ (2005) Angew Chem Int Ed 44:7450
142. Compain P, Martin OR, Boucheron C, Godin G, Yu L, Ikeda K, Asano N (2006) ChemBioChem 7:1356

9.2 Carbasugars: Synthesis and Functions

Yoshiyuki Kobayashi
Daiichi Sankyo Research Institute,
4250 Executive Square, La Jolla, California 92037, USA
ykobayashi@daiichisankyo-us.com

Abstract

It is well recognized that glycosidase inhibitors are not only tools to elucidate the mechanism of a living system manipulated by glycoconjugates but also potential clinical drugs and insecticides by inducing the failure of glycoconjugates to perform their function. In this chapter, the syntheses and functions of natural glycosidase inhibitors (cyclophelitol, allosamidine, and trehazoilin), which possess highly oxygenated and functionalized cyclohexanes or cyclopentanes in their structures and are defined as carbasugars, and the structure and activity relationships (SAR) of their derivatives are described. Also, recently much attention has been focused on

In: *Glycoscience*. Fraser-Reid B, Tatsuta K, Thiem J (eds)
Chapter-DOI 10-1007/978-3-540-30429-6_49:

neuraminidase inhibitors as anti-influenza drugs since relenza, which was derived from sialic acid, and also, tamiflu, which is the artificial carbasugar designed as a transition state analogue in the hydrolysis pathway of substrates by neuraminidase, were launched in the market. Herein, the medicinal chemistry efforts to discover tamiflu and some efficient syntheses applicable to process chemistry are described. Finally, useful synthetic methodologies for carbasugar formation from sugars are also introduced in this chapter.

Keywords

Carbasugars; Cyclitols; Glycosidase inhibitors; Cyclophelitol; Allosamidin; Trehazolin; Neuraminidase inhibitor; Tamiflu; Intramolecular [3+2] cycloaddition; The Ferrier reaction

Abbreviations

CDMA	catalytic desymmetrization of *meso*-aziridine
Chx	cyclohexyl
DBU	1,8-diazabicyclo[5.4.0]undec-7-ene
DDMPO	diisobutylaluminum 2,6-di-*tert*-butyl-4-methoxyphenoxide
DEIPS	diethylisopropylsilyl
DIBALH	diisobutylaluminum hydride
DIPEA	diisopropylethylamine
DMAP	*N*,*N*-dimethylaminopyridine
DMF	*N*,*N*-dimethylformamide
2,2-DMP	2,2-dimethoxypropane
DMSO	dimethylsulfoxide
HA	hemagglutinin
HPLC	high pressure liquid chromatography
KHMDS	potassium (1,1,1,3,3,3)-hexamethyldisilazide
mCPBA	*m*-chloroperbenzoic acid
MOM	methoxymethyl
NA	neuraminidase
NBA	*N*-bromoacetamide
NBS	*N*-bromosucciimide
NIS	*N*-iodosucciimide
NMO	4-methylmorpholine *N*-oxide
PPTS	pyridinium *p*-toluenesulfonate
RCM	ring-closing metathesis
SEMCl	2-(trimethylsilyl)ethoxymethyl chloride
TBAI	tetrabutylammonium iodide
TBAF	tetrabutylammonium fluoride
TBDMS	*t*-butyldimethylsilyl
TBDPSCl	*t*-butyldiphenylsilyl chloride
TFA	trifluoroacetic acid
TFAA	trifluoroacetic anhydride
THAB	tetrahexylammonium bromide
TMSCN	trimethylsilyl cyanide

1 Introduction

Recently, a great deal of attention has been focused on the glycosidase inhibitors under the premise that glycoconjugates such as oligosaccharides, glycolipids, and glycoproteins play pivotal roles in living systems, and recently a scientific field called "glycogenomics" has been developed as a part of genomics science. Glycosidase inhibitors possess interesting enzyme-specific inhibitory activities, therefore, they are expected not only to be tools to elucidate the mechanisms of a living system manipulated by the glycoconjugates but also to be potential clinical drugs and insecticides by inducing the failure of glycoconjugates to perform their function: anti-obesity drugs, anti-diabetics, anti-fungal, and anti-viral including substances active towards the human immunodeficiency virus (HIV) [1,2,3,4,5,6,7].

Most of the glycosidase inhibitors are isolated from natural sources, and they possess interesting structures in their molecules. Some of them possess the highly functionalized and oxygenated cyclohexane or cyclopentane moieties. In general, these moieties are called cyclitols, and also in this chapter, glycosidase inhibitors possessing cyclitols in their molecular structures are defined as carbasugars. Biochemically, carbasugars and cylitols themselves are recognized as the pseudo-sugars in a living system, and they show interesting biological activities based on the structure similarity to sugars. In the meantime, chemically, such highly functionalized molecules are quite challenging targets for synthetic studies. The most interesting and significant points for the synthesis of glycosidase inhibitors possessing cyclitols are how we can form the frameworks of the cyclitols and how we can introduce the functional groups essential to generate their specific and interesting biological activities.

In the last few decades new generations of glycosidase inhibitors, e. g., cyclophelitol, allosamidin, mannnostatin, and trehazolin, have been isolated from natural sources, and the appearance of these natural products has contributed to the development of the new methodologies needed to perform the aforementioned synthetic tasks. Also, due to technical developments in the field of structure biology, the interaction between enzymes and inhibitors can now be visualized. This is helpful in the design of new scaffolds and elucidation of the detailed mechanisms of action involved. The creation of tamiflu, which is an artificially designed carbasugar showing neuraminidase (NA) inhibitory activity, is a good example favored by structure biology, and it has already been approved by US Food and Drug Administration (FDA) as an orally available anti-influenza drug.

Herein, I will describe a variety of synthetic strategies directed to total syntheses and chemical modifications of these glycosidase inhibitors and the structure and activity relationships of their derivatives. Also, the methodologies of carbasugar formation from sugars will be described.

2 Overview

One of the key points in the synthesis of carbasugars is the formation of the scaffolds of the cyclitol moieties, and three types of synthetic strategy can be considered:

(a) Transformation of sugars to cyclitols;
(b) Synthesis from nonsugar substrates by using asymmetric synthetic methodologies;
(c) Utility of the framework of intact cyclitols.

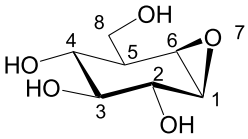
Also, in addition, much attention should be focused on the structure-activity relationships (SAR) on enzyme inhibitory activities influenced by their stereochemistry and their functional groups. Therefore, the synthetic strategies should be applicable to the synthesis of their stereoisomers and congeners, and, in practice, some of the synthetic strategies have also been designed in consideration of this point. In this chapter, case studies on total synthesis of cyclophelitol, tamiflu, allosamidin, and trehazolin, including the synthesis of their stereoisomers and congeners, will be described, and SAR of these compounds will be discussed. Finally, $SnCl_4$-promoted cyclitol formation and SmI_2-mediated cyclitol formation will be described as useful methodologies for the conversion of sugars to cyclitols as well as SmI_2-mediated carbasugar formation.

3 Synthesis of Carbasugars: New Generations of Glycosidase Inhibitors

3.1 New Generations of Glycosidase Inhibitors Bearing 6-Membered Cyclitol

3.1.1 Cyclophelitol: A β-Glucosidase Inhibitor

Cyclophelitol (**1**) is a β-glucosidase inhibitor isolated from a culture filtrate of the mushroom strain, *Phellinus* sp., and shows inhibitory activity towards almond β-glucosidase. In general, a series of β-glucosidase inhibitors such as castanospermine and 1-deoxynojirimycin have been reported to inhibit syncytium formation and infection with human immunodeficiency virus (HIV), possibly by perturbing the gp 120-linked glycan structure [8]. The structural feature of cyclophelitol is the fully oxygenated cyclohexane, which possesses the stereochemistry of the hydroxyl groups corresponding to that of D-glucose and the β-epoxy group at the C1, C6 positions. This β-epoxy group is recognized to be the equivalent of a β-glycoside moiety by enzyme, and it is considered that this is possibly the reason for the inhibitory activities exhibited towards β-glucosidases [9,10].

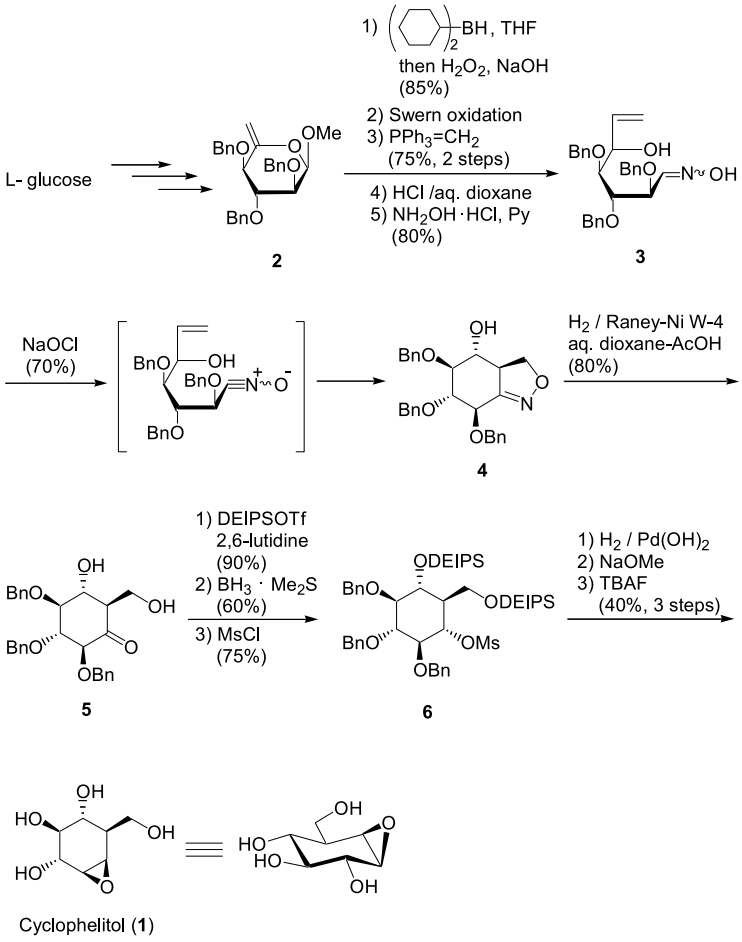


Cyclophelitol (**1**)

This interesting structure encouraged many organic chemists to undertake synthetic projects to develop appropriate synthetic methodologies, and several elegant syntheses of this compound have been reported.

The first total synthesis of cyclophelitol was achieved by Tatsuta's group in 1990 [11]. The key point of their synthetic strategy is the construction of this highly oxygenated cyclohexane via intramolecular [3+2] cycloaddition of a nitrile oxide to an alkene.

Oxime **3**, which was derived from L-glucose via stereoselective hydroboration of the *exo*-olefin of L-*xylo*-hex-5-enopyranoside **2** [12] with dicyclohexyl borane, Swern oxidation of the

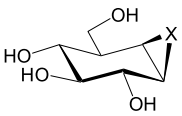


Scheme 1

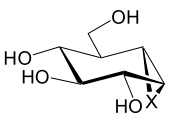
corresponding alcohol, subsequent Wittig alkenylation of the resulting unstable aldehyde with $Ph_3P{=}CH_2$, and acid hydrolysis, was treated with NaOCl to obtain isoxazoline **4** as a single product via the corresponding nitrile oxide [13,14]. The ring opening of isoxazoline **4** was achieved with H_2 and Raney nickel W-4 in aqueous dioxane in the presence of AcOH to afford the corresponding keto-diol **5**.

After diethylisopropylsilylation [15], which can be easily removed under hydrogenolysis conditions using $Pd(OH)_2$ on carbon, the resulting ketone was reduced with $BH_3{\cdot}SMe_2$ stereoselectively, to furnish the desired α-alcohol. Mesylation of this α-alcohol provided the labile mesylate **6**, which was subjected to hydrogenolysis with $Pd(OH)_2$ on carbon to give the de-*O*-benzylated compound. Epoxidation of this debenzylated compound with NaOMe and subsequent removal of the silyl groups of the corresponding epoxide with TBAF afforded cyclophelitol (**1**) (*Scheme 1*).

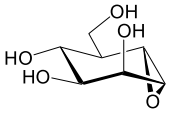
For further studies regarding structure-activity relationships on the enzyme inhibitory activities influenced by the stereochemistry of cyclophelitols, Tatsuta's group also synthesized a series of cyclophelitol-related compounds.



1:X=O (Cyclophelitol)
9:X=NH
11: X=S

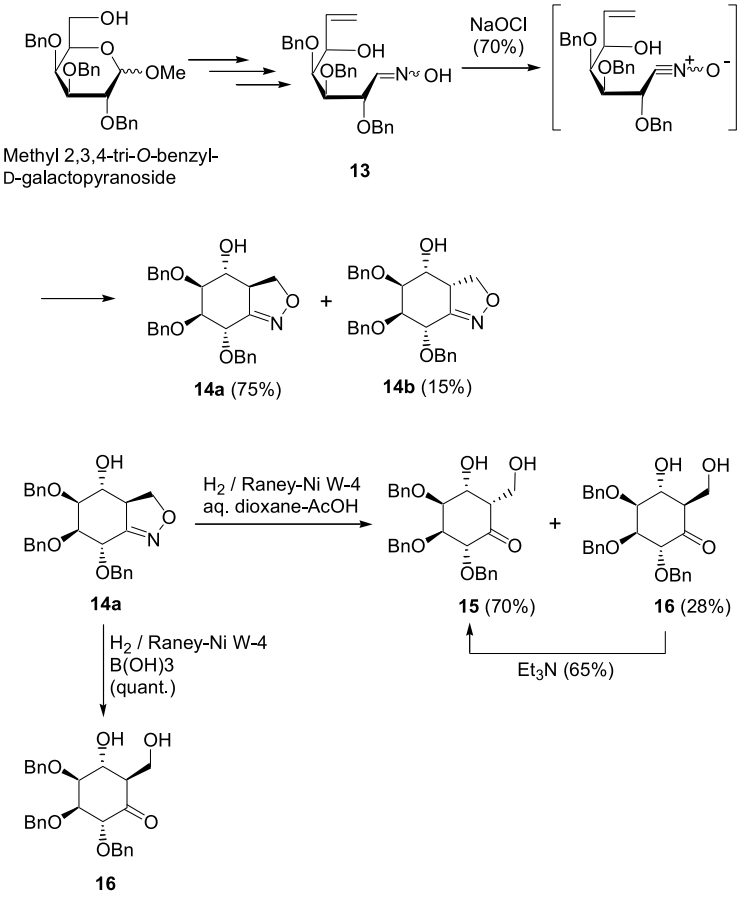


7: X=O (1,6-*epi*-Cyclophelitol)
10: X=NH
12: X=S

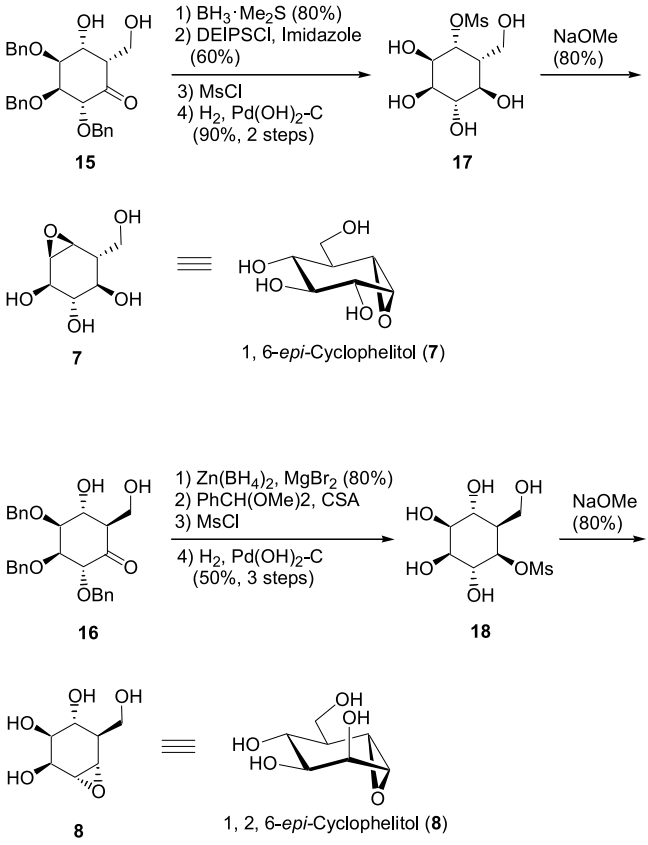


8: 1,2, 6-*epi*-Cyclophelitol
α-mannno-type analogue

The syntheses of 1,6-epi-cyclophelitol (**7**) and the 1,2,6--cyclophelitol *α-manno*-type analogue (**8**) are outlined in ❯ *Scheme 2* and ❯ *Scheme 3* [16]. These compounds were synthesized from



Scheme 2



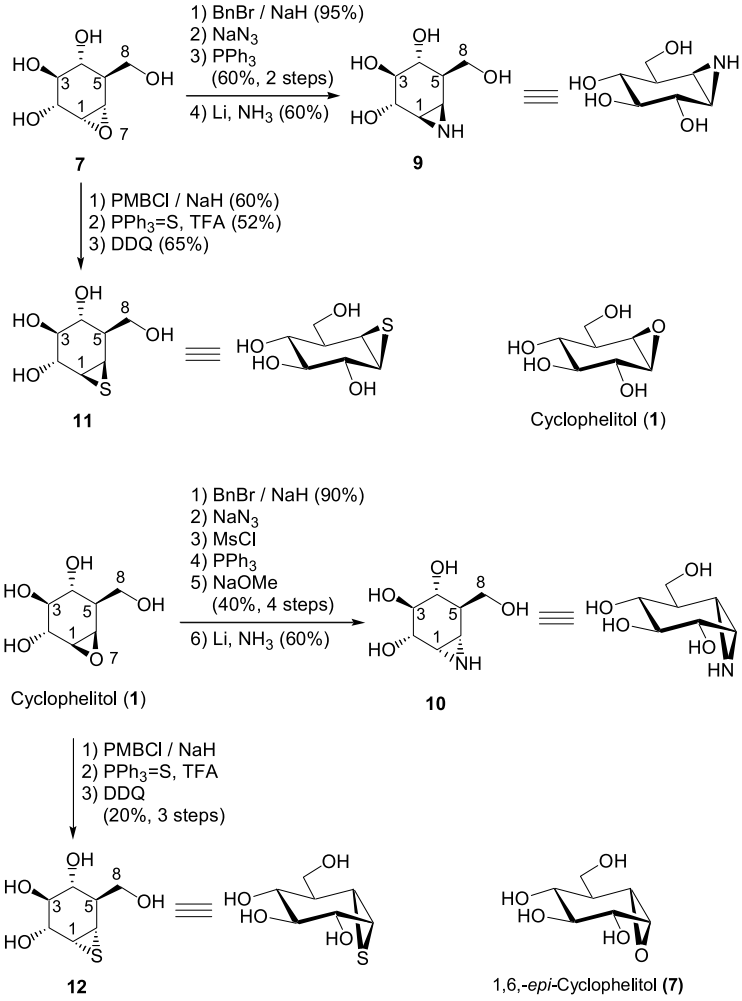
Scheme 3

a common intermediate, isoxazoline **14a**, which was derived from methyl 2,3,4-tri-*O*-benzyl-α-D-galactopyranoside via the above-mentioned intramolecular [3+2] cycloaddition. Interestingly, in the case of the synthesis of 1,6-*epi*-cyclophelitol (**7**), isoxazoline **14a** although substantially undesired to produce compound **7**, which is derived from oxime **13**, was subjected to acidic hydrogenolysis with Raney nickel W-4 to afford the desired keto-alcohol **15** with epimerization at the C1 position.

In contrast, in the case of the synthesis of the α-*manno*-type analogue, hydrogenolysis of isoxazoline **14a** which is desired to obtain compound **8** was conducted using Raney nickel and $B(OH)_3$ to furnish the desired keto-alcohol **16** in a quantitative yield (*Scheme 2*).

These keto-alcohols **15** and **16** were converted into 1,6-*epi*-cyclophelitol (**7**) and the 1,2,6-*epi*-cyclophelitol α-*manno*-type analogue (**8**), respectively, according to the synthetic route shown in *Scheme 3*.

Furthermore, on the basis of these syntheses, this group also synthesized the β-azirizine analogue (**9**) from 1,6-*epi*-cyclophelitol (**7**) and the α-azirizine analogue (**10**) from cyclophelitol (**1**), respectively [17,18]. In addition, the thiirane analogues (**11**) and (**12**) were also synthe-



Scheme 4

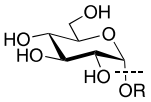
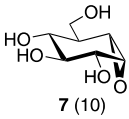
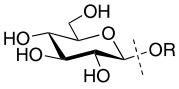
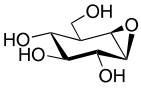
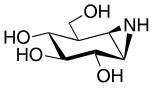
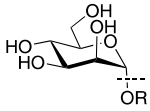
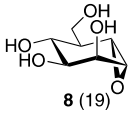
sized from 1,6-*epi*-cyclophelitol (**7**) and cyclophelitol (**1**), respectively, by treatment of their *O*-PMB derivatives with $Ph_3P{=}S$ and trifluoroacetic acid, followed by de-*O*-*p*-methoxybenzylation with DDQ (*Scheme 4*) [18].

As shown in *Table 1*, glycosidase inhibitory activities of these cyclophelitol-related compounds were evaluated [19].

In contrast to cyclophelitol (**1**) with inhibitory activity toward only β-D-glucosidase at an IC_{50} of 0.8 µg/ml, 1,6-*epi*-cyclophelitol (**7**) exhibited inhibitory activity toward only α-D-glucosidase at an IC_{50} of 10 µg/ml and α-*manno*-type analogue (**8**) showed inhibitory activity toward α-mannosidase at an IC_{50} of 19 µg/ml. Also, the β-azirizine analogue (**9**) exhibited high inhibitory activity toward β-glucosidase at an IC_{50} of 0.22 µg/ml, while the α-azirizine analogue (**10**) showed little α-glucosidase inhibitory activity. Interestingly, both thiirane ana-

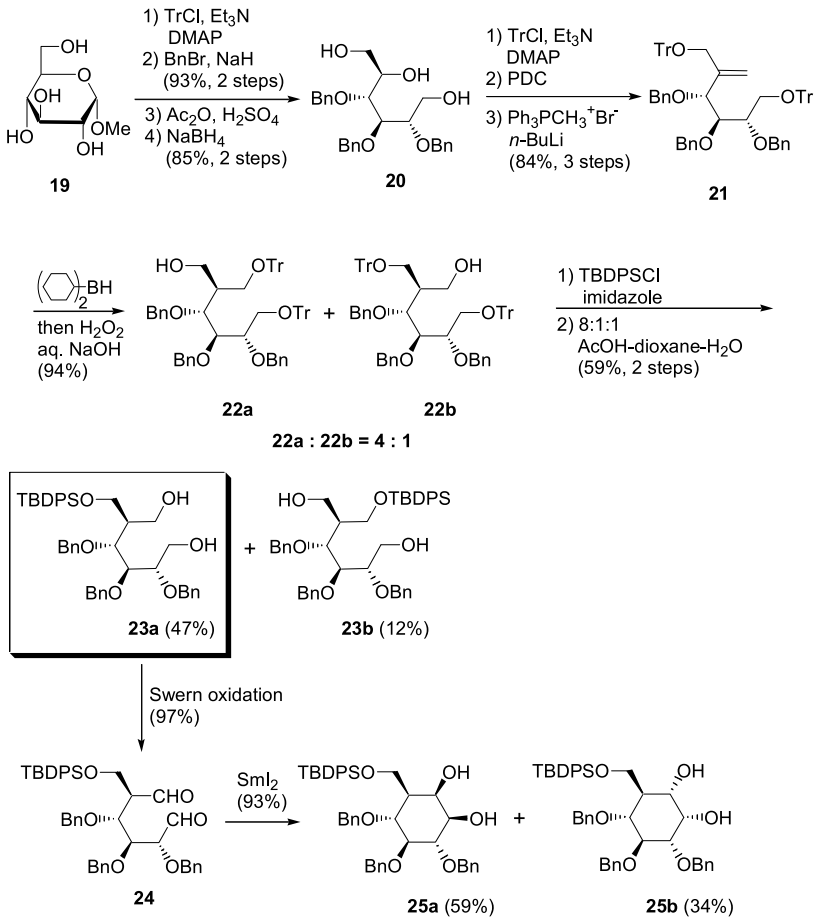
Table 1
Inhibitory activity of cyclophelitol (1) and its related compounds 7–9 towards glycosidases

Inhibitory activity of cycliphelitol (**1**) and its related compounds **7- 9** towards glycosidases

Glycosidase	Inhibitors (IC_{50} : μg/ml)		
α-D-glucosidase[a]		 **7** (10)	
β-D-glucosidase[b]		 **1** (0.8)	 **9** (0.22)
α-D-mannosidase[c]		 **8** (19)	

[a]Baker's yeast α-D-glucosidase; [b]Almond β-D-glucosidase;
[c]Snail β-D-mannosidase

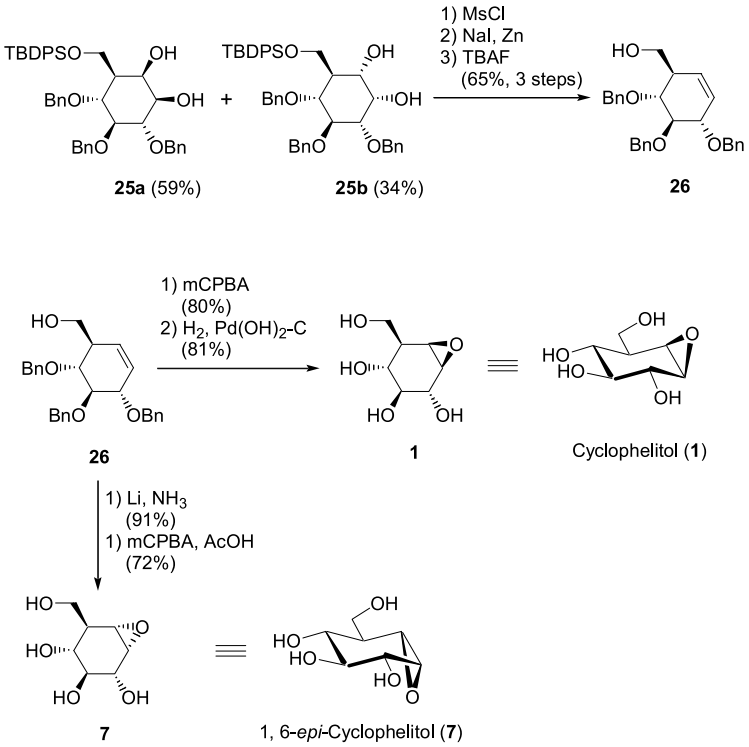
logues (**11**) and (**12**) showed no significant activities. With respect to this SAR on glycosidase inhibitory activities of these compounds, Tatsuta's group concluded the following;
Structurally, cyclophelitol (**1**) and its *β*-azirizine analogue (**9**) have *quasi*-equatorially oriented C1–O and C1–N bonds, which correspond to the equatorial C1–O bond *β*-glucopyranosides, whereas 1,6-*epi*-cyclophelitol (**7**) and *α*-*manno*-type analogue (**8**) have *quasi*-axial C1–O bonds corresponding to the axial C1–O bond of *α*-D-glycopyranosides. Their glycosidase inhibitory activities emphasized that the *α*- and *β*-glycosidase recognized especially the C-1 positions and the residual portions as corresponding to those of *α*- and *β*-glycopyranosides. Consequently, these glycosidase inhibitors (**1**), (**7**), (**8**), and (**9**) serve as antagonists of the corresponding *α*- and *β*-D-glycopyranosides.
Nakata et al. reported the straightforward and practical synthesis of cyclophelitol (**1**) and 1,6-*epi*-cyclophelitol (**7**) from D-glucose via SmI_2-mediated reductive coupling [20]. Hydroboration of olefin **21**, which was derived from methyl *α*-D-glucopyranoside (**19**) in seven steps, with dicyclohexylborane afforded the mixture of alcohols **22** (**22a**:**22b** = 4:1). After silylation with TBDPSCl, followed by selective acid hydrolysis of the Tr group, which furnished both the desired diol **23a** and the undesired one **23b** in 47 and 12% yield, respectively, subsequent Swern oxidation of the desired alcohol **23a** afforded dialdehyde **24**. This dialdehyde **24** was treated with SmI_2 [21,22,23,24,25], to give separable *cis*-diols **25a** and **25b** (*Scheme 5*).
Each diol was converted to the same olefin **26** by mesylation, reduction, and subsequent desilylation. Compound **26** was transformed to cyclophelitol (**1**) via stereoselective epoxidation, and on the other hand, debenzylation of **26** followed by epoxidation provided 1,6-*epi*-cyclophelitol (**7**) (*Scheme 6*).



Scheme 5

Fraser-Reid's group reported the synthesis of Tatsuta's penultimate intermediate, the mesylates for cyclophelitol (**1**) and 1,6-*epi*-cyclophelitol (**7**) via a 6-*exo-dig* radical cyclization of 2-deoxy-2-iodo-6-alkenyl glycoside and completed the formal total synthesis of cyclophelitol (**1**) and 1,6-*epi*-cyclophelitol (**7**) [26]. Swern oxidation of glucal **27**, alkynylation of the corresponding aldehyde and subsequent acetylation of the resulting alcohols afforded an epimeric mixture **28**. Thiem's iodoalkylation of **28** [27,28] provided the 2-deoxy-2-iodo-6-alkynyl glycoside **29**, and the radical cyclization of **29** with Bu_3SnH and AIBN [29,30,31,32] gave a quantitative yield of material which, after deacetylation, could be fractionated into two sets of isomers comprised of diastereomeric [2.2.2]oxabicycloglycoside **30** and 2-deoxy epimer **31**. Finally, compound **30** was converted into mono alcohol **35** in four steps (*Scheme 7*).

On the basis of the configuration at C1 (cyclophelitol numbering), the components of this diastereomeric mixture were divided into **35a** and **35b**, which were correlated to Tatsuta's intermediates. Compound **35a** was transformed to Tatsuta's mesylate **39** for 1,6-*epi*-

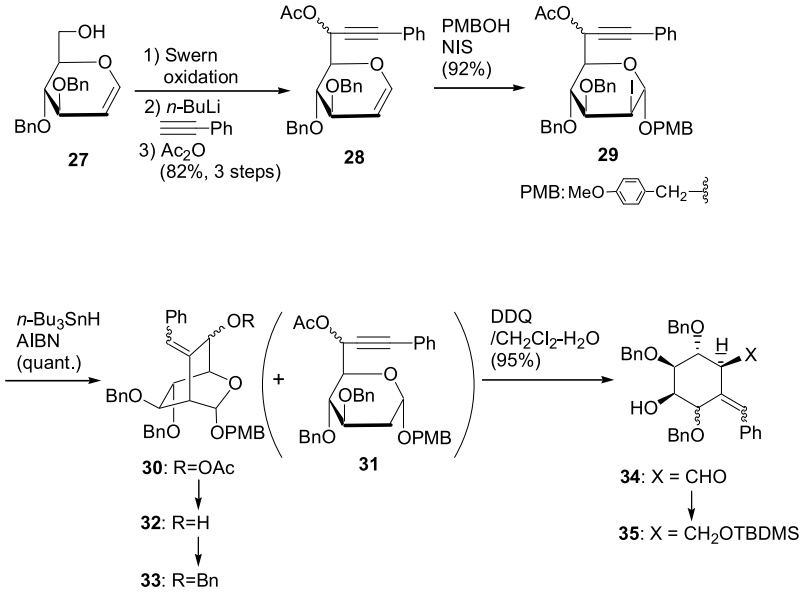


Scheme 6

cyclophelitol (**7**) in seven steps including C2–OH inversion via oxidation/reduction with the Dess–Martin reagent [33]/$NaBH_4$, ozonolysis followed by $BH_3{\cdot}SMe_2$ reduction, and mesylation of the corresponding alcohol. On the other hand, **35b** was also converted into Tatsuta's mesylate **47** for cyclophelitol (**1**) in ten steps involving ozonolysis followed by stereocontrolled reduction with $NaBH(OAc)_3$ [34,35] via complex **40** and C2–OH inversion via oxidation/reduction with the Dess–Martin reagent/$BH_3{\cdot}SMe_2$ (*Scheme 8*).

Ferrier reaction [36,37] is well-known as a representative of the synthetic methodologies for the conversion of intact sugars into cyclitols. It is not exceptional that Ferrier reaction was also used for total synthesis of cyclophelitol as a key step for cyclitol formation.

The synthesis reported by Sato et al. [38] is shown in *Scheme 9*. The features of this synthesis are the stereoselective introduction of a dichloromethyl function to C2-ketone **48** [39], which was derived from D-glucose, and the conversion of **49** with $NaBH_4$ in DMSO into the 2-deoxy-2-*C*-hydroxymethyl α-D-glucopyranoside derivative **50**, and furthermore, Ferrier reaction of *exo*-olefin **51**, which was derived from **50** in two steps, with $HgCl_2$ in acetone/H_2O, and subsequent treatment of the product with MsCl and Et_3N provided the desired α,β-unsaturated ketone **52**. Stereoselective 1,2-reduction of **52** with $NaBH_4$ and $CeCl_3{\cdot}7H_2O$, protection of the corresponding alcohol with a TBDMS group, and subsequent cleavage of acyl groups under the basic conditions afforded triol **53**. The epoxidation to the olefin part of triol **53** was proceeded stereoselectively favored by the bulky TBDMS group, to afford epoxide **54** which

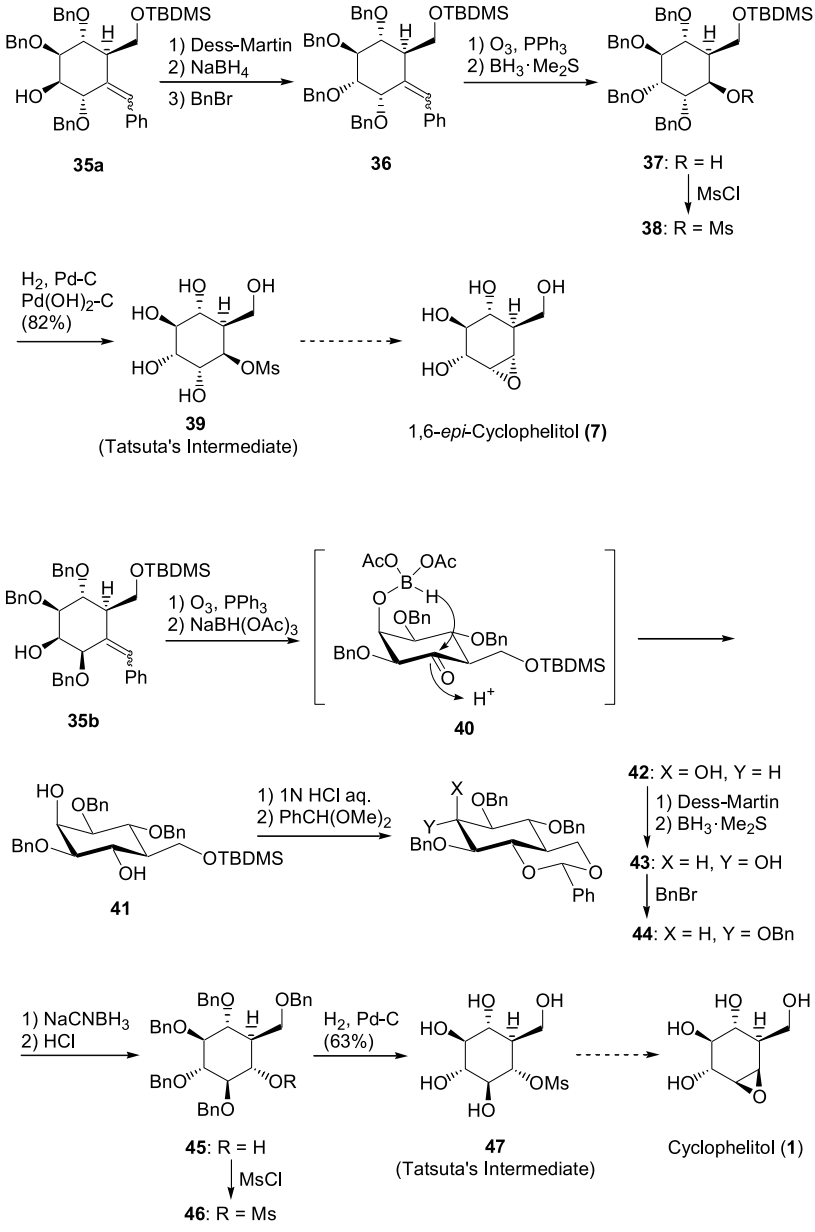


Scheme 7

possesses the desired stereochemistry, and subsequent acid desilylation furnished cyclophelitol (**1**) (*Scheme 9*).

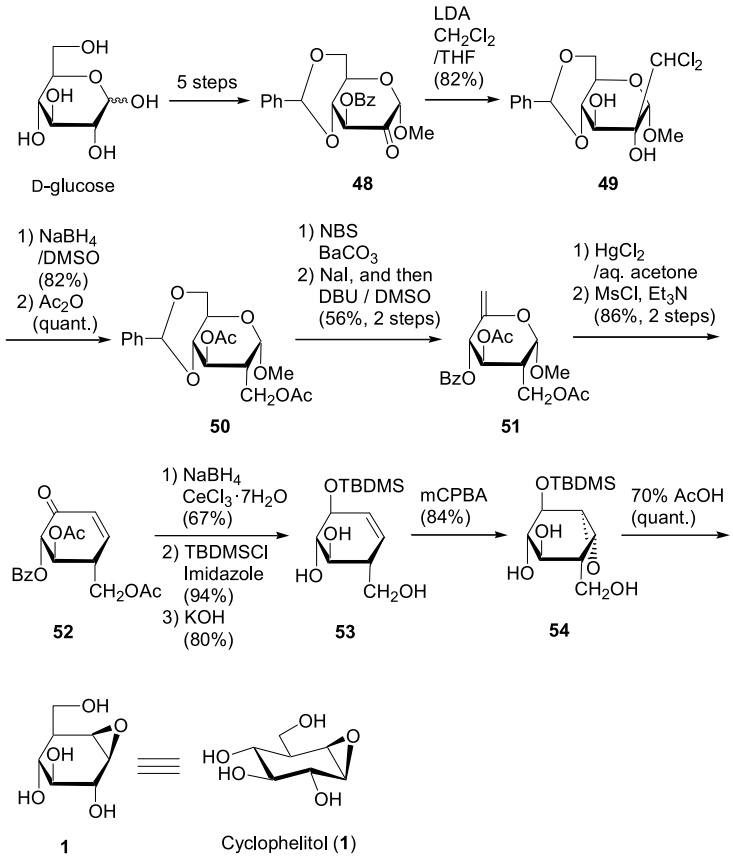
Also, Jung et al. reported the total synthesis of cyclophelitol (**1**) and 1,6-*epi*-cyclophelitol (**7**) using Ferrier reaction as its key steps for cyclitol formation [40]. The known methyl benzylidenemannoside **55** [41], which was derived from D-mannose in two steps, was converted into alkene **56** via regioselective benzylation of the equatorial hydroxyl group of **55** [42,43], Swern oxidation of the remaining hydroxyl group, and subsequent Wittig reaction. Hydroboration of **56** was expected to occur preferentially from the axial direction, *anti* to the sterically hindering axial methoxy group. However, unexpectedly, the reaction furnished a 1:1 mixture of two isomeric hydroxymethyl compounds **57a** and **57b**. Even with a sterically bulky borane such as 9-BBN the reaction afforded a 1:2 mixture in favor of **57b**. The problem of the production of exclusively the equatorial isomer was solved by Swern oxidation of the mixture, subsequent quantitative equilibration to the β-aldehyde in a mild base, and reduction of the corresponding α-aldehyde to give only the desired equatorial hydroxymethyl compound **57a**. Benzylation, reductive opening of the benzylidene acetal, iodination, and elimination furnished *exo*-olefin **58**. Ferrier reaction of **58** and subsequent elimination gave the desired enone **59**. 1,2-reduction of **59** to the corresponding allylic alcohol and subsequent benzoylation gave compound **60**, and epoxidation of **60** with mCPBA provided a mixture of the epoxides. Cleavage of all protective groups produced cyclophelitol (**1**) and 1,6-*epi*-cyclophelitol (**7**) (*Scheme 10*).

Recently, ring-closing metathesis (RCM) using Grubbs catalyst [44] became popular in the field of the synthesis of natural products. This methodology was also used in the synthesis of cyclophelitols. The synthesis using RCM was reported by Ziegler et al. and started from D-xylose as shown in *Scheme 11* [45]. Didithioacetal **61**, which was derived from D-xylose,



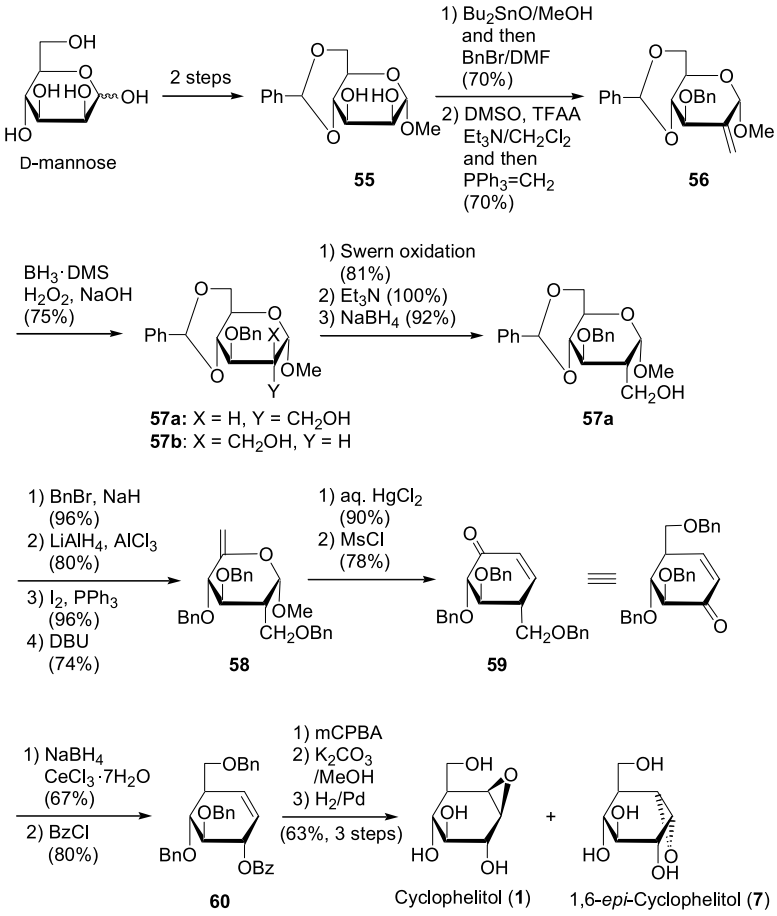
Scheme 8

was converted into aldehyde **63** in three steps that were silylation of the primary alcohol, benzylation of all the remaining secondary hydroxyl groups, and subsequent removal of dithioacetal with HgO and $HgCl_2$. Methylenation to aldehyde **63** was performed with Tebbe's reagent in place of the Wittig reaction because the basic methylenation with Wittig reaction gave sev-



Scheme 9

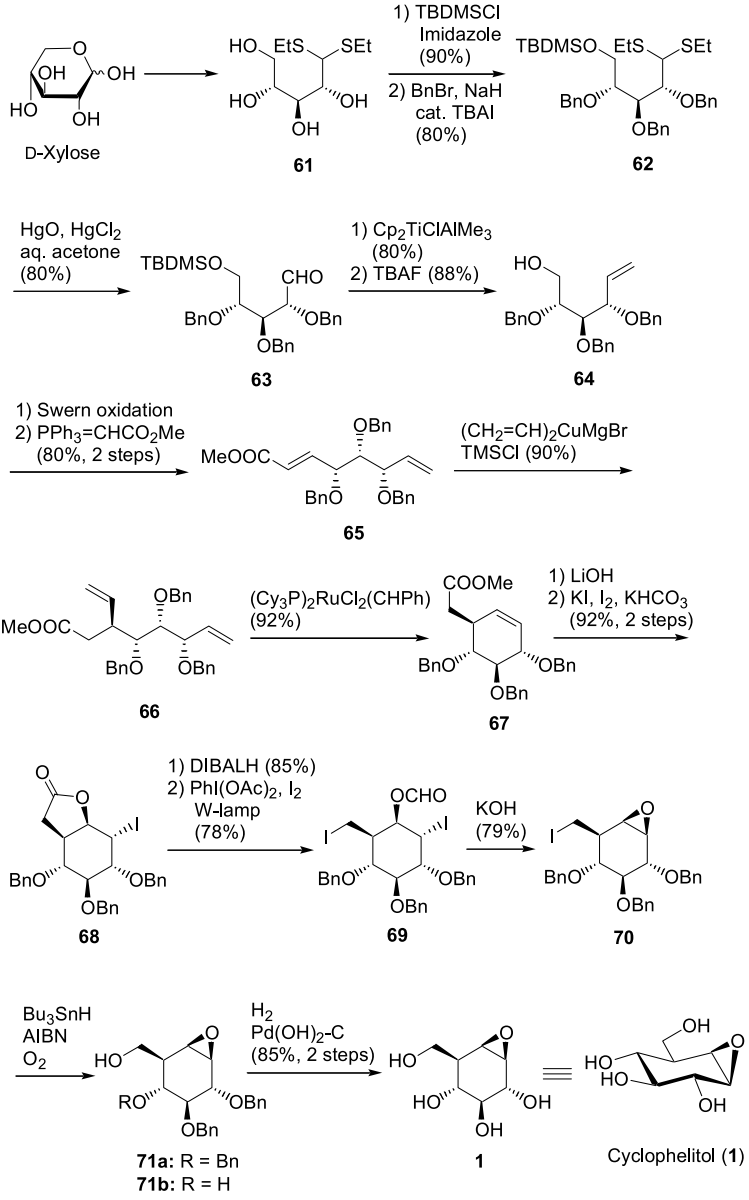
eral by-products such as the unsaturated aldehyde and its diene derived from β-elimination of benzyl alcohol from aldehyde **63** as well as the desired olefin **64**. After Swern oxidation of the remaining primary hydroxyl group, Wittig reaction of the corresponding aldehyde with PPh_3=$CHCO_2Me$ afforded α, β-unsaturated ester **65**. One of the problems to be solved here was the stereoselective 1,4-addition to α, β-unsaturated ester **65**. In this case, the best method for this objective is the magnesium-based vinyl cuprate, which was employed under the protocol reported by Hanessian [46] to give **66** with both high yield and high stereoselectivity. Afterwards, RCM to **66** using Grubbs catalyst provided cyclohexene **67** efficiently, and subsequent iodolactonization to **67** afforded iodide **68**. Iodide **68** was converted into the β-epoxide **70** according to the Saurez procedure [47,48]. After reduction of **68** with DIBALH, the resulting lactol efficiently afforded diiodideformate **69** upon irradiation in the presence of $PhI(OAc)_2/I_2$ [49]. Subsequently, base treatment of **69** gave iodoepoxide **70**. The transformation of iodine to a hydroxyl group was accomplished by radical oxygenation to afford a mixture of epoxyalcohol **71a** and epoxydiol **71b**. Finally, hydrogenolysis of a mixture of the epoxides **71a** and **71b** removed the benzyl groups to afford cyclophelitol (**1**) (*Scheme 11*).



Scheme 10

Kornienko's group reported the formal synthesis of cyclophelitol (**1**) via RCM to the diene derived from D-xylose [50]. Their synthetic strategy was interestingly focused on the latent plane of chirality present in D-xylose as shown in *Fig. 1*, and the enantiodivergent synthesis of (+)- and (–)-cyclophelitol from D-xylose was achieved (*Fig. 1*).

Their enantiodivergent strategy relies on the synthesis of both enantiomeric forms of Ziegler's enoate **72 (=65)**. They have reported a straightforward preparation of enoate **72** from 2,3,4-tri-*O*-benzyl-D-xylopyranose by way of Wittig methylenation, Swern oxidation, and reaction of the resulting aldehyde with $PPh_3{=}CHCO_2Me$ [51]. On the other hand, the synthesis of *ent*-**72** by directly reversing the order of the two olefination steps was not always successful because of an enoate reactivity problem. To circumvent this problem, the anomeric position of 2,3,4-tri-*O*-benzyl-D-xylopyranose was protected to form ethyl thioacetal, and subsequent Swern oxidation of the remaining primary hydroxyl group, followed by the intermediate treatment of the corresponding aldehyde with $PPh_3{=}CH_2$, deprotection of ethyl thioacetal, and olefination of the corresponding aldehyde with $PPh_3{=}CHCO_2Me$ cleanly provided the desired



Scheme 11

enantiomeric enoate *ent*-**72** in excellent yield. With the synthetic route to **1** and *ent*-**1** available, the feasibility of accessing each cyclophelitol enantiomer was demonstrated by completing the synthesis of the (+)-antipode. Addition of a vinylcopper reagent to enoate **72** provided ester **73** in excellent yield and exclusive *anti* selectivity. Treatment of potassium enolate of ester

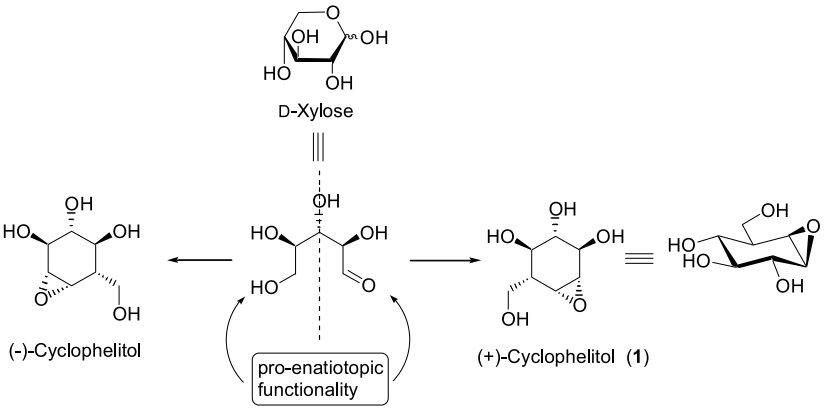
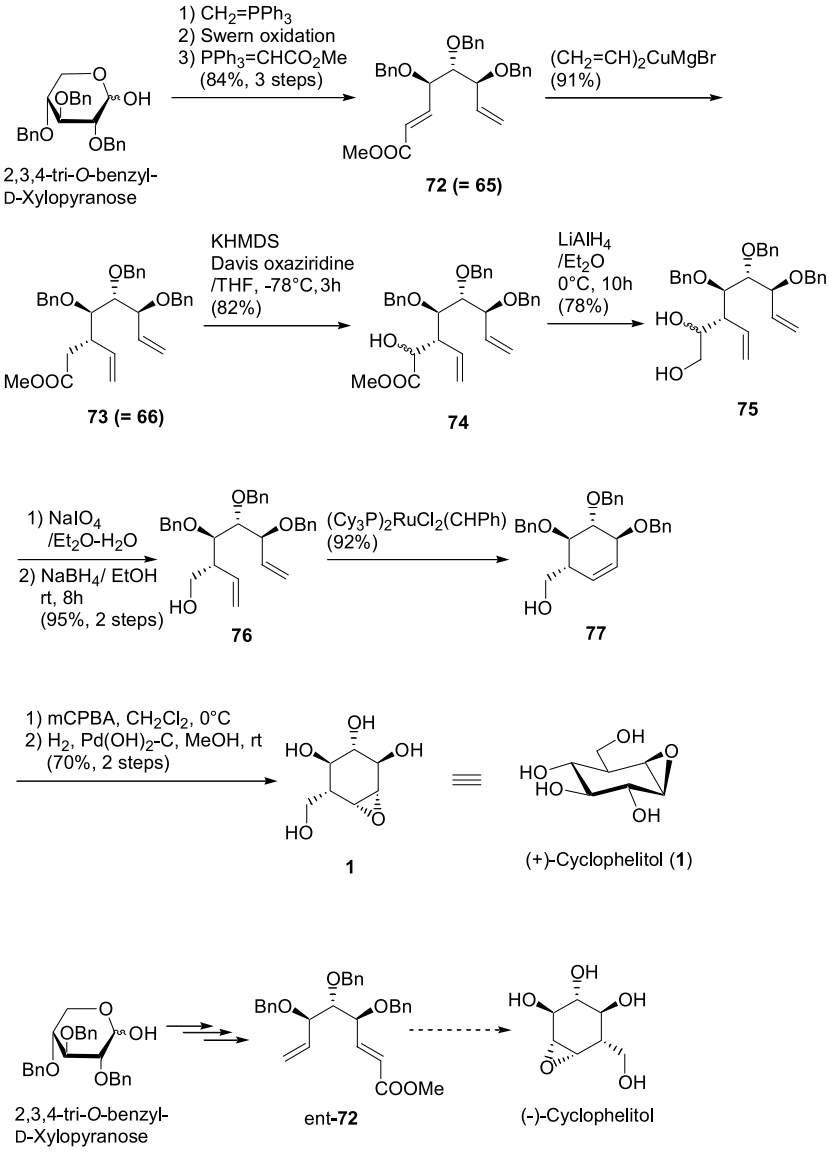


Figure 1
"Latent symmetry" approach for the synthesis of cyclitol enantiomers

73 with Davis oxaziridine gave a 1:1 mixture of α-hydroxylated derivatives **74**, which was reduced to diols **75**. Cleavage of the vicinal diol functionality with $NaIO_4$ followed by treatment of the corresponding aldehyde with $NaBH_4$ gave alcohol **76**. Finally, RCM of alcohol **76** with the first generation Grubbs' ruthenium catalyst resulted in conduritol analogue **77**. Since the transformation of **77** to (+)-cyclophelitol (**1**) by way of directed epoxidation and hydrogenolytic deprotection was reported by Trost's group [52,53], this pathway constituted a formal synthesis of the natural product (*Scheme 12*).

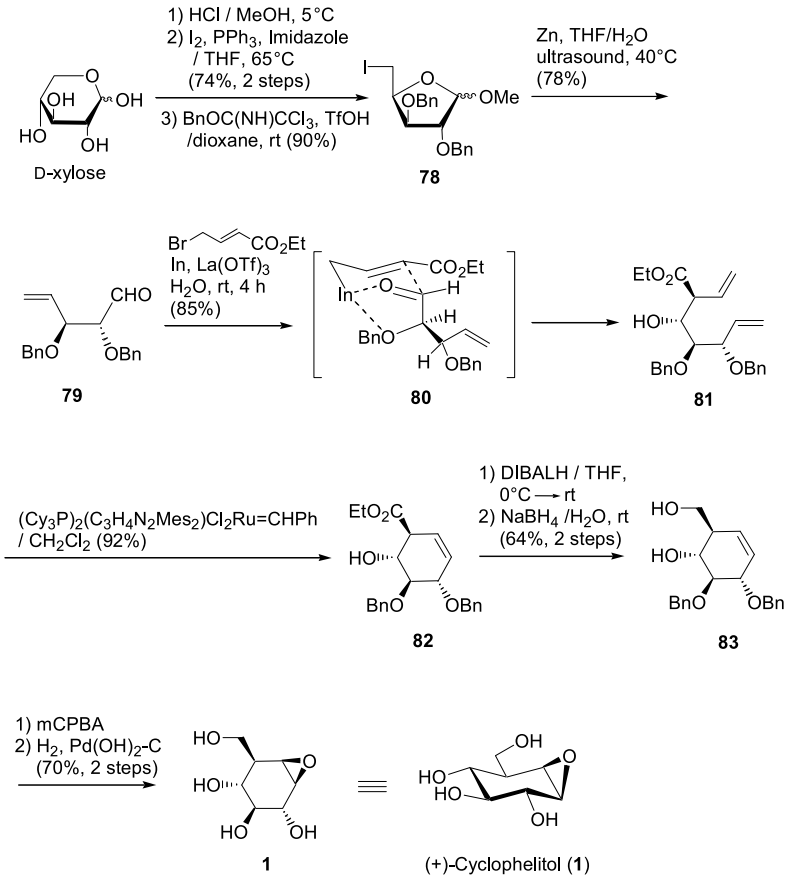
Madsen et al. also reported the synthesis of cyclophelitol (**1**) from D-xylose via RCM [54]. The key transformations involve a zinc-mediated fragmentation of benzyl protected methyl 5-deoxy-5-iodo-xylofuranoside (**78**) followed by a highly diastereoselective indium-mediated coupling with ethyl 4-bromocrotonate as shown in *Scheme 13*. Compound **78**, which was derived from D-xylose in three steps including acidic methyl glycosylation, iodination to primary alcohol, and subsequent benzylation to the remaining hydroxyl groups with benzyl trichloroacetimidate under acidic conditions, was sonicated with zinc to afford unsaturated aldehyde **79**. The indium-mediated addition reaction to **79** with ethyl 4-bromocrotonate and 60 mesh indium powder (99.999% pure) in the presence of $La(OTf)_3$ [55,56,57] provided the desired product **81** as only one diastereomer. The generation of the only desired diastereomer **81** can be explained by invoking a chelated intermediate such as **80**. Diene **81** was converted into the corresponding cyclohexene **82** by RCM with Grubbs' second generation catalyst [58]. The ester functionality was reduced with DIBALH to afford diol **83**. This reduction was quite sluggish and afforded minor amounts of the intermediate aldehyde as a byproduct. Using a larger excess of DIBALH or a higher temperature did not solve the problem. It is interesting that water and $NaBH_4$ instead were added in the workup to drive the reaction to completion. Epoxidation of **83** with mCPBA was completely stereoselective to afford the desired epoxide without another isomer, and finally the corresponding epoxide was deprotected by hydrogenolysis to furnish cyclophelitol (**1**) (*Scheme 13*).

An interesting approach to cyclophelitol (**1**) is its synthesis via the nonracemic Diels–Alder reaction to form the scaffold of cyclophelitol as performed by Schlessinger's group, shown



Scheme 12

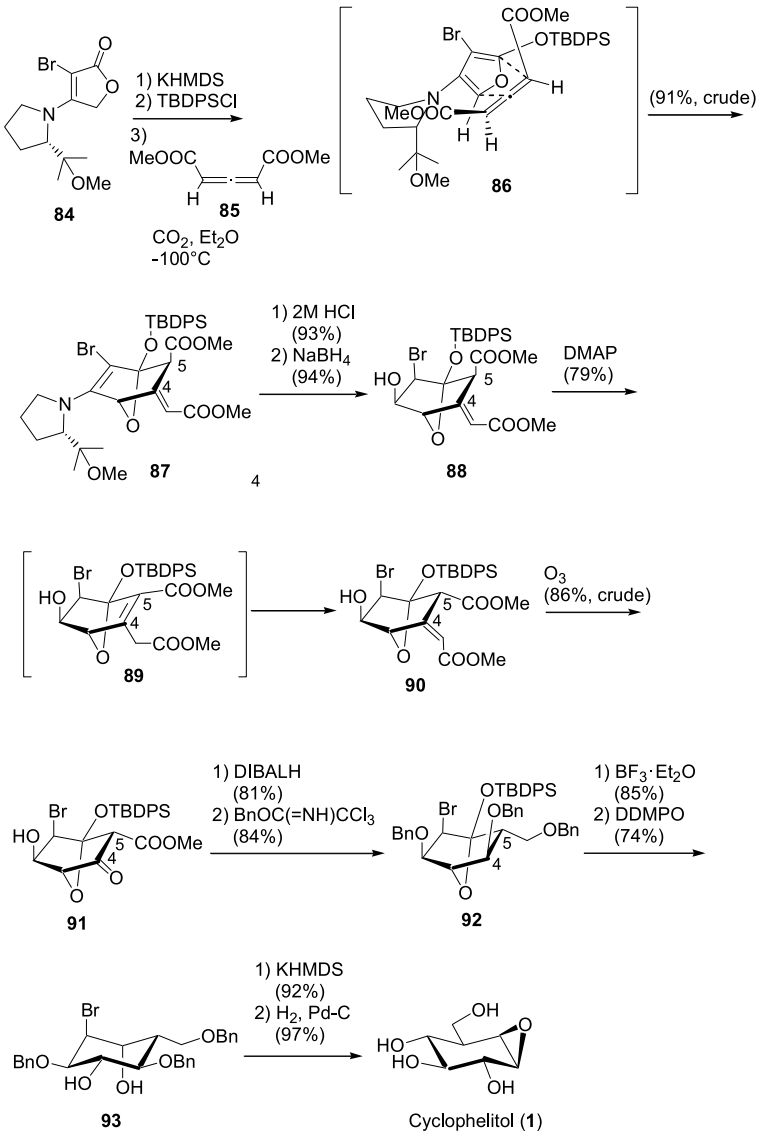
in *Scheme 14* [59]. To the nonracemic furan which was prepared by deprotonation of the vinylogous lactone **84**, followed by treatment of the resulting potassium enolate with TBDPSCl, dimethyl 2,3-pentadienedioate **85** was added [60], and the resulting compound **87**, which contained the requisite elements of structure and functionality for conversion to cyclophelitol (**1**), was obtained as a crude product via the proposed transition state **86**. Acid hydrolysis of **87**, and subsequent reduction of the corresponding bromoketone with $NaBH_4$ gave the *syn*-



Scheme 13

bromohydrin **88**, which possesses the correct functionality and stereochemistry at C1 and C2, respectively, as required by the structure (**1**). To reform C4 and C5 of **88** for the structure (**1**), firstly the stereochemistry of the axial ester at C5 was adjusted by treatment with DMAP via intermediate **89**, to afford the epimerized C5 equatorial ester **90**. Ozonolysis of **90**, reduction of the resulting ketone of **91** with DIBALH, and subsequent benzylation of the corresponding alcohol under acidic conditions afforded compound **92** possessing the correct stereochemistry at C4 for cyclophelitol synthesis. The oxabicycloketal moiety of tribenzyl ether **92** was cleaved smoothly with $BF_3{\cdot}Et_2O$ to afford the tractable cyclohexanone. Diisobutylaluminum 2,6-di-*t*-butyl-4-methylphenoxide (DDMPO) [61,62] treatment of this cyclohexanone gave **93** stereoselectively. After the transformation of bromohydrin **93** to the epoxide with KHMDS, hydrogenolysis of the corresponding, fully protected epoxide furnished cyclophelitol (**1**) (*Scheme 14*).

Finally for this section, as another asymmetric synthesis of cyclophelitol (**1**), the synthetic effort of Trost's group will be introduced [52]. Their synthesis was conducted based on the palladium-catalyzed kinetic resolution of racemic conduritol B tetraacetate (±)-**94**. Since racemic



Scheme 14

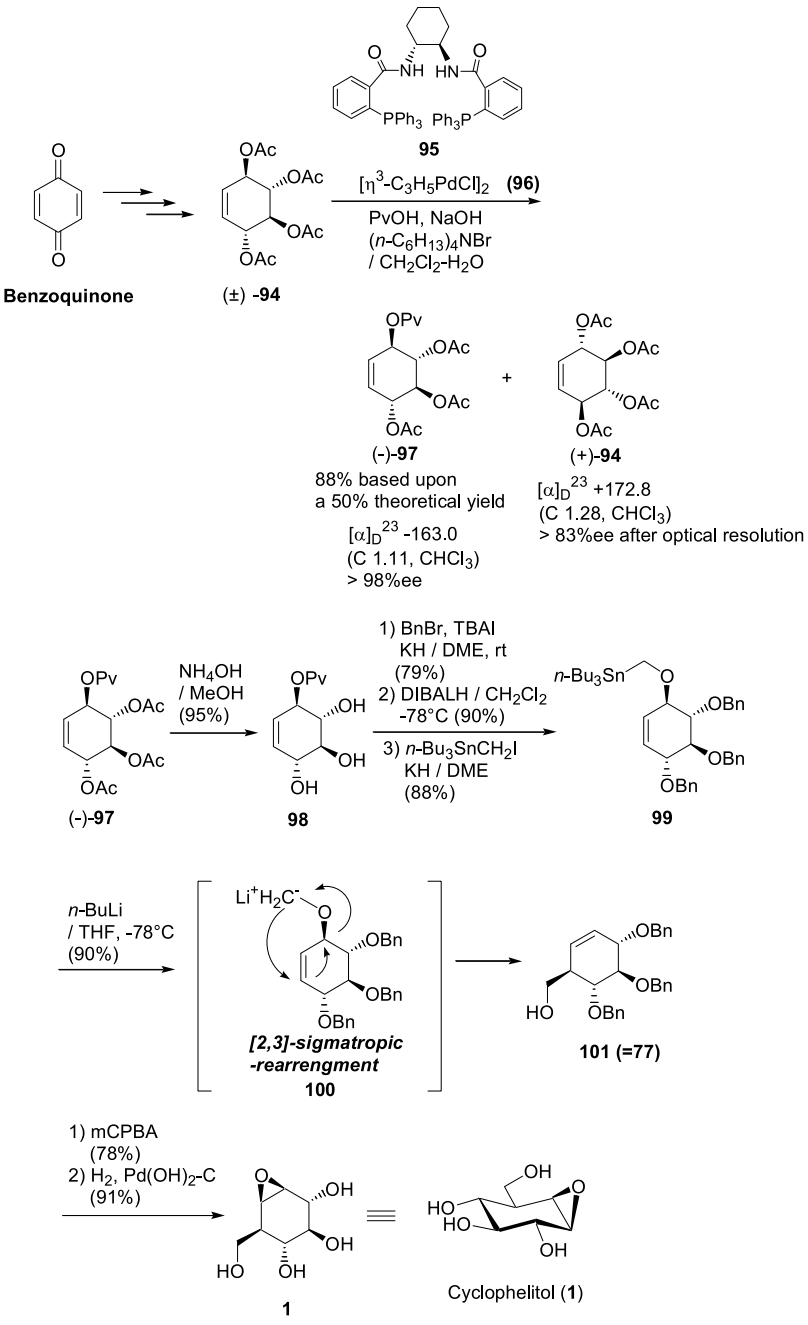
conduritol B is easily accessed from benzoquinone, two issues needed to be addressed: (1) how could racemic conduritol B easily be resolved and (2) how could the hydroxyl groups be readily differentiated. Trost's group considered that asymmetric palladium catalysts offer a simple solution to both of these issues. On the basis of their preliminary study on a palladium-catalyzed kinetic resolution of the racemic C_2-symmetric tetraacetate (±)-**94**, they expected that it should be possible using the chiral ligand (*R*, *R*)-**95**, since, with respect to the ligand, one enantiomer of (±)-**94** would provide a "matched" substrate for ionization while the oth-

er would be "mismatched" [63]. In the meantime, a pivalate was chosen as the carboxylate nucleophile for the resolution because the resultant allyl pivalate was anticipated to ionize much more slowly than the starting material and the pivalate should provide a means for easy differentiation of the alcohol protecting groups later on. The kinetic resolution of (±)-**94**, which was synthesized from in three steps from benzoquinone, was carried out using 0.65 equiv. of sodium pivalate (formed in situ from 0.80 equiv. of PvOH and 0.65 equiv. of NaOH) with 1 mol% of **96** and 3 mol% of **95** at 0.5 M in a two-phase methylene chloride/water system with tetrahexylammonium bromide (THAB) as a phase transfer catalyst. The reaction stopped cleanly at 50% conversion providing a quantitative yield (based on 50% theoretical yield) of recovered tetraacetate (+)-**94** and an 88% yield (based upon a 50% theoretical yield) of monopivalate (–)-**97** with only 1% of dipivalate isolated. Chiral HPLC analysis of (–)-**97** showed it to have 97% ee. Cleavage of Ac groups of (–)-**97** in the presence of the Pv group was straightforward with NH_4OH in MeOH to provide triol **98**. With access to enantiomerically pure triol **98**, the conversion of **98** into cyclophelitol (**1**) was carried out as summarized in ❯ *Scheme 15*. They made the synthetic strategy using 2,3-sigmatropic rearrangement of an alkoxymethyl anion [64] as the key step, and use of this rearrangement required protection of the free hydroxyl groups as the benzyl ethers. After pivalate cleavage of the corresponding benzyl ethers, the tin derivative **99** was readily formed. Tin-lithium exchange promoted the 2,3-sigmatropic rearrangement to put the hydroxymethyl group in place with correct regio- and diastereochemistry with production of **101 (=77)** via transition state **100**. Epoxidation with mCPBA, directed by the homoallylic alcohol, gave a 78% yield of the desired epoxide, with a 7% yield of the diastereomeric epoxide. Completion of the synthesis by the hydrogenation procedure published by Ziegler et al., gave (+)-cyclophelitol (**1**) (❯ *Scheme 15*).

In this section, a variety of the synthetic efforts on cycliphelitol (**1**) were introduced. From these efforts, useful synthetic methodologies to synthesize highly functionalized cyclitols were developed, and, in practice, they favored SAR analysis on glycosidase inhibitory activities of cyclophelitol (**1**) and its stereoisomers. This means that the appearance of cyclophelitol (**1**) paved the way for a new wave of drug discovery based on the science of glycosidase inhibitors and this will contribute to the further development of glycoscience and the development of new synthetic methodologies for total synthesis.

3.1.2 Tamiflu: A Neuraminidase Inhibitor

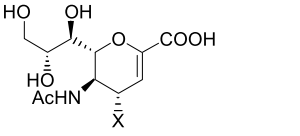
Despite considerable progress in elucidating the molecular mechanism and cellular biology of influenza virus, influenza infection continues to be the most serious respiratory disease both in terms of morbidity and mortality [65]. Amantidine and its analogue rimantidine represent the compounds licensed for the treatment and prophylaxis of influenza A infection. However, these compounds are not effective against influenza B viruses, and their clinical use has been limited by side effects and the rapid emergence of resistant viral strains [66]. On the other hand, vaccine development has been partially successful in the control of influenza epidemics due to the highly variable mutation of the influenza virus [67,68]. This means that promising anti-influenza drugs don't always exist in medical practice. Recently, a great deal of attention in this field has been focused on the unique replication mechanism of the influenza virus, and this has allowed the relevant scientists to identify a number of potential molecular targets for drug design. Those targets include hemagglutinin (HA) [69,70], neuraminidase [71,72], M2


95
[η3-C3H5PdCl]2 (96)
PvOH, NaOH
(n-C6H13)4NBr
/ CH2Cl2-H2O
Benzoquinone
(±) -94
(-)-97
88% based upon
a 50% theoretical yield
[α]D23 -163.0
(C 1.11, CHCl3)
> 98%ee
(+)-94
[α]D23 +172.8
(C 1.28, CHCl3)
> 83%ee after optical resolution
NH4OH
/ MeOH
(95%)
98
1) BnBr, TBAI
KH / DME, rt
(79%)
2) DIBALH / CH2Cl2
-78°C (90%)
3) n-Bu3SnCH2I
KH / DME
(88%)
99
n-BuLi
/ THF, -78°C
(90%)
[2,3]-sigmatropic
-rearrengment
100
101 (=77)
1) mCPBA
(78%)
2) H2, Pd(OH)2-C
(91%)
1
Cyclophelitol (1)

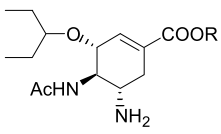
Scheme 15

protein [73], and endonuclease [74]. Hemagglutinin and neuraminidase are two major surface glycoproteins expressed by both influenza A and B viruses. HA is known to mediate binding of viruses to target cells via terminal sialic acid residue in glycoconjugates. This binding is the first step of viral infection. In contrast to HA activity, NA catalyzes removal of terminal sialic acids linked to glycoproteins and glycolipids. Although the biological consequences of this activity are not completely understood, it has been postulated that NA activity is necessary in the elution of newly formed viruses from infected cells by digesting sialic acid in the HA receptor [75,76]. NA may also promote viral movement through respiratory tract mucus, thus enhancing viral infectivity [74. 75]. Therefore, NA has been considered to be a suitable target for designing new types of anti-influenza drugs.

In earlier studies, 2,3-didehydro-2-deoxy-*N*-acetylmeuraminic acid (Neu5Ac2en, **102**) was found to be an influenza NA inhibitor with a *Ki* of 4 μM [78,79]. Biochemical studies [80] indicate **102** is considered a transition state-like analogue binding to the active site of NA [81,82]. Afterward, on the basis of structural information generated from X-ray crystallographic study of **102** complexed with NA, the following rationally designed NA inhibitors were prepared: 2,3-didehydro-2, 4-dideoxy-4-amino-*N*-acetylneuraminic acid (4-amino-Neu5Ac2en, **103**) and its guanidine analogue (4-guanidino-Neu5Ac2en, **104**) [81,83,84]. In comparison to **102**, both **103** and **104** are more potent NA inhibitors with *Ki* values of 10^{-8} M and 10^{-10} M, respectively. Both the amino group in **103** and the guanidino group in **104** are suggested to form salt bridges with Glu119 in the NA active site, while the latter adds a strong charge-charge interaction with Glu227 [84]. In addition, **104** also exhibited potent antiviral activity against a variety of influenza A and B strains in the cell culture assay [85]. Eventually, **104** was approved by the FDA, and it has been launched as an anti-influenza drug called relenza [86,87,88,89]. However, poor oral bioavailability and rapid excretion precluded **104** as a potential oral agent against influenza infection and **104** has to be administered by either intranasal or inhaled routes. In the case of an influenza epidemic, oral administration may be a more convenient and economical method for treatment and prophylaxis. Therefore, it would be desirable to have a new class of orally active NA inhibitors as potential agents against influenza infection. The research group of Gilead Science tried to discover a new class of compounds by using a carbocyclic template in place of the dihydropyran ring of the Neu5Ac2en system, and it was expected that the carbocyclic ring would be chemically more stable than the dihydropyran ring and easier to modify for optimization of antiviral and pharmacological properties. Consequently, they succeeded in identifying GS-4071 (**105**) and its prodrug GS-4104 (**106**), which is called tamiflu and has been launched as an orally available neuraminidase inhibitor for the treatment of influenza [90]. In this section, I will describe the story of the discovery of tamiflu. This represents a success story of an artificial carbasugar launched as an actual drug.



102: X = OH
103: X = NH_2
104: X = –HN–C(=NH)–NH_2 (Relenza)



105: R = H (GS-4071)
106: R = Et (GS-4104, Tamiful)

Design NA has been classified into nine subtypes for type A influenza virus strains according to their serological properties. However, there are no subtypes in the B-type virus. Although amino acid sequence homology among NA from both type A and type B virus strains has been found to be only 30% [91,92], the enzyme activity of NA among the different strains is the same, indicating the highly conserved nature of the active site of the enzyme [93]. The X-ray crystallographic structures of NA have been determined from three influenza subtypes: A/Tokyo [94], A/Tern [95], and B/Beijing [96]. The structures displayed a symmetrical folding pattern of six four-stranded antiparallel β-sheets arranged like blades of a propeller. NA exists as a mushroom-shaped spike with a boxlike head on top of a long stalk containing a hydrophobic region by which it is embedded in the viral membrane. Crystallographic studies of NA reveal that, in the active site, the amino acids which line and surround the walls of the binding pocket are highly conserved among all influenza strains examined so far. The high-resolution crystallographic structure of sialic acid (**107a**) complexed with NA revealed that sialic acid binds the enzyme in a considerably deformed conformation due to the strong ionic interactions between the carboxylate of the substrate and Arg118, -292, and -371 in the active site of the enzyme [97]. In solution the carboxylate of sialic acid is axial, but the deformation of the ring on binding places the carboxylate into a pseudoequatorial position. This binding mode is similar to that found in the X-ray crystal structure of Neu5Ac2en (**102**) complexed with NA [95]. In this case, the double bond of Neu5Ac2en constrains the pyranose ring of the sugar into a planar structure around the ring oxygen. On the basis of this structural information, it has been proposed that the catalytic mechanism for the cleavage of sialic acid from glycoconjugates (**107b**) implicates the formation of the C_2 carbonium cation (**108**) which is stabilized by the neighboring oxygen atom as shown in ❯ *Fig. 2* [81,82]. Furthermore, kinetic isotope studies have provided convincing evidence for the C_2 carbonium cation formation as an intermediate structurally similar to the transition state in the sialic acid cleavage by NA [80]. In general, transition-state mimics frequently are potent inhibitors for the catalyzing enzyme. The concept of structural similarity to the transition state has found wide application in drug design over the years. The multitude of enzyme-inhibitor interactions are governed by steric as well as electronic factors. In theory, compounds that closely resemble the transition-state structure should give high binding affinity toward the target enzyme [98,99], and the cases of neuraminidase inhibitors are not exceptional. Using intermediate **108** as a key transition-state mimic is a reasonable approximation in view of the X-ray crystallographic studies described above. Their earlier attempts to mimic the intermediate **108** with a completely flat benzene ring did not lead to potent NA inhibitors [100,101], suggesting the importance of the stereochemistry of substituents around the ring in the design of inhibitors with high affinity for NA. Therefore, considering the flat oxonium cation in **108** as an isoster of the double bond and the fact that the carbocyclic system would be expected to be chemically versatile for the manipulation of side chains attached to the ring, Gilead's group selected the cyclohexene scaffold as a replacement for the oxonium ring of **108** which would keep the conformational changes to a minimum (❯ *Fig. 2*).

Syntheses of the Prototypes In their initial stage of designing carbocyclic NA inhibitors, the olefinic isomers that are shown in structures **109** (Type [**I**]) and **110** (Type [**II**]) were considered as two possible prototypes of transition-state analogues. The isomer **109** is structurally closer to transition-state **108** than isomer **110**. However, Gilead's group expected that it would be

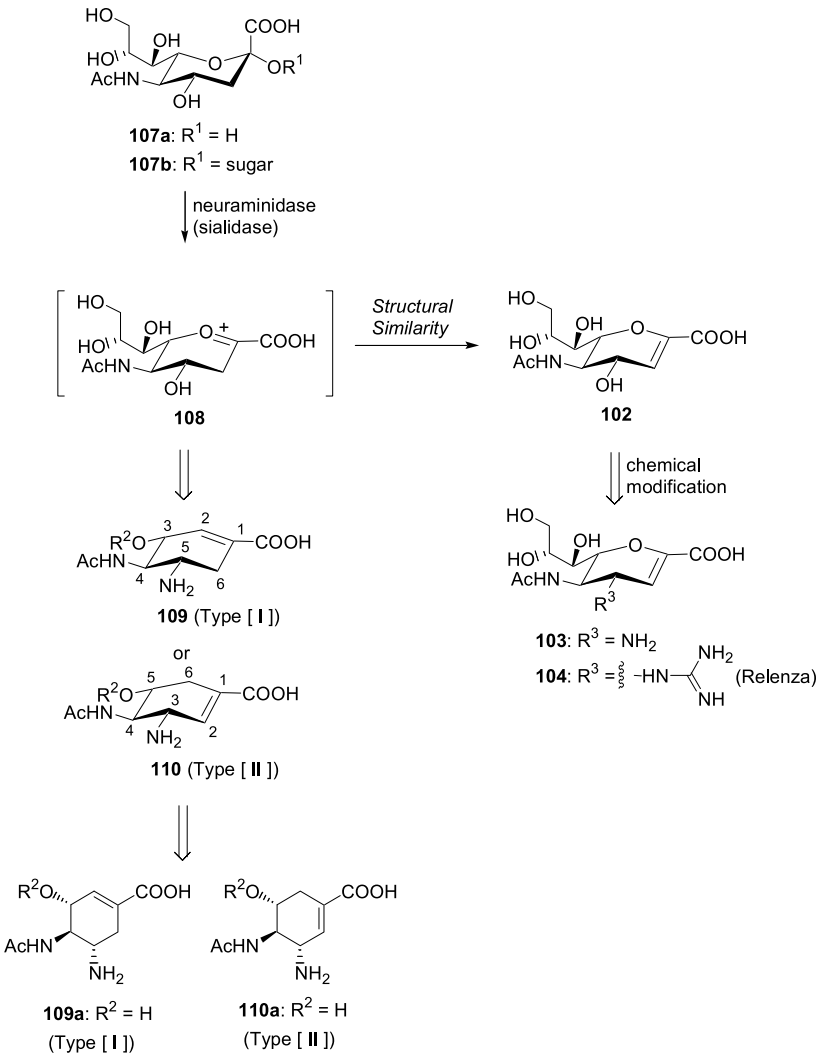
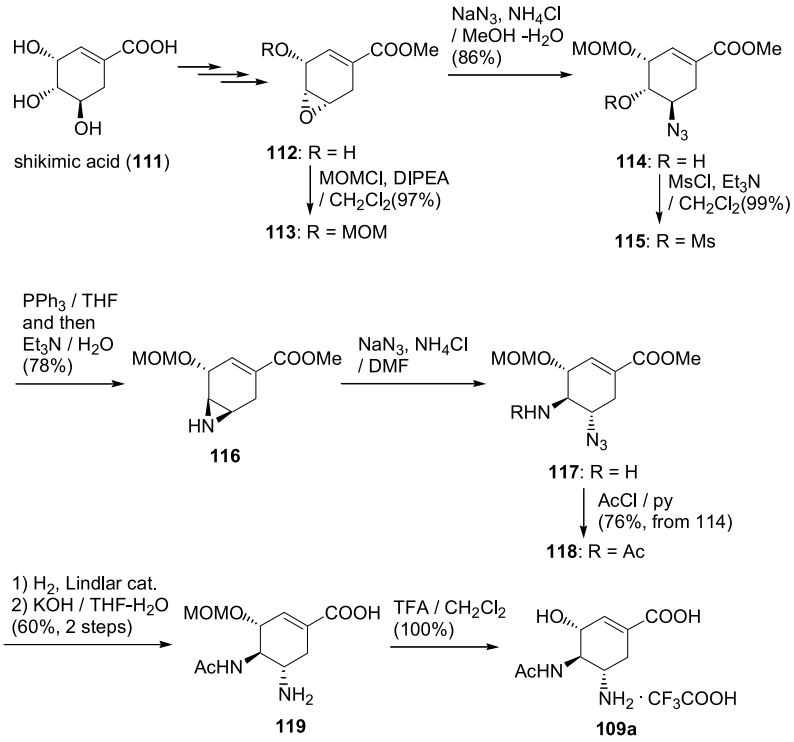


Figure 2
Rational design of carbocyclic transition-state analogues

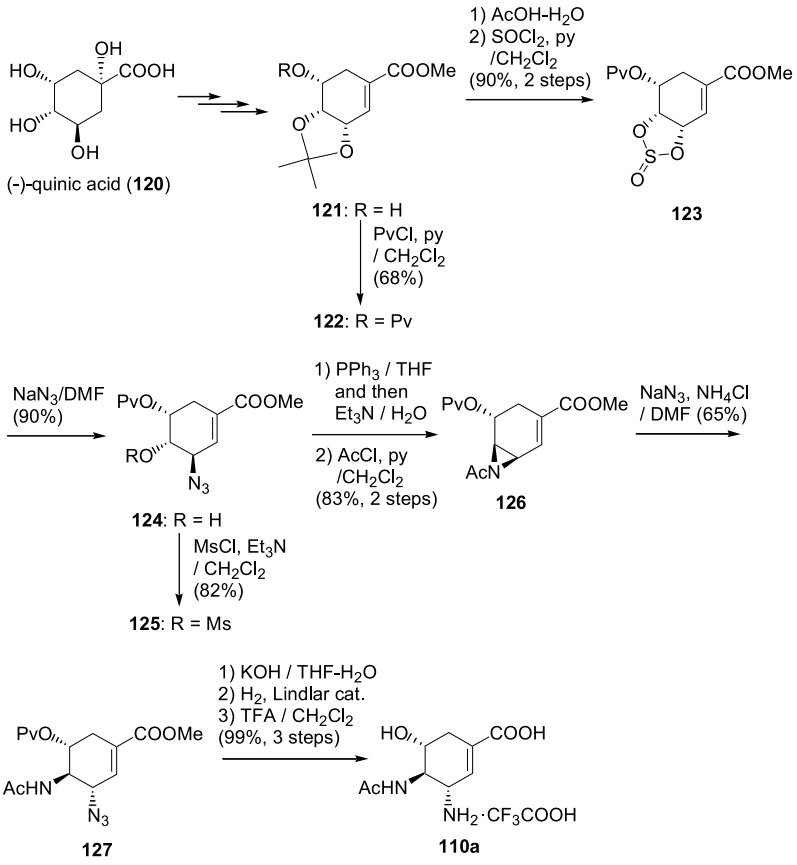
difficult to assess a priori which isomer would be a preferred NA inhibitor, especially, in light of the potent NA inhibitory activity displayed by Neu5Ac2en (**102**) and its analogues **103** and **104**, in which the double bond was located in the position corresponding to isomer **110** and the fact that Chandler et al. reported carbocyclic analogues have some interesting NA inhibitory activity [102]. In the meantime, the molecular modeling of isomers **109** and **110** indicated that these two molecules overlay well. Therefore, Gilead's group synthesized the simple compounds **109a** and **110a**, and evaluated which would be a more appropriate prototype for further extensive structure-activity relationships studies (*Fig. 2*).



Scheme 16

A comparison of structural similarities revealed that (–)-shikimic acid (**111**) and **109a** share common structural features. However, the conversion of **111** to **109a** requires the effective stereochemical control for transforming the trans C_4 and C_5 hydroxyls of **111** to the trans C_4 and C_5 amino groups of **109a**. The approach selected for this transformation, as outlined in *Scheme 16*, relied on the conversion of **111** to aziridine **116** followed by azide ion attack in a regio- and stereospecific manner as a key reaction.

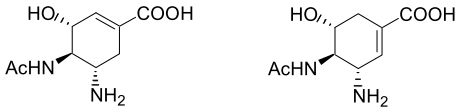
The synthesis began with the preparation of epoxide **112** from **111** as described in the literature [103]. Nucleophilic ring opening of MOM-protected epoxide **113** with NaN_3 in the presence of NH_4Cl generated azido alcohol **114**. The ring-opening of the epoxide was both region- and stereospecific as depicted in **114**, and this could be attributed to the steric and electronegative inductive influence of the MOM group in **113**. Conversion of azide **114** to aziridine **116** was efficiently accomplished via a two-step sequence: (1) mesylation of the hydroxyl group in **114** and (2) reduction of the azide functionality in **115** with PPh_3 in the presence of Et_3N and H_2O. The aziridine ring opening of **116** with NaN_3 gave **117** exclusively. This selective ring opening was again a consequence of the favored azide ion attack at the C_5 position due to the steric and electronegative inductive effects of the MOM group. Finally, acetylation of **117** followed by reduction of the azide group in **118** and saponification of the methyl ester and deprotection of the MOM group in **119** with TFA provided **109a** (*Scheme 16*).



Scheme 17

On the other hand, **110a** was synthesized from (–)-quinic acid (**120**). The cyclohexene intermediate **122**, readily available from **121** by literature methods [104], possesses considerable structural similarity to **110a**. The C_5 hydroxyl of **121** was protected as a pivaloyl ester, and mild acid hydrolysis of **122** furnished the corresponding diol. Exposure of the diol to $SOCl_2$ in the presence of Et_3N generated cyclic sulfite **123**. The sulfite ring opening of **123** with NaN_3 gave azide **124** as a single product. This complete regiospecific ring opening is a consequence of the favored azide ion attack at allylic C_3 position of **123**. The β-hydroxy azide moiety in **124** was then converted in a three-step sequence to aziridine **126**, analogous to the conversion of **114** to **116**. Ring opening of acetylated aziridine **126** with NaN_3 was completely regiospecific, giving azide **127** followed by reduction of azide functionality which provided **110a** (*Scheme 17*).

109a and **110a** showed interesting inhibitory activities in a NA enzymatic assay (*Fig. 3*). While **109a** proved to be a potent NA inhibitor with an IC_{50} of 6.3 μM, **110b** did not exhibit inhibitory activity at concentrations up to 200 μM.



109a **110a**

Influenza Neuraminidase Inhibition (IC_{50}, μM)	6.3	> 200

Figure 3
Influenza neuraminidase inhibition by prototypes 109a and 110a

It was concluded that this result demonstrated that the double bond position in the design of carbocyclic NA inhibitors plays an important role in NA activity; however, further structural investigation is required to illustrate the binding differences of **109a** and **110a** in the NA active site. Eventually, based on this result, the structure of **109** was selected as the prototype for further drug discovery attempts.

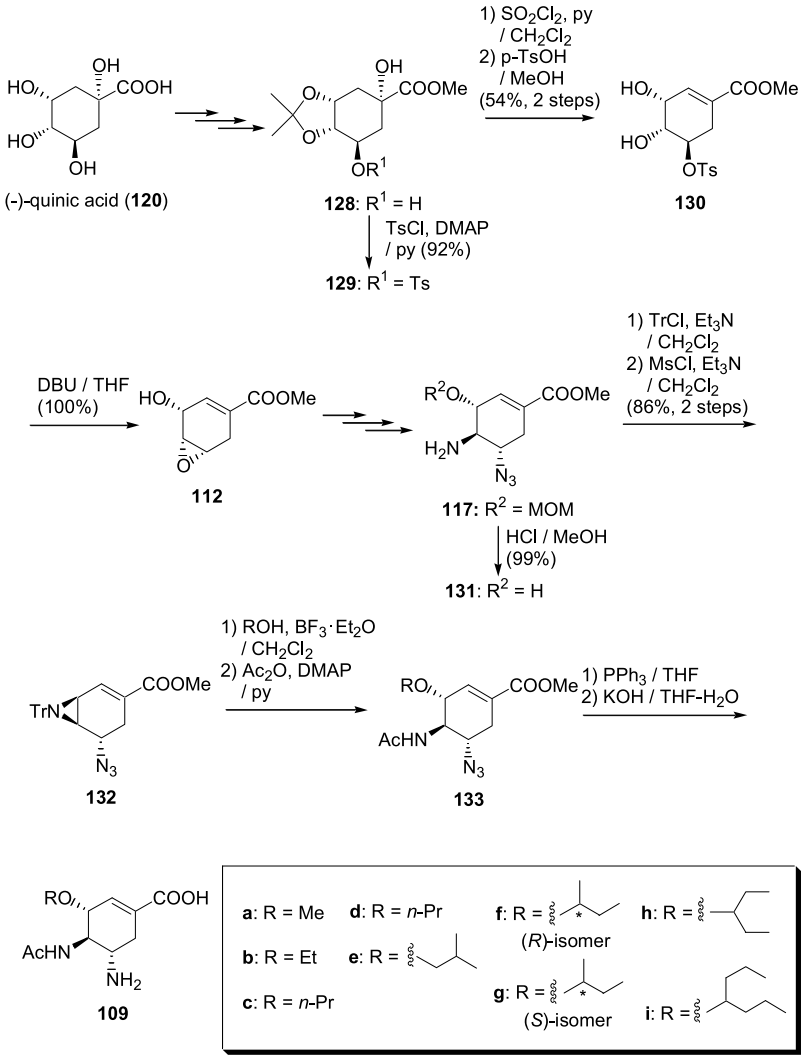
SAR Analysis Crystallography studies of Neu5Ac and its analogues bound to NA appear to indicate that the C_7 hydroxyl of the glycerol side chain does not interact with any amino acids of the NA active site [72]. This suggested that the C_7 hydroxyl could be eliminated from the carbocyclic analogues without losing binding affinity to NA. Furthermore, it was also realized that, in the transition-state intermediate **108**, the oxonium double bond is highly polarized and electron deficient. Considering these features, in the carbocyclic structure, the CHOH group at the C_7 position of the glycerol side chain in the Neu5Ac system was replaced with an oxygen atom as shown in structure **109**. This C_3 oxygen atom would reduce the electron density of the double bond via the σ bond electronegative inductive effect. In addition to these rationales, Gilead's decision to have the C_3 oxygen atom was based on the synthetic practicality of modifying R^2 groups to optimize NA inhibitory activity and pharmacological properties (*Fig. 2*).

X-ray crystallographic structures of Neu5Ac and its analogues complexed with NA show that two terminal hydroxyls of the glycerol side chain form a bidentate interaction with Glu276 [94,96]. However, it is also noted that C_8 of the glycerol chain makes hydrophobic contacts with the hydrocarbon chain of Arg224 [72]. Therefore, it was expected that the optimization of this hydrophobic interaction would lead to new NA inhibitors with increased lipophilicity while maintaining potent NA inhibitory activity. This consideration was especially important for designing orally bioavailable drugs since balancing lipophilicity and water solubility could be as critical as the size of the molecule for its absorption from the intestinal tract. For optimization of hydrophobic interactions, the dimensions of the spacer would be expected to play an important role in binding affinity to the enzyme, as factors such as length, geometry, and conformational mobility. On the basis of this premise, a systematic modification of the R^2 portion in **109** with various aliphatic side chains was conducted. Previous studies [72] suggested that the carboxylate of Neu5Ac2en (**102**) and its analogues form strong ionic interactions with three guanidino groups of Arg118, -292, and -371. The acetamido

moiety at C_5 of Neu5Ac2en and its analogues interacts with Arg152 and Glu227, and the methyl group fits nicely into a hydrophobic pocket formed by Trp180, Ile224, and Arg226. The acetamido group was also demonstrated to be optimal for antiviral activity [105]. The amino and guanidino groups at the C4 position of the Neu5Ac2en system were optimal for NA activity as described above [79,84]. Therefore, in the carbocyclic structure **109**, the C_1 carboxylate, C_4 acetamido, and C_5 amino groups were kept constant, while the C_3 aliphatic group was optimized for antiviral activity.

For structure-activity relationship studies of carbocyclic analogues **109**, a general and efficient route to introduce various alkyl ethers at the C_3 position was required. Also, although shikimic acid as a chiral starting material is convenient because of easy conversion to the key epoxy intermediate **112**, its high cost and low availability in large quantities made it impractical for scale-up of **112**. This meant that this would become a formidable challenge when multi-kilograms of material were needed for evaluation of clinical candidates for further drug development as well as frequent consumption of common synthetic intermediate **112** for SAR studies. In this stage, the development of the synthetic route, including the solutions of these issues, was achieved. With respect to the improvement of the synthetic route of **112**, (–)-quinic acid (**120**), which was used as a starting material for the synthesis of **110a** as shown in ❯ *Scheme 17*, appeared to be an ideal starting material because of its low cost and commercial availability. The focus on the synthesis is the selective dehydration of the C_2 hydroxyl in **129**. Tosylate **129** was reacted with sulfuryl chloride in pyridine, followed by acetonide cleavage in refluxing methanol in the presence of *p*-TsOH. Interestingly, the desired diol **130** was directly crystallized out of the reaction mixture in 54% overall yield, and in the dehydration and acetonide removal sequence for the conversion of **129** to **130**, the other olefinic regioisomer of **130** was aromatized under the reaction condition and easily separated by crystallization. Gilead's group reported that this procedure would be amenable to a several hundred gram scale. Finally, epoxide **112** was produced in quantitative yield by treatment of **130** with DBU in THF, and the synthesis of epoxide **112** from quinic acid has been achieved in good overall yield without column chromatography. Also, regarding a general and efficient route to introduce various alkyl ethers at the C_3 position, the approach, as outlined in ❯ *Scheme 18*, relied on the aziridine opening of **132** with alcohols, which should in principle be highly regioselective due to the preferred nucleophilic attack at the C_3 allylic position. The requisite aziridine **132** was derived from the *trans* amino alcohol **131** by the two-step, one-pot process: (1) selective protection of the amino functionality with TrCl and (2) mesylation of the hydroxyl in the presence of Et_3N. Under these conditions, the mesylate intermediate was converted into aziridine **132**. Treatment of **132** with various alcohols in the presence of $BF{\cdot}Et_2O$ followed by acetylation of the crude product provided the ether **133** in 55–88% yield. This aziridine ring opening proceeded regio- and stereospecifically, and no other regio- and stereoisomer were observed. Finally, reduction of the azide functionality and saponification of the ester group in **133** gave **109**. Thus, the convergent and efficient synthesis of various ether analogues of **109** has been achieved from intermediate **112** which is readily available from quinic acid as well as shikimic acid (❯ *Scheme 18*).

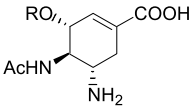
The NA inhibitory activities of the carbocyclic analogues were evaluated in two assay systems (❯ *Table 2*). The intrinsic activity of each compound was assessed by measuring the inhibition of enzymatic activity, and compounds that exhibited potent NA inhibitory activity were further evaluated in cell culture by a plaque reduction assay using an influenza A (H1N1) strain.



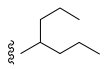
Scheme 18

As shown in *Table 2*, the length, size of branching, and geometry of the alkyl groups in **109** profoundly influence the NA inhibitory activity. In a series of linear alkyl analogues (**109b–e**), steady increases in the enzyme inhibitory activity were observed up to the *n*-propyl analogue **109d**. The over 20-fold increase in the NA inhibitory activity for **109d** compared to the methyl counterpart **109b** implicated a significant hydrophobic interaction of the *n*-propyl group with amino acids in the active site. Branching at the β-carbon of the *n*-propyl group (compound **109f**) resulted in no enhancement of NA inhibitory activity compared to that of **109d**. In contrast, branching at the α-carbon of the *n*-propyl group resulted in almost 20-fold enhance-

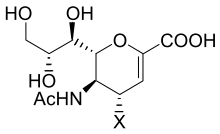
Table 2
Influenza neuraminidase inhibition and plaque reduction by carbocyclic analogues



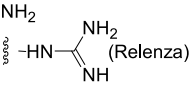
109

R	Compound	enzyme[a] IC_{50} (nM)	plaque[b] EC_{50} (nM)
H	**109a**	6300	ND [c]
Me	**109b**	3700	ND
Et	**109c**	2000	ND
n-Pr	**109d**	180	ND
n-Bu	**109e**	300	ND
	**109f**	200	ND
 (*R*)-isomer	**109g**	10	80
 (*S*)-isomer	**109h**	9	135
	**109i**	1	16
	**109j**	16	ND
	103	150	2500
	104	1	15

a: neuraminidase (NA)
b: H1N1, A/ws
c: ND = not determined



103: X = NH_2

104: X = 

ment of NA inhibitory activity. This improved potency undoubtedly arose from additional hydrophobic interactions in the NA active site generated by the α-methyl group. Remarkably, the absolute stereochemistry of the α-methyl group (**109g** and **109h**) did not influence the NA inhibitory activity as they exhibited almost equal potency. Gilead's group concluded that this result means it can be speculated that the methyl and ethyl groups in the *sec*-butyl analogues **109g** and **109h** contribute almost the same binding energy by an equal degree of interaction with amino acids in the active site. This hypothesis was further supported by NA activity of 3-pentyl analogue **109i (=105)**, which possesses two identical branched ethyl groups and shows the NA inhibitory activity increase of almost 10-fold that of either isomers **109g** and **109h**. Further extension of the two ethyl groups to the 4-heptyl analogue **109j** resulted in a decrease of NA inhibitory activity, suggesting that the 3-pentyl group in **109i (=105)** might provide the optimum hydrophobic contact with the NA active site. Interestingly, the enzymatic and plaque reduction activities of **109i (=105)** were comparable to those of **104**, which exhibited over 100-fold increase in activity compared to that of amino analogue **103**. Replace-

ment of the amino group in **109i (=105)** with the guanidino moiety resulted in a significant increase in its enzymatic and cell culture activity compared to **109i (=105)**. Thus, carbocyclic NA inhibitors represented by **109i (=105)** and its guanidino analogue are more potent NA inhibitors than any inhibitors including compounds **103** and **104**. Finally, ethyl ester of **109i (=105)** (designated as GS-4104, **106**) exhibited good oral bioavailability in several animals (mice, rats, and dogs) and demonstrated oral efficacy in the mouse and ferret influenza model. On the basis of potent in vitro/in vivo activity and favorable pharmacological properties, GS-4104 (**106**) has been selected as a clinical candidate for the oral treatment and prophylaxis of influenza infection.

Synthetic Method for Process Chemistry After GS-4104 (**106**) was identified as a clinical candidate, a practical kilogram-scale preparation of **106** was necessary to supply an accelerated program of clinical and toxicological studies. The key structural feature in **106**, in terms of both pharmacological activity and synthetic challenge, was the 3-pentyl ether group. The aforementioned discovery synthesis of **106** constructed this 3-pentyl ether by acid-catalyzed opening of the tritylaziridine **132** with 3-pentanol (Route [A], ❯ *Fig. 4*). The tritylaziridine **132** was prepared stereospecifically from either (–)-quinic acid (**120**) or (–)-shikimic acid (**111**), but both sequences required double inversion and repeated protection/deprotection of the (*R*)-3-hydroxyl group. The generality of the tritylaziridine opening for a wide variety of alcohols and other nucleophiles facilitated the rapid creation of structural diversity in the discovery program. However, identification of the lead compound **106** allowed for the design of a more efficient approach. Rohloff et al. developed a new 12-step synthesis of **105** utilizing a novel and efficient reductive ketal opening of **134** to construct the 3-pentyl ether (Route [B], ❯ *Fig. 4*) [106]. This new process is highly applicable to kilogram-scale synthesis in that it fea-

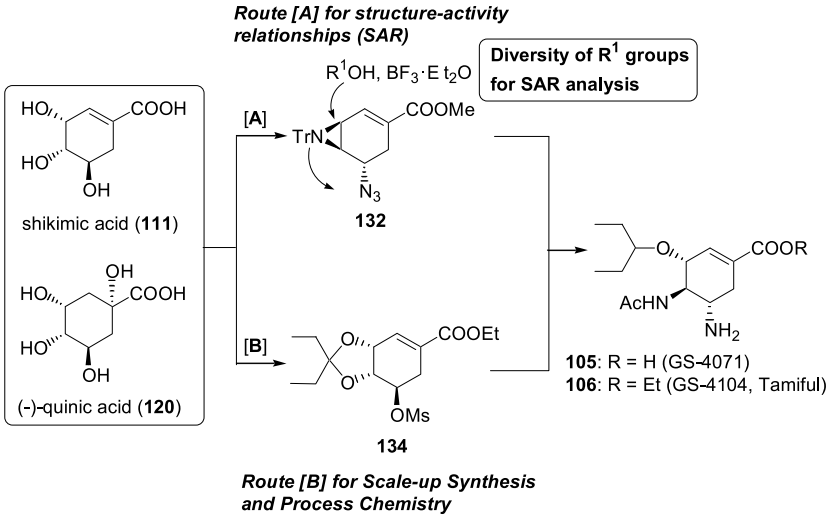
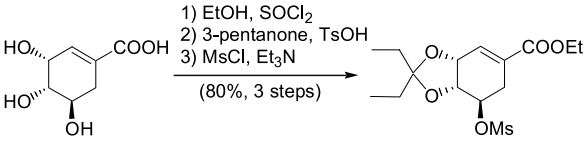


Figure 4
Synthetic strategies of 105 and 106

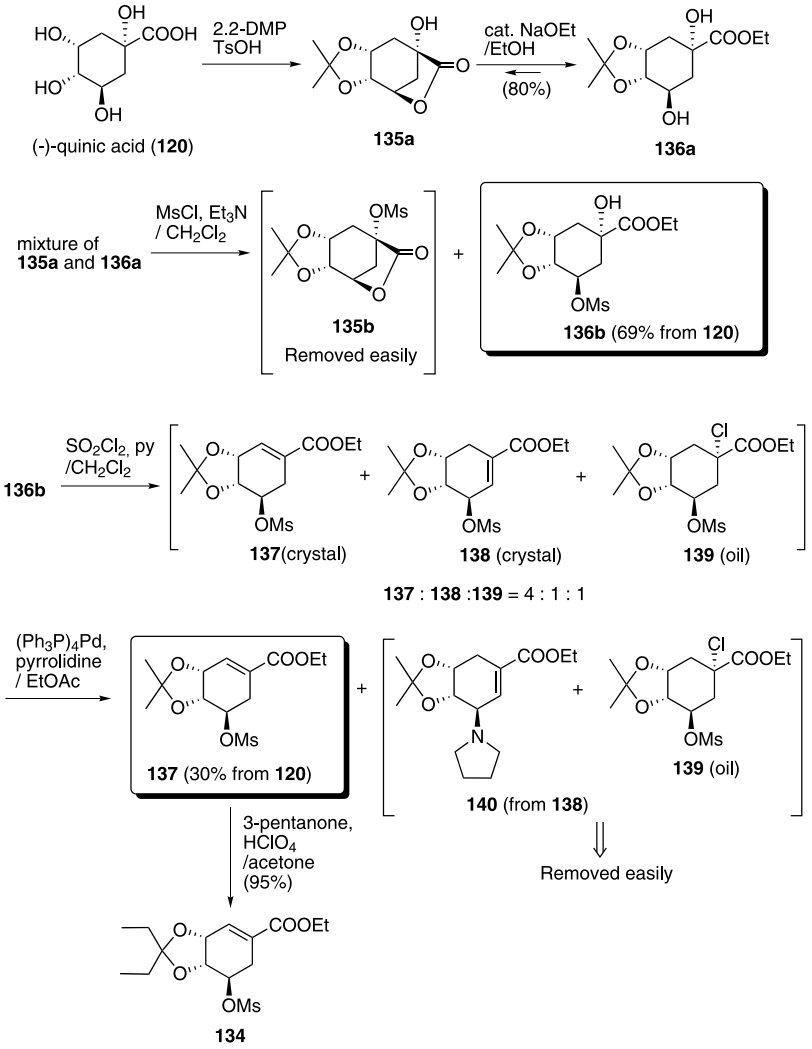


(-)-shikimic acid (**111**) **134**

Scheme 19

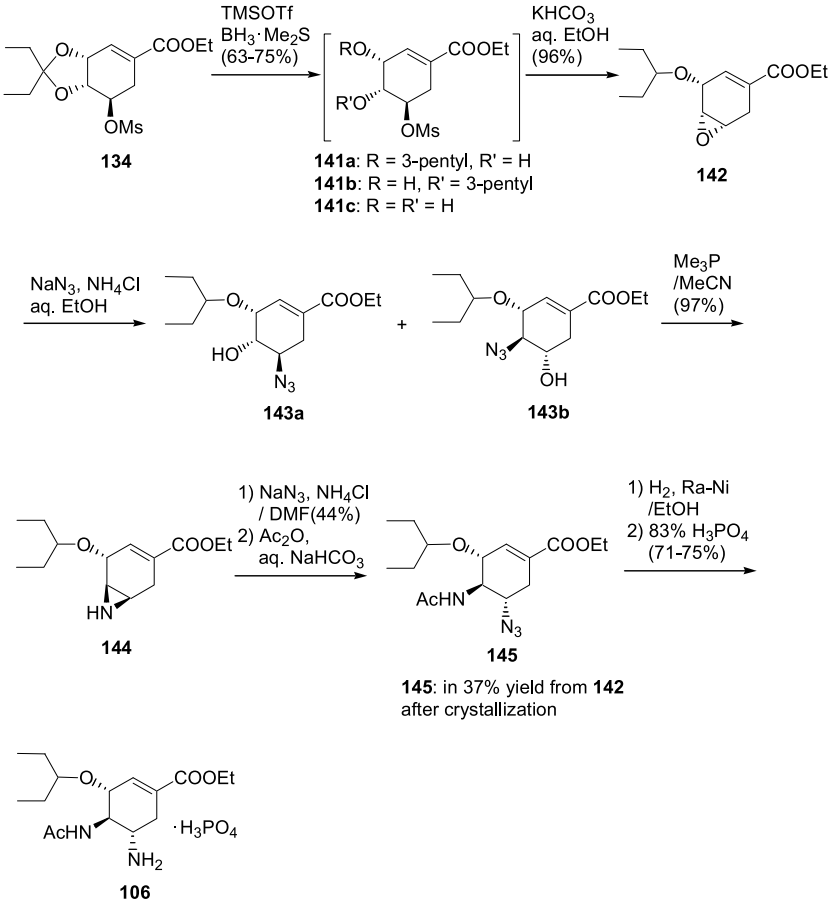
tures only three isolated crystalline intermediates and requires no chromatography (*Fig. 4*). Compound **134** was synthesized in 80% yield from natural shikimic acid (**111**) (*Scheme 19*). However, due to the limited commercial availability of (–)-shikimic acid (111), an efficient preparation of ketal 134 from (–)-quinic acid (120) was developed as shown in *Scheme 20*. The quinic lactone acetonide **135a** was prepared in 90% yield from **120** by modification of the method of Shing [107], and this was converted to a 1:5 equilibrium mixture of lactones : hydroxyl ester **135a**:**136a** in anhydrous EtOH (1 M) containing catalytic NaOEt. Separation of the **135a**:**136a** mixture by fractional crystallization was found to be inefficient on the kilogram scale, and instead, the crude **135a**:**136a** mixture was treated with MsCl (1.1 equivalents) in CH_2Cl_2 in the presence of Et_3N to chemoselectively afford the monomesylates **135b**:**136b** in an unchanged ratio of 1:5. Undesired, but highly crystalline, lactone mesylate **135b** was readily removed by filtration of an EtOAc slurry of the **135b**:**136b** mixture. Upon evaporation of the filtrate, crude oily **136b** was isolated in 69% overall yield from (–)-quinic acid (120). Dehydration of **136b** to form the unsaturated "shikimic" ring system was accomplished with SO_2Cl_2 and pyridine in CH_2Cl_2 at –20°. A mixture of 1,2- and 1,6-olefin regioisomers **137**:**138** (ratio 4:1) was obtained in 60% yield along with 10–15% of oily α-chloro compound **139**. Because of the high crystallinity of both olefin isomers, a high-throughput fractional crystallization could not be accomplished. Instead, it was found that treatment of the crude **137**:**138**:**139** mixtures with pyrrolidine and catalytic $(Ph_3P)_4Pd$ in EtOAc led to selective conversion of the undesired 1,6-olefin (*allylic* mesylate) **138** into pyrrolidine substitution product **140** [108,109]. This 1,6-olefin adduct was readily removed by aqueous H_2SO_4 extraction. The pure 1,2-olefin isomer **137** was then isolable by crystallization from EtOAc/hexane in 30% overall yield from **120**. In order to complete the link to the (–)-shikimic acid route, the quinic acid-derived acetonide **137** was transketalized using catalytic perchloric acid in 3-pentanone at ambient temperature with continuous vacuum distillation of acetone. The 3,4-pentyliden ketal **134** was isolated in nearly quantitative yield and was identical in all respects to the ketal **134** prepared from (–)-shikimic acid (**111**). A more concise route based on initial formation of the 3,4-pentylidene ketal analogues of **135a**, **136a**, and **136b** from (–)-quinic acid (**120**) was found to be impractical for scale-up due to the lack of crystallinity in the 3,4-pentylidene series (*Scheme 20*).

Treatment of **134** with TMSOTf and $BH_3{\cdot}Me_2S$ in CH_2Cl_2 under modified Hunter conditions [110], gratifyingly afforded a 10:1:1 mixture of isomeric pentyl ethers **141a**:**141b** and diol **141c** in 75% yield. Separation of the **141a**:**141b**:**141c** mixture was not possible through fractional crystallization or chromatography, but heating the crude mixture in aqueous EtOH in the presence of $KHCO_3$ selectively converted **141a** into the alkane-soluble epoxide **142**. Heptane extraction gave crystalline epoxide **142** in 60% overall yield from ketal **134**. Epoxide



Scheme 20

142 was heated at 70°C with NaN_3 and NH_4Cl in aqueous EtOH to afford an oily 10:1 mixture of isomeric azido alcohols **143a**:**143b**. Intramolecular reductive cyclization of the crude **143a**, **b** mixture with Me_3P in anhydrous MeCN at 35°C cleanly afforded a single aziridine **144** of ca. 70% purity. Aziridine opening proceeded smoothly at 80°C with NaN_3 and NH_4Cl in DMF affording the corresponding azidoamine, which was directly acylated with Ac_2O. The resulting azidoacetamide **145** (m.p 137–8°C) was isolated, after recrystallization in 37% overall yield from epoxide **142**. Because of its stability and high crystallinity, azidoacetamide **145** has proven to be an excellent final intermediate for drug substance synthesis. To complete

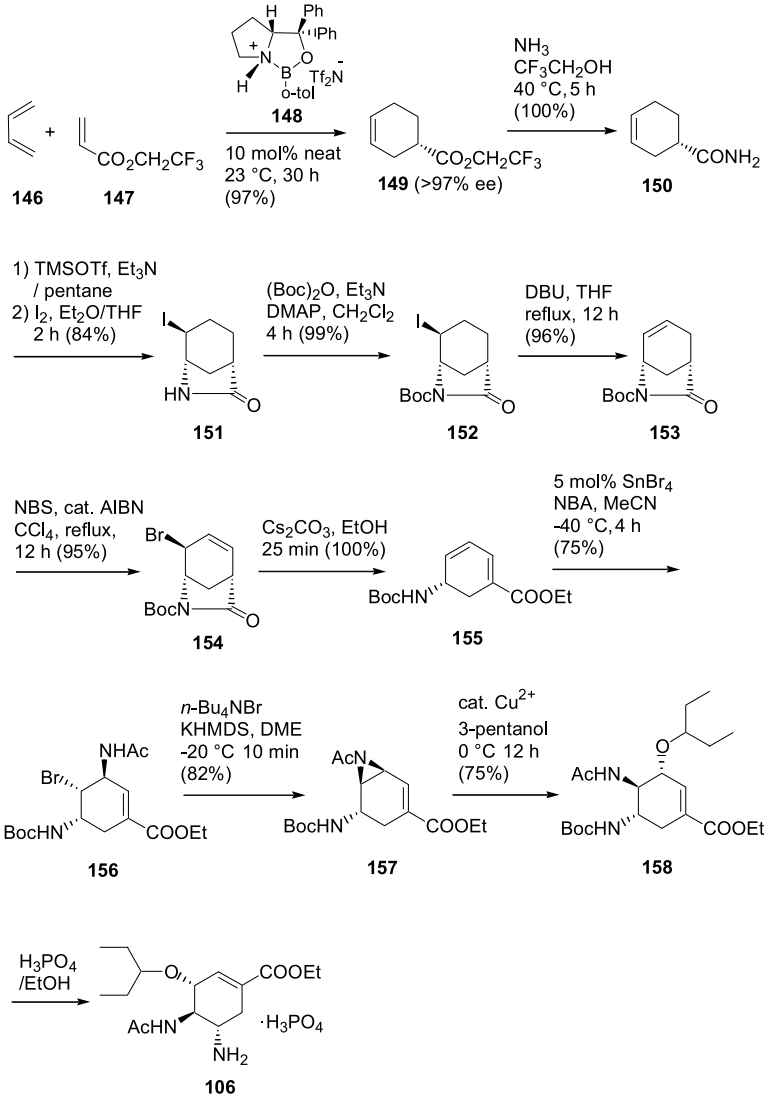


Scheme 21

the synthesis, azide reduction of substrate **145** was accomplished using catalytic hydrogenation with Raney Ni in EtOH (1 atm H_2) at 35°C. After removal of catalyst by filtration, 85% phosphoric acid was added. The salt **106**·H_3PO_4 crystallized as feathery needles, and it was isolated in 71% yield from **145**. Consequently, overall, **106** was prepared in 12 steps and 4.4% yield from (–)-quinic acid (**120**) (Average 77.2% yield/step), and (formally) in ten steps and 21% yield from (–)-shikimic acid (**111**), and kilogram quantities of drug substance for clinical and toxicological studies have been prepared by this route (*Scheme 21*).

Novel Synthetic Routes The recent emergence of the avian virus H5N1 raises the possibility of a pandemic wave of life-threatening influenza that requires prompt action. Currently, a four-pronged effort to avert widespread disease is now underway, and the effort consists of the following components: (1) worldwide surveillance of both wild and domesticated birds with quick culling of the latter, (2) development of recombinant vaccines against the H5N1

virus and its mutated forms whose production can be scaled up rapidly, (3) procedure for quarantine, and (4) ramped up production of the orally effective, synthetic neuraminidase inhibitor tamiflu (**106**). Especially, the stock of tamiflu (**106**) is becoming a serious social problem, and it is true that many nations have planned to stock a significant amount of **106** to protect against a possible influenza outbreak. The current commercial synthetic route of **106** uses naturally occurring (–)-shikimic acid (**111**) or quinic acid (**120**) as a starting material. However, as mentioned in previous sessions, these starting materials are complex, relatively expensive, and of



Scheme 22

limited availability, therefore, a more reliable source is desired for constant and large-scale supply of **106**.

In 2006, two groups independently reported the novel asymmetric synthesis of tamiflu (**106**). Corey et al. reported a short enantioselective pathway for the synthesis of **106** from 1,3-butadiene and acrylic acid shown in ❯ *Scheme 22* [111]. The key steps of the synthesis are (1) Diels–Alder reaction of 1,3-butadiene (**146**) and trifluoroethyl acrylate (**147**) in the presence of chiral ligand **148** developed in the laboratory [112], (2) the introduction of two amino groups in tamiflu (**106**) without using potentially hazardous and explosive azide reagents, and (3) a novel $SnBr_4$-catalyzed bromoacetamidation.

The initial Diels–Alder step easily provided adduct **149** at room temperature on multigram scale in 97% yield and with >97% ee; recovery of the chiral ligand corresponding to **148** is simple and efficient. Ammonolysis of **149** produced amide **150** quantitatively, and subsequent iodolactamization of **150** using the Knapp protocol [113,114] generated lactam **151**, which was transformed by *N*-acylation with $(Boc)_2O$ into Boc derivative **152** in very high yield. Dehydroiodination of **152** occurred cleanly with DBU to give **153**, which was allylically brominated using NBS to generate **154** very efficiently. Treatment of **154** with $CsCO_3$ in EtOH afforded the diene ethyl ester **155** quantitatively. The next step in the synthetic sequence was a novel $SnBr_4$-catalyzed bromoacetamidation reaction which was completely region- and stereoselective using *N*-bromoacetamide (NBA) in MeCN at −40°C that converted the diene **155** to the bromodiamide **156**. It is surmised that this bromoacetamidation process involves the transfer of Br^+ from a $SnBr_4$-NBA complexed to the γ, δ-bond of the diene ester **155** followed by nucleophilic attack on the intermediate bromonium ion. Cyclization of **156** to the *N*-acetylaziridine was rapid and efficient using in situ generated tetra-*n*-butylammonium hexamethyldisilazane and provided the bicyclic product **157**. Reaction of **157** in 3-pentanol solution containing a catalytic amount of cupric triflate at 0°C occurred regioselectively to generate the ether **158**. Finally, removal of the Boc group and salt formation with H_3PO_4 in EtOH afforded **106**·H_3PO_4 (tamiflu) (❯ *Scheme 22*).

On the other hand, Shibasaki et al. reported the synthesis of **106** utilizing a general catalytic enantioselective ring-opening of *meso*-aziridines with $TMSN_3$ (❯ *Fig. 5*, ❯ *Scheme 23*) [115].

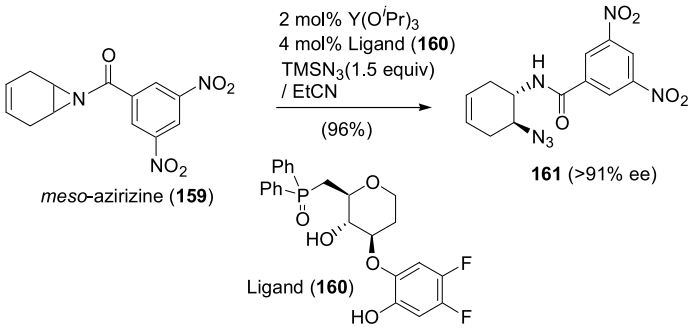
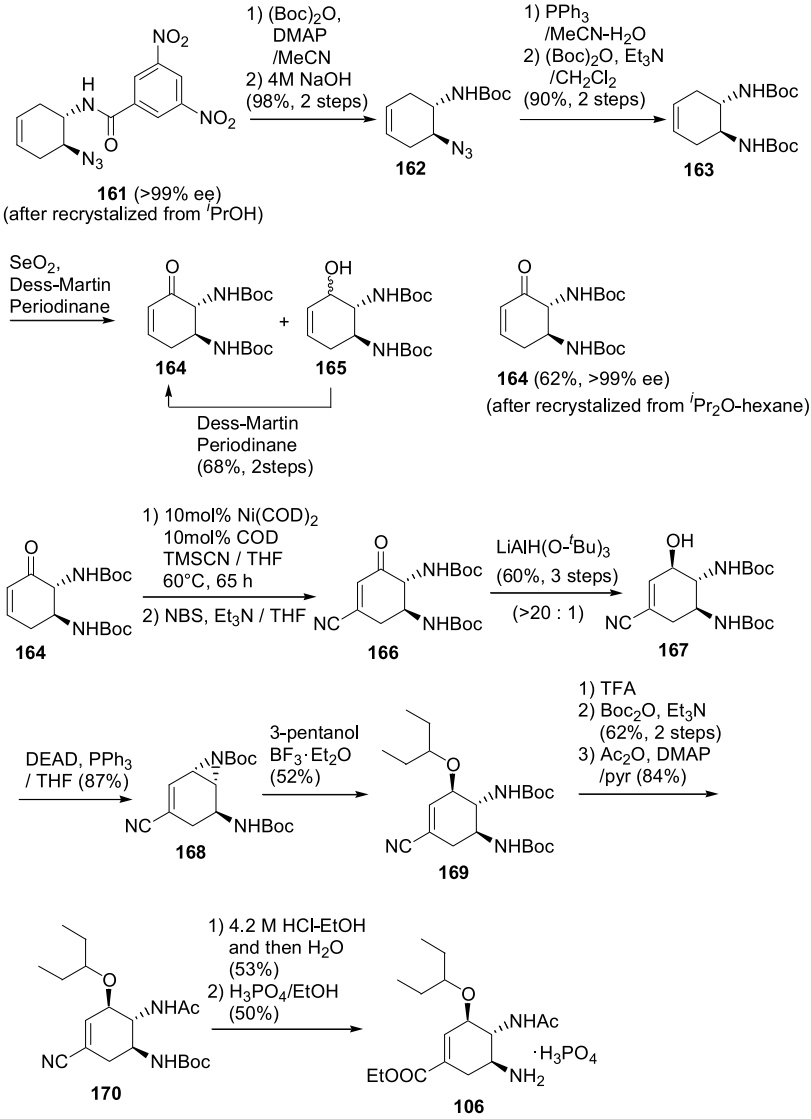


Figure 5
Key step of asymmetric synthesis of tamiflu (106) utilizing a general catalytic enantioselective ring-opening of *meso*-aziridines with $TMSN_3$



Scheme 23

The basic concept of this synthesis is on the extension of catalytic desymmetrization of *meso*-azridines (CDMA) with TMSCN using a poly Gd complex derived from ligand **160** [116], and this synthesis also includes the investigation of the possibility of extending this catalysis to using $TMSN_3$ as the nucleophile. As the results of optimization of the reaction conditions, *N*-3, 5-dinitrobenzoyl aziridines were selected as the substrates and the screening of rare earth alkoxides as the catalyst metal source indicated that $Y(O^iPr)_3$ was optimum. On the basis of these results, *meso*-azirizine **159** was treated with 2 mol% of $Y(O^iPr)_3$, 4 mol% of ligand **160**,

and 1.5 equivalents of $TMSN_3$ in EtCN to afford the product **161** in 96% yield with > 91% ee (*Fig. 5*). It can be predicted that the reaction should proceed through a mechanism similar to the CDMA with TMSCN [116]: generation of a reactive yttrium azide from $TMSN_3$ through transmetalation [117] and intramolecular transfer of the azide to an activated acylaziridine by a Lewis acidic yttrium in the same poly Y catalyst [118]. Finally, products were converted to optically active C_2 symmetric 1,2-diamines in excellent yield.

The main tasks required for the synthesis of **106** from **161** were the introduction of an oxygen functionality at the allylic position and an ethoxycarbonyl group at the olefin. After obtaining enantiomerically enriched **161** by recrystallization, C_2 symmetric diamide **163** was synthesized in four steps. Allylic oxidation of **163** with SeO_2 in the presence of Dess–Martin periodinane produced a mixture of enone **164** and allylic alcohol **165** (ca. 2:3), which was treated without purification by Dess–Martin periodinane, giving **164** in 68% yield. Enantiomerically pure (> 99% ee) **164** was obtained at this stage by recrystallization. A 1,4-addition of TMSCN in the presence of 10 mol% of $Ni(COD)_2$ followed by treatment with NBS and Et_3N produced γ-keto nitrile **166**, which was selectively reduced with bulky aluminum reagent, $LiAlH(O\text{-}^tBu)_3$ to give alcohol **167**. Aziridine formation of **167** under Mitsunobu conditions furnished **168** and, subsequently, $BF_3{\cdot}Et_2O$-mediated aziridine opening of **168** with 3-pentanol afforded **169**. Treatment of **169** with TFA, followed by protection of the sterically less hindered amine with a Boc group and acetylation afforded **170**, and conversion of nitrile **170** to ethoxycarbonyl in acidic EtOH concomitant with removal of the Boc group, and H_3PO_4 salt formation afforded tamiflu (**106**) (*Scheme 23*).

Recent advances in structural biology have led to new drug discovery processes with the neuraminidase inhibitors such as tamiflu (**106**) and relenza (**104**) being notable achievements from this scientific area. It should be emphasized that tamiflu (**106**) does not originate from a natural substrate but is an artificially designed glycosidase inhibitor based on structural biological data, which has already been prescribed for influenza patients in practice. A variety of drugs designed using structural biology, in addition to tamiflu, will appear in the clinical market in the near future.

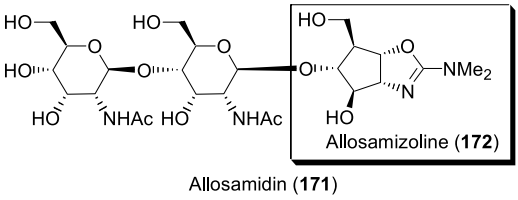
3.2 New Generations of Glycosidase Inhibitors Bearing 5-Membered Cyclitols

Among the glycosidase inhibitors, there is a type containing 5-membered cyclitol, which also exhibits interesting enzyme-specific inhibitory activities. Interestingly, most of the 5-membered cyclitols are aminocyclitols. For instance, mannostatine, which is a 5-membered aminocyclitol derivative, shows β-mannosidase-specific inhibitory activity, and allosamidin possessingthe [3.3.0] bicyclic aglycone called allosamizoline, which consists of a 5-membered aminocyclitol, exhibits inhibition towards chitinase specifically. Furthermore, trehazolin possessing the [3.3.0] bicyclic aglycone called trehalamine, which also consists of a 5-membered aminocyclitol, shows the trehalase-specific inhibitory activity. One of the scientific focuses on the researches of allosamidin and trehazolin is the elucidation of structure and activity relationships on the stereochemistry of their aglycons and their enzyme-specific inhibitory activities. . Therefore, while a variety of synthetic methodologies to form their scaffolds and to introduce the functional groups focused on total synthesis of the aforementioned glycosidase

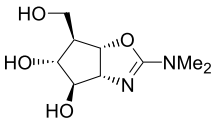
inhibitors, have been developed, the general application of the methodologies for the synthesis of their related compounds to investigate the SAR have also been reported. Herein, the new generations of glycosidase inhibitors, allosamidin and trehazolin, which possess the 5-membered aminocyclitol unit and exhibit interesting tri- or disaccharide hydrase-specific inhibitory activities, will be introduced and a variety of syntheses of the aglycones of allosamidin, allosamizoline, and trehazolins will be described.

3.2.1 Allosamizoline: The Aglycone of the Chitinase-Specific Glycosidase Inhibitor, Allosamidin

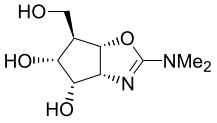
The importance of chitin as one of the main structural components of insect cuticles [119,120] and fungal cell walls [121,122] generates interest in discovering agents that may interact with its biosynthesis. With this in mind, allosamidin was discovered as expected. Allosamidin (**171**) is a novel pseudotrisaccharide as shown in Structure 4; it was isolated from the mycelial extract of *Streptomyces* sp. no. 1713, and shows the chitinase-specific inhibitory activity [123]. Its structure was deduced from hydrolysis experiments which produce 2 equivalents of D-allosamine and 1 equivalent of the aglycone moiety, allosamizoline (**172**). In particular, its aglycone moiety, allosamizoline has an interesting structure; it is a [3.3.0] bicyclic compound possessing aminocyclitol and dimethyl aminooxazoline parts. At first, the structure of allosamizoline was suggested to be the *cis*-diol (*allo*-isomer) **173** [124] and later revised to the *trans*-diol **172** [123. 124]. In the course of its correct structural assignment, the necessity for the synthesis of a series of allosamizolines (**172**–**175**) appeared and many chemists directed their efforts to the construction of these fascinating structures.



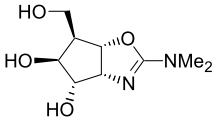
Allosamidin (**171**)



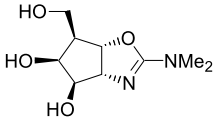
Allosamizoline (**172**)



Allo-isomer (**173**)



Gulo-isomer (**174**)



Galacto-isomer (**175**)

Trost's group and Danishefsky's group reported the total syntheses of allosamizoline independently and almost simultaneously. Both groups chose the racemic cyclopetenediol deriva-

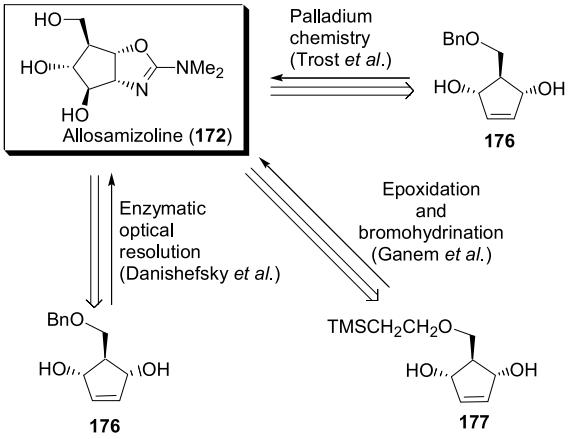


Figure 6
Diversity of the syntheses of allosamizoline (172) from cyclopentendiols 176 and 177

tive **176**, which was derived from cyclopentadienylthallium [127], as the starting material, and also Ganem's group also reported the synthesis of allosamizoline from the relevant cyclopentenediol derivative **177** later (*Fig. 6*).

First, Trost's allosamizoline synthesis will be discussed [128,129]. The basic concept of their synthesis is shown in Fig. *Fig. 7*. They have studied the use of Pd(0)-catalyzed reactions to provide the aminocyclitols with varying regio- and stereoselectivity. Also, they envisioned that the *cis*-vicinal aminocyclitols might be derived either from epoxides in a single step as in *Fig. 7* (Eq. (1), path a) [130,131] or from their synthetic equivalents such as 2-alkene-1,4-diols as illustrated in the one-pot sequence of *Fig. 7* (Eq. (1), path b) [132,133,134], with the bisurethane being generated in situ, and, in both cases, the regio- and diastereoselectivity might be assured by covalent tethering of the nitrogen nucleophile to the substrate. Moreover, they have found that significant levels of enantioselectivity may be achieved in the ionization process of the Pd(0)-catalyzed alkylation illustrated in *Fig. 7* (Eq. (2)) and have developed a class of easily prepared chiral phosphine ligands which allow the prediction of the absolute stereochemistry of the oxazolidinone product based on the stereochemistry of the ligand precursors [133,134,135,136,137,138,139,140,141].

On the basis of their preliminary synthetic concept, the allosamizolines **172–175** were synthesized as shown in *Scheme 24*, 25, and 26. The first synthesized allosamizolines were racemates and were separated as optically pure components after glycosidation with a sugar moiety, to give di-D-allosamine derivatives. Furthermore, they are continuing development of the asymmetric version by using chiral phosphine ligands. The cyclopentenediol **176**, which was derived from cyclopentadienylthallium, was treated with TsNCO and a Pd(0) catalyst to afford the oxazolidinone **178**. In a further study of this step, the use of 3 mol% of $(dba)_3Pd_2CHCl_3$ and 6 mol% of (–)-BINAPO at ambient temperature gave a 91% yield of oxazolidinone **178** with 59–65% ee [134]. After reductive de-*N*-tosylation of **178**, subsequent *O*-methylation of oxazolidinone **179** with MeOTf and exposure of the resultant imino ether to dimethylamine furnished the aminooxazoline **180**. Treatment of aminooxazoline **180** with

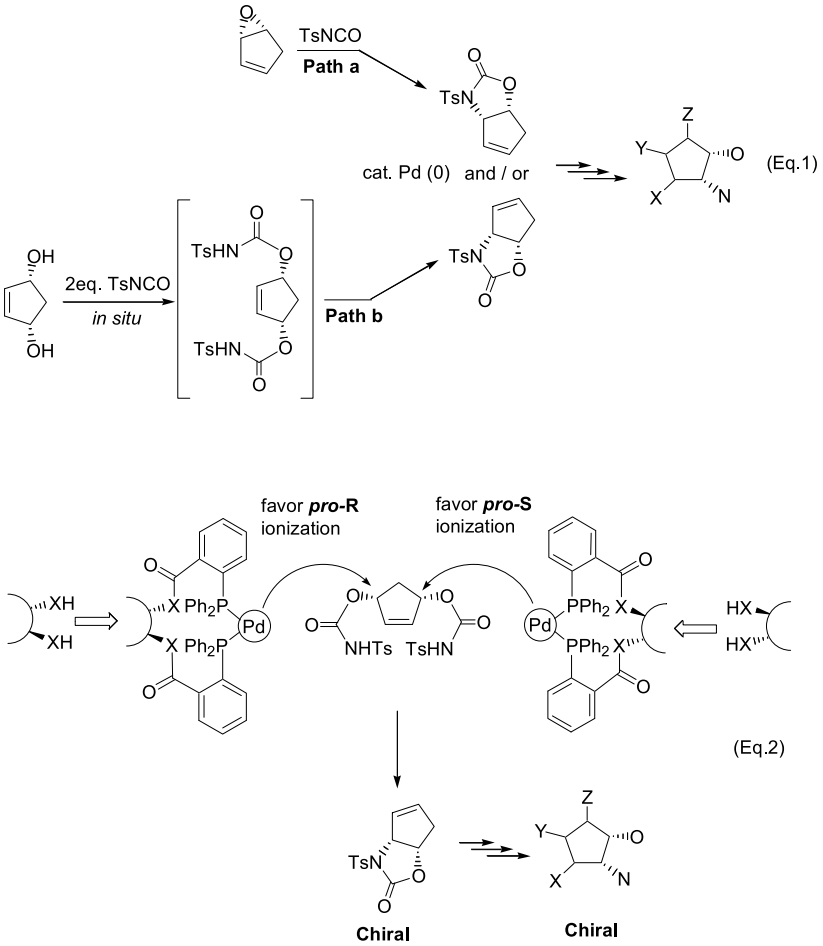
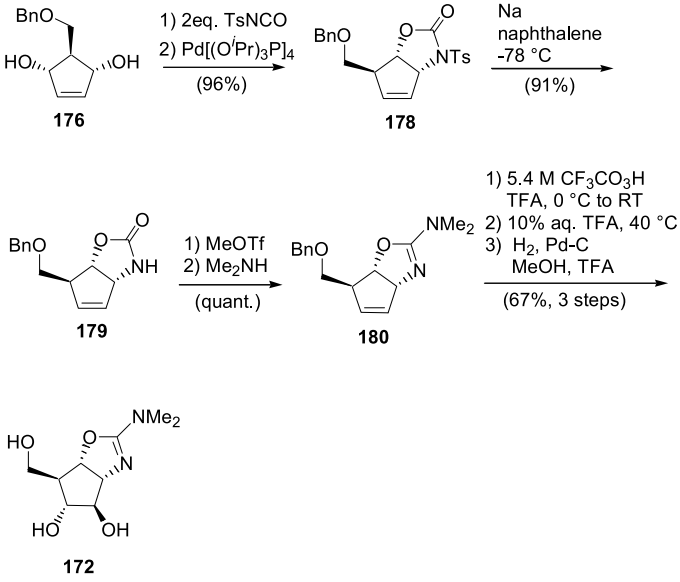
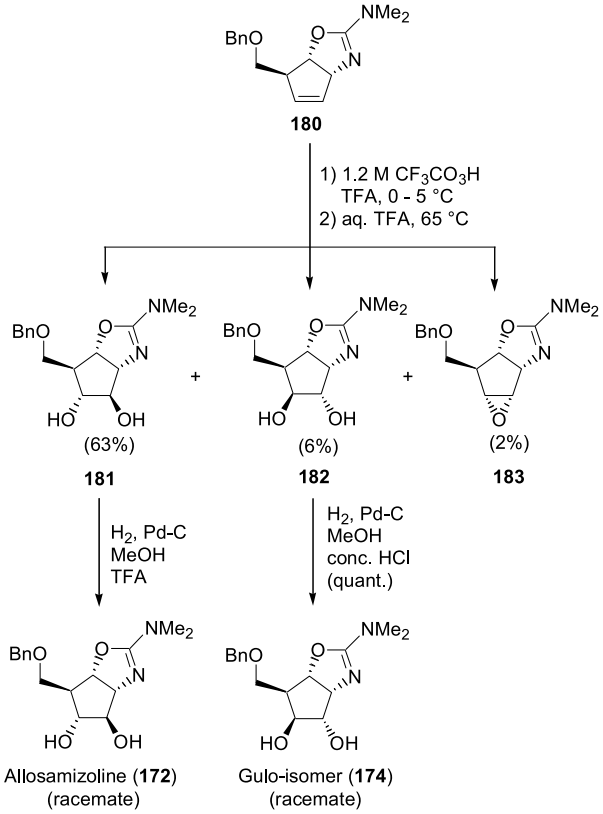


Figure 7
Basic concept of Pd(0)-mediated optical resolution for the synthesis of allosamizoline (172) reported by Trost's group

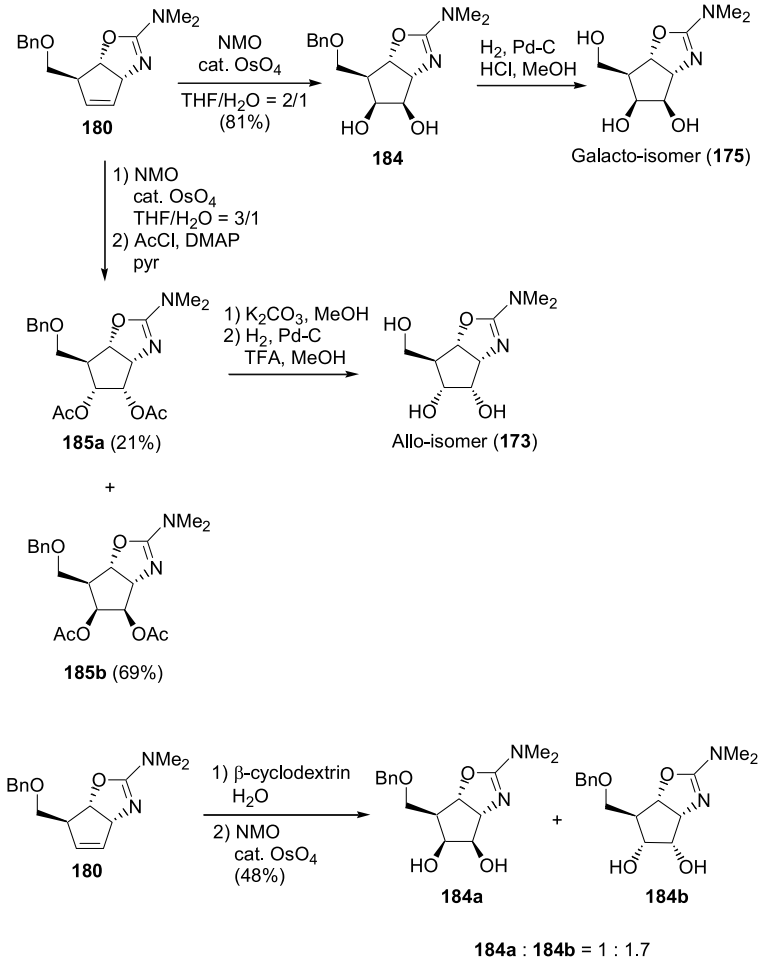
5.4 M CF_3CO_3H, exposure of the corresponding epoxide to 10% aqueous TFA at 40°C, and subsequent hydrogenolysis afforded allosamizoline **172** (*Scheme 24*).
In the meantime, the *gulo*-isomer **174** was also derived from the intermediate **180**, via minor product **182** after the treatment with 1.2 M CF_3CO_3H and subsequent exposure of the corresponding epoxide to 10% aqueous TFA at 65°C (*Scheme 25*).
On the other hand, the initially proposed structure, the *cis*-diol (*allo*-isomer) **173** and its *galacto*-type isomer **175** were synthesized from the same aminooxazoline **180** via dihydroxylation with OsO_4 as shown in *Scheme 26*.
Also, the concept and outcomes for this synthetic study were applied to the syntheses of mannostatine A [129] and valienamine [142].



Scheme 24



Scheme 25



Scheme 26

Next, Danishefsky's allosamizoline synthesis will be described [143,144]. The key point of their synthetic strategy is the utilization of enzymatic optical resolution to the racemic substrate. As illustrated in *Fig. 8*, there are two approaches for the enzymatic optical resolution. One is the enzymatic hydrolysis of a diester [145,146,147], and the other is the enzymatic transacylation of the *meso*-diol [148,149,150] (*Fig. 8*). In Danishefsky's group, the former route was chosen as the key step. Treatment of diacetate **186** with electric eel acetylcholinesterase provided the monoacetate **187**, which was reported by Deardorrf et al. [147]. This work was also applied to the synthesis of PG F_{2a} in Danishefsky's laboratory [151]. On the basis of the success of their synthesis of PG F_{2a}, diacetate **188**, which was derived from the 2-alkene-1,4-diol derivative **176**, was treated with electric eel acetylcholinesterase as well. Interestingly, this treatment provided the unexpected monoacetate **189** in 95% yield, > 95% ee (*Fig. 8*).

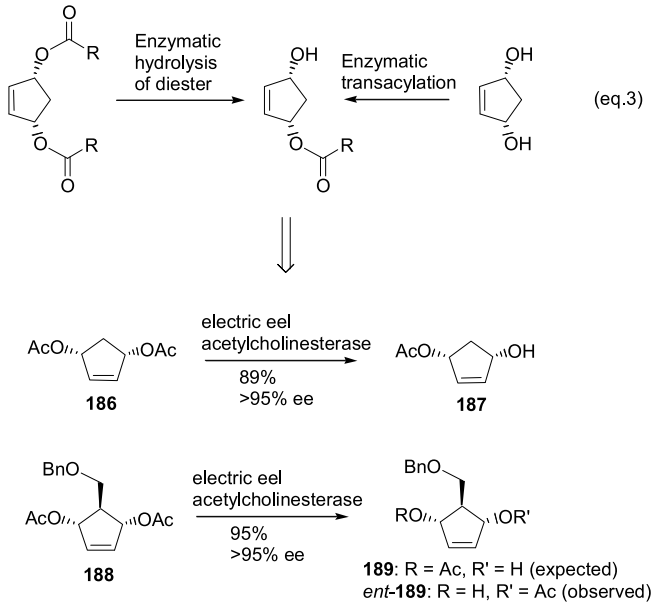
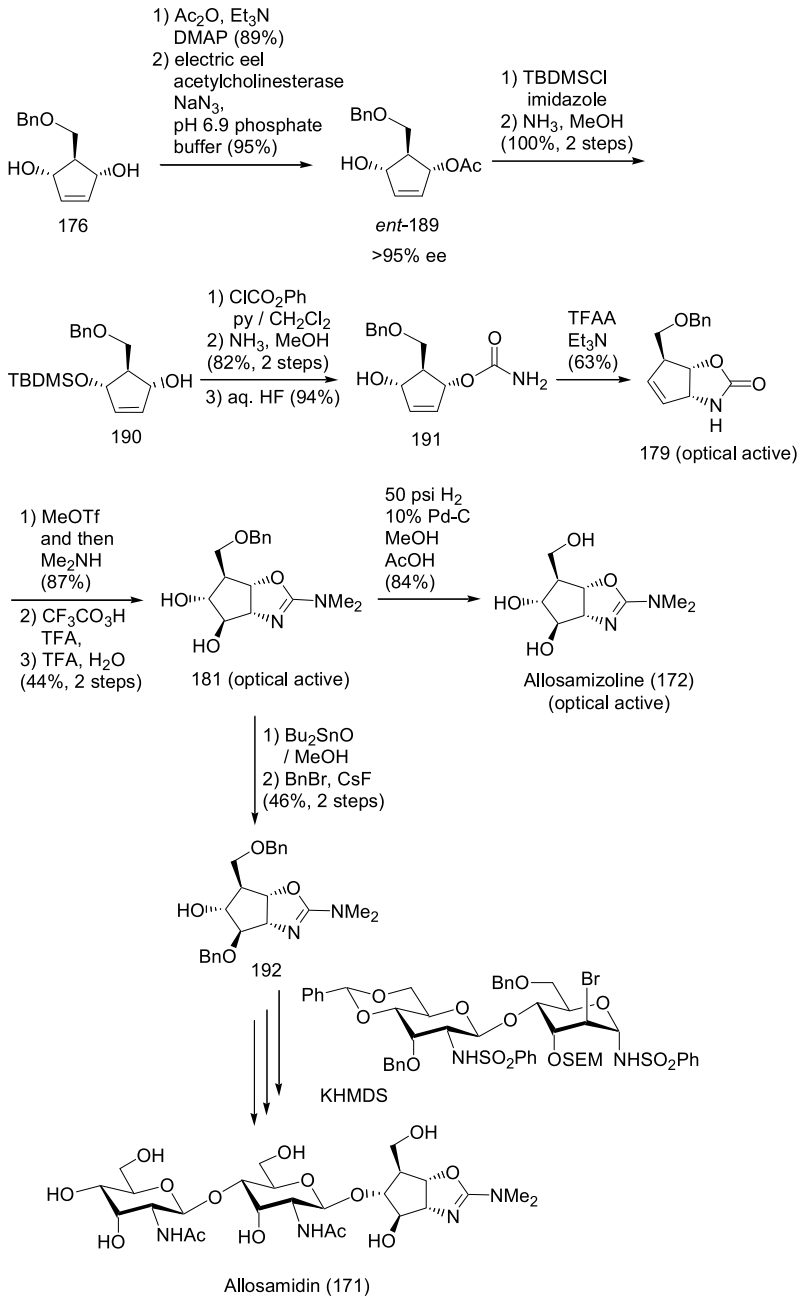


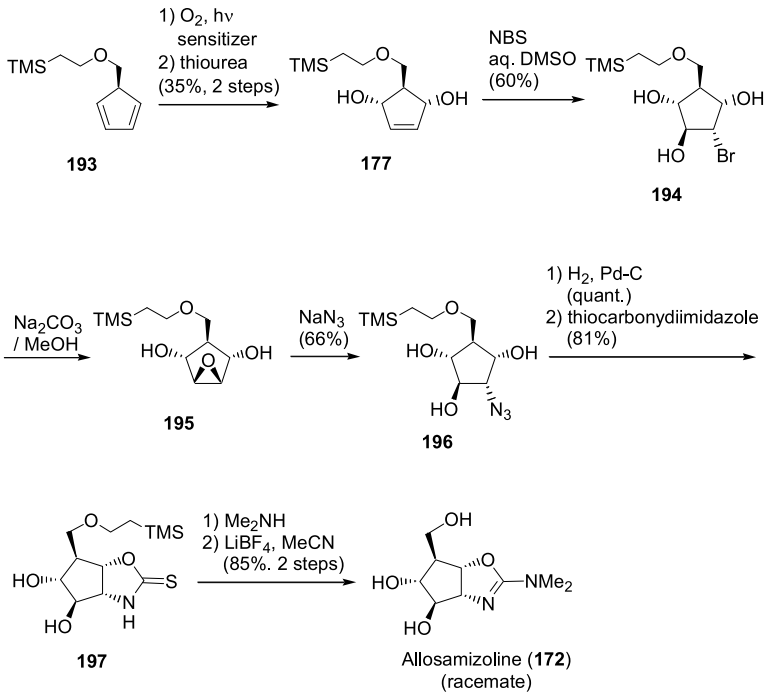
Figure 8
Basic concept of enzymatic optical resolution for the synthesis of allosamizoline (172) reported by Danishefsky's group

This result initially discouraged Danishefky's group to complete the synthesis of the intact allosamizoline **172**. However, eventually, Danishefsky et al. succeeded in the conversion of this undesired monoacetate *ent*-**189** to the intact allosamizoline (**172**) shown in *Scheme 27*. After silylation of the corresponding hydroxyl group of *ent*-**189**, the remaining Ac group was removed with NH_3 in MeOH to give the alcohol **190**. After introduction of a carbamate group by reaction with $ClCO_2Ph$ and NH_3, the TBDMS group was cleaved to provide the alcohol **191**. Treatment of **191** with TFAA and Et_3N furnished the optically active oxazolidinone **179**, which was the intermediate in Trost's synthesis. *O*-Methylation of the optical active oxazolidinone **179** with MeOTf and exposure of the resultant imino ether to dimethylamine furnished the corresponding optical active aminooxazoline. This aminooxazoline was treated with CF_3CO_3H and the corresponding epoxide was exposed to 10% aqueous TFA; subsequent hydrogenolysis under acidic conditions afforded the optical active allosamizoline (**172**). In the meantime, the resulting dibenzyl ether **192**, which was regioselectively benzylated after treatment of **181** with Bu_2SnO, was converted into allosamidine (**171**) via the KHMDS-promoted coupling with bromosulfonamide-disaccharide, which is by way of the interesting glycosidation developed by this group (*Scheme 27*).

As mentioned above, Ganem's group also synthesized allosamizoline (**172**) from the 2-alkene-1,4-diol derivative **177** [152]. They derived the 2-alkene-1,4-diol derivative **177** from cyclopentadienylthallium [127] via alkylation with SEMCl, cycloaddition of the corresponding cyclopentadiene **193** with singlet oxygen, and subsequent reduction in situ of the transient endoperoxide. This 2-alkene-1,4-diol derivative **177** was treated with NBS in DMSO to afford


1) Ac2O, Et3N
DMAP (89%)
2) electric eel
acetylcholinesterase
NaN3,
pH 6.9 phosphate
buffer (95%)
1) TBDMSCl
imidazole
2) NH3, MeOH
(100%, 2 steps)
BnO
HO
OH
OAc
176
ent-189
>95% ee
1) ClCO2Ph
py / CH2Cl2
2) NH3, MeOH
(82%, 2 steps)
3) aq. HF (94%)
TFAA
Et3N
(63%)
TBDMSO
NH2
190
191
OBn
179 (optical active)
1) MeOTf
and then
Me2NH
(87%)
2) CF3CO3H
TFA,
3) TFA, H2O
(44%, 2 steps)
181 (optical active)
50 psi H2
10% Pd-C
MeOH
AcOH
(84%)
NMe2
Allosamizoline (172)
(optical active)
1) Bu2SnO
/ MeOH
2) BnBr, CsF
(46%, 2 steps)
192
Ph
Br
NHSO2Ph
OSEM
KHMDS
NHAc
Allosamidin (171)

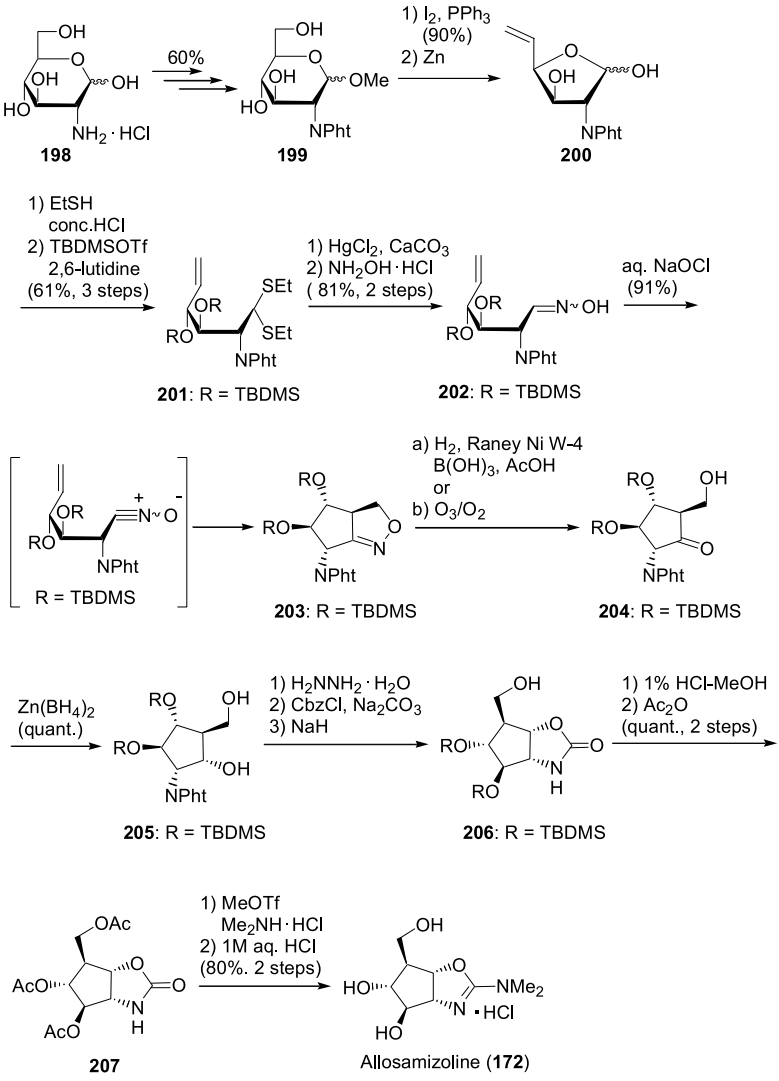
Scheme 27



Scheme 28

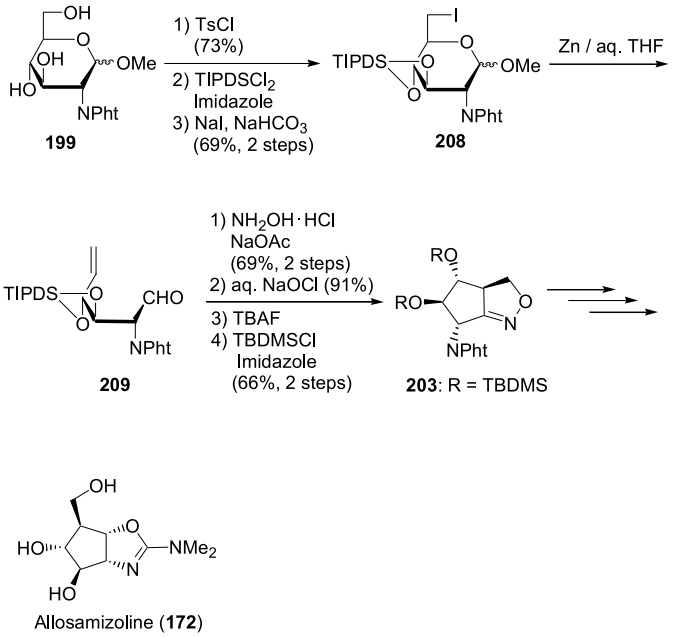
a single bromohydrin **194**. The structure was confirmed by cyclization with Na_2CO_3-MeOH to furnish exclusively the epoxide **195** possessing the desired stereochemistry. Ring-opening with NaN_3 furnished the racemic azidotriol **196**, and subsequently compound **196** was reduced to corresponding aminotriol. Cyclization with thiocarbonyldiimidazole produced the thiooxazolidinone **107**. After a one-step construction of the dimethylaminooxazoline ring had been achieved by heating with Me_2NH-MeOH, deprotection of the silyl group gave the racemic allosamizoline (**172**) (*Scheme 28*).

Allosamizoline (**172**) was also synthesized from natural sugar. Tatsuta et al. reported the synthesis of allosamizoline (**172**) from D-glucosamine using an intramolecular [3+2] cycloaddition as a key step [153]. On the other hand, Kitahara's group also independently achieved the synthesis of allosamizoline using the same key step [154]. First, Tatsuta's allosamizoline synthesis will be described. Iodination of methyl 2-deoxy-2-phthalimide-D-glucopyranoside **199**, which was readily derived from D-glucosamine **198**, and subsequent reductive β-elimination of the corresponding iodide gave the 5-enofuranose **200** concomitant with the reductively dehalogenated C6 product. Treatment of this mixture with EtSH and concentrated HCl, and subsequent silylation, furnished dithioacetal **201**. Dedithioacetalization of **201** with $HgCl_2$ and $CaCO_3$ provided the corresponding aldehyde, and, subsequently, the aldehyde was treated with $NH_2OH{\cdot}HCl$ to afford oxime **202**. Intramolecular [3+2] cycloaddition of oxime **202** with aqueous NaOCl furnished the isoxazoline **203** via nitrile oxide [13,14]. Isoxazoline opening



Scheme 29

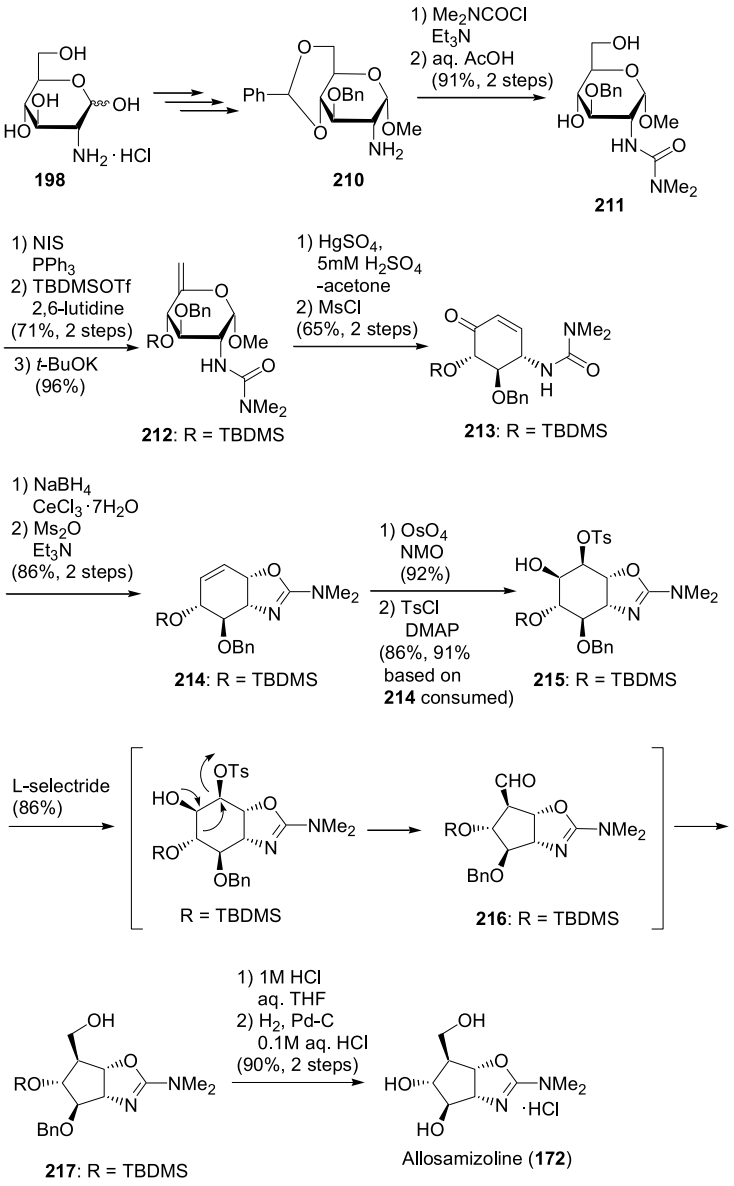
of **203** was a troublesome step, however, and as a result of screening the reaction conditions, it was completed by hydrogenolysis with Raney-Ni W4 in the presence of $B(OH)_3$ and AcOH or ozonolysis [155] to give the β-hydroxy ketone **204**. The reduction of **204** with $Zn(BH_4)_2$ proceeded stereoselectively and afforded a single diol **205**. After de-*N*-phthaloylation of **205**, benzyloxycarbonylation followed by base treatment gave the oxazolidinone **206**. Acid desilylation and subsequent acetylation afforded the triacetate **207**. Then *O*-methylation of **207** with MeOTf, followed by the treatment with $Me_2NH \cdot HCl$ and Et_3N, as well as further acidic deacetylation furnished allosamizoline (**172**) (*Scheme 29*).



Scheme 30

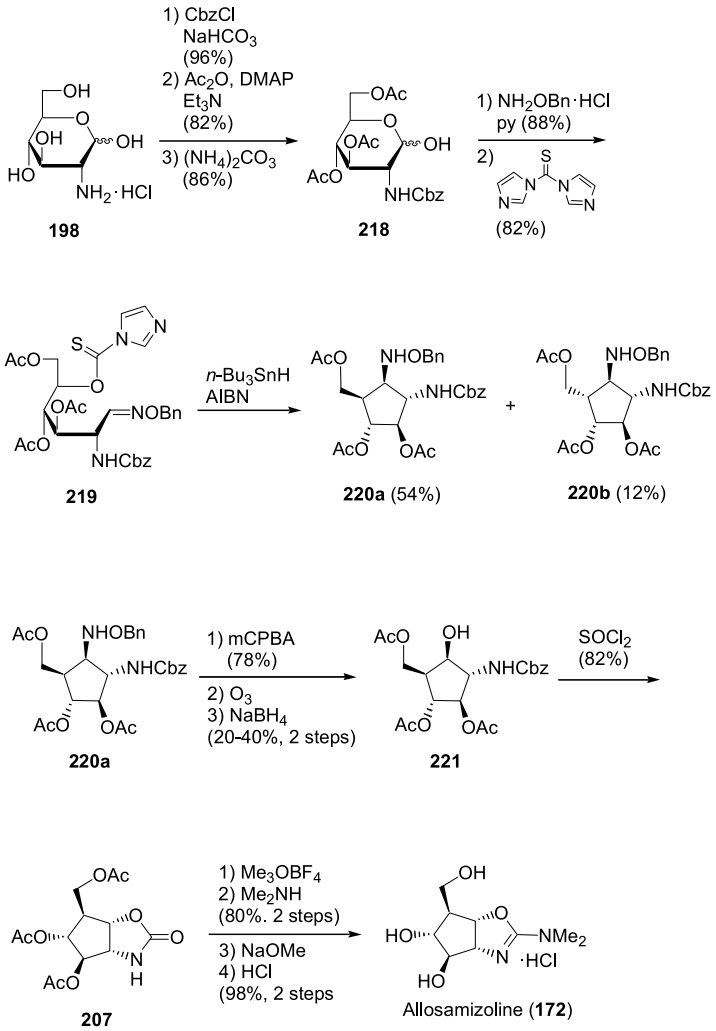
As shown in *Scheme 30*, Kitahara et al. also derived allosamizoline (**172**) from D-glucosamine **198** via Tatsuta's intermediate, isoxazoline **203**, as well [154].

As described in the session on cyclophelitol, the Ferrier reaction is a very useful method to transform sugars into carbasugars, and Kuzuhara's group used this reaction in the synthesis of allosamizoline (**172**) shown in *Scheme 31* [156]. A further interesting point in their synthesis is the transformation of the [4.3.0] bicyclic structure, which was derived from D-glucosamine **198** via the Ferrier reaction, to the [3.3.0] bicyclic scaffold of allosamizoline (**172**). It was performed by reductive ring contraction to the α-hydroxy tosylate under the action of L-selectride. Selective iodination of the *N*, *N*-dimethylurea-diol derivative **211** with NIS and PPh_3 [157], whereby **211** was derived from D-glucosamine **198** via the 4, 6-*O*-benzylidene compound **210**, silylation of the secondary hydroxyl group, and subsequent dehydroiodination of the corresponding iodide with *t*-BuOK gave the enol ether **212**. For conversion of the ring system from pyranose to cyclohexane, the modified Ferrier reaction with $HgSO_4$ [158] was employed and, furthermore, β-elimination of the resulting ketol with MsCl-pyridine provided the enone **213**. 1,2-Reduction of this enone **213** with $NaBH_4$ and $CeCl_3{\cdot}7\ H_2O$ proceeded stereoselectively and subsequent treatment of the corresponding allyl alcohol with Ms_2O and Et_3N furnished the dimethylaminooxazoline ring **214** simultaneously [159]. Dihydroxylation of **214** with OsO_4 proceeded exclusively from the convex face and, fortunately, the subsequent sluggish monotosylation gave the desired monotosylate **215**. Ring contraction of **215** in basic media was regarded as the key step in this synthesis. As the expected rearrangement product **216** bearing an aldehyde group was assumed to lack stability in basic media,



Scheme 31

an appropriate basic reducing agent [160] was screened for the ring contraction followed by immediate reduction of aldehyde. As the result, L-selectride was selected to give the desired cyclopentanemethanol **217**. Finally, compound **217** was converted into allosamizoline (**172**) by deprotection of the TBDMS and Bn groups (*Scheme 31*).



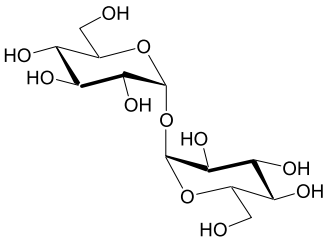
Scheme 32

Another synthesis of allosamizoline (**172**) from sugar was reported by Simpkins' group, and the key step of their synthesis is a radical cyclization to construct the 5-membered aminocyclitol [161,162]. So far, radical cyclization for the conversion of carbohydrates to carbasugars has been problematic because of the suitability of carbon-centered radicals for the preparation of the highly functionalized 5-membered ring [163]. However, significant contributions in this field, most notably by Rajanbabu [164], the possibility of using an aldehyde as the radical acceptor in the key cyclization reported by Fraser-Reid [165,166], and the cyclizations of radicals onto oxime ether reported by Bartlett [167] have paved the way to the formation of the highly functionalized 5-membered rings. Finally, Simpkins et al., focusing their attention

on a report by Bartlett [167], accomplished the synthesis of allosamizoline (**172**). In a way, Simpkins' group has verified the usefulness of the concept of radical cyclization for the formation of the highly functionalized cyclitols through the synthesis of allosamizoline (**172**) shown in *Scheme 32*. Treatment of the *N*-Cbz-tri-*O*-acetyl sugar **218**, which was derived from D-glucosamine **198** in three steps, with the *O*-benzyl ether of hydroxylamine, followed by derivatization of the secondary alcohol with thiocarbonylimdazole gave the compound **219**. This oxime derivative was treated with Bu_3SnH and AIBN according to Bartlett's protocol, to afford a mixture of diastereomeric products **220**. Judging from the configuration of the major product **220a**, obtained as a mixture of C1 epimers, it was expected to be convertible into allosamizoline (**172**). Oxidation of the mixture of benzyloxyamines with mCPBA afforded the corresponding oxime [168], and ozonolysis of the resulting oxime, followed by direct reductive work up gave the alcohol **221**. Treatment of **221** with $SOCl_2$ provided the oxazolidinone **207**, and the reaction proceeded with inversion at C1. *O*-Methylation of **207**, followed by the exposure to Me_2NH, and deacetylation gave allosamizoline (**172**) (*Scheme 32*).
Interestingly, in each case of the synthesis of allosamizoline (**172**) from sugar, D-glucosamine **198** was chosen as the starting material, and, in fact, there is the notion that D-glucosamine **198** is considered to be a biosynthetic precursor of allosamizoline (**172**). In a way, these syntheses from D-glucosamine might pave the way to elucidate the biosynthetic route of allosamizoline (**172**).

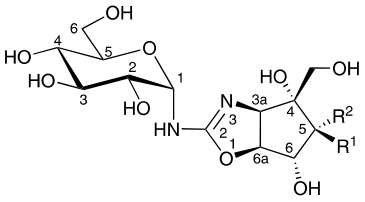
3.2.2 Trehazolin, Trehalamine and Its Aminocyclitol Moiety, Trehazolamine: A Trehalase-Specific Glycosidase Inhibitor

Trehalose (**222**) is a structurally interesting disaccharide in which two D-glucoses are linked by an $\alpha(1\text{–}1)\alpha$ glycosyl bond.



Trehalose (**222**)

Biologically, it is considered that D-glucose degraded from trehalose (**222**) with trehalase, which is a kind of glycosidase and is possibly used as an energy source by insects, and a series of glycosidase inhibitors exhibiting inhibitory activity towards this enzyme, would be expected to be a new type of potential insecticide. In 1991, Ando et al. reported the isolation of trehazolin (**223**) from the culture broth of *Micromonospora* sp. strain SANK 62390 [169,170]. This compound is a unique natural pseudodisaccharide showing a strong trehalase-specific inhibitory activity. Nearly concomitantly, the isolation of a compound named trehalostatin (**224**) from the culture broth of *Amicoratopsis trehalostatica* was also reported [171].



R^1 = OH, R^2 = H Trehazolin (**223**)

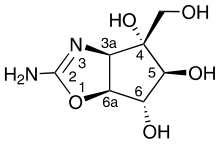
IC_{50} = 0.011 μg/ml (Silkworm trehalase)
IC_{50} = 0.006 μg/ml (Porcine trehalase)

$[\alpha]^{25}_D$ +99.5 ‹ (*C* 0.44, H_2O)

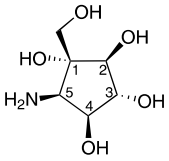
R^1 = H, R^2 = OH Trehalostatin (**224**)

IC_{50} = 0.68 ng/ml (Blowly trehalase)

$[\alpha]^{25}_D$ +115 ‹ (*C* 1.0, H_2O)

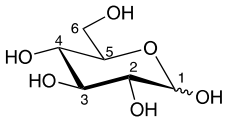


Trehalamine (**225**)

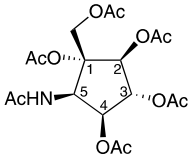


Aminocyclitol moiety
(Trehazolamine) (**226**)

Ac_2O, DMAP
py, rt, 24 h



D-Glucose

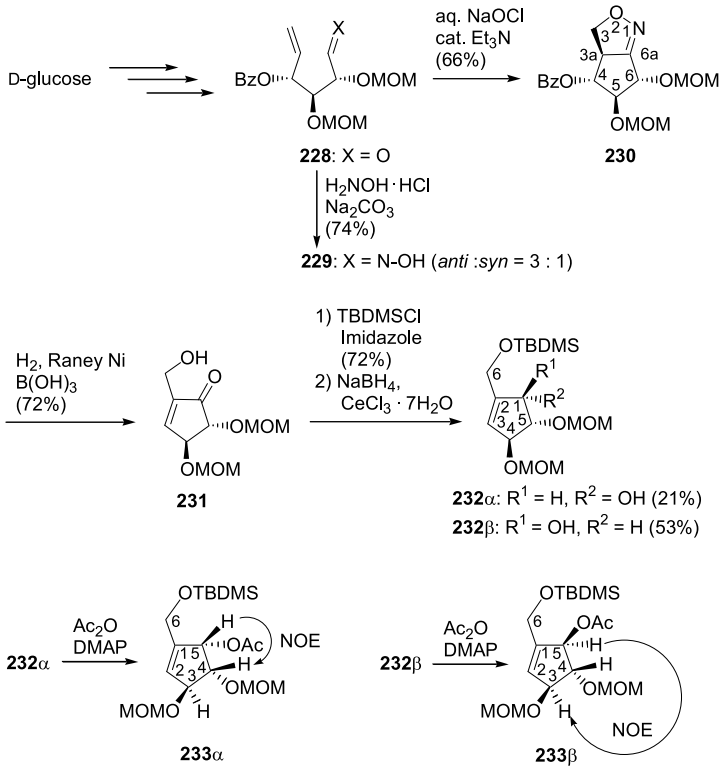


Aminocyclitol moiety
(Trehazolamine)
hexaacetate (**227**)

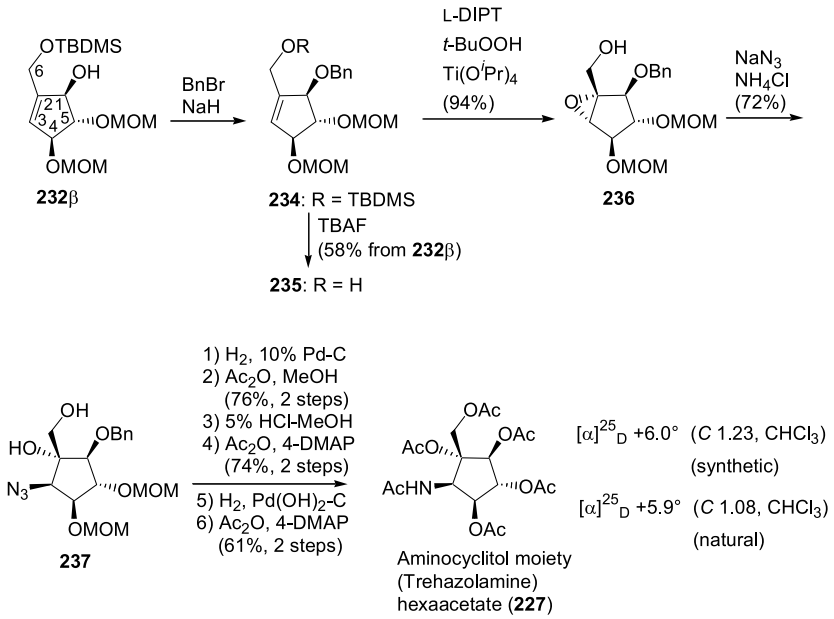
Trehalostatin (**224**) was reported to be the C5 epimer of the aglycone of trehazolin (**223**). According to the reported physical and biological data, trehalostatin (**224**) seems to be identical to trehazolin (**223**). Because both compounds are noncrystalline, X-ray crystallography could not be used to determine the correct stereochemistry; hence, the stereochemistry at the C5 position of the aglycones remains unknown. Therefore, a series of degradation studies of trehazolin (**223**) was attempted to try to obtain degradation products from which the correct stereochemistry could be deduced. Hydrochloric acid degradation provided degradation products: the aglycone moiety trehalamine (**225**), D-glucose, and the aminocyclitol moiety (**226**), from which we next obtained the acetylated product **227**. However, all products were also noncrystalline. Consequently, they could not be used to determine the correct stereochemistry. In the meantime, further NMR analysis of trehazolin (**223**) and a series of biochemical studies to elucidate its trehalase-specific inhibitory activity were undertaken, and on the basis of

these studies, it was surmised that the actual structural resemblance between trehazolin (**223**) and trehalose (**222**) may bear on the generation of trehazolin's activity towards various trehalases, and the absolute configuration of its aminocyclitol moiety (**226**) was hypothesized as [1*R*-(1α, 2β, 3α, 4β, 5β)]. Then, in order to verify the proposed correct structure of trehazolin (**223**), a series of synthetic studies were performed. The first enantioselective synthesis of trehazolin (**223**) and its components was accomplished by Kobayashi and Shiozaki et al. [172,173,174]. It was performed from D-glucose, and intramolecular [3 + 2] cycloaddition was used as the key step to form the 5-membered scaffold of the aminocyclitol moiety. As the first stage of the synthesis, the hexaacetate of the aminocyclitol moiety was synthesized to determine the absolute configuration of the aminocyclitol unit shown in *Scheme 33* and *Scheme 34* [173,174].

Intramolecular [3+2] cycloaddition of the oxime **229**, derived from D-glucose according to the method developed by Bernet and Vasella [175], furnished the corresponding isoxazoline **230**. In general, hydrogenolysis of isoxazolines **230** with Raney Ni in the presence of boric acid [176] results in conversion to a hydroxymethyl ketone. However, in this case, by virtue of the electron-withdrawing effect of the benzoyl group, β-elimination of the benzoyloxy group by the generated ketone was induced and the corresponding hydroxymethyl



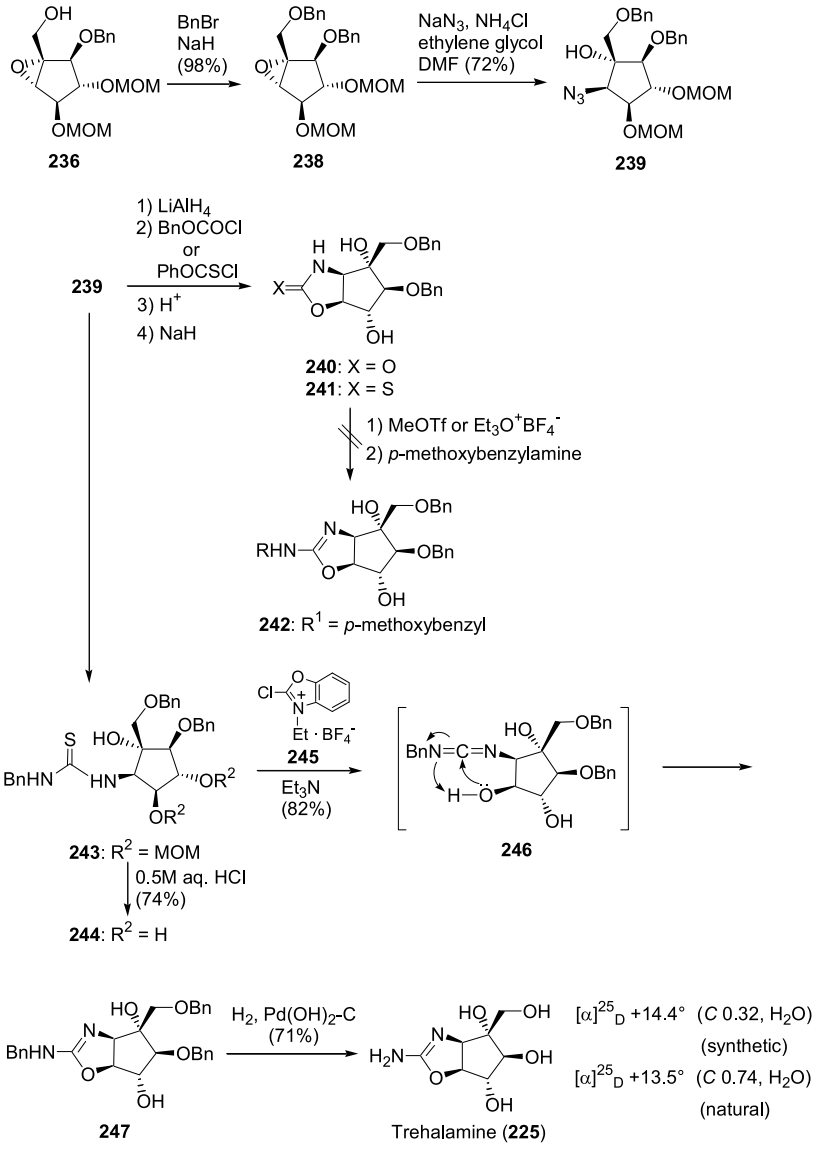
Scheme 33



Scheme 34

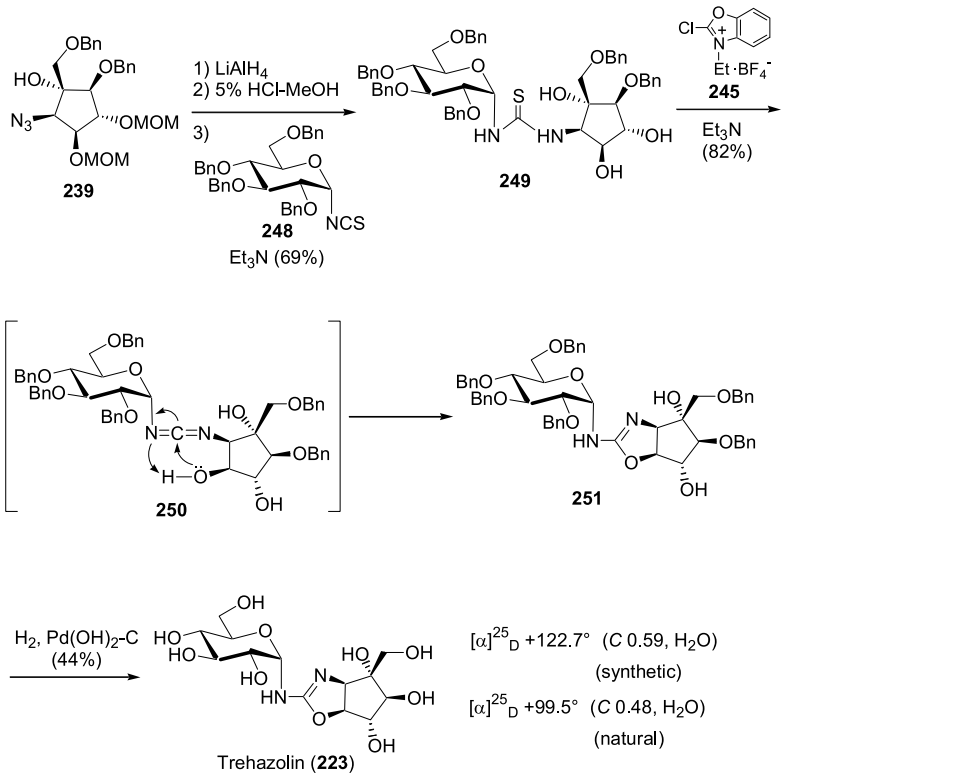
enone **231** was obtained. Silylation of the enone **231** and subsequent 1,2-reduction of the corresponding ketone afforded a separable 1:2.5 mixture of alcohols **232**. The configurations at the C1 positions of these alcohols **232** were determined by analysis of the ^{1}H-NMR data of the acetates **233**; that is, NOEs were observed between C5-H and C4-H of compound **233***α* and between C5-H and C3-H of compound **233***β* (*Scheme 33*).

Benzylation of the alcohol **232*β***, possessing the desired configuration, and subsequent removal of the TBDMS group of compound **234** afforded the corresponding allyl alcohol **235**. Several types of epoxidation towards the allyl alcohol **235** were attempted; Sharpless' epoxidation using diisopropyl L-tartrate [177] furnished the desired epoxide **236** as a single isomer. Finally, the synthesis of hexaacetate of the aminocyclitol moiety (**227**) was accomplished by regiospecific azide opening towards the chiral epoxide **236**, deprotection, and subsequent complete acetylation. This synthetic aminocyclitol hexaacetate was identical in all respects to the aminocyclitol hexaacetate obtained from the degradation product of natural trehazolin, and the absolute configuration was found to be [1*R*-(1*α*, 2*β*, 3*α*, 4*β*, 5*β*)] as expected (*Scheme 34*). Next, the synthesis of trehazolin aglycone, trehalamine (**225**) was conducted as shown in *Scheme 35*. The key point of this synthesis was the method by which the aminooxazoline ring was formed. Initially, the aminooxazoline formation, which was used for allosamizoline synthesis [128,129,143,144,153,161,162], was attempted via *O*-methylation of **240** or **241** derived from epoxide **236**. The *O*-alkylated compounds derived from **240** and **241** were exposed to benzylamine derivatives, but this method did not yield the desired product **242**. Finally, the cyclization of thiourea alcohol via the carbodiimide alcohol was attempted. The aminooxazoline ring can be considered as an equivalent of the *cis*-carbodiimide alcohol,



Scheme 35

and a number of methods to generate carbodiimides were investigated with the method of Mukaiyama [178,179] being selected. This synthetic method to derive carbodiimides from thiourea alcohol uses 2-chloro-3-ethylbenzoxazolium tetrafluoroborate **245** and related reagents, and was also applied to the synthesis of indolemycine [180]. The thiourea alcohol **244** derived from the azide alcohol **239** was treated with 2-chloro-3-ethylbenzoxazolium tetrafluoroborate **245**, furnishing the aminooxazoline derivative **247** via transformation of



Scheme 36

the *cis*-thiourea alcohol **244** to the corresponding carbodiimide alcohol **246** and subsequent cyclization of **246**. Finally, the aminooxazoline **247** was hydrogenolyzed to cleave three benzyl groups and to give trehalamine (**225**). This synthetic trehalamine was identical in all respects to natural trehalamine (*Scheme 35*).

On the basis of the synthesis of trehalamine (**225**), the synthesis of trehazolin (**223**) was undertaken. Coupling between the amino alcohol derived from the azide compound **239**, and the α-D-glucopyranosyl isothiocyanate derivative **248**, as synthesized by Camarasa [181], afforded the α-D-glucopyranosylthiourea derivative **249**. Treatment of this thiourea **249** with 2-chloro-3-ethylbenzoxazolium tetrafluoroborate **245** furnished the aminooxazoline derivative **251** via carbodiimide alcohol **250** as the intermediate. Finally, this aminooxazoline **251** was hydrogenolyzed to cleave the benzyl groups and to generate trehazolin (**223**). This synthetic trehazolin was identical to the natural trehazolin in all respects, including biological activities (*Scheme 36*).

Also, Ogawa's group reported the synthesis of racemic trehazolin aminocyclitol derivatives, as well as trehazolin and its diastereoisomers from *myo*-inositol, and thus contributed independently to the determination of the correct stereochemistry of trehazolin [182,183,184,185,186]. The successful total synthesis of trehazolin (**223**) encouraged Kobayashi and Shiozaki et al. to investigate structure-activity relationships regarding the inhibitory activities towards vari-

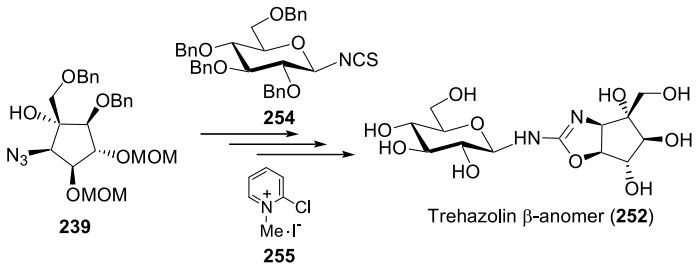
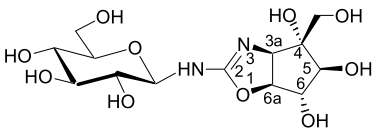
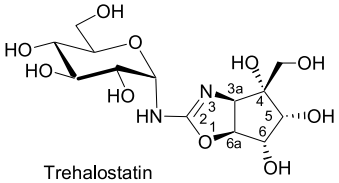
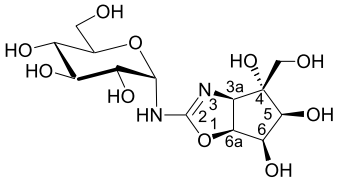


Figure 9
Synthesis of the trehazolin β-anomer

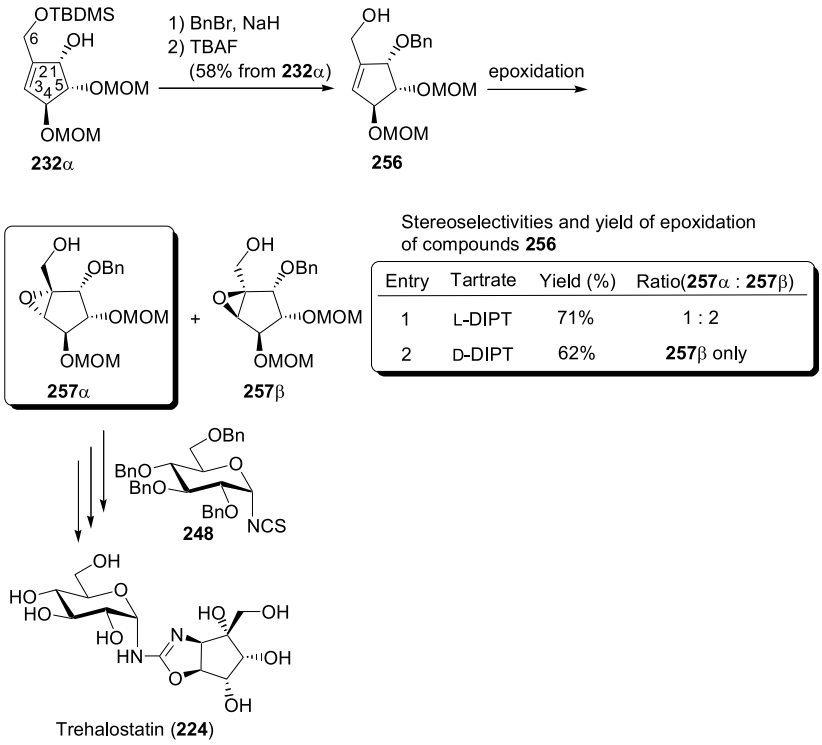
ous α-glucosidases resulting from the stereochemistry of trehazolin. They then designed the trehazolin stereoisomers.







To investigate the influence of the stereochemistry of the anomeric position on the inhibitory activities, the β-anomer of trehazolin (**252**) was synthesized from the azide **239** and the β-D-glucopyranosyl isothiocyanate derivative **254** [187]. To avoid anomerization and contamination of the α-anomer, 2-chloro-1-methylpyridinium iodide **255** was used in place of 2-chloro-3-ethylbenzoxazolium tetrafluoroborate **245** for the ring formation (*Fig. 9*).
At the step of the 1,2-reduction of the hydroxymethylenone **231**, the allyl alcohol **232**α was also obtained. Trehalostatin (the C5-epimer of trehazolin) (**224**) was then synthesized from this allyl alcohol **232**α according to the synthetic route to trehazolin (**223**) [188]. The physical data, including the ^{1}H-NMR spectrum, of trehalostatin (**224**), were quite close to those of trehazolin (**223**) itself, but the inhibitory activities of this compound towards trehalases were much weaker than those of trehazolin (**223**). As a result, the argued stereochemistries of the aminocyclitol moiety of trehazolin (**223**) and trehalostatin (**224**) were determined to be [1*R*-(1α, 2β, 3α, 4β, 5β)] and [1*R*-(1α, 2α, 3α, 4β, 5β)] by these synthetic studies, respectively (*Scheme 37*).



Scheme 37

Next, in order to investigate further the structure-activity relationships regarding the inhibitory activities towards various α-glucosidases resulting from the stereochemistry of the aminocyclitol moiety, the C-6 epimer of trehazolin (**253**) was synthesized [189,190]. In this synthesis, the challenging tandem aldol-Wittig type reaction was performed to construct the enone **[II]**. As shown in *Fig. 10*, it was expected that treatment of the silylenol lactone **[III]**, which was derived from D-ribonolactone, with the α-lithiated phosphorane [191] would give the enone **[II]** via the aldol reaction and subsequent intramolecular Wittig reaction in one pot (*Fig. 10*). *Scheme 38* shows the practical synthesis of the C-6 epimer of trehazolin (**253**). Treatment of the silylenol lactone **259**, which was derived from D-ribonolactone **258** in three steps, with the lithiated phosphorane furnished the cyclopentenone **260** in moderate yield. This reaction should thus have synthetic utility for the one-step synthesis of cyclic α, β-unsaturated ketones from cyclic enol ester-type derivatives. Afterwards, the synthesis proceeded basically according to the trehazolin synthesis, including 1,2-reduction of the enone **260**, stereoselective epoxidation, regiospecific azide-opening of the epoxide **262**, coupling between the amine derived from compound **263** and the isothiocyanate **248**, and cyclization to form the aminooxazoline ring **265**. Finally, compound **265** was hydrogenolyzed to cleave the benzyl groups and to furnish trehazolin C-6 epimer (**253**) (*Scheme 38*).

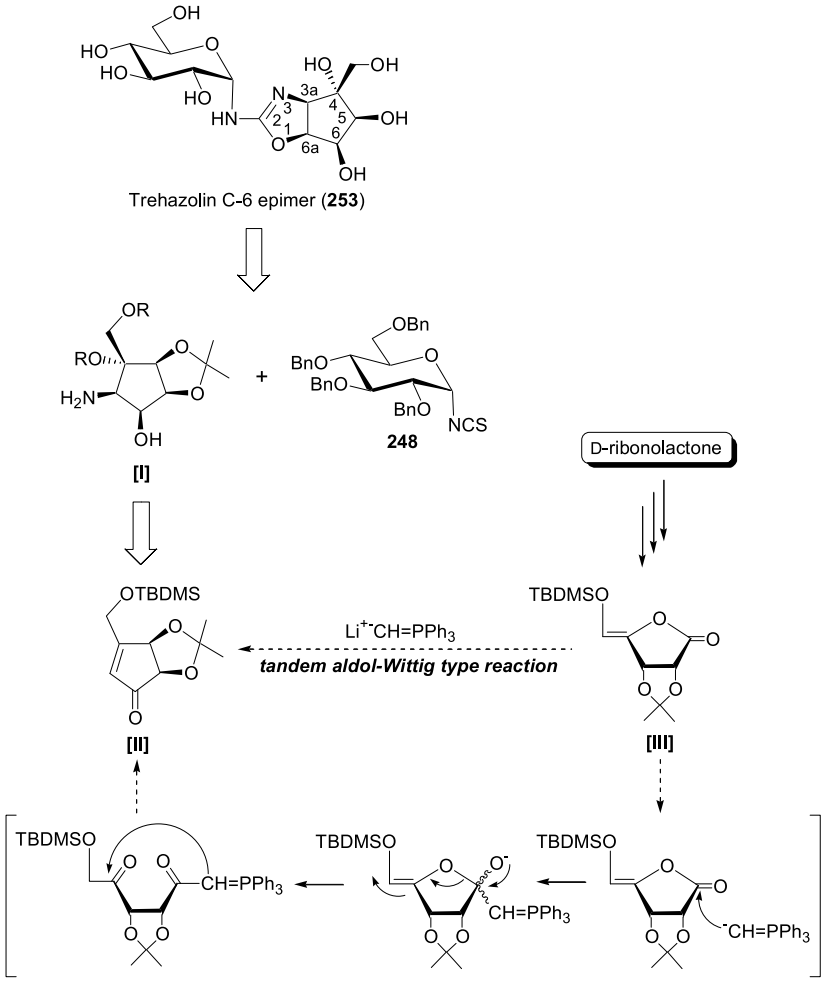
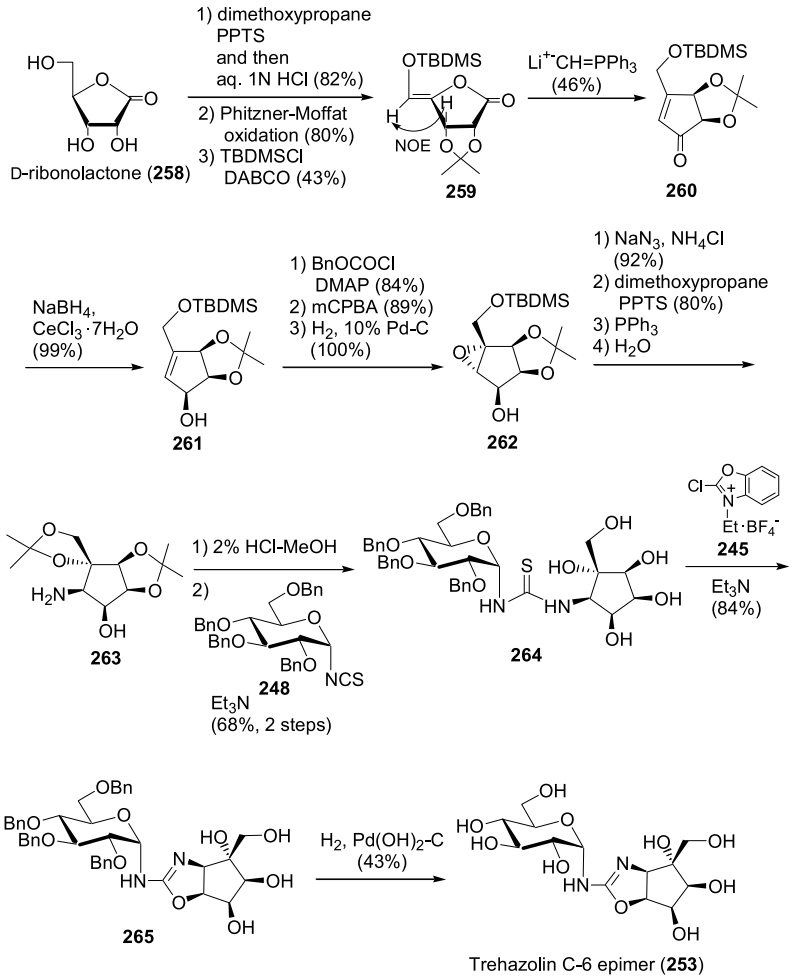


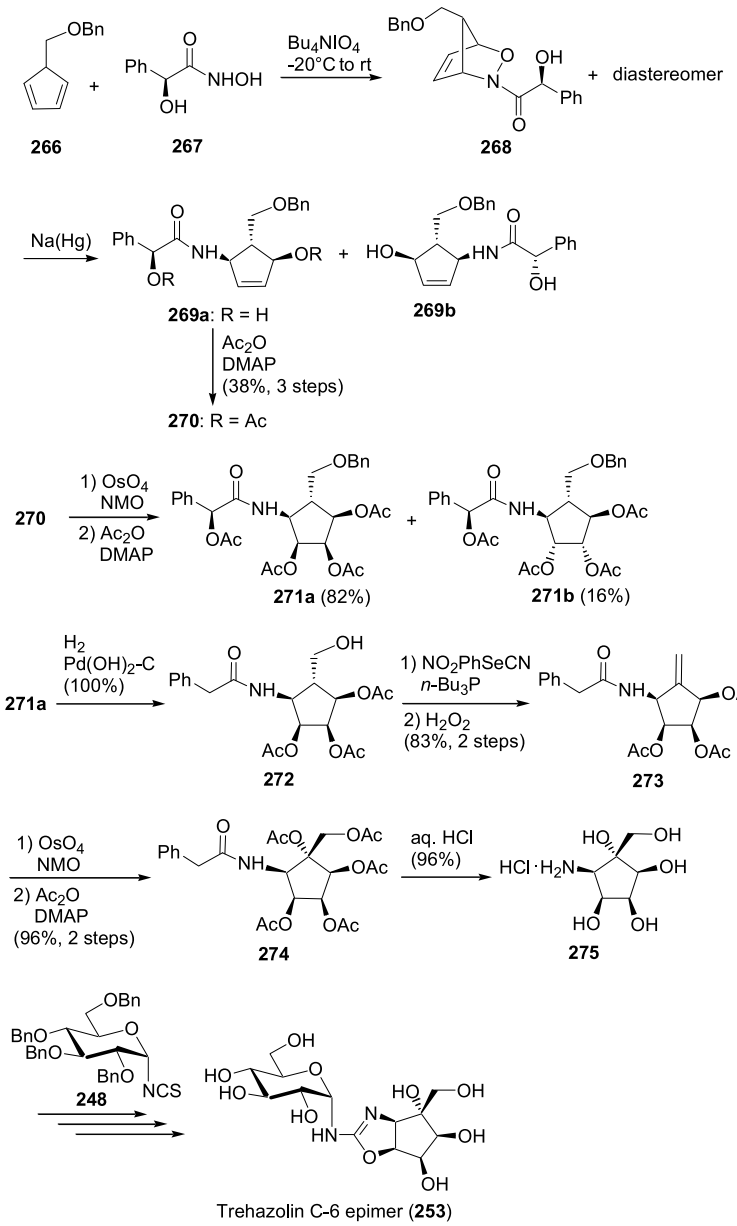
Figure 10
Synthetic strategy of trehazolin C-6 epimer (253) via tandem aldol-Wittig reaction

Ganem et al. also reported the synthesis of the trehazolin C-6 epimer (**253**) and the formal total synthesis of trehazolin (**223**) [192,193]. The synthetic strategy used was related to the synthesis of allosamizoline, and involved the cyclopentadiene derivative **266**, which was also used for the synthesis of allosamizoline, as the starting material. Treatment of the cyclopentadiene **266** [194] with (*S*)-mandelohydroxamic acid **267** in the presence of Bu_4NIO_4 led to a mixture of the desired cycloadduct **268** and its diastereomers. The inseparable mixture was reduced using Na(Hg) for separation to afford the pure cyclopentene **269a** and the minor cycloadduct **269b**. Compound **269a** was converted to acetate **270** for characterization of the structure. Dihydroxylation of **270** with OsO_4 favored *syn* addition, and subsequent acetylation gave compound **271a** selectively and quantitatively. The next key point in



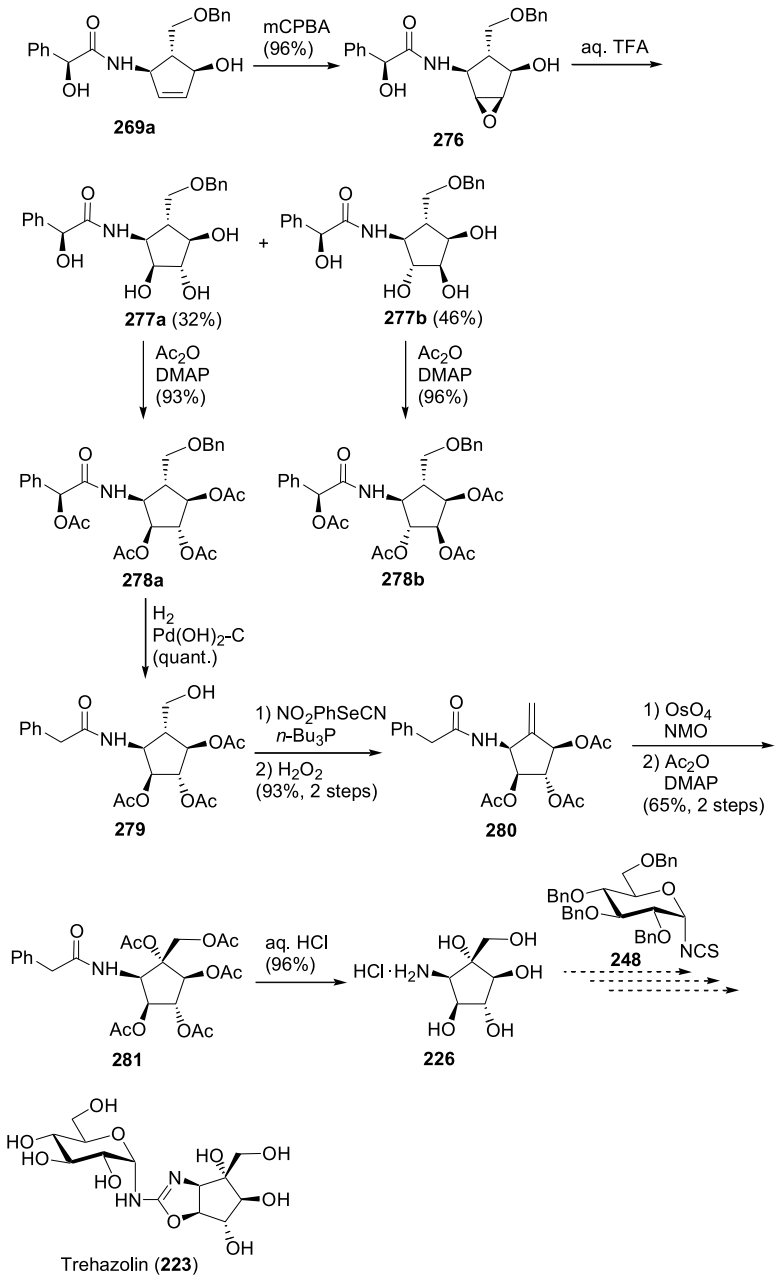
Scheme 38

the synthesis was the introduction of a quaternary stereocenter. Hydrogenolysis of both benzylic ether and Ac groups of **271a** using $Pd(OH)_2$ on carbon afforded the alcohol **272**, and nitrophenylselenation followed by in situ oxidative elimination according to the method of Grieco et al. [195] cleanly converted **272** into alkene **273**. Flanked by two allylic substituents shielding the top face of the 5-membered ring, the exocyclic alkene increment in **273** underwent vicinal hydroxylation exclusively from the opposite face and subsequent acetylation furnished acetate **274**. Exhaustive acid hydrolysis of **274** provided the aminocyclitol **275** required for the synthesis of the C6-epimer of trehazolin (**253**). This aminocyclitol hydrochloride **275** was converted into the C6 epimer of trehazolin (**253**) in analogy to the syntheses of trehazolin (**223**) and its stereoisomers **224**, **252**, and **253** reported by Kobayashi and Shiozaki et al. [172,173,174,187,188,189,190] and Ogawa's group [182,183,184,185,186] (*Scheme 39*).

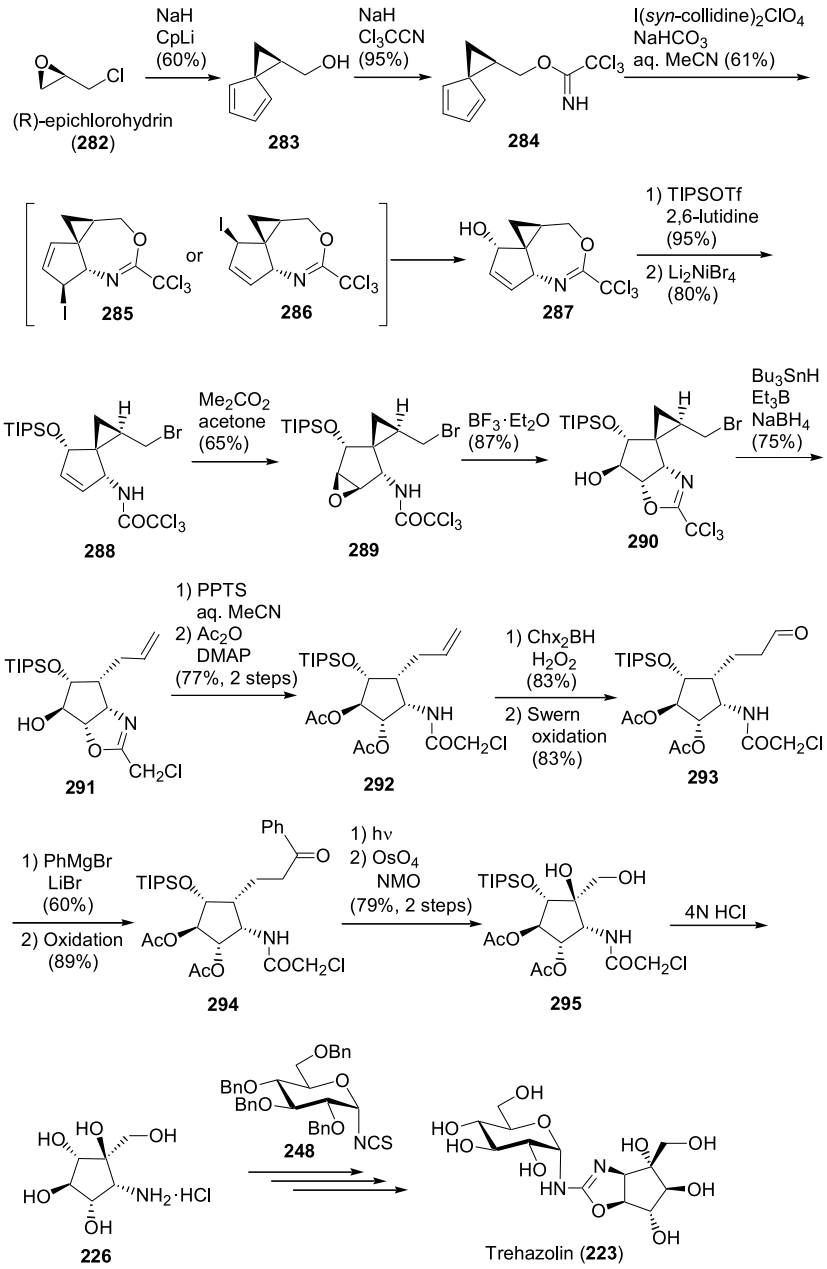


Scheme 39

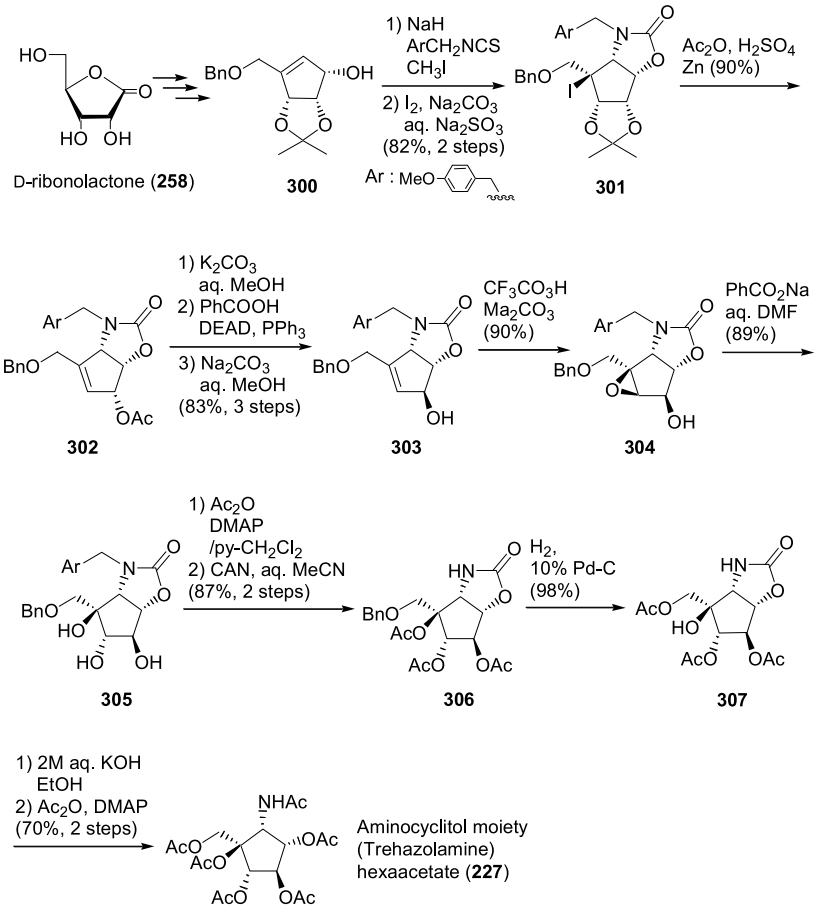
Moreover, these authors prepared the aminocyclitol **226** needed for the synthesis of trehazolin (**223**) from the allyl alcohol **269a**. Epoxidation of **269a** with mCPBA gave the desired epoxide **276** exclusively. The epoxide **276** was exposed to 2:1 H_2O-TFA to open the epoxy ring, and afforded the alcohols **277**. They were transformed to peracetates **278a** and **278b**,


mCPBA
(96%)
aq. TFA
269a
276
277a (32%)
277b (46%)
Ac2O
DMAP
(93%)
Ac2O
DMAP
(96%)
278a
278b
H2
Pd(OH)2-C
(quant.)
279
1) NO2PhSeCN
n-Bu3P
2) H2O2
(93%, 2 steps)
280
1) OsO4
NMO
2) Ac2O
DMAP
(65%, 2 steps)
281
aq. HCl
(96%)
226
248
NCS
Trehazolin (223)

Scheme 40


NaH
CpLi
(60%)
NaH
Cl3CCN
(95%)
I(syn-collidine)2ClO4
NaHCO3
aq. MeCN (61%)
(R)-epichlorohydrin
(282)
283
284
or
285
286
287
1) TIPSOTf
2,6-lutidine
(95%)
2) Li2NiBr4
(80%)
Me2CO2
acetone
(65%)
BF3·Et2O
(87%)
Bu3SnH
Et3B
NaBH4
(75%)
288
289
290
1) PPTS
aq. MeCN
2) Ac2O
DMAP
(77%, 2 steps)
1) Chx2BH
H2O2
(83%)
2) Swern
oxidation
(83%)
291
292
293
1) PhMgBr
LiBr
(60%)
2) Oxidation
(89%)
1) hν
2) OsO4
NMO
(79%, 2 steps)
4N HCl
294
295
248
NCS
226
Trehazolin (223)

Scheme 41



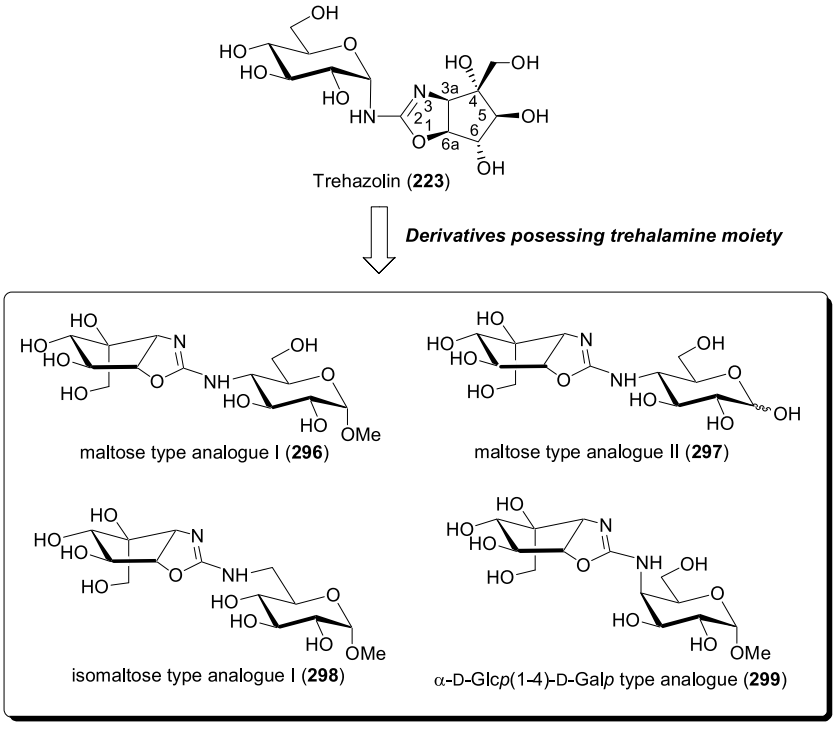
Scheme 42

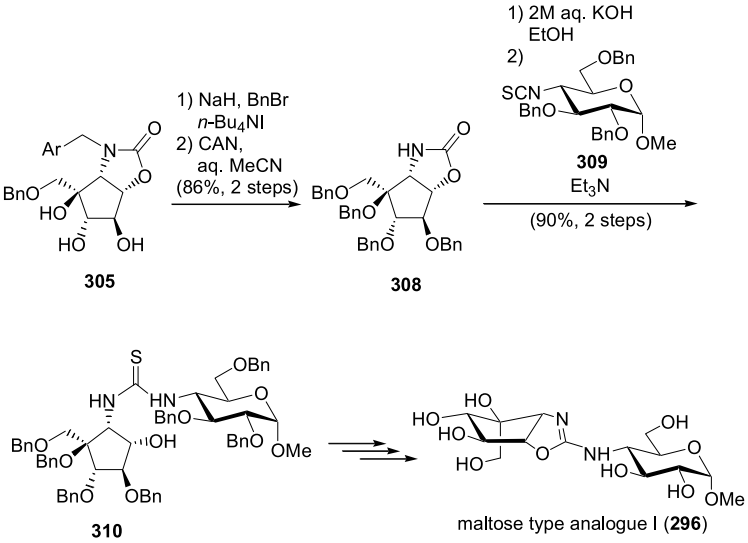
respectively, for further characterization of the structure of the alcohols **277**. The desired peracetate **278a** was converted into the aminocyclitol **226** necessary for the synthesis of trehazolin (**223**), according to the synthetic procedure for the C-6 epimer of trehazolin (**253**), and thus completed the formal total synthesis of trehazolin (**223**) (*Scheme 40*).

On the other hand, Carreira et al. also reported the total synthesis of trehazolin from the optically active spirocycloheptadiene [196], which was prepared from the (*R*)-epichlorohydrin **282** and lithium cyclopentadienide, shown in *Scheme 41*. Treatment of lithium cyclopentadienide (CpH+BuLi) with (*R*)-epichlorohydrin **282** afforded the optically active spirocycloheptadiene **283** in 91% ee. Compound **283** was converted into trichloroacetimidate **284** by treatment of NaH and Cl_3CCN [197], and subsequent treatment of **284** with I(*syn*-collidine)$_2ClO_4$ gave the alcohol **287** via the unstable intermediates **285** and **286**. After silylation of the secondary hydroxy group of **287**, the imidate underwent nucleophilic opening upon treatment of the corresponding silyl-protected imidate with Li_2NiBr_4 to yield the cyclopropylcarbinyl bro-

mide **288** [198]. The alkene moiety in **288** was converted into the epoxide **289** with dimethoxydioxorane [199], and subsequent treatment of the resulting epoxide **289** with $BF_3 \cdot Et_2O$ afforded the alcohol **290**. Conversion of **290** to **291** was performed under free radical conditions, which also resulted in partial reduction of the trichloromethyl moiety. Oxazoline **291** was directly converted into **292** by treatment with aqueous PPTS, followed by acetylation. Hydroboration of the terminal olefin unit in **292**, followed by Swern oxidation of the resulting primary alcohol, provided aldehyde **293** [200,201]. Conversion of **293** to the corresponding phenyl ketone **294** was achieved by treatment of **293** with PhMgBr and oxidation of the resulting secondary alcohol. The aryl ketone **294** underwent Norrish type II cleavage upon irradiation through a Pyrex filter in degassed PhH, giving the corresponding alkene, which, without purification, was reacted with catalytic OsO_4 to yield diol **295** as a single diastereomer. Exhaustive acid hydrolysis of **295** provided the aminocyclitol hydrochloride **226** necessary for the synthesis of trehazolin (**223**), and this corresponding aminocyclitol **226** was transformed to trehazolin (**223**) according to the preceding synthesis (*Scheme 41*).

After the total synthesis of trehazolin (**223**) and its stereoisomers, the further focus of synthetic studies on trehazolins was directed to the design and exploration of the other glucosidase-specific inhibitors possessing the trehalamine moiety, because trehalamine is considered to be a kind of pseudo-D-glucose. Compound **296** was designed by Knapp et al. as a pseudo-maltose and it was expected to have a potential for maltase inhibition [202]. On the other hand, compound **296** and its related compounds **297**, **298** and **299** were independently designed by Kobayashi and Shiozaki et al. [203].





Scheme 43

Knapp et al. achieved the synthesis of the trehazolin aminocyclitol moiety (trehazolamine) as the hexaacetate **227** in the process of the synthesis of a maltose-type of trehazolin derivative **296**. The allylic alcohol **300**, available from D-ribonolactone **258** in several steps according to the method of Marquez [204], was converted to its thiocarbamate derivative by condensation with *p*-methoxybenzyl isothiocyanate, which in turn was cyclized with I_2 to afford the iodooxazolidinone **301**. This is a rare example of formal anti-Markovnikov iodocyclization and may result from the kinetic preference for 5-membered (rather than 6-membered) and fused (rather than bridged) ring formation. One-pot treatment of **301** with Ac_2O and H_2SO_4 followed by activated Zn provided allylic acetate **302**, whose configuration was inverted at the C3 position (trehazolin aminocyclitol numbering), by the Mitsunobu procedure [205], and the allyl alcohol **303** was obtained. Epoxidation of the resulting alcohol **303** gave the desired epoxide **304** as a single stereoisomer detected, and then hydrolysis of **304** at the less substituted position (C2 position, trehazolin aminocyclitol numbering) led to the triol **305**. The stereochemistry of the cyclitol chemistry was proven by the conversion of **305** to the aminocyclitol hexaacetate **227** (*Scheme 42*).

Moreover, the triol **305** was converted into the *O*-benzylated compound **308**, which was transformed to compound **296** via coupling with D-glucopyranosyl isothiocyanate **309** [206], formation of aminooxazoline ring, and complete removal of protecting groups (*Scheme 43*).

Kobayashi and Shiozaki et al. synthesized the trehazolin-related compounds **296**–**299** from aminocyclitol **226** degraded from natural intact trehazolin (**223**) [203]. The purpose of the synthesis was to discover the pseudosaccharides with interesting α-glucosidase inhibitory activities, and especially, compounds **296** and **297** were designed as intestinal maltase inhibitors and **298** was designed as an intestinal isomaltase inhibitor, and these inhibitors were directed towards potent therapies for noninsulin-dependent diabetes. The general synthetic route is summarized in *Fig. 11*, and this route was established, based on the experience that cycliza-

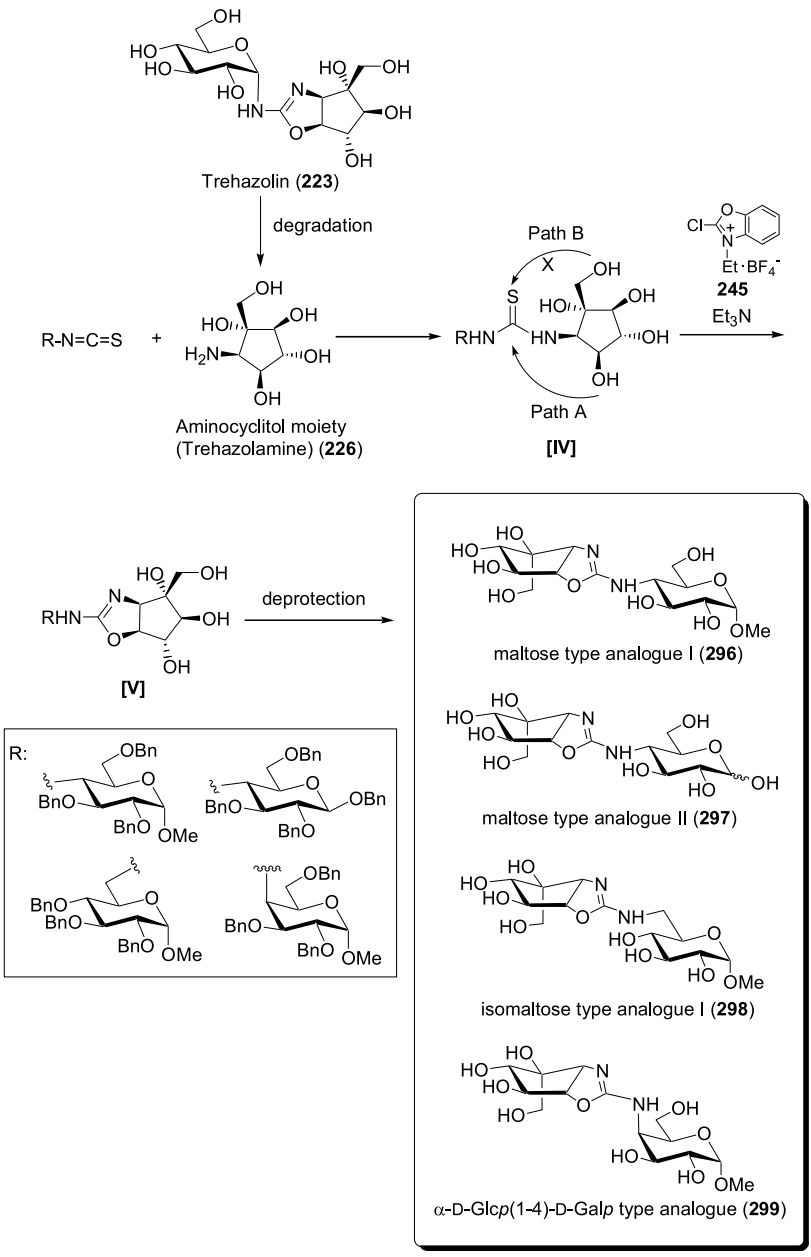
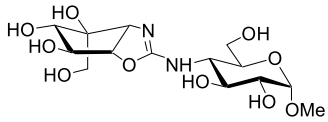


Figure 11
General synthetic route of trehazolin-related derivatives 296–299

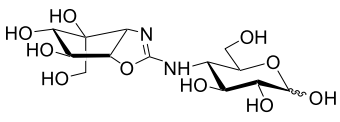
Table 3
Inhibitory activity of trehazolin-related compounds 296–299 toward glucosidases

enzyme	origin	trehazolin (**223**)	(**296**)	(**297**)	(**298**)	(**299**)
trehalase	silkworm	0.011	>100	>100	>100	>100
trehalase	porcine	0.006	>100	0.245	92	30
maltose	rat	76	>100	>100	9	>100
isomaltose	rat	3.9	>100	>100	55	72
sucrase	rat	76	>100	>100	10	>100

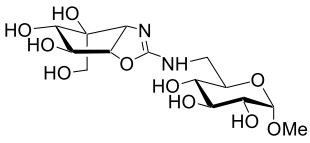
IC_{50}: μg/ml



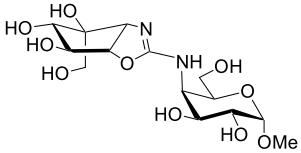
maltose type analogue I (**296**)



maltose type analogue II (**297**)



isomaltose type analogue I (**298**)



α-D-Glc*p*(1-4)-D-Gal*p* type analogue (**299**)

tion of compound **[IV]** proceeded through path A to form thermodynamically stable [3, 3, 0] bicyclic structure in their synthesis of trehazolin C-6 epimer (**253**) [189,190].

Compounds **296**–**299** were synthesized according to this general synthetic route successfully and inhibitory activities of these compounds are shown in *Table 3*.

Interestingly, the structure-activity relationships of these compounds based on structural similarities couldn't be found. Compound **298**, which was designed as a pseudo-isomaltose, inhibited maltase and sucrase more potently than trehazolin (**223**); whereas compounds **296**, **297**, and **299** did not exhibit inhibitory activities toward maltase, isomaltase, and sucrase, compared to trehazolin (**223**). On the other hand, while none of the derivatives inhibited silkworm trehalase at a concentration of 100 μg/ml, only compound **297** possessed inhibitory activity toward porcine trehalase, with an IC_{50} value of 0.245 μg/ml. These results suggest that the interaction between an α-glucosidase and a glucose unit would be rather diverse among various glucosidases even though they can catalyze common substrates, and particularly, in the case of α-glucosidase reacting to disaccharide specifically, it was concluded that accurate structural analyses of the complexes consisting of the inhibitors and target enzymes are necessary for the design of the inhibitors with specific α-glucosidase inhibitory activity.

4 New Methods for Conversion of Sugars to Carbasugars

The definition of carbasugars can be described as highly oxygenated and functionalized carbocycles. Also, these types of carbocycle components are included in the biologically significant molecules, such as a variety of enzyme inhibitors represented by glycosidase inhibitors, nucleoside mimetics, and natural products. Therefore, the methodologies of carbocycle formation should be designed to be applied to a wide range of the synthesis.
In general, as mentioned earlier, the Ferrie reaction, intramolecular [3 + 2] cycloaddition of nitrile oxide and cyclization of the diene with Grubbs' catalysts are powerful methods as the key reactions for conversion of sugars to carbasugars. In addition to these, regarding the synthesis of 6-membered carbasugars, transformation of furanolactones to carbasugars via C1 unit prolongation and Lewis acid-mediated aldol-like cyclization, is also a useful synthetic method [207,208,209]. Also, regarding the transformation of a 6-membered to a 5-memberd ring, the development of methodologies has received much attention. Most of the methodologies involve a free-radical approach in which Bu_3SnH mediated a well-known 5-hexenyl-type radical cyclization from bromide, iodide, or thiocarbonyl imidazole precursors [210,211,212]. In the meantime, the chemical characters of Sm and Zr have been a topic of interest and their utility in synthetic chemistry continues to be examined. In particular, synthetic exploitation of the reagent SmI_2 has rapidly become one of the most significant fields in organic chemistry [213], and a great number of important synthetic transformations with SmI_2 involve the one-electron reduction of ketones and aldehydes to a samariumketyl radical anion which can promote cyclizations, deoxygenations, and reductions. In fact, during this decade, some new methodologies for the transformation of pyranose derivatives to 5-memberd carbasugars using Sm and Zr reagents have been reported. In this section, novel synthetic methodologies for conversion of sugars to carbasugars will be introduced.

4.1 Carbasugar Formation via $SnCl_4$-Promoted Intramolecular Aldol Condensation

Tatsuta et al. developed the method of carbasugar formation using $SnCl_4$-promoted intramolecular aldol condensation as a key step. The methodology is shown in ❯ *Fig. 12*, and a key precursor of carbasugar, α-hydroxymethyl enone **[XI]** was achieved according to the following steps: (1) 1,2-addition of lithiated $MeSO_2Ph$ to furanolactones containing a dimethyl acetal **[VII]**, (2) transformation of compound **[VIII]** to linear silyl enol ether **[IX]**, (3) $SnCl_4$-promoted aldol-like cyclization of **[IX]** to α-phenylsulfonyl cyclohexenone **[X]** and (4) conversion of the corresponding α-phenylsulfonyl cyclohexenone **[X]** to α-hydroxymethyl cyclohexenone **[XI]** through the Michael-type addition of tributylstannyl-lithium followed by trapping with formaldehyde and desulfonylation. This transformation is ideally suited to the synthesis of carbocycle-containing natural products and carbasugars, since the core skeleton arises after appropriate replacement of the phenylsulfonyl group, and, in practice, this method was applied to total synthesis of progesterone receptor ligands, (–)-PF1092A, B, and C [207], and (–)-glyoxalase inhibitor and its precursor (–)-KD16-U1 (❯ *Fig. 12*) [208].
As shown in ❯ *Scheme 44*, this method was also applied to total synthesis of (+)-valienamine (**320**) and (+)-validamine (**324**) [209]. Silyl enol ether **314** derived from D-xylose in nine steps

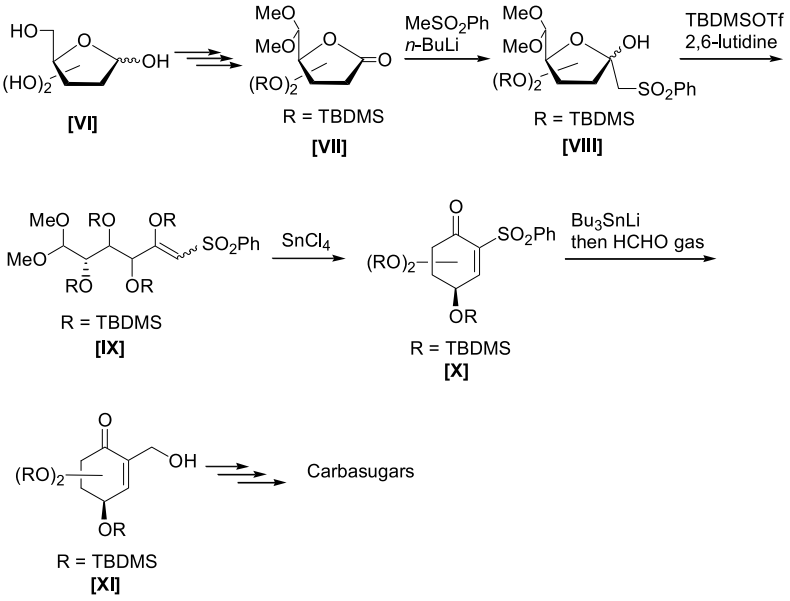
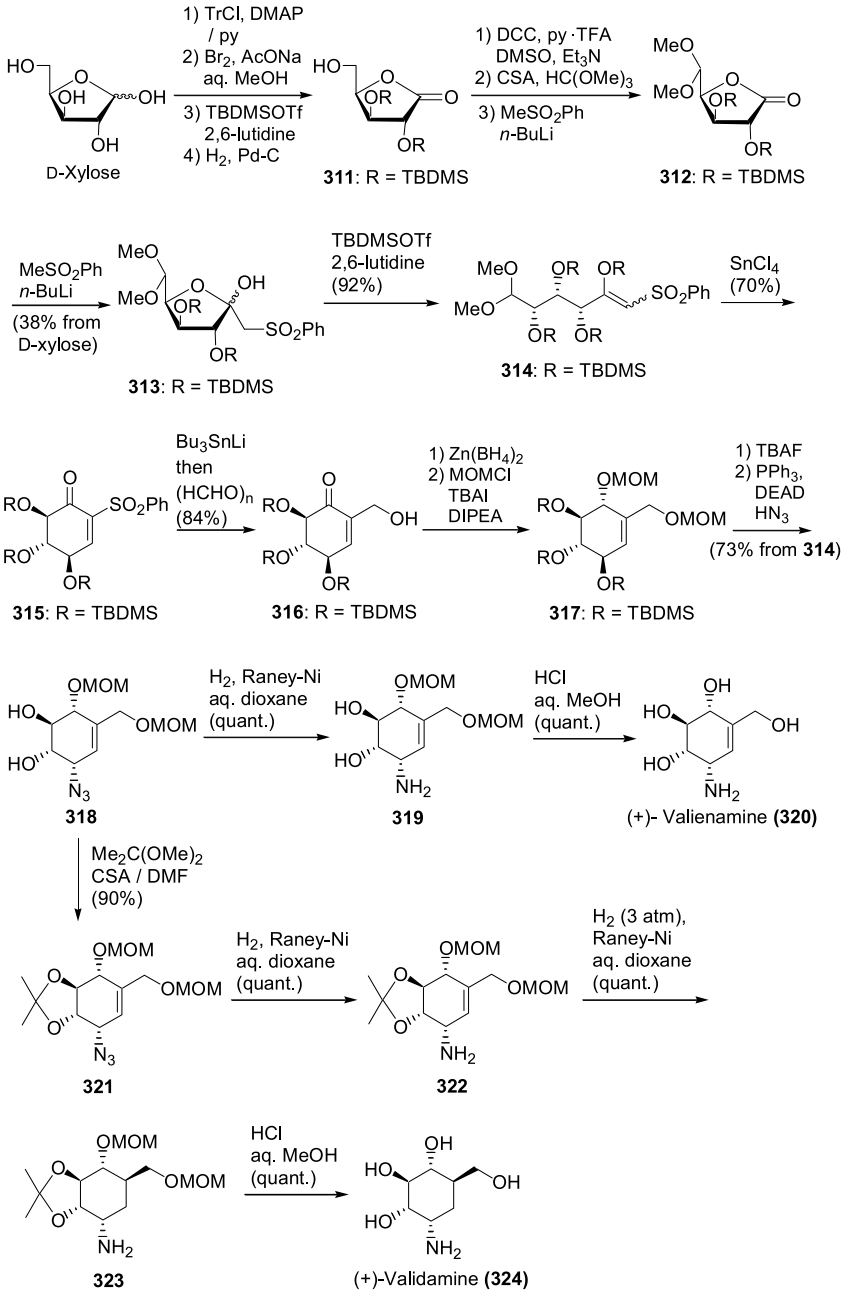


Figure 12
Key steps of the synthesis of the 6-membered carbasugar using $SnCl_4$-promoted intramolecular aldol condensation

was converted to α-hydrixymethyl cyclohexenone **316** via $SnCl_4$-promoted intramolecular aldol condensation and Michael-aldol reaction with tributylstannyl-lithium and formaldehyde in good yield. Cyclohexenone **316** was converted to (+)-valienamine (**320**) and (+)-validamine (**324**) via the common intermediate, azide **318** in good yield (*Scheme 44*).
It can be expected that this transformation would be one of the methods useful for the synthesis of 6-membered carbasugars.

4.2 SmI_2-Mediated Carbasugar Formation

First, the intramolecular cyclizations between carbonyl compounds and olefins will be described. This field can be classified into three types as illustrated in *Fig. 13*. One is a type of cyclization between carbonyl compounds and α, β-unsaturated esters (type [A]), the next is a type of cyclization between carbonyl compounds and simple olefins (type [B]), and the last is a type of cyclization between carbonyl compounds and oximes (type [C]) (*Fig. 13*).
The general feature of these reactions is as follows: When treated with SmI_2, a reductive cyclization between the carbonyl compound and the β-carbon of the olefin of **[XIII]** derived from sugars **[XII]** gives the desired polyhydroxylated cyclopentane **[XIV]**, in which a new C1/C5 bond (carbohydrate numbering) is formed between sp^2 centers. In the overall sequence, the sp^3 alcohol stereocenter at C5 is destroyed when oxidation occurs to form the carbonyl moieties and subsequently reinstated, upon treatment with SmI_2, to form a new hydroxy-


1) TrCl, DMAP / py
2) Br2, AcONa aq. MeOH
3) TBDMSOTf 2,6-lutidine
4) H2, Pd-C
D-Xylose
311: R = TBDMS
1) DCC, py·TFA DMSO, Et3N
2) CSA, HC(OMe)3
3) MeSO2Ph n-BuLi
312: R = TBDMS
MeSO2Ph n-BuLi (38% from D-xylose)
313: R = TBDMS
TBDMSOTf 2,6-lutidine (92%)
314: R = TBDMS
SnCl4 (70%)
Bu3SnLi then (HCHO)n (84%)
315: R = TBDMS
316: R = TBDMS
1) Zn(BH4)2 2) MOMCl TBAI DIPEA
317: R = TBDMS
1) TBAF 2) PPh3, DEAD HN3
(73% from 314)
318
H2, Raney-Ni aq. dioxane (quant.)
319
HCl aq. MeOH (quant.)
(+)- Valienamine (320)
Me2C(OMe)2 CSA / DMF (90%)
321
H2, Raney-Ni aq. dioxane (quant.)
322
H2 (3 atm), Raney-Ni aq. dioxane (quant.)
323
HCl aq. MeOH (quant.)
(+)-Validamine (324)

Scheme 44

bearing stereocenter. In contrast, the existing 5-hexenyl free radical methods for the cyclization of carbohydrates do not allow for the incorporation of this additional sp^3 alcohol center which provides further functionality for subsequent synthetic manipulations. This is a general difference between the traditional radical cyclization and the SmI_2-mediated radical cyclization [214].

4.2.1 Cyclization between Carbonyl Compounds and α, β-Unsaturated Esters

Enholm et al. reported the SmI_2-mediated construction of carbocycles from carbohydrate templates in 1989 [214]. This is the example of the type [A] reaction in *Fig. 13*. This strategy, which permits the conversion of pyranose sugars into highly oxygenated cyclopetanes, is illustrated in *Fig. 14* (Eq. (1)). Treatment of carbohydrate templates **[XVI]**, which are derived from sugars **[XV]** via Wittig reaction with stabilized ylides and subsequent PDC oxidation, with the one-electron reducing reagent SmI_2 prompts the intramolecular coupling of two sp^2-hybridized carbon centers and finally constructs a highly oxygenated cyclopentane

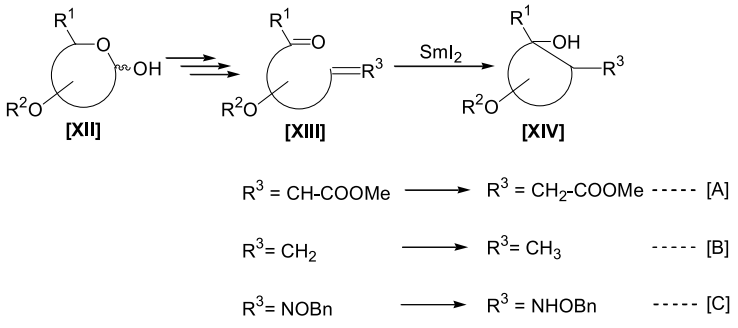


Figure 13
SmI_2-mediated carbasugar formation

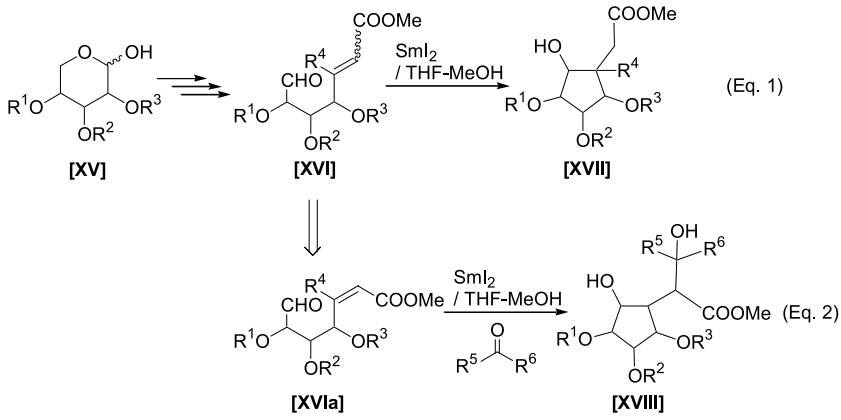
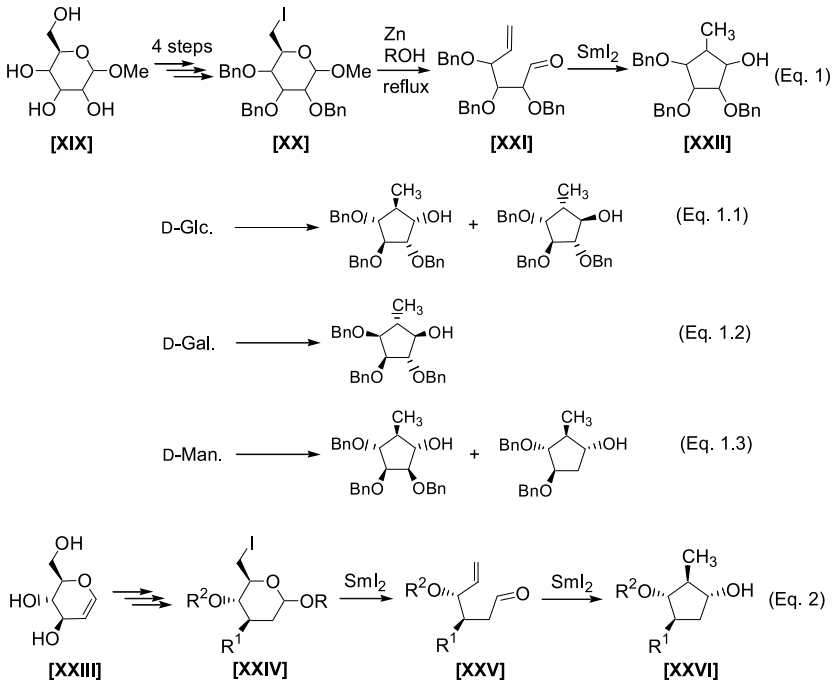


Figure 14
Cyclization between carbonyl compounds and α, β-unsaturated esters

ring **[XVII]**. Interestingly, the geometry of the olefin unit of the substrates influences the stereochemistry at C1 and C5 (carbohydrate numbering) of the products, and there is the tendency that *Z*-olefins give the *anti*-alcohols and *E*-olefins provide the *syn*-alcohols exclusively. Also, this method was applied to the synthesis of the natural product, anguidine [215]. Moreover, Enholm also reported the sequential SmI_2-promoted one- and two-electron reactions as an application of this concept (● *Fig. 14*, Eq. (2)) [216]. Treatment of the compounds **[XVIa]** and carbonyl compounds with SmI_2 afforded a highly modified carbocycle. In the domino key step, SmI_2 promoted a sequential one-electron (radical) cyclization of an aldehyde and an alkene, followed by a two-electron intermolecular carbonyl addition reaction to afford compounds **[XVIII]** (● *Fig. 14*).

4.2.2 Cyclization between Carbonyl Compounds and Simple Olefins

As the example of the type [B] reaction, the reports by Holzapfel et al. and Chiara et al. who independently reported this type of reaction are described. Initially, Holzapfel et al. reported the SmI_2-mediated radical cyclization of the ring-opened hex-5-enals **[XXI]** derived from pyranoses **[XIX]** via the step of the reductive elimination of C6 iodopyranose derivatives **[XX]** with active Zn [217] (● *Fig. 15*, Eq. (1)). Afterwards, they also reported the one-pot SmI_2-



● Figure 15
Cyclization between carbonyl compounds and simple olefins (1)

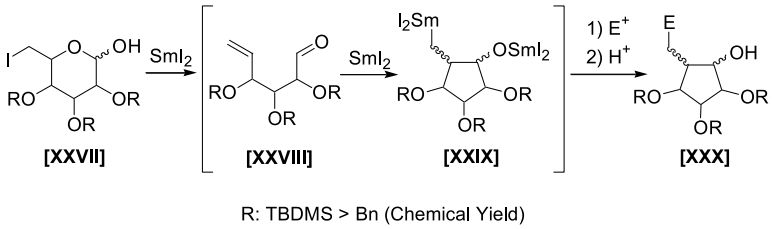


Figure 16
Cyclization between carbonyl compounds and simple olefins (2)

mediated radical cyclization of 6-deoxy-6-iodohexopyranose derivatives **[XXIV]** which were derived from glycals **[XXIII]** [218] (*Fig. 15*, Eq. (2)).
In each case, the cyclization itself proceeds typically with above 65% efficiency, and the stereochemistry of C1 and C5 (carbohydrate numbering) of the products shows an *anti*-configuration. However, as an exception, in case of the *manno*-type derivative of **[XXI]**, the compound arises via initial SmI_2-mediated elimination of the α-substituent (benzyloxy group) to the carbonyl (*Fig. 15*, Eq. (1.3)).
On the other hand, Chiara et al. also reported outcomes similar to Holzapfel's results [219]. They attempted the one-pot, SmI_2-mediated cyclization of fully functionalized 6-deoxy-6 iodohexopyranose derivatives and investigated the influence of the stereochemistry of the substrates as well as the effects of the protecting groups in this type of reaction (*Fig. 16*).
Basically, their results are consistent with those of Holzapfel et al., especially with respect to the stereochemical arguments given above and the chemical yield of the products. According to these studies, the TBDMS group is suitable for the reaction from the viewpoints of the ease of separation of the products and the avoidance of the β-elimination seen in the reports of Holzapfel et al. Also, they reported that 6-deoxy-hexopyranose derivatives were obtained as the by-products in the range of yields of 10–50%.

4.2.3 Cyclization between Carbonyl Compounds and Oximes

Chiara et al. also reported the type [C] reaction, the cyclization between carbonyl compounds and oximes [220,221]. As illustrated in *Fig. 17*, three types of substrate were synthesized from hexopyranoses and pentopyranoses, and the terminal *O*-benzylformaldoxime-ketones and the terminal *O*-benzylformaldoxime-α, β-unsaturated esters were synthesized from a series of hexopyranoses, and meanwhile the terminal *O*-benzylformaldoxime-aldehydes were derived from a series of pentopyranoses (*Fig. 17*).
The goal of this attempt is the stereochemistry of the hydroxy group and amino group generated after cyclization. In cases of cyclization of the terminal *O*-benzylformaldoxime-ketones and the terminal *O*-benzylformaldoxime-α, β-unsaturated esters, the stereochemistry is mainly the *anti*-configuration as is also the case for the cyclization described above.
However, in contrast to the latter substrates, in the case of cyclization of the terminal *O*-benzylformaldoxime-aldehydes, the realized stereochemistry depends on the configuration or the protecting groups.

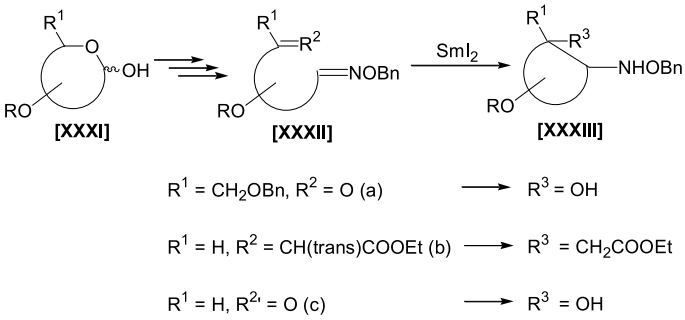


Figure 17
Cyclization between carbonyl compounds and oximes

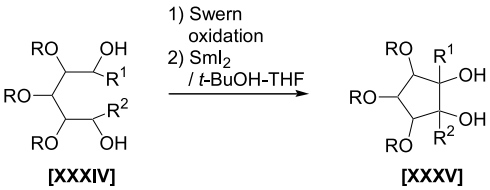


Figure 18
SmI_2-mediated Pinacol coupling

4.3 SmI_2-Mediated Pinacol Coupling

SmI_2-mediated pinacol coupling is also a useful synthetic method to form carbasugars possessing a *cis*-diol moiety from sugars. In fact, as mentioned above, Nakata et al. utilized this method in the synthesis of cyclophelitols, and it is one of the good examples to verify the synthetic utility of SmI_2-mediated pinacol coupling [20]. This coupling reaction has also been utilized in the field of the transformation of sugars to highly functionalized 5-membered carbocycles (*Fig. 18*).

Initially, application of this method to the formation of 5-membered carbocycles was reported by Sinäy et al. [222]. In their case, a dialdehyde was used as the substrate, and cyclization occurred with a comparable overall yield of carbocyclic *cis*-diols but with much lower diastereoselectivity. Later, Iadonisi et al. reported the results of the application using keto-aldehydes as the substrates [223]. Compared to Sinay's results, cyclization of the keto-aldehydes occurred with a relatively high stereoselectivity, and it should be suggested that, in particular, electrostatic interactions play an important role in the stereocontrol of this reaction.

4.4 SmI_2- or Zirconium-Mediated Ring Contraction of Hexapyranoside Derivatives to 5-Membered Carbosugar

As mentioned in the above sections on the syntheses of natural glycosidase inhibitors, the ring contraction reaction of a 6-membered ring system to a 5-membered ring system is definitely

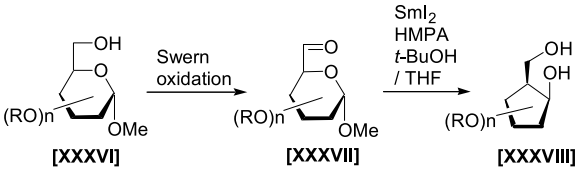


Figure 19
SmI_2-mediated ring contraction of hexapyranoside derivatives to 5-membered carbasugars

a significant aspect in synthetic chemistry, and the further development of this type of method should pave the way to the syntheses of the complex natural products and artificial chemicals including clinical drugs as well as a series of glycosidase inhibitors.

Here, the SmI_2- or zirconium-mediated ring contraction of hexapyranoside derivatives to 5-membered ring carbosugars will be described.

Sinäy et al. presented the SmI_2-mediated ring contraction of methyl 6-unprotected-hexopyranoside derivatives to fully functionalized 5-membered carbosugars via the intermediate aldehyde shown in *Fig. 19* [224]. The aldehyde **[XXXVII]**, which was derived from the methyl 6-unprotected-hexopyranoside derivative **[XXXVI]** via Swern oxidation, was treated with SmI_2 (5 equiv.) in THF in the presence of HMPA and *t*-BuOH to afford the *cis*-cyclopentane **[XXXVIII]** exclusively. Interestingly, the *cis*-configuration between C1 and C5 (carbohydrate numbering), which is the new stereocenter generated in this reaction, does not depend on the substrates, and each compound can be obtained in moderate yield (*Fig. 19*).

Sinäy et al. anticipated the mechanism of the reaction illustrated in *Fig. 20*. The first equivalent of SmI_2 reduces the aldehyde **326**, which was derived from **325**, to the samariumketyl **327**. The second equivalent of SmI_2 reduces the corresponding ketyl **327** to disamarium species **328**, which then undergoes ring opening followed by methoxide elimination to give the key intermediate **330**. Subsequently, this reaction uniquely generates a system which is ideally suited for a following aldol cyclization reaction involving intramolecular nucleophilic attack of the samarium enolate onto aldehyde through a 5-enol *exo-exo-trig* process. With respect to the stereoselectivity of this reaction, they concluded that it was expected to ensue from a samarium-linked, medium-sized chelate, from which the carbon–carbon bond formation would take place as a ring contraction (*Fig. 20*).

Recently, Aurrecoechea et al. reported the synthesis of 2-vinylcyclopentanols by SmI_2/Pd(0)-promoted carbohydrate ring-contraction [225]. On the other hand, Ito and Taguchi et al. independently presented the $Cp_2Zr/BF_3{\cdot}Et_2O$-mediated carbohydrate ring-contraction [226,227]. The contrast regarding the stereochemistry of products in each methodology is interesting as illustrated in *Fig. 21*.

First, the SmI_2/Pd(0)-promoted carbohydrate ring-contraction will be described. As illustrated in *Fig. 21*, substrates **[XXXIX]**, which were derived from methyl 6-unprotected hexapyranoside derivatives via Swern oxidation followed by Wittig olefination, were treated with SmI_2 and a catalytic amount of $Pd(PPh_3)_4$ to afford the 2-vinylcyclopentanols **[XXXX]** possessing the *trans*-stereochemistry between the vinyl and the newly generated hydroxyl group. The authors speculated on the mechanism of this reaction and the cause of the predominant *trans*-stereoselectivity of the resulting compounds as illustrated in *Fig. 22*. The for-

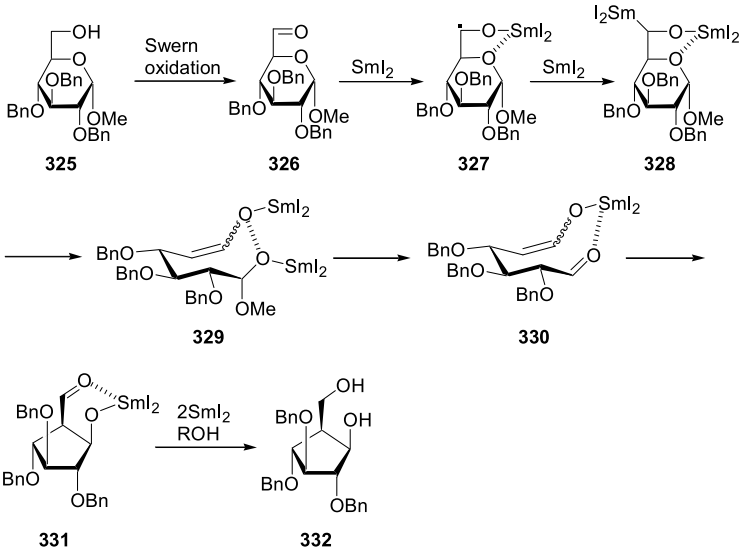


Figure 20
Proposed mechanism of SmI_2-mediated ring contraction

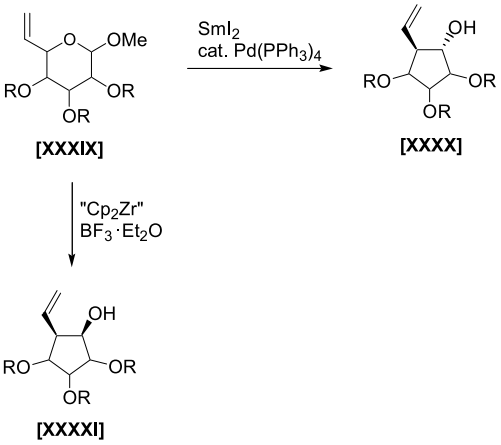


Figure 21
SmI_2-Pd(0)- or Zr-mediated ring contraction of hexapyranoside derivatives to 5-membered carbasugars

mation of cyclopentane products can be rationalized by a Pd(0)-promoted ring-opening of **[XXXIX]** leading to an intermediate **[XXXXII]** that contains both a *p*-allylpalladium complex and an aldehyde moiety. Reduction of the palladium complex by SmI_2 and carbonyl addition of the resulting allylsamarium species **[XXXXIII]** would then lead to the observed products **[XXXX]**. With respect to the stereochemistry of the products observed in these cyclizations, the predominant *trans*-stereoselectivity is surprising when compared to the relat-

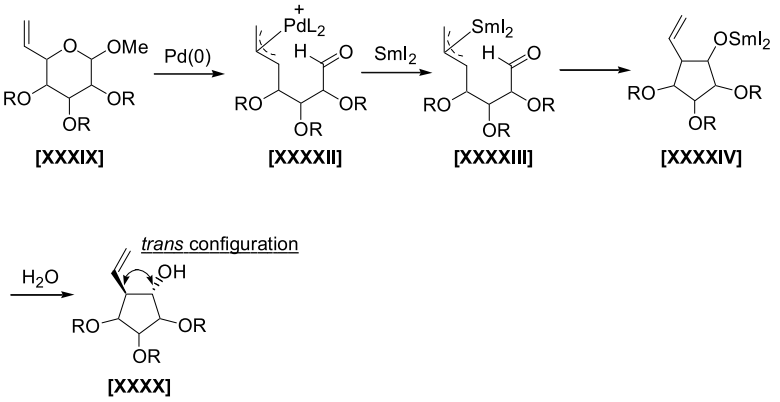


Figure 22
Proposed mechanism of SmI_2-Pd(0)-mediated ring contraction

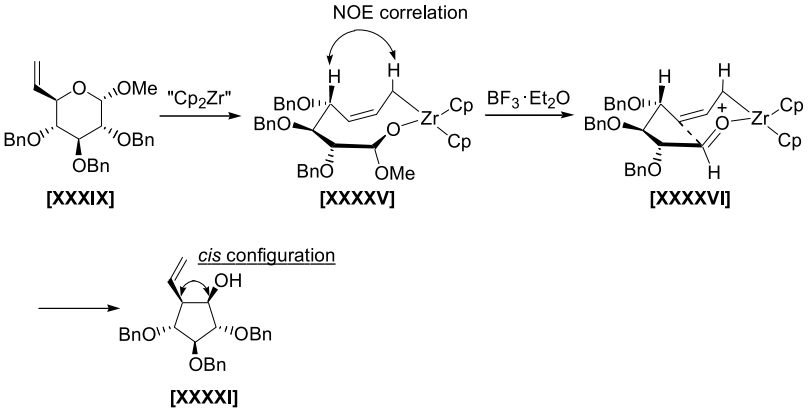


Figure 23
Proposed mechanism of Zr-mediated ring contraction

ed SmI_2/Pd(0)-promoted intramolecular cyclization of a ketone containing a pendent vinyloxirane moiety. This latter case is also assumed to proceed through allylsamarium intermediates, but affords preferentially a *cis*-product, presumably through a chelated cyclic transition structure. If a similar mechanism is indeed operating, the different stereochemical outcomes described above could be due to a combination of factors derived from differences in reaction conditions and substrate structures. Interestingly, in addition to the preferred *trans*-relationship observed between the vinyl and hydroxyl groups, the glucose- and mannose-derived substrates also displayed preference for a *trans*-relationship between the hydroxyl group and the adjacent benzyloxy group. Whereas for the more selective galacto derivatives, the opposite *cis*-relationship was found (*Fig. 22*).

On the other hand, the zirconium-mediated ring contraction of hexapyranoside derivatives, which was reported by Ito and Taguchi et al. afforded the 2-vinylcyclopentanols, and the

resulting 2-vinylcyclopentanols possess *cis*-stereochemistry between the vinyl group and the newly generated hydroxyl group. They explained that the derived stereochemistry would be induced through the intermediates illustrated in Fig. 23. Treatment of **[XXXIX]** with Cp_2Zr in THF provided intermediate **[XXXXV]**, which was confirmed by NMR analysis. The proof of **[XXXXV]** as an intermediate in the reaction was further confirmed by the conversion of **[XXXXV]** to **[XXXXI]** upon addition of $BF_3{\cdot}Et_2O$. It is probable that $BF_3{\cdot}Et_2O$ accelerates the elimination of the methoxy group through coordination to the methoxy oxygen of **[XXXXV]** to form the sterically favored oxocarbenium ion **[XXXXVI]**. Finally, this intermediate was cyclized to afford the resulting *cis*-2-vinylcyclopentaol **[XXXXI]** (Fig. 23).

5 Conclusion

As mentioned above, a variety of natural glycosidase inhibitors have been isolated, and they show interesting, strong, and enzyme-specific inhibitory activities. Recently, it was recognized that these compounds are important molecules to elaborate the pivotal roles of glycoconjugates in living systems, and have potential use as clinical drugs, and, in the field of biochemistry, the synthesis and supply of the glycosidase inhibitors and their analogues occupies a quite significant position. In this chapter, a variety of syntheses of glycosidase inhibitors were described, and also applications and developments of the new synthetic methodologies for transformation of sugars to carbasugars, ranging from the typical Ferrier reaction to the recent SmI_2/Pd(0)-mediated ring-contraction, could be introduced through the synthesis of a variety of glycosidase inhibitors. I hope that a variety of studies in glycoscience using carbasugars will pave the way to further elucidation of the functions of glycosidases and glycoconjugates in living systems.

Acknowledgements

Finally, I thank Dr. Masao Shiozaki, currently at RIKEN, for helpful discussions and suggestions during the preparation of this chapter.

References

1. Springer TA (1990) Nature 346:425
2. Stoolman LM (1989) Cell 56:907
3. Sharon N, Lis H (1989) Science 246:227
4. Clark AE, Wilson IA (1988) Carbohydrate-protein interactions. Springer, Berlin Heidelberg New York
5. Elbein AD (1987) Ann Rev Biochem 56:497
6. Fellows LE (1989) New Sci 123:45
7. Fellows LE (1989) Chem Br 23:847
8. Atsumi S, Umezawa K, Iinuma H, Naganawa H, Nakamura H, Iitaka Y, Takeuchi T (1990) J Antibiot 43:49
9. Atsumi S, Iinuma H, Nosaka C, Umezawa K (1990) J Antibiot 43:1579
10. Withers SG, Umezawa K (1991) Biochem Biophys Res Commun 177:532
11. Tatsuta K, Niwata Y, Umezawa K, Toshima K, Nakata M (1990) Tetrahedron Lett 31:1171
12. Semenia D, Philippe M, Delaumeny J-M, Sepulchre A-M, Gero SD (1983) Synthesis 710
13. Kozikowski AP, Stein PD (1984) J Org Chem 49:2301
14. Bernet B, Vasella A (1979) Helv Chim Acta 62:2411

15. Toshima K, Tatsuta K, Kinoshita M (1988) Bull Chem Soc Jpn 61:2369
16. Tatsuta K, Niwata Y, Umezawa K, Toshima K, Nakata M (1990) Carbohydr Res 222:189
17. Tatsuta K, Niwata Y, Umezawa K, Toshima K, Nakata M (1991) J Antibiot. 44:912
18. Tatsuta K, Niwata Y, Umezawa K, Toshima K, Nakata M (1993) J Antibiot. 46:1919
19. Tatsuta K (1996) Pure and Appl Chem. 68:1341
20. Ohba K, Suzaki K, Nakata M (1996) Carbohydr Lett 2:175
21. Chiara JL, Cabri W, Hanessian S (1991) Tetrahedron Lett 62:1125
22. Chiara JL, Marti-Lomas M (1994) Tetrahedron Lett 35:2969
23. Guidot JP, Le Gall T, Mioskowski C (1994) Tetrahedron Lett 35:6671
24. Chiara JL, Valle N (1995) Tetrahedron Asymmetry 6:1895
25. Sawada T, Shirai R, Iwasaki S (1996) Tetrahedron Lett 37:885
26. McDevitt RE, Fraser-Reid B (1994) J Org Chem 59:3250
27. Thiem J, Karl H (1978) Tetrahedron Lett 4999
28. Thiem J, Karl H, Schwenter J (1978) Synthesis:996
29. Alonso RA, Bugey CS, Rao BV, Vite GD, Vollerthum R, Zottola MA, Fraser-Reid B (1993) J Am Chem Soc 115:6666
30. Lopez JC, Gomez AM, Fraser-Reid B (1993) J Chem Soc Chem Commun 762
31. Tsang R, Frase-Reid B (1992) J Org Chem 57:1065
32. Alonso RA, Vite GD, McDevitt RE, Fraser-Reid B (1992) J Org Chem 57:573
33. Dess DB, Martin JC (1983) J Org Chem 43:4155
34. Saksena AK, Mangiaracina P (1983) Tetrahedron Lett 24:273
35. Turnbull MD, Matter G, Ledgerwood DE (1984) Tetrahedron Lett 25:5449
36. Ferrier RJ (1979) J Chem Soc Perkin Trans 1 1455
37. Ferrier RJ, Middleton S (1993) Chem Rev 93:2779 and literature cited therein
38. Sato K, Bokura M, Moriyama H, Igarashi T (1994) Chem Lett 37
39. Sato K, Suzuki K, Ueda M, Katayama M, Kajiwara M (1991) Chem Lett 1469
40. Jung ME, Choe WT (1995) J Org Chem 60:3280
41. Buchanan JG, Schwarz JCP (1962) J Chem Soc 4770
42. Nashed MA, Anderson L (1976) Tetrahedron Lett 3503
43. Nashed MA (1978) Carbohydr Res 60:200
44. Schwab P, France MB, Ziller JW, Grubbs RH (1995) Angew Chem Int Ed Engl 34:2039
45. Ziegler FE, Wang y (1998) J Org Chem 63:426
46. Hanessian S, Gai Y, Wang W (1996) Tetrahedron Lett 37:7473
47. Suarez E, Freire R, Marrero JJ, Rodriguez MS (1986) Tetrahedron Lett 27:383
48. Suarez E, Armas P, Francisco CG (1992) Angew Chem Int Ed Engl 31:772
49. Oh J (1997) Tetrahedron Lett 38:3249
50. Kireev AS, Breithaupt AT, Collins W, Nadein ON, Kornienko A (2005) J Org Chem 70:742
51. Nadien ON, Kornienko A (2004) Org Lett 6:831
52. Trost BM, Hember EJ (1999) Tetrahedron Lett 40:219
53. Trost BM, Patterson DE, Hember EJ (2001) Chem Eur J 7:3768
54. Hansen FG, Bundgaard E, Madsen R (2005) J Org Chem 70:10139
55. Loh T-P, Cao G-Q, Pei J (1998) Tetrahedron Lett 39:1453
56. Diana S-CH, Sim K-Y, Loh T-P (1996) Synlett 263
57. Wang R, Lim C-M, Tan C-H, Lim B-K, Sim K-Y, Loh T-P (1995) Tetrahedron: Asymmetry 6:1825
58. Scholl M, Ding S, Lee CW, Grubbs RH (1999) Org Lett 1:953
59. Schlessinger RH, Bergstorm CP (1995) J Org Chem 60:16
60. Bryson TA, Dolak TM (1977) Org Synth 57:62
61. Haubenstock H (1990) Tetrahedron 46:6633
62. Brunne J, Hoffmann N, Scharf H-D (1994) Tetrahedron 50:6819
63. Trost BM, Van Vranken DL (1996) Chem Rev 96:395
64. Still WC, Mitra MG (1978) J Am Chem Soc 100:1972
65. Lui K-J, Kendal AP (1987) Am J Public Health 77(6):712
66. Saul H (1995) New Scientist:26
67. Couch RB, Six HR (1986) In: Mills J, Corey L (eds) Antiviral Chemotherapy: New Direction for Clinical Application and Research. Elsevier Science Publishing, Oxford, p. 50
68. Hayden FG, Belshe RB, Clover RD, Hay AJ, Oakes MG, Soo W (1989) N Engl J Med 321:1696
69. Sauter NK, Hansen JE, Glick GD, Brown JH, Crowther RL, Park SJ, Skehel JJ, Wiley DC (1992) Biochemistry 31:9609

70. Mammen M, Dahmann G, Whitesides GJ (1995) J Med Chem 38:4179
71. Colman PM (1989) In: Krug RM (ed) The Influenza Viruses: Influenza Virus Neuraminidase, Enzyme and Antigen. Plenum Press, New York, p. 175
72. Colman PM (1994) Protein Sci 3:1687
73. Hay AJ, Wolstenholme AJ, Skehel JJ, Smith MH (1985) EMBO J 4:3021
74. Hastings JC, Selnick H, Wolanski B, Tomassini JE (1996) Antimicrob Agents Chemother 40:1304
75. Palse P, Jobita K, Ueda M, Compans RW (1974) Virology 61:397
76. Liu C, Eichelberger MC, Compans RW, Air GM (1995) J Virol 69:1099
77. Klenk HD, Rott R (1988) Adv Virus Res 34:247
78. Meinal P, Bodo G, Palese P, Schulman J, Tuppy H (1975) Virology 58:457
79. Holzer CT, von Itzstein M, Jin B, Pegg MS, Stewart WP, Wu W-Y (1993) Glycoconjigate J 10:40
80. Chong AKJ, Pegg MS, von Itzstein M (1991) Biochem Int 24:165
81. Taylor NR, von Itzstein M (1994) J Med Chem 37:616
82. Janakiraman MN, White CL, Laver WG, Air GM, Luo M (1994) Biochemistry 33:8172
83. von Itzstein M, Wu W-Y, Jin B (1994) Carbohydr Res 259:301
84. von Itzstein M, Wu W-Y, Kok GB, Pegg MS, Dayson JC, Jin B, Phan TV, Smythe ML, White HF, Oliver SW, Colman PM, Varghese JN, Ryan DM, Wood JM, Bethell RC, Hotham VJ, Cameron JM, Penn CR (1993) Nature (London) 368:418
85. Wood JM, Bethell RC, Coates JA, Healy N, Hiscox SA, Pearson BA, Ryan DM, Ticehurst J, Tilling J, Walcott SM, Penn CR (1993) Antimicrob Agents Chemother 37:1473
86. Ryan DM, Ticehurst J, Demsey MH, Penn CR (1994) Antimicrob Agents Chmother 38: 2270
87. Hayden FG, Treanor JJ, Betts RF, Lobo M, Esinhart J, Hussey EK (1996) J Am Med Assoc 275:295
88. Hayden FG, Treanor JJ, Esinhart J, Eason CU, Hussey EK (1995) Antiviral Res 26(3) A 300 (Abst. 140)
89. Fromtling RA, Castaner J (1996) Drugs Future 21(4):375
90. Choung UK, Willard L, Matthew AW, Hongtao L, Lijun Z, Swaminathan S, Norbert B, Ming SC, Dirk BM, Chun YT, W Graeme L, Raymond CS (1997) J Am Chem Soc 119:681
91. Air GM, Laver WG, Luo M, Stray SJ, Legrone G, Webster RG (1990) Virology 177:578
92. Blok J, Air GM, Laver WG, Ward CW, Lilley GG, Wood EF, Roxburgh CM, Inglis AS (1982) Virology 119:578
93. Colman PM (1984) Pept Protein Rev 4:215
94. Varghese JH, Colman PM (1991) J Mol Biol 221:473
95. Bossart-Whitaker P, Carson M, Babu YS, Smith CD, Laver WG, Air GM (1993) J Mol Biol 232:1069
96. Burmeister WP, Ruigrok RW, Cusak S (1992) EMBO J 11:49
97. Vargese JN, McKimm-Bresckhin L, Caldwell JB, Kortt A, Colman PM (1992) Proteins: Struct Funct Genet 14:327
98. Pauling L (1946) Chem Eng News 24:1375
99. Bartlett PA (1985) Stud Org Chem 20:439
100. Tedrzejas MJ, Singh S, Brouillette WJ, Laver GW, Air GM, Luo M (1995) Biochemistry 34:3144
101. Williams M, Bischofberger N, Swaminathan S, Kim CU (1995) Bioorg Med Chem Lett 5(9):2551
102. Chandler M, Conroy R, Cooper AWJ, Lamont RB, Scicinski JJ, Smart JE, Storer R, Weir NG, Wilson RD, Wyatt PG (1995) J Chem Soc Perkin Trans 1:1189
103. McGowan DA, Berchtold GA (1981) J Org Chem 46:2381
104. Ulibarri G, Nadler W, Skrydtrup T, Audrian H, Chiaroni A, Riche C, Grierson DS (1995) J Org Chem 60:2753
105. Smith PW, Starkey ID, Howes PD, Sollis SL, keeling SP, Cherry PC, von Itzstein M, Wu WY, Jin B (1996) Eur J Med Chem 31:143
106. Rohloff JC, Kent KM, Postich MJ, Becker MW, Chapman HH, Kelly DE, Lew W, Louie MS, McGee LR, Prisbe EJ, Schultze LM, Yu RH, Zhang L (1998) J Org Chem 63:4545
107. Shing TKM, Tang Y (1990) Tetrahedron 46:6575
108. Trost BM, Kelnan E (1978) J Am Chem Soc 100:7779
109. Genet JP, Balabane M, Backvall JE, Nystrom JE (1983) Tetrahedron Lett 24:2745
110. Hunter R, Bartels B (1993) J Org Chem 58:6756
111. Yeung YY, Hong S, Corey EJ (2006) J Am Chem Soc 128:6310
112. Ryu DH, Corey EJ (2003) J Am Chem Soc 125:6388

113. Knapp S, Gibson FS (1998) Organic Syntheses. Wiley, New York
114. Knapp S, Levorse AT (1988) J Org Chem 53:4006
115. Fukuta Y, Mita T, Fukuda N, Kanai M, Shibasaki M (2006) J Am Chem Soc 128:6312
116. Mita T, Fujimori I, Wada R, Wen J, Kanai M, Shibasaki M (2005) J Am Chem Soc 127:11252
117. Tosaki S-Y, Tsuji R, Ohshima T, Shibasaki M (2005) J Am Chem Soc 127:2147
118. Kanai M, Kato N, Ichikawa E, Shibasaki M (2005) Synlett 1491
119. Kramer KJ, Dziadik-Turner C, Koga D (1995) In: Kerkut GA, Gilbert LI (eds) Comprehensive Insect Physiology Biochemistry and Pharmacology,vol 3. Pergamon Press, New York
120. Anderson SD (1979) Annu Rev Entomol 24:29
121. Barett-Bee K, Hamilton MJ (1984) Gen Microbiol 130:1857
122. Bartmicki-Garcia S (1968) Annu Rev Microbial 22:87
123. Sakuda S, Isogai A, Matsumoto S, Suzuki A (1987) J Antibiot 40:296
124. Sakuda S, Isogai A, Matsumoto S, Suzuki A, Koseki K (1986) Tetrahedron Lett 27:2475
125. Sakuda S, Isogai A, Makita T, Matsumoto S, Koseki K, Kodama H, Suzuki A (1987) Agric Biol Chem 51:3251
126. Sakuda S, Isogai A, Matsumoto S, Suzuki A, Koseki K, Kodama H, Yamada Y (1988) Agric Biol Chem 52:1615
127. Corey EJ, Weinshenker NM, Schaef TK, Huber W (1969) J Am Chem Soc 91:5675
128. Trost BM, Van Vranken DL (1990) J Am Chem Soc 112:1261
129. Trost BM, Van Vranken DL (1993) J Am Chem Soc 115:444
130. Trost BM, Sudhakar AR (1988) J Am Chem Soc 110:7933
131. Trost BM, Sudhakar AR (1987) J Am Chem Soc 109:3792
132. Hayashi T, Yamamoto A, Ito Y (1987) Tetrahedron Lett 28:4837
133. Consiglo G, Waymouth R (1989) Chem Rev 89:257
134. Trost BM, Murphy DJ (1985) Organomet 4:1143
135. Auburn PR, Mackenzie PB, Bosnich B (1985) J Am Chem Soc 107:2033
136. Hayashi T (1988) Pure Appl Chem 60:7
137. Fiaud JC, Legros J-Y (1990) J Org Chem 55:4840
138. Yamaguchi M, Shima T, Yamagishi T, Hida M (1991) Tetrahedron Asymmetry 2:663
139. Muller D, Umbricht G, Weber B, Pfaltz A (1991) Helv Chim Acta 74:232
140. Okada Y, Minami I, Yamamoto T, Ichikawa J (1992) Chem Lett 547
141. Trost BM, Van Vranken DL (1992) Angew Chem Int Ed Engl 31:228
142. Trost BM, Chupak LS, Lubbers T (1998) J Am Chem Soc 120:1732
143. Griffith DA, Danishefsky SJ (1991) J Am Chem Soc 113:5863
144. Griffith DA, Danishefsky SJ (1996) J Am Chem Soc 118:9526
145. Wang Y, Chen C, Girdaukas G, Shi CJ (1984) J Am Chem Soc 106:3695
146. Lauman K, Schneider M (1984) Tetrahedron Lett 25:5875
147. Deardorff DP, Matthews AJ, McMeekim DS, Craney CL (1986) Tetrahedron Lett 27: 1255
148. Thiel F, Ballschuh S, Schick A, Haupt M, Hafner B, Schwarz S (1988) Synthesis 540
149. Jommi G, Orsini F, Sisti M, Verotta L (1988) Gazz Chim Ital 118:863
150. Babiak KA, Ng JS, Dygos JH, Weyker CL, Yang YE, Wong CH (1990) J Org Chem 55:3377
151. Danishefsky SJ, Cabal MP, Chow K (1989) J Am Chem Soc 111:3456
152. Ganem B, Goering BK (1994) Tetrahedron Lett 35:6997
153. Nakata M, Akagawa S, Kitamura S, Tatsuta K (1991) Tetrahedron Lett 32:5363
154. Kitahara T, Suzuki N, Koseki K, Mori K (1993) Biosci Biotech Biochem 57:1906
155. Kozikowski AP, Adamczyk M (1982) Tetrahedron Lett 23:3123
156. Takahashi S, Terayama H, Kuzuhara H (1991) Tetrahedron Lett 32:5123
157. Hanessian S, Lavallee P (1976) Methods Carbohydr Chem 7:49
158. Barton DH, Augy-Dorey S, Camara J, Dalko P, Delaumeny JM, Gero SD, Quiclet-Sire B, Stuts P (1990) Tetrahedron 46:215
159. Kobayashi S, Kamiyama K, Ohno M (1990) J Org Chem 55:1169
160. Baer HH, Astles DJ, Chin H-C, Siemsen L (1985) Can J Chem 63:432
161. Simpkins NS, Stokes S, Whittle AJ (1992) Tetrahedron Lett 33:793
162. Simpkins NS, Stokes S, Whittle AJ (1992) J Chem Soc Prekin Trans 1 2471
163. Gries B (1986) Radicals in Organic Synthesis: Formation of Carbon-Carbon Bonds. Pergamon Press, Oxford

164. Rajanbabu TV, Fukunaga T, Reddy GS (1989) J Am Chem Soc 111:1759 and literature cited therein
165. Tsang R, Fraser-Reid B (1986) J Am Chem Soc 108:8102
166. Tsang R, Dickson JK, Pak H, Walton R, Fraser-Reid B (1987) J Am Chem Soc 109:3484
167. Bartlett PA, McLaren KL, Ting PC (1988) J Am Chem Soc 110:1633
168. Corey EJ, Pyne SG (1983) Tetrahedron Lett 24:2821
169. Ando O, Satake H, Itoi K, Sato A, Nakajima M, Takahashi S, Haruyama H, Ohkuma Y, Kinoshita T, Enokita R (1991) J Antibiot 44: 1165
170. Ando O, Nakajima M, Hamano K, Itoi K, Takahashi S, Takamatsu Y, Sato A, Enokita R, Haruyama H, Kinoshita T (1993) J Antibiot 46:1116
171. Nakayama T Amachi T, Murao S, Sakai T, Shin T, Kenny PT, Iwashita T, Zagorsky M, Komura H, Nomoto K (1991) J Chem Soc Chem Commun 919
172. Kobayashi Y, Miyazaki H, Shiozaki M (1992) J Am Chem Soc 114:10065
173. Kobayashi Y, Miyazaki H, Shiozaki M (1994) J Org Chem 59:813
174. Kobayashi Y, Miyazaki H, Shiozaki M (1993) Tetrahedron Lett 34:1505
175. Bernet B, Vasella A (1979) Helv Chim Acta 62:1990
176. Curan DP (1983) J Am Chem Soc 105:5826
177. Katsuki T, Sharpless KB (1980) J Am Chem Soc 102:5826
178. Shibanuma T, Shiono M, Mukaiyama T (1977) Chem Lett 575
179. Mukaiyama T (1979) Angew Chem Int Ed Engl 18:707
180. Takeda T, Mukaiyama T (1980) Chem Lett 163
181. Camarasa MJ, F-Resa P, Garcialopez MT, G delas Heras F, M-Castrillon PP, Felix AS (1984) Synthesis 509
182. Ogawa S, Uchida C, Yuming Y (1992) J Chem Soc Chem Commun 886
183. Ogawa S, Uchida C (1992) J Chem Soc Perkin Trans 1 1939
184. Ogawa S, Uchida C (1993) Chem Lett 173
185. Uchida C, Yamagishi T, Ogawa S (1993) Chem Lett 971
186. Uchida C, Yamagishi T, Ogawa S (1994) J Chem Soc Perkin Trans 1 559
187. Kobayashi Y, Shiozaki M (1994) J Antibiot 47:243
188. Kobayashi Y, Miyazaki H, Shiozaki M, Haruyama H (1994) J Antibiot 47:932
189. Shiozaki M, Kobayashi Y, Arai M, Haruyama H (1994) Tetrahedron Lett 35:887
190. Shiozaki M, Arai M, Kobayashi Y, Kasuya A, Miyamoto S, Furukawa Y, Takauama T, Haruyama H (1994) J Org Chem 59:4450
191. Corey EJ, Kang J, Kyler K (1985) Tetrahedron Lett 26:555
192. Goering BK, Li J, Ganem B (1995) Tetrahedron Lett 36:8905
193. Li J, Lang F, Ganem B (1998) J Org Chem 63:3403
194. Corey EJ, Koelliker U, Neuffer J (1971) J Am Chem Soc 93:1489
195. Grieco PA, Gilman S, Nishizawa M (1976) J Org Chem 41:1485
196. Ledford BE, Carreira EM (1995) J Am Chem Soc 117:11811
197. Overman LE (1974) J Am Chem Soc 96: 597
198. Dawe RD, Molinski TF, Turner JV (1984) Tetrahedron Lett 25:2061
199. Adam W, Bialas J, Hadjiarapoglou L (1991) Chem Ber 124:2377
200. Zweifel G, Ayyangar NR, Brown HC (1963) J Am Chem Soc 85:2072
201. Mancuso AJ, Huang S-L, Swern D (1975) J Org Chem 43:2485
202. Knapp S, Purandare A, Rupitz K, Withers SG (1994) J Am Chem Soc 116:7461
203. Kobayashi Y, Shiozaki M, Ando O (1995) J Org Chem 60:2570
204. Marquez VE, Lim M, Khan MS, Kaskar B (1991) Nucleic Acid Chem 4:27
205. Hughes DL (1992) Org React 42:335
206. Knapp S, Naughton ABJ, Murali Dhar TG (1992) Tetrahedron Lett 33:1025
207. Tatsuta K, Yasuda S, Kurihara K, Tanabe K, Shinei R, Okonogi T (1997) Tetrahedron Lett 38:1439
208. Tatsuta K, Yasuda K, Araki N, Takahashi M, Kamiya Y (1998) Tetrahedron Lett 39: 401
209. Tatsuta K, Mukai H, Takahashi M (2000) J Antibiot 53:430
210. Curran DP (1988) Synthesis 489
211. Ramaiah M (1987) Tetrahedron 43:3541
212. Neumann WP (1987) Synthesis 665
213. Molander GA, Harris CR (1996) Chem Rev 96:307 and literature cited therein
214. Enholm EJ, Trivellas A (1989) J Am Chem Soc 111:6463

215. Enholm EJ, Satici H, Trivellas A (1989) J Org Chem 54:5841
216. Enholm EJ, Trivellas A (1989) Tetrahedron Lett 35:1627
217. Grove JJC, Holzapfel CW, Williams DBC (1996) Tetrahedron Lett 37:1305
218. Grove JJC, Holzapfel CW, Williams DBC (1996) Tetrahedron Lett 37:5817
219. Chiara JL, Martinez S, Bernabe M (1996) J Org Chem 61:6468
220. Chiara JL, Marco-Contelles J, Khiar N, Gallego P, Destabel C, Bernabe M (1995) J Org Chem 60:6010
221. Marco-Contelles J, Gallego P, Rodriguez-Fernandez M, Khiar N, Destabel C, Bernabe M, Martinez-Grau A, Chiara JL (1997) J Org Chem 62:7397
222. Perrin E, Mallet J-M, Sinay P (1996) Carbohydr Lett 1:215
223. Adinolfi M, Barone G, Iadonisi A, Mangoni L (1998) Tetrahedron Lett 39:2021
224. Chenede A, Pothier P, Sallogoub M, Fairbanks AJ, Sinay P (1995) J Chem Soc Chem Commun 1373
225. Aurrecoechea JM, Lopez B (1998) Tetrahedron Lett 39:2857
226. Ito H, Motoki Y, Taguchi T, Hanzawa Y (1993) J Am Chem Soc 115:8835
227. Hanzawa Y, Ito H, Taguchi T (1995) Synlett 299 and literature cited therein

9.3 Sulfur-Containing Glycomimetics

Andreas Steiner, *Arnold Stütz*, *Tanja Wrodnigg*
Glycogroup, Institut für Organische Chemie, Technische Universität Graz,
8010 Graz, Austria
stuetz@tugraz.at

Abstract

Sugars with heteroatoms other than oxygen in the ring are unusual but highly interesting relatives of "normal" sugars. In particular, 4-thiofuranoses as well as 5-thiopyranoses have attracted considerable attention. Side by side with their counterparts with nitrogen in the ring,

In: *Glycoscience*. Fraser-Reid B, Tatsuta K, Thiem J (eds)
Chapter-DOI 10-1007/978-3-540-30429-6_50:

the iminosugars, these thiosugars feature exciting challenges for both the chemists synthesizing them and the biologists investigating their biological properties as glycomimetics. This chapter is aimed at giving an introduction to the world of sugars with sulfur in the ring, their syntheses as well as their chemical and biochemical properties as known to date.

Keywords

Thiosugars; 4-Thioaldofuranoses; 5-Thioaldopyranoses; 5-Thioketofuranoses; 6-Thioketopyranoses; 6-Thioaldoseptanoses; Thiosugar-related natural products; Thiodisaccharides

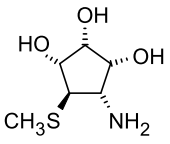
Abbreviations

DMF *N*,*N*-dimethylformamide
RAMA rabbit muscle aldolase

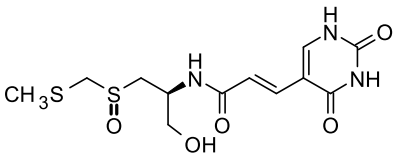
1 Introduction

Displacement of the ring oxygen in a sugar by other heteroatoms such as nitrogen or sulfur leads to pronounced changes of the physico-chemical properties and, thus, in the "information content" of the modified carbohydrate towards biochemical systems. Consequently, the altered biological properties of such a derivative with sugar-modifying enzymes, for example glycosidases or glycosyl transferases, as well as other carbohydrate recognizing biological entities can lead to useful applications as a diagnostic or therapeutic agent. Sugar analogs with nitrogen instead of oxygen in the ring have attracted considerable attention due to their pronounced biological activities as inhibitors of glycosidases. Similar effects have been found with carbohydrates bearing sulfur in the ring. Other elements introduced into this position, for example phosphorus or selenium, have had lesser impact on carbohydrate chemistry and glycomimetics research.

In the periodic table, sulfur is positioned in the same group next to oxygen. Despite such a close resemblance of the outer shell orbitals, sulfur can also occur in oxidation states other than −2, namely +2, +4 as well as +6. Thus, a large variety of organo-sulfur compounds exhibiting a wide range of different properties is found in nature.



1



2

In addition to quite a few sulfur-containing amino acids, many sulfur containing organic natural products exhibiting notable biological activities are known to date, such as the vitamin biotin, the *Allium* components which are responsible for the characteristic taste and smell of onion, garlic, and related plants, the antibiotic penicillins and lincomycins as well as the

powerful mannosidase inhibitor mannostatin (**1**) [1]. Compounds with sulfur in a higher oxidation state include, the antiviral agent sparsomycin (**2**) [2], a sulfoxide as well as taurin (2-aminoethanesulfonic acid, the sulfonic acid equivalent of glycine) just to mention a few well-known examples. Consequently, sugars, sugar analogs, as well as carbohydrate-related natural products with sulfur in the ring have attracted considerable interest, both in terms of synthetic challenges as well as potential biological activities.

2 Thiosugars

2.1 Physico-Chemical Properties

A sulfur-containing hemiacetal ring is more stable [3,4,5] as was concluded from the distinctly reduced mutarotation exhibited by 5-thioaldoses when compared to their oxygenated parent compounds. On the other hand, it was discovered that glycosides of such thiosugars are more prone to hydrolysis than their natural counterparts [6].

Because of the higher nucleophilicity of sulfur as compared to oxygen, the ring size is generally dependent on the position of the thiol group. In 5-thiosugars the pyranoid tautomers prevail over five-membered ring systems whereas in 4-thiosugars the furanoid form is found to be preferred over the pyranoid ring containing O-5. As with natural sugars, for thermodynamic reasons, pyranoid 5-thiosugars have been found to be more stable than furanoid 4-thiosugars [7]. As a consequence of the smaller C–S–C bond angle in 5-thio-D-glucopyranose as compared to the C–O–C bond angle in D-glucopyranose, the former was found to be more puckered [8].

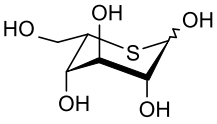
Anchimeric assistance by the ring sulfur during nucleophilic displacement reactions at various ring positions have frequently been reported and could also be exploited for defined and desired structural alterations.

Sulfoxides and sulfones have been prepared by oxidation of thiosugars but have only found very limited use as carbohydrate mimics.

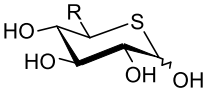
On the other hand, biologically active sulfonium-containing species have been discovered and will be discussed in the following.

2.2 Early Developments

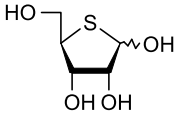
As early as 1961/62 the first examples of 5-thiopyranoses were reported in the literature. Adley and Owen [4] described the synthesis of 5-thio-L-idopyranose (**3**) by thiirane ring opening of 5,6-dideoxy-5,6-epithio-1,2-*O*-isopropylidene-β-L-idofuranose with potassium acetate in a mixture of glacial acetic acid and acetic anhydride and subsequent deprotection.



3

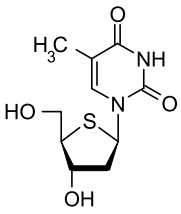


4 R = H
5 R = CH_2OH

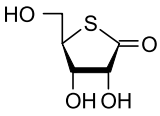


6

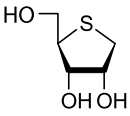
5-Thio-D-xylopyranose (**4**) was prepared by the same group and also by Schwarz and Yule [5] as well as by Whistler and co-workers [9,10]. 5-Thio-D-glucopyranose (**5**) was first synthesized by Feather and Whistler [11].



7



8



9

The first sulfur-containing furanose rings were described by Reist and co-workers [12,13], who prepared 4-thio-D- (**6**) and -L-ribofuranose via nucleophilic displacement of the tosylate in 2,3-*O*-isopropylidene-4-*O*-toluenesulfonyl-α-D- and -α-L-lyxopyranosides, respectively, with potassium thiobenzoate in DMF. Subsequently, several nucleosides of 4-thiofuranoses were synthesized, for example, the 4′-thio analogue [14] of natural thymidine (**7**) or the corresponding derivative of cytidine [14].

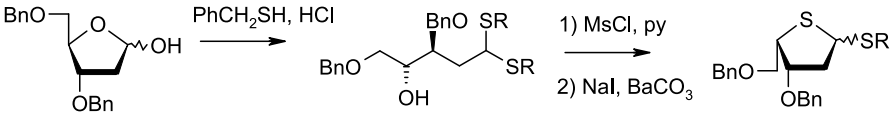
2.3 Furanoid Systems

Replacement of the ring oxygen in naturally occurring as well as synthetic nucleosides for potentially useful therapeutic reasons has been a major task of carbohydrate chemistry in context with the topic under consideration. Thus, starting in the early 1960s, all diastereomeric 4-thio analogs of natural pentofuranoses as well as some of their enantiomers were prepared and based on these approaches a wide range of synthetic 4′-thionuclosides has been reported.

2.3.1 4-Thioaldopentoses and Derivatives

4-Thio-D- as well as -L-ribofuranosides, synthesized by Reist and co-workers [12,13], were amongst the first furanoid thiosugars with sulfur in the ring. The synthesis of methyl 4-thio-D-ribofuranoside from α-L-lyxose was also reported by Whistler and co-workers [15]. The corresponding per-*O*-acetylated sugar served as an intermediate in the preparation of various purine and pyrimidine nucleosides [14]. A synthesis of free 4-thio-D-ribofuranose was achieved by Imbach and co-workers [16] who also prepared protected derivatives suitable for nucleoside syntheses [17]. Protected 4-thio-D-arabinofuranose as well as the corresponding D-*xylo* configured sugar were prepared by the Reist group [18]. Nayak and Whistler [19] reported the synthesis of methyl 4-thio-β-D-arabinofuranoside and subsequently employed this compound for nucleoside syntheses [20].

4-Thio analogs of D-ribono- (**8**) as well as L-lyxonolactones were synthesized by Varela and Zunszain [21]. Syntheses of 1,4-anhydro-4-thio-D- (**9**) and -L-ribitol were recently described by Altenbach and Merhof [22,23]. 1,4-Anhydro-4-thio-D-arabinitol was employed in the synthesis of a potential inhibitor of glycosyl transferases [24]. Thiofuranosides



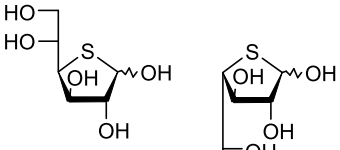
10

Scheme 1

of 4-thiopentofuranoses [25], prepared by intramolecular displacement of a sulfonate by a C-1 thioacetal sulfur atom, as well as the corresponding 2-deoxyderivatives such as **10** [26], including epimers thereof and C-3-branched analogues [27] as well as an O-glycoside [28] have been employed mainly en route to 4-thionucleosides.

2.3.2 4-Thioaldohexoses and Derivatives

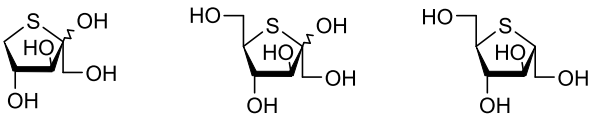
In the series of the aldohexoses, 4-thio derivatives of D-glucose, **11** [29], D-mannose [30] and, more recently, D-galactose **12** [31] have been synthesized. In addition, the 6-deoxy analogs of D-*gluco* [32], D-*gulo* [33] as well as D-*ido* [34] configured 4-thioaldohexoses have been prepared. Per-*O*-acetylated 6-deoxy-4-thio-D-galactofuranose was obtained by acetolysis of methyl 6-deoxy-4-thio-D-galactopyranoside [35]. The synthesis of methyl α-D-talofuranoside was reported by Garegg and co-workers [36]. 4-Thiofuranoside derivatives of D-galactosamine are also known [37].



11 **12**

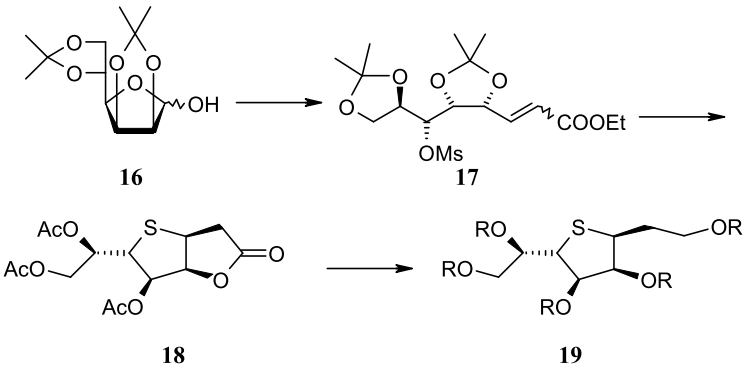
2.3.3 5-Thioketopentoses, 5-Thioketohexoses, and Derivatives

A de novo approach to 5-thio-D-*threo*-2-pentulofuranose (5-thio-D-xylulose, **13**) from 2-mercaptoacetaldehyde with the aid of rabbit muscle aldolase (RAMA, EC 4.1.2.13) as well as yeast transketolase (EC 2.2.1.1) was reported by Effenberger and co-workers [38]. This approach was recently extended in combination with the glucose isomerase (EC 5.3.1.5) catalyzed isomerization of compound **13** in an interesting synthesis [39] of 5-thio-D-xylose (**4**), a precursor for antithrombotic drugs [40].



13 **14** **15**

5-Thio-D-fructofuranose (**14**) was synthesized by Chmielewski and Whistler. The 2-deoxy derivative of 5-thio-β-D-fructofuranose, 2,5-anhydro-5-thio-D-mannitol (**15**), as well as the corresponding L-*ido* configured compound were prepared by ring-opening reactions of D-mannitol derived 1,2:5,6-diepoxides and were found to be poor inhibitors of D-glucosidases [41]. Various protected bicyclic 2,5-anhydro derivatives of 5-thio-D-mannitol were obtained by ring opening of trans-annular thiiranium ring-opening reactions of methyl 2,3,4-tri-*O*-acetyl-6-*O*-mesyl-5-thio-α-D-glucopyranoside [42].



16 **17** **18** **19**

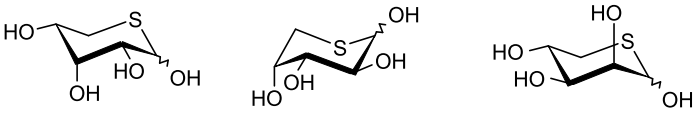
Scheme 2

Starting from 2,3:5,6-di-*O*-isopropylidene-D-mannose (**16**, *Scheme 2*), by chain elongation to enoate **17** and subsequent thiol introduction and ring closure, 6,3-anhydro-6-thio-octanoates **18** as well as the corresponding 6-thioanhydrooctitols **19** could be prepared [43]. Not unexpectedly, they proved ineffective in the attempted inhibition of a range of standard glycosidases.

2.4 Pyranoid Systems

2.4.1 5-Thioaldopentoses, 5-Thioaldohexoses, 6-Thioketohexoses, and Derivatives

5-Thio-D-xylose (**4**) was prepared by Adley and Owen [4], Scharz and Yule [5], as well as by Whistler and co-workers [9,10].

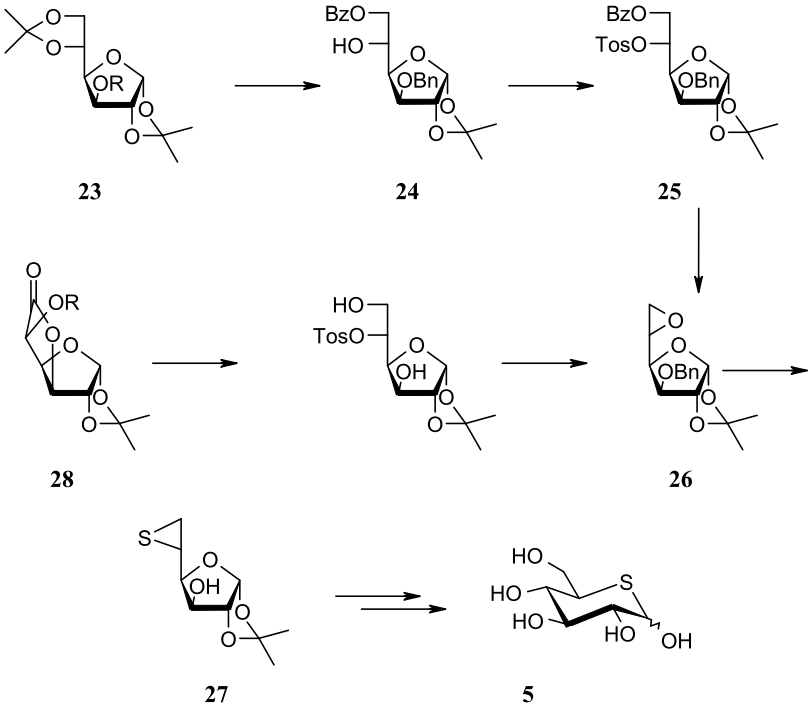


20 **21** **22**

5-Thio-D-ribose (**20**) was first reported by Clayton and Hughes [44,45], an improved synthesis from D-ribono-1,4-lactone was published in 2002 [46]. Isopropylidene acetals [47] as well as 1,4-anhydro derivatives thereof [48] have also been made available.

5-Thio-D-arabinose (**21**) as well as 5-thio-D-lyxose (**22**) were made by Hughes and Munkombwe [49] from 1,2-*O*-isopropylidene-5-*O*-tosyl-D-arabinofuranose by reaction with thioacetate and subsequent deprotection. A synthesis from 5-*O*-trityl-D-arabinofuranose was reported in 1996 [50].

5-Thio-D-glucose (**5**) was one of the first examples of sugars with sulfur in the ring [11].



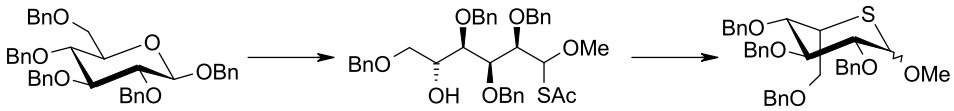
Scheme 3

In a typical synthetic approach (*Scheme 3*) [51], 1,2:5,6-di-*O*-isopropylidene-α-D-glucofuranose (**23**) was employed as the starting material.

By a sequence of standard protecting group manipulations, O-3 was protected as the benzyl ether, the 5,6-acetal was removed by acid hydrolysis and O-6 was subsequently regioselectively protected with a benzoyl group to give **24**. Activation of C-5 was achieved by *O*-tosylation to give compound **25**. Upon base treatment, this intermediate was converted into the corresponding L-*ido* configured 5,6-oxirane **26** by intramolecular nucleophilic displacement of the tosylate. Thiourea treatment of **26** gave the corresponding thiirane derivative **27** with inversion of configuration at C-5. This was regioselectively opened at C-6 with acetate. Removal of the protecting groups gave 5-thio-D-glucose (**5**) in good overall yields after ten steps.

An improved synthesis based on the same strategy was devised by Driguez and Henrissat [52]. These workers took advantage of the regiospecifically available position C-5 in 1,2-*O*-isopropylidene-α-D-glucofuranurono-6,3-lactone (**28**). This easily accessible starting material was O-sulfonylated with tosyl chloride followed by reduction of the lactone moiety and

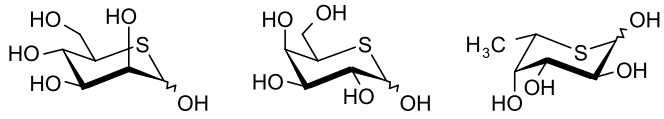
subsequent 5,6-oxirane formation, the remaining steps to compound **5** being the same as in Whistler's approach. Recently, Uenishi and Ohmiya provided an interesting threose-based approach featuring a two-carbons chain extension followed by Sharpless epoxidation of the primary allylic alcohol thus obtained and conversion of the oxirane into a 5-thio 5,6-cyclic xanthate intermediate from which – depending on the configuration of the threose – 5-thio-D- and L-glucose as well as the corresponding L-altro epimers and their 1,6-anhydro derivatives could be made available in enantiomerically pure forms [53].



29 **30** **31**

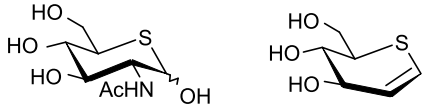
Scheme 4

Another chemical method for the preparation of 5-thioaldopyranosides was chosen by Hashimoto and co-workers [54]. They converted per-*O*-benzylated methyl β-D-glucopyranoside **29** into the open-chain mixed acetal **30** which, in turn, was treated with the Mitsunobu system, triphenylphosphine and diethyl azodicarboxylate, in the presence of benzoic acid to yield protected 5-thio-L-idopyranosides **31**.



32 **33** **34**

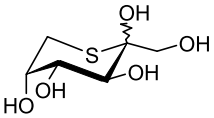
Most of the other important 5-thiopyranoses such as 5-thio-D-mannose (**32**) [55], 5-thio-D-galactose (**33**) [56], and 5-thio-D-allose [57] have been synthesized.
5-Thio-L-fucopyranose (**34**) was synthesized by Hashimoto and co-workers [58,59] combining the synthetic approach depicted in *Scheme 3* with a reductive thiirane ring-opening procedure previously reported in the synthesis of 6-deoxy-5-thio-D-glucose [42].
The 5-thio analogue of *N*-acetyl-D-glucosamine (**35**) was synthesized from suitably protected furanoid [60,61,62] or open-chain [63] D-glucosamine derivatives employing the general approach via the thiirane route as well as from 5-thio-D-glucal (**36**) [64] via azidonitration [65,66].



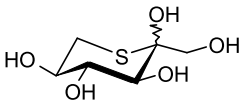
35 **36**

Access to thioketoses with sulfur in the ring could also conveniently be provided by an aldolase-based strategy. This involves rabbit muscle aldolase catalyzed carbon–carbon bond

formation of a 3-thioglyceraldehyde with dihydroxyacetone phosphate leading, for example, to compounds such as 6-thio-D-fructose (**37**) [38] or 6-thio-L-sorbose (**38**) [38,67].



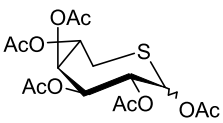
37



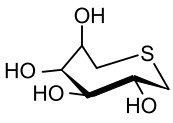
38

2.5 Septanoses and Derivatives

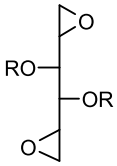
Despite the strong nucleophilicity of the thiol substituent, septanoses have not been found to generally form spontaneously from 6-thiosugars [7]. As with their natural counterparts, in cases where the formation of a pyranoid ring is not possible, the 6-thiofuranose is favored over the 6-thioseptanose system [18,68].



39



40



41

Early synthetic approaches to the 6-thioseptanose system were provided by Cox and Owen [68] as well as by Whistler and Campbell [69]. The latter workers prepared a protected open-chain 6-thio dithioacetal of D-galactose which upon deprotection at C-1 and S-6 could be cyclized to form the corresponding 1,2,3,4,5-penta-*O*-acetyl-D-galacto-6-thioseptanose (**39**) from which several derivatives such as both anomers of the thioseptanosyl chloride as well as the methyl thioseptanosides were synthesized. It was found that the methyl thioseptanoside is more prone to hydrolysis than the corresponding sulfur-containing pyranoid glycoside.
1,6-Anhydroalditol derivatives such as **40** were recently obtained by Le Merrer from D-mannitol via di-oxiranes such as **41** [41].

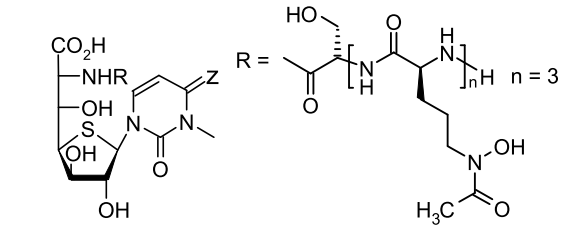
2.6 Summary

From the selected examples mentioned, it is clearly evident that sufficiently versatile methodology has been made available to allow for the synthesis of practically any sugar analog with sulfur in the ring. On this sound basis, a range of interesting and potentially valuable sophisticated natural products and synthetic analogs have recently been targeted.

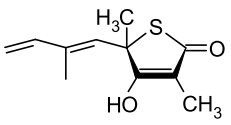
3 Examples of Glycomimetics with Sulfur in the Sugar Ring

3.1 Natural Products

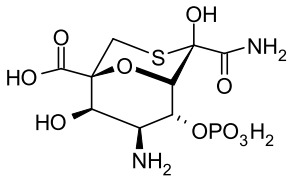
Only a few sugars and related compounds with sulfur in the ring have been reported as natural products to date. In the following, some selected, structurally quite dissimilar examples will be discussed.



42



43



44

3.1.1 Albomycin

As early as 1952, the antibiotic albomycin (**42**) was reported in a Polish journal [70]. It was found to be an iron-binding siderophore whose structure was determined by a Bayer group in collaboration with Paulsen and co-workers more than 30 years later [71]. Novobiocin is a synonym also found in the literature [72].

3.1.2 Thiolactomycin

The antibiotic thiolactomycin (**43**), a fermentation product from a *Nocardia* species containing an unusual thiolactone moiety was patented as antibiotic "no. 2-200" and subsequently reported in the literature in 1982 [73,74]. It resembles a sugar-derived α,β-unsaturated 4-thioglycono-1,4-lactone and was found to be a broad-spectrum antibiotic [75] by interference with the fatty acid metabolism of bacteria and also inhibited inducible β-lactamases [76]. A de novo synthesis of the racemate was reported by a Du Pont group in 1984 [77]. Chambers and Thomas [78] reported the synthesis of the (5*S*)-enantiomer in 1989 and concluded from its optical rotation that the natural product is the (5*R*)-enantiomer.

As an inhibitor of mycolic acid biosynthesis, compound **43** exhibits activity against strains of *Mycobacterium tuberculosis*, the causative agent of tuberculosis [79,80,81]. Furthermore,

thiolactomycin was also found to be active against various parasites including Trypanosoma [82]. Structure-activity considerations have recently been reported [83].
Structurally closely related is the antibiotic thiotetromycin [84,85].

3.1.3 Tagetitoxin

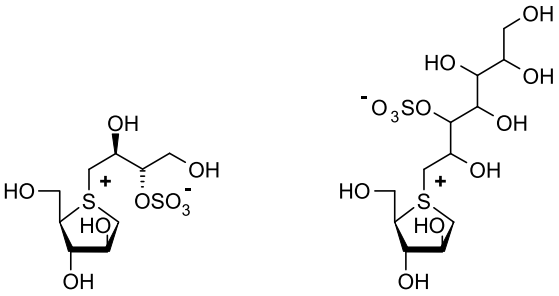
Tagetitoxin (**44**) is a metabolite of *Pseudomonas syringae*, first reported in 1981 [86] and, based on NMR investigations, was assumed to be a highly functionalized eight-membered ring system resembling a sulfur-containing keto-octonic amide [87]. A revised structure was published in 1989 [88] suggesting a 9-oxa-3-thiabicyclo[3.3.1]nonane core. Early synthetic studies providing simplified analogs were conducted by Furneaux and co-workers [89]. Recently, successful completion of a partially protected intermediate en route to the bicyclic system was reported [90].
It was noted that this compound is an inhibitor of RNA synthesis in plants by specifically interfering with RNA polymerases of chloroplasts [91] suggesting tagetitoxin as a novel herbicidal agent.
Moreover, tagetitoxin was found to inhibit RNA polymerase III by a unique, as yet not fully clarified mechanism [92] thus providing interesting biochemical and medicinal aspects for this compound.

3.1.4 5-Thio-D-Mannose

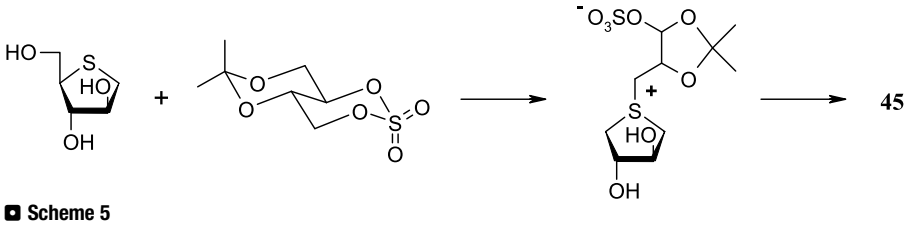
5-Thio-D-mannose (**32**), discovered as a metabolite of the marine sponge *Clathria pyramida* (Lendenfeld) is still unique representing the only example of a free 5-thioaldopyranose in nature [93]. Shortly after this discovery a synthesis from methyl 2,3:5,6-di-*O*-isopropylidene-α-D-mannofuranoside was reported [55] which essentially followed the principle of the approach depicted in *Scheme 3*. Before, 5-thio-D-mannose had been available by molybdate-catalyzed epimerization [94,95] of 5-thio-D-glucose [51].

3.1.5 Salacinol and Kotalanol

An interesting thiosugar, salacinol **45**, containing sulfur atoms in two different oxidation states was isolated and found to be the active principle of the Indian plant *Salacia reticulata* which is used in local traditional medicine for treatment of diabetes [96]. The compound was shown to be a potent inhibitor of intestinal α-glucosidases comparable in its activity with the powerful aminosugar derivative acarbose [97]. Kotalanol (**46**), a chain-extended homolog of salacinol, was discovered around the same time [98].
Triggered by their novel structures and consequential synthetic challenges as well as by their interesting biological properties, synthetic efforts were made to make salacinol and kotalanol as well as numerous synthetic derivatives available. Two teams of researchers, Yoshikawa and co-workers as well as Pinto and his group have provided significant contributions to aspects of synthesis as well as biological activities and structure-activity relationships in this area.
Apparently, the combination of both the sulfonium as well as the sulfate moieties is vital for the biological properties of salacinol and kotalanol [96].



45 46



Scheme 5

The first synthesis of salacinol by a multi-step procedure featuring the reaction of 1,4-anhydro-4-thio-D-arabinitol with the cyclic sulfate derived from 1,3-isopropylidene D-erythritol was reported in 2000 [99]. A similar approach was taken advantage of by Pinto and co-workers shortly thereafter [100]. Exploiting hexafluoroisopropanol as the solvent, the efficacy of the method could be further improved [101]. Improvements and extensions to other configurations of the five-membered ring system were also provided [102,103]. Other modifications include deoxy derivatives [104], chain extended analogues [105,106] as well as 2,5-anhydro-5-thio-L-iditol containing [107] and ring-enlarged [108,109] relatives whose inhibitory properties towards maltase-glucoamylase were screened as potential anti-Diabetes type II agents [110,111] and with a view to active-site mapping of this enzyme [111]. Some of the analogs prepared exhibit excellent inhibitory activities in the single figure micromolar range and below.

Employing *Drosophila* retaining Golgi α-mannosidase II, the binding mode of salacinol was compared with the interactions of swainsonine and kifunensine with the active site [112,113]. To date, to the best of our knowledge, the configuration of the heptitol moiety in kotalanol is still not elucidated.

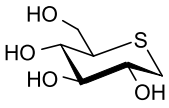
Salacinol derivatives with selenium instead of sulfur in the ring have also been prepared and will be discussed in *Section 4*.

3.2 Synthetic Glycomimetics

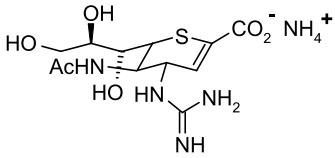
A wide variety of synthetic sugar and nucleoside mimicking compounds with sulfur in the ring have been prepared for basic research as well as with a view to biologically active compounds with improved biochemical properties.

3.2.1 Glycosidase Inhibitors

*5-Thio-*D*-Xylose, 5-Thio-*D*-Glucose* As early as 1965 the inhibitory activity of 5-thio-D-xylopyranose (**4**) against β-D-xylosidases was observed and a fairly high K_i value of 2 mM was reported for this reversible inhibition [114].



47



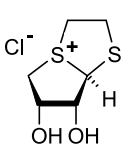
48

In-depth investigations [115,116,117] into the biological properties of 5-thio-D-glucose (**5**) rapidly led to the conclusion that the replacement of the ring oxygen with sulfur has a pronounced influence on the sugar metabolism. Despite being non-toxic with an LD_{50} of 14 g/kg, diabetogenic effects of **5** were reported by Hoffman and Whistler [115]. These workers found rapidly developing glucosuria and hyperglycemia in rats which had received this D-glucose mimic. Other effects observed were diminished glycogen contents in the liver and reduced membrane transport of D-glucose. Furthermore, inhibition of spermatogenesis in mice [118] and rats [119,120] as well as reduction of proliferation of trypanosoma [121] by this compound were observed. Biological activities included inhibition of growth of cancer cells [122]. 5-Thio-D-glucose was also found to inhibit yeast α-glucosidase with a K_i value of 750 µM but is a poor inhibitor of the β-specific enzyme from sweet al.monds [123].
Weak inhibitory effects on the former enzyme have also been observed with the corresponding methyl α-pyranoside [123] as well as the 1-deoxy derivative **47** [41,124]. None of these sugars or related derivatives [123,124,125] have significant effects on β-glucosidases. No pronounced biological effects were reported for 6-thio-D-fructopyranose (**37**). Instead, this compound is unusually sweet and essentially non-toxic in mice [126]. In addition to the chemical synthesis [127], an enzymatic preparation from 6-thio-D-glucose employing glucose isomerase (EC 5.3.1.5) was reported [128]. In keeping with the general properties of sulfur-containing pyranoid ring systems, the 6-thiofructopyranose is strongly favored in the equilibrium allowing the isolation of this product in over 90% yield.

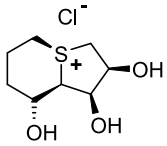
*5-Thio-*L*-Fucose* 5-Thio-L-fucose (**34**) was characterized as a fairly potent α-L-fucosidase inhibitor with K_i values of 42 µM and 84 µM against the enzymes from bovine epididymis and kidneys, respectively. Nevertheless, it is about five orders of magnitude less powerful than 1-deoxy-L-fuconojirimycin [129], the corresponding iminoalditol featuring a basic nitrogen instead of sulfur in the ring. Recently, the same group reported the chemical synthesis of fucosidase inhibiting disaccharides containing the 5-thio-α-L-fucopyranosyl moiety [130]. These were found to exhibit similar inhibitory power as **34**.

Neuraminic Acid Analogs Non-natural mimics of neuraminic acids which inhibit viral sialidase have recently emerged as interesting anti-influenza drugs [131]. In this context, 6-thio analogs have been targeted by von Itzstein and co-workers [132] based on the chain exten-

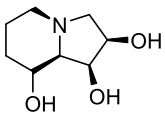
sion of suitably protected 3-thio-*N*-acetyl-D-glucosamine derivatives as outlined by Mack and Brossmer [133,134]. Gratifyingly, one of the derivatives with sulfur in the ring, **48**, inhibited influenza virus sialidase with an IC_{50} value in the low nanomolar range.



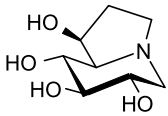
49



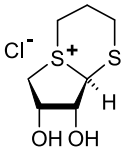
50



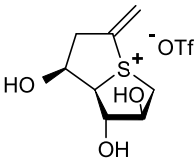
51



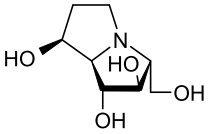
52



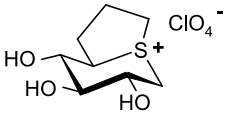
56



55



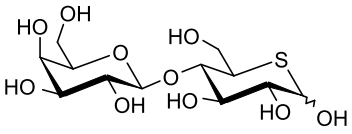
54



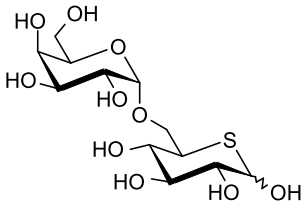
53

Analogs of Glycosidase Inhibiting Polyhydroxyindolizidine and Pyrrolizidine Alkaloids On the basis of an early finding by a French-British team of workers [135] who discovered that bicyclic D-erythrose-derived dithioacetal **49** is a selective inhibitor of human liver α-mannosidases, and supported by synthetic technology developed for the preparation of salacinol analogs, Pinto and co-workers have also provided access to the thio analog **50** of the powerful mannosidase-inhibiting alkaloid swainsonine (**51**) and related structures [136,137]. Furthermore, amongst others, some variations on the basic structure of the glucosidase inhibitors castanospermine **52** such as **53** [138,139] and australine **54**, for example **55** [140] have been synthesized by the same workers.

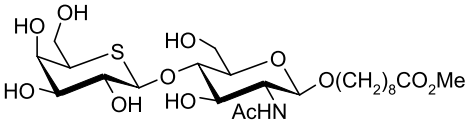
Another swainsonine-related compound (**56**) was also reported by Siriwardena and his group [141] in an extension of their seminal publication [135].



57

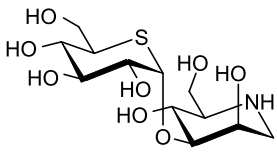


58

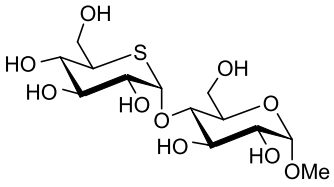


59

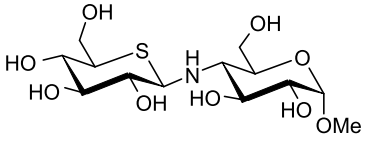
5-Thiodisaccharides and Oligosaccharides The first examples of disaccharides containing a 5-thiopyranosyl unit were reported between 1991 and 1993 by Wong and his group [142] for lactose analogs **57** and **58** featuring the thiosugar at the reducing end, the Thiem team [143] for a 1,1′-linked disaccharide containing 5-thiogentosamine (3-aminodeoxy-D-xylose) and galactose as well as Hindsgaul and co-workers [144] for a 5′-thiolactosamine analog **59** with the thiosugar as the non-reducing sub-unit. Such 5′-thiolactosaminides were observed to exhibit increased stability towards β-galactosidase from *E. coli* [144].
Interestingly, all these workers exploited biocatalytic approaches taking advantage of galactosyltransferase-mediated glycosyl transfer.
Chemical 5-thioglycosylation has largely relied on the trichloroacetimidate procedure as exemplified by Pinto and co-workers of 1,2- as well as 1,6-connected gluco-disaccharides with the thiosugar at the non-reducing end [145]. Several examples of classical Koenigs–Knorr methodology have also been reported, just to mention the most frequently employed approaches.
Targeted compounds include analogs of the H-type 2 [146,147] and Lewis^X [148,149] antigens as well as a wide range of potential endo-glycosidase inhibitors.



60

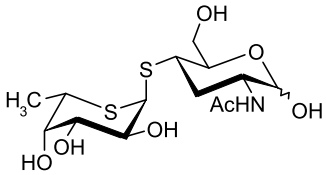


61

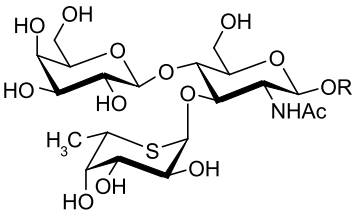


62

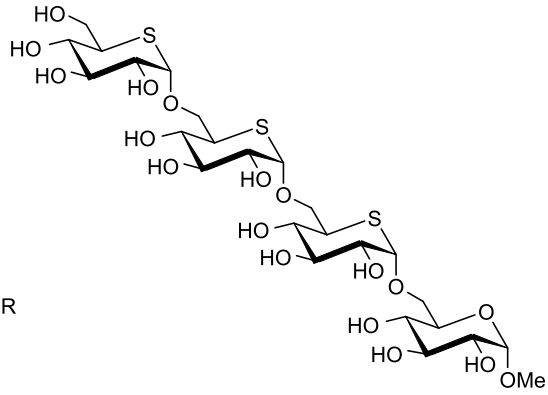
For example, the 3-*O*-(5-thio-α-D-glucopyranosyl) derivative of 1-deoxymannojirimycin (**60**), a powerful D-mannosidase inhibitor, was synthesized as a potential inhibitor of *endo*-α-D-mannosidase of glycoprotein trimming [150].
The weak glycosidase inhibitory activities usually observed for monosaccharides with sulfur in the ring when compared to the corresponding natural substrates have frequently also been found with di- and oligosaccharides.
Maltose analogs **61** featuring the thiosugar at the non-reducing end are only weakly inhibiting glucoamylase G2 [151], whereas disaccharides with sulfur in the ring of the non-reducing sugar and a basic nitrogen as interglycosidic linkage such as **62** were potent glucoamylase inhibitors [152]. Because of the sensitivity of fucosidases, 5-thiofucose containing thio-bridged disaccharides, for example **63**, inhibit fucosidase from bovine epididymis with K_i-values in the low micromolar range [153].



63

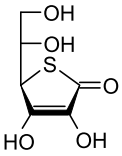


64

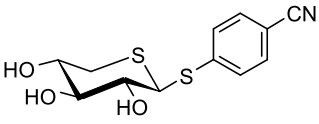


65

Trisaccharides such as compound **64** featuring a 5-thiopyranosidic linkage were successfully synthesized with a view to sulfur-containing Lewis^X analogues [154].
A maltotetraoside **65** containing three 5-thioglucose units in the non-reducing positions was prepared by the Hashimoto group who exploited the trichloroacetimidate approach in a linear sequence of glycosylation steps [155].

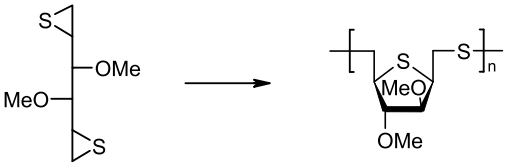


66



67

Others 4-Thioascorbic acid (**66**) was prepared by Stachel and co-workers who could demonstrate that this analogue is a stronger reducing agent than the parent compound [156].
Orally administered 4-cyanophenyl 1,5-dithio-β-D-xylopyranoside (**67**) was recently reported to exhibit powerful antithrombotic activity [157]. Consequently, syntheses of several related sugar derivatives of various chain lengths and configurations have been reported since [158,159,160].



68 **69**

Scheme 6

Polymers An interesting new concept for the synthesis of thiosugar oligo- and polymers was recently introduced taking advantage of open-chain meso alditol derived *O*-methyl protected 1,2;5,6-bisthiiranes (**68**) which upon cationic as well as anionic polymerization gave polymers containing 2,5-anhydro-1,5-dithioalditol subunits (**69**) as the main components [161,162]. Regio- and stereoselectivities of the processes have recently been discussed [163].

3.2.2 Nucleosides

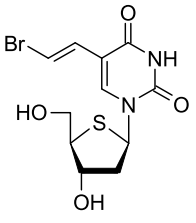
Early synthetic approaches to nucleosides with sulfur in the furanose ring were reported by Reist and co-workers [13,18] as well as Whistler and his group [14,20,164], mainly aiming at antibacterial agents [165,166]. In the early 1990s, Walker and co-workers synthesized, amongst others, 4′-thio derivatives of BVDU [*E*-5-(2-bromoethenyl)-2′-deoxyuridine] such as **70** as anti-viral compounds [167,168]. 4′-Thio-AZT (**71**) and other thio analogs of antiviral nucleosides were prepared by Secrist [169] and other workers [170] but subsequently were found to be toxic [171].

This was also the case with 2′-deoxy-4′-thio analogs of purine nucleosides [172]. On the other hand, the United States Food and Drug Administration has approved the 3′-thia analogue of L-cytidine (**72**) [173] as an anti-HIV agent. This was the first example of an L-enantiomer being more potent than the naturally configured compound. 4′-Thioarabinonucleosides have been discovered to exhibit high activity against herpes and cytomegalo viruses [174].

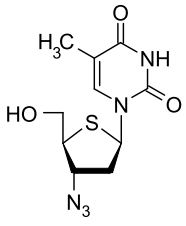
Amongst others, both enantiomers of the 4′-thio analog **73** of the powerful anti-herpes agent L-FMAU ("2′-fluoro-5-methyl-β-L-arabinofuranosyluracil") have been prepared in this context [175,176].

A variety of nucleosides have been prepared by Wirsching and co-workers [177] as well as other groups. Activities, albeit, were found not suitable for further development.

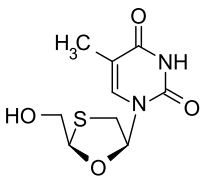
Thionucleotides have also been prepared to be incorporated into oligonucleotides [178].



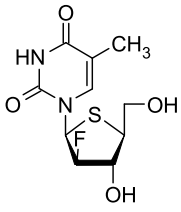
70



71



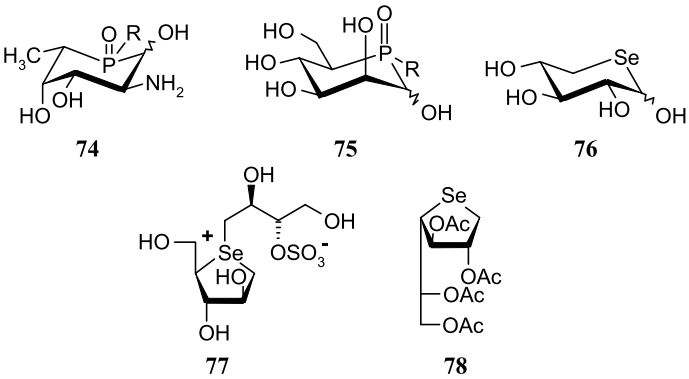
72



73

4 Other Changes in the Ring

Sugar analogs with ring heteroatoms other than nitrogen and sulfur have only raised limited interest due to their lack [179] of noteworthy biological properties. Sugars with phosphorus in the ring have been reviewed by Yamamoto and Inokawa [179], by Witczak and Whistler [180], who have also treated selenosugars [181], as well as by Yamamoto and Hanaya [182].



Some examples in the phosphahexose series include synthetic approaches to analogs of L-fucosamine (**74**) [183] and D-mannopyranose (**75**) [184]. Selenosugars such as **76** have been found to be highly sensitive [7,185] to air oxidation and, thus, are strongly limited in their potential as carbohydrate mimics. An interesting exception to this statement has been provided by Pinto and co-workers who synthesized blintol **77** [186,187], the seleno analog of salacinol which turned out to be stable and to exhibit, albeit mediocre glucosidase inhibiting properties [186]. Other compounds including fully protected selenogalactofuranose derivative **78** featuring selenium species in the ring have also been investigated by this group [188,189].

5 Further Reading

Many aspects of synthesis as well as biological activities and drug design associated with thiosugars of the structural types under consideration here have been competently and extensively reviewed. Amongst others, noteworthy reviews have recently been provided by Witczak [190], Carvalho [191] as well as by Robina and Vogel [192] which will provide additional pieces of information adding more "color" to the selected examples and outlines presented above.

References

1. Aoyagi T, Yamamoto T, Kojiri K, Morishima H, Nagai N, Hamada M, Takeuchi T, Umezawa H (1989) J Antibiot 42:883
2. Wiley PF, MacKellar FA (1976) J Org Chem 41:1858
3. Grimshaw CE, Whistler RL, Cleland WW (1979) J Am Chem Soc 101:1521
4. Adley TJ, Owen LN (1961) Proc Chem Soc 418
5. Schwarz JCP, Yule KC, (1961) Proc Chem Soc 417
6. Whistler RL, van Es T (1963) J Org Chem 28:2303
7. Paulsen H, Todt K (1968) Adv Carbohydr Chem Biochem 23:115
8. Lambert JB, Wharry SM (1981) J Org Chem 46:3193
9. Whistler RL, Feather MS, Ingles DL (1962) J Am Chem Soc 84:122
10. Ingles DL, Whistler RL (1962) J Org Chem 27:3896
11. Feather MS, Whistler RL (1962) Tetrahedron Lett 3:667
12. Reist EJ, Gueffroy DE, Goodman L (1963) J Am Chem Soc 85:3715
13. Reist EJ, Gueffroy DE, Goodman L (1964) J Am Chem Soc 86:5658
14. Urbas B, Whistler RL (1966) J Org Chem 31:813

15. Whistler RL, Dick WE, Ingle TR, Rowell RM, Urbas B (1964) J Org Chem 29:3723
16. Bellon L, Barascut J-L, Imbach J-L (1992) Nucleos Nucleot 11:1467
17. Leydier C, Bellon L, Barascut J-L, Deydier J, Maury G, Pelicano H, El Alaoui MA, Imbach J-L (1994) Nucleos Nucleot 13:2035
18. Reist EJ, Fisher LV, Goodman L (1968) J Org Chem 33:189
19. Nayak UG, Whistler RL (1970) J Org Chem 35:519
20. Whistler RL, Doner LW, Nayak UG (1971) J Org Chem 36:108
21. Varela O, Zunszain PA (1993) J Org Chem 58:7860
22. Altenbach H-J, Merhof GF (1996) Tetrahedron: Asymm 7:3087
23. Altenbach H-J, Brauer DJ, Merhof GF (1997) Tetrahedron 53:6019
24. Yuasa H, Kajimoto T, Wong C-H (1994) Tetrahedron Lett 35:8243
25. Tiwari KN, Secrist III JA, Montgomery (1994) JA Nucleosides Nucleotides 13:1819
26. Hasegawa J, Hamada M, Miyamoto T, Nishide K, Kajimoto T, Uenishi J, Node M (2005) Carbohydr Res 340:2360
27. Braanalt J, Kvarnström I, Niklasson G, Svensson SCT, Classon B, Samuelsson B (1994) J Org Chem 59:1783
28. Braanalt J, Kvarnström I, Svensson SCT, Classon B, Samuelsson B (1994) J Org Chem 59:4430
29. Vegh L, Hardegger E (1973) Helv Chim Acta 56:208
30. Shah RH, Bose JL, Bahl OP (1979) Carbohydr Res 77:107
31. Varela O, Cicero D, de Lederkremer RM (1989) J Org Chem 54:1884
32. Owen LN, Ragg PL (1966) J Chem Soc C 1291
33. Boigegrain RA, Gross B (1975) Carbohydr Res 41:135
34. Gross B, Oriez FX (1974) Carbohydr Res 36:385
35. Cicero D, Varela O, de Lederkremer RM (1990) Tetrahedron 46:1131
36. Classon B, Garegg PJ, Samuelsson B, Liu Z (1987) J Carbohydr Chem 6:593
37. Fernández-Bolãnos JG, Zafra E, García S, Fernández-Bolãnos J, Fuentes J (1998) Carbohydr Res 305:33
38. Effenberger F, Straub A, Null V (1992) Liebigs Ann Chem 1297
39. Charmantray F, Dellis P, Hélaine V, Samreth S, Hecquet L (2006) Eur J Org Chem 24:5526
40. Bellamy F, Horton D, Millet J, Picard F, Samreth S, Chazan JB (1993) J Med Chem 36:898
41. Le Merrer I, Fuzier M, Dosbaa I, Foglietti M-J, Depezay J-C (1997) Tetrahedron 53:16731
42. Bozó E, Boros S, Kuszmann J, Gács-Baitz E (1996) Carbohydr Res 290:159
43. Riedner J, Robina I, Fernández-Bolanos JG, Gómez-Bujedo S, Fuentes J (1999) Tetrahedron Asymm 10:3391
44. Clayton CJ, Hughes NA (1967) Carbohydr Res 4:32
45. Clayton CJ, Hughes NA (1973) Carbohydr Res 27:89
46. Lalot J, Stasik I, Demailly G, Beaupere D (2002) Carbohydr Res 337:1411
47. Hughes NA, Wood CJ (1976) Carbohydr Res 49:225
48. Fleetwood A, Hughes NA (1999) Carbohydr Res 317:204
49. Hughes NA, Munkombwe NM (1985) Carbohydr Res 136:397
50. Izumi M, Tsuruta O, Hashimoto H (1996) Carbohydr Res 280:287
51. Whistler RL, Lake WC (1972) Methods Carbohydr Chem 6:286
52. Driguez H, Henrissat B (1981) Tetrahedron Lett 22:5061
53. Uenishi J, Ohmiya H (2003) Tetrahedron 59:7011
54. Hashimoto H, Kawanishi M, Yuasa H (1996) Carbohydr Res 282:207
55. Yuasa H, Izukawa Y, Hashimoto H (1989) J Carbohydr Chem 8:753
56. Shin JEN, Perlin AS (1979) Carbohydr Res 76:165
57. Al-Masoudi NAL, Hughes NA (1986) Carbohydr Res 148:25
58. Hashimoto H, Fujimori T, Yuasa H (1990) J Carbohydr Chem 9:683
59. Hashimoto H, Izumi M (1992) Chem Lett 25
60. Hasegawa A, Kawai Y, Kasugai H, Kiso M (1978) Carbohydr Res 63:131
61. Guthrie RD, O'Shea K (1981) Aust J Chem 34:2225
62. Bognar R, Herczegh P, Whistler RL, Madumelu EB (1981) Carbohydr Res 90:138
63. Tanahashi E, Kiso M, Hasegawa A (1983) Carbohydr Res 117:304
64. Korytnyk W, Angelino N, Dodson-Simmons O, Hanchak M, Madson M, Valentekovic-Horvath S (1983) Carbohydr Res 113:166

65. Lemieux RU, Ratcliffe RM (1979) Can J Chem 57:1244
66. Csuk R, Glänzer BI (1986) J Chem Soc Chem Commun 343
67. Chou W-C, Chen L, Fang J-M, Wong C-H (1994) J Am Chem Soc 116:6191
68. Cox JM, Owen LN (1967) J Chem Soc C 1121
69. Whistler RL, Campbell CS (1966) J Org Chem 31:816
70. Jonczyk J (1952) Farm Polska 8:466
71. Benz G, Born L, Brieden M, Grosser R, Kurz J, Paulsen H, Sinnwell V, Weber B (1984) Liebigs Ann Chem 1408
72. Brock TD (1956) J Bacteriol 72:320
73. Oishi H, Noto T, Ssaki H, Suzuki K, Hayashi T, Okazaki H, Ando K, Sawada M (1982) J Antibiot 35:391
74. Sasaki H, Oishi H, Hayashi T, Matsuura I, Ando K, Sawada M (1982) J Antibiot 35:396
75. Noto T, Miyakawa S, Oishi H, Endo H, Okazaki H (1982) J Antibiot 35:401
76. Miyakawa S, Suzuki K, Noto T, Harada Y, Okazaki H (1982) J Antibiot 35:411
77. Wang C-LJ, Salvino JM (1984) Tetrahedron Lett 25:5243
78. Chambers MS, Thomas EJ (1989) J Chem Soc Chem Commun 23
79. Heath RJ, White SW, Rock CO (2001) Progress Lipid Res 40:467
80. Schaeffer ML, Agnihotri G, Volker C, Kallender H, Brennan PJ, Lonsdale JT (2001) J Biol Chem 276:47029
81. Kremer L, Douglas JD, Baulard AR, Morehouse C, Guy MR, Alland D, Dover LG, Lakey JH, Jacobs WR Jr, Brennan PJ, Minnikin DE, Besra GS (2000) J Biol Chem 275:16857
82. Roberts CW, McLeod R, Rice DW, Ginger M, Chance ML, Goad LJ (2003) Mol Biochem Parasitology 126:129
83. Kim P, Zhang Y-M, Shenoy G, Nguyen Q-A, Boshoff HI, Manjunatha UH, Goodwin MB, Lonsdale J, Price AC, Miller DJ, Duncan K, White SW, Rock CO, Barry CE III, Dowd CS (2006) J Med Chem 49:159
84. Omura S, Iwai Y, Nakagawa A, Iwata R, Takahashi Y, Shimizu H, Tanaka H (1983) J Antibiot 36:109
85. Omura S, Nakagawa A, Iwata R, Hatano A (1983) J Antibiot 36:1781
86. Mitchell RE, Durbin RD (1981) Physiol Plant Pathol 18:157
87. Mitchell RE, Hart PA (1983) Phytochemistry 22:1425
88. Mitchell RE, Coddington JM, Young H (1989) Tetrahedron Lett 30:501
89. Dent BR, Furneaux RH, Gainsford GJ, Lynch GP (1999) Tetrahedron 55:6977
90. Plet JRH, Porter MJ (2006) Chem Commun 1197
91. Mathews DE, Durbin RD (1990) J Biol Chem 265:493
92. Vassylyew DG, Svetlov V, Vassylyeva MN, Perederina A, Igarashi N, Matsugaki N, Wakatsuki S, Artsimovitch I (2005) Nature Struct Mol Biol 12:1086
93. Capon R, MacLeod JK (1987) J Chem Soc Chem Commun 1200
94. Bilik V (1972) Chem Zvesti 26:76
95. Serianni AS, Vuorinen T, Bondo PB (1990) J Carbohydr Chem 9:513
96. Yoshikawa M, Murakami T, Shimada H, Matsuda H, Yamahara J, Tanabe G, Muraoka O (1997) Tetrahedron Lett 38:8367
97. Legler G (1990) Adv Carbohydr Chem Biochem 48:319
98. Yoshikawa M, Murakami T, Yashiro K, Matsuda H (1998) Chem Pharm Bull 46:1339
99. Yuasa H, Takada J, Hashimoto H (2000) Tetrahedron Lett 41:6615
100. Ghavami A, Johnston BD, Pinto BM (2001) J Org Chem 66:2312
101. Ghavami A, Sadalapure KS, Johnston BD, Lobera M, Snider BB, Pinto BM (2003) Synlett 1259
102. Gallienne E, Benazza M, Demailly G, Bolte J, Lemaire M (2005) Tetrahedron 61:4557
103. Kumar NS, Pinto BM (2005) Carbohydr Res 340:2612
104. Muraoka O, Yoshikai K, Takahashi H, Minematsu T, Lu G, Tanabe G, Wang T, Matsuda H, Yoshikawa M (2006) Bioorg Med Chem 14:500
105. Johnston BD, Jensen HH, Pinto BM (2006) J Org Chem 71:1111
106. Liu H, Sim L, Rose DR, Pinto BM (2006) J Org Chem 71:3007
107. Nasi R, Pinto BM (2006) Carbohydr Res 341:2305
108. Szczepina MG, Johnston BD, Yuan Y, Svensson B, Pinto BM (2004) J Am Chem Soc 126:12458
109. Liu H, Pinto BM (2006) Can J Chem 84:497
110. Rossi EJ, Sim L, Kuntz DA, Hahn D, Johnston BD, Ghavami A, Szczepina MG, Kumar NS, Sterchi EE, Nichols BL, Pinto BM, Rose DR (2006) FEBS Journal 273:2673

111. Nasi R, Sim L, Rose DR, Pinto BM (2007) J Org Chem 72:180
112. Kuntz DA, Ghavami A, Johnston BD, Pinto BM, Rose DR (2005) Tetrahedron Assym 16:25
113. Wen X, Yuan Y, Kuntz DA, Rose DR, Pinto BM (2005) Biochemistry 44:6729
114. Claeyssens M, De Bruyne CK (1965) Naturwissenschaften 52:515
115. Hoffman DJ, Whistler RL (1968) Biochemistry 7:4479
116. Whistler RL, Lake WC (1972) Biochem J 130:919
117. Graham LL, Whistler RL (1976) Biochemistry 15:1189
118. Zysk JR, Bushway AA, Whistler RL, Carlton WW (1975) J Reproduct Fertil 45:69
119. Nakamura M, Hall PF (1976) Biochim Biophys Acta 447:474
120. Nakamura M, Hall PF (1977) J Reproduct Fertil 49:395
121. Bushway A, Keenan TW (1978) Biochem Biophys Res Commun 81:305
122. Bushway A, Whistler RL (1975) J Carbohydr Nucleos Nucleot 2:399
123. Kajimoto T, Liu KK-C, Pederson RL, Zhong Z, Ichikawa Y, Porco JA, Wong C-H (1991) J Am Chem Soc 113:6187
124. Cubero II, López-Espinosa MTP, Richardson AC, Ortega MDS (1993) Carbohydr Res 242:109
125. Ermert P, Vasella A (1993) Helv Chim Acta 76:2687
126. Pitts MJ, Chemielewski M, Chen MS, Abd El-Rahman MMA, Whistler RL (1975) Arch Biochem Biophys 169:384
127. Feather M, Whistler RL (1963) J Org Chem 28:1567
128. Chmielewski M, Chen M-S, Whistler RL (1976) Carbohydr Res 49:479
129. Fleet GWJ, Shaw AN, Evans SV, Fellows LE (1985) J Chem Soc Chem Commun 841
130. Izumi M, Tsuruta O, Harayama S, Hashimoto H (1997) J Org Chem 62:992
131. von Itzstein M, Wu W-Y, Kok GB, Pegg MS, Dyason JC, Jin B, Van Phan T, Smythe ML, White HF, Oliver SW, Colman PM, Varghese JN, Ryan DM, Woods JM, Bethell RC, Hotham VJ, Cameron JM, Penn CR (1993) Natureb 363:418
132. Kok GB, Campbell M, Mackey B, von Itzstein M (1996) J Chem Soc, Perkin Trans I 2811
133. Mack H, Brossmer R (1987) Tetrahedron Lett 28:191
134. Mack H, Brossmer R (1998) Tetrahedron 54:4521
135. Siriwardena AH, Chiaroni A, Riche C, El-Daher S, Winchester B, Grierson DS (1992) J Chem Soc Chem Commun 1531
136. Kumar NS, Pinto BM (2006) J Org Chem 71:1262
137. Kumar NS, Pinto BM (2006) Carbohydr Res 341:1685
138. Svansson L, Johnston BD, Gu J-H, Patrick B, Pinto BM (2000) J Am Chem Soc 122:10769
139. Johnson MA, Jensen MT, Svensson B, Pinto BM (2003) J Am Chem Soc 125:5663
140. Kumar NS, Pinto BM (2006) J Org Chem 71:2935
141. Siriwardena A, Strachan H, El-Daher S, Way G, Winchester B, Glushka J, Moremen K, Boons G-J (2005) ChemBioChem 6:845
142. Wong C-H, Krach T, Gautheron-Le Narvor C, Ichikawa Y, Look GC, Gaeta F, Thomson D, Nicolaou KC (1991) J Am Chem Soc 113:8137
143. Nishida Y, Wiemann T, Thiem J (1993) Tetrahedron Lett 34:2905
144. Yuasa H, Hindsgaul O, Palcic MM (1992) J Am Chem Soc 114:5891
145. Mehta S, Jordan KL, Weimar T, Kreis UC, Batchelor RJ, Einstein FWB, Pinto BM (1994) Tetrahedron Asymm 5:2367
146. Izumi M, Tsuruta O, Hashimoto H, Yazawa S (1996) Tetrahedron Lett 37:1809
147. Izumi M, Tsuruta O, Kajihara Y, Yazawa S, Yuasa H, Hashimoto H (2005) Chem Eur J 11:3032
148. Ichikawa Y, Lin Y-C, Dumas DP, Shen G-J, Garcia-Junceda E, Williams MA, Bayer R, Ketcham C, Walker LE, Paulson JC, Wong C-H (1992) J Am Chem Soc 114:9283
149. Tsuruta O, Yuasa H; Hashimoto H, Kurono S, Yazawa S (1999) Bioorg Med Chem Lett 9:1019
150. Ding Y, Hindsgaul O (1998) Bioorg Med Chem Lett 8:1215
151. Mehta S, Andrews JS, Johnston BM, Svensson B, Pinto BM (1995) J Am Chem Soc 117:9783
152. Andrews JS, Weimar T, Frandsen TP, Svensson B, Pinto BM (1995) J Am Chem Soc 117:10799
153. Witczak ZJ (1995) Bioorg Med Chem Lett 5:2169
154. Tsuruta O, Yuasa H, Hashimoto H, Sujimo K, Otter A, Li H, Palcic MM (2003) J Org Chem 68:6400
155. Matsuda H, Ohara K, Morii Y, Hashimoto M, Miyairi K, Okuno T (2003) Bioorg Med Chem Lett 13:1063

156. Stachel H-D, Schachtner J, Lotter H (1993) Tetrahedron 49:4871
157. Bellamy F, Barberousse V, Martin N, Masson P, Millet J, Samreth S, Sepulchre C, Théveniaux J, Horton D (1995) J Med Chem 38:101
158. Bozó E, Boros S, Kuszmann J (1997) Carbohydr Res 301:23
159. Bozó E, Boros S, Kuszmann J (1997) Carbohydr Res 302:149
160. Bozó E, Boros S, Kuszmann J (1997) Carbohydr Res 304:271
161. Satoh T, Kitazawa D, Nonokawa R, Kamada M, Yokota K (2000) Macromol 33:5303
162. Satoh T, Imai T, Sugie N, Nonokawa R, Yokota K, Kakuchi T (2002) J Polymer Sci A 40:965
163. Satoh T, Imai T, Sugie N, Hashimoto H, Kakuchi T (2005) J Polymer Sci A 43:4118
164. Bobek M, Whistler RL, Bloch A (1970) J Med Chem 13:411
165. Bobek M, Whistler RL, Bloch A (1972) J Med Chem 15:168
166. Bobek M, Bloch A, Parthasarathy R, Whistler RL (1975) J Med Chem 18:784
167. Dyson MR, Coe PL, Walker RT (1991) J Chem Soc Chem Commun 741
168. Dyson MR, Coe PL, Walker RT (1991) J Med Chem 34:2782
169. Secrist JA, Riggs RM, Tiwari KN, Montgomery (1992) J Med Chem 35:533
170. Tber B, Fahmi N-E, Ronco G, Villa P, Ewing DF, Mackenzie G (1995) Carbohydr Res 267:203
171. Uenishi J, Takahashi K, Motoyama M, Akashi H, Sasaki T (1994) Nucleos Nucleot 13:1347
172. van Draanen NA, Freeman GA, Short SA, Harvey R, Jansen R, Szczech G, Koszalka GW (1996) J Med Chem 39:538
173. Beach JW, Jeong LS, Alves AJ, Pohl D, Kim HO, Chang C-N, Doong S-L, Schinazi RF, Cheng Y-C, Chu CK (1992) J Org Chem 57:2217
174. Yoshimura Y, Watanabe M, Satoh H, Ashida N, Ijichi K, Sakata S, Machida H, Matsuda A (1997) J Med Chem 40:2177
175. Jeong LS, Moon HR, Yoo SJ, Lee SN, Chun MW, Lim Y-H (1998) Tetrahedron Lett 39:5201
176. Watts JK, Sadalapure K, Choubar N, Pinto BM, Damha MJ (2006) J Org Chem 71:921
177. Wirsching J, Voss J, Giesler A, Kopf J, Adiwidjaja G, Balzarini J, De Clercq E (2003) Nucleos Nucleot Nucl Acids 22:1867
178. Haeberli P, Berger I, Pallan PS, Egli M (2005) Nucl Acids Res 33:3965
179. Yamamoto H, Inokawa S (1984) Adv Carbohydr Chem Biochem 42:135
180. Witczak ZJ, Whistler RL (1983) J Carbohydr Chem 2:351
181. Witczak ZJ, Whistler RL (1982) Heterocycles 19:1719
182. Yamamoto H, Hanaya T (1990) In: Atta-ur-Rhaman (ed) Studies in Natural Products Chemistry. Elsevier, Amsterdam, vol 6, p. 351
183. Hanaya T, Yamamoto H, Kawamoto H, Armour MA, Hogg AM (1992) Bull Chem Soc Jpn 65:2922
184. Hanaya T, Hirose K, Yamamoto H (1993) Heterocycles 36:2557
185. van Es T, Whistler RL (1967) Tetrahedron 23:2849
186. Johnston BD, Ghavami A, Jensen MT, Svensson, Pinto BM (2002) J Am Chem Soc 124:8245
187. Liu H, Pinto BM (2005) J Org Chem 70:753
188. Veerapen N, Yuan Y, Sanders D, Pinto BM (2004) Carbohydr Res 339:2205
189. Veerapen N, Taylor SA, Walsby CJ, Pinto BM (2006) J Am Chem Soc 128:227
190. Witczak ZJ, Culhane JM (2005) Appl Microbiol Biotechnol 69:237
191. Borges de Melo E, da Silveira Gomes A, Carvalho I (2006) Tetrahedron 62:10277
192. Robina I, Vogel P (2002) Curr Org Chem 6:471

9.4 *C*-Glycosyl Analogs of Oligosaccharides

Boris Vauzeilles, *Dominique Urban*, *Gilles Doisneau*, *Jean-Marie Beau*
Laboratoire de Synthèse de Biomolécules, Institut de Chimie Moléculaire et des Matériaux associé au CNRS, Université Paris-Sud, 91405 Orsay Cedex, France
bvauzeil@icmo.u-psud.fr, domurban@icmo.u-psud.fr, gdoisneau@icmo.u-psud.fr, jmbeau@icmo.u-psud.fr

Abstract

This chapter covers the synthesis of a large collection of "*C*-oligosaccharides", synthetic analogs of naturally occurring oligosaccharides in which a carbon atom replaces the anomeric, interglycosidic oxygen atom. These non-natural constructs are stable to chemical and enzymatic degradation, and are primarily devised to probe carbohydrate-based biological processes. These mainly target carbohydrate-protein interactions such as the modulation of glycoenzyme (glycosylhydrolases and transferases) activities or the design of ligands for lectin Carbohydrate Recognition Domains. The discussion is based on the key carbon–carbon bond assembling steps on carbohydrate templates: ionic (anionic and cationic chemistries, sigmatropic rearrangements) or radical assemblage, and olefin metathesis. Synthetic schemes in which at least one of the monosaccharide units is constructed by total synthesis or by cyclization of acyclic

In: *Glycoscience*. Fraser-Reid B, Tatsuta K, Thiem J (eds)
Chapter-DOI 10-1007/978-3-540-30429-6_51:

chiral chains are presented separately in a "partial de novo synthesis" section. The review also provides comments, when they are known, on the conformational and binding properties of these synthetic analogs, as well as their biological behavior when tested.

Keywords

C-Glycosyl compounds; *C*-Linked oligosaccharides; Sugar organometallic; Glycosyl lithium; Glycosyl samarium; Glycosyl radical; Radical cyclization; Nitro sugars; Olefin metathesis

Abbreviations

9-BBN	9-borabicyclo[3.3.1]nonane
ABCN	1,1′-azobis-1-cyclohexanenitrile
AIBN	2,2′-azobisisobutyronitrile
BtOH	1-hydroxybenzotriazole
CAN	ceric ammonium nitrate
COD	1,5-cyclooctadienyl
CSA	10-camphorsulfonic acid
DBAD	di-*tert*-butyl azodicarboxylate
DCC	1,3-dicyclohexylcarbodiimide
DDQ	2,3-dichloro-5,6-dicyano-1,4-benzoquinone
DEAD	diethyl azodicarboxylate
DHP	3,4-dihydro-2*H*-pyran
DIBALH	diisobutylaluminum hydride
DLP	dilauroyl peroxide
DMAP	4-(dimethylamino)pyridine
DMSO	dimethyl sulfoxide
DPPA	diphenylphosphoryl azide
Im	1-imidazolyl or imidazole
LDA	lithium diisopropylamide
LHMDS	lithium hexamethyldisilazide
***m*-CPBA**	*m*-chloroperoxybenzoic acid
MMTr	4-methoxyphenyldiphenylmethyl
NaHMDS	sodium hexamethyldisilazide
NIS	*N*-iodosuccinimide
NMO	4-methylmorpholine *N*-oxide
NMP	1-methyl-2-pyrrolidinone
PCC	pyridinium chlorochromate
PMB	4-methoxybenzyl
Py	2-pyridyl or pyridine
RCM	ring-closing metathesis
Suc	succinimidyl
TBAF	tetrabutylammonium fluoride
TBAI	tetrabutylammonium iodide
TBS	*t*-butyldimethylsilyl

TBTU	2-(1*H*-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium tetrafluoroborate
TDS	thexyldimethylsilyl
TES	triethylsilyl
TFA	trifluoroacetic acid
THP	tetrahydropyranyl
TMEDA	*N,N;N′,N′*-tetramethyl-1,2-ethylenediamine
TMS	trimethylsilyl

1 Introduction

1-C-Linked glycosyl compounds are carbohydrate analogs in which a carbon atom replaces the anomeric oxygen atom or nitrogen atom of glycosides (simple glycosides, oligosaccharides, nucleosides, *O*- or *N*-glycosyl amino acids and peptides, or glycolipids). In recent years, much effort has been devoted to developing efficient methods for their syntheses. Chap. 3.8 in Vol. 1 as well as numerous reviews [1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23] present the wide range of methodologies currently available.

Historically, the synthesis of *C*-glycosyl compounds was confined to the preparation of naturally occurring carbon-linked nucleosides [1,2,5] or *C*-glycosyl arenes [2,13]. More recently this field has expanded enormously into the construction of mimics of glycostructures that are relevant to glycobiology such as oligomers, glycopeptides, glycolipids, metabolic intermediates [11], or "clustered" structures of various kinds. The main idea is to create molecular species, stable to chemical and enzymic degradation, which conserve the configurational and conformational features of the naturally occurring parents. These analogs may serve as glycoenzyme regulators (glycosylhydrolases and transferases) or artificial ligands that can be useful in probing cellular interactions.

Due to space limitations, we have restricted this overview to the synthesis of carbon-linked mimics of natural oligosaccharides that contain the correct number of linking carbons. This, of course, does not presuppose that the chemistry or the potential usefulness of these mimics is more interesting than the ones associated with other types of *C*-linked analogs (*pseudo-C*-glycosyl analogs). Because of this necessary limitation, one should also realize that a good deal of important chemistry relevant to the field of functionalized tetrahydropyrans or tetrahydrofurans will not appear in this chapter, as for example, the precursory work of Vasella on the condensation of anomeric nitro sugars [24] or the chemistry developed for the total synthesis of a variety of natural products containing cyclic ether fragments. Four very representative examples are the total syntheses of palytoxin-carboxylic acid [25], tunicamycin V [26], brevetoxin B [27], or altohyrtin C [28] among many others. The *C*-glycosyl compounds, including C-dimers, associated with carbasugars or imino sugars ("aza" sugars) have been reviewed in two preceding chapters.

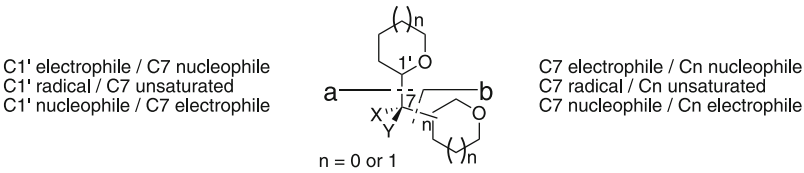
In the following, we have divided the discussion, mostly centered on synthetic approaches, according to the key carbon–carbon bond-forming steps (ionic or radical assemblage, olefin metathesis). Syntheses of oligosaccharide mimics in which one of the monosaccharide units is constructed by total synthesis or by cyclization of acyclic chiral chains are presented separately (partial de novo synthesis). Where appropriate, brief comments on the conformational and binding properties of these synthetic constructs will be specified, as well as the biological

behavior when tested. Finally, we have systematically used the chair conformational drawings of the tetrahydropyranyl ring systems although it is well recognized that a great many synthetic intermediates adopt a conformation very different from the chair. Most of the final deprotected forms of the C-analogs will however adopt a conformation identical or very close to the naturally occurring glycoside in the cyclic substructures.

2 The Ionic Assemblage

The desired stereochemical features of a mimic, e. g., a glycopyranosyl or a glycofuranosyl dimer, may be most attractively derived from the parent monomers if they are readily available. To add the missing carbon, all possible key assembling steps have been employed which are summarized in *Scheme 1*.

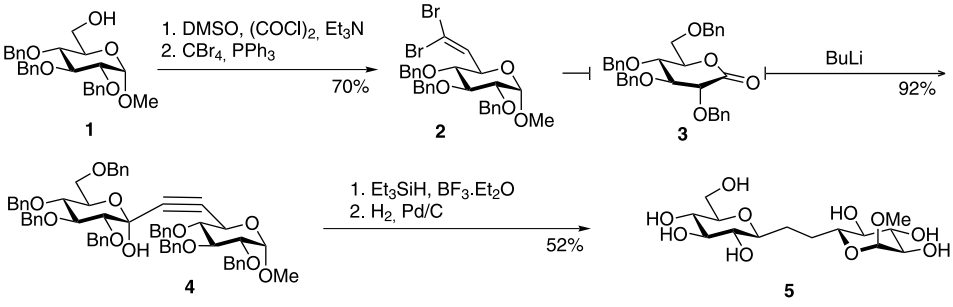
Disconnection **a** implies that a stereoselective carbon–carbon bond formation at the anomeric center of the "non-reducing" unit has to be identified (C1′ electrophile, radical, or nucleophile) using the appropriate "reducing" unit carrying the additional carbon. A simple, one-carbon *C*-glycosyl compound is rather utilized in disconnection **b**, having as a coupling "reducing" partner a nucleophilic, an electrophilic, or an unsaturated species.



Scheme 1

2.1 C5 (or C4) and C1 Alkynyl Anions

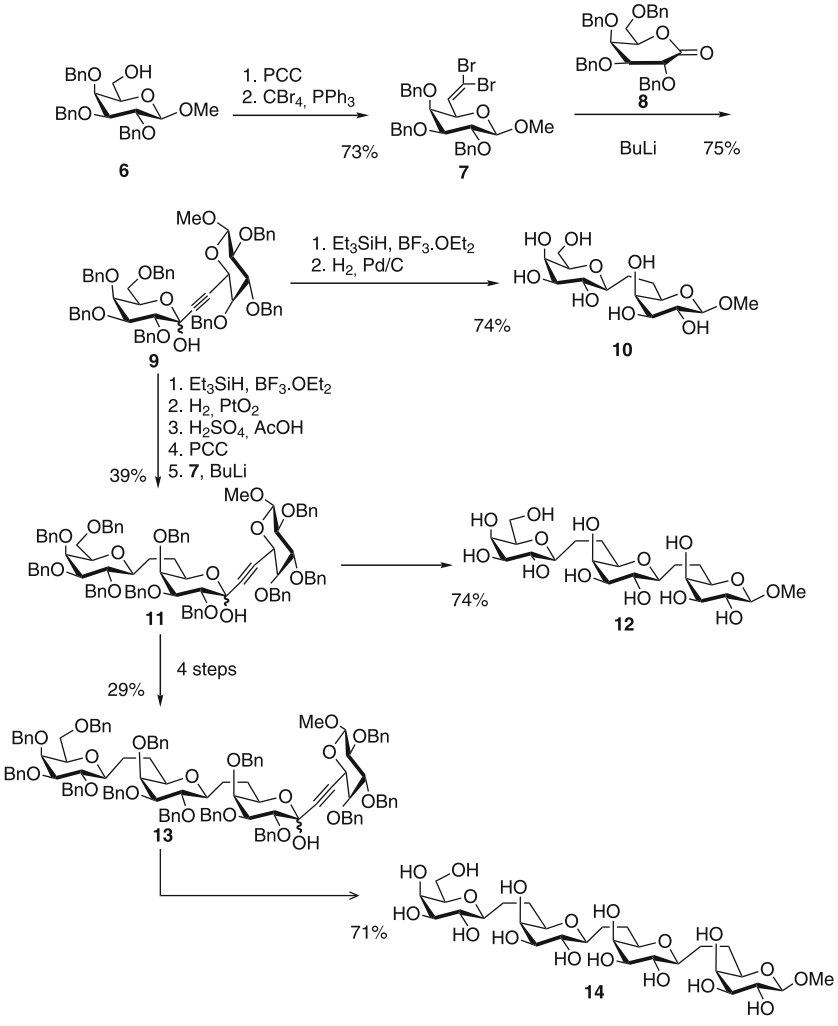
As briefly discussed above, every synthesis that will be described involves the addition of one linking carbon between two cyclic carbohydrate structures in order to replace the anomeric oxygen. The easiest route to accomplishing this task is to target a (1→6) linkage because the



Scheme 2

organometallic partner can be built up in a simple way and only one asymmetric center has to be created. It is then no surprise that the first synthesis of a *C*-analog of a natural disaccharide was reported by Rouzaud and Sinaÿ in 1983 following this route [29] (*Scheme 2*).

In their work they took advantage of the previously reported syntheses of *C*-alkenyl- and *C*-alkynyl-β-D-glycosides by Kishi [30] and Sinaÿ [31], using vinyl dibromide **2** easily prepared in two steps from glucose derivative **1**. Conversion of dibromide **2** to the acetylenic lithium reagent with BuLi, followed by treatment with glucopyranolactone **3** leads to hemiacetal **4** in a high yield. Stereospecific reduction of **4** by hydride delivery on the α-face of



Scheme 3

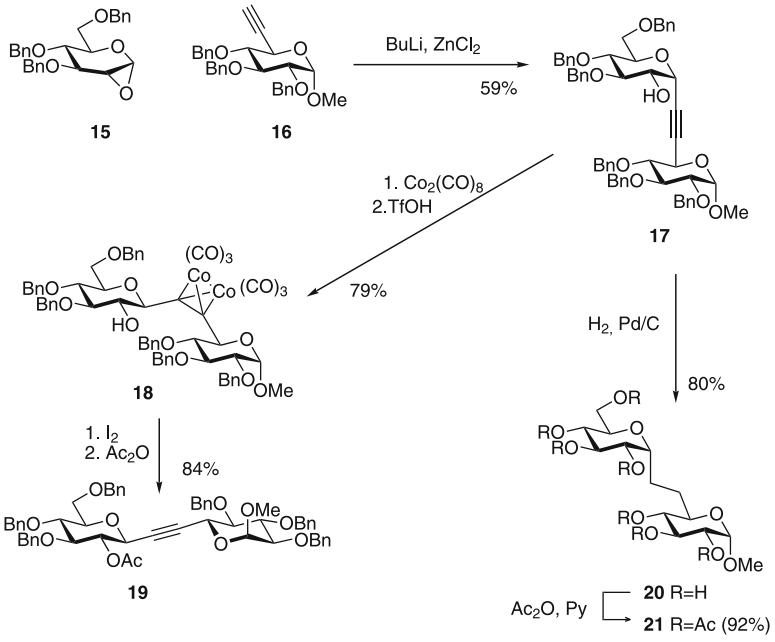
a cyclic oxycarbenium ion using the $Et_3SiH/BF_3{\cdot}Et_2O$ system affords the β-D-Glc(1→6)-D-Glc *C*-disaccharidic mimic **5** after catalytic hydrogenolysis. The efficiency of this approach led the Sinaÿ group to integrate this reaction sequence in an iterative process for the construction of the β-(1→6) *C*-oligomers **10**, **12**, and **14** through dimer **9**, trimer **11**, and tetramer **13** by an elongation through the reducing end [32] (*Scheme 3*).

These *C*-analogs show behaviors identical to the natural haptens in their binding to monoclonal antigalactan immunoglobulins [33], indicating the absence of hydrogen bonds in the carbohydrate-protein interaction involving the interunit oxygen atoms of the natural ligands.

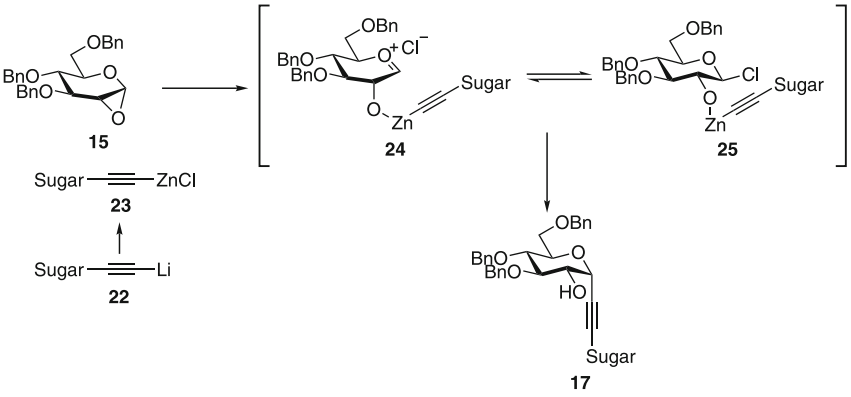
With the same acetylenic lithium reagent derived from the C6-alkynyl D-glucopyranose derivative **16**, van Boom and Sinaÿ developed a stereoselective $ZnCl_2$-assisted ring-opening of 1,2-epoxide **15**, establishing a ready access to *C*-isomaltoside **20** [34] by the route shown in *Scheme 4*.

This interesting result may be rationalized by transmetallation with $ZnCl_2$ of the lithium acetylide to **23** and formation of an alkynyl-zinc complex **24** with the 1,2-epoxide **15** (*Scheme 5*).

Intramolecular delivery of the alkynyl moiety in ion pair **24** or β-chloride **25** via the Zn tether would lead stereoselectively to α-*C*-glycosyl compound **17**. The Nicholas reaction [35] with triflic acid-induced isomerization [36] of the dicobaltohexacarbonyl complex of alkyne **17** provides cleanly the corresponding β-*C*-glucoside complex **18** (*Scheme 4*). Decomplexation with iodine and acetylation then gives β-*C*-disaccharide **19**.



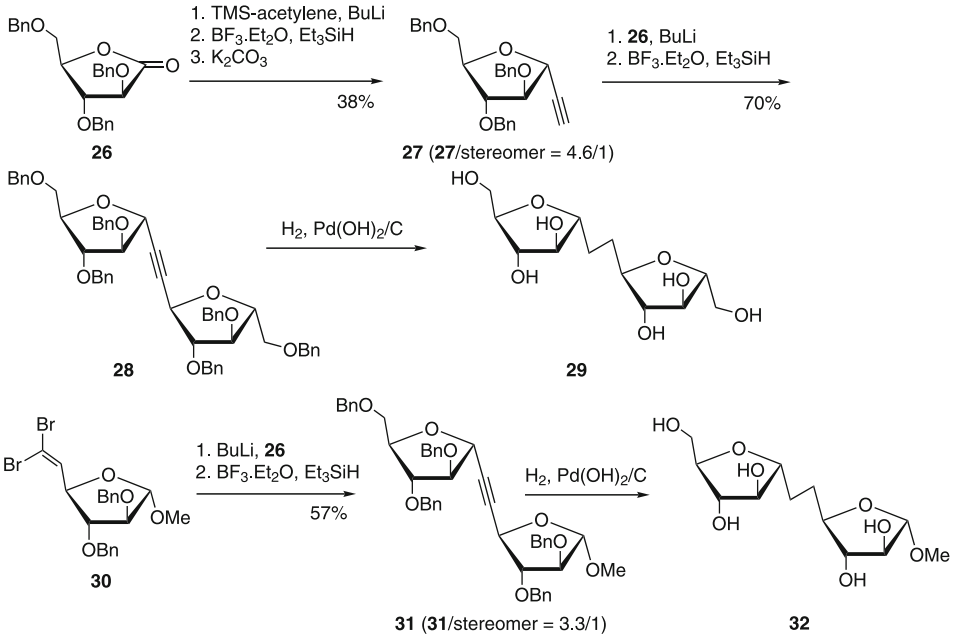
Scheme 4



Scheme 5

Different synthetic approaches to *C*-analogs related to the important α-D-arabinofuranosyl-(1→5)-α-D-arabinofuranose structural motif of mycobacterial arabinan will be discussed in this chapter. Routes involving acetylenic intermediates were reported by Wightman et al. [37] (*Scheme 6*).

The strategy is based on the pseudo C_2-symmetric feature of the target. Lactone **26** prepared from D-arabinose, serves as the precursor of acetylenic derivative **27** and as the substrate in the



Scheme 6

coupling reaction to furnish symmetric dimer **29** after selective silane reduction and protecting group removal. The other alternative in preparing methyl glycoside **32** is to start from dibromoalkene **30**, available from methyl α-D-arabinofuranoside. Reaction of the lithio derivative with lactone **26** and reduction with $Et_3SiH/BF_3{\cdot}Et_2O$ provide acetylenic derivative **31** as the major product, transformed to *C*-disaccharidic mimic **32** after catalytic hydrogenolysis and triple bond reduction.

2.2 C5 (or C4) and C1 Methylenephosphonium Ylides

Another route to generate (1→6) linkages, reported by Dondoni in 1996 [38], is to exploit the Wittig olefination between 1-*C*-glycopyrano(or furano)syl formaldehydes, available through the thiazole-based formylation of glyconolactones [39], and phosphonium ylides at position 6 of glycopyranosides. A representative example is the construction of the *C*-analog of the β-D-Gal(1→6)-D-Gal disaccharide (*Scheme 7*).

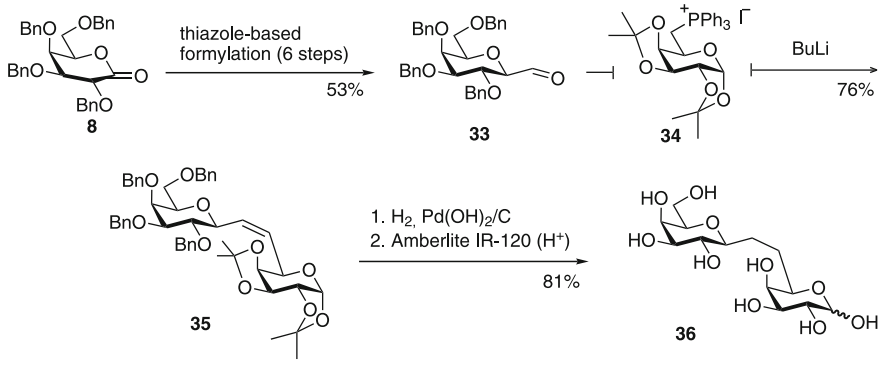
Condensation of the phosphonium ylide derived from the galactose phosphonium salt **34** and aldehyde **33** affords alkene **35** which, by reduction and removal of protecting groups, leads to the final *C*-dimer **36**.

A similar sequence of reactions starting from aldehyde **37** provides efficiently the *C*-pentamer **40** in a 33% overall yield from **37** [40] (*Scheme 8*).

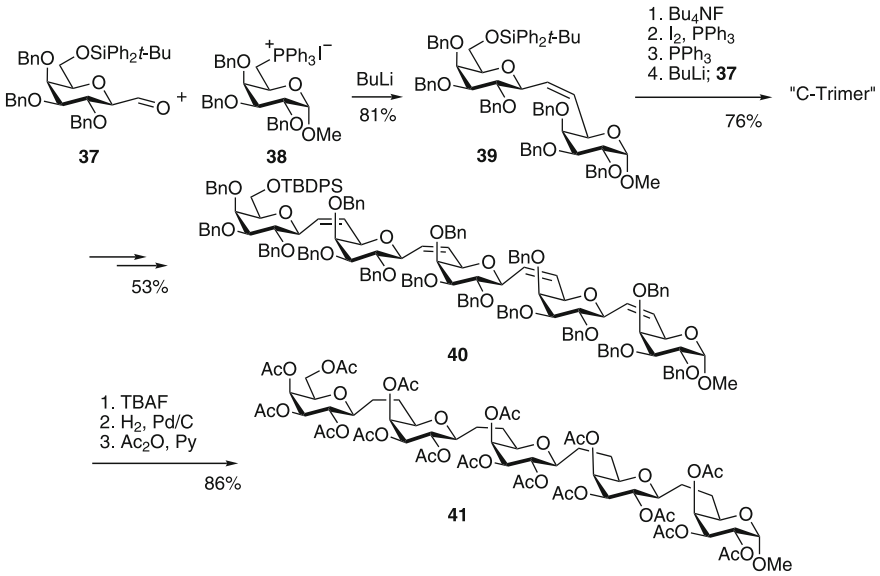
Protecting group removal and hydrogenation of the carbon–carbon double bonds furnish a *C*-mimic of the β-(1→6) pentagalactoside, characterized as the corresponding *O*-acetyl derivative **41**.

A set of ten carbon-linked (1→6) dimers was prepared in this way, starting from the *C*-glycosyl formaldehydes **33** and **42–47** and the phosphonium salts **34** and **48** [41] (*Scheme 9*).

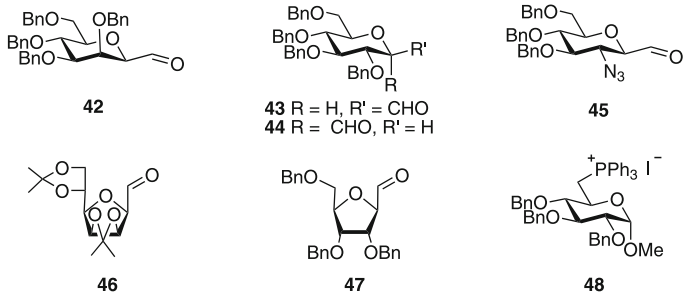
This route requires, however, the use of the thermodynamically more stable aldehydes; α-*C*-glucosyl precursor **44** will provide, for example, the corresponding β-*C*-dimer by isomerization under the basic conditions of the Wittig reaction. An exception is the use of aldehyde **50** obtained selectively from the corresponding thiazolyl α-*C*-glycoside **49**, provid-



Scheme 7



Scheme 8



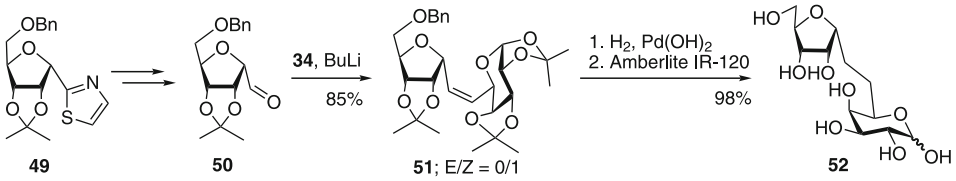
Scheme 9

ed by deoxygenation of the thiazolylketose acetate using TMSOTf-triethylsilane [42]. Having such a compound, α-(1→6)-*C*-disaccharide **52** was then synthesized (*Scheme 10*).

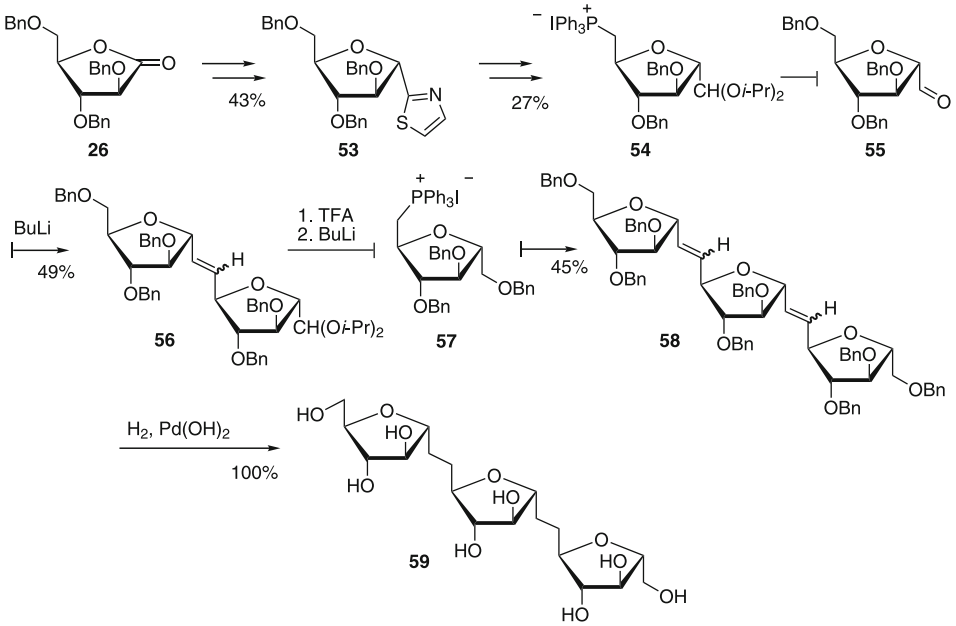
Aldehyde **50** is stable in the course of the Wittig olefination with phosphonium salt **34** and provides only the *Z*-alkene **51**. Reduction of the double bond and benzyl group removal, followed by an acidic treatment then afford α-(1→6)-*C*-disaccharide **52**.

C-analogs of the α-(1→5)-arabinofuranosyl repeating motifs found in the polysaccharidic chains of mycobacterial cell wall were also prepared using this approach by Dondoni and Marra [43]. This route is based on the use of a bifunctional unit **54** with regeneration of the formyl group in one step after the Wittig coupling reaction (*Scheme 11*).

Both partners **54** and **55** of the Wittig reaction are derived from the same thiazolyl intermediate **53**, itself prepared following the thiazole-based formylation of lactone **26**. The resulting



Scheme 10

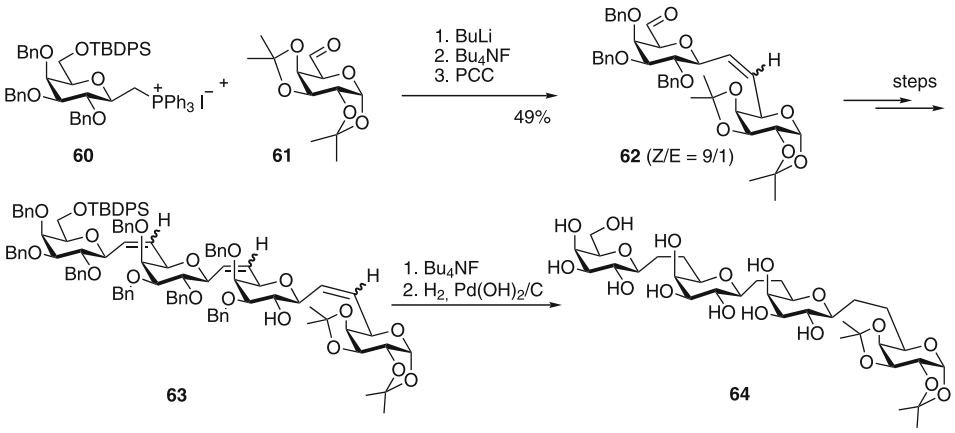


Scheme 11

alkenes **56** are then subjected to acidic treatment leading to the corresponding aldehydes, electrophilic partners in a second Wittig reaction. Hydrogenation of the double bonds provides the *C*-analog of α-(1→5)-*C*-arabinofuranotriose **59**. In this synthesis the α-configuration is also maintained during the elongation of the chain.

The reversed strategy has also been reported by condensation of C6 aldehydo sugars with β-linked methylenephosphonium ylides [44] (*Scheme 12*).

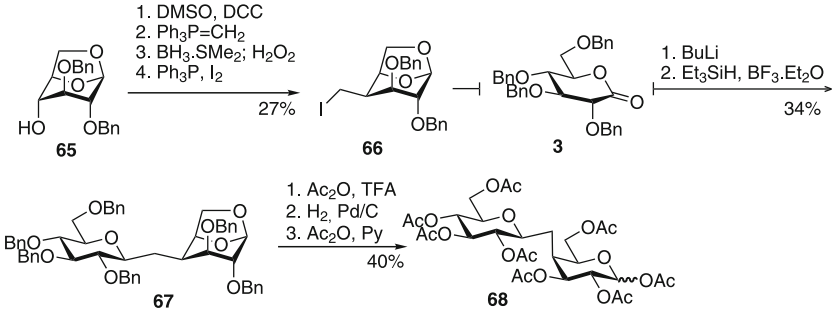
Condensation of the D-galactopyranosyl-derived phosphonium salt **60** and aldehyde **61** provides the unsaturated dimer **62** after desilylation and oxidation. Iteration of the sequence twice with phosphonium salt **60** leads to tetramer **63** with, however, low coupling efficiencies (36% for the trimer and 11% for the tetramer) mostly due to the base-induced α,β-unsaturation of the aldehydic substrates. Protecting group removal and hydrogenation of the carbon–carbon double bonds furnish tetramer **64**, a *C*-mimetic of the β-(1→6) tetragalactoside similar to the one shown in *Scheme 3*.



Scheme 12

2.3 C4 Lithiomethyl Sugars

Sequences similar to the ones detailed in *Sect. 2.1* in which the anionic carbon is present at positions other than C6 have also been described such as the C4-branched lithium species derived from iodide **66** reported by Preuss and Schmidt [45] (*Scheme 13*).
Stereoselective hydroboration-oxidation of the olefin obtained in two steps from the 1,6-anhydroglucose **65** followed by iodination and lithium-halogen exchange generates a lithio reagent that adds to gluconolactone **3** to give a hemiketal, which is stereoselectively reduced to the *C*-linked disaccharide **67**. Opening of the 1,6-anhydro ring and deprotection provides the *C*-linked β-D-Glc(1→4)-D-Gal disaccharide **68**, after acetylation. The same sequence using galactonolactone **8** leads to the *C*-linked β-D-Gal(1→4)-D-Gal dimer.

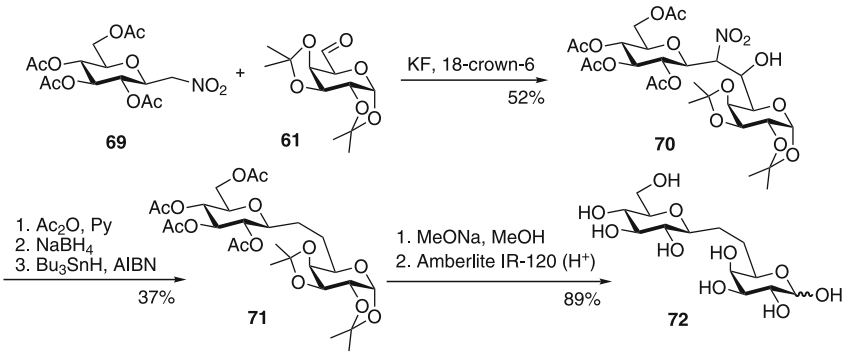


Scheme 13

2.4 C1 Nitromethyl Sugars

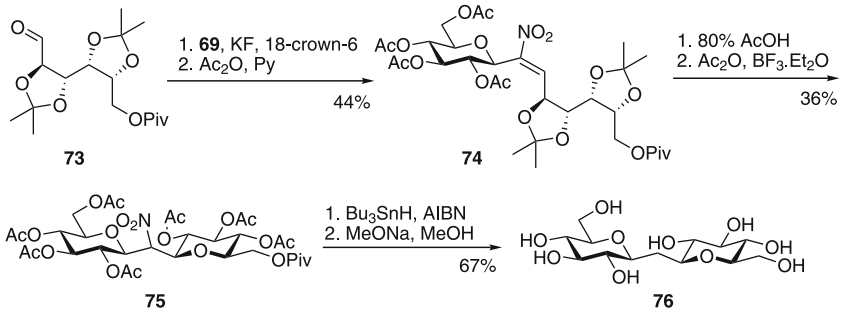
One way to generate *C*-linked dimers is to start from a preformed *C*-linked monomer in which the "missing" carbon is carried as either the electrophilic partner (for example, C1-aldehydo sugars, see *Sect. 2.2*, or the nucleophilic partner. In the latter case β-elimination must be controlled. The nitronate anions derived from *C*-glycosyl nitromethanes are species of the latter kind, and Martin and Lai [46] reported a short preparation of the *C*-analogs of β-D-Glc(1→6)-D-Gal and non-natural β,β-trehalose [β-D-Glc(1→1)-β-D-Glc] using as a key assembling step, a nitroaldol condensation (Henry reaction), a well-known method for chain extension of carbohydrates [24,47]. Fluoride-mediated condensation of *C*-glucosyl nitromethane **69**, obtained in two steps from glucose, with the D-galactose-derived aldehyde **61**, provides nitroaldols **70**. Acetylation-elimination to nitroalkene, reduction of the double bond, and radical denitration readily afford the protected *C*-linked disaccharide **71**, deprotected in two steps to the *C*-linked β-D-Glc(1→6)-D-Gal mimetic **72** (*Scheme 14*).

Extension of the same methodology to the aldehydo-glucose derivative **73** yields, after the elimination step, the nitroalkene **74** (*Scheme 15*).



69 **61** **70** **71** **72**

Scheme 14



73 **74** **75** **76**

Scheme 15

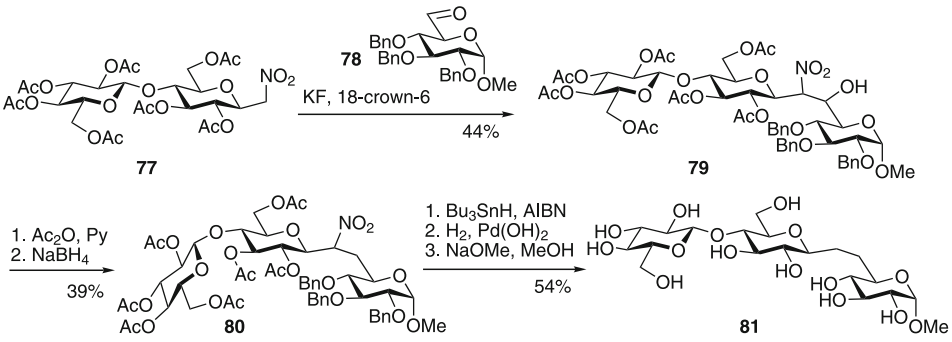
Upon removal of the isopropylidene groups, Michael addition occurs providing the *C*-linked disaccharide **75** after acetylation, together with the tetrahydrofuranyl isomer (25% yield). Radical denitration and deacylation afford the *C*-linked β,β-trehalose **76**.

The same fluoride condensation was used for the synthesis of a mixed *O*,*C*-trisaccharide analog of methyl β-D-glucopyranosyl-(1→4)-β-D-glucopyranosyl-(1→6)-α-D-glucopyranoside [48] (*Scheme 16*).

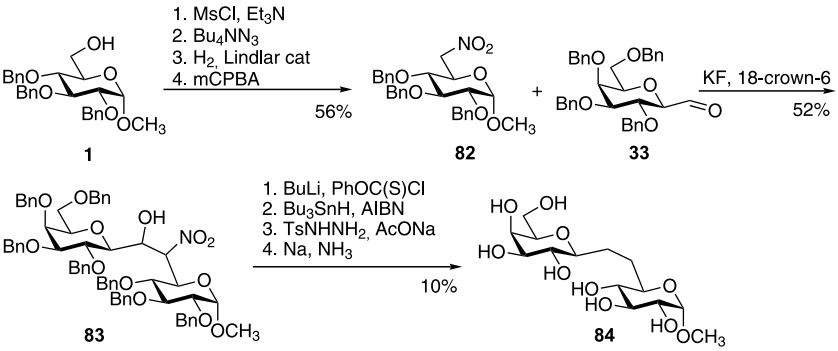
Reaction of the nitromethyl derivative **77** obtained in four steps from cellobiose peracetate, with the glucose-derived aldehyde **78** and potassium fluoride/18-crown-6 leads to the expected condensation product **79**. The acetylation-elimination-reduction sequence (→**80**) followed by reductive denitration and total deprotection affords the *O*,*C*-trimer mimic **81**.

A reverse option of the nitro-aldol condensation is available, as in the synthesis of the *C*-analog of β-D-Gal-(1→6)-D-Glc, also called allolactose, viewed by the authors as a potential inducer of the Lac repressor protein that is not susceptible to β-galactosidase cleavage [49] (*Scheme 17*).

The nitronate anion derived from nitro sugar **82**, available from glucopyranoside **1** using standard methods, condenses with aldehyde **33** to afford adduct **83**. Reduction with diimide of the *E*-olefin, produced in low yield by radical elimination of the nitroaldol, and removal of the



Scheme 16

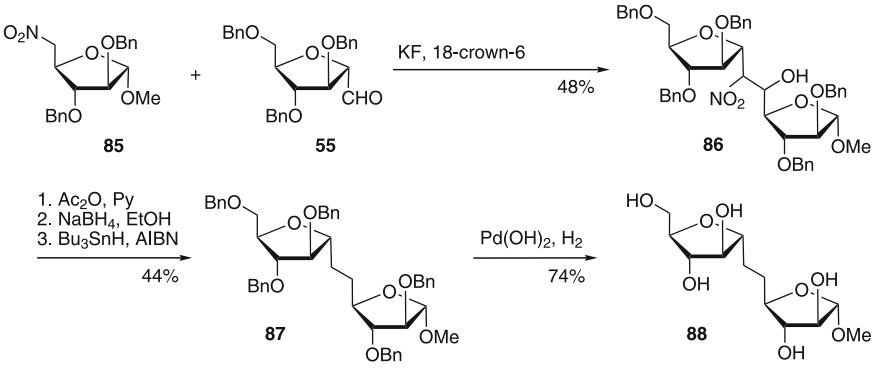


Scheme 17

benzyl groups provide the *C*-analog **84** of methyl allolactoside. 6-Deoxy-6-nitro sugars are also good substrates in intramolecular radical reactions for the preparation of *C*-linked dimers (see *Sect. 3.1*).

Using a similar approach, Gurjar described the synthesis of *C*-analog **88** of the disaccharide α-ara*f*-(1→5)-ara*f* present in motif C of the arabino-galactan portion of *Mycobacterium tuberculosis* [50] (*Scheme 18*).

Fluoride-induced condensation of the 5-deoxy-5-nitro furanoside **85** with the α-*C*-furanosyl formaldehyde **55** obtained by ring contraction of glucosamine and subsequent functional group manipulation lead to a diastereomeric mixture of the nitroaldols **86**. An acetylation-elimination sequence followed by reduction of the alkene, denitration and hydrogenolysis provide the *C*-disaccharide **88**.

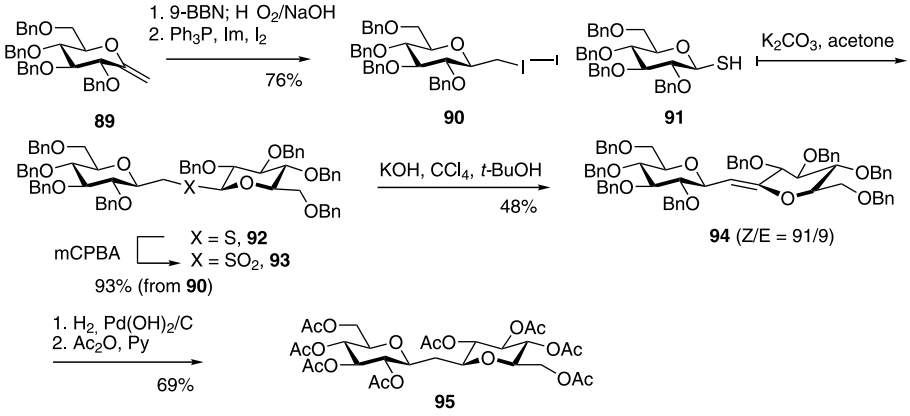


Scheme 18

2.5 The Ramberg–Bäcklund Route

The groups of Taylor [51a] and Franck [51b] independently reported in 1998 the interesting use of the Ramberg–Bäcklund rearrangement for preparing substituted exo-glycals (1-exomethylene sugars) from anomeric sulfones. In this reaction, α-halogenated sulfones are transformed to regio-defined olefins on treatment with base. The potential of this approach in synthetic applications has been reviewed [51c]. This procedure for preparing *C*-linked disaccharides has mainly been developed by the Taylor group. They took advantage of their initial studies on the exo-glycal preparation using this rearrangement [51a,52,53]. In their strategy the two monosaccharides are coupled through a thioetherification reaction and the resulting *S*-linked disaccharide is oxidized to the required sulfone for the Ramberg–Bäcklund rearrangement. This was successfully applied to the synthesis of a carba analog of isotrehalose (β-D-glucopyranosyl-(1→1)-β-D-glucopyranose), a disaccharide in which residues are linked through their anomeric center [54,55] (*Scheme 19*).

Exoglycal **89**, obtained by either methylenation of the gluconolactone using (Cp_2TiMe_2) [56] or the Ramberg–Bäcklund rearrangement of the anomeric methylsulfone [51a], is converted to iodomethylglucoside **90** in a two-step procedure: stereoselective hydroboration-oxida-

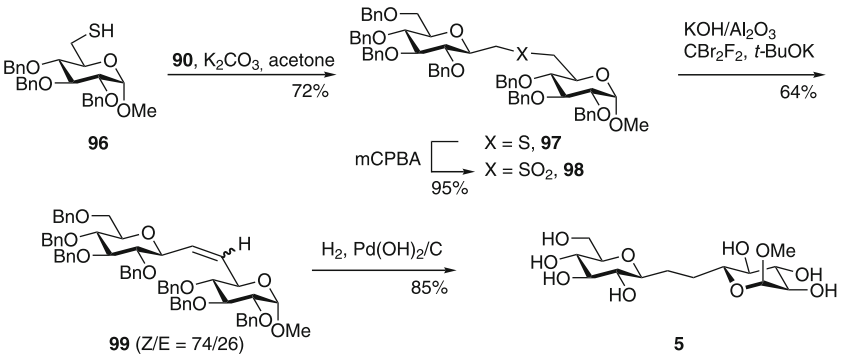


Scheme 19

tion and iodination using the Garegg protocol. Thioetherification with the protected thioglucose **91**, followed by oxidation of the resulting sulfide with mCPBA, provides sulfone **93** in 93% yield. The Ramberg–Bäcklund rearrangement under the Meyer's conditions [57] affords enol ether **94** predominantly as the Z-isomer (Z/E = 91/9). Hydrogenolysis-hydrogenation in a methanol/ethanol mixture and acetylation provide selectively the β,β-*C*-dimer **95**.

The same approach was used for the synthesis of the β-(1→6) *C*-disaccharide **5**, analog of methylgentiobioside [54,55] (*Scheme 20*).

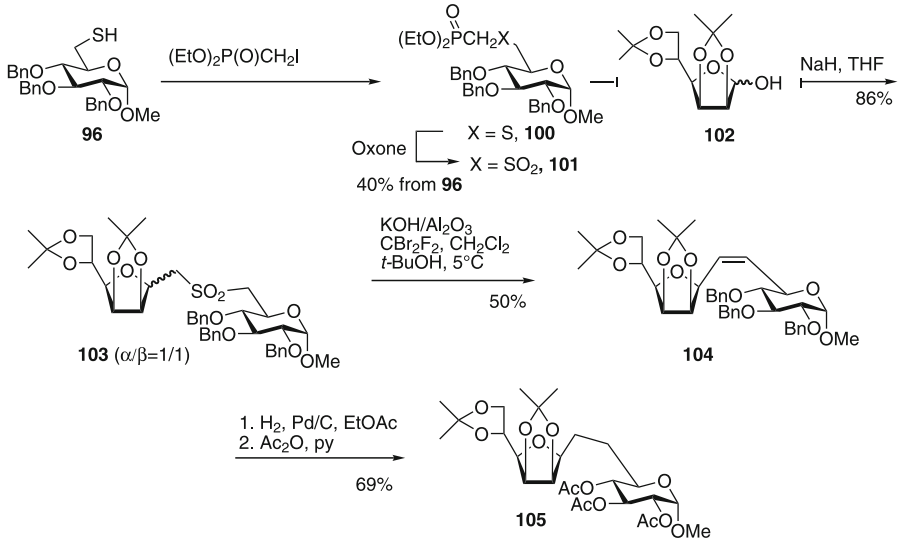
The thioetherification-oxidation sequence from thiol **96** delivers sulfone **98** in 68% yield. The sulfone functionality of this substrate is centrally positioned and displays, on both sides, a methylene linked to a monosaccharide. The Ramberg–Bäcklund rearrangement under the Chan's halogenation conditions [58] (KOH/Al_2O_3 and CBr_2F_2) provides the unsaturated *C*-disaccharide **99** in 64% yield as a mixture of isomers (Z/E = 70/30). Hydrogenation-hydrogenolysis finally gives the methyl-*C*-gentiobioside **5**.



Scheme 20

Taylor's group also showed that the Horner–Wadsworth–Emmons (HWE)/conjugate-addition sequence can be an efficient alternative to the thioetherification-oxidation reactions for the preparation of the required sulfones [59]. This was illustrated in the synthesis of the β-(1→6)-mannofuranose-α-Glu *C*-dimer [60] (Scheme 21).

Alkylation of thiol **96** with diethyliodomethylphosphonate followed by oxidation produces the HWE reagent **101**. The HWE-conjugate addition sequence with diisopropylidene mannofuranose **102** proceeds in 86% yield to sulfone **103** as a α/β mixture (1/1). The Ramberg–Bäcklund rearrangement under Chan's conditions applied to this mixture provides the *E*-alkene **104** in 50% yield, with total β selectivity. Hydrogenolysis and concomitant hydrogenation of the alkene give, after acetylation, *C*-disaccharide **105** in 69% yield.

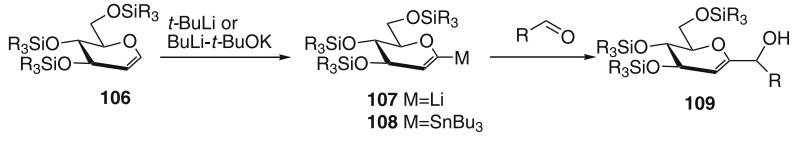


Scheme 21

2.6 C1-Lithiated Glycals

Three independent reports in 1986 showed that glycals are deprotonated by *tert*-butyllithium [61,62] or butyllithium/potassium *tert*-butoxide [63] when silyl ethers are used as protecting groups (Scheme 22).

Vinylic lithium reagents **107** or the corresponding 1-tributylstannyl glycals **108** have provided useful routes to C1-substituted glycals and *C*-glycosyl compounds [10]. The procedure has



Scheme 22

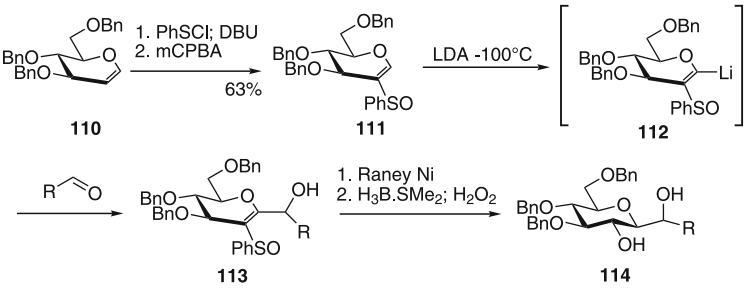
not yet been applied, however, to the synthesis of true mimics of oligosaccharides. A phenylsulfinyl substituent at C2, as shown by Schmidt and coworkers, facilitates deprotonation at C1 of glycals so that LDA is a base strong enough for C1-lithiation [64] (Scheme 23).

The 2-phenylsulfinyl glucal **111** is available from glucal **110** by addition of phenylsulfenyl chloride, DBU elimination and oxidation. The C1-substituted glycal **113**, obtained by addition of the lithium reagent **112** to aldehydes, is transformed to a β-D-*C*-glucosyl derivative **114** by reductive removal of the phenylsulfinyl group with Raney nickel and stereoselective hydration of the C/C double bond by hydroboration-oxidation.

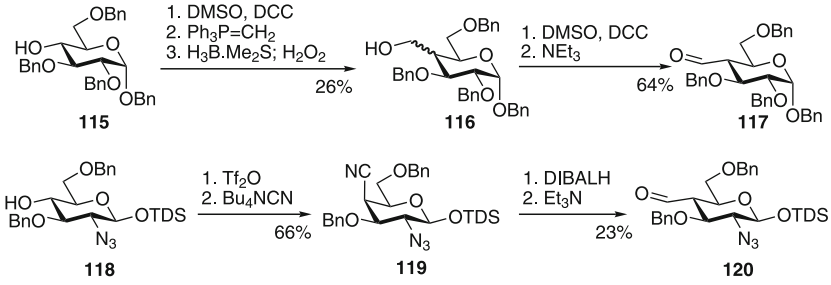
The linking carbon in the aldehydo-sugar partner is usually introduced by a Wittig reaction (or Tebbe's reaction) using a sequence of reactions starting from the appropriate glycoside as exemplified by glucoside **117** [65,66] (Scheme 24).

When this route is inefficient, as with the 2-azido-2-deoxy-D-gluco derivative **118**, the hydroxy group may be replaced by a cyano group (as in **119**), which is reduced by a diisobutylaluminum hydride treatment and isomerized with triethylamine to produce the C4-formyl glucoside **120**. Similar one-carbon homologation of uloses has also been reported by Armstrong [68] and Vasella [69].

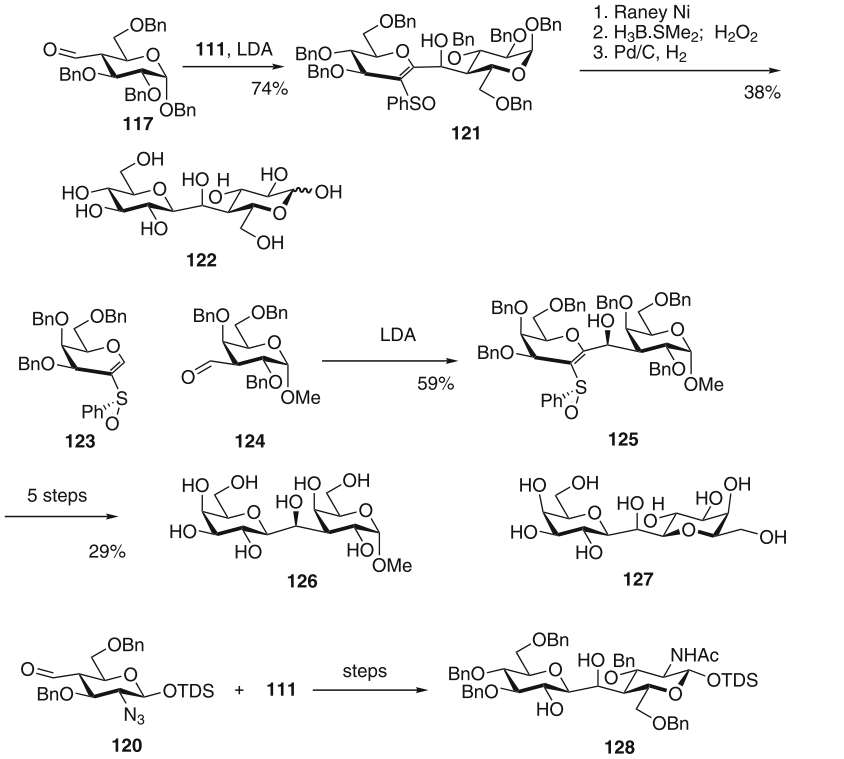
Several hydroxymethylene *C*-linked disaccharides have been prepared using this approach including the *C*-linked mimics of cellobiose **122** [65], β-D-Gal-(1→3)-D-Gal **126** [66], "galacto-β,β-trehalose" [β-D-Gal-(1→1)-β-D-Gal] **127** [70], and *N*-acetyllactosamine **128** [67] (Scheme 25).



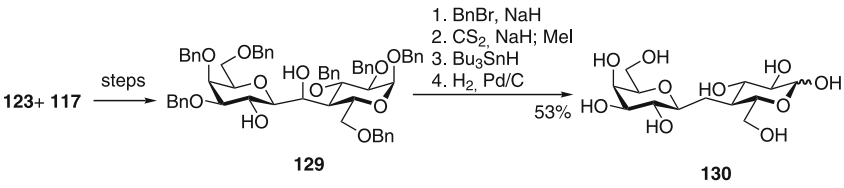
Scheme 23



Scheme 24



Scheme 25



Scheme 26

There is one report of the hydroxymethylene bridge being deoxygenated using the Barton procedure, leading to the *C*-analog of lactose **130** [71] (*Scheme 26*).

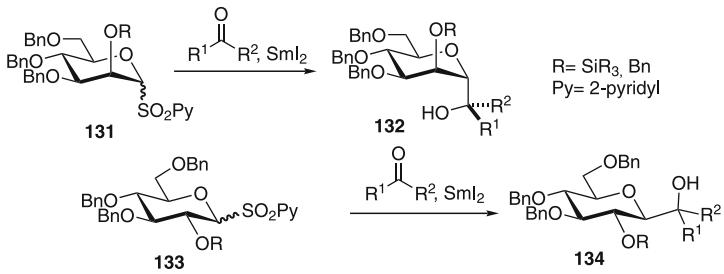
2.7 Glycosyl Samarium Reagents or Related Species

Tetrahedral anomeric lithium reagents are [10], initially restricted to the 2-deoxy sugar series in order to avoid the facile β-elimination of the substituent at position 2, is usable if a protective metallation at the C2-substituent is carried out before the metallation event at C1 by reductive lithiation or transmetallation of the corresponding stannanes [10]. This process has

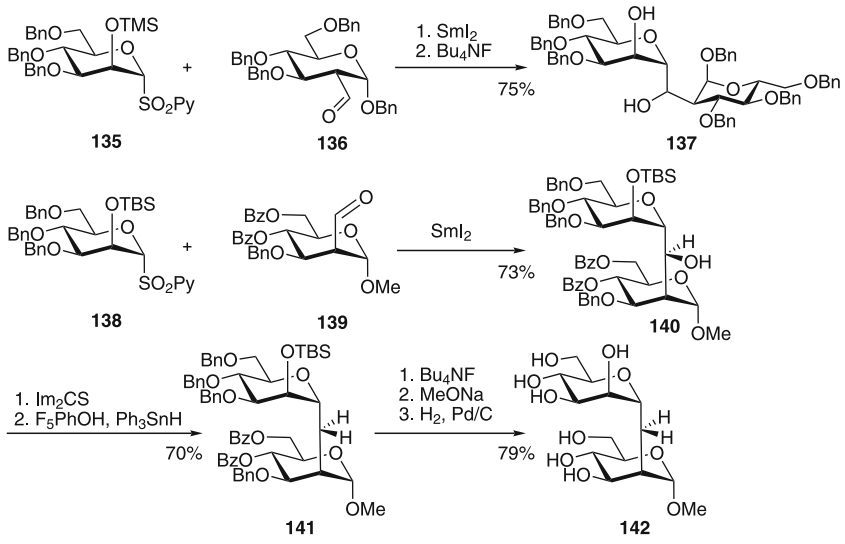
not been applied to the construction of *C*-analogs of disaccharides. An unexpected solution to this problem reported by Skrydstrup and Beau in 1995 [72] showed that reductive samariation (SmI_2, THF without an additive) of manno (and gluco) pyranosyl 2-pyridyl sulfones **131** and **133**, respectively, produces stereospecifically and instantaneously, in the presence of carbonyl compounds under Barbier conditions, the corresponding 1,2-*trans*-*C*-glycosyl compounds in good (*manno* series, **132**) to fair (*gluco* series, **134**) yields (*Scheme 27*).

This mild and simple procedure has been extended to the synthesis of *C*-linked disaccharides such as the *C*-linked mimic **137** of the α-D-mannopyranosyl(1→2)-D-glucopyranoside from sulfone **135** and aldehyde **136** [72,73] and the methyl α(1→2)-*C*-mannobioside **142** from sulfone **138** and aldehyde **139** [74,75] (*Scheme 28*).

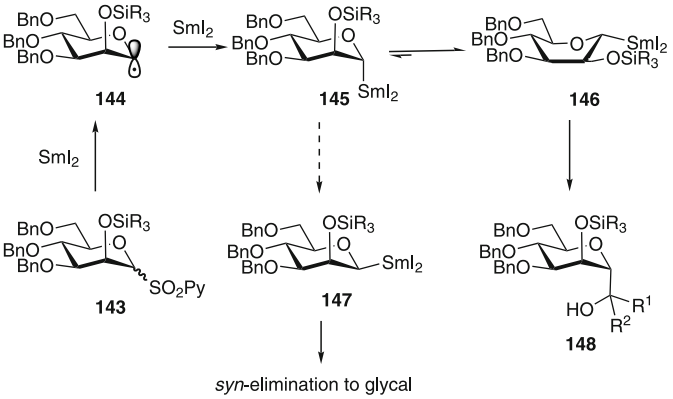
The production of a kinetic α-organosamarium(III) species **145** by reductive samariation of the corresponding anomeric radical **144** explains the stereospecific formation of α-*C*-mannosyl compounds (*Scheme 29*).



Scheme 27



Scheme 28

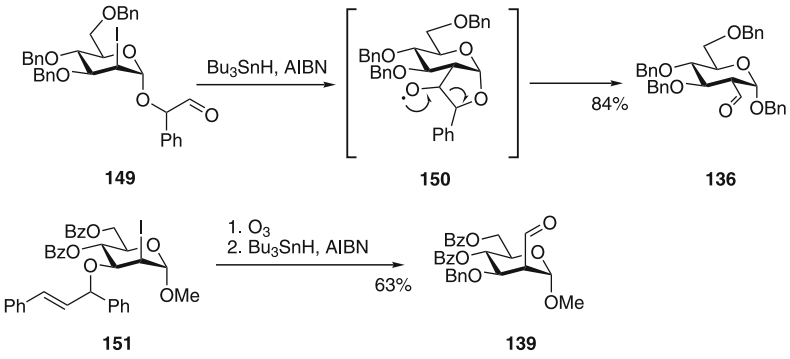


Scheme 29

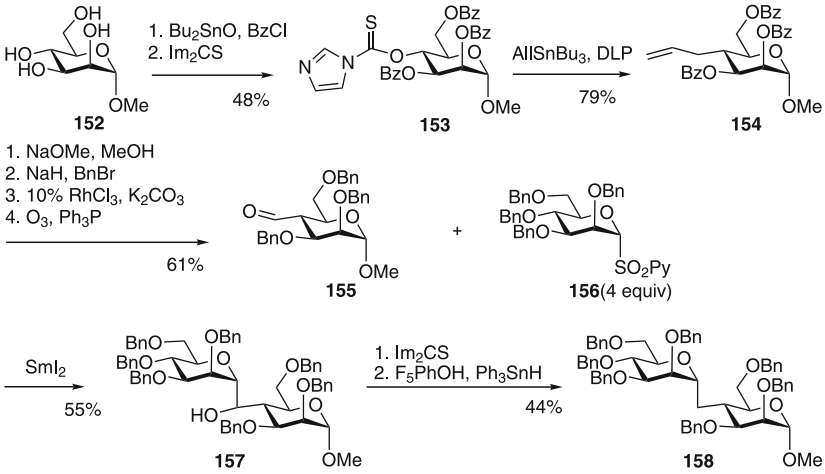
Rather than isomerization to β-samarium species **147**, reagent **145** probably undergoes a facile conformational change in which the C1 and C2 substituents are placed in more stable equatorial orientations. The isomerized species **147**, if any, apparently undergoes a *syn* β-elimination at room temperature. The synthesis of aldehydes **136** and **139** merits a special comment because a versatile formyl group transfer originally proposed by Jung and Choe [76] is used (*Scheme 30*).

The authors showed that iodo-glycoside **149** undergoes, under Bu_3SnH treatment, a 5-*exo* radical cyclization-fragmentation to the more stable benzyl radical resulting in a stereospecific formyl group transfer to C2, leaving a standard benzyl protecting group at the anomeric position. Delivery of the formyl group from the C3 position, as with iodo glycoside **151**, produces rather the isomeric axial C2-formyl manno derivative **139** as shown by Skrydstrup and Beau [74,75].

For the assembly of the *C*-linked disaccharide α-D-Man-(1→4)-D-Man, Mikkelsen and Skrydstrup followed the same route using this SmI_2-mediated coupling of pyridyl sulfone **156** with



Scheme 30



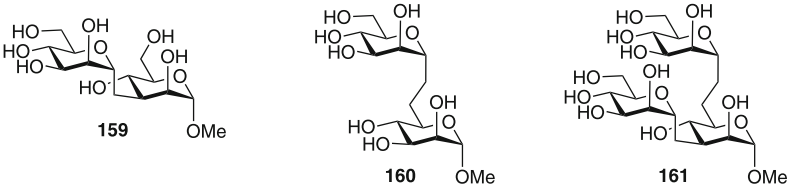
Scheme 31

aldehyde **155**, affording the *C*-disaccharide derivative **157** with complete stereocontrol at the two new stereogenic centers (*Scheme 31*) [77].

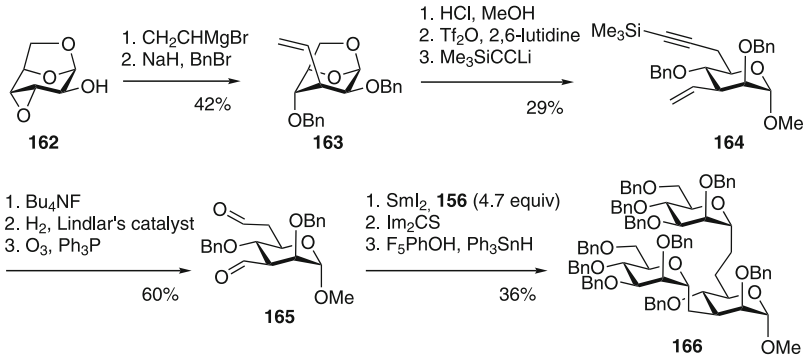
In this synthesis, the C4-formyl monosaccharide **155** was prepared from the benzoylated thionocarbamate **153** using a stereoselective radical based allylation. The exclusive equatorial selectivity observed in the *manno* series is remarkable and can probably be explained by the unfavorable axial approach of the incoming allyl group due to steric interactions with the axially oriented C2-substituent. Exchange of the protecting groups, rhodium-catalyzed double bond migration and reductive ozonolysis then produced the aldehyde **155**.

Methyl α-1,3- and α-1,6-*C*-mannobiosides **159** and **160**, as well as the branched *C*-trimannose analog **161** of the common core structure of asparagine-linked oligosaccharides have also been synthesized using the same sulfone **156** and, for the *C*-trimer, the monosaccharide dialdehyde **165** (*Scheme 32* and *Scheme 33*) [78,79,80].

The conformational preferences of these *C*-linked dimers and trimer have been analyzed by comparing with their natural counterparts, using a combination of NMR spectroscopy and time-averaged restrained molecular dynamics [81]. It was notably found that a major conformational distinction between the natural trisaccharide and the glycomimetics concerns the ω-(1,6) torsion angle around the α-(1,6)-linkage, with a unique *gt* rotamer for the *C*-trimer **161**. The *C*-glycosyl trimer analog is recognized by three mannose-binding lectins, as shown



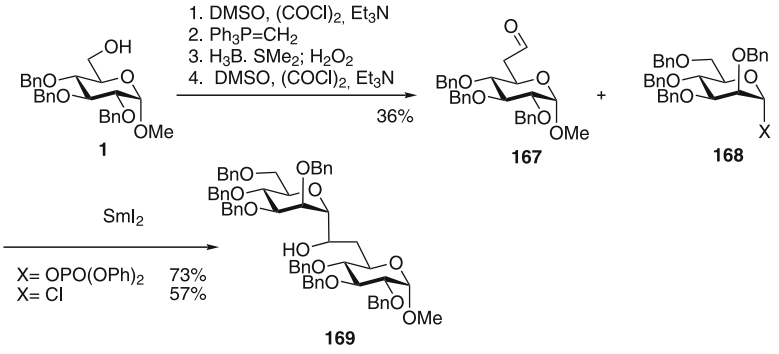
Scheme 32



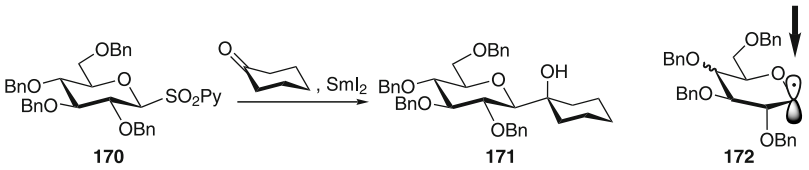
Scheme 33

by NMR (line broadening, TR-NOE, and STD) and surface plasmon resonance (SPR) methods, in a binding mode similar to the one whereby the lectins recognize the natural compound. Anomeric chlorides or phosphates may also be used as reducible groups in the reaction as shown in the synthesis of the $\alpha(1\rightarrow2)$ *C*-linked disaccharide **169** of mannose [82,83] (*Scheme 34*).

For the synthesis of 1,2-trans *diequatorial* *C*-glycosyl compounds (e. g., with the glucopyranosyl series), this *C*-glycosylation procedure was not efficient enough to be used in the construction of complex *C*-oligomers. The major drawback was a too high level of the competing β-elimination (an elimination/C–C bond formation ratio of approximately 1/1) that could be minimized only by using a bulky protecting group at O2 (R in **133** = *t*-butyldimethylsilyl, *Scheme 27*) [72,73]. An unpredicted solution was reported in 2000 by incorporating catalytic amounts of nickel(II) iodide (1 mol%) with the reducing samarium salt [84]. Catalysis by transition metals of some reactions mediated by samarium diiodide was introduced by Namy and Kagan [85]. This dramatic effect of catalytic nickel is seen in the high coupling efficiency of sulfone **170** and cyclohexanone at 0 °C with a solution of SmI_2 containing 1 mol% of nickel(II) iodide (*Scheme 35*).



Scheme 34



Scheme 35

These results suggest a possible mechanism in which, during the samariation step, the intermediate anomeric radical **172** is primarily reduced under the control of steric factors by a bulky catalytic species containing low valent nickel.

This procedure provides a fast synthesis of *C*-glycosyl disaccharides with the anomeric sulfones derived from the D-gluco, D-galacto and L-fuco series and this is presented using aldehyde **175**, obtained by the five-step sequence of reactions shown in *Scheme 36*.

The *C*-linked $\beta(1\rightarrow6)$-disaccharides **176**, **178** and **180** are all obtained in good yields (83 to 89%). The addition of the organosamarium reagents to the carbonyl group of aldehyde **175** also occurred with a high facial selectivity (diastereomeric ratio of about 95:5), exclusively controlled by the asymmetry of the *C*-glycosyl donor.

This chemistry is also possible with per-*O*-trimethylsilyl glycosyl iodides [86] or, if a multistep construction of more complex *C*-glycosyl compounds is considered, with the more practical per-*O*-benzyl iodides derived from D-glucose, D-galactose or L-fucose [87] (*Scheme 37*).

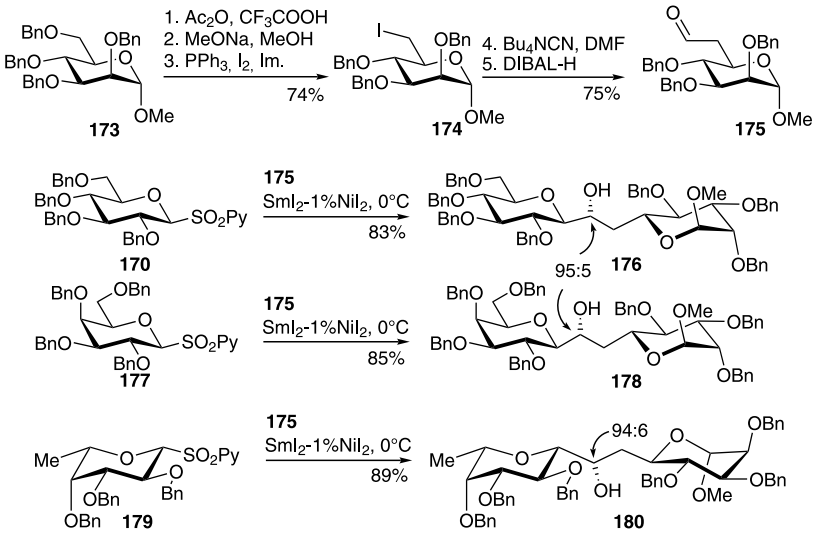
For example, reductive samariation of benzyl glycopyranosyl iodide **182** in the presence of aldehyde **175** provides the corresponding 1,2-*trans* *C*-glycosyl disaccharide **176**, transformed to the methylene-linked dimer **183**. A striking difference from the results obtained with the anomeric 2-pyridyl sulfones is the decrease or the absence of the β-elimination reaction when using SmI_2 alone, with the efficiency marginally improved by incorporating catalytic amounts of NiI_2 with SmI_2. This is not yet explained but it is possible that there is a change in the electron transfer mechanism ongoing from anomeric pyridyl sulfones to anomeric iodides inducing a change in the product distribution (C–C bond formation versus elimination).

This SmI_2-induced *C*-glycosylation was extended to the synthesis of *C*-linked glycosides of 2-acetamido-2-deoxy sugars [88,89,90]. Reductive samariation of the 2-pyridyl sulfone of *N*-acetylglucosamine **184** in the presence of aldehyde **167** leads selectively to the α-*C*-glycosyl dimer **185** [90] (*Scheme 38*).

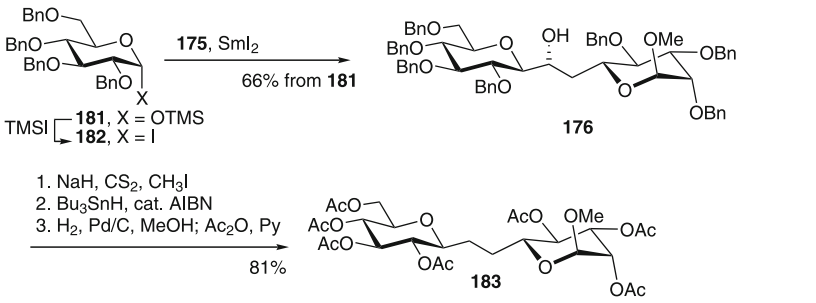
The transformation is only moderately stereoselective in this situation and the stereochemical outcome of the reaction is explained by a complexation of the samarium atom by the acetamido group in the kinetic α-Sm(III) species **186**, slowing down the isomerization to the β-species. The anomeric selectivity is completely lost when operating with the acetylated sulfone **187**.

A solution to restore a selective reaction and to access to *C*-glycosides mimicking the biologically ubiquitous β-GlcNAc motif, relies on this intermolecular samarium-Barbier reaction followed by an oxidation-isomerization sequence [91] (*Scheme 39*).

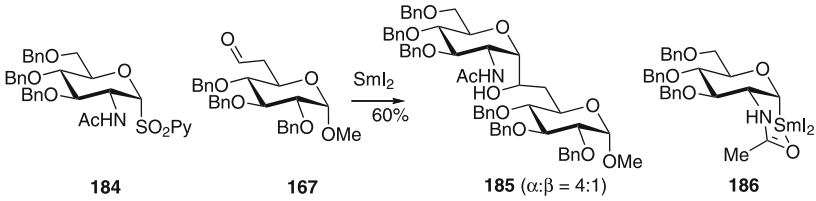
Coupling of sulfone **187** with aldehyde **175** provides a mixture of four isomers which, after oxidation, epimerization and reacetylation produce the β-linked *C*-disaccharide **188**, a protected analog of the β-D-GlcNAc-$(1\rightarrow6)$-D-Man motif of *tri*- and *tetra*-antennary complex-type



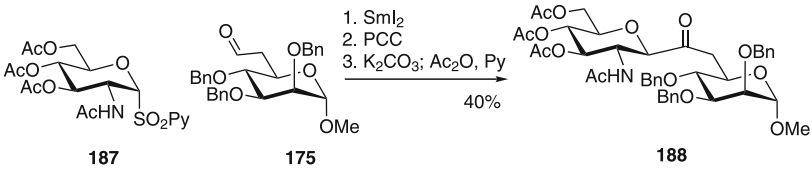
Scheme 36



Scheme 37



Scheme 38



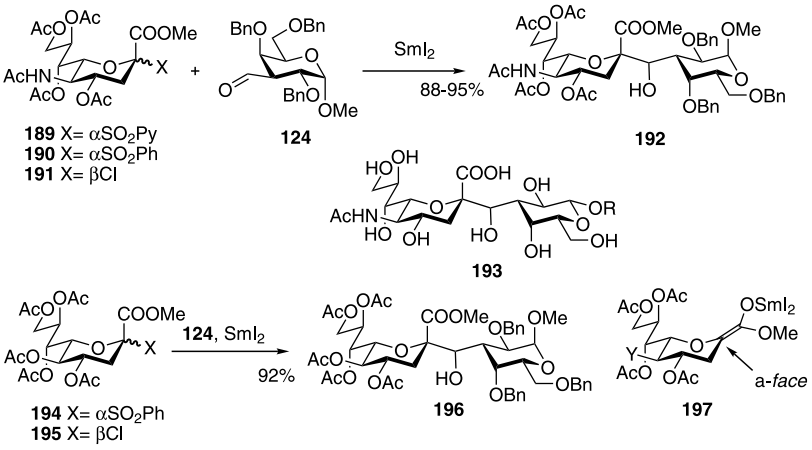
Scheme 39

N-glycans. The procedure also capitalizes on the observation that these anionic conditions remarkably tolerate the presence of standard *O*-acetyl protecting groups as noted previously [72], as well as acidic protons in acetamido groups.

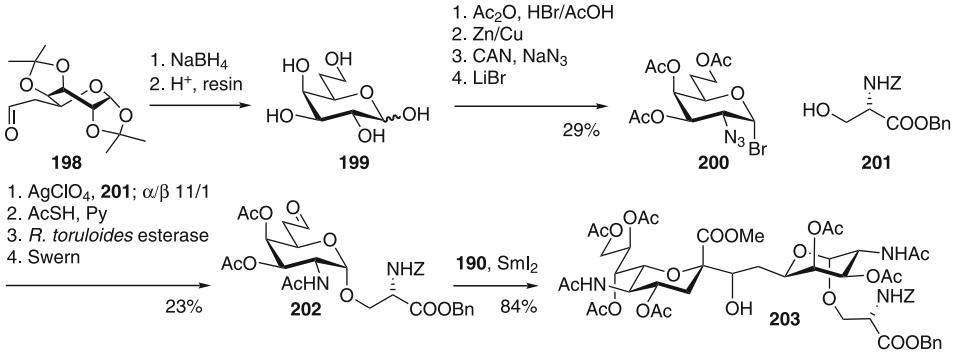
This method has also been applied to a high-yielding stereoselective synthesis of *C*-linked *N*-acetylneuraminic acid (Neu5Ac)-containing disaccharides [92] starting from either the 2-pyridyl (phenyl) sulfones **189** and **190** or the anomeric chloride **191** (*Scheme 40*).

The reductive samariation of phenyl sulfone **190** or chloride **191** is possible without HMPA because the initial homolytic cleavage of the C–S bond, leading to a radical stabilized by the carbomethoxy group, is easier. We note again the exceptional efficiency of this coupling process with acetyl protecting groups. The high α-stereoselectivity can certainly be explained by the formation of a samarium enolate intermediate **197**, which accepts the electrophile from the less sterically hindered α-face. Similar results were reported with other ulosonic acids such as KDN (3-deoxy-D-glycero-D-galacto-2-nonulopyranosylonic acid), a natural analog of NeuAc in which the acetamido group is replaced by a hydroxy group, using phenyl sulfone **194** or chloride **195** [93] (*Scheme 40*).

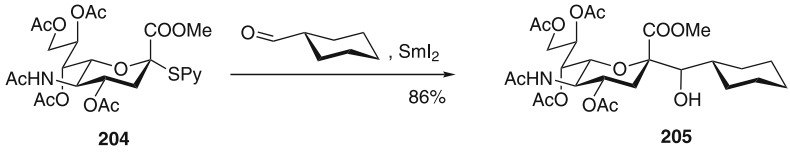
The conformational analysis of deprotected *C*-glycosyl analog of sialyl-α-(2→3)-D-galactose **193** show that the population distribution of conformers around the pseudo-glycosidic linkages is mainly controlled by steric interactions [94].



Scheme 40



Scheme 41



Scheme 42

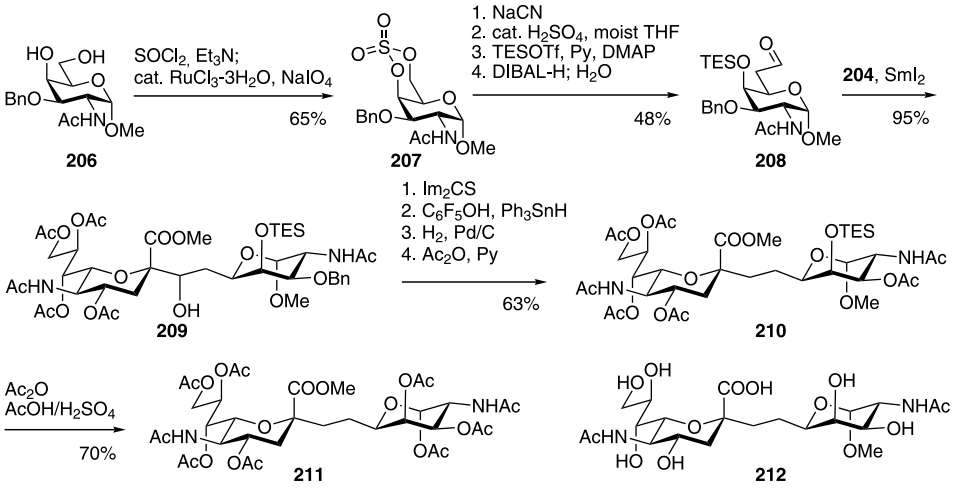
Phenyl sulfone **190** was also the *C*-glycosyl donor used in a synthesis of a *C*-linked mimic **203** of the sialylTn antigen [Neu5Ac-α(2→6)-GalNAc-α(1→*O*)Ser] [95] (*Scheme 41*).

Aldehyde **198**, available in four steps from D-galactose, is the starting point of a ten-step synthesis of the key α-glycosyl-serine aldehyde **202**, partner in the reductive samariation procedure with sulfone **190**. This provides efficiently the expected fully protected *C*-analog **203** of this important antigen that could not, however, be deoxygenated to the methylene-linked analog [95].

It was more recently shown that the stable and crystalline 2-pyridyl sulfide of Neu5Ac derivative **204** is also an excellent precursor of the anomeric organometallic species, in this samarium-Reformatsky procedure [96] (*Scheme 42*).

The efficiency of this approach is seen in a prompt reductive coupling of sulfide **204** with cyclohexane carbaldehyde to afford *C*-glycosyl derivative **205**. This was applied to a fast synthesis of the same disaccharidic component of the sialylTn antigen [96] (*Scheme 43*).

The synthesis was designed in such a way that an activated disaccharidic block would be available for modular attachment to a wide variety of acceptors aiming at its incorporation into a synthetic vaccine. The preparation of the required aldehyde **208** started from *N*-acetylgalactosamine derivative **206**, obtained from the inexpensive *N*-acetyl-D-glucosamine. It was converted to 4,6-cyclic sulfate **207**, following the procedure of Gao and Sharpless [97]. Regioselective ring opening of cyclic sulfate **207** at C-6 [98] by sodium cyanide provided in high yield a cyanide sulfate, which was transformed in three steps into aldehyde **208**. The samarium-Reformatsky key-coupling step proceeds in high yield to afford the *C*-linked dimer **209**, transformed to the methylene-bridge analog **210**. The high stability of these *C*-linked mimics



Scheme 43

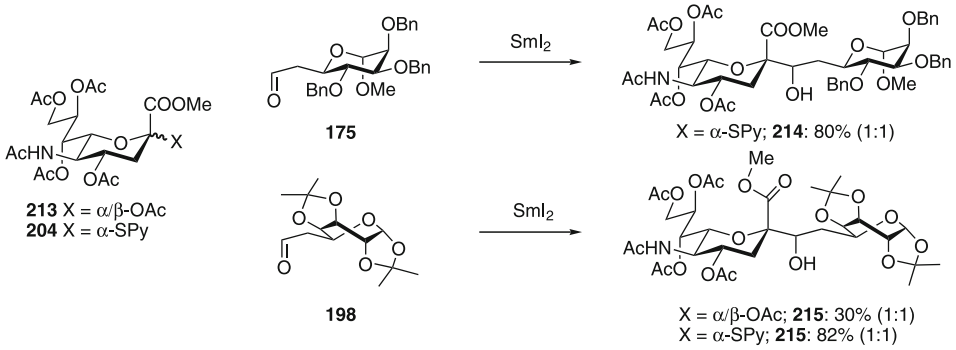
allows for modifications under conditions unacceptable with a native *O*-linkage. Thus, acetolysis of the methyl glycoside in dimer **210** provides anomeric acetate **211**, easily exchangeable to other anomeric substituents usable in a block synthesis.

The clear advantage in using the 2-pyridyl motif (sulfone or sulfide here) over the phenyl sulfonyl anomeric substituent reported in the work of Linhardt [95] concerns the amount of the reducing agent. While 2.5 equiv. are needed with the SPy (or SO_2Py) substituent (theoretical amount, 2 equiv.), a large excess (more than 5–8 equiv.) is necessary with the SO_2Ph substituent because of the competitive deoxygenation of the leaving phenylsulfinate as was previously pointed out [116].

The conformational analysis of the deprotected analog **212**, performed by using NMR and molecular mechanics, indicates a drastic difference from the major conformation found with the natural compound present in solution, especially around the Φ and ω angles [99]. The low energy barriers for conformer interconversion might nevertheless allow easy access to other conformers that can bind to natural receptors without major entropy penalties.

In an even more simple approach, the reductive samariation of the readily available anomeric acetates **213** is possible in the absence of any additive that increases the reducing power of the samarium salt [100] (*Scheme 44*).

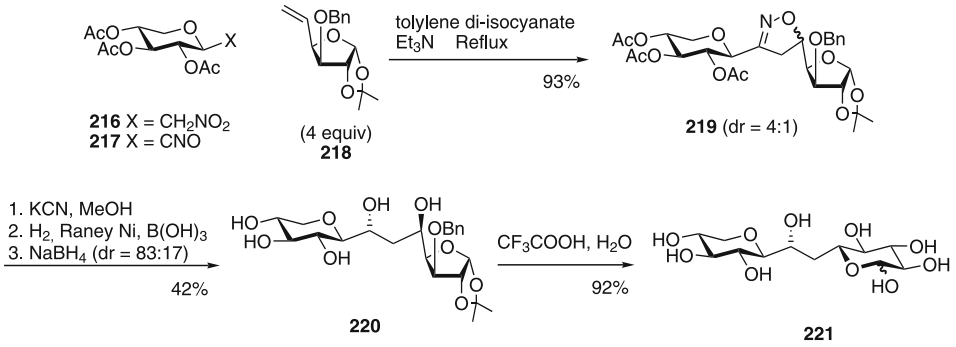
The procedure is remarkably effective for the coupling with cyclic ketones and provides a fast access to α-*C*-ketosides of Neu5Ac. The coupling efficiency is however moderate in the synthesis of *C*-disaccharides, as shown with the D-galactose-derived aldehyde **198** (*Scheme 44*). Acetates **213** provide the expected compound **215** but much less efficiently than sulfide **204** (30 versus 82% yield, respectively). The moderate yield using acetates **213** is due to the competitive pinacol coupling of aldehyde **198** because the rate of the reductive metallation of the anomeric acetates is too slow.



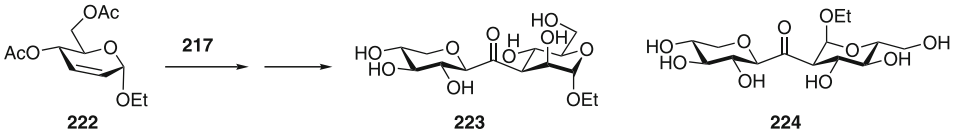
Scheme 44

2.8 Dipolar Cycloaddition

The key carbon–carbon bond formation may be carried out under very different conditions using pericyclic reactions. Paton used the 1,3-dipolar cycloaddition of nitrile oxides derived from pyranoses with unsaturated sugars [15]. Nitrile oxide **217**, generated in situ by dehydration of the nitromethyl derivative **216** in the presence of an excess of alkene **218** at reflux provides regioselectively isoxazoline cycloadducts **219** (dr = 4:1) [101] (*Scheme 45*).
Deacetylation, reductive hydrolysis, and carbonyl reduction of the major isomer yield diastereomers, with one (**220**) being hydrolyzed to the hydroxymethylene-bridged analog of the β-D-Xyl(1→6)-D-Glc dimer **221**. Another example starting from an excess of 2,3-unsaturated sugar **222** and nitrile oxide **217** generates, after similar steps, an equimolar mixture of β-(1→3) and β-(1→2) mimics **223** and **224**, respectively [102] (*Scheme 46*). One should notice that *C*-dimer analog **224** arises from the isomerization of the bulky acyl substituent at C2 to the more stable equatorial position.
The substituents of the eno-pyranoside control well the facial selectivity, however without regiocontrol. With a terminal, non-cyclic alkene such as **218** (*Scheme 45*), the controlling features are reversed (good regioselectivity with lower facial selectivity).



Scheme 45



Scheme 46

3 The Radical Assemblage

3.1 Intermolecular Reactions

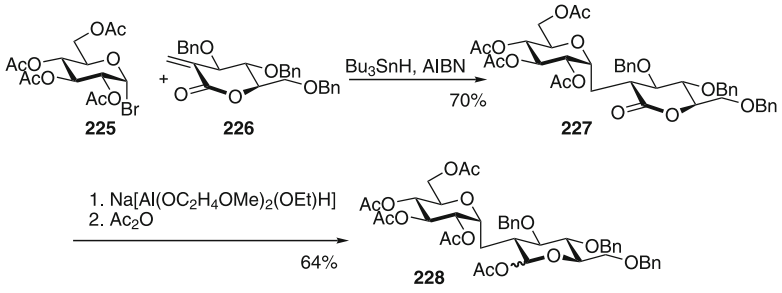
Instead of relying on ionic reactions to execute the crucial assembling of two glycosidic units, one may profitably consider radical reactions [12,16]. An anomeric radical, being an alkoxyalkyl radical, will behave as a nucleophile, and its high-lying SOMO will better interact with the LUMO of an electron-poor alkene, typically α,β-unsaturated carbonyl compounds or similar derivatives. In other words, in intermolecular reactions, the radical acceptor should be activated enough to compete efficiently with an H-donor.

Giese showed that an anomeric radical, obtained by reaction of acetobromoglucose **225** with tributyltin hydride, adds to sugar-derived α-methylene-γ-lactones such as **226** with good yields and high α stereoselectivity [103] (*Scheme 47*).

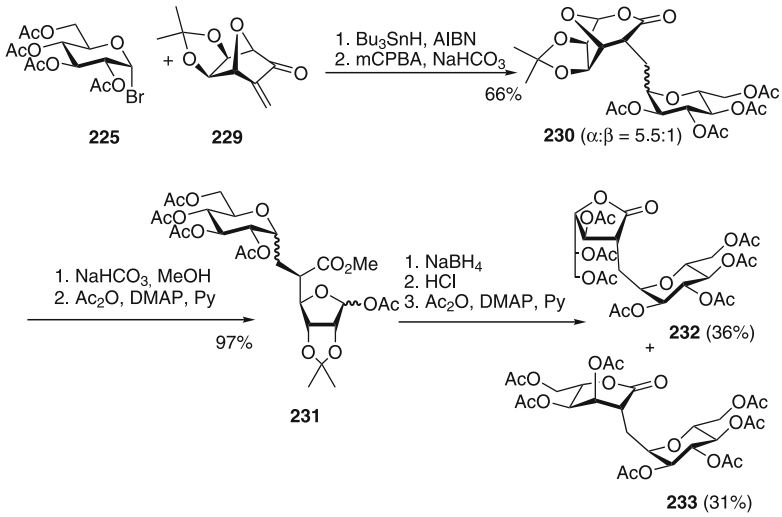
The resulting lactone is further modified to the corresponding *C*-linked disaccharide **228** (*C*-kojibiose).

Similar reactions were described by Vogel, using a "naked sugar" as a radical acceptor. The resulting product has to be further elaborated before it becomes a "true" *C*-linked disaccharide, with the key-step being a Baeyer–Villiger ring expansion [104,105] (*Scheme 48*).

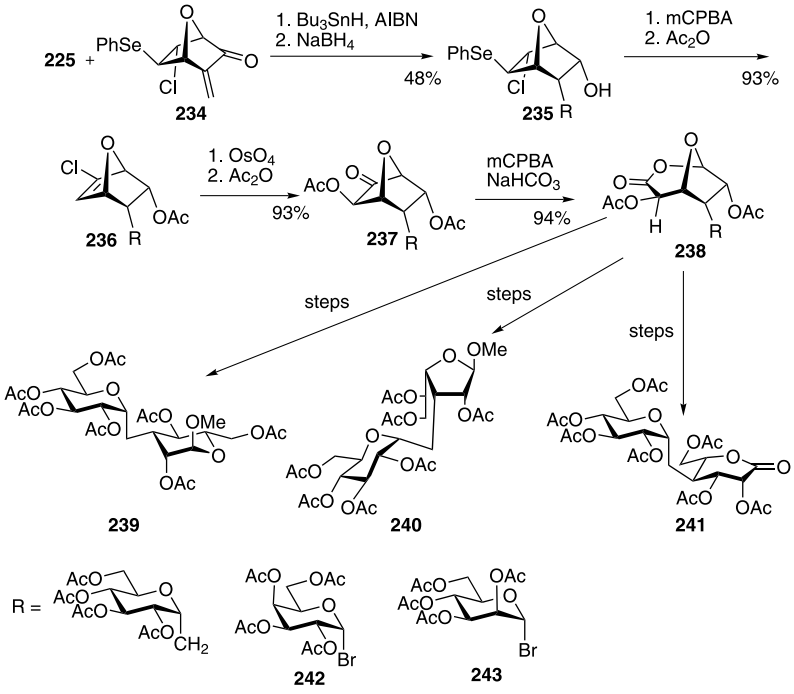
Variations of the radical acceptor substitution pattern and use of acetobromogalactose **242** [106] or acetobromomannose **243** [107,108] as a radical precursor has resulted in the synthesis of various *C*-dimers including the *C*-linked analog of Man-α(1→3)-GalNAc, which proved to be an inhibitor of several glycosidases, as well as human α-1,3-fucosyltransferase VI [108] (*Scheme 49*).



Scheme 47

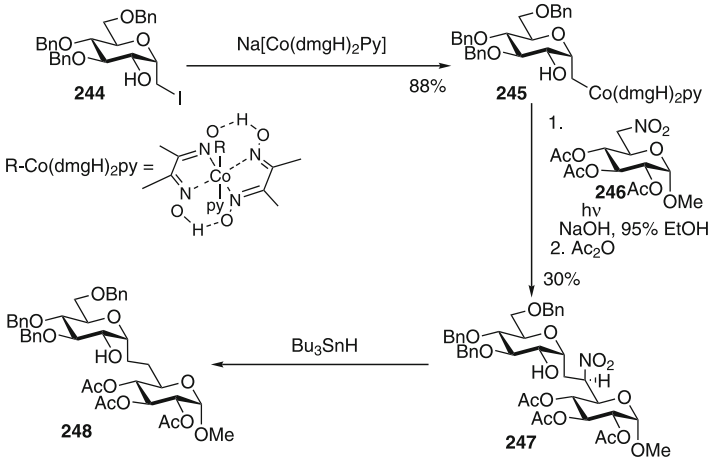


Scheme 48

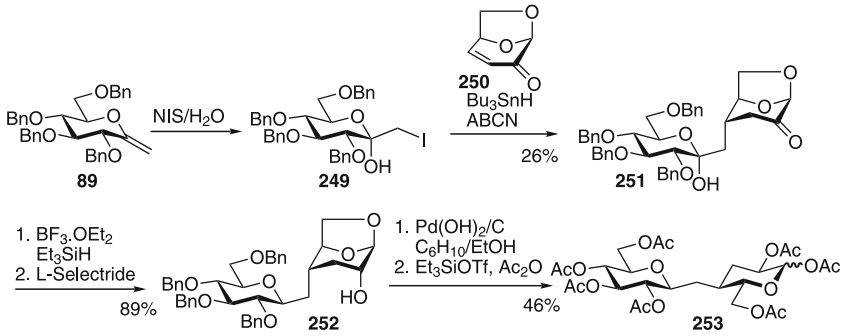


Scheme 49

In a very different approach, Martin used an α-glycosylmethyl cobaloxime **245**, prepared from iodide **244**, to photochemically generate the corresponding radical [109] (*Scheme 50*).
This radical adds to the nitronate ion derived from 6-nitro sugar **246**. Tin-hydride mediated denitration leads to an α-(1→6) *C*-linked disaccharide **248**.
Witczak also used a methylene "extended" glucose derivative **249** as a radical precursor [110] (*Scheme 51*).
Stereoselective radical addition on the *exo*-face of levoglucosenone **250** leads to the *C*-linked dimer **251** in a modest 26% yield. The expected competing reaction is the direct hydrogen abstraction by the initial radical species. Stereoselective hemiketal and ketone reductions afford the 1,6-anhydro derivative **252**, which is deprotected and opened to give the final *C*-analog of a β-(1→4) disaccharide **253**.



Scheme 50



Scheme 51

3.2 Intramolecular Reactions

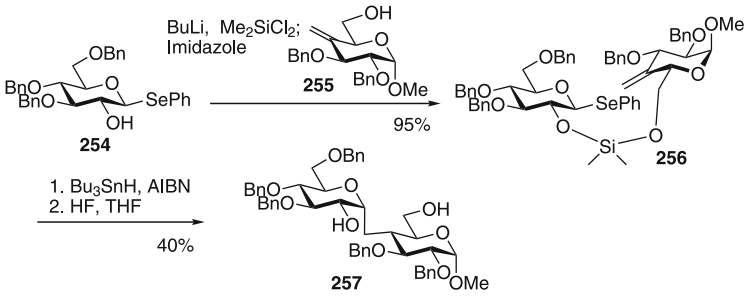
Olefin activation by electron-withdrawing groups is no longer necessary when an intramolecular trapping is programmed because radical cyclizations occur faster than do intermolecular additions. In such cases, the stereoselectivity of the cyclization and its efficiency will depend on several parameters such as ring size, the type of atoms, and the nature of asymmetric centers present in the ring.

This strategy was mostly exploited by Sinaÿ and his coworkers, enabling them to prepare various *C*-linked disaccharides in an expeditious fashion, via radical cyclization of temporarily tethered monosaccharidic precursors [111]. Medium-sized ring intermediates are obtained in reasonable yields, leading to the desired compounds after removal of the tether. The prototype of this approach is described below [112a] (*Scheme 52*).

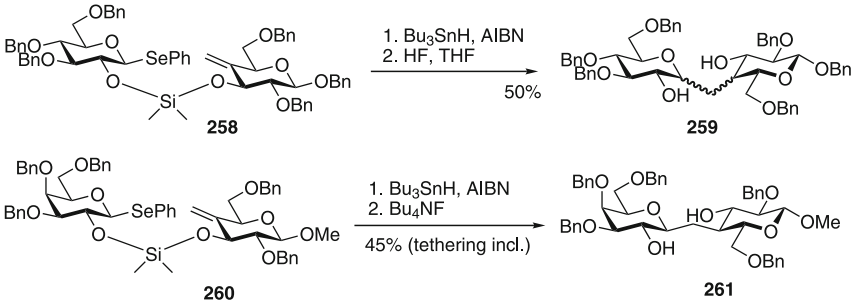
The two glucose derivatives **254** and **255** are linked via a silaketal function. Radical 9-*endo*-trig cyclization of intermediate **256** leads, after removal of the tether, to a single *C*-dimer **257** isolated in 40% yield. Anomeric radical trapping occurs exclusively from the α side of the initial species, while H-abstraction by the resulting intermediate radical leads to the equatorially substituted product. As a result, the isolated compound is the *C*-linked analog of methyl maltoside **257**. When this methodology was extended to the *galacto* series, the cyclization yield increased to 60%, although three different *C*-dimers were obtained via the 9-*endo*-trig cyclization [112b].

The nature of the attacking sugar radical, the size of the intermediate ring, and the choice of the tethering positions have a great influence on selectivities, and consequently on product distributions. For example, while the 8-*endo*-trig cyclization of the gluco compound **258** yields an inseparable mixture of all four possible products [112a], the same reaction performed in the *galacto* series leads, in 45% yield, to the *C*-linked analog of methyl β-lactoside **261** [113] (*Scheme 53*).

It is noteworthy that this is an exceptional reported situation where the major isolated compound resulted from an attack of the anomeric radical on its β face, leading to a β-*C*-linked dimer. This unique behavior is probably caused by a strong conformational bias in the intermediate species.

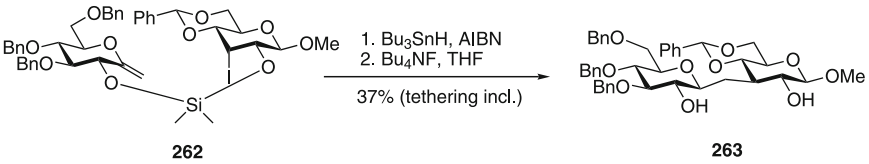


Scheme 52



258 **259** **260** **261**

Scheme 53



262 **263**

Scheme 54

Interestingly, when the role of the partners is reversed, e. g., when the radical is formed on a non-anomeric position of one of the sugars as in **262** (Scheme 54), the addition occurs equatorially to an anomeric exomethylene, resulting in an anomeric radical that abstracts hydrogen from the α face.

Consequently, the β-*C*-linked disaccharide **263**, analog of a methyl laminaribioside derivative, is isolated [114]. Other β-*C*-linked disaccharides, namely analogs of cellobiose, Gal-β-(1→3)-Glu, β,β and α,β trehaloses have been prepared using this strategy.

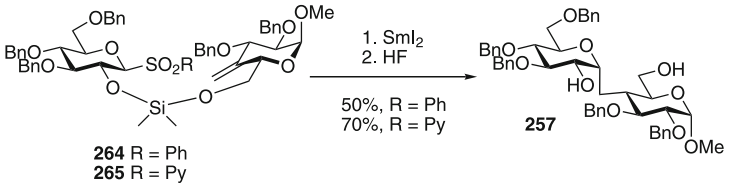
As an alternative to the Bu_3SnH-phenylseleno glycoside chemistry, radicals may also be generated by samarium diiodide reduction of anomeric arylsulfonyl groups [115], (R=Ph) (Scheme 55).

The use of the 2-pyridylsulfonyl group introduced by Skrydstrup and Beau [116], more easily reduced than the phenylsulfonyl group, allows a significant increase in the cyclization yield [117].

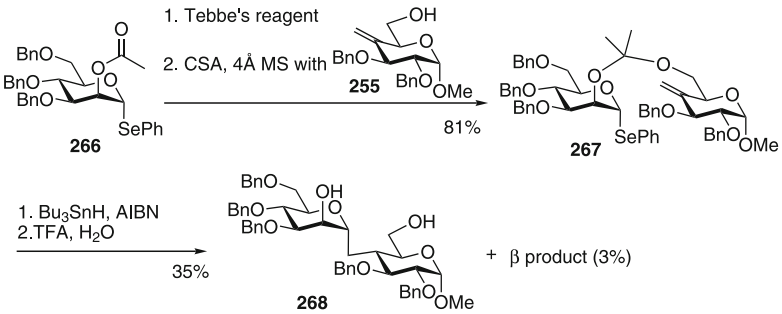
Other temporary connections, such as ketal tethers were also tested. Tebbe olefination of acetate **266** and ketal formation with sugar alcohol **255** lead to tethered intermediate **267**, which is cyclized to a *C*-analog derivative **268** of the α-D-Man(1→4)-D-Glc disaccharide after ketal cleavage [118] (Scheme 56).

Similarly, DDQ oxidation of the *p*-methoxybenzyl group of **269** in the presence of alcohol **270**, a strategy reported by Ito and Ogawa [119] in *O*-glycoside synthesis, provides benzylidene acetal **271** [120] (Scheme 57).

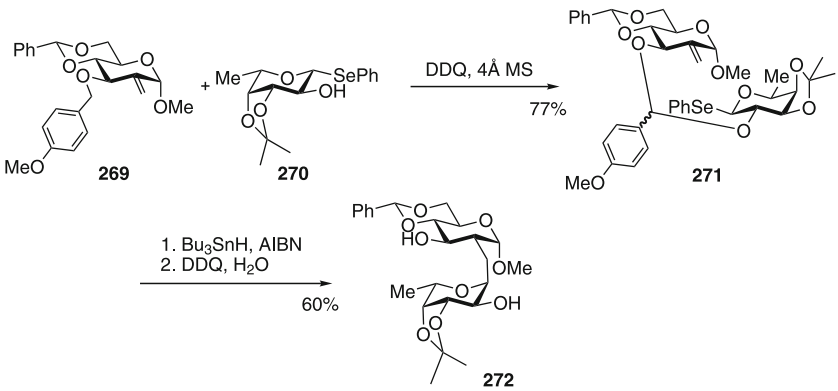
8-*Endo* radical cyclization and removal of the acetal generate a *C*-linked derivative **272** of the α-L-Fuc(1→2)-D-Glc dimer.



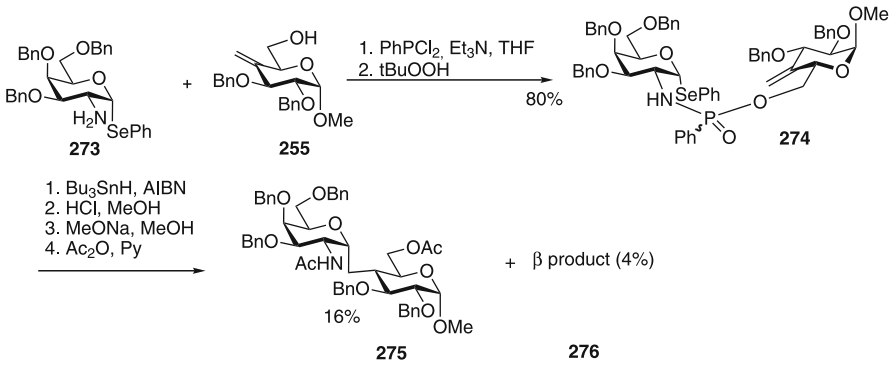
Scheme 55



Scheme 56



Scheme 57



Scheme 58

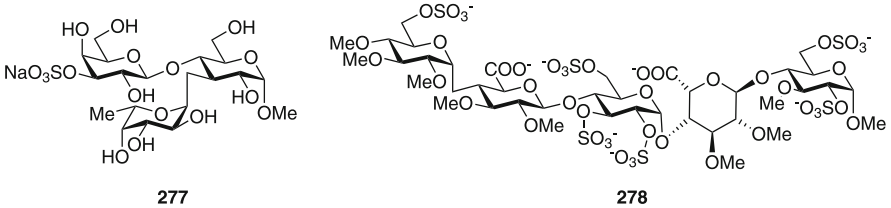
A phosphoramidic linker was also used for tethering via the amino group of a 2-amino-2-deoxy sugar [121] (*Scheme 58*).

Once again, 9-*endo* cyclization of intermediate **274**, obtained in two steps from galactosamine derivative **273** and unsaturated alcohol **255**, followed by detethering and acetylation of the amino group leads to *C*-dimers **275** and **276** in an expeditious way.

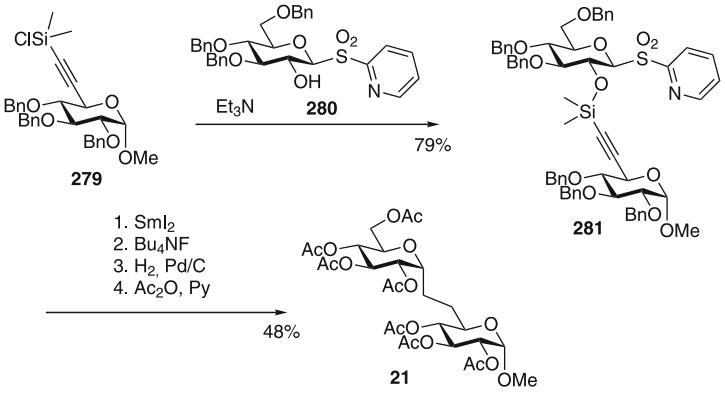
We have seen in the preceding examples the efficiency and potential generality of this type of *C*-glycosylation procedure. The monosaccharidic precursors are relatively easy to obtain, allowing these reactions to be performed on a multigram scale. The resulting *C*-linked disaccharides can consequently be incorporated in more complex structures, like mixed C/O-linked oligosaccharides. This approach has been successfully used in the synthesis of several *C*-linked glycosides of potential biological interest such as a sialyl Lewis^x analog **277** [122], and a "C/O-linked pentasaccharidic" analog of the antithrombin III binding region of heparin **278** [123] (*Scheme 59*).

Another intramolecular approach to *C*-linked disaccharides has been described by Skrydstrup and Beau [116] using 2-pyridyl sulfones to generate anomeric radicals. Tethering sulfone **280** with chlorosilane **279** provides silyl ether **281** (*Scheme 60*).

Samarium diiodide-promoted 5-*exo* radical cyclization and further transformation lead only to the *C*-analog derivative **21** of methyl isomaltoside.



Scheme 59

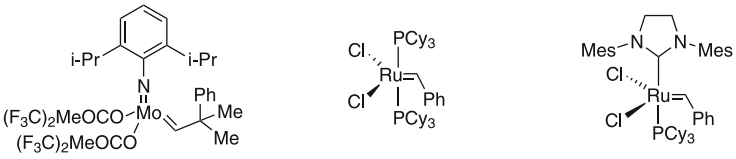


Scheme 60

4 The Olefin Metathesis Route

In the last decade, olefin metathesis has emerged as a powerful method for the formation of carbon–carbon bonds under mild conditions. This reaction can be defined as the exchange of alkylidene fragments between two olefins promoted by metal–carbene complexes. The synthetic development in this field happened with the discovery of catalysts of high activity and functional group tolerance. Initial use of the air-sensitive Schrock's molybdenum catalyst **282** [124], was followed by the use of Grubbs first generation ruthenium carbene **283** (Grubbs I), which is more stable to air and moisture [125]. Metathesis protocols now rely heavily upon the utilization of the highly stable second generation Grubbs catalyst **284** (Grubbs II) bearing *N*-heterocyclic carbene ligands [126] or similar catalysts (*Scheme 61*).

The potential of this reaction with polyfunctional substrates has provided new synthetic solutions in carbohydrate chemistry. Among the three main metathesis versions (cross metathesis, ring-opening or closing metathesis and enyne metathesis), the ring-closing metathesis (RCM) is widespread because, in this case, intramolecular reaction shows an entropic advantage compared to cross metathesis. The Postema group illustrated the value of RCM in the synthesis of *C*-analogs of natural structures [23]. In an initial study, they showed that esterification of an appropriate olefin alcohol followed by a methylenation reaction and ring-closing metathe-



Scheme 61

sis allowed the preparation of C1-substituted glycals in reasonable yields [127]. Selective hydroboration-oxidation of the C–C double bond of these substrates delivers the *C*-glycosides [128] (Scheme 62).

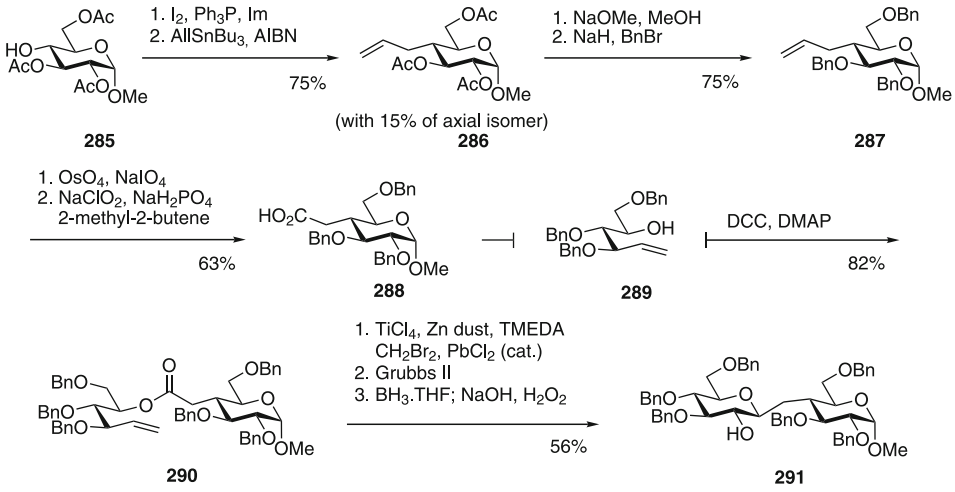
In the example shown in Scheme 62, the C4-allylated sugar **286** is prepared by a radical allylation from the C4-iodide. Exchange of protecting groups, oxidative cleavage of the alkene and Pinnick oxidation of the major aldehyde provide the corresponding acid **288**, which is esterified with olefin alcohol **289** to ester **290**. Takai's methylenation furnishes the acyclic enol ether, substrate for the RCM in presence of Schrock or Grubbs II catalysts. Hydroboration-oxidative quench provides selectively the β-*C*-*gluco* disaccharide isomer **291**. The overall yield of the three-step sequence (methylenation-RCM-hydroboration) is largely improved (56%) using a one-pot procedure with the Grubbs II catalyst.

This approach was extended to the synthesis of $\beta(1\rightarrow1)$,$\beta(1\rightarrow2)$,$\beta(1\rightarrow3)$ and $\beta(1\rightarrow6)$, [129] linked *C*-disaccharides in good overall yield (Scheme 63).

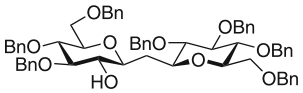
All these *C*-disaccharides were deprotected and tested against various glycosidase enzymes. Only the $\beta(1\rightarrow1)$-linked *C*-disaccharide derived from **292** showed a modest inhibitory activity with a K_i of 126 µM against almond β-glucosidase.

More recently, the same methodology was applied to the synthesis of β-*C*-trisaccharide **298** [130] and tetrasaccharide **301** [131] from diacid **296** and triacid **299**, respectively (Scheme 64).

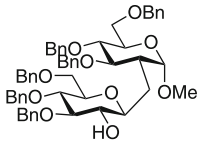
In contrast to RCM, cross metathesis is more challenging to implement because undesired homodimerization of olefins can occur. However, Lowary and co-workers took advantage of this side-reaction to access to potential arabinosyltransferase inhibitors. They reported the synthesis of a *C*-analog of α-D-Ara*f*-(1→5)-α-D-Ara*f*, a motif present in the mycobacteria cell wall [132] (Scheme 65).



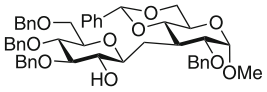
Scheme 62



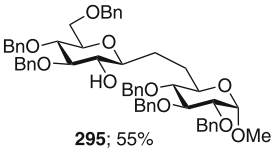
292; 59%



293; 53%

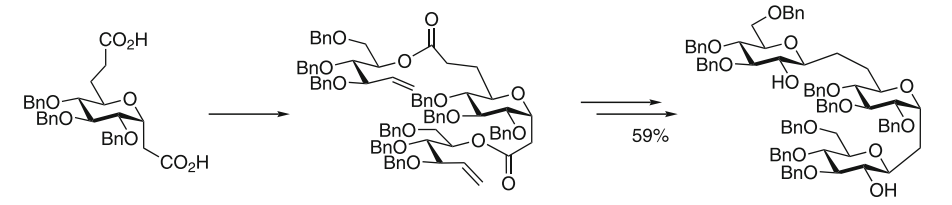


294; 59%

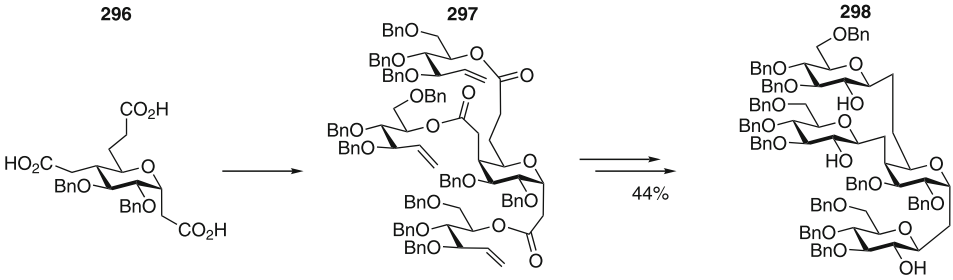


295; 55%

Scheme 63

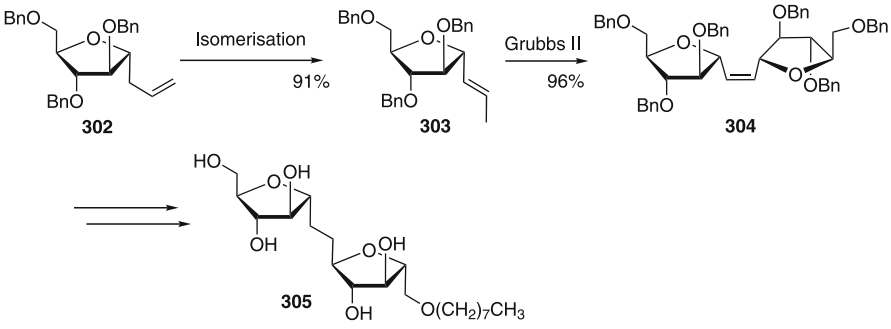


296 **297** **298**



299 **300** **301**

Scheme 64



302 **303** **304** **305**

Scheme 65

Isomerization of the *C*-allyl glycoside **302** provides the corresponding *C*-(1-propenyl) derivative **303** as a major *E*-isomer. Cross metathesis proceeds in high yield only with the Grubbs II catalyst to deliver the homodimer **304**, transformed to *C*-disaccharide analog **305** of the mycobacterial motif.

5 Partial de Novo Synthesis

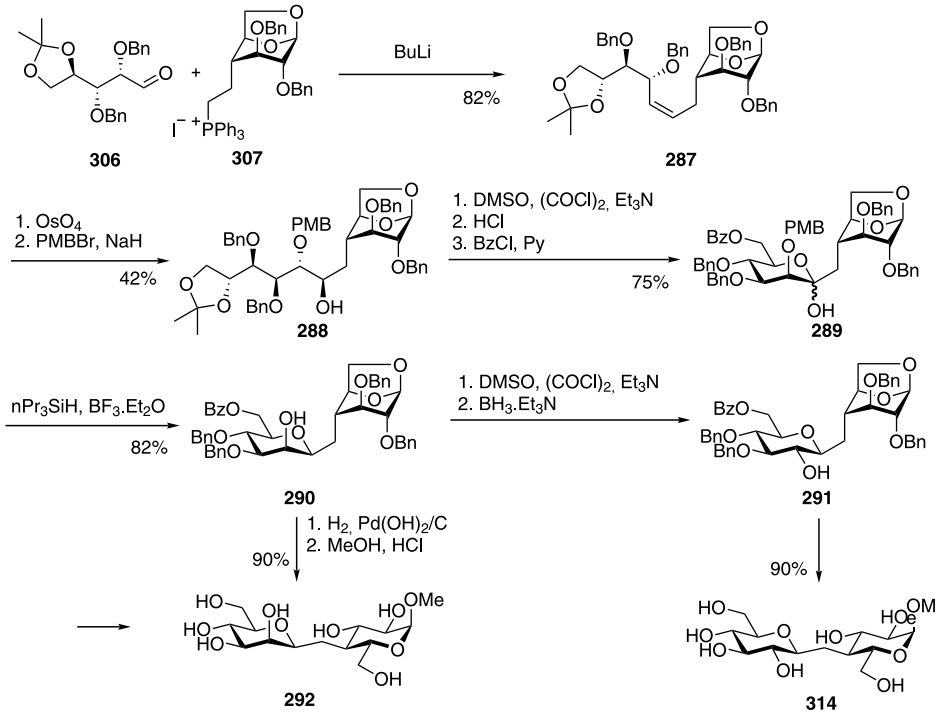
Instead of using a readily accessible natural sugar that can ultimately be transformed to a sugar analog, a de novo synthesis of one or several carbohydrate units in the oligomer can be programmed. One of the advantages of the de novo design is that a given synthetic plan towards a particular target can easily be redesigned to accommodate stereostructural variants using the appropriate diastereo(or enantio-)selective transformations currently available, particularly in acyclic systems. Much of the work in this area was pioneered by Kishi and his coworkers who prepared a collection of *C*-glycosides and *C*-oligomers, a synthetic program primarily designed to probe their solution conformation for purposes of comparison with the corresponding parent glycosides. This work was also closely connected to the stereochemical assignment and total synthesis of palytoxin [25], a polyhydroxylated marine natural product containing no less than seven functionalized tetrahydrofurans and tetrahydropyrans which can be considered as substructural *C*-glycosyl motifs. One should note the various facets of the synthetic technology developed by Kishi and coworkers including, to name only a few, axial and equatorial nucleophilic *C*-allylation (allenylation) of electrophilic sugars, $CrCl_2$/$NiCl_2$-mediated coupling of vinylic halides and aldehydes, stereo-directed *cis*-dihydroxylation of olefins, cyclization strategies to stereo-defined tetrahydropyranyl units, that have been a rich source of inspiration to many other investigators in this field.

A common feature of these partial de novo syntheses of *C*-linked dimers or oligomers is that an intact (or slightly modified) sugar unit is used while the other is being appended and constructed by total (or partial) synthesis.

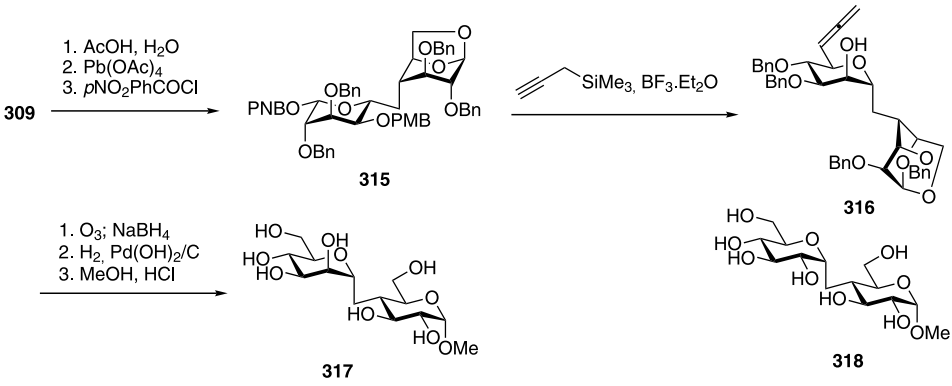
The first example of the Kishi series reported in 1987 [133] involved a stereocontrolled synthesis of four *C*-dimers which include the *C*-analogs of cellobiose and maltobiose. Construction of the *C*-disaccharides starts by the exclusive *cis*-olefination of aldehyde **306**, available from D-arabinose and the ylide generated from phosphonium salt **307** (*Scheme 66*).

Stereoselective dihydroxylation, monoprotection of the diol (→**309**), oxidation, and cleavage of the isopropylidene group lead to a cyclic hemiacetal which is benzoylated on the primary hydroxy group (→**310**). Hemiacetal reduction to the *C*-disaccharidic derivative **311** is stereoselective only with the *n*-Pr_3SiH reagent. Deprotection and methyl glycoside formation provides the *C*-analog of the β-D-Man-(1→4)-D-Glc disaccharide **313**. High yield conversion of the *manno* series **311** to the *gluco* series **312** is feasible using an oxidation-reduction sequence as shown, leading ultimately to the *C*-analog of methyl cellobioside **314**. The isomeric axially substituted α-*C*-disaccharides are also available from the same intermediate **309** by inversion of the hydroxy-carbonyl functions in the acyclic segment (*Scheme 67*).

This is readily done by isopropylidene removal, diol cleavage, and formation of the anomeric *p*-nitrobenzoate **315**. Axial introduction of the allenyl moiety yields stereoselectively, as predicted, the α-*C*-manno derivative **316** after chair inversion. Degradation of the allenyl group to



Scheme 66



Scheme 67

a primary hydroxy group, deprotection, and methyl glycoside formation provides the *C*-mimic of the α-D-Man-(1→4)-D-Glc disaccharide **317**. A similar sequence of reactions also provides methyl *C*-maltobioside **318**.

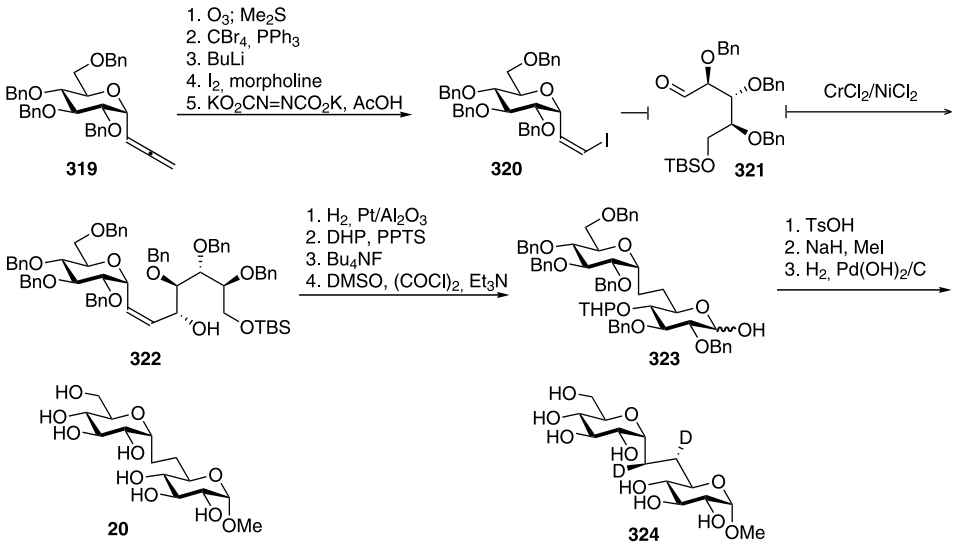
The stereoselective formation of allylic alcohol **322** (15:1 ratio) by the $CrCl_2/NiCl_2$-mediated coupling reaction [134] of vinyl iodide **320**, generated from *C*-allenyl glucoside **319**, and aldehyde **321**, available in six steps from L-xylose, is the key assembling step in the synthesis of methyl *C*-isomaltoside **20** [135] (● *Scheme 68*).

The remaining sequence includes a four-step transformation of the acyclic segment to a D-glucopyranosyl unit in **323** that was deprotected to the final *C*-analog **20**. The unsaturated intermediate **322** is also very useful to stereoselectively deuterate the analog (*cis*-deuteration by D_2, Pt on Al_2O_3) as in the d_2-derivative **324**, for conformational studies.

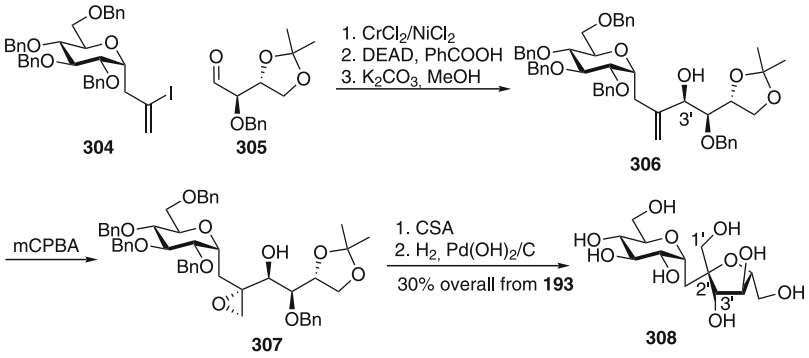
The same strategy was used in a concise synthesis of the *C*-analog of sucrose (and ultimately its C2′, C3′ diastereomers) starting from iodide **325** and aldehyde **326** [136] (● *Scheme 69*).

The anti-selective (10:1 ratio) $CrCl_2/NiCl_2$ coupling reaction produces an allylic alcohol (C3′-diastereomer of **327**) that is inverted to the desired allylic alcohol **327** by a Mitsunobu procedure. *m*CPBA oxidation of **327** gives selectively the *anti*-epoxy alcohol **328** that cyclizes stereospecifically in a 5-*exo* fashion under acid catalysis, by exclusive attack of the secondary hydroxy group at the quaternary center of the epoxide. Hydrogenolysis of the protecting groups completes this efficient synthesis of *C*-sucrose **329** in an overall yield of 30% from vinyl iodide **325**.

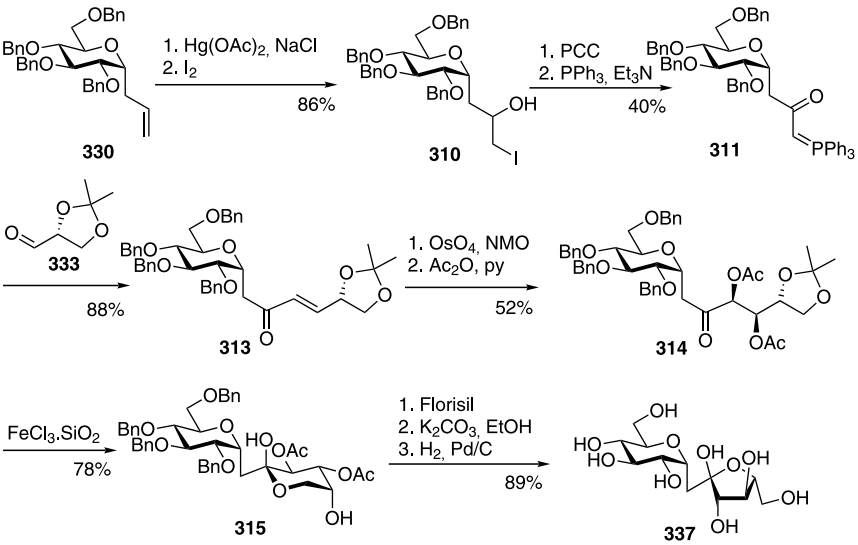
Another synthesis of a *C*-sucrose analog, missing however C1′, was reported by Nicotra from stabilized ylide **332**, generated in four steps from *C*-allyl glucose **330** [137] (● *Scheme 70*).



Scheme 68



Scheme 69

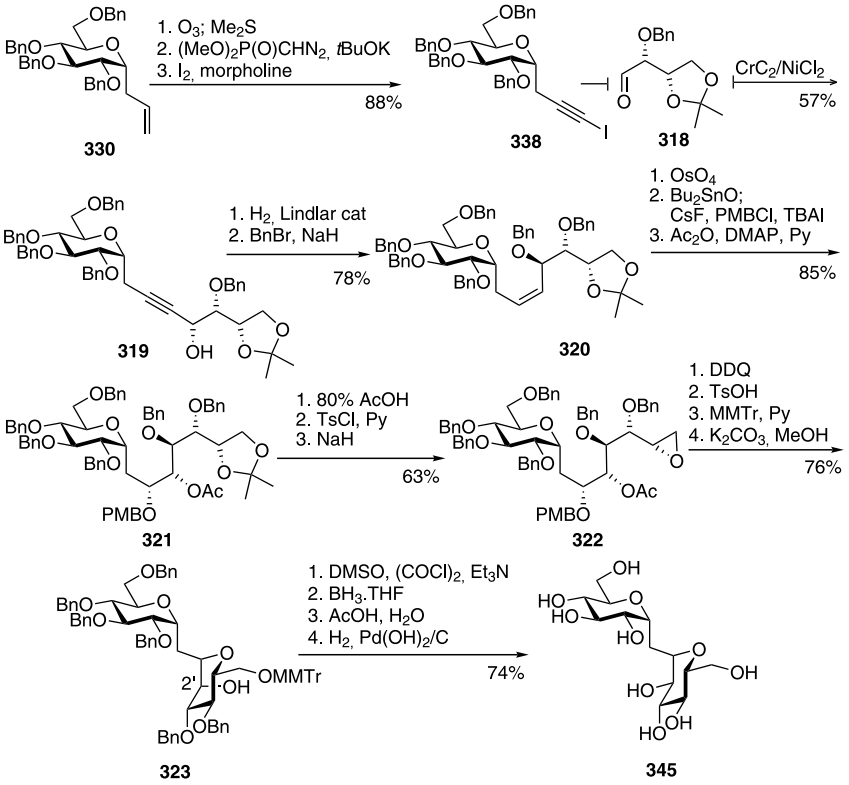


Scheme 70

Condensation of **332** with the D-glyceraldehyde derivative **333**, stereoselective dihydroxylation (60% de) and acetylation afford diacetate **335**. Acidic treatment provides rather the pyranose form **336** which isomerizes to the furanose form upon Florisil treatment. Deprotection in two steps leads to the *C*-analog of sucrose **337**, in equilibrium with the pyranose form.

Another synthesis including a Cr(II)/Ni(II)-mediated coupling reaction as a key coupling step was proposed by Wei and Kishi in 1994 [138]. They reported a single strategy for the stereocontrolled synthesis of the *C*-analog of α,α-trehalose **345** [α-D-glucopyranosyl $(1\rightarrow1')$-α-D-glucopyranoside] as well as its α,β- and β,β-diastereomers (*Scheme 71*).

Coupling of alkynyl iodide **338**, obtained in three steps from allyl glucose **330** with aldehyde **339**, prepared from L-xylose, furnishes the *anti*-propargylic alcohol **340** (stereoselectivity of 4 to 1) which can be reduced to either the *E*-olefin with bis(methoxyethoxy)aluminum hydride

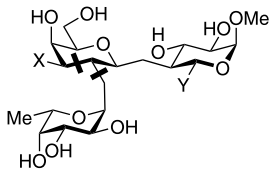


Scheme 71

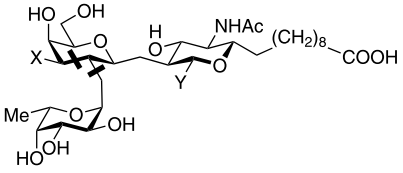
(Red-Al) or the *Z*-olefin **341** after benzylation. Stereoselective dihydroxylation and regioselective protection of the diol provides protected polyol **342** that is transformed in three steps to terminal epoxide **343**. Removal of the *p*-methoxybenzyl group followed by acid-catalyzed epoxide opening yields exclusively tetrahydropyran **344** after protection of the primary alcohol and deacetylation. Inversion of the C2′ stereocenter by an oxidation-reduction sequence and deprotection furnishes the *C*-mimic of α,α-trehalose **345**.

An aldol route was rather adopted by Kishi [139,140] for the construction of the central unit of trimer analogs **346–353** of the type II(H) cellular antigen trisaccharide α-L-Fuc-(1→2)-β-D-Gal-(1→4)-β-D-GlcNAc (*Scheme 72*).

This strategy is flexible enough to accommodate specific structural modifications. The aldol disconnections, presented in *Scheme 72*, call for a glucosamine (or glucose) unit with a three-carbon appendage at position-4 and a C1-formyl fucose unit connected by a four-carbon acyclic segment carrying two asymmetric centers of the future central galactose unit. The synthesis of the glucosamine unit **358**, used in the construction of *C*-analog **350**, starts with epoxide **354** transformed to the C4-branched epoxide **355** by a Cu(I)-catalyzed Grignard reaction followed by base treatment [140] (*Scheme 73*).

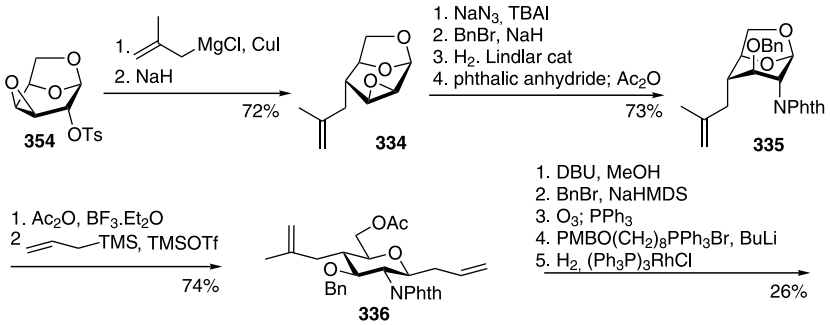


346 X = OH, Y = CH_2OH
326 X = H, Y = CH_2OH
327 X = OH, Y = H
328 X = Y = H



329 X = OH, Y = CH_2OH
330 X = H, Y = CH_2OH
331 X = OH, Y = H
353 X = Y = H

Scheme 72

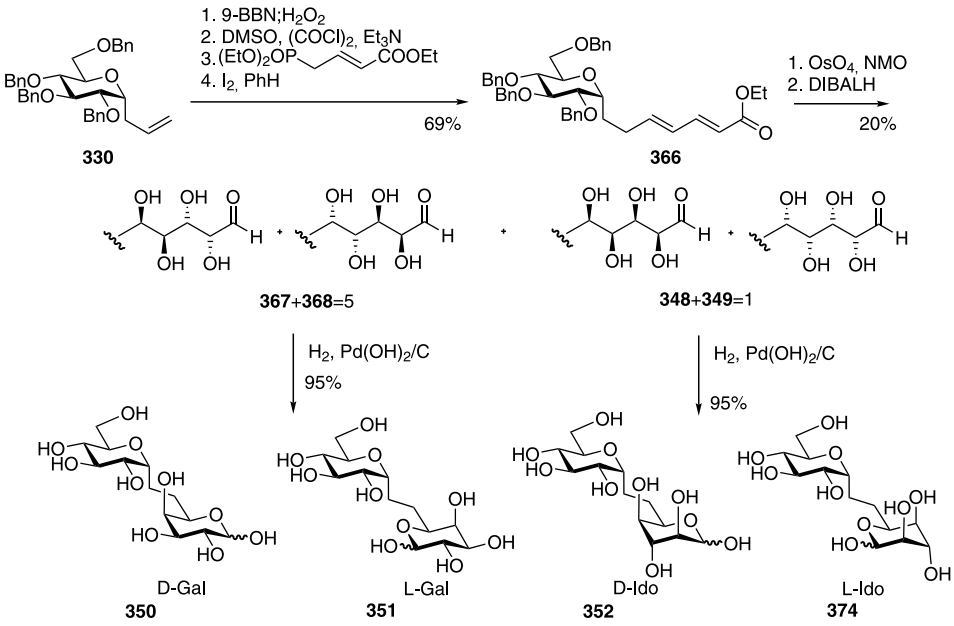


Scheme 73

Azide opening of epoxide **355**, benzylation, azide reduction, and phthalimide formation provide the 1,6-anhydro derivative **356** that is opened and stereoselectively β-allylated to **357**, an interesting stereochemical outcome, probably directed by the C2-phthalimido group. The first aldol condensation between methyl ketone **358**, generated in five steps from **357** and aldehyde **359**, provides smoothly aldol products **360**, converted to the *C*-disaccharide ketone **361** by silyl removal, hemithioketal formation/tin hydride reduction, and oxidation. The second aldol reaction between **361** and 1-*C*-fucopyranosyl formaldehyde **362** needs optimal equilibrating conditions (LHDMS, TMEDA, $MgBr_2$) to produce the equatorial aldols **363**. Conjugate and ketone reductions of enones **364**, generated by mesylation and base elimination on **363**, lead to *C*-trimer **365** transformed to the final polyol **350** in the form of a lithium carboxylate by a five-step sequence of routine transformations.

Armstrong and Sutherlin applied a combinatorial concept, by keeping one sugar unit constant while varying the structure of others to include all possible stereomers. This is exemplified by the synthesis of four *C*-linked D-glucopyranosyl α-(1→6) hexopyranoses [141] (*Scheme 74*).

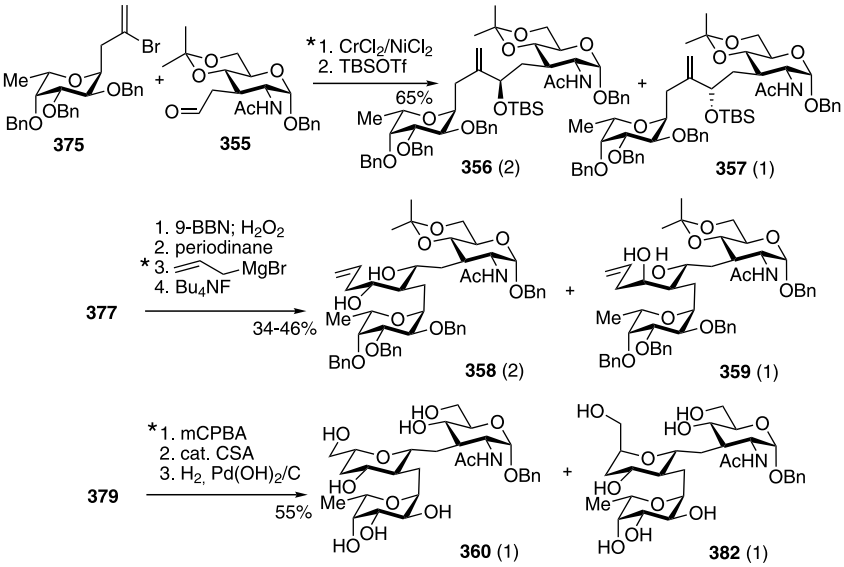
Elaboration of *C*-allyl glucose **330** to diene ester **366** by the four steps shown followed by a dihydroxylation-reduction sequence affords a mixture of the lactols **367–370** (represented in acyclic forms) having the D/L-galacto and the D/L-ido configurations in a 5:1 ratio. Hydrogenolysis provides the corresponding *C*-dimers **371–374**, four out of the sixteen possible diastereomers.



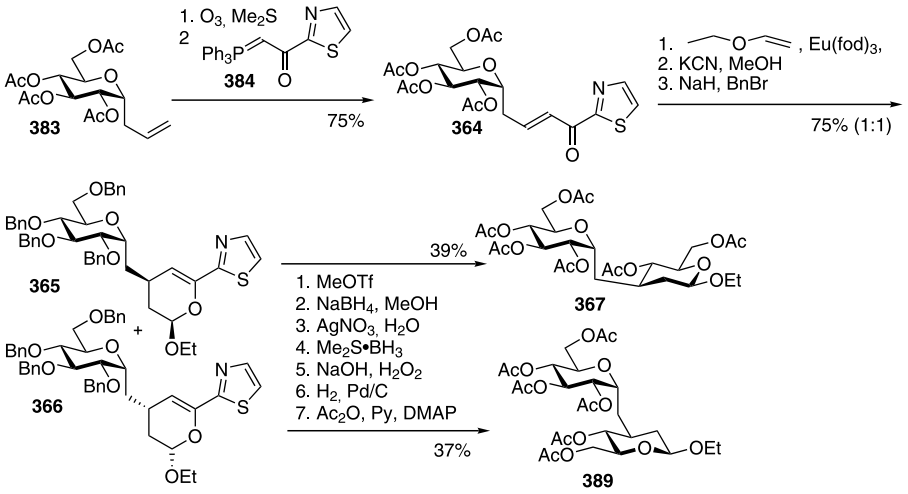
Scheme 74

Another example is the synthesis of *C*-trimer analogs of the Lewis type I blood group determinant having the general formula α-L-Fuc-(1→2)-hexose-(1→3)-D-GlcNAc in which the structural diversity in the central hexose unit is achieved by performing either non-selective transformations to create a mixture of diastereomers, or predictably selective transformations to construct a single diastereomer [142]. The approach begins with a $CrCl_2/NiCl_2$ coupling [134] of vinyl bromide **375**, available in three steps from L-fucose, and aldehyde **376**, generated in four steps from *N*-acetylglucosamine, leading after silylation to diastereomers **377** and **378** (2:1 ratio) [142] (● *Scheme 75*).

Diastereoselective hydroboration-oxidation and periodinane oxidation on major isomer **377** give an aldehyde which is homologated by the addition of allylmagnesium bromide and desilylation to diols **379** and **380** (2:1 ratio). The major isomer **379** is in turn unselectively oxidized with *m*CPBA and subsequently cyclized under acid catalysis providing the two *C*-trisaccharide analogs **381** and **382** (1:1 ratio) after removal of all protecting groups with the exception of the anomeric benzyl group on GlcNAc. This strategy has resulted in the synthesis of eleven *C*-analogs of a hexose trimer. In the case where inseparable diastereomeric mixtures are generated by this type of synthetic sequence, access to a single component of the mixture is possible by replacing a non-selective step (marked by asterisks in ● *Scheme 75*) with a selective option, an operation referred to as a "recursive stereochemical deconvolution" strategy by the authors. Kniezo and collaborators developed [143] a short and efficient de novo synthesis of α-(1→3) *C*-linked disaccharides starting from *C*-allyl glucose **383**, which, subjected to ozonolysis prior to direct reaction of the resulting ozonide with a stabilized phosphorus ylide **384**, leads to **385** (● *Scheme 76*).



● **Scheme 75**



Scheme 76

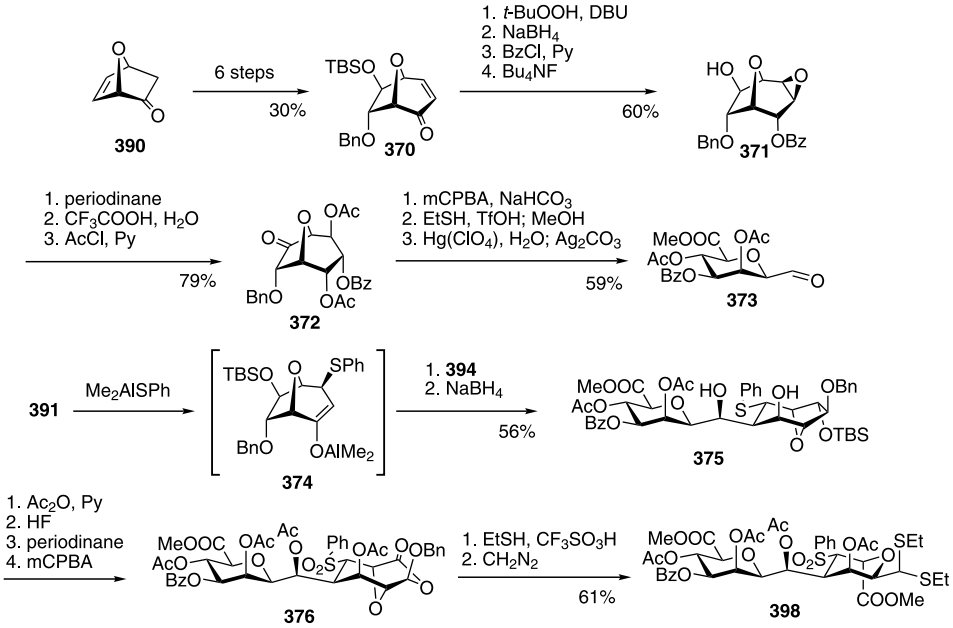
Compound **385** undergoes a $Eu(fod)_3$ catalyzed hetero-Diels–Alder cycloaddition with ethyl vinyl ether, leading to an equimolar mixture of the two *endo*-cycloadducts which prove chromatographically separable after deacetylation and benzylation. Each of the resulting products (**386** and **387**) was subjected to the same series of transformations (conversion of the thiazole ring into a formyl group [144], simultaneous reduction of the aldehyde and hydroboration-oxidation of the alkene from the less hindered face) with a final protecting group exchange (Bn → Ac), leading to two *C*-linked disaccharides **388** and **389**.

Another very different conceptual de novo design was developed by Vogel and coworkers [14], which is based on stereoselective transformations of the rigid oxanorbornane bicyclic system leading to *C*-linked sugar units. Some illustrations are given in *Sect. 3.1.*.

In addition, Gerber and Vogel have described an ionic assemblage of a dimer using optically pure ketone **390** [145] (*Scheme 77*).

Enone **391**, precursor of both units of the dimer and containing the future tetrahydropyranyl ring, is stereoselectively transformed in seven steps to highly oxygenated intermediates **392** and **393**. Regioselective Baeyer–Villiger oxidation, opening of the uronolactone, and dithioacetal hydrolysis provide uronic acid derivative **394**, the electrophilic partner of the coupling step. The aldol condensation between the aluminum enolate **395** derived from enone **391** and aldehyde **394** furnishes the precursor of the *C*-linked dimer **396** after carbonyl reduction. Transformation of **396** by a six-step sequence similar to the sequence leading to **394** provides a protected form of a *C*-linked β-D-mannuronic-(1→3) dimer **398**, a starting point towards *C*-linked oligomers, obtainable by an iterative process.

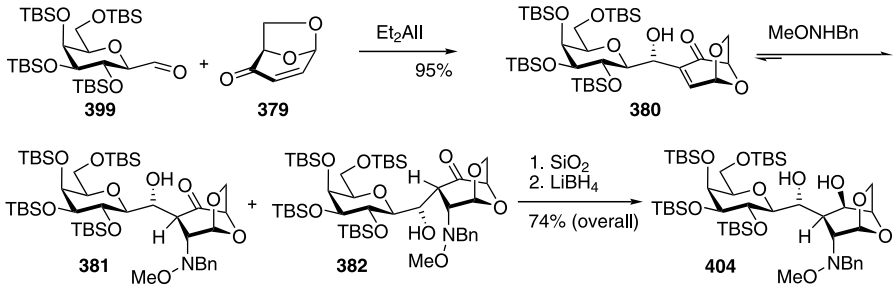
Vogel and coworkers also developed entries into a variety of *C*-linked disaccharides and analogs based on the *Oshima–Nozaki* condensation of levoglucosenone and isolevoglucosenone with monosaccharide derived aldehydes [146]. For instance, a series of *C*-linked β-(1→3)-glucopyranosides of 2- and 4-deoxy-D-hexoses were prepared by elaboration of the



Scheme 77

condensation products obtained by Me_2AlSPh or Et_2AlI induced coupling of glucose-derived aldehyde **43** and isolevoglucosenone or levoglucosenone [147].

This strategy recently culminated in the preparation of a *C*-linked disaccharide analog of the tumor-associated Thomson–Friedereich (T) epitope α-*O*-conjugated to L-serine [148]. Interestingly, final elaboration and coupling to serine proved to be a challenge since the *C*-linked disaccharide did not behave as expected based on the chemistry known for simpler galactose derivatives and *O*-linked disaccharides [149,150]. The final, successful synthetic route is depicted in *Scheme 78*.



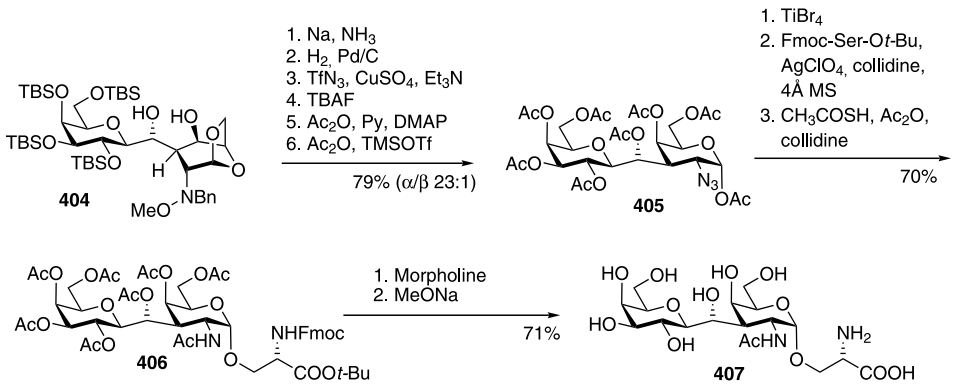
Scheme 78

In the initial C–C bond-forming step, *Oshima–Nozaki* condensation of the persilylated galactose-derived aldehyde **399** with isolevoglucosenone **400**, in the presence of diethylaluminum iodide, leads to enone **401** with 95% yield. Conjugate addition of the highly nucleophilic MeONHBn without solvent or catalyst gives a 1:1 mixture of diastereomeric adducts **402** and **403**, which are isomerized on silica gel allowing the recovery of pure **403** in 82% yield. Diol **404** is finally isolated in 90% yield following hydride addition ($LiBH_4$) on the *exo* face of the bicyclic system. It then proved necessary to transform the methoxy(benzyl)amino group since attempts to pursue the synthesis in its presence led exclusively to β-selective glycosylation of the protected serine derivative. Birch reduction followed by catalytic hydrogenolysis liberates the free amine which is converted into the corresponding azide in 81% yield, with trifluoromethanesulfonyl azide in the presence of cupric sulfate [151] (*Scheme 79*).

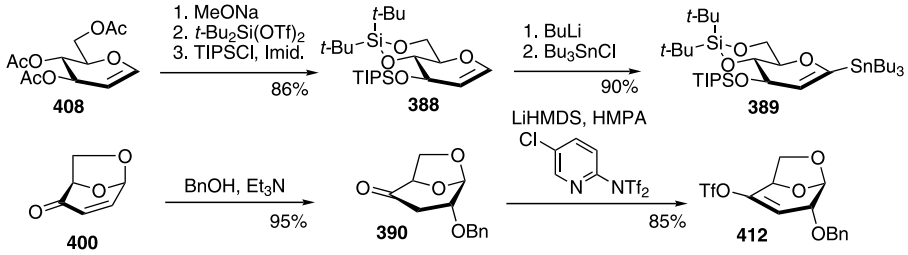
The synthesis was further conducted by desilylation, peracetylation, ring-opening of the 1,6-anhydrogalac-

tose moiety, conversion into the anomeric bromide, and Königs–Knorr glycosylation with the protected serine derivative, under the control of the kinetic anomeric effect (α/β selectivity, 5/1). Final deprotection leads to **407**.

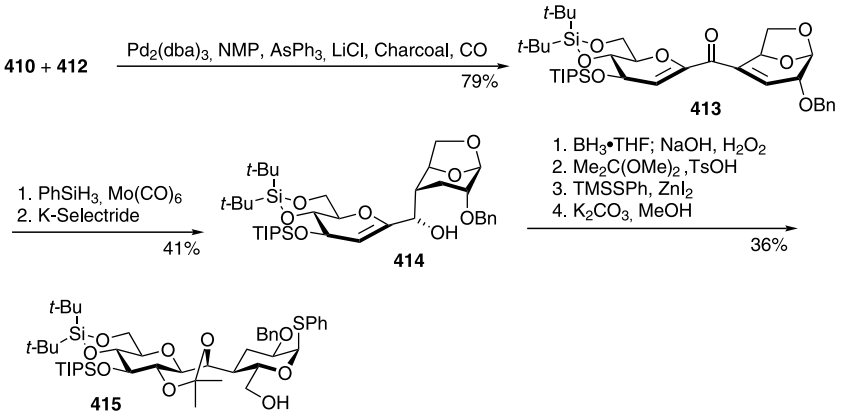
Another approach developed by the Vogel group relies on a carbonylative Stille cross-coupling between stannylated glucal derivative **410** and enol triflate **412** derived from isolevoglucosenone [152] (*Scheme 80*).



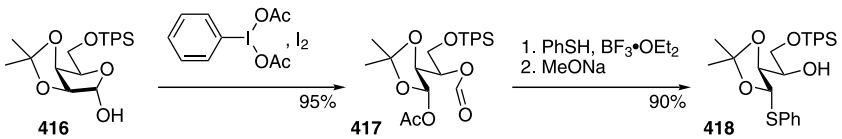
Scheme 79



Scheme 80



Scheme 81



Scheme 82

Coupling of these two readily available precursors leads to bis-enone **413** elaborated to **414** by *chemo-* and *stereo*selective hydrogenation of the more strained, bicyclic alkene with $PhSiH_3$ in the presence of catalytic $Mo(CO)_6$, and K-selectride diastereoselective reduction (*Scheme 81*).

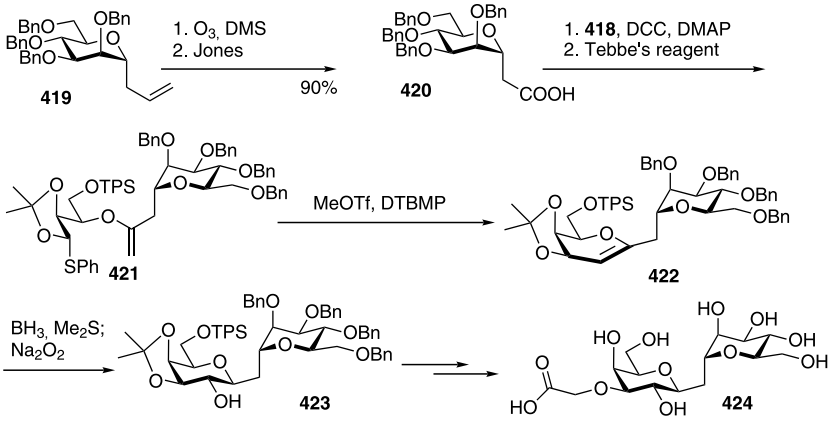
Hydroboration-oxidation occurs on the probably less hindered α-face of the glucal, leading to a diol which is protected as an acetonide prior to conversion of the 1,6-anhydrohexose ring into the corresponding thioglycoside **415**, a potential glycosyl donor.

Thioacetal **418** was used by Mootoo to access several *C*-linked disaccharides [153,154] including the *C*-linked analog of a Sialyl Lewis x mimetic [155] (*Scheme 82*).

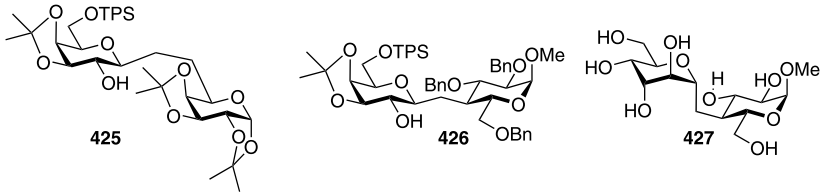
Thioacetal **418** can easily be prepared on a large scale from D-lyxose via intermediate **416**, which is oxidized into **417** by the diacetoxy-iodobenzene/I_2 system [156]. Treatment of **417** with thiophenol and $BF_3{\cdot}OEt_2$ affords **418** after methanolysis of the formate group.

Oxidative cleavage of *C*-allyl mannose **419** [157] following Wong's procedure [158] leads to acid **420** which is esterified with **418** prior to methylenation with Tebbe's reagent (*Scheme 83*).

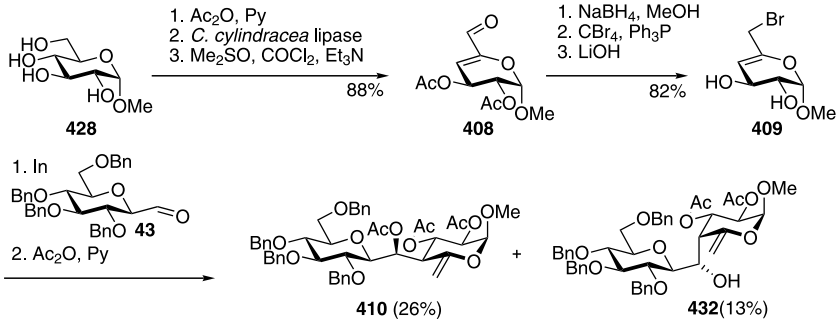
The resulting enol ether intramolecularly traps the oxycarbenium ion obtained by treatment of the thioacetal with TMSOTf, leading after regioselective deprotonation, to the substituted glycal **422**, which gives, after hydroboration-oxidation, *C*-linked glycoside **423**. Further elaboration leads to the desired molecule **424**. This derivative, as well as its *O*-linked counterpart (and more constrained *C*-linked analogs) were tested for P-selectin binding, but they did not show significant variations in activity [159].



Scheme 83

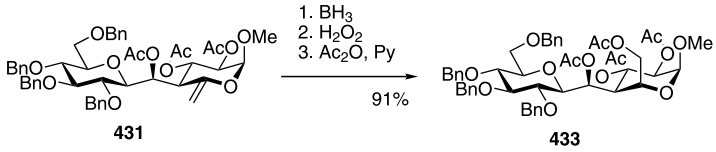


Scheme 84

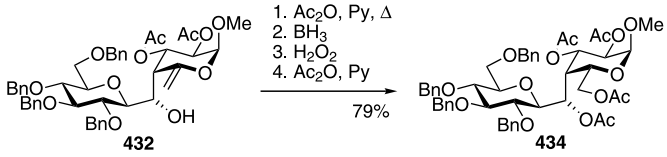


Scheme 85

Other *C*-linked disaccharides have been obtained using this strategy, including Gal-$\beta(1\rightarrow6)$-Gal **425**, Gal-$\beta(1\rightarrow4)$-Glc **426** [153] and Alt-$\alpha(1\rightarrow4)$-Glc [154] **427** (*Scheme 84*).
Lubineau et al. prepared two (1→4) *C*-linked disaccharides via an indium promoted allylation in aqueous medium [160] (*Scheme 85*).
The starting glucose-derived allylic bromide **430** is obtained in 6 steps (acetylation, lipase-catalyzed selective deacetylation of the primary position, Swern oxidation leading directly to



Scheme 86



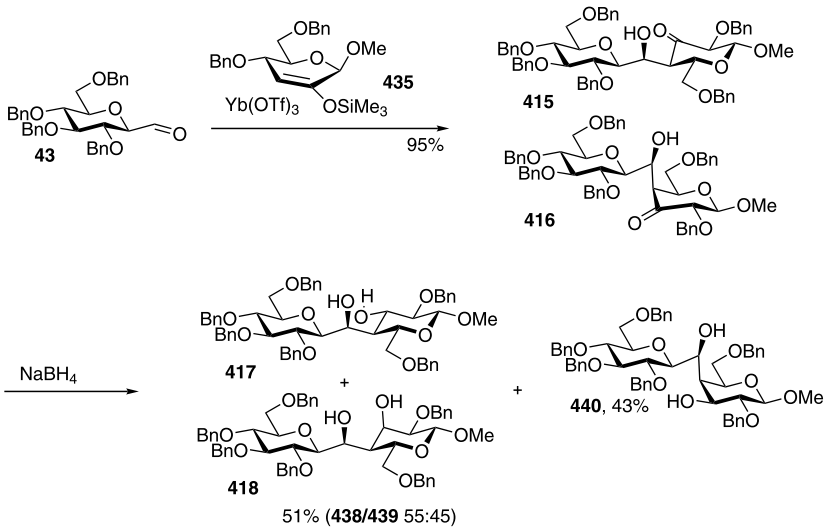
Scheme 87

the 4,5-unsaturated aldehyde, reduction, bromination and final saponification) from α-methyl glucoside **428** in 72% overall yield. Allylation of aldehyde **43** in the presence of indium in a THF/phosphate buffer (2/1) leads to a mixture of *C*-linked disaccharide precursors which can be separated after acetylation (**432** proved to be only partially acetylated). After complete acetylation of **432**, requiring more demanding conditions, the two derivatives were subjected to the same series of transformations, leading after hydroboration-oxidation and acetylation to two *C*-linked disaccharides **433** and **434** (*Scheme 86* and *Scheme 87*).

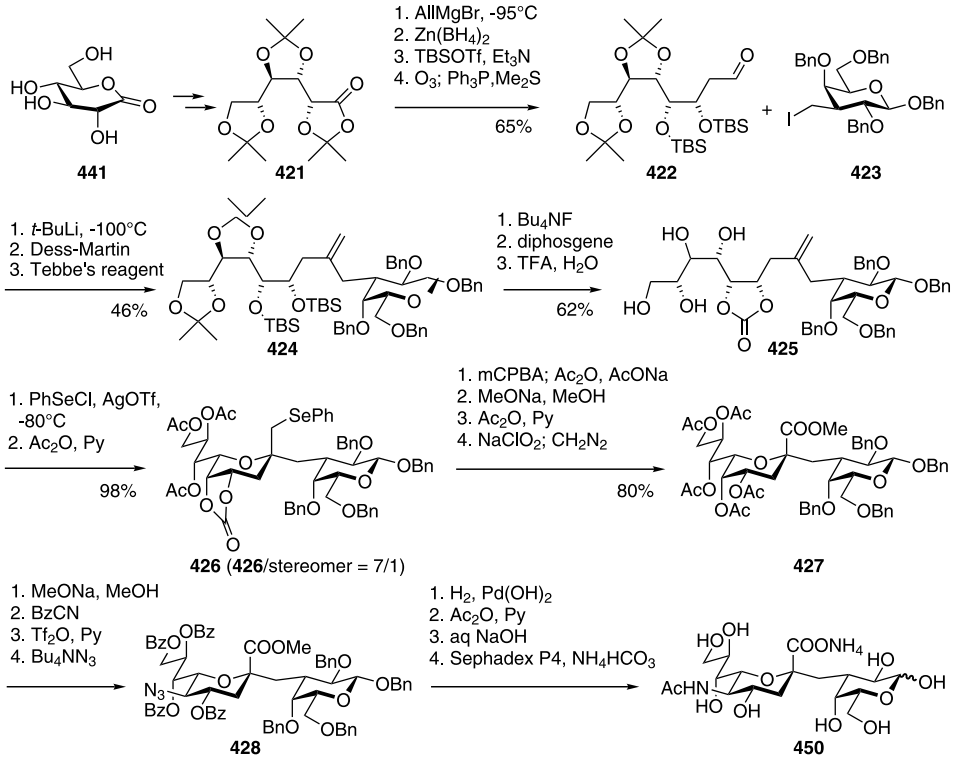
The same group described the $Yb(OTf)_3$ catalyzed Mukaiyama aldol reaction of aldehyde **43** with silyl enol ether **435** in an aqueous medium [161] (*Scheme 88*). This results in the formation of two non-separable diastereomeric aldols **436** and **437** in a 95% combined yield. Finally, $NaBH_4$ reduction leads to isolation of pure **440** and a mixture of **438** and **439**.

Instead of using "Neu5Ac-*C*-glycosyl donors" as reported in *Sect. 2.7* for the synthesis of *C*-linked *N*-acetylneuraminic acid-containing disaccharides, another alternative, described by Schmidt's group, is a de novo construction of the Neu5Ac moiety. This is exemplified by the synthesis of a *C*-glycosyl analog of sialyl-α-(2→3)-D-galactose starting from D-gluconolactone [162] (*Scheme 89*).

The key step of the strategy is a low temperature phenylselenyl-mediated 6-*exo-trig* electrophilic cyclization of acyclic precursor **446**, elaborated by the coupling between the lithium reagent derived from iodide **444** and aldehyde **443**, itself produced from D-glucono-δ-lactone **441**. A selena-Pummerer rearrangement, oxidation and esterification sequence delivers methyl ester **448** which, after deacetylation, is regioselectively benzoylated with benzoyl cyanide. Triflate activation and nucleophilic displacement by azide ion afforded azide **449**, which was reduced with hydrogen and Pearlman's catalyst with simultaneous removal of the benzyl ethers. The following three steps successfully complete this de novo synthesis of the *C*-linked Neu5Ac-α-(2→3)-D-Gal disaccharide.



Scheme 88



Scheme 89

Acknowledgement

Financial support from the Ministère de l'Enseignement Supérieur et de la Recherche and the Centre National de la Recherche Scientifique is gratefully acknowledged.

References

1. Hanessian S, Pernet AG (1976) Adv Carbohydr Chem Biochem 33:111
2. Hacksell U, Daves GD, Jr (1985) Prog Med Chem 22:1
3. Herscovici J, Antonakis K (1992) In: Atta-ur-Rahman (ed) Studies in Natural Products Chemistry, vol 10. Elsevier, Amsterdam
4. Postema MHD (1992) Tetrahedron 48:8545; Postema MHD (1995) *C*-Glycoside Synthesis. CRC Press, London
5. Watanabe KA (1994) In: Townsend LB (ed) Chemistry of Nucleosides and Nucleotides, vol 3. Plenum Press, New York
6. Jaramillo C, Knapp S (1994) Synthesis 1
7. Casiraghi G, Zanardi F, Rassu G, Spanu P (1995) Chem Rev 95:1677
8. Levy DE, Tang C (1995) The Chemistry of *C*-Glycosides. Pergamon, Oxford
9. Bertozzi C, Bednarski M (1996) In: Khan SH, O'Neill RA (eds) Modern Methods in Carbohydrate Synthesis. Harwood Academic, Amsterdam
10. Beau J-M, Gallagher T (1997) Topics Curr Chem 187:1
11. Nicotra F (1997) Topics Curr Chem 187:55; Nicotra F (1998) In: Chapleur Y (ed) Carbohydrate Mimics. Wiley-VCH, Weinheim, p 67
12. Giese B, Zeitz HG (1997) In: Hanessian S (ed) Preparative Carbohydrate Chemistry. Marcel Dekker, New York, p 507
13. Suzuki K, Matsumoto T (1997) In: Hanessian S (ed) Preparative Carbohydrate Chemistry. Marcel Dekker, New York, p 527
14. Vogel P, Ferritto R, Kraehenbuehl K, Baudat A (1998) In: Chapleur Y (ed) Carbohydrate Mimics. Wiley-VCH, Weinheim, p 19
15. Paton RM (1998) In: Chapleur Y (ed) Carbohydrate Mimics. Wiley-VCH, Weinheim, p 49
16. Togo H, He W, Waki Y, Yokoyama M (1998) Synlett 700
17. Du Y, Linhardt RJ, Vlahov IR (1998) Tetrahedron 54:9913
18. Skrydstrup T, Vauzeilles B, Beau J-M (2000) In: Ernst B, Sinaÿ P, Hart G (eds) Oligosaccharides in Chemistry and Biology - A Comprehensive Handbook, Vol. 1. Wiley-VCH, Weinheim, p 495
19. Dondoni A, Marra A (2000) Chem Rev 100:4395
20. Somsák L (2001) Chem Rev 101:81-136
21. Liu L, McKee M, Postema MHD (2001) Curr Org Chem 5:1133
22. Yuan X, Linhardt RJ (2005) Curr Top Med Chem 5:1393
23. Postema MHD, Piper JL, Betts RL (2005) Synlett 9:1345
24. Aebischer B, Vasella A, Weber H-P (1982) Helv Chim Acta 65:621; Aebischer B, Vasella A, Weber H-P (1983) Helv Chim Acta 66: 789; Meuwly R, Vasella A (1986) Helv Chim Acta 69:751
25. Armstrong RW, Beau J-M, Cheon SH, Christ WS, Fujioka H, Ham W-H, Hawkins LD, Jin H, Kang SH, Kishi Y, Martinelli MJ, McWhorter WW, Jr, Mizumo M, Nakata M, Stutz AE, Talamas FX, Taniguchi M, Tino JA, Ueda K, Uenishi J-I, White JB, Yonaga M (1989) J Am Chem Soc 111:7525, 7530
26. Myers AG, Gin DY, Rogers DH (1994) J Am Chem Soc 116:4697
27. Nicolaou KC (1996) Angew Chem Int Ed Engl 35:589
28. Evans DA, Trotter BW, Côté B, Coleman PJ, Dias LC, Tyler AN (1997) Angew Chem Int Ed Engl 36:2744
29. Rouzaud D, Sinaÿ P (1983) J Chem Soc Chem Commun 1353
30. Lewis MD, Cha JK, Kishi Y (1982) J Am Chem Soc 104:4976
31. Lancelin J-M, Amvam Zollo PH, Sinaÿ P (1983) Tetrahedron Lett 24:4833
32. Xin Y-C, Zhang Y-M, Mallet J-M, Glaudemans CPJ, Sinaÿ P (1999) Eur J Org Chem 471
33. Wang J, Kovác P, Sinaÿ P, Glaudemans CPJ (1998) Carbohydr Res 308:191
34. Leeuwenburg MA, Timmers CM, van der Marel GA, van Boom JH, Mallet J-M, Sinaÿ PG (1997) Tetrahedron Lett 38:6251

35. Nicholas KM (1987) Acc Chem Res 20:207
36. Tanaka S, Isobe M (1994) Tetrahedron 50:5633
37. Aslam T, Fuchs MGG, Le Formal A, Wightman RH (2005) Tetrahedron Lett 46:3249
38. Dondoni A, Boscarato A, Zuurmond H (1996) Tetrahedron Lett 37:7587
39. Dondoni A, Scherrmann M-C (1994) J Org Chem 59:6404
40. Dondoni A, Mizuno M, Marra A (2000) Tetrahedron Lett 41:6657; Dondoni A, Marra A, Mizuno M, Giovannini PP (2002) J Org Chem 67:4186
41. Dondoni A, Zuurmond H, Boscarato A (1997) J Org Chem 62:8114
42. Dondoni A, Formaglio P, Marra A, Massi A (2001) Tetrahedron 57:7719
43. Dondoni A, Marra A (2003) Tetrahedron Lett 44:4067
44. Dondoni A, Kleban M, Zuurmond H, Marra A (1998) Tetrahedron Lett 39:7991; Dondoni A, Marra A, Mizuno M, Giovannini PP (2002) J Org Chem 67:4186
45. Preuss R, Schmidt RR (1991) J Carbohydr Chem 10:887
46. Martin OR, Lai W (1990) J Org Chem 55:5188
47. Sowden JC (1951) Adv Carbohydr Chem 6:291; Aebisher B, Bieri JH, Prewo R, Vasella A (1982) Helv Chim Acta 65:2251
48. Spak SJ, Martin OR (2000) Tetrahedron 56:217
49. Kobertz WR, Bertozzi CR, Bednarski MD (1996) J Org Chem 61:1894
50. Gurjar MK, Nagaprasad R, Ramana CV (2002) Tetrahedron Lett 43:7577
51. (a) Griffin FK, Murphy PV, Paterson DE, Taylor RJK (1998) Tetrahedron Lett 39:8179; (b) Belica PS, Franck RW (1998) Tetrahedron Lett 39:8225; (c) Taylor RJK, McAllister GD, Franck RW (2006) Carbohydr Res 341:1298
52. Griffin FK, Paterson DE, Murphy PV, Taylor RJK (2002) Eur J Org Chem 130
53. Alcaraz M-L, Griffin FK, Paterson DE, Taylor RJK (1998) Tetrahedron Lett 39:8183
54. Griffin FK, Paterson DE, Taylor RJK (1999) Angew Chem Int Ed 38:2939
55. Paterson DE, Griffin FK, Alcarez M-L, Taylor RJK (2002) Eur J Org Chem 1323
56. Csuk R, Glänzer BI (1991) Tetrahedron 47:1655
57. Meyers CY, Malte AM, Matthews WS (1969) J Am Chem Soc 91:7510
58. Chan T-L, Fong S, Li Y, Man T-O, Poon C-D (1994) J Chem Soc Chem Commun 1771
59. Jeanmart S, Taylor RJK (2005) Tetrahedron Lett 46:9043
60. McAllister GD, Paterson DE, Taylor RJK (2003) Angew Chem Int Ed 42:1387
61. Lesimple P, Beau, J-M, Jaurand G, Sinaÿ P (1986) Tetrahedron Lett 27:6201
62. Nicolaou KC, Hwang C-K, Duggan ME (1986) J Chem Soc Chem Commun 925
63. Hanessian S, Martin M, Desai RC (1986) J Chem Soc Chem Commun 926
64. Preuss R, Schmidt RR (1989) Liebigs Ann Chem 429
65. Schmidt RR, Preuss R (1989) Tetrahedron Lett 30:3409; Preuss R, Schmidt RR (1991) J Carbohydr Chem 10:887
66. Schmidt RR, Beyerbach A (1992) Liebigs Ann Chem 983
67. Eisele T, Ishida H, Hummel G, Schmidt RR (1995) Liebigs Ann Chem 2113
68. Daley SM, Armstrong RW (1989) Tetrahedron Lett 30:5713
69. Alzeer J, Cai C, Vasella A (1995) Helv Chim Acta 78:242
70. Patro B, Schmidt RR (1998) Synthesis 1731
71. Dietrich H, Schmidt RR (1994) Liebigs Ann Chem 975
72. Mazéas D, Skrydstrup T, Beau, J-M (1995) Angew Chem Int Ed Engl 34:909
73. Skrydstrup T, Jarreton O, Mazéas D, Urban D, Beau J-M (1998) Chem Eur J 5:430
74. Jarreton O, Skrydstrup T, Beau J-M (1996) J Chem Soc Chem Commun 1661
75. Jarreton O, Skrydstrup T, Espinosa J-F, Jiménez-Barbero J, Beau J-M (1999) Chem Eur J 5:430
76. Jung ME, Choe SWT (1993) Tetrahedron Lett 34:6247
77. Mikkelsen LM, Skrydstrup T (2003) J Org Chem 68:2123
78. Krintel SL, Jiménez-Barbero J, Skrydstrup T. (1999) Tetrahedron Lett 40:7565
79. Mikkelsen LM, Krintel SL, Skrydstrup T (2000) Chem Commun 2319
80. Mikkelsen LM, Krintel SL, Jiménez-Barbero J, Skrydstrup T (2002) J Org Chem 67:6297
81. Mikkelsen LM, Hernaíz, MJ, Martín-Pastor M, Skrydstrup T, Jiménez-Barbero J (2002) J Am Chem Soc 124:14940
82. Hung S-C, Wong C-H (1996) Tetrahedron Lett 37:4903
83. Hung S-C, Wong C-H (1996) Angew Chem Int Ed Engl 35:2671
84. Miquel N, Doisneau G, Beau J-M (2000) Angew Chem Int Ed 39:4111

85. Girard P, Namy JL, Kagan HB (1980) J Am Chem Soc 102:2693; Machrouhi F, Hamann B, Namy JL, Kagan HB (1996) Synlett 633
86. Uchiyama T, Hindsgaul O (1996) Synlett 499
87. Miquel N, Doisneau G, Beau J-M (2000) Chem Commun 2347
88. Urban D, Skrydstrup T, Riche C, Chiaroni A, Beau J-M (1996) J Chem Soc Chem Commun 1883
89. Urban D, Skrydstrup T, Beau J-M (1998) J Org Chem 63:2507
90. Andersen L, Mikkelsen LM, Beau J-M, Skrydstrup T (1998) Synlett 1393
91. Palmier S, Vauzeilles B, Beau J-M (2003) Org Biomol Chem 1:1097
92. Vlahov IR, Vlahova PI, Linhardt RJ (1997) J Am Chem Soc 119:1480; Du Y, Linhardt RJ (1998) Carbohydr Res 308:161; Polat T, Du Y, Linhardt RJ (1998) Synlett 1195
93. Du Y, Polat T, Linhardt RJ (1998) Tetrahedron Lett 39:5007
94. Bazin HG, Du Y, Polat T, Linhardt RJ (1999) J Org Chem 64:7254; Poveda A, Asensio JL, Polat T, Bazin H, Linhardt RJ, Jiménez-Barbero J (2000) Eur J Org Chem 1805
95. Kuberan B, Sikkander SA, Tomiyama H, Linhardt RJ (2003) Angew Chem Int Ed 42:2073
96. Abdallah Z, Doisneau G, Beau J-M (2003) Angew Chem Int Ed 42:5209; Malapelle A, (2006) PhD Thesis, Université Paris-Sud Orsay
97. Gao Y, Sharpless KB (1988) J Am Chem Soc 110:7538
98. Vargas-Berenguel A, Santoyo-González F, Calvo-Asín JA, Calvo-Flores FG, Expósito-López JM, Hernández-Mateo F, Isac-García J, Giménez Martínez JJ (1998) Synthesis 1778; Dagron F, Lubineau A (2000) J Carbohydr Chem 19:311
99. García-Aparicio V, Malapelle A, Abdallah Z, Doisneau G, Santos I, Asensio JL, Cañada FJ, Beau J-M, Jiménez-Barbero J (2007) Carbohydr Res 342:1974
100. Malapelle A, Abdallah Z, Doisneau G, Beau J-M (2006) Angew Chem Int Ed 45:6016, Malapelle A, Coslovi A, Doisneau G, Beau J-M (2007) Eur J Org Chem 3145
101. Paton RM, Penman KJ (1994) Tetrahedron Lett 35:3163
102. Dawson IM, Johnson T, Paton RM, Rennie RAC (1988) J Chem Soc Chem Commun 1339
103. Giese B, Witzel T (1986) Angew Chem Int Ed Engl 25:450; Giese B, Hosh M, Lamberth C, Schmidt RR (1988) Tetrahedron Lett 29:1375
104. Bimwala RM, Vogel P (1991) Tetrahedron Lett 32:1429
105. Bimwala RM, Vogel P (1992) J Org Chem 57:2076
106. Ferrito R, Vogel P (1994) Tetrahedron Asymmetry 29:2077
107. Pasquarello C, Demange R, Vogel P (1999) Bioorg Med Chem Lett 9:793
108. Pasquarello C, Picasso, S, Demange R, Malissard M, Berger EG, Vogel P (2000) J Org Chem 65:4251
109. Martin OR, Xie F, Kakarla R, Benhamza R (1993) Synlett 165
110. Witczak ZJ, Chhabra R, Chojnacki J (1997) Tetrahedron Lett 38:2215
111. Sinaÿ P (1997) Pure & Appl. Chem 69: 459; Sinaÿ P (1998) Pure & Appl Chem 70:407
112. (a) Xin Y-C, Mallet J-M, Sinaÿ P (1993) J Chem Soc Chem Commun 864; (b) Mallet A, Mallet J-M, Sinaÿ P (1994) Tetrahedron Asymmetry 5:2593
113. Rubinstenn G, Mallet J-M, Sinaÿ P (1998) Tetrahedron Lett 39:3697
114. Vauzeilles B, Sinaÿ P (2001) Tetrahedron Lett 42:7269
115. Chénedé A, Rekaï E, Perrin E, Sinaÿ P (1994) Synlett 420
116. Mazéas D, Skrydstrup T, Doumeix O, Beau J-M (1994) Angew Chem Int Ed Engl 33:1383; Skrydstrup T, Mazéas D, Elmouchir M, Doisneau G, Riche C, Chiaroni A, Beau J-M (1997) Chem Eur J 3:1342
117. Chénedé A (1995) PhD Thesis, Université Pierre et Marie Curie
118. Vidal P, Vauzeilles B, Blériot Y, Sollogoub M, Sinaÿ P, Jiménez–Barbero J, Espinosa JF (2007) Carbonhydr Res 342:1910
119. Ito Y, Ogawa T (1994) Angew Chem Int Ed Engl 33:1765; Dan A, Ito Y, Ogawa T (1995) J Org Chem 60:4680
120. Rekaï ED, Rubinstenn G, Mallet J-M, Sinaÿ P (1998) Synlett 831
121. Rubinstenn G, Esnault J, Mallet J-M, Sinaÿ P (1997) Tetrahedron Asymmetry 8:1327
122. Rubinstenn G (1996) PhD Thesis, Université Pierre et Marie Curie
123. Helmboldt A, Mallet J-M, Petitou M, Sinaÿ P (1997) Bull Soc Chim Fr 134:1057; Helmboldt A, Petitou M, Mallet J-M, Hérault J-P, Lormeau J-C, Driguez PA, Herbert J-M, Sinaÿ P (1997) Bioorg Med Chem Lett 7:1507; Petitou M, Hérault J-P, Lormeau J-C, Helmboldt A,

Mallet J-M, Sinaÿ P, Herbert J-M (1998) Bioorg Med Chem 6:1509
124. Bazan GC, Khosravi E, Schrock RR, Feast WJ, Gibson VC, O'Regan MB, Thomas JK, Davis WM (1990) J Am Chem Soc 112:8378; Schrock RR, Murdzeck JS, Bazan GC, Robbins J, DiMare M, O'Regan M (1990) J Am Chem Soc 112:3875
125. Schwab P, France MB, Ziller JW, Grubbs RH (1995) Angew Chem Int Ed 34:2039; Schwab P, Grubbs RH, Ziller JW (1996) J Am Chem Soc 118:100
126. Scholl M, Ding S, Lee CW, Grubbs RH (1999) Org Lett 1:953
127. Calimente D, Postema MHD (1999) J Org Chem 64:1770
128. Postema MHD, Piper JL, Liu L, Shen J, Faust M, Adreana P (2003) J Org Chem 68:4748; Liu L, Postema MHD (2001) J Am Chem Soc 123:8602
129. Postema MHD, Calimente D, Liu L, Behrmann TL (2000) J Org Chem 65:6061; Postema MHD, Calimente D (1999) Tetrahedron Lett 40:4755
130. Postema MHD, Piper JL, Komanduri V, Liu L (2004) Angew Chem Int Ed 43:2915
131. Piper JL, Postema MHD (2004) J Org Chem 69:7395
132. Chang GX, Lowary TL (2006) Tetrahedron Lett 47:4561
133. Babirad SA, Wang Y, Kishi Y (1987) J Org Chem 52:1372
134. Jin H, Uenishi J, Christ WJ, Kishi Y (1986) J Am Chem Soc 108:5644; Takai K, Tagashira M, Kuroda T, Oshima K, Utimoto K, Nozaki H (1986) J Am Chem Soc 108:6048
135. Goekjian PG, Wu T-C, Kang H-Y, Kishi Y (1987) J Org Chem 52:4823
136. Dyer UC, Kishi Y (1988) J Org Chem 53:3384
137. Carcano M, Nicotra F, Panza L, Russo G (1989) J Chem Soc Chem Commun 642
138. Wei A, Kishi Y (1994) J Org Chem 59:88
139. Haneda T, Goekjian PG, Kim SH, Kishi Y (1992) J Org Chem 57:490
140. Wei A, Haudrechy A, Audin C, Jun H-S, Haudrechy-Bretel N, Kishi Y (1995) J Org Chem 60:2160
141. Armstrong RW, Sutherlin DP (1994) Tetrahedron Lett 35:7743
142. Sutherlin DP, Armstrong RW (1996) J Am Chem Soc 118:9802; Sutherlin DP, Armstrong RW (1997) J Org Chem 62:5267
143. Stepanek P, Kniezo L, Dvorakova H, Vojtisek P (2003) Synlett 963; Stepanek P, Vich O, Kniezo L, Dvorakova H, Vojtisek P (2004) Tetrahedron Asymmetry 15:1033
144. Dondoni A, Marra A, Schermann M-C, Bertolasi V (2001) Chem Eur J 7:1371
145. Gerber P, Vogel P (1999) Tetrahedron Lett 40:3165
146. Awad L, Demange R, Zhu Y-H, Vogel P (2006) Carbohydr Res 341:1235
147. Demange R, Bühlmann C, Vogel P (2003) Helv Chim Acta 86:361
148. Awad L, Riedner J, Vogel P (2005) Chem Eur J 11:3565
149. Zhu Y-H, Vogel P (2001) Synlett 79
150. Demange R, Awad L, Vogel P (2004) Tetrahedron Asymmetry 15:3573
151. Nyffeler PT, Liang C-H, Koeller KM, Wong C-H (2002) J Am Chem Soc 124:10773
152. Steunenberg P, Jeanneret V, Zhu Y-H, Vogel P (2005) Tetrahedron Asymmetry 16:337
153. Khan N, Cheng X, Mootoo D (1999) J Am Chem Soc 121:4918
154. Denton RW, Mootoo DR (2003) J Carbohydr Chem 22:671
155. Cheng X, Khan N, Mootoo D (2000) J Org Chem 65:2544
156. De Armas P, Francisco CG, Suarez E (1992) Angew Chem Int Ed Engl 31:772
157. Hosomi A, Sakata Y, Sakurai H (1984) Tetrahedron Lett 25:2383
158. Wong CH, Moris-Varas F, Hung S-C, Marron TG, Lin C-C, Gong KW, Weitz-Schmidt G (1997) J Am Chem Soc 119:8152
159. Denton RW, Cheng X, Tony KA, Dilhas A, Hernandez JJ, Canales A, Jimenez-Barbero J, Mootoo DR (2007) Eur J Org Chem 645
160. Levoirier E, Canac Y, Norsikian S, Lubineau A (2004) Carbohydr Res 339:2737
161. Zeitouni J, Norsikian S, Merlet D, Lubineau A (2006) Adv Synth Catal 348:1662
162. Notz W, Hartel C, Waldscheck B, Schmidt RR (2001) J Org Chem 66:4250

9.5 Oligosaccharide Mimetics

Hans Peter Wessel, *Susana Dias Lucas*
Pharmaceutical Research, Discovery Chemistry, F. Hoffmann-La Roche Ltd.,
4070 Basel, Switzerland
hans_p.wessel@roche.com

Abstract

The important roles of oligosaccharides in physiological and pathophysiological processes have spurred the development of mimetics. Oligosaccharide mimetics discussed in this chapter may possess a linker of two or more atoms such as amide or urea groups that may lead to isosteric linkage replacements but mostly do not. Larger groups that replace a full sugar unit we refer to as spacers and have grouped molecules with flexible acyclic spacers and more rigid cyclic spacers. The employment of pharmacophore models has led to oligosaccharide mimetics with only one sugar unit left or finally without any saccharide unit as exemplified in mimotopes.

Keywords

Glycomimetics; Linked sugars; Spaced sugars; Functional mimetics; Saccharide-peptide hybrids; Sugar amino acids; Solid phase synthesis; Sialyl Lewis^x; Carbohydrate medicinal chemistry

Abbreviations

CDMT 2-chloro-4,6-dimethoxy-1,3,5-triazine
Gum glucosyl–uronic acid–methyl amine
M6P mannose 6-phosphate

1 Introduction

In this chapter mainly molecules will be discussed that contain a sugar linking unit which deviates from the usual linking pattern, the one-atom linkage via oxygen, or in which sugars

In: *Glycoscience*. Fraser-Reid B, Tatsuta K, Thiem J (eds)
Chapter-DOI 10-1007/978-3-540-30429-6_52:

Scheme 1

have been replaced by a non-sugar structural unit. In a simplified representation (*Scheme 1*) these are sugar molecules containing a linker consisting of more than one atom or a spacer group.

Not discussed in this but in previous chapters are the *structural* mimetics: sugars, in which formally minor changes have been made, such as replacement of the ring oxygen by a nitrogen atom (imino sugars), a sulfur atom or carbon atom (carbasugars) or the replacement of the glycosidic oxygen by a carbon atom (*C*-glycosides) or sulfur atom (thioglycosides). In contrast, the mimetics under discussion here, such as the "spaced sugars", are usually planned to be *functional* oligosaccharide mimetics, i. e., molecules which typically display the activity or property of a carbohydrate.

Many natural products will essentially not be regarded here although bis-glycosides may be viewed as spaced sugars in which the aglycone is the spacer. Interestingly, it is in most cases not clear which role the aglycone plays and whether this non-sugar moiety is possibly a sugar mimetic. Examples may be the saponins from medicinal foodstuffs [1] or the triterpenoid glycosides from medicinal plants [2].

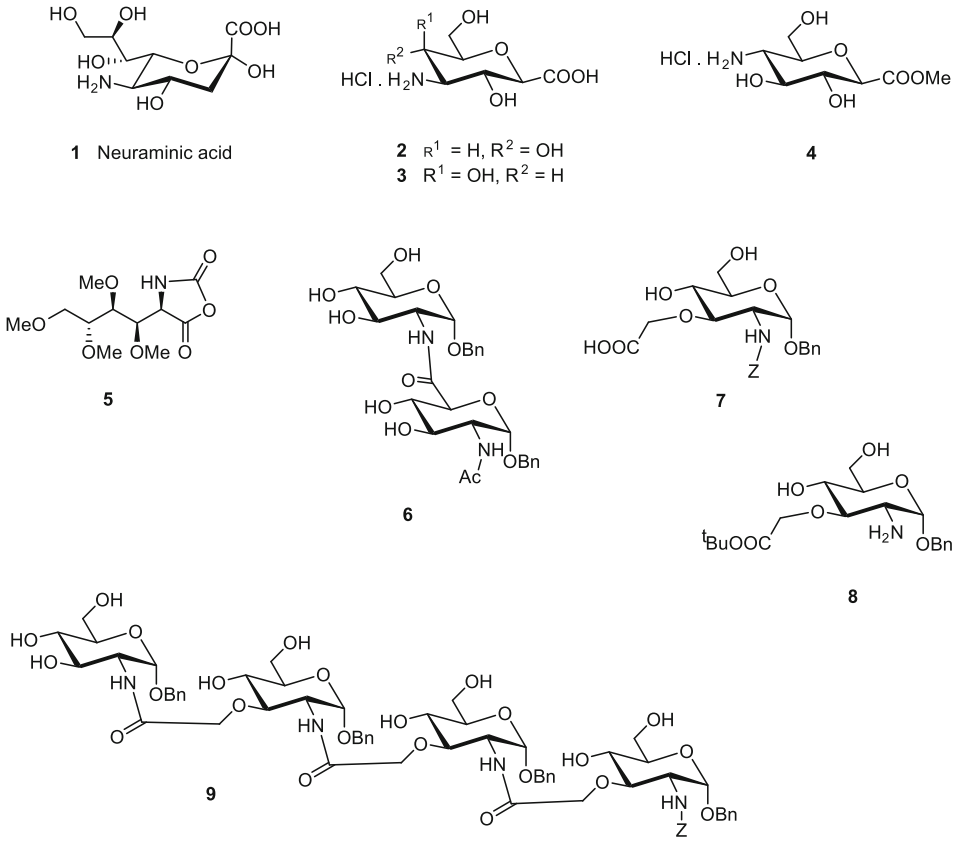
2 More-Atom Linked Sugars

A linker we call a relatively small group of at least two atoms between two sugar units. A "more-atom linked sugar" is thus a spaced saccharide in which the spacer is different in length from the interglycosidic oxygen or its isosteric replacements as in *C*- or *S*-glycosides and definitely smaller than a sugar unit. Phosphate-linked will not be in the scope of this review.

2.1 Amide-Linked Sugars

A formal replacement of the interglycosidic oxygen atom by an amide group leads to amide-linked sugars (*Scheme 2*). This type of structure is suited to build up oligomers or polymers as oligosaccharide or polysaccharide mimetics, respectively, the monomer unit being a sugar amino acid. Sugar amino acids occur in Nature; one example is neuraminic acid **1**

Scheme 2



Scheme 3

(*Scheme 3*) abundant on glycoproteins or certain complex nucleoside antibiotics such as ezomycin [3].

First synthetic reports on amide-linked sugars stem from the mid-1970s. Fuchs and Lehmann [4] prepared the amino-D-*glycero*-L-*manno*- and -D-*glycero*-D-*gulo*- heptonic acids **2** and **3** (*Scheme 3*) and pointed out their potential for polymerization. The isomeric 5-amino-2,6-anhydro-5-deoxy-D-*glycero*-D-*gulo*-heptonic acid methyl ester **4** [5] was submitted to oligomerization under basic conditions and resulted, depending on the reaction time, in oligomers or polymers which, however, were not individually characterized. At about the same time, Yoshimura et al. [6] synthesized amide-linked disaccharides such as **6** as analogues of cell wall components or certain antibiotics. 3,4-Di-*O*-acetyl-6-*O*-carboxymethylglucosaminyl bromide was polymerized with diphenylphosphoryl azide and trimethylamine to give, after deblocking and hydrolysis of the bromide, a water-soluble reducing polymer with a mean molecular mass of 150 kD [7], and the *N*-carboxyanhydride derivative of a methylated 2-amino-2-deoxy-gluconic acid **5** was polymerized in the presence of triethylamine to give polyamides of an average molecular mass of 10 kD [8].

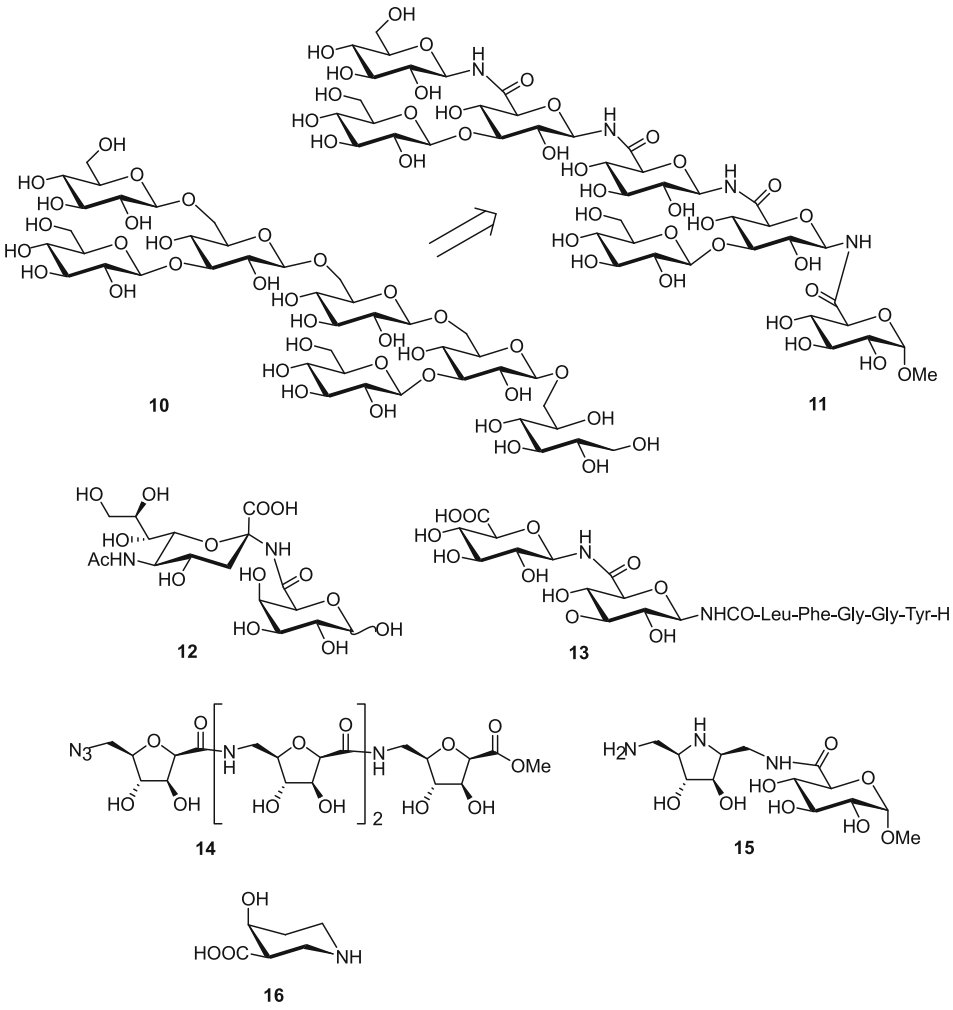
With an interest in oligosaccharide mimetics for pharmaceutical applications, a Roche group prepared the first amide-linked oligomers in a controlled fashion from suitably protected sugar amino acids to construct a tetramer in a [2 + 2] block synthesis. A standard peptide synthesis strategy in solution was applied in which **7** and **8** (*Scheme 3*), both equipped with an acetic acid linker and prepared in five steps from readily available glucosamine, were coupled using a mixed anhydride. The tetramer **9** with four-atom linkers was obtained after activation of a dimer with CDMT [9]. Notably, no protection of hydroxyl groups was required employing this approach.

This type of carbohydrate mimetic was coined "saccharide-peptide hybrid" to point out that carbohydrate properties as well as peptide properties are represented. The mimetics still retain the carbohydrate epitopes available for binding, the pyranose or furanose ring plus hydroxy groups, as well as the peptide function.

A number of groups became interested in this class of compounds, and the saccharide-peptide hybrids were coined "carbopeptoids" [10], "glycotides" [11], "peptidosaccharides" [12], or "saccharopeptides" [13]. In comparison to the formation of oligosaccharides, the complexity in the synthesis of saccharide-peptide hybrids is reduced since no α/β-anomers can be formed. In addition, the oligomerization can be carried out without protecting the hydroxy groups [14] since the amino function is more nucleophilic. We have demonstrated that the oligomerization can also be carried out on a solid phase [14], which opened the way to combinatorial approaches [15,16,17].

If the sugar amino acid is a glycuronopyranosylamine, the oligomers are isosters of (1→6)-linked sugars. This type of mimetic has been investigated in various contexts. A very interesting example has been contributed by van Boom's group [18] who have prepared a branched saccharide-peptide hybrid **11** analogous to the natural phytoalexin elicitor **10** (*Scheme 4*). Disappointingly, this mimetic did not display biological activity, which was attributed to the decreased flexibility at the amide bonds. Sabesan [12] targeted the α-D-Neu5Ac-(2→6)-β-D-Gal*p* disaccharide, which occurs naturally in glycoproteins and glycolipids and mimics of which may be neuraminidase inhibitors. He prepared the amide-linked sialoside **12** as well as analogues replacing the galactose unit. With a view to more selective analgesics, Drouillat et al. [19] worked on glucuronic acid modified enkephalins; a solid-phase synthesis yielded an amide-linked glucuronic acid disaccharide mimetic and its peptide conjugate **13**. In a mouse model, this sugar dimer modified leucin enkephalinamide was slightly more active than the Leu- enkephalinamide standard. D'Onofrio et al. [20] prepared by solid-phase synthesis a tetrasaccharide mimetic composed of two (1→6)-amide-linked lactose units conjugated to an oligonucleotide; the presence of this mimetic at the 3′-end of antisense 18-mer sequences enhanced the stability of the oligonucleotides in bovine fetal serum without negatively interfering with their ability to form stable duplexes with complementary DNA strands. The Fleet group has investigated various saccharide-peptide hybrids based on furanoid sugars; their tetramer **14** [21] can be seen as an analogue of (1→6)-linked hexofuranosides. An analogue of glycosidase inhibitors of the imino sugar family, the (1→6)-amide-linked pyrrolidine disaccharide mimetic **15**, was reported [22].

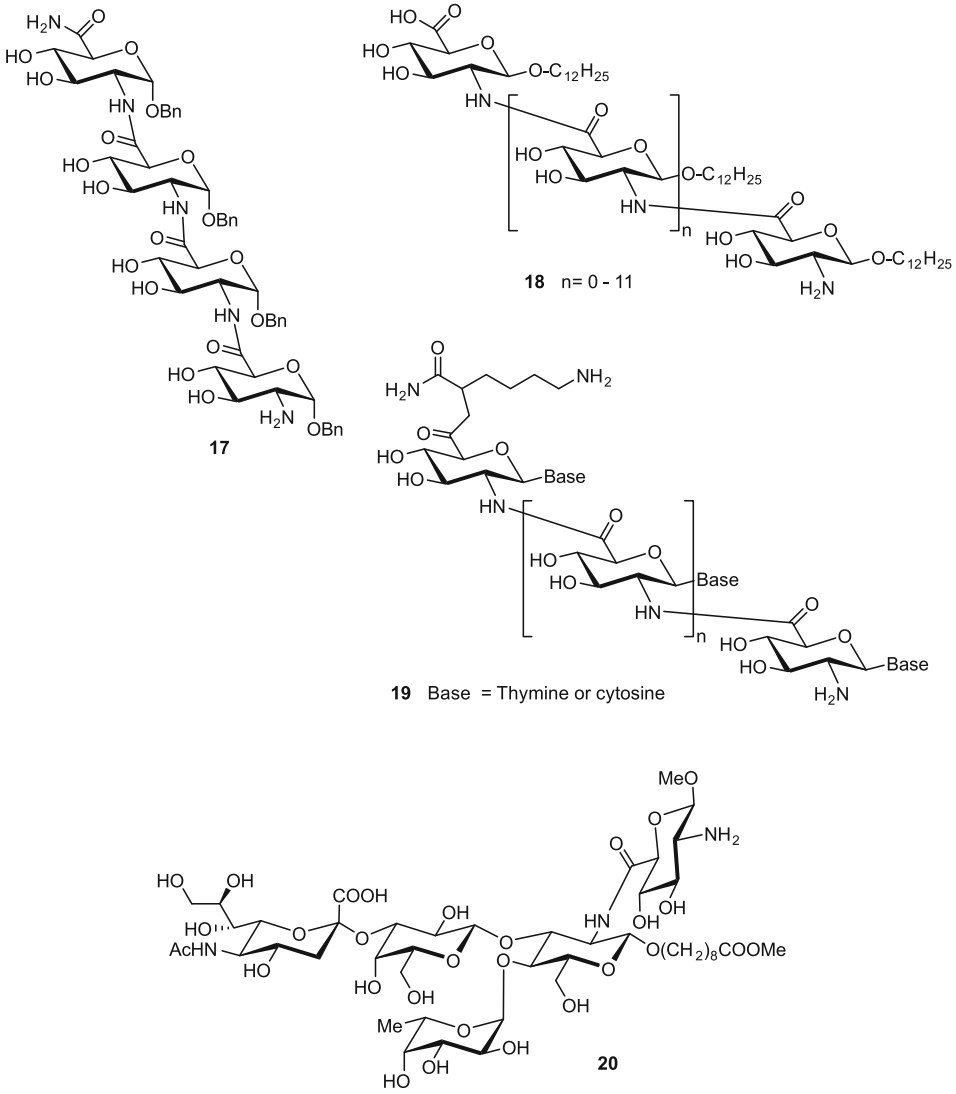
Also, piperidines have been employed for oligomerization, notably 4-hydroxypiperidine-3-carboxylic acid **16** [23] (*Scheme 4*); the ring nitrogen has been used as the amine component, and amide linkage led to mimetics of (1→3)-linked sugars with a one-atom linker (the carbonyl carbon atom). Reacting quinic acid derivatives with sugar amines, analogues



Scheme 4

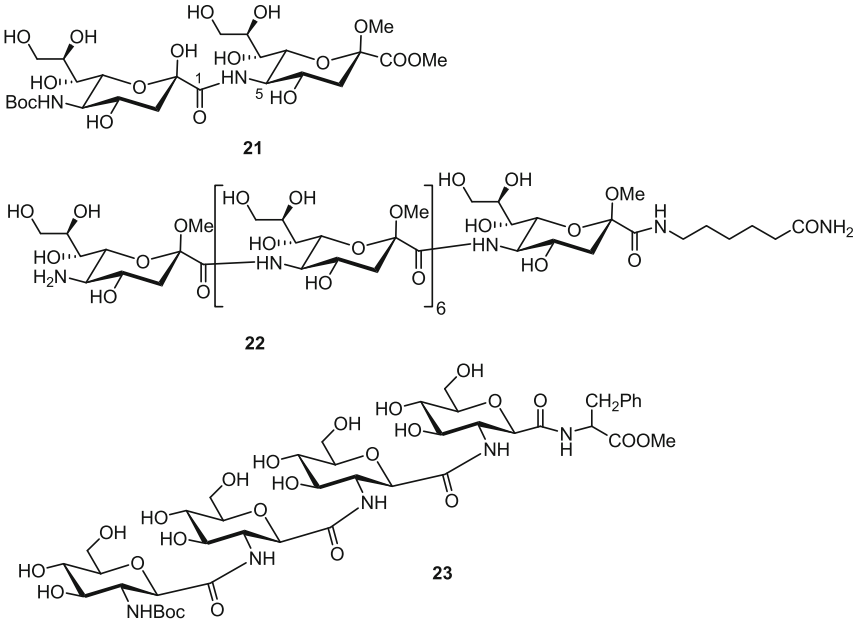
of cardiac glycoside disaccharide moieties with a three-atom amide linker were synthesized [24].

Glucosamine as a readily available and inexpensive sugar is an obvious starting point to prepare sugar amino acids. Amide coupling of 2-amino-2-deoxy-glucuronic acids leads to a (2→6)-linkage with a two-atom linker, i. e. one atom more than in the usual interglycosidic linkage, as in disaccharide mimetic **6** (*Scheme 3*) and the tetramer **17** [14] (*Scheme 5*), the latter was prepared by solid phase synthesis. As a consequence of this (2→6)-linkage, the anomeric center is free and can be employed for further derivatization, e. g., to increase the hydrophobic character as in **17** with the benzyl groups. Another application has been the glycosylation of the sugar amino acid with long chain alcohols to pre-orient the alkyl groups in an oligomer **18** which was synthesized as a mixture of compounds with differing degrees of poly-



Scheme 5

merization [25]. With a nucleo-base at the anomeric center [26], polymers were synthesized on a solid phase as pyranose-based antisense agents [27]. A 13-mer **19** is depicted in *Scheme 5*. Analogues of the natural oligosaccharides sialyl Lewis^a and sialyl Lewis^x have been prepared [28,29], replacing the acetamido group of the central glucosamine by a methyl glycopyranosiduronamido group. This replacement still allowed the construction of these molecules using the enzymes 2,3-sialyltransferase, $\beta(1\rightarrow4)$-galactosyltransferase [29] and fucosyltransferase III and IV. Compound **20** is the glucuronamide analogue of sialyl Lewis^a, the biological properties of these derivatives have not been reported.



Scheme 6

With neuraminic acid, a sugar amino acid widely occurring naturally as *N*-acetylneuraminic acid, an amide-linked (1→5)-dimer **21** (*Scheme 6*) and the corresponding 2,3-ene analogue were prepared [30]. On a solid phase, even an octamer **22** was available [31].

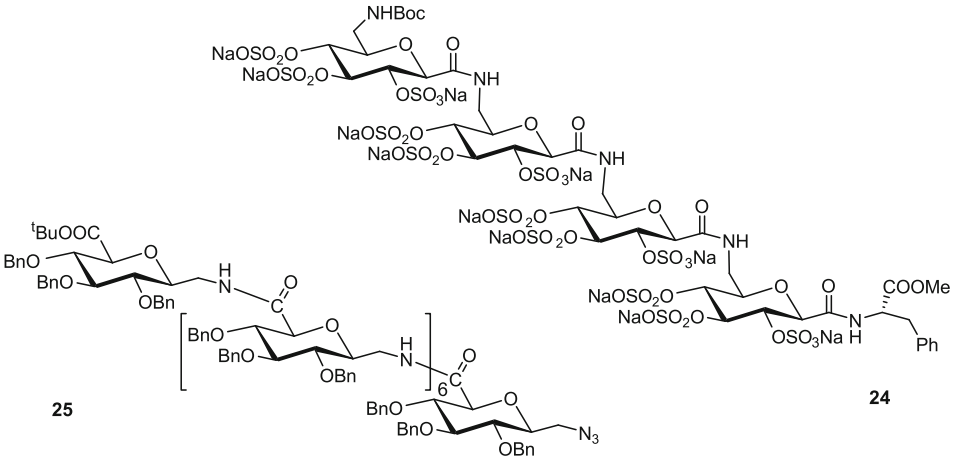
A mimetic of a (1→2)-linked glycoside was devised by Ichikawa's group [32] using a 3-amino-2,6-anhydro-3-deoxy-heptonic acid building block to arrive at tetramer **23**. A two-atom linker replacing the interglycosidic oxygen also characterizes this saccharide-peptide hybrid. A sulfated derivative of **23** blocked syncytium formation caused by HIV infection to CD4 cells at 50 µM concentration.

Oligomerization of a 7-amino-2,6-anhydro-7-deoxy-heptonic acid results in mimetics of (1→6)-linked sugars with a three-atom instead of a two-atom linker. The sulfated tetramer **24** (*Scheme 7*) turned out to exhibit μ-molar activity in the protection of MT2 cells from HIV infection [33], and was also shown to provide heparanase inhibitory activity [34]. With different location of the amide bond, a still benzyl protected β(1→6)-linked octamer of glucosyl–uronic acid–methyl amine (Gum [35]) **25** was synthesized [36].

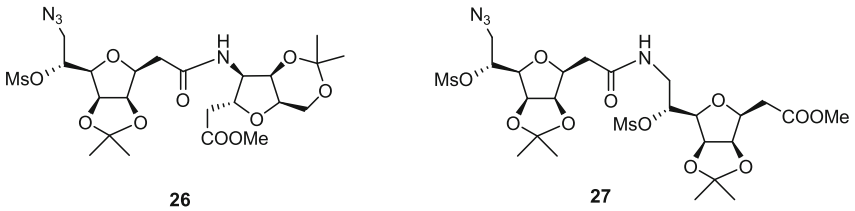
Using a *C*-glycosidic furanoid sugar amino acid, McDevitt and Lansbury [11] prepared a saccharide-peptide hybrid with a three-atom linker, the dimer **26**. The homodimer of this *C*-glycosidic sugar amino acid gave rise to a five-atom spacer as in compound **27** (*Scheme 7*); these derivatives have not been reported in deprotected form.

An interesting example of a branched trimer containing β(1→6)- and β(1→2)-amide-linked sugars was supplied by Sicherl and Wittmann who devised mimetics of amino glycosides [37].

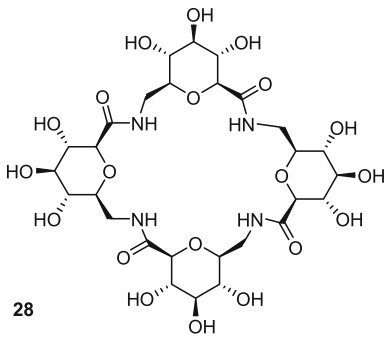
Pioneered by the group of Kessler [38] pyranoid sugar amino acid templates were described to represent turn mimetics and model peptides [39,40] including cyclic peptides [41,42]. As



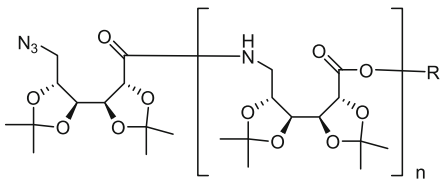
25 24



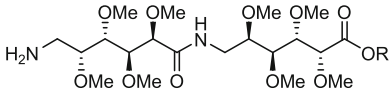
26 27



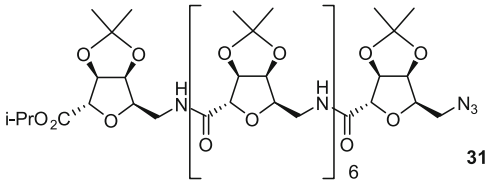
28



29 n = 1, 3, 5, 7



30



31

Scheme 7

a logical extension, also cyclic homooligomers of sugar amino acids were devised, first using Gum (see above) as a monomeric unit thus creating three-atom linkages as in cyclic tetramer **28** (*Scheme 7*) [43]. Further examples of cyclodextrin analogues based on furanoid sugar amino acids with three-atom [44] or four-atom linkers [45] as well as oxetane based cyclic homooligomers with three-atom linker [46] were described.

More recently, also oligomers of open chain sugar amino acids were described such as galactonates **29** (*Scheme 7*) of which also cyclic analogues were prepared [47,48,49,50]. This is an extension of work directed at polymers composed completely of amide-linked open chain sugars as hydroxylated analogues of polyamides (nylon) with the main goal to achieve biodegradability [51,52,53]. Instead of starting from monomers, also amide-linked dimers such as **30** have been employed to arrive at well-defined, enantiomerically pure and stereoregular polyhydroxylated polymers [54,55].

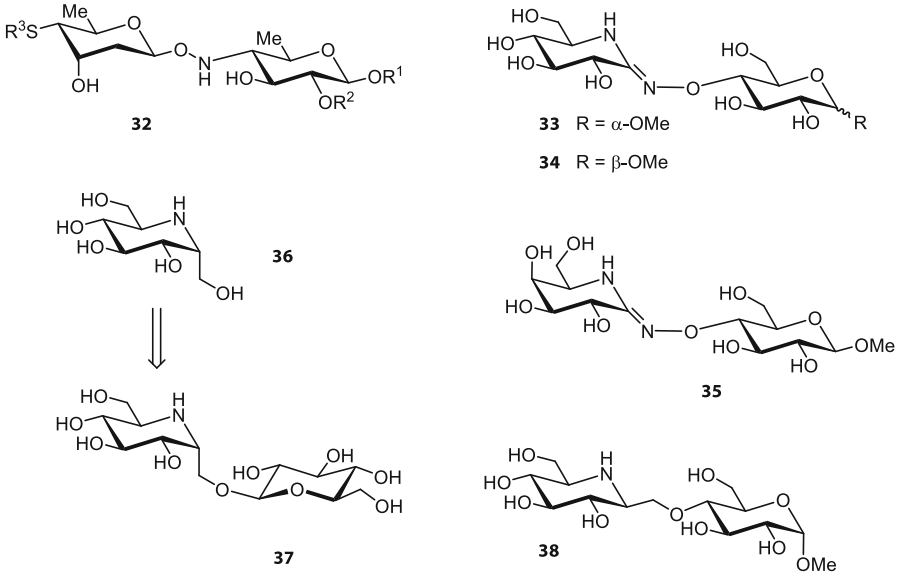
Saccharide-peptide hybrids have attracted particular attention because of their conformational properties. Oligomers of sugar amino acids have a tendency to adopt a compact conformation, a type of oligomer termed "foldamer" by Gellman [56]. Kessler's group [38] had already demonstrated that sugar amino acids may induce specific peptide conformations. Combined ^{1}H-NMR and CD spectroscopic evidence suggested that (1→5)-amide linked sialic acid oligomers in water adopt a defined secondary structure from the size of a tetramer on [31]. Similarly, a β(1→2)-amide-linked pyranoid sugar, a decamer of compound **23**, formed a right-handed helix [57]. The acyl-protected furanoid oligomer **14** prepared by Fleet's group [58] was shown by calculation and ^{1}H-NMR spectroscopy in organic solvents to exhibit a β-turn secondary structure. The investigation of further analogues showed that most higher oligomers with a 2,5-*cis* stereochemistry across the tetrahydrofuran ring adopt a right-handed helix conformation composed of repeating β-turns and with NH_i-O_{i-2} hydrogen bonds [59,60]. A secondary structure was however not found in the D-*galacto*-configured analogues [61].

For the 2,5-*trans* furanoid octamer **31** with D-*talo* building block a left-handed helix with NH_i-O_{i-3} interresidue hydrogen bonds was described [62] while the corresponding tetramer adopts a rigid but not hydrogen bonded conformation [63]. The van Gunsteren group, based on molecular dynamics simulations, has cautioned that NOEs can not be interpreted for single conformations but that clusters of various stable conformations are to be expected [64,65]. A conformation classification system based on various spectroscopic parameters including chiroptical spectroscopy was proposed [66].

In the D-*manno* series, an unprotected octamer showed an ordered structure in solution based on circular dichroism [67], but not the lower unprotected [67] or protected homooligomers [68]. Initial X-ray and NMR evidence for a secondary structure was found in a new series of tetrafuran based L-*ribo* configured γ-amino acids [69]. The conformational investigation of oxetane based oligomers revealed that a β-amino acid hexamer formed a left-handed helix [70], but δ-amino acid hexamers did not exhibit hydrogen bonded interactions but some regularity on steric grounds [71].

2.2 Non-Amide-Linked Sugars

There are numerous interglycosidic linkers other than amide linkers known in also natural products—Nature has already invented saccharide mimetics! One example are the calicheam-

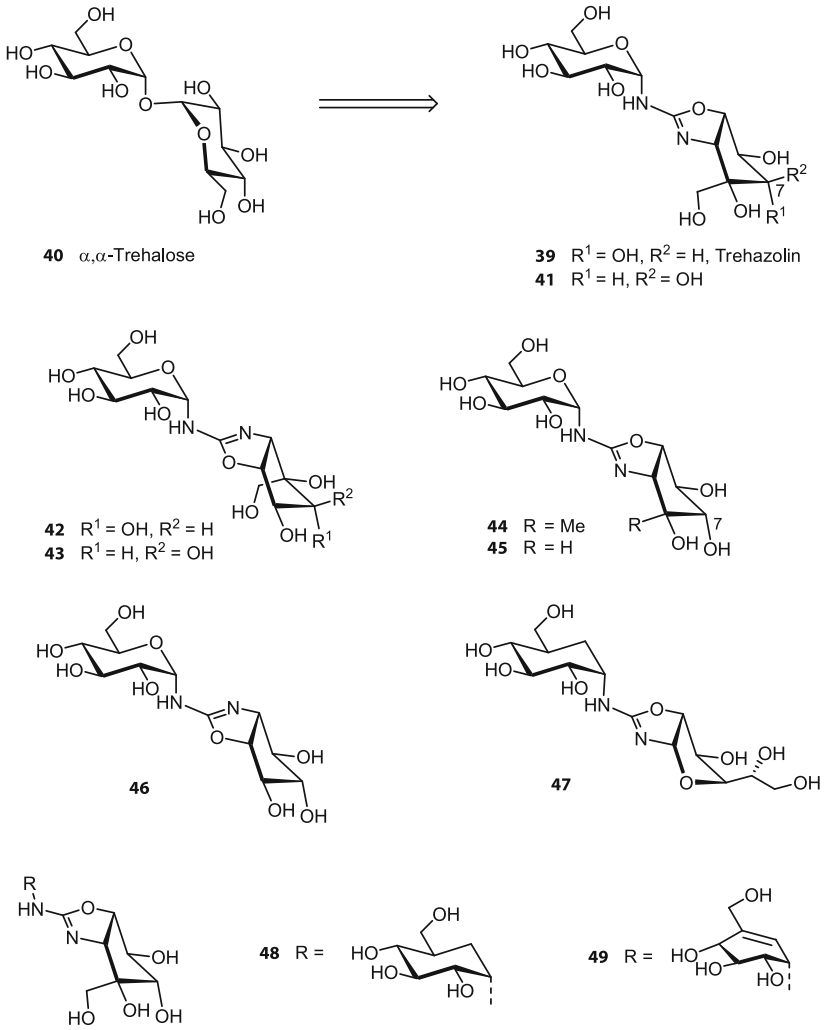


Scheme 8

icin [72,73] and esperamycin [74,75] endiine antibiotics sharing the characteristic disaccharide unit **32** (*Scheme 8*) with its two-atom *N*,*O*-interglycosidic bond [76]. Several groups have approached this interesting structure synthetically [77,78,79,80,81]. With a similar two-carbon linker, glycosides of sugar lactam oximes were prepared as selective glycosidase inhibitors [82]; both anomeric disaccharide mimetics **33** and **34** are good inhibitors of the β-glucosidase from *C. saccarolyticum*, the lactoside analogue **35** is a strong inhibitor of the *E. coli* β-galactosidase with an inhibition constant of $K_i = 0.1$ µMol.

Homo-*aza*-sugars such as α-homonojirimycin **36** have been devised to mimic the lengthening of the glycosidic bond in the transition-state in a glycosidase [83]. Interestingly, this mimetic **36** has also been isolated from a moth (*Urania fulgens*) and from its larval food plant. The β-D-glucoside of **36**, the disaccharide mimetic **37** (MDL 25,637 [84]), turned out to be a potent inhibitor of rat intestinal sucrase in vitro and had in vivo activity in a mouse hyperglycemia model. This mimetic is a homo-analogue of a trehalose mimetic, i. e. (1→1)-linked. A corresponding (1→4)-linked mimetic, the homo-*aza* analogue **38** of methyl α-cellobioside, was synthesized by Martin's group [85].

Trehazolin (**39**, *Scheme 9*) is a natural compound with a cyclic urea three-atom linker which has been first isolated only in 1991 [86], before the proof of identity it was also called trehalostatin [87]. Trehazolin is an inhibitor of the enzyme α,α-trehalose glucosidase which cleaves trehalose, the blood sugar and reserve carbohydrate of many insects. The aminocyclopentitol oxazoline thus mimics one glucosyl unit of α,α-trehalose (**40**), or, more likely, the postulated glucopyranosyl cation intermediate. The synthesis of trehazolin was first achieved by S. Ogawa et al. [88], other groups offered synthetic alternatives [89,90,91]. To ascertain the structure of trehazolin, the Ogawa group has also prepared the 7-epimer **41** and the diastereomers **42** and **43** of trehazolin with inverted stereochemistry



Scheme 9

in the aminocyclopentitol moiety [92]; these derivatives were, however, not active against trehalase. The deoxy- and dehydroxymethyl trehazolin derivatives **44** and **45** were clearly less active against silkworm trehalase than trehazolin, but the dehydroxymethyl derivative of **42** with the inverted configuration, compound **46**, lost activity only by a factor of about three [93]. The glucofuranosyl analogue **44** [93] with a carbasugar moiety replacing one glucosyl unit had sacrificed its trehalase activity. Simple oxazoline derivatives, i. e. fragments of trehazolin and analogues missing the cyclitol moiety, had no trehalase activity [94].

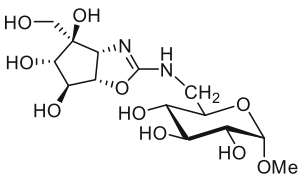
Replacements of the α-glucosyl moiety of trehazolin by α-mannopyranosyl, 3-deoxy-*ribo*-hexopyranosyl, α-galactopyranosyl, and 6-deoxy-glucopyranosyl [95] residues abolished

the trehalase activity demonstrating that the glucosyl moiety is sensitive towards variation. Changes in stereochemistry, particularly at the anomeric center, are not tolerated, and the hydroxyl groups in positions 3 and 6 seem to be involved in trehalase binding as judged by the deoxy derivatives. In contrast, the α-carbaglucosyl analogue **48** gave the same activity towards silkworm trehalase as trehazolin. The unsaturated valienamine derivative **46** had again no activity [95].

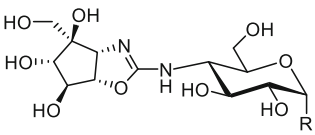
Modifications in the attachment of the glucosyl amine moiety have been devised by a Sankyo group [96]: they prepared the isomaltose analogue **50** (◉ *Scheme 10*), the maltose analogues **51** and **52**, and aminogalactosyl derivative **53**. Derivatives **51**–**53** were inactive towards maltase, isomaltase, and sucrase demonstrating that the aminocyclopentitol oxazoline does not generally replace glucose. Compound **52** displayed weak activity against porcine trehalase ($IC_{50} = 245$ ng/ml). Interestingly, compound **50** had some inhibitory activity against maltase and sucrase. Compound **51** was also independently synthesized by Knapp et al. [97] and shown to display weak yeast α-glucosidase activity. Another interesting member of this imidazol family of compounds is the 1-deoxynojirimycin (**54**) derivative formally fused with trehalamine (**55**) resulting in dimer **56** [98], a blend of two different types of glycosyl mimetics. During the last deprotection step (removal of benzyl groups from the deoxynojirimycin moiety with $Pd(OH)_2$ on carbon in methanol) also the rearranged product **57** was obtained. Both products **56** and **57** come with two-carbon linkers.

Formal opening of the oxazoline ring leads to urea or the analogous thiourea derivatives. Glycosyl thiourea **58** [89] or differently protected species [90,92] and analogues [95,96,97] have been prepared as precursors of trehazolin. These derivatives may be viewed as sugar mimetics with a three-atom linker or, in the case of the 1-deoxynojirimycin analogue **59** with one of the thiourea nitrogens in the ring, with a two-carbon linker. The amino glycoside antibiotic glycocinnasperimicin D (**60**) [99] family isolated from *Nocardia* strains contains urea linked aminodeoxy sugars, and urea and thiourea linked sugars have received increasing attention as oligosaccharide mimetics. Glycosyl ureas (as well as thioureas and guanidines) have been built up e. g. from carbodiimide-linked glycosides [100,101,102,103]. Those are themselves oligosaccharide mimetics with a three-carbon linker, and can be obtained via an *aza*-Wittig reaction such as compound **61** [104]. Following that methodology urea-linked nojirimycin glycosides such as **62** [105] and related Calystegine B analogues [106] were synthesized as glycosidase inhibitors (◉ *Scheme 10*). Urea-linked α- and β-glycosides have been synthesized from glycosyl isocyanates [107,108,109,110,111] or by oxazolidinone opening reaction with amines [112]. We have obtained unsymmetrical urea-linked sugars like **63** (◉ *Scheme 10*) in the lithium aluminum hydride reduction of *N*-benzyloxycarbonyl-protected glucosamine derivatives which are thought to be formed also via an intermediate 2-isocyanato derivative [113].

Urea-, thiourea- and guanidine-linked sugars have been investigated as phosphate binders, the guanidine-linked sugars showing the most favorable association constants [114]. In the nucleotide area, the research on antisense oligonucleotides has led to non-phosphorous phosphate replacements to modify parameters like RNA binding affinity, stability, especially against nucleases, or uptake into cells [115,116]. In this context, formacetal, thioformacetal, *N*-methylhydroxylamine, guanidino, sulfamate, urea, thiourea [117], carbamate, thiocarbamate [118], as well as different amide or thioamide [119] groups have been applied as linkers between 2-deoxyribofuranosyl derivatives.

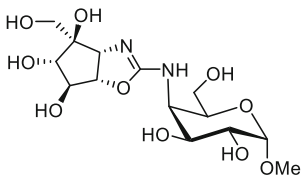


50

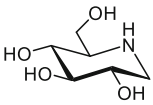


51 R = OMe

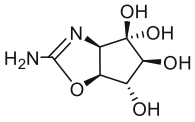
52 R = α/β-OH



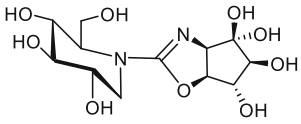
53



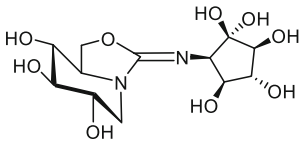
54



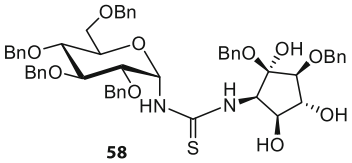
55



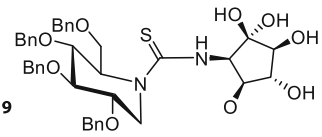
56



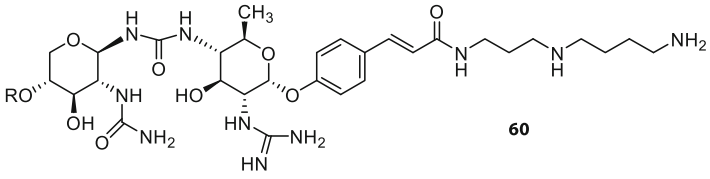
57



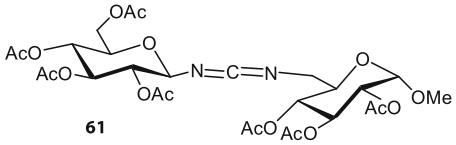
58



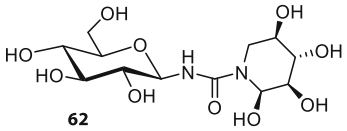
59



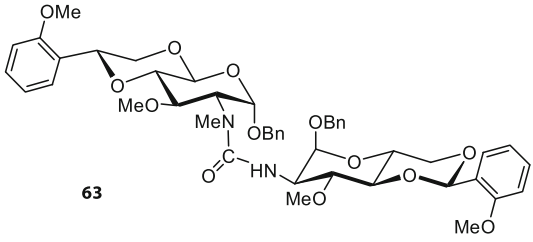
60



61

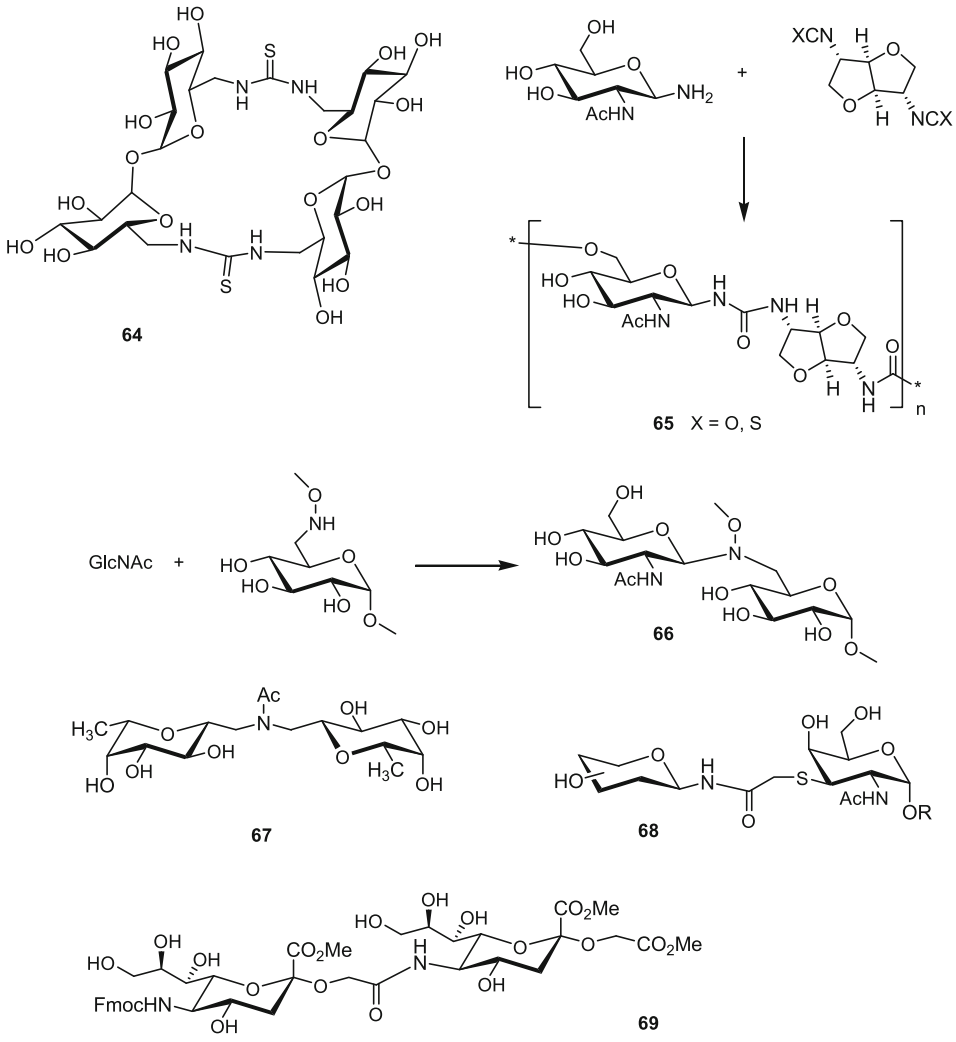


62



63

Scheme 10



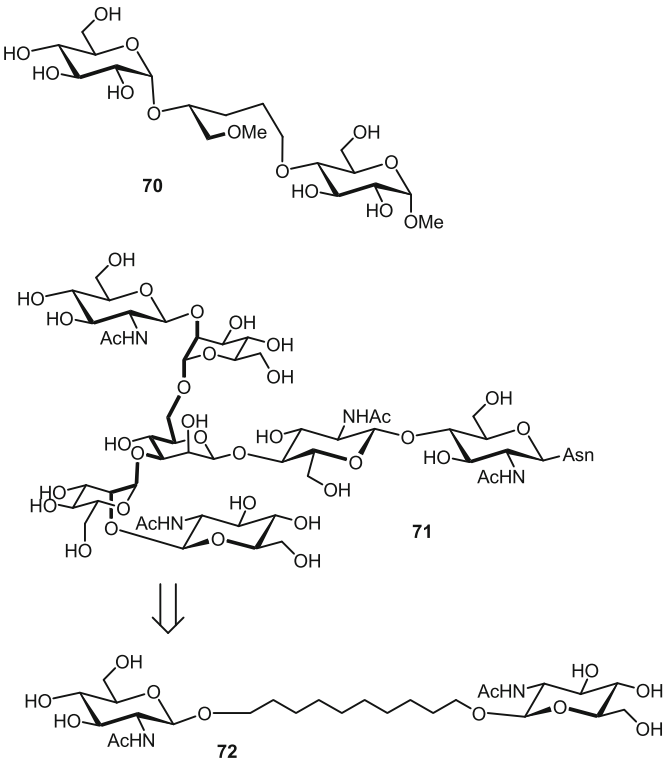
Scheme 11

Thiourea (6→6)-linked sugars were prepared [120] to arrive at a macrocycle **64** with alternating (1→1) and (6→6)-linkages which binds Cs^+ ions and forms strong complexes with Cu^{2+} ions. Urea linkers have also been employed to bridge cyclodextrins [121,122] (*Scheme 11*) and to build up template structures for the construction of glycoclusters [123]. Remarkable carbohydrate-based polyurethanes such as **65** (*Scheme 11*) containing urea- or thiourea-linked sugars as well as carbamate linkages were synthesized from diisocyanates or dithioisocyanates and amino sugars [124]. Defined glycosyl carbamates were synthesized starting from glycosyl isocyanides followed by oxidation and reaction with alcohols [125].

Condensation of methoxyamino sugars with free sugars resulted in the interesting oxyamino linked sugars **66** as mimetics of (1→6)-linked sugars [126]. While these di- and trisaccharide mimetics [127] are isosteric to the natural saccharides the conformational behavior was however found to be different [128].
A tail-to-tail, thus not C-glycosidically, acetylene linked pseudodisaccharide was described by the Beaupère group [129]. The Gross group [130] found a new *C*-glycosidic three-atom methylene amino linker in the course of the catalytic reduction of glycosyl cyanides (see compound **67**). 3-Thiogalactosamine was coupled with *N*-bromoacetamido sugars to result in **68** with a four-atom linker to mimic mucin-type oligosaccharides [131]. The Hindsgaul group employed an acetic acid linker amide coupled to neuraminic acids, a motif **69** that was found in marine sources and is thought to mediate sperm-egg recognition [132]. The same linker was utilized to tether C-6 and C-2′ of lactosamine thus constraining the conformation [133]. Acetic acid extended glycosides were used as building block in Ugi multicomponent reactions [134,135]. A new synthetic approach to this linker was devised by Queneau through opening of carboxymethylglycosyl lactones with amines [136]. An analogous *C*-glycosidic spacer was built up by oxidation of allyl *C*-glycosides, amide coupling furnished linear and cyclic homooligomers [137]; the amides were also reduced to the corresponding amino-linked compounds.

3 Spaced Sugars with Acyclic Spacers

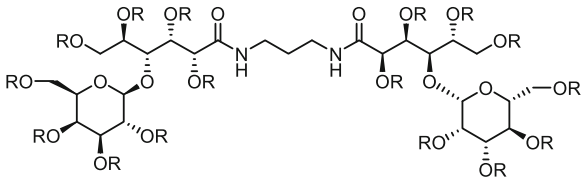
Acyclic spacers have been used to replace sugar units in a rather flexible way. In a first report by Jegge and Lehmann [138], an alkyl spacer-linked mimetic **70** (*Scheme 12*) of methyl α-maltotrioside was prepared, the six-carbon spacer would mimic one (1→4)-linked glycosyl residue. One motivation at that time was to enable with the spacer the introduction of a reactive group for photoaffinity labeling of enzymes [139]. Mimetic **70** turned out to be an inhibitor of porcine α-amylase. As a mimetic of the biantennary core structure **71** of *N*-linked glycoconjugates, the compound **72** with a long ten-carbon alkyl spacer replacing the three central mannose residues was devised [140]. This substrate could still be galactosylated with a membrane-bound galactosyl transferase from a microsomal fraction of liver cells. A binding model with two separate binding sites for the outer GlcNAc residues was supported by the lower activity of analogues with shorter alkyl spacers than in **72** [141] (*Scheme 12*).
Penta- and hexasaccharide mimetics related to fragments of the capsular polysaccharide of *Streptococcus pneumoniae* type 14, {6)-[β-D-Gal*p*-(1→4)]-β-D-GlcNAc*p*-(1→3)-β-D-Gal*p*-(1→4)-β-D-Glc*p*-(1}$_n$, were synthesized in which one or two galactosyl units were replaced by a three-carbon alkyl spacer. These linear oligosaccharide mimetics served as good acceptor substrates for bovine milk β-1,4-galactosyltransferase to enzymatically introduce the galactosyl branches resulting in hexa- and octasaccharide mimetics, respectively [142].
Head-to-head β,β-linked spaced galactosides were obtained by glycosylation of 1,3-propanediol, 1,6-hexanediol, and of 1,4-butynediol [143]. Roy's group [144,145] presented an approach to alkenyl-spaced (and with that implicitly alkyl-spaced) saccharides employing homodimerization of suitable glycosides such as allyl galactosides by olefin metathesis, similarly *C*-glycosidically (1→1)-linked β-galactosides with an alkenyl spacer were obtained. Also employing olefin metathesis, head-to-head linked [146], tail-to-tail linked [147] or



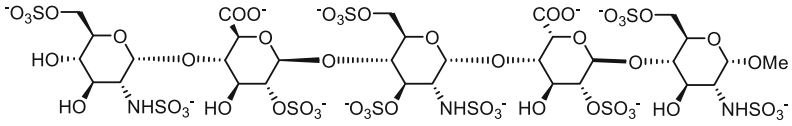
Scheme 12

macrocyclic [148] alk(en)yl spaced amino and other [149,150] glycosides were prepared. Alkenyl spaced sugars were further synthesized by Stille coupling of halo-*exo*-glycals [151] and by palladium-catalyzed glycosylation with allylic biscarbonates [152]. Spacers were also attached via amides [153,154,155,156] or ureas [103,157].

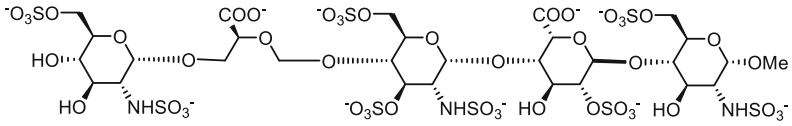
A prominent example of a spaced saccharide is Aprosulate **73** (*Scheme 13*), a symmetrical sulfated lactobionic acid amide. This compound was developed as an antithrombotic agent to mimic heparin, a glucosaminoglycan which is employed clinically as an anticoagulant since decades. Different spacer lengths were studied varying from five to eleven atoms, as well as a branched four-atom *N*,*N*′-dimethylethylenediamine spacer [158]. The five-atom propylenediamine spacer compound yielded the most promising anticoagulant profile. This spacer may replace one glycosyl unit so that **73** may mimic a yet unknown natural pentasaccharide. Aprosulate sodium probably acts by catalyzing thrombin inhibition by heparin cofactor II [159]. Aprosulate had reached phase I clinical trials and was efficacious, but being associated with increased transaminase levels [160,161], the further clinical development has been suspended. On a more rational basis, open-chain analogues of the heparin pentasaccharide **74** have been prepared. This heparin pentasaccharide [162] was identified as the unit which binds to antithrombin (AT) with high affinity resulting in the acceleration of the inhibition of coagulation factor Xa (AT-mediated anti- Xa activity: 700 U/mg). As a reference compound,



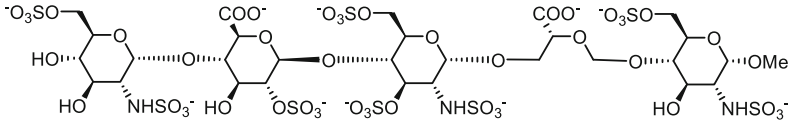
73 R = SO_3Na



74



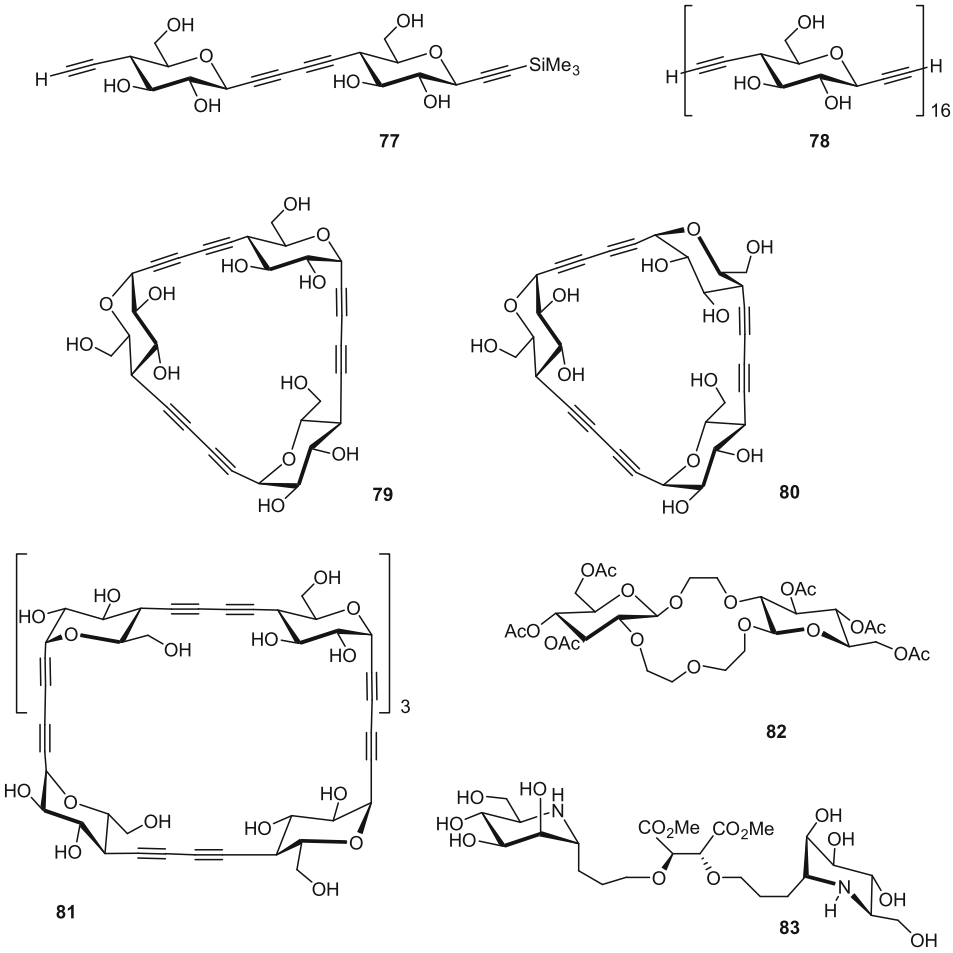
75



76

Scheme 13

a derivative additionally sulfated at the 3-position of the glucosamine unit at the reducing end was chosen with even higher activity (AT-mediated anti-Xa activity: 1250 U/mg) [163]. Both uronic acids of this pentasaccharide were replaced by open chain fragments, the D-glucuronic acid by an (*S*)-glyceric acid oxymethylene residue to give **75** [163], and the L-iduronic acid by an (*R*)-glyceric acid oxymethylene residue to give **76** [164]. These are six-atom spacers with the same atom count as the (1→4)-linked pyranuronic acid. While **75** was inactive, pentasaccharide **76** (AT-mediated anti- Xa activity: 150 U/mg) retained some 12% of the activity of the reference compound. The methylene acetal in **76** adopted a linear conformation [165]. Since the conformational flexibility of L-iduronic acid is well documented, the pyranoside exists in a chair/skew boat equilibrium, it is not surprising that a more flexible unit is tolerated. In contrast, a flexible replacement of the rigid D-glucuronic acid seems to be not allowed. Accordingly, the pentasaccharide in which both the D-glucuronic acid unit and the L-iduronic acid unit have been replaced by glyceric acid oxymethylene residues (structure not shown here) was inactive.



Scheme 14

Heparan sulfate and hyaluronan pentenyl disaccharides were coupled by metathesis or reductive amination reactions resulting in 8- or 12-atom spacers [166,167].

Vasella et al. developed the concept of acetylenosaccharides as oligo- and polysaccharide analogues to study intra- and intermolecular interactions [168]. A first representative was the dimer **77** (*Scheme 14*) [169] with a rigid four-carbon butadiynyl spacer. Higher oligomers up to the β-*C*-glucosyl hexadecamer **78** were assembled [170]. The same spacer concept was applied to devise cyclic acetylenosaccharides as analogues of cyclodextrins. Starting with an α-*C*-linked building block, two different cyclotrimers were synthesized, the C_3-symmetric **79** with (1→4)-linkages exclusively [171], and **80** with one (1→1)-linkage [172]. In addition, syntheses of larger rings like several cyclotetramers and, of bigger ring sizes, up to the cyclooctamer **81** were achieved [173]. Other acetyleno cyclodextrins contained mannose [174] or glu-

cosamine [175] units, or only one butadiynyl [176], pentadiynyl or hexadiynyl spacer [177]. By intramolecular attack of the butadyine triple bonds the spacer could be shortened to three- or two-carbon length [178]. In the context of a study on oligonucleoside foldamers ethynyl-linked nucleosides were investigated [179]. Cyclic dinucleosides were prepared by bridging with butylene spacers thus creating a 24-membered ring [180].

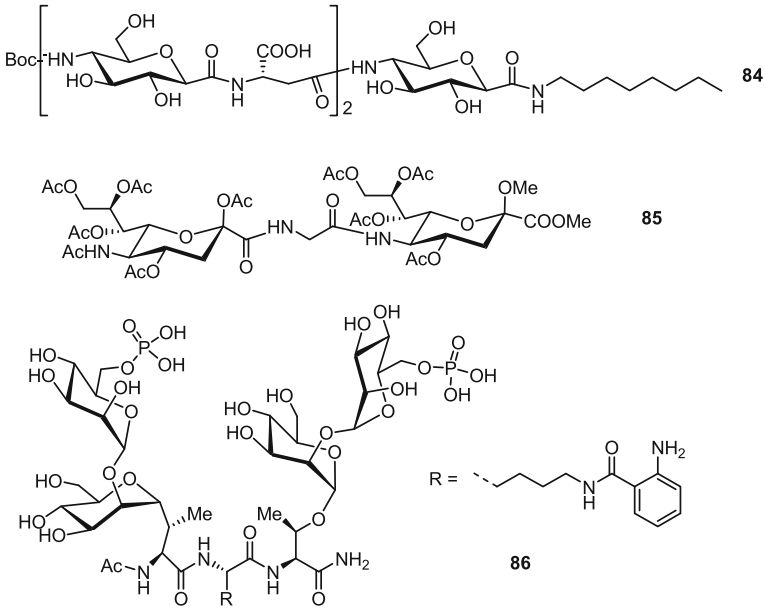
Carbohydrate crown ethers were obtained with ethylene spacers; from a crown ether point of view, the carbohydrate vicinal diols are replacing one ethylene glycol unit [181,182]. Cyclic compounds synthesized include bis-gluco-15-crown-5 **82**, bis-gluco-21-crown-7, and tetra-gluco-24-crown-8 (❯ *Scheme 14*). These chiral macrocycles could serve as catalysts in the asymmetric Michael addition of methyl α-phenylacetate to methyl acrylate. With the goal to study molecular interactions, β,β-, β,α-, and α,α-bis-maltosides with aliphatic two-, three-, or four-carbon spacers were synthesized [183]. Spaced cyclodextrins were prepared to study their supramolecular properties [184,185].

Also spaced glycoside mimetics have been described, e. g., analogues of trehazolin **39** (cf. ❯ *Sect. 2.2*, ❯ *Scheme 9*) with two- and three-carbon alkyl spacers [186]; they turned out to be inactive as yeast α-glucosidase inhibitors [187]. Further examples are head-to-head β,β- or α,α-linked 1-deoxymannojirimycins with eight carbon alkyl spacers or those with more polar centers like dipropylamine (seven atoms) or a ten-atom spacer as in **83** [188] (❯ *Scheme 14*). Glycosyltransferase inhibitors (reviewed in [189]) with micromolar activity were obtained with iminosugars spaced to other sugars or nucleosides with the spacer acting as phosphate mimetic.

Glycoproteins (cf. ❯ Chap. 8) and natural or synthetic glycopeptides as well as glycoconjugates can be viewed as "spaced sugars" but in their entity are not in the scope of this chapter; only a limited number of oligosaccharide mimetics are discussed here. Suhara et al. [190] have synthesized alternating sugar amino acids and α-amino acids, and the oligomers were examined in two in vitro systems measuring inhibition of chemotaxis and invasion in highly metastatic tumor cell lines. A trimer containing D-aspartate and a pentamer **84** (❯ *Scheme 15*) containing L-aspartate showed considerable activity. Molecular dynamics calculations of these oligomers indicated stable conformations due to intramolecular hydrogen bonding. Ramamoorthy and Gervay [13] described an acetyl protected neuraminic acid-glycine-neuraminic acid trimer **85**.

Mannosyl oligosaccharides with two terminal mannose 6-phosphate (M6P) residues have been suggested to be the natural ligand for the mannose-6-phosphate receptor, which is involved in inflammatory processes. Bidentate structures with 6-*O*-phosphorylated mannose disaccharides on a peptide template with three to five amino acids have been created. A compound with a tripeptide spacer **86** was distinctly more active in a binding inhibition assay than a suggested natural ligand oligosaccharide [191]. A much simplified mimetic with a four-carbon alkyl spacer between two mannose units with phosphate replacements showed appreciable affinity for the M6P/insulin-like growth factor II receptor [192]. Glycopeptide mimetics of mammalian Man_9NAc_2 were designed and tested for their ability to inhibit binding of mannan-binding protein to mannan from *Saccharomyces cerevisiae* [193]. Further linear [194] and cyclic [195,196,197,198,199] alternating sugar amino acids/amino acids were studied [200].

Another group of spaced sugars is found in the growing class of glycodendrimers [201,202,203,204], prepared as neoglycoconjugates [205,206] or cluster glycosides [207,208,209,210] and to address the issue of multivalency [211,212]—this topic is covered in detail in



Scheme 15

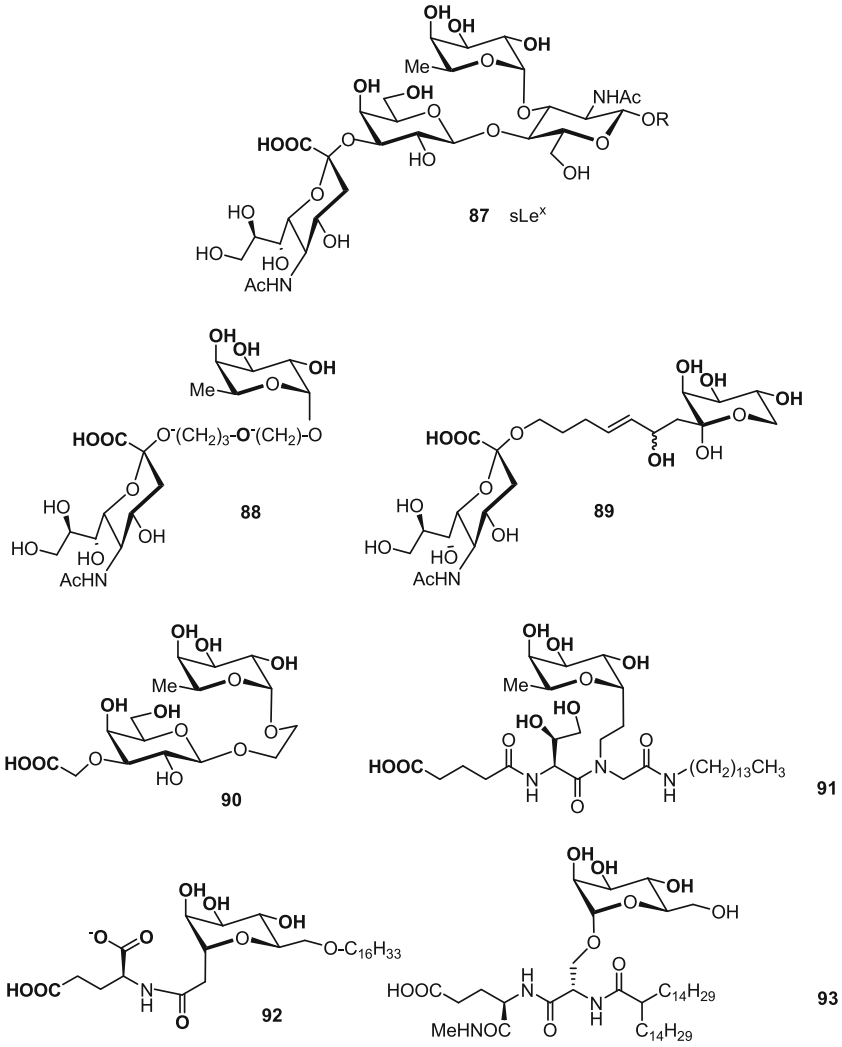
Chap. 12.4. Also in these cases sugars have been linked "unnaturally" as amides [213, 214,215,216] or thioureas [217,218].

Early examples of neoglycoconjugates are sulfated heparin oligosaccharides linked by a more than 50-atom spacer to replace ten or more glycosyl units and to inhibit the coagulation factor thrombin [219] or the head-to-head linked *N*-acetylglucosamines with linear spacers with up to 66 atoms which have been prepared as model compounds for more complex sugars [220].

With respect to physicochemical properties, sugars with sufficiently long spacers, such as bolaamphiphiles [221] or gemini molecules [222], may form supramolecular structures and have surfactant character.

Active research programs in a number of pharmaceutical companies and academic laboratories have been devoted to the identification of mimetics of sialyl Lewis x (sLe^x, **87**) [223,224], cf. *Scheme 16*. sLe^x glycosides are found on the outer surface of glycolipids and glycoproteins and are the key recognition elements of the selectins, a group of cell surface proteins with carbohydrate recognition domains classified as E-, P-, and L-selectins according to their occurrence on endothelial cell, platelets, and lymphocytes [225]. Selectin-carbohydrate interactions occur at an early stage of inflammatory reactions or metastasis. The hope is that with sLe^x mimetics it would be possible to intervene in acute and chronic inflammatory diseases (asthma, arthritis, myocardial infarction, lung injury) and to find new anticancer agents [226,227]. Towards this goal the full knowledge on carbohydrate mimetics described above has been employed in this practical exercise of carbohydrate medicinal chemistry to simplify the complex sLe^x saccharidic structure.

As indicated in the structural representation of **87** (*Scheme 16*) some groups (marked in bold print) are important groups in the binding to E-selectin. In early examples, the galactose



Scheme 16

residue was substituted by alkyl or hydroxylated alkyl chains [228,229]. Both the galactose and glucosamine residues have been replaced by non-cyclic spacers as shown in fucoside **88** [230] or branched fucoside **89** [231]. Other spacers were simple alkyl chains [230], pentaerythritol [232], or higher hydroxylated or keto analogues of **89** [231]. These extensive flexible sugar replacements abolished, however, the biological activity.

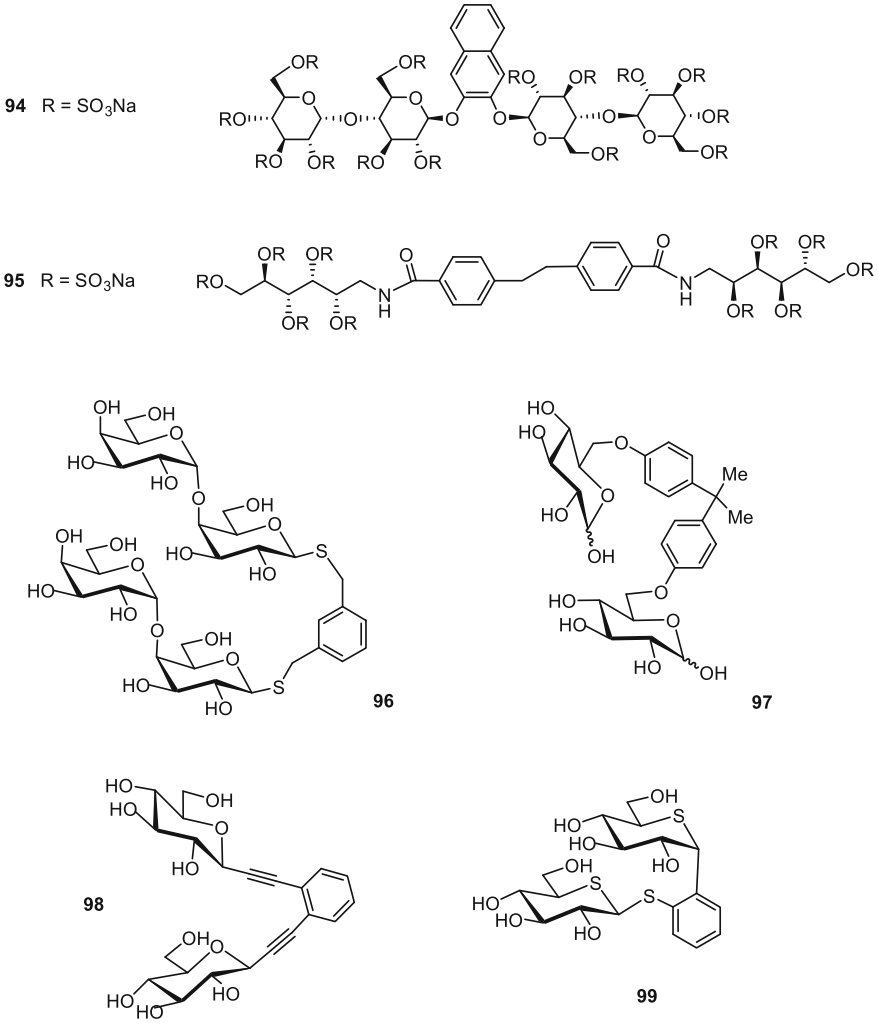
Once it was realized that the most important group of the neuraminic acid is the carboxylate, this saccharide was frequently replaced by a, in some cases substituted, hydroxyacetic acid group. The glucosamine unit was substituted by a simple glycol as in **90** [233] or a substituted glycol, e. g., the C_2-symmetrical 2,3- butanediol group [234,235], also *C*-glycosidic mimetics

were prepared [236,237]. Among these, the glycol derivative **90** retained the highest activity (0.8 mM against E-selectin) being ca. three times less active than sLe^x itself. An alternative substitution of the Neu5Ac-(1→3)-Gal disaccharide was brought about by alkylated malonic acid derivatives giving rise to sLe^x-like activities [238], 2- and 3-malonate substituted galactosides were designed resulting in lead compounds more potent than sLe^x for P-selectin [239]. Further trimming of the inhibitors by replacement of also the galactosyl unit derivatives with only one saccharide unit were devised. *O*- and *C*-glycosidically linked fucosyl peptides with hydroxylic or peptidic side chains were synthesized [240,241,242,243,244], and a marked increase in activity was observed with the attachment of lipid moieties as in **91** ($IC_{50} = 37\,\mu M$, E-selectin) [241]. Similarly, 6-amino-6-deoxy-L-galactose [245] and mannose [246] with the same arrangement of hydroxy groups as in fucose have been employed; it was found that the essential hydroxy groups of the galactose moiety can be replaced by a carboxylate: compound **92** had an $IC_{50} = 40\,\mu M$ against E-selectin [247]. sLe^x mimetics with glycosylated β-turn dipeptides showed to be potent but unselective selectin antagonists, for instance mannosylated compound **93** had $IC_{50} = 3.5$, 0.45 and 4.0 µM for E-, P- and L-selectin, respectively [249,249]. Combinatorial approaches have been chosen to build up peptidic sLe^x mimetics employing either the four-component Ugi condensation with a *C*-fucosyl, -mannosyl, or -galactosyl building block [250,251] or parallel chemistry on a solid phase [252].

The synthesis of polymers containing carbohydrate monomers has received attention to enhance their hydrophilicity and biodegradability [253], the synthesis of chiral polyamides has been reviewed [254]. Thiem and coworkers synthesized various polymers with aliphatic spacers from dianhydro sugars [255,256] or sugar diamino derivatives [257], the Kiely group started from various aldaric acids to synthesize polyhydroxy nylons [258,259,260,261]. On the basis of aldaric acids, novel carbohydrate-segmented silicone polyamides were synthesized [262]. Stereoregular polyamides with monomers of D-*manno* or L-*ido* configurations [263,264] and other AABB-type polymers [265,266] were prepared. The Fleet group built up macrocyclic lactams from open chain sugars spaced with alkyl groups [267].

4 Spaced Sugars with Cyclic Spacers

Cyclic spacers are more rigid than acyclic ones and present the attached sugar molecules at more defined exit vectors. Early example are the sulfated spaced oligosaccharides described by Wessel et al. [268] that were developed as heparinoid mimetics from sulfated trehalose oligosaccharides [269,270] with heparin-like antiproliferative activity. Contrary to the oligosaccharides, the spaced disaccharides could be completely sulfated to arrive at chemically well defined compounds. The spacer compound **94** (*Scheme 17*) had a 20% higher antiproliferative activity than the polysaccharide heparin itself along with good selectivity against antithrombin-mediated anticoagulant effects [271]. A number of other aromatic spacers could be used advantageously, but non-aromatic and particularly acyclic spacers gave distinctly lower activities; the aromatic spacer contributed to the binding and thus increased the activity. In an extension of that work, the disaccharides were replaced by open-chain sugars that were attached to aromatic spacers via an amide, sulfonamide, ether, amine, urea, or thiourea linker. This type of heparinoid mimetic was readily accessible; in the simplest case glucamine (1-amino-1-deoxy-glucitol) was coupled to aromatic bis-carboxylic acids



Scheme 17

followed by sulfation to furnish, e. g. **95**. This specific sulfated mimetic had a 50% higher antiproliferative activity than heparin in a smooth muscle cell growth assay with concomitant high selectivity with regard to its anticoagulant effects in a clotting assay.

Sulfated heparin oligosaccharides were coupled to aromatic spacers by reductive amination to investigate their binding to platelets [272,273].

Structurally similar bis-*O*- and *S*-glycosides were prepared to study the mode of presentation of the sugar binding epitopes in carbohydrate-protein recognition, galabiose (α-D-Gal*p*-(1→4)-D-Gal*p*) as in **96** (*Scheme 17*) and globotriose (α-D-Gal*p*-(1→4)-β-D-Gal*p*-(1→4)-D-Glc*p*) were employed as saccharide units [274]. Replacement of the core disaccharide

βGal(1→3)GalNAc of the GQ1bα ganglioside by a biphenyl spacer as in **97** led to antagonists of the myelin-associated glycoprotein [275]. C_2-Symmetric phenylene bridged glycosyl ureides were reported to display good Leishmanicidal activity [276]. Divalent aromatically spaced tetragalloylglucoses inhibited lipopolysaccharide-induced tumor necrosis facor-α secretion by more than 50% [277]. Other divalent structures, glycosyl amides with restricted conformations, were prepared to influence receptor clustering [278,279].

Bis-*C*-aryl- and bis-*C*-alkynylaryl-glycosides such as **98** were introduced by Vasella and collaborators [168,280]. Aromatic ethynyl spacers have been built up by Sonogashira coupling [281,282].

Bicyclic aromatic [283,284] or triazole [285] spacers were used to arrange oligosaccharides in a parallel fashion to investigate mimicry of cellulose I and II. Interesting aryl *C*,*S*-bisglycosides of type **89** (● *Scheme 17*) have been obtained during the work on antithrombotic 1,5-dithiopyranosides [286].

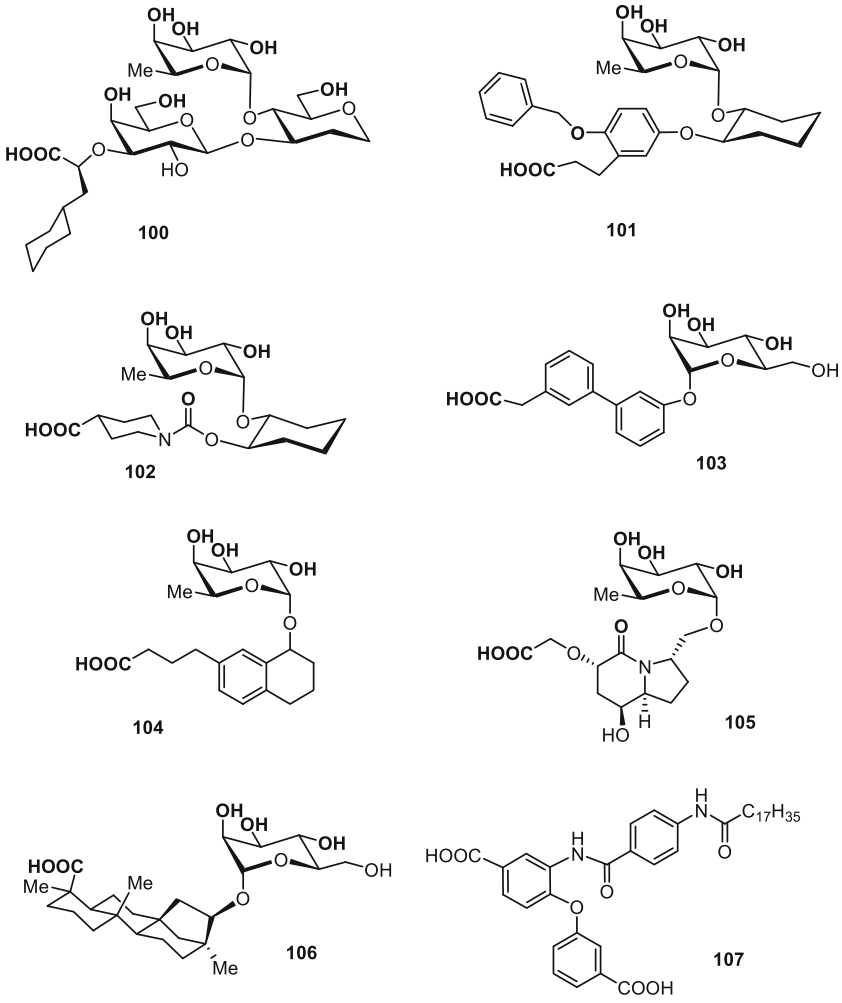
In the area of glycosidase inhibitors, a cyclic analogue of the spaced imino sugar **83** (● *Scheme 14*) was reported with replacement of the tartaric acid element by a cyclohexanediol [188]. Triazole spaced saccharides were devised as glycosidase inhibitors in which the trizole mimicked a valienamine unit [287]. Thanks to the ready access via 1,3-dipolar addition—the copper-catalyzed Huisgen reaction is the basis of the so-called "click chemistry"—triazoles have become popular to construct linear [288,289,290] or cyclic [177,291] spaced saccharides, likewise oxazole spacers were constructed [292]. Other ways to construct a cyclic spacer from glycosyl building blocks was by yne-ene cross metathesis followed by Diels–Alder reaction [293] or by Biginelli polycondensation [294].

Carboranes are particular cyclic spacers that have been introduced for boron neutron capture therapy to be taken up by tumor cells [295,296,297]. Aromatic spacers have been integrated in cyclodextrin analogues to modify the guest binding properties [298,299,300]. Also amide-linked polymeric open chain sugars with aromatic spacers were prepared [257,264,301].

Sugars with cyclic spacers are also found in Nature [302], well known examples are antibiotics such as calicheamicin [303,304] or vancomicin [305].

The last paragraphs of this chapter shall be devoted again to mimetics of the sLe^x tetrasaccharide **87** (● *Scheme 16*). Since it was realized that the glucosamine moiety of **87** does not considerably contribute to the selectin binding, this sugar was replaced by a cyclohexyl group without loss of affinity towards E- and P-selectin [228]. Cyclic analogues of the glycol derivative **90**, in which the neuraminic acid was reduced to a hydroxyacetic acid, were synthesized using, for example, a cyclohexyl, phenyl [235], or pyrrolidine [236,306] spacer. While these replacements alone were disappointing, a derivative with a cyclohexylmethyl-substituted hydroxyacetic acid replacing the neuraminic acid had good activity against E-selectin [307,308], the hydroxymethylated analogue **100** (● *Scheme 18*) was 30 times more active than **87** and showed to be efficient in a murine model of acute inflammation ($ED_{50} = 15$ mg/kg) [309,310].

In the series of peptidic fucosyl derivatives, Wong's group also presented analogues of sLe^x with a cyclohexyl group replacing the glucosamine unit [311,312]. The galactose and glucosamine units have also been substituted individually by carbocycles derived from D-quinic acid [313]. Mimicking the Gal-GlcNAc disaccharide with xylene derivatives [314] or a rigid spiroketal [315] did not improve the activity, the use of bis-*C*-glycosylated diphenylmethane to mimic the same units led to low micromolar L- and P-selectin antagonistic activity [316].



Scheme 18

Finally, three glycosyl units were replaced by a carboxylic acid and a suitable spacer, leaving only a fucosyl unit. In **101**, the spacer is an aryl cyclohexyl ether [317]; similar structures were prepared by the groups at Hoechst [318], and, incorporating their cyclohexylmethyl-substituted hydroxyacetic acid, Novartis [319]. Toepfer and Kretschmar with, e. g., **102** have presented [318,320] piperidine carbamates as a galactose replacement. Kogan et al. [321,322] employed various biphenyl derivatives as disaccharide mimetics as exemplified in **103** with a seven-fold higher activity towards E-selectin than sLe^x. Murphy et al. [323,324] incorporated tetralin and naphthalene ring systems as in **104** to obtain sLe^x-like activity. Hanessian et al. [325] introduced γ-lactams and indolizidinones as β-turn mimetics to achieve with for example **105** activity against P-selectin in the low µM range. With activities in µM range

β-linked aryl *C*-glycosides [326] and also the interesting α-mannosyl terpenoid **106** [327] proved to be P-selectin antagonists.

More recently, multivalent approaches [328,329,330] were reported. Based on molecular modeling approaches, non-sugar sLe^x mimetics [331,332,333] such as **107** ($IC_{50} = 86$, 6.1, and 30 µM for E-, P- and L-selectins respectively) [331] were created. A peptide mimicking the sialyl $Lewis^a$ oligosaccharide has been identified [334], this concept is discussed in more detail the next chapter.

5 Peptides as Carbohydrate Mimetics

As discussed in *Sect. 3*, peptides have been employed successfully as spacers in carbohydrate mimetics. Here we cover cases where peptides alone are able to mimic oligosaccharides and particularly carbohydrate antigens on the surface of cells. This parallels the finding that anti-idiotypic antibodies could act as immunogens and thus mimic carbohydrates [335,336]. Carbohydrates are T cell-independent and thus only weakly immunogenic, a problem which had been addressed by the formation of carbohydrate-protein conjugates. Now it has been demonstrated that peptides can mimic the meningococcal group C capsular polysaccharide, carbohydrate epitopes on breast adenocarcinoma cell lines [337,338], and those expressed on the envelope protein of the human immunodeficiency virus [339,340] and the Galα1→3Gal epitope which is of importance in hyperacute rejection [341]. The screening of suitable peptides has been greatly facilitated by the phage display technology [342,343]. Using this approach, peptides mimicking the bacterial saccharide *O*-antigen [344] of the *Shigella flexneri* serotype 5a lipopolysaccharide [345], the polysaccharide capsule of *Cryptococcus neoformans* [346] or glycosphingolipids [347] were prepared. Also peptide ligands for endogenous galactose-specific lectins (galectins) that are involved in various physiological and pathophysiological pathways including tumor growth regulation were identified [348,349,350].

Peptidic oligosaccharide mimetics that not only compete with the native protein for antibody binding but are also able to induce antibodies with the same biological activity were coined mimotopes [351]. Early hopes that this concept will lead to new vaccines directed against carbohydrate epitopes were supported by successful immunization of tumor bearing mice leading to complete elimination of tumor cells with co-administration of interleukin 12 [352].

Structural knowledge has been derived from X-ray structures of protein–carbohydrate complexes [353]; conformational preferences of carbohydrate mimicking peptides were studied by NMR [354]. The computer program LUDI [355] has been employed to epitope map the anti-Lewis Y antibody combining site, correlating peptide reactivity patterns [356]. The mimicry of specific carbohydrates or carbohydrate sequences was found to be associated with specific peptide sequences, for example, glucose with trypsin-arginine-tyrosine, α(2→9)sialic acid with tyrosine-tyrosine-arginine-tyrosine-aspartic acid, and the Lewis Y antigen [Fucα1→2Galβ1→4(Fucα1→3)GlcNAc] with proline-trypsin-leucine-tyrosine [337]. Mainly aromatic-aromatic and hydrophobic interactions were found to drive the peptide–protein binding.

References

1. Yoshikawa M, Murakami T, Komatsu H, Yamahara J, Matsuda H (1998) Heterocycles 47:397
2. Anam EM (1997) Indian J Chem Sect B 36:901
3. Knapp S (1995) Chem Rev 95:1859
4. Fuchs E-F, Lehmann J (1975) Chem Ber 108:2254
5. Fuchs E-F, Lehmann J (1976) Carbohydr Res 49:267
6. Yoshimura J, Ando H, Sato T, Tsuchida S, Hashimoto H (1976) Bull Chem Soc Jpn 49:2511
7. Tokura S, Ikeuchi Y, Nishimura SI, Nishi N (1983) Int J Biol Macromol 5:249
8. García Martín MdG, Violante de Paz Báñez M, Galbis Pérez JA (1993) Carbohydr Res 240:301
9. Wessel HP, Mitchell CM, Lobato CM, Schmid G (1995) Angew Chem 107:2920, Angew Chem Int Ed Engl 34:2712
10. Nicolaou KC, Flörke H, Egan MG, Barth T, Estevez VA (1995) Tetrahedron Lett 36:1775
11. McDevitt JP, Lansbury Jr. PT (1996) J Am Chem Soc 118:3818
12. Sabesan S (1997) Tetrahedron Lett 38:3127
13. Ramamoorthy PS, Gervay J (1997) J Org Chem 62:7801
14. Müller C, Kitas E, Wessel HP (1995) J Chem Soc Chem Commun 2425
15. Schweizer F (2002) Angew Chem Int Ed 41:230 and references cited
16. Bornaghi LF, Wilkinson BL, Kiefel MJ, Poulsen S-A (2004) Tetrahedron Lett 45:9281
17. Edwards AA, Ichihara O, Murfin S, Wilkes R, Whittaker M, Watkin DJ, Fleet GWJ (2004) J Comb Chem 6:230
18. Timmers CM, Turner JJ, Ward CM, van der Marel GA, Kouwijzer MLCE, Grootenhuis PDJ, van Boom JH (1997) Chem Eur J 3:920
19. Drouillat B, Kellam B, Dekany G, Starr MS, Toth I (1997) Bioorg Med Chem Lett 7:2247
20. D'Onofrio J, de Champdoré M, De Napoli L, Montesarchio D, Di Fabio G (2005) Bioconjugate Chem 16:1299
21. Smith MD, Long DD, Marquess DG, Claridge TDW, Fleet GWJ (1998) J Chem Soc Chem Commun 2039
22. McCort I, Duréault A, Depezay J-C (1998) Tetrahedron Lett 39:4463
23. Byrgesen E, Nielsen J, Willert M, Bols M (1997) Tetrahedron Lett 38:5697
24. Gremyachinskiy DE, Samoshin VV, Gross PH (2003) Tetrahedron Lett 44:6587
25. Nishimura SI, Nomura S, Yamada K (1998) J Chem Soc Chem Commun 617
26. Goodnow Jr. RA, Richou A-R, Tam S (1997) Tetrahedron Lett 38:3195
27. Goodnow Jr. RA, Tam S, Pruess DL, McComas WW (1997) Tetrahedron Lett 38:3199
28. Baisch G, Öhrlein R (1998) Carbohydr Res 312:61
29. Baisch G, Öhrlein R (1998) Bioorg Med Chem 6:1673
30. Gervay J, Flaherty TM, Nguyen C (1997) Tetrahedron Lett 38:1493
31. Szabo L, Smith BL, McReynolds KD, Parrill AL, Morris ER, Gervay J (1998) J Org Chem 63:1074
32. Suhara Y, Hildreth JEK, Ichikawa Y (1996) Tetrahedron Lett 37:1575
33. Suhara Y, Izumi M, Ichikawa M, Penno MB, Ichikawa Y (1997) Tetrahedron Lett 38:7167
34. Suhara Y, Yamaguchi Y, Collins B, Schnaar RL, Yanagishita M, Hildreth JEK, Shimada I, Ichikawa Y (2002) Bioorg Med Chem 10: 1999
35. Graf von Roedern E, Kessler H (1994) Angew Chem 106:684, Angew Chem Int Ed Engl 33:687
36. Durrat F, Xie J, Valéry J-M (2004) Tetrahedron Lett 45:1477
37. Sicherl F, Wittmann V (2005) Angew Chem Int Ed 44:2096
38. Graf von Roedern E, Lohof E, Hessler G, Hoffmann M, Kessler H (1996) J Am Chem Soc 118:10156
39. Kessler H, Gratias R, Hessler G, Gurrath M, Müller G (1996) Pure Appl Chem 68:1201
40. Gruner SAW, Locardi E, Lohof E, Kessler H (2002) Chem Rev 102:491 and references cited
41. van Well RM, Overkleeft HS, Overhand M, Carstenen EV, van der Marel GA, van Boom JH (2000) Tetrahedron Lett 41:9331
42. Chakraborty TK, Roy S, Koley D, Dutta SK, Kunwar AC (2006) J Org Chem 71:6240, and references cited
43. Locardi E, Stöckle M, Gruner S, Kessler H (2001) J Am Chem Soc 123:8189
44. Chakraborty TK, Srinivasu P, Bikshapathy E, Nagaray R, Vairamani M, Kumar SK, Kunwar AC (2003) J Org Chem 68:6257

45. van Well RM, Marinelli L, Erkelens K, van der Marel GA, Lavecchia A, Overkleeft HS, van Boom JH, Kessler H, Overhand M (2003) Eur J Org Chem 2303
46. Fleet GWJ, Johnson SW, Jones JH (2006) J Peptide Sci 12:599
47. Hunter DFA, Fleet GWJ (2003) Tetrahedron Asymm 14:3831
48. Mayes BA, Stetz RJE, Watterson MP, Edwards AA, Ansell CWG, Tranter GE, Fleet GWJ (2004) Tetrahedron Asymm 15:627
49. Mayes BA, Stetz RJE, Ansell CWG, Fleet GWJ (2004) Tetrahedron Lett 45:153
50. Mayes BA, Simon L, Watkin DJ, Ansell CWG, Fleet GWJ (2004) Tetrahedron Lett 45:153
51. de Gracia Garcia-Martin M, Pérez RR, Hernandez EB, Galbis JA (2001) Carbohydr Res 333:95
52. Mancera M, Roffé I, Rivas M, Galbis JA (2003) Carbohydr Res 338:1115
53. de Gracia Garcia-Martin M, Hernandez EB, Pérez RR, Alla A, Munoz-Guerra S, Galbis JA (2004) Macromolecules 37:5550
54. Romero Zaliz CL, Varela O (2005) Tetrahedron Asymm 16:97
55. Romero Zaliz CL, Erra-Balsells R, Nonami H, Sato Y, Varela O (2005) Arkivoc (xii):76
56. Gellman SH (1998) Acc Chem Res 31:173
57. Suhara Y, Kurihara M, Kittaka A, Ichikawa Y (2006) Tetrahedron 62:8207
58. Smith MD, Claridge TDW, Tranter GE, Sansom MSP, Fleet GWJ (1998) J Chem Soc Chem Commun 2041
59. Smith MD, Claridge TDW, Sansom MSP, Fleet GWJ (2003) Org Biomol Chem 1:3647
60. Chakraborthy TK, Srinivasu P, Madhavendra SS, Kumar SK, Kunwar AC (2004) Tetrahedron Lett 45:3573
61. Brittain DEA, WatersonMP, Claridge TDW, Smith MD, Fleet GWJ (2000) J Chem Soc Perkin Trans 1 3655
62. Claridge TDW, Long DD, Hungerford NL, Aplin RT, Smith MD, Marquess DG, Fleet GWJ (1999) Tetrahedron Lett 40:2199
63. Baker CM, Odell B, Grant GH, Edwards AA, Tranter GE, Fleet GWJ, Smith MD (2005) J Org Chem 70:2082
64. Baron R, Bakowies D, van Gunsteren WF (2004) Angew Chem Int Ed 43:4055
65. Baron R, Bakowies D, van Gunsteren WF (2005) J Peptide Sci 11:74
66. Edwards AA, Fleet GWJ, Tranter GE (2006) Chirality 18:265
67. Chakraborthy TK, Jayaprakash S, Srinivasu P, Govardhana Chary M, Diwan PV, Nagaraj R, Ravi Sankar A, Kunwar AC (2000) Tetrahedron Lett 41:8167
68. Smith MD, Long DD, Marquess DG, Claridge TDW, Fleet GWJ (1999) Tetrahedron Lett 40:2191
69. Edwards AA, Sanjayan GJ, Hachisu S, Tranter GE, Fleet GWJ (2006) Tetrahedron 62:7718
70. Claridge TDW, Goodman JM, Moreno A, Angus D, Barker SF, Taillefumier C, Watterson MP, Fleet GWJ (2001) Tetrahedron Lett 42:4251
71. Johnson SW, Jenkinson SF, Pérez-Victoria I, Edwards AA, Claridge TDW, Tranter GE, Fleet GWJ, Jones JH (2005) J Peptide Sci 11:517
72. Lee MD, Dunne TS, Siegel MM, Chang CC, Morton GO, Borders DB (1987) J Am Chem Soc 109:3464
73. Lee MD, Dunne TS, Chang CC, Ellestad GA, Siegel MM, Morton GO, McGahren WJ, Borders DB (1987) J Am Chem Soc 109:3466
74. Golik J, Clardy J, Dubay G, Groenewold G, Kawaguchi H, Konishi M, Krishnan B, Ohkuma H, Saitoh K-i, Doyle TW (1987) J Am Chem Soc 109:3461
75. Golik J, Dubay G, Groenewold G, Kawaguchi H, Konishi M, Krishnan B, Ohkuma H, Saitoh K-i, Doyle TW (1987) J Am Chem Soc 109:3462
76. Nicolaou KC, Dai W-M (1991) Angew Chem 103:1453, Angew Chem Int Ed Engl 30:1387
77. Nicolaou KC, Groneberg RD (1990) J Am Chem Soc 112:4085
78. Nicolaou KC, Groneberg RD, Miyazaki T, Stylianides NA, Schulze TJ, Stahl W (1990) J Am Chem Soc 112:8193
79. Yang D, Kim SH, Kahne D (1991) J Am Chem Soc 113:4715
80. Halcomb RL, Wittman MD, Olson SH, Danishefsky SJ (1991) J Am Chem Soc 113:5080
81. Bamhaoud T, Lancelin J-M, Beau J-M (1992) J Chem Soc Chem Commun 1494
82. Vonhoff S, Heightman TD, Vasella A (1998) Helv Chim Acta 81:1710
83. Liu PS (1987) J Org Chem 52:4717
84. Anzeveno PB, Creemer LJ, Daniel JK, King C-HR, Liu PS (1989) J Org Chem 54:2539
85. Saavedra OM, Martin OR (1996) J Org Chem 61:6987
86. Ando O, Satake H, Itoi K, Sato A, Nakajima M, Takahashi S, Haruyama H, Ohkuma Y, Kinoshita T, Enokita R (1991) J Antibiot 44:1165
87. Murao S, Sakai T, Gibo H, Nakayama T, Shin T (1991) Agric Biol Chem 55:895

88. Ogawa S, Uchida C (1993) Chem Lett 173
89. Kobayashi Y, Miyazaki H, Shiozaki M (1994) J Org Chem 59:813
90. Ledford BE, Carreira EM (1995) J Am Chem Soc 117:11811
91. Boiron A, Zillig P, Faber D, Giese B (1998) J Org Chem 63:5877
92. Uchida C, Yamagishi T, Ogawa S (1994) J Chem Soc Perkin Trans 1 589
93. Uchida C, Kitahashi H, Yamagishi T, Iwaisaki Y, Ogawa S (1994) J Chem Soc Perkin Trans 1: 2775
94. Uchida C, Ogawa S (1996) Bioorg Med Chem 4:275
95. Uchida C, Kitahashi H, Watanabe S, Ogawa S (1995) J Chem Soc Perkin Trans 1:1707
96. Kobayashi Y, Shiozaki M, Ando O (1995) J Org Chem 60:2570
97. Knapp S, Purandare A, Rupitz K, Whithers SG (1994) J Am Chem Soc 116:7461
98. Shiozaki M, Ubukata O, Haruyama H, Yoshiike R (1998) Tetrahedron Lett 39:1925
99. Nishiyama T, Isobe M, Ichikawa Y (2005) Angew Chem Int Ed 44:4372
100. García Fernández JM, Ortiz Mellet C, Díaz Pérez VM, Fuentes J, Kovács J, Pintér I (1997) Tetrahedron Lett 38:4161
101. Díaz Pérez VM, Ortiz Mellet C, Fuentes J, García Fernández JM (2000) Carbohydr Res 326:161
102. Kovács J, Ósz E, Györgydeák (2002) Carbohydr Res 337:1171
103. Bianchi A, Ferrario D, Bernardi A (2006) Carbohydr Res 341:1438
104. García Fernández JM, Ortiz Mellet C, Díaz Pérez VM, Fuentes J, Kovács J, Pintér I (1997) Carbohydr Res 304:261
105. García-Moreno MI, Ortiz Mellet C, García Fernández JM (1999) Tetrahedron Asymm 10:4271
106. García-Moreno MI, Ortiz Mellet C, García Fernández JM (2004) Eur J Org Chem 1803 and references cited
107. Jochims JC, Seeliger A (1965) Tetrahedron 21:2611
108. Wiley PF, McMichael DL, Koert JM, Wiley VH (1976) J Antibiotics 29:1218
109. Ichikawa Y, Matsukawa Y, Nishiyama T, Isobe M (2004) Eur J Org Chem 586 and references cited
110. Ichikawa Y, Nishiyama T, Isobe M (2004) Tetrahedron 60:2621
111. Prosperi D, Ronchi S, Lay L, Rencurosi A, Russo G (2004) Eur J Org Chem 395
112. Ichikawa Y, Matsukawa Y, Nishiyama T, Isobe M (2006) J Am Chem Soc 128:3934 and references cited
113. Wessel HP, Plessis I, Cassel S, Arnold W (1997) J Carbohydr Chem 16:789
114. Jiménez Blanco JL, Bootello P, Benito JM, Ortiz Mellet C, García Fernández JM (2006) J Org Chem 71:5136 and references cited
115. Beaucage SL, Iyer RP (1993) Tetrahedron 49:6123
116. De Mesmaeker A, Häner R, Martin P, Moser HE (1995) Acc Chem Res 28:366
117. Dempcy RO, Browne KA, Bruice TC (1995) J Am Chem Soc 117:6140
118. Waldner A, De Mesmaeker A (1995) Synlett 108
119. De Mesmaeker A, Jouanno C, Wolf RM, Wendeborn S (1997) Bioorg Med Chem Lett 7:447
120. García Fernández JM, Jiménez Blanco JL, Ortiz Mellet C, Fuentes J (1995) J Chem Soc Chem Commun 57
121. López Ó, Maza S, Maya I, Fuentes J, Fernández-Bolaños JG (2005) Tetrahedron 61:9058
122. Charbonnier F, Marsura A, Roussel K, Kovács J, Pintér I (2001) Helv Chim Acta 84:535
123. Patel A, Lindhorst TK (2001) J Org Chem 66:2674
124. Bachmann F, Ruppenstein M, Thiem J (2001) J Polymer Sci: Part A: Polymer Chemistry 39:2332
125. Prosperi D, Ronchi S, Panza L, Rencurosi A, Russo G (2004) Synlett 9:1529
126. Peri F, Cipolla L, La Ferla B, Nicotra F (2003) C R Chimie 6:635
127. Peri F, Nicotra F (2004) Chem Commun 623
128. Peri F, Jiménez-Barbero J, García-Aparicio V, Tvaroška I, Nicotra F (2004) Chem Eur J 10:1433
129. Gourlain T, Wadouachi A, Beaupère D (1999) Synthesis 290
130. Phiasivongsa P, Gallagher J, Chen C-N, Jones PR, Samoshin VV, Gross PH (2002) Org Lett 4:4587
131. Marcaurelle LA, Bertozzi CR (2001) J Am Chem Soc 123:1587
132. McAuliffe JC, Rabuka D, Hindsgaul O (2002) Org Lett 4 :3067
133. Galan MC, Venot AP, Glushka J, Imberty A, Boons J-G (2003) Org Biomol Chem 1:3891
134. Lockhoff O (1998) Angew Chem Int Ed Engl 37:3436
135. Westermann B, Dörner S (2005) Chem Commun 2116

136. Le Chevalier A, Pierre R, Kanso R, Chambert S, Doutheau A, Queneau Y (2006) 47 :2431
137. Menand M, Blais J-C, Hamon L, Valéry J-M, Xie J (2005) J Org Chem 70:4423
138. Jegge S, Lehmann J (1984) Carbohydr Res 133:247
139. Lehmann J, Schmidt-Schuchardt M (1994) Meth Enzymology 247:265
140. Lehmann J, Petry S (1990) Carbohydr Res 204:141
141. Ats SC, Lehmann J, Petry S (1994) Carbohydr Res 252:325
142. Niggemann J, Kamerling JP, Vliegenthart JFG (1998) J Chem Soc Perkin Trans 1:3011
143. Patch RJ, Chen H, Pandit CR (1997) J Org Chem 62:1543
144. Dominique R, Das SK, Roy R (1998) J Chem Soc Chem Commun 2437
145. Roy R, Dominique R, Das SK (1999) J Org Chem 64:5408
146. Kirschning A, Chen G (1999) Tetrahedron Lett 40:4665
147. Kirschning A, Chen G, Jaunzems J, Jesberger M, Kalesse M, Lindner M (2004) Tetrahedron 60:3505
148. Jaunzems J, Oelze B, Kirschning A (2004) Org Biomol Chem 2:3448 and references cited
149. Blackwell HE, O'Leary DJ, Chatterjee AK, Washenfelder RA, Bussmann DA, Grubbs RH (2000) J Am Chem Soc 122:58
150. Schmidtmann FW, Benedum TE, McGarvey GJ (2005) Tetrahedron Lett 46:4677
151. Gómez AM, Barrio A, Amurrio I, Valverde S, Jarosz S, López JC (2006) Tetrahedron Lett 47:6243
152. Sinou D, Lhoste P, Pichon N, Krycka B, Porwanski S, Zawisza A (2002) 21:545
153. Thiele G, Rottmann A, Germer A, Kleinpeter E, Spindler K-D, Synstad B, Eijsink VGH, Peter MG (2002) J Carbohydr Chem 21:471
154. Germer A, Peter MG, Kleinpeter E (2002) J Org Chem 67:6328
155. Choudhury AK, Kitaoka M, Hayashi K (2003) Eur J Org Chem 2462
156. Pitt N, Duane RM, Bradley H, Wilson SJ, O'Boyle KM, Murphy PV (2004) Carbohydr Res 339:1873
157. Nishiyama T, Ichikawa Y, Isobe M (2003) Synlett 47
158. Klauser RJ, Meinetsberger E, Raake W (1991) Semin Thromb Hemost 17 (1):118
159. Hoppenstedt D, Ahsan A, Walenga JM, Fareed J, Schumacher H (1988) Blood 72:198a
160. Papoulias UE, Wyld PJ, Haas S, Stemberger A, Jeske W, Hoppensteadt D, Kämmereit A (1993) Thromb Res 72:99
161. Anon (1994) Drugs of the Future 19:721
162. van Boeckel CAA, Petitou M (1993) Angew Chem 105:1741, Angew Chem Int Ed Engl 32:1671
163. Lucas H, Basten JEM, van Dinther TG, Meuleman DG, van Aelst SF, van Boeckel CAA (1990) Tetrahedron 46:8207
164. van Boeckel CAA, Basten JEM, Lucas H, van Aelst SF (1988) Angew Chem 100:1217, Angew Chem Int Ed Engl 27:1177
165. van Boeckel CAA, Wagenaars GN, Mellema J-R (1988) Recl Trav Chim Pays-Bas 107:649
166. Rele SM, Iyer SS, Chaikof EL (2003) Tetrahedron Lett 44:89
167. Iyer SS, Rele SM, Baskaran S, Chaikof EL (2003) Tetrahedron 59:631
168. Vasella A (1998) Pure Appl Chem 70:425
169. Alzeer J, Vasella A (1995) Helv Chim Acta 78:177
170. Bohner TV, Becker O-S, Vasella A (1999) Helv Chim Acta 82:198
171. Bürli R, Vasella A (1997) Helv Chim Acta 80:1027
172. Bürli R, Vasella A (1997) Helv Chim Acta 80:2215
173. Bürli R, Vasella A (1997) Angew Chem 109:1945, Angew Chem Int Ed Engl 36:1852
174. Stichler-Bonaparte J, Vasella A (2001) Helv Chim Acta 84:2355
175. Stichler-Bonaparte J, Bernet B, Vasella A (2002) Helv Chim Acta 85:2235
176. Hoffmann B, Zanini D, Ripoche I, Bürli R, Vasella A (2001) Helv Chim Acta 84:1862
177. Hoffmann B, Bernet B, Vasella A (2002) Helv Chim Acta 85:265
178. Miao Z, Xu M, Hoffmann B, Bernet B, Vasella A (2005) Helv Chim Acta 88:1885
179. Vasella A (2005) Chimia 59:785 and references cited
180. Zhong S, Mondon M, Pilard S, Len C (2006) Tetrahedron Lett 47:6221
181. Vicent C, Martín-Lomas M, Penadés S (1989) Tetrahedron 45:3605
182. Vicent C, Jiménez-Barbero J, Martín-Lomas M, Penadés S, Cano FH, Foces-Foces C (1991) J Chem Soc Perkin Trans 2 905
183. Tsuzuki M, Tsuchiya T (1998) Carbohydr Res 311:11
184. Yamamura H, Yamada S, Kohno K, Okuda N, Araki S, Kobayashi K, Katakai R, Kano K,

Kawai M (1999) J Chem Soc Perkin Trans 1 2943
185. Bistri O, Lecourt T, Mallet J-M, Sollogoub M, Sinaÿ (2004) Chem Biodiversity 1:129
186. Shiozaki M, Ubukata O, Haruyama H, Yoshiike R (1998) Tetrahedron Lett 39:1925
187. Shiozaki M, Yoshiike R, Ando O, Ubukata O, Haruyama H (1998) Tetrahedron 54:15167
188. Johns BA, Johnson CR (1998) Tetrahedron Lett 39:749
189. Compain P, Martin O (2001) Bioorg Med Chem 9:3077
190. Suhara Y, Ichikawa M, Hildreth JEK, Ichikawa Y (1996) Tetrahedron Lett 37:2549
191. Christensen MK, Meldal M, Bock K, Cordes H, Mouritsen S, Elsner H (1994) J Chem Soc Perkin Trans 1 1299
192. Berkowitz DB, Maiti G, Charette BD, Dreis CD, MacDonald RG (2004) Org Lett 6:4921
193. Franzyk H, Meldal M, Paulsen H, Thiel S, Jensenius JC, Bock K (1996) Bioorg Med Chem 4:1881
194. Chakraborty TK, Roy S, Kumar SK, Kunwar AC (2005) Tetrahedron Lett 46:3065
195. Stöckle M, Voll G, Günther R, Lohof E, Locardi E, Gruner S, Kessler H (2002) Org Lett 4:2501
196. van Well RM, Overkleeft HS, van der Marel GA, Bruss D, Thibault G, de Groot PG, van Boom JH, Overhand M (2003) Bioorg Med Chem Lett 13:331
197. Raunkjær M, El Oualid F, van der Marel GA, Overkleeft HS, , Overhand M (2004) Org Lett 6:3167
198. Katajisto J, Lönnberg H (2005) Eur J Org Chem 3518
199. Billing JF, Nilsson UJ (2005) Tetrahedron Lett 46:991
200. Gruner SAW, Truffault V, Voll G, Locardi E, Stöckle M, Kessler H (2002) Chem Eur J 8:4365
201. Peerlings HWI, Nepogodiev SA, Stoddart JF, Meijer EW (1998) Eur J Org Chem 1879
202. André S, Pieters RJ, Vrasidas I, Kaltner H, Kuwabara I, Liu F-T, Liskamp RMJ, Gabius H-J (2001) ChemBioChem 2:822
203. Boysen MMK, Elsner K, Sperling O, Lindhorst TK (2003) Eur J Org Chem 4376
204. Sato M, Furuike T, Sadamoto R, Fujitani N, Nakahara T, Niikura K, Monde K, Kondo H, Nishimura S-I (2004) J Am Chem Soc 126:14013
205. Lee YC, Lee RT (1994) In: Lee YC, Lee RT (eds) Neoglycoconjugates: Preparation and Applications. Academic Press, San Diego
206. Kieburg C, Sadalapure K, Lindhorst TK (2000) Eur J Org Chem 2035
207. Lindhorst TK, Ludewig M, Thiem J (1998) J Carbohydr Chem 17:1131
208. Liu B, Roy R (2002) Chem Commun 594
209. Amaya T, Tanaka H, Takahashi T (2004) Synlett 503
210. Walter M, Wiegand M, Lindhorst TK (2006) Eur J Org Chem 719
211. Roy R (1997) Top Curr Chem 187:241
212. Gardiner JM (1998) Exp Opin Invest Drugs 7:405
213. García-López JJ, Santoyo-González F, Vargas-Berenguel A, Giménez-Martínez JJ (1999) Chem Eur J 5:1775
214. Fulton DA, Pease AR, Stoddart JF (2000) Israel J Chem 40:325
215. Larpent C, Laplace A, Zemb T (2004) Angew Chem Int Ed 43:3163
216. Patel A, Lindhorst TK (2006) Carbohydr Res 341:1657
217. Vrasidas I, André S, Valentini P, Böck C, Lensch M, Kaltner H, Liskamp RMJ, Gabius H-J, Pieters RJ (2003) Org Biomol Chem 1:803
218. Walter M, Lindhorst TK (2006) Synthesis 952
219. Westerduin P, Basten JEM, Broekhoven MA, de Kimpe V, Kuijpers WHA, van Boeckel CAA (1996) Angew Chem 108:339, Angew Chem Int Ed Engl 35:331
220. Lubineau A, Escher S, Alais J, Bonaffé D (1997) Tetrahedron Lett 38:4087
221. Shimizu T, Masuda M (1997) J Am Chem Soc 119:2812
222. Castro MJL, Kovensky J, Fernández Cirelli A (1997) Tetrahedron Lett 38:3995
223. Simanek EE, MacGarvey GJ, Jablonowski JA, Wong C-H (1998) Chem Rev 98:833
224. Kaila N, Thomas IV BE (2002) Med Res Rev 22:566
225. Unger FM (2001) Adv Carbohydr Chem Biochem 57:207
226. Sears P, Wong C-H (1998) J Chem Soc Chem Commun 1161
227. Magnani JL (2004) Archives Biochem Biophys 426:122
228. Toepfer A, Kretschmar G, Bartnik E (1995) Tetrahedron Lett 36:9161
229. Heskamp BM, Veeneman GH, van der Marel GA, van Boeckel CAA, van Boom JH (1995) Rec Trav Chim Pays-Bas 114:398
230. Dekany G, Wright K, Ward P, Toth I (1996) J Carbohydr Chem 15:383

231. Allenson NM, Davidson AH, Floyd CD, Martin FM (1994) Tetrahedron Asymmetry 5:2061
232. Hanessian S, Prabhanjan H (1994) Synlett 868
233. Ragan JA, Cooper K (1994) Bioorg Med Chem Lett 4:2563
234. Prodger JC, Bamford MJ, Gore PM, Holmes DS, Saez V, Ward P (1995) Tetrahedron Lett 36: 2339
235. Bamford MJ, Bird M, Gore PM, Holmes DS, Priest R, Prodger JC, Saez V (1996) Bioorg Med Chem Lett 6:239
236. Uchiyama T, Vassilev VP, Kajimoto T, Wong W, Huang H, Lin C-C, Wong C-H (1995) J Am Chem Soc 117:5395
237. Carrel F, Giraud S, Spertini O, Vogel P (2004) Helv Chim Acta 87:1048
238. Toepfer A, Kretschmar G, Schuth S, Sonnentag M (1997) Bioorg Med Chem Lett 7:1317
239. Marinier A, Martel A, Bachand C, Plamondon S, Turmel B, Daris J-P, Banville J, Lapointe P, Ouellet C, Dextraze P, Menard M, Wright JJK, Alford J, Lee D, Stanley P, Nair X, Todderud G, Tramposch KM (2001) Bioorg Med Chem 9:1395
240. Wu SH, Shimazaki M, Lin C-C, Moree WJ, Weitz-Schmidt G, Wong C-H (1996) Angew Chem 108:106, Angew Chem Int Ed Engl 35:88
241. Woltering TJ, Weitz-Schmidt G, Wong C-H (1996) Tetrahedron Lett 37:9033
242. Kiyoi T, Kondo H (1998) Bioorg Med Chem Lett 8:2845
243. Huwe CM, Woltering TJ, Jiricek J, Weitz-Schmidt G, Wong C-H (1999) Bioorg Med Chem 7:773
244. Kaila N, Thomas IV BE, Thakker P, Alvarez JC, Camphausen RT, Crommie D (2001) Bioorg Med Chem Lett 11:151
245. Cappi MW, Moree WJ, Qiao L, Marron TG, Weitz-Schmidt G, Wong C-H (1996) Angew Chem 108:2501, Angew Chem Int Ed Engl 35:2346
246. Marron TG, Woltering TJ, Weitz-Schmidt G, Wong C-H (1996) Tetrahedron Lett 37:9037
247. Wong C-H, Moris-Varas F, Hung SC, Marron TG, Gong KW, Weitz-Schmidt G (1997) J Am Chem Soc 119:8152
248. Tsukida T, Moriyama H, Kurokawa K, Achiha T, Inoue Y, Kondo H (1998) J Med Chem 41:4279
249. Kurokawa K, Kumihara H, Kondo H (2000) Bioorg Med Chem Lett 10:1827
250. Sutherlin DP, Stark TM, Hughes R, Armstrong RW (1996) J Org Chem 61:8350
251. Tsai C-Y, Park WKC, Weitz-Schmidt G, Ernst B, Wong C-H (1998) Bioorg Med Chem Lett 8:2333
252. Lampe TFJ, Weitz-Schmidt G, Wong C-H (1998) Angew Chem 110:1761; Angew Chem Int Ed Engl 37:1711
253. Thiem J, Bachmann F (1994) Trends Polym Sci 2:425
254. Varela O, Orgueira HA (2000) Adv Carbohydr Chem Biochem 55:137
255. Thiem J, Lüders H (1986) Makromol Chem 187:2775
256. Bachmann F, Reimer J, Ruppenstein M, Thiem J (2001) Macromol Chem Phys 202:3410
257. Thiem J, Bachmann F (1993) Makromol Chem 194:1035
258. Kiely DE, Chen L, Lin T-H (1994) J Am Chem Soc 116 :571
259. Morton DW, Kiely DE (2000) J Polym Sci, Part A: Polym Chem 38:3085
260. Carter A, DW, Kiely DE (2000) J Polym Sci, Part A: Polym Chem 38:3892
261. Styron SD, Kiely DE, Ponder G (2003) J Carbohydr Chem 22:123
262. Henkensmeier D, Abele BC, Candussio A, Thiem J (2004) Polymer 45:7053
263. Orgueira HA, Varela O (2001) J Polym Sci, Part A: Polym Chem 39:1024
264. Mancera M, Roffé I, Al-Kass SSJ, Rivas M, Galbis JA (2003) Macromolecules 36:1089
265. Mancera M, Zamora F, Roffé I, Bermúdez M, Alla A, Muñoz-Guerra S, Galbis JA (2004) Macromolecules 37:2779
266. Garcia-Martín MG, Hernández EB, Pérez RR, Alla A, Muñoz-Guerra S, Galbis JA (2004) Macromolecules 37:5550
267. Mayes BA, Cowley AR, Ansell CWG, Fleet GWJ (2004) Tetrahedron Lett 45:163
268. Wessel HP, Chucholowski A, Fingerle J, Iberg N, Märki HP, Müller R, Pech M, Pfister-Downar M, Rouge M, Schmid G, Tschopp T (1998) From glycosaminoglycans to heparinoid mimetics with antiproliferative activity. In: Chapleur Y (ed) Carbohydrate mimics: concepts and methods. Verlag Chemie, Weinheim, p 417
269. Wessel HP, Tschopp TB, Hosang M, Iberg N (1994) Bioorg Med Chem Lett 4:1419
270. Wessel HP, Iberg N, Trumtel M, Viaud M-C (1996) Bioorg Med Chem Lett 6:27
271. Wessel HP (1997) Top Curr Chem 187:215
272. Koshida S, Suda Y, Sobel M, Kusumoto S (2001) Tetrahedron Lett 42:1289

273. Koshida S, Suda Y, Arano A, Sobel M, Kusumoto S (2001) Tetrahedron Lett 42:1293
274. Hansen HC, Magnusson G (1998) Carbohydr Res 307:243
275. Schwizer D, Gäthje H, Kelm S, Porro M, Schwardt O, Ernst B (2006) Bioorg Med Chem 14:4944
276. Tewari N, Ramesh, Mishra RC, Tripathi RP, Srivastava VML, Gupta S (2004) Bioorg Med Chem Lett 14:4055
277. Feldman KS, Wilson SL, Lawlor MD, Lang CH, Scheuchenzuber WJ (2002) Bioorg Med Chem 10:47
278. Bradley H, Fitzpatrick G, Glass WK, Kunz H, Murphy PV (2001) Org Lett 3:2629
279. Murphy PV, Bradley H, Tosin M, Pitt N, Fitzpatrick G, Glass WK (2003) J Org Chem 68:5692
280. Xu J, Egger A, Bernet B, Vasella A (1996) Helv Chim Acta 79:2004
281. Roy R, Das SK, Santoyo-González F, Hernández-Mateo F, Dam TK, Brewer CF (2000) Chem Eur J 6:1757
282. Dondoni A, Marra A, Zampolli MG (2002) Synlett 1850
283. Xu J, Vasella A (1999) Helv Chim Acta 82:1728
284. Murty KVSN, Xie T, Bernet B, Vasella A (2006) Helv Chim Acta 89:675
285. Marmuse L, Nepogodiev AS, Field RA (2005) Org Biomol Chem 3:2225
286. Bozó É, Boros S, Kuszmann J (1997) Carbohydr Res 304:271
287. Périon R, Ferrières V, García-Moreno MI, Mellet CO, Duval R, Garcia Fernández JM, Plusquellec D (2005) Tetrahedron 61: 9118
288. Hotha S, Kashyap S (2006) J Org Chem 71:364
289. Oliveira RN, Sinou D, Srivastava RM (2006) J Carbohydr Chem 25:407
290. Temelkoff DP, Zeller M, Norris P (2006) Carbohydr Res 341:1081
291. Bodine KD, Gin DY, Gin MS (2005) Org Letters 7:4479
292. Calvo-Flores FG, Isac-García J, Hernández-Mateo F, Pérez-Balderas F, Calvo-Asín JA, Sanchéz-Vaquero E, Santoyo-González F (2000) Org Lett 2:2499
293. Schürer SC, Blechert S (1999) Chem Commun 1203
294. Dondoni A, Massi A, Sabbatini S, Bertolasi V (2002) J Org Chem 67:6979
295. Giovenzana GB, Lay L, Monti D, Palmisano G, Panza L (1999) Tetrahedron 55:14123
296. Tietze LF, Griesbach U, Schuberth I, Bothe U, Marra A, Dondone A (2003) Chem Eur J 9:1296
297. Ronchi S, Prosperi D, Thimon C, Morin C, Panza L (2005) Tetrahedron Asymm 16:39
298. Bürli R, Vasella A (1999) Helv Chim Acta 82:485
299. Bornaghi LF, Wilkinson BL, Kiefel MJ, Poulsen S-A (2004) Tetrahedron Lett 45:9281
300. Velasco-Torrijos T, Murphy PV (2005) Tetrahedron:Asymm 16:261
301. Moreno-Vargas AJ, Fernández-Bolaños JG, Fuentes J, Robina I (2001) Tetrahedron Lett 42:1283
302. Furuya T, Asada Y, Mizobata S, Matsuura Y, Hamada H (1998) Phytochemistry 49:109
303. Hamann PR, Hinman LM, Beyer CF, Lindh D, Upeslacis J, Flowers DA, Bernstein I (2002) Bioconjugate J 13:40
304. Giorgio E, Tanaka K, Ding W, Krishnamurthy G, Pitts K, Ellestad GA, Rosini C, Berova N (2005) Bioorg Med Chem Lett 13:5072
305. Yoshida O, Yasukata T, Sumino Y, Munekage T, Narukawa Y, Nishitani Y (2002) Bioorg Med Chem Lett 12:3027
306. Huang H, Wong C-H (1995) J Org Chem 60:3100
307. Kolb HC, Ernst B (1997) Chem Eur J 3:1571
308. Kolb HC, Ernst B (1997) Pure Appl Chem 69:1879
309. Bänteli R, Herold P, Bruns C, Patton JT, Magnani JL, Thoma G (2000) Helv Chim Acta 83:2893
310. Thoma G, Magnani JL, Patton JT (2001) Bioorg Med Chem Lett 11:923
311. Wang R, Wong C-H (1996) Tetrahedron Lett 37:5427
312. Lin C-C, Shimazaki M, Heck M-P, Aoki S, Wang R, Kimura T, Ritzèn H, Takayama S, Wu SH, Weitz-Schmidt G, Wong C-H (1996) J Am Chem Soc 118:6826
313. Hanessian S, Reddy GV, Huynh HK, Pan J, Pedatella S, Ernst B, Kolb HC (1997) Bioorg Med Chem Lett 7:2729
314. Kaila N, Yu H-A, Xiang Y (1995) Tetrahedron Lett 36:5503
315. Birkbeck AA, Ley SV, Prodger JC (1995) Bioorg Med Chem Lett 5:2637
316. Kuribayashi T, Gohya S, Mizuno Y, Shimojima M, Ito K, Satoh S (1999) Synlett 6:737
317. Liu A, Dillon K, Campbell RM, Cox DC, Huryn DM (1996) Tetrahedron Lett 37:3785

318. Kretschmar G (1998) Tetrahedron 54:3765
319. Bänteli R, Ernst B (1997) Tetrahedron Lett 38:4059
320. Toepfer A, Kretschmar G (1997) Bioorg Med Chem Lett 7:1311
321. Kogan TP, Dupré B, Keller KM, Scott IL, Bui H, Market RV, Beck PJ, Voytus JA, Revelle BM, Scott D (1995) J Med Chem 38: 4976
322. Dupré B, Bui H, Scott IL, Market RV, Keller KM, Beck PJ, Kogan TP (1996) Bioorg Med Chem Lett 6:569
323. Murphy PV, Hubbard RE, Manallack DT, Montana JJ, Taylor RJK (1998) Tetrahedron Lett 39:3273
324. Murphy PV, Hubbard RE, Manallack DT, Wills RE, Montana JJ, Taylor RJK (1998) Biooorg Med Chem 6:2421
325. Hanessian S, Huynh HK, Reddy GV, McNaughton-Smith G, Ernst B, Kolb HC, Magnani J, Sweeley C (1998) Bioorg Med Chem Lett 8:2803
326. Kuribayashi T, Ohkawa N, Satoh S (1998) Bioorg Med Chem Lett 8:3307
327. Ikeda T, Kajimoto T, Kondo H, Wong C-H (1997) Bioorg Med Chem Lett 7:2485
328. Geyer A, Gege C, Schmidt RR (2000) Angew Chem Int Ed 39:3246
329. Sasaki K, Nishida Y, Tsurumi T, Uzawa H, Kondo H, Kobayashi K (2002) Angew Chem Int Ed 41:4463.
330. Ali M, Hicks AER, Hellewell PG, Thoma G, Norman KE (2004) Faseb J 18:152
331. Moriyama H, Hiramatsu Y, Kiyoi T, Achiha T, Inoue Y, Kondo H (2001) Bioorg Med Chem 9:1479
332. Fukunaga K, Tsukida T, Moriyama H, Kondo H (2001) Bioorg Med Chem Lett 11:2365
333. Kaila N, Somers WS, Thomas BE, Thakker P, Janz K, DeBernardo S, Tam S, Moore WJ, Yang R, Wrona W, Bedard PW, Crommie D, Keith jr JC, Tsao DHH, Alvarez JC, Ni H, Marchese E, Patton JT, Magnani JL, Camphausen RT (2005) J Med Chem 48:4346
334. O I, Kieber-Emmons T, Otvos L, Blaszczyk-Thurin M (2000) Biochem Biophys Res Commun 268:106
335. Cunto-Amesty G, Luo P, Monzavi-Kabassi B, Kieber-Emmons T (2001) Intern Rev Immunol 20:157
336. Sen G, Chakraborty M, Foon KA, Reisfeld RA, Bhattacharya-Chatterjee M (1998) J Immunother 21:75
337. Kieber-Emmons T (1998) Immunol Res 17:95
338. Prinz DM, Smithson SL, Westerink MAJ (2004) J Immunol Methods 285:1
339. Pashov AD, Plaxco J, Kaveri SV, Monzavi-Kabassi B, Harn D, Kieber-Emmons T (2006) J Biol Chem 281:29675
340. Agadjanyan M, Luo P, Westerink MAJ, Carey LA, Hutchins W, Steplewski Z, Weiner DB, Kieber-Emmons T (1997) Nature Biotechnol 15:547
341. (1997) Exp Opin Ther Patents 7:1345
342. Ishikawa D, Taki T (1999) Trends Glycosci Glycotech 11:277
343. Brett PJ, Tiwana H, Feavers IM, Charalambous BM (2002) J Biol Chem 277:20468
344. Weintraub A (2003) Carbohydr Res 338:2539
345. Phalipon A, Folgori A, Arondel J, Sgaramella G, Fortugno P, Cortese R, Sansonetti PJ, Felici F (1997) Eur J Immunol 27:2620
346. Beenhouwer DO, May RJ, Valadon P, Scharff MD (2002) J Immun 169:6992
347. Taki T, Ishikawa D, Hamasaki H, Handa S (1997) FEBS Lett 418:219
348. Arnusch CJ, André S, Valentini P, Lensch M, Russwurm R, Siebert H-C, Fischer MJE, Gabius H-J, Pieters RJ (2004) Bioorg Med Chem Lett 14:1437
349. Zou J, Glinsky VV, Landon LA, Matthews L, Deutscher SL (2005) Carcinogenesis 26:309
350. André S, Arnusch CJ, Kuwabara I, Russwurm R, Kaltner H, Gabius H-J, Pieters RJ (2004) Bioorg Med Chem Lett 14:1437
351. Meloen RH, Puijk WC, Slootstra JW (2000) J Mol Recognition 13:352
352. Monzavi-Kabassi B, Luo P, Jousheghany F, Torres-Quiñones M, Cunto-Amesty G, Artaud C, Kieber-Emmons T (2004) Cancer Res 64:2162
353. Hayashida M, Fujii T, Hamasu M, Ishiguro M, Hata Y (2003) J Mol Biol 334:551
354. Johnson MA, Jaseja M, Zou W, Jennings HJ, Copié V, Pinto BM, Pincus SH (2003) J Biol Chem 278:24740
355. Böhm H-J (1992) J Comput Aided Mol Drug Des 6:593
356. Luo P, Canziani G, Cunto-Amesty G, Kieber-Emmons T (2000) J Biol Chem 275:16146

Part 10
Key Technologies and Tools for Functional Glycobiology

10.1 Key Technologies and Tools for Functional Glycobiology: Introduction

Shin-Ichiro Nishimura
Laboratory of Advanced Chemical Biology, Graduate School of Advanced Life Science, Frontier Research Center for the Post-Genome Science and Technology, Hokkaido University and Drug-Seeds Discovery Research Laboratory, National Institute of Advanced Industrial Science and Technology, Sapporo, Japan
shin@glyco.sci.hokudai.ac.jp

Keywords

Glycome; Glycomics; Glycoproteomics; Functional glycomics; Microarray; Mass spectrometry; Chemical probe; Molecular imaging; Bioinformatics

Abbreviations

MS mass spectrometry
ORFs open reading frames
PCR polymerase chain reaction

1 Overview: New Tools for Glycobiology

Various glycans that are located at the cellular surfaces as glycosphingolipids, glycoproteins, and proteoglycans, deposited in the extracellular matrices and bound to secretory proteins have a variety of crucial roles in the phenotypic expression of cellular genotypes. General technologies such as chemical synthesis, spectrometric analysis using UV, fluorescence, CD, NMR,

In: *Glycoscience*. Fraser-Reid B, Tatsuta K, Thiem J (eds)
Chapter-DOI 10-1007/978-3-540-30429-6_53:

and SPR have greatly contributed to the insight into the molecular basis of structural features and functional roles of many glycoconjugates. Although there have been substantial advances in our understanding of the effects of glycosylation on some biological systems, we still do not fully understand the significance and mechanism of glycoform alteration observed widely in many human diseases. As a result, the therapeutic potential of complex glycans has not been well exploited with a few notable exceptions. This chapter describes background and provides a brief introduction of recent advances in new technologies and tools for glycomic approaches that should greatly accelerate functional glycobiology and their clinical applications.

2 Proteomics and Glycomics: Glycan Expression is not Template-Driven

Glycosylation is one of the most important posttranslational modifications of proteins in eukaryotes. This step is essential to modulate a wide range of protein (and lipid) functions both on the cellular surfaces and within the cells [1]. Therefore, the long-term challenge of glycomics is enormous; to define the identities, quantities, relationship between glycoforms and functions of the glycoconjugates, and to characterize how these properties vary in different cellular contexts. Proteomics, the analysis of genomic complements of proteins, has burst onto the broad scientific fields over the past few years, perhaps befitting a discipline that can enjoy the virtually instantaneous conversion of the genome sequence into a set of predicted proteins [2]. The development of method for parallel proteomics has relied on the rapid identification of open reading frames (ORFs) and their facile cloning and expression protocols. Cloning of a genomic set of ORFs and the chemical synthesis of gene-specific primers that are suitable for amplification by the polymerase chain reaction (PCR) and for subsequent insertion of the PCR products into appropriate plasmids permits a widespread use of recombination-based biotechnologies. Genomic-scale PCR amplification using mouse and human ORFs has also been carried out towards large-scale mammalian proteomics. Herein, it must be noted that primers can be made by high-throughput synthetic methods with high fidelity, at reasonable cost, and in 96-well format amenable for robotic manipulation.

Glycomics, a term defining the sequence analysis or profiling of glycome, must be integrated with the results (database) obtained by proteomics because the glycosylation (glycoforms) greatly influences protein functions and structures. However, it should be emphasized that individual glycoforms (profiled glycan structures or carbohydrate structures) of the glycoconjugates of interest are impossible to predict based on gene expression patterns as described above because the biosynthetic process of the glycans is not template-driven and is subject to multiple sequential and competitive enzymatic pathways [3]. There is no PCR-like glycan amplification technology for glycomics. While proteomics can use satisfactory amounts of the proteins of interest by means of the above-mentioned ORF-PCR-based recombination technology, glycomics need an enrichment process of glycans from highly complicated mixtures such as serum, cells, and tissues. One of the key bottlenecks in structural and functional glycomics is on the difficulty that the purification of trace amounts of oligosaccharides requires extremely tedious and time-consuming multi-step processes. This is because crude sample mixtures prepared by enzymatic digestion from cells, organs, serum, etc., usually contain large amounts of impurities such as peptides, lipids, and salts. These technical problems in the sequenc-

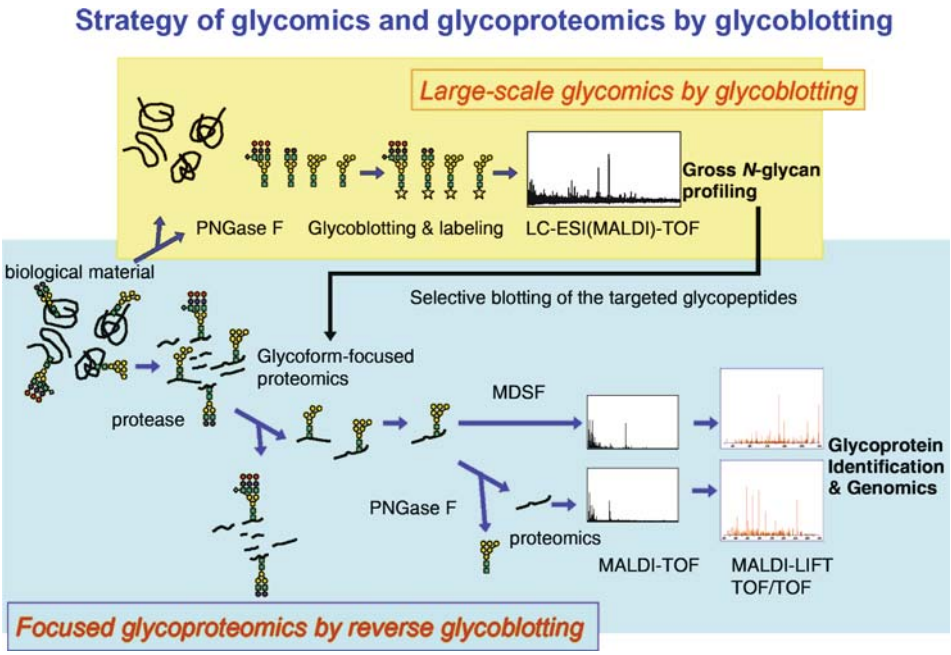


Figure 1
Strategy of glycomics and glycoproteomics by glycoblotting

ing of carbohydrates make it impossible to achieve high-throughput protein glycomics. Our recent efforts have been directed to the establishment of practical methods for glycan enrichment, namely "glycoblotting" [4,5,6,7,8]. The optimized protocol needs only 5–20 microliter of human serum for the quantitative profiling of 30–40 kinds of major glycoforms within 5–8 hours [9]. This novel technology will allow for large-scale glycomics and the clinical glycomic/glycoproteomic approach toward discovery research of new biomarkers (*Fig. 1*).

3 Emerging Glycomics Technologies

3.1 Mass Spectrometry (MS)

Mass spectrometry (MS) has become an indispensable tool for proteomics that in general deals with the large-scale determination of gene and cellular function directly at the protein level. It seems that the ability of MS to identify and to precisely quantify thousands of proteins from complex samples can be expected to impact broadly on cell biology and medicine. As in proteomics and metabolomics, MS-based analytical methods have come to the fore as a powerful tool for highly sensitive and definitive glycan analysis. Ann Dell (Imperial College, London) describes (Chap. 10.4) the state of the art in modern MS applications in the structural glycomics.

3.2 Microarrays

Microarrays of glycans (glycoarrays) displayed on solid surfaces are expected to be a convenient tool for advancing the glycomics field because these arrays will greatly facilitate high-throughput analysis and determination of the glycan-protein interactions. Ten Feizi (Imperial College, London) describes designer arrays for deciphering the glycocode that are based on an established platform of glycolipids and neoglycolipids arrayed on nitrocellulose-coated glass slides [10]. This technique permits the construction of a microarray displaying a library of 300 diverse sequence-defined glycans and the investigation of the binding of various glycan-binding proteins at femto mole levels of sensitivity.

3.3 Chemical Probes and Monitoring

Rapid and efficient methods for probing and identifying cell surface glycans in living animals are of growing importance because dynamic structural changes in glycosylation are known to accompany cell differentiation and disease progression. A metabolic glycan engineering strategy based on ManNAc analogs has been recognized as a promising approach for allowing the glycocalyx of living cells to be remodeled [11]. Kevin J. Yarema (Johns Hopkins University, Baltimore) focuses on recent advances in the metabolic installation of versatile thiol functional groups into cell surface glycans in order to modulate stem cell adhesion [12].

3.4 Bioinformatical Tools

Posttranslational glycosylations of proteins dramatically enhance their structural and functional diversity. Glycome provides vast information space at minimum genetic cost, because biosynthesis of glycans is not encoded through a template-driven system in the information flow from the genome to the glycome. With a goal of building a systems biology approach to glycan structure-function relationships, diverse datasets generated from different glycomics technologies should be integrated into versatile bioinformatics platforms. Claus-Wilhelm von der Lieth (Chap. 10.5) (German Cancer Research Center) describes the importance of relational bioinformatics databases and international plans to establish a federated network of distributed databases for glycobiology.

References

1. a) Dwek RA (1996) Chem Rev 96:683; b) Zachara NE, Hart GW (2002) Chem Rev 102:431
2. Aebersold R, Mann M (2002) Nature 422:198
3. Taylor ME, Drickamer K (2003) Introduction to Glycobiology. Oxford University Press, Oxford
4. a) Nishimura S-I, Niikura K, Kurogochi M, Matsushita T, Fumoto M, Hinou H, Kamitani R, Nakagawa H, Deguchi K, Miura N, Monde K, Kondo H (2005) Angew Chem Int Ed 44:91; b) Kurogochi M, Amano M, Fumoto M, Takimoto A, Kondo H, Nishimura S-I (2007) Angew Chem Int Ed 46:8808
5. Niikura K, Kamitani R, Kurogochi M, Uematsu R, Shinohara Y, Nakagawa H, Deguchi K, Monde K, Kondo H, and Nishimura S-I (2005) Chem Eur J 11:3825
6. Shimaoka H, Kuramoto H, Furukawa J, Miura Y, Kurogochi M, Kita Y, Hinou H, Shinohara Y, Nishimura S-I (2007) Chem Eur J 13:1664

7. Miura Y, Shinohara Y, Furukawa J, Nagahori N, Nishimura S-I (2007) Chem Eur J 13:4797
8. Kita Y, Miura Y, Furukawa J, Nakano M, Shinohara Y, Ohno M, Takimoto A, Nishimura S-I (2007) Mol Cell Proteomics 6:1437
9. Miura Y, Nishimura S-I et al. (2007) Mol Cell Proteomics (in press)
10. Feizi T, Chai WG (2004) Nat Rev Mol Cell Biol 5:582
11. Kayser H, Zeitler R, Kannicht C, Grunow D, Nuck R, Reutter W (1992) J Biol Chem 267: 16934
12. Sampathkumar S-G, Li AV, Jones MB, Sun Z, Yarema KJ (2006) Nat Chem Biol 2:149

10.2 Microarrays – A Key Technology for Glycobiology

Yan Liu, *Ten Feizi*
Glycosciences Laboratory, Faculty of Medicine, Imperial College London, Northwick Park and St Mark's Campus, Harrow, Middlesex HA1 3UJ, UK
t.feizi@imeprial.ac.uk

Abstract

Carbohydrate chains of glycoproteins, glycolipids, and proteoglycans can mediate processes of biological and medical importance through their interactions with complementary proteins. The unraveling of these interactions is a priority therefore in biomedical sciences. Carbohydrate microarray technology is a new development at the frontiers of glycomics that has revolutionized the study of carbohydrate-protein interactions and the elucidation of their specificities in endogenous biological processes, immune defense mechanisms, and microbe-host interactions. In this chapter we briefly touch upon the principles of numerous platforms since the introduction of carbohydrate microarrays in 2002, and we highlight platforms that are beyond proof-of-concept, and have provided new biological information.

Keywords

Carbohydrate-binding proteins; Glycolipid; Microarray; Neoglycolipid; Oligosaccharide

Abbreviations

AO-NGLs	NGLs prepared *via* oxime ligation with lipid tag AOPE
AOPE	*N*-aminooxyacetyl-1,2-dihexadecyl-*sn*-glycero-3-phosphoethanolamine
CFG	Consortium for Functional Glycomics
DH-NGLs	NGLs prepared *via* reductive amination with lipid tag DHPE
DHPE	1,2-dihexadecyl-*sn*-glycero-3-phosphoethanolamine
HA	hemagglutinin
HIV	human immunodeficiency virus
MIC-1	micronemal protein-1

In: *Glycoscience*. Fraser-Reid B, Tatsuta K, Thiem J (eds)
Chapter-DOI 10-1007/978-3-540-30429-6_54:

NGL neoglycolipid
SPR surface plasmon resonance

1 Overview of Carbohydrate Microarray Technologies

The diverse oligosaccharides that 'decorate' glycoproteins, glycolipids, proteoglycans, and polysaccharides are potentially a vast source of information, and could harbor a 'glycocode' that is waiting to be deciphered in various contexts of biological and medical importance [16,45]. It is desirable therefore to develop a knowledge base of biological systems that operate through oligosaccharide recognition. However, detailed analyses of carbohydrate-protein interactions has been a challenging area due to the structural complexities of oligosaccharides, the limited amounts that can be isolated, the difficulties in their characterization, and in general, the low affinities of their interactions with cognate proteins. Oligosaccharides cannot be cloned and the diverse repertoire of oligosaccharides is difficult to access by chemical synthesis. For all these reasons, the advent of carbohydrate microarrays is a welcome development for elucidating carbohydrate-binding specificities in a wide variety of contexts: cell-cell, cell-protein, and host pathogen interactions. The main advantage of microarray analysis is that a broad range of sugars (tens or hundreds, eventually thousands) can be immobilized and simultaneously interrogated. The multivalent display of arrayed ligands can usually mimic cell surface display and is ideal for detecting the generally very low affinities of interactions that involve carbohydrates. The miniaturization in microarrays is particularly well suited for investigations in glycomics, as the amounts of oligosaccharide probes arrayed are relatively small and thousands of binding events can be assessed in parallel on a single microarray slide. A number of recent reviews have covered in some depth different methodologies for constructing carbohydrate microarrays [11,17,37,40]. With the exception of polysaccharides [41,46,48] and natural glycolipids [16], chemical derivatizations are required for immobilizing carbohydrates for microarrays, as hydrophilicities of oligosaccharides preclude their direct noncovalent immobilization on solid matrices. One principle is to conjugate natural or chemically synthesized oligosaccharides to lipid by reductive amination to generate neoglycolipid (NGL) probes with amphipathic properties for arraying [20]. The use of reductive amination has also been described for preparing fluorescent oligosaccharide derivatives which contain a primary amine for array generation [49]. Most other mono- or oligosaccharide probes generated for printing have been de-novo synthesized chemically or chemoenzymatically, requiring substantial chemical expertise for access to defined structures that incorporate specific functional groups, e. g. thiol [1,7,36], maleimide [34], amine [4,12,14], azide [15], and cylcopentadiene [22] functionalities, or proteins as tags [1,31] for covalent attachment to solid matrices. Oligosaccharides with lipid [8] or fluorous tags [30] for noncovalent immobilization have also been described. These methods are promising due to compatibilities with advanced chemical or enzymatic synthesis of oligosaccharides [24,25,35]. However, relatively few structures can be obtained by current synthetic approaches [39]. Moreover, multi-step manipulations limit applications to the small quantities of oligosaccharides that can be isolated from natural sources. As natural oligosaccharides are crucial for discoveries of hitherto unknown oligosaccharide ligands [16,17], carbohydrate microarrays should ideally encompass oligosaccharides from both synthetic and natural sources. Approaches

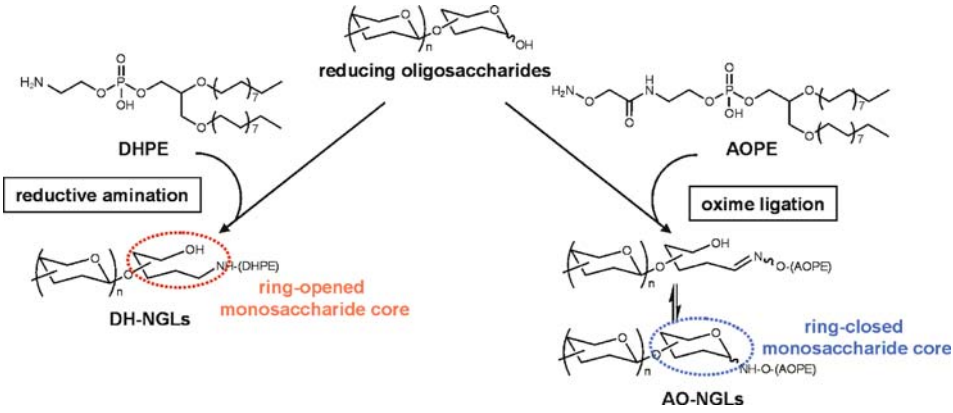
have been described for printing and covalent attachment of unmodified mono and short oligosaccharides onto aminooxy- or hydrazide-modified glass slides [26,53] and heparin-derived sulfated oligosaccharides onto hydrazide-modified gold surfaces [52]. Disadvantages of direct covalent surface immobilization methods are the high concentrations of the saccharides required and the potential variation of the immobilization efficiencies depending on the nature of saccharides. In most cases there is no way of knowing how much sample is attached at each spot. Fluorescence-based measurement is, to date, the prevalent detection principle for detecting binding to carbohydrate microarrays. An alternative approach described recently is the use of surface plasmon resonance (SPR) imaging for real-time, label-free analysis of protein interactions with microarrays of biotinylated glycans (printed onto neutravidin-coated gold chips) [23].

Many of the aforementioned carbohydrate microarray strategies have been used for observations on proteins of known specificities and their interactions with sugars. We highlight below several oligosaccharide microarray systems that are beyond proof-of-concept and have provided new biological insights. We dwell, in some detail, on the NGL-based oligosaccharide microarray platform with which we have first hand experience.

2 Sequence Defined Carbohydrate Microarrays

Encouraged by the observations of Wang and colleagues [46] that polysaccharides and glycoproteins can be satisfactorily immobilized on nitrocellulose by noncovalent interaction, Feizi and colleagues adapted their long established [44] and validated NGL technology [19] to generate microarrays of lipid-linked oligosaccharide probes using nitrocellulose as the matrix [16,20,21,38]. The NGL technology involves conjugating oligosaccharides by microscale reductive amination to an aminolipid, 1,2-dihexadecyl-*sn*-glycero-3-phosphoethanolamine (DHPE), as depicted in ❯ *Scheme 1*. This allows minute amounts (nmol scale) of oligosaccharides released from *O*- and *N*-glycosylated proteins, glycosaminoglycans, and polysaccharides to be converted into lipid-linked probes (DH-NGLs). The amphipathic property conferred by introduction of the lipid tag enables immobilization in clustered display on solid matrices such as silica-gel plates, plastic microtiter wells, and nitrocellulose. This mode of display is optimal for detecting binding signals with the majority of carbohydrate-binding antibodies and receptors, as the affinities of binding are low. In addition, the NGL technology provides a way to resolve, by high performance (HP) TLC, mixtures of oligosaccharide probes, and to perform carbohydrate-binding experiments on TLC plates in conjunction with oligosaccharide sequence determination by mass spectrometry in situ [10]. The NGL principle has led to discoveries of unsuspected oligosaccharide sequences on glycoproteins [50], as well as new oligosaccharide ligands for carbohydrate-binding proteins [51] (other examples are reviewed in [18]).

Other than ring-opening of the monosaccharide residues at the reducing ends, oligosaccharides remain intact in DH-NGLs. Those derived from tri- or larger oligosaccharides have performed well for the majority of carbohydrate-recognition systems which have the peripheral or backbone regions of oligosaccharides as recognition motifs [16]. However, ring-opening of the monosaccharides at the reducing-end may affect biological activities of short oligosaccharides. For instance, the DH-NGL of Lewis^x (Le^x) trisaccharide is not bound by anti-Le^x



Scheme 1
Principles of the preparation of NGL probes from reducing oligosaccharides by reductive amination and oxime ligation

antibodies [43], and that of sialyl-Le^x tetrasaccharide is not bound by the selectins [27]. To overcome this limitation and enhance the applicabilities of NGL probes derived from short oligosaccharides which are the most accessible via chemical synthesis, a novel type of NGLs with ring-closed monosaccharide cores has been introduced [28,29]. These are prepared from a variety of reducing sugars by conjugation with an aminooxy-functionalized DHPE (AOPE) via microscale oxime ligation (without reduction). The key features of these NGLs (AO-NGLs) have been demonstrated by microarray analyses: there is efficient presentation of short oligosaccharides, such as Le^x trisaccharide to anti-Le^x-antibodies and sialyllactose analogs to receptors of the immune system known as siglecs; the core-monosaccharide of a fucosylated *N*-glycan is also preserved as a ligand for the plant lectin from *Pisum sativum* (pea lectin) which recognizes the unmodified core.

The NGL-based microarray platform schematically presented in *Fig. 1*, currently has over 300 robotically arrayed, sequence-defined saccharide probes, which comprise NGLs from both natural and synthetic oligosaccharides, as well as natural and synthetic glycolipids, and are expanding in number. Included are: *N*-glycans (neutral and acidic, high-mannose and complex types), *O*-glycans and blood group-related sequences (A, B, H, Le^a, Le^b, Le^x, and Le^y) on linear or branched backbones, and their sialylated and/or sulfated analogs, gangliosides, glycosaminoglycans, homo-oligomers of sialic acid and oligosaccharide fragments of other polysaccharides, ranging in size from 2 to 20 monosaccharide units. This platform has all the attributes of the NGL technology with provision for generating 'designer' microarrays from targeted tissues and macromolecules [20,32,33]. It is advantageous that there is the provision for 'deconvolution', by resolving arrays of NGL mixtures by HPTLC before ligand-binding experiments in conjunction with mass spectrometry in situ for sequence assignment. This platform shows considerable promise as a novel approach to surveying entire glycomes and proteomes for the molecular definition of recognition systems.

The first generation NGL-based microarray system using nitrocellulose membranes provided insights to the specificity of an antibody (CS-56) to chondroitin sulfates A and C [20], and

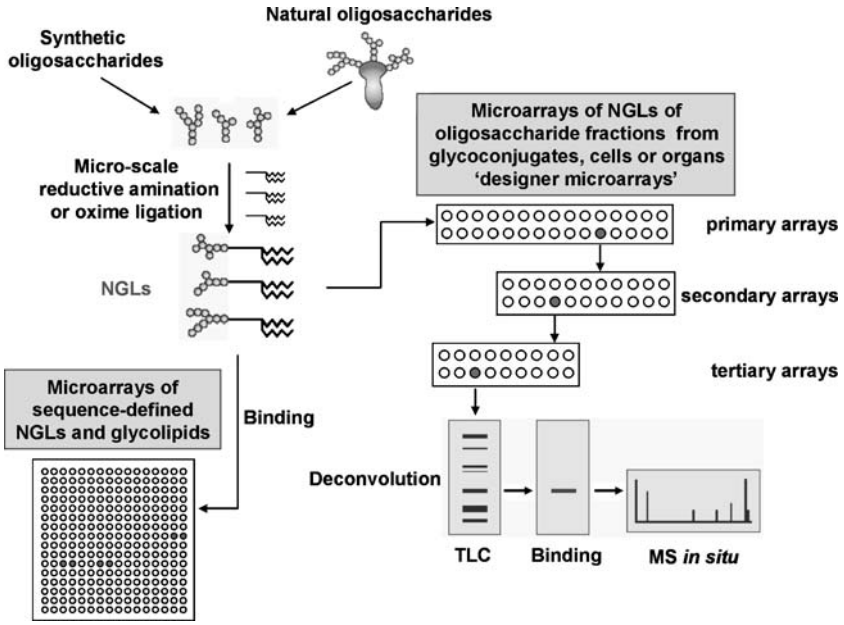


Figure 1
The NGL-based oligosaccharide microarray platform

enabled the identification of ligands for receptors of the immune system known as SIGN-R1, SIGN-R3, and langerin [21]. Assignment could also be made of a carbohydrate-binding function to domain 5 of the cation-independent mannose 6-phosphate receptor [38]. In addition, the applicability of NGL-based microarray to mixtures of oligosaccharides obtained from biological sources such as cells and whole organs has been shown [20,32].

With the second generation advanced NGL-based microarrays, using nitrocellulose-coated glass slides, the influence of sulfation on carbohydrate recognition by the siglecs has been shown [9]. Recent carbohydrate ligand assignments made by this microarray system have revealed, at one extreme a broad range of sialyl motifs, among *N*- and *O*-glycans and glycolipids to be recognized by the micronemal protein-1 (MIC-1) of *Toxoplasma gondii* (*Fig. 2* and *Fig. 3*) [5]. This accounts for the unique adaptation of this protozoan parasite to infect a wide range of hosts, including virtually all warm-blooded animals and up to 50% of the world population. At the other extreme, the ligands for Decin-1, a receptor of the innate immune system, were found to be highly restricted, and could be identified only by generating designer microarrays from the fungal and bacterial polysaccharide glycomes that the receptor targets (*Fig. 4*) [33].

Seeberger and colleagues have applied microarrays containing synthetic oligosaccharides to address several biological problems. Mainly two surface chemistries have been adopted for the preparation of the microarrays. The first involves maleimide-functionalized surfaces to form a covalent bond between the slides and the thiol containing synthetic carbohydrates (*Fig. 5a*). By this means, high-mannose microarrays were constructed using a panel of

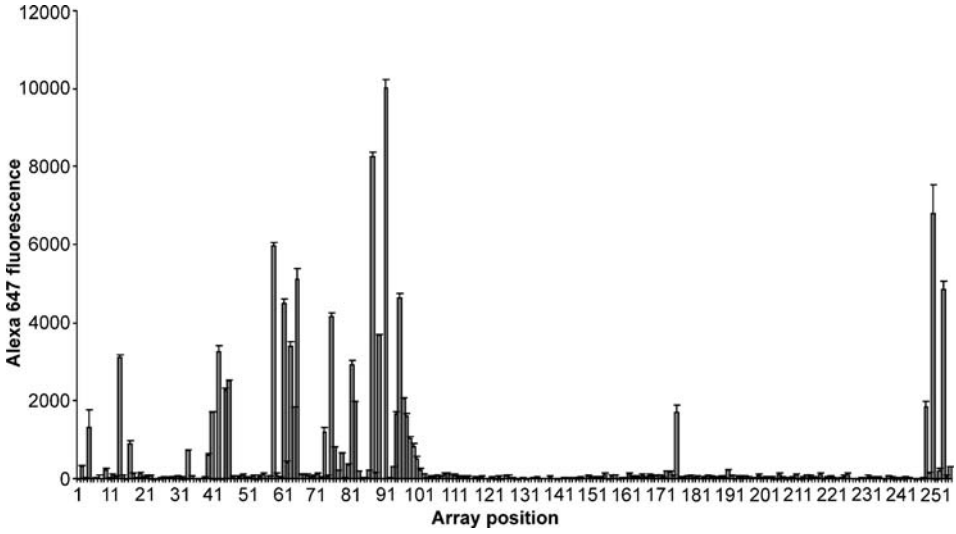


Figure 2

Microarray analyses of the binding of the *T.gondii* protein MIC-1. Two hundred and eighteen sequence-defined lipid-linked oligosaccharide probes were printed on nitrocellulose-coated slides [5]. A total of 256 oligosaccharide positions are shown, as some probes were arrayed at more than one position. Binding is plotted as relative fluorescence, means, with error bars, of duplicate spots minus backgrounds, at 7 fmol/spot. Binding was observed only to sialylated probes, which are described in *Fig. 3* below

synthetic mannose-containing oligosaccharides for identification of human immunodeficiency virus (HIV) vaccine candidate antigens [1]. The binding profiles were described for several proteins that bind to gp120 of HIV. These include the receptor of immune system known as DC-SIGN, the antibody 2G-12, and the bacterial proteins Cyanovirin-N and Scytovirin. Microarrays of natural or chemically modified glycoproteins were also included in this study and provided additional information regarding the importance of the underlying peptide context in which glycans are presented to their binding partners. A microwave-assisted two-step procedure was recently reported to introduce a thiol linker at the reducing end of carbohydrates for the construction of arrays on maleimide-functionalized surfaces [7]. For the second surface chemistry, amine linkers were employed to covalently attach the sugars onto amine-reactive slides (*Fig. 5b*). By this approach, monosaccharide microarrays were assembled and applied for adhesion studies on pathogenic bacteria [14], and microarrays displaying synthetic heparin oligosaccharides were used to study the binding profiles of heparin-binding proteins [12,47]. Microarrays of aminoglycosides were similarly constructed by random covalent immobilization onto amine-reactive glass for the evaluation of aminoglycoside antibiotics [13].

The Consortium for Functional Glycomics (CFG, http://www.functionalglycomics.org) has developed two generations of oligosaccharide microarrays [2]. The first is built on streptavidin-coated microtiter plates to which a library of biotinylated synthetic or natural glycans is attached [3]. This has now been replaced by the second generation, which is built using microarray printing technology to spot amine-terminating glycans on commercially available NHS-activated glass slides [4]. The key features in the construction of the printed arrays are

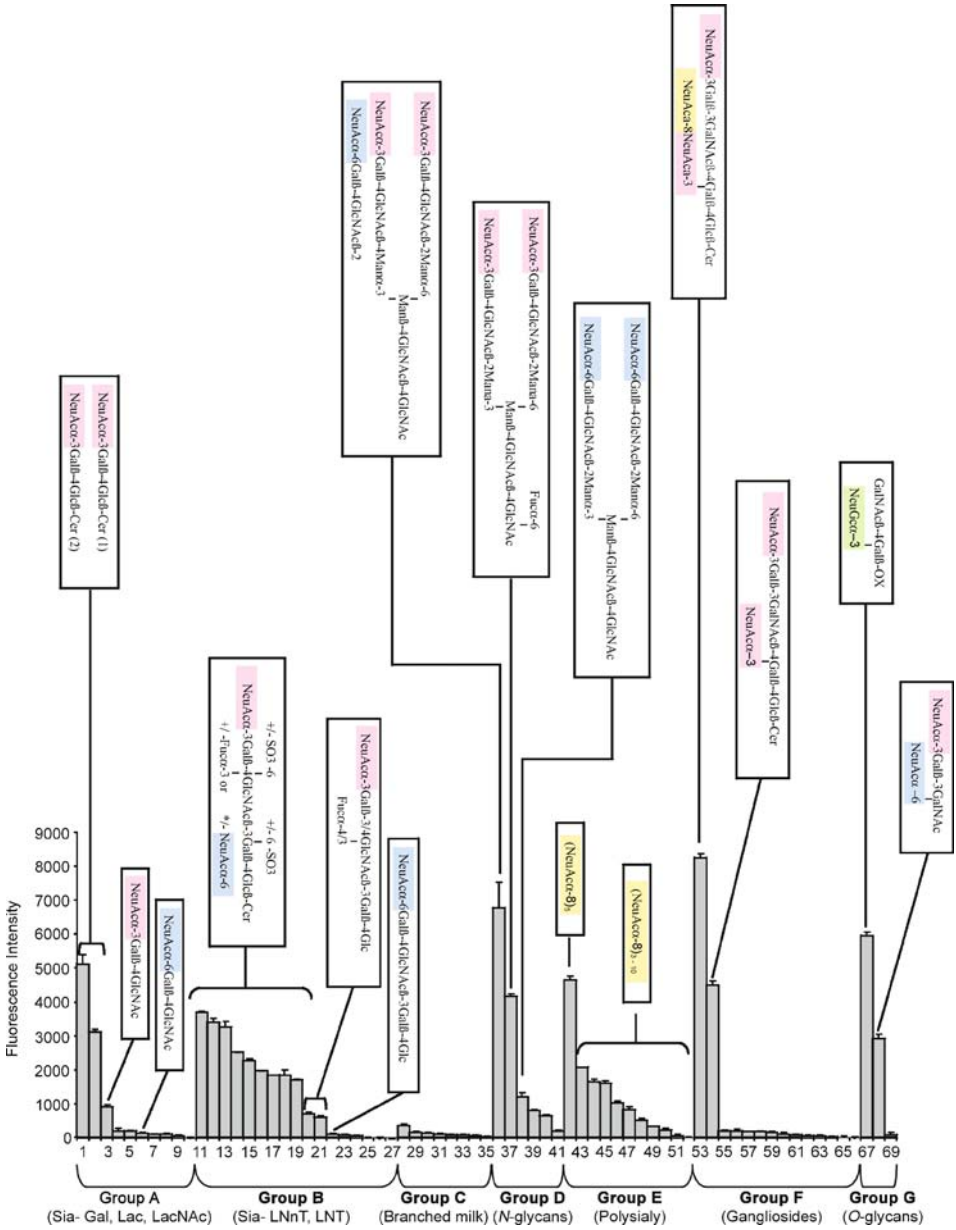


Figure 3

Numerical scores are shown of the binding signals with the *T.gondii* protein MIC-1: means of duplicate spots at 7 fmol/spot (with error bars) for the 59 sialyl oligosaccharide probes examined. These are classified into seven groups according to back-bone sequence. Sixty-nine positions are shown as ten of the probes were printed at two positions. Twenty-two selected probes are annotated, with designations of Neu5Acα2–3Gal linkage as pink; Neu5Acα2–6Gal, blue and Neu5Acα2–8 linkage yellow. The scores for the non-sialylated probes in the microarray are not shown; none gave binding signals above 150. Reprinted from ref. [5] with permission from the EMBO Journal

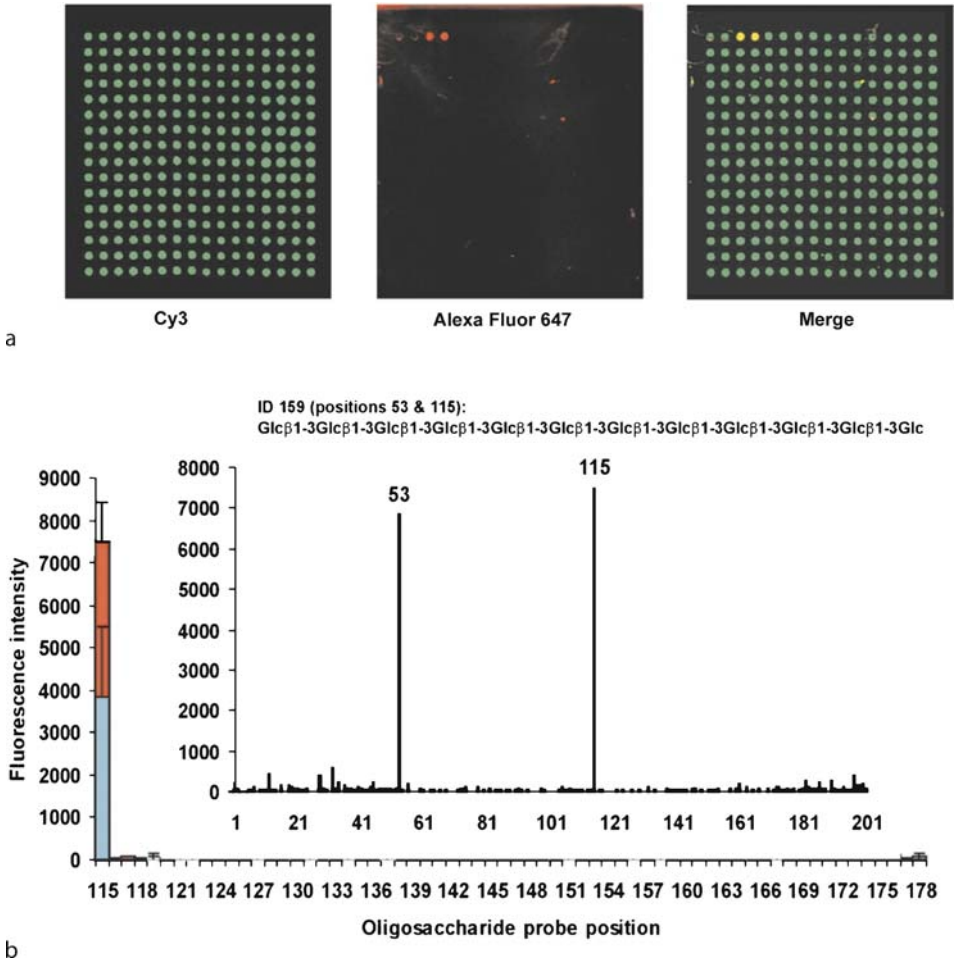


Figure 4

Microarrays of lipid-linked oligosaccharide probes showing Dectin-1 binding exclusively to the 'designer' NGL probe derived from the bacterial polysaccharide curdlan. (a) Oligosaccharide probes were printed in duplicate on nitrocellulose-coated glass slides at 2 and 7 fmol/spot with Cy3 dye included as a marker (green emission). Binding was detected with Alexa Fluor 647-labeled streptavidin (red emission). The image is that of a subset of 256 arrayed spots of 64 oligosaccharide probes. (b) Numerical scores are shown of the binding signals (means of duplicate spots) at 2 (blue) and 7 fmol/spot (red) with error bars for the 64 probes at positions 115–178. The inset shows numerical scores of the binding signals at 7 fmol/spot for all of the oligosaccharide probes arrayed. The signals at positions 53 and 115 correspond to those elicited by a probe derived from curdlan (β1,3-linked glucose 12mer). Reprinted from ref. [33] with permission from Journal of Biological Chemistry

introduction of short neutral azido-linkers or *N*-Fmoc protected amino acids chemically to the monosaccharides, chemical or enzymatic elongation of these monosaccharides to obtain various oligosaccharide structures and conversion of the azido group of the linkers and the amino acid-moiety into amine functionality for covalent immobilization on NHS activated slides

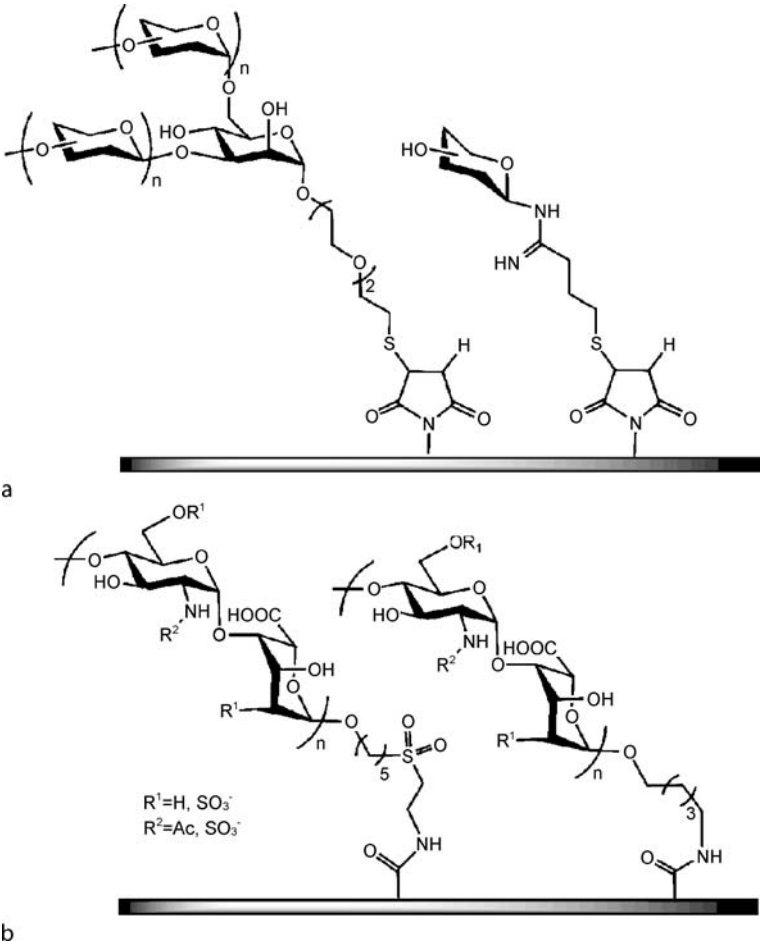


Figure 5
Covalent immobilization of chemically synthesized carbohydrates onto derivatized surfaces by Seeberger group. (a) Attachment of thiol containing carbohydrates to maleimide-functionalized glass slides; (b) attachment of amine-terminated heparin oligosaccharides onto amine-reactive glass slides

(*Fig. 6*). A growing library of more than 300 structurally defined glycans, predominantly synthetic, ranging in size from monosaccharides to undecasaccharides is included in the current microarrays. A recent development in this microarray platform is the preparation of glycan probes by one-step derivatization of free reducing glycans using a bi-functional spacer which contains both an aminooxy functionality and an amino group [6]. The binding specificities of a wide variety of carbohydrate-binding proteins, including C type lectins, siglecs, galectins, anti-carbohydrate antibodies, and lectins from plants, have been examined [4]. This system has also been applied to detect anti-carbohydrate antibody specificities in human serum [4]. In addition, glycan array analyses have been carried out on hemagglutinin (HA) from the

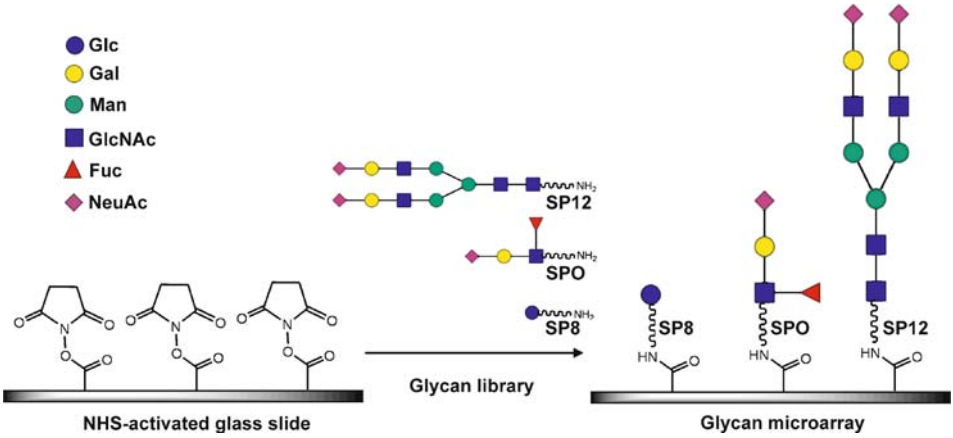


Figure 6
Covalent immobilization of glycans printed onto NHS-activated glass slides. Linkers SP0 = -$(CH_2)_2NH_2$, SP8 = -$(CH_2)_3NH_2$, and SP12 = Asparagine. Symbols for presentation of monosaccharides in glycans are as designated by CFG

1918 pandemic human influenza virus and the HA mutants with changed carbohydrate-binding specificities [42]. Microarray analysis with whole influenza virus was also shown to be possible [4].

3 Comments

Carbohydrate microarrays are coming of age. Their advent has transformed the studies of carbohydrate-protein interactions, leading to high throughput analyses of biomedically important systems that operate through carbohydrate recognition. It can be anticipated that many new and hitherto unsuspected carbohydrate-binding proteins and their ligands will be identified by these approaches.

4 Acknowledgments

The authors acknowledge support from the UK Medical Research Council and UK Research Councils Basic Technology Grant (GR/S79268, 'Glycoarrays').

References

1. Adams EW, Ratner DM, Bokesch HR, McMahon JB, O'Keefe BR, Seeberger PH (2004) Chem Biol 11:875
2. Alvarez RA, Blixt O (2006) Methods Enzymol 415:292
3. Blixt O, Collins BE, van dN, I, Crocker PR, Paulson JC (2003) J Biol Chem 278:31007
4. Blixt O, Head S, Mondala T, Scanlan C, Huflejt ME, Alvarez R, Bryan MC, Fazio F, Calarese D, Stevens J, Razi N, Stevens DJ, Skehel JJ, Van D,

I, Burton DR, Wilson IA, Cummings R, Bovin N, Wong CH, Paulson JC (2004) Proc Natl Acad Sci U S A 101:17033
5. Blumenschein TMA, Friedrich N, Childs RA, Saouros S, Carpenter EP, Campanero-Rhodes MA, Simpson P, Chai W, Koutroukides T, Blackman MJ, Feizi T, Soldati-Favre DaMS (2007) EMBO J 26:2808
6. Bohorov O, Andersson-Sand H, Hoffmann J, Blixt O (2006) Glycobiology 16:21C
7. Brun MA, Disney MD, Seeberger PH (2006) Chembiochem 7:421
8. Bryan MC, Plettenburg O, Sears P, Rabuka D, Wacowich-Sgarbi S, Wong CH (2002) Chem Biol 9:713
9. Campanero-Rhodes MA, Childs RA, Kiso M, Komba S, Le Narvor C, Warren J, Otto D, Crocker PR, Feizi T (2006) Biochem Biophys Res Comm 344:1141
10. Chai W, Cashmore GC, Carruthers RA, Stoll MS, Lawson AM (1991) Biol Mass Spectrom 20:169
11. Culf AS, Cuperlovic-Culf M, Ouellette RJ (2006) Omics-A J Integ Biol 10:289
12. de Paz JL, Noti C, Seeberger PH (2006) J Am Chem Soc 128:2766
13. Disney MD, Magnet S, Blanchard JS, Seeberger PH (2004) Angew Chem Int Ed 43:1591
14. Disney MD, Seeberger PH (2004) Chem Biol 11:1701
15. Fazio F, Bryan MC, Blixt O, Paulson JC, Wong CH (2002) J Am Chem Soc 124:14397
16. Feizi T, Chai W (2004) Nat Rev Mol Cell Biol 5:582
17. Feizi T, Fazio F, Chai W, Wong CH (2003) Curr Opin Struct Biol 13:637
18. Feizi T, Lawson AM, Chai W (2003) Neoglycolipids: identification of functional carbohydrate epitopes. In: Wong C-H (ed) Carbohydrate-based Drug Discovery: from the Laboratory to the Clinic. Wiley-VCH, Weinheim, chap 27, p 747
19. Feizi T, Stoll MS, Yuen C-T, Chai W, Lawson AM (1994) Methods Enzymol 230:484
20. Fukui S, Feizi T, Galustian C, Lawson AM, Chai W (2002) Nat Biotechnol 20:1011
21. Galustian C, Park CG, Chai W, Kiso M, Bruening SA, Kang YS, Steinman RM, Feizi T (2004) Int Immunol 16:853
22. Houseman BT, Mrksich M (2002) Chem Biol 9:443
23. Karamanska R, Clarke J, Blixt O, MacRae JI, Zhang JQ, Crocker PR, Laurent N, Wright A, Flitsch SL, Russell DA, Field RA (2007) Glycoconjugate J [Epub ahead of print]:
24. Koeller KM, Wong CH (2000) Glycobiology 10:1157
25. Lee JC, Greenberg WA, Wong CH (2007) Nat Protocols 1:3143
26. Lee M, Shin I (2005) Organic Letters 7:4269
27. Leteux C, Stoll MS, Childs RA, Chai W, Vorozhaikina M, Feizi T (1999) J Immunol Methods 227:109
28. Liu Y, Chai W, Childs RA, Feizi T (2006) Methods Enzymol 415C:326
29. Liu Y, Feizi T, Campanero-Rhodes MA, Childs RA, Zhang Y, Mulloy B, Evans PG, Osborn HM, Otto D, Crocker PR, Chai W (2007) Chem Biol 14:847
30. Mamidyala SK, Ko KS, Jaipuri FA, Park G, Pohl NL (2006) J Fluor Chem 127:571
31. Manimala JC, Roach TA, Li ZT, Gildersleeve JC (2006) Angew Chem Int Ed 45:3607
32. Osanai T, Feizi T, Chai W, Lawson AM, Gustavsson ML, Sudo K, Araki M, Araki K, Yuen CT (1996) Biochem Biophys Res Commun 218: 610
33. Palma AS, Feizi T, Zhang Y, Stoll MS, Lawson AM, Diaz-Rodríguez E, Campanero-Rhodes AS, Costa J, Brown GD, Chai W (2006) J Biol Chem 281:5771
34. Park S, Shin I (2002) Angew Chem Int Ed 41:3180
35. Plante OJ, Palmacci ER, Seeberger PH (2001) Science 291:1523
36. Ratner DM, Adams EW, Disney MD, Seeberger PH (2004) Chembiochem 5:1375
37. Ratner DM, Adams EW, Disney MD, Seeberger PH (2004) Chembiochem 5:1375
38. Reddy ST, Chai W, Childs RA, Page JD, Feizi T, Dahms NM (2004) J Biol Chem 279:38658
39. Seeberger PH, Werz DB (2005) Nat Rev Drug Discov 4:751
40. Shin I, Park S, Lee MR (2005) Chemistry 11:2894
41. Shipp EL, Hsieh-Wilson LC (2007) Chem Biol 14:195
42. Stevens J, Blixt O, Glaser L, Taubenberger JK, Palese P, Paulson JC, Wilson IA (2006) J Mol Biol 355:1143
43. Streit A, Yuen C-T, Loveless RW, Lawson AM, Finne J, Schmitz B, Feizi T, Stern CD (1996) J Neurochem 66:834
44. Tang PW, Gooi HC, Hardy M, Lee YC, Feizi T (1985) Biochem Biophys Res Commun 132: 474
45. Taylor ME, Drickamer K (2006) Introduction to Glycobiology, 2nd edn. Oxford University Press, New York

46. Wang D, Liu S, Trummer BJ, Deng C, Wang A (2002) Nat Biotechnol 20:275
47. Werz DB, Seeberger PH (2005) Chem-A Eur J 11:3194
48. Willats WG, Rasmussen SE, Kristensen T, Mikkelsen JD, Knox JP (2002) Proteomics 2:1666
49. Xia BY, Kawar ZS, Ju TZ, Alvarez RA, Sachdev GP, Cummings RD (2005) Nat Methods 2: 845
50. Yuen C-T, Chai W, Loveless RW, Lawson AM, Margolis RU, Feizi T (1997) J Biol Chem 272:8924
51. Yuen C-T, Lawson AM, Chai W, Larkin M, Stoll MS, Stuart AC, Sullivan FX, Ahern TJ, Feizi T (1992) Biochemistry 31:9126
52. Zhi ZL, Powell AK, Turnbull JE (2006) Anal Chem 78:4786
53. Zhou XC, Zhou JH (2006) Biosens Bioelectron 21:1451

10.3 Non-Natural Sugar Analogues: Chemical Probes for Metabolic Oligosaccharide Engineering

Udayanath Aich, *Kevin J. Yarema*
Department of Biomedical Engineering, The Johns Hopkins University, Baltimore, MD 21218, USA
kyarema1@jhu.edu

In: *Glycoscience*. Fraser-Reid B, Tatsuta K, Thiem J (eds)
Chapter-DOI 10-1007/978-3-540-30429-6_55:

Abstract

Metabolic oligosaccharide engineering (MOE) is a rapidly growing technology emerging from the field of chemical biology that allows novel chemical functionalities to be biosynthetically installed into the carbohydrates of living cells and animals. Since pioneering efforts to modulate sialic acid display through the use of non-natural *N*-acetyl-D-mannosamine (ManNAc) analogues were reported 15 years ago, monosaccharide probes have been developed to manipulate *N*-acetyl-D-galactosamine (GalNAc), *N*-acetyl-D-glucosamine (GlcNAc), and fucose-containing glycans. The 'first generation' of analogues, comprised of a series of ManNAc derivatives with elongated *N*-acyl chains, demonstrated pathway permissivity and the ability of this methodology to impinge on biological processes ranging from pathogen binding to gene expression and cell adhesion. Later analogues have incorporated chemical function groups including ketones, azides, thiols, and alkyne not normally found in carbohydrates. These groups serve as 'tags' for the subsequent use of chemoselective ligation reactions to further elaborate the chemical properties of the cell surface and thereby greatly expand the potential of MOE technology to offer control over biological processes.

Keywords

Carbohydrates; Chemoselective ligation; Neoglycoconjugates

Abbreviations

Ac acetyl
AFM atomic force microscopy
BSA bovine serum albumin
CDG congenital disorder of glycosylation
CMP cytidine monophosphate,
DLT dose limiting toxicity
FITC fluorescein isothiocyante
GAG glycosylaminoglycan
GNE UDP-GlcNAc 2-epimerase/ManNAc 6-kinase
GDP guanidine diphosphate
hEBD human embryoid body-derived
MB (+)-biotinyl-3-maleimidopropionamidyl-3,6-dioxaoctanediamine
MOE metabolic oligosaccharide engineering
NMR nuclear magnetic resonance
OGT *O*-glycosyltransferase
PCR polymerase chain reaction
PSA polysialic acid
RNA ribonucleic acid
Ser serine
SCFA short chain fatty acid
SPPS solid phase synthesis
TCEP tris-(2-carboxyethyl) phosphine hydrochloride
Thr threonine
UDP uridine diphosphate

1 Introduction

Carbohydrates are the most abundant organic molecules found in Nature. By mass, polysaccharides such as cellulose in plant walls and chitosan in the shells of marine organisms vastly outweigh other biological polymers. These polysaccharides are typically assembled from homopolymers of monosaccharides or repeating disaccharides units that can be several hundred residues in length. They play major structural roles in fungi, crustaceans, and plants but are also often subject to post-synthetic modifications, especially in animals, that greatly increase their chemical diversity and biological activities. In higher organisms, carbohydrates are quantitatively less abundant and also of smaller size, often occurring as oligosaccharide structures of only 20 or fewer—sometimes only one or two—residues. Despite their relatively small size, mammalian oligosaccharides have critical biological functions that derive from two fundamental differences compared to the other three major classes of biomacromolecules (nucleic acids, proteins, and lipids); carbohydrates can be highly branched molecules and their monomeric units may be connected to one another by many different linkage types. For example, four different amino acids can be combined to form 24 different tetrapeptides while four different monosaccharides can be combined to give 35,560 distinct tetrasaccharides [1,2]. This complexity allows even relatively small oligosaccharides to exist in a profuse number of structural variations with immense information carrying capacity so that, even when they are present in vanishingly small quantities, they have a profound impact on modulating the structure and function of the proteins and lipids.

Although they can be present without being attached to other molecules, most mammalian carbohydrates are attached to proteins or lipids and are generically referred to as 'glycoconjugates' or 'glycans'. According to the survey performed by Spiro, a wealth of sugar–protein linkages are known, involving 13 different monosaccharides and eight amino acids leading to a total of at least 41 bonds [3]. These diverse glycans, when present on the cell surface, largely determine the interactions of the host cell with other cells or with its environment. These functional roles, initiated by protein–carbohydrate or carbohydrate–carbohydrate interactions, include numerous biological recognition processes including, viral and bacterial infection, cancer metastasis, inflammatory response, innate and adaptive immunity, and many other receptor-mediated signaling processes [4,5,6,7]. In addition to serving as a surface 'sugar code' [8] that identifies a cell to the outside world, oligosaccharides can also modify the intrinsic properties of the specific proteins and lipids to which they are conjugated. For proteins, these properties include catalytic activity, resistance to proteolytic attack, solubility and conformational aspects including folding [9,10,11]. For lipids, an outstanding example of the role of sugar is the dramatic differences in biological activities exhibited by small differences, the acetylation status, of the sialic acid residue of ganglioside GD3 in sensitizing or providing resistance to apoptosis [12]. Of note, although glycosylation has long been regarded as an extracellular phenomenon, carbohydrate-specific biological activities—exemplified by the nuclear and cytosolic *O*-GlcNAc protein modification [13,14,15]—are gaining in importance. On the basis of the ever-growing list of biological processes influenced by glycosylation, there is intense interest in modulating glycan display for purposes ranging from study of function, cell engineering, to treatment of disease. Methods akin to the techniques for nucleic acid and proteins have been hindered by the fact that glycans are not primary gene products and there is not a one-to-one correspondence between a template and the ultimate structure the way

that a nucleic acid sequence specifies the primary amino acid structure of a protein. Moreover, because glycans are not template-based, methods similar to PCR for amplification of minor components do not exist. As a consequence, structure-function correlation is a difficult problem because glycan chains of glycoproteins are complex, flexible, and microheterogenous in nature making simple structural characterization challenging. However, with the advent of sophisticated analytical, spectral, and biological techniques, the 'glycomics' analyses of cellular glycans is becoming possible [16]. Even when glycoconjugate structures are identified, however, in situ regulation of glycan production remains difficult with few methods available to precisely up- or down-regulate the biosynthesis of a targeted oligosaccharide structure. In some cases, elegant chemical or enzymatic methods are available for the synthesis of glycoconjugates of biological interest that allow glycans to be produced outside a cell. Rarely, however, can exogenous glycans be added to a cell to reproduce the effects seen upon endogenous production (an important exception is glycosphingolipids that can be synthesized and added exogenously).

Considering the immense challenges confronting the glycobiologist seeking to manipulate glycans in living cells, a technology developed over the past decade by the chemical biology community now generally referred to as 'metabolic oligosaccharide engineering' (MOE) offers a major advance. MOE technology—wherein non-natural, exogenously supplied metabolic precursors intercept biosynthetic pathways and are incorporated into cellular glycans—was largely pioneered by Reutter and colleagues for sialic acid [17,18] and is now proving to be remarkably versatile, having now been demonstrated for several different monosaccharides. ❯ *Sect. 2* of this report will provide a general overview of MOE technology giving information on the basic technology, limitations and opportunities, and scope in living systems. Then, ❯ *Sect. 3* will give a detailed analysis of the specific glycosylation pathways that have been targeted successfully and will describe the monosaccharide analogues that serve as chemical probes for specific pathways. After that, ❯ *Sect. 4* will describe how, after the glycosylation machinery of a cell has been exploited for metabolic incorporation of the analogues into mature glycans, further chemical manipulation of the cell surface is possible via chemoselective ligation reactions. Overall, the focus of this chapter is on the use of 'chemical biology' methods to manipulate glycosylation with a detailed discussion given to the chemical intermediates used to intercept metabolic pathways. In addition, examples of changes to 'downstream' biological processes altered by MOE—either directly or after further chemical manipulation of cell surface displayed sugars—are provided throughout, along with relevant references for the interested reader who seeks in depth information.

2 Basics of Metabolic Oligosaccharide Engineering

2.1 MOE—the Incorporation of Non-Natural Monosaccharides into Glycans

The basic technology behind MOE involves the uptake of the building blocks of glycan biosynthesis—i. e., monosaccharides—followed by interception of the targeted biosynthetic pathway and ultimate incorporation into cellular glycans in the place of endogenously supplied metabolic substrates. MOE has proven to be remarkably broad-based, with metabolic incor-

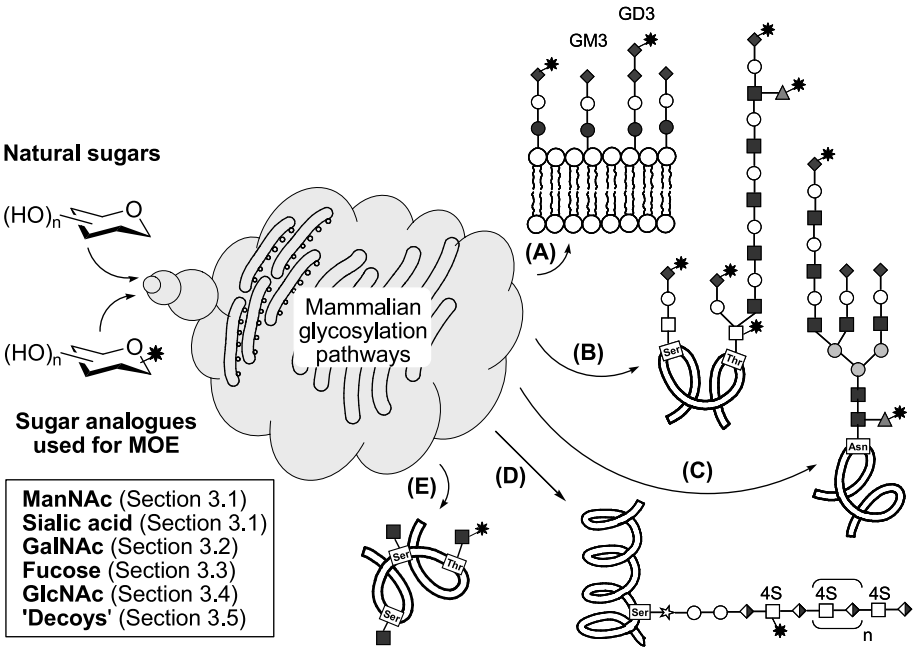


Figure 1

MOE in mammalian cells. Non-natural analogues of monosaccharides intercept glycosylation pathways in place of the natural sugars and are biosynthetically incorporated into various classes of carbohydrates including (A) glycosphingolipids, (B) *O*-linked glycoproteins, (C) *N*-linked glycoproteins, (D) glycosoaminoglycans, and (E) *O*-GlcNAc-modified nuclear and cytosolic proteins. Non-natural residues incorporated into these glycans are indicated with a *starburst* and the *symbols* used to represent sugars are given in *Fig. 2*

poration of non-natural sugar residues already demonstrated for several of the major classes of mammalian carbohydrates (*Fig. 1*). These include *N*-linked glycoproteins [19,20], mucins (*O*-linked glycoproteins) [21,22,23], glycosphingolipids [24,25], glycosoaminoglycans (GAGs) [26], and *O*-GlcNAc-modified nuclear and cytosolic proteins [27]. In addition to the well-studied mammalian systems, pilot studies have shown incorporation of carbohydrate analogues into bacterial lipooligosaccharides [28,29] and cell wall components [30,31].

2.2 Glycosylation—a Catalogue of 'Building Blocks' and Biosynthetic Elements

Despite the complexity of glycan production, the biosynthetic components that comprise the 'glycosylation machinery' have become well defined and the resulting knowledge is critical for ongoing progress in MOE. Basic information needed for MOE includes information on the genes responsible for the metabolic processing of the nine common monosaccharides found in mammalian oligosaccharides (Glc, Gal, Man, Xyl, GlcNAc, GalNAc, GlcN, Fuc and sialic acid, **1–9**). In addition, GlcA and IdoA (**10** and **11**) residues found in GAGs and certain sugars, exemplified by ManNAc (**12**) in sialic acid-based MOE, play important roles as metabolic

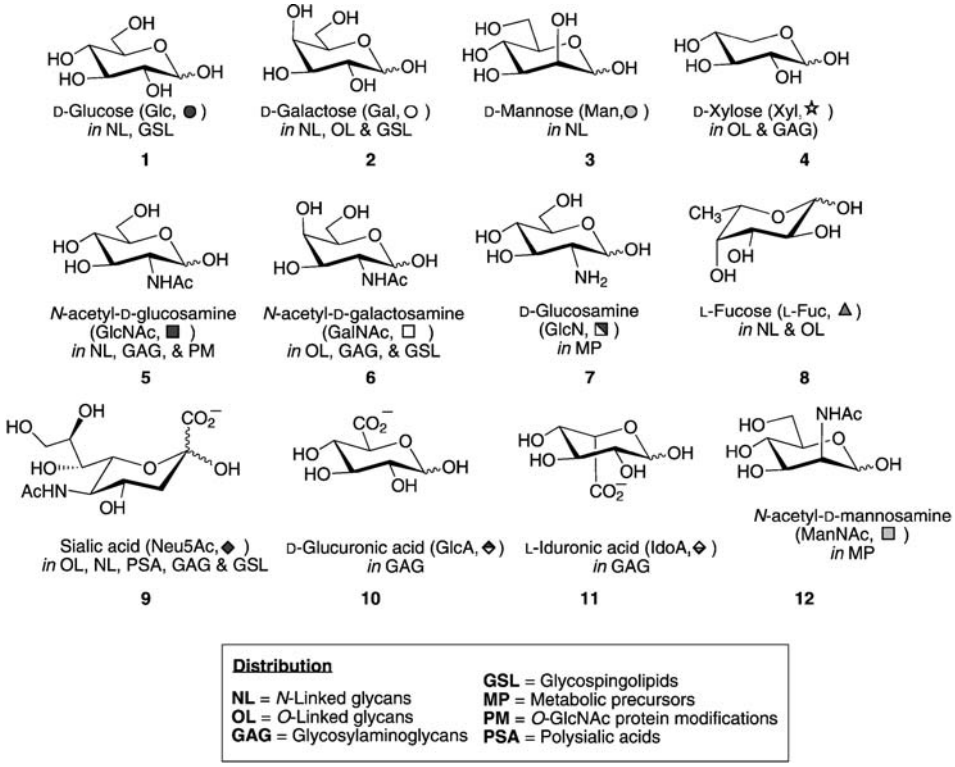


Figure 2
Monosaccharide structures. The chemical structures, symbols specified in "Essentials of Glycobiology" (http://grtc.ucsd.edu/symbol.html), and predominant glycan distribution of common mammalian sugars is given

intermediates [32]. These monosaccharides, shown in *Fig. 2*, are converted to nucleotide sugar 'building blocks' and assembled into glycoconjugates by the sequential and concerted actions of cytosolic processing enzymes, membrane transporters, glycosyltransferases, and glycosidases. Current estimates are that the protein components of the 'glycosylation machinery' collectively comprise 1 to 3% of the human genome [33] resulting in a complex set of metabolic networks [32].

Next, a detailed knowledge of the sequential metabolic intermediates found along the biosynthetic pathways responsible for glycan production is critical in MOE undertakings to ensure that the intended pathway is successfully targeted. From a historical perspective, in many cases the characterization of pathway metabolites significantly predated recently obtained genomic information on the proteins responsible for enzyme and transporter activities. Much of the biochemistry of the sialic acid pathway, for example, was determined in the 1960s by Roseman and colleagues [34,35,36,37] and even earlier for GlcN [38]. As described in detail in *Sect. 3*, a thorough knowledge of metabolic networks allows intelligent choices to be made when deciding where to intercept a biosynthetic pathway for the targeted replacement of a sugar. This decision process is exemplified by MOE manipulation of sialic acid where the

biosynthetic process can be directly modulated by sialic acid analogues but is more commonly controlled through ManNAc derivatives and less frequently by GlcNAc or CMP-sialic acid analogues; as discussed in *Sect. 3.1*, the biochemistry of the pathway stipulates efficient and rationale selection of one of these four possible interception points.

2.3 Overcoming the Plasma Membrane Barrier to Analogue Up-Take by Cells

2.3.1 Non-Natural Sugars Are not Membrane Permeable

A common problem that plagues MOE experiments is the slow initial uptake of the sugar analogues by a cell. Although families of transporters exist for common dietary sugars such as Glc [39], they are usually refractory to modified derivatives such as the GlcNAc analogues (*Sect. 3.4*) now used in MOE. A similar uptake problem exists for ManNAc as no membrane transporters have been reported for even the natural form of this sugar [40] thereby posing a significant hindrance to sialic acid pathway intervention. Consequently, millimolar quantities of mannosamine analogs such as ManNProp (**13**) [17] or ManNLev (**14**) [20], which are presumably taken into cells primarily by diffusion or pinocytosis, are required to maximize metabolic incorporation into sialic acids.

Non-receptor mediated uptake of carbohydrates into a cell by diffusion is severely hampered by the ubiquitous hydroxyl groups of sugars and experimental evidence suggests that the number of hydroxyl groups in the glycan determines the rate and extent of uptake. Because a strong correlation exists between membrane permeability and partitioning of solutes into an organic solvent, one strategy for improving the cellular uptake of monosaccharides has been to increase the hydrophobicity of the sugar by appending lipophilic groups. The difficulty of passing the polar hydroxyl groups through the interior of a membrane, which resembles a low dielectric solvent, was illustrated by the pentoside Xyl-β-*O*-2-naphthol (**15**), which diffuses into CHO cells more readily than the hexoside Gal-β-*O*-2-naphthol (**16**), containing the same aglycone (*Fig. 3A*) [41]. The comparison of these compounds showed that even one additional hydroxyl group creates an obstacle to uptake and the multiple hydroxyl groups of a disaccharide constitute an even larger energy barrier to membrane diffusion. Consequently, naphthol-derivatization successfully used for monosaccharides did not improve the uptake of the Gal-β-1,3-Gal-β-*O*-2-naphthol disaccharide (**17**) [41]. In theory, a further increase in the size and hydrophobicity of the aglycon moiety should enhance uptake but in practice strongly amphipathic compounds act like detergents. Thus, pyrene derivatives of glycosides such as **18** [42] or sugars containing alkyl chains of 15 carbons (**19**) result in cell lysis (*Fig. 3B*).

2.3.2 Short Chain Fatty Acids Increase Cellular Uptake and Modulate Biological Function

To circumvent problems encountered when attaching a single, highly lipophilic moiety to ferry a sugar across a cellular membrane (*Fig. 3B*), modified analogues containing short chain fatty acids (SCFAs) were developed to increased cellular uptake. SCFAs represent a compromise between extremely lipophilic groups that damage the membrane and hydrophilic nature of the sugar that prevents membrane uptake. Moreover, SCFAs, most often acetates, have been

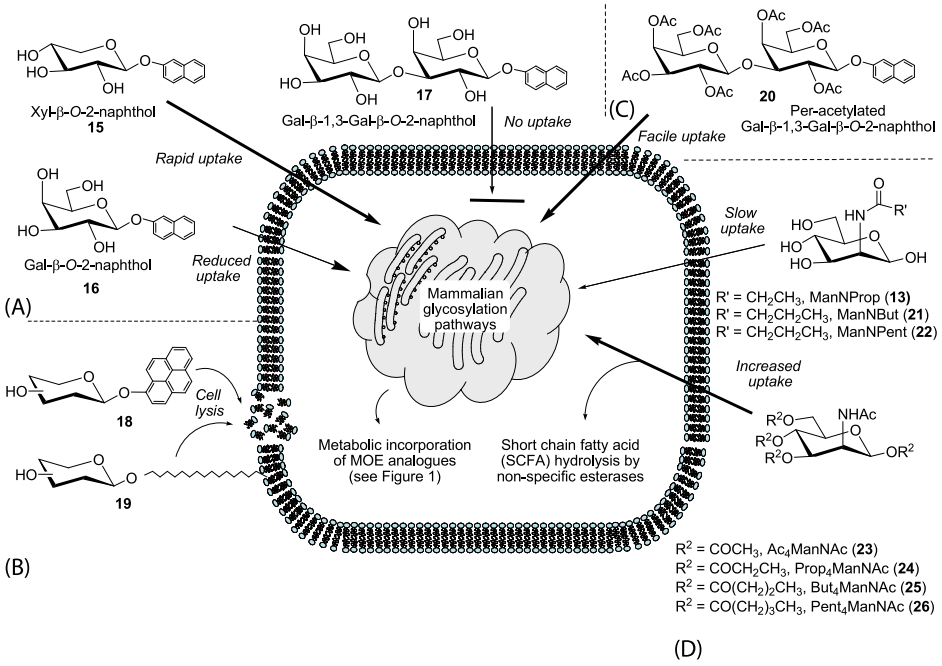
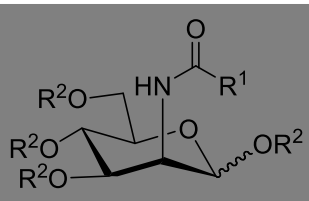


Figure 3
Strategies to improve membrane permeability of MOE analogues. The use of (A) naphthol derivatives or (B) highly lipophilic groups has been attempted to increase cellular uptake of small carbohydrates. (C) Peracetylation of 17 (to form 20) dramatically facilitated uptake leading to the general strategy of using SCFAs for the delivery of MOE substrates including ManNAc analogues (D). Inside a cell, the SCFAs are hydrolyzed by esterases, regenerating the parent monosaccharide that in turn intercepts the targeted glycosylation pathway

used to improve bioavailability of drugs (e. g., acetylsalicylic acid), second messengers (e. g., dibutyryl cAMP), and glycosidase inhibitors (e. g., carbonoyloxy analog of swainsonine) [41]. Cells have numerous esterases, some of which have broad substrate specificity, that can remove the acetates and convert the pro-drugs into their active forms. On the basis of this precedent, it was proposed that the cellular uptake of MOE analogues would be improved by the addition of acetyl esters to the hydroxyl groups of sugars. This hypothesis was demonstrated over a decade ago when the Esko group showed that peracetylation enhanced the uptake of Gal-β-1,3-Gal-β-*O*-2-naphthol (**20**) [41,43] (● *Fig. 3C*). This strategy was then applied to the 'first generation' MOE analogues ManNProp (**13**), ManNBut (**21**), and ManNPent (**22**) where per-acetylation increased metabolic utilization by an estimated ~200-fold in the Jurkat cell line [44] (● *Fig. 3D*); a similar increase upon peracetylation was subsequently observed for ManNLev (**14**) and several other ManNAc analogues (see ● *Table 1* for ManNAc analogues) [45]. In a separate study, a thorough quantitative comparison of peracetylated ManNAc compared to its free monosaccharide counterpart revealed from 300- to 900-fold increased efficiency across cell lines [46]. Together, these studies established that the strategy of using SCFAs to increase cellular uptake applied for multiple analogues across various cell types and can therefore be regarded as generally useful for MOE experiments.

Table 1

ManNAc analogues used in MOE are categorized as having (A) unbranched alkyl substituents, (B) glycolyl or thiolglycolyl groups, (C) branched or bulky groups, (D) ketone groups, and (E) azides at the *N*-acyl position. Representative references are provided for each compound; in most cases additional citations are provided in the text



ManNAc analogues

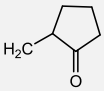
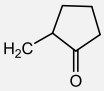
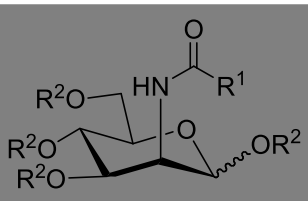
	R^1	R^2	Name	No	Ref(s)
A	CH_3	H	ManNAc	**12**	
	CH_2CH_3	H	ManNProp	**13**	[17,18]
	$CH_2CH_2CH_3$	H	ManNBut	**21**	[17,18]
	$CH_2CH_2CH_2CH_3$	H	ManNPent	**22**	[17,18]
	$CH_2CH_2CH_2CH_2CH_3$	H	ManNHex	**30**	[29]
	$CH_2CH_2CH_2CH_2CH_2CH_3$	H	ManNHept	**56**	[29]
	$CH_2CH_2CH_2CH_2CH_2CH_2CH_3$	H	ManNOct	**57**	[29]
	CH_3	$COCH_3$	Ac_4ManNAc	**23**	[46,47]
	CH_2CH_3	$COCH_3$	Ac_4ManNProp	**35**	[47,63]
	$CH_2CH_2CH_3$	$COCH_3$	Ac_4ManNBut	**58**	[47,63]
	$CH_2CH_2CH_2CH_3$	$COCH_3$	Ac_4ManNPent	**59**	[44,47]
	$CH_2CH_2CH_2CH_2CH_3$	$COCH_3$	Ac_4ManNHex	**60**	[47,63]
	$CH_2CH_2CH_2CH_2CH_2CH_3$	$COCH_3$	Ac_4ManNHept	**61**	[47,63]
	$CH_2CH_2CH_2CH_2CH_2CH_2CH_3$	$COCH_3$	Ac_4ManNOct	**62**	[63]
B	CH_2OH	H	ManNGc	**63**	
	CH_2OH	$COCH_3$	Ac_5ManNGc	**65**	[103]
	CH_2SCOCH_3	$COCH_3$	Ac_5ManNTGc	**68**	[104]
	$CH_2CH_2SCOCH_3$,	$COCH_3$	Ac_5ManNTProp	**70**	[unpub.]
	$CH_2CH_2CH_2SCOCH_3$	$COCH_3$	Ac_5ManNTBut	**71**	[unpub.]
C	CH_2Ph	H	ManNPhAc	**72**	[25,119]
	$CH(CH_3)_2$	H	ManNiBu	**73**	[25,119]
	$C(CH_3)$	H	ManNPiv	**74**	[25,119]
	Ph	H	ManNBz	**75**	[25,119]
	CH_2CF_3	H	ManNTFP	**76**	[25,119]
	$CH(CH_3)CH_2COCH_3$	H	–	**77**	[63]
	$CH(CH_3)CH_2COCH_3$	$COCH_3$	–	**78**	[63]
		H	–	**79**	[63]
		$COCH_3$	–	**80**	[63]
	$CH{=}CHCH_3$	H	ManNCrot	**81**	[18]

Table 1
(continued)

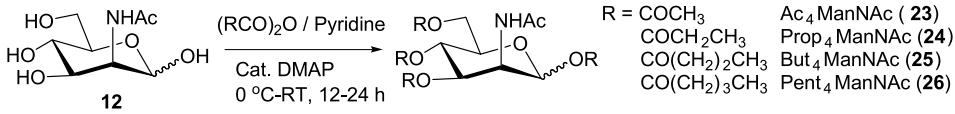


ManNAc analogues

	R^1	R^2	Name	No	Ref(s)
D	CH_2COCH_3	H	–	**82**	[63]
	$CH_2CH_2COCH_3$	H	ManNLev	**14**	[20,63]
	$CH_2CH_2COCH_2CH_3$	H	ManNHomoLev	**83**	[47]
	$CH_2(CH_2)_2COCH_3$	H	ManNOxoHex	**31**	[47]
	$CH_2(CH_2)_3COCH_3$	H	ManNOxoHept	**84**	[47]
	$CH_2(CH_2)_4COCH_3$	H	ManNOxoOct	**85**	[47]
	CH_2COCH_3	$COCH_3$		**86**	[63]
	$CH_2CH_2COCH_3$	$COCH_3$	Ac_4ManLev	**87**	[44,63]
	$CH_2CH_2COCH_2CH_3$	$COCH_3$	Ac_4ManNHomoLev	**88**	[47,63]
	$CH_2(CH_2)_2COCH_3$	$COCH_3$	Ac_4ManNOxoHex	**89**	[47,63]
	$CH_2(CH_2)_3COCH_3$	$COCH_3$	Ac_4ManNOxoHept	**90**	[47,63]
	$CH_2(CH_2)_4COCH_3$	$COCH_3$	Ac_4ManNOxoOct	**91**	[47,63]
E	CH_2N_3	H	ManNAz	**33**	[23,60]
	CH_2N_3	$COCH_3$	Ac_4ManNAz	**34**	[23,60]

Because the increase in metabolic efficiency observed upon peracetylation of a mono- or disaccharide analogue is attributed to enhanced membrane diffusion due to increased lipophilicity, our laboratory tested whether SCFAs longer than acetate would further improve cellular uptake. Upon the synthesis of a series of SCFA-ManNAc hybrid molecules (*Scheme 1*), we found additional increases in metabolic efficiency for propionate-derivatized ManNAc (**24**) (~1,800 fold compared to the free hydroxyl form of this sugar) as well as for fully butyrate-derivatized ManNAc (**25**) (~2,100) [47]. These analogues also provided an opportunity to investigate a potential downside of SCFA-derivatized sugars; which is their tendency to be toxic. Although toxicity is not a major concern because it usually occurs at concentrations greater than required for robust metabolic incorporation [46,48], and has not been reported to be significant in several in vivo studies [49,50], its mechanistic basis is nevertheless of interest. One possibility—with scant experimental evidence to date—is that the increased analogue uptake exacerbates latent sugar effects. A second possibility—supported by several studies—is that toxic effects are due to the short chain fatty acid itself.

Experimental evidence that acetates are not silent moieties that only function as delivery vehicles to increase membrane uptake came from a comparison of the effects of ManNAc and Ac_4ManNAc (**23**) on mRNA levels for the enzymes involved in sialic acid biosynthesis. In these studies, the addition of exogenous ManNAc to the culture medium had no effect on these enzymes [51] whereas supplementation with **23** at levels that increased metabolic flux



Scheme 1
Synthesis of SCFA-derivatized ManNAc

through the pathway to a similar level changed the expression of several genes [52]. The ability of SCFAs, including acetate [53], to function as histone deacetylase inhibitors (HDCAi) and thereby influence gene expression [54] provided a tentative mechanistic explanation for these results. To confirm that sugar-delivered SCFAs had the widely observed biological activities expected from these compounds, we evaluated the effects of But_4ManNAc (**25**)—a hybrid molecule consisting of butyrate and ManNAc—on HDAC inhibition and gene expression. In these experiments we substituted butyrate for the acetate to exacerbate SCFA-specific cellular responses because this 4-carbon SCFA has the highest level of biological activity. The ability of **25** to function as an HDACi was tested in several human cancer cells lines and the expected suite of SCFA responses including inhibition of proliferation, induction of apoptosis, alteration of histone acetylation, and up-regulation of $P21^{WAF1/Cip1}$ expression were observed [55].
Several important implications arose from the finding that SCFA 'protecting groups' influence biological activity. First, until the complex interactions between SCFAs and sugar metabolism are more thoroughly understood, the use of these protecting groups in MOE experiments may give rise to unanticipated biological activities that complicate the interpretation of the results. Second, once these responses are adequately understood, SCFA-derivatization will prove a design opportunity to extend the sugar-specific effects of MOE analogues over biological systems. Such opportunities were vividly illustrated by experiments that compared fully butyrated ManNAc (But_4ManNAc, **25**) and mannose (But_5Man, **27**) and demonstrated that the core sugar had a powerful ability to tune the activity of the SCFA moiety. In particular, although both compounds exhibited canonical SCFA activities such as modulation of histone acetylation status, **25** elicited significantly different patterns of gene expression than **27**. The former compound also induced apoptosis in cancer cell lines at concentrations where **27**, as well as the stereoisomer But_4GlcNAc (**28**), only provided transient cell cycle arrest [55]. Together, these experiments demonstrate that acetate and butyrate are necessary but not sufficient for toxicity and that biological responses to SCFAs are exquisitely tuned by the core sugar.

2.4 Monitoring Surface and Metabolic Flux Responses to MOE Analogues

2.4.1 Measurement of Surface Display of Non-Natural Sugar Residues

The surface display of non-natural sugar residues requires methods to assess the degree of replacement of natural sugars with the corresponding non-natural residues. The characterization of analogues with inert alkyl chains, exemplified by the original set of MOE analogues ManNProp (**13**), ManNBut (**21**), and ManNPent (**22**) developed by the Reutter group

(see ❯ *Sect. 3.1*) for surface display as modified sialic acids, is relatively time consuming, laborious, and generally cannot be performed on live cells. Glycosidases often do not function on the modified sugars, eliminating a straightforward route to analysis wherein these enzymes are used to remove modified residues from the surfaces of intact cells. In some cases, lectins can be used to detect surface changes upon analogue incorporation [56]; other times antibodies can be generated [57] that can be used—at least in theory—as imaging agents [58]. While neither lectins nor antibodies have yet proved to be highly quantitative and reproducible for probing modified glycan display, they remain valuable because they can provide conclusive qualitative evidence for metabolic incorporation of an analogue.

Radiolabeled probes constituted an early strategy to quantify analogue incorporation into cellular glycans and follow incorporation into membrane fractions [17]. Rigorous determination of analogue incorporation into glycans, however, required isolation of the carbohydrate-containing components by methods not compatible with physiological conditions such as acid hydrolysis followed by chromatographic analysis [50]. Similarly, separation methods including thin layer chromatography that have been employed for detection of sialic acids found in gangliosides [59] can not be performed on living cells. Consequently, the tandem development of analogues with orthogonal chemistries not naturally found in the glycocalyx, including ketones [20], azides [60], and thiols [61] along with complementary chemoselective ligation strategies compatible with physiological conditions have greatly facilitated the detection and quantification of surface displayed non-natural sugars (this topic is discussed in detail in ❯ *Sect. 4*).

2.4.2 Measurement of Intracellular Flux and Metabolite Levels

Several factors have motivated the development of methods to quantify intracellular metabolites to complement measurements of surface display of non-natural sugars. First, in cases when the detection of modified sugars such as Sia5Prop (**29**) on the cell surface is difficult, the monitoring of intracellular flux provides a rough measure of metabolic utilization of the precursor analogue. For example, it is relatively straightforward to measure the production of sugar intermediates by colorimetric assays, especially for the sialic acid pathway where the periodate-resorcinol assay [62] has been widely applied. Consequently, the metabolic uptake and pathway incorporation of analogues such as ManNProp (**13**) or ManNBut (**21**) that support a large increase in flux can be monitored easily by such colorimetric assays [48,63]. By contrast certain analogues that lead to surface modifications, such as ManNHex (**30**) or ManNOxoHex (**31**), do not measurably change intracellular metabolite levels. In these cases the flux of non-natural intermediates through the pathway is real but quantitatively small compared to endogenous metabolites and simple colorimetric assays are not informative.

In situations where both the surface display and intracellular metabolites resulting from analogue metabolism are refractory to rapid analysis, a benchmark of utilization can be obtained by analysis of analogue uptake from the medium. Although uptake occurs in advance of pathway incorporation and may not always directly correspond to metabolic utilization, the analysis of conditioned media by functional assays as described by Jones and coworkers is valuable for probing early steps (i. e., cell uptake) in the MOE process [46]. This method is mainly of value for SCFA-derivatized sugars where a large proportion (> 50%) of total analogue is removed from the culture medium compared. By contrast, the margin of error in this analy-

sis of +/− ~5% can mask uptake of free hydroxyl monosaccharides where a much smaller fraction (often less than 1%) of total analogue is removed from the culture medium by the cells. Although requiring significantly more effort, detailed analytical characterization of intracellular metabolites or media components, to test pathway interception or membrane transit, respectively, is possible by chromatographic methods [64] or by using radioactive tracers [63]. In conclusion, monitoring intracellular metabolic intermediates serves the dual purposes of verifying intake into the targeted glycosylation pathway as well as identifying 'bottlenecks' that hinder efficient metabolic processing, a topic that will be discussed at more length in Sect. 3.1.7.

2.5 Scope of MOE in Living Systems

2.5.1 Mammalian Cell Lines

To date, MOE technology been most extensively applied to mammalian cells; Jurkat cells in particular have been employed in many experiments and HeLa and HL-60 cells have also been well studied. Although analogue utilization appears to be a universal feature of human cells, the extent of incorporation varies several-fold based on several factors. In some cases, size differences between cells can largely explain differences in surface display; this situation is exemplified by the Jurkat and HeLa lines where the 3- to 4- fold difference in total surface display of Sia5Lev (**32**) upon ManNLev (**14**) exposure roughly corresponded to the relative surface areas of these lines [65]. In other cases, exemplified by human chondrocytes, lower incorporation is hypothesized to result from the copious secretion of extracellular matrix components that sequester the analogue and hinder its up-take by the cells (K. J. Yarema, unpublished results). In addition, analogues are utilized differently based on subtle differences in their chemical structures; this topic is discussed in depth in Sect. 3.1.7. Finally, species-specific differences in analogue metabolism can be particularly pronounced, with rodent cells showing dramatically lower (e. g., ~20-fold) surface display of ketones when incubated with **14** [65] and azides when incubated with ManNAz (**33**) [23] compared to human lines.

On the basis of these considerations, as a general guideline for initial experiments where novel analogues undergo evaluation in untested cell lines, peracetylated analogs should be tested over the range of 0 to ~250 μM. Metabolic incorporation usually does not increase significantly above 250 μM for peracetylated analogues but the 'real' upper limit for media supplementation is typically determined by dose limiting toxicity (DLT), which can range between 50 to > 700 μM depending on the cell line [48], the *N*-acyl substituent [47], and cell density [46]. Generally, DLT is not observed for free hydroxyl monosaccharides and concentrations up to 50 mM [65] or higher [55] result in increasing surface display and intracellular metabolite production. When analogue concentrations approach 100 mM, however, further maximization of metabolic uptake is not possible due to the deleterious effects of such high levels of solute.

2.5.2 Animal Tests

The evaluation of MOE technology in vivo has taken place almost entirely in rodent models with several mouse and rat studies having been reported. A striking feature shared amongst many studies is that analogue metabolization is highly organ- and tissue-specific for analogs

ranging from ManNProp (**13**), ManNBut (**21**), and ManNPent (**22**) [17,66] to Ac_4ManNAz (**34**) [67]. It remains unclear if pharmacokinetic factors are primarily responsible for differences in incorporation where organs that simply receive more analogue experience higher levels of metabolic replacement. For example, the heart receives high levels of blood flow suggesting it would be exposed to high levels of analogue, which is consistent with the high levels of sialic acid replacement observed for this organ in a recent study by Gagiannis and coworkers [50]. In this study the brain—hypothesized to have low incorporation because it is a privileged compartment due to the blood brain barrier [22,67]—had very low incorporation of Sia5Prop (**29**). This study, however, revealed the complexity of analogue metabolism in vivo by showing that PSA production was down-regulated in brain Ac_4ManNProp (**35**), thereby indicating that the analogue was not excluded from this organ after all. Instead, the brain was able to avoid the most common biological response to this analogue [i. e., glycan incorporation of Sia5Prop (**29**)] while directing a less common outcome (i. e., inhibition of PSA production [68]). As more information becomes available, including the ability of rodent B cell glycoproteins to become robustly labeled with Ac_4GalNAz (**36**) exposure while T cell glycoproteins were not [22], it is becoming clear that fundamental differences in the glycosylation machinery or metabolism exist between even very similar cell types.

2.5.3 Bacteria, Plants, Insects, and Yeast

Bacteria, yeast, plants, or insect cells have not been extensively investigated as hosts for MOE. However, there are hints—and in some cases pilot studies—that each of these cell types may be amenable to incorporation of non-natural sugars. For example efforts to 'humanize' glycoprotein production in yeast [69] and insect cells [70], to metabolically engineer carbohydrate production in bacteria [71], and to investigate sialic acid production in plants [72] delve in depth into pathway analysis of the type required to design analogues to intercept specific pathways at defined points (see *Sect. 3*). Already there have been reports that sialylation of recombinant proteins can be increased in insect cells through supplementation with ManNAc analogs [73]. The manipulation of carbohydrates in bacteria has also been achieved through the ingenious use of MOE. To briefly highlight two prokaryotic results, Schilling and coworkers demonstrated the biosynthesis of lipooligosaccharides in *Haemophilus ducreyi* with modified sialosides bearing extended *N*-acyl moieties. In *H. ducreyi*, unlike mammalian cells, analogue incorporation was dependent on exogenous sialic acids and could not be achieved with modified mannosamines [28,29]. In a second pilot study targeting oligosaccharide display on the surfaces of bacteria, Nishimura's group synthesized UDP-MurNAc pentapeptide derivatives that were incorporated into cell wall components of *Lactobacilli plantarum* [30,31,74].

3 Glycosylation Pathways and Analogues Used in MOE

This section will systematically discuss the various glycosylation pathways that have been targeted by MOE to date, using the well-studied sialic acid pathway to illustrate several aspects of analogue design and utilization. After that, efforts to extend MOE to GalNAc, fucose, and GlcNAc will be summarized followed by a brief discussion of glycosylation inhibitors that share mechanistic similarities with 'conventional' MOE analogues.

3.1 Metabolic Precursors for the Biosynthesis of Cell Surface Sialosides

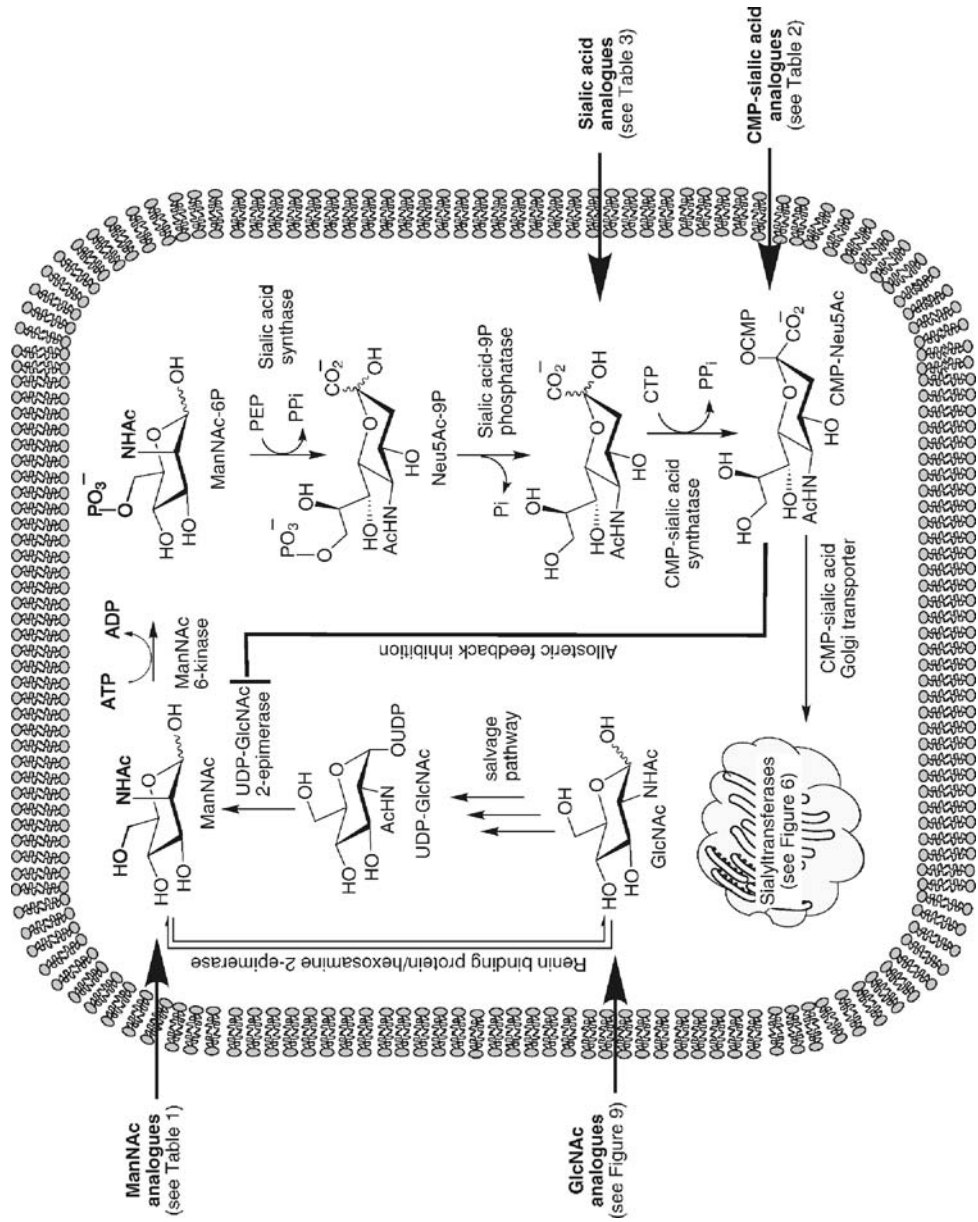
3.1.1 Overview of Sialic Acid Biosynthesis and Biological Function

The sialic acid biosynthetic pathway, which consists of the enzymes, transporters, and metabolites shown in *Fig. 4* and described in detail elsewhere [18,75], was exploited to pioneer both the basic concept of MOE [17] as described in *Sect. 3* as well as the subsequent development of the cell surface chemoselective ligation technology [20,76] described in *Sect. 4*. As a brief background, sialic acids are attractive candidates for MOE for several reasons. First, these sugars are usually terminal residues situated at the non-reducing ends of oligosaccharides. Consequently, they are located at the outer periphery of the glycocalyx and are ideally positioned to interact with other cells, signaling molecules such as Siglecs or cytokines, pathogens, or exogenously supplied chemoselective ligation reagents. The sialic acid biosynthetic pathway is also attractive for MOE because it is unusually permissive to substrate alterations, perhaps because Nature already exploits sialic acid biosynthesis with a type of MOE where over 50 different naturally occurring structural variations are introduced into this sugar (*Fig. 5*) [77]. The most abundant sialic acid in man is 5-acetamido-D-*glycero*-D-*galacto*-2-nonulosonic acid (Neu5Ac, **9**). The 5-glycolylamido derivative *N*-glycolyl-D-neuraminic acid (Neu5Gc, **37**) is common in most other animals and the non-aminated 3-deoxy-D-*glycero*-D-*galacto*-2-nonulosonic acid (KDN, **38**) is found in many other biological systems [78,79]. Besides of these three common sialic acids, additional natural modifications typically involve mono-acetylation of hydroxyl groups, most often at the C-9 position, although di- and tri-*O*-acetylated derivatives also occur along with other variations that range from lactoylation or phosphorylation at C-9 to methylation or sulfation at C-8 [80].

Several functional consequences of Nature's MOE of sialic acid are already well known; in particular, biological responses to the selective display of Neu5Ac (**9**) and Neu5Gc (**37**) are profound. Differential incorporation of these two sugars, which vary only by the absence or presence of a hydroxyl group appended to the *N*-acyl chain, have been suggested to play roles in human evolution both far in the past by affecting brain development [81] and more recently through persistently changing patterns of Siglec recognition [82]. MOE technology extends the repertoire of chemical changes that Nature uses to tune the chemical and functional properties of sialic acid. For example, in the early experiments, ManNProp (**13**) was used to substitute the hydroxyl of **37** with a methyl group [17]. Even though the resulting structural change to sialic acid structure was modest, **13** elicits a gamut of biological responses not found with natural sialic acids (as discussed below in *Sect. 3.1.4*).

A second layer of complexity that both confounds and increases the potential ability of MOE experiments to modulate biological responses lies in the various glycosidic linkages of sialic acid (*Fig. 6*). In humans, which only express a subset of these linkages (α-2,3-, α-2,6-, and α-2,8-), 20 sialyltransferases (STs [51]) work in parallel to install sialic acids onto glycoproteins and lipids. The structural importance of each linkage is illustrated by binding preference

Figure 4
The sialic acid biosynthetic pathway. An outline of the sialic acid biosynthetic pathway is diagrammed along with four options for intercepting this pathway with MOE analogues


ManNAc analogues (see Table 1)
GlcNAc analogues (see Figure 9)
Sialic acid analogues (see Table 3)
CMP-sialic acid analogues (see Table 2)
Renin binding protein/hexosamine 2-epimerase
ManNAc
ManNAc-6P
ManNAc 6-kinase
ATP
ADP
UDP-GlcNAc 2-epimerase
UDP-GlcNAc
GlcNAc
salvage pathway
Allosteric feedback inhibition
PEP
PPi
Sialic acid synthase
Neu5Ac-9P
Sialic acid-9P phosphatase
Pi
CMP-sialic acid synthatase
CTP
PPi
CMP-Neu5Ac
CMP-sialic acid Golgi transporter
Sialyltransferases (see Figure 6)

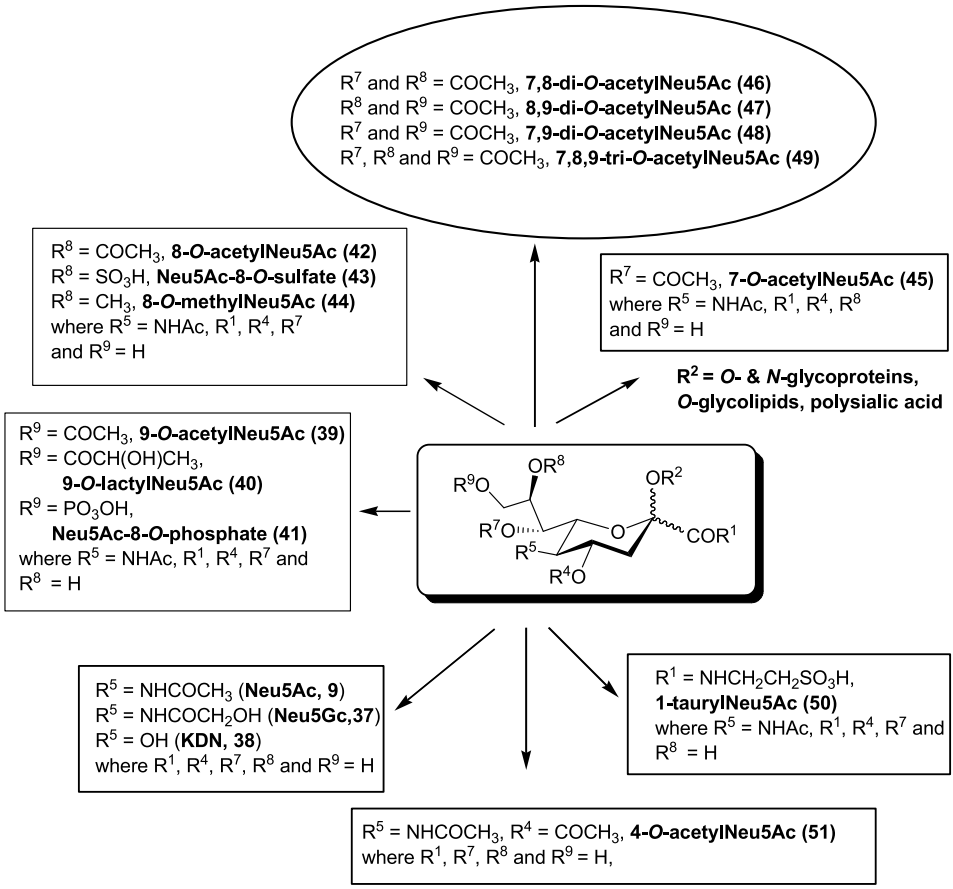


Figure 5
Structural variations found in natural sialic acids

of influenza viruses that show exquisite discrimination between α-2,3- and α-2,6-linked sialic acids [83,84] and can rapidly evolve binding specificity for one linkage over the other [85]. Because of the biological repercussion of sialic acid display in these different linkages, it is likely that certain MOE applications will require careful control to partition metabolic analogues between the various glycosidic linkages. Achieving this objective will be challenging

Figure 6
Glycosidic linkages of sialic acid. In mammals the nucleotide sugar donors 52 and 53 are used to supply sialic acids 9 and 37 that either α-2,3- or α-2-6-linked to Gal and GalNAc or α-2-8-linked to an underlying sialic acid. The human sialyltransferases responsible for each linkage are listed with detailed information given elsewhere [51]. Linear homopolymers of 9 and 37, either α-2,8-, α-2,9-, or alternating α-2,8-/α-2,9-linked, have been found in glycoproteins of embryonic neural membranes and neural cell adhesion molecule, fish eggs, and certain bacteria. Polymers of 37 have also been found in sea urchin eggs linked through the glycolyl hydroxyl group resulting in α-2,5-linkages

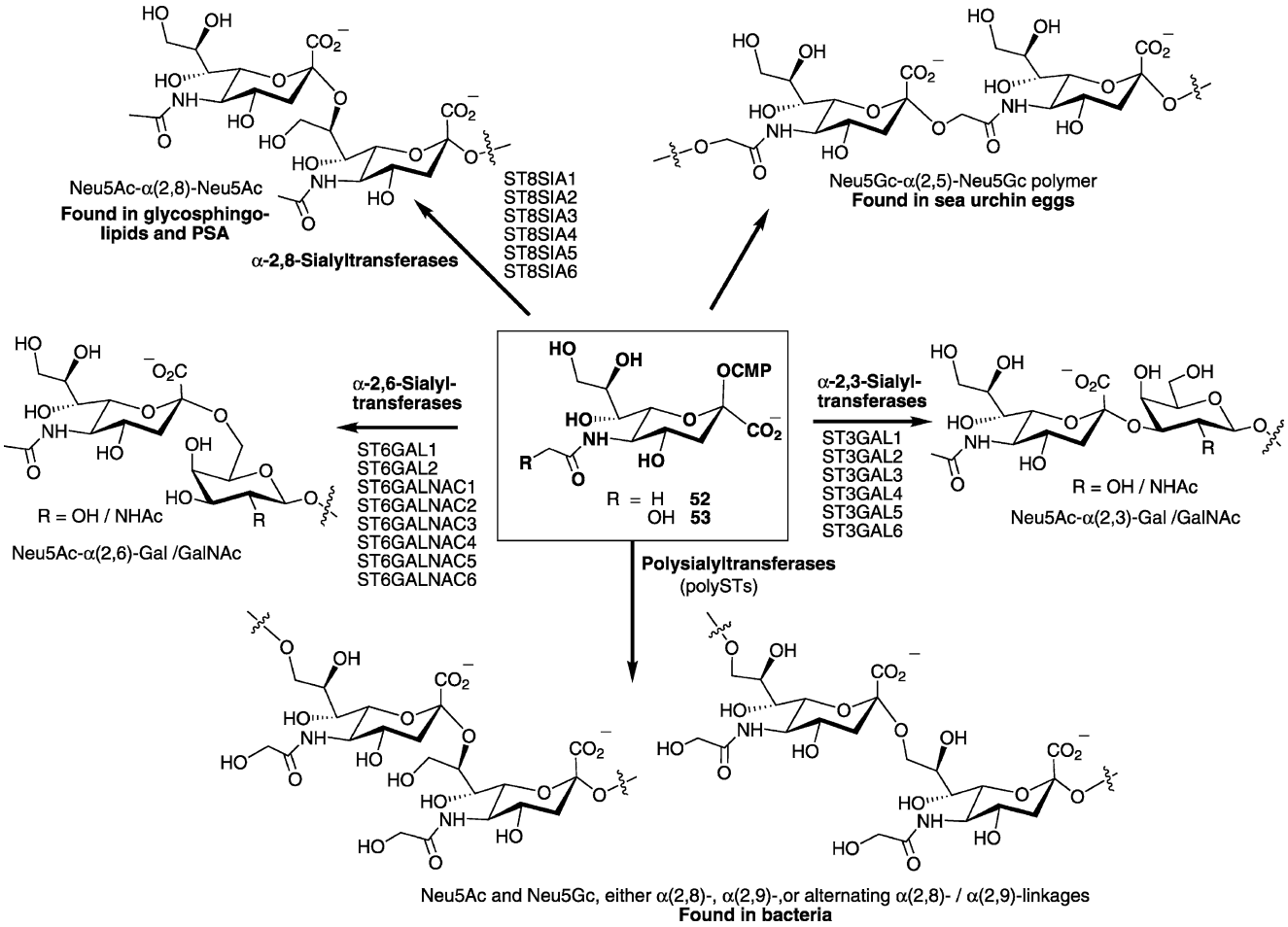
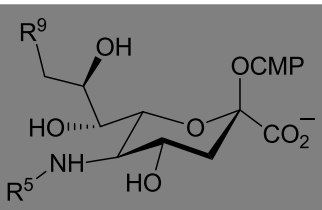

Neu5Ac-α(2,8)-Neu5Ac
Found in glycosphingo-lipids and PSA
α-2,8-Sialyltransferases
ST8SIA1
ST8SIA2
ST8SIA3
ST8SIA4
ST8SIA5
ST8SIA6
Neu5Gc-α(2,5)-Neu5Gc polymer
Found in sea urchin eggs
α-2,6-Sialyl-transferases
ST6GAL1
ST6GAL2
ST6GALNAC1
ST6GALNAC2
ST6GALNAC3
ST6GALNAC4
ST6GALNAC5
ST6GALNAC6
R = OH / NHAc
Neu5Ac-α(2,6)-Gal /GalNAc
OCMP
R = H 52
OH 53
α-2,3-Sialyl-transferases
ST3GAL1
ST3GAL2
ST3GAL3
ST3GAL4
ST3GAL5
ST3GAL6
R = OH / NHAc
Neu5Ac-α(2,3)-Gal /GalNAc
Polysialyltransferases
(polySTs)
Neu5Ac and Neu5Gc, either α(2,8)-, α(2,9)-,or alternating α(2,8)- / α(2,9)-linkages
Found in bacteria

Table 2
CMP-sialic acid analogues used in MOE



	CMP-sialic acid analogues				
	R^9	R^5	Name	No	Ref
A	N_3	$COCH_3$	CMP-9-Azido-Sia5Ac	**98**	[112]
	NH_3^+	$COCH_3$	CMP-9-Amino-Sia5Ac	**99**	[112,107]
	NHAc	$COCH_3$	CMP-9-Acetamido-Sia5Ac	**100**	[112,107]
	$NHCSCH_3$	$COCH_3$	CMP-9-ThioacetamidoSia5Ac	**101**	[108]
	NHCOPh	$COCH_3$	CMP-9-benzamidoSia5Ac	**102**	[107]
	N-fluoresceinyl) thioureide	$COCH_3$	CMP-9-deoxy-9-*N*-*N*-fluoresceinyl) thioureido-Sia5Ac	**103**	[108]
	$NHCO(CH_2)_4CH_3$	$COCH_3$	CMP-9-hexanoylamidoSia5Ac	**104**	[107]
B	OH	$COCH_2NH_3$	CMP-5-*N*- aminoacetylNeu5Ac	**105**	[108]
	OH	$CSCH_3$	CMP-5-*N*-thioacetylSia5Ac	**106**	[108]
	OH	CHO	CMP-5-*N*-formylSia5Ac	**107**	[108]

because of several factors that could, in theory, skew analogue distribution. For example, the SCFA groups used to increase metabolic uptake of ManNAc alter patterns of ST expression [52]. In addition, the Km of each ST for various nucleotide sugar donors (i. e., for CMP-Neu5Ac (**52**) vs. CMP-Neu5Gc (**53**) vs. CMP-Neu5Prop (**54**), or for the other sialosides shown in ❯ *Table 2*) may be different resulting in predominant or disproportional display of a non-natural sialic acid in a particular linkage or on a subset of proteins or lipids. In the future, once these complex effects are better understood, it may be possible to exploit them to gain exquisite control over biological function. At present, in the absence of definitive information, we include this discussion as a cautionary note if MOE results in otherwise inexplicable biological outcomes.

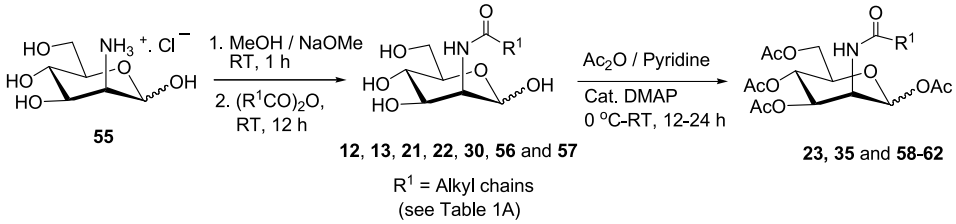
3.1.2 ManNAc Analogues Exploit Feedback Inhibition of GNE for Efficient Sialoside Display

The epimerization of UDP-GlcNAc to ManNAc (**12**) by UDP-GlcNAc 2-epimerase/ManNAc 6-kinase (GNE) is the first committed step in sialic acid biosynthesis (❯ *Fig. 4*) [86]. Because there are no other known fates for **12** (with the possible exception of conversion to GlcNAc in cells over-expressing recombinant renin-binding protein [87] that was later found to have C-2 epimerase activity for hexosamines [88]), sialic acid biosynthesis is uniquely targeted by ManNAc analogues. Importantly from a practical perspective, ManNAc analogues feed into the biosynthetic pathway directly downstream of the point of feedback inhibition, allowing

incorporation with reduced competition from the incoming flux of natural **12**. The importance of diminished metabolic competition is illustrated by experiments where exogenously added **12** effectively out-competed non-natural analogues for transit through the pathway [65]. Similarly, the sialuria form of GNE [89], which is not subject to stringent feedback inhibition and therefore allows a high level of UDP-GlcNAc to be converted to **12** [90], ablates surface expression of sialic acid analogues from non-natural ManNAc precursors [91]. Conversely, cells with inactive forms of GNE that are consequently unable to produce endogenous **12** are superior hosts for replacement of surface sialic acids with their non-natural counterparts [92,93].

3.1.3 The 'First Generation' ManNAc Analogues had Elongated *N*-Acyl Alkyl Chains

The application of MOE to sialic acid was pioneered by Reutter and coworkers who synthesized ManNProp (**13**), ManNBut (**21**), and ManNPent (**22**), the three, four, and five carbon homologues of ManNAc, respectively (*Scheme 2* and *Table 1*A) [17,24]. These compounds were converted into their respective sialic acids by living cells without affecting their growth rate or viability and into the sialosides of rodents without any notable toxicity [50]. In subsequent experiments, limits to the permissivity of the sialic acid pathway in mammalian cells were probed by using the extended *N*-acyl chain analogues ManNHex (**30**), ManNHept (**56**), and ManNOct (**57**) (analogues with longer alkyl chains than **57** tend to form micelles under aqueous conditions). In human cells, longer chain derivatives beyond the straight chain 'Hex' substituent, as well as analogues with branching side chains [25,63], show negligible flux through the sialic acid pathway [47] or incorporation into sialoglycoconjugates [63]. Interestingly, however, the bulky *N*-phenylacetyl form of ManNAc (**72**) was efficiently used by human cancer cells and incorporated into ganglioside GM3 [25].



Scheme 2
Syntheses of mannosamines with modified *N*-acyl groups

3.1.4 Biological Responses Supported by ManNProp, ManNBut, and ManNPent

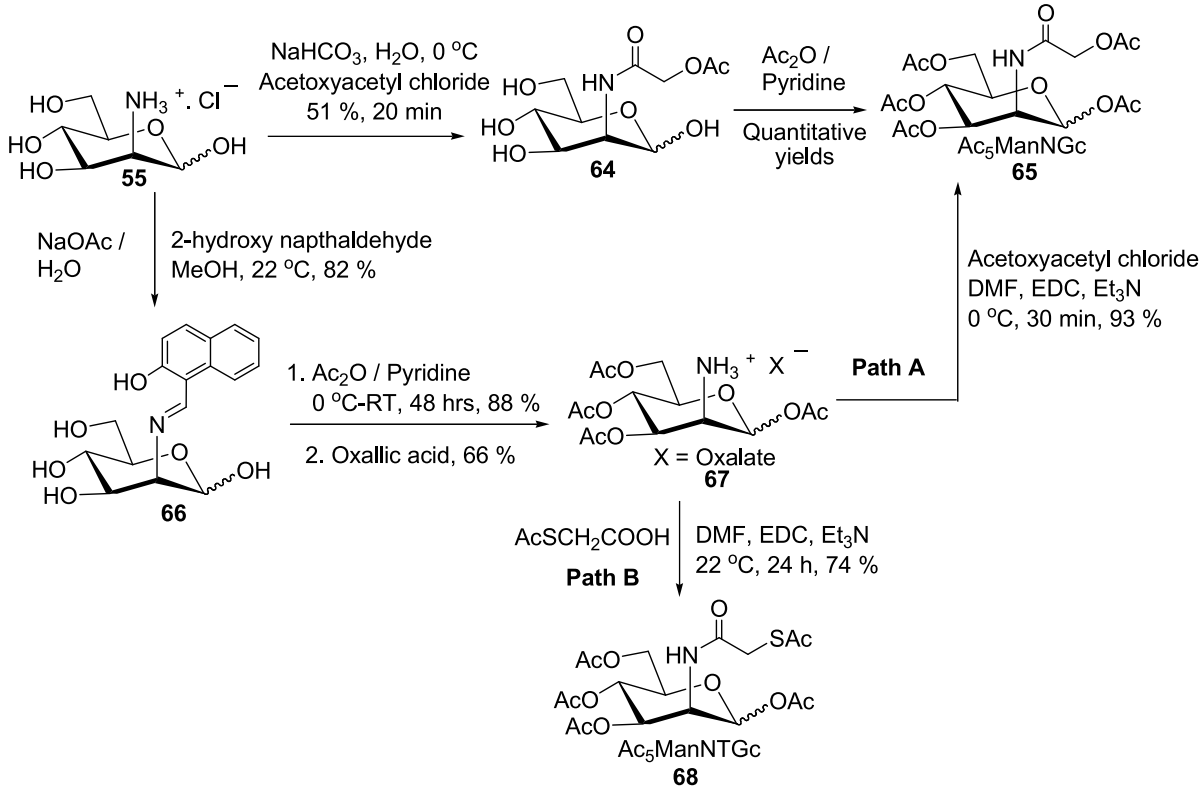
Although the primary focus of this paper is not to provide an in depth analysis of the surfeit of biological responses to analogues, the ability of these compounds to intervene in many aspects of cell biology is critical to propelling the field forward and to justifying continued research in this area. Consequently, by using the seminal ManNProp (**13**), ManNBut (**21**), and ManNPent (**22**) analogs as exemplars, we briefly survey biological responses to these compounds.

First, considering that many viruses, including HIV and influenza, have binding specificities determined by sialic acid, **13**, **21**, and **22** offer broad—and usually inhibitory—perturbation of viral binding and infectivity that may provide therapeutic options against the looming threat of global pandemics [18,94]. Certain analogs, for example **13**, have also shown potential as anticancer drugs through their incorporation into polysialic acid [95] or gangliosides [25], where they function as tumor-specific antigens [57]. Analogues **21** and **22**, by virtue of their ability to inhibit PSA production [96] by incorporation of non-natural sialosides into the nascent polymer that in turn terminates further elongation [68], may also find roles in cancer therapy. Although these therapeutic applications remain in the early stages of development, ManNAc analogues have already become valuable research tools. For example, the reversible inhibition of PSA expression with these small molecules [68] is a powerful tool for studying the role of this epitope in neuronal development [96,97,98] and tumor-related processes [95]. Similarly, fibroblasts cultured in the presence of **13** experienced suppression of contact-dependent growth inhibition [99] providing early evidence of sialic acid's role in linking cell adhesion with signaling.

Many of the profound biological consequences of *N*-acyl modification of sialic acids are anticipated. For example, the snug fit of the *N*-acyl side chain into the binding pocket of hemagluttin of influenza virus [100] portended the results—predicted by thermodynamic considerations [18]—where increased steric bulk at this submolecular site inhibited influenza A virus infection by as much as 80% [101]. In other cases, biological responses to MOE analogues have been unanticipated; for example, **13**, **21**, or **22** sometimes increased viral binding or infectivity. Even more puzzling have been reports that **13** can alter the expression of genes such as integrins at the mRNA level [102]. One possible explanation is that the transit of non-natural metabolites through the nucleus affects transcription. A second possibility is that cross-talk between sialic acid on the cell surface and signaling pathways upstream of gene expression are modulated by this analogue. Either way, solving this puzzle is sure to provide added insights into the biological roles of sialic acid and may provide novel methods to control the relevant molecular systems.

3.1.5 Introduction of Neu5Gc and Neu5TGc into Glycans

As mentioned earlier, every day Nature in effect uses MOE technology in humans who consume meat or animal products to restore man's lost ability [82] to biosynthetically install the Neu5Gc (**37**) form of sialic acid into human glycans. The cell surface display of dietary **37** demonstrates that the substitution of a hydroxyl group on the *N*-acyl side chain is tolerated by late steps in the pathway. In the laboratory, MOE experiments show that the installation of **37** onto human cells can also be accomplished with Ac_5ManNGc (**65**), a compound that can be synthesized by the route shown by Path A in ❯ *Scheme 3* [52]. In neurons, the presence of **37** abrogated binding of the myelin-associated glycoprotein [103], an event implicated in the inhibition of nerve regeneration after injury, thereby opening the door to new approaches for the regeneration of damaged nerve tissue. On the basis of the precedent established by dietary **37** as well as by **65** for pathway permissivity for hydrophilic groups located at the termini of the *N*-acyl chain, our laboratory synthesized the *N*-thioglycolyl counterpart Ac_5ManNTGc (**68**) (❯ *Scheme 3*, Path B) [52,104]. A key feature of Neu5TGc (**69**), the surface product of **69** is that, unlike the alkyl chain and hydroxyl analogues discussed above, an orthogonal chemical



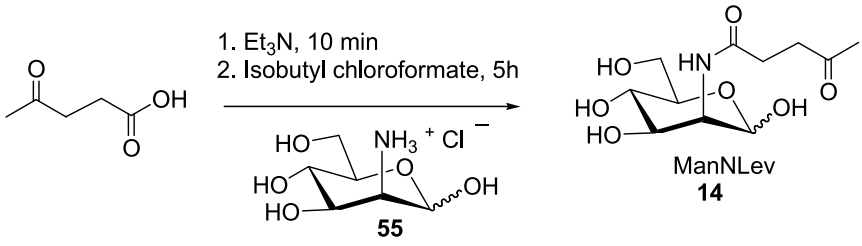
Scheme 3
Synthesis of glycol- and thiol-containing ManNAc analogues

reactivity—the thiol group—is installed in the unique location of the glycocalyx. As such, **68** builds on the work initiated by the Bertozzi lab and described below in more detail to introduce new chemical functionalities onto the cell surface that can be exploited via chemoselective ligation reactions.

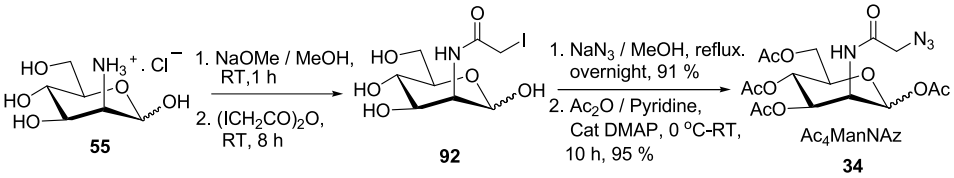
3.1.6 Introduction of Orthogonal Chemical Functional Groups into ManNAc Analogues

The 'first generation' ManNAc analogues used in MOE displayed relatively simple alkyl chain extensions devoid of chemical reactivity. In follow-up experiments, structural diversity was increased by the use of branching and ring structures (● *Table 1*C), as well as chemical functional groups already common in sugars such as the hydroxyl of Neu5Gc (**37**). While these compounds were valuable for diverse purposes ranging from cancer immunotargeting [25,57] to altering receptor-binding interactions [56,103], the efforts of the Bertozzi group to exploit metabolic processes to incorporate orthogonal, reactive functionality into cell-surface glycoproteins added an important new dimension to MOE. In the groundbreaking experiments, a mannosamine analogue derivatized with a levulinoyl side chain (ManNLev, **14**) was synthesized from the commercially available mannosamine hydrochloride (● *Scheme 4*) [20]. The key feature of ManNLev was the presence of the ketone group in the *N*-acyl levulinoyl ('Lev') moiety that is foreign to the cell surface and will readily undergo selective reaction with chemoselective ligation agents such as aminooxy, hydrazide, or related functional groups to form stable covalent adducts. Following the synthesis of ManNLev, several additional ketone-derivatized ManNAc analogues (**31**, **82–91**, shown in ● *Table 1*D) have been reported.

In the continuing expansion of MOE, additional ManNAc analogues containing orthogonal functional groups for glycocalyx display include the aforementioned thiol groups of Neu5TGc and the azide groups of ManNAz (**33**) and the peracetylated counterpart Ac_4ManNAz (**34**) (● *Scheme 5*). The display of azides via the analogues shown in ● *Table 1*E is particularly attractive from a chemical standpoint because unlike ketones (that exist inside a cell) or thiols (that occur ubiquitously in a cell with the exception of the glycocalyx) these groups are completely abiotic and can be exploited for orthogonal reactions intracellularly, on the cell surface, or even in the extracellular matrix. To recap, the display of ketones, azides, or thiols is fundamentally different than alkyl chains or hydroxyl groups. While the latter groups have the remarkable ability to influence sialic acid-specific biological activities, the analogues containing functional groups not naturally present in sugars install new chemical reactivity onto the cell surface.



Scheme 4
Synthesis of ketone-derivatized ManNAc analogues

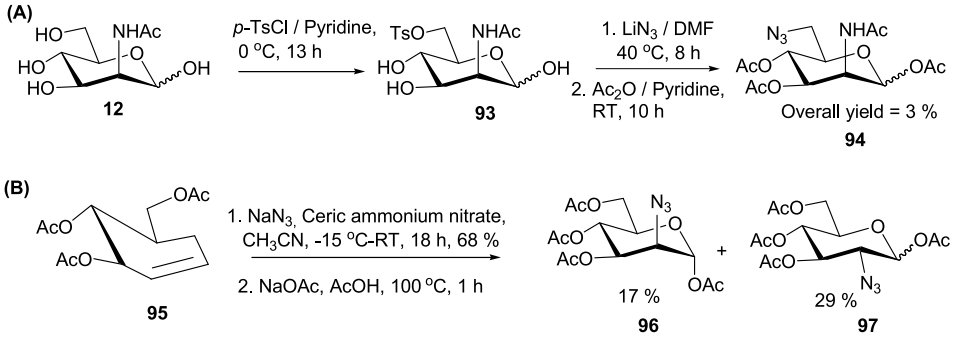


Scheme 5
Synthesis of Ac_4ManNAz (34)

3.1.7 Structural and Biochemical Considerations that Determine Flux Through the Pathway

As briefly discussed above, and as described in more detail in our publications where a panel of analogues with ketone-containing *N*-acyl groups that varied in length or steric bulk were evaluated [47,63], the sialic acid pathway has an impressive but ultimately limited ability to accommodate non-natural metabolic intermediates. As a consequence of these limitations, analogue design requires careful consideration; to illustrate this point, when installing orthogonal functional groups, azide groups at the *N*-acyl position of ManNAc (i. e., ManNAz, **33**) are more efficiently accommodated by the pathway than the ketone groups of ManNLev (**14**). Another critical aspect of analogue design is that the functional groups can only be accommodated at certain positions on the monosaccharide backbone. This point is illustrated by comparing the effects of an azide group attached to the C-6 carbon of ManNAc or directly appended to the C-2 carbon instead of being displayed on the *N*-acyl chain of **33**. The two former compounds, 6-azido ManNAc (**94**) and 2-azido mannose (**96**), were synthesized by Saxon and coworkers as shown in *Scheme 6*, A and B, respectively [23] and resulted in greatly reduced or negligible surface display of azide groups compared to **33** [23].

An understanding of the biochemical conversions imposed on each metabolite along a biosynthetic pathway can lend insight into the mechanism of inhibition and suggest alternative strategies to bypass problematic steps. For example, after UDP-GlcNAc is converted to ManNAc, phosphate is installed on the C-6 hydroxyl group by GNE. Analogues such as

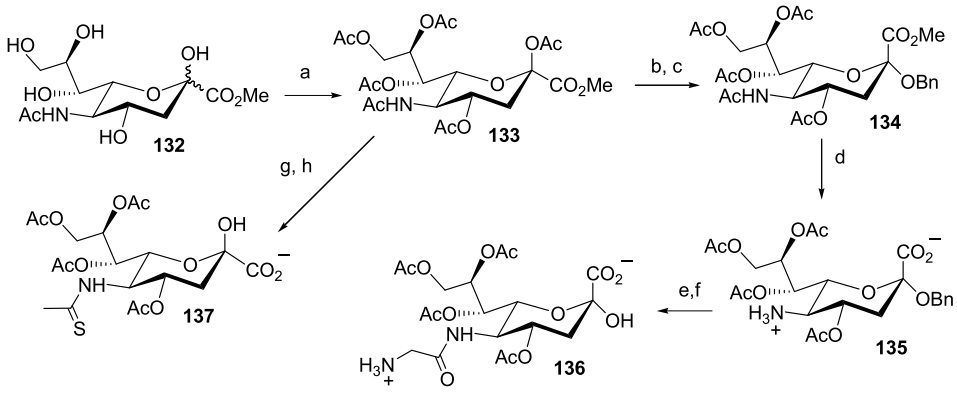


Scheme 6
Synthesis of C-6 and C-2 azide-modified mannose derivatives

6-azido ManNAc cannot undergo this reaction because of a lack of a C-6 hydroxyl, and combined with the enhanced conversion of phosphorylated intermediates to sialic acid (non-phosphorylated analogs have ~30-fold lower conversion rates to sialic acid [105]), it is not surprising that overall flux of such compounds through the pathway is low. Of note, Jacobs and coworkers showed that even when the C-6 hydroxyl is present, phosphorylation of this group remains the rate limiting step of the pathway [63]. Identification of ManNAc 6-kinase processing as a bottleneck for the sialic acid pathway provides options, discussed next and outlined in *Fig. 4*, for intercepting the pathway at other points to increase overall flux.

3.1.8 Sialic Acid Analogues—Bypassing Metabolic Pathway Bottlenecks

Once early steps in sialic acid biosynthesis were implicated as rate-limiting bottlenecks in the metabolic pathway, an obvious strategy to increase the flux of recalcitrant analogues was to intercept the pathway downstream of these chokepoints. Taking advantage of the ability of sialyltransferases to utilize CMP-9-fluoresceinyl-Neu5Ac (**103**) [106] and other modified CMP-Neu5Ac groups including those with photoactivatable appendages (*Table 2*) [107,108,109,110], interception of the pathway at the nucleotide sugar stage (*Fig. 4*) allowed highly permissive substitutions to be made in living cells [111,112]. Importantly, the use of CMP-sialic acid analogues bypassed the C-6 phosphorylation requirement of ManNAc analogues, allowing metabolic engineering of the C-9 position of sialic acids. A serious drawback to the use of nucleotide sugars for MOE was the membrane impermeability of these compounds that limited initial uptake into a cell as well as the subsequent subcellular localization to the lumen of the ER or Golgi where glycosyltransferases reside. The requirement for Golgi localization necessitated the use of permeabilized cells for experiments with non-natural CMP-sialic acids [111,112], which severely compromised cell viability and precluded in vivo applications.

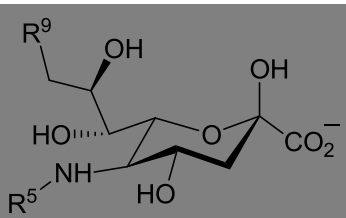


(a) Ac_2O / pyridine, 4 °C, 24 h, 91 % (b) Dry HCl, 4 °C, 12 h (c) Benzyl alcohol, 3 A Molecular seives RT, 12 h, 84 % (d) Aq MeOH, $Ba(OH)_2$, 18 h, 85 % (e) *N*-Z-glycine nitrophenyl ester, dry DMF, 93 % (f) Aq MeOH, PdO, H_2, 4h, 95 % (g) Sat aq $Ba(OH)_2$, 16 h, 31 % (h) Et_3N, Methyl dithioacetate, 0 °C-RT, 12 h, 86 %

Scheme 7
Synthesis of C-5 modified sialic acids

To avoid the difficulties inherent in using CMP-sialic acids for MOE, sialic acid analogues that can be synthesized with modifications at the C-5 (› *Scheme 7*, [57]) or other (› *Table 3*) positions, were tested. Because sialic acids including both **9** [113] and **37** [114] are efficiently taken up by cells [56], the introduction of analogues at the sialic acid stage of the pathway (see › *Fig. 4*) offered the dual advantages of bypassing rate limiting steps early in the pathway and avoiding the need to permeabilize cells for uptake. Illustrating the metabolic advantages of sialic acid analogues a specific example, the ability of Ac_4ManNLev (**87**) to inhibit sialic acid production in human cells [47] implied that the use of keto sialic acids would result in a higher

Table 3
Sialic acid analogues used in MOE



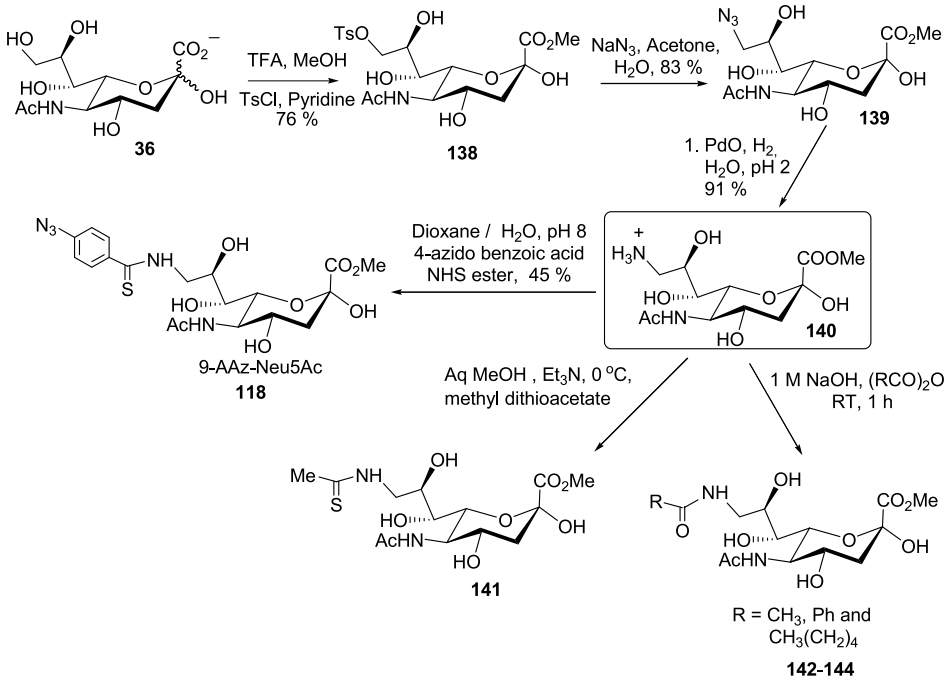
Sialic acid analogues

	R^9	R^5	Name	No	Ref
A	OH	$COCH_3$	Neu5Ac	**9**	[56]
	H	$COCH_3$	9-deoxy-Neu5Ac	**108**	[56]
	NH_2	$COCH_3$	9-amino-Neu5Ac	**109**	[56]
	$NHCOCH_3$	$COCH_3$	9-acetamido-Neu5Ac	**110**	[56]
	$NHCOCH_2NH_2$	$COCH_3$	9-N-Gly-Neu5Ac	**111**	[56]
	$NHCO(CH_2)_2COOH$	$COCH_3$	9-N-Succ-Neu5Ac	**112**	[56]
	I	$COCH_3$	9-Iodo-Neu5Ac	**113**	[56]
	SH	$COCH_3$	9-Thio-Neu5Ac	**114**	[56]
	SCH_3	$COCH_3$	9-SCH_3-Neu5Ac	**115**	[56]
	SO_2CH_3	$COCH_3$	9-SO_2CH_3-Neu5Ac	**116**	[56]
	$NHCOPhN_3$	$COCH_3$	9-AAz-Neu5Ac	**117**	[116]
B	OH	$COCH_2F$	5-N-Fluoroac-Neu	**118**	[56]
	OH	$COCF_3$,	5-N-Trifluoroac-Neu	**119**	[56]
	OH	$COCH_2NH_2$	5-N-Gly-Neu	**120**	[56]
	OH	$CO(CH_2)_2COOH$,	5-N-Succ-Neu	**121**	[56]
	OH	$CSCH_3$	5-N-thioac-Neu	**122**	[56]
	OH	$CO(CH_2)_2COCH_3$	Sia5Lev	**32**	[23]
	OH	$CO(CH_2)_3COCH_3$	Sia5OxoHex	**123**	[23]
	OH	$CO(CH_2)_4COCH_3$	Sia5OxoHept	**124**	[23]
	OH	$COCH_2N_3$	Sia5Az	**125**	[23]
	OH	$COCH_2PhN_3$	Sia5PhAz	**126**	[23]
	OH	$COCH_2OH$	Neu5Gc	**37**	[28]
	OH	$COCH_2CH_3$	Sia5Prop	**29**	[29]
	OH	$CO(CH_2)_2CH_3$	Sia5But	**127**	[29]
	OH	$CO(CH_2)_3CH_3$	Sia5Pent	**128**	[29]
	OH	$CO(CH_2)_4CH_3$	Sia5Hex	**129**	[29]
	OH	$CO(CH_2)_5CH_3$	Sia5Hept	**130**	[29]
	OH	$CO(CH_2)_6CH_3$	Sia5Oct	**131**	[29]

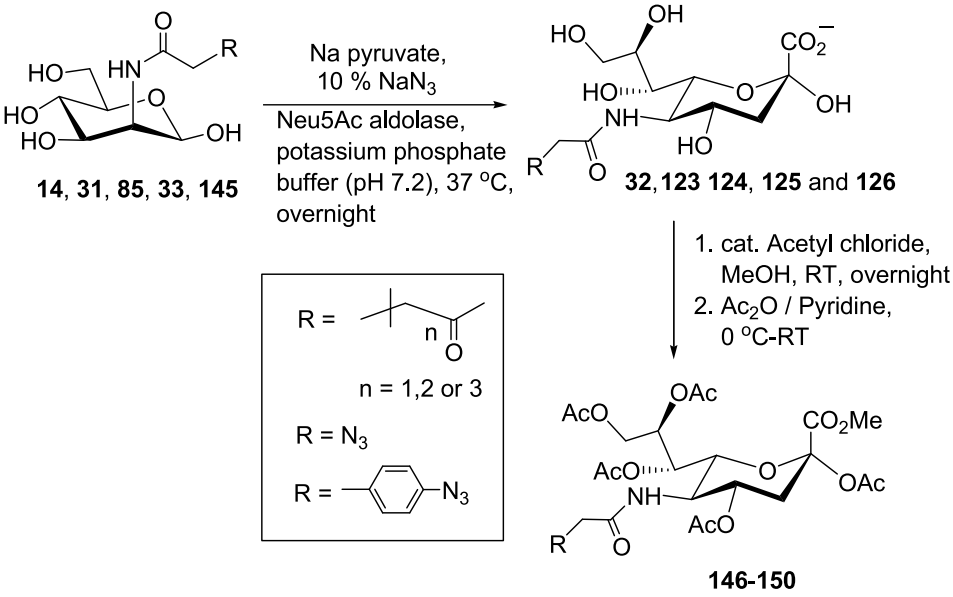
level of cell-surface ketones than the corresponding keto mannosamine analogues. Upon testing of the levulinoyl analogues, the sialic acid derivatives **32**, **123**, and **124** produced three times more cell-surface ketones than the corresponding ManNAc analogues **14**, **31**, and **84**. In contrast to the trend observed with these ketone derivatives, the azido-bearing mannosamine and sialic acid compounds (**125** and **126**) showed similar metabolic efficiency. This finding might derive from the fact that ManNAz is processed by sialoside biosynthetic enzymes considerably more efficiently than ManNLev [23] thereby masking any further gains realized by pathway interception at the sialic acid stage. Finally, even though sialic acid analogues are taken up by cells with reasonable efficiency [56], peracetylated sialic acids are utilized by cells between eight- and 20-fold more efficiently than their unprotected counterparts [115], suggesting that SCFA-derivatization (*Sect. 2.3*) is valuable for MOE experiments with this class of compounds.

3.1.9 ManNAc vs. Sialic Acid Analogues—Practical Considerations Prefer ManNAc Usage

As discussed above, the interception of the sialic acid pathway at the CMP-sialic acid stage offers tremendous versatility for functional group incorporation but is not compatible with living cells. By contrast, sialic acid analogues are efficiently taken up by cells and generally offer superior flux compared to the corresponding ManNAc analogues (although in some cases, such



Scheme 8
Synthesis of C-9 modified sialic acids



Scheme 9
Synthesis of modified sialic acids by aldolase-catalyzed condensation

as the azido derivatives, rates of metabolic processing are similar). Despite the apparent superiority of sialic acid analogues, a perusal of the literature reveals that ManNAc analogues have been used more often in MOE experiments. One answer for this apparent disparity between theory and practice is the ready availability and facile chemical synthesis of ManNAc analogs (as shown by examples given in Schemes 1 to 6) compared to the relatively difficult de novo chemical synthesis of the corresponding sialic acids.

In some cases, the direct synthesis of sialic acid analogs, such as that of 9-AAz-Neu5Ac (**117**) is fairly straight-forward (*Scheme 8*) [116]. More often, however, the synthesis of sialic acids is laborious, as exemplified by the 16-step synthesis required to produce modified forms of this sugar from the readily available starting material D-Glc [117]. Not surprisingly, the difficulty of purely synthetic approaches spurred efforts to obtain sialic acid analogues by neuraminic acid (Neu5Ac) aldolase-catalyzed condensation of pyruvate and the corresponding ManNAc analogue (*Scheme 9*) [118]; many of the mannosamine derivatives shown in *Table 1*D and *Table 1*E [119] were derived by this route. A difficulty with chemoenzymatic syntheses of sialic acid analogues, however, is the limited substrate permissivity of Neu5Ac aldolase as illustrated in attempts to prepare nitrogen isosteres, such as **152**, of Neu5Ac [117] from a series of C-3 nitrogen-functionalized ManNAc derivatives (e. g., **151**) [120]. In these experiments, as well as in other studies [121], C-3-substituted ManNAc derivatives did not serve as substrates for Neu5Ac aldolase. In summary, unless sialic acid analogues are required for organisms such as *H. ducreyi* that do not convert ManNAc into sialic acids [28,29], the modest gain in metabolic efficiency generally does not justify the extra effort required for the production of modified sialic acids.

3.2 Metabolic Precursors for the Substitution of GalNAc in Mucins and GAGs

3.2.1 Overview of Mucin Biology

N-Acetyl-D-galactosamine (GalNAc, **6**) is abundant in many natural glycoconjugates including those from higher eukaryotes where this sugar is attached to the hydroxyl group of Ser/Thr side chains by α-linkage and provides the foundation for 'mucin-type' glycoproteins. 'Muco-proteins' and 'mucopolysaccharides' have been studied since the early 1800s when the term 'mucin' was first used by Nicolas Theodore de Saussure in 1835 to describe substances isolated from mucus [122]. By 1865 Eichwald had determined that mucins were a combination of protein and carbohydrate with the sugar components often encompassing more than 50% of total mass. More recently, in 1969 Weissman and Hinrichsen reported that the carbohydrates of mucins were attached to the peptide by an α-GalNAc linkage by analyzing purified α-*N*-acetylgalactosaminidase from bovine liver [123]. Subsequent structural studies using atomic force microscopy (AFM) [124], light scattering [125], and nuclear magnetic resonance (NMR) spectroscopy [126] have demonstrated that clustered *O*-linked glycosylation results in extended protein structures up to hundreds of nanometers in length [127]. In the past three decades, more than 150 different mucin-type *O*-linked glycans have been isolated and identified [128].

3.2.2 Metabolic Pathways Responsible for Replacement of GalNAc with Analogues

An ever-increasing number of biological processes are now attributed to *O*-linked glycans ranging from lymphocyte homing [129,130] to roles in diseases from Wiskott–Aldrich syndrome [127], to hematological disorders [131] and cancer [132]. Considering the biological importance of *O*-linked glycans, together with the limited specificity of lectins for binding the diverse array of *O*-linked glycan produced by mammalian cells, proteomic analysis of *O*-linked glycosylation remains difficult. Consequently, the metabolic incorporation of non-natural monosaccharides that bear unique functional group handles such as ketones or azides has gained tremendous attention as a method to analyze *O*-linked glycans. In order to accomplish this objective through MOE, a detailed knowledge of the metabolic pathways that process GalNAc, similar to the considerations discussed above for sialic acid modulation, is of value when planning the point of pathway interception.

Mucin-type *O*-linked glycoproteins result from the biosynthetic efforts of a family of GalNAc transferases (ppGalNAcTs) that use UDP-GalNAc as the nucleotide donor for attachment of the sugar to a protein substrates [133]. UDP-GalNAc can be obtained for use in glycan assembly through epimerization from UDP-GlcNAc as well as via salvage pathways (*Fig. 7*). Consequently, two options exist for exogenous supplementation with GalNAc-targeting analogues. The first, and poorer, option is to feed analogues into the hexosamine (see *Fig. 9*, below) or salvage pathways that lead to UDP-GlcNAc production. This alternative, however, may not be desired because UDP-GlcNAc has multiple metabolic fates in addition to epimerization to UDP-GalNAc which include sialic acid biosynthesis (*Sect. 3.2*) and *O*-GlcNAc protein modification (*Sect. 3.5*). The second option is to target GalNAc salvage pathways

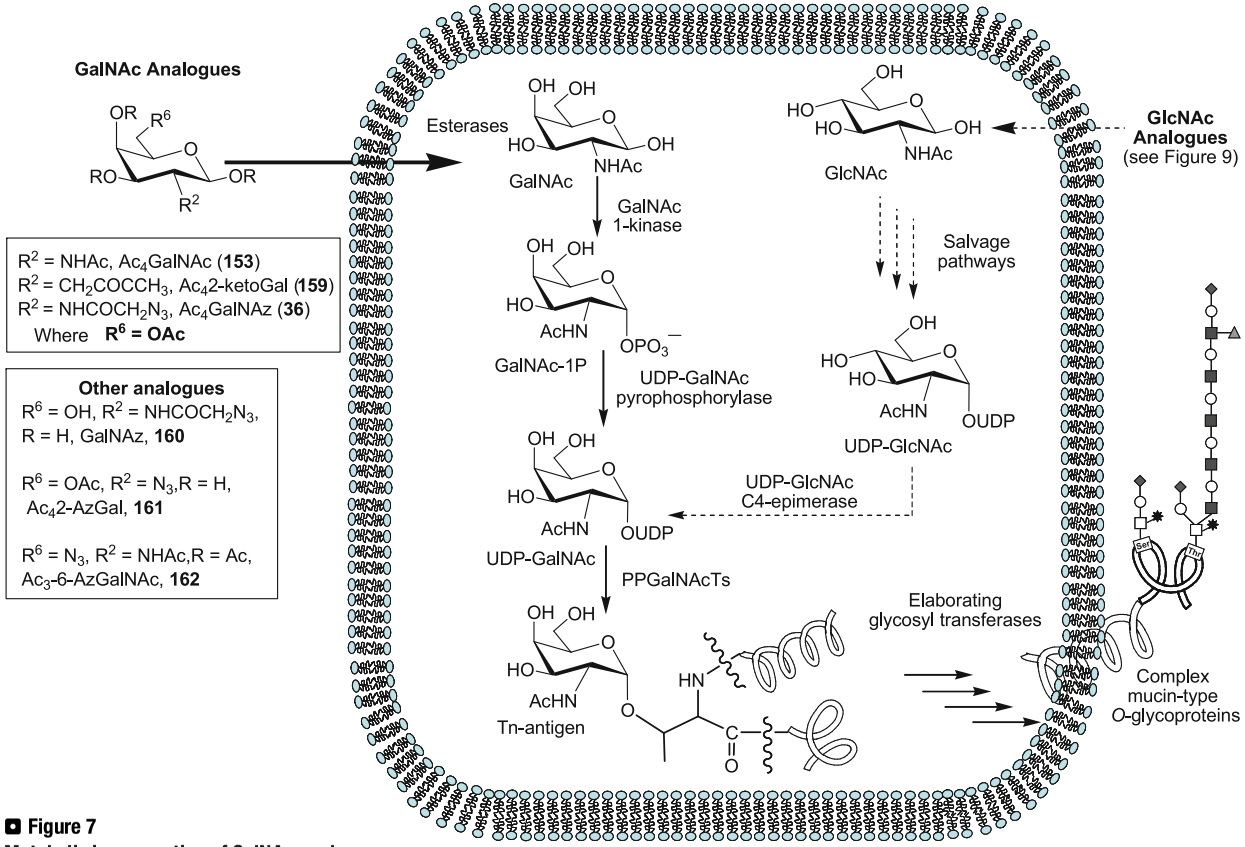
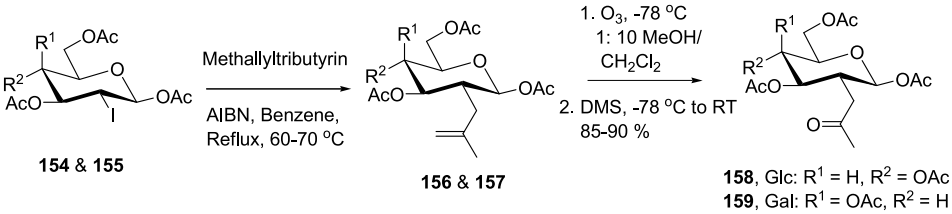


Figure 7
Metabolic incorporation of GalNAc analogues



Scheme 10
Synthesis of '2-ketosugars'

directly. Using this strategy, Hang and Bertozzi synthesized (*Scheme 10*) '2-ketosugars', which are C2-carbon isosterates of the 2-acetamidosugars of GlcNAc (**158**) and GalNAc (**159**), and demonstrated incorporation into *O*-linked glycans, and possibly into GalNAc-containing GAGs (*Fig. 1*) [26]. These analogues possessed a ketone group, similar to keto ManNAc analogues that target sialic acid for metabolic replacement (*Sect. 3.1.6*), that allows chemoselective reactions with aminooxy- or hydrazide-bearing reagents (*Sect. 4.2*). In addition, the GalNAc analogues discussed in the next section can be exploited for tagging mucin-type *O*-linked glycans with azides, which then can be selectively reacted with phosphine reagents for proteomic analysis (*Sect. 4.3*).

3.2.3 Optimizing GalNAc Analogue Structure

Although fewer studies have been done to optimize GalNAc analogue metabolism compared to the detailed exploration of the substrate specificity of the ManNAc pathway outlined above, a rudimentary indication of pathway permissivity has been gained from a comparison of azido GalNAc analogues. In particular, the metabolic incorporation of *N*-azidoacetyl-D-galactosamine (GalNAz, **160**) or the peracetylated counterpart 1,3,4,6-tetra-*O*-acetyl-*N*-azidoacetylgalactosamine (Ac_4GalNAz, **35**), into mucin-type *O*-linked glycans in a variety of mammalian cell lines has been demonstrated in vitro and in animals [22]. By comparison, 1,3,4,6-tetra-*O*-acetyl-2-azido-2-deoxy-D-galactose ($Ac_4$2-AzGal, **161**), and 1,3,4-tri-*O*-acetyl-6-azido-6-deoxy-*N*-acetyl-D-galactosamine (Ac_3-6-AzGalNAc, **162**) were synthesized and their metabolic incorporation was compared with the former compounds in CHO cells. In these studies, cells treated with **35** demonstrated a 30-fold higher fluorescence than untreated cells whereas **161**- and **162**-treated cells did not exhibit fluorescence above the level of untreated cells suggesting that these two analogues are not effective substrates for the GalNAc salvage pathway.

3.3 L-Fucose Analogues

3.3.1 Overview of Fucose Pathways and Biology

Similar to sialic acid, L-fucose is usually displayed as a terminal sugar in glycans allowing it to readily participate in cell–cell interactions and modulate cell motility and migration processes connected with fertilization, embryogenesis, lymphocyte trafficking, immune responses, and cancer metastasis [134,135,136]. To date, many of the specific roles of fucose have been traced

to the fact that it is a critical component of the Lewis X (Le^X) trisaccharide epitope as well as the tetrasaccharide sialyl Lewis X (sLe^X) [137]. Over-expression of sLe^X is characteristic of many cancers [138] and this structure is also a determinant of the selectin ligands that mediate leukocyte-endothelial interactions at sites of inflammation [139]. Glycoproteins displaying Le^X promote embryonic cell adhesion and have important roles in neurogenesis [134]. Fucose is also a terminal sugar of many tumor-associated antigens and the multitude of fucosylated antigens up-regulated in cancerous tissue provides an opportunity to exploit these glycans as potential cancer biomarkers [140]. On the basis of these, and many additional, biological activities, the investigation of fucosylated glycans in health and disease is of great interest and the development of specific tagging methods to visualize and quantify fucosylated glycoconjugates in situ has become a subject of keen interest [16,141,142,143].

3.3.2 Targeting Fucose for Metabolic Replacement

Akin to the incorporation of dietary Neu5Gc sialic acid into cellular sialoglycans that presaged the use of sialic acid analogues in MOE, the finding that the congenital disorder of glycosylation (CDG) leukocyte adhesion deficiency II (LADII) could be corrected by oral supplementation [144] provided a foundation for applying MOE to fucose. An analysis of glycosylation pathways responsible for fucose display indicates two possibilities for the introduction of L-fucose analogues into glycoconjugates (*Fig. 8*). First, mannose can be converted to GDP-D-mannose and subsequently to GDP-L-fucose, suggesting that mannose derivatives have the potential to serve as fucose precursors. This possibility, however, was discounted by the lack of surface display of the 2-azido form of mannose (**96**) discussed in *Sect. 3.1.7*. The second option was the direct supply of fucose analogues, an approach that has now been successfully demonstrated by two laboratories. In one set of experiments, Bertozzi's group synthesized fully acetylated 6-azido (**164**), 2-azido (**165**), and 4-azido (**166**) analogues of peracetylated fucose [145]. All three compounds were evaluated for incorporation into glycoproteins in the human T lymphoma cell line Jurkat by labeling the analogue-treated cells with phosphine-bearing FLAG peptide [146] followed by staining with a FITC-conjugated α-FLAG antibody and analysis by flow cytometry. In these experiments significant fluorescence was only observed after treatment with **164** while compounds **165** and **166** did not label cell surface glycans. In a second set of experiments reported by the Wong laboratory [147], the 6-azido analogue was successfully incorporated along with the alkyne-bearing analogue **167**; in these studies new fluorescent probes were utilized to detect surface display in a one-step procedure as described in more detail in *Sect. 4.6*.

3.4 GlcNAc Analogues as Tools to Study *O*-GlcNAc Protein Modification

3.4.1 Surface GlcNAc Residues Are Refractory to Metabolic Replacement

Of the three common hexosamines found in mammalian cells, analogues of ManNAc and GalNAc are readily incorporated into surface glycans (as discussed in *Sect. 3.2* and *Sect. 3.3*, respectively). By contrast, GlcNAc analogues are largely refractory to incorporation into glycoconjugates with GlcNProp (**168**) [17], GlcNLev, **169** [65], and GlcNAz (**170**) [23] all showing dramatically lower surface display than the corresponding mannosamines (*Fig. 9*).

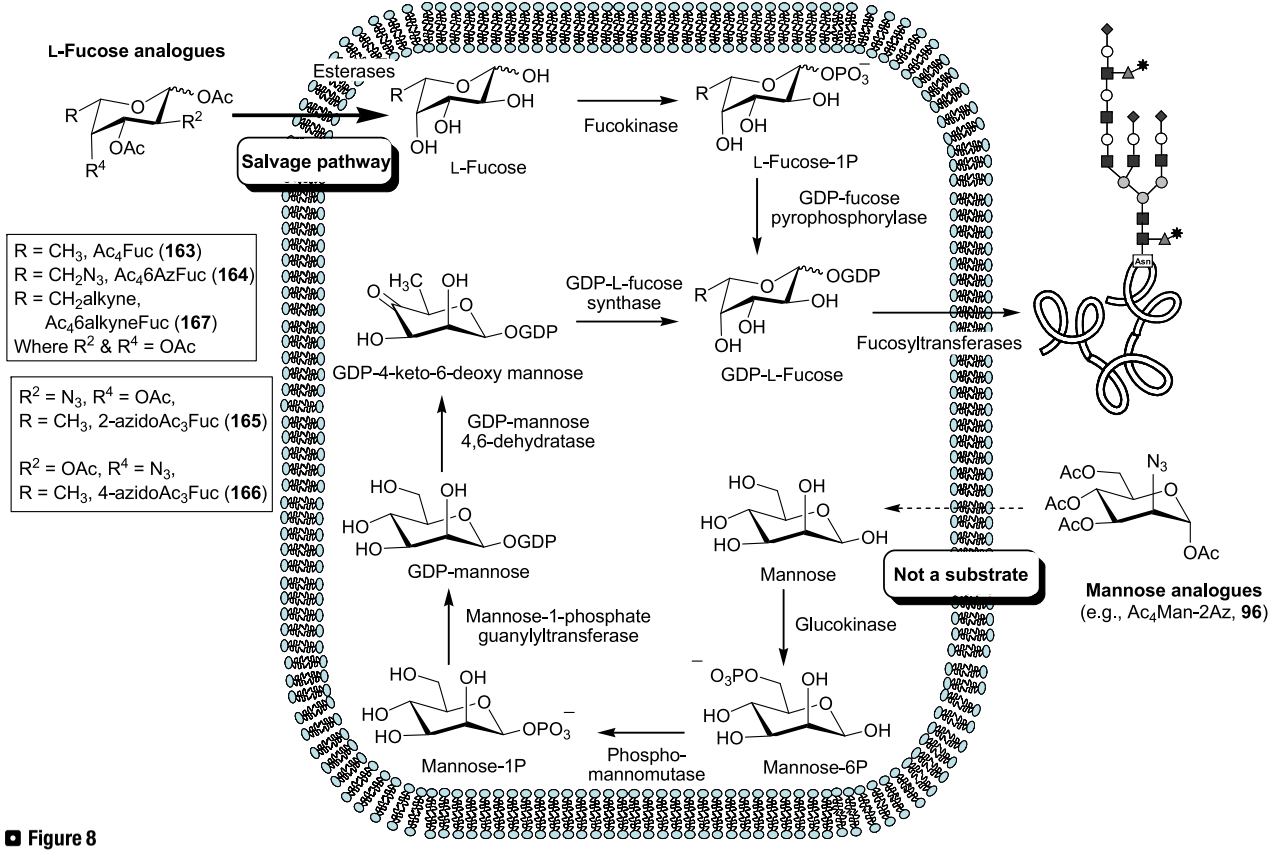


Figure 8
Metabolic incorporation of fucose analogues

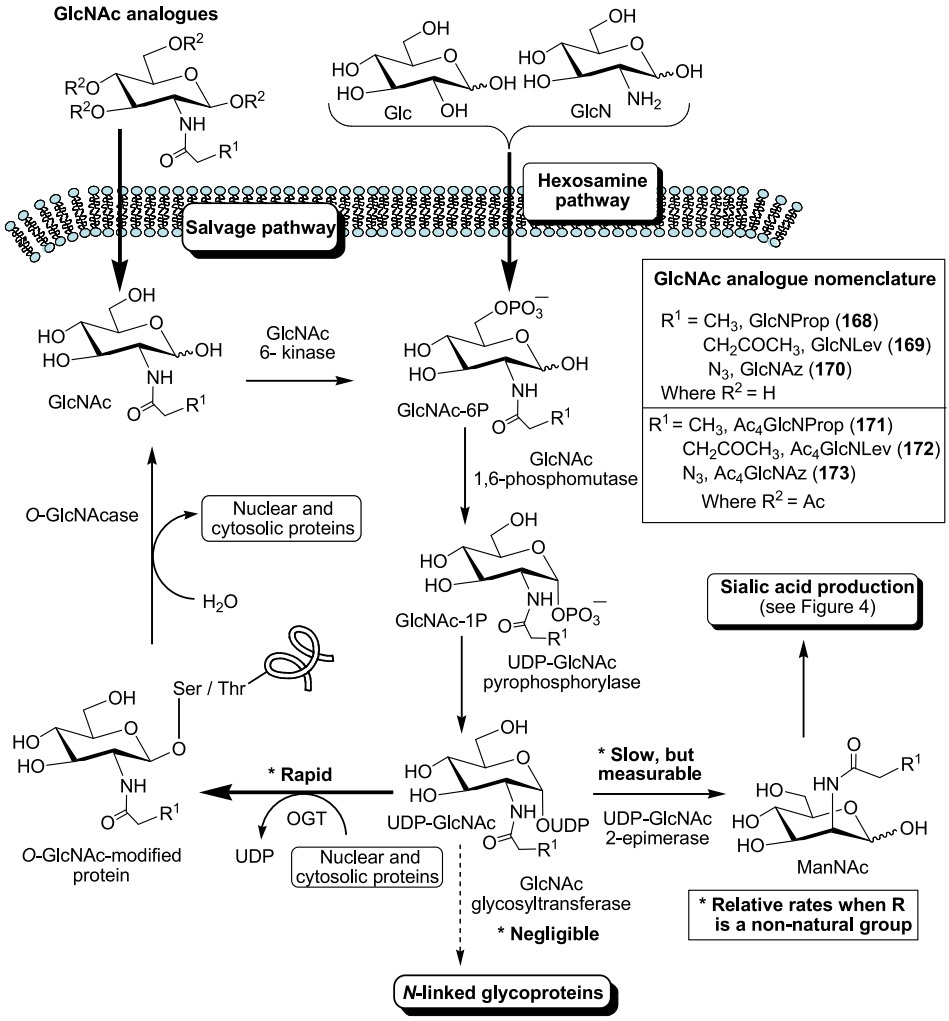


Figure 9
Metabolic incorporation of GlcNAc analogues

One possibility for the incorporation of GlcNAc analogues into glycans can be envisaged by considering the ability of the renin binding protein/GlcNAc 2-epimerase to interconvert GlcNAc to ManNAc [88] as shown in *Fig. 4*; in this case, the surface display of GlcNAc analogues would occur via incorporation into the sialic acid pathway. Evidence for surface display by this route has been shown for Ac_4GlcNAz (**173**), which was thought to be converted to ManNAz (**33**) after deacetylation followed by further conversion to Sia5Az (**125**) [23]. The relatively low level surface display of azide groups upon exposure to exogenous **173** compared to Ac_4ManNAz (**34**) suggested that the epimerization of **170** to ManNAz **33** was severely rate limiting. In theory, a second route to surface display of GlcNAc analogues is through the

salvage pathway after phosphorylation at the C-6 position [148] that, if amenable to non-natural substitutions, would produce UDP-GlcNAc analogues that could be incorporated into *N*-linked glycans directly or epimerized to UDP-GalNAc analogues for *O*-glycan display as diagrammed in ❯ *Fig. 9*.

The lack of incorporation of metabolites of **173** into cell surface glycoproteins via UDP-GlcNAc intermediates established that the GlcNAc salvage pathway was refractory to MOE [23]. Several explanations can be offered why a UDP-GlcNAc analogue such as UDP-GlcNAz (**174**) is not incorporated into surface glycans: (*i*) **174** may not be efficiently transported into the lumen of organelles of the secretory pathway, (*ii*) GlcNAc transferases may not tolerate the *N*-acyl modification of **174**, or (*iii*) **174** may be diverted into other pathways. Considering the latter possibility, UDP-GlcNAc can be converted to UDP-GalNAc or to ManNAc (❯ *Fig. 9*), thereby offering alternative routes to surface display. With the exception of possible conversion to ManNAc and glycan incorporation as sialosides, these options have been largely eliminated as efficient routes to the surface display of GlcNAc analogues. Interestingly, the surface inaccessibility of GlcNAc analogues has prompted the exploration of alternative uses for these MOE agents and they are now gaining acceptance as probes for *O*-GlcNAc protein modification. The significance of this emerging application is given below by briefly describing *O*-GlcNAc protein modification (❯ *Sect. 3.4.2*) and relevant analogue probes (❯ *Sect. 3.4.3*).

3.4.2 UDP-GlcNAc Is Used for *O*-Glycosylation of Cytosolic and Nuclear Proteins

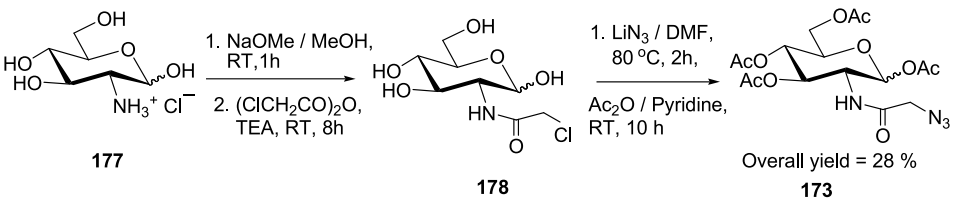
In 1984, Torres and Hart described a post-translational modification where Ser and Thr residues found in nuclear and cytoplasmic proteins were *O*-linked to GlcNAc through β-linkage [149,150]. *O*-GlcNAc was novel for two reasons: first, *O*-GlcNAc was not elongated into a more complex structure, with the possible exception of plant nuclear pore proteins [151]; second, it had a nucleocytoplasmic distribution, whereas 'traditional' glycoproteins had been localized to the cell surface and topologically equivalent intracellular compartments, such as the lumens of the endoplasmic reticulum and Golgi apparatus [152]. *O*-GlcNAc shares many similarities with protein phosphorylation in cellular regulation; for example, like phosphorylation, *O*-GlcNAc is highly dynamic with rapid cycling in response to cellular signals or cellular stages and often occurs at the same sites on the protein backbone that are modified by protein kinases [14]. Unlike protein phosphorylation, however, where 650 genetically distinct enzymes regulate the addition and removal of phosphate, just two catalytic polypeptides catalyze the turnover of *O*-GlcNAc; the uridine diphospho-*N*-acetyl-D-glucosamine: peptide β-*N*-acetylglucosaminyl transferase (OGT) adds the sugar residue to proteins and *O*-β-*N*-acetylglucosaminidase hexosaminidase (*O*-GlcNAcase) reverses the modification [15] (❯ *Fig. 9*).

In most organisms, *O*-GlcNAc protein modification is vital for life. It occurs ubiquitously across species with only the yeasts *Saccharomyces* and *Pombe* lacking biochemical or genetic evidence for *O*-GlcNAc protein modification [15]. Clues to *O*-GlcNAc function come from estimates that hundreds of proteins are subject to this modification, with substrates including cytoskeletal proteins, nuclear pore proteins, chromatin associated proteins, RNA polymerase II (RNA Pol II) and its transcription factors, hnRNPs, proto-oncogenes, tumor suppressors, hormone receptors, phosphatases, and kinases [14]. In essence, *O*-GlcNAc acts as a glucose

sensor, through the conversion of glucose to glucosamine and then to UDP-GlcNAc via the hexosamine biosynthetic pathway [13], that tune the activities of this diverse set of cellular players to environment conditions of the cell.

3.4.3 Chemically Tagged GlcNAc Analogues Can Be Used to Probe *O*-GlcNAc Modification

Despite numerous studies implicating *O*-GlcNAc modification of proteins as a modulator of cellular processes of critical importance to diseases such as diabetes [153], many aspects of *O*-GlcNAc modification including the complete repertoire of modified proteins and the specific sites of modification remain largely unknown. Indeed, although hundreds or even thousands of proteins are *O*-GlcNAc modified, relatively few have been identified and fewer still have had the specific glycosylated residues identified. Furthermore, no consensus sequence directing the *O*-GlcNAc modification of proteins to a particular Ser or Thr residue has been established on the basis of known modification sites. To help overcome these limitations, MOE methods based on bioorthogonal non-natural GlcNAc analogues are now proving to be important tools to assist in the characterization of *O*-GlcNAc and to study these protein modifications in cells. In particular, based on observations that compounds such as *p*-nitrophenyl 2-acetamido-2-deoxy-β-D-glucopyranoside (*p*NP-GlcNAc, **175**) act as a substrate for OGT with only a 35% decrease in the value of the apparent second-order rate constant and a 2-fold decrease in the value of *V*max relative to the native substrate, Vocadlo and coworkers synthesized peracetylated GlcNAz (Ac_4GlcNAz, **173**, ❯ *Scheme 11*) and the control compound *p*-nitrophenyl 2-azidoacetamido-2-deoxy-β-D-glucopyranoside (PNP-GlcNAz, **176**) [27]. The ability of OGT to tolerate the azido group remarkably well allowed the identification of specific proteins by the chemoselective ligation strategies discussed in ❯ *Sect. 4.3.2* that are undergoing continued development in a variety of 'tagging via-substrate' (TAS) methods [141,154,155,156] that complement more conventional proteomics tagging schemes used to identify *O*-GlcNAc modifications [157].



Scheme 11
Synthesis of GlcNAc analogues

3.5 MOE Offers Opportunities for Inhibition of Glycan Biosynthesis

MOE technology—which has been described so far as a method to provide alternative flux options through metabolic pathways and incorporate new sugar residues into cellular glycans—can also provide opportunities to *inhibit* glycan biosynthesis via three distinct mechanisms. The first is 'conventional' inhibition, wherein the activities of one or more path-

way elements are reduced. The second is the incorporation of non-natural sugar residues into a nascent glycan structure, which then terminates or retards subsequent biosynthesis. Finally, non-natural sugars can function as 'decoys' for synthesis. Each of these strategies is discussed next.

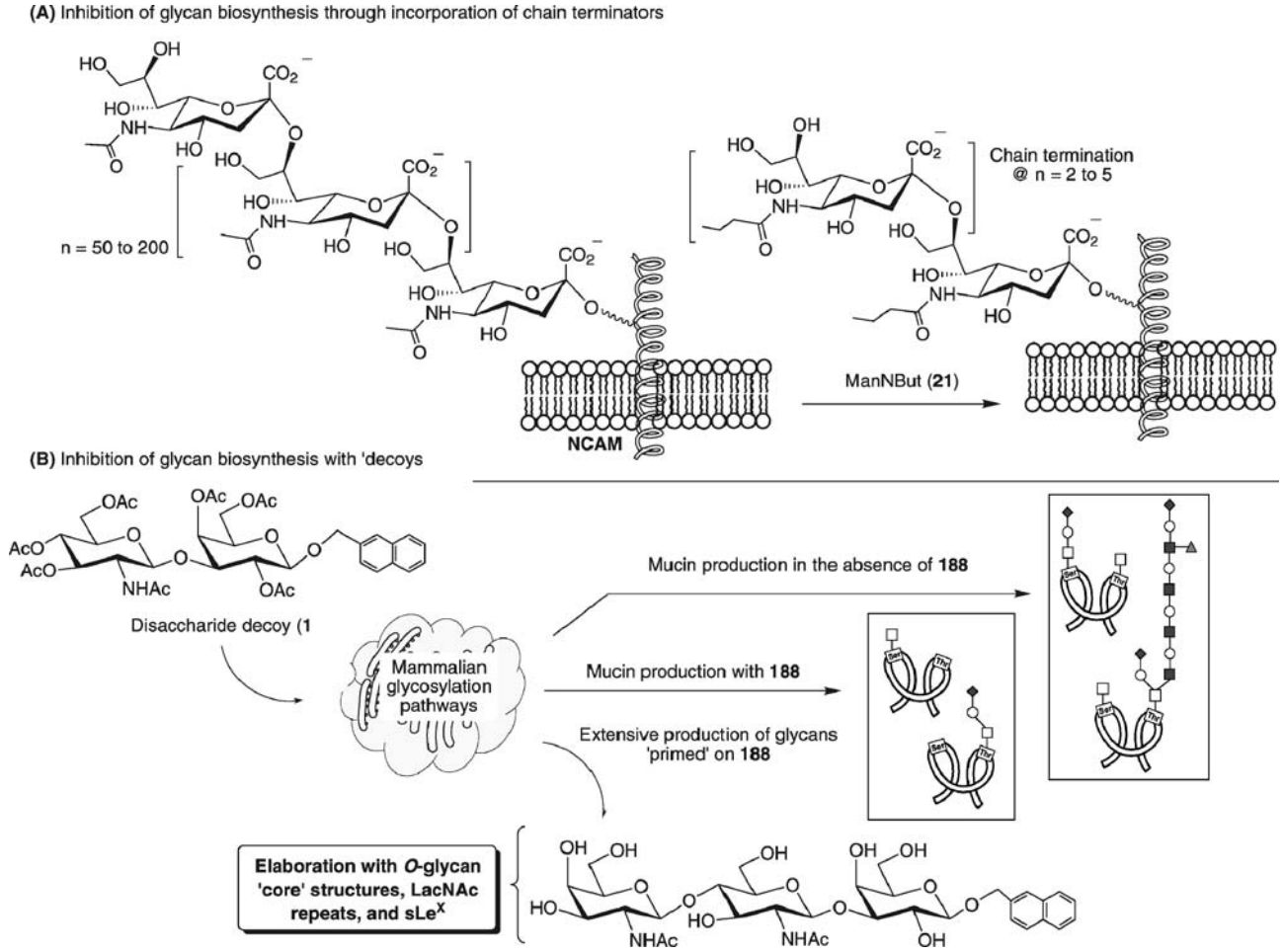
3.5.1 MOE Grew Out of Attempts to Inhibit Glycosylation Pathways

Taking a historical perspective, one foundation of today's MOE technologies grew from the efforts by Brossmer and colleagues about two decades ago, as discussed in *Sect. 3.1.8*, to exploit the substrate permissivity of sialyltransferases. Interestingly, another basis for MOE had already been established in living cells at that time from efforts to inhibit glycosylation pathways that date back to the 1970s. Even then it was clear that abnormal glycan production, most often over-expression, was associated with the transformed and malignant cells, leading to efforts to block oligosaccharide biosynthesis as a potential cancer therapy. Early attempts to use fluorine-modified and acetylated ManNAc analogues to block sialic acid production, however, resulted in the unplanned metabolic conversion of these non-natural compounds into the corresponding sialic acids [158,159]. As described in *Sect. 3.1*, Reutter and colleagues built on these findings by deliberately exploiting the substrate permissivity of sialic acid enzymes to achieve surface display on non-natural sugars [17]. The ketone-containing analog ManNLev (**14**) however, which was the prototype analogue used to install new chemical functionalities onto the cell surface [20], actually inhibits flux through the sialic acid pathway when used in its peracetylated form [47]. On the basis of these connections between MOE technologies and efforts to inhibit glycosylation, we will briefly describe ongoing developments in this area.

3.5.2 Inhibition of PSA via Metabolic Incorporation of ManNAc Analogues

As mentioned above, the Sartorelli group attempted sialic acid pathway inhibition over two decades ago by the conventional strategy of substituting protons and hydroxyl groups of the metabolic substrate, in this case ManNAc, with fluorine [158,159]. Even if these original attempts to inhibit flux through the sialic acid pathway had met their objective, they may not have translated into useful cancer therapies due to the many critical roles of sialic acid in healthy cells that would have been affected adversely. Consequently, subsequent efforts have attempted to achieve highly targeted effects on glycan production, not through inhibition of flux through the linear portion of the sialic acid pathway (*Fig. 4*), but rather by targeting a subset of sialyltransferases (*Fig. 6*). The strategy of targeting specific glycan epitopes is a daunting task but is reaching fruition for the fetal-oncogenic marker polysialic acid (PSA). In this case, incorporation of a run of up to five consecutive Sia5But (**127**) or Sia5Pent (**128**) residues into a nascent PSA chain as a consequence of ManNBut (**21**) or ManNPent (**22**) uptake, respectively, blocks further chain extension [68,96] (*Fig. 10A*). These analogues

Figure 10
Inhibition of glycan production by MOE. (A) PSA production is terminated upon the sequential incorporation of Sia5But (shown) or Sia5Pent residues. (B) 'Decoys' function as inhibitors of glycosylation by hijacking the metabolic machinery resulting in glycan synthesis on these non-natural initiators rather than on the normal complement of cellular proteins or lipids


(A) Inhibition of glycan biosynthesis through incorporation of chain terminators
n = 50 to 200
Chain termination
@ n = 2 to 5
ManNBut (21)
NCAM
(B) Inhibition of glycan biosynthesis with 'decoys
Disaccharide decoy (1
Mammalian
glycosylation
pathways
Mucin production in the absence of 188
Mucin production with 188
Extensive production of glycans
'primed' on 188
Elaboration with O-glycan
'core' structures, LacNAc
repeats, and sLe^x

have already served as important research tools for the discovery of the function of PSA in neuronal development [97,98]; in the future ablation of PSA expression in cancer [160] may prove equally significant.

3.5.3 Inhibition with Decoys

The concept of a 'decoy' inhibitor shares the basic similarity with MOE in that a non-natural sugar derivative introduced to a cell intercepts the glycosylation process. Glycan decoys were first demonstrated in 1973 when Okayama and coworkers showed that β-D-xylosides act as substrates for the assembly of glycosaminoglycans and thereby stimulate the synthesis of free GAG [161]. Subsequently, D-xylose, β-D-xylosides, and β-D-galactosides have been used as artificial initiators of chondroitin sulfate synthesis that enable polysaccharide-chain synthesis to be studied as an event separate from the synthesis of intact proteoglycan [162]. As a consequence of the biosynthetic efforts of a cell being diverted to non-naturally-primed glycan synthesis, these 'decoys' competitively inhibit the biosynthesis of glycosaminoglycan chains on their usual substrates, the endogenous proteoglycan core proteins (*Fig. 10B*). Interestingly, α- and β-xylosides can also prime the production of GM3-like compounds and thereby inhibit glycolipid biosynthesis [163]. In a similar way, aryl-*N*-acetyl-α-galactosaminides were later discovered to prime mucin-like oligosaccharides and alter the assembly of *O*-linked oligosaccharides on glycoproteins resulting in 1-benzyl-2-acetamido-2-deoxy-α-D-galactopyranoside (GalNAc-α-*O*-Bn, **179**) now being arguably the most well-known and most-utilized inhibitor of glycosylation [164].

Recently, efforts in decoy development have turned to the inhibition of specific glycoepitopes, such as the sLe^X ligand involved in key steps of inflammation and tumor metastasis. In the past, GalNAc-α-*O*-Bn (**179**) was used to inhibit the sialic acid or fucose components of sLe^X but, while this compound had the desired effect of suppressing both α-2,3-sialylation and α-1,3-fucosylation, it also dramatically increased α-2,6-sialylation and led to massive accumulations of truncated oligosaccharides [165] indicating an unacceptable lack of specificity. In subsequent experiments, increased specificity was achieved through the use of peracetylated disaccharides including Gal-β-1,3-GlcNAc (**180** and **181**), Gal-β-1,4-GlcNAc (**182** and **183**), GlcNAc-β-1,3-Gal (**184** and **185**), and Gal-β-1,3-GalNAc (**186** and **187**) with naphthyl or naphthyl-methene aglycons [41]. Later, peracetylated GlcNAc-β-1–3Gal-naphthalenemethanol (GlcNAcβ1–3Gal-NM, **188**) was found to have even greater potency and will inhibit sLe^X-dependent cell adhesion [43]. Thus, peracetylated disaccharides have distinct advantages over monosaccharide primers such as **179**, including greater specificity and efficacy at lower concentration that derives from the ability of these derivatives to efficiently cross the hydrophobic barrier of the cell and Golgi membranes and down regulate the expression of sLe^X ligands. Moreover, the ability of these compounds to suppress cellular adhesion events implicated in chronic inflammation and metastatic cancer eventually could be useful for treating immune disorders and ameliorating tumor progression.

4 Cell Surface Chemistry—Chemoselective Ligation Reactions

Section 3 of this chapter described strategies based on the metabolic engineering of several monosaccharides to achieve cell surface display of chemical functional groups including

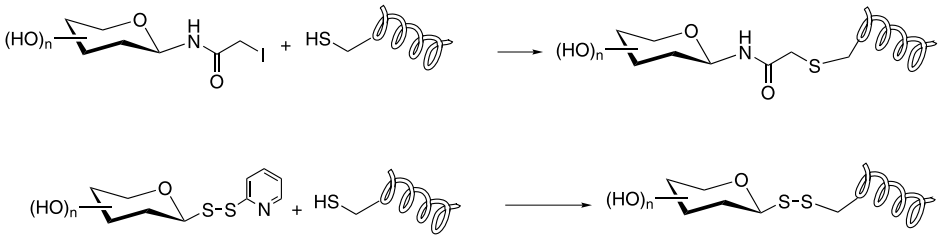
ketones, azides, and thiols that are not normally found in the glycocalyx. Once situated in this novel cellular location, these functionalities can be exploited via the chemoselective ligation strategies described here in ❯ *Sect. 4* for an ever-expanding complement of purposes. On the basis of the growing number of chemoselective coupling strategies, including a recent proliferation of bioorthogonal reactions designated for azides [166], this discussion is intended to be a guide to assist the non-specialist in selecting appropriate ligation chemistry.

4.1 Requirements of Chemoselective Ligation, Previous Applications, and Use in MOE

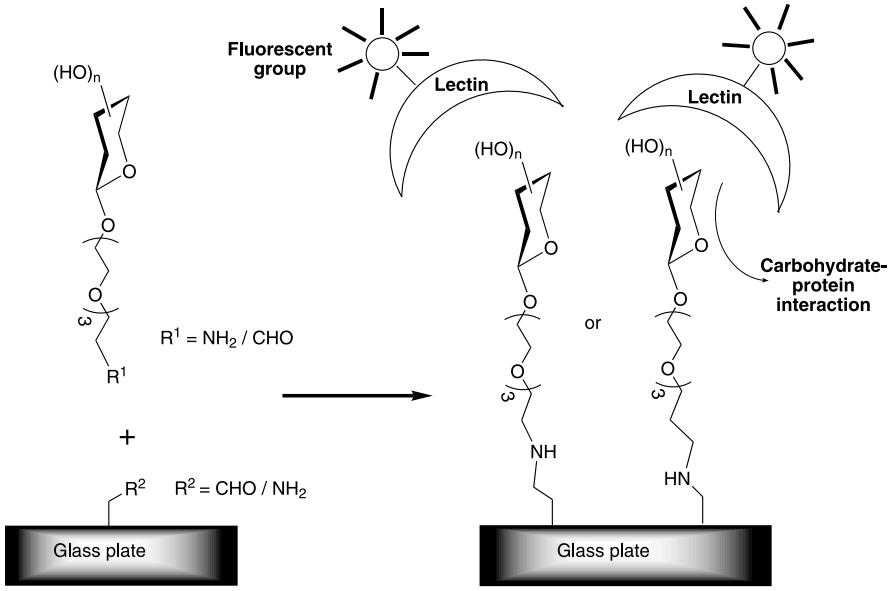
Chemoselective ligation reactions are designed to allow the selective reaction of two complementary chemical functional groups even in a complex milieu of a biological system such as a living cell [76,167]. Ideally, the two reactive partners are abiotic, undergo non-reversible covalent bond formation under physiological conditions, and recognize only each other while remaining impervious to their cellular surroundings [60]. An early example of ketone- and aldehyde-based chemoselective ligation is the work of Rideout and colleagues who reported the self-assembly of inactive prodrugs into cytotoxic agents inside cells [168,169].

The chemoselective ligation strategy was soon applied to the synthesis of neoglycoconjugates because bioorthogonal coupling provided an attractive alternative to previous methods. For example, in a common synthetic strategy used for glycopeptides, sugars were attached to the peptide backbone during solid phase synthesis (SPPS) by means of glycosyl amino acid building blocks. This 'linear' synthetic approach was laborious, required extensive use of protecting groups, and only allowed the preparation of one glycoform per synthesis. A second, convergent approach that suffered from similar limitations was based on either the enzymatic elongation of the oligosaccharide chain of a pre-synthesized glycopeptide or glycosyl amino acid or on the conjugation of a fully elaborated, complex saccharide to short synthetic peptides. As reviewed in detail elsewhere [170,171,172,173], chemoselective ligation reactions provide attractive alternative strategies that overcomes several synthetic stumbling blocks to neoglycoconjugate synthesis. An example of this strategy is shown in ❯ *Fig. 11A* where neoglycoconjugates are synthesized through the coupling of a thiol-derivatized peptide to thiol-specific reactants conjugated to glycans. Similar technology is also proving to be of great utility in the construction of oligosaccharide microarrays (❯ *Fig. 11B*), which are finding various uses in emerging high-throughput 'glycomics' technologies [16,142,174,175].

To date, several chemoselective coupling strategies have been exploited in MOE. In this chapter five different bioorthogonal ligation reactions that have been applied to surface carbohydrates on living cells or in animals will be discussed briefly. The first one is the reaction of carbonyl groups (ketones or aldehydes) with strong nucleophiles, the second is the addition of sulfur nucleophiles to other electrophiles, the third is the modified Staudinger ligation of azides with phosphines, the fourth is photo-activated functional groups, and finally, the fifth example is provided by the cycloaddition between diene and dienophiles or the (3+2) cycloaddition between azides and alkynes. In addition to the chemistry involved, a sampling of applications will be given to provide the reader with a small window into the enormous potential of this emerging technology to modulate biology.



(A) Chemoselective ligation in neoglyconjugate synthesis



(B) Chemoselective ligation for carbohydrate microarray analysis of protein-carbohydrate interactions

Figure 11
Carbohydrate-based examples of chemoselective ligation reactions

4.2 Glycan-Displayed Ketones and Aldehydes: Establishing Chemoselective Ligation for MOE

4.2.1 ManNLev, keto2Gal, and GlcNLev are Ketone-Derivatized Metabolic Precursors

Mild periodate treatment of mammalian cells introduces aldehydes into surface sialic acids. The chemical reactivity of periodate-produced aldehydes has been exploited for purposes ranging from investigation of sLe^X-mediated lymphocyte homing [176] to the development of new modes of carbohydrate-based cell attachment [177]. Although periodate treatment of surface

sugars provided an important precedent for MOE because of the similar chemical properties of aldehydes and ketones found in sugar analogues, limitations to the periodate method included transient surface display of modified sugars due to membrane recycling, significant loss of cell viability, and incompatibility with in vivo conditions. By comparison, glycan-displayed ketones installed by MOE can be maintained indefinitely on the cell surface [65], harm to the host cells—even for toxic analogues—can be avoided by controlled delivery rates and control of cell density [46], and MOE methodology can be used in animals [17].

The Bertozzi group pioneered the introduction of orthogonal chemical functionalities onto the cell surface with a spate of monosaccharides derivatized with ketone groups. The first such compound, ManNLev (**14**), was a precursor for sialic acid biosynthesis as have been several later ManNAc analogues that include the peracetylated forms of ManNLev (Ac_4ManNLev, **87**), ManNHomoLev (Ac_4ManNHomoLev, **88**) [47], and ManNOxoHex (Ac_4ManNOxoHex, **89**) [63]. In follow-up experiments Hang and coworkers showed that a ketone-containing GalNAc mimic (*Sect. 3.2*) could also be used to install ketones into *O*-linked glycans and possibly GAGs, although fewer ketones were expressed relative to ManNLev [26]. Treatment of mammalian cells with a ketone-containing GlcNAc analogue GlcNLev (**169**) resulted in the incorporation of ketones into *O*-GlcNAc-modified nuclear and cytosolic proteins (*Sect. 3.4*). Finally, ketones have been incorporated into the cell walls of lactic acid bacteria via metabolic substrate-based engineering [30,31].

4.2.2 Exploitation of Ketone-Bearing Glycans

After the metabolic installation of a ketone group into a cellular glycan, chemoselective ligation reactions can be used towards several ends. In particular, there are two complementary chemical functional groups—the aminooxy and hydrazide—that selectively reactive with ketones to the exclusion of other surface components (*Fig. 12*). The choice of reactant depends on several factors including expense (hydrazide conjugation reagents are commercially available at lower costs), reaction kinetics (aminooxy ligation occurs more rapidly at physiological pH), or the desired stability of the bond. To expound on the last point, hydrazide-coupling reactions result in hydrazone linkages, which are pH sensitive and upon internalization to low pH endosomal vesicles can be reversed. By this process, ‘cargo’ attached to the hydrazide group can be released in an internal compartment and the ketone ‘tag’ can return to the cell surface to be reused; this method in effect results in the shuttling of agents such as drugs into a cell [178]. By contrast, the oxime linkage that results from aminooxy ligation is pH insensitive and preferred when stable conjugation products are desired, such as the delivery of MRI contrast agents to non-natural sialosides [44].

In addition to drug delivery, surface ketones installed via MOE have been exploited for the development of MRI contrast reagents [44], for gene transfer [179], for glycan remodeling [65], and for the targeting of cancer cells with toxic proteins [20] (*Fig. 13*). Probably the most commonly utilized ketone-based reactions have been to quantify surface expression of non-natural sialic acids. A facile method for quantification of surface-displayed functional groups has been developed for ketones and involves a two-step process [20,65]. The first step involves covalent conjugation with one of many commercially available hydrazide- or aminooxy-biotin probes. In the second step, the biotinylated cells are stained with FITC-avidin followed by flow cytometry quantification. A drawback of this methodology is that detection

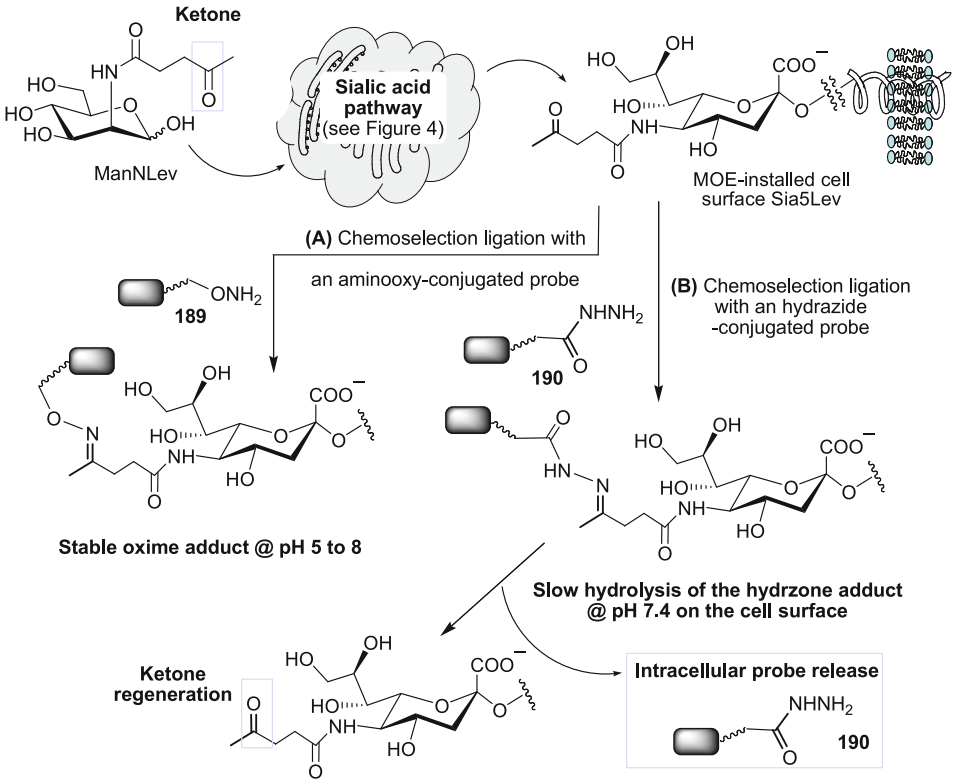


Figure 12
Ketone-based chemoselective ligation reactions that exploit cell surface displayed Sia5Lev (32)

is not quantitative insofar as the labeling efficiency depends on pH and biotin hydrazide concentration; for hydrazides, pH ranges where fast conjugation occurs with (i. e., at ~pH 5) are not compatible with cell survival. Another limiting factor is that biotin hydrazide is not soluble above ~5 mM whereas labeling increases throughout the soluble range, indicating that if higher concentrations were possible, additional conjugation would occur [180]. Nonetheless, this labeling method—contingent upon reagent concentrations and reaction times being maintained constant from experiment to experiment—provides a rapid means to compare cell lines, concentrations, time of incubation with analogue, and so on. Further, even with these limitations, up to 10^7 surface functionalities can be ligated under physiological conditions [180], a number that compares favorably with tumor-associated antigens targeted in immunotherapy where only ~10^4 to 10^5 surface epitopes are present, which is often not sufficient to achieve adequate levels of drug delivery [181].

4.2.3 Shortcomings of Ketone-Based Ligation Strategies

The ketone-based chemoselective ligation reactions described above initially demonstrated the versatility and potential of MOE to modulate biology through changes in cell surface

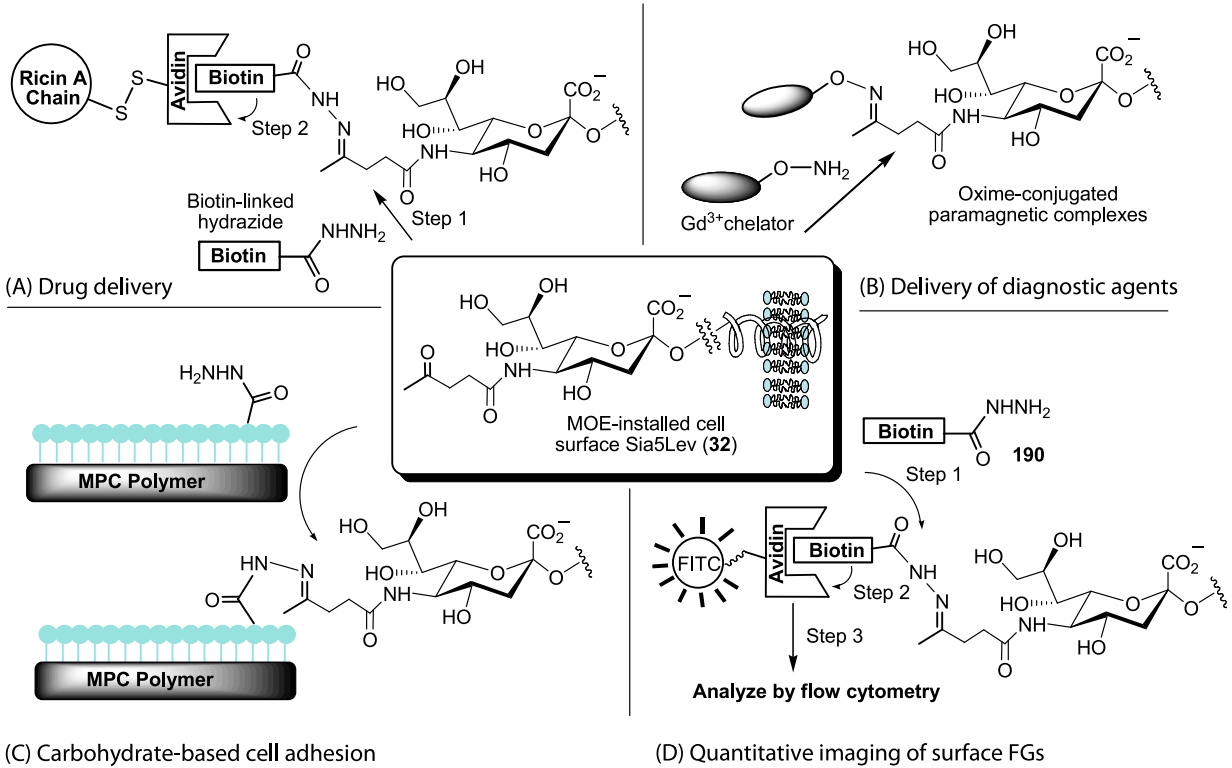


Figure 13
Applications of chemoselective ligation reactions that exploit Sia5Lev

chemistry. These reactions, however, have limited intracellular use owing to competition with endogenous keto-metabolites and they can be compromised by pH requirements that do not match physiological conditions. In addition, keto-analogs, exemplified by ManNLev, do not transit metabolic pathways efficiently (in fact, Ac_4ManNLev inhibits flux through the sialic acid pathway [47]). Therefore, to augment the potential of these pioneering analogues in the study and manipulation of cellular components, the ongoing development of additional cell-compatible chemoselective ligation reactions is of fundamental importance and has led in several directions outlined below in ❯ *Sect. 4.3* to ❯ *Sect. 4.6*.

4.3 Glycan-Displayed Azides

4.3.1 Azides and the Modified Staudinger Reaction

An exciting chemoselective ligation technique based on the conjugation between an azide and a phosphine was developed when the Bertozzi group modified the Staudinger ligation to be compatible with physiological conditions. The conventional Staudinger reaction is hydrolyzed under aqueous conditions (❯ *Fig. 14A*) whereas the physiological reaction forms a stable amide bond by coupling of an azide and a specifically engineered triarylphosphine [60]. This 'second generation' MOE chemical functional group is superior to ketones in that azides are completely abiotic and can be exploited for chemoselective ligation reactions inside a cell as well as on the surface, furthermore they do not suffer from the pH limitations of ketone-hydrazide coupling. As an added benefit, the bevy of azide-bearing monosaccharide analogues now includes ManNAc, GlcNAc, GalNAc, and fucose derivative that typically are metabolized by glycosylation pathways more efficiently than their keto counterparts (as discussed in ❯ *Sect. 3.1*.).

4.3.2 Azide-Derivatized Analogues Are Being Used in Proteomics

Similar to the labeling scheme shown in ❯ *Fig. 13* for ketones, the azide group provides a convenient handle for quantification of metabolic incorporation (❯ *Fig. 14C*) [60]. In ongoing efforts to move beyond whole cell analysis by flow cytometry to the identification of individual molecules, azide-modified analogues are gaining extensive use in emerging 'glycoproteomics' efforts. Particularly fruitful have been GlcNAz analogues that, upon incorporation into *O*-GlcNAc-modified nuclear and cytoplasmic proteins (❯ *Sect. 3.4*), allow facile detection by biotin- or pFLAG-based probes (❯ *Fig. 15*). Metabolic incorporation of GlcNAz analogues also enables rapid enrichment, analysis, and identification of modified proteins as demonstrated by a recent study where more than 199 putative *O*-GlcNAc-modified proteins were isolated from analogue-treated HeLa cells. Among the first proteins identified in this experiment, 10 were previously reported and 13 were newly identified as bearing the *O*-GlcNAc modification [156]. In a variation of this method, enzymatic tagging of *O*-GlcNAc with keto-analogues allowed the identification of modified proteins from the brain [141]. Although the methodology is not yet as well developed, similar proteomics approaches are under development for the metabolic tagging, isolation, and characterization of GalNAc- and fucose-bearing glycans. Finally, it is worth noting that azide detection methods apply to bacteria [182], potentially opening the door to similar glycomics approaches to prokaryotes where nascent MOE experiments have been reported [28,30].

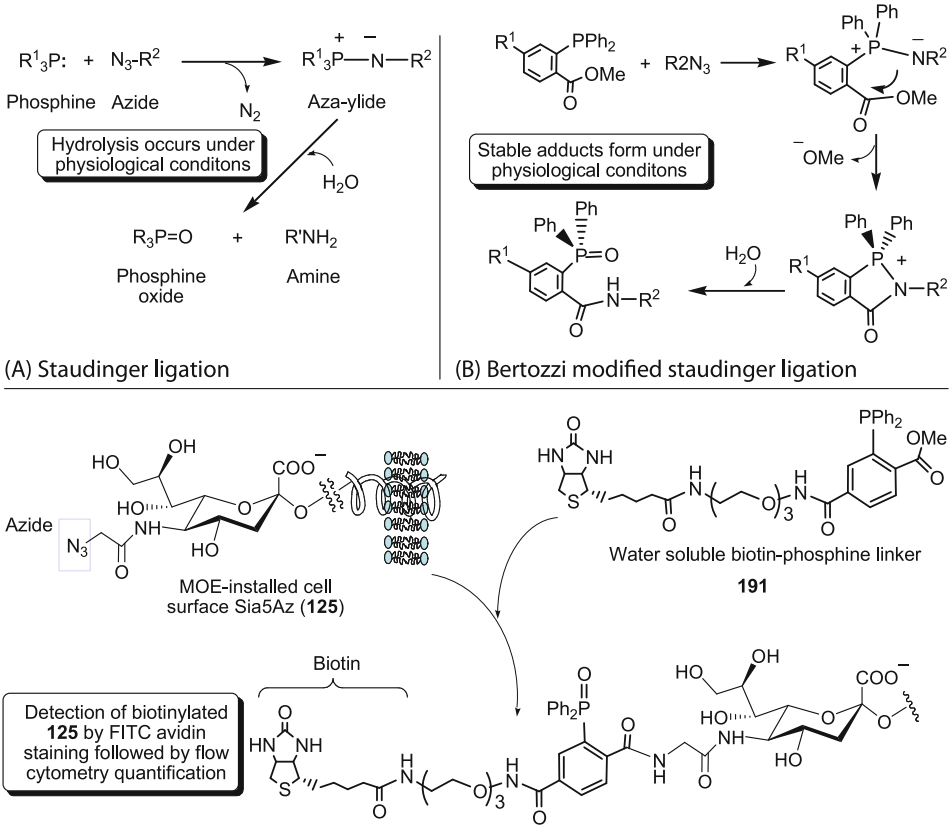


Figure 14
Use of the Staudinger ligation in MOE. (A) The original ligation product of the original Staudinger reaction undergoes hydrolysis under physiological conditions. (B) The Bertozzi modification of the Staudinger ligation [60] produces stable adducts in water thereby adapting this reaction for MOE and rendering it appropriate for applications including quantification of surface display of Sia5Az (C)

4.4 Glycan-Displayed Thiols

4.4.1 Thiols Have Versatile Chemistry but Are not Unique to the Cell Surface

Recently, our laboratory synthesized thiol-derivatized ManNAc analogues [52] corresponding to the glycolyl sugars used by the Schnaar group to install Neu5Gc (**37**) onto human cells [103]. The impetus behind the production of the thiol-derivatized ManNAc analogue Ac_5ManNTGc (*Fig. 16A*, **68**) used to install Sia5TGc on the cell surface was the versatile chemistry of thiol groups, along with the wealth of commercial reagents available for conjugation, that would extend the possibilities of ketone and azide groups. A potential downside to thiol display is that these groups are not unique to a cell. They occur in small molecules such

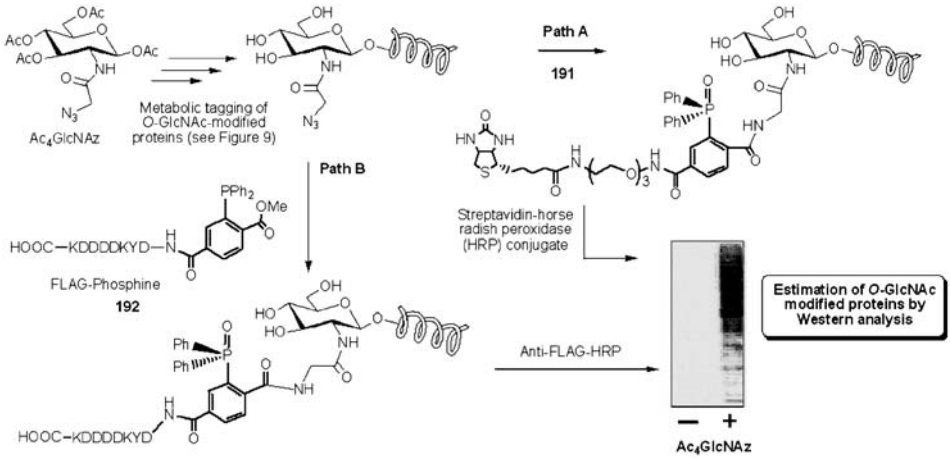


Figure 15

GlcNAc analogues can be exploited for proteomics. The GlcNAc analogue Ac_4GlcNAz tags *O*-GlcNAc-modified proteins with an azide group that can then be analyzed through Western analysis after visualization by strategies shown in either Path A or Path B

as glutathione and cysteine inside a cell as well as on proteins both inside a cell and on the surface. However, the bulk of surface thiols participate in disulfide bonds, are complexed with metals, or are shielded by the glycocalyx (*Fig. 16*), providing sialic acid-displayed thiols with unique accessibility to chemical reagents and growth surfaces in the cell's microenvironment.

4.4.2 Sialic Acid Displayed Thiols Are Highly Accessible to Outside Reactions

The first indication that sialic acid-displayed thiols endowed the cell surface with unique properties was the clustering of Jurkat cells upon incubation with **68** [104]. Typically, Jurkat cells grow in suspension as single cells but the presence of thiols on the cell surface apparently led to disulfide formation between adjacent cells creating cell aggregates containing a few thousand cells reminiscent of hepatocyte spheroids created for use in bioartificial livers. The expression of sialic acid-displayed thiols was confirmed by adapting the chemoselective labeling strategy used for ketone detection shown in *Fig. 13D* and for sialic acid-displayed azides shown in *Fig. 14C*. In this case, because maleimide is a chemoselective ligation partner for thiols, the analogue-treated cells were incubated with (+)-biotinyl-3-maleimidopropionamidyl-3,6-dioxaoctanediamine (MB) followed by staining with FITC-conjugated avidin and flow cytometric quantification. In these experiments, cells directly labeled with MB after exposure to **68** (*Fig. 16B*) showed at least an order of magnitude greater staining than controls, confirming that glycocalyx-displayed thiols were able to interact with the exogenous labeling reagents much better than sulfurs located in membrane proteins [104]. When the cells incubated with **68** were treated with tris-(2-carboxyethyl) phosphine hydrochloride (TCEP) a reducing agent for disulfide bonds prior to MB labeling (*Fig. 16C*), the fluorescent signal increased by another order of magnitude, indicating that Sia5TGc-displayed thiols predomi-

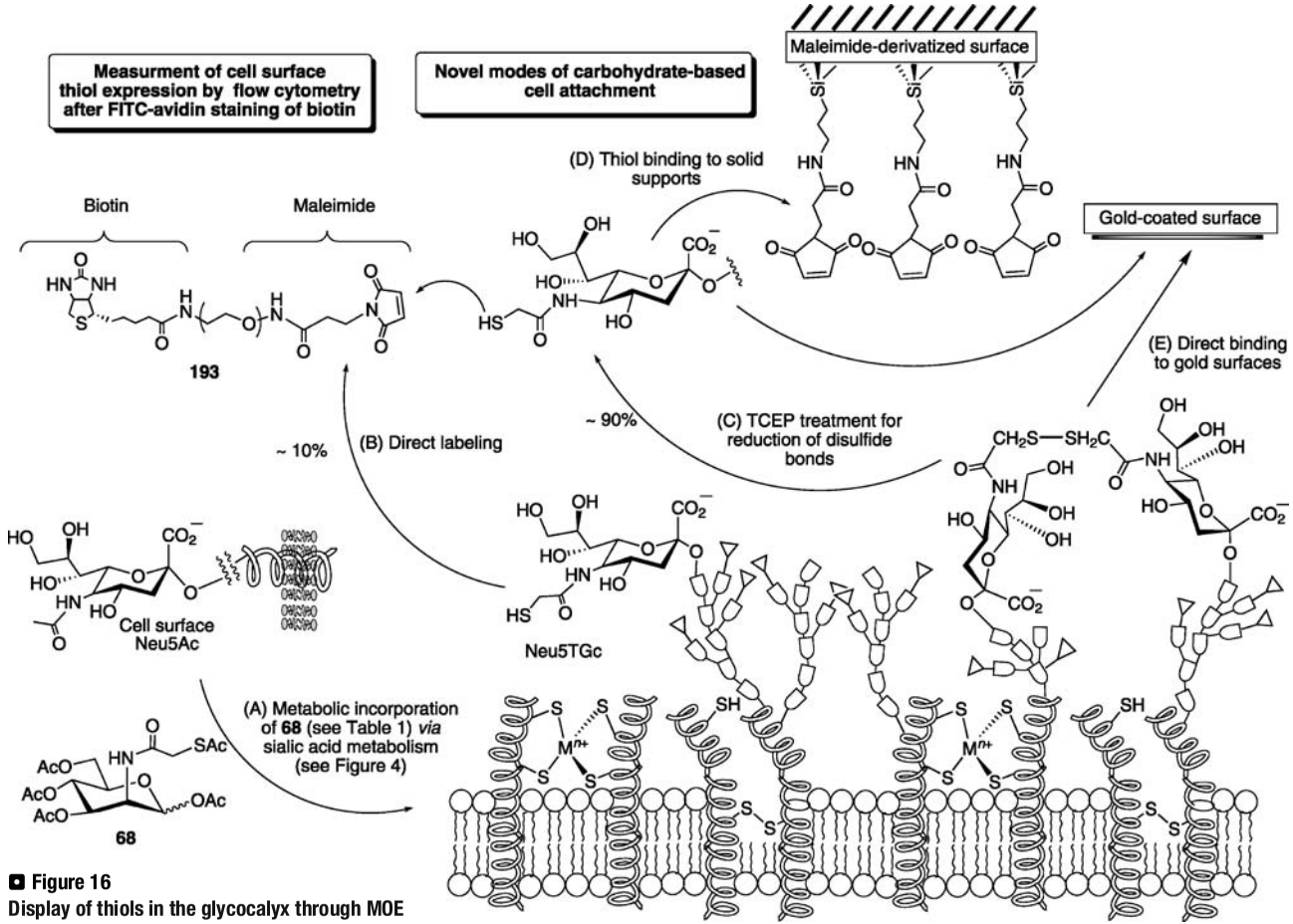


Figure 16
Display of thiols in the glycocalyx through MOE

nantly exist as disulfides, an observation consistent with the slightly oxidative nature of tissue culture medium.

4.4.3 Sialic Acid Displayed Thiols Are Useful for the Control of Stem Cell Fate

Similar to sugar-displayed aldehydes and ketones that have been exploited to attach cells to chemically compatible growth surfaces or polymers [177,183], Sia5TGc has proven useful for novel modes of carbohydrate-based cell adhesion. In one set of experiments, Jurkat cells treated with **68** selectively attached to a maleimide functionalized surface [104] **68** (*Fig. 16D*). In follow-up experiments, the human embryoid body-derived (hEBD) LVEC cell line was incubated **68** and grown on gold to take advantage of the high-affinity interaction expected from glycocalyx-displayed thiols with this metal (*Fig. 16E*). These growth conditions activated the Wnt signaling pathway, led to β-catenin accumulation, distinct changes in cellular morphology, nestin staining, and re-organization of the cytoskeleton indicative of differentiation to a neural cell lineage [104]. By contrast, only negligible changes in morphology were seen for cells treated with Ac_5ManNGc, the oxy-analogue of **68**, or the solvent vehicle ethanol, indicating that the thiol moiety of **68**, or an intracellular metabolic intermediate arising from this compound, had a critical role in the analogue-mediated differentiation to neurons.

4.5 The Use of Photo-Affinity Cross-Linkers in MOE

The Paulson group devised an ingenious method to exploit MOE analogues derivatized with photo-activated crosslinkers to identify the in situ binding partners of CD22 [116]. In these experiments the sialic acid analogue 9-AAz-NeuAc, which was synthesized as shown in *Scheme 9* and bears an aryl-azide moiety at the C-9 position, was incubated with B-cells. Surface display of the aryl-azide functionality was verified by labeling the cells with a biotin probe that was ligated by the Bertozzi–Staudinger reaction and visualized by FITC-avidin staining. Interestingly, despite the bulk of the 9-AAz group, the modified glycan remained a binding partner for CD22, allowing in situ interactions to be captured by photo-activated cross-linking (*Fig. 17*). The importance of devising new methods for probing biological function, such as this approach to understanding CD22, was illustrated by the outcome of this experiment wherein CD22 was found to assemble into homo-multimeric complexes—a result that differed from previous experiments that used more highly invasive methods to explore CD22 interactions.

4.6 Cycloaddition Ligation Reactions

4.6.1 Cycloaddition Reactions Are Compatible with Carbohydrates Under Physiological Conditions

Cycloaddition reactions between azides and alkynes—exemplified by the Huisgen 1,3-dipolar cycloaddition—have tremendous potential for the development of biomolecules and have been employed for conjugation of sugars to peptides [184]. Drawbacks of the Huisgen cycloadditions, however, are poor regioselectivity and incompatibility with physiological conditions. These limitations, which hinder MOE applications in living cells, were first overcome by cop-

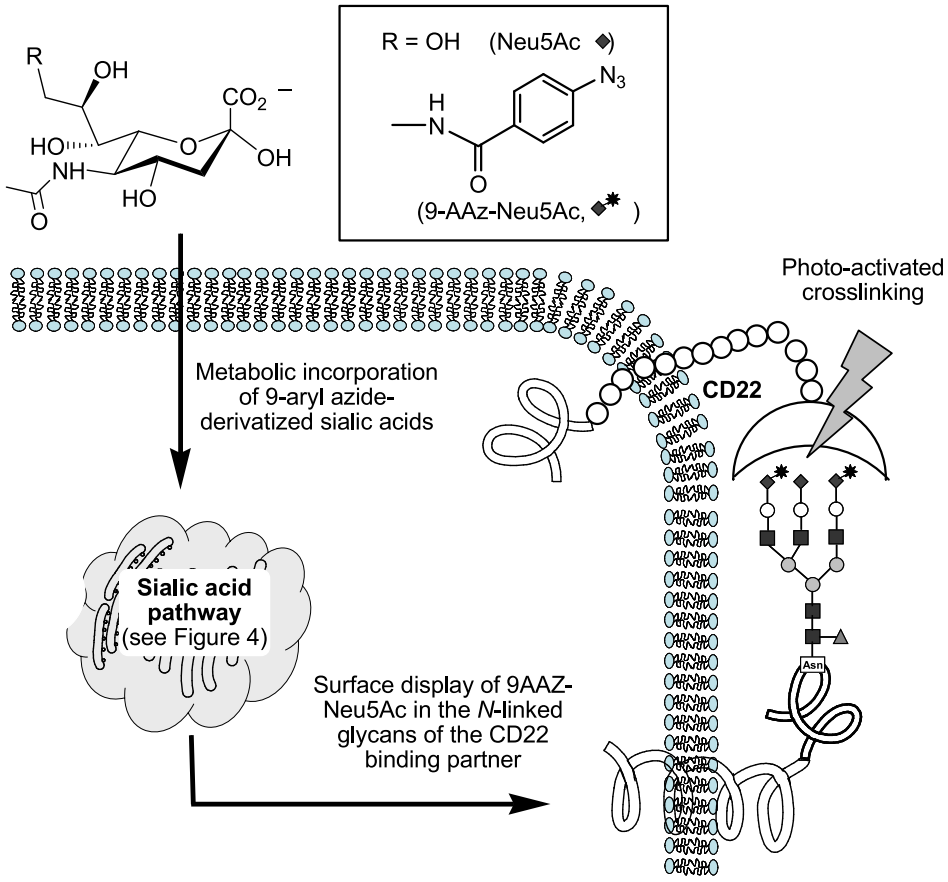


Figure 17
Photo-activated cross-linking of MOE sialic acids to identify CD22 binding partners

per [Cu (I)] catalyzed versions of the Huisgen 1,3-dipolar cycloaddition of azides and terminal alkynes reported independently by the Sharpless and Meldal groups (*Fig. 18A*) [184,185]. In the first experiments, grafting of protected and unprotected carbohydrates onto a polymer was demonstrated via an azide located at either the C-6 or anomeric position by Cu(I)-catalyzed 'click chemistry', providing a simple and efficient route to synthetic glycopolymers. The (3+2) cycloaddition chemistry has now been extended to MOE applications by using non-natural L-Fuc derivatives derivatized with azide or alkyne groups for direct ligation of fluorescent tags [147] (*Fig. 18B*). Although compatible with physiological conditions, considerable copper-attributed toxicity has been reported in azide-alkyne coupling, which suggests this chemistry will be more appropriate for proteomics rather than live cell applications. To avoid copper toxicity, Agard and coworkers have recently reported a family of cyclooctynes that were effective for tagging azides on live cells in the absence of copper through strain-promoted [3+2] cycloaddition [166] (*Fig. 18C*) that may find roles in MOE-proteomics (*Fig. 18D*).

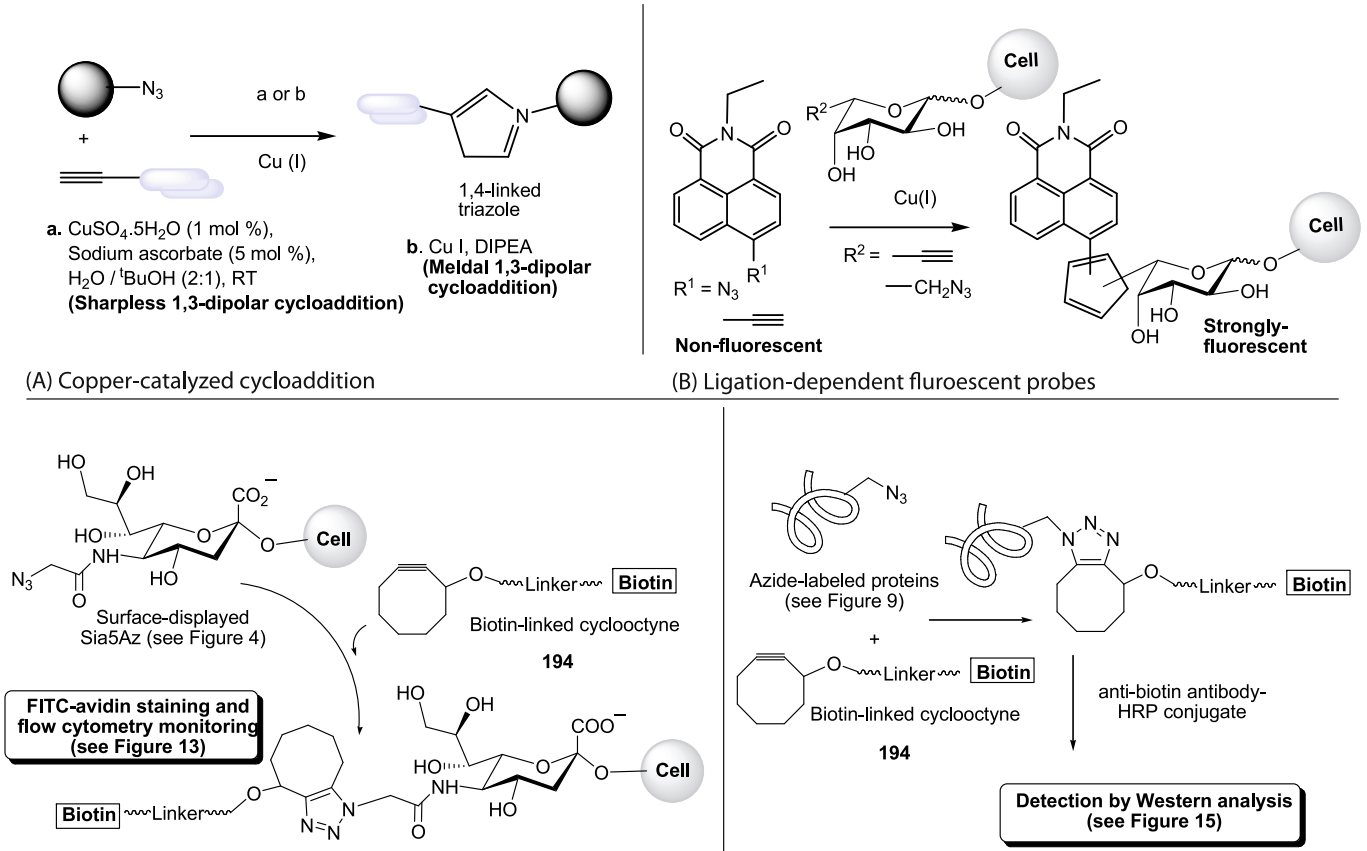


Figure 18
Cycloaddition ligation reactions used in MOE

4.6.2 Development of Fluorogenic Reactions for Azide-Derivatized Glycans

To date, the quantification of MOE functional groups on the cell surface has generally used secondary fluorescent dyes for visualization. This strategy can be compromised by difficult-to-eliminate background noise making attractive the use of fluorogenic reactions, where the ligated product generates a strong detectable signal while the unreacted reagent remains traceless. On the basis of this concept, two click-activated fluorescent probes based on 1,8-naphthalimide (*Fig. 18B*) that can selectively label azido- or acetylene-modified L-fucose analogs by a Cu(I)-catalyzed azide-alkyne ligation, which triggers their fluorescence, were used in a fluorogenic labeling technique sufficiently sensitive to visualize fucosylated glycoproteins in intact cells [147].

4.6.3 Extending the Use of Cycloaddition Reactions in MOE

Cycloaddition reactions between dienes and dienophiles provide a versatile method in synthetic organic chemistry for converting simple molecules to structures of increased complexity. In 2002, Pozsgay and coworkers developed a method for chemoselective bioconjugation of carbohydrates to proteins by using Diels–Alder cycloaddition [186]. Of potential relevance to MOE, maleimidoyl-conjugated BSA (**196**) was coupled with L-rhamnose containing oligosaccharides (particularly the *O*-specific polysaccharide of *Shigella dysenteriae* type 1, **195** that contain a diene) under physiological conditions (*Fig. 19*). These investigators found that the Diels–Alder conjugation methods—while favored at pH 5.7 in phosphate-borate buffer—proceeded with equal efficiency at neutral pH in ion-exchanged water. Prospects of extending this chemistry to metabolically installed dienes—although to our knowledge not yet demonstrated—should be possible based on the transit of a single alkene though the sialic acid pathway (ManNCrot, *Table 1*C) [18], combined with the use of alkyl chains of sufficient length, such as ManNHex or ManNHept (*Table 1*A) [63], to accommodate a diene.

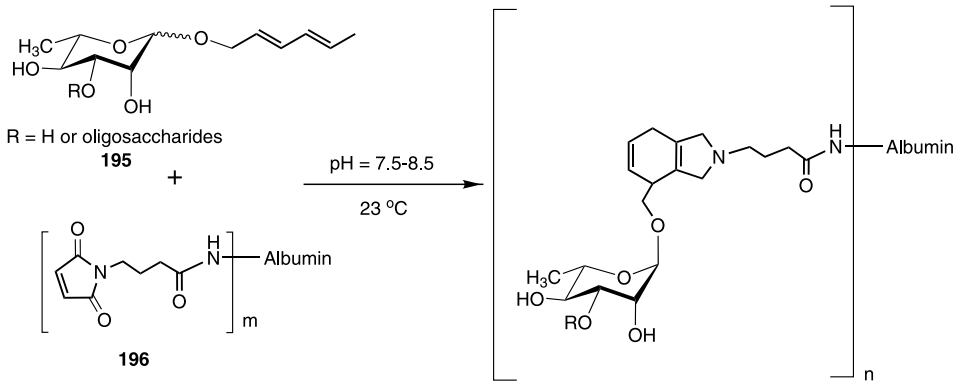


Figure 19
Chemoselective bioconjugation of carbohydrates to proteins by using Diels–Alder cycloaddition chemistry

5 Concluding Comments and Future Prospects

In conclusion, this report has provided an overview of MOE technology by outlining the various biological systems that accept non-natural sugar analogues and metabolically incorporate them into cellular glycoconjugates. Although novel analogues are constantly under development, the lists of compounds described in *Sect. 3* designed to target sialic acid, GalNAc, GlcNAc, and fucose metabolism provide a broad-based 'tool-kit' to manipulate cell surface display on oligosaccharides in living cells and animals. An especially exciting aspect of MOE technology is the seamless merging of chemistry and biology, wherein novel chemical functionalities introduced into cellular components can be further exploited by a growing list of chemoselective ligation reactions. As can be deduced from the brief sampling of applications of MOE referenced in this report, the potential impact of MOE on biological research and medicine has already been significant and is expected to gain in importance as it becomes widely adopted by the research community.

Acknowledgement

The authors would like to thank the National Institutes of Health (R01 CA112314–01A1), the National Science Foundation (QSB-0425668), and The Susan B. Komen Breast Cancer Foundation (BCTR0503768) for financial support.

References

1. Sharon N, Lis H (1993) Sci Am 268:82
2. Rudd PM, Dwek RA (1997) Crit Rev Biochem Mol Biol 32:1
3. Spiro RG (2002) Glycobiology 12:43R
4. Varki A (1993) Glycobiology 3:97
5. Sears P, Lin CC, Hung SC, Takayama S, Witte KL, Alper PB, Wong CH (1998) Chem Commun 11:1161
6. Dwek RA (1996) Chem Rev 96:683
7. Fukuda M (2000) Cell surface carbohydrates: cell type-specific expression. In: Fukuda M, Hindsgaul O (eds) Molecular and Cellular Glycobiology. Oxford University Press, Oxford, UK, pp 1
8. Gabius HJ, Siebert HC, André S, Jiménez-Barbero J, Rüdiger H (2004) ChemBioChem 5:740
9. Imperiali B, Rickert KW (1995) Proc Natl Acad Sci USA 92:97
10. Wormald MR, Dwek RA (1999) Structure 7:R155
11. Luo BH, Springer TA, Takagi J (2003) Proc Natl Acad Sci USA 100:2403
12. Chen HY, Varki A (2002) J Exp Med 196:1529
13. Hanover JA (2001) FASEB J 15:1865
14. Zachara NE, Hart GW (2002) Chem Rev 102:431
15. Zachara NE, Hart GW (2006) Biochim Biophys Acta 1761:599
16. Campbell CT, Yarema KJ (2005) Gen Biol 6:236
17. Kayser H, Zeitler R, Kannicht C, Grunow D, Nuck R, Reutter W (1992) J Biol Chem 267:16934
18. Keppler OT, Horstkorte R, Pawlita M, Schmidt C, Reutter W (2001) Glycobiology 11:11R
19. Kornfeld R, Kornfeld S (1985) Annu Rev Biochem 54:631
20. Mahal LK, Yarema KJ, Bertozzi CR (1997) Science 276:1125
21. Strous JG, Dekker J (1992) Crit Rev Biochem Mol Biol 27:57
22. Dube DH, Prescher JA, Quang CN, Bertozzi CR (2006) Proc Natl Acad Sci USA 103:4819
23. Saxon E, Luchansky SJ, Hang HC, Yu C, Lee SC, Bertozzi CR (2002) J Am Chem Soc 124:14893
24. Kayser H, Geilen CC, Paul C, Zeitler R, Reutter W (1992) FEBS Lett 301:137

25. Chefalo P, Pan Y, Nagy N, Guo Z, Harding CV (2006) Biochemistry 45:3733
26. Hang HC, Bertozzi CR (2001) J Am Chem Soc 123:1242
27. Vocadlo DJ, Hang HC, Kim EJ, Hanover JA, Bertozzi CR (2003) Proc Natl Acad Sci USA 100:9116
28. Schilling B, Goon S, Samuels NM, Gaucher SP, Leary JA, Bertozzi CR, Gibson BW (2001) Biochemistry 40:12666
29. Goon S, Schilling B, Tullius MV, Gibson BW, Bertozzi CR (2003) Proc Natl Acad Sci USA 18:3089
30. Sadamoto R, Niikura K, Ueda T, Monde K, Fukuhara N, Nishimura SI (2004) J Am Chem Soc 126:3755
31. Sadamoto R, Nishimura SI (2005) In: Yarema KJ (ed) Handbook of Carbohydrate Engineering. Francis & Taylor/CRC Press, Boca Raton, FL, pp 495
32. Goon S, Bertozzi CR (2002) J Carbohydr Chem 21:943
33. Apweiler R, Hermjakob H, Sharon N (1999) Biochim Biophys Acta 1473:4
34. Comb DG, Roseman S (1960) J Biol Chem 235:2529
35. Ghosh S, Roseman S (1965) J Biol Chem 240:1525
36. Jourdian GW, Swanson A, Watson D, Roseman S (1966) Meth Enzymol 8:205
37. Roseman S (2001) J Biol Chem 276:41527
38. Roseman S (1957) J Biol Chem 226:115
39. Wood IS, Trayhurn P (2003) Br J Nutr 89:3
40. Chen H, Wang Z, Sun Z, Kim EJ, Yarema KJ (2005) In: Yarema KJ (ed) Handbook of Carbohydrate Engineering. Francis & Taylor/CRC Press, Boca Raton, FL, pp 1
41. Sarkar AK, Fritz TA, Taylor WH, Esko JD (1995) Proc Natl Acad Sci USA 92:3323
42. Fritz TA, Lugemwa FN, Sarkar AK, Esko JD (1994) J Biol Chem 269:300
43. Sarkar AK, Rostand KS, Jain RK, Matta KL, Esko JD (1997) J Biol Chem 272:25608
44. Lemieux GA, Yarema KJ, Jacobs CL, Bertozzi CR (1999) J Am Chem Soc 121:4278
45. Jacobs CL, Yarema KJ, Mahal LK, Nauman DA, Charters N, Bertozzi CR (2000) Meth Enzymol 327:260
46. Jones MB, Teng H, Rhee JK, Baskaran G, Lahar N, Yarema KJ (2004) Biotechnol Bioeng 85:394
47. Kim EJ, Sampathkumar SG, Jones MB, Rhee JK, Baskaran G, Yarema KJ (2004) J Biol Chem 279:18342
48. Kim EJ, Jones MB, Rhee JK, Sampathkumar SG, Yarema KJ (2004) Biotechnol Prog 20:1674
49. Prescher JA, Dube DH, Bertozzi CR (2003) Glycobiology 13:894
50. Gagiannis D, Gossrau R, Reutter W, Zimmermann-Kordmann M, Horstkorte R (2007) Biochim Biophys Acta 1770:297
51. Wang Z, Sun Z, Li AV, Yarema KJ (2006) J Biol Chem 281:27016
52. Sampathkumar SG, Li AV, Yarema KJ (2006) Nature Protocols 1:2377
53. Basson MD, Liu YW, Hanly AM, Emenaker NJ, Shenoy SG, Gould Rothberg BE (2000) J Gastrointest Surg 4:501
54. Miller SJ (2004) Mini Rev Med Chem 4:839
55. Sampathkumar SG, Jones MB, Meledeo MA, Campbell CT, Choi SS, Hida K, Gomutputra P, Sheh A, Gilmartin T, Head SR, Yarema KJ (2006) Chem Biol 13:1265
56. Oetke C, Brossmer R, Mantey LR, Hinderlich S, Isecke R, Reutter W, Keppler OT, Pawlita M (2002) J Biol Chem 277:6688
57. Chefalo P, Pan YB, Nagy N, Harding C, Guo ZW (2004) Glycoconjug J 20:407
58. Lemieux GA, Bertozzi CR (2001) Chem Biol 8:265
59. Schnaar RL, Needham LK (1994) Meth Enzymol 230
60. Saxon E, Bertozzi CR (2000) Science 287:2007
61. Sampathkumar SG, Jones MB, Yarema KJ (2006) Nature Protocols 1:1840
62. Jourdian GW, Dean L, Roseman S (1971) J Biol Chem 246:430
63. Jacobs CL, Goon S, Yarema KJ, Hinderlich S, Hang HC, Chai DH, Bertozzi CR (2001) Biochemistry 40:12864
64. Tomiya N, Ailor E, Lawrence SM, Betenbaugh MJ, Lee YC (2001) Anal Biochem 293:129
65. Yarema KJ, Mahal LK, Bruehl RE, Rodriguez EC, Bertozzi CR (1998) J Biol Chem 273:31168
66. Kayser H, Geilen CC, Paul C, Zeitler R, Reutter W (1993) Experientia 49:885
67. Prescher JA, Dube DH, Bertozzi CR (2004) Nature 430:873
68. Horstkorte R, Mühlenhoff M, Reutter W, Nöhring S, Zimmermann-Kordmann M, Gerardy-Schahn R (2004) Exp Cell Res 298:268
69. Wildt S, Gerngross TU (2005) Nat Rev Microbiol 3:119
70. Betenbaugh MJ, Tomiya N, Narang S, Hsu JTA, Lee YC (2004) Curr Opin Struct Biol 14:601

71. Ruffing A, Chen RR (2006) Microbial Cell Factories 5:25
72. Shah MM, Fujiyama K, Flynn CR, Joshi L (2003) Nat Biotechnol 21:1470
73. Viswanathan K, Lawrence S, Hinderlich S, Yarema KJ, Lee YC, Betenbaugh M (2003) Biochemistry 42:15215
74. Sadamoto R, Niikura K, Sears PS, Liu H, Wong CH, Suksomcheep A, Tomita F, Monde K, Nishimura SI (2002) J Am Chem Soc 124:9018
75. Tanner ME (2005) Bioorg Chem 33:216
76. Lemieux GA, Bertozzi CR (1998) Trends Biotechnol 16:506
77. Angata T, Varki A (2002) Chem Rev 102:439
78. Schauer R, Kelm S, Reuter G, Roggentin P, Shaw L (1995) In: Rosenberg A (ed) Biology of the Sialic Acids. Plenum Press, New York
79. Schauer R (2000) Glycoconjug J 17:485
80. Schauer R (2004) Arch Biochem Biophys 426:132
81. Chou HH, Hayakawa T, Diaz S, Krings M, Indriati E, Leakey M, Paabo S, Y S, Takahata N, Varki A (2002) Proc Natl Acad Sci USA 99:11736
82. Brinkman-Van der Linden EC, Sjoberg ER, Juneja LR, Crocker PR, Varki N, Varki A (2000) J Biol Chem 275:8633
83. Shinya K, Ebina M, Yamada S, Ono M, Kasai N, Kawaoka Y (2006) Nature 440:435
84. Stevens J, Blixt O, Tumpey TM, Taubenberger JK, Paulson JC, Wilson IA (2006) Science 312:404
85. Stevens J, Blixt O, Glaser L, Taubenberger JK, Palese P, Paulson JC, Wilson IA (2006) J Mol Biol 355:1143
86. Keppler OT, Hinderlich S, Langner J, Schwartz-Albiez R, Reutter W, Pawlita M (1999) Science 284:1372
87. Takahashi S, Inoue H, Miyake Y (1992) J Biol Chem 267:13007
88. Luchansky SJ, Yarema KJ, Takahashi S, Bertozzi CR (2003) J Biol Chem 278:8036
89. Seppala R, Lehto VP, Gahl WA (1999) Am J Hum Genet 64:1563
90. Weiss P, Tietze F, Gahl WA, Seppala R, Ashwell G (1989) J Biol Chem 264:17635
91. Yarema KJ, Goon S, Bertozzi CR (2001) Nat Biotechnol 19:553
92. Hinderlich S, Berger M, Keppler OT, Pawlita M, Reutter W (2001) Biological Chemistry 382:291
93. Mantey LR, Keppler OT, Pawlita M, Reutter W, Hinderlich S (2001) FEBS Lett 503:80
94. Keppler OT, Stehling P, Herrmann M, Kayser H, Grunow D, Reutter W, Pawlita M (1995) J Biol Chem 270:1308
95. Liu T, Guo Z, Yang Q, Sad S, Jennings HJ (2000) J Biol Chem 275 32832
96. Mahal LK, Charter NW, Angata K, Fukuda M, Koshland DE, Jr., Bertozzi CR (2001) Science 294:380
97. Charter NW, Mahal LK, Koshland DE, Jr., Bertozzi CR (2002) J Biol Chem 277:9255
98. Faure C, Chalazonitis A, Rhéaume C, Bouchard G, Sampathkumar SG, Yarema KJ, Gershon MD (2007) Developmental Dynamics 236:44
99. Wieser JR, Heisner A, Stehling P, Oesch R, Reutter W (1996) FEBS Lett 395:170
100. Matrosovich M, Klenk HD (2003) Rev Med Virol 13:85
101. Keppler OT, Herrmann M, von der Lieth CW, Stehling P, Reutter W, Pawlita M (1998) Biochem Biophys Res Commun 253:437
102. Verras M, Sun Z (2005) Cancer Lett epub
103. Collins BE, Fralich TJ, Itonori S, Ichikawa Y, Schnaar RL (2000) Glycobiology 10:11
104. Sampathkumar SG, Li AV, Jones MB, Sun Z, Yarema KJ (2006) Nat Chem Biol 2:149
105. Lawrence SM, Huddleston KA, Pitts LR, Nguyen N, Lee YC, Vann WF, Coleman TA, Betenbaugh MJ (2000) J Biol Chem 275:17869
106. Gross HJ, Brossmer R (1988) Eur J Biochem 177:583
107. Gross HJ, Rose U, Krause JM, Paulson JC, Schmid K, Feeny RE, Brossmer R (1989) Biochemistry 28:7386
108. Brossmer R, Gross HJ (1994) Meth Enzymol 247:153
109. Brossmer R, Gross HJ (1994) Meth Enzymol 247:177
110. Gross HJ, Brossmer R (1995) Glycoconjug J 12:739
111. Gross HJ (1992) Eur J Biochem 203:269
112. Kosa RE, Gross HJ (1993) Biochem Biophys Res Commun 190:914
113. Oetke C, Hinderlich S, Brossmer R, Reutter W, Pawlita M, Keppler OT (2001) Eur J Biochem 268:4553
114. Bardor M, Nguyen DH, Diaz S, Varki A (2005) J Biol Chem 280:4228
115. Luchansky SJ, Goon S, Bertozzi CR (2004) ChemBioChem 5:371
116. Han S, Collins BE, Bengtson P, Paulson JC (2005) Nat Chem Biol 1:93
117. Baumberger F, Vasella A, Schauer R (1988) Helvetica Chimica Acta 71:429

118. Liu JLC, Shen GJ, Ichikawa Y, Rutan JF, Zapata G, Vann WF, Wong CH (1992) J Am Chem Soc 114:3901
119. Pan Y, Ayani T, Nadas J, Wen S, Guo Z (2004) Carbohydr Res 339:2091–2100
120. Kok GB, Campbell M, Mackey BL, von Itzstein M (2001) Carbohydr Res 332:133
121. Fitz W, Schwark JR, Wong CH (1995) J Org Chem 60:3663
122. Meyer K (1945) In: Anson ML, Edsall JT (eds) Advances in Protein Chemistr. Academic Press, New York, NY, pp 249
123. Weissmann B, Hinrichsen DF (1995) Biochemistry 8:2034
124. McMaster TJ, Berry M, Corfield AP, Miles MJ (1999) Biophys J 77:533
125. Shogren R, Gerken TA, Jentoft N (1989) Biochemistry 28:5525
126. Gerken TA, Butenhof KJ, Shogren R (1989) Biochemistry 28:5536
127. Tsuboi S, Fukuda M (2001) BioEssays 23:46
128. Julenius K, Molgaard A, Gupta R, Brunak S (2005) Glycobiology 15:153
129. Rudd PM, Elliorr T, Cresswell P, Wilson IA, Dwek RA (2001) Science 291:2370
130. Kannagi R (2002) Curr Opin Struct Biol 12:599
131. Berger EG (1999) Biochim Biophys Acta 1455:255
132. Brockhausen I (1999) Biochim Biophys Acta 1473:67
133. Ten Hagen KG, Fritz TA, Tabak LA (2003) Glycobiology 13:1R
134. Becker DJ, Lowe JB (2003) Glycobiology 13:41R
135. Haltiwanger RS, Lowe JB (2004) Annu Rev Biochem 73:491
136. Dube DH, Bertozzi CR (2005) Nat Rev Drug Discov 4:477
137. Yarema KJ, Bertozzi CR (1998) Curr Opin Chem Biol 2:49
138. Kim YJ, Varki A (1997) Glycoconjug J 14:569
139. Rosen SD (2004) Annu Rev Immunol 22:129
140. Fuster MM, Esko JD (2005) Nat Rev Cancer 5:526
141. Khidekel N, Ficarro SB, Peters EC, Hsieh-Wilson LC (2004) Proc Natl Acad Sci USA 101:13132
142. Ratner DM, Adams EW, Disney MD, Seeberger PH (2004) ChemBioChem 5:1375
143. Prescher JA, Bertozzi CR (2005) Nat Chem Biol 1:13
144. Wild MK, Lühn K, Marquardt T, Vestweber D (2002) Cells Tissues Organs 172:161
145. Rabuka D, Hubbard SC, Laughlin ST, Argade SP, Bertozzi CR (2006) J Am Chem Soc 128:12078
146. Laughlin ST, Agard NJ, Baskin JM, Carrico IS, Chang PV, Ganguli AS, Hangauer MJ, Lo A, Prescher JA, Bertozzi CR (2006) Meth Enzymol 415:230
147. Sawa M, Hsu TL, Itoh T, Sugiyama M, Hanson SR, Vogt PK, Wong CH (2006) Proc Natl Acad Sci USA 103:12371
148. Hinderlich S, Berger M, Schwarzkopf M, Effertz K, Reutter W (2000) Eur J Biochem 267:3301
149. Torres CR, Hart GW (1984) J Biol Chem 259:3308
150. Hanover JA, Cohen CK, Willingham MC, Park MK (1987) J Biol Chem 262:9887
151. Heese-Peck A, Cole RN, Borkhsenious ON, Hart GW, Raikhel NV (1995) Plant Cell 7:1459
152. Hart GW (1992) Curr Opin Cell Biol 4:1017
153. Zachara NE, Hart GW (2004) Biochim Biophys Acta 1673:13
154. Kho Y, Kim SC, Jiang C, Barma D, Kwon SW, Cheng J, Jaunbergs J, Weinbaum C, Tamanoi F, Falck J, Zhao Y (2004) Proc Natl Acad Sci USA 101:12479
155. Sprung R, Nandi A, Chen Y, Chan Kim S, Barma D, Falck JR, Zhao Y (2005) J Proteome Res 4:950
156. Nandi A, Sprung R, Barma DK, Zhao Y, Kim SC, Falck JR, Zhao Y (2006) Anal Chem 78:452
157. Wells L, Vosseller K, Cole RN, Cronshaw JM, Matunis MJ, Hart GW (2002) Mol Cell Proteomics 1:791
158. Hadfield AF, Mella SL, Sartorelli AC (1983) J Pharm Sci 72:748
159. Schwartz EL, Hadfield AF, Brown AE, Sartorelli AC (1983) Biochim Biophys Acta 762:489
160. Seidenfaden R, Krauter A, Schertzinger F, Gerardy-Schahn R, Hildebrandt H (2003) Mol Cell Biol 23:5908
161. Okayama M, Kimata K, Suzuki S (1973) Biochem J 74:1069
162. Robinson HC, Brett MJ, Tralaggan PJ, Lowther DA, Okayama M (1975) Biochem J 148:25
163. Freeze HH, Sampath D, Varki A (1993) J Biol Chem 268:1618
164. Gouyer V, Leteurtre E, Zanetta JP, Lesuffleur T, Delannoy P, Huet G (2001) Front Biosci 6:d1235
165. Zanetta JP, Gouyer V, Maes E, Pons A, Hemon B, Zweibaum A, Delannoy P, Huet G (2000) Glycobiology 10:565

166. Agard NJ, Baskin JM, Prescher JA, Lo A, Bertozzi CR (2006) ACS Chemical Biology 1:644
167. Winans KA, Bertozzi CR (1998) Chem Biol 5:R313
168. Rideout D (1986) Science 233:561
169. Rideout D, Calogeropoulou T, Jaworski J, McCarthy M (1990) Biopolymers 29:247
170. Grogan MJ, Pratt MR, Marcaurelle LA, Bertozzi CR (2002) Annu Rev Biochem 71:593
171. Pratt MR, Bertozzi CR (2005) Chem Soc Rev 34:58
172. Sears P, Wong CH (2001) Science 291:2344
173. Guo Z, Shao N (2005) Med Res Rev 25:655
174. Love KR, Seeberger PH (2002) Angew Chem Int Ed 41:3583
175. Biskup MB, Müller JU, Weingart R, Schmidt RR (2005) ChemBioChem 6:1007
176. Puri KD, Chen S, Springer TA (1998) Nature 392:930
177. De Bank PA, Kellam B, Kendall DA, Shakesheff KM (2003) Biotechnol Bioeng 81:800
178. Nauman DA, Bertozzi CR (2001) Biochim Biophys Acta 1568:147
179. Lee JH, Baker TJ, Mahal LK, Zabner J, Bertozzi CR, Wiemar DF, Welsh MJ (1999) J Biol Chem 274:21878
180. Yarema KJ (2002) In: Al-Rubeai M (ed) Cell Engineering 3. Glycosylation. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp 171
181. Sampathkumar SG, Yarema KJ (2007) In: Kumar CSSR (ed) Nanomaterials for Cancer Diagnosis. Wiley, Hoboken, NJ, pp 1
182. Link AJ, Vink MK, Tirrell DA (2004) J Am Chem Soc 126:10598
183. Iwasaki Y, Tabata E, Kurita K, Akiyoshi K (2005) Bioconjug Chem 16:567
184. Tornøe CW, Christensen C, Meldal M (2002) J Org Chem 67:3057
185. Rostovtsev VV, Green LG, Fokin VV, Sharpless KB (2002) Angew Chem Int Ed 41:2596
186. Pozsgay V, Vieira NE, Yergey A (2002) Org Lett 4:3191

10.4 Glycomics and Mass Spectrometry

Anne Dell, *Jihye Jang-Lee*, *Poh-Choo Pang*, *Simon Parry*, *Mark Sutton-Smith*, *Berangere Tissot*, *Howard R. Morris*, *Maria Panico*, *Stuart M. Haslam*
Biopolymer Mass Spectrometry Group, Division of Molecular Biosciences, Imperial College London, SW7 2AZ London, UK
a.dell@imperial.ac.uk

In: *Glycoscience*. Fraser-Reid B, Tatsuta K, Thiem J (eds)
Chapter-DOI 10-1007/978-3-540-30429-6_56:

Abstract

There is an increasing body of evidence indicating that glycans are implicated in numerous biological processes such as cell–cell interactions, intracellular signaling, and immune response. Glycomics emerges from the necessity to understand the mechanisms underlying the interactions responsible for these activities. The term glycomics is used to describe experimental approaches to studying the structure and function of the glycomes of fluids, cells, tissues, organs etc. Glycomics embraces a variety of technologies amongst which mass spectrometry (MS) plays a pivotal role because it is the method of choice for defining the primary structures of glycopolymers. This chapter provides an insight into MS-based structural strategies that are best suited to studying complex glycomes.

Keywords

Glycomics; Glycopeptides; Glycoproteomics; Mass spectrometry; Nano liquid chromatography; *N*-glycans; *O*-glycans; Profiling; Structural analysis; Tandem mass spectrometry

Abbreviations

Asn	asparagine
CAD	collisionally activated decomposition
CAD ESI-MS/MS	collisionally activated decomposition tandem MS
CID	collisionally induced decomposition
CHCA	α-cyano-4-hydroxycinnamic acid
Da	Dalton
DE	delayed extraction
DHB	2,5-dihydroxybenzoic acid / gentisic acid
ECD	electron capture dissociation
EI	electron impact/ionization
ESI	electrospray ionization
ESI-MS	electrospray ionization mass spectrometry
eV	electron volt
FAB	fast atom bombardment
FAB-MS	fast atom bombardment mass spectrometry
FT-ICR	Fourier transform ion cyclotron resonance
Fuc	fucose
Gal	galactose
GalNAc	*N*-acetylgalactosamine
GC-MS	gaschromatography–mass spectrometry
Glc	glucose
GlcNAc	*N*-acetylglucosamine
Hex	hexose
HexNAc	*N*-acetylhexosamine
HF	hydrofluoric acid
Hz	Hertz
kDa	kilo-Dalton

kV	kilo-Volt
LC	liquid chromatography
LC (ESI)-MS	liquid chromatography (electrospray ionization) mass spectrometry
m/z	mass / charge
MALDI	matrix-assisted laser desorption ionization
MALDI-MS	matrix-assisted laser desorption ionization mass spectrometry
MALDI-TOFMS	matrix-assisted laser desorption ionization– time-of-flight mass spectrometry
Man	mannose
MS	mass spectrometry
MS/MS	tandem mass spectrometry
MS^n	multiple-stage mass spectrometry
µl	micro-Liter
NeuAc	*N*-acetylneuraminic acid
NeuGc	*N*-glycolylneuraminic acid
Nd-YAG	neodymium-doped yttrium aluminum garnet
PNGase F	peptide: *N*-glycosidase F / *N*-glycanase
ppm	part per million
Q-TOF	quadrupole/orthogonal acceleration time-of-flight
Ser	serine
TFA	trifluoroacetic acid
Thr	threonine
TOF	time-of-flight

1 Overview

The repertoire of glycans in an organism or part thereof constitutes its glycome. The term glycomics is used to describe experimental approaches to studying the structure and function of the glycomes of fluids, cells, tissues, organs etc. Glycomics embraces a variety of technologies amongst which mass spectrometry (MS) plays a pivotal role because it is the method of choice for defining the primary structures of glycopolymers. Ultra-high sensitivity, coupled with the ability to characterize individual components within a complex mixture, are key features of mass spectrometry that are especially advantageous in glycomic research. The first part of this chapter provides an overview of the MS techniques that are widely used in glycan analysis. Subsequent sections provide an insight into MS-based structural strategies that are best suited to studying complex glycomes. Informatic tools that are being developed to support glycomic research are described in the following chapter.

1.1 Components of a Mass Spectrometer

At a fundamental level, mass spectrometers generate ions, separate them according to their mass to charge ratio, and measure their abundance. Hence, the three main components of mass spectrometers are (i) an ionization source, where samples are ionized, (ii) one or more

mass analyzers, where ions are separated according to their mass to charge ratio (m/z), and (iii) a detector, where signals are recorded, generating a mass spectrum consisting of the relative abundance and m/z ratio. Additionally, the instruments may include collision cells for producing fragment ions from selected parent ions.

1.2 Ionization Techniques

From as far back as the 1960s, structural studies on glycans were carried out using electron impact (EI) and chemical ionization (CI) techniques. Limitations of these techniques were that they are largely restricted to the analysis of volatile, low molecular weight molecules (less than 1000 Da). Nevertheless, the EI method coupled with gas chromatography remains an indispensable tool for compositional and linkage analyses [1] and is an integral component of many glycomic strategies.

During the 1970s, the development of field desorption (FD) ionization [2], allowed scientists to study non-volatile compounds by mass spectrometry for the first time. Moreover, the coupling of this ionization technique with a new generation of high mass analyzers [3] allowed the analyses of biological compounds, including glycoconjugates, whose masses were significantly in excess of 1000 Da [4]. However, FD was a very challenging technique and very few laboratories had the necessary expertise for its successful exploitation. Not surprisingly, therefore, FD rapidly became obsolete when more tractable methods for ionizing non-volatile substances were developed. The first of these was fast atom bombardment (FAB) which was developed at the end of the 1970s [5] and soon revolutionized the structural analyses of complex glycans [6,7,8,9]. MS sequencing of large glycans was feasible for the first time leading to an exponential increase in the knowledge of glycan structures. Subsequently matrix-assisted laser desorption ionization (MALDI) [10,11] and electrospray ionization (ESI) [12,13] became the ionization methods of choice for analyzing glycans. The principles of FAB, MALDI, and ESI are outlined in the following sections.

1.2.1 Fast Atom Bombardment (FAB)

In the FAB experiment the sample is suspended in a non-volatile, viscous liquid such as glycerol or thioglycerol referred to as the matrix, and is bombarded with a high energy atom or ion beam. The purpose of the matrix is (i) to replenish the surface and thereby minimize rapid destruction of the sample, and (ii) to dissipate excess energy from the atom or ion beam preventing extensive fragmentation of the sample. The sample-matrix mixture is loaded on a metal target that is attached on a probe tip and inserted directly into the FAB source, which is held at high vacuum. A beam of accelerated atoms (argon or xenon) or ions (cesium) is fired at the target from an atom or ion gun. This impact transfers energy to the sample molecules in the matrix allowing them to sputter into the high vacuum of the ion source. During this process, the majority of the molecules couple with ions such as H^+, Na^+, and K^+ generating singly charged species, e. g., $[M+H]^+$ and $[M+Na]^+$. FAB-MS is capable of analyzing glycans of up to about 6000 Da yielding both molecular and fragment ions in a single experiment. Thus, it affords an efficient and rapid method for mapping both terminal epitopes and molecular ions within a complex mixture of glycans. In recent years, due to advances in MALDI and ESI instrumentation (see below), FAB-MS has been phased out in many laboratories.

Nevertheless an appreciation of FAB research is very helpful because the fragmentation pathways observed in FAB spectra are recapitulated in MALDI and ESI experiments. Moreover, many current glycomic protocols have been adapted from FAB-based methodologies developed in the 1980s and 1990s to study complex mixtures of glycans isolated from cells and tissues [14,15,16,17,18].

1.2.2 Matrix-Assisted Laser Desorption Ionization (MALDI)

In MALDI-MS, the sample is co-crystallized with a low molecular weight ultraviolet-absorbing matrix on a metal target which is introduced into the high-vacuum MALDI source. The sample/matrix mixture is irradiated by laser pulses that result in the accumulation of energy in the condensed phase and cause desorption of ionized molecules. The ionization process is still poorly understood. However, it is widely believed that the matrix allows the energy from the laser to be dissipated and assists the formation of ions by proton transfer and chemical processes between the photo-excited matrix and the analyte [19]. Commonly used matrices include 2,5-dihydroxybenzoic acid (DHB) which is ideal for native and derivatized glycans, α-cyano-4-hydroxy cinnamic acid (CHCA) for peptides/glycopeptides, and sinapinic acid for the analysis of larger proteins and for glycoprotein analysis. Most commercial MALDI-MS instruments use ultraviolet lasers, with the nitrogen laser, which has a wavelength of 337 nm and pulse rates of 2–20 Hz [20], being the most widely used due to its low cost and effectiveness. More recently, a diode-pumped neodymium-doped yttrium aluminum garnet (Nd-YAG) laser operating at 355 nm between 20 Hz and 2 kHz [21] has been introduced in MALDI instruments.

MALDI-MS is now a firmly established and highly sensitive ionization technique used in the study of glycans. It is a "softer" ionization method than FAB and yields a series of singly charged molecular ions with minimal fragmentation, making it an ideal method for obtaining an overall glycan profile, i. e., a fingerprint of glycans in a sample mixture. The upper mass limit for successful analysis by MALDI is considerably higher than for FAB and glycopolymers well in excess of 10,000 Da are amenable to this technology [22,23].

1.2.3 Electrospray Ionization (ESI)

The electrospray ionization (ESI) method was originally described by Malcolm Dole and co-workers [24] but it took another 20 years to develop a practical system. Yamashita and Fenn [25,26] and Alexandrov and co-workers [12] pioneered these later developments. The ESI technique is based on producing a spray by applying a high electrical field, under atmospheric pressure, to a liquid passing through a capillary tube. Applying a potential difference of 3–6 kV between the capillary and the counter electrode generates an electric field, which induces a charge accumulation at the liquid surface situated at the end of the capillary. This causes the surface to break and form an aerosol of highly charged micro droplets. These droplets then traverse a series of skimmers where they encounter a drying gas. As the solvent evaporates, the charged micro droplets decrease in size whilst the charge density at the surface of the droplet increases until they reach the Raleigh stability limit (1.1×10^5 charges). At this point, droplets explode due to the mutual repulsion of charges. This process is repeated until the net effect is the production of gaseous ions devoid of solvent. In contrast to the singly

charged molecular ions that are formed in FAB and MALDI experiments, the charge status of ESI-generated ions mirror charges prior to spraying. Consequently multiply charged molecular ions are observed in ESI spectra of most biological molecules.

The nano-electrospray (nanoES) source is essentially a miniaturized version of the ES source. This technique allows very small amounts of sample to be ionized efficiently at nanoliters per minute flow rates and it involves loading sample volumes of 1–2 µl into a gold-coated capillary needle, which is introduced to the ion source. Alternatively for on-line nanoLC-MS experiments the end of the nanoLC column serves as the nanospray needle. The nanoES source disperses the liquid analyte entirely by electrostatic means [27] and does not require assistance such as solvent pumps or nebulizing gas. This improves sample desolvation and ionization and sample loading can be made to last 30 minutes or more. Also, the creation of nanodroplets means a high surface area to volume ratio allowing the efficient use of the sample without losses. Additionally, the introduction of the Z-spray ion source on some instruments has enabled an increase in sensitivity. In a Z-spray ion source, the analyte ions follow a Z-shaped trajectory between the inlet tube to the final skimmer which differs from the linear trajectory of a conventional inlet. This allows ions to be diverted from neutral molecules such as solvents and buffers, resulting in enhanced sensitivity.

1.3 Mass Analyzers

Once ions are formed, mass analyzers are required to separate the ions according to their mass to charge ratio. A number of mass analyzers are commonly used in biopolymer mass spectrometry including double focusing analyzers, quadrupoles, time-of-flight (TOF), quadrupole ion traps, and Fourier transform ion cyclotron resonance (FT-ICR). The majority of FAB research was carried out on instrumentation employing double focusing analyzers which consist of a magnetic sector, which separates ions according to their mass to charge ratio, and an electrostatic analyzer that focuses ions according to their kinetic energy. These analyzers provided excellent mass accuracy and resolution but, like FAB, they have been largely superseded by instrumentation described in the following sections which is cheaper and/or more sensitive and/or more versatile.

1.3.1 Time-of-Flight (TOF) Analyzers

In time-of-flight (TOF) analyzers, as the name implies, the separation and measurement of ions is based on the fact that ions of different mass and the same kinetic energy will travel through a field-free region at different velocities. Most MALDI instruments have TOF analyzers because the pulsed method of ionization is ideally suited to TOF analysis. Ions formed in the source are simultaneously accelerated and extracted into a linear, field-free flight tube. The mass-to-charge ratios are determined by the time that ions take to travel through this field-free region. A major drawback of a linear TOF instrument is its poor mass resolution due to the flight time variations of molecules with the same m/z ratio. There are several factors contributing to this phenomenon; one of them is the inherent length of the ion formation process resulting in ions being introduced in the TOF tube at various times and thus introducing a significant variability on the time of flight measurement. In addition, the ionization process also adds an initial kinetic energy before acceleration leading to a variation of the final kinetic

energy of the ions. Moreover, the varying spatial positions of each ion within the ionization chamber can also lead to errors in the time of flight calculation. To ameliorate the problem of poor resolution "Delayed Extraction (DE)" of ions from the MALDI source was introduced. Under DE conditions a delay between hundreds of nanoseconds and several microseconds is allowed after the laser pulse, prior to applying a voltage pulse to extract the ions from the source. Applying the pulse after a delay transmits more energy to ions that have remained for a longer period of time in the source allowing them to catch up with the more energetic ions at the entry to the flight tube.

Another great improvement of TOF analyzers is the use of a series of electrostatic mirrors, also called a reflectron. The reflectron is used to focus ions of the same m/z but with different kinetic energies, thus acting as an ion "retarding field". Ions with the same m/z ratio but a higher kinetic energy arrive at the reflectron earlier, penetrate more deeply, and leave later than ions of the same m/z ratio but less kinetic energy. The result is ions of the same m/z but different kinetic energies reach the detector simultaneously, therefore significantly improving the mass resolution. High mass, slow moving ions cannot penetrate the reflectron so the high resolution reflectron mode of a TOF analyzer is restricted to analyses below masses of about 10,000 Da.

The main strengths of the TOF are its almost unlimited mass range, and its very high sensitivity achieved because scanning of the analyzer is not required for m/z selection. Also hundreds of samples can be robotically loaded on each MALDI plate thus enabling high throughput analysis. Indeed, MALDI-TOF is an indispensable tool in glycomic analyses. More recently, tandem MALDI-TOF/TOF mass spectrometers have been developed which are proving to be enormously powerful for glycomics because of their sequencing capabilities.

1.3.2 Quadrupole Analyzers

In quadrupole analyzers the ions are separated according to their m/z ratio using the stability of their trajectories in oscillating electrical fields. A quadrupole analyzer consists of four rods with circular or hyperbolic cross-section arranged perfectly in parallel to each other. A direct current voltage (U) and an oscillating radio-frequency (V) are applied to each pair of diagonally opposite rods creating an electric field. Rods adjacent to each other have opposite current polarity. Scanning the voltages allows ions of successively higher or lower m/z to follow a stable path through the analyzer. Therefore, the rods in a quadrupole mass analyzer act as mass filters. Quadrupole mass analyzers have several advantages including fast scanning speed, high robustness, and low cost. They can be operated at relatively high pressure and are therefore especially suitable for ESI instrumentation. Also, EI sources for GC-MS analysis are usually coupled with quadrupole analyzers. A potential weakness of quadrupole mass analyzers is their upper m/z limit of about 4000. Fortunately this is not usually a problem for ESI-MS because larger molecules tend to carry more charges allowing their m/z values to fall within this range.

1.3.3 Quadrupole Ion Traps

The operating principle of a quadrupole ion trap mass analyzer is similar to that of a standard quadruople but the geometry is different. A trap consists of three hyperbolic electrodes comprising a ring and two endcap electrodes. Applying voltages to these electrodes, results

in a 3D quadrupole within the trap. As the name implies, ions are held or "trapped" in the small area between the end caps and the center of the ring electrode. Ions pass into the trap, are stored for a period of time (usually μsec), and are ejected out to the detector by varying the voltages on the entrance and exit of the trap electrodes. The ability to "trap" ions means that a sufficient number of ions can be stored for subsequent MS/MS analysis with high detection sensitivity. However, in practice, the volume in which ions are trapped tends to be small and storage of an excessive number of ions can cause them to leave the trap without being detected. Recent developments in ion trap technology are helping to overcome this problem by the incorporation of linear ion traps [28,29]. The linear geometry provides a larger space in which ions can be accumulated. An increased space decreases the charge-state effect, thus allowing significant gain of sensitivity. Recent studies using linear ion trap instrumentation are clearly demonstrating its potential importance for glycomic research [30].

1.3.4 Fourier Transform Ion Cyclotron Resonance (FT-ICR)

The Fourier transform ion cyclotron resonance analyzer is the most powerful ion trap employed in mass spectrometry. It determines *m*/*z* of ions based on the cyclotron frequency of ions trapped within the magnetic field of a static superconducting high field magnet (typically 4.7 to 13 Tesla). When the ions pass from the source into the ICR cell (also called a Penning trap) they are bent by the Lorentz Force into a circular motion in a plane perpendicular to the field. They are prevented from processing out of the cell by trapping plates at each end. Trapped ions are then excited to a large and coherent cyclotron orbit by an oscillating electric field perpendicular to the magnetic field. An image current generated by the ions as they drop back to their natural orbit can be measured and a Fourier transform is applied to give a mass spectrum. FT-ICR instruments can achieve higher resolution and higher mass accuracy than any other mass spectrometers and, like ion traps, multiple stages of MS can be achieved. However they are hugely expensive and the maintenance of the high magnetic field can be difficult. Therefore this technology has been limited to only a few specialist laboratories and is not yet widely exploited for the structural analyses of glycans.

1.4 Tandem Mass Spectrometry (MS/MS)

MALDI-MS and ESI-MS experiments yield very few fragment ions because the ionization process is very gentle. Although this means they are highly suited to mass fingerprinting of glycan mixtures which is very valuable for initial stages of glycomic analysis, a lack of fragment ions can be a problem for rigorous sequencing. Fortunately, this can be solved with appropriate instrumentation by making the molecular ions more energetic thereby inducing fragmentation. This is achieved by employing two mass analyzers in tandem (MS/MS) or by using ion trap instruments. In the former case, molecular ions are selected by the first analyzer and passed into a pressurized collision cell containing an inert gas such as argon or xenon. The selected ions undergo collision induced dissociation (CID) [also known as collisionally activated decomposition (CAD)]. This promotes fragmentation of the ion of interest and the resulting fragments are subsequently separated by the second mass analyzer. A great variety of tandem instruments have been employed in biopolymer research including triple quadrupoles [31], quadrupole/orthogonal acceleration time-of-flight (Q-TOF) [3], and TOF/TOF instrumenta-

tion [32]. The latter two are currently the analyzers of choice for ESI- and MALDI- MS/MS, respectively.

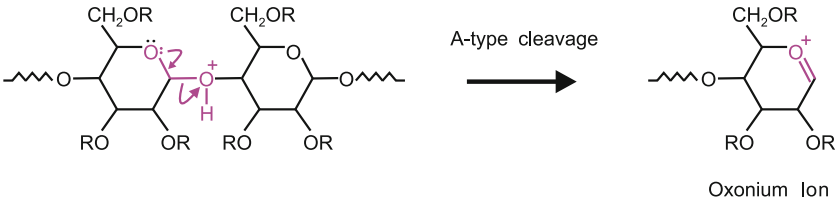
When ion trap instrumentation is used for MS/MS analysis a second mass analyzer is not required since the trap serves as the analyzer for both the parent and fragment ions and as the collision chamber. After selection of the parent ion the trap is pressurized with an inert gas, typically helium. Fragment ions produced by CID are trapped and then sequentially ejected from the trap to produce a mass spectrum. One of the advantages of the ion trap is that the nature of trapping and ejecting allows fragment ions to be selectively isolated and further fragmented allowing so called MS^n experiments. This can be helpful for the assignment of linkages and branching patterns of glycans. Another advantage of ion trap instrumentation for MS/MS experiments is the applicability of alternative methods to CID for energizing the parent ions. These include irradiating trapped ions with an infra-red multi-photon laser, and electron capture dissociation (ECD) in which one of the positive charges on a multiply charged ion is quenched by an electron beam resulting in an energetic radical cation. These methods often produce different fragment ions from those afforded by CID which is helpful for certain structural studies. A limitation of many ion traps is that a low-mass cut-off point is created according to the different parameters used. No ions below that cut-off point will be trapped and therefore important fragment ions at low mass may not be observed. Ions above the cut-off point will be trapped but with a decrease in trapping efficiency for larger *m*/*z* values. In addition, traditional ion traps have a limited space, hence a low ion capacity, which results in poor sensitivity compared to tandem instruments like the Q-TOF. However, the geometry of the new generation of linear ion traps (see above) eliminates this cut-off problem and this instrumentation is therefore promising for future glycomic research.

1.5 Derivatization of Glycans for Structural Analyses

Purified, intact glycans can be analyzed directly by MS, but, because they do not ionize as efficiently as other molecules such as peptides [33], derivatization prior to MS analysis is the preferred strategy in glycomic analysis due to the enhanced sensitivity achievable with suitable derivatives [34,35,36,37]. Permethylation gives the smallest increase in molecular weight of the sample, and the greatest enhancement in sensitivity. In addition, permethylated glycans fragment very selectively resulting in a limited number of structurally diagnostic fragment ions which are relatively easy to interpret [38], and ions arising from single and multiple cleavage events can be readily discriminated. Moreover, permethylation of glycans is a prerequisite for linkage analysis [1]. Hence permethylation steps are integral to many glycomic investigations.

1.6 Fragmentation of Carbohydrates

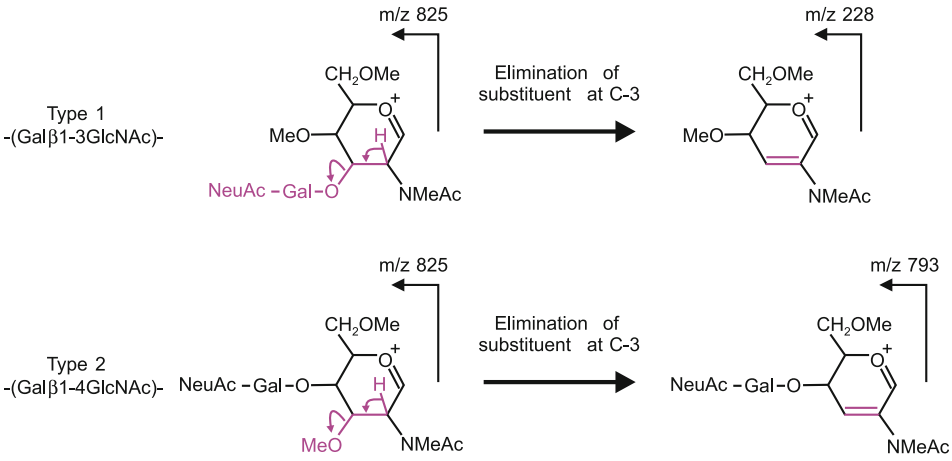
To interpret spectra produced in MS/MS experiments, it is necessary to understand the fragmentation behavior of glycans. The majority of fragment ions are produced by cleavage on either side of glycosidic bonds to yield non-reducing and reducing-end fragment ions, respectively. These ions provide sequence and branching information. Ring fragmentation can also occur especially for smaller glycans, provided that MS/MS experiments are carried out under conditions that impart sufficient internal energy, for example by high energy collisional acti-



Scheme 1
Glycosidic cleavage to form an oxonium ion

vation. Ring cleavages are helpful for assigning linkages. The most common fragmentation pathways are:

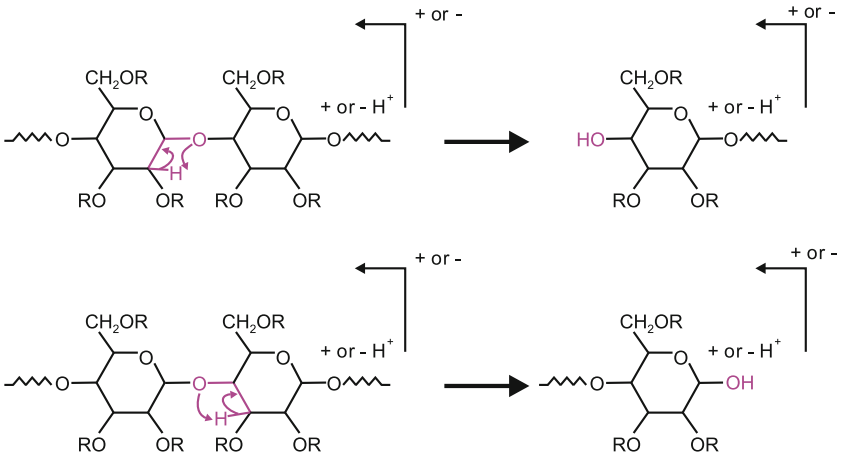
(1) Cleavage of positively charged parent ions on the non-reducing side of the glycosidic bond to form an oxonium ion which may carry sodium if the parent ion was sodiated prior to cleavage (*Scheme 1*). This type of fragmentation is often referred to as A-type cleavage. A-type cleavage is favored at HexNAc residues and is often accompanied by secondary fragmentation involving β-elimination of the substituent at position three on the ring (*Scheme 1.1*).



Scheme 1.1
***β*-Elimination from the 3-position of oxonium ions**

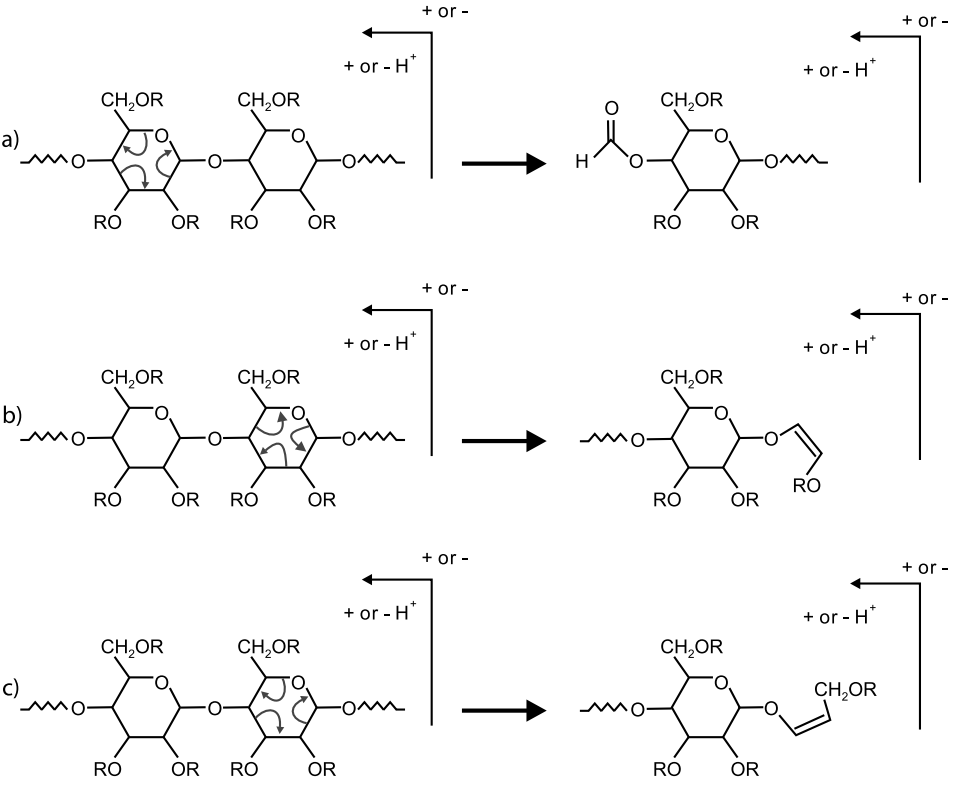
(2) Glycosidic cleavage accompanied by a hydrogen transfer with charge residing on either the reducing or non-reducing ends depending on which bond to the glycosidic oxygen is cleaved. This fragmentation is often referred to as β-cleavage and occurs in both positive and negative modes (*Scheme 2*).

(3) Ring cleavage. The charge can remain on either the reducing or the non-reducing end, depending on the nature of the sample and whether positive or negative ions are being detected (*Scheme 3*).



Scheme 2

Glycosidic cleavages with hydrogen transfer, charge retained on the reducing end (upper) or non-reducing end (lower); sometimes called β-cleavages. In the positive mode, the ions may be protonated as shown or sodiated



Scheme 3

Examples of ring cleavages

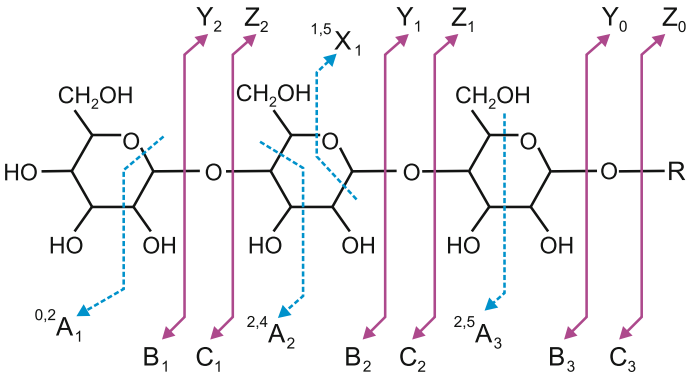


Figure 1
The systematic nomenclature introduced by Domon and Costello [39]. A, B, and C refer to non-reducing terminal fragments whereas X, Y and Z represent reducing terminal fragments

Systematic nomenclature for the fragment ions shown in the above Schemes is depicted in *Fig. 1* [39].

1.7 Enzymatic and Chemical Treatments of Glycans

Another important tool in the structural analyses of glycans is the use of enzymatic and chemical treatments in addition to MS screening and MS/MS sequencing experiments. These treatments provide important information on structural features such as the stereochemistry, monosaccharide unit identity, and linkage. Prior to carrying out these experiments, data derived from glycan profiling and MS/MS analyses are used as a guide for the choice of the appropriate enzyme(s) and/or chemical treatment(s).

Enzymatic degradations are relatively simple experiments from which a significant amount of structural information is obtained. There are various enzymes available. These include exo-glycosidases that remove non-reducing end terminal sugar residues, endoglycosidases which can cleave internal sugar residues, and glycosyltransferases that add sugars to non-reducing positions. These enzymes are usually specific in terms of sugar type, linkages, and anomeric conformation of the sugars they remove or extend. Thus, comparing glycan profiles before and after enzymatic reactions will identify components with specified terminal sugars, or defined internal sequences in the case of endoglycosidases. Chemical treatments compatible with facile MS analysis include hydrofluoric (HF) acid and trifluoroacetic (TFA) acid treatments. HF removes $\alpha 1 \rightarrow 3$ linked fucose residues rapidly whereas $\alpha 1 \rightarrow 2$, 4 and 6 linked fucose are released at a slower rate [40]. Treatment with TFA releases acid labile sugars including sialic acids.

Although some commercially available enzymes are linkage specific, the majority of linkages cannot be defined solely using this approach and/or using the chemical approaches. Hence, linkage analysis experiments are often necessary for rigorous structure analysis.

1.8 Linkage Analyses by Gas Chromatography Mass Spectrometry (GC-MS)

Linkage analysis of partially methylated alditol acetates by GC-MS, coupled with the techniques described earlier, facilitates the assignment of all aspects of the primary structure of glycans. Most importantly, it allows the identification of monosaccharide constituents and glycosidic bond positions. In this technique, glycans are first permethylated in order to derivatize hydroxyl and amide groups. The methyl groups act as permanent tags for hydroxyl groups that are not involved in ring formation or linkage with another sugar residue. The derivatized glycans are then subjected to acid hydrolysis releasing the partially methylated monosaccharides which are subsequently converted to alditols. Hydroxyl groups are then *O*-acetylated thereby providing labels that indicate the positions of the former linkages. The partially methylated alditol acetate (PMAA) sugar derivatives are then analyzed by GC-MS. Monosaccharides can be identified by comparison of their retention times on the GC column and the EI-MS spectra with those of known standards, and quantitation is achieved by comparing peak areas in the ion chromatograms. Linkage analysis is an ideal experiment to carry out after MS and MS/MS analyses of permethylated glycans providing enough material is available.

2 MS Strategies for Glycomics

As pointed out at the beginning of this chapter, the term "glycome" is used to denote the complement of glycans in a cell line, organism etc. Thus, strictly speaking, the term "glycomics" should mean the study of the full complement of glycans from a defined source. In practice, because glycomic investigations are still in their infancy, they are usually confined to studies of subsets of glycans, for example the neutral and sialylated *N*-, *O*-, or lipid-linked glycans present in a particular cell-type or tissue. Over the past decade high sensitivity MS strategies have been devised for screening populations of *N*- and *O*-glycans isolated from a diverse range of biological material, including body fluids, secretions, organs and cultured cell lines. More recently these protocols have been adapted to include glycolipid glycans [41,42]. These glycomic strategies are delineated in ❯ *Fig. 2*. First, glycolipids and glycoproteins are extracted from the biological sample and separately partitioned. Glycans are liberated from the glycolipid fraction using ceramide glycanase, and *N*- and *O*-glycans are released sequentially from the glycoprotein fraction by PNGase F and reductive elimination, respectively. A portion of each glycan pool is then analyzed by MS after permethylation. Putative structures are assigned to each molecular ion based on the usually unique glycan composition for a given mass and prior knowledge of biosynthetic pathways. This is referred to as "mass mapping" and is most conveniently carried out using MALDI ionization because of its very high sensitivity. An algorithm called Cartoonist has been developed to assist data interpretation ([43]; see ❯ Chap. 10.5). Assignments are confirmed and further refined by MS/MS experiments in which molecular ions are selected for collisional activation. Further structural information is provided by linkage analyses and by MS and MS/MS experiments on chemical and enzymatic degradations, the choice of which is guided by the sequence information provided by the initial mass mapping and MS/MS experiments.

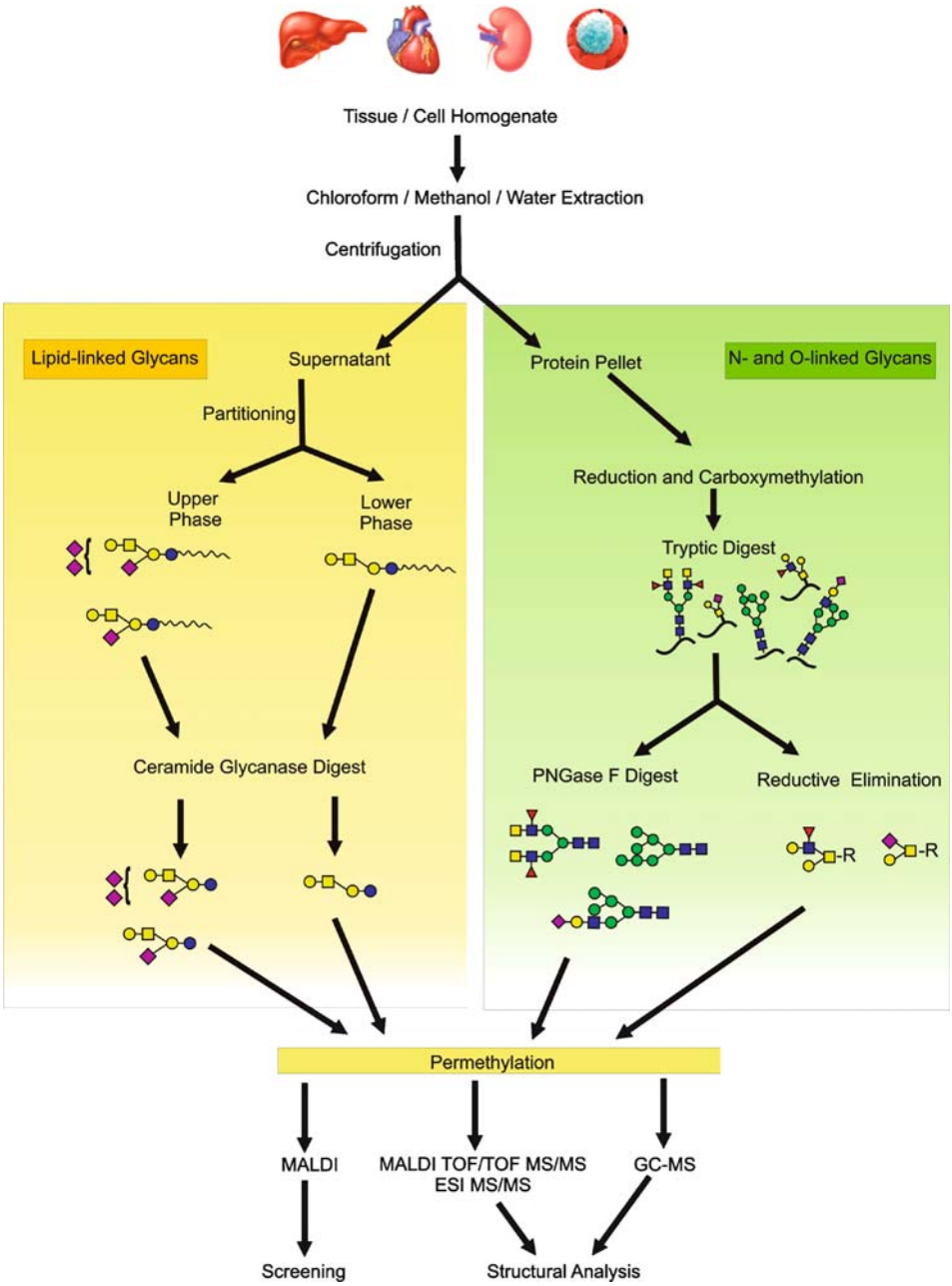


Figure 2

Schematic representation of MS strategies for glycomics. Glycosphingolipids and glycoproteins are extracted from homogenates of biological samples. Glycans are released from the glycosphingolipid fraction using ceramide glycanase and *N*- and *O*-glycans are released from the glycoprotein fraction by PNGase F digestion and reductive elimination, respectively. An aliquot of each glycan pool is permethylated and initially screened by MALDI-TOF MS. R: reduced

These methodologies are illustrated by data from a glycomics analysis of the mouse spleen which is described below.

2.1 Example 1: Characterization of Spleens from Wild-Type and Mgat4a Knock-Out Mice

2.1.1 MALDI-TOF MS Profile

The gene *Mgat4a* encodes the enzyme GlcNAcT-IVa. This enzyme transfers a GlcNAc residue via a $\beta 1 \rightarrow 4$ linkage to the 3-arm of the *N*-glycan trimannosyl core subsequently generating tri- and tetra-antennary *N*-glycans. Knocking out this gene is thought to preclude the formation of tri- and tetra-antennary structures. The MALDI-TOF profiles of the *N*-glycans derived from wild-type and *Mgat4a* knockout mouse spleens are shown in ❯ *Fig. 3* (key symbols are shown in ❯ *Table 1*). In both cases, the molecular ions observed are consistent with high mannose and bi-, tri-, and tetra-antennary complex *N*-glycans. The complex *N*-glycans are predominantly core fucosylated and their terminal structures contain NeuAc/NeuGc residues and Gal-α-Gal epitopes. The most abundant complex structure is a bi-antennary core fucosylated *N*-glycan carrying NeuGc residues on both of the antennae (*m/z* 3026, $NeuGc_2Fuc_1Hex_5HexNAc_4$). On the basis of the MALDI-TOF profiles, the gene that has been knocked out, *Mgat4a*, appears to be compensated for by at least one other gene leading to the very similar *N*-glycan fingerprints.

2.1.2 MALDI-TOF/TOF MS/MS Experiment (*m/z* 2489)

To further characterize the tentative structures predicted from MALDI-MS data, components observed in the wild-type and knockout mouse spleen were subjected to MS/MS analyses using MALDI-TOF/TOF instrumentation. These experiments provide information on features like branching patterns, non-reducing end terminal structures, and sometimes linkage positions. ❯ *Figure 4* illustrates the MALDI-TOF/TOF spectrum of the molecular ion at *m/z* 2489. Major assignments are annotated and fragmentation patterns are shown schematically.
The main signals observed following fragmentation of this molecular ion correspond to independent β-cleavages of HexNAc (*m/z* 2230), HexHexNAc (*m/z* 2026), and $Hex_2HexNAc$ (*m/z* 1822) illustrating the main non-reducing end terminal epitopes. These assignments are further corroborated by the observation of the non-reducing end fragment ions at *m/z* 486 (HexHexNAc) and *m/z* 690 ($Hex_2HexNAc$). Furthermore, the fragment ion at *m/z* 1577 corresponds to the single β-cleavage of $Hex_2HexNAc_2$, confirming a tandem HexHexNAc repeat. The signal at *m/z* 472 indicates that the terminal Hex residue on the $Hex_2HexNAc$ non-reducing epitope is linked to a sub-terminal Hex residue. This is consistent with the presence of Gal$\alpha 1 \rightarrow$ 3Gal$\beta 1 \rightarrow$ 4GlcNAc epitopes. Overall, the data indicate the presence of a mixture of different isomers, as shown in ❯ *Fig. 4*. Diagnostic signals for each isomer are indicated.

2.1.3 ESI-MS/MS Experiment (*m/z* 2651)

ESI-MS/MS is another powerful tool for performing tandem mass spectrometry experiments. To get sequence information on the *N*-glycan at *m/z* 2651 ($FucHex_7HexNAc_4$; ❯ *Fig. 2*), the

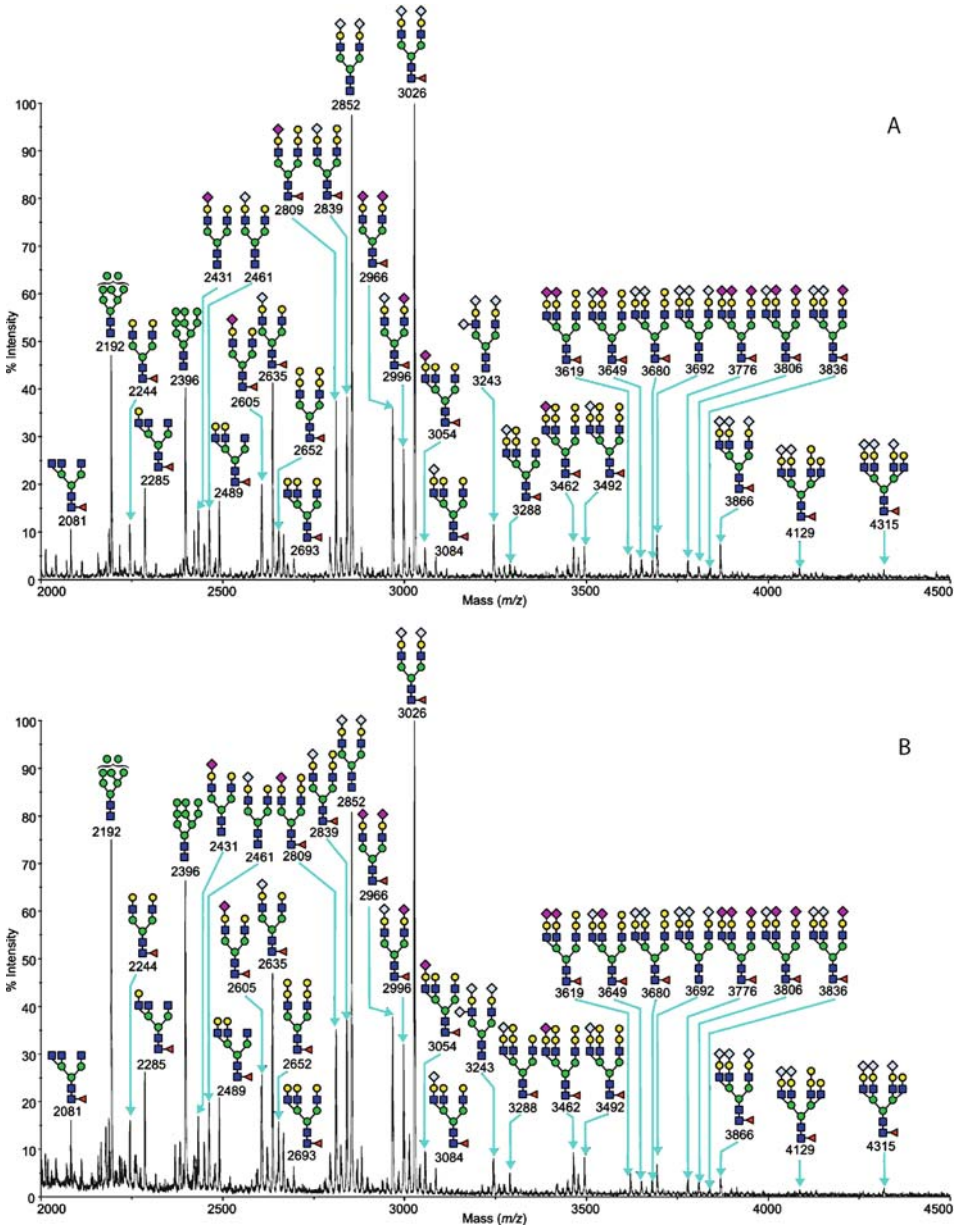


Figure 3

MALDI-MS profiles of the wildtype (panel A) and knockout (panel B) mouse spleen. Putative structures based on knowledge of *N*-glycan biosynthetic pathways and glycan composition are shown. For convenience only one branching pattern for tri-antennary structures is shown. Monosaccharide symbols are shown in *Table 1*

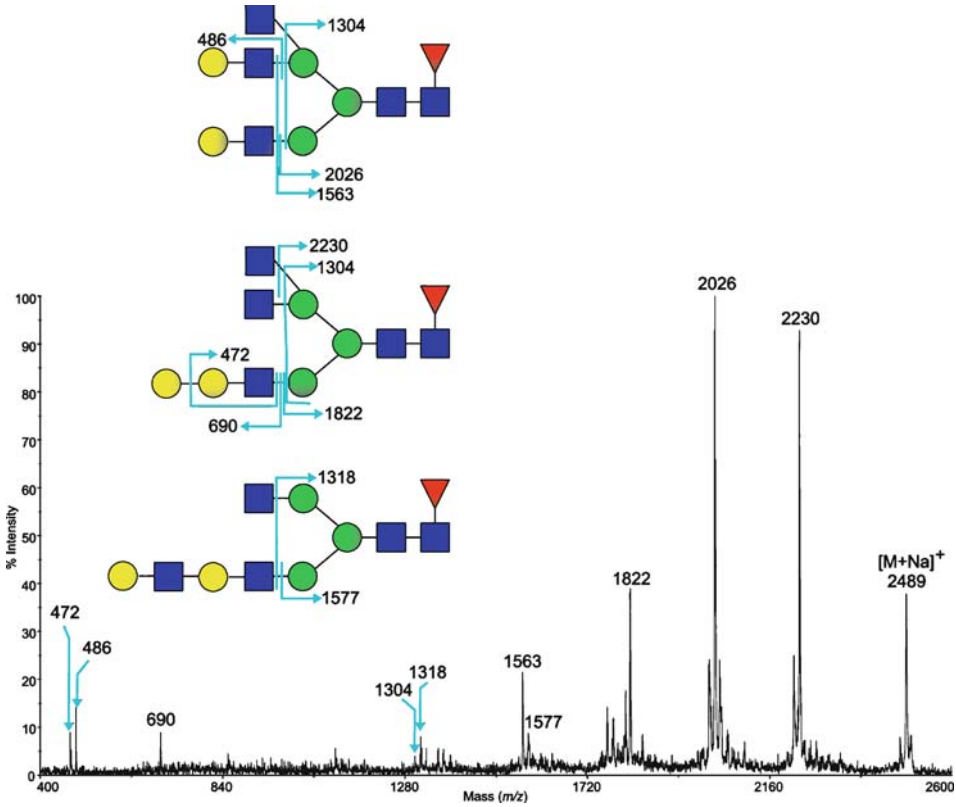


Figure 4
MALDI-TOF/TOF spectrum of the molecular ion at *m/z* 2489. For key to symbols see *Table 1*

doubly charged sodiated molecular ion $[M+2Na]^{2+}$ (m/z 1337^{2+}) was chosen for MS/MS analyses (*Fig. 5*).

Following fragmentation of the molecular ion, a major signal at m/z 690 corresponding to the non-reducing end structure $Hex_2HexNAc$ established that this is the major non-reducing end epitope of the *N*-glycan. Additionally, the presence of m/z 463 (Hex_2) indicates that the two Hex residues are present in tandem. The signal at m/z 533 represents a cross-ring cleavage and is indicative of a 4-linkage between the sub-terminal Hex and the HexNAc residue. The data therefore provide good evidence for Galα1 → 3Galβ1 → 4GlcNAc epitopes.

Major signals corresponding to a single β-cleavage of $Hex_2HexNAc_1$ (m/z 1984 and 1003^{2+}) were observed. Additionally, a fragment ion produced as a result of a double β-cleavage of two $Hex_2HexNAc_1$ (m/z 1318) epitopes was detected indicating that this is a bi-antennary *N*-glycan carrying $Hex_2HexNAc_1$ on each antenna. Hence, it can be concluded that this *N*-glycan has a core fucose (shown by m/z 474) and two antennae, each carrying a Galα1 → 3Galβ1 → 4GlcNAc epitope.

Monosaccharide	Symbol	Permethylated mass	Perdeuteromethylated mass
Deoxyhexose			
Fucose (Fuc)	▼	174	180
Hexose			
Mannose (Man)	●	204	213
Galactose (Gal)	●	204	213
Glucose (Glu)	●	204	213
N-acetylhexosamine			
N-acetylgalactosamine (GalNAc)	■	245	254
N-acetylglucosamine (GlcNAc)	■	245	254
Sialic acid			
N-acetylneuraminic acid (NeuAc)	◆	361	376
N-glycolylneuraminic acid (NeuGc)	◇	391	409
Reduced sugars			
Hexitol	○-R	220	223
HexNAcitol	□-R	261	264

Table 1

The increment mass of each sugar. Sugar symbols used throughout this chapter are those employed by the Consortium for Functional Glycomics (http://functionalglycomics.org). R: reduced

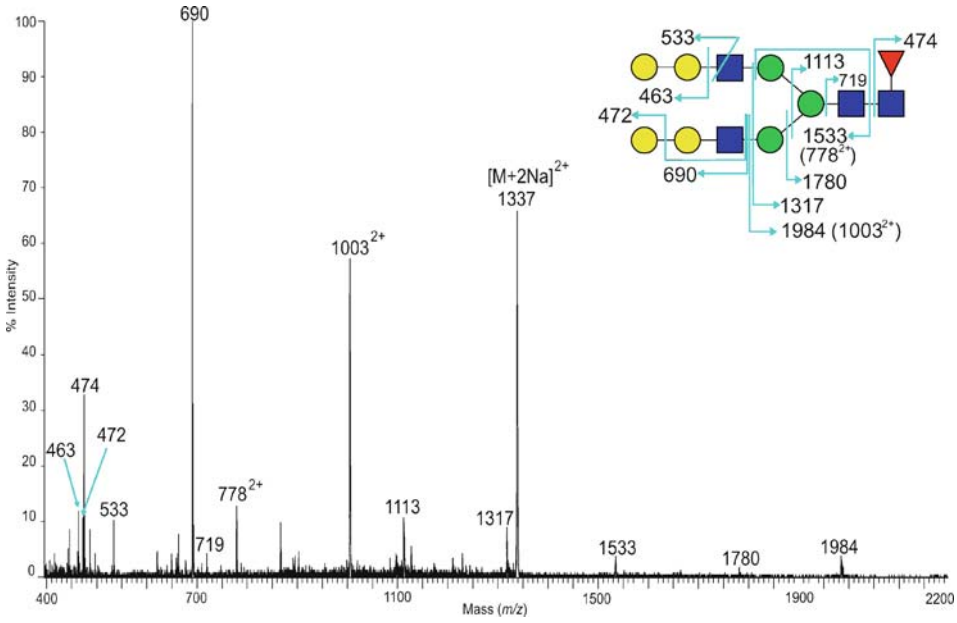


Figure 5

ESI-MS/MS spectrum of the molecular ion at *m/z* 2651. Monosaccharide symbols are shown in *Table 1*

2.1.4 Enzyme Digest—α-Galactosidase Digestion

To confirm anomeric configurations and the nature of monosaccharide constituents, enzymatic digestions were carried out. In this case an α-Galactosidase digestion was performed. α-Galactosidase is an exoglycosidase that removes terminal α-linked Gal residues from the non-reducing end of glycans. After α-Galactosidase treatment, an aliquot of the digested sample was permethylated and analyzed by MALDI-MS. *Figure 6* illustrates the spectra before (panel A) and after (panel B) digestion.

The signals at *m/z* 2448, 2652, 2665, 2809, 2839, 2897, and 3101, which were present in the glycan mixture before digestion, were absent in the spectrum of the digestion products. These results confirmed the presence of terminal α-linked Gal residue(s) in the structures corresponding to these signals. The remaining complex type structures were insensitive to digestion consistent with their tentative assignments.

2.1.5 Linkage Analysis

Partially methylated alditol acetates were prepared from the permethylated *N*-glycans of the wild-type and knockout mouse spleens, and subsequently analyzed by GC-MS (*Table 2*). The data shows both samples are composed of similar components and are consistent with the structural assignments shown in *Fig. 3*. In particular, 2-linked Man is present at high levels, consistent with the high abundance of bi-antennary *N*-glycans whereas 2,4- and 2,6-linked Man support the presence of tri- and tetra-antennary *N*-glycans. Bisected *N*-glycans are minor components of the mouse spleen, corroborated by the low levels of 3,4,6-linked GlcNAc. The absence of Le^x structures is further supported by the absence of 3,4-linked GlcNAc, while the 3-linked Gal is attributed to Gal$\alpha 1 \rightarrow$ 3Gal epitopes as well as sialylated glycans. The presence of 6-linked Gal indicates that some of the sialic acids are $\alpha 2 \rightarrow 6$ linked to subterminal galactose.

2.1.6 Summary

The strategy described above demonstrates the power of mass spectrometry as a tool for glycomics. From a single mouse spleen, *N*-glycans can be purified, derivatized, and profiled by MALDI-MS to gain information on their compositions. Based on these compositions and the well-defined biosynthetic pathways of *N*-glycans putative structures can be proposed for each molecular ion. Further evidence from linkage analysis (GC-MS), tandem mass spectrometry (MALDI-TOF/TOF MS/MS and/or ESI-MS/MS), and enzyme digests can be used to further refine these putative assignments. The example above is a comparison between a spleen from a wild-type mouse and a spleen from a mouse which has had the *Mgat4a* gene knocked out. The strategy outlined here quickly demonstrated that the *N*-glycans derived from the two spleens were very similar, leading to the conclusion that other genes can compensate for the action of *Mgat4a* in spleen.

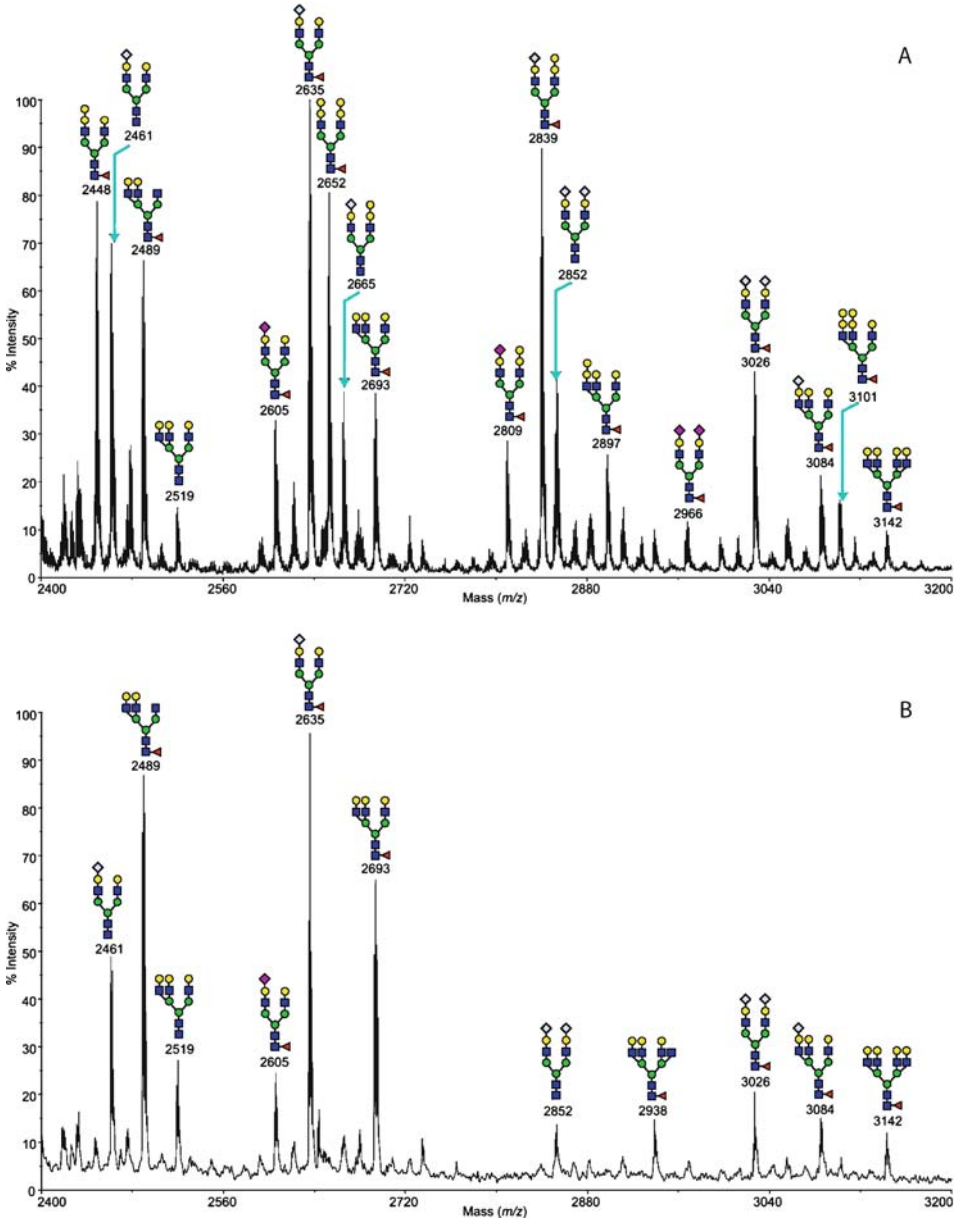


Figure 6

MALDI-MS spectra of released *N*-glycans before (panel A) and after (panel B) α-Galactosidase digestion. Monosaccharide symbols are shown in *Table 1*

Table 2
GC-MS linkage analysis of partially methylated alditol acetates derived from the 50% (v:v) aqueous acetonitrile fractions of the permethylated *N*-glycans derived from the wildtype and Mgat4a knockout mouse spleen

Elution time (min.) (WT)	Elution time (min.) (KO)	Characteristic Fragment Ions	Assignments	RA (WT)	RA (KO)
17,501	17,462	102, 115, 118, 131, 162, 175	Terminal Fucose	0,22	0,24
18,989	18,969	102, 118, 129, 145, 161, 205	Terminal Mannose	1	0,9
19,324	19,236	102, 118, 129, 145, 161, 205	Terminal Galactose	0,14	0,18
20,143	20,125	129, 130, 161, 190, 204, 234	2-linked Mannose	0,44	1
20,436	20,418	118, 129, 161, 202, 234	3-linked Mannose	0,03	0,05
20,576	20,558	118, 129, 161, 203, 234, 277	3-linked Galactose	0,1	0,3
20,958	20,939	99, 102, 118, 129, 162, 189, 233	6-linked Galactose	0,1	0,18
21,404	21,299	87, 88, 99, 113, 130, 190, 233	2,4-linked Mannose	0,03	0,04
21,739	21,711	87, 88, 129, 130, 189, 190	2,6-linked Mannose	0,04	0,09
21,888	21,872	118, 129, 189, 202, 234	3,6-linked Mannose	0,39	0,75
22,376	22,326	118, 139, 259, 333	3,4,6-linked Mannose	0,03	0,06
22,861	22,827	117, 129, 145, 205, 247	Terminal GlcNAc	0,01	0,03
23,711	23,694	117, 159, 346	4-linked GlcNAc	0,15	0,38
25,058	24,531	117, 159, 261	4,6-linked GlcNAc	0,02	0,04

RA, relative abundance; WT, wildtype; KO, knockout

3 Glycoproteomics

Whereas glycomics creates an inventory of all the glycans present in a tissue or an organ with no respect to the molecules carrying these glycans, glycoproteomics determines the position and the structure of each glycan carried by a specific protein. In order to achieve this level of structural characterization several experiments are required. The glycoprotein of interest is first isolated from a cell lysate, tissue or a protein mixture by methods such as affinity or lectin chromatography, immunoprecipitation, or electrophoresis. If sufficient material is available, direct MALDI-MS or ESI-MS profiling of the intact glycoprotein can be performed to determine the level of glycosylation. However, this profiling may require a lot of sample and the quality of the results largely depends on factors such as the molecular weight and heterogeneity of the glycoprotein and its ability to ionize efficiently. Therefore digestion to form glycopeptides of manageable size is the preferred route for analysis.

Putative *N*- and *O*-glycosylation sites of a protein of interest can be determined from the amino acid sequence. *N*-glycosylation sites can be identified by the consensus sequences Asn-Xaa-Ser/Thr (where Xaa cannot be Pro), whereas prediction of likely *O*-glycosylation sites is possible in favorable cases using specialist NetOGlyc software (http://www.cbs.dtu.dk/services/NetOGlyc/). Once these putative sites have been identified, a convenient enzyme digestion strategy is designed to produce glycopeptides with molecular weights that are convenient for mass spectrometric analyses (approximately 2000–5000 Da). Once the enzymatic strategy has been decided, the glycoprotein is digested and glycopeptides are analyzed using online

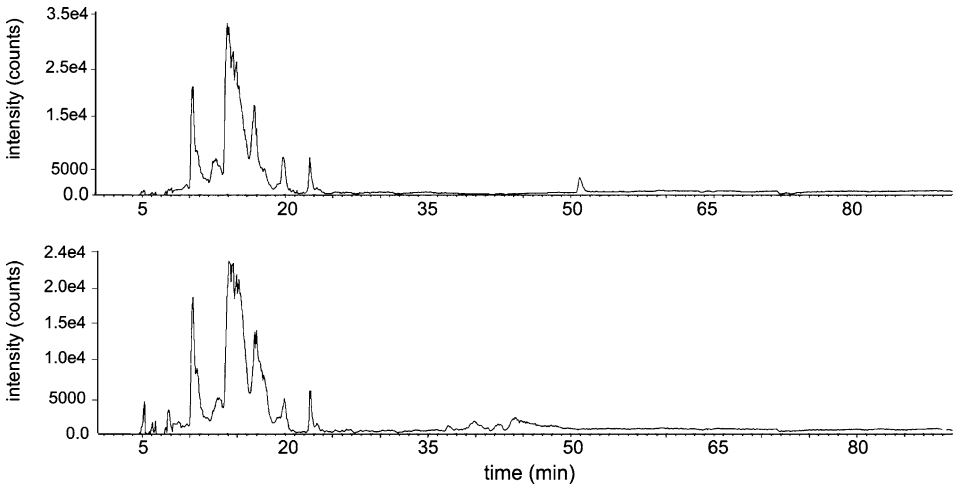


Figure 7

NanoLC-ESI-MS of tryptically digested human IgG. Extracted chromatograms for the two main diagnostic glycan ions, HexHexNAc (*m/z* 366; upper panel) and HexNAc (*m/z* 204; lower panel) are shown. The chromatograms of both ions indicate that most of the glycopeptides elute between 10 and 18 minutes

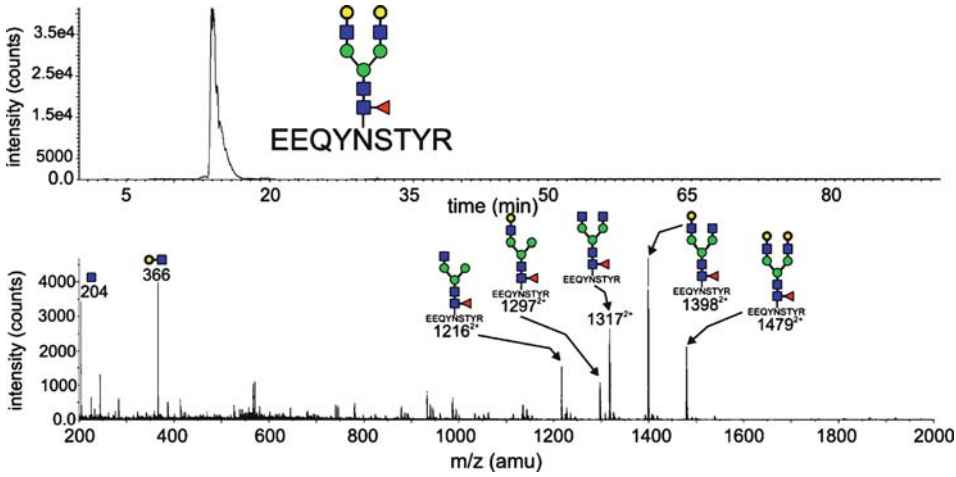
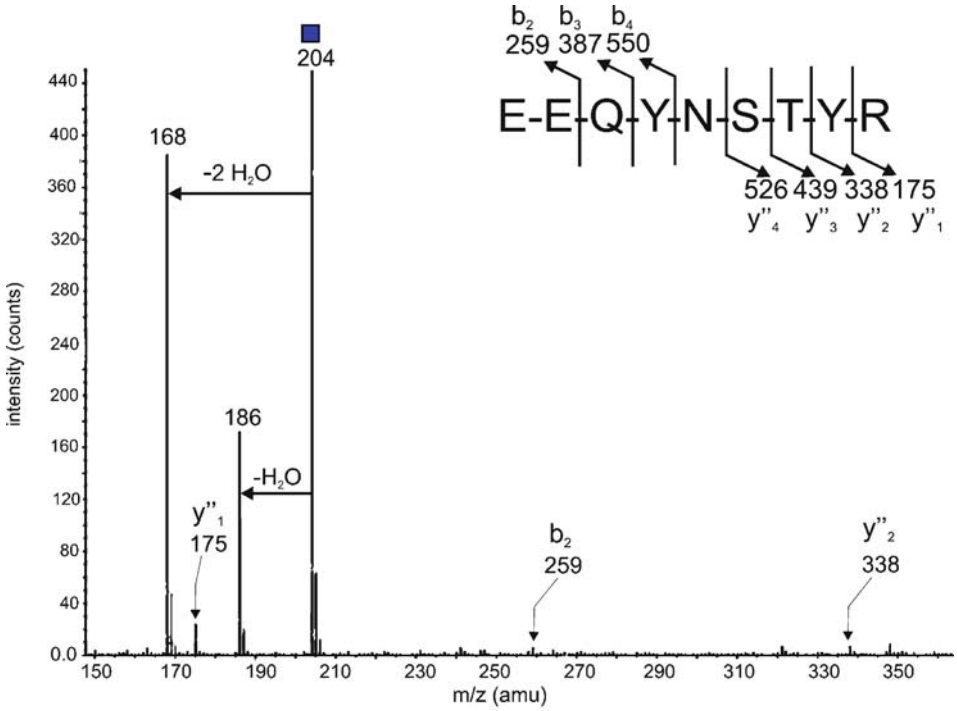


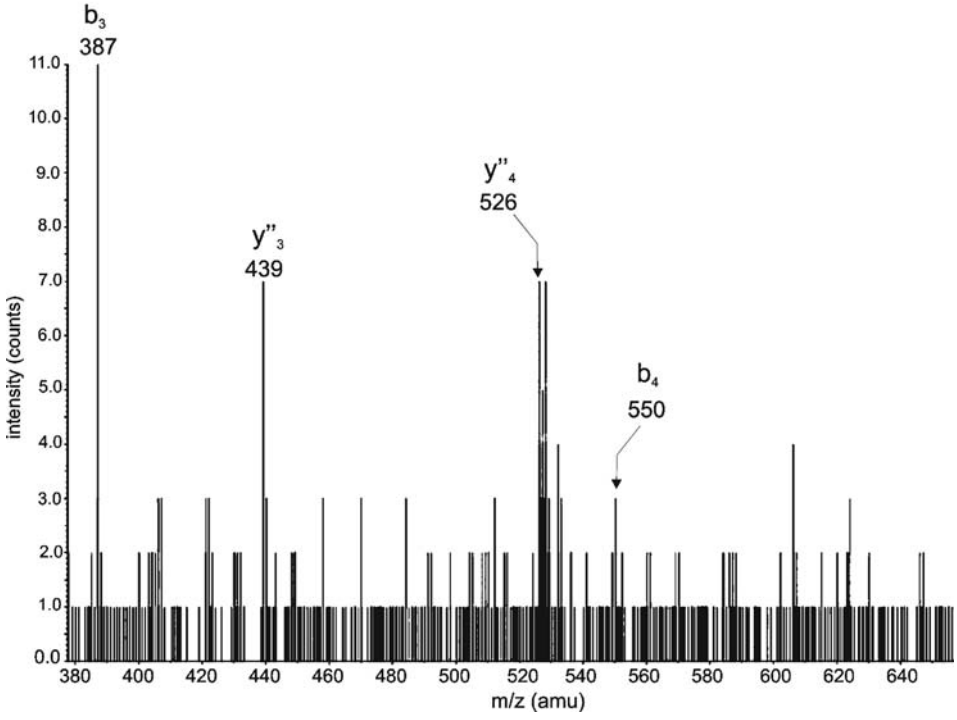
Figure 8

NanoLC-ESI-MS of tryptically digested human IgG. The extracted chromatogram of the ion at *m/z* 1479 shows a signal at 15 minutes (upper panel). The lower panel shows the summation of all spectra acquired between 14 and 16 min. The resulting MS spectrum demonstrates the presence of several glycoforms of the peptide EEQYNSTYR. Monosaccharide symbols are shown in ***Table 1***

Figure 9

Magnified views of the nanoLC-ESI-MS/MS of the glycopeptide signal at *m/z* 13982⁺. The b- and y-ions were used to sequence the peptide (summarised in top panel)


b2 b3 b4
259 387 550
E-E-Q-Y-N-S-T-Y-R
526 439 338 175
y''4 y''3 y''2 y''1
204
168
-2 H2O
186
-H2O
y''1
175
b2
259
y''2
338
intensity (counts)
m/z (amu)


b3
387
y''4
526
y''3
439
b4
550
intensity (counts)
m/z (amu)

nano-LC chromatography coupled to ESI-MS and ESI-MS/MS instrumentation. This type of experiment allows the determination of:

- the occupancy of each of the predicted sites;
- the molecular composition of each glycan at each occupied site and consequently the heterogeneity of glycosylation;
- the complete or partial mapping of the glycoprotein sequence.

Although these methodologies are highly sophisticated, they cannot rigorously define glycan features such as stereochemistry, linkage, and branching patterns. To achieve this detail, the glycans must be released from the purified protein and glycomic analyses such as those described in the previous section must be undertaken. To illustrate a typical glycoproteomic strategy, an analysis of human immunoglobulin G (IgG) is detailed below.

3.1 Example 2: Glycoproteomics Analysis of Human IgG

The putative *N*-glycosylation sites of IgG were determined by searching the amino acid sequence for Asn-Xaa-Ser/Thr motifs. Because of the convenient location of lysine and arginine residues in the regions flanking these sites, trypsin was chosen as the most amenable enzyme for producing glycopeptides that are a convenient size for mass spectrometry. Accordingly, the IgG sample was reduced and carboxymethylated to protect S–S bridges, digested by trypsin and analyzed by nanoLC-ESI-MS(/MS). When glycopeptides are subjected to ESI, partial fragmentation of the glycan moiety results in high abundance $[M+H]^+$ fragment ions at *m/z* 204 (HexNAc) and *m/z* 366 (HexHexNAc). This particular feature is extremely useful for the determination of the time window during which glycopeptides are eluted. *Figure 7* shows the extracted ion chromatograms of *m/z* 204 (HexNAc) and 366 (HexHexNAc) of the nanoLC-ESI-MS experiment.

The region in the chromatogram where the diagnostic fragment ions were observed (between 10 and 18 minutes) was then manually searched for glycopeptides and two major species were detected. A signal at *m/z* 1398^{2+} corresponds to the peptide EEQYNSTYR carrying an *N*-glycan with a molecular composition of $FucHex_4HexNAc_4$, whereas the signal found at *m/z* 1479^{2+} (*Fig. 8*) corresponds to the same peptide carrying an *N*-glycan with a molecular composition of $FucHex_5HexNAc_4$.

To confirm the structure of the peptide carrying the two different glycans, nanoLC-ESI-MS/MS experiments were performed. *Figure 9* shows magnification of the regions of the ESI-MS/MS spectrum where series of y″ and b-ions corresponding to peptide fragments were observed. By combining the information gained from the y″-type ions at *m/z* 175, 338, 439, and 526, and the b-type ions at *m/z* 259, 387, and 550 the peptide backbone of the glycopeptide was determined.

To complement the glycoproteomics data, the IgG *N*-glycans were released, and profiled by MALDI-TOF MS using the glycomics strategy described above (*Fig. 10*).

The most abundant molecular ions were detected at *m/z* 1836, 2040, and 2244 corresponding to core-fucosylated bisected glycans with compositions of $FucHex_3HexNAc_4$, $FucHex_4HexNAc_4$, and $FucHex_5HexNAc_4$, respectively. The high abundance of the glycans at *m/z* 2040 and 2244 correlates with the data obtained from the glycoproteomics study.

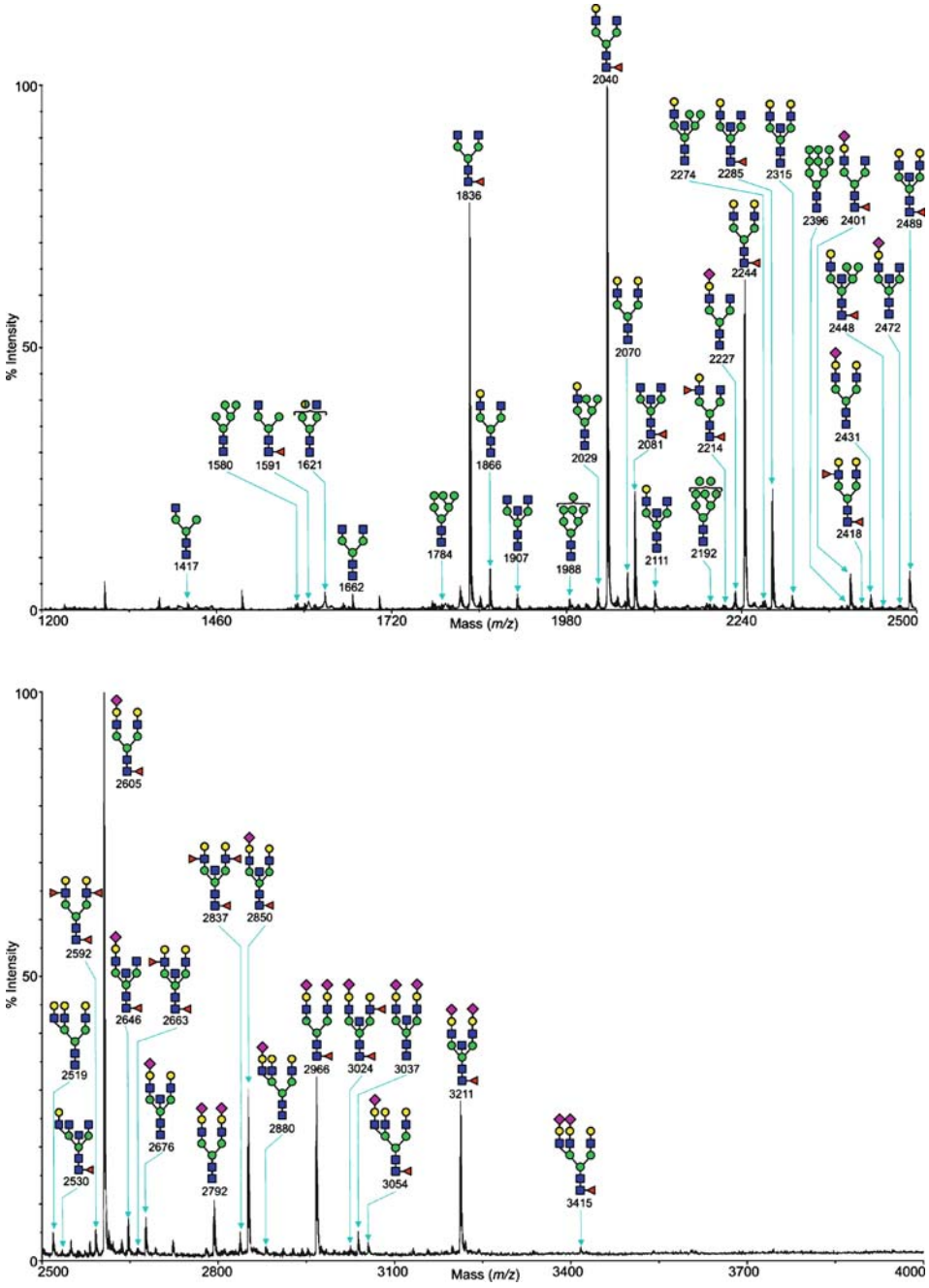


Figure 10

MALDI-MS spectrum of released *N*-glycans from human IgG. Monosaccharide symbols are shown in *Table 1*

4 Conclusion

Over the past decade, mass spectrometry has revolutionized our knowledge of biological macromolecules. Our understanding of the structures and diversity of glycans present in cells and tissues has increased enormously due to the high sensitivity and versatility of these instruments. Today, mass spectrometers are smaller and more user-friendly than ever before and are becoming de rigueur in university departments and pharmaceutical companies throughout the world. This surge in structural information has been enormously beneficial to the glycobiological community, but has created a new problem. The challenge in the future will be how to interpret and present the huge volumes of mass spectrometric data in a useful and timely fashion. Already, the Consortium for Functional Glycomics has begun to catalogue some of this data (http://www.functionalglycomics.org/), and the following chapter will deal with the role of bioinformatic tools in glycomic analysis.

References

1. Albersheim P, Nevins DJ, English PD, Karr A (1967) Carbohydr Res 5:340–345
2. Beckey HD (1969) J. Mass Spectrom Ion Phys 2:500
3. Morris HR, Paxton T, Dell A, Langhorne J, Berg M, Bordoli RS, Hoyes J, Bateman RH (1996) Rap Comm Mass Spec 10:889
4. Linscheid M, D'Angona J, Burlingame AL, Dell A, Ballou CE (1981) Proc Natl Acad Sci USA 78:1471
5. Barber M, Bordoli RS, Garner GV, Gordon DB, Sedgwick RD, Tetler LW, Tyler AN (1981) Biochem J 197:401
6. Ballou CE, Dell A (1985) Carbohydr Res 140:139
7. Dell A, Ballou CE (1983) Biomed Mass Spectrom 10:50
8. Dell A, Morris HR, Egge H, Strecker G (1983) Carbohydr Res 115:41
9. Egge H, Peter-Katalini J (1987) Mass Spectrom Rev 6:331
10. Karas M, Hillenkamp F (1988) Anal Chem 60:2299
11. Tanaka K, Waki H, Ido Y, Akita S, Yoshida Y, Yoshida T (1988) Rap Comm Mass Spec 2:151
12. Alexandrov ML, Gall LN, Krasnov NV, Nikolaev VI, Pavlenko VA, Shkurov VA (1984) Dokl Akad Nauk SSSR 277:379
13. Fenn JB, Mann M, Meng CK, Wong SF, Whitehouse CM (1989) Science 246:64
14. Haslam SM, North SJ, Dell A (2006) Curr Opin Struct Biol 16:584
15. Fukuda M, Dell A, Oates JE, Fukuda MN (1984) J Biol Chem 259:8260
16. Fukuda M, Spooncer E, Oates JE, Dell A, Klock JC (1984) J Biol Chem 259:10925
17. Fukuda MN, Dell A, Oates JE, Fukuda M (1985) J Biol Chem 260:6623
18. Sutton-Smith M, Morris HR, Dell A (2000) Tetrahedron: Asymmetr 11:363
19. Zenobi R, Knochenmuss R (1998) Mass Spectrom Rev 17:337
20. Medzihradszky KF, Campbell JM, Baldwin MA, Falick AM, Juhasz P, Vestal ML, Burlingame AL (2000) Anal Chem 72:552
21. Vestal ML, Campbell JM, Hayden KM, Juhasz PJ (2000) Proc 48th ASMS Conf, Palm Beach, CA
22. Lee BS, Krishnanchettiar S, Lateef SS, Lateef NS, Gupta S (2005) Rapid Commun Mass Spectrom 19:2629
23. Terada M, Khoo KH, Inoue R, Chen CI, Yamada K, Sakaguchi H, Kadowaki N, Ma BY, Oka S, Kawasaki T, Kawasaki N (2005) J Biol Chem 280:10897
24. Dole M, Mack LL, Hines RL, Mobley RC, Ferguson LD, Alice MB (1968) J Chem Phys 49:2240
25. Yamashita M, Fenn JB (1984) J Phys Chem 88:4451
26. Yamashita M, Fenn JB (1984) J Phys Chem 88:4671
27. Wilm M, Mann M (1996) Anal Chem 68
28. Schwartz JC, Senko MW, Syka JE (2002) J Am Soc Mass Spectrom 13:659

29. Song Q, Kothari S, Senko MA, Schwartz JC, Amy JW, Stafford GC, Cooks RG, Ouyang Z (2006) Anal Chem 78:718
30. Wang Y, Wu Sl, Hancock WS (2006) Glycobiology:514
31. Yost RA, Enke CG (1978) J Am Chem Soc 100:2274
32. Vestal ML, Campbell JM (2005) Methods Enzymol 402:79
33. Harvey DJ (1999) Mass Spectrom Rev 18:349
34. Chalabi S, Panico M, Sutton-Smith M, Haslam SM, Patankar MS, Lattanzio FA, Morris HR, Clark GF, Dell A (2006) Biochemistry 45:637
35. Dell A (1990) Methods Enzymol 193:647
36. Dell A, Reason AJ, Khoo KH, Panico M, McDowell RA, Morris HR (1994) Methods Enzymol 230:108
37. Jang-Lee J, North SJ, Sutton-Smith M, Goldberg D, Panico M, Morris H, Haslam S, Dell A (2006) Methods Enzymol 415:59
38. Dell A (1987) Adv Carbohydr Chem Biochem 45:19
39. Domon B, Costello C (1988) Glycoconj J 5:397
40. Haslam SM, Khoo KH, Houston KM, Harnett W, Morris HR, Dell A (1997) Mol Biochem Parasitol 85:53
41. Manzi AE, Norgard-Sumnicht K, Argade S, Marth JD, van Halbeek H, Varki A (2000) Glycobiology 10:669
42. Parry SK, Ledger V, Tissot B, Haslam SM, Dell A (2007) Glycobiology 17:646
43. Goldberg D, Sutton-Smith M, Paulson J, Dell A (2005) Proteomics 5:865

10.5 Informatics Tools for Glycomics: Assisted Interpretation and Annotation of Mass Spectra

Alessio Ceroni[1], *Hiren J. Joshi*[2], *Kai Maaß*[3], *René Ranzinger*[2], *Claus-W. von der Lieth (deceased)*
[1] Imperial College London, Division of Molecular Biosciences, Biopolymer Mass Spectrometry Group, London SW7 2AZ, UK
[2] German Cancer Research Center, Central Spectroscopy B090, 69120 Heidelberg, Germany
[3] Institute of Biochemistry, Faculty of Medicine, University of Giessen, 35392 Giessen, Germany
a.ceroni@imperial.ac.uk, h.joshi@dkfz.de, Kai.maass@biochemie.med.uni-giessen.de, r.ranzinger@dkfz.de

In: *Glycoscience*. Fraser-Reid B, Tatsuta K, Thiem J (eds)
Chapter-DOI 10-1007/978-3-540-30429-6_57:

Abstract

Recent years have seen an increase in both the development and use of informatics tools and databases in glycobiology-based research. Mass spectrometric methods, which are capable of detecting oligosaccharides in the low pico- to femtomole range, are fundamental technologies used in glycan analysis. The availability of robust and reliable algorithms to automatically interpret MS spectra is critical to many glycomic projects. Unfortunately, the current state-of-the-art in glycoinformatics is characterized by the existence of disconnected and incompatible islands of experimental data, resources, and proprietary applications. The development of tools for the robust automatic assignment of glycans on the basis of MS measurements is often hampered by the paucity of available MS data. Here, we review the methodologies for semi-automatic interpretation of MS spectra of glycans, based upon current technology. Three promising approaches are highlighted: (a) combinatorial approaches to the automatic assignment of possible monosaccharide superclass composition—*Glyco-Peakfinder*, (b) the scoring of a set of identified structures with theoretically calculated fragments–*GlycoWorkbench* and (c) the correlation of experimental masses to a database of theoretical fragment masses, in a technique known as Glycofragment Mass Fingerprinting.

Keywords

Informatics; Glycomics; Mass spectrometry; Algorithms; Automatic interpretation; Assignment of fragments

Abbreviations

BMP	windows Bitmap
CID	collision induced dissociation
CFG	Consortium for Functional Glycomics
ESI-MS	electrospray ionization-mass spectrometry
GMF	glycan fragment mass fingerprinting
GLYDE-II	glycan data exchange version II
HTML	hypertext markup language
LID	laser induced dissociation
MALDI	matrix-assisted laser desorption ionization
MS	mass spectrometry
NMR	nuclear magnetic resonance
PDF	portable document format
PMF	peptide mass fingerprinting
SVG	scalable vector graphics
XML	extensible markup language

1 Introduction

1.1 Definitions

The term '*glycomics*' describes the scientific endeavor to identify, characterize, and study all glycan molecules [1,2,3]. Functional glycomics is defined as the systematic study of glycan interaction with other biological macromolecules, often using modern array technologies [2,4,5,6]. More formally, the term '*glycome*' encompasses the molecular families of glycoproteins, proteoglycans, glycolipids, peptidoglycans, lipopolysaccharides, and any other glycoconjugates synthesized by an organism over its lifetime. '*Glycoproteomics*'—the study of the carbohydrates attached to proteins—is regarded as a branch of proteomics, and is often discussed in combination with other post-translational modifications [7,8]. The analytical techniques to analyze glycans are fundamentally similar to those found in proteomics: electrophoresis and chromatography, mass spectrometry, NMR, and bioinformatics.

1.2 Analytical Techniques

Several intrinsic factors have hampered the development of techniques for the structural analysis of carbohydrates. As the biosynthesis of glycans is not a template-driven process, no biological amplification methods exist. Consequently, carbohydrates either have to be analyzed at their physiological concentration, or time-consuming enrichment procedures such as electrophoresis and HPLC have to be applied. MS methods capable of detecting oligosaccharides in the low pico- to femtomole range are currently the workhorse technologies used in glycan profiling and detailed sequencing [7,9,10,11,12,13,14,15,16,17].
NMR methods enable a complete and unambiguous assignment of all structural features of glycans—from the stereochemistry of monosaccharide units, to the type of linkage between connected units and even down to conformational preferences. However, the large amount of purified material required normally excludes NMR from being used in glycomics projects, where only a low amount of material is available [18].

1.3 Implications of the Structural Complexity of Carbohydrates for the Development of Informatics for Glycomics

The buildings blocks of complex carbohydrates consist of monosaccharides, which often have the same molecular weight and chemical constitution (superclasses), differing only in the stereochemistry of the attached hydroxyl groups. MS techniques are not able to distinguish between different stereoisomers like galactose, glucose, and mannose, which manifest the same residue mass. Consequently, seven main mass constituents of glycoprotein oligosaccharides need to be considered in the majority of mammalian analyses for mammals: Hex (mannose, glucose, galactose), HexNAc (*N*-acetylglucosamine, *N*-acetylgalactosamine), dHex (fucose, rhamnose), NeuAc (*N*-acetyl neuraminic acid), NeuGc (*N*-glycolyl neuraminic acid), HexA (glucuronic, galacturonic, and iduronic acids), and S (sulfate).

This intrinsic property of carbohydrates considerably limits the success of any de novo assignment of monosaccharide composition, although such a technique is feasible for amino acids in proteins (contingent on accurate mass values of the peptide precursor ion and a small number of accurate fragment ion mass values) [19].

In proteomics, 19 of the 20 different naturally occurring amino acids can be identified by their mass. Additionally, the diversity of atoms which constitute monosaccharides is restricted such that C-, O-, and H-atoms occur more frequently, while *N*- and *S*-atoms only rarely occur. This further limits the diversity in elemental composition of glycan fragments when viewed in comparison to those of peptides.

The detection of linkages between monomeric units is a cumbersome process, dependent on the presence of key data points, and so other sources of information such as biosynthetic pathways are incorporated in a structural assignment.

Nevertheless, MS-based studies analyzing the *N*-glycan repertoire often provide a list of fully characterized *N*-glycan structures [9,10], where data from other analytical techniques are included as justification for the assignment.

Supplemental information involved in this detailed assignment is the knowledge of *N*-glycosylation pathways in mammalian systems, from which a number of inferences can be made on the unknown structure.

Therefore, the assignment of explicit *N*-glycan structures can be justified when these two independent sources of information are taken into consideration in parallel.

The structural heterogeneity seen in glycans, for example attached to proteins (glycoforms), is enormous. Although the same glycosylation machinery is available to all proteins in a given cell, most glycoproteins emerge with characteristic glycosylation patterns and heterogeneous populations of glycans at each glycosylation site. More than 100 glycoforms have been reported for a single glycosylation site on a protein, and more than glycoforms are often reported [20,21].

1.4 Structural Diversity of Glycans

Since the biosynthesis of glycans is a non-template-driven process, the catalogue of carbohydrate structures is dependent on the activity of the glycosylation machinery at any point of time. Consequently, all glycans existing in a given species can currently only be coarsely estimated. Hence, the only way to describe the range of expressed glycans is to curate collections of previously characterized structures. In contrast to the genomic and proteomic areas, no comprehensive, currently maintained collections summarizing results from literature have been compiled so far [4,22].

2 (Semi)-Automatic Interpretation of Mass Spectra

Interpretation of glycan mass spectrometric data is still an emerging field. Bioinformatic scientists all over the world are attempting to develop a tool to provide complete structural information (i. e. composition, sequence, branching, linkage, and anomeric state) from mass spectra, without the requirement for additional information derived from specific structural experiments. Up till now, several tools have been developed in different laboratories, mostly with the

Table 1
Important URLs

Web-tools (in bold freely available services)			
Name	**Functionality**	**URL**	**Ref.**
GlycoMod	Composition	http://www.expasy.ch/tools/glycomod	[24]
GlycoComp	Composition	http://www.glycosuite.com	
Glyco-Peakfinder	Composition	http://www.eurocarbdb.org/applications/ms-tools/	[25]
GlycoFragment	Annotation Fragmentation	http://www.glycosciences.de/tools/GlycoFragments	[50]
GlycoFrag	Annotation Fragmentation	http://www.glycosuite.com	
GlycanBuilder	Structure-Editor	http://www.eurocarbdb.org/applications/ms-tools/	
GlycosidIQ	GMF	http://www.glycosuite.com	[29]
Glyco-Search-MS	GMF	http://www.glycosciences.de/sweetdb/start.php?action=form_ms_search	[27]
Stand-alone program			
GlycoWorkbench	Annotation Fragmentation	http://www.eurocarbdb.org/applications/ms-tools	
Other mentioned URLs			
CFG Symbolic Representation	http://www.functionalglycomics.org/glycomics/molecule/jsp/carbohydrate/carbMoleculeHome.jsp		
CFG- Glycan Profiling MS	http://www.functionalglycomics.org/glycomics/publicdata/glycoprofiling.jsp		
CFG-Home Page	http://www.functionalglycomics.org/fg/index.shtml		
GLYCOSCIENCES.de	http://www.glycosciences.de		
EUROCarbDB	http://www.eurocarbdb.org		
GlycoCT	http://www.eurocarbdb.org/recommendations/encoding/		
GLYDE-II	http://lsdis.cs.uga.edu/projects/glycomics/		
MonosaccharideDB	http://www.monosaccharidedb.org		

aim to solve the specific problems of the working group, thus limiting the type of structures that can be resolved. Additionally, the tools often lack user-friendly interfaces and therefore require expert knowledge to be utilised. The tools summarized in this section are categorized into two families—one which uses data from single mass spectrometry measurements and the other which additionally uses information from fragmentation experiments (*Table 1*).

2.1 Compositional Analysis

Elucidation of glycan composition is a combinatorial problem where the number of compositions that must be tested scales exponentially with the number of different monomers that can form the solution. Therefore, the composition for molecules of high mass can take a very long time to be derived. The number of solutions will also grow at a similar ratio, thus offering an enormously large amount of alternatives to the user. Furthermore, when adducts and other mass spectrometric losses are taken into account, a number of compositions may be indistinguishable within a relatively low mass delta threshold (less than 0.05 Da). Therefore, it is often

useful to introduce constraints to the calculations to make the problem tractable. Taxonomic and biosynthetic information can be used to limit the compositional searches, eliminating any residue classes that are known not to occur in the system or requiring the presence of essential saccharides in any solution composition.

Several tools capable of calculating the sugar composition corresponding to a target mass have been developed. These tools usually employ a combination of dynamic programming and greedy algorithms to solve the combinatorial problem and adopt different approaches for restricting the number of displayed results, showing only relevant data. *GlycoComp* is a commercial software tool developed by Proteome Systems. In *GlycoComp* the derived composition is used in tandem with statistical information from *GlycoSuiteDB* [23] (also from Proteome Systems) to rank compositions, thus presenting solutions deemed to be biologically likely. *GlycoMod* [24] is a web-based tool for compositional analysis. Composition can be restricted by selecting the type of monomers present and the type of glycan (*N*- or *O*-linked). Also, several derivatizations and reducing-end modifications can be taken into account, and the tool is able to compute glycopeptide composition—given a protein sequence and the type of enzymatic digestion. In addition, the derived composition can be used to search for matching structures in *GlycoSuiteDB* [23]. *Glyco-Peakfinder* [25] is the most recent addition to this category of algorithms. Composition can by derived by taking into account a comprehensive list of saccharides, reducing-end modifications, derivatizations, and glycoconjugates (peptides and lipids). The biological relevance of fragments derived with *Glyco-Peakfinder* can be checked by searching if such a composition exits in a structure database. *Glyco-Peakfinder* supports such a search in an open access database. In contrast to other tools, *Glyco-Peakfinder* can also estimate the composition of fragment ions from their masses, giving useful information for sequence analysis. *Glyco-Peakfinder* is described in detail in a following section.

A different paradigm to composition prediction is adopted by the *Cartoonist* tool [26]. *Cartoonist* generates all the *N*- and *O*-linked glycans that could be possibly synthesized by mammalian cells using a set of archetypal structures, and a set of rules for the modification of said structures. The archetypes and rules have been compiled by a group of experts, and represent the current knowledge about biosynthetic pathways in mammalian organisms. By incorporating this pathway information, and using constructive techniques in building structures, the numbers of matching structures are greatly reduced, and only biologically feasible molecules are suggested.

2.2 Sequence Analysis

In contrast to compositional analysis, sequence analysis requires further data derived from additional experiments. Modern MS techniques are capable of producing mass spectra of fragmented carbohydrate molecules, and suitable algorithms have been developed to exploit this information to resolve glycan structure. Unfortunately, even MS^n spectra can be populated by isobaric ions with distinct structures, given that fragments with equal composition share the same mass. However, cross-ring fragmentation can be useful to distinguish among different branching and linkage positions. Sequencing tools can be divided into three categories: semi-automatic, library based, and de novo.

2.2.1 Semi-Automatic Sequencing

Semi-automatic sequencing tools assist the interpretation of a mass spectrum by evaluating a set of structures proposed by the user against a spectrum. Each structure is used to generate a theoretical peak list from its theoretical fragments. The theoretical peak list is then matched with the real peak list obtained from the mass spectra. A number of semi-automatic sequencing tools exist. *GlycoFrag* is one such tool produced by Proteome Systems. Similarly, *GlycoFragments* [27] is a tool that computes all the theoretical fragments of a given structure. Like *GlycoFrag*, *GlycoFragment* displays all the fragments with their associated masses and allows for derivatized structures. Away from web-based tools, *Sweet Substitute* [28] is a Windows-based software package, complete with a graphical interface for building glycan structures. The software is restricted to consider only mammalian *N*-glycans, with no derivatization and reducing-end modifications. Matching of fragments with the given peak list is done automatically. Finally, *GlycoWorkbench* has an easy to use interface for building glycans that displays the structure using symbolic notations. The tool features a complete list of glycan types, derivatizations, monosaccharides, substituents and reducing-end modifications. Raw spectra can be directly loaded into the tool so the user can pick the interesting peaks directly on the graph. The fragments are automatically matched with the given peak list, allowing for multiple cleavages and cross-ring fragments. More details will be given about the tool in later sections.

2.2.2 Library-Based Sequencing Tools

Library-based sequencing tools identify the glycan sequence by matching the unassigned spectra with data derived from known glycan structures. In mass fingerprinting approaches, a database of known structures is used to generate synthetic peak lists which are then matched to the unassigned spectra. This approach is conceptually similar to the peptide mass fingerprinting (PMF) procedures for protein sequencing. Thus far, a single tool has been proposed that automatically evaluates all the glycans from a structure database on the given MS^n spectra, *GlycosidIQ* [29]. The main feature of *GlycosidIQ* is a proprietary scoring function that intelligently ranks the candidate structures. More details about the algorithm and its scoring function are given in the following sections. A restriction of structure library-based methods is that the structures found by mass fingerprinting methods are limited to those already contained in databases. As mentioned, this can be problematic since no exhaustive collection of experimentally derived glycan sequences exists. A different approach to library-based sequencing is through matching against a library of experimentally determined fragment spectra, instead of synthetic peak lists, as described in detail in [30]. Essential to this method is a comprehensive library of experimental MS^2 and MS^3 spectra—which are obtained in a meticulous fashion. A search algorithm is used to identify structures whose library MS^2 spectrum exhibits high similarity to the unassigned spectrum. Afterwards, the unassigned spectrum is analyzed concerning the potential of its MS^3 ions to function as diagnostic ions between candidate structures. The most promising MS^2 ion is selected for additional fragmentation, and the resulting MS^3 spectrum is matched with the library. This approach is highly dependent on experimentally derived data, and since no public collection of assigned MS^n spectra of pure glycans exists, the general availability of this technique is limited. Furthermore, since experimental conditions must be very similar to those under which the library of spectra was recorded, experimental flexibility is limited.

2.2.3 De Novo Glycan Sequencing

De novo glycan sequencing tools are not restricted to previously characterized structures, and thus represent the most promising family of techniques. Unfortunately, many approaches have been proposed, but no single one has demonstrated the capability to deliver the desired accuracy and flexibility. *STAT* [31] prompts the user to select a composition from amongst the alternatives compatible with the precursor mass. The algorithm then calculates all the possible structural topologies having the given composition. The structures thus produced are evaluated against the given peak list, and ranked accordingly. Like *STAT*, *Oscar* [32] generates candidate structures from an estimated composition. *Oscar* uses the information contained in fragmentation pathways derived from MS analysis of permethylated oligosaccharides as a basis for restriction of the set of possible results. All possible result structures must contain the common *N*-glycan mammalian core. For each peak in the provided MS^n fragmentation pathways, a series of structural inference rules are applied to filter the candidate glycan set by eliminating structurally infeasible results, thus retaining only those candidates that are consistent with all input pathways. *StrOligo* [33] proceeds in the opposite way: the differences between fragment masses are used to estimate the loss of known moieties and to produce a candidate composition for the precursor ion. Only glycosidic cleavages are considered, and the structure must not be derivatized. Given the estimated composition of the precursor, *StrOligo* generates a set of structures by applying biosynthetic rules specific to mammalian *N*-glycans. The topology of the proposed structures is limited by the type of terminals that are tested (lactosamine repeats, capping sialic acid, terminal and core fucosylation). *GLYCH* [34] is a dynamic programming algorithm derived from de novo peptide sequencing programs, modified to allow branching in the polymer structure (only binary branching is considered). The *GLYCH* algorithm performs a maximization of the number of assigned peaks by generating a series of b-ions starting from the leaves of the glycan tree structure. The complete structure is generated from the top-level b-ion. In a final re-evaluation phase, the top scoring results from the optimization procedure are re-ranked according to double cleavages. Lastly, Kameyama and co-workers [35] describe a method for matching the unassigned spectra with simulated fragment mass spectra of arbitrary non-characterized *N*-glycan structures. This method is an improvement over library-based spectral matching algorithms and is not limited by the availability of experimentally derived spectra. The simulated fragment spectra are generated from the observation of characteristic fragmentation patterns for every possible branch type of *N*-linked oligosaccharides. The fragmentation patterns have been previously derived by performing collision-induced dissociation experiments of sets of oligosaccharides complementarily labeled with $^{13}C_6$-D-galactose.

3 Approaches with a Potentially Broad Appeal

In the previous section we have given an overview of approaches and tools for assisted interpretation of MS spectra of glycans. It is obvious that unlike for proteomics, currently no widely available and accepted solution exists for the rapid identification of glycans. Here, we describe in more detail three approaches to assisted interpretation and their implementation in more detail. In the opinion of the authors, these three techniques have the promise to become widely adopted. Two of the techniques, embodied in *Glyco-PeakFinder* and *GlycoWorkbench*, are

freely accessible through the Internet. MS technique agnostic, they provide support for the basic daily work needs of scientists when annotating MS spectra of glycans. The glycan fragment mass fingerprinting (GMF) approach, although unable to demonstrate its full potential because of the lack of comprehensive glycan structure databases, seems also to have a high potential to be broadly, and routinely used as soon as free services will become available. An important prerequisite for the success of these tools is that an efficient and user-friendly integration of the currently independent tools is realized.

3.1 Automatic Compositional Analysis of Glycan MS Spectra

All the necessary information to calculate compositions of glycans (i. e. masses of the residues, mechanisms of bond cleavages, ion masses, etc.) is well known. However, as mentioned earlier, efficient solutions to deal with the combinatorial problems are required. The knowledge of the sequence of the underlying glycan in question is not mandatory for performing a compositional analysis. This can easily be understood by comparing the structure of carbohydrates to those of hydrocarbons. For hydrocarbons, although the number of isomers increases with the number of carbon atoms, all structural isomers of one hydrocarbon have the same atomic composition. Accordingly, carbohydrate structures can be treated as linear sequences of monomers, which allows the calculation of their masses by simple incremental addition of the masses of the residues (mass of monosaccharides minus mass of water molecule, i. e. Hex m = 162 u, HexNAc m = 203 u, dHex m = 146 u, etc.). On the basis of this "core", the masses of whole glycans for MS profiles, different fragment types (A, B, C, X, Y, Z: for definition see ❯ *Fig. 1*) in fragment spectra, reducing-end modifications, derivatized structures can be calculated by adding or subtracting further incremental masses.

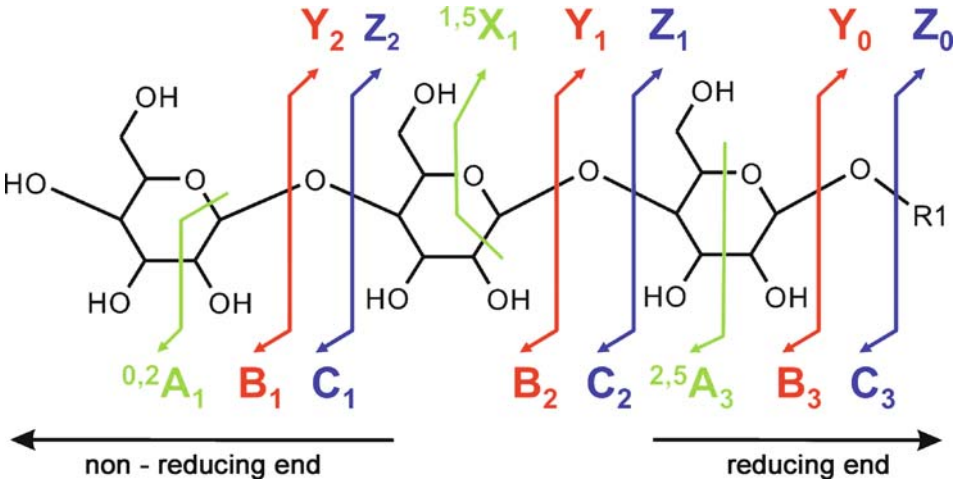


Figure 1
Nomenclature of fragments of carbohydrates as defined by Domon and Costello [47]

Modern MS techniques, such as ESI-MS, often result in multiply charged ions, mostly with ions of different charge states. A fully fledged de novo compositional analysis algorithm should be able to analyze the whole set of signals given from a mixed-stage MS spectra without charge deconvolution in the preprocessing steps.

3.1.1 Glyco-Peakfinder

Glyco-Peakfinder is able to determine the composition for a mass signal independent to the source of spectral data (MALDI, ESI) or fragmentation technique (no fragmentation, LID, CID, ion-trap, etc). This program is able to take into account multiple charges (−4 to +4), multiple cleavages (up to four), all known fragmentation schemes (B-, C-, Y-, Z- and cross ring cleavages A, X: see *Fig. 1*) and modifications (either at the reducing end, or for the whole structure, such as permethylation). The whole calculation is a complete de novo concept and does not need the support of structural databases or any precalculated mass values from (sub-) structures.

Peak lists from mass spectrometric experiments can be uploaded to the web interface of *Glyco-Peakfinder* in several data formats. To decrease the calculation time and number of solutions for each mass value, further options can be set before calculation. The estimated accuracy of the mass signals—representing the experimental technique used—can be given, as well as a choice between monoisotopic and average mass calculation. Depending on the experimental conditions, *Glyco-Peakfinder* can optimize its calculations towards a selection of either glycan profile or fragment spectrum data. Further inputs allow the specification of monosaccharide quantities, including their proposed minimum and maximum occurrence in the structure. As well as the default residues (*Table 2*), up to three new user-specified residues can be used in the calculation.

Advanced settings for the calculation can also be specified in further dialogs. The "ion/charge" settings (see *Fig. 2*, example data were taken from a pyridylamino-oligosaccharide $Hex_3HexNAc_5dHex$-PA fraction obtained from batroxobin of *Bothrops moojeni* venom [36]) refer to experimental technique and fragmentation style: for example, for experimental techniques other than MALDI, ions bearing multiple combinations of charges have to be calculated. Different fragmentation techniques result in different choices for the fragment types and maximum number of cleavages used in the calculation. Finally, the modifications settings enable the calculation of either manually modified structures (i. e. permethylation, reduction, pyridylamination, etc.) or of naturally occurring glycoconjugates, such as glycolipids or glycopeptides.

Table 2
***Glyco-Peakfinder*, a selection of residues**

Classification	Superclasses
Unsubstituted sugars	hexose (galactose, glucose, mannose), pentose (xylose), heptose
Amino sugars	2-aminohexose, 2-*N*-acetylaminohexose (GalNac, GlcNAc)
Desoxy sugars	6-desoxyhexose (fucose), 5-desoxypentose, 4,6-didesoxyhexose (tyvelose)
Others	3-methylhexose, 4-methylhexose, 6-hexuronic acid
Sialic acids	*N*-acetyl sialic acid, *N*-glycolyl sialic acid, KDN, KDO, muramic acid
Substituents	sulfate, phosphate, phosphocholine, pyruvate, acetate

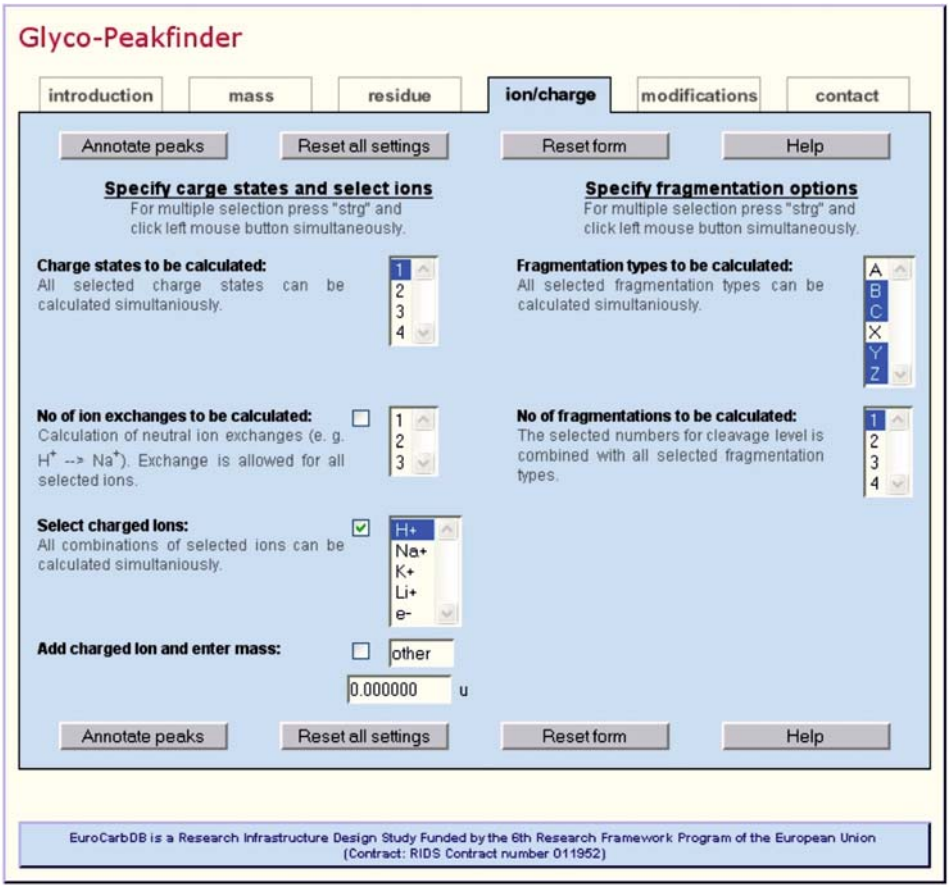
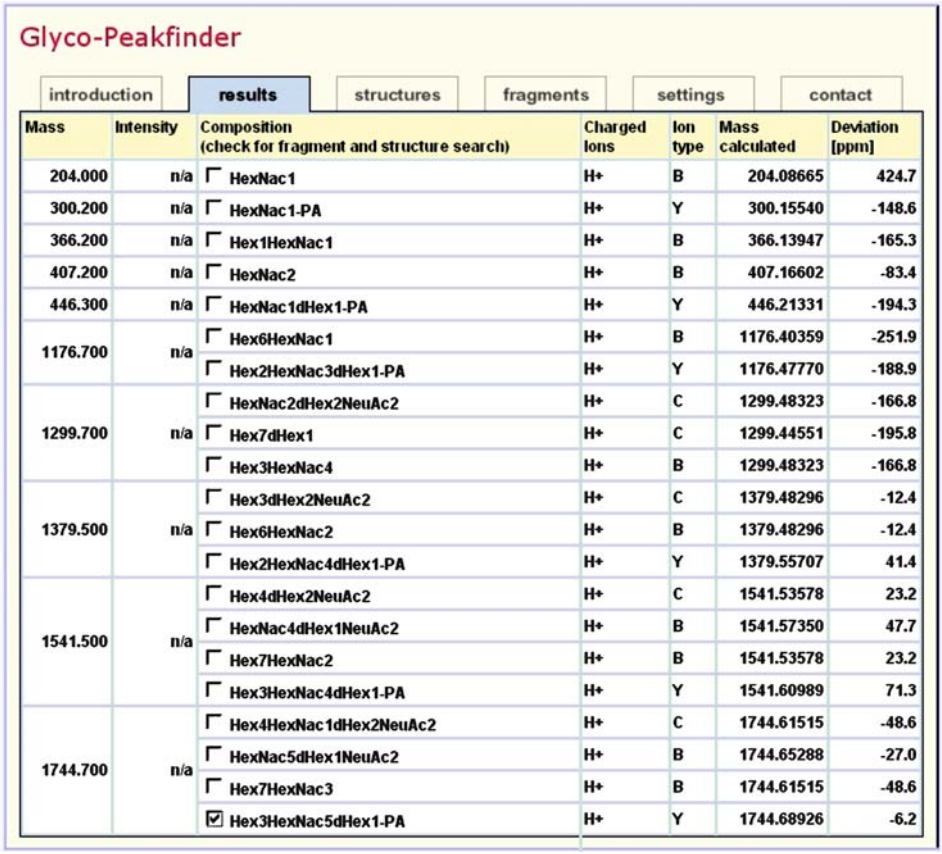


Figure 2

***Glyco-Peakfinder*, ion/charge window. *Glyco-Peakfinder* is a web-based tool for automatic composition analysis of glycan MS spectra. The *Glyco-Peakfinder* ion/charge window allows the specification of parameters, such as charge state or fragment types, to be used for the calculation of possible compositions from the given mass list. The settings in the example were used to calculate compositions from a mass spectrum of a pyridylamino-oligosaccharide $Hex_3HexNAc_5dHex$-PA fraction obtained from batroxobin of *Bothrops moojeni* venom [36]. *Glyco-Peakfinder* can be accessed through the URL: http://www.eurocarbdb.org/applications/ms-tools/**

Starting the calculation leads to a catalogue of possible compositions for all peaks in the mass list according to the specified parameters. Within the results, composition, charged ions, fragment type, and the agreement between measured and calculated masses are displayed (see *Fig. 3*, for example data see [36]).

All calculations up to this point were performed without a structural database. To receive an idea of possible structures belonging to the derived compositions a database search is available (see *Fig. 4*, example data see [36]). The retrieved results from the database can be used for further investigations in *GlycoWorkbench*.



Mass	Intensity	Composition (check for fragment and structure search)	Charged Ions	Ion type	Mass calculated	Deviation [ppm]
204.000	n/a	☐ HexNac1	H+	B	204.08665	424.7
300.200	n/a	☐ HexNac1-PA	H+	Y	300.15540	-148.6
366.200	n/a	☐ Hex1HexNac1	H+	B	366.13947	-165.3
407.200	n/a	☐ HexNac2	H+	B	407.16602	-83.4
446.300	n/a	☐ HexNac1dHex1-PA	H+	Y	446.21331	-194.3
1176.700	n/a	☐ Hex6HexNac1	H+	B	1176.40359	-251.9
		☐ Hex2HexNac3dHex1-PA	H+	Y	1176.47770	-188.9
1299.700	n/a	☐ HexNac2dHex2NeuAc2	H+	C	1299.48323	-166.8
		☐ Hex7dHex1	H+	C	1299.44551	-195.8
		☐ Hex3HexNac4	H+	B	1299.48323	-166.8
1379.500	n/a	☐ Hex3dHex2NeuAc2	H+	C	1379.48296	-12.4
		☐ Hex6HexNac2	H+	B	1379.48296	-12.4
		☐ Hex2HexNac4dHex1-PA	H+	Y	1379.55707	41.4
1541.500	n/a	☐ Hex4dHex2NeuAc2	H+	C	1541.53578	23.2
		☐ HexNac4dHex1NeuAc2	H+	B	1541.57350	47.7
		☐ Hex7HexNac2	H+	B	1541.53578	23.2
		☐ Hex3HexNac4dHex1-PA	H+	Y	1541.60989	71.3
1744.700	n/a	☐ Hex4HexNac1dHex2NeuAc2	H+	C	1744.61515	-48.6
		☐ HexNac5dHex1NeuAc2	H+	B	1744.65288	-27.0
		☐ Hex7HexNac3	H+	B	1744.61515	-48.6
		☑ Hex3HexNac5dHex1-PA	H+	Y	1744.68926	-6.2

Figure 3
***Glyco-Peakfinder* results window. The *Glyco-Peakfinder* results window shows the calculated compositions for mass signals from a mass spectrum of a pyridylamino-oligosaccharide $Hex_3HexNAc_5dHex$-PA fraction obtained from batroxobin of *Bothrops moojeni* venom**

3.2 *GlycoWorkbench*: Assisted Annotation of Fragment Mass Spectra

The previous sections have explained that completely automated annotation of generic glycan mass spectra is still unfeasible. Expert knowledge is fundamental for the correct interpretation of spectra but this is not yet available in the form of machine readable data collections. Tools like *Glyco-Peakfinder* can help to determine the possible candidates for a certain mass value but these hypothetical structures must then be assessed by MS^n fragmentation experiments. Manual annotation of fragment spectra comprises a series of tedious and repetitive steps whose automation is straightforward, and can result in a substantial decrease in the time needed for sequencing a structure. The *GlycoWorkbench* is a suite of software tools designed to assist the experts during the annotation of glycan fragment spectra. The graphical interface of *Glyco-*

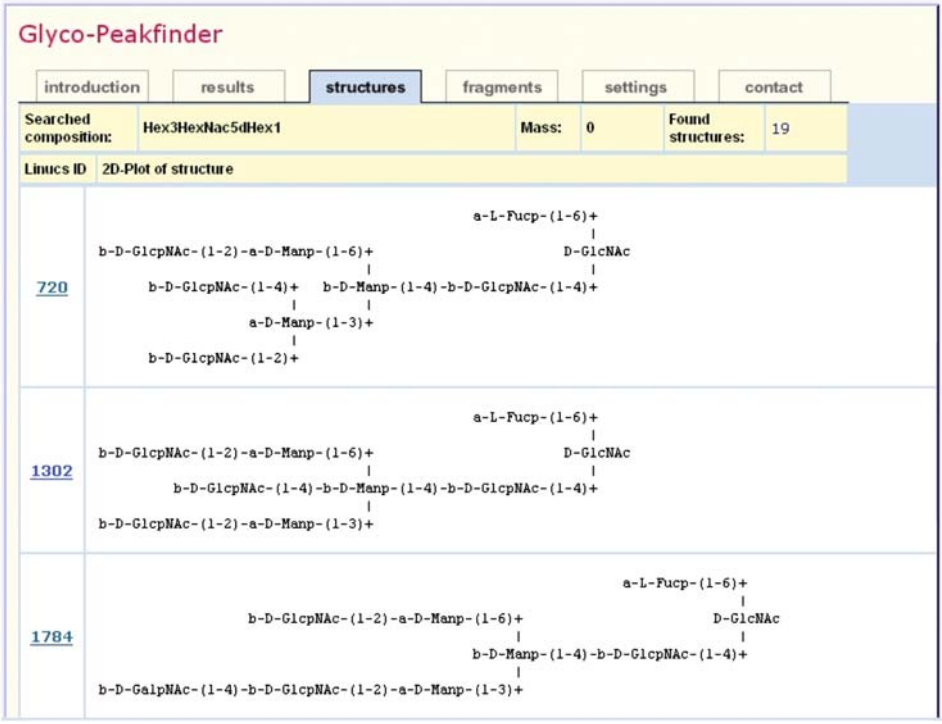


Figure 4
***Glyco-Peakfinder* structures window. The *Glyco-Peakfinder* structures window shows structures in *GLYCOSCIENCES.de* [48] stored for the $Hex_3HexNAc_5dHex$ composition. The retrieved structures are displayed as a 2D graph using an IUPAC-CarbBank-like encoding for complex carbohydrates. LINUCS-ID [49] is an identifier for unique descriptions of glycan structures as used in the *GLYCOSCIENCES.de* portal**

Workbench (see *Fig. 5*) provides an environment in which structure models can be rapidly assembled, automatically matched with MS^n data, and compared to assess the best candidate. The screenshots of *GlycoWorkbench* shown in this chapter are taken from a semi-automatic assignment of a pyridylamino-oligosaccharide $Hex_3HexNAc_5dHex$-PA fraction obtained from batroxobin of *Bothrops moojeni* venom [36], the same experimental setup used for demonstrating the *Glyco-Peakfinder* tool. The three structural models have been retrieved from the result of a composition search with *Glyco-Peakfinder* in a structure database. This demonstrates a possible complete workflow for semi-automatic assignment involving the two tools.

3.2.1 *GlycanBuilder*

The main component of *GlycoWorkbench* is the *GlycanBuilder* [37], a rapid and flexible visual editor of glycan structures (see *Fig. 5*). Carbohydrates mostly present tree-like non-sequential structures, and their constituents exhibit great diversity. Therefore, the input of a structure in a computer readable format is not as straightforward as writing a sequence of characters, as

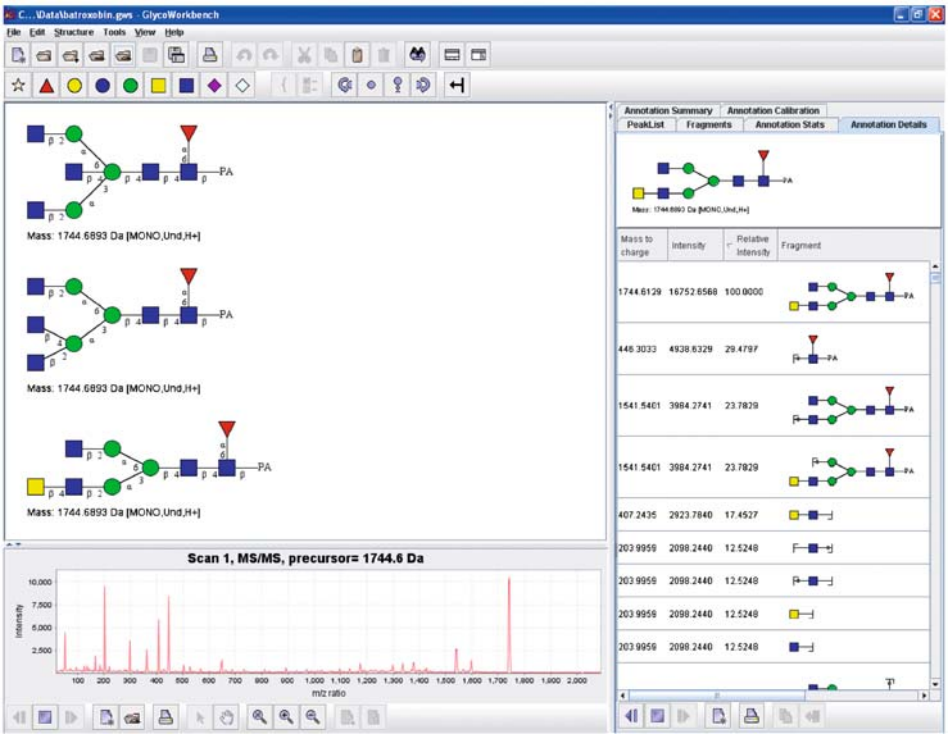


Figure 5

***GlycoWorkbench*, main window. *GlycoWorkbench* is an integrated suite of software tools for assisting the annotation of glycan fragment mass spectra. All the tools are accessible from a common user interface. Here the *GlycoWorkbench* interface with the *GlycanBuilder* (*left top*), the "spectra viewer" (*left bottom*), and the "fragmentation tool" (*right*) are shown. The commonly used cartoon-like representation as recommended by the US Consortium for Functional Glycomics is used to display the oligosaccharide structures. The three structures shown here have been identified with the help of the *Glyco-Peakfinder* structures window (*Fig. 4*) from the *GLYCOSCIENCES.de* database. *GlycoWorkbench* can be downloaded from the URL: http://www.eurocarbdb.org/applications/ms-tools/**

for DNA, RNA, and peptide sequences. Additionally, numerous alternative notations are commonly adopted to graphically represent glycan structures. Two types of structure builders have been developed so far, mainly for searching databases: template-based editors, where the user can only choose among a fixed set of templates and extensions (e. g. Consortium for Functional Glycomics [38], Bacterial Carbohydrate Structure Data Base [39]) and totally graphical editors, where the user can freely position the residues on a drawing board, often manually adjusting the image for optimal layout [40,41]. In the *GlycoWorkbench* editor, the structure is assembled simply by selecting the point of attachment of the residue to add. The spatial placement of the new residue is automatically determined according to the rules for the given notation. Therefore, the display of a glycan is dependent only on its structure and the chosen notation: the software always knows how to represent a structure, and a new notation can

be applied without the need for user intervention. The structure is directly available for use both in bioinformatics computations and for publication purposes. In addition, the structure can be exported into standard encoding formats like *GlycoCT*, or *GLYDE-II* (see *Table 1*) or graphical formats like SVG, PDF, or BMP. The popular notations for glycan representation from the Consortium for Functional Glycomics and the Oxford Glycobiology Institute [42] are available as notation styles. The available constituents comprise an exhaustive list of saccharides, substituents, reducing-end markers, and saccharide modifications. All the stereo-chemical information about a saccharide, like anomeric conformation, chirality, ring configuration, and linkage position, can be specified.

3.2.2 *In Silico* Fragmentation

After a set of structures has been created with the *GlycanBuilder*, the remaining tools can be used to derive their fragments, compute the fragment masses, build a peak-list, and annotate it. Computation of fragments and their masses from the intact structure is a central step for the annotation of CID MS^n spectra. The "fragmentation tool" creates all topologically possible fragmentations of the precursor molecular ion, applying both multiple glycosidic cleavages and cross-ring fragments. The list of possible ring fragments is exhaustive, and is available for all the residues supported by *GlycoWorkbench*. Partial or total loss of labile substituents such as sulfates and phosphates is automatically taken into account. Alternatively, a fragment editor is available where the user can specify which positions the cleavages are occurring on the displayed structure in order to reproduce an already known fragment molecule. Given a fragment structure, m/z values resulting from several types and quantities of charges can be calculated both for native and derivatized structures (permethylated / peracetylated).

3.2.3 Annotation of Peaks

The next step in the annotation process is the definition of a peaklist. In *GlycoWorkbench* a peaklist can either be loaded from a tab-separated text file, thus allowing for import from peak-picking software, or it can be created by typing mass and intensity values directly into the application (see *Fig. 6* and *Fig. 7*). Alternatively, the whole spectrum can be loaded from several standard XML or vendor-specific data formats (supported through the use of the ProteomeCommons IO library [43]). The raw mass spectrum is displayed and can be panned or zoomed as in a normal spectrum viewer. The user can then select mass values directly from the curve, add them to the peak-list or check them on-the-fly for matching against fragment masses. Once the peak-list is ready, the fragmentation tool is used to generate all the possible fragments of each input structure and their m/z values are tested for matches against every peak given the desired accuracy. The resulting annotated peak-list can then be viewed using various panels that show its different aspects. Each panel is based around a spreadsheet-like table view, whose cell values can be sorted by each column, and can be copied into spreadsheet software. The detailed view (see *Fig. 7*) lets the user check a comprehensive list of fragment-peak matches for each of the structures, showing the fragment structure, mass, m/z value, distribution of charges, and annotation accuracy. The annotation can also be reviewed by removing the matches that are not satisfactory. The summary view (see *Fig. 6*) lets the user compare the annotations for the different structures back-to-back in the same table. The

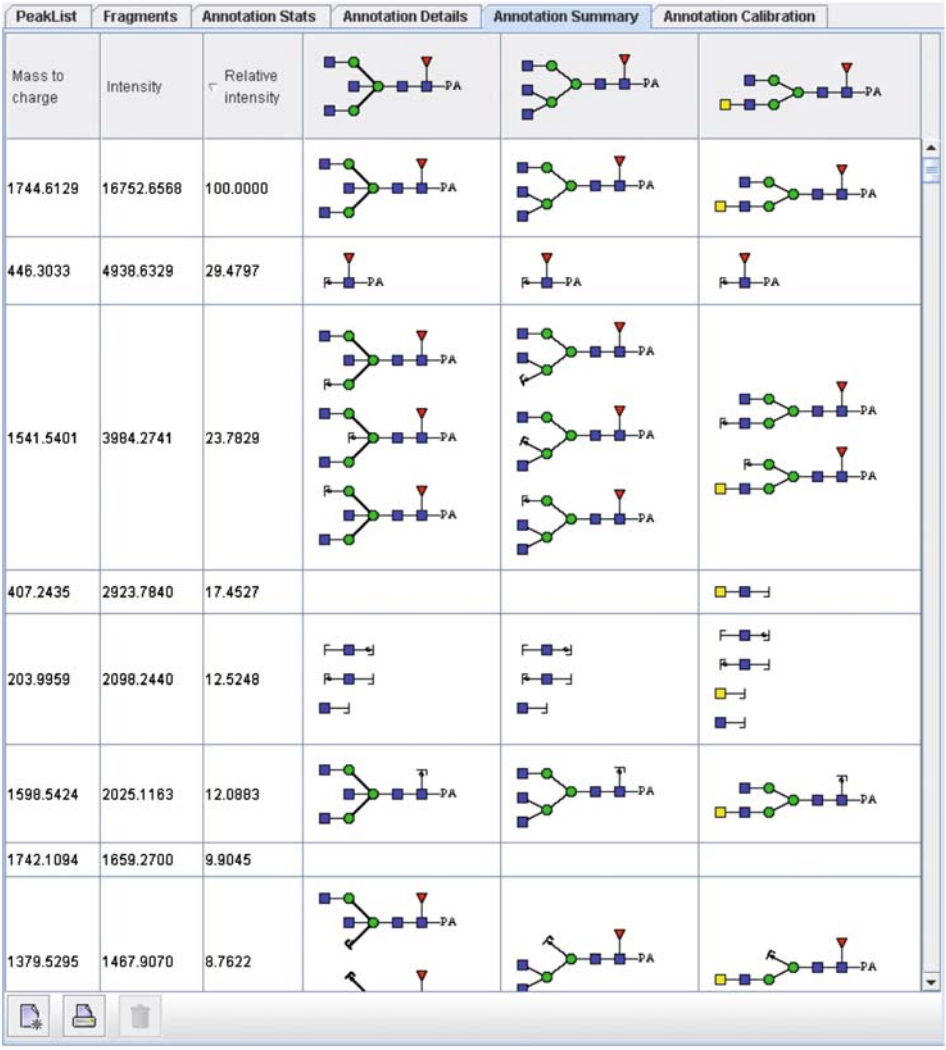


Figure 6

***GlycoWorkbench*, annotation summary view. The *summary* view enables a direct comparison of peak assignments to fragments for various user-selected structures back-to-back in the same table. The matching fragments from different structures are shown in adjacent columns, with each row corresponding to a single peak. The user can thus identify signals which clearly distinguish the correct annotation from the other hypothetical models**

matching fragments from different structures are shown in adjacent columns, with each row corresponding to a single peak. In this way, signals that clearly distinguish the correct annotation from the other hypothetical models can be identified. The statistics view lets the user perform a quantitative comparison between the annotations, by showing the number of assigned

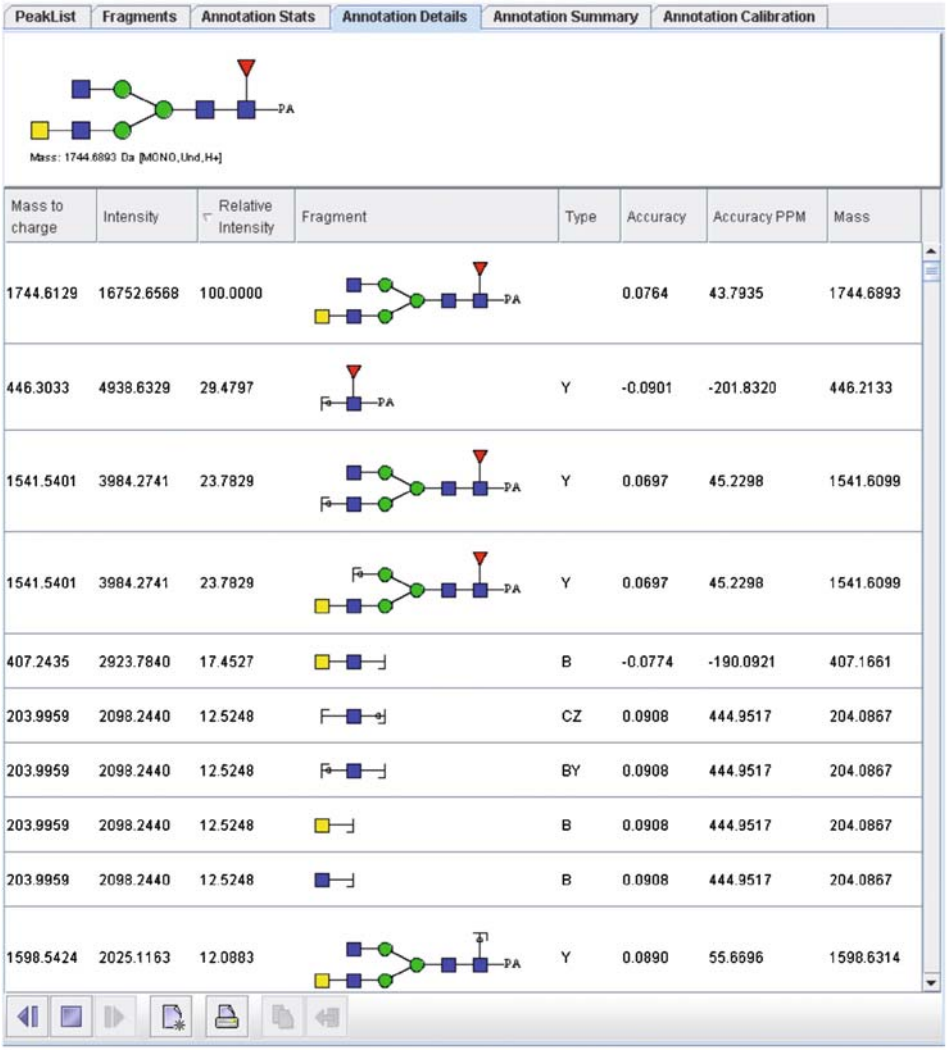


Figure 7

***GlycoWorkbench*, annotation detailed view. The *detailed* view enables users to check for each individual structure all the fragments that match at least one measured peak, showing the associated fragment structure, theoretical mass, m/z value, distribution of charges, and annotation accuracy. In this panel, the annotation can also be reviewed by removing the matches that are not satisfactory**

peaks at different thresholds of relative peak intensity, the root mean square deviation between peak and fragment *m*/*z* values and the average intensity of assigned peaks. Finally, a calibration graph shows the annotation accuracies at the various *m*/*z* values, allowing the user to check the correct calibration of the mass spectra.

3.3 Glycofragment Mass Fingerprinting

Glycofragment Mass Fingerprinting [27,29] is a bioinformatic technique analogous to Peptide Mass Fingerprinting [44,45], suggesting structures for a particular glycan fragmentation mass spectrum. In GMF, experimental masses are matched up against a database of theoretical fragment masses, and the best matches are presented as results.

Fundamental to GMF, a comprehensive library of structures serves as the source of theoretical fragmentation data. Each of the structures within the library is transformed into a set of fragments by performing a virtual cleavage of bonds. This produces a set of fragmentations for each structure that is representative of all possible fragmentations that could occur on the glycan. Each fragment has a mass associated with it that can be considered as a virtual peak mass for the fragment.

The underlying GMF technique can be split into two stages—with the first being the matching of spectra against the library of known fragmentations, and the second being the scoring and ranking of candidate structures.

3.3.1 Matching of Spectra

An important preprocessing step in performing the search is to ensure that all spectra being used in the algorithms are normalized, and no mass spectrometric or experimentally specific modifications to the spectrum are reflected in the data used in the algorithm execution. To achieve this, many GMF tools normalize all m/z values by resolving the charge state and removing any adducts. This functionality is controlled through the setting of various experimental parameters (such as charge state, adducts, and derivatizations). Certain tools can also accept spectral data in XML format allowing for automatic normalization of spectra.

Structures which match up with the input structure are then selected using a voting algorithm, where each fragment derived from a structure will count as a vote for that structure if the virtual peak mass matches with a normalized peak mass. Those structures having at least one vote (or matched fragment) are candidate structures for the match. Each normalized peak mass may not necessarily match up with a single fragment mass from a structure, as many fragmentations from a structure can share the same mass.

3.3.2 Scoring of Candidate Structures

The scoring stage of GMF is the most important part of the matching process whereby candidate structures are ranked according to the suitability of match to the input spectrum. A number of approaches can be taken in this stage, from simple counting of matched peaks to algorithms that perform an analysis of the evidence presented by the matched peaks.

There are no standard sets of scoring methods, so the methods available in each tool differ, however, it is generally accurate to say that current scoring algorithms do not provide a single reliable quantitative measure for the match. Instead, combinations of scores are used as an aid to the further interpretation of data using expert knowledge.

GMF excels at distinguishing between structural isomers, as match information from a search can be important in identifying diagnostic ions that support one structure over another. Important diagnostic ions are often related to cross-ring fragmentation, which will yield information related to the linkages on a structure.

3.3.3 Limitations of GMF

Limitations of GMF approaches are related to the data that comes in to the algorithms. Since tandem MS data is not as information rich as a set of multiple-stage MS spectra, it may not be possible to distinguish between two structures based on the one spectrum alone. In fact, it is often the case that supplemental evidence is required to make a confident assertion that structures have been assigned correctly.

The quality of the GMF match that occurs is dependent on the quality of the library of structures. When the library of structures is incomplete, there may be a chance that the structure in question may not be part of the library, and an incomplete or incorrect assignment may result. This can be mitigated in part through the use of intelligent scoring algorithms that can rank structures that are similar to the actual (unknown) structure highly.

The term mass fingerprinting is a slight misnomer when applied to glycan fragmentation, as the structures do not necessarily yield unique "fingerprint" spectra when fragmented. In fact, due to the nature of the fragmentation of some structures, many spectra are isobaric. This problem is in fact inherent to tandem MS spectra where high levels of fragmentation are desired. Although it is possible to deduce the combination of fragmentations that have occurred to produce a fragment in optimal experimental conditions, you often cannot discern the order of fragmentation, or the pathway of fragmentation needed to produce that fragment. It is for this reason that one-off spectral matching will not work well for larger glycan structures. However, spectral matching has been shown to work well when used to identify smaller structures, working well enough to even distinguish the stereochemistry for residues. Used in combination with multiple-stage mass spectrometry, it is possible to control the fragmentation pathway so that all fragmentations are well characterized. This MS^n [32,35,46] approach to analysis is a fundamentally different approach to GMF, yielding significantly richer data, and obviating the need for sophisticated scoring algorithms. The key to the success of MS^n approaches is the ability to select ions for further fragmentation.

For spectra where it is not possible to distinguish between two structural isomers, clarification of structure can be found through the further fragmentation of the structure. To do this, algorithms pick out the peaks corresponding to the glycan fragments containing the sub-structures in question. This is a recursive approach to sequencing, where ion selection and further fragmentation occurs until the structure is fully elucidated. Coupled with technologies to control the ion selection on the mass spectrometer, this technique shows promise for fully automated identification and sequencing of structures.

4 Summary and Conclusion

Investigations into the analytical techniques of glycans showed clearly that mass spectrometry is still the key technology for research involving the analysis of carbohydrate structure. Despite its popularity, the automatic interpretation of MS spectra is an analytical bottleneck in the rapid and reliable interpretation of MS data in high-throughput glycomics projects, and robust solutions are of critical importance. Given the importance of establishing such solutions, it is not surprising that various experimentally oriented groups have been developing software solutions and algorithms to bypass this bottleneck.

A side-effect of the multiple efforts being put into place to solve this bottleneck is that the current situation in glycoinformatics can be characterized by the existence of multiple disconnected and incompatible islands of experimental data, data resources, and specific applications, each managed by various consortia, institutions, or local groups. This is in no small part due to the closed nature of much of the software—only a limited amount of software is freely available to be shared by various projects.

The coverage and capability of existing tools is somewhat varied—compositional analysis is well served by the current number of tools available, encompassing a number of approaches to the analysis. For MS tools, the existing software serves the needs less well for general researchers, primarily due to the availability of the programs. However, with respect to functionality, the tools are capable of performing semi-automatic interpretation, functioning as an additional source of information in the elucidation of structure.

Recent developments have raised the hope that in the near future freely available software suites for compositional analysis as well as for annotation of spectra (such as *Glyco-Peakfinder* and *GlycoWorkbench*) will form the basis of software which can support easy and reliable semi-automatic assignment of complete MS spectra to glycans.

Other library-based approaches (such as GMF) would benefit from free availability of databases, containing structural information as well as spectrometric data. The use of widely available free databases would enhance the use of informatics in glycomics to a level comparable to proteomics, where the use of databases is part of daily work in nearly all laboratories.

A long-term goal of informatics for glycomics is to enable high-throughput analysis of data. To do this, each piece of software must be able to automatically interact, functioning as an analysis pipeline. The issue of interoperability is not only important between tools, but is a crucial issue with respect to databases also. Maturity in the glycomic software area will allow for data to freely move between databases and software, providing maximum flexibility in analysis workflows. With the agreement to accept *GLYDE-II* (see NIH White Paper: "*Urgent requirements for the development of informatics for glycomics and glycobiology*") as the central format to exchange structural data, a fundamental issue surrounding interoperability has been solved.

It is becoming clear that the optimal way to provide semi-automatic, and even wholly automatic sequencing of structures is through a pluralistic methodology, using multiple analytical techniques such as HPLC, NMR, glycan arrays, and GC MS. This multi-technique approach to analysis, including the automated integration of data from multiple sources yields the best possibility for achieving a fully automatic unambiguous structural determination.

Acknowledgement

Glyco-Peakfinder and *GlycoWorkbench* are developed as part of the EUROCarbDB project, a Research Infrastructure Design Study Funded by the 6th Research Framework Program of the European Union (Contract: RIDS Contract number 011952). The development of the *GLYCOSCIENCES.de* portal at the DKFZ (German Cancer Research Center) was supported by a Research Grant from the German Research Foundation (DFG BIB 46 HDdkz 01–01) within the digital library program as well as the president fond of the Helmholtz society.

References

1. von der Lieth CW, Bohne-Lang A, Lohmann KK, Frank M (2004) Brief Bioinform 5:164
2. Raman R, Raguram S, Venkataraman G, Paulson JC, Sasisekharan R (2005) Nat Methods 2:817
3. Packer NH (2006) Proteomics 6:6121
4. von der Lieth CW, Lutteke T, Frank M (2006) Biochim Biophys Acta 1760(4):568–577
5. Comelli EM, Head SR, Gilmartin T, Whisenant T, Haslam SM, North SJ, Wong NK, Kudo T, Narimatsu H, Esko JD, Drickamer K, Dell A, Paulson JC (2006) Glycobiology 16:117
6. Paulson JC, Blixt O, Collins BE (2006) Nat Chem Biol 2:238
7. Geyer H, Geyer R (2006) Biochim Biophys Acta – Proteins and Proteomics 1764:1853–1869
8. Morelle W, Canis K, Chirat F, Faid V, Michalski JC (2006) Proteomics 6:3993
9. Jang-Lee J, North SJ, Sutton-Smith M, Goldberg D, Panico M, Morris H, Haslam SM, Dell A (2006) Methods Enzymol 415:59
10. Haslam SM, North SJ, Dell A (2006) Curr Opin Struct Biol 16:584
11. Geyer H, Geyer R (1998) Acta Anat 161:18
12. Harvey DJ (2005) Expert Rev Proteomics 2:87
13. Harvey DJ (2006) Mass Spectrom Rev 25:595
14. Dell A, Morris HR (2001) Science 291:2351
15. Peter-Katalinic J (2005) Methods Enzymol 405:139
16. Zaia J (2004) Mass Spectrom Rev 23:161
17. Wuhrer M, Catalina MI, Deelder AM, Hokke CH (2006) J Chromatogr B Analyt Technol Biomed Life Sci 849(1–2):115–128
18. Manzi AE, Norgard-Sumnicht K, Argade S, Marth JD, van Halbeek H, Varki A (2000) Glycobiology 10:669
19. Spengler B (2004) J Am Soc Mass Spectrom 15:703
20. Rudd PM, Dwek RA (1997) Crit Rev Biochem Mol Biol 31:1
21. Chalabi S, Panico M, Sutton-Smith M, Haslam SM, Patankar MS, Lattanzio FA, Morris HR, Clark GF, Dell A (2006) Biochemistry 45:637
22. von der Lieth CW (2004) J Carbohyd Chem 23:277
23. Cooper CA, Joshi HJ, Harrison MJ, Wilkins MR, Packer NH (2003) Nucleic Acids Res 31:511
24. Cooper CA, Gasteiger E, Packer NH (2001) Proteomics 1:340–349
25. Maass K, Ranzinger R, Geyer H, von der Lieth CW, Geyer R (2007) Proteomics 7:4435–4444
26. Goldberg D, Sutton-Smith M, Paulson J, Dell A (2005) Proteomics 5:865
27. Lohmann KK, von der Lieth CW (2004) Nucleic Acids Res 32:W261
28. Clerens S, Van den Ende W, Verhaert P, Geenen L, Arckens L (2004) Proteomics 4:629
29. Joshi HJ, Harrison MJ, Schulz BL, Cooper CA, Packer NH, Karlsson NG (2004) Proteomics 4:1650
30. Kameyama A, Kikuchi N, Nakaya S, Ito H, Sato T, Shikanai T, Takahashi Y, Takahashi K, Narimatsu H (2005) Anal Chem 77:4719
31. Gaucher, SP, Morrow, J, Leary, JA (2000) Anal. Chem. 72: 2331
32. Lapadula AJ, Hatcher PJ, Hanneman AJ, Ashline DJ, Zhang H, Reinhold VN (2005) Anal Chem 77:6271
33. Ethier M, Saba JA, Spearman M, Krokhin O, Butler M, Ens W, Standing KG, Perreault H (2003) Rapid Commun Mass Sp 17:2713
34. Tang H, Mechref Y, Novotny MV (2005) Bioinformatics 21:i431
35. Kameyama A, Nakaya S, Ito H, Kikuchi N, Angata T, Nakamura M, Ishida HK, Narimatsu H (2006) J Proteome Res 5:808
36. Lochnit G, Geyer R (1995) Eur J Biochem 228:805
37. Ceroni A, Dell A, Haslam SM (2007) "The GlycanBuilder: a fast, intuitive and flexible software tool for building and displaying glycan structures", Source Code for Biology and Medicine, vol 2. p 3
38. Raman R, Venkataraman M, Ramakrishnan S, Lang W, Raguram S, Sasisekharan R (2006) Glycobiology 16:82R
39. Toukach P, Joshi HJ, Ranzinger R, Knirel Y, von der Lieth CW (2007) Nucleic Acids Res 35:D280
40. Hashimoto K, Kawano S, Goto S, Aoki-Kinoshita KF, Kawashima M, Kanehisa M (2005) Genome Inform 16:214
41. Hashimoto K, Goto S, Kawano S, Aoki-Kinoshita KF, Ueda N, Hamajima M, Kawasaki T, Kanehisa M (2006) Glycobiology 16:63R
42. Royle L, Dwek RA, Rudd PM (2006) In: Coligan J, Dunn B, Speicher D, Wingfield P (eds) Current Protocols in Protein Science, vol UNIT 12.6. Wiley, New York

43. Falkner JA, Falkner JW, Andrews PC (2007) Bioinformatics 23(2):262–3
44. Thiede B, Hohenwarter W, Krah A, Mattow J, Schmid M, Schmidt F, Jungblut PR (2005) Methods 35:237
45. Aebersold R, Mann M (2003) Nature 422: 198
46. Ashline D, Singh S, Hanneman A, Reinhold VN (2005) Anal Chem 77:6250
47. Domon B, Costello CE (1988) Glycoconjugate 5:397–409
48. Lutteke T, Bohne-Lang A, Loss A, Goetz T, Frank M, von der Lieth CW (2006) Glycobiology 16:71R
49. Bohne-Lang A, Lang E, Forster T, von der Lieth CW (2001) Carbohydr Res 336:1
50. Lohmann KK, von der Lieth CW (2003) Proteomics 10:2028

Part 11
Biosynthesis and Degradation

11.1 Biosynthesis and Degradation of Mono-, Oligo-, and Polysaccharides: Introduction

Iain B. H. Wilson
Department für Chemie, Universität für Bodenkultur
(University of Natural Resources and Applied Life Sciences),
Muthgasse 18, 1190 Wien, Austria
iain.wilson@boku.ac.at

Abstract

Glycomolecules, whether they be mono-, oligo-, or polysaccharides or simple glycosides, are—as any biological molecules—the products of biosynthetic processes; on the other hand, at the end of their lifespan, they are also subject to degradation. The beginning point, biochemically, is the fixation of carbon by photosynthesis; subsequent metabolism in plants and other organisms results in the generation of the various monosaccharides. These must be activated—typically as nucleotide sugars or lipid-phosphosugars—before transfer by glycosyltransferases can take place in order to produce the wide variety of oligo- and polysaccharides seen in Nature; complicated remodelling processes may take place—depending on the pathway—which result in partial trimming of a precursor by glycosidases prior to the addition of further monosaccharide units. Upon completion of the 'life' of a glycoconjugate, glycosidases will degrade the macromolecule finally into monosaccharide units which can be metabolized

In: *Glycoscience*. Fraser-Reid B, Tatsuta K, Thiem J (eds)
Chapter-DOI 10-1007/978-3-540-30429-6_58:

or salvaged for incorporation into new glycan chains. In modern glycoscience, a wide variety of methods—genetic, biochemical, analytical—are being employed in order to understand these various pathways and to place them within their biological and medical context. In this chapter, these processes and relevant concepts and methods are introduced, prior to elaboration in the subsequent more specialized chapters on biosynthesis and degradation of mono-, oligo-, and polysaccharides.

Keywords

Glycans; Glycosidases; Glycosyltransferases; NDP-sugars

Abbreviations

ATP	adenosine triphosphate
CDG	congenital disorder of glycosylation
CDP	cytidine 5′-diphospho-
CMP	cytidine 5′-monophospho-
Dol-P	dolichol phosphate
dTDP	deoxythymidine 5′-diphospho-
EGF	epidermal growth factor
$EtNPO_4$	ethanolamine phosphate
GDP	guanosine 5′-diphospho-
GlcNAc-T	*N*-acetylglucosaminyltransferase
GPI	glycosylphosphatidylinositol
NAD	nicotinamide adenine dinucleotide
NADP	nicotinamide adenine dinucleotide phosphate
NDP	nucleoside diphosphate
PAPS	3′-phosphoadenosine 5′-phosphosulfate
PEP	phosphoenolpyruvate.
TSR	thrombospondin type 1 repeat
UDP	uridine 5′-diphosphate

1 General Principles

The term ‘glycoconjugate’ encompasses a wide range of biological molecules containing at least one monosaccharide unit combined with other moieties, which can be either glycosidic or non-glycosidic in nature. Correspondingly, the breadth of molecular transformations of these molecules—leading to their formation or degradation—is also large and great strides have been made in the past two or three decades as regards our understanding of the structures themselves and the various biochemical pathways in which they are involved. Most recently, molecular biology has revolutionized the study of the enzymes responsible for these biotransformations and has made larger quantities of protein available for enzymatic and structural characterizations. These enzymes fall into a number of categories—those involved in synthesizing the building blocks of glycoconjugates (i. e., a variety of enzymes required for

monosaccharide synthesis and modification), those involved in generating glycosides (glycosyltransferases, which transfer one or more sugars to another molecule) and those involved in breaking the bonds made by glycosyltransferases (primarily glycosidases). Furthermore, transporters and flippases necessary for transferring glycomolecules across biological membranes are also important in completing the vast biochemical network required for the biosynthesis and degradation of mono-, oligo-, and polysaccharides.

The starting point for this biochemical journey from monosaccharide to glycoconjugate and back again is photosynthesis and, indeed, glucose is a central player in the ability of cells to generate glycoconjugates. Although knowledge about glycoconjugates has increased vastly in the last century, it is still of interest, in this context, to quote Armstrong from his 1919 monograph [1]:

"The interest attaching to the [carbohydrates] may be said to centre around glucose, this carbohydrate being the first to arise in the plant and the unit group from which substances such as cane sugar, maltose, starch and cellulose are derived; it is also of primary importance in animal metabolism, as the main bulk of the carbohydrate in our food materials enters the circulation in the form of glucose."

It is not the intention here to discuss the synthesis of glucose and glucosidic polymers such as starch or glycogen and their central role in energy storage and production. However, every monosaccharide is 'descended' in some way from glucose; thus, any discussion of biosynthesis and degradation of glycoconjugates must give some attention to the transformation of glucose into other sugars and, thereby, Armstrong's comment is still relevant. A number of steps including epimerization, decarboxylation, dehydration and reduction are necessary to turn glucose into the range of monosaccharides observed as components of glycoproteins with their *N*- and *O*-linked glycans, proteoglycans, glycolipids, peptidoglycans, lipophosphoglycans, simple glycosides, nucleic acids and other sugar-containing molecules and polymers.

Another important general principle is that, prior to the joining of a sugar to the 'acceptor', which can be another sugar or another type of molecule, the monosaccharide must be activated (other than in the case of transglycosylation). This reflects the requirement for energy in the formation of glycosidic bonds; this energy originates from ATP (the synthesis of which is itself driven by the oxidative phosphorylation of glucose). In turn, this means that, in general, sugars are present in the form of nucleoside mono- or diphosphosugars or lipid-linked phosphosugars; it is these activated forms which are the 'donors' for the reactions catalysed by glycosyltransferases (EC 2.4.1.-, 2.4.2.- and 2.4.99.-). In some cases, glycosidases can be used to also generate glycosidic bonds, but these reactions are often transglycosylations resulting from the use of these enzymes under very artificial conditions or of genetically engineered forms ('glycosynthases'); although enzymes which naturally transglycosylate do exist, e. g., the *trans*-sialidase of *Trypanosoma cruzi* or plant xyloglucan endo-transglycosylases [2,3], the vast majority of enzymes of glycoconjugate biosynthesis are glycosyltransferases requiring nucleotide sugars and it is these which are the focus of much attention in glycobiology.

In 'opposition' to the enzymes forming glycosidic bonds are those that degrade them, primarily members of the vast families of glycosidases (EC 3.2.1.-). In contrast to the glycosyltransferase reaction, the hydrolysis of a glycosidic bond is thermodynamically favourable, with ΔG° values of approximately –10 kJ mol^{-1} [4]; at the same time, activation energies of perhaps +40–80 kJ mol^{-1}, as in the case of mammalian lysosomal hexosaminidases [5], indicate that these reactions require a catalyst. It is also interesting that glycosidases are often chemically and structurally well-described, due to the relative ease of purification, recombinant expression,

assaying and handling of these enzymes as compared to glycosyltransferases which are considered 'finicky' to purify or express and expensive to assay. Despite this 'advantage', the biological role of catabolic glycosidases in higher organisms is often neglected as compared to the glycosyltransferases, particularly those involved in *N*-linked glycan biosynthesis. However, their clinical relevance is shown particularly well by the various 'storage' diseases which result from mutations in the genes encoding either lysosomal glycosidases themselves or the proteins required for their targeting [6]. Indeed, many 'knock-outs' (natural or artificial) demonstrate that the proper biosynthesis and degradation of a range of mono- and oligosaccharides is a prerequisite for biological viability.

2 Components and Types of Glycoconjugates in Mammals

A major focus of modern glycoscience is on the oligosaccharides of mammalian glycoconjugates. Mammals express a range of these macromolecules—proteins with either *N*-linked, *O*-linked or *C*-linked oligosaccharides, glycosylphosphatidylinositol membrane anchors, glycolipids and glycosaminoglycans (*Fig. 1*) as well as the storage polymer glycogen [7,8]. The *N*-linked oligosaccharides or *N*-glycans all share a common pentasaccharide core consisting of two *N*-acetylglucosamine and three mannose residues; this core is elaborated in a number of ways depending on the class of oligosaccharide. The categories of oligomannosidic (or 'high mannose'), hybrid and complex oligosaccharides have been known for many years and will be discussed in more detail in Chap. 11.2; it is sufficient to say that the range of structures is very broad and that, for mammalian *N*-glycans, the most important monosaccharides are *N*-acetylglucosamine, mannose, galactose, glucose, *N*-acetylgalactosamine, glucuronic acid, fucose and sialic acids. Modifications with sulfate and phosphate are also known.

The *O*-glycans, on the other hand, have a wide range of possible core structures—not only should the typical 'mucin-type' *O*-linked sugars (*N*-acetylgalactosamine linked to serine or threonine) be considered; furthermore, as recently reviewed by Spiro [8], other linkages of fucose, galactose, glucose, xylose, mannose and *N*-acetylglucosamine are also present. Some of these have important biological functions. For example, cytosolic/nuclear '*O*-GlcNAc' is an element, 'competing' with phosphorylation, in controlling protein function [9], while *O*-fucose functions as a modifier of Notch signalling in development [10]. Furthermore, *O*-mannose is a form of glycosylation of dystroglycan with roles in certain forms of muscular dystrophy [11] and *O*-linked xylose is the first sugar in chondroitin and heparan sulfate glycosaminoglycan chains [12]. Whereas 'mucin-type' *O*-glycans contain monosaccharides familiar as modifications of *N*-glycans (*N*-acetylgalactosamine, *N*-acetylglucosamine, galactose, fucose and sialic acids), the chondroitin and heparan sulfates contain a core of GlcAβ1,3Galβ1,3Galβ1,4-Xylβ1-O-Ser which is then modified by either GalNAcβ1,4GlcAβ1,3 (chondroitin) or GlcNAcα1,4GlcAβ1,4 (heparan) repeats. These sugars can then be modified by sulfation or, in the case of glucuronic acid, epimerization to iduronic acid. In the case of epidermal growth factor-like (EGF) and thrombospondin type 1 repeat (TSR) domains of, e. g., blood clotting factors, the modifications by Xylα1,3Glcβ-O-Ser and *O*-fucose must also be considered [13], whereas hydroxylysine residues of collagen and collagen-like domains can carry Glcα1,2Gal structures [14]. A more unusual modification is the *C*-mannosylation of tryptophan also present on thrombospondin repeats of certain proteins [15].

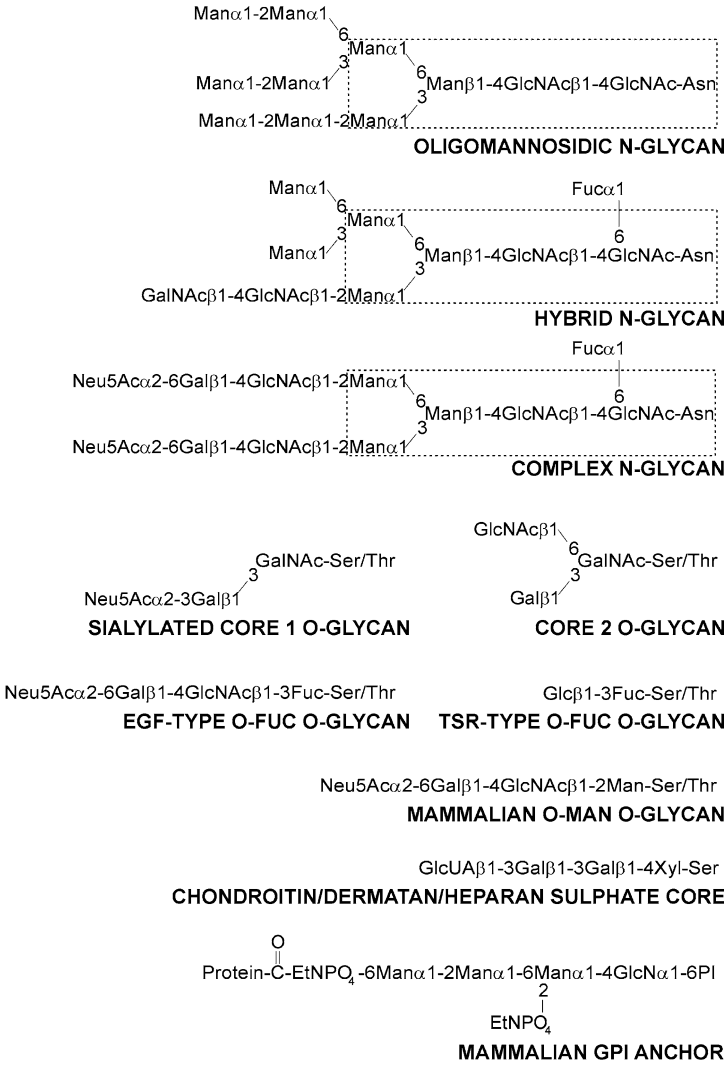


Figure 1

The main types of mammalian protein-linked glycoconjugate. A selection of typical mammalian protein-linked glycans (*N*- and *O*-linked), as well as a glycosylphosphatidylinositol (GPI) anchor, are shown. As typical *N*-glycans, an oligomannosidic structure ('Man9'), a hybrid structure found in a mannosidase double-knockout mouse and a sialylated core fucosylated biantennary structure are depicted; the common pentasaccharide core is boxed. 'Mucin-type' core 1 and core 2 *O*-glycans, the *O*-fucose-type glycans of EGF and TSR repeats, the *O*-mannose-type glycan of, e. g., dystroglycan and the core tetrasaccharide region of chondroitin, dermatan and heparan sulfate glycosaminoglycan chains are also shown. In the case of mammalian GPI anchors, the inositol portion of the phosphatidylinositol (PI) moiety is typically acylated; other side chains on the glycan moiety are possible. Other modifications such as cytoplasmic *O*-GlcNAc, collagen-type *O*-glycans, *O*-glucosylation of EGF domains and *C*-mannosylation of tryptophan are not included. See text for references

lacto	Galβ1-3GlcNAcβ1-3Galβ1-4Glcβ1-Cer
lactoneo	Galβ1-4GlcNAcβ1-3Galβ1-4Glcβ1-Cer
globo	GalNAcβ1-3Galβ1-4Galβ1-4Glcβ1-Cer
isoglobo	GalNAcβ1-4Galβ1-4Galβ1-4Glcβ1-Cer
ganglio	Galβ1-3GalNAcβ1-4Galβ1-4Glcβ1-Cer
muco	Galβ1-3Galβ1-4Galβ1-4Glcβ1-Cer
gala	Galα1-4Galβ1-Cer
sulphatide	3-*O*-sulpho-Galβ1-Cer
arthro	GalNAcβ1-4GlcNAcβ1-3Manβ1-4Glcβ1-Cer

Figure 2
Major categories of glycosphingolipid. The six basic cores of mammalian glucosylceramides, two types of mammalian galactosylceramide and the core of an insect/nematode *arthro*-type glycosphingolipid are shown [7,17]

In the case of glycosylated lipids, the glycosylphosphatidylinositol (GPI) membrane anchors and the glycosphingolipids represent distinct categories of glycoconjugate. The GPI anchors replace traditional *C*-terminal transmembrane domains on a number of proteins such as decay accelerating factor (DAF) and a form of the neural cell adhesion molecule (N-CAM-120); they are based, as the name suggests, on a phosphatidylinositol lipid moiety elongated with a $GlcNMan_3EtNPO_4$ structure; the *C*-terminus of the mature protein is linked through the ethanolamine moiety via an amide bond [16]. Glycosphingolipids, on the other hand, are not covalently linked to protein, contain a core glucose or galactose residue linked to ceramide and often act as receptors for bacteria or their toxins. These two monosaccharyl lipids are modified by a variety of structures—the mammalian glucosylceramides fall into six categories (*lacto*-, *lactoneo*-, *globo*-, *isoglobo*-, *ganglio*- and *muco*-series; *Fig. 2*), whereas the generally simpler galactosylceramides contain a *gala*-core or are sulfatides (sulfatoglycosphingolipids) [7,17]. Another mammalian sulfoglycolipid is, on the other hand, a glycerolipid: the seminolipid sulfo-3-monogalactosylalkylacylglycerol [18]. Finally, polysaccharides also occur in mammals—glycogen, which indeed can be seen as being a tyrosine-linked *O*-glycan [19], and the glycosaminoglycan hyaluronic acid, a 'free' glycosaminoglycan [20].

3 Components and Types of Glycoconjugates in Other Eukaryotes

It is obvious that much glycan-related research has centred on mammals; however, the glycosylation of many other eukaryotes has been studied, either due to curiosity or due to the clinical and biotechnological relevance of these systems. One major point is that the glycosylation of unicellular organisms, plants, fungi and invertebrates often shows significant differences to that of mammals and of vertebrates in general. These differences are then considered 'foreign' by mammalian immune systems and result in immune reactions to these glycans. A classic instance is the antiserum raised against the plant glycoprotein horseradish peroxidase; the antiserum recognizes predominantly the xylose and core α1,3-fucose residues on plant *N*-glycans [21]. These features are shared with many other plants and invertebrates—thus resulting in cross-reactivity; for instance, this same antiserum recognizes the neural tissue of

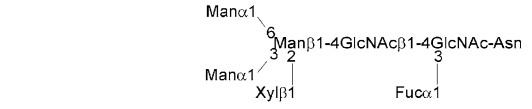
many invertebrates and this is also due to the presence of core α1,3-fucose in, e. g., fruitfly and nematode glycoproteins [22]. Furthermore, this same structural motif is often an epitope for the immunoglobulin E (IgE) from patients with plant food, pollen or insect venom allergies [23].

The example of anti-horseradish peroxidase highlights the significance of the presence, in plant and invertebrate *N*-glycans, of residues or linkages not found in the *N*-glycans of mammals; conversely, there are residues in mammalian *N*-glycans which are not, or perhaps only rarely, found in plant, fungal and invertebrate glycans—sialic acids being a prime example [24]. Indeed, the 'complex' *N*-glycans found in vertebrates (e. g., those with up to five outer antennae with sialic acids, bisecting GlcNAc and possibly core α1,6-fucose) are not found in 'lower' species. Although the pentasaccharide core is conserved, the tendency is that the *N*-glycans are simpler than in mammals, with many truncated 'paucimannosidic' structures being present, and that the most 'complex' *N*-glycans carry fucose residues (core and outer arm) as well as, sometimes, galactose and *N*-acetylgalactosamine (❯ *Fig. 3*). Indeed, some of the structures are 'complex' in the sense of having a 'complicated' structure, even if the modifications do not resemble those of complex glycans found in vertebrates. For instance, in plants, the most complex *N*-glycans carry Lewis a epitopes in addition to core fucose and xylose [25], in the trematode parasite *Schistosoma* species Lewis x, core fucose and xylose are present [26], whereas the fruitfly *Drosophila melanogaster* and the honeybee *Apis mellifera* have glycans with two core fucose residues [27,28] and the nematode worm *Caenorhabditis elegans* expresses *N*-glycans carrying up to four fucose residues and methyl groups [29]. Further modifications with phosphorylcholine in nematodes [30] and phosphorylethanolamine in an insect species [31] have also been reported. All these species also express the oligomannosidic *N*-linked glycans found in mammals.

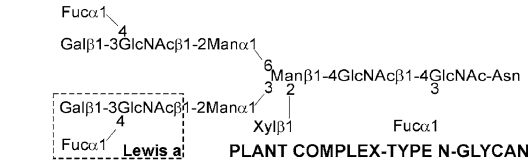
In fungi, the tendency is to have *N*-glycans containing mainly mannose residues, which are sometimes phosphorylated. 'Hypermannosylation', as well as phosphomannosylation, is a common feature in yeasts [32], which limits the biotechnological potential of fungal species such as *Saccharomyces cerevisiae* and *Pichia pastoris* as regards the production of glycoprotein pharmaceuticals—although re-engineering of glycosylation pathways has emerged as a technology to circumvent such problems [33]. Occasionally, residues such as galactose and pyruvate are also found in yeast, such as in *Schizosaccharomyces pombe* [34]; furthermore, antigenic galactofuranose and β1,2-linked mannose residues, respectively, are found on the *N*-glycans of the pathogens *Aspergillus fumigatus* and *Candida* spp [35,36]. In unicellular organisms, such as trypanosomatids, both oligomannosidic and complex structures are found [37], whereas in the slime mould (or social amoeba) *Dictyostelium discoideum*, core fucose and 'intersecting' GlcNAc residues are also present [38].

In the case of *O*-glycans, the situation is more variable—*O*-linked mannosylation being a typical feature of yeasts and fungi [39]; insects possess short 'mucin-type' *O*-glycans [40], whereas the *O*-fucosylation and *O*-mannosylation pathways have been shown to be developmentally important in the fruitfly [41,42]. Nematodes also have 'mucin-type' *O*-glycans [43], with some unusual modifications with glucose, fucose and glucuronic acid being reported. Also, *O*-linked chondroitin and heparan sulfate glycosaminoglycan chains are present in the proteoglycans of invertebrates [44].

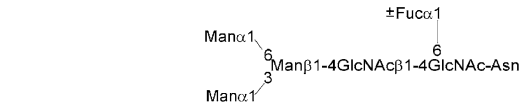
More unusual forms of *O*-glycosylation, such as linkages of fucose and GlcNAc via phosphate to serine and of a pentasaccharide to a hydroxyproline of the cytoplasmic Skp1 protein,



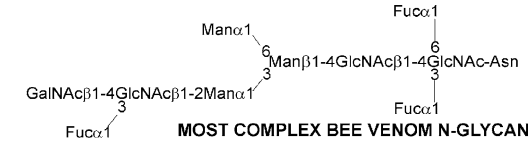
PLANT VACUOLAR/TRUNCATED N-GLYCAN



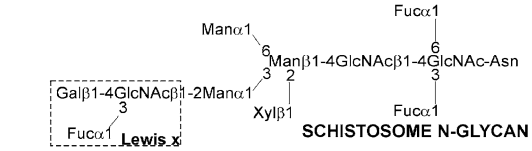
PLANT COMPLEX-TYPE N-GLYCAN



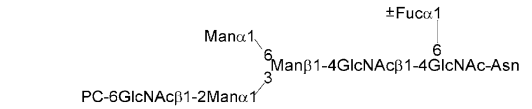
INSECT/NEMATODE PAUCIMANNOSIDIC N-GLYCAN



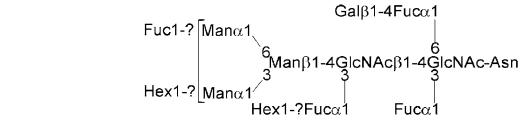
MOST COMPLEX BEE VENOM N-GLYCAN



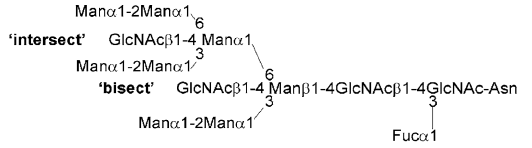
SCHISTOSOME N-GLYCAN



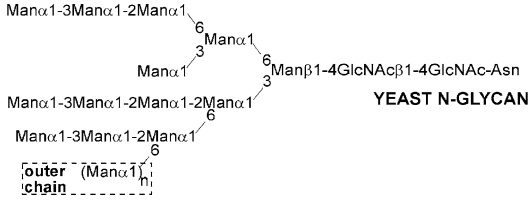
NEMATODE PC-MODIFIED N-GLYCAN



NEMATODE N-GLYCAN WITH CORE MODIFICATIONS



PUTATIVE SLIME MOULD N-GLYCAN



YEAST N-GLYCAN

Figure 3
***N*-glycans from non-mammalian eukaryotic species. Selected *N*-linked oligosaccharides are shown from plants (vacuolar-type as found on, e. g., horseradish peroxidase and a 'Lewis-type' found on, e. g., laccase; Ref. [25]), insects (a 'paucimannosidic-type', the most common *N*-glycan in *Drosophila melanogaster*, and the most complex structure found on bee venom phospholipase A_2; Refs. [27] and [28]), the trematode parasite *Schistosoma mansoni* (a structure with Lewis x, core fucose and xylose modifications; Ref. [26]), nematodes (a phosphorylcholine-containing structure of the type found in many nematodes and a putative structure with incompletely characterized modifications from *Caenorhabditis elegans*; Refs. [29] and [30]), the slime mould *Dictyostelium discoideum* (a putative structure with bisecting and intersecting GlcNAc and core fucose; based on Ref. [38] and unpublished data) and the yeast *Saccharomyces cerevisiae* (the outer chain is absent from *S. cerevisiae* intracellular glycoproteins such as invertase, but in other cases can contain up to n = 200 residues [32]; similar structures with different numbers of mannose residues sometimes in different linkages are found in other yeast species). Oligomannosidic *N*-glycans of the same type as in mammals (see *Fig. 1*) are also present in a range of 'lower' eukaryotes**

have been found in the slime mould [45,46,47]; various novel *O*-glycan types have also been observed in trypanosomatids—such as the 'mucins' of *Trypanosoma cruzi* containing *O*-linked GlcNAc (rather than GalNAc) [48] and the phosphoglycans of *Leishmania* [49]. In plants, quite different *O*-glycans are found which contain arabinose and galactose residues [50]. Plants are also a rich source of various simple glycosides, including glucosylated and rhamnosylated forms of flavonoids and hormones, which are products of the large family of UDP-glycosyltransferases [51].

Glycolipids of 'lower' eukaryotes, too, display major differences as compared to mammals. In insects and nematodes, the *arthro*-series of glycosphingolipids possess a GalNAc-β1,4GlcNAcβ1,3Manβ1,4Glc core attached to ceramide [52,53] and can act as 'receptors' for bacterial toxins [54]; fungi have glycoinositolphosphosphingolipids containing mannose and sometimes other sugars [55]; plants have galactosyldiacylglycerols and sulfoquinovosyldiacylglycerol in chloroplast membranes [56,57]; trypanosomatids have a wide-range of species-specific GPI-like glycolipids [58]. Those protein-bound GPI anchors from 'lower' eukaryotes have, regardless of source, the same core as in mammals, although the decorations are different [16]. Finally, the picture is completed by the vast world of polysaccharides: chitins (poly-GlcNAc) and glucans in some invertebrates and yeasts [59,60], the glucuronoxylomannan of the pathogenic fungus *Cryptococcus neoformans* [61], starch [62] and the various cell wall components of plants, such as cellulose (also a feature of amoebae), hemicelluloses (e. g., xyloglucans) and pectins (e. g., rhamnogalacturonans), which contain a very wide range of monosaccharides not found elsewhere [63]. Probably, very many other oligo- and polysaccharides remain to be discovered or be fully characterized.

4 Components and Types of Glycoconjugates in Prokaryotes

The glycoworld of prokaryotes (bacteria and archaea) is quite different from that of eukaryotes. The cell walls of prokaryotes contain a bewildering number of monosaccharide components—indeed, peptidoglycans, lipopolysaccharides, capsular polysaccharides, *S*-layer glycoproteins, mycobacterial lipoglycans and arabinogalactans as well as Gram-positive bacterial secondary cell wall polymers contain many novel sugar residues: therefore, there are also many unusual donor substrates found in prokaryotes, with the CDP-glucose [64] and dTDP-glu-

cose [65] pathways being important in the synthesis of various bacterial antigens containing residues such as tyvelose (a 3,6-dideoxyhexose) and rhamnose (6-deoxy-L-mannose). Another interesting feature is that the prokaryotic genes required for biosynthesis of donor substrates, as well as those encoding the glycosyltransferases themselves, are often grouped into gene clusters [66].

What has also become very obvious in the past few years is that bacteria can indeed glycosylate proteins, even if this is not the case in standard laboratory strains of *Escherichia coli*: the sugar-protein linkages are often unlike those of eukaryotes. For instance, in the case of the pathogen *Campylobacter jejuni*, a glycan containing bacillosamine (2,4-diacetamido-2,4,6-trideoxyglucose) with further *N*-acetylgalactosamine residues is transferred to asparagine [67], whereas the archaeal species *Methanococcus voltae* has an *N*-glycan of the form Man*p*NAcA6Thr-β1,4Glc*p*NAc3NAcAβ1,3Glc*p*NAc [68]. Various *O*-glycans with different novel linkages to protein are also known, including addition of heptoses to specific proteins in wild-type strains of *E. coli* [69], of FucNAc to the pilin of *Pseudomonas aeruginosa* [70], of glucose to tyrosine in the *S*-layer glycoprotein of *Thermoanaerobacterium thermosaccharolyticum* [71] and of *O*-linked mannose (just as in eukaryotes) to mycobacterial proteins [72]. Furthermore, the glycosylation of eukaryotic proteins by prokaryotic enzymes (e. g., glucosylation and ADP-ribosylation) is important in the mechanism of some bacterial toxins [73,74]. The reader is directed to specialist reviews on the various classes of prokaryotic glycoconjugates [75,76,77,78,79,80,81,82,83,84].

5 The Biosynthesis of Activated Monosaccharides

The pre-requisites for the formation of a glycosidic bond are a relevant glycosyltransferase, an acceptor and an activated donor sugar. These glycosyltransferase donors are most commonly nucleotide sugars (NDP-sugars) such as UDP-Glc, GDP-Fuc or CMP-NeuAc [7,85]; however, also important are the lipid-linked (dolichol-linked in eukaryotes) donors such as Dol-P-Man and Dol-P-Glc [86]. Furthermore, some glycosyltransfer reactions require an oligosaccharyl donor: in eukaryotes, a dolicholpyrophosphoryloligosaccharide is the donor for *N*-glycosylation [87,88,89,90] and similar lipid-linked (polyisoprenyl) donors are a feature of glycoconjugate biosynthesis in bacteria [91,92]. However, the oligosaccharide donors themselves contain residues which originated as nucleotide sugars.

In order for a monosaccharide to be biosynthetically active, it is generally phosphorylated. In the case of glucose, the free sugar obtained by hydrolysis of various polymers is brought into the hexose phosphate pool by the action of hexokinase [7,93]. The resulting glucose-6-phosphate can be converted into fructose-6-phosphate, which can then enter the glycolysis pathway. Other fates of these hexose-6-phosphates exist: fructose-6-phosphate can be acted upon by phosphomannose isomerase to form mannose-6-phosphate, thus beginning the pathway necessary for formation of mannose and fucose donors [94]; an amidotransferase and an acetyltransferase, using glutamine and acetyl coenzyme A as the respective 'donors', facilitate the formation of GlcNAc [95] and, indirectly, GalNAc [96] and sialic acid [97] donors; glucose-1-phosphate is also required for UDP-Glc synthesis and the latter can be epimerized to form UDP-Gal [98] or oxidized to UDP-GlcA, which can then be decarboxylated to form UDP-Xyl [99]. Thus, we can account for the formation of the typical monosaccharide com-

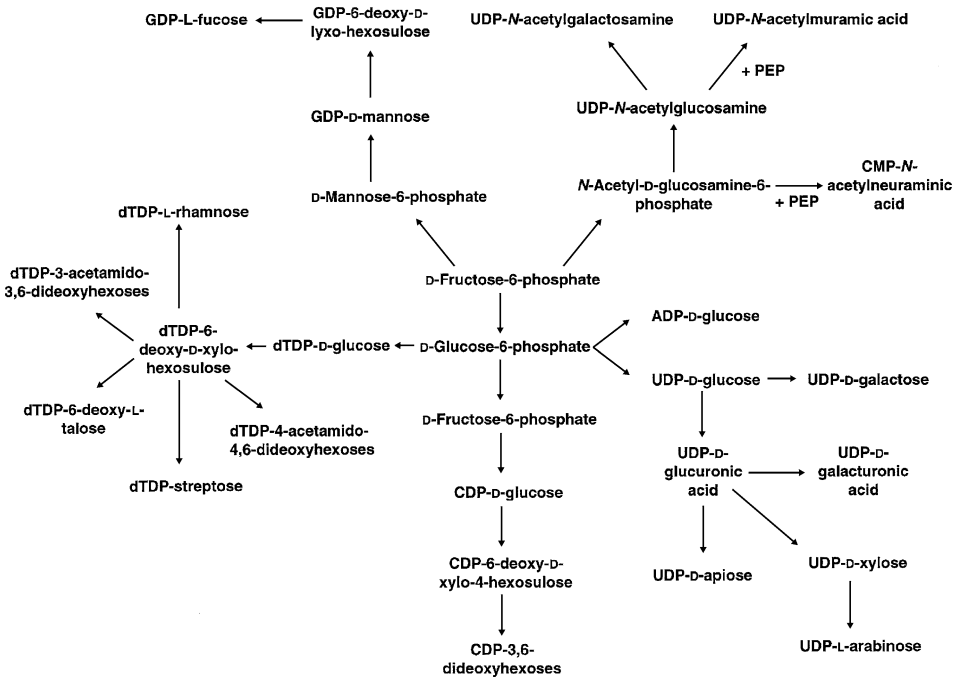


Figure 4

Pathways in the formation of NDP-sugars. A selection of monosaccharide conversions occurring at the phosphate ester and nucleotide-monosaccharide level in animals, plants and bacteria are shown; all pathways radiate from fructose-6-phosphate, indicating the central role of this metabolite. The dTDP, CDP and UDP-*N*-acetylmuramic pathways are peculiar to bacteria, whereas ADP-D-glucose, UDP-D-apiose and UDP-L-arabinose are found in plants. For reasons of simplicity, other pathways, e. g., to UDP-L-rhamnose in plants, GDP-D-rhamnose in *Pseudomonas aeroginosa* and GDP-D-arabinose in trypanosomatids are not included. The figure is reproduced from Chap. 6.4 of the first edition of this book

ponents of glycoconjugates found in mammals: *N*-acetylglucosamine, mannose, galactose, glucose, *N*-acetylgalactosamine, glucuronic acid, fucose, sialic acids and xylose (see *Fig. 4* for a selection of NDP-sugar biosynthetic pathways). Similar pathways are obviously necessary in invertebrates, whereas the large number of cell wall modifications in plants require formation of other donors, such as those for arabinosyl-, rhamnosyl- and galacturonyltransferases [100]. As mentioned above, a wide range of monosaccharides and relevant NDP-sugar biosynthesis pathways are present in bacteria [66]. Many of the enzymes involved in the formation of NDP-sugars, such as the two enzymes converting GDP-Man to GDP-Fuc or the UDP-hexose-4-epimerases, are members of the short-chain dehydrogenase/reductase family, whose *N*-terminal regions bind the cofactors NAD(H) or NADP(H) [101]. Any substitutions, such as methylation (e. g., in plants, nematodes, gastropods and some bacteria), sulfation and phosphorylation of non-reducing termini, are assumed to be a result of modification of the completed glycans by relevant transferases using activated donors such as *S*-adenosylmethionine [102], 3′-phosphoadenosine 5′-phosphosulfate [103] or NDP-donors.

In addition to the various de novo pathways for activated monosaccharide formation, 'salvage' pathways exist in which monosaccharides, present in the bloodstream or released by lysosomal degradation of macromolecules, can re-enter the biosynthetic pathways via phosphorylation [104,105,106,107]; thus, various monosaccharide kinases exist to facilitate the synthesis of new activated sugars from 'old' components. These salvage pathways are of interest since they can be used to 'rescue' defects in de novo biosynthesis, such as the mutation in phosphomannose isomerase in the Congenital Disorder of Glycosylation, or CDG, Type Ib [108] or a fucosylation defect in the slime mould *Dictyostelium discoideum* [109]. Finally, as reviewed in the next chapter, a large number of other monosaccharides (such as the ribose present in nucleotides and nucleic acids) can be derived from routes related to photosynthesis and glycolysis via the hexose and pentose interconversion pathways; in 'lower' organisms, a wide range of secondary metabolic pathways can also be based upon free monosaccharides.

6 The Transport of Activated Monosaccharides

Activated monosaccharides are generally, but not always, synthesized in the cytoplasm. This means that they are normally in the 'wrong' compartment as compared to the localization, in the ER or Golgi, of most of the glycosyltransferases which utilize them as donor substrates [110]. Thus, either transport or flipping is necessary for them to cross the endomembranes (in eukaryotes) of the endoplasmic reticulum and Golgi apparatus. In the case of dolichol-linked monosaccharides (Dol-P-Glc and Dol-P-Man), a flipping event occurs whose exact nature is still uncertain; for nucleotide sugars as well as the sulfate donor 3′-phosphoadenosine 5′-phosphosulfate (PAPS), though, there exist a number of transporters which facilitate the entry of the donor substrates in exchange for the relevant nucleoside monophosphate (e. g., GMP or UMP), which 'returns' to the cytoplasm [111]. A number of these transporters have been characterized recently—in all cases thus far, they contain multiple transmembrane domains and all are members of the solute carrier family 35 (SLC35) [112]. A number of defects in man and in model organisms (as well as farm animals) have been found to be due to mutations in SLC35 gene family members; examples include human Leukocyte Adhesion Deficiency II (also known as CDG IIc), which is caused by a defect in GDP-Fuc transport and, so, results in a lack of the fucosylated structures required for lymphocyte homing [113], and bovine Complex Vertebral Malformation syndrome, which correlates with a defective UDP-GlcNAc transporter gene in descendents of a single elite bull [114].

7 Heterogeneity of Oligo- and Polysaccharides

The pathways for the biosynthesis of the various glycoconjugates in prokaryotes and eukaryotes are often complex—furthermore, these pathways are, unlike nucleic acid or protein biosynthesis, not template-driven. Indeed, most "purified" glycoconjugates are mixtures—resulting in so-called microheterogeneity; in the case of glycoproteins, the native isolated material consists of a series of glycoforms [115]. Even a simple glycoprotein such as bovine pancreatic ribonuclease B can carry, on its single *N*-glycosylation site, structures with anything between five and nine mannose residues [116]; even ignoring the potential isomers of oligomannosidic *N*-glycans, this means that a population of ribonuclease molecules is

a mixture of at least five isoforms (i. e., glycoforms), with some molecules bearing *N*-glycans with five mannose residues, some with six, and so on. Thus, the potential for a high degree of heterogeneity becomes very large in the case of proteins with many glycosylation sites and, in the whole population of molecules, a variety of glycan structures. For proteoglycans and plant polysaccharides, the potential for microheterogeneity is even larger due to the sheer size of the glycan chains.

The glycoforms observed in any mixture purified from a natural eukaryotic source are the products of the interplay between the various enzymes (whether these are glycosyltransferases or 'processing' glycosidases) localized, except when we think of nucleocytoplasmic glycosylation, within the secretory pathway—the relative amounts expressed, their localization, their exact substrate specificities—and the accessibility and availability of the substrates whether acceptor or donor [117,118]. On a glycoprotein, the folding of the peptide backbone can affect the outcome of the various glycomodification steps due to steric effects [119] or the presence, in some special cases, of sequence or structural motifs [120]. Therefore, it is no wonder that the biosynthesis of these molecules results in heterogeneous products. It may be that this is 'desired', in order that not all glycoproteins have the same glycosylation—the addition of variation beyond that directly encoded by the genome may confer advantages in terms of having a spectrum of activities and binding specificities. Certainly, though, the lack of a template means that glycoscientists have a major challenge not just in determining the structures present, but also as to the metabolic routes necessary for their synthesis.

The major 'workhorses' in the various biosynthetic pathways are, of course, the glycosyltransferases [7]. However, in some pathways (e. g., *N*-glycosylation), selective trimming of glycans opens up new routes—thus, the $Glc_3Man_9GlcNAc_2$ structure transferred by eukaryotic oligosaccharyltransferases (with the exception of some protists which transfer shorter glycans to protein) from a dolichol-linked precursor to Asn-Xaa-Ser/Thr/Cys motifs of nascent polypeptides (i. e., those still being synthesized by ribosomes bound to the endoplasmic reticulum) is initially not the substrate for glycosyltransferases, but for glucosidases, mannosidases and sometimes an endomannosidase [121]. Indeed, it is only when the *N*-glycan is trimmed down by 'processing' (rather than 'degradative') glycosidases to $Man_5GlcNAc_2$ that we come to a point at which glycosyltransferases in most eukaryotes (but not yeasts) begin to add more residues to this core structure [89].

Even the description so far of these biochemical reactions indicates that just a single glycosylation pathway (*N*-glycosylation in eukaryotes) is highly complex. This impression is even more pronounced when one considers that different species, different cell types or different developmental, pathogenic or activation states may result in different glycosylation patterns even of the same polypeptide. This fact explains the concept of 'glycotype' [115], which was first described to account for the glycosylation patterns of Thy-1, a protein expressed in both brain and thymocytes found to carry different *N*-glycans when isolated from different tissues in the same species [122]. This concept becomes more complicated when considering that different species do not have the same numbers of genes encoding glycosyltransferases and glycosidases, although the overall number is generally about 1% of the identified genes in a species [123]. Therefore, the glycosylation potential of a mouse is different from that of a human or a worm, never mind that changes in sialic acid metabolism have the consequence that our own sialylation is different even from that of a chimpanzee [124]. These differences also have an impact on our ability to make conclusions from the use of model organisms in gly-

cobiological research or in the use of heterologous systems for the production of recombinant glycoproteins.

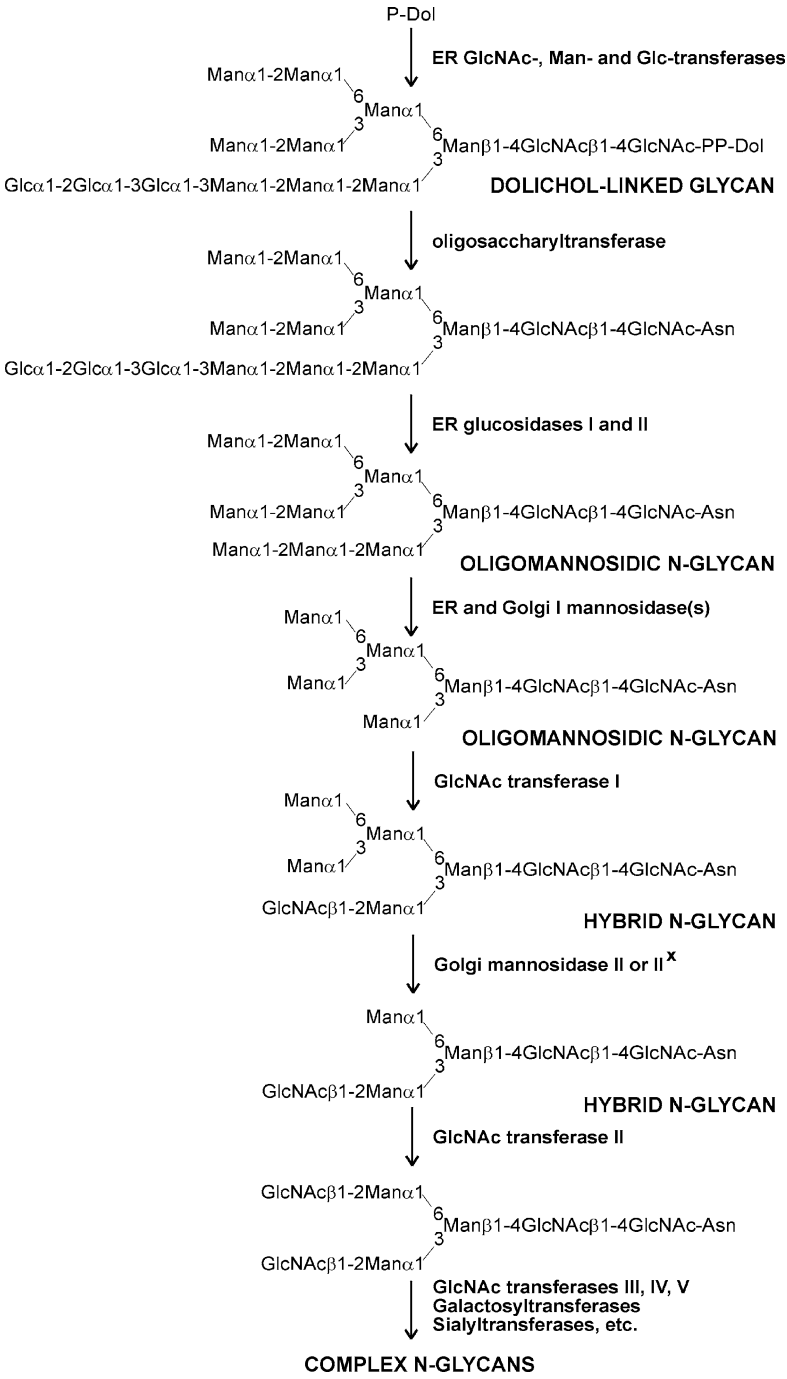
8 Glycosyltransferases and Glycosidases in Biosynthesis of Glycans in Eukaryotes

All the heterogeneity observed in glycoconjugates results from the different abilities of glycosyltransferases and 'processing' glycosidases to modify glycan substrates. In the case of *N*-glycosylation, the process in most eukaryotes begins with the synthesis of a dolichol-linked tetradecasaccharide (the aforementioned $Glc_3Man_9GlcNAc_2$ structure; see *Fig. 5*). The first steps in the synthesis of this structure are the result of the action of two *N*-acetylglucosaminyltransferases and five mannosyltransferases which utilize UDP-GlcNAc and GDP-Man as donor substrates and whose catalytic domains are on the cytoplasmic face of the endoplasmic reticulum [89]; then in a process seemingly mediated by a protein encoded by the so-called *RFT1* gene [125], the resulting heptasaccharide, still linked to dolichol, is flipped from the cytoplasmic to the lumenal face of the endoplasmic reticulum. The final four mannose and three glucose residues are 'donated' by Dol-P-Man and Dol-P-Glc; thereafter, oligosaccharyltransferase catalyses the transfer to protein [87,89]. In higher eukaryotes, the majority of protein-linked *N*-glycans are trimmed, by a series of glycosidases in the endoplasmic reticulum and early Golgi, down to $Man_5GlcNAc_2$ [121], which is then the substrate for an *N*-acetylglucosaminyltransferase (GlcNAc-TI or GnTI; EC 2.4.1.101) in the Golgi apparatus, with UDP-GlcNAc being the donor substrate [126]. In yeast, though, trimming does not progress beyond the $Man_8GlcNAc_2$ stage [127]; however, this decasaccharide is the substrate for subsequent mannosylation reactions in the yeast Golgi apparatus, commencing with transfer of an α1,6-linked mannose by the Och1p '**o**uter **ch**ain' mannosyltransferase [32,128,129]. In other unicellular organisms, other trimming and decoration events can occur, although oligomannosidic structures may predominate.

In mammals, the action of GlcNAc-TI 'heralds' the beginning of all the routes leading to hybrid and complex oligosaccharides [117,130]. First the two outer mannose residues from the hybrid oligosaccharide $Man_5GlcNAc_3$ (sometimes called Man5Gn) are trimmed by Golgi mannosidases II or II^x [130,131]; the remaining non-substituted α1,6-linked mannose is then the target for the next *N*-acetylglucosaminyltransferase known as GlcNAc-TII or GnTII [126], whose action generates a complex *N*-glycan. A wide variety of other modifications by branching and bisecting *N*-acetylglucosaminyltransferases (GlcNAc-TIII-VI) are subsequently possible [117,130], depending on the tissue, species and protein; then, most commonly, galactosyltransferases, fucosyltransferases and sialyltransferases may act, with sulfation, glu-

Figure 5

An outline of *N*-glycan biosynthesis in mammals. The summarized progression of the biosynthetic pathway of *N*-glycans in mammals is shown, starting with the formation of the dolichol-linked $Glc_3Man_9GlcNAc_2$ precursor in the endoplasmic reticulum (ER) and ending with the formation of complex *N*-glycans in the Golgi apparatus. The steps as far as the transfer to protein by oligosaccharyltransferase are generally conserved also in eukaryotic micro-organisms, plants and animals. For reasons of clarity, the whole multitude of potential pathways, by-pass reactions and branching and elongation events are not included. See text for references


P-Dol
ER GlcNAc-, Man- and Glc-transferases
Manα1-2Manα1
6
3
Manα1
Manα1-2Manα1
6
Manβ1-4GlcNAcβ1-4GlcNAc-PP-Dol
3
Glcα1-2Glcα1-3Glcα1-3Manα1-2Manα1-2Manα1
DOLICHOL-LINKED GLYCAN
oligosaccharyltransferase
Manα1-2Manα1
6
3
Manα1
Manα1-2Manα1
6
Manβ1-4GlcNAcβ1-4GlcNAc-Asn
3
Glcα1-2Glcα1-3Glcα1-3Manα1-2Manα1-2Manα1
ER glucosidases I and II
Manα1-2Manα1
6
3
Manα1
Manα1-2Manα1
6
Manβ1-4GlcNAcβ1-4GlcNAc-Asn
3
Manα1-2Manα1-2Manα1
OLIGOMANNOSIDIC N-GLYCAN
ER and Golgi I mannosidase(s)
Manα1
6
3
Manα1
Manα1
6
Manβ1-4GlcNAcβ1-4GlcNAc-Asn
3
Manα1
OLIGOMANNOSIDIC N-GLYCAN
GlcNAc transferase I
Manα1
6
3
Manα1
Manα1
6
Manβ1-4GlcNAcβ1-4GlcNAc-Asn
3
GlcNAcβ1-2Manα1
HYBRID N-GLYCAN
Golgi mannosidase II or II x
Manα1
6
Manβ1-4GlcNAcβ1-4GlcNAc-Asn
3
GlcNAcβ1-2Manα1
HYBRID N-GLYCAN
GlcNAc transferase II
GlcNAcβ1-2Manα1
6
Manβ1-4GlcNAcβ1-4GlcNAc-Asn
3
GlcNAcβ1-2Manα1
GlcNAc transferases III, IV, V
Galactosyltransferases
Sialyltransferases, etc.
COMPLEX N-GLYCANS

curonylation or addition of *N*-acetylgalactosamine also occurring. This means that mammalian cells require also transporters capable of facilitating the import of UDP-GlcNAc, UDP-Gal, GDP-Fuc, CMP-Sia, UDP-GlcUA, UDP-GalNAc and PAPS into the Golgi lumen [110]. In some cases, though, it is the oligomannosidic structures which are modified in mammals—the addition of GlcNAc-1-phosphate to the C6-hydroxyl of certain non-reducing terminal mannose residues is a pre-requisite for the formation of the 'mannose-6-phosphate' marker for lysosomal enzymes [132]. The entire process of *N*-glycan maturation may, according to pulse-chase experiments, take about 30 minutes, prior to final delivery of a glycoprotein to the cell surface [90].

In plants and invertebrates, the *N*-glycome (i. e., the sum of different *N*-glycan structures) tends to be simpler—dominated by either oligomannosidic structures or oligosaccharides carrying fucose residues. However, the range of possibilities is huge and it is remarkable that some 'small' organisms such as the nematode *Caenorhabditis elegans* seemingly have a more complex *N*-glycome [29,133,134] than a 'larger' animal such as the fruitfly *Drosophila melanogaster* [27,135]. In invertebrate species, the processing of *N*-glycans first requires the transfer of one GlcNAc residue by GlcNAc-TI; then mannosidase II, core α1,6-fucosyltransferase and, optionally, core α1,3-fucosyltransferase can act [24,136]. Subsequently, the GlcNAc residue originally transferred by GlcNAc-TI is often removed by a processing hexosaminidase present in the secretory pathway [137,138]. In plants, the action of GlcNAc-TI is also important for the modification of the core by fucose and xylose [139,140]; however, the common plant-type glycan structure lacking non-reducing terminal GlcNAc residues results from the action of vacuolar hexosaminidases [141]. Alternatively, a β1,3-galactosyltransferase and an α1,4-fucosyltransferase generate Lewis a epitopes on the proteins secreted by plant cells [24].

For *O*-glycosylation, the overall concept of modification is different from that of *N*-glycans; in this case, regardless of the type of *O*-glycan formed, the residues are transferred to the protein one-at-a-time. There is no lipid-linked oligosaccharide precursor: the transfer of the first sugar to the peptide is followed by sequential addition of the subsequent residues. Furthermore, since there is no larger precursor, there are probably no '*O*-glycan-processing' glycosidases; only a wide variety of glycosyltransferases are involved in *O*-glycan biosynthesis. Since there are a number of different types of *O*-glycosylation, a large number of enzymes can initiate the process (again in contrast to the 'single' oligosaccharyltransferase in *N*-glycan biosynthesis). Indeed, even for one type of linkage, there may be multiple initiating enzymes; for instance, in the case of mucin-type *O*-glycosylation in vertebrates and invertebrates, there are about twenty polypeptide-modifying *N*-acetylgalactosaminyltransferases (EC 2.4.1.41) encoded by the human genome [142]; thereafter 'core 1' β1,3-galactosyltransferase and various β-*N*-acetylglucosaminyltransferases may act to form the different *O*-glycan cores known in mammals [117,130].

In the case of *O*-fucose, there are two types of *O*-fucosyltransferase (EC 2.4.1.221), specific for either EGF or TSR domains, whose action is followed by, respectively, the β1,3-*N*-acetylglucosaminyltransferase known as Fringe or a β1,3-glucosyltransferase [143]. In mammals, there are two polypeptide-modifying xylosyltransferases (EC 2.4.2.26) which are the first of a series of glycosyltransferases required for the synthesis of the chondroitin and heparan sulfates [144]. The first residue of *O*-mannosyl glycans is transferred by a heterodimeric peptide-*O*-mannosyltransferase in mammals [145], whereas in yeast there are a number of

potential mannosyltransferase subunits [146]. The glycosphingolipids are also built up one glycosyl residue at a time [147,148]; in contrast, the GPI anchors are 'pre-assembled' prior to the transamidation reaction which results in the formation of the bond to the new *C*-terminus of the acceptor protein [16,149]. Some further details about these reactions will be discussed in Chap. 11.2.

Obviously, all these various glycosylation events require an enzymatic machinery, but our understanding of it has grown only relatively slowly. The beginnings of glycosyltransferase research were in the 1960s with the first activities being detected and the first enzyme purifications being attempted. The basic problem was the low amounts in which the majority of glycosyltransferases are expressed. Bovine β1,4-galactosyltransferase (which acts as lactose synthase when complexed with α-lactalbumin; EC 2.4.1.22) was one of the first glycosyltransferases to be partially purified [150] and was for many years the most studied due to the large amounts present in milk; this was an advantage that the vast majority of glycosyltransferases did not have. Another problem is that biological extracts may contain a mixture of activities with overlapping specificities: this is something which we realize only in retrospect. Indeed, the 'one linkage, one enzyme' hypothesis of Hagopian and Eylar [151] has to be modified in the light of the use of molecular biology, which has revealed a higher degree of complexity, than was first imagined, as regards the glycogenomic capacities of various organisms. Often, in the case of the terminal modifications, multiple genes encode enzymes with subtly different activities. The cloning of glycosyltransferase cDNAs, which began in the 1980s also with the β1,4-galactosyltransferase [152,153], has proved most valuable in gaining large amounts of recombinant enzymes for characterization in terms of substrate specificities and crystal structures. Most recently, the revolution in molecular glycobiology has been in terms of reverse genetics—targeted knock-outs and knock-downs have yielded new perspectives as to the biological function of glycosylation [154,155].

9 The Degradation of Oligo- and Polysaccharides

Biological macromolecules such as proteins and carbohydrates are rarely long-lasting, but are subject to turnover; thus, degradation is an integral part of the 'life' of an oligosaccharide (see also Chap. 11.3). Indeed, a deficiency in degradation can be as problematic as a mutation in a biosynthetic pathway: this is particularly obvious with the lysosomal storage diseases of humans and domesticated animals, for which the defective genes in a relatively long list of diseases correspond to a number of steps in glycosaminoglycan and glycolipid catabolism [6,7,156]. The degradation of many glycoconjugates is not, though, a simple reversal of their biosynthesis. Furthermore, the enzymes responsible, primarily the glycosidases, are highly numerous, although normally not showing perhaps the same exact linkage and substrate specificities displayed by glycosyltransferases.

In the case of *N*-glycans, degradation in the mammalian lysosome involves the peptide being subject to proteolysis and the release of the glycan by a peptide:*N*-glycosidase, followed by an endo-β-*N*-acetylglucosaminidase in some species, then by removal of the non-reducing terminal sugars [157]. There are also pathways to degrade glycans from glycoproteins which have failed the 'quality control process' in the endoplasmic reticulum [158,159] or from dolichyl-linked oligosaccharides [158]. For glycolipids, there are some unique glycosidases includ-

ing an enzyme releasing the reducing-terminal residue from ceramide and, in some cases, the glycosidases require 'sphingolipid activator proteins' (saposins) in order to cleave their substrate [148,160]; in the case of glycosaminoglycan chains, endoglycosidases, exoglycosidases and sulfatases are important in the degradation process [161].

Plants, particularly their vacuoles, are a rich source of glycan-degrading enzymes: examples, used in laboratories to determine glycan structure, include the peptide:*N*-glycosidase A from almonds [162] and the various jack bean exoglycosidases [163,164,165]. For microorganisms and fungi, on the other hand, it is not just the turnover of the endogenous glycans or the ability to degrade glycoconjugates for purposes of gaining nutrients which are to be considered: glycosidases also play roles in pathogenicity. For instance, the influenza virus has a neuraminidase, which is a drug target and is important for movement and release of viruses [166].

Finally, two other groups of glycan-degrading enzymes are the polysaccharide lyases (E.C. 4.2.2.-) and the phosphorylases. Examples of the first are the pectate lyases, which generate unsaturated galacturonosyl residues at their non-reducing ends: these enzymes are produced by both plants and micro-organisms and have roles, respectively, in fruit ripening and pathological cell wall damage [167]. Other lyases from bacteria, specific for glycosaminoglycans and, therefore, known as chondroitinases or heparinases [168,169], are used in structural determination studies. The phosphorylases, though, are primarily of interest due to their roles in metabolism: phosphorolysis of starch or glycogen within cells results in the production of glucose-1-phosphate, which can then directly be used in hexose metabolism; although the importance of starch phosphorylase (EC 2.4.1.7) in plants is currently not clear [170], human glycogen phosphorylase (EC 2.4.1.1) is a potential therapeutic target in diabetes [171].

10 Concluding Remarks

In the previous pages I have sought to summarize a number of basic concepts as regards the biosynthesis and degradation of mono-, oligo- and polysaccharides, with an emphasis on those pathways involved in the formation of eukaryotic glycoconjugates. In the next two chapters on 'Molecular Basis for the Biosynthesis of Oligo-, and Polysaccharides' and 'Polysaccharide Degradation', each of these areas will be discussed in more depth. It is sufficient to state that glycoconjugate metabolism is a complex web of biochemical reactions with a large number of species- and tissue-specific factors and that modern methods of molecular and structural biology have in recent years greatly expanded our knowledge of these pathways. This knowledge, which still needs to be supplemented, is certain to have a major impact in biotechnology and biomedicine in the future.

Acknowledgement

I thank Drs. Paul Messner and Christina Schäffer (Universität für Bodenkultur Wien) for updating me on aspects of prokaryotic glycosylation and Drs. Frank Unger (Universität Wien), Harry Schachter (Hospital for Sick Children, Toronto), Erika Staudacher, Katharina Paschinger and Dubravko Rendić (Universität für Bodenkultur Wien) for their comments on the chapter as a whole.

References

1. Armstrong EF (1919) The simple carbohydrates and the glucosides, 3rd edn. Longmans, Green and Co., London
2. Colli W (1993) FASEB J 7:1257
3. Campbell P, Braam J (1999) Trends Plant Sci 4:361
4. Tewari YB, Goldberg RN (1989) J Biol Chem 264:3966
5. Pérez LF, Tutor JC (1998) Clin Chem 44:226
6. Vellodi A (2004) Br J Haematol 128:413
7. Varki A, Cummings R, Esko J, Freeze H, Hart G, Marth J (1999) Essentials of Glycobiology. Cold Spring Harbour Press, New York
8. Spiro RG (2002) Glycobiology 12:43R
9. Wells L, Vosseller K, Hart GW (2001) Science 291:2376
10. Haltiwanger RS (2002) Curr Opin Struct Biol 12:593
11. Endo T, Toda, T (2003) Biol Pharm Bull 26:1641
12. Wilson IBH (2004) Cell Mol Life Sci 61:794
13. Harris RJ, Spellman MW (1993) Glycobiology 3:219
14. Cunningham LW, Ford JD (1968) J Biol Chem 243:2390
15. de Beer T, Vliegenthart JF, Löffler A, Hofsteenge J (1995) Biochemistry 34:11785
16. Ikezawa H (2002) Biol Pharm Bull 25:409
17. Macher BA, Sweeley CC (1978) Methods Enzymol 50:236
18. Alvarez JG, Storey BT, Hemling ML, Grob RL (1990) J Lipid Res 31:1073
19. Smythe C, Caudwell FB, Ferguson M, Cohen P (1988) EMBO J 7:2681
20. Laurent TC, Fraser JR (1992) FASEB J 6:2397
21. Kurosaka A, Yano A, Itoh N, Kuroda Y, Nakagawa T, Kawasaki T (1991) J Biol Chem 266:4168
22. Paschinger K, Fabini G, Schuster D, Rendić D, Wilson IBH (2005) Acta Biochim Pol 52:629
23. Altmann F (2007) Int Arch Allergy Immunol 142:99
24. Wilson IBH (2002) Curr Opin Struc Biol 12:569
25. Wilson IBH, Zeleny R, Kolarich D, Staudacher E, Stroop CJ, Kamerling JP, Altmann F (2001) Glycobiology 11:261
26. Khoo KH, Huang HH, Lee KM (2001) Glycobiology 11:149
27. Fabini G, Freilinger A, Altmann F, Wilson IBH (2001) J Biol Chem 276:28058
28. Kubelka V, Altmann F, Staudacher E, Tretter V, März L, Hård K, Kamerling JP, Vliegenthart JFG (1993) Eur J Biochem 213:1193
29. Haslam SM, Dell A (2003) Biochimie 85:25
30. Haslam SM, Houston KM, Harnett W, Reason AJ, Morris HR, Dell A (1999) J Biol Chem 274:20953
31. Hård K, Van Doorn JM, Thomas-Oates JE, Kamerling JP, Van der Horst DJ (1993) Biochemistry 32:766
32. Kukuruzinska MA, Bergh MLE, Jackson BL (1987) Annu Rev Biochem 56:915
33. Wildt S, Gerngross TU (2005) Nat Rev Microbiol 3:119
34. Andreishcheva EN, Kunkel JP, Gemmill TR, Trimble RB (2004) J Biol Chem 279:35644
35. Morelle W, Bernard M, Debeaupuis JP, Buitrago M, Tabouret M, Latgé JP (2005) Eukaryot Cell 4:1308
36. Suzuki A, Shibata N, Suzuki M, Saitoh F, Oyamada H, Kobayashi H, Suzuki S, Okawa Y (1997) J Biol Chem 272:16822
37. Parodi AJ (1993) Glycobiology 3:193
38. Sharkey DJ, Kornfeld R (1991) J Biol Chem 266:18485
39. Willer T, Valero MC, Tanner W, Cruces J, Strahl S (2003) Curr Opin Struct Biol 13:621
40. Lopez M, Tetaert D, Juliant S, Gazon M, Cerutti M, Verbert A, Delannoy P (1999) Biochim Biophys Acta 1427:49
41. Okajima T, Irvine KD (2002) Cell 111:893
42. Lyalin D, Koles K, Roosendaal SD, Repnikova E, Van Wechel L, Panin VM (2006) Genetics 172:343
43. Guerardel Y, Balanzino L, Maes E, Leroy Y, Coddeville B, Oriol R, Strecker G (2001) Biochem J 357:167
44. Yamada S, Okada Y, Ueno M, Iwata S, Deepa SS, Nishimura S, Fujita M, Van Die I, Hirabayashi Y, Sugahara K (2002) J Biol Chem 277:31877
45. Srikrishna G, Wang L, Freeze HH (1998) Glycobiology 8:799
46. Mreyen M, Champion A, Srinivasan S, Karuso P, Williams KL, Packer NH (2000) J Biol Chem 275:12164
47. West CM, Van Der Wel H, Sassi S, Gaucher EA (2004) Biochim Biophys Acta 1673:29
48. Jones C, Todeschini AR, Agrellos OA, Previato JO, Mendonca-Previato L (2004) Biochemistry 43:11889

49. Ilg T (2000) Parasitol Today 16:489
50. Léonard R, Petersen BO, Himly M, Kaar W, Wopfner N, Kolarich D, van Ree R, Ebner C, Duus JO, Ferreira F, Altmann F (2005) J Biol Chem 280:7932
51. Bowles D, Isayenkova J, Lim EK, Poppenberger B (2005) Curr Opin Plant Biol 8:254
52. Seppo A, Moreland M, Schweingruber H, Tiemeyer M (2000) Eur J Biochem 267:3549
53. Lochnit G, Dennis RD, Ulmer AJ, Geyer R (1998) J Biol Chem 273:466
54. Griffitts JS, Aroian RV (2005) Bioessays 27:614
55. Jennemann R, Geyer R, Sandhoff R, Gschwind RM, Levery SB, Gröne HJ, Wiegandt H (2001) Eur J Biochem 268:1190
56. Benning C, Ohta H (2005) J Biol Chem 280:2397
57. Frentzen M (2004) Curr Opin Plant Biol 7:270
58. Guha-Niyogi A, Sullivan DR, Turco SJ (2001) Glycobiology 11:45R
59. Cohen E (2001) Pest Manag Sci 57:946
60. Klis FM, Mol P, Hellingwerf K, Brul S (2002) FEMS Microbiol Rev 26:239
61. Bose I, Reese AJ, Ory JJ, Janbon G, Doering TL (2003) Eukaryot Cell 2:655
62. Martin C, Smith AM (1995) Plant Cell 7:971
63. Lerouxel O, Cavalier DM, Liepman AH, Keegstra K (2006) Curr Opin Plant Biol 9:621
64. Koropatkin NM, Holden HM (2004) J Biol Chem 279:44023
65. Zuccotti S, Zanardi D, Rosano C, Sturla L, Tonetti M, Bolognesi M (2001) J Mol Biol 313:831-843
66. Reeves PR, Hobbs M, Valvano MA, Skurnik M, Whitfield C, Coplin D, Kido N, Klena J, Maskell D, Raetz CRH, Rick PD (1996) Trends Microbiol 4:495
67. Wacker M, Linton D, Hitchen PG, Nita-Lazar M, Haslam SM, North SJ, Panico M, Morris HR, Dell A, Wren BW, Aebi M (2002) Science 298:1790
68. Chaban B, Voisin S, Kelly J, Logan SM, Jarrell KF (2006) Mol Microbiol 61:259
69. Benz I, Schmidt MA (2001) Mol Microbiol 40:1403
70. Castric P, Cassels FJ, Carlson RW (2001) J Biol Chem 276:26479
71. Schäffer C, Dietrich K, Unger B, Scheberl A, Rainey FA, Kählig H, Messner P (2000) Eur J Biochem 267:5482
72. Van der Ven BC, Harder JD, Crick DC, Belisle JT (2005) Science 309:941
73. Aktories K, Just I (1995) Trends Cell Biol 5:441
74. Holbourn KP, Shone CC, Acharya KR (2006) FEBS J 273:4579
75. Benz I, Schmidt MA (2002) Mol Microbiol 45:267
76. Briken V, Porcelli SA, Besra GS, Kremer L (2004) Mol Microbiol 53:391
77. Caroff M, Karibian D (2003) Carbohydr Res 338:2431
78. Crick DC, Mahapatra S, Brennan PJ (2001) Glycobiology 11:107R
79. Hitchen PG, Dell A (2006) Microbiology 152:1575
80. Messner P, Schäffer C (2003) Progr Chem Org Nat Prod 85:51
81. Samuel G, Reeves P (2003) Carbohydr Res 338:2503
82. Schäffer C, Messner P (2005) Microbiology 151:643
83. Upreti RK, Kumar M, Shankar V (2003) Proteomics 3:363
84. Whitfield C (2006) Annu Rev Biochem 75:39
85. Leloir LF (1964) Biochem J 91:1
86. Waechter C, Lennarz WJ (1976) Annu Rev Biochem 45:95
87. Castro O, Movsichoff F, Parodi AJ (2006) Proc Natl Acad Sci USA 103:14756
88. Krag SS (1998) Biochem Biophys Res Commun 243:1
89. Kornfeld R, Kornfeld S (1985) Annu Rev Biochem 54:631
90. Hubbard SC, Ivatt RJ (1981) Annu Rev Biochem 50:555
91. Bugg TDH, Brandish PE (1994) FEMS Microbiol Lett 119:255
92. Wacker M, Feldman MF, Callewaert N, Kowarik M, Clarke BR, Pohl NL, Hernandez M, Vines ED, Valvano MA, Whitfield C, Aebi M (2006) Proc Natl Acad Sci USA 103:7088
93. Stamford NPJ (2001) In: Fraser-Reid BO, Tatsuta K, Thiem J (eds) Glycoscience - Chemistry and Chemical Biology vol II, 1st edn. Springer, Berlin Heidelberg New York, p 1215
94. Tonetti M, Sturla L, Bisso A, Zanardi D, Benatti U, De Flora A (1998) Biochimie 80:923
95. Milewski S, Gabriel I, Olchowy J (2006) Yeast 23:1
96. Ishiyama N, Creuzenet C, Lam JS, Berghuis AM (2004) J Biol Chem 279:22635
97. Angata T, Varki A (2002) Chem Rev 102:439
98. Frey P (1996) FASEB J 10:461
99. Oka T, Jigami Y (2006) FEBS J 273:2645
100. Reiter W-D, Vanzin GF (2001) Plant Mol Biol 47:95

101. Kallberg Y, Oppermann U, Jörnvall H, Persson B (2002) Eur J Biochem 269:4409
102. Geelen D, Leyman B, Mergaert P, Klarskov K, Van Montagu M, Geremia R, Holsters M (1995) Mol Microbiol 17:387
103. Negishi M, Pedersen LG, Petrotchenko E, Shevtsov S, Gorokhov A, Kakuta Y, Pedersen LC (2001) Arch Biochem Biophys 390:149
104. Ishihara H, Massaro DJ, Heath EC (1968) J Biol Chem 243:1103
105. Aronson NN, Docherty PA (1982) J Biol Chem 258:4266
106. Rome LH, Hill DF (1986) Biochem J 235:707
107. Alton G, Hasilik M, Niehues R, Panneerselvam K, Etchison JR, Fana F, Freeze HH (1998) Glycobiology 8:285
108. Niehues R, Hasilik M, Alton G, Korner C, Schiebe-Sukumar M, Koch HG, Zimmer KP, Wu R, Harms E, Reiter K, von Figura K, Freeze HH, Harms HK, Marquardt T (1998) J Clin Invest 101:1414-20
109. Gonzalez-Yanes B, Mandell RB, Girard M, Henry S, Aparicio O, Gritzali M, Brown RD Jr, Erdos GW, West CM (1989) Dev Biol 133:576
110. Hirschberg CB, Snider MD (1987) Annu Rev Biochem 56:63
111. Berninsone PM, Hirschberg CB (2000) Curr Opin Struct Biol 10:542
112. Ishida N, Kawakita M (2004) Pflugers Arch 447:768
113. Yakubenia S, Wild MK (2006) FEBS J 273:4390
114. Thomsen B, Horn P, Panitz F, Bendixen E, Petersen AH, Holm LE, Nielsen VH, Agerholm JS, Arnbjerg J, Bendixen C (2006) Genome Res 16:97
115. Rademacher TW, Parekh RB, Dwek RA (1988) Annu Rev Biochem 57:785
116. Fu D, Chen L, O'Neill RA (1994) Carbohydr Res 261:173
117. Schachter H (1986) Biochem Cell Biol 64:163
118. Varki A (1998) Trends Cell Biol 8:34
119. Faye L, Sturm A, Bollini R, Vitale A, Chrispeels MJ (1986) Eur J Biochem 158:655
120. Baenziger JU (1994) FASEB J 8:1019
121. Herscovics A (1999) Biochim Biophys Acta 1473:96
122. Parekh RB, Tse AG, Dwek RA, Williams AF, Rademacher TW (1987) EMBO J 6:1233
123. Davies GJ, Gloster TM, Henrissat B (2005) Curr Opin Struct Biol 15:637
124. Chou HH, Hayakawa T, Diaz S, Krings M, Indriati E, Leakey M, Paabo S, Satta Y, Takahata N, Varki A (2002) Proc Natl Acad Sci USA 99:11736
125. Helenius J, Ng DT, Marolda CL, Walter P, Valvano MA, Aebi M (2002) Nature 415:447
126. Narasimhan S, Stanley P, Schachter H (1977) J Biol Chem 252:3926
127. Byrd JC, Tarentino AL, Maley F, Atkinson PH, Trimble RB (1982) J Biol Chem 257:14657
128. Nakayama K, Nakanishi-Shindo Y, Tanaka A, Haga-Toda Y, Jigami Y FEBS Lett 412:547
129. Dean N (1999) Biochim Biophys Acta 1426:309
130. Schachter H (2000) Glycoconj J 17:465
131. Akama TO, Nakagawa H, Wong NK, Sutton-Smith M, Dell A, Morris HR, Nakayama J, Nishimura S, Pai A, Moremen KW, Marth JD, Fukuda MN (2006) Proc Natl Acad Sci USA 103:8983
132. Kudo M, Bao M, D'Souza A, Ying F, Pan H, Roe BA, Canfield WM (2005) J Biol Chem 280:36141
133. Paschinger K, Rendić D, Lochnit G, Jantsch V, Wilson IBH (2004) 279:49588
134. Hanneman AJ, Rosa JC, Ashline D, Reinhold VN (2006) Glycobiology 16:874
135. North SJ, Koles K, Hembd C, Morris HR, Dell A, Panin VM, Haslam SM (2006) Glycoconj J 23:345
136. Paschinger K, Staudacher E, Stemmer U, Fabini G, Wilson IBH (2005) Glycobiology 15:463
137. Zhang W, Cao P, Chen S, Spence AM, Zhu S, Staudacher E, Schachter H (2003) Biochem J 372:53
138. Léonard R, Rendić D, Rabouille C, Wilson IBH, Préat T, Altmann F (2006) J Biol Chem 281:4867
139. Johnson KD, Chrispeels MJ (1987) Plant Physiol 84:1301
140. Strasser R, Stadlmann J, Svoboda B, Altmann F, Glössl J, Mach L (2005) Biochem J 387:385
141. Vitale A, Chrispeels MJ (1984) J Cell Biol 99:133
142. Ten Hagen KG, Fritz TA, Tabak LA (2003) Glycobiology 13:1R
143. Luo Y, Nita-Lazar A, Haltiwanger RS (2006) J Biol Chem 281:9385
144. Voglmeir J, Voglauer R, Wilson IBH (2007) J Biol Chem 282:5984
145. Manya H, Chiba A, Yoshida A, Wang X, Chiba Y, Jigami Y, Margolis RU, Endo T (2004) Proc Natl Acad Sci USA 101:500
146. Girrbach V, Strahl S (2003) J Biol Chem 278:12554

147. Kolter T, Proia RL, Sandhoff K (2002) J Biol Chem 277:25859
148. Sandhoff K, Kolter T (2003) Philos Trans R Soc Lond B Biol Sci 358:847
149. Ferguson MAJ, Brimacombe JS, Brown JR, Crossman A, Dix A, Field RA, Güther ML, Milne KG, Sharma DK, Smith TK (1999) Biochim Biophys Acta 1455:327
150. Babad H, Hassid WZ (1966) J Biol Chem 241:2672
151. Hagopian A, Eylar EH (1968) Arch Biophys Biochem 128:422
152. Shaper NL, Shaper JH, Meuth JL, Fox JL, Chang H, Kirsch IR, Hollis GF (1986) Proc Natl Acad Sci USA 83:1573
153. Narimatsu H, Sinha S, Brew K, Okayama H, Qasba PK (1986) Proc Natl Acad Sci USA 83:4720
154. Lowe JB, Marth JD (2003) Annu Rev Biochem 72:643
155. Nishihara S, Ueda R, Goto S, Toyoda H, Ishida H, Nakamura M (2004) Glycoconj J 21:63
156. Jolly RD, Walkley SU (1997) Vet Pathol 34:527
157. Winchester B (2005) Glycobiology 15:1R
158. Cacan R, Duvet S, Kmiecik D, Labiau O, Mir AM, Verbert A (1998) Biochimie 80:59
159. Suzuki T, Lennarz WJ (2003) Biochem Biophys Res Commun 302:1
160. Li YT, Li SC (1999) Anal Biochem 273:1
161. Fuller M, Chau A, Nowak RC, Hopwood JJ, Meikle PJ (2006) Glycobiology 16:318
162. Altmann F, Paschinger K, Dalik T, Vorauer K (1998) Eur J Biochem 252:118
163. Li YT (1967) J Biol Chem 242:5474
164. Li SC, Li YT (1970) J Biol Chem 245:5153
165. Li SC, Mazzotta MY, Chien SF, Li YT (1975) J Biol Chem 250:6786
166. Gubareva LV, Kaiser L, Hayden FG (2000) Lancet 355:827
167. Marín-Rodríguez MC, Orchard J, Seymour GB (2002) J Exp Bot 53:2115
168. Prabhakar V, Raman R, Capila I, Bosques CJ, Pojasek K, Sasisekharan R (2005) Biochem J 390:395
169. Shaya D, Tocilj A, Li Y, Myette J, Venkataraman G, Sasisekharan R, Cygler M (2006) J Biol Chem 281:15525
170. Smith AM, Zeeman SC, Thorneycroft D, Smith SM (2003) J Exp Bot 54:577
171. Baker DJ, Timmons JA, Greenhaff PL (2005) Diabetes 54:2453

11.2 Molecular Basis for the Biosynthesis of Oligo- and Polysaccharides

Iain B. H. Wilson[1], *Christelle Breton*[2], *Anne Imberty*[2], *Igor Tvaroška*[3]
[1] Department für Chemie, Universität für Bodenkultur (University of Natural Resources and Applied Life Sciences), Muthgasse 18, 1190 Wien, Austria
[2] Molecular Glycobiology, CERMAV-CNRS (affiliated with Université Joseph Fourier), Grenoble 38041, France
[3] Institute of Chemistry, Centre for Glycomics, Slovak Academy of Sciences, 845 38 Bratislava, Slovak Republic
iain.wilson@boku.ac.at, christelle.breton@cermav.cnrs.fr, anne.imberty@cermav.cnrs.fr, chemitsa@savba.sk

In: *Glycoscience*. Fraser-Reid B, Tatsuta K, Thiem J (eds)
Chapter-DOI 10-1007/978-3-540-30429-6_59:

Abstract

The oligo- and polysaccharide structures observed in samples of biological origin represent the products of a large number of enzymes: the workhorses of the biosynthetic pathways involved are the glycosyltransferases, although some steps necessary for the final glycoconjugate structures result from glycosidase or transglycosylase activities. Glycosyltransferases are, as the name suggests, enzymes which transfer sugars from an activated donor (normally a nucleotide sugar or lipid-linked phosphosugar) to an acceptor, which can be a polypeptide, a lipid or another saccharide. These enzymes have been studied by a wide variety of approaches—biochemical, genetic, chemical, crystallographic—in order to determine their occurrence, substrate specificity, biological function, mechanism and three-dimensional structure.

Keywords

Glycans; Glycosyltransferases; Crystallography; Mechanism

Abbreviations

alg	asparagine-linked glycosylation (mutant)
ALG	asparagine-linked glycosylation (gene)
Alg	asparagine-linked glycosylation (protein)
ATP	adenosine triphosphate
CAZy	carbohydrate active enzyme (database)
CDG	congenital disorder of glycosylation
CMP	cytidine 5′-monophospho-
Dol-P	dolichol phosphate
EDEM	ER degradation-enhancing α-mannosidase-like protein
EGF	epidermal growth factor
ER	endoplasmic reticulum
ERAD	ER-associated degradation
GDP	guanosine 5′-diphospho-
GlcNAc-T	*N*-acetylglucosaminyltransferase
GPI	glycosylphosphatidylinositol
GT(s)	glycosyltransferase(s)
LBHB	low-barrier hydrogen bond
NDP	nucleoside diphosphate
PAPS	3′-phosphoadenosine 5′-phosphosulfate
QM/MM	quantum mechanics/molecular mechanics
RNAi	RNA interference
siRNA	small interfering RNA
TSR	thrombospondin type 1 repeat
UDP	uridine 5′-diphospho-

1 Introduction

Glycosylation is quantitatively the most significant biochemical reaction on earth and the range of oligo- and polysaccharides observed in nature is the result of a large number of biochemical reactions catalyzed by a multitude of enzymes. The key players in the biosynthesis of these structures are the glycosyltransferases, although 'processing' glycosidases are also important in the *N*-glycosylation pathway. It is estimated that, normally, 1–3% of the genes of any organism are involved in synthesis and breakdown of glycoconjugates [1]; this figure includes glycosyltransferases and glycosidases, but excludes necessary 'accessory proteins' such as nucleotide sugar transporters or nucleotide sugar synthases and reductases. This indicates that organisms make a significant genomic investment in glycoconjugate biosynthesis, which in turn suggests that glycoconjugates have important biological roles.

As the name may suggest, glycosyltransferases are the major synthetic enzymes in glycoconjugate metabolism; they can be defined as enzymes which transfer glycomoieties from a donor substrate to an acceptor molecule. As mentioned in ❯ Chap. 11.1, the synthesis of glycosidic bonds requires a substantial input of energy; in biological systems, this means that phosphodiester bonds are cleaved. In the case of glycosyltransferases, the donor molecules are glycosides of either nucleotides or lipid phosphates (❯ *Eq. 1*):

$$XMP - Y + HO - R \rightarrow XMP + Y - O - R \tag{1}$$

In this generalised reaction, X can be a nucleotide (such as UMP, GMP), nucleoside (such as C) or polyprenol (such as dolichol), Y is the sugar linked through a phosphodiester bond destined to be transferred to the acceptor, and R is the lipid, protein or saccharide acceptor. Examples of this type of reaction include:

$$UDP - Xyl + Peptide \rightarrow UDP + Xyl - O - Peptide \tag{2}$$

which is the reaction catalyzed by peptide *O*-xylosyltransferase (EC 2.4.2.26), as the first step of chondroitin or heparan sulfate biosynthesis, and:

$$Dol - P - Man + Man_5GlcNAc_2 - PP - Dol \rightarrow Dol - P + Man_6GlcNAc_2 - PP - Dol \tag{3}$$

which is the reaction catalyzed by the sixth mannosyltransferase (Alg3p) during the biosynthesis of the lipid-linked precursor of *N*-linked oligosaccharides in most eukaryotes.

Other than the distinctions in terms of the types of donors and acceptors, glycosyltransferases fall into two basic categories: the inverting and retaining enzymes. In the case of an inverting enzyme, a monosaccharide α-linked to its donor is β-linked in the final product (or vice versa), whereas a retaining enzyme involves no change in the anomericity of the transferred sugar (see ❯ *Sect. 8*). Furthermore, glycosyltransferases fall into a number of categories based on primary sequence—the Carbohydrate-Active Enzyme Database [2] currently contains 86 families of glycosyltransferases (see ❯ *Sect. 7.1*). Biochemically, the Enzyme Commission has numbered many glycosylation reactions (with numbers of the form EC 2.4.-), but this list is highly incomplete and sometimes inaccurate; on the other hand, many proteins predicted to be glycosyltransferases on the basis of sequence homology are not yet proven to catalyse a specific

reaction. Certainly, the progress in cloning glycosyltransferase genes, primarily as a result of the many genome projects, has outpaced the ability to determine their biochemical and biological significance, never mind that the 3D-structural information on glycosyltransferases is based on a mere handful of crystal structures.

This chapter will concentrate on the biological, biochemical, molecular modelling and crystallographic approaches to the study of the glycosyltransferases, particularly those involved in the biosynthesis of the major classes of eukaryotic glycoconjugates: *N*-glycans, *O*-glycans, glycolipids (including glycosylphosphatidylinositols) and glycosaminoglycans. The biosynthesis of plant polysaccharides was well summarized in the previous edition of this book [3] and will only be briefly considered here, whereas the basics of biosynthesis of *N*-glycans in mammals has been well-reviewed over the years [4,5]; indeed, this chapter is not intended to be exhaustive, but, building upon ❯ Chap. 11.1, will concentrate on the aspects related to the most important modern methods in the field. For brief accounts of the major mammalian glycosyltransferases, the reader is referred to the relatively recent *Handbook of Glycosyltransferases and Related Genes* [6].

2 *N*-Glycan Biosynthesis in Eukaryotes

Probably the best-studied glycans are those linked to asparagine residues in eukaryotes (*N*-linked glycans), especially those of the yeast *Saccharomyces cerevisiae* and of mammals. As summarized in ❯ Chap. 11.1, the initial steps in *N*-glycan biosynthesis are those catalyzed by a series of glycosyltransferases (see ❯ *Fig. 1*) in the endoplasmic reticulum (ER) that lead to the synthesis of the tetradecasaccharide $Glc_3Man_9GlcNAc_2$ linked to dolichol through a pyrophosphate 'bridge' [4,5]; this type of glycosylation intermediate was first reported by Luis Leloir and his colleagues [7]. In some eukaryotes (seemingly mainly protozoan parasites), smaller dolichol-linked oligosaccharides are produced and efficiently transferred to protein: this is associated with the apparent loss of a number of the relevant *ALG* (asparagine-linked glycosylation) genes [8]. These genes have been well-studied in *S. cerevisiae* and are conserved between yeast and mammals; indeed the functional and sequence conservation is sufficient to allow mammalian *ALG* genes to rescue relevant *alg* mutants in yeast. This ability to transform mammalian genes into yeast has been valuable in the identification of mannosyl- and glucosyltransferases which are mutated in various Congenital Disorders of Glycosylation (Type I CDGs) [5,9]. The most recent advances in determining the molecular basis for the various reactions leading to the formation of $Glc_3Man_9GlcNAc_2$, thus leading to the identification of all the necessary glycosyltransferases, have been the discovery of the Alg13p/Alg14p complex required for the transfer of the second GlcNAc residue [10,11] and the demonstration of the dual activities of recombinant *Saccharomyces* Alg2p and Alg11p [12]. The formation of the lipid-linked oligosaccharide can be blocked by the action of tunicamycin, a transition-state analogue for the first reaction (the formation of GlcNAc-PP-Dol) catalyzed by Alg7p and this reagent can be used to inhibit *N*-glycosylation in cell culture [13]. Embryos with a homozygous deletion in the murine *ALG7* homologue die 4–5 days after fertilisation [14], which is an indication of the importance of *N*-glycosylation.

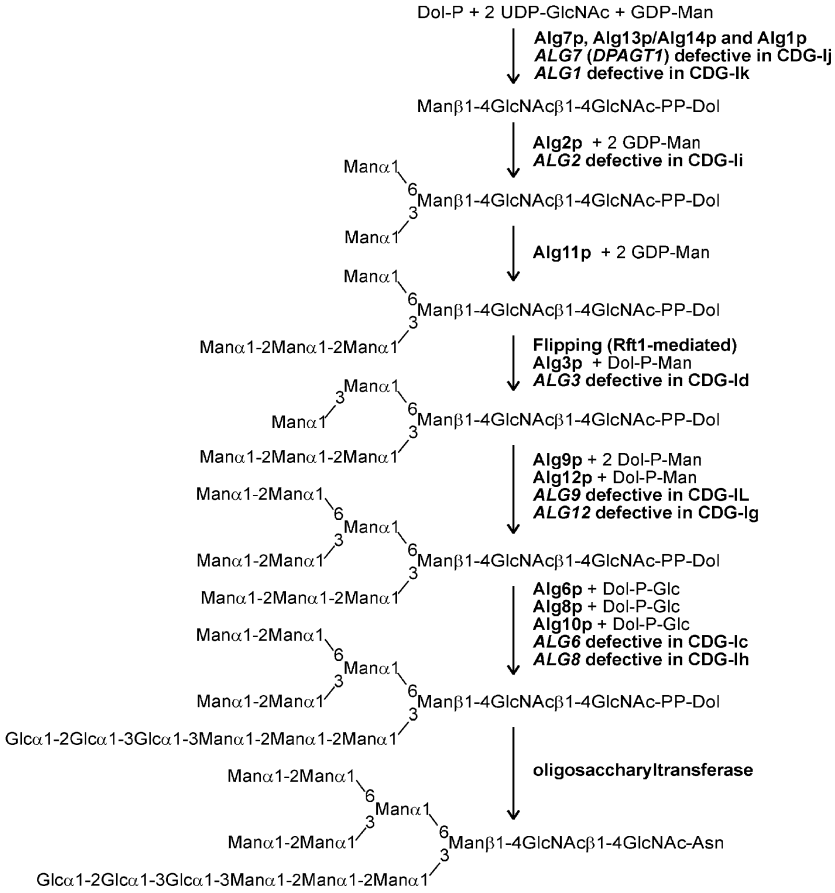


Figure 1

The formation of the dolichol-linked tetradecasaccharide in most eukaryotes. The series of GlcNAc-, Man- and Glc-transfer events (a total of 14 reactions catalyzed by 11 glycosyltransferases, each being in yeast an AlgXp protein) occur in the endoplasmic reticulum. Many of the relevant *ALG* genes are mutated in a corresponding human Type 1 Congenital Disorder of Glycosylation (CDG-I)

2.1 Post-Transfer Processing of Oligomannosidic *N*-Glycans

The completion of the lipid-linked $Glc_3Man_9GlcNAc_2$ is followed by transfer of this structure to nascent proteins (i. e., co-translationally), which is mediated by the multi-subunit complex known as oligosaccharyltransferase [15]. Subsequently, the $Glc_3Man_9GlcNAc_2$ structure is subject to a variety of trimming reactions (see *Fig. 2*), catalyzed by glycosidases such as glucosidases I and II, endoplasmic reticulum mannosidase I and endomannosidase [16,17,18]. The action of both glucosidases I and II can be inhibited by castanospermine or forms of deoxynojirimycin, whereas ER mannosidase I activity is blocked by deoxymannojirimycin [19]. This class I mannosidase (CAZy glycohydrolase family 47) is highly conserved

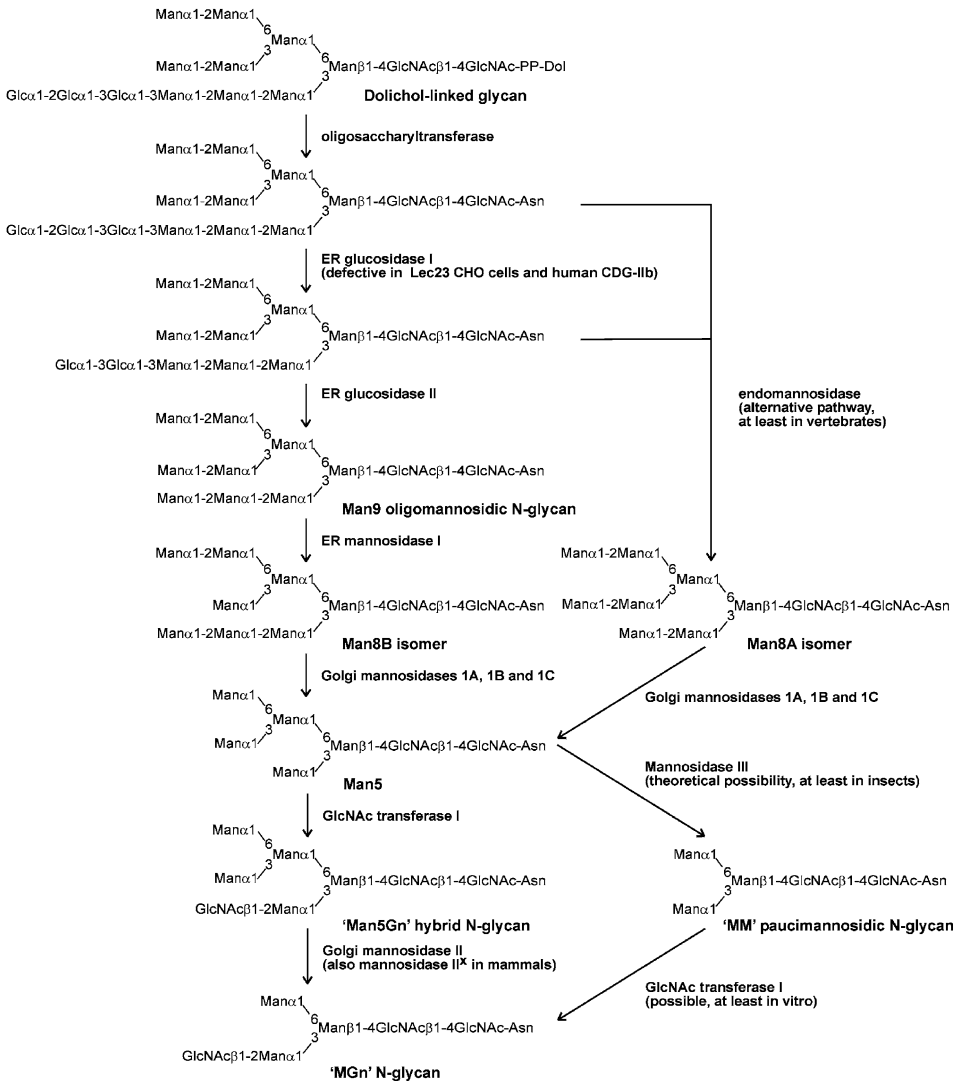


Figure 2

Post-transfer processing of oligomannosidic *N*-glycans. In general, eukaryotes transfer $Glc_3Man_9GlcNAc_2$ to proteins and this is the target of ER glucosidases I and II as well as of ER mannosidase. In yeast, the pathway is then continued by Golgi mannosyltransferases (not shown), whereas in multicellular eukaryotes, the classical pathway proceeds via the action of Golgi class I mannosidases, GlcNAc-TI and Golgi mannosidase II. However, alternative pathways exist or are hypothesised: the Golgi endomannosidase is known to operate in vertebrates and a route from Man5 to MGn via mannosidase III may exist in insects and potentially at a low level in nematodes, but, as judged by the most recent work with knockout organisms, is most probably not occurring in mammals and plants

in all eukaryotes and is the last trimming enzyme found in yeast; it catalyses the removal of a single mannose residue resulting in the formation of one isomer of $Man_8GlcNAc_2$ [20,21], whereas endomannosidase, found in vertebrates and molluscs, is a Golgi enzyme whose activity removes $Glc_{1-3}Man$ from *N*-glycans to form another isomer of $Man_8GlcNAc_2$ [18,22].

The trimming of the *N*-glycans in the endoplasmic reticulum is closely associated with the process of protein folding [23,24]: certainly, it is remarkable that most secreted proteins are glycosylated and this led, over 40 years ago, to Eylar proposing that glycosylation was a form of 'chemical passport' allowing proteins to exit cells [25]. This concept can now be re-considered in the light of the role of the lectins calnexin and calreticulin which recognise the glycans of unfolded proteins [23,24]; these may be reglucosylated in many eukaryotes by a UDP-Glc:glycoprotein glucosyltransferase in order that the proteins can re-enter the folding process [26]. If this fails, then the protein is recognised by lectins with homology to class I mannosidases (EDEMs) which direct the proteins towards ER-associated degradation (ERAD) [27,28], a process involving deglycosylation by a peptide:*N*-glycosidase and proteasome-mediated degradation of the polypeptide [29]. The free *N*-glycan is then, at least in vertebrates, the target of a cytosolic endo-β-*N*-acetylglucosaminidase [30] and a cytosolic mannosidase encoded by the *MAN2C1* gene [31].

If a glycoprotein, however, 'survives' the endoplasmic reticulum, it then enters the Golgi. In mammals, depending on the protein and cell-type, some further mannosidase-mediated trimming, specifically by class I mannosidases, takes place [32]. Many species have multiple Golgi mannosidase I isoforms; in mammals, mannosidases IA, IB and IC have subtly different activities, but can digest a $Man_9GlcNAc_2$ structure to $Man_5GlcNAc_2$ [33]; they display homology to the ER mannosidase I and indeed one amino acid substitution can alter the specificity of the ER enzyme so that it can also digest down to the $Man_5GlcNAc_2$ structure [34]. Mannosidase IB is essential for neonatal survival in mice, as shown by a recently published knock-out study [35].

Similar Golgi mannosidase-mediated processing probably takes place in all multicellular organisms; however, as mentioned above, in yeast only one mannose residue is removed in the endoplasmic reticulum and the maturation of *N*-glycans in the Golgi is, primarily, the result of addition of extra mannose residues [36]. Other 'destinies' occur in other species: for instance, the major *N*-glycan of the slime mould *Dictyostelium discoideum* is a $Man_8GlcNAc_2$ structure decorated with extra *N*-acetylglucosamine residues with or without a core fucose [37].

In mammals, if a glycan is not, or only partially, processed by mannosidases, then it remains oligomannosidic and is generally not subject to any further modification, unless the glycan is attached to an enzyme destined for the lysosome. Such glycans become phosphorylated in a two-stage process catalyzed by a GlcNAc-1-phosphotransferase and a specific 'uncovering' *N*-acetylglucosaminidase [38]; the result is the mannose-6-phosphate moiety recognised by receptors which enable correct targeting of lysosomal hydrolases. Indeed, defects in the two genes encoding the α/β and γ subunits of the GlcNAc-1-phosphotransferase are respectively associated with the lysosomal storage diseases mucolipidoses II and III (otherwise known as I-cell disease and pseudo-Hurler syndrome); the defective targeting resulting from these mutations causes a build-up of macromolecules in the lysosomes and hence to various pathogenic symptoms [39,40].

2.2 The Formation of Hybrid and Branched *N*-Glycans

On the other hand, if the glycan attains a $Man_5GlcNAc_2$ structure, then a major decision point in the 'life' of the glycan has been reached: it is then potentially the target of *N*-acetylglucosaminyltransferase I (GlcNAc-TI or GnTI; EC 2.4.1.101; see *Fig. 2* and *Fig. 3*), which is a resident of the *medial*-Golgi [41]. This enzyme appears to be a 'hallmark' of multicellular organisms and necessary for the formation of so-called 'complex', 'hybrid' and 'paucimannosidic' *N*-glycans (see Chap. 11.1 for an introduction). Certainly, the relevant *MGAT1* genes can be found in plant and animal species. GlcNAc-TI is essential for viability in mammals [42,43], although in plants and nematodes this enzyme is less important [44,45], but still probably required for 'normal' existence. For example, recent results indicate the knocking-out of GlcNAc-TI isoforms in the model nematode *Caenorhabditis elegans* results in differences in the pathogenic effects of bacteria on the worms [46]. In the case of the fruit fly *Drosophila melanogaster*, the effect of knocking-out GlcNAc-TI is a malformation of the fly brain and an unwillingness or incapability of male flies to reproduce, even though the sperm count is normal [47]. These results illustrate the importance of modern genetic methods in the determination of the function of glycans and glycosyltransferases, but also show the problems in extrapolating results from 'lower' to 'higher' organisms. They further indicate that complex *N*-glycans have acquired more and more important functions during evolution. On the other hand, the fact that cell-cell interactions are irrelevant to the survival of mammalian cells in culture may explain why they do not require GlcNAc-TI, as evidenced by the existence of, e. g., Lec1 (lectin resistant) Chinese hamster ovary cells that are null for GlcNAc-TI [48].

Analogously, knock-outs of the next enzyme in the series, Golgi mannosidase II (EC 3.2.1.114) show a similar 'problem' in inter-species comparisons: mutants of this enzyme in plants (*Arabidopsis*) and nematodes (*Caenorhabditis*) grow normally under standard laboratory conditions, even though significant changes in the *N*-glycome occur, which result in the appearance of non-wild-type hybrid structures [49,50]. However, in mammals there are two Golgi mannosidase II genes: one encoding the 'classical' mannosidase II [51] and one encoding mannosidase II^x [52]. The mouse mannosidase II-null mutant displayed an autoimmune phenotype and tissue-dependent loss of complex *N*-glycans, compatible with the known 'gatekeeper' role of mannosidase II [53,54]. However, in some tissues, the presence of complex glycans led to the suggestion that another mannosidase ('mannosidase III'; see *Fig. 2*) digested $Man_5GlcNAc_2$ directly to $Man_3GlcNAc_2$ without the need for the prior action of GlcNAc-TI; mammalian forms of GlcNAc-TI can accept $Man_3GlcNAc_2$, thus, theoretically (following the action of mannosidase III) facilitating the subsequent action of GlcNAc-TII (see *Fig. 3*) and, so, the formation of complex *N*-glycans [53]. Such an idea was built upon previous data suggesting that a Co^{2+}-stabilised mannosidase activity was present in the Golgi [55]; a recombinant enzyme with similar properties and a Golgi localisation has been studied in insects [56]. However, the results with the double mannosidase II/II^x knock-out mouse appear to have 'buried' the notion that a mannosidase III has any (significant) role in mammalian *N*-glycan biosynthesis or that mannosidase II^x is equivalent to mannosidase III: only hybrid *N*-glycans were found in the double mutant embryos [57]. Homozygous double knock-out offspring die either in utero or neonatally [58], whereas the single mannosidase II^x knock-out has an obvious defect 'only' in spermatogenesis [59]. It is interesting that hybrid *N*-glycans are relatively rare in mammals, but are plentiful on, e. g., chicken egg glycoproteins such as ovalbumin [60], suggesting either

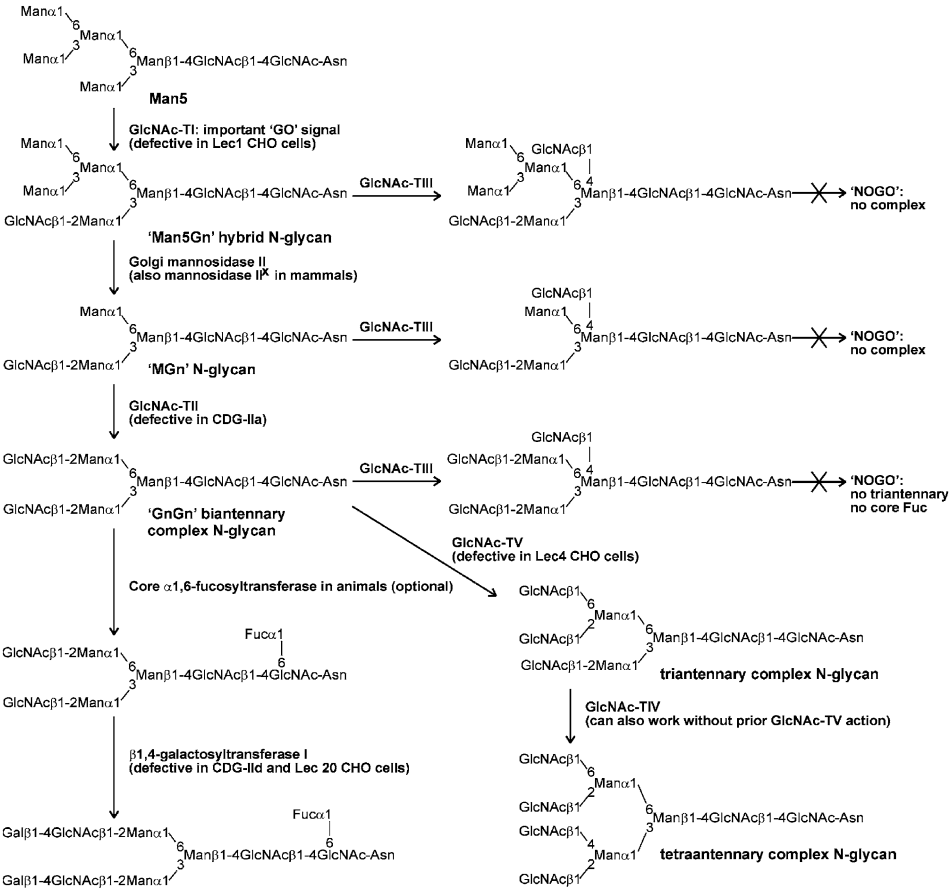


Figure 3

Processing of hybrid and complex branched or bisected *N*-glycans in vertebrates. The $Man_5GlcNAc_2$ structure which results from the action of Golgi class I mannosidases is the target of GlcNAc-TI; the action of this enzyme is a 'GO' signal for mannosidase II in multicellular eukaryotes and also for GlcNAc-TIII in vertebrates. If GlcNAc-TIII has modified an *N*-glycan, this is a 'NOGO' signal for mannosidase II, GlcNAc-TII, GlcNAc-TIV and GlcNAc-TV. In vertebrates chain elongation is generally initiated by β1,4-galactosyltransferase(s); such galactosylation, in turn, prevents or reduces subsequent modification by GlcNAc transferases. For simplicity not all possible core modifications, branching events and 'NOGO' signals are shown

a lack of mannosidase II in the chicken oviduct or that the glycan is poorly accessible to this enzyme or that high levels of GlcNAc-TIII (which transfers the 'bisecting' GlcNAc, thereby preventing mannosidase II action; see *Fig. 3*) or β1,4-galactosyltransferase compete for the same substrate [61].

Based on its primary sequence, mannosidase II is a member of the class II mannosidase family (CAZy glycohydrolase family 38, which includes also the lysosomal mannosidase involved in glycoprotein catabolism). Like the lysosomal enzyme, mannosidase II is inhibited by swainsonine [62], thus this reagent can be used to block the formation of complex *N*-glycans. The

structure of the *Drosophila* mannosidase II has been studied using enzyme crystallised in the presence and absence of this and other inhibitors [63,64]; the enzyme has also been shown to be a resident of the *medial*-Golgi [51,65] and it may form a complex with GlcNAc-TI [66].

Once mannosidase II has acted, the resulting $GlcNAcMan_3GlcNAc_2$ structure (sometimes known as 'MGn') is the substrate for GlcNAc-TII (EC 2.4.1.143; see ❯ *Fig. 3*). This is the key enzyme enabling cells to produce complex *N*-glycans and a mutation in the relevant *MGAT2* gene is associated with a Congenital Disorder of Glycosylation (CDG-IIa, formerly known as carbohydrate-deficient glycoprotein syndrome II) [67,68]. GlcNAc-TII can only act once GlcNAc-TI and mannosidase II have performed their task. Thus, as described by Harry Schachter [69], the action of GlcNAc-TI and mannosidase II constitute so-called 'GO' signals for this enzyme; in turn GlcNAc-TII is a 'GO' signal for the next 'branching' enzyme GlcNAc-TV (EC 2.4.1.155; encoded by the *MGAT5* gene). On the other hand, the prior action of GlcNAc-TIII (EC 2.4.1.144; encoded by the *MGAT3* gene) is a 'NOGO' signal, since the transfer of a bisecting GlcNAc to the core β1,4-linked mannose residue of *N*-glycans prevents the subsequent action of both GlcNAc-TII and GlcNAc-TV. Both these enzymes are also present in *C. elegans* [70,71], whereas GlcNAc-TII, but not GlcNAc-TIII and GlcNAc-TV, is found in plants and insects [72,73]. Novel 'intersecting' and 'bisecting' GlcNAc transferase activities, which take oligomannosidic oligosaccharides as acceptors, have been detected in the slime mould *Dictyostelium* [74]; the relevant genes may be amongst the various β1,4-GlcNAc transferase homologues in this organism. Two other 'branching' enzyme activities are also known: GlcNAc-TIV (EC 2.4.1.145) and the homologous GlcNAc-TVI, which are both β1,4-GlcNAc transferases acting respectively on the α1,3- and α1,6-antennae. The former apparently occurs in two isoforms in mammals [75], whereas the latter is found in birds [76] and its penta-antennary products are found particularly on proteins of chicken egg white [77]. Although potential products of GlcNAc-TIV are found in *Drosophila* [78], any enzymatic activity of the relevant homologues has not yet been described. On the other hand, as judged by glycan analyses, the many potential GlcNAc-TIII homologues in plants do not obviously modify *N*-glycans in vivo.

The role of the bisecting GlcNAc-TIII has attracted attention because of its 'NOGO' potential, particularly as regards the action of GlcNAc-TV [69,79]. The latter is often considered to be relevant to metastasis; certainly, the β1,6-branches initiated by the action of this enzyme are often the starting point for elongation into poly-*N*-acetyllactosaminoglycan chains [80]. Thus, once GlcNAc-TV has acted, there is an increased potential for the formation of very large *N*-glycans, the presence of which has long been considered to be correlated with transformation of cells; an accompanying increase in the average size of glycans was first noted by Warren and Glick [81] and a relationship between GlcNAc-TV levels and metastasis has been observed [82,83]. However, the evidence that that induction of increased GlcNAc-TIII activity (thereby inhibiting the action of GlcNAc-TV) may decrease metastasis has been contradictory, since GlcNAc-TIII is increased in some leukaemias [84]. Ablation of GlcNAc-TIII and GlcNAc-TV shows that neither gene is essential [85,86], but alterations in certain models of metastasis have been observed [84,86,87] and *mgat5* deficient mice display signs of an autoimmune phenotype [88]; interpretation of the apparent non-essential nature of GlcNAc-TV may, though, be complicated in some tissues by the identification of a second human enzyme (known as GnT-Vb or GlcNAc-TIX) with an overlapping in vitro activity [89]. Certainly, though, the 'NOGO' function of GlcNAc-TIII has potential in biotechnologically relevant

re-engineering of glycosylation - particularly since modifications of the core by fucosyltransferases in mammalian cells and plants (see below) are also inhibited by GlcNAc-TIII [90,91]. Two further GlcNAc transferases (VII and VIII) have been postulated to exist in mammals, due to the occurrence of a GlcNAc linked either α/β1,6 to the penultimate Man-linked core GlcNAc residue (LEC18) or β1,2 to the β-linked core mannose (LEC14) in two 'gain-of-function' Chinese hamster ovary cell lines [92].

2.3 Modification of the *N*-Glycan Core Region

Modifications of the core region of *N*-glycans are, in general, widespread: in plants, core α1,3-fucosylation of the reducing-terminal GlcNAc and β1,2-xylosylation of the core β1,4-linked mannose residue occur [93]. These modifications are immunogenic and are key epitopes for antisera generated in rabbits against plant glycoproteins, such as horseradish peroxidase [94]. 'Anti-HRP' is indeed an often-used reagent since it recognises neuronal glycans from invertebrate species [95,96,97]. Thus, this is an indication that xylose and/or α1,3-fucose are present in invertebrates. Interestingly, schistosomal and snail glycoproteins can have both types of modification [98,99], whereas insects and nematodes express core α1,3-fucosylated glycans [100,101]. Such a core α1,3-fucose residue is also present on honeybee venom glycoproteins [102] and constitutes an epitope for the IgE from allergic patients [103], a phenomenon also associated with other allergies [104]. Other modifications of the core involving fucose include the presence of this fucose on the second (or distal) GlcNAc residue of the core, as is the case in the parasite *Haemonchus contortus* [105] and galactosylation of the core fucose in octopus and in *Caenorhabditis* [106,107]; whether these are immunogenic is unknown.

In mammals, the core is not modified with α1,3-linked fucose; in contrast, α1,6-fucosylation is present on many glycoproteins, such as IgG. The relevant enzyme encoded by the *FUT8* gene (EC 2.4.1.68) is dependent on the prior action of GlcNAc-TI (see *Fig. 3*), but will not act if GlcNAc-TIII or β1,4-galactosyltransferase have already modified the *N*-glycan [69]. This type of fucosylation is indeed common throughout the animal kingdom and relevant genes have been cloned from mammals [108,109], *Caenorhabditis* and *Drosophila* [110]; double fucosylation of the core by both α1,3- and α1,6-fucosyltransferases (see *Fig. 4*) is a feature of nematodes and insects [100,101,102]. Knocking-out the α1,6-fucosyltransferase gene in mice has a number of effects, including semi-lethality and dysregulation of growth factor signalling [111,112]. In cell culture a true knock-out of the gene or the mere downregulation of its transcripts by siRNA enables the production of non-fucosylated IgG, which is more efficient in antibody-dependent cellular cytotoxicity (ADCC) assays [113,114].

Subsequent to the action of the various GlcNAc transferases and core fucosyltransferases, further modifications of *N*-glycans occur in a manner dependent on the type of organism. In plants, secreted glycoproteins such as sycamore cell laccase are modified by β1,3-galactosyltransferase(s) and α1,4-fucosyltransferase, thus resulting in formation of Le^a-structures [115]. The relevant α1,4-fucosyltransferase gene has been identified in a number of species [116,117], whereas an *Arabidopsis* β1,3-galactosyltransferase has only recently been identified, amongst twenty potential homologues, as having a role in *N*-glycan modification [118]. However, in plants there exists a different modification pathway for vacuolar glycoproteins; these have one

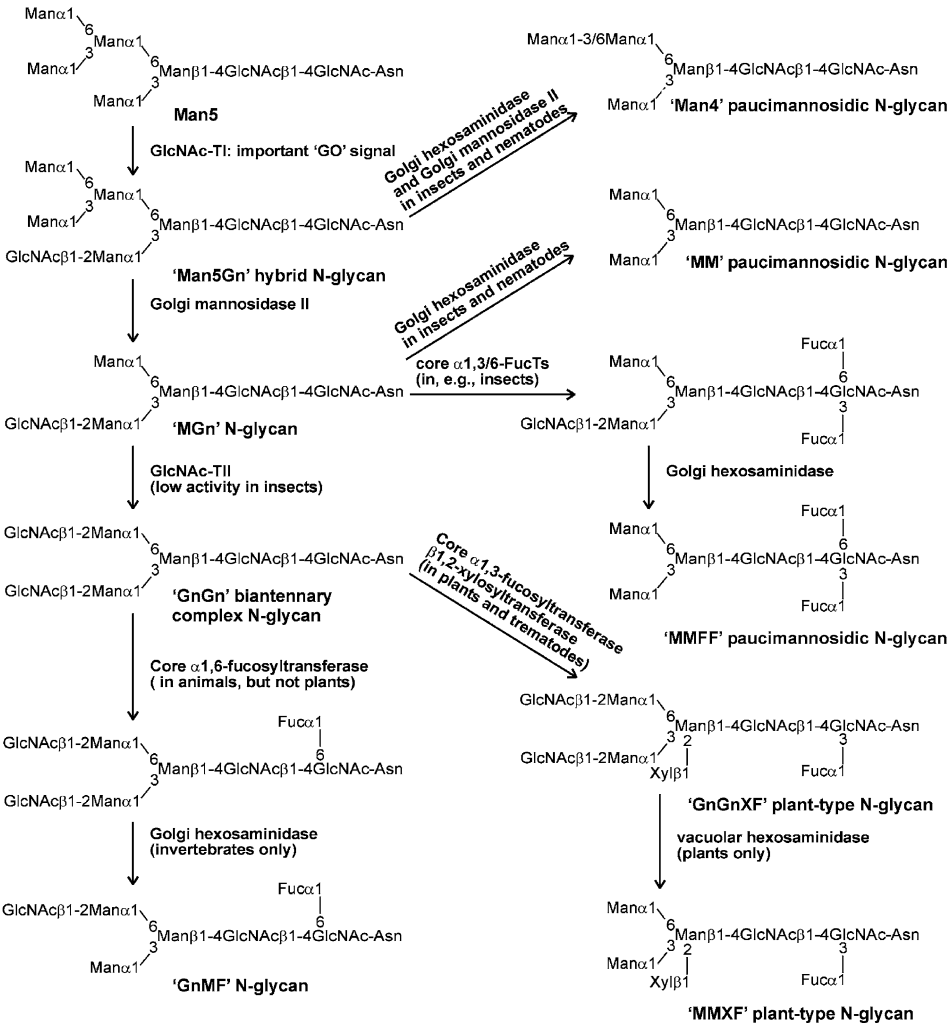


Figure 4

Selected *N*-glycan modification pathways in plants and invertebrates. The initial stages of *N*-glycan maturation in plants and invertebrates are akin to those of vertebrates (*Fig. 3*). However, after the action of GlcNAc-TI in invertebrates, a Golgi hexosaminidase may act which results in the formation of MM and Man4 paucimannosidic glycans. MGn may also be fucosylated (in nematodes only by the α1,6-fucosyltransferase, in insects and trematodes by α1,6- and/or α1,3-fucosyltransferases) prior to the action of this hexosaminidase. It is also possible that this removal of GlcNAc from the α1,3-antenna occurs after the action of GlcNAc-TII. In plants, though, the hexosaminidase acts putatively in the vacuole after the action of core α1,3-fucosyltransferase (EC 2.4.1.214) and β1,2-xylosyltransferase (EC 2.4.2.38); the latter type of enzyme is also present in trematodes and snails. For simplicity, the route in insects and nematodes to solely α1,6-fucosylated ('MMF') glycans, the pathways in nematodes including the GlcNAc-TI-independent core α1,3-fucosylation and the GlcNAc-TI-dependent transfer of phosphorylcholine as well as the addition of Lewis a (Le^a) epitopes in plants and less common antennal modifications in insects are not shown

or both of their non-reducing terminal GlcNAc residues removed by vacuolar hexosaminidases [119], thus resulting in the core-modified *N*-glycans terminating in mannose (see *Fig. 4*) which are familiar from plantproteins such as horseradish peroxidase [120].

A seemingly similar hexosaminidase-mediated modification of *N*-glycans occurs in nematodes and insects; however, there is a difference. The relevant enzyme (see *Fig. 4*) is postulated to occur in the Golgi and removes specifically the GlcNAc from the α1,3-antenna [121,122]; its action thus accounts for the fact that the most abundant glycans in these species are the so-called paucimannosidic structures with a trimannosyl core with and without fucose moieties [78,93,100,101]. Recently, the relevant *Drosophila* hexosaminidase has been identified and shown to be present throughout the secretory pathway [123]; genes encoding other enzymes with a similar specificity have been found in *Caenorhabditis* [124]. The significance of identifying hexosaminidase genes from invertebrates (especially insects) is that the presence of the encoded enzymes prevents elongation of the antennae of the *N*-glycans of these species, thus reducing their ability to produce, in a biotechnological context [125,126], proteins with 'human-like' glycosylation (i. e., with terminal galactose and sialic acid residues). Furthermore, cloning of these genes enables the examination of the biological significance of truncated *N*-glycans in these species and an unravelling of the 'mystery' as to how glycans can carry fucose residues (which generally, but not always, requires the prior action of GlcNAc-TI), yet lack any non-reducing terminal GlcNAc residues. Interestingly, in the fruit fly, the *fdl* hexosaminidase mutant has a 'fused lobes' phenotype [127] similar to that seen in the fly GlcNAc-TI mutant [47], suggesting that correct 'fly-type' glycosylation is required for normal neural development in *Drosophila*. However, only the GlcNAc-TI mutant lacks neural anti-HRP epitopes [47] and this result is compatible with the requirement of the fly core α1,3-fucosyltransferase (FucTA) for a non-reducing terminal GlcNAc on the α1,3-antenna [100,128]. The conclusion is that, in the fly, the *fdl* hexosaminidase acts subsequent to both core fucosyltransferases, thus accounting for the observed fucosylated paucimannosidic *N*-glycans; furthermore, although there are four potential α1,3-fucosyltransferase homologues in the fly, only the use of FucTA double-stranded RNA results in RNAi-mediated diminution of anti-HRP epitopes in a *Drosophila* neural cell line [128], indicating that a single fucosyltransferase gene is responsible for generation of these epitopes in the fly.

In *C. elegans*, the exact mechanisms for core α1,3-fucosylation are a bit different. Unlike the GlcNAc-TI-dependent fly core α1,3-fucosyltransferase (FucTA), the nematode core α1,3-fucosyltransferase (FUT-1) acts only on *N*-glycans that lack the GlcNAc-TI-dependent GlcNAcβ1,2Manα1,3 moiety [101]. It is concluded that, in the worm, the residue transferred by GlcNAc-TI must be removed by a hexosaminidase before FUT-1 can act. The worm core α1,6-fucosyltransferase (FUT-8), however, requires prior action of GlcNAc-TI and, therefore, the synthesis of difucosylated paucimannosidic *N*-glycans involves the sequential action of α1,6-fucosyltransferase, hexosaminidase and α1,3-fucosyltransferase. As expected, worms lacking an intact core α1,3-fucosyltransferase (*fut-1*) gene do not express anti-HRP epitopes; however, these epitopes are still observed in GlcNAc-TI-null worms lacking all three GlcNAc-TI isoforms [101] consistent with the presence of potentially anti-HRP antibody reactive fucosylated $Hex_{4-9}GlcNAc_2$ *N*-glycans in the GlcNAc-TI-null worm [45,46]. If a core fucosyltransferase does not act, the 'MGn' structure (*Fig. 4*) is processed primarily to the trimannosyl core [46]. MGn has, also, various other fates, e. g., incorporation of phosphorylcholine. This GlcNAc-TI-dependent modification is absent from GlcNAc-TI-null worms. In

the context of *N*-glycans, the phosphorylcholine epitope is seemingly nematode-specific and is associated with immunomodulation mediated by certain nematode glycoproteins [129].

2.4 Terminal Decorations of *N*-Glycans

In contrast to the hexosaminidase-mediated modification of *N*-glycans in 'lower' multicellular eukaryotes, the glycosylation pathways of mammals are, after the action of mannosidase II, based on addition, rather than further removal, of glycan moieties. The terminal GlcNAc residues are acceptors for β1,4- and β1,3-galactosyltransferases, although the former modification is the most familiar (see ❯ *Fig. 3*). There are a number of enzymes in mammals with β1,4-galactosyltransferase activity, but the major form is β4Gal-TI [130]. This enzyme is the same as the lactose synthase found in milk and is probably one of the most studied glycosyltransferases; it has a relatively broad substrate specificity and can tolerate even other nucleotide sugar donors, such as UDP-GalNAc [131], although it is most efficient with UDP-Gal. The enzyme displays an 'ordered sequential' mechanism by which UDP-Gal, in the presence of Mn(II) ions, first binds the enzyme; thereafter the acceptor substrate binds, the transition state forms, the saccharide product is released and finally the UDP 'by-product' dissociates [132]. In the Golgi, this UDP is cleaved to UMP [133], which is then returned to the cytoplasm via the UDP-Gal/UMP antiporter [134]. In the presence of α-lactalbumin, β4Gal-TI has a different acceptor specificity and prefers glucose over GlcNAc-terminating substrates [135]. Due to its role in producing the resulting disaccharide (lactose; Galβ1,4GlcNAc), the enzyme is, as a result of activation of a special transcription initiation site, overexpressed in the mammary gland; in most tissues, though, it is expressed constitutively and is considered to be a 'housekeeping' gene [136].

Knocking-out the β4GalT-I gene results in semi-lethality and other defects [137], but it is not absolutely essential, perhaps due to the possible compensatory action of five other β1,4-galactosyltransferases [130]. It is deficient in the human CDG-IId, which is characterised by hydrocephalus, myopathy, and blood-clotting defects [138]. Two major isoforms of β4Gal-TI are found and result from the use of alternative translation start sites: the longer form with an extended cytosolic domain is localised, for instance, in the plasma membrane of spermatozoa, where it may play a role in fertilisation perhaps acting as a lectin [139]; the function of this longer cytoplasmic tail in directing cell surface expression is apparently regulated by phosphorylation [140]. The shorter form is found in the Golgi, but is sometimes cleaved from the membrane and thereby secreted into the milk or into the serum [141,142].

The knowledge that this enzyme, like many Golgi glycosyltransferases, is active in the absence of its *N*-terminal cytoplasmic, transmembrane and stem regions (it is a typical Type II transmembrane protein with an uncleaved signal sequence) has facilitated its preparation as a soluble recombinant enzyme in *E. coli*, yeast (*S. cerevisiae* and *P. pastoris*) and insect systems [143,144,145,146,147,148,149]. β4Gal-TI was the first eukaryotic glycosyltransferase to have its cDNA cloned [150,151] and the first to be crystallised [152]. Its activity in the absence of its natural *N*-terminus does not mean that this region is without function: indeed, this region is required for the proper localisation of the membrane-bound form to the *trans*-Golgi [153].

When the β4Gal-TI gene is knocked-out, there is an increase in the amount of β1,3-galactosylated glycans [154]; in vitro data from two different laboratories implicate β3Gal-TI or

β3Gal-TII in the modification of *N*-glycans [155,156]. Another alternative form of modification of non-reducing terminal GlcNAc residues is the addition of β1,4-linked GalNAc, thus forming the GalNAcβ1,4GlcNAc moiety ('LacdiNAc'). β1,4-GalNAc-T activity has been associated with invertebrates due to a Tyr→Ile substitution found in insect homologues of β4Gal-TI [157]; GalNAc residues are found, for instance, on the *N*-glycans of honeybee venom phospholipase A_2 [102], although β1,4-galactosylation is also known in invertebrates such as the fruit fly *Drosophila* and the nematode *Ascaris suum* [78,158]. In the case of mammals, β1,4-GalNAc residues attached to GlcNAc are also found; in some cases, i. e., those of pituitary glycoprotein hormones, an amino acid sequence within the protein to which the *N*-glycan is attached acts as a signal for the relevant GalNAc transferase [159]. The GalNAc residues on these glycans are then sulfated, which is important for control of hormone levels in the bloodstream [160]. Another modification by β1,4-GalNAc is that observed in the CT antigen, which has the terminal structure GalNAcβ1,4(Siaα2,3)Gal-R also seen in the human Cad/Sd^a epitope (see *Fig. 5*); the transfer of GalNAc by the relevant enzyme to terminal galactose is dependent on prior α2,3-sialylation [161].

Sialylation is indeed a modification very familiar to those who study mammalian glycosylation. However, the levels of sialylation in non-vertebrates are in the range 'low' or 'no'; an absolute 'yes/no' is not possible since proving the absence of sialylation is as difficult as demonstrating conclusively that it is present at low levels. The relevant glycosyltransferases are all members of the sialyltransferase family, which utilise CMP-NeuAc or CMP-NeuGc as donors [162,163]; the latter cannot be made by humans de novo, due to an evolutionarily recent mutation in the relevant CMP-sialic acid hydroxylase gene [164]. The most familiar member of the sialyltransferase family is the α2,6-sialyltransferase, now known as ST6GalI, which modifies terminal galactose residues [165]; recently, a second such enzyme was found [166]. ST6GalI is an enzyme of the late Golgi [167], but can be cleaved by β-secretase, thus resulting in its secretion [168].

α2,3-Sialyltransferases may also transfer to terminal galactose with some preferring *N*-glycan substrates, whereas others modify *O*-glycans or glycolipids [169,170]. In special cases, the terminal sialic acid is then the target of up to six differentially expressed α2,8-sialyltransferases - thus initiating the formation of polysialic acid on *N*-glycans, *O*-glycans or on glycolipids [171,172]. Polysialic acid is particularly associated with embryonic development, during which the 'anti-adhesive' properties of this modification may aid the flexibility required during neurite migration; its expression is reduced in adults. A recent study indicates that induced expression of the relevant enzymes may be a promising approach to neural repair in later life [173]. Targeted ablation of different sialyltransferase genes results in immunological or neural defects [174,175,176]; a complete ablation of sialylation by knocking-out the UDP-GlcNAc 2-epimerase gene, required for CMP-sialic acid biosynthesis, indicates the overall essential role for sialic acid in mammalian development [177], whereas a defect in the Golgi CMP-sialic acid transporter is associated with human CDG-IIf [178].

Other terminal modifications of *N*-glycans include the poly-*N*-acetyllactosaminoglycans, the HNK-1 epitope and fucosylation (for examples, see *Fig. 5*). The first is a modification often associated, as mentioned above, with the branch initiated by GlcNAc-TV. Poly-*N*-acetyllactosaminoglycans (poly-LacNAc) are repeats of Galβ1,4GlcNAcβ1,3, i. e., LacNAc, hence the name. Their synthesis is initiated by a β1,3-GlcNAc transferase known as i-extension enzyme and the new GlcNAc terminus is then the target of 'standard' β1,4-galactosyltransferases; this

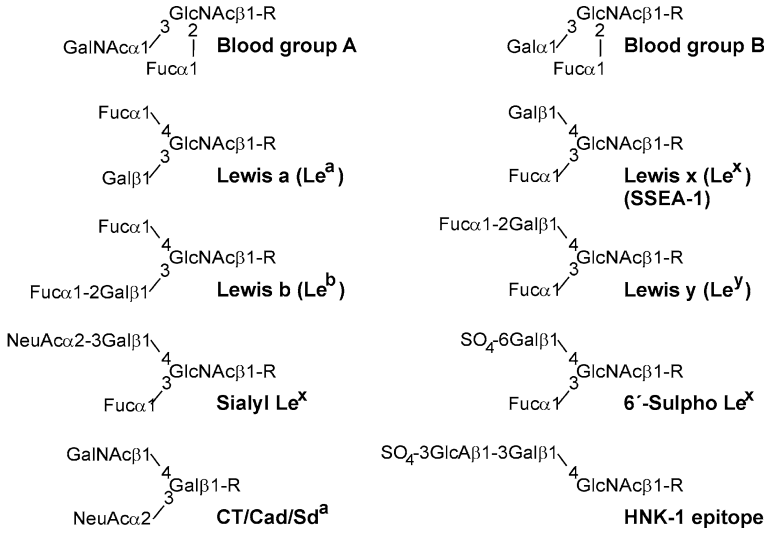


Figure 5
Examples of terminal modifications of mammalian glycans. Shown are the blood group A and B structures, variations of Lewis structures, CT/Cad/Sd^a antigen and the HNK-1 epitope

process can be repeated until the glycan chain is completed [179]. Branching of these chains, initiated by certain β1,6-GlcNAc transferases and thereby forming the so-called I antigen, is also possible [180,181]; the occurrence of poly-*N*-acetyllactosaminoglycans is not just confined to *N*-glycans - they are also present on some *O*-glycans and glycolipids.

Another modification which can be present regardless of the core is the addition of the HNK-1 epitope, SO_4-3GlcAβ1,3Galβ1,4GlcNAc (see *Fig. 5*), which is formed by the action of a glucuronyltransferase and a sulfotransferase [182,183]. Similarly, various blood Groups (Lewis and ABO) are also present on a range of glycoconjugates: the pre-requisite is a Galβ1,3GlcNAc ('type I') or Galβ1,4GlcNAc ('type II') non-reducing terminus. The H- and Se-type α1,2-fucosyltransferases, encoded in man by the *FUT1* and *FUT2* genes present on chromosome 19 [184], can modify these sequences, thus forming the O blood group. The further modification by the AB transferase then follows with the transfer of either an α1,3-linked GalNAc (blood group A; see *Fig. 5*) or a Gal (blood group B) depending on the sequence of the AB transferase gene. Four nucleotide differences alter the substrate specificity of the encoded enzyme and thus determine the blood group of the individual [185]; those with one A-type allele and one B-type have an AB phenotype, whereas individuals with an O blood group have a premature stop codon in both alleles of the ABO gene.

The Lewis-type blood groups have attracted attention, particularly in their sialylated or sulfated forms, due to their role in leukocyte homing, e. g., the selectin-mediated rolling associated with inflammation [186]. In mammals, a number of α1,3/4-fucosyltransferases can catalyse the formation of Lewis epitopes (Le^a, Le^b, Le^x, Le^y; see *Fig. 5*); Le^x is considered equivalent to the stage-specific embryonic antigen 1 (SSEA-1). The human *FUT3* gene encodes an enzyme which can accept Galβ1,3GlcNAc-R or Galβ1,4GlcNAc-R in order to form either

Galβ1,3(Fucα1,4)GlcNAc-R (Le^a) or Galβ1,4(Fucα1,3)GlcNAc-R (Le^x) [187], whereas the products of the *FUT4*, *FUT5*, *FUT6*, *FUT7* and *FUT9* genes are solely or primarily acting as α1,3-fucosyltransferases with subtly different substrate specificities [186,188]. Particularly, Fuc-TVII (i. e., the enzyme encoded by the *FUT7* gene) is involved in synthesis of sulfo- and sialyl-Le^x and is associated with the formation of selectin ligands necessary for normal leukocyte trafficking [189]; a further indication of the role of fucose in such processes is that human leukocyte adhesion deficiency II (or CDG-IIc) is a result of a defect in the GDP-Fuc transporter [190].

Not all mammals possess as many fucosyltransferase genes as are present in the human genome: indeed, *FUT3, FUT5* and *FUT6* are evolutionarily relatively new genes. Cattle only possess one such gene [191]; on the other hand, two mammalian fucosyltransferase homologues (Fuc-TX and Fuc-TXI) have, as yet, no known enzymatic function [192]. Lewis-type fucosylation is also a feature of invertebrates, such as the honeybee [193] and the trematode parasite *Schistosoma* [110], as well as fish [194] and the pathogen *Helicobacter pylori* [195], although, in the honeybee and in trematodes, the Lewis-type structures are all/partly fucosylated LacdiNAc moieties, i. e., GalNAcβ1,4(Fucα1,3)GlcNAc-R. In the case of schistosomes, these structures may play a role in the immune response to the parasite [196].

Another terminal modification, in this case restricted to *N*-glycans, is the addition of α1,3-galactose, otherwise known as the Galili epitope: this occurs in many mammals and also in one nematode parasite which infects deer [197]. This epitope, though, is absent from 'old world' monkeys and apes, including ourselves and other primates [198,199]. The relevant α1,3-galactosyltransferase gene is a pseudogene in humans. This correlates with high levels of naturally occurring 'anti-α-Gal' antibodies and is one barrier (of many) to the use of, e. g., pigs as a source of organs for xenotransplantation; the recent knocking-out of the α1,3-galactosyltransferase gene in pigs is only one step of many necessary in this area [200].

3 *O*-Glycan Biosynthesis in Animals

The term *O*-glycan has often been considered to be synonymous with the presence of a GalNAc linked to a serine or threonine residue. However, this 'mucin-type' of glycosylation is only the tip of an iceberg (see also Chap. 11.1), which includes *O*-linked fucose, *O*-linked mannose and *O*-linked GlcNAc in mammals and a range of other linkages in non-vertebrates and prokaryotes [201]. *O*-linked fucose is associated with 'Notch' signalling during development [202], *O*-linked mannose is present on α-dystroglycan and defects in the *O*-mannosyl-oligosaccharide pathway are associated with forms of congenital muscular dystrophy [203], whereas *O*-linked GlcNAc in the cytoplasm is important for signal transduction and is essential for mammalian cells [204].

3.1 Mucin-Type *O*-Glycans

In animals, the typical 'mucin-type' glycosylation is initiated by the transfer of GalNAc by a number of different enzymes; about twenty polypeptide-modifying GalNAc transferases

(ppGalNAcT) are known with various exact substrate specificities [205]. Some accept 'naked' peptides, others will only act if another residue on the acceptor peptide is already modified by a GalNAc and a final examination of their enzymatic functions still requires large-scale screening of substrates in order to understand how densely glycosylated mucin sequences can arise [206]. This multitude of initiating enzymes may explain, in retrospect, the difficulty in defining a consensus sequence for *O*-glycosylation [207]. In fruit flies, one specific ppGalNAcT encoded by the *pgant35A* gene is essential [208] and, in humans, mutations in the *GALNT3* gene are associated with familial tumoural calcinosis [209]; however, the overlapping tissue expression and substrate specificities of the mammalian enzymes make it difficult to produce 'knock-out' mice with a clear phenotype [210].

The situation is different with the next major modification of peptide-linked GalNAc: the addition of a β1,3-linked galactose to form the so-called 'core 1' structure Galβ1,3GalNAc otherwise known as the T antigen (see ❯ *Fig. 6*). In mammals, there appears to be a single relevant enzyme [211] which has a potential role in IgA nephropathy [212] and, as judged by the results with a knock-out mouse, in angiogenesis [213]. For its expression as an active enzyme, the core 1 galactosyltransferase requires the presence of a specific chaperone [214]. The relevant chaperone gene, *COSMC*, is X-linked in mammals; a deficiency in this gene is associated with a number of diseases, such as the autoimmune disease Tn syndrome [215], so-called since the Tn antigen (unmodified GalNAc-α-Ser/Thr) is present. The corresponding enzyme, however, in insects and nematodes does not require such a chaperone [216,217].

The core 1 structure is the target of a number of enzymes in mammals, such as α2,3-sialyltransferases modifying the galactose residue [218], α2,6-sialyltransferases modifying the GalNAc [219], a 'core 1 extension' β1,3-GlcNAc transferase required for synthesis of the MECA-79 epitope [Galβ1,4(SO_4-6)GlcNAcβ1,3Galβ1,3GalNAc] which forms part of L-selectin ligands [220] and β1,6-GlcNAc transferases also modifying the GalNAc. The latter type of enzyme catalyses the formation of the 'core 2' *O*-glycan, which is one of eight mucin-type 'cores' identified in mammals (❯ *Fig. 6*). The human genome has three genes encoding β1,6-GlcNAc transferases forming the core 2 structure [181,221,222]. This structure can be modified by poly-*N*-acetyllactosaminoglycans, which may be fucosylated to create Lewis groups, such as the *O*-glycans of P-selectin glycoprotein ligand 1 (PSGL1) [223]; knock-out mice lacking one of the core 2 GlcNAc transferase genes have a partial deficiency in selectin ligands and reduced neutrophil recruitment to sites of inflammation [224]. Core 2 structures are also associated with control of the immune system and are a marker of activated T-cells; the abnormal expression of core 2 structures in AIDS and Wiskott–Aldrich syndrome results in T-cells which are not truly active also carrying the modification [225]. In the fruit fly, there are no core 2 structures, but in *Caenorhabditis elegans*, a putative β1,6-glucosyltransferase has been found [226], which is probably involved in the formation of the rather unusual *O*-glycans found in this nematode [227].

An alternative to the core 1 structure is the core 3 structure GlcNAcβ1,3GalNAc, present in a restricted range of tissues. The relevant β1,3-GlcNAc transferase is known as β3Gn-T6 and is a member of the large β1,3-glycosyltransferase family [228]; it apparently suppresses the metastatic potential of carcinoma cells [229]. Core 4 [GlcNAcβ1,6(GlcNAcβ1,3)GalNAc] is formed from core 3 by the action of the so-called 'mucin-type' core 2 synthesising enzyme (C2GnT-M) [181]. Four other mucin core structures are known (cores 5, 6, 7 and 8), but information as to the genetic basis for their formation is currently absent.

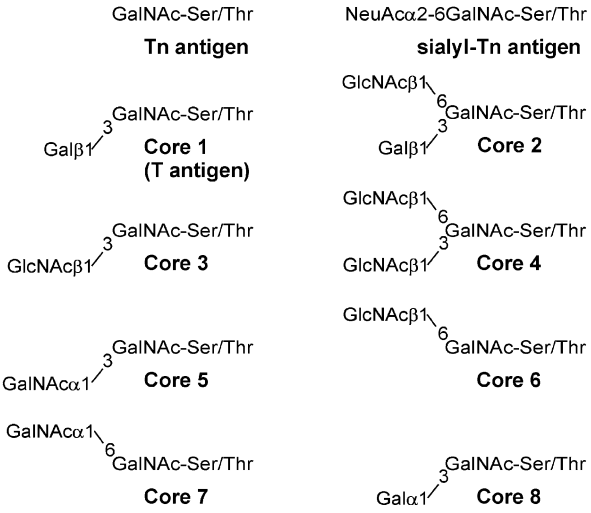


Figure 6
Core structures of mammalian mucin-type *O*-glycans

3.2 *O*-Fucosylation and *O*-Glucosylation

O-fucosylation of polypeptides is catalyzed by one of two E*R*-localised enzymes (the protein-*O*-fucosyltransferases encoded by the *POFUT1* and *POFUT2* genes), which modify, respectively, EGF (epidermal growth factor) or TSR (thrombospondin type 1 repeat) domains [230,231]. These enzymes recognise not just a peptide, but the actual 3D-structure of the relevant domain. Furthermore, the enzymes modifying the peptide-bound fucose residue (in the case of EGF domains Golgi-localised β1,3-GlcNAc transferase(s) and for TSR domains an ER-localised β1,3-glucosyltransferase) also recognise the domain structure [232,233,234,235]. Another modification specific for EGF-domains is *O*-glucosylation, which is associated with proteins such as clotting factors VII and IX and Notch receptor proteins [236]; the protein-bound glucose is, in turn, the acceptor for an α1,3-xylosyltransferase, which was recently purified from bovine liver [237].

Mutations in the *POFUT1* and *O*-linked fucose-modifying β1,3-GlcNAc transferase genes are associated with, for instance, mid-gestation lethality in mice [238] and abnormal wing development in the fruit fly [232,239]; indeed, the name of the β1,3-GlcNAc transferase mutant is *fringe*, reflects the nature of the morphogenetic defect in the fruit fly, which has similarities to the *notch* phenotype ('notched' wings in the fly). This is not surprising considering that *O*-fucosyltransferase I and Fringe modify the Notch protein and that this glycosylation pathway is important for Notch-mediated signalling required for embryonic development in animals [202]. The GlcNAc transferred by the Fringe enzymes is, in mammals, the target of β4Gal-TI [240] and the galactose can then be sialylated. Unlike flies, mammals have three *fringe* homologues encoding β1,3-GlcNAc transferases, specific for fucosylated EGF domains, known as lunatic fringe, manic fringe and radical fringe [233]; a mutation in

the human *lunatic fringe* gene has been found in one patient with spondylocostal dystosis, a disease associated with vertebral defects [241]. In zebrafish, use of antisense morpholino-oligonucleotides to target *lunatic fringe* transcripts indicates a role for the encoded enzyme in hypochord development [242].

3.3 *O*-Mannosylation

Another *O*-glycosylation event that has recently attracted attention is *O*-mannosylation, a modification once considered to be yeast-specific, but which is found also in mammals [203] and, most recently, in *Mycobacteria* [243]. In mammals and in fruit flies, two related proteins, POMT1 and POMT2 are required to be co-expressed in order for protein *O*-mannosylation to occur [244,245]; they require Dol-P-Man as the donor substrate. Defects in man result in a form of muscular dystrophy known as Walker–Warburg syndrome [246], whereas in flies the *twisted (tw)* and *rotated abdomen (rt)* mutations are associated with, as the names suggest, a twisted abdomen phenotype [247]. In yeast, the essential process of protein *O*-mannosylation [248] is followed by addition of further mannose residues in the Golgi [249], whereas in mammals, the *O*-mannose residue is then the acceptor for a GlcNAc residue transferred by protein *O*-mannose *N*-acetylglucosaminyltransferase I (POMGnTI), a homologue of GlcNAc-TI, which is defective in Muscle-Eye-Brain disease [250]. Indeed, the absence of correct *O*-glycosylation of α-dystroglycan is associated with a number of congenital neuromuscular diseases with mutations in glycosyltransferase-like genes (e. g., *LARGE*, *fukutin* and *FKRP*) encoding proteins of unknown activity [5,203,251].

4 Glycolipid Biosynthesis in Animals

In mammals, the term ‘glycolipid’ is often used to mean ganglioside - but these are only a subset of the glycolipids found in eukaryotes. As summarized in Chap. 11.1, gangliosides are a form of glycosphingolipid, which are glycosylated forms of ceramide. The first step in their biosynthesis is the transfer of glucose to ceramide to form glucosylceramide. This reaction takes place on the cytoplasmic face of the endomembrane system [252]; subsequent to flipping of glucosylceramide to the lumenal face of the endomembrane, a number of Golgi-localised glycosyltransferases act to make the range of gangliosides observed, with galactosyltransferases, GalNAc transferases and sialyltransferases being particularly important [253]. For mammalian glycolipids, the first modification of glucosylceramide is mediated by β1,4-galactosyltransferase VI [254]; the resulting lactosylceramide is then modified by either (i) sialylation with GM3 synthase (also known as ST3GalV) [255] and/or addition of GalNAc by GM2/GD2 synthase [256] to form members of the *ganglio*-series, (ii) α1,4-galactosylation yielding the P^k/Gb3/CD77 antigen of the *globo*-series [257], (iii) α1,3-galactosylation by the iGb3 synthase initiating the *isoglobo*-series [258] and (iv) transfer by β1,3GlcNAc-TV in the formation of the *lacto*- and *neolacto*-series [259]. In the case of the galactolipids, including sulfatide and seminolipid, the first step is catalyzed by galactosylceramide synthase [260] on the lumenal side of the endomembrane system.

Ablation of genes of glycolipid biosynthesis is often associated with neurological defects [253]. In humans, a loss-of-function mutation in the aforementioned ST3GalV/GM3 synthase

(encoded by the *SIAT9* gene) has been identified in Amish infantile epilepsy syndrome [261]; however, as an example of the problems of comparing mice and men, the corresponding knockout mouse shows no neurological symptoms, but displays insulin hypersensitivity [262]. As judged by recent tissue-specific knockout mice, the formation of glucosylceramide is also important for the maintenance of the barrier function of the epidermis and for brain maturation [263,264]. On the other hand, the relevant glucosylceramide synthase has attained increased interest due to it being a target for 'substrate reduction therapy' in lysosomal storage diseases. In this approach for treatment of type I Gaucher's disease, the intention is to reduce the pathological build-up of non-catabolised glucosylceramide in the lysosome; this is achieved by use of the glucosylceramide synthase inhibitor *N*-butyldeoxynojirimycin (also an ER glucosidase II inhibitor), which reduces the ability of cells to produce more glycolipids, including gangliosides, that are eventually catabolised to glucosylceramide during the normal physiological turnover of these compounds [265].

In invertebrates, different forms of glycosphingolipids are found; those of nematodes and insects, i. e., *arthro*-series lipids, are probably the best examined. Although the first reaction in their biosynthesis is the glucosylation of ceramide, the next enzymes to act are a β1,4-mannosyltransferase and a β1,3-GlcNAc transferase. Developmental defects in *Drosophila* are associated with defects in these two enzymes and this is reflected in the names of the corresponding mutants - *egghead* and *brainiac* [266,267]. The elongation of glycolipids in insects is not fully understood, but as judged by the structures observed [268,269] includes the action of a β1,4-GalNAc transferase [270], an α1,4-GalNAc transferase [271] and other enzymes such as a phosphoethanolamine transferase which is yet to be identified. Glycolipids often act as receptors for bacterial toxins and the series of *Caenorhabditis bre* (***B****acillus thuringiensis* toxin **re**sistant) mutants, displaying deletions in the glycolipid biosynthesis pathway, are associated with resistance to the Cry5B crystal toxin. The corresponding enzymes BRE-1, BRE-2, BRE-3, BRE-4 and BRE-5 are, respectively, a GDP-mannose dehydratase (required for GDP-Fuc biosynthesis), a β1,3-galactosyltransferase, an Egghead-like β1,4-mannosyltransferase, a β1,4-GalNAc transferase and a Brainiac-like β1,3-GlcNAc transferase [272,273,274] required for biosynthesis of fucosylated *arthro*-series glycolipids in the nematode. Some glycolipids in *C. elegans* and other nematodes (e. g., *Ascaris*), though, also carry phosphorylcholine residues and display immunomodulatory activity [275].

Another type of glycolipid is involved in anchoring some membrane proteins via a glycosylphosphatidylinositol (GPI). These types of membrane anchor were first found attached to the variant surface glycoprotein (VSG) of the African sleeping sickness parasite *Trypanosoma brucei* [276], but GPI-anchored proteins are widespread in other eukaryotes [277]; examples in mammals include various brush border enzymes, Thy-1 and decay accelerating factor (DAF). All GPI anchors consist of core based upon an inositol phospholipid modified by the sequential addition of an *N*-acetylglucosamine (which is then de-*N*-acetylated), three mannose residues, and an ethanolamine-phosphate which is linked to the mature *C*-terminus of the anchored protein. GPI biosynthesis begins on the cytoplasmic face of the endoplasmic reticulum; following a flipping event, similar to that during the formation of dolichyl-linked oligosaccharides, and final modifications of the core, the GPI-anchor is transferred in a transamidase-mediated reaction to the protein [277,278]. Obviously, a number of glycosyltransferases and auxiliary proteins are involved: these are encoded by *PIG* genes in mammals and *GPI* genes in yeast and trypanosomes. A somatic, rather than inherited, defect in the *PIG-A* gene in humans results in

the absence of decay accelerating factor from erythrocyte membranes and, thus, to paroxysmal nocturnal haemoglobinuria [279]. *GPI* genes are essential for the survival of the bloodstream form of *Trypanosoma brucei* [280] and subtle differences in the GPI-biosynthetic pathways of mammals and trypanosomatids have facilitated the development of inhibitors, which specifically target parasitic GPI biosynthesis [281]. GPI biosynthesis is also important for viability of the human fungal pathogen *Candida albicans* [282].

5 Glycosaminoglycan Biosynthesis

Animals synthesise several different types of glycosaminoglycans (GAGs) consisting of disaccharide repeats: (i) keratan sulfate in the form of sulfated poly-*N*-acetyllactosaminoglycans present on 'normal' *N*- and *O*-glycans [283], (ii) chondroitin, dermatan and heparan sulfate chains attached via xylose to serine residues of proteoglycan core proteins [284,285] and (iii) the free glycosaminoglycan known as hyaluronic acid [286]. These types of glycoconjugates have been known for some fifty or more years [287,288]; however, only recently has molecular biology made its impact on 'GAGobiology'. The biosynthesis of chondroitin, dermatan and heparan sulfates is initiated by the transfer of xylose to serine residues of a relevant core protein (see ❯ *Fig. 7*); generally, an acidic patch followed by a Ser-Gly sequence is required for xylosyltransferase to act [289]. Whereas most invertebrates have one xylosyltransferase gene [290,291], mammals have two such enzymes, XT-I and XT-II, which are functionally similar [292,293,294]; these xylosyltransferases display homology to the core 2 GlcNAc transferases involved in formation of some mammalian *O*-glycans. RNAi-feeding of *Caenorhabditis* with bacteria that express double-stranded RNA targeting the xylosyltransferase showed that embryonal cytokinesis depends on the presence of this protein [295].

Following xylosylation of the peptide core, the action of three other enzymes (Gal-TI, Gal-TII and GlcA-TI in the nomenclature of 'GAGolgists' and not to be confused with actual gene names) results in formation of the tetrasaccharide core (GlcAβ1,3Galβ1,3Galβ1,4Xyl-*O*-Ser) common to all chondroitin, dermatan and heparan sulfate chains; these enzymes are respectively encoded by the *B4GALT7, B3GALT6* and *B3GAT3* genes. The first are members of the familiar β1,4- and β1,3-glycosyltransferase families [296,297,298] and deficiencies of *B4GALT7* gene are associated with some cases of Ehlers–Danlos syndrome [299].

After the action of the β1,3-glucuronyltransferase, a 'decision' must be made: the action of an α1,4-GlcNAc transferase is required for initiation of heparan sulfate [300], whereas the transfer of a β1,4-GalNAc determines the beginning of chondroitin sulfate synthesis (dermatan sulfate being a modification of a chondroitin-type backbone) [301]. The subsequent formation of GlcAβ1,3GalNAcβ1,4- and GlcAβ1,4GlcNAcα1,4-disaccharide repeats is dependent on the actions of multiple enzymes. Some of these have two active sites and may form 'co-polymerases'. In mammals, the genetic basis is particularly complex, since five genes (*EXT1*, *EXT2*, *EXTL1*, *EXTL2* and *EXTL3*) are involved in the initiation and polymerisation of heparan sulfate [302]. The gene names reflect the role of mutations in heparan sulfate biosynthesis in the formation of exostoses, a benign bone tumour. Interestingly, exostoses patients only ever have one defective allele; in mice, homozygous mutants do not survive [303]. In the case of chondroitin sulfate, various 'chondroitin synthases' or 'chondroitin polymerisation factors' have been identified and combinations of these proteins may be involved in chondroitin poly-

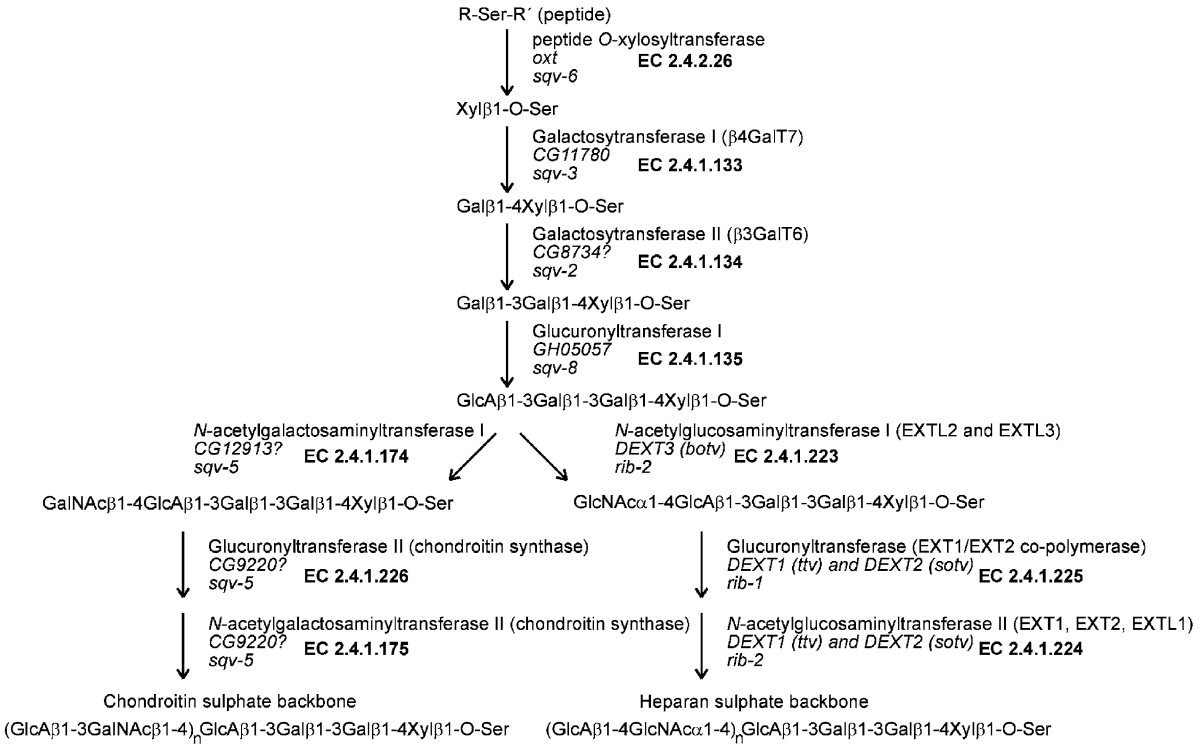


Figure 7

The initial stages of chondroitin and heparan sulfate biosynthesis. Each enzyme is defined by its shorthand name(s) as well as by the abbreviations of the relevant known or proposed corresponding *Drosophila* and *Caenorhabditis* genes

merisation [304]. Defects in all enzymes required for chondroitin synthesis in *Caenorhabditis elegans*, including those involved in formation of the tetrasaccharide core region, are associated with so-called *sqv* (**sq**uashed **v**ulva) phenotypes, indicative of the role of all the glycosyltransferases involved, as well as a transporter and UDP-sugar synthesising enzymes, in morphogenesis [305]. The two genes involved in nematode heparan sulfate biosynthesis are defective in the *rib-1* and *rib-2* mutants [306], whereas the corresponding *Drosophila* EXT-like genes are known as *ttv* (*tout-velu* or all-hairy), *botv* (brother of *ttv*) and *sotv* (sister of *ttv*) [307]. The phenotypes of all three mutant flies involve defects in growth factor signalling indicating that the fly also requires the formation of heparan sulfates for normal development to occur [308].

Once the chondroitin, heparan or keratan backbone has been formed, various modifications take place, particularly *O*-sulfation, *N*-deacetylation/*N*-sulfation and epimerisation to iduronic acid [309,310,311,312], which result in the final sulfated forms of these macromolecules found in nature. RNAi- or morpholino-induced knock-down of heparan sulfotransferases in the fruit fly, worm or zebrafish and a knockout of chondroitin-4-sulfotransferase 1 in the mouse are associated with developmental defects [313,314,315,316]. In humans, defects in the *CHST6* (**c**arbo **h**ydrate **s**ulfo **t**ransferase **6** or corneal GlcNAc-6-sulfotransferase) gene, associated with abnormal keratan sulfate levels, are the apparent cause of macular corneal dystrophy [317], a disease which results in progressive bilateral loss of vision; furthermore, Omani-type spondyloephiphyseal dysplasia, a skeletal disease, is associated with a mutation in the *CHST3* (chondroitin-6-sulfotransferase) gene [318]. Certainly, many growth factors bind to glycosaminoglycan chains and so this accounts for the defects observed in their absence; for instance, heparan sulfate is required for the non-classical secretion of fibroblast growth factor 2 potentially by generating a concentration gradient across the plasma membrane [319].

6 Biosynthesis of Plant Polysaccharides

Akin to the animal proteoglycans, the plant arabinogalactans are considered to be important in plant development [320]; as the name suggests, these glycoconjugates contain arabinose-modified branched galactose polymers attached to protein, which is often attached to the membrane via a GPI anchor. Despite the biological importance, knowledge of the full structures and the relevant enzymology is lacking. This is actually a general phenomenon in terms of our knowledge about plant polysaccharides in general and has been exacerbated by the lack of defined glycosyltransferase acceptor substrates as well as decent quantities of the necessary nucleotide sugars. For instance, only recently has the enzyme required for UDP-arabinofuranose biosynthesis (a mutase) been found [321]. However, there has been much progress in recent years, which has been aided by the sequencing of the *Arabidopsis* genome and by the use of forward and reverse genetics in this plant [322].

In terms of quantity, probably cellulose (Chap. 6.3) is the most important plant polysaccharide and has a key role in maintaining the cell wall. Enzyme complexes, with perhaps 36 subunits, produce this polymer and a number of the relevant *CESA* (cellulose synthase) genes have been identified; defects in some *CESA* genes are associated with some *irx* (**ir**regular **x**ylem) phenotypes [323]. Related proteins, encoded by cellulose synthase-like A (*CslA*) genes, are apparently β-mannan synthases [324]. Mannans are only one of the many hemicellu-

losic polysaccharides which associate with cellulose: xyloglucans and xylans are other examples, for which a number of relevant enzymes have been identified. Xyloglucans contain a poly-β1,4-glucose chain decorated with α1,6-xylose residues, which may be further modified by β1,2-galactose and α1,2-fucose. Examples of all three types of required enzyme have been identified: two xylosyltransferases [325,326], a fucosyltransferase encoded by the *MUR2* gene [327,328] and a galactosyltransferase encoded by the *MUR3* gene [329]. Xylans, on the other hand, contain a poly-β1,4-xylose backbone modified with glucuronic acid and 4-*O*-methylglucuronic acid residues. Relevant *Arabidopsis* mutants, *fra8* (i. e., *fragile fibre 8*), *irx8* and *irx9*, have a 'dwarf plant' phenotype and a defective glucuronoxylan composition compatible with the homology of the corresponding predicted gene products to glycosyltransferases [330,331].

Two other major classes of plant polysaccharides are the homogalacturonans and rhamnogalacturonans of pectin. The *quasimodo1* (*qua1*) mutant also has a 'dwarf' phenotype and shows a reduced homogalacturonan content, although the actual function of the protein encoded by the *QUA1* gene is yet to be proven [332]. Similarly, the exact function of the *ARAD1* (*ARABINAN DEFICIENT 1*) gene has not yet been defined, but it is putatively required for the formation of α1,5-arabinan side chains on rhamnogalacturonan [333]. Recently, an α1,4-galacturonyltransferase (GAUT1) involved in homogalacturonan synthesis [334] and a putative glucuronyltransferase (GUT1), with homology to animal *EXT* heparan synthases, required for the developmentally important borate cross-linking of rhamnogalacturonan-II [335] have been also identified. However, it is obvious that a multitude of genes and enzymes involved in polysaccharide synthesis are still unknown.

7 Structural and Functional Aspects of Glycosyltransferases

As introduced above, glycosyltransferases (GTs) are enzymes which catalyse the transfer of an activated donor sugar, usually a nucleotide-sugar, to an appropriate acceptor molecule that can be another sugar, lipid, protein or small molecule. They are present in both prokaryotes and eukaryotes, and they generally display exquisite specificity for both the glycosyl donor and the acceptor substrates. The vast majority of mammalian GTs are endoplasmic reticulum- and Golgi-resident membrane proteins. Whereas some ER glycosyltransferases possess multiple transmembrane domains, a typical Golgi-resident GT has the topology of a type II membrane protein consisting of a short *N*-terminal cytoplasmic tail followed by a transmembrane domain, a stem region of variable length and a large *C*-terminal catalytic domain that extends into the Golgi lumen. The topology of Golgi GTs may be more complex since some GTs were shown to have additional non-catalytic modules (such as a lectin domain) that could potentially mediate substrate recognition and/or protein-protein interaction. GTs involved in the synthesis of alternating polysaccharides, such as heparin, hyaluronan, are examples of bi-functional enzymes harbouring two separate catalytic domains.

7.1 Classification of Glycosyltransferases

The development of bioinformatics, together with new or improved cloning strategies, led in the last decade to the identification of a wide number of GTs. These enzymes are primarily

classified on the basis of their donor, acceptor and product specificity, according to the recommendations of the International Union of Biochemistry and Molecular Biology (IUBMB). In addition, they are considered as either "inverting" or "retaining" enzymes depending on whether the stereochemistry is changed or conserved during glycosyl transfer. Some of these enzymes have a strict requirement of divalent cations for activity, whereas many others are metal cation-independent or may just be activated by metal cations (such as Mn^{2+}). GTs are also classified into sequence similarity-based families [336] which are kept updated on the Carbohydrate Active enZyme database (CAZy, available at http://afmb.cnrs-mrs.fr/CAZY/). In addition to nucleotide-sugar-dependent GTs, the database also integrates enzymes that utilise dolichol-phospho-sugars, sugar-1-phosphates, and lipid diphospho-sugars as activated donors. At the time of writing (February 2007), the CAZy database comprises over 24,000 known and putative GT sequences that have been divided into 86 GT families (denoted as GTx), the vast majority of these sequences being uncharacterised open reading frames. Although there are families numbered GT1 to GT87, the GT36 family has been reclassified as a glycohydrolase family (GH94) since these are phosphorolytic enzymes, such as a chitobiose phosphorylase, whose 3D structure does not resemble any glycosyltransferase, but rather one of a glycosidase [337].

Large differences in the number and function of GTs are observed among families. Some families are monospecific and comprise only a few sequences whereas other families contain a huge number of sequences from various sources with diverse functions. For these polyspecific families, sequence similarities are mostly restricted to a portion of the catalytic domain. Therefore if a putative GT sequence is related to a large polyspecific family, its precise biochemical function cannot be reliably predicted and one of the major challenges facing glycobiologists, particularly those working on bacterial and plant GTs, is to determine the donor and acceptor specificities of the many hypothetical GT sequences identified through the systematic sequencing of genomes. Eukaryotic genomes encode a wide variety of GTs from many of the sequence-derived families (e. g., 230 human GT sequences that are found in 42 out of 86 GT families). In contrast, GT sequences of archaeal genomes show a different picture since they fall only into two GT families, GT2 and GT4, which represent the two largest inverting and retaining enzyme families, respectively. This sequence-based classification is assumed to integrate both structural and mechanistic features within each family [336]. When applied to glycosylhydrolases, this classification scheme was shown to correlate well with enzyme mechanism (inversion or retention of the anomeric configuration). This may not apply so safely to all GT families, because distant similarities between inverting and retaining CAZy families have been noted, suggesting that these families may share a common ancestor [338,339]. A similar observation has been made for the retaining GT27 and GT78 families which display sequence similarities with the large inverting GT2 family [340,341,342].

7.2 Crystal Structures of Glycosyltransferases

Structure-function studies on GTs have been hampered for many years by the relative paucity of structural data for this important class of enzymes. To date, crystal structures for 39 GTs have been solved and structural information is now available for 25 GT families (*Table 1* and *Table 2*). The past 3 years have seen significant advances in structural glycobiology

with the first three-dimensional structures of sialyltransferases and fucosyltransferases as well as the resolution of several members of the family of UDP-GalNAc:polypeptide α-*N*-acetylgalactosaminyltransferases (ppGalNAcTs) which are involved in the first step of mucin-type *O*-glycosylation. The latter enzyme family is rather unique among glycosyltransferases, containing both catalytic and lectin domains.

In contrast to glycosyl hydrolases, which display a myriad of topologies, GT structures have, thus far, revealed only two folds (and variants thereof), termed GT-A and GT-B [341,343], and which were first observed in the original SpsA and β-glucosyltransferase (BGT) structures [344,345]. The two GT folds now determined consist primarily of $\alpha/\beta/\alpha$ sandwiches, similar or very close to the Rossmann-type fold, a classical structural motif found in many nucleotide-binding proteins. It is striking to note that the GT-A and GT-B folds are also shared by non-GT enzymes, such as nucleotidyltransferases and sugar epimerases [346]. As shown on ❯ *Table 1* and ❯ *Table 2*, 3D structures are available for inverting and retaining GTs originating from eukaryotes and prokaryotes. Most of these proteins have been crystallised in complex with either the donor or/and acceptor substrates (or analogues). Of the 25 GT families that have a three-dimensional structural representative, 13 and 9 families display the canonical GT-A and GT-B folds, respectively. Three of the remaining families can be considered as variants of these two folds (see below). However, it must be emphasised that the two recent structures of peptidoglycan transglycosidases (members of the GT51 family) are not related to either the GT-A or GT-B fold, but they show notable structural similarity to the bacteriophage lambda lysozyme, a glycosidase, both in the overall fold and in the catalytic centre [347,348]; therefore, they are not considered further in this chapter. There is no correlation between the overall fold of a GT and its mechanism of action, since both inverting and retaining enzymes of both fold types are known [341]. A database that gathers structural information concerning GTs is now available (http://www.cermav.cnrs.fr/glyco3D). It contains useful links that permit retrieval of bibliographic information as well as atomic coordinates at the Protein Data Bank, and images are provided to illustrate the details of protein-substrate interactions.

The recent burst of reported GT structures has shed light on their substrate binding and catalytic mechanisms. GTs are thought to use mechanistic strategies that directly parallel those used by the well characterised glycosyl hydrolases. However, while the mechanism of inverting GTs appears to be widely accepted, the mechanism of retaining enzymes remains a topic of considerable debate. These aspects will be discussed later in the chapter.

7.2.1 The GT-A Fold

The GT-A fold is best described as a single $\alpha/\beta/\alpha$ domain that is formed by a central β-sheet flanked by α-helices (❯ *Fig. 8a,b*; see also list in ❯ *Table 1*). The mixed β-sheet typically consists of seven β-strands of topology 3214657 where strand 6 is antiparallel. The central β-sheet is flanked by a smaller one, and the association of both creates the active site. All of the metal ion-dependent GTs for which structures have been determined display this type of fold. Until recently, a general feature of all members of the GT-A superfamily was considered to be the presence of a common motif, the DxD motif, and their requirement for a divalent cation for activity. The DxD motif is a degenerate sequence that is shown in all crystal structures to interact primarily with the phosphate groups of the nucleotide donor through the coordination of a divalent cation, typically Mn^{2+}. However, the recently determined X-ray structure of

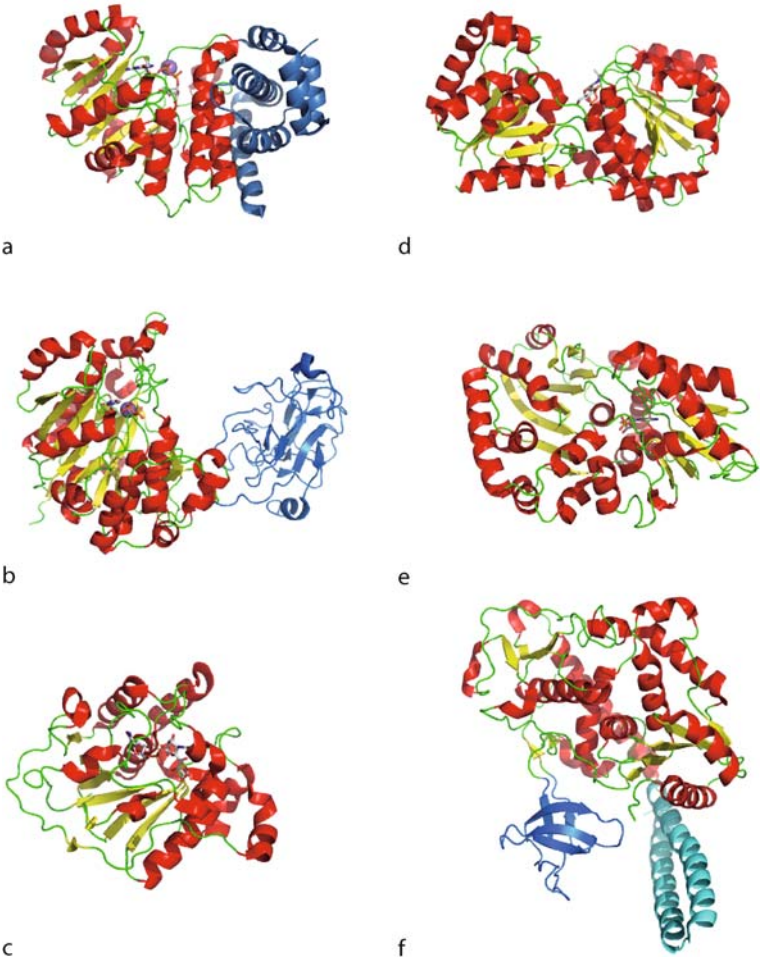


Figure 8
Graphical representation of representative crystal structures of glycosyltransferases. Catalytic domains are coloured yellow and red, and additional domains blue. a Mannosylglycerate synthase (PDB code 2BO8 [342]), b ppGalNAc-transferase 10 (code 2D7I [349]), c sialyltransferase CstII (code 1RO7 [350], d GlcNAc-transferase MurG (code 1NLM [351]), e sialyltransferase PmSTI (code 2EX1 [352], f α6-fucosyltransferase (code 2DE0 [353]). Structures a and b have GT-A folds, d and e are of the GT-B type, whereas c and f are variants of GT-A and GT-B folds, respectively

murine leukocyte type Core 2 β6-*N*-acetylglucosaminyltransferase (C2GnT-L, a key enzyme in the biosynthesis of branched *O*-glycans; see *Sect. 3.1*) shows that this enzyme possesses the canonical GT-A fold but it lacks the characteristic metal ion binding DxD motif [354]. The role played by the metal ion in other metal ion-dependent GT-A structures is served in C2GnT-L by two basic residues located in the *C*-terminus of the catalytic domain. The use of such basic amino acids to counteract the positive charge of the diphosphate moiety is striking-

Table 1
Crystal structures of glycosyltransferases with GT-A (or GTA-like) fold

CAZY family	Function	Name	Organism	R/I[a]	Ref[b]
		GT-A fold			
GT2	Putative glycosyltransferase	SpsA	*Bacillus subtilis*	I	[344]
GT6	α3-Galactosyltransferase	α3GalT	*Bos taurus*	R	[359]
	α3-Galactosyltransferase B	GTB	*Homo sapiens*	R	[360]
	α3-GalNAc-transferase A	GTA	*Homo sapiens*	R	[360]
GT7	β4-Galactosyltransferase I	β4GalT1	*Bos taurus*	I	[152]
	β4-Galactosyltransferase I	β4GalT1	*Homo sapiens*	I	[361]
GT8	α4-Galactosyltransferase	LgtC	*Neisseria meningitidis*	R	[357]
	α-Glucosyltransferase	Glycogenin	*Oryctolagus cuniculus*	R	[362]
GT13	β2-GlcNAc-transferase I	GnT1	*Oryctolagus cuniculus*	I	[363]
GT14	β6-GlcNAc-transferase (core 2)	C2GnT	*Mus musculus*	I	[354]
GT15	α2-Mannosyltransferase	Kre2p/Mnt1p	*Saccharomyces cerevisiae*	R	[364]
GT27	Polypeptide GalNAc-transferase 1	ppGalNAcT1	*Mus musculus*	R	[365]
	Polypeptide GalNAc-transferase 2	ppGalNAcT2	*Homo sapiens*	R	[366]
	Polypeptide GalNAc-transferase 10	ppGalNAcT10	*Homo sapiens*	R	[349]
GT31	β3-GlcNAc-transferase (fringe)	Mfng	*Mus musculus*	I	[367]
GT43	β3-GlcA-transferase I	GlcAT I	*Homo sapiens*	I	[368]
	β3-GlcA-transferase P	GlcAT-P	*Homo sapiens*	I	[369]
GT44	α-Glucosyltransferase	Toxin B	*Clostridium difficile*	R	[370]
GT64	α4-GalNAc-transferase	EXTL2	*Mus musculus*	R	[371]
GT78	Mannosylglycerate synthase	MGS	*Rhodothermus marinus*	R	[342]
		GT-A like fold			
GT42	α2,3/8-Sialyltransferase	CstII	*Campylobacter jejuni*	I	[350]

[a]Mechanism, R: retaining, I: inverting [b]Only the first solved crystal structure is cited for each glycosyltransferase family

ly similar to that observed in a number of GT-B fold GTs and this suggests a convergence of catalytic mechanism.

Comparison of the catalytic domains of enzymes of the GT-A superfamily revealed the presence of two regions that are structurally well conserved in both inverting and retaining enzymes [341]. This suggests that common structural elements are necessary for the glycosyl transfer reaction, irrespective of the stereochemistry of the reaction. The first region mostly corresponds to the Rossmann-type nucleotide binding domain, encompassing the first 100–120 residues of the catalytic domain, and that is often terminated by the aforementioned DxD motif. The key amino acids that interact with nucleotide donor are mostly located at equivalent positions. The *C*-terminal part of the catalytic domain is mostly dedicated to the recognition of the acceptor and it generally shows higher variability. However, a common structural motif, that is part of the active site, is observed in the acceptor binding domains of

all GT-A enzymes. It comprises residues that were shown in some crystal structures to interact with both the donor sugar and the sugar acceptor [355,356,357,358].

As mentioned above, the family of UDP-GalNAc:polypeptide α-*N*-acetylgalactosaminyltransferases (ppGalNAcTs, GT27; see *Fig. 8b*) is rather unique among glycosyltransferases, in possessing a C-terminal, ricin-type lectin domain of ~130 residues containing three putative carbohydrate-binding sites. Biochemical analyses suggest that this domain functions in the transfer of GalNAc to glycopeptide but not peptide substrates [372,373]. The recently solved 3D structures of three protein members (ppGalNAc-T1, -T2 and -T10) reveal that the association of ppGalNAcT catalytic and lectin domains can be dynamic and also reveal the molecular bases of substrate recognition [349,365,366]. These studies highlighted the importance of the peptide region linking the catalytic and lectin domains and that is thought to modulate acceptor substrate specificity [349]. This "flexible tether" model proposed by Fritz et al. [366], suggests that lectin domain mobility endows certain ppGalNAcTs with an even greater capacity to adapt to and capture glycosylated substrates, thus ensuring the high density of glycosylation characteristic of mucin domains.

The first structure determined for a sialyltransferase (CstII from *Campylobacter jejuni* of family GT42; see *Fig. 8c*) revealed an unusual variant of the GT-A fold [350]. The protein displays a similar type of fold as the canonical GT-A fold, but with some differences in the connectivity of β-strands (parallel β-sheet of topology 8712456) and it has no DxD motif. Therefore the CstII structure represents a new type of fold. Another prokaryotic sialyltransferase has, though, a GT-B fold (see *Table 2*).

Structural studies on several GT-A enzymes have shown that one or two flexible loops undergo a marked conformational change from an open to a closed conformation upon nucleotide sugar binding, as first observed with GlcNAc-TI [363]. This provides a structural basis for the ordered binding of the donor and acceptor to the enzyme, i. e., the "ordered sequential" binding mechanism (UDP-sugar on first, UDP off last) deduced from earlier kinetic studies on β4Gal-TI and GlcNAc-TI [132,374]. On binding of the metal ion, if required, and the nucleotide sugar, a disordered loop in the free enzyme becomes ordered and moves to act as a lid covering the buried donor substrate. This conformational change creates a pocket that will accommodate the acceptor substrate. After catalysis, the glycosylated product is released first, before the nucleotide. Another significant feature is the observed distorted conformation of the bound nucleotide sugar in the active site that may be important in the catalytic mechanism (for a review, see [375]).

7.2.2 The GT-B Fold

The GT-B fold is characterised by two separate Rossmann-type domains with a connecting linker region and a catalytic site located between the domains. Both domains show an $\alpha/\beta/\alpha$ structure formed by a central parallel β-sheet of general topology 321456 (*Fig. 8d,e*; see also list in *Table 2*). The similarity between the two domains is high enough to propose that they are the result of gene duplication [376]. There is an excellent structural conservation between protein members of the GT-B family, particularly in the *C*-terminal domain which corresponds to the nucleotide binding region. Variations are much more pronounced in the *N*-terminal domains, in the loops and α-helices which point towards the active site, which have evolved to accommodate very different acceptors. The GT-B superfamily is particularly

Table 2
Crystal structures of glycosyltransferases with GT-B (or GTB-like) fold

CAZY family	Function	Name	Organism	I/R[a]	Ref[b]
		GT-B fold			
GT1	β-epi-Vancosaminyltransferase	GtfA	*Amycolatopsis orientalis*	I	[379]
	β-Glucosyltransferase	GtfB	*Amycolatopsis orientalis*	I	[376]
	β-Vancosaminyltransferase	GtfD	*Amycolatopsis orientalis*	I	[380]
	Flavonoid β-glucosyltransferase	UGT71G1	*Medicago truncatula*	I	[381]
	Flavonoid β-glucosyltransferase	VvGT1	*Vitis vinifera*	I	[382]
GT4	α3-Glucosyltransferase	WaaG	*Escherichia coli*	R	[383]
	Putative glycosyltransferase	aviGT4	*Streptomyces viridochromogenes*	R	[383]
GT5	α4-Glucosyltransferase (glycogen synthase)	AtGS	*Agrobacterium tumefaciens*	R	[384]
	α4-Glucosyltransferase (glycogen synthase)	PaGS	*Pyrococcus abyssi*	R	[385]
GT9	Heptosyltransferase	WaaC	*Escherichia coli*	I	[386]
GT20	Trehalose-6-phosphate synthase	OtsA	*Escherichia coli*	R	[387]
GT28	β4-GlcNAc-transferase	MurG	*Escherichia coli*	I	[388]
GT63	β-Glucosyltransferase	BGT	*Phage T4*	I	[345]
GT72	α-Glucosyltransferase	AGT	*Phage T4*	R	[389]
GT80	α2,3-Sialyltransferase	PmST1	*Pasteurella multocida*	I	[352]
		GT-B like fold			
GT10	α3-Fucosyltransferase	FucT	*Helicobacter pylori*	I	[390]
GT23	α6-Fucosyltransferase	FUT8	*Homo sapiens*	I	[353]

[a]Mechanism, R: retaining, I: inverting [b]Only the first solved crystal structure is cited for each glycosyltransferase family

remarkable for the diversity of products its members produce, particularly those that classify into GT1 family and which include most of the prokaryotic and plant enzymes that glycosylate secondary metabolites, and a large number of known or putative glucuronosyltransferases that glycosylate potentially toxic lipophilic compounds [377]. Although divalent cations may be required for full activity of GT-B enzymes [377,378], there is no evidence of a bound metal ion associated with catalysis. In the GT-B family, either the positive helix dipole effect of an $\alpha/\beta/\alpha$ structural motif or positively charged side chains of basic residues are found to counteract the negative charge of the diphosphate moiety.

Common themes in GT-B enzymes concern the presence of a conserved peptide motif termed GPGTF ("Glycogen Phosphorylase/Glycosyltransferase" motif) as defined by Wrabl and Grishin [391] located in the nucleotide binding domain and that is involved in binding the glycosyl donors, the involvement of a glutamate residue (that is part of the GPGTF motif) in ribose binding, the presence of a long *C*-terminal α-helix which interacts with the *N*-terminal domain, a conformational change upon sugar donor binding, and the phenomenon that

most of the protein contacts with the acceptor involve residues of the *N*-terminal domain. However, so far, GT-B enzymes do not appear to share any strictly conserved residues and exceptions to the above rules have been recently documented. Glycogen synthases (GT5) lack the glutamate residue that specifically anchors the ribose moiety (GtfA and GtfD that use dTDP also lack the Glu residue) [379,380]. The bacterial sialyltransferase of *Pasteurella multocida* (PmST1, GT80) lacks the typical signature of the GT-B fold of the *C*-terminal α-helix crossing over to the *N*-terminal domain (◉ *Fig. 8e*) [352]. The phage T4 AGT enzyme (GT72) presents an unusual feature since it predominantly binds donor and acceptor substrates via its *C*-terminal domain [389]. Domain movements that would bring the two substrates to correct positions for catalysis are expected for GT-B enzymes. Such conformational changes have been observed upon nucleotide sugar donor binding in BGT (GT63), MurG (GT28), and PMST1 (GT80) [351,352,378]. In contrast, WaaC, a heptosyltransferase of *E. coli* (GT20) does not undergo such domain rotation upon substrate binding, but the substrate analogue (ADP-2F-heptose) and the reaction product (ADP) are found to adopt remarkably distinct conformations inside the active site [386].

Sequence similarities between GT families are extremely low. However, phylogenetic trees based upon sequence alignments derived from the structural superimposition of *C*-terminal domains of 18 GT-B structures (3D representatives of families GT1, GT4, GT5, GT9, GT20, GT28, GT35, GT63, GT72, GT80) intriguingly resulted in a clear separation of inverting and retaining enzymes [383]. The sole "outlier" is the inverting GT80 bacterial sialyltransferase which appears on the "retaining" side of the tree. The GT80 enzyme is also unusual in many other ways in that it uses a nucleoside monophosphate donor and has also been reported to catalyse transglycosylation with retention (in addition to glycosyl transfer with inversion) [352]. Furthermore, the structural overlap of all 18 GT-B enzymes also shows a clear structural basis for the inverting versus retaining enzymes based upon the topology of the loops interacting with the donor sugar itself [383].

The recently determined first two 3D-structures of the fucosyltransferases revealed that these enzymes adopt folds that are variants of the GT-B fold. The bacterial α3-fucosyltransferase of *Helicobacter pylori* [390] belongs to the GT10 family that also comprises the eukaryotic α3- and α3/4-fucosyltransferases. It displays the same 2-domain architecture as other GT-B enzymes but with some differences in the connectivity of β-strands. Since the authors crystallised a truncated and poorly active form of the protein, and also because the overall conformation appeared to be more open than other GT-B enzymes, detailed interactions observed in the crystal structures of the enzyme complexed with GDP and GDP-fucose should be interpreted cautiously. The second fucosyltransferase to have been recently crystallised is the human α6-fucosyltransferase (FUT8) of family GT23, which adds a fucose to the reducing terminal GlcNAc of the core structure of an Asn-linked oligosaccharide [353]. This enzyme displays a quite unusual modular architecture consisting of an *N*-terminal coiled coil region followed by a catalytic domain and a SH3 domain (◉ *Fig. 8f*). SH3 domain-containing proteins are typically localised in the cytosol where they mediate numerous signal transduction pathways via critical protein-protein interactions. The α6-FucT is a type II membrane GT located in the Golgi and thus the function of its SH3 domain, that appears to be unique in GTs, is not yet understood. The catalytic domain is comprised of two sub-structures, an open sheet α/β structure and a Rossmann domain. The three conserved peptide regions that were previously identified in all α2 and α6-fucosyltransferases, as well as in protein-*O*-fucosyltransferases,

are found in the Rossmann domain [392,393]. Although the structure of the *N*-terminal part is not a typical Rossmann fold, the catalytic domain of α6-FucT was found to be similar to the GT-B enzymes, particularly to the ADP-heptose lipopolysaccharide heptosyltransferase of family GT9.

Little information is known about the detailed catalytic mechanism of GT-B enzymes. Sequential ordered as well as random kinetic reactions have been described for enzymes belonging to the GT-B superfamily. For instance, MurG utilises a sequential ordered mechanism in which the nucleotide sugar (UDP-GlcNAc) binds first [351], whereas in the case of BGT, each substrate (UDP-Glc and DNA) can bind randomly [394]. In both cases, large conformational changes that bring the two domains together are observed upon nucleotide sugar binding.

7.3 Molecular Modelling and Fold Recognition Studies

Since only a few topologies have been observed for the GTs that have been crystallised to date, the use of fold recognition methods seems to be a very promising approach. Fold recognition is a theoretical approach which allows alignment of one sequence with one structure by a process referred as "threading" [395]. When performed on selected sequences representing CAZy GT families having no structural representative, such threading analyses predicted that many other GT families should adopt the GT-A or GT-B fold [338,341,346]. The picture that emerged from these studies was the prediction of a GT-A fold and a GT-B fold for 22 and 13 additional families, respectively, and several among them were later confirmed by crystal structures. A fold similar to the one of the bacterial sialyltransferase CstII was predicted with confidence for GT29, the enzyme family which comprises the eukaryotic sialyltransferases [341]. Some families gave weak or moderate scores in fold recognition thus suggesting that novel folds for GTs are likely to be discovered. Fold recognition was also used to identify new GT sequences in the fully sequenced genome of *Mycobacterium tuberculosis* [396]. Probabilistic methods of database searching, such as PSI-BLAST [397] and HMMs [398], have also been used to predict the fold of GTs [339,399].

Molecular modelling has been applied only recently in the field of glycosyltransferases. This approach can be particularly useful to decipher structure-function relationships of this class of enzymes, particularly for a better understanding of the molecular bases governing donor and acceptor specificities, and to identify the key active site residues. Once a sequence has been attributed a fold, it is possible to predict the secondary structure elements, namely α-helices and β-strands, and to align it with sequences of known 3D structure. Ideally, the model must be built using templates that present a maximum of sequence similarities with the candidate protein. However, molecular modelling of GTs presents peculiar difficulties. Sequence similarities between GT families, and even within one family, are extremely low, and this makes sequence alignment particularly difficult. The ratio of loops to secondary structure elements is high in GTs and many crystal structures do not describe the entire catalytic domain, because several loops are flexible and do not present clear electron density. Therefore, the 3D models will have a low confidence index in these regions. Docking of substrates also appears to be a difficult task owing to the flexibility of the nucleotide sugar and the presence of phosphate and divalent cation [400]. Difficulties also come from the fact that large movements of loops and domains are observed in GTs upon substrate binding, and the enzymes can exist in an

"open" or "closed" conformation, therefore complicating the modelling procedure. When the target and the template have sufficient identity, the models are accurate and allow for docking of nucleotide sugar and acceptor. This approach was used to model several GTs belonging to GT6 (a family which comprises the mammalian α3-GalTs and α3-GalNAcTs and which adopts a GT-A fold) [401]. Specificity towards the sugar donor and sugar acceptor was shown to be determined by a few critical residues in the binding site. In addition, a closed active conformation of the modelled enzymes has been proposed that may complement data from the crystal structures of the blood group A and B transferases of family GT6 that have been determined in an open (inactive) conformation [360]. Homology modelling has also been applied to identify key active site residues of plant GTs that adopt a GT-B fold [402,403,404].

The conclusion from this review on the 3D structures of GTs is that a sequence-based classification spreads GTs into a large number of families, but only two general folds, called GT-A and GT-B, have been observed to date. In addition, threading analyses indicate that the majority of as yet uncharacterised GT families could adopt one of these two folds. However, the small variety of folds observed for GTs is compensated by a large structural variability in the acceptor binding domain, thus conferring some functional plasticity which allows fine tuning with respect to the acceptor. Compared to glycosyl hydrolases, our detailed knowledge of catalytic mechanisms of GTs is extremely poor and much work remains.

8 Structural Insights into the Catalytic Mechanism Using Ab Initio Molecular Modelling

Complementary to the crystallographic and kinetic approaches for the examination of glycosyltransferases, ab initio molecular modelling has also increased our understanding of the catalytic mechanism of these enzymes and has opened up the possibility of new approaches in the design of glycosylation inhibitors.

8.1 The Catalytic Mechanism of Glycosyltransferases

The transfer of a glycosyl residue, from an activated donor to an acceptor catalyzed by glycosyltransferases, can be regarded as a nucleophilic displacement of the functional group at the anomeric carbon, e. g. nucleoside diphosphate, of the transferred saccharide residue of a donor by a hydroxyl group of a specific acceptor. The glycosylation reaction can mechanistically proceed with either inversion or retention of stereochemistry at the anomeric carbon of the donor sugar. Glycosyltransferases are, therefore, classified as either retaining or inverting, depending on the stereochemical outcome (see *Fig. 9*).

The catalytic mechanisms of glycosidases and glycosyltransferases were commonly assumed to be rather similar, the main difference being the acceptor molecule to which the glycosyl residue is transferred**:** a water molecule in the case of glycosidases and a free hydroxyl group of a specific saccharide acceptor in the case of glycosyltransferases [405,406,407,408]. With the increase of structural information on glycosyltransferases, some distinct differences between the mechanistic strategies employed by glycosidases and glycosyltransferases are becoming apparent [341,409,410,411,412]. Many aspects of the functions and the catalytic mechanisms of glycosyltransferases, however, are still unresolved. Theoretical studies of enzyme-catalyzed

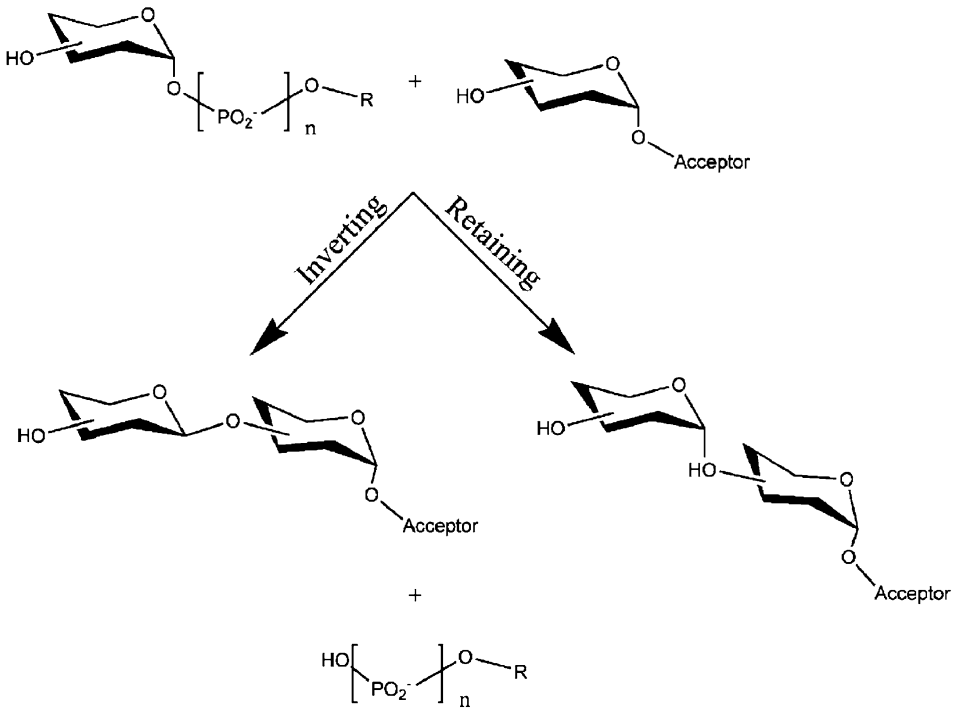


Figure 9
Schematic representation of the overall reaction catalyzed by glycosyltransferases

reactions have recently received a great deal of attention because they can lead to a deeper understanding of enzymatic catalysis. Specifically, hybrid methodologies [413,414] have allowed for the inclusion of the protein environment into the study of the catalytic reaction and represent a promising approach that may offer more detailed insights into the catalytic mechanism of enzymes than is possible by experimental means alone [415,416,417].

8.2 Inverting Glycosyltransferases

The catalytic mechanism of inverting glycosyltransferases has emerged from the extensive theoretical investigations on models of different complexities. The modelling results [418,419] showed that the proton transfer to the glycosidic oxygen is not required for the breaking of the C1–O1 bond, and they imply that, in contrast to glycosidases, glycosyltransferases do not require a second catalytic residue in their active site for the catalytic reaction to proceed. The proposed mechanism, illustrated in *Fig. 10*, suggests a concerted S_N2-type mechanism featuring a nearly simultaneous nucleophilic addition of the oxygen nucleophile from the acceptor to the C1 carbon of the donor sugar, as well as dissociation of the C1–O1 bond of a nucleoside diphosphate accompanied by the proton transfer from nucleophile to the catalytic base.
This type of mechanism is supported by the available crystal structures of inverting glycosyltransferases. In fact, none of the completed structures of glycosyltransferases, either inverting or retaining, has so far revealed the presence of two catalytic amino acids in the active site of

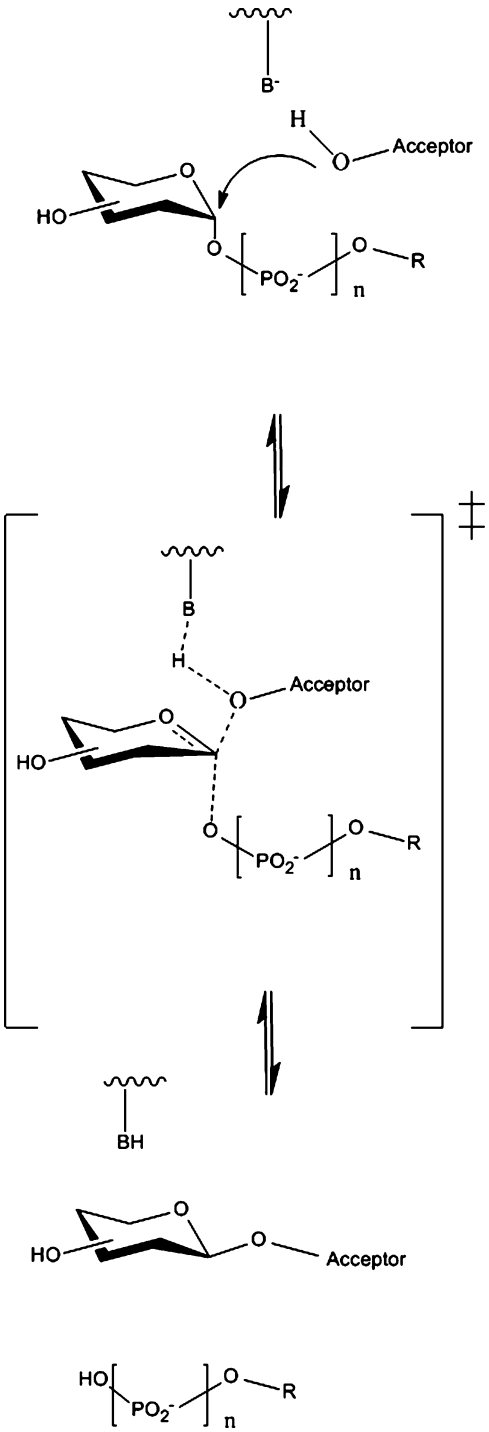

B-
H
O
Acceptor
HO
O
O
PO2-
O
R
n
B
H
Acceptor
BH
HO
Acceptor

Figure 10
Schematic representation of the predicted S_N2-like mechanism for inverting glycosyltransferases, with the enrolment of only a catalytic base inferred from the calculated potential energy surfaces of the catalytic reaction and supported by available X-ray structures

the enzyme. Moreover, the recently solved crystal structure of β1,3-glucuronyltransferase 1 (GlcAT1) in the presence of donor and acceptor substrates [368] provided direct support for the mechanism involving just one catalytic residue.

A recent investigation [420] of the catalytic mechanism of GlcNAc-TI provided additional important information on the structures of the active site models in the Michaelis (ES) and transition state (TS) complexes, respectively. These calculations used a hybrid QM/MM methodology and a model, based on the crystal structure of GlcNAc-TI [363], which contains a whole enzyme and its substrates. The structure of the Michaelis complex model (*Fig. 11*) revealed several features that are relevant for the catalytic reaction. The crucial aspect for the catalytic mechanism is the position of a nucleophile, the C2 hydroxyl group. In the ES, a distance of 2.97 Å between the nucleophile O_{ac} and the anomeric carbon C1 of the donor is such that O_{ac} oxygen is in an appropriate position for the nucleophilic attack. The O_{ac} hydroxyl group also makes the key hydrogen bond with the carboxylate oxygen O_B of the catalytic base [363]. The calculated O_{ac}–O_B and H2–O_B distances of 2.592 Å and 1.677 Å, respectively, indicate a short, strong hydrogen bond (low-barrier hydrogen bond, LBHB) [421]. The LBHB interaction facilitates deprotonation of the nucleophile and thereby increases the nucleophilicity of O_{ac}. The C6 hydroxyl of the transferred sugar makes hydrogen bonds to both the β-phosphate oxygen of the UDP-GlcNAc and to the water molecule HOH14 placed in the active site of the enzyme [363]. The ES structure also shows several hydrogen bonds between the acceptor hydroxyl groups and surrounding enzyme that are supported by kinetic studies on substrate analogues [422]. These experiments showed that either removal or methylation of some hydroxyl groups leads to a loss or a decrease of the enzyme activity.

In the procession from the ES complex to the TS, as expected, the major movements can be observed for the acceptor and for the transferred GlcNAc. The structure of the active site in the TS model is displayed in *Fig. 11*. The TS features nearly simultaneous nucleophilic addition and dissociation steps, as evidenced by the O_{ac}–C1 and C1–O1 bond distances of 1.912 Å and 2.542 Å, respectively. The nucleophile hydrogen atom H2 is located at the oxygen O_B of the catalytic base. The glucopyranose ring adopts the $^4H_3$ conformation in the TS model. The sum of the bond angles surrounding C1 is 354° and shows that C1 is essentially planar. As a result, the anomeric hydrogen H1 is positioned in the plane defined by the C1, C2 and O5 atoms. This arrangement facilitates coincident interactions of C1 with the leaving O1 and the nucleophile O_{acc}. The C1–O5 distance of 1.333 Å points to a significant delocalisation of the O5 lone pairs to C1 and results in the partial double-bond character of the bond between the ring oxygen and the anomeric carbon. This interaction stabilises the partial positive charge on the anomeric carbon C1 within the oxocarbenium ion-like character of the TS. The cleavage of the C1–O1 glycosidic bond is accompanied by the 17° rotation of the β-phosphate oxygen upon going from the ES to the TS complex, while the remaining part of UDP does not show any relevant changes. One can only speculate whether this kind of conformational change is characteristic for inverting glycosyltransferases, since a similar conformational change of the β-phosphate oxygen during a cleavage of the anomeric linkage was observed [423] for β-1,4-galactosyltransferase I.

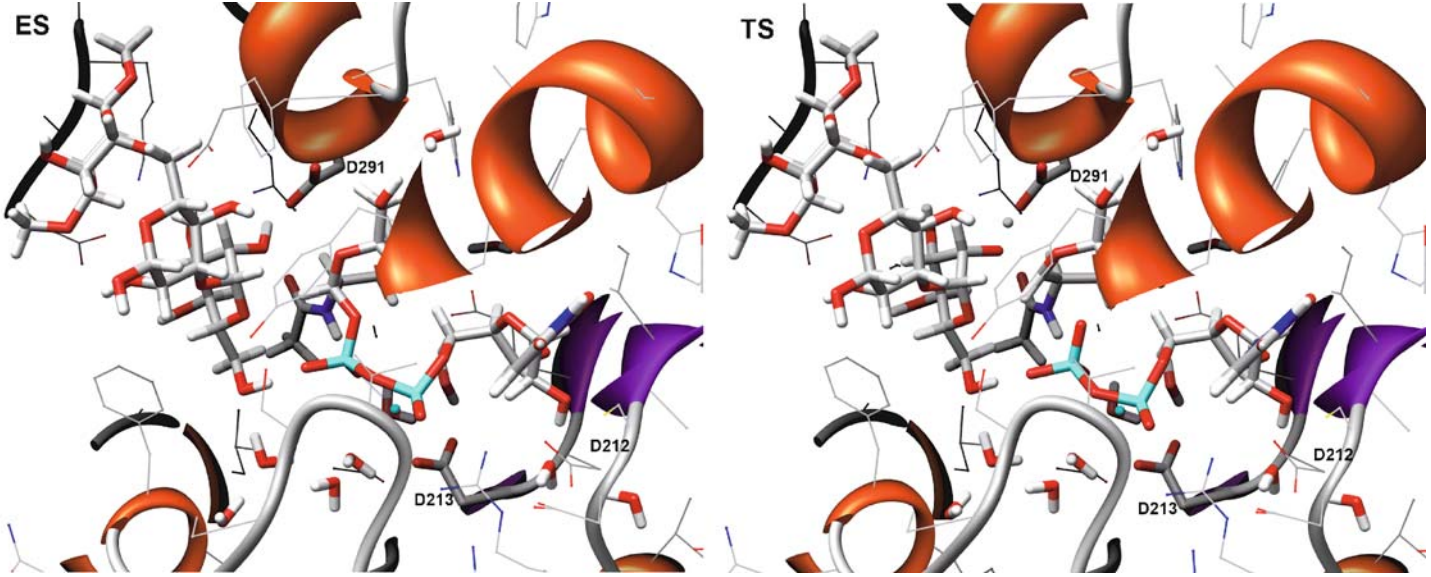


Figure 11

The active site model for the Michaelis complex (ES) and transition-state complex (TS) of GlcNAc-TI, obtained using the QM(DFT)/MM method. Red indicates oxygen, blue-green phosphorus, blue is nitrogen; Mn is a small sphere below the phosphate, close to the loop and D213

The calculated energy barrier for the proposed mechanism ES → TS → PC (*Fig. 12*) is 18.7 kcal mol^{-1} and is consistent with the experimentally observed barriers of 15–25 kcal mol^{-1} (based on published k_{cat} values [424,425]), though this coincidence is likely to be superficial. Further support for the proposed mechanism came from the calculated α-deuterium kinetic isotope effect (KIE). Experimental values of observed α-^{2}H KIE for inverting glycosyltransferases varies in the interval $D_{V/K}$ = 1.050–1.210 [425,426,427,428,429]. The calculated α-^{2}H KIE value of 1.060 is in the range of those experimental determined; however, it is slightly higher than values of measured KIE for the S_N2 mechanism and is considerably lower than KIE values for the reactions corresponding to the S_N1 mechanism. The calculated value may suggest, that in the catalytic reaction, the nucleophilic addition and dissociation of the glycosidic bonds occurs nearly synchronously, as in the S_N2 mechanism, without forming a covalent intermediate with an enzyme. Similar mechanistic features were observed for a dissociative S_N2 mechanism of the purine nucleoside phosphorylase [430], in which the distances of the making and breaking bond are not symmetrical in the transition state, and the length of the breaking bond is longer than the forming bond.

Understanding the structural origin of catalytic efficiency of enzymes is a fundamental goal and challenge in biological science. A comparison of individual energetic contributions from the QM/MM energies revealed that the energetic contribution from the enzyme decreases the overall energetic barrier by 8.8 kcal mol^{-1} and indicates that the GlcNAc-TI environment is favouring the TS complex over the ES complex. Analysis of the calculated structural features indicates several ways in which the GlcNAc-TI environment is involved in the catalytic reaction. In the ES complex, the acceptor is placed in the active site in such a way that the nucleophile oxygen O2 is properly positioned for the nucleophilic attack and for activation by the catalytic base. The O_{ac} atom is at a distance of 3.0 Å from C1 of the donor and 2.6 Å from the catalytic base oxygen. Moreover, the orientation of the nucleophile and the leaving group oxygen atoms represents almost the ideal angle for the nucleophilic attack. In the TS complex, two specific interactions that could be responsible for a specific stabilisation in the TS model were identified, namely, LBHB between the H2 and O_B and strong interactions between the hydrogen atom H6 and β-phosphate oxygen (indicated by a distance of 1.66 Å) that together with the hydrogen from the S322 side chain hydroxyl group stabilise the building of a negative charge on the β-pyrophosphate.

8.3 Retaining Glycosyltransferases

Theoretical investigations [431] of various possible reaction pathway models, describing a double-displacement mechanism for retaining glycosyltransferases, clearly showed that the "classical" mechanism involving two catalytic residues is, from an energy point of view, very doubtful. The calculations showed that the energetically most favourable pathway requires the presence of only one catalytic acid in the active site. This led to a proposed double-displacement mechanism (*Fig. 13*), in which the first step consists of a nucleophilic attack of a catalytic base on the anomeric carbon of the donor and the formation of the covalent glycosyl-enzyme complex with the inverted (β-) configuration at the anomeric carbon C1. In the second step, the hydroxyl group of the acceptor attacks the anomeric carbon of the glycosyl-enzyme complex, with the UDP functioning as a general base to deprotonate the

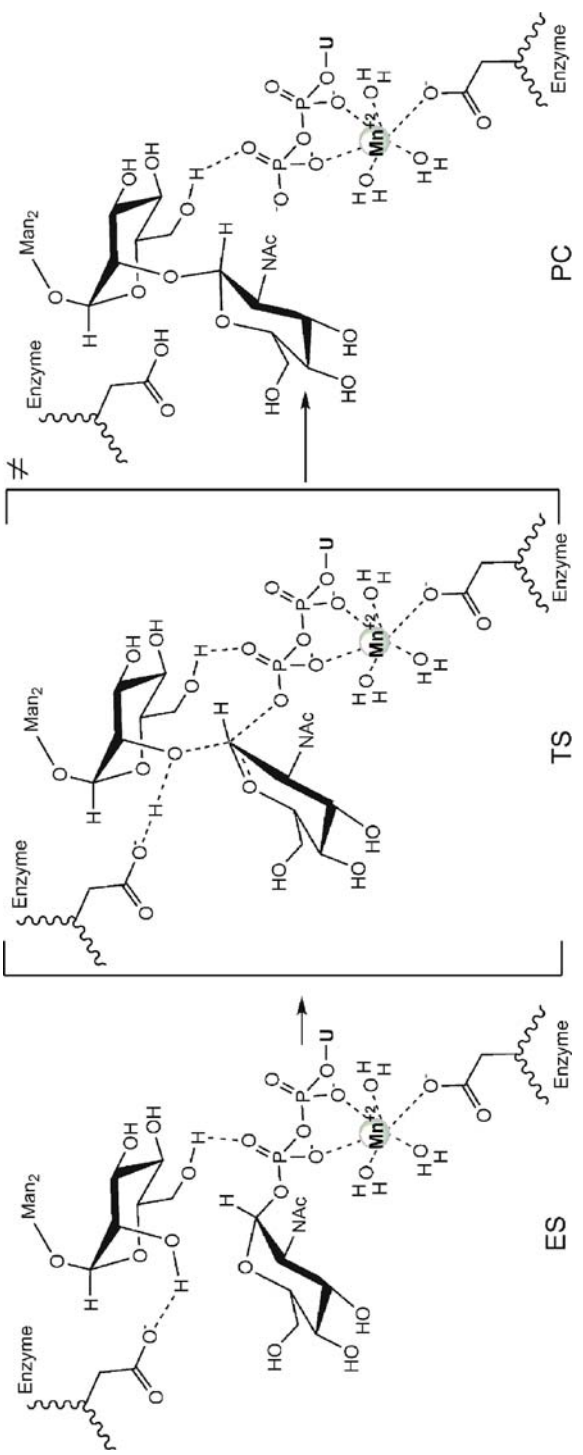


Figure 12
Schematic representation of the proposed mechanism for the transfer of GlcNAc catalyzed by GlcNAc-TI. ES, enzyme-substrate complex; TS, transition state complex; PC, product complex

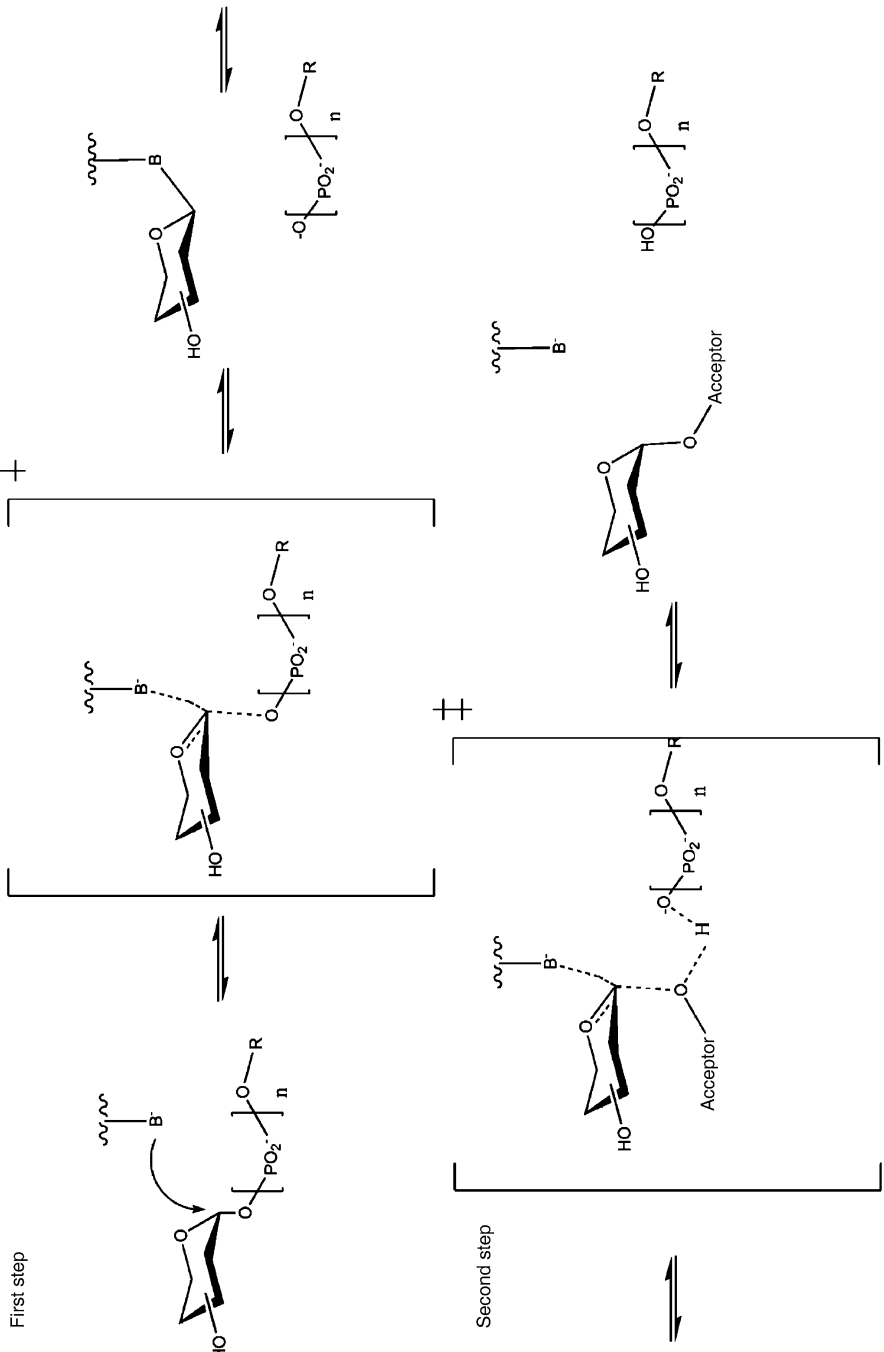


Figure 13
Schematic representation of the proposed double displacement mechanism for retaining glycosyltransferases, inferred from the calculated potential energy surfaces for various mechanisms of retaining glycosyltransferases

hydroxyl group. The second inversion at the anomeric centre occurs during this step, with the resulting glycoside having the α-configuration. The crystal structures of several retaining glycosyltransferases [152,355,359,360,362] support this type of mechanism as they show only one catalytic acid in the active site of the enzyme. However, the conclusive evidence for this mechanism remains to be seen since covalent intermediates have not yet been reported for these enzymes.

Two retaining glycosyltransferases, namely lipopolysaccharyl α-galactosyltransferase C (LgtC) from *Neisseria meningitides* [357] and α1,4-*N*-acetylglucosaminyltransferase EXTL2 [371], represent a very specific case and, therefore, their mechanism was extensively explored both experimentally [357,432,433] and theoretically [434,435,436]. The crystal structures of LgtC and EXTL2 solved in the presence of both the donor and acceptor substrates revealed a catalytic site in which the only functional group appropriately placed for the catalytic reaction was glutamine or arginine in the case of LgtC and EXTL2, respectively. The analysis of thorough investigations [357,432] that focused on whether these and other amino acids can function as possible catalytic nucleophiles was inconclusive. The results for LgtC pointed toward another mechanism that might not require any participation of a catalytic residue, and they led to the suggestion [357] that LgtC glycosyltransferase might instead proceed through a single front-side displacement reaction, also known as a S_{Ni} mechanism.

Molecular modelling examinations [434] of various reaction hypotheses of the galactosyl transfer catalyzed by LgtC using a model based on the crystal structure of LgtC [357] shed some light on this mechanistic puzzle. Calculations predicted that the preferred reaction pathway is a one-step mechanism in which the O_{acc} nucleophile from the acceptor attacks the anomeric carbon of the galactopyranose residue of the donor UDP-Gal from a face of the leaving group with the simultaneous proton transfer to phosphate oxygen. The proposed mechanism, illustrated in *Fig. 14*, proceeds via a front attack similar to the proposed S_{Ni} mechanism [357], but does not directly involve any catalytic acid in the reaction mechanism. Moreover, the calculations clearly showed [434] that for this kind of mechanism the specific orientation of both the donor and acceptor sugar is crucial. From these results, it is apparent that the active site constraints, imposed by the enzyme on the substrates, change the conformation of the donor UDP-Gal and the acceptor to the one that is suitable for the nucleophilic attack from the leaving group side. However, a covalent intermediate recently observed [433] for the Gln198Glu mutant revived discussion of the catalytic mechanism of LgtC, but more experimental evidence is clearly needed to fully understand the catalytic mechanism of retaining glycosyltransferases.

The calculated TS structure of LgtC shows several interesting geometrical features (*Fig. 15*). The most interesting one is that the nucleophilic attack by the acceptor oxygen O_{ac} and the leaving group departure occur on the same face of the transferred galactose [434]. The transition state structure is also characterised by C1–O1 and C1–O_{ac} distances of 2.66 Å and 2.34 Å. The oxygen from the phosphate and the acceptor hydroxyl sandwich the nucleophile proton assisting in the formation of the transition state. This proton is located at the O_{ac}–H and O1–H distances of about 1.09 Å and 1.38 Å, respectively. The distance between the O_{ac} and O1 oxygen is about 2.4 Å and suggests that the LBHB interaction provides, as similarly found for inverting glycosyltransferases, the stabilising energy for the transition state. One can only speculate whether the LBHB interactions represent a general feature employed by glycosyltransferases to facilitate their catalytic action. The ring conformation of the transferred galac-

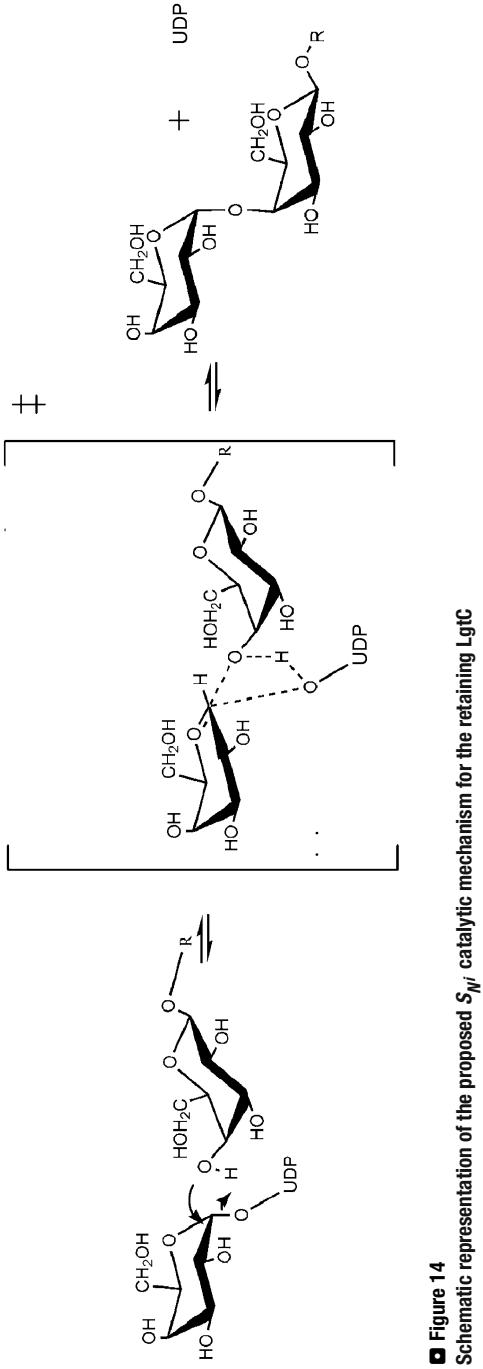


Figure 14
Schematic representation of the proposed S_Ni catalytic mechanism for the retaining LgtC

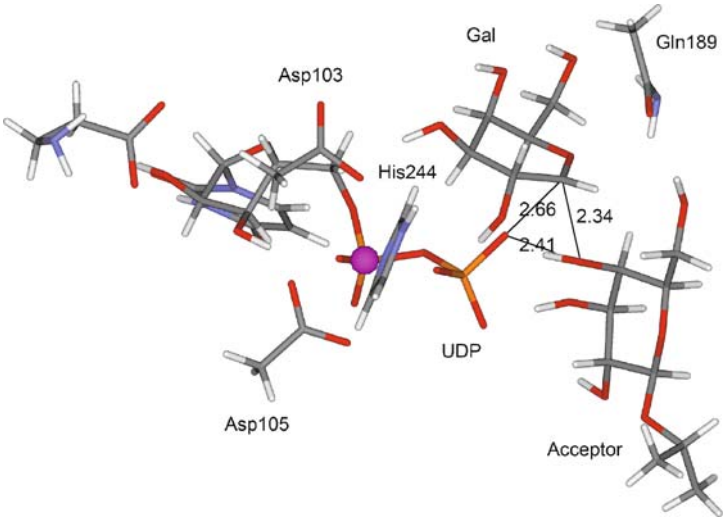


Figure 15
The calculated models for the transition state complex (TS) of LgtC, obtained using the DFT method

topyranose is a distorted ^4E envelope. It is evident that the geometry of this transition state structure estimated by the ab initio method is unique and differs from the transition state structure models previously determined for inverting or retaining glycosyltransferases.

8.4 Transition State Analysis

Transition states of catalytic reactions are high energy intermediates on reaction paths between reactants and products with lifetimes of $\sim 10^{-13}$ s, and their structure and location in catalytic sites can be only be estimated. In the case of glycosyltransferases, the "entire" transition state consists of both a deformed nucleotide sugar and acceptor oligosaccharide linked in a specific arrangement. Unfortunately, crystal structures of complexes with transition state analogues that could provide some information about the nature of the transition states are not available. Thus, model structures of transition states represent important outcomes of the calculated potential energy surfaces.

Investigations of various models of catalytic mechanism for glycosyltransferases of the GT-A superfamily [418,419,420,431,434] revealed the following general characteristics of the transition state models: a) the ring conformation resembles a half-chair conformation with an oxo-carbenium character at the anomeric carbon; b) the distance of C1–O1 is longer than the normal C–O bond length; c) a new β-glycosidic bond is being created with the bond length larger than the normal bond length; both the forming and breaking bonds are oriented almost perpendicularly with respect to the plane defined by the C2-C1-O5 atoms.

Calculations on glycosyltransferases [418,419,420,431,434] have led to 23 different transition state models. These belong to various reaction pathways having different energetics. However, these TS models provide pertinent information on possible structural variations of transi-

tion states for glycosyltransferases. The structural analysis revealed significant variations in the C1–O_{ac} and C1–O1 bond lengths that can be as large as 1.3 Å and 1.7 Å, respectively. It appeared that calculated TS structures can be clustered using geometrical criteria based on the values of two reaction coordinates [434,435,436]. Indeed, clustering of the TS structures based on the C1–O_{ac} and C1–O1 distances led to three distinct groups that may represent canonical models of transition states. Structural characteristics of these groups are illustrated in ❯ *Fig. 16*. The first group (❯ *Fig. 16a*) is characterised by long C1–O_{ac} bonds within the range of 2.4–2.7 Å and short C1–O1 distances between 1.5 and 2.1 Å. The geometrical parameters of this transition state model are similar to the initial reactants and this canonical form has been termed "early transition state". The second group (❯ *Fig. 16b*) corresponds to the "intermediate transition state" structures where both C1–O_{ac} (2.1–2.4 Å) and C1–O1 (2.5–2.7 Å) distances are elongated compared to their initial values, but the structures did not yet reach the final arrangement observed in the products. The third group (❯ *Fig. 16c*) is characterised by short C1–O_{ac} bonds within the range of 1.4–1.6 Å and long C1–O1 distances between 2.8 Å and 3.2 Å. The geometry of this canonical form is close to that of the final products and therefore has been named "late transition state". The three canonical structures might represent a framework of variations occurring in transition state structures of glycosyltransferases. Of course, the transition state structure of a particular enzyme depends on the structure of the catalytic site which, in turn, imposes constraints on substrates to generate a required transition state structure for a given enzyme. For example, the TS model inferred from QM/MM calculations on GlcNAc-TI corresponds to the intermediate transition state.

8.5 Rational Design of Transition State Analogues

Different strategies have been used in order to identify potent inhibitors of glycosyltransferases [437,438,439,440,441,442,443,444]. In spite of an increasing number of potential inhibitors, only a few of them have exhibited significant activity [443,444]. Though transition state analogues are valued tools for drug discovery as potent and specific inhibitors of enzymes, to date, the ability to generate transition state analogues of glycosyltransferases has lagged behind [442,443,444]. The TS models obtained from molecular modelling may provide blueprints for the design of the transition state analogue inhibitors. A transition state analogue is a stable compound that structurally resembles the three-dimensional structure and charge distribution of the unstable transition state. These analogues have the potential to bind more tightly than the substrates, proportional to the catalytic rate enhancement imposed by the enzyme. With an enzymatic rate enhancement of 10^{10} – 10^{15}, transition state analogues have the potential for near-irreversible binding on the biological scale.

It is clear that, due to the above described characteristics of TS, it would not be possible to design a stable molecule that mimics the TS precisely. However, even crude TS analogues consisting of relevant features are expected to be excellent reversible inhibitors. To achieve this, the TS analogue should contain the main structural features of the TS from both the acceptor and donor. Recently, based on the calculated information, a new scaffold for TS analogues has been designed [445], with an attempt to mimic structural features around the reaction centre. The proposed structural mimic of TS is represented by the (tetrahydro-2-(methylthio)furan-2-yl)methyl phosphate dianion (**1**) shown in ❯ *Fig. 17*.

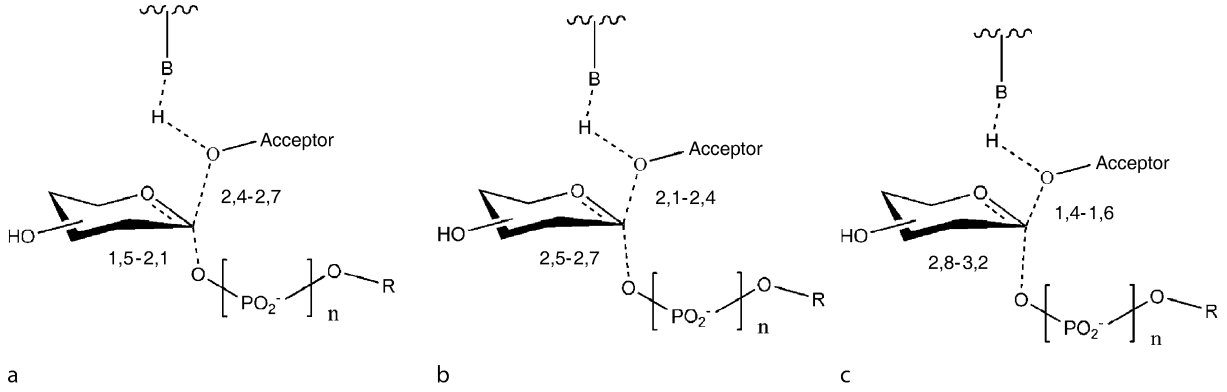
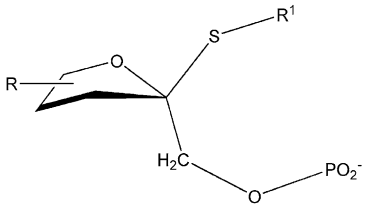


Figure 16
Schematic representation of canonical forms of the transition states for glycosyltransferases. Canonical forms are defined by similarities in their C1–O1 and C1–O_{acc} distances. (a) Early transition state, (b) Intermediate transition state, and (c) Late transition state



R = OH, NHAc, ...etc.
R^1 = Et, Phe, ...etc.

Figure 17
Schematic representation of the proposed TS scaffold for the (tetrahydro-2-(methylthio)furan-2-yl)methyl phosphate dianion (1)

In this scaffold, the enlarged C1–O1 distance presented in TS is accomplished by linking a methyl phosphate group to the C1. In this moiety, the C1–O distance between the anomeric carbon and the phosphate oxygen is 2.4 Å, which is very close to the corresponding distance in the TS; the length of the C1–S5 bond is 1.9 Å, similar to the distance between the anomeric carbon and the attacking oxygen in the TS. The acceptor is mimicked by a thiol linkage, in which the C1–C distance of 2.9 Å corresponds to the prolong C1–O_{ac} distance in the TS. A deformed monosaccharide ring in the TS is represented by a five-membered ring. The structural similarity between the transition state and its analogue is a key factor in determining the binding properties of TS analogues. Generally, low energy conformers of (**1**) closely overlap with the corresponding parts of the TS; the furanose ring oxygen with the pyranose ring oxygen, sulfur with the attacking oxygen, methyl carbon with the acceptor carbon atom, and a phosphate oxygen with the phosphate oxygen of the leaving group of the donor [445]. A superposition also revealed that a phosphate dianion in **1** is properly oriented to interact with a bivalent metal cofactor presented in the active site of retaining glycosyltransferases. While these kinds of electrostatic interactions are very strong, it remains to be seen whether these interactions are sufficient to mimic the UDP part of the donor. Calculations clearly showed that proposed scaffolds for potential transition state analogues exhibit very similar structural features as those predicted for the TS of reactions catalyzed by inverting glycosyltransferases. To further explore binding characteristic of this scaffold, its derivatives were docked into the active site of the TS model of GlcNAc-TI [420]. It appeared that the binding mode of the proposed TS analogues in GlcNAc-TI is very similar to that of the TS [446] as illustrated in *Fig. 18*, which support its utilisation as a TS mimetic.

However, just how potent and specific the inhibitors will be, based on this scaffold, remains to be seen. It is obvious that this scaffold is only a basis for further refinement, which might involve various modifications of substituents on the furanose ring to mimic a transferred monosaccharide, and a substitution of the group linked to the sulfur atom by various larger substituents mimicking the acceptor monosaccharide. In particular, the later variations might be crucial for determining the potency and specificity of design inhibitors, since it is well known that there is a limited number of donor nucleotides and a very large number of possible acceptors that likely define the specificity of glycosyltransferases.

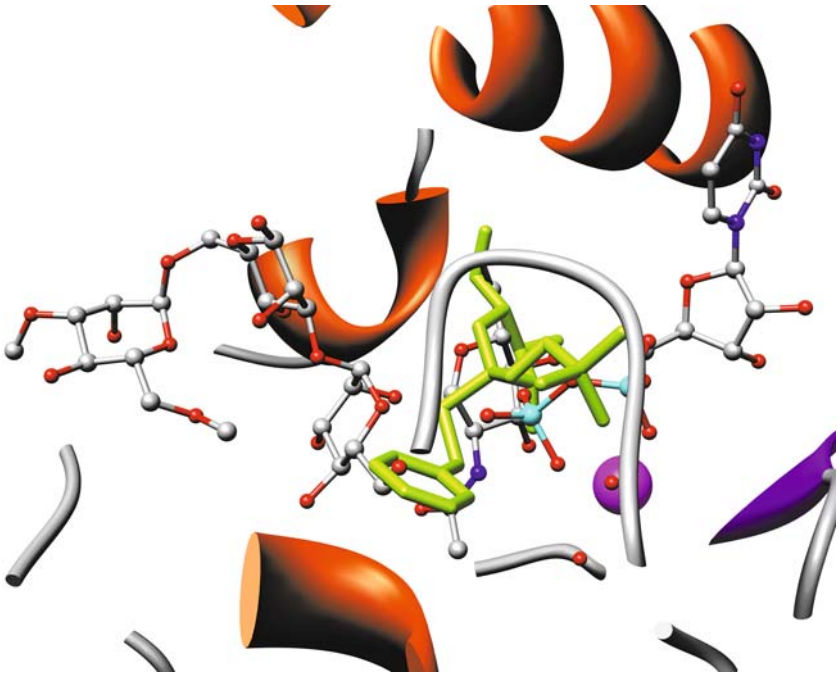


Figure 18
The (tetrahydro-2-(phenylthio)furan-2-yl)methyl phosphate dianion (green) docked into the active site of GnT-I. For comparison, the location of the active site of the TS model is also given

A detailed understanding of glycosyltransferase function provides a basis for their application in the development of new classes of therapeutic agents. It is also clear, however, that more structural, enzymological, and computational work remains to be done to fully understand the catalytic machinery of glycosyltransferases.

9 Concluding Remarks

In our account of the biochemical, biological and enzyme structural aspects of the biosynthesis of oligo- and polysaccharides, we have concentrated on selected aspects of this vast field. It is obvious that the past few years have been very fruitful in the field of structural glycobiology. In particular, the use of recombinant enzymes, genetic ablation techniques and the deployment of mass spectrometric glycan analysis have greatly increased our knowledge about glycosyltransferases and processing glycosidases. Indeed the sheer complexity of glycan biosynthesis has become even clearer. In the future one may expect that further advances will be made through the use of interdisciplinary approaches in determining the function of glycans and the pathways for their biosynthesis; the many questions remaining will occupy glycobiologists and glycochemists in the years ahead.

Acknowledgement

We thank Dr. Harry Schachter (Hospital for Sick Children, Toronto) for his comments on the chapter as a whole and Drs. Inka Brockhausen (Queen's University, Kingston, Ontario), Vlad Panin (Texas A & M University), Terry Butters (University of Oxford) and Terry Smith (University of Dundee) for their advice on latest developments in *O*-glycan, glycosphingolipid and GPI biosynthesis as well as Katharina Paschinger for proof-reading.

References

1. Davies GJ, Gloster TM, Henrissat B (2005) Curr Opin Struct Biol 15:637
2. Henrissat B, Coutinho P M (1999) In: Gilbert HJ, Davies GJ, Henrissat B, Svensson B (eds) Recent Advances in Carbohydrate Bioengineering, The Royal Society of Chemistry, p 3
3. Stone BA, Svensson B (2001) In: Fraser-Reid BO, Tatsuta K, Thiem J (eds), Glycoscience – Chemistry and Chemical Biology Vol. III, 1st edn., Springer, p 1905
4. Kornfeld R, Kornfeld S (1985) Annu Rev Biochem 54:631
5. Freeze HH (2006) Nat Rev Genet 7:537
6. Taniguchi N, Honke K, Fukuda M (eds) (2002), Handbook of Glycosyltransferases and Related Genes, Springer
7. Parodi A, Behrens NH, Leloir LF, Carminatti H (1972) Proc Natl Acad Sci USA 69:3268
8. Samuelson J, Banerjee S, Magnelli P, Cui J, Kelleher DJ, Gilmore R, Robbins PW (2005) Proc Natl Acad Sci USA 102:1548
9. Aebi M, Hennet T (2001) Trends Cell Biol 11:136
10. Bickel T, Lehle L, Schwarz M, Aebi M, Jakob CA (2005) J Biol Chem 280:34500
11. Gao X-D, Tachikawa H, Sato T, Jigami Y, Dean N (2005) J Biol Chem 280:36254
12. O'Reilly MK, Zhang G, Imperiali B (2006) Biochemistry 45:9593-9603
13. Tkacz JS, Lampen JO (1975) Biochem Biophys Res Commun 65:248
14. Marek KW, Vijay IK, Marth JD (1999) Glycobiology 9:1263
15. Kelleher DJ, Gilmore R (2006) Glycobiology 16:47R
16. Moremen K, Trimble RB, Herscovics (1994) Glycobiology 4:113
17. Herscovics A (1999) Biochim Biophys Acta 1473:96
18. Roth J, Ziak M, Zuber M (2003) Biochimie 85:287
19. Elbein AD (1991) FASEB J 5:3055
20. Gonzalez DS, Karaveg K, Vandersall-Nairn AS, Lal A, Moremen KW (1999) J Biol Chem 274:21375
21. Tremblay LO, Herscovics A (1999) Glycobiology 9:1073
22. Dairaku K, Spiro RG (1997) Glycobiology 7:579
23. Parodi A (2000) Biochem J 348:1
24. Helenius A, Aebi M (2004) Annu Rev Biochem 73:1019
25. Eylar EH (1965) J Theoret Biol 10:89
26. Caramelo JJ, Castro OA, Alonso LG, de Prat-Gay G, Parodi AJ (2003) Proc Natl Acad Sci USA 100:86
27. Braakman I (2001) EMBO Rep 2:666
28. Hirao K, Natsuka Y, Tamura T, Wada I, Morito D, Natsuka S, Romero P, Sleno B, Tremblay LO, Herscovics A, Nagata K, Hosokawa N (2006) J Biol Chem 281:9650
29. Suzuki T, Lennarz WJ (2003) Biochem Biophys Res Commun 302:1
30. Suzuki T, Yano K, Sugimoto S, Kitajima K, Lennarz WJ, Inoue S, Inoue Y, Emori Y (2002) Proc Natl Acad Sci USA 99:9691
31. Suzuki T, Hara I, Nakano M, Shigeta M, Nakagawa T, Kondo A, Funakoshi Y, Taniguchi N (2006) Biochem J 400:33
32. Herscovics A (2001) Biochimie 83:757
33. Tremblay LO, Herscovics A (2000) J Biol Chem 275:31655
34. Romero PA, Vallée F, Howell PL, Herscovics A (2000) J Biol Chem 275:11071
35. Tremblay LO, Nagy Kovacs E, Daniels E, Wong NK, Sutton-Smith M, Morris HR, Dell A, Marcinkiewicz E, Seidah NG, McKerlie C, Herscovics A (2007) J Biol Chem 282:2558

36. Dean N (1999) Biochim Biophys Acta 1426:309
37. Schiller B, Voglmeir J, Pöltl G, Geyer R, Wilson IBH (2006) unpublished data
38. Varki A, Kornfeld S (1980) J Biol Chem 255:8398
39. Tiede S, Storch S, Lübke T, Henrissat B, Bargal R, Raas-Rothschild A, Braulke T (2005) Nature Med 11:1109
40. Raas-Rothschild A, Cormier-Daire V, Bao M, Genin E, Salomon R, Brewer K, Zeigler M, Mandel H, Toth S, Roe B, Munnich A, Canfield WM (2000) J Clin Invest 105:673
41. Burke J, Pettitt JM, Schachter H, Sarkar M, Gleeson PA (1992) J Biol Chem 267:24433
42. Metzler M, Gertz A, Sarkar M, Schachter H, Schrader JW, Marth JD (1994) EMBO J 13:2056
43. Ioffe E, Stanley P (1994) Proc Natl Acad Sci USA 91:728
44. von Schaewen A, Sturm A, O'Neill J, Chrispeels MJ (1993) Plant Physiol 102:1109
45. Zhu S, Hanneman A, Reinhold VN, Spence AM, Schachter H (2004) Biochem J 382 :995
46. Shi H, Tan J, Schachter H (2006) Methods Enzymol 417:359
47. Sarkar M, Leventis PA, Silvescu CI, Reinhold VN, Schachter H, Boulianne GL (2006) J Biol Chem 281:12776
48. Stanley P, Narasimhan S, Siminovitch L, Schachter H (1975) Proc Natl Acad Sci USA 72:3323
49. Strasser R, Schoberer J, Jin C, Glössl J, Mach L, Steinkellner H (2006) Plant J 45:789
50. Paschinger K, Hackl M, Gutternigg M, Kretschmer-Lubich D, Stemmer U, Jantsch V, Lochnit G, Wilson IBH (2006) J Biol Chem 281:28265
51. Moremen KW, Robbins PW (1991) J Cell Biol 115:1521
52. Misago M, Liao YF, Kudo S, Eto S, Mattei MG, Moremen KW, Fukuda MN (1995) Proc Natl Acad Sci USA 92:11766
53. Chui D, Oh-Eda M, Liao YF, Panneerselvam K, Lal A, Marek KW, Freeze HH, Moremen KW, Fukuda MN, Marth JD (1997) Cell 90: 157
54. Chui D, Sellakumar G, Green R, Sutton-Smith M, McQuistan T, Marek K, Morris H, Dell A, Marth J (2001) Proc Natl Acad Sci USA 98:1142
55. Bonay P, Roth J, Hughes RC (1992) Eur J Biochem 205:399
56. Kawar Z, Karaveg K, Moremen KW, Jarvis DL (2001) J Biol Chem 276:16335
57. Hato M, Nakagawa H, Kurogochi M, Akama TO, Marth JD, Fukuda MN, Nishimura S (2006) Mol Cell Proteomics 5:2146
58. Akama TO, Nakagawa H, Wong NK, Sutton-Smith M, Dell A, Morris HR, Nakayama J, Nishimura S, Pai A, Moremen KW, Marth JD, Fukuda MN (2006) Proc Natl Acad Sci USA 103:8983
59. Akama TO, Nakagawa H, Sugihara K, Narisawa S, Ohyama C, Nishimura S, O'Brien DA, Moremen KW, Millan JL, Fukuda MN (2002) Science 295:124
60. Harvey DJ, Wing DR, Küster B, Wilson IBH (2000) J Am Soc Mass Spectrom 11:564
61. Allen SD, Tsai D, Schachter H (1984) J Biol Chem 259:6984
62. Tulsiani DRP, Harris TM, Touster O (1982) J Biol Chem 257:7936
63. van den Elsen JM, Kuntz DA, Rose DR (2001) EMBO J 20:3008
64. Kawatkar SP, Kuntz DA, Woods RJ, Rose DR, Boons GJ (2006) J Am Chem Soc 128:8310
65. Rabouille C, Kuntz DA, Lockyer A, Watson R, Signorelli T, Rose DR, van den Heuvel M, Roberts DB (1999) J Cell Sci 112 :3319
66. Nilsson T, Rabouille C, Hui N, Watson R, Warren G (1996) J Cell Sci 109:1975
67. Tan J, Dunn J, Jaeken J, Schachter H (1996) Am J Hum Genet 59:810
68. Wang Y, Tan J, Sutton-Smith M, Ditto D, Panico M, Campbell RM, Varki NM, Long JM, Jaeken J, Levinson SR, Wynshaw-Boris A, Morris HR, Le D, Dell A, Schachter H, Marth JD (2001) Glycobiology 11:1051
69. Schachter H (1986) Biochem Cell Biol 64:163
70. Chen S, Tan J, Reinhold VN, Spence AM, Schachter H (2002) Biochim Biophys Acta 1573:271
71. Warren CE, Krizus A, Roy PJ, Culotti JG, Dennis JW (2002) J Biol Chem 277:22829
72. Strasser R, Steinkellner H, Boren M, Altmann F, Mach L, Glössl J, Mucha J (1999) Glycoconj J 16:787
73. Tsitilou SG, Grammenoudi S (2003) Biochem Biophys Res Commun 312:1372
74. Sharkey DJ, Kornfeld R (1989) J Biol Chem 264:10411
75. Oguri S, Yoshida A, Minowa MT, Takeuchi M (2006) Glycoconj J 23:473
76. Sakamoto Y, Taguchi T, Honke K, Korekane H, Watanabe H, Tano Y, Dohmae N, Takio K, Horii A, Taniguchi N (2000) J Biol Chem 275:36029

77. Yamashita K, Kamerling JP, Kobata A (1982) J Biol Chem 257:12809
78. Aoki K, Perlman M, Lim JM, Cantu R, Wells L, Tiemeyer M (2007) J Biol Chem 282:9127
79. Sasai K, Ikeda Y, Eguchi H, Tsuda T, Honke K, Taniguchi N (2002) FEBS Lett 522:151
80. van den Eijnden DH, Koenderman AHL, Schiphorst WECM (1988) J Biol Chem 263:12461
81. Buck CA, Glick MC, Warren L (1971) Science 172:169
82. Murata K, Miyoshi E, Kameyama M, Ishikawa O, Kabuto T, Sasaki Y, Hiratsuka M, Ohigashi H, Ishiguro S, Ito S, Honda H, Takemura F, Taniguchi N, Imaoka S (2000) Clin Cancer Res 6:1772
83. Yamamoto H, Swoger J, Greene S, Saito T, Hurh J, Sweeley C, Leestma J, Mkrdichian E, Cerullo L, Nishikawa A, Ihara Y, Taniguchi N, Moskal JR (2000) Cancer Res 60:134
84. Taniguchi N, Miyoshi E, Ko JH, Ikeda Y, Ihara Y (1999) Biochim Biophys Acta 1455:287
85. Priatel JJ, Sarkar M, Schachter H, Marth JD (1997) Glycobiology 7:45
86. Granovsky M, Fata J, Pawling J, Muller WJ, Khokha R, Dennis JW (2000) Nature Med 6:306
87. Bhaumik M, Harris T, Sundaram S, Johnson L, Guttenplan J, Rogler C, Stanley P (1998) Cancer Res 58:2881
88. Demetriou M, Granovsky M, Quaggin S, Dennis JW (2001) Nature 409:733
89. Kaneko M, Alvarez-Manilla G, Kamar M, Lee I, Lee JK, Troupe K, Zhang W, Osawa M, Pierce M (2003) FEBS Lett 554:515
90. Ferrara C, Brünker P, Suter T, Moser S, Püntener U, Umaña P (2006) Biotechnol Bioeng 93:851
91. Rouwendal GJA, Wuhrer M, Florack DEA, Koeleman CAM, Deelder AM, Bakker H, Stoopen GM, van Die I, Helsper JPFG, Hokke CH, Bosch D (2007) Glycobiology 17: 334
92. Raji TS, Stanley P (1998) J Biol Chem 273:14090
93. Wilson IBH (2002) Curr Opin Struct Biol 12:569
94. Wilson IBH, Harthill JE, Mullin NP, Ashford DA, Altmann F (1998) Glycobiology 8:651
95. Jan LY, Jan YN (1982) Proc Natl Acad Sci USA 79:2700
96. Siddiqui SS, Culotti JG (1991) J Neurogenet 7:193
97. Haase A, Stern M, Wachtler K, Bicker G (2001) Dev Genes Evol 211:428
98. Faveeuw C, Mallevaey T, Paschinger K, Wilson IBH, Fontaine J, Mollicone R, Oriol R, Altmann F, Lerouge P, Capron M, Trottein F (2003) Eur J Immunol 33:1271
99. Gutternigg M, Ahrer K, Grabher-Meier H, Burgmayr S, Staudacher E (2004) Eur J Biochem 271:1348
100. Fabini G, Freilinger A, Altmann F, Wilson IBH (2001) J Biol Chem 276:28058
101. Paschinger K, Rendić D, Lochnit G, Jantsch V, Wilson IBH (2004) J Biol Chem 279:49588
102. Kubelka V, Altmann F, Staudacher E, Tretter V, März L, Hård K, Kamerling JP, Vliegenthart JFG (1993) Eur J Biochem 213:1193
103. Tretter V, Altmann F, Kubelka V, März L, Becker WM (1993) Int Arch Allergy Immunol 102:259
104. Pöltl G, Ahrazem O, Paschinger K, Ibañez MD, Salcedo G, Wilson IBH (2007) Glycobiology 17:220
105. Haslam SM, Coles GC, Munn EA, Smith TS, Smith HF, Morris HR, Dell A (1996) J Biol Chem 271:30561
106. Zhang Y, Iwasa T, Tsuda M, Kobata A, Takasaki S (1997) Glycobiology 7:1153
107. Hanneman AJ, Rosa JC, Ashline D, Reinhold VN (2006) Glycobiology 16:874
108. Uozumi N, Yanagidani S, Miyoshi E, Ihara Y, Sakuma T, Gao CX, Teshima T, Fujii S, Shiba T, Taniguchi N (1996) J Biol Chem 271:27810
109. Yanagidani S, Uozumi N, Ihara Y, Miyoshi E, Yamaguchi N, Taniguchi N (1997) J Biochem 121:626
110. Paschinger K, Staudacher E, Stemmer U, Fabini G, Wilson IBH (2005) Glycobiology 15:463
111. Wang X, Inoue S, Gu J, Miyoshi E, Noda K, Li W, Mizuno-Horikawa Y, Nakano M, Asahi M, Takahashi M, Uozumi N, Ihara S, Lee SH, Ikeda Y, Yamaguchi Y, Aze Y, Tomiyama Y, Fujii J, Suzuki K, Kondo A, Shapiro SD, Lopez-Otin C, Kuwaki T, Okabe M, Honke K, Taniguchi N (2005) Proc Natl Acad Sci USA 102: 15791
112. Wang X, Gu J, Miyoshi E, Honke K, Taniguchi N (2006) Methods Enzymol 417:11
113. Yamane-Ohnuki N, Kinoshita S, Inoue-Urakubo M, Kusunoki M, Iida S, Nakano R, Wakitani M, Niwa R, Sakurada M, Uchida K, Shitara K, Satoh M (2004) Biotechnol Bioeng 87:614
114. Mori K, Kuni-Kamochi R, Yamane-Ohnuki N, Wakitani M, Yamano K, Imai H, Kanda Y, Niwa R, Iida S, Uchida K, Shitara K, Satoh M (2004) Biotechnol Bioeng 88:901

115. Fitchette-Lainé AC, Gomord V, Cabanes M, Michalski JC, Saint Macary M, Foucher B, Cavelier B, Hawes C, Lerouge P, Faye L (1997) Plant J 12:1411
116. Wilson IBH (2001) Glycoconj J 18:439
117. Léonard R, Costa G, Darrambide E, Lhernould S, Fleurat-Lessard P, Carlué M, Gomord V, Faye L, Maftah A (2002) Glycobiology 12:299
118. Strasser R, Bondili JS, Vavra U, Schoberer J, Svoboda B, Glössl J, Léonard R, Stadlmann J, Altmann F, Steinkellner H, Mach L (2007) Plant Cell 19:2278
119. Vitale A, Chrispeels MJ (1984) J Cell Biol 99:133
120. Kurosaka A, Yano A, Itoh N, Kuroda Y, Nakagawa T, Kawasaki T (1991) J Biol Chem 266:4168
121. Altmann F, Schwihla H, Staudacher E, Glössl J, März L (1995) J Biol Chem 270:17344-
122. Zhang W, Cao P, Chen S, Spence AM, Zhu S, Staudacher E, Schachter H (2003) Biochem J 372:53
123. Léonard R, Rendić D, Rabouille C, Wilson IBH, Préat T, Altmann F (2006) J Biol Chem 281:4867
124. Gutternigg M, Kretschmer-Lubich D, Paschinger K, Rendic D, Hader J, Geier P, Ranftl R, Jantsch V, Lochnit G, Wilson IBH (2007) J Biol Chem 282:27825
125. Jarvis DL (2003) Virology 310:1
126. Tomiya N, Narang S, Lee YC, Betenbaugh MJ (2004) Glycoconj J 21:343
127. Boquet I, Hitier R, Dumas M, Chaminade M, Préat T (2000) J Neurobiol 42:33
128. Rendić D, Linder A, Paschinger K, Borth N, Wilson IBH, Fabini G (2006) J Biol Chem 281:3343
129. Harnett W, Harnett MM (2001) Biochim Biophys Acta 1539:7
130. Furukawa K, Sato T (1999) Biochim Biophys Acta 1473:54
131. Palcic MM, Hindsgaul O (1991) Glycobiology 1:205
132. Khatra BS, Herries DG, Brew K (1974) Eur J Biochem 44:537
133. Brandan E, Fleischer B (1982) Biochemistry 21:4640
134. Milla ME, Clairmont CA, Hirschberg CB (1992) J Biol Chem 267:103
135. Brew K, Vanaman TC, Hill RL (1968) Proc Natl Acad Sci USA 59:491
136. Rajput B, Shaper NL, Shaper JH (1996) J Biol Chem 271:5131
137. Asano M, Furukawa K, Kido M, Matsumoto S, Umesaki Y, Kochibe N, Iwakura Y (1997) EMBO J 16:1850
138. Hansske B, Thiel C, Lubke T, Hasilik M, Honing S, Peters V, Heidemann PH, Hoffmann GF, Berger EG, von Figura K, Korner C (2002) J Clin Invest 109:725
139. Evans SC, Youakim A, Shur BD (1995) BioEssays 17:261
140. Hathaway HJ, Evans SC, Dubois DH, Foote CI, Elder BH, Shur BD (2003) J Cell Sci 116:4319
141. Babad H, Hassid WZ (1966) J Biol Chem 241:2672
142. Fujita-Yamaguchi Y, Yoshida A (1981) J Biol Chem 256:2701
143. Boeggeman EE, Balaji PV, Sethi N, Masibay AS, Qasba PK (1993) Protein Eng 6:779
144. Shibatani S, Fujiyama K, Nishiguchi S, Seki T, Maekawa (2001) J Biosci Bioeng 91:85
145. Borsig L, Berger EG, Malissard M (1997) Biochem Biophys Res Commun 240:586
146. Malissard M, Zeng S, Berger EG (2000) Biochem Biophys Res Commun 267:169
147. Bencúrová M, Rendić D, Fabini G, Kopecky EM, Altmann F, Wilson IBH (2003) Biochimie 85:413
148. Zhou D, Malissard M, Berger EG, Hennet T (2000) Arch Biochem Biophys 374:3
149. Guo S, Sato T, Shirane K, Furukawa K (2001) Glycobiology 11:813
150. Shaper NL, Shaper JH, Meuth JL, Fox JL, Chang H, Kirsch IR, Hollis GF (1986) Proc Natl Acad Sci USA 83:1573
151. Narimatsu H, Sinha S, Brew K, Okayama H, Qasba PK (1986) Proc Natl Acad Sci USA 83:4720
152. Gastinel LN, Cambillau C, Bourne Y (1999) EMBO J 18:3546
153. Masibay AS, Balaji PV, Boeggeman EE, Qasba PK (1993) J Biol Chem 268:9908
154. Kotani N, Asano M, Iwakura Y, Takasaki S (2001) Biochem J 357:827
155. Amado M, Almeida R, Carneiro F, Levery SB, Holmes EH, Nomoto M, Hollingsworth MA, Hassan H, Schwientek T, Nielsen PA, Bennett EP, Clausen H (1998) J Biol Chem 273:12770
156. Hennet T (2002) Cell Mol Life Sci 59:1081
157. Ramakrishnan B, Qasba PK (2007) J Mol Biol 365:570
158. Pöltl G, Kerner D, Paschinger K, Wilson IBH (2007) FEBS J 274:714
159. Mengeling BJ, Manzella SM, Baenziger JU (1995) Proc Natl Acad Sci USA 92:502

160. Manzella SM, Dharmesh SM, Beranek MC, Swanson P, Baenziger JU (1995) J Biol Chem 270:21665
161. Smith PL, Lowe JB (1994) J Biol Chem 269:15162
162. Tsuji S (1996) J Biochem 120:1
163. Harduin-Lepers A, Mollicone R, Delannoy P, Oriol R (2005) Glycobiology 15:805
164. Chou HH, Hayakawa T, Diaz S, Krings M, Indriati E, Leakey M, Paabo S, Satta Y, Takahata N, Varki A (2002) Proc Natl Acad Sci USA 99:11736
165. Weinstein J, Lee EU, McEntee K, Lai PH, Paulson JC (1987) J Biol Chem 262:17735
166. Takashima S, Tsuji S, Tsujimoto M (2002) J Biol Chem 277:45719
167. Chen C, Ma J, Lazic A, Backovic M, Colley KJ (2000) J Biol Chem 275:13819
168. Kitazume S, Tachida Y, Oka R, Shirotani K, Saido TC, Hashimoto Y (2001) Proc Natl Acad Sci USA 98:13554
169. Kono M, Ohyama Y, Lee YC, Hamamoto T, Kojima N, Tsuji S (1997) Glycobiology 7:469
170. Okajima T, Fukumoto S, Miyazaki H, Ishida H, Kiso M, Furukawa K, Urano T, Furukawa K (1999) J Biol Chem 274:11479
171. Nakayama J, Fukuda MN, Fredette B, Ranscht B, Fukuda M (1995) Proc Natl Acad Sci USA 92:7031
172. Takashima S, Ishida H, Inzau T, Ando T, Ishida H, Kiso M, Tsuji S, Tsujimoto M (2002) J Biol Chem 277:24030
173. El Maarouf A, Petridis AK, Rutishauser U (2006) Proc Natl Acad Sci USA 103:16989
174. Hennet T, Chui D, Paulson JC, Marth JD (1998) Proc Natl Acad Sci USA 95:4504
175. Priatel JJ, Chui D, Hiraoka N, Simmons CJT, Richardson KB, Page DM, Fukuda M, Varki NM, Marth JD (2000) Immunity 12:273
176. Weinhold B, Seidenfaden R, Rockle I, Muhlenhoff M, Schertzinger F, Conzelmann S, Marth JD, Gerardy-Schahn R, Hildebrandt H (2005) J Biol Chem 280:42971
177. Schwarzkopf M, Knobeloch KP, Rohde E, Hinderlich S, Wiechens N, Lucka L, Horak I, Reutter W, Horstkorte R (2002) Proc Natl Acad Sci USA 99:5267
178. Martinez-Duncker I, Dupré T, Piller V, Piller F, Candelier JJ, Trichet C, Tchernia G, Oriol R, Mollicone R (2005) Blood 105:2671
179. Ujita M, McAuliffe J, Hindsgaul O, Sasaki K, Fukuda MN, Fukuda M (1999) J Biol Chem 274:16717
180. Bierhuizen MFA, Mattei M-G, Fukuda M (1993) Genes Dev 7:468
181. Yeh J-C, Ong E, Fukuda M (1999) J Biol Chem 274:3215
182. Shimoda Y, Tajima Y, Nagase T, Harii K, Osumi N, Sanai Y (1999) J Biol Chem 274:17115
183. Bakker H, Friedmann I, Oka S, Kawasaki T, Nifant'ev N, Schachner M, Mantei N (1997) J Biol Chem 272:29942
184. Rouquier S, Lowe JB, Kelly RJ, Fertitta AL, Lennon GG, Giorgi D (1995) J Biol Chem 270:4632
185. Yamamoto F, Clausen H, White T, Marken J, Hakomori S (1990) Nature 345:229
186. Lowe JB (2002) Immunol Rev 186:19
187. Kukowska-Latallo JF, Larsen RD, Nair RP, Lowe JB (1990) Genes Dev 4:1288
188. Kudo T, Ikehara Y, Togayachi A, Kaneko M, Hiraga T, Sasaki K, Narimatsu H (1998) J Biol Chem 273:26729
189. Malý P, Thall A, Petryniak B, Rogers CE, Smith PL, Marks RM, Kelly RJ, Gersten KM, Cheng G, Saunders TL, Camper SA, Camphausen RT, Sullivan FX, Isogai Y, Hindsgaul O, von Andrian UH, Lowe JB (1996) Cell 86: 643
190. Lübke T, Marquardt T, Etzioni A, Hartmann E, von Figura K, Körner C (2001) Nature Genet 28:73
191. Oulmouden A, Wierinckx A, Petit J-M, Costache M, Palcic MM, Mollicone R, Oriol R, Julien R (1997) J Biol Chem 272:8764
192. Baboval T, Smith FI (2002) Mamm Genome 13:538
193. Rendić D, Klaudiny J, Stemmer U, Schmidt J, Paschinger K, Wilson IBH (2007) Biochem J 402:105
194. Kageyama N, Natsuka S, Hase S (1999) J Biochem 125:838
195. Ma B, Audette GF, Lin S, Palcic MM, Hazes B, Taylor DE (2006) J Biol Chem 281:6385
196. Okano M, Satoskar AR, Nishizaki K, Harn DA (2001) J Immunol 167:442
197. Duffy MS, Morris HR, Dell A, Appleton JA, Haslam SM (2006) Glycobiology 16:854
198. Galili U, Swanson K (1991) Proc Natl Acad Sci USA 88:7401
199. Koike C, Uddin M, Wildman DE, Gray EA, Trucco M, Starzl TE, Goodman M (2007) Proc Natl Acad Sci USA 104:559
200. Milland J, Christiansen D, Sandrin MS (2005) Immunol Cell Biol 83:687
201. Spiro RG (2002) Glycobiology 12:43R

202. Haltiwanger RS (2002) Curr Opin Struct Biol 12:593
203. Endo T, Toda T (2003) Biol Pharm Bull 26:1641
204. Wells L, Vosseller K, Hart GW (2001) Science 291:2376
205. Ten Hagen KG, Fritz TA, Tabak LA (2003) Glycobiology 13:1R
206. Pratt MR, Hang HC, Ten Hagen KG, Rarick J, Gerken TA, Tabak LA, Bertozzi CR (2004) Chem Biol 11:1009
207. Wilson IBH, Gavel Y, von Heijne G (1991) Biochem J 275:529
208. Tian E, Ten Hagen KG (2007) J Biol Chem 282:606
209. Topaz O, Shurman DL, Bergman R, Indelman M, Ratajczak P, Mizrachi M, Khamaysi Z, Behar D, Petronius D, Friedman V, Zelikovic I, Raimer S, Metzker A, Richard G, Sprecher E (2004) Nature Genet 36:579
210. Hennet T, Hagen FK, Tabak LA, Marth JD (1995) Proc Natl Acad Sci USA 92:12070
211. Ju T, Brewer K, D'Souza A, Cummings RD, Canfield WM (2002) J Biol Chem 277:178
212. Li GS, Zhang H, Lv JC, Shen Y, Wang HY (2007) Kidney Int 71:448
213. Xia L, Ju T, Westmuckett A, An G, Ivanciu L, McDaniel JM, Lupu F, Cummings RD, McEver RP (2004) J Cell Biol 164:451
214. Ju T, Cummings RD (2002) Proc Natl Acad Sci USA 99:16613
215. Ju T, Cummings RD (2005) Nature 437:1252
216. Müller R, Hulsmeier AJ, Altmann F, Ten Hagen K, Tiemeyer M, Hennet T (2005) FEBS J 272:4295
217. Ju T, Zheng Q, Cummings RD (2006) Glycobiology 16:947
218. Kojima N, Lee Y-C, Hamamoto T, Kurosawa N, Tsuji S (1994) Biochemistry 33:5772
219. Lee Y-C, Kaufmann M, Kitazume-Kawaguchi S, Kono M, Takashima S, Kurosawa N, Liu H, Pircher H, Tsuji S (1999) J Biol Chem 274:11958
220. Yeh J-C, Hiraoka N, Petryniak B, Nakayama J, Ellies LG, Rabuka D, Hindsgaul O, Marth JD, Lowe JB, Fukuda M (2001) Cell 105:957
221. Bierhuizen MF, Fukuda M (1992) Proc Natl Acad Sci USA 89:9326
222. Schwientek T, Yeh JC, Levery SB, Keck B, Merkx G, van Kessel AG, Fukuda M, Clausen H (2000) J Biol Chem 275:11106
223. Wilkins PP, McEver RP, Cummings RD (1996) J Biol Chem 271:18732
224. Ellies LG, Tsuboi S, Petryniak B, Lowe JB, Fukuda M, Marth JD (1998) Immunity 9:881
225. Tsuboi S, Fukuda M (2001) Bioessays 23:46
226. Warren CE, Krizus A, Partridge EA, Dennis JW (2002) Glycobiology 12:8G
227. Guerardel Y, Balanzino L, Maes E, Leroy Y, Coddeville B, Oriol R, Strecker G (2001) Biochem J 357:167
228. Iwai T, Inaba N, Naundorf A, Zhang Y, Gotoh M, Iwasaki H, Kudo T, Togayachi A, Ishizuka Y, Nakanishi H, Narimatsu H (2002) J Biol Chem 277:12802
229. Iwai T, Kudo T, Kawamoto R, Kubota T, Togayachi A, Hiruma T, Okada T, Kawamoto T, Morozumi K, Narimatsu H (2005) Proc Natl Acad Sci USA 102:4572
230. Wang Y, Shao L, Shi S, Harris RJ, Spellman MW, Stanley P, Haltiwanger RS (2001) J Biol Chem 276:40338
231. Luo Y, Koles K, Vorndam W, Haltiwanger RS, Panin VM (2006) J Biol Chem 281:9393
232. Moloney DJ, Panin VM, Johnston SH, Chen J, Shao L, Wilson R, Wang Y, Stanley P, Irvine KD, Haltiwanger RS, Vogt TF (2000) Nature 406:369
233. Luo Y, Nita-Lazar A, Haltiwanger RS (2006) J Biol Chem 281:9385
234. Kozma K, Keusch JJ, Hegemann B, Luther KB, Klein D, Hess D, Haltiwanger RS, Hofsteenge J (2006) 281:36742
235. Sato T, Sato M, Kiyohara K, Sogabe M, Shikanai T, Kikuchi N, Togayachi A, Ishida H, Ito H, Kameyama A, Gotoh M, Narimatsu H (2006) Glycobiology 16:1194
236. Shao L, Luo Y, Moloney DJ, Haltiwanger RS (2002) Glycobiology 12:763
237. Ishimizu T, Sano K, Uchida T, Teshima H, Omichi K, Hojo H, Nakahara Y, Hase S (2007) J Biochem 141:593
238. Shi S, Stanley P (2003) Proc Natl Acad Sci USA 100:5234
239. Sasamura T, Sasaki N, Miyashita F, Nakao S, Ishikawa HO, Ito M, Kitagawa M, Harigaya K, Spana E, Bilder D, Perrimon N, Matsuno K (2003) Development 130:4785
240. Chen J, Moloney DJ, Stanley P (2001) Proc Natl Acad Sci USA 98:13716
241. Sparrow DB, Chapman G, Wouters MA, Whittock NV, Ellard S, Fatkin D, Turnpenny PD, Kusumi K, Sillence D, Dunwoodie SL (2006) Am J Hum Genet 78:28
242. Appel B, Marasco P, McClung LE, Latimer AJ (2003) Dev Dyn 228:281

243. Van derVen BC, Harder JD, Crick DC, Belisle JT (2005) Science 309:941
244. Manya H, Chiba A, Yoshida A, Wang X, Chiba Y, Jigami Y, Margolis RU, Endo T (2004) Proc Natl Acad Sci USA 101:500
245. Ichimiya T, Manya H, Ohmae Y, Yoshida H, Takahashi K, Ueda R, Endo T, Nishihara S (2004) J Biol Chem 279:42638
246. Akasaka-Manya K, Manya H, Endo T (2004) Biochem Biophys Res Commun 325:75
247. Lyalin D, Koles K, Roosendaal SD, Repnikova E, Van Wechel L, Panin VM (2006) Genetics 172:343
248. Gentzsch M, Tanner W (1996) EMBO J 15:5752
249. Lussier M, Sdicu AM, Bussereau F, Jacquet M, Bussey H (1997) J Biol Chem 272:15527
250. Yoshida A, Kobayashi K, Manya H, Taniguchi K, Kano H, Mizuno M, Inazu T, Mitsuhashi H, Takahashi S, Takeuchi M, Herrmann R, Straub V, Talim B, Voit T, Topaloglu H, Toda T, Endo T (2001) Dev Cell 1:717
251. Martin PT (2006) Nat Clin Pract Neurol 2:222
252. Kolter T, Proia RL, Sandhoff K (2002) J Biol Chem 277:25859
253. Furukawa K, Tokuda N, Okuda T, Tajima O, Furukawa K (2004) Semin Cell Dev Biol 15:389
254. Nomura T, Takizawa M, Aoki J, Arai H, Inoue K, Wakisaka E, Yoshizuka N, Imokawa G, Dohmae N, Takio K, Hattori M, Matsuo N (1998) J Biol Chem 273:13570
255. Ishii A, Ohta M, Watanabe Y, Matsuda K, Ishiyama K, Sakoe K, Nakamura M, Inokuchi J, Sanai Y, Saito M (1998) J Biol Chem 273:31652
256. Yamashiro S, Haraguchi M, Furukawa K, Takamiya K, Yamamoto A, Nagata Y, Lloyd KO, Shiku H, Furukawa K (1995) J Biol Chem 270:6149
257. Furukawa K, Iwamura K, Uchikawa M, Sojka BN, Wiels J, Okajima T, Urano T, Furukawa K (2000) J Biol Chem 275:37752
258. Keusch JJ, Manzella SM, Nyame KA, Cummings RD, Baenziger JU (2000) J Biol Chem 275:25308
259. Togayachi A, Akashima T, Ookubo R, Kudo T, Nishihara S, Iwasaki H, Natsume A, Mio H, Inokuchi J, Irimura T, Sasaki K, Narimatsu H (2001) J Biol Chem 276:22032
260. Schulte S, Stoffel W (1993) Proc Natl Acad Sci USA 90:1026
261. Simpson MA, Cross H, Proukakis C, Priestman DA, Neville DC, Reinkensmeier G, Wang H, Wiznitzer M, Gurtz K, Verganelaki A, Pryde A, Patton MA, Dwek RA, Butters TD, Platt FM, Crosby AH (2004) Nature Genet 36: 1225
262. Yamashita T, Hashiramoto A, Haluzik M, Mizukami H, Beck S, Norton A, Kono M, Tsuji S, Daniotti JL, Werth N, Sandhoff R, Sandhoff K, Proia RL (2003) Proc Natl Acad Sci USA 100:3445
263. Jennemann R, Sandhoff R, Langbein L, Kaden S, Rothermel U, Gallala H, Sandhoff K, Wiegandt H, Gröne H-J (2007) J Biol Chem 285:3083
264. Jennemann R, Sandhoff R, Wang S, Kiss E, Gretz N, Zuliani C, Martin-Villalba A, Jager R, Schorle H, Kenzelmann M, Bonrouhi M, Wiegandt H, Gröne H-J (2005) Proc Natl Acad Sci USA 102:12459
265. Butters TD, Dwek RA, Platt FM (2005) Glycobiology 15:43R
266. Goode S, Melnick M, Chou TB, Perrimon N (1996) Development 122:3863
267. Wandall HH, Pizette S, Pedersen JW, Eichert H, Levery SB, Mandel U, Cohen SM, Clausen H (2005) J Biol Chem 280:4858
268. Helling F, Dennis RD, Weske B, Nores G, Peter-Katalinic J, Dabrowski U, Egge H, Wiegandt H (1991) Eur J Biochem 200:409
269. Seppo A, Moreland M, Schweingruber H, Tiemeyer M (2000) Eur J Biochem 267: 3549
270. Sasaki N, Yoshida H, Fuwa TJ, Kinoshita-Toyoda A, Toyoda H, Hirabayashi Y, Ishida H, Ueda R, Nishihara S (2007) Biochem Biophys Res Commun 354:522
271. Mucha J, Domlatil J, Lochnit G, Rendić D, Paschinger K, Hinterkörner G, Hofinger A, Kosma P, Wilson IBH (2004) Biochem J 382:67
272. Barrows BD, Haslam SM, Bischof LJ, Morris HR, Dell A, Aroian RV (2007) J Biol Chem 282:3302
273. Griffitts JS, Haslam SM, Yang T, Garczynski SF, Mulloy B, Morris H, Cremer PS, Dell A, Adang MJ, Aroian RV (2005) Science 307:922
274. Kawar ZS, Van Die I, Cummings RD (2002) J Biol Chem 277:34924
275. Lochnit G, Dennis RD, Ulmer AJ, Geyer R (1998) J Biol Chem 273:466
276. Ferguson MAJ, Homans SW, Dwek RA, Rademacher TW (1988) Science 239:753
277. Ikezawa H (2002) Biol Pharm Bull 25:409
278. Kinoshita T, Inoue N (2000) Curr Opin Chem Biol 4:632
279. Rosti V (2000) Haematologica 85:82

280. Nagamune K, Nozaki T, Maeda Y, Ohishi K, Fukuma T, Hara T, Schwarz R, Schneider P (2000) Proc Natl Acad Sci USA 97:10336
281. Smith TK, Crossman A, Brimacombe JS, Ferguson MAJ (2004) EMBO J 23:4701
282. Grimme SJ, Colussi PA, Taron CH, Orlean P (2004) Microbiology 150:3115
283. Funderburgh JL (2000) Glycobiology 10:951
284. Sugahara K, Mikami T, Uyama T, Mizuguchi S, Nomura K, Kitagawa H (2003) Curr Opin Struct Biol 13:612
285. Lin X (2004) Development 131:6009
286. Stuhlmeier KM (2006) Wien Med Wochenschr 156:563
287. Weissmann B, Meyer K, Sampson P, Linker A (1954) J Biol Chem 208:417
288. Muir H (1958) Biochem J 69:195
289. Wilson IBH (2004) Cell Mol Life Sci 61:794
290. Wilson IBH (2002) J Biol Chem 277:21207
291. Hwang H-Y, Olson S, Brown JR, Esko JD, Horvitz HR (2003) J Biol Chem 278:11735
292. Götting C, Kuhn J, Zahn R, Brinkmann T, Kleesiek K (2000) J Mol Biol 304:517
293. Cuellar K, Chuong H, Hubbell SM, Hinsdale ME (2007) J Biol Chem 282:5195
294. Voglmeir J, Voglauer R, Wilson IBH (2007) J Biol Chem 282:5984
295. Wang H, Spang A, Sullivan MA, Hryhorenko J, Hagen FK (2005) Mol Biol Cell 16:4202
296. Almeida R, Levery SB, Mandel U, Kresse H, Schwientek T, Bennett EP, Clausen H (1999) J Biol Chem 274:26165
297. Bai X, Zhou D, Brown JR, Crawford BE, Hennet T, Esko JD (2001) J Biol Chem 276:48189
298. Kitagawa H, Tone Y, Tamura J, Neumann KW, Ogawa T, Oka S, Kawasaki T, Sugahara K (1998) J Biol Chem 273:6615
299. Seidler DG, Faiyaz-Ul-Haque M, Hansen U, Yip GW, Zaidi SH, Teebi AS, Kiesel L, Gotte M (2006) J Mol Med 84:583
300. Kitagawa H, Shimakawa H, Sugahara K (1999) J Biol Chem 274:13933
301. Uyama T, Kitagawa H, Tamura JJ, Sugahara K (2002) J Biol Chem 277:8841
302. Kim BT, Kitagawa H, Tamura J, Saito T, Kusche-Gullberg M, Lindahl U, Sugahara K (2001) Proc Natl Acad Sci USA 98:7176
303. Zak BM, Crawford BE, Esko JD (2002) Biochim Biophys Acta 1573:346
304. Izumikawa T, Uyama T, Okuura Y, Sugahara K, Kitagawa H (2007) Biochem J 403:545
305. Hwang HY, Olson SK, Esko JD, Horvitz HR (2003) Nature 423:439
306. Kitagawa H, Izumikawa T, Mizuguchi S, Dejima K, Nomura KH, Egusa N, Taniguchi F, Tamura JI, Gengyo-Ando K, Mitani S, Nomura K, Sugahara K (2007) J Biol Chem 282:8533
307. Izumikawa T, Egusa N, Taniguchi F, Sugahara K, Kitagawa H (2006) J Biol Chem 281:1929
308. Takei Y, Ozawa Y, Sato M, Watanabe A, Tabata T (2004) Development 131:73
309. Kusche-Gullberg M, Kjellén L (2003) Curr Opin Struc Biol 13:605
310. Crawford BE, Olson SK, Esko JD, Pinhal MAS (2001) J Biol Chem 276:21538
311. Maccarana M, Olander B, Malmström J, Tiedemann K, Aebersold R, Lindahl U, Li JP, Malmström A (2006) J Biol Chem 281:11560
312. Yu H, Chen X (2007) Org Biomol Chem 5:865
313. Kamimura K, Fujise M, Villa F, Izumi S, Habuchi H, Kimata K, Nakato H (2001) J Biol Chem 276:17014
314. Kinnunen T, Huang Z, Townsend J, Gatdula MM, Brown JR, Esko JD, Turnbull JE (2005) Proc Natl Acad Sci USA 102:1507
315. Bink RJ, Habuchi H, Lele Z, Dolk E, Joore J, Rauch G-J, Geisler R, Wilson SW, den Hertog J, Kimata K, Zivkovic D (2003) J Biol Chem 278:31118
316. Klüppel M, Wight TN, Chan C, Hinek A, Wrana JL (2005) Development 132:3989
317. Akama TO, Nishida K, Nakayama J, Watanabe H, Ozaki K, Nakamura T, Dota A, Kawasaki S, Inoue Y, Maeda N, Yamamoto S, Fujiwara T, Thonar EJ, Shimomura Y, Kinoshita S, Tanigami A, Fukuda MN (2000) Nat Genet 26:237
318. Thiele H, Sakano M, Kitagawa H, Sugahara K, Rajab A, Höhne W, Ritter H, Leschik G, Nürnberg P, Mundlos S (2004) Proc Natl Acad Sci USA 101:10155
319. Zehe C, Engling A, Wegehingel S, Schäfer T, Nickel W (2006) Proc Natl Acad Sci USA 103:15479
320. Majewska-Sawka A, Nothnagel EA (2000) Plant Physiol 122:3
321. Konishi T, Takeda T, Miyazaki Y, Ohnishi-Kameyama M, Hayashi T, O'Neill MA, Ishii T (2007) Glycobiology 17:345
322. Lerouxel O, Cavalier DM, Liepman AH, Keegstra K (2006) Curr Opin Plant Biol 9:621
323. Somerville C (2006) Annu Rev Cell Dev Biol 22:53
324. Liepman AH, Wilkerson CG, Keegstra K (2005) Proc Natl Acad Sci USA 102:2221
325. Faik A, Price NJ, Raikhel NV, Keegstra K (2002) Proc Natl Acad Sci USA 99:7797

326. Cavalier DM, Keegstra K (2006) J Biol Chem 281:34197
327. Faik A, Bar-Peled M, DeRocher AE, Zeng W, Perrin RM, Wilkerson C, Raikhel NV, Keegstra K (2000) J Biol Chem 275:15082
328. Vanzin GF, Madson M, Carpita NC, Raikhel NV, Keegstra K, Reiter W-D (2002) Proc Natl Acad Sci USA 99:3340
329. Madson M, Dunand C, Li X, Verma R, Vanzin GF, Caplan J, Shoue DA, Carpita NC, Reiter W-D (2003) Plant Cell 15:1662
330. Zhong R, Peña MJ, Zhou GK, Nairn CJ, Wood-Jones A, Richardson EA, Morrison WH, Darvill AG, York WS, Ye Z-H (2005) Plant Cell 17:3390
331. Peña MJ, Zhong R, Zhou GK, Richardson EA, O'Neill MA, Darvill AG, York WS, Ye Z-H (2007) Plant Cell 19:549
332. Leboeuf E, Guillon F, Thoiron S, Lahaye M (2005) J Exp Bot 56:3171
333. Harholt J, Jensen JK, Sørensen SO, Orfila C, Pauly M, Scheller HV (2006) Plant Physiol 140:49
334. Sterling JD, Atmodjo MA, Inwood SE, Kumar Kolli VS, Quigley HF, Hahn MG, Mohnen D (2006) Proc Natl Acad Sci USA 103:5236
335. Iwai H, Hokura A, Oishi M, Chida H, Ishii T, Sakai S, Satoh S (2006) Proc Natl Acad Sci USA 103:16592
336. Coutinho PM, Deleury E, Davies GJ, Henrissat B (2003) J Mol Biol 328:307
337. Hidaka M, Honda Y, Kitaoka M, Nirasawa S, Hayashi K, Wakagi T, Shoun H, Fushinobu S (2004) Structure 12:937
338. Franco OL, Rigden DJ (2003) Glycobiology 13:707
339. Liu J, Mushegian A (2003) Protein Sci 12:1418
340. Breton C, Imberty A (1999) Curr Opin Struct Biol 9:563
341. Breton C, Šnajdrová L, Jeanneau C, Koča J, Imberty A (2006) Glycobiology 16:29R
342. Flint J, Taylor E, Yang M, Bolam DN, Tailford LE, Martinez-Fleites C, Dodson EJ, Davis BG, Gilbert HJ, Davies GJ (2005) Nat Struct Mol Biol 12:608
343. Davies GJ, Gloster TM, Henrissat B (2005) Curr Opin Struct Biol. 15:637
344. Charnock SJ, Davies GJ (1999) Biochemistry 38:6380
345. Vrielink A, Rüger W, Driessen HPC, Freemont PS (1994) EMBO J 13:3413
346. Breton C, Heissigerova H, Jeanneau C, Moravcova J, Imberty A (2002) Biochem Soc Symp 69:23
347. Yuan Y, Barrett D, Zhang Y, Kahne D, Sliz P, Walker S (2007) Proc Natl Acad Sci USA in press
348. Lovering AL, de Castro LH, Lim D, Strynadka NC (2007) Science 315:1402
349. Kubota T, Shiba T, Sugioka S, Furukawa S, Sawaki H, Kato R, Wakatsuki S, Narimatsu H (2006) J Mol Biol 359:708
350. Chiu CP, Watts AG, Lairson LL, Gilbert M, Lim D, Wakarchuk WW, Withers SG, Strynadka NC (2004) Nat Struct Mol Biol 11:163
351. Hu Y, Chen L, Ha S, Gross B, Falcone B, Walker D, Mokhtarzadeh M, Walker S (2003) Proc Natl Acad Sci USA 100:845
352. Ni L, Sun M, Yu H, Chokhawala H, Chen X, Fisher AJ (2006) Biochemistry 45:2139
353. Ihara H, Ikeda Y, Toma S, Wang X, Suzuki T, Gu J, Miyoshi E, Tsukihara T, Honke K, Matsumoto A, Nakagawa A, Taniguchi N (2007) Glycobiology 17:455
354. Pak JE, Arnoux P, Zhou S, Sivarajah P, Satkunarajah M, Xing X, Rini JM (2006) J Biol Chem 281:26693
355. Boix E, Zhang Y, Swaminathan GJ, Brew K, Acharya KR (2002) J Biol Chem 277:28310
356. Pedersen LC, Darden TA, Negishi M (2002) J Biol Chem 277:21869
357. Persson K, Ly HD, Dieckelmann M, Wakarchuk WW, Withers SG, Strynadka NCJ (2001) Nature Struct Biol 8:166
358. Ramakrishnan B, Balaji PV, Qasba PK (2002) J Mol Biol 318:491
359. Gastinel LN, Bignon C, Misra AK, Hindsgaul O, Shaper JH, Joziasse DH (2001) EMBO J. 20:638
360. Patenaude SI, Seto NO, Borisova SN, Szpacenko A, Marcus SL, Palcic MM, Evans SV (2002) Nature Struct. Biol. 9:685
361. Ramasamy V, Ramakrishnan B, Boeggeman E, Ratner DM, Seeberger PH, Qasba PK (2005) J Mol Biol 353:53
362. Gibbons BJ, Roach PJ, Hurley TD (2002) J Mol Biol 319:463
363. Ünligil UM, Zhou S, Yuwaraj S, Sarkar M, Schachter H, Rini JM (2000) EMBO J 19:5269
364. Lobsanov YD, Romero PA, Sleno B, Yu B, Yip P, Herscovics A, Howell PL (2004) J. Biol. Chem. 279:17921
365. Fritz TA, Hurley JH, Trinh LB, Shiloach J, Tabak LA (2004) Proc Natl Acad Sci USA 101:15307
366. Fritz TA, Raman J, Tabak LA (2006) J Biol Chem 281:8613

367. Jinek M, Chen YW, Clausen H, Cohen SM, Conti E (2006) Nat Struct Mol Biol 13:945
368. Pedersen LC, Tsuchida K, Kitagawa H, Sugahara K, Darden TA, Negishi M (2000) J Biol Chem 275:34580
369. Kakuda S, Shiba T, Ishiguro M, Tagawa H, Oka S, Kajihara Y, Kawasaki T, Wakatsuki S, Kato R (2004) J Biol Chem 279:22693
370. Reinert DJ, Jank T, Aktories K, Schulz GE (2005) J Mol Biol 351:973
371. Pedersen LC, Dong J, Taniguchi F, Kitagawa H, Krahn JM, Pedersen LG, Sugahara K, Negishi M (2003) J Biol Chem 278:14420
372. Hassan H, Reis CA, Bennett EP, Mirgorodskaya E, Roepstorff P, Hollingsworth MA, Burchell J, Taylor-Papadimitriou J, Clausen H (2000) J Biol Chem 275:38197
373. Tenno M, Saeki A, Kezdy FJ, Elhammer AP, Kurosaka A (2002) J Biol Chem 277:47088
374. Nishikawa Y, Pegg W, Paulsen H, Schachter H (1988) J Biol Chem 263:8270
375. Qasba PK, Ramakrishnan B, Boeggeman E (2005) Trends Biochem Sci 30:53
376. Mulichak AM, Losey HC, Walsh CT, Garavito RM (2001) Structure 9:547
377. Hu Y, Walker S (2002) Chem Biol 9:1287
378. Moréra S, Larivière L, Kurzeck J, Aschke-Sonnenborn U, Freemont PS, Janin J, Ruger W (2001) J Mol Biol 311:569
379. Mulichak AM, Losey HC, Lu W, Wawrzak Z, Walsh CT, Garavito RM (2003) Proc Natl Acad Sci USA 100:9238
380. Mulichak AM, Lu W, Losey HC, Walsh CT, Garavito RM (2004) Biochemistry 43:5170
381. Shao H, He X, Achnine L, Blount JW, Dixon RA, Wang X (2005) Plant Cell 17:3141
382. Offen W, Martinez-Fleites C, Yang M, Kiat-Lim E, Davis BG, Tarling CA, Ford CM, Bowles DJ, Davies GJ (2006) EMBO J 25:1396
383. Martinez-Fleites C, Proctor M, Roberts S, Bolam DN, Gilbert HJ, Davies GJ (2006) Chem Biol 13:1143
384. Buschiazzo A, Ugalde JE, Guerin ME, Shepard W, Ugalde RA, Alzari PM (2004) EMBO J 23:3196
385. Horcajada C, Guinovart JJ, Fita I, Ferrer JC (2006) J Biol Chem 281:2923
386. Grizot S, Salem M, Vongsouthi V, Durand L, Moreau F, Dohi H, Vincent S, Escaich S, Ducruix A (2006) J Mol Biol 363:383
387. Gibson RP, Turkenburg JP, Charnock SJ, Lloyd R, Davies GJ (2002) Chem Biol 9:1337
388. Ha S, Walker D, Shi Y, Walker S (2000) Protein Sci 9:1045
389. Larivière L, Sommer N, Moréra S (2005) J Mol Biol 352:139
390. Sun HY, Lin SW, Ko TP, Pan JF, Liu CL, Lin CN, Wang AH, Lin CH (2007) J Biol Chem 282:9973
391. Wrabl J, Grishin N (2001) J Mol Biol 314:365
392. Breton C, Oriol R, Imberty A (1998) Glycobiology 8:87
393. Martinez-Duncker I, Mollicone R, Candelier JJ, Breton C, Oriol R (2003) Glycobiology 13:1C
394. Larivière L, Moréra S (2004) J Biol Chem 279:34715
395. Godzik A (2003) Methods Biochem Anal 44:525
396. Wimmerová M, Engelsen SB, Bettler E, Breton C, Imberty A (2003) Biochimie 85:691
397. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Nucleic Acids Res 25:3389
398. Eddy SR (1998) Bioinformatics 14:755
399. Kikuchi N, Kwon YD, Gotoh M, Narimatsu H (2003) Biochem Biophys Res Commun 310: 574
400. Petrova P, Monteiro C, Herve du Penhoat C, Koča J, Imberty A (2001) Biopolymers 58:617
401. Heissigerová H, Breton C, Moravcová J, Imberty A (2003) Glycobiology 13:377
402. Botté C, Jeanneau C, Šnajdrová L, Bastien O, Imberty A, Breton C, Maréchal E (2005) J Biol Chem 280:34691
403. Hans J, Brandt W, Vogt T (2004) Plant J 39:319
404. Thorsøe KS, Bak S, Olsen CE, Imberty A, Breton C, Møller BL (2005) Plant Physiol. 139:664
405. Sinnott ML (1990) Chem Rev 90:1171
406. Legler G (1993) Carbohydr Res 250:vii
407. Withers SG, Aebersold R (1995) Protein Sci 4:361
408. Davies G, Sinnott ML, Withers SG (1998) In: Sinott ML (ed), Comprehensive biological catalysis: a mechanistic reference, chapter 3. Academic Press, New York, p 119
409. Zechel DL, Withers SG (1999) In: Poulter CD (ed), Mechanisms of glycosyl transfer, Comprehensive Natural Products Chemistry, chapter 5, Elsevier, New York, p 279
410. Zechel DL, Withers SG. (2000) Acc Chem Res 33:11
411. Pedersen LC, Darden TA, Yoshimitsu K, Negishi M (2001) Trends Glycosci Glycotech 13:121

412. Lairson LL, Withers SG (2004) Chem Commun 2243
413. Gao DQ, Pan YK, Byun K, Gao JL (1998) J Am Chem Soc 120:4045
414. Maseras F, Morokuma K (1995) J Comp Chem 16:1170
415. Gao J, Truhlar DG (2000) Ann Rev Phys Chem 53:467
416. Garcia-Viloca M, Gao J, Karplus M, Truhlar DG (2004) Science 303 :186
417. Truhlar DG, Gao J, Alhambra C, Garcia-Viloca M, Corchado J, Sanchez ML, Villa J (2002) Acc Chem Res 35:341
418. Tvaroška I, André I, Carver JP (2000) J Am Chem Soc 122:8762
419. Tvaroška I, André I, Carver JP (2003) Glycobiology 13:559
420. Kozmon S, Tvaroška I (2007) J Am Chem Soc 51:16921
421. Frey PA, Whitt SA, Tobin JB (1994) Science 264:1927
422. Schachter H, Reck F, Paulsen H (2003) Method Enzymol 363:459
423. Ramakrishnan B, Ramasamy V, Qasba PK (2006) J Mol Biol 357:1619
424. Seto NOL, Compston CA, Evans SV, Bundle DR, Narang SA, Palcic MM (1999) Eur J Biochem 259:770
425. Murray BW, Takayama S, Schultz J, Wong CH (1996) Biochemistry 35:11183
426. Kim SC, Singh AN, Raushel FM (1988) J Biol Chem 263:10151
427. Kim SC, Singh AN, Raushel FM (1988) Arch Biochem Biophys 267:54
428. Murray BW, Wittmann V, Burkart MD, Hung SC, Wong CH (1997) Biochemistry 36:823
429. Werner RM, Stivers JT (2000) Biochemistry 39:14054
430. Lewandowicz A, Schramm VL (2004) Biochemistry 43:1458
431. André I, Tvaroška I, Carver JP (2003) Carbohydr Res 338:865
432. Ly HD, Lougheed B, Wakarchuk WW, Withers SG (2002) Biochemistry 41:5075
433. Lairson LL, Chiu CPC, Ly HD, He S, Wakarchuk WW, Strynadka NCJ, Withers SG (2004) J Biol Chem 279:28339
434. Tvaroška I (2004) Carbohydr Res 339:1007
435. Tvaroška I (2005) Trends Glycoscience Glycotechnology 17:177
436. Tvaroška I (2006) In: Vliegenthart JFG, Woods RJ (eds), NMR mpectroscopy and computer modeling of carbohydrates (ACS Symposium Series), vol 930. American Chemical Society, Washington DC, p 285
437. Look GC, Fotsch CH, Wong CH (1993) Acc Chem Res 26:182
438. Sears P, Wong CH (1996) Proc Natl Acad Sci USA 93:12086
439. Wang R, Steensma DH, Takaoka Y, Yun JW, Kajimoto T, Wong CH (1997) Bioorg Med Chem 5:661
440. Sears P, Wong CH (1999) Angew Chem Int Ed Engl 38:2300
441. Elhalabi JM, Rice KG (1999) Curr Med Chem 6:93
442. Waldscheck B, Streiff M, Notz W, Kinzy W, Schmidt RR (2001) Angew Chem Int Ed Engl 40:4007
443. Compain P, Martin OR (2001) Bioorg Med Chem 9:3077
444. Compain P, Martin OR (2003) Curr Med Chem 3:541
445. Raab M, Kozmon S, Tvaroška I (2005) Carbohydr Res 340:1007
446. Kozmon S, Tvaroška I unpublished results.

11.3 Polysaccharide Degradation

Bruce A. Stone[1], *Birte Svensson*[2], *Michelle E. Collins*[3], *Robert A. Rastall*[3]
[1] Department of Biochemistry, La Trobe University, Bundoora, VIC 3083, Australia
[2] Department of Chemistry, Carlsberg Laboratory, 2500 Valby, Denmark
[3] Department of Food Biosciences, The University of Reading, Reading RG6 6AP, UK
b.stone@latrobe.edu.au, m.e.collins@reading.ac.uk, r.a.rastall@reading.ac.uk

In: *Glycoscience*. Fraser-Reid B, Tatsuta K, Thiem J (eds)
Chapter-DOI 10-1007/978-3-540-30429-6_60:

Abstract

An overview of current and potential enzymes used to degrade polysaccharides is presented. Such depolymerases are comprised of glycoside hydrolases, glycosyl transferases, phosphorylases and lyases, and their classification, active sites and action patterns are discussed. Additionally, the mechanisms that these enzymes use to cleave glycosidic linkages is reviewed as are inhibitors of depolymerase activity; reagents which react with amino acid residues, glycoside derivatives, transition state inhibitors and proteinaceous inhibitors. The characterization of various enzymes of microbial, animal or plant origin has led to their widespread use in the production of important oligosaccharides which can be incorporated into food stuffs. Sources of polysaccharides of particular interest in this chapter are those from plants and include inulin, dextran, xylan and pectin, as their hydrolysis products are purported to be functional foods in the context of gastrointestinal health. An alternative use of degraded polysaccharides is in the treatment of disease. The possibility exists to treat bacterial exopolysaccharide with lyases from bacteriophage to produce oligosaccharides exhibiting bioactive sequences. Although this area is currently in its infancy the knowledge is available to investigate further.

Keywords

Action patterns; Adhesion domains; Affinity and mechanism-based inhibitors; Catalytic domains; Classification schemes; Hydrolases; Lyases; Mechanism of action; Phosphorylases; Proteinaceous inhibitors

1 Introduction

Polysaccharides, particularly cellulose and chitin, are the most abundant carbon compounds in the biosphere. They are functionally important as structural and/or protecting materials and are the predominant storage form of carbohydrate in bacteria, protistans, algae, fungi, plants, and animals. They are also important precursors for the enzymatic manufacture of a range of bioactive oligosaccharides with application in nutrition and disease management. In this chapter enzymes involved in polysaccharide depolymerization and biosynthesis will be reviewed with special reference to their specificity, mechanism of action, and structural organization. Applications of such enzymes to the manufacture of bioactive oligosaccharides will also be discussed.

Enzymes with over one hundred different specificities have been recognized to be involved in the reactions of polysaccharide degradation and synthesis. It has not been possible to address each of them. However, particular enzymes or groups of enzymes are discussed to highlight important principles or aspects.

The first section of this chapter deals with enzymes depolymerizing polysaccharides. The reactions they catalyze are essential steps in the provision of assimilable mono- or oligosaccharides from exogenous polysaccharide substrates. The breakdown products are substrates for metabolism and energy production by bacteria, archaeans, single-celled eukaryotes, animals, and fungi. Structural polysaccharides are enzymically degraded in the course of morphogenetic changes, when cells or tissues grow and differentiate, for example, during plant cell wall modification or in the turnover of vertebrate connective tissue and arthropod exoskeletons. Polysaccharide depolymerizing enzymes produced by bacterial, fungal, insect, and nematode pathogens of plants or animals are used to gain access to host tissues. Polysaccharide hydrolases and lyases find many applications in industry, particularly the food industry. These will not all be reviewed here but the potential for manufacture of functional oligosaccharides by enzymatic degradation of polysaccharides will be discussed.

2 Polysaccharide Depolymerizing Enzymes

2.1 Introduction

Glycoside-cleaving enzymes are a diverse group, which more or less selectively, hydrolyze the wide variety of *O*-glycosidic linkages found in glycosides, oligo- and polysaccharides, and glycoconjugates. An account of one group of glycoside hydrolases; those acting on oligogalactosides, is to be found in ❯ Chap. 5.4. This section deals specifically with enzymes depolymerizing interglycosidically, *O*-linked polysaccharides, although, where relevant, information from *O*-glycoside-cleaving enzymes is discussed.

The chemistry of enzymic glycoside cleavage is a focus of attention for carbohydrate chemists and biochemists, and a considerable body of mechanistic information has accumulated, especially with regard to glycoside and polysaccharide hydrolyzing enzymes [1,2,3,4,5,6]. In parallel, there has been an explosion of DNA cloning and direct amino acid sequence determination for glycosidases [7]. In addition, the three-dimensional structures of a growing number of glycoside and polysaccharide-degrading enzymes have been determined.

2.2 Types of Glycoside Depolymerizing Enzymes

Enzyme-catalyzed glycoside cleavages (and syntheses) are considered by Hehre [6,8] to be variants of the same reaction (❯ *Eq. 1*), catalyzed by a single type of enzyme, the glycosylases, comprising glycoside hydrolases, glycosyl transferases, phosphorylases and certain lyases.

$$\text{Glycosyl} - \text{X} + \text{H} - \text{X}' = \text{Glycosyl} - \text{X}' + \text{H} - \text{X} \qquad (1)$$

This description reduces all carbohydrase reactions to an interchange between a glycosyl group and a proton (hydrogen atom) about sites X and X′ in the various potential substrates.

Thus, X may be a saccharide, a 1,2-glycal, a phosphate residue, a nucleoside diphosphate, or a polyprenol diphosphate and X′ the OH of water, an alcohol, and often a saccharide. It is convenient, however, to consider enzymes cleaving polysaccharide chains in one of four classes: hydrolases, glycosyl transferases, lyases, and phosphorylases. These follow the general reactions discussed below.

2.2.1 Hydrolysis

$$\text{Glycosyl} - \text{OR(polysaccharide)} + \text{HOH} = \text{Glycosyl} - \text{OH} + \text{HOR} \tag{2}$$

Some polysaccharide hydrolases also catalyze glycosyl transfer.

2.2.2 Glycosyl Transfer

The glycosyl transferases catalyze redistribution of glycosidic linkages between di-, oligo-, or polysaccharide donor substrates and acceptor oligo- or polysaccharides.

$$\begin{aligned}\text{Glycosyl} - \text{OR}'\text{(di-, oligo- or polysaccharide)} + \text{HOR}''\text{(oligo- or polysaccharide)} \\ = \text{Glycosyl} - \text{OR}'' + \text{HOR}\end{aligned} \tag{3}$$

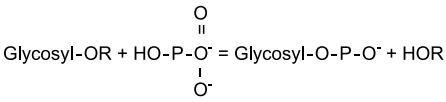
They are considered separately from the polysaccharide hydrolases, but are closely related to them, both mechanistically and structurally. In glycosyl transfer the acceptor water is replaced by a hydroxy group provided by an oligo- or polysaccharide.

2.2.3 Lyase Action (Glycal Formation)

$$\text{Glycosyl} - \text{OR} = 1,2 - \text{Glycal} + \text{HOR} \tag{4}$$

2.2.4 Phosphorolysis

Glycosidic linkages in polysaccharides may also be cleaved by phosphorolysis.
The phosphorolysis reaction is shown in ❯ *Structure 1*.



Structure 1

2.3 Criteria for Classification

2.3.1 Specificity

The Nomenclature Committee of IUB [9] classifies polysaccharide depolymerizing enzymes on the basis of the reaction catalyzed, into *O*-glycoside hydrolases [EC 3.2.1.-], glycosyl transferases [EC 2.4.1.-], lyases [EC 4.2.2.-], and the phosphorylases in [EC 2.4.1.-]. Within each

group, the individual enzymes are sub-classified according to their substrate and product specificity. The Enzyme Nomenclature list is revised continually as new enzymes are discovered and previously included enzymes are better characterized.
Each sub-category of polysaccharide depolymerizing enzymes shows specificity with respect to the configuration of the bond in the glycosidic linkage cleaved, the ring size and configuration of the glycone portion of the substrate, and most often the "aglycone" portion of the substrate. The regioselectivity for the linkage cleaved may in some cases be narrow and in others quite broad. There are also more subtle distinctions between individual enzymes of the same sub-class. Thus differences may occur in the region of the polymeric substrate attacked, i. e., the non-reducing end, the reducing end, or the interior region, the size and state of solubility of the preferred substrate [10], or the structure of the products formed. In addition, and quite fundamentally, the classification in some cases distinguishes different *O*-glycosidases on the basis of the mechanism of cleavage of the glycosidic linkage.

2.3.2 Structure

An additional classification has been introduced that disposes polysaccharide depolymerizing enzymes into glycosidase families based on their amino acid sequence relationships, made possible by direct sequencing or derived from the numerous cloned cDNA or gene sequences for the enzymes [7,11,12,13]. At present, this classification is based on about 5000 known sequences and includes 82 hydrolase, 48 glycosyl transferase families (including phosphorylases) and the ten lyase families. The determination of the three-dimensional structures, at high resolution, of a growing number of these enzymes has permitted amalgamation of some hydrolase and glycosyl transferase families into clans. Members of a clan have common structural features with respect to overall fold motifs, topology of active site residues and catalytic mechanisms, even when the enzymes are without recognized sequence identity [14].

2.4 Active Sites

The catalysis of glycoside cleavage occurs when a region of the polymeric substrate chain is bound at the active site of the enzymes to form an enzyme-substrate complex. The active sites of all types of enzymes involved in polysaccharide depolymerization are formed by catalytic amino acid residues that are directly involved in the covalent bond-breaking or -making processes and by amino acids involved in binding of substrate along an extended binding site. The former category comprises aspartic and glutamic acids in the hydrolases and glycosyl transferases [4,13] and probably arginine in at least some lyases [15]. The typical binding residues can be polar: arginine, histidine, lysine, aspartic and glutamic acids, asparagine, glutamine, serine, threonine or aromatic: tyrosine, phenylalanine, and tryptophan [4,16,17].
The substrate-binding region of the enzyme surface can be envisaged as a series of consecutively arranged binding subsites, complementary to monosaccharide residues in the substrate chain. These are often distributed asymmetrically about the site of glycosidic bond cleavage. In a widely accepted nomenclature [18] describing the monosaccharide-binding subsites in polysaccharide hydrolases, but applying equally well to glycosyl transferases, lyases, and phosphorylases, the subsites are labeled from $-n$ to $+n$ (where n is an integer) and where n

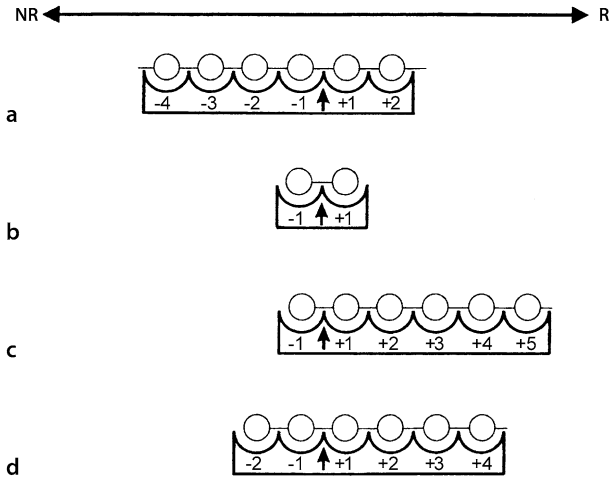


Figure 1
Schematic drawing of monosaccharide binding sites in various glycosyl hydrolases. (a) Subsites of an *endo*-hydrolase, e. g., hen egg white lysozyme, (b) subsites of a disaccharide hydrolase, e. g., ⟨α-glucosidase, (c) subsites of a monosaccharide-releasing *exo*-hydrolase, e. g., glucoamylase, (d) subsites of a disaccharide-releasing *exo*-hydrolase, e. g., β-amylase (after [18])

represents the non-reducing end and $+n$ the reducing end, with cleavage taking place between subsites -1 and $+1$ (*Fig. 1*) [18].

The physical form of the binding subsites can be seen directly from crystallographic analyses of the enzymes in complex with non-reactive substrate analogues or of inactive, mutated enzymes in complex with substrates. This allows the identification of those amino acid residues that are involved in hydrogen bond, or ionic and van der Waals interactions with the bound substrate [19]. This approach was first used for hen egg white lysozyme (reviewed in [20]) and has since been widely applied to other hydrolases, glycosyl transferases, polyuronate lyases, and phosphorylases. Information on stoichiometry and binding thermodynamics, binding kinetics, and conformational changes for substrate analogues can be obtained from solution studies using equilibrium dialysis or gel filtration, surface plasmon resonance and isothermal titration calorimetry, time-resolved fluorescence or transfer nOe NMR spectroscopy, stopped-flow spectrophotometric and temperature jump relaxation methods [21,22,23]. Models of enzyme-substrate complexes can be constructed from the three-dimensional structures using docking procedures involving conformational energy calculations and molecular dynamics. They provide useful insights into the relationship between structure and function [23,24,25,26].

Characterization of the extended substrate-binding site requires information on the number of subsites, the individual subsite binding affinities, bond cleavage frequencies and the location of the catalytic site, which can be determined by analyzing the amount and types of product formed from terminally labeled oligosaccharides [27,28]. The possibility of side reactions such as transglycosylation, and the validity of assumptions that the subsites are independent and that the intrinsic bond cleavage rate is constant, have been addressed in recent applications of the theory [29,30]. In the multiple subsite model it is envisaged that a single substrate and enzyme molecule react at one time, however, it is known that multimolecular substrate reac-

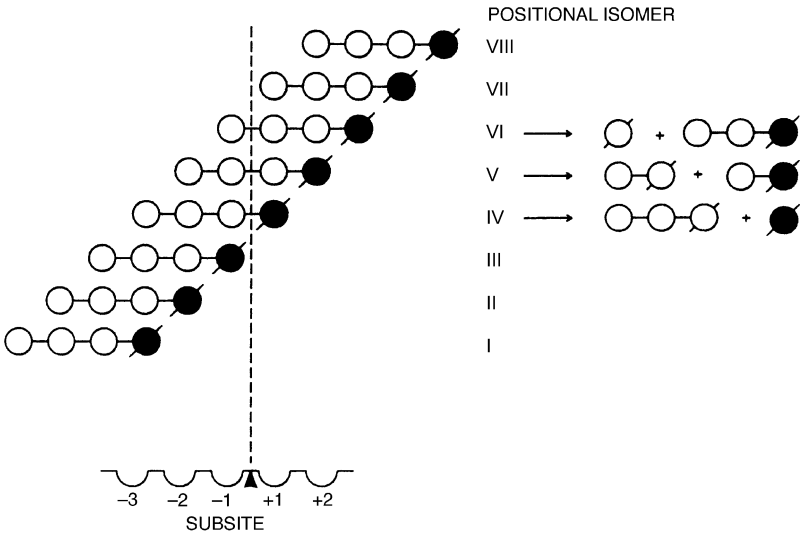


Figure 2
Positional isomers of a tetrasaccharide on a five subsite endo-hydrolase. Of the eight possible positional isomers only IV, V, and VI are productive, i. e., lie over the catalytic site. The cleavage products are shown: subsite on enzyme; position of catalytic site; monosaccharide residue; residue bearing the reducing hemiacetal in substrate; residue bearing the reducing hemiacetal in products (after [28])

tions may occur. These reactions involve, e. g., facilitated binding of one substrate molecule elicited by binding of another. Practical methods are available to assess and minimize these complications [30,31,32].

In the multiple subsite model (*Fig. 2*) [28] an oligosaccharide substrate can interact at the subsites in a number of ways. Thus, a tetrasaccharide may form eight positional isomers with a five subsite enzyme, but only the three in which substrate lies over the catalytic site can be productive. The population of the different complexes is dictated, in principle, by the energetics of interaction of the substrate monomer units with the respective subsites. Each depolymerase has its own subsite map for oligosaccharide substrates, depicting the number of subsites, the location of the catalytic site, and the binding energies contributed by the individual subsites. The latter can be calculated from the bond cleavage frequencies and Michaelis parameters determined for a range of oligosaccharide substrate lengths. A subsite map for two porcine pancreatic α-amylases deduced from kinetic analysis is shown in *Fig. 3* [33].

For this enzyme, kinetic evidence shows several weak oligosaccharide-binding sites in addition to the active site [34], in agreement with the observation that oligosaccharides in the crystal are also bound outside the active site area [35]. Thus, it is important to note that kinetically derived subsite maps may only qualitatively reflect key features of the three-dimensional structure of enzyme-substrate complexes and vice versa. However, the length of the binding site and the position of the catalytic site, as detected from crystallographic or modeled enzyme-ligand complexes generally agree with the subsite map [24,36]. Moreover, conspicuous aromatic stacking onto substrate rings, or H-bonds between charged groups of the enzyme and substrate OH groups tend to reflect high subsite affinity [19].

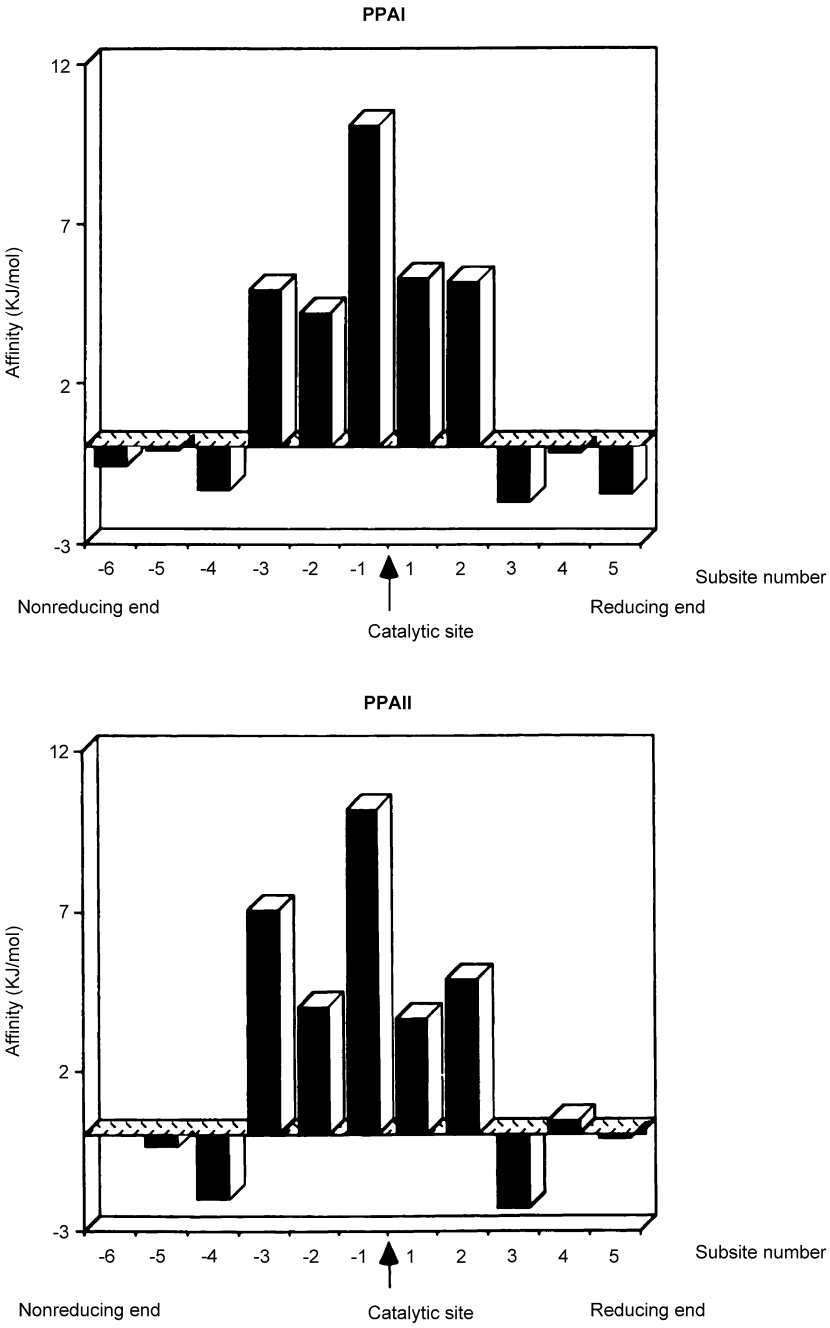


Figure 3
Porcine pancreatic α-amylase I and II energetic profiles. *Arrow* indicates the catalytic site (from [33])

2.5 Action Patterns

2.5.1 *Endo*-Action

In *endo*-cleavage the enzyme first attacks at an interior glycosidic linkage in the substrate. The *endo*-hydrolases are widely represented among the polysaccharide hydrolases and lyases. The cleavage may represent a single productive encounter with the substrate, so that both polymer fragments are released from the enzyme. This action pattern has been described as *multi-chain* attack (◉ *Fig. 4*) [37].

In other *endo*-cleavages several catalytic events may occur after the first encounter in which one polymer fragment is retained and translocated along the substrate binding site to allow

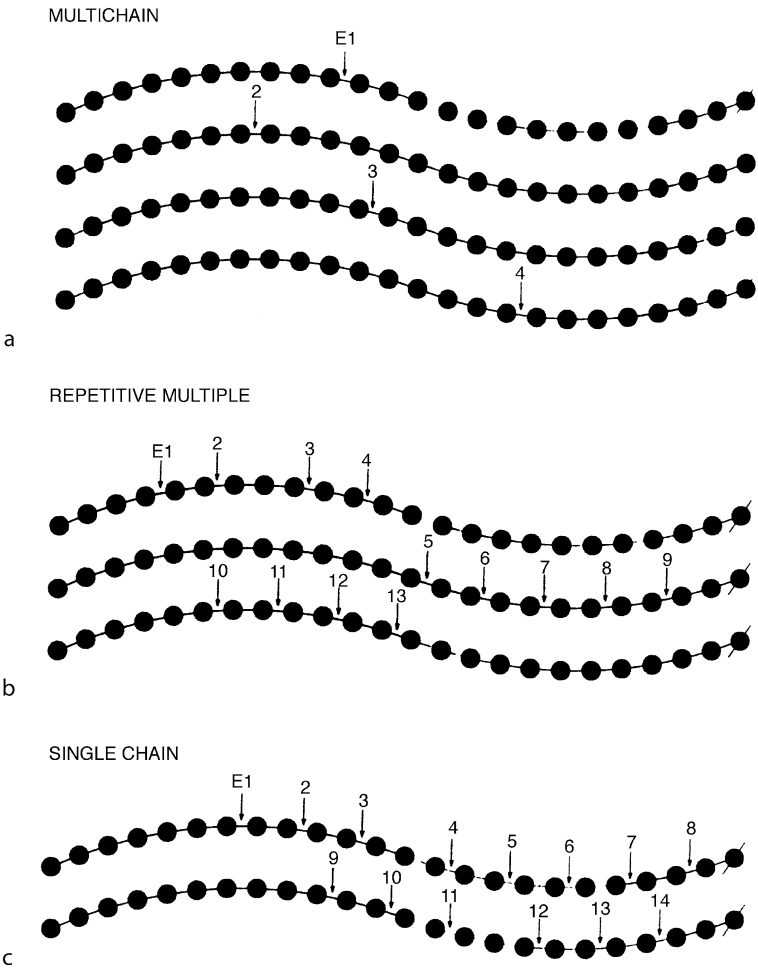


Figure 4
Types of attack patterns for polysaccharide *endo*-hydrolases (after [37])

further cleavages at the catalytic site. This mode of attack is known as *repetitive* or *multiple* attack and the action involves substrate *processivity*. Multiple attack has also been described as *endo-* and *exo*-action (see below). It was first demonstrated for α-amylases [37] but has also been demonstrated for several cellulolytic enzymes [38,39]. In the extreme case, if the enzyme-substrate complex does not dissociate, the bound polysaccharide substrate would degrade to oligosaccharides in a *single chain* attack. However, examples of this mode of attack are not known for polysaccharide degrading enzymes. The degree of multiple attack (DMA) is described by the number of cleavages following the first cleavage of the polymeric substrate. A quantitative analysis of DMA has been developed for α-amylases [37], based on the ratio between the number of oligomeric and polymeric molecules produced from amylose, a linear (1→4)-α-glucan substrate. Average values of DMA vary from 1 through 6 for Taka-amylase A, barley α-amylase, and porcine pancreatic α-amylase [37,40]. The method of Robyt and French [37] provides no information on the chain polarity of the multiple attack, i. e., which cleavage product is retained in the active site. This may be deduced either from the knowledge of the enzyme mechanism or from experiments with end-labeled oligosaccharide substrates. For *endo*-acting enzymes with a retaining mechanism one may assume that, after

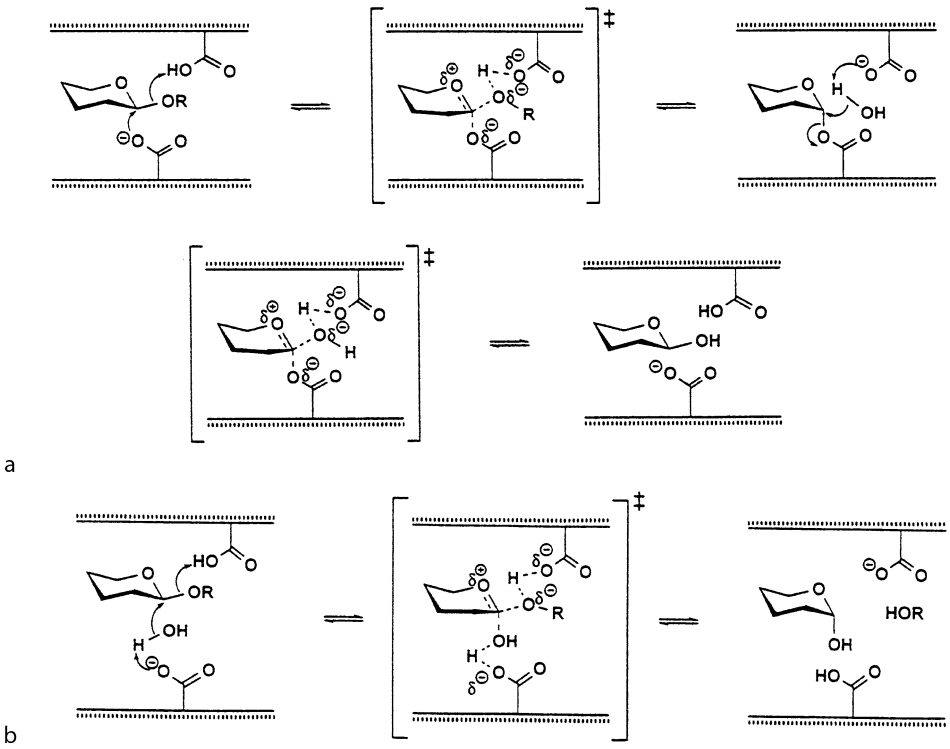


Figure 5
(a) Proposed catalytic mechanism for a retaining β-glycosidase, (b) proposed catalytic mechanism for an inverting β-glycosidase

release of the leaving product, the covalently bound glycosyl intermediate released by attack of water (Fig. 5a) is subsequently translocated in the active site for the next cleavage in the multiple attack reaction. It has been noted that the type of attack may change as the length of the substrate chains decreases during the course of the depolymerization or be dependent on reaction conditions, e. g., pH [41].

2.5.2 *Exo*-Action

Exo-action is well-recognized among the (1→4)-α-, (1→6)-α-, (1→3)-β-, and (1→4)-β-glucan hydrolases, amylolyases, and glycogen and starch phosphorylases. *Exo*-attack generally occurs from the non-reducing end of the substrate chain and is blocked by modification of the terminal residue, e. g., by periodate oxidation in the case of a (1→3)-β-glucan substrate [42,43,44]. *Exo*-cleaving enzymes, depending on the source, may release mono-, di-, or longer oligosaccharides, up to octasaccharides, from the polysaccharide substrate [45,46,47,48,49,50,51,52,53]. The active sites of *exo*-enzymes can also be conveniently described using the now generally accepted nomenclature (see Fig. 1). Thus, the active site of an *exo*-hydrolase, releasing disaccharides, is described as −2, −1, +1, +2, +3, indicating that two glycosyl binding subsites, −2 and −1, accommodate the future product, and that three binding subsites, +1, +2, and +3, interact with glycosyl residues toward the reducing end of the substrate.

Exo-attack may follow the same action patterns as described for *endo*-attack. Thus, β-amylases catalyze the liberation of maltose from the non-reducing end of starch polymer chains in a multiple attack mechanism [54]. Modeling of substrate complexes, for this enzyme, based on its three-dimensional structure, supports the view that the new non-reducing end of the (longer) chain remaining after removal of maltose is translocated to subsite −2 for subsequent attack [55]. The DMA value for the (1→3)-β-exoglucanase from *Sporotrichum pruinosum* QM 806 is 4 [42]. In rare cases, e. g., some cellulases, the *exo*-attack has been described as occurring from the reducing chain end. Thus, the cellulase from *Thermomonospora fusca* E4 [56], which corresponds to the cellobiohydrolase I from *Trichoderma reesei*, displays a preferred release of ^{18}O-labeled cellobiose from the reducing end of ^{18}O-labeled cellopentaose. In contrast, for cellobiohydrolase II the preferred attack between subsites −1 and +1 is at the non-reducing end of cellopentaose [57]. In this connection it should be noted that cellobiohydrolase II (Cel6A) from *Trichoderma reesei*, earlier believed to be an *exo*-enzyme, has been shown to exhibit *endo*-attack through its ability to cleave a bifunctionalized, fluorogenic tetrasaccharide [58]. Thus, it will be important to know whether the *T. fusca* enzyme is a true *exo*-enzyme. It may be envisaged that the attack described as *exo* from the reducing end of the substrate chain may very well be the result of an initial *endo*-action by a retaining enzyme followed by multiple *exo*-attack [38].

2.6 Mechanism of Glycosidic Linkage Cleavage

2.6.1 Hydrolases

The mechanism of enzyme-catalyzed hydrolysis of glycosidic linkages is analogous to acid-catalyzed hydrolysis. Cleavage is between the glycosyl anomeric carbon (C1) and the gly-

cosidic oxygen [59]. The general reaction pathways were first recognized by Koshland [60] who categorized glycoside hydrolases mechanistically on the basis of whether the hydrolysis occurs with retention or inversion of configuration at the newly exposed anomeric center (❯ *Fig. 5*). A distinction can also be made for hydrolases and glycosyl transferases with pyranosyl substrates depending on whether the enzymes cleave equatorial or axial glycosides [2]. Thus, four stereochemical types of glycosidase action, $e \rightarrow e, e \rightarrow a, a \rightarrow a$, and $a \rightarrow e$, can be recognized.

The most probable route for the retaining glycoside hydrolase reaction is through a general acid/base catalysis that involves two crucial carboxylic acid residues at the enzyme's catalytic site. In retaining enzymes the reaction path shown in ❯ *Fig. 5a* involves a double-displacement in which the glycosidic oxygen is protonated by one of the carboxylic acids. The other carboxylic acid acts as a nucleophile, i. e., forms a glycosyl-enzyme and the aglycone departs. A water molecule is then deprotonated and makes a second nucleophilic attack at the anomeric carbon generating a product with the same stereochemistry as the substrate. The dual role of the general acid (acid/base) catalyst in the double displacement reaction involving a covalent glycosyl-enzyme intermediate (❯ *Fig. 5a*) was demonstrated in the *Bacillus circulans* xylanase by the oscillation of its pK value with the individual steps in the reaction pathway [61]. The distance between the two catalytic carboxyls in retaining enzymes is 4.5–5.5 Å (❯ *Fig. 5a*) [62] allowing a close approach of the nucleophile to the anomeric carbon. In the action of retaining enzymes it is proposed that either an oxacarbenium ion-like intermediate or a covalent glycosyl intermediate is involved (❯ *Fig. 5a*). Both states may participate. The covalent intermediate is a carboxylic ester of the catalytic nucleophile in which the glycosyl residue is in an anomeric configuration opposite to that of the substrate. Stable covalent glycosyl intermediates have been isolated by designing substrates with good leaving groups. In inverting enzymes the reaction pathway (❯ *Fig. 5b*) involves a single direct displacement in which the glycosidic oxygen is protonated by one of the pair of carboxylic acids and the aglycone departs concomitantly with the attack by a water molecule activated by the second carboxylic acid residue. In these enzymes the two catalytic acids are separated by 9–9.5 Å (❯ *Fig. 5b*) [62] allowing the water and substrate to bind simultaneously.

There is now substantial direct evidence from crystallographic and mutagenesis studies, and from the use of mechanism-based inhibitors that the two postulated catalytic groups do indeed exist for glycoside hydrolases and glycosyl transferases. The proton donor and catalytic nucleophile are invariably carboxyl groups of aspartyl or glutamyl residues. Further evidence for the proposed reaction schemes has been adduced from kinetic isotope effects [63,64]. The bell-shaped activity-pH profiles for glycoside hydrolases (and glycosyl transferases), with optima generally in the range 4.5–8.0, are consistent with the presence of ionizing groups with pK values below and above the optimum pH. The value of the optimum is determined largely by the nature and environment of the catalytic nucleophile and acid-base catalytic carboxylic acids in the active site [65], but cases with pH optima above and below the normal range have been recorded. Using site-directed mutagenesis, the pK_a values of catalytic residues have been changed with consequential alteration of the pH-activity profile [66,67,68]. However, no correlations were found between altered pH-activity dependence, and the calculated electrostatic fields in *Bacillus licheniformis* α-amylase [67]. Although examples exist of mutated enzymes with altered pH-activity dependence, rational engineering of desired shifts in the pH-activity profiles is considered to have only a limited possibility of being successful.

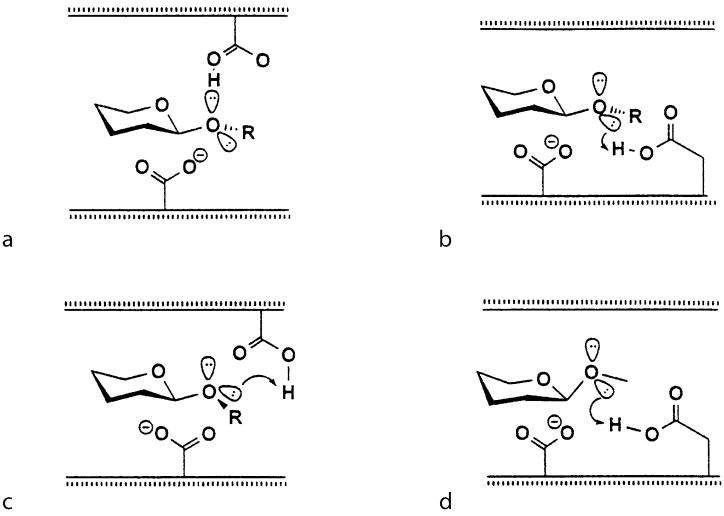


Figure 6

Protonation trajectories. (a) Protonation perpendicular to the plane of the ring, (b) *anti*-protonation, (c) *syn*-protonation, (d) lateral protonation moves the fissile C–O bond into a pseudoaxial position (after [69])

The direction of approach of the proton to the glycosidic oxygen shown in the mechanistic scheme in *Fig. 6a* [69] suggests that its trajectory is from above or below, and perpendicular to the plane of the glycone. It is now clear, however, from studies with inhibitors, analysis of the structures of enzymes in complex with ligands and by modeling, that protonation in many, if not all, glycosidases, does not follow such trajectories [69,70]. In contrast, the direction of protonation is lateral, or side-on, and in the plane of the glycosyl ring (at subsite −1) (*Fig. 6b, c*). Thus, the lateral trajectory of the proton may follow one of two directions with respect to the plane defined by O1, C1, and H1, i. e., it may be *anti* or *syn* to the pyranoside endocyclic O5–C1 bond as shown in *Fig. 6b, c*. The lateral protonation moves the fissile C–O bond into a pseudoaxial position (*Fig. 6d*). Heightman and Vasella [69] state that "It makes sense that the catalytic carboxyl group is located close to the main plane of the pyranoside ring and not above it since the departure of the aglycone in pseudo-axial direction would be restricted if the carboxyl group was located above the plane of the ring" (*Fig. 6a*). Stereoelectronic considerations require that protonation of the glycosidic oxygen atom be accompanied by a lengthening of the fissile C–O bond in a pseudo-axial direction as the incipient oxacarbenium ion develops. This is supported by two high-resolution crystal structures of the cellulases from *Fusarium oxysporum* [71] (Family 7, *syn*-protonation) and *Bacillus agaradherans* (Family 5, *anti*-protonation) [64]. They show that the pyranose ring is distorted to a skew conformation which places the glycosidic bond in a pseudo-axial orientation. A similar observation has been reported for two chitin-degrading enzymes of families 18 and 20 [72]. Similarly, the water molecule which subsequently attacks the glycosyl ester intermediate must have access from the pseudo-axial direction. The recognition of the *anti* or *syn* trajectories of approach of the proton allows families of glycosidases to be further sub-divided [69], emphasizing the relatedness of the members of the sub-divisions in respect to active site structures and mechanisms.

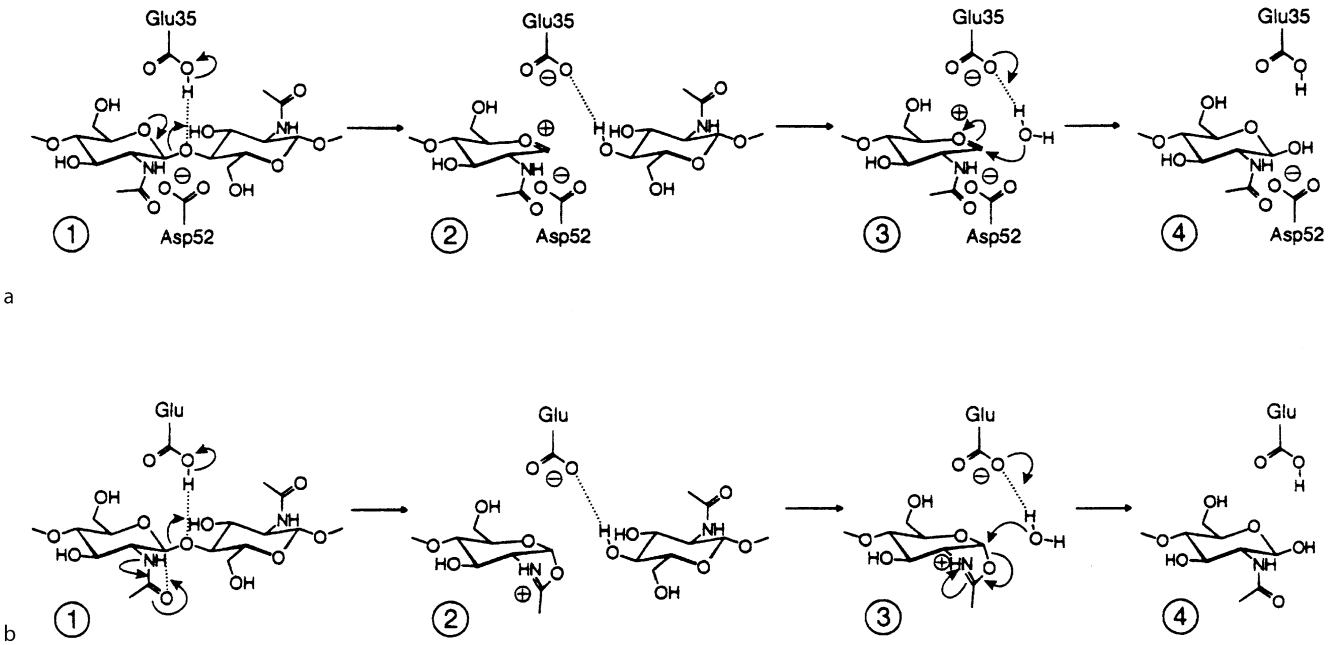


Figure 7

(a) Glycosyl hydrolysis by hen egg white lysozyme following general acid-base catalysis. Glu35 is the acid protonating the glycosidic oxygen and Asp52 is the nucleophile proposed to stabilize the reaction intermediate electrostatically. The two enzyme carboxylates have respective pKa values of 6 and 4. (b) Scheme for retaining chitinases where stabilization is proposed to occur via a covalent oxazolinium ion intermediate. The nucleophile is the *N*-acetyl group of the substrate itself. For stereochemical reasons the scheme is limited to retaining enzymes (from [72])

For a number of retaining polysaccharide hydrolases structural studies have revealed, contrary to expectation, a single obvious candidate for the catalytic acid but no candidate for the catalytic nucleophile near the anomeric carbon. These enzymes include family 18 and 20 chitinases [73], muramidase/transglycosylase [74,75,76], and some lysozymes [77,78]. To explain this it was proposed that in chitinase action the acetamido group of the *N*-acetyl-D-glucosamine monomer in subsite −1 plays the role of the catalytic nucleophile in the attack on the glycosyl intermediate (● *Fig. 7*) [72]. This was supported by X-ray crystallography of the enzyme in complex with allosamidin (see ● *Fig. 17k*), a pseudotetrasaccharide derivative substituted at its reducing end with an oxazoline-like structure.

2.6.2 Lyases

Lyases Acting on Uronic Acid-Containing Polysaccharides These enzymes, classified in [EC 4.2.2.-], are produced by bacteria, bacteriophages, marine gastropods, and a few fungi. Their substrates are all uronic acid-containing polymers with α- or β-(1→4) linkages, such as pectin and pectate (polygalacturonan), alginate, chondroitin, heparin, heparin sulfate, dermatan sulfate, hyaluronan, gellan, the side chain of xanthan and the Vi antigen of enteric bacteria [79]. The lyases so far have been classified into 10 families based on sequence homologies and folding patterns. The bacterial pectate lyases have a pH optimum around 9.5 and require Ca^{2+} for activity. The Ca^{2+} is bound in the active site [80]. The *Erwinia chrysanthemi* pectate lyase C (PelC) has been studied in great detail and generates a trigalacturonide product with a ^{4,5}-unsaturated bond in the galacturonosyl residue at its non-reducing end [81]. The β-elimination reaction (● *Fig. 8*, [82]) in pectolytic cleavage can formally be envisaged as involving three processes: neutralization of the carboxyl group adjacent to the glycosidic linkage cleaved, abstraction of the C5 proton, and transfer of the proton to the glycosidic oxygen. An arginine and an aspartic acid are presumed to play a crucial role in catalysis [15]. The complex of PelC crystallized with a Gal*p*A pentasaccharide shows the substrate fragment bound in a cleft and interacting primarily with positively charged groups: either lysine or

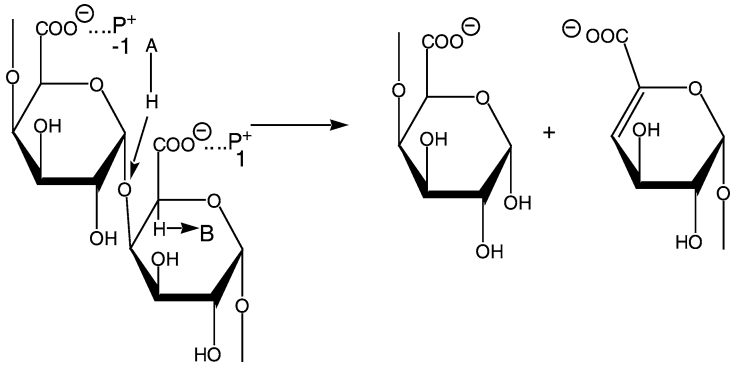


Figure 8
Proposed β-elimination reaction for cleavage of an oligogalacturonide in a pectate lyase (after [82])

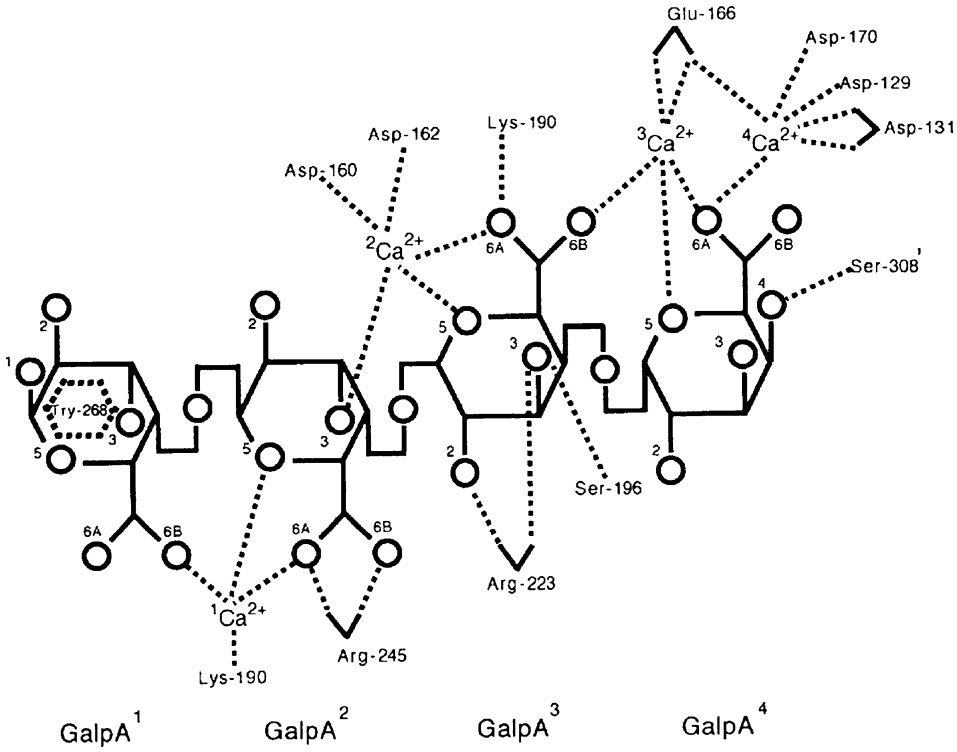


Figure 9

Schematic representation of the Ca^{2+} ion interaction with a tetraGal*p*A in the active site of the mutated (R218K, not shown) pectate lyase (PelC) showing the interactions primarily with positively charged groups: either lysine or arginine residues and four Ca^{2+} ions. GalpA^1 is the reducing saccharide, and GalpA^4 is the non-reducing terminus. The bond cleaved is most likely between Gal*p*A^3 and Gal*p*A^4. Gal*p*A^3 has the most interactions with the enzyme surface. The interactions are shown with *dotted lines*. Oxygen atoms are represented by *circles*, with the corresponding number, and the carbon atoms are assumed at the intersection of the bonds shown as *boldface lines*. Water molecules interacting with tetraGal are not shown (from [83])

arginine on the enzyme and the four Ca^{2+} ions found in the complex. The interaction of PelC with tetraGal*p*A as deduced from the crystal structure is shown in ❯ *Fig. 9* [83].

The protein-oligosaccharide interactions shown in ❯ *Fig. 9* provide a functional explanation for the occurrence of many of the invariant or conserved amino acids in this family of proteins [84]. An arginine residue (Arg218) is proposed to act as a general base in the catalysis. This is suggested by the proximity to the C5 of a Gal*p*A residue, its lower than normal pK_a, 9.5, compared with the calculated pK_a of 12.0–12.5 for all other arginines, and the impairment of activity when the residue is mutated to a lysine [15]. Members of the polysaccharide lyase superfamily (clan) are suggested to have the same reaction mechanism.

Lyases Acting on Neutral Polysaccharides Only one group of enzymes of this type is known, these are the amylolyases or (1→4)-α-glucan (starch) lyases [EC 4.2.2.13] that specifically

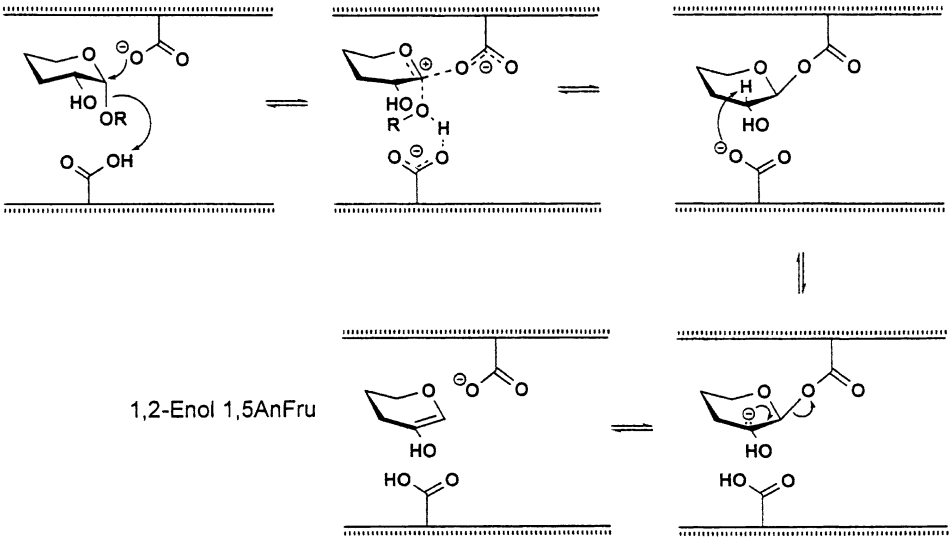


Figure 10
Proposed reaction mechanism of a (1→4)-⟨α-glucan lyase (amylolyase). The non-reducing glucosyl residue of the (1→4)-⟨α-glucan substrate chain is shown, the remainder of the chain is represented by R (after [51]). It has been suggested, however, that the catalytic nucleophile shifts position and becomes protonated by the H-C(2) in an E_{21} or E_{24} elimination (A. Vasella, personal communication)

cleave (1→4)-α-glucosidic linkages in starch or glycogen in an *exo*-action, converting the glucose residues at the non-reducing ends of the substrate to 1,5-AnFru (anhydrofructose), and ultimately liberating the reducing end residue of the substrate as free glucose [51] (● *Fig. 10*). The amylolyases are formally related to the glycosidases that catalyze the formation of 1,2-glycals from 2-deoxy-hexopyranosides detected by the reverse reaction in which a 1,2-glycal is hydrated (termed 'glycal hydrolysis') [85,86]. Amylolyases have been found in the cytosol of some fungi and in the stroma of chloroplasts of red algae, where they may be involved in the depolymerization of algal (Floridean) starch. Sequence alignment of the amylolyases has revealed that they are related to glycoside hydrolase Family 31 that includes α-glucosidases and sucrase-isomaltases. Seven well-conserved segments comprise functional residues including the catalytic groups identified in members of Family 31 [51]. This is the first case of a lyase and a hydrolase belonging to the same family and as such it may illuminate facets of the catalytic reaction mechanisms of both. Apart from sharing substrates with the amylases, the transition state analogue inhibitors, such as, 1-deoxynojirimycin (see ● *Fig. 17b*), and acarbose (see ● *Fig. 17g*), are effective with both types of enzyme. However, other common glycoside hydrolase inhibitors like glucono-1,5-lactone and the cyclodextrins, and the affinity labeling reagents, conduritol-B epoxide (● *Fig. 11a*) and 2′,3′-epoxypropyl-α-D-glucoside, do not inhibit the amylolyases. The mechanism of action [51,87] is likely to resemble that of α-glucosidases to a certain extent (● *Fig. 10*).

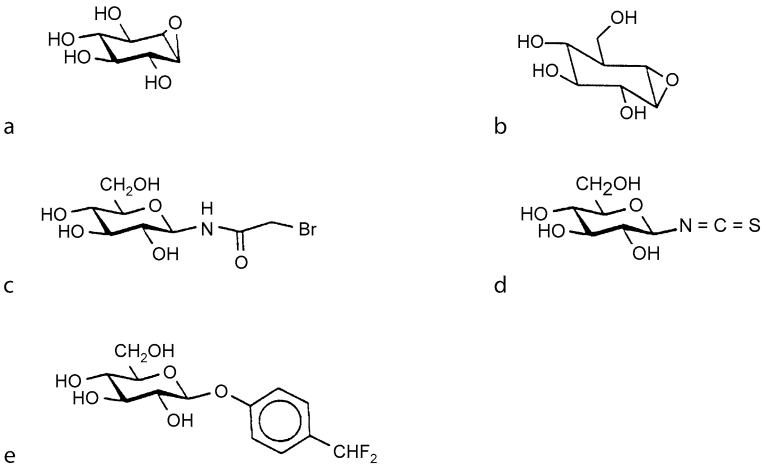


Figure 11
Structures of affinity- and mechanism-based inhibitors of glycoside hydrolases. (a) Conduritol B epoxide, (b) cyclophellitol, (c) *N*-bromoacetyl-glucosylamine, (d) glucosyl isothiocyanate, (e) *p*-difluoromethylaryl β-glucoside (mechanism-based)

2.6.3 Phosphorolases

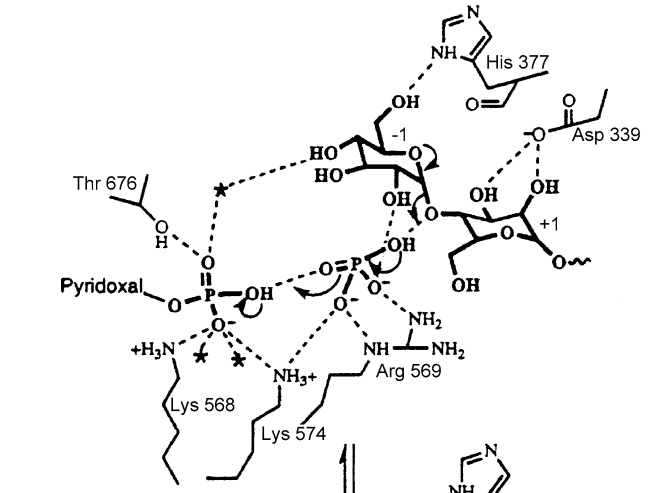
Phosphorolysis of glycosidic linkages in α- and β-glucans (*Structure 1*) is an alternative to depolymerization by hydrolysis and is used in mobilizing cellular storage polysaccharides. The glucose-1-phosphate product is a direct intermediate in hexose metabolism. Thus, the phosphorylases are found intracellularly together with their substrates: glycogen in vertebrates, invertebrates, and bacteria; starch polymers in higher plants; and (1→3)-β-glucans (paramylon) in euglenoid protozoa. Bacteria that are able to transport maltodextrins across their cell membranes depolymerize them with a maltodextrin phosphorylase and comparable eukaryotic, bacterial, and archaean phosphorylases for laminaribiose [88], cellobiose, and cellodextrins [89] have been described. The (1→4)-α-glucan phosphorylases from mammalian, plant, and bacterial sources are *exo*-enzymes catalyzing the reversible phosphorolysis of glucosidic linkages at the non-reducing ends of glucosyl chains to produce αGlc-1-phosphate, i. e., with retention of configuration. In the reverse direction, the (1→4)-α-oligoglucoside or glucan acceptor is lengthened by the transfer of the glucosyl moiety from α-glucose 1-phosphate, e. g., to its non-reducing end with the concomitant production of inorganic phosphate. The equilibrium constant for the reaction expressed as the ratio of [Pi]/[Glc-1-P] is 3.6 at pH 6.8. All phosphorylases are dependent for activity on the presence of metal ions, typically Mg^{2+} or Mn^{2+}, and have an absolute requirement for pyridoxal 5′ phosphate as a co-factor [90]. The (1→4)-α-glucan phosphorylases are classified in glycosyl transferase Family 36 and probably have similar kinetic mechanisms.

Phosphorylases from animal, bacterial, and plant sources can be distinguished by their substrate preferences. Animal and bacterial phosphorylases prefer the short outer branches of highly branched α-glucans, such as glycogen. The phosphorylase from *Corynebacterium calluna*, which accumulates a starch-like polysaccharide, is exceptional among bacterial enzymes

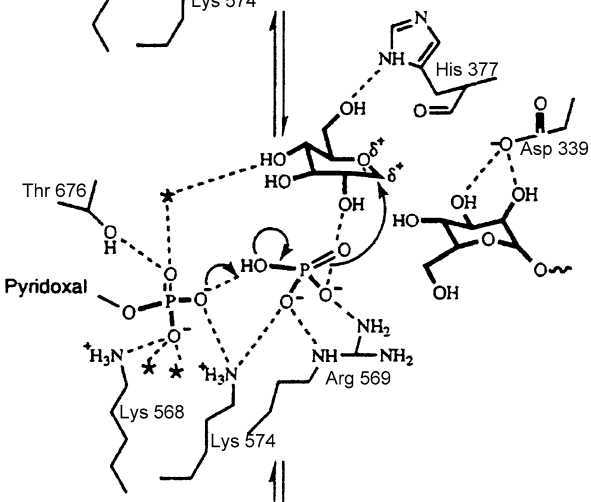
in its preference for long chain substrates [91]. *E. coli* possesses, in addition to a genuine glycogen phosphorylase, a maltodextrin phosphorylase [92] which has a high affinity for linear oligosaccharides, and less than 1% activity against glycogen. There are two isoforms of the plant phosphorylase; Pho1, from the cytosol, with a high affinity for amylopectin and glycogen, and Pho2, from plastids, that prefers amylose and maltodextrins. The three-dimensional structures of the mammalian [93,94] and yeast [95,96,97] enzymes reveals the active site in a tunnel with the Glc-1-P situated in the −1 subsite, at its end, far removed from the bulk solvent and adjacent to the essential pyridoxal cofactor [98,99]. The mammalian phosphorylases have not been crystallized in complex with their oligosaccharide substrates. However, the *E. coli* maltodextrin phosphorylase has 98% sequence identity with the mammalian enzyme in the region of the catalytic site and its structure has been determined in complex with bound maltopentaose [100] and the pseudo-oligosaccharide, acarbose (see ❯ *Fig. 17g*) [53,92]. The catalytic site is situated in the center of the large subunit and is accessible through a 20-Å-long channel that forms the oligosaccharide binding site. The non-reducing terminal maltosyl portion of the bound maltopentaose and the acarviosine portion of acarbose occupy subsites +1 and +2 and the reducing terminal maltosyl of acarbose, subsites +3 and +4.

In the *E. coli* maltodextrin phosphorylase, structural analysis of the binary complex with maltopentaose, 4-*S*-α-D-glucopyranosyl-4-maltotetraose (G*S*G4) and the ternary complex with GSG4 and phosphate show the pentasaccharide binding along the catalytic site and occupying subsites −1 to +4 with the phosphate group poised to attack the glycosidic bond and promote phosphorolysis [53]. In all three complexes, the pentasaccharide exhibits an altered conformation across subsites −1 and +1 (❯ *Fig. 12*), from the preferred conformation of (1→4)-α-linked glucosyl polymers. The oligosaccharides can find their way into the catalytic site tunnel only by alteration of the glycosidic torsion angles. Similar observations have been made with porcine pancreatic and barley α-amylase [36,101] and soybean β-amylase [55]. In starch phosphorylase [Pho2(L)] a 78 residue insertion is located beside the mouth of the active-site cleft and presumably sterically hinders the approach of the large highly branched glucan to the active site [102,103].

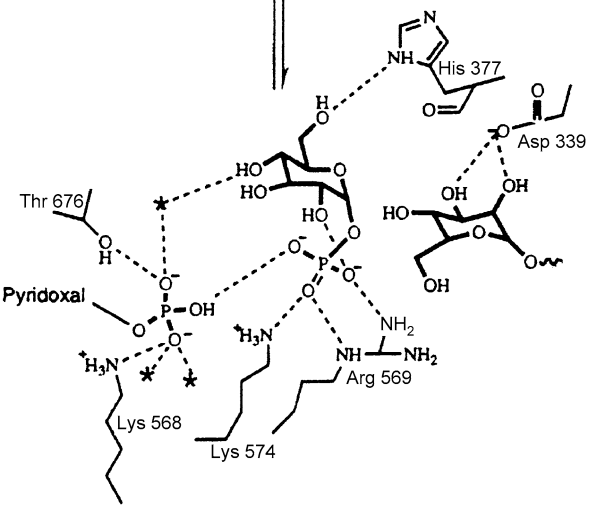
Muscle glycogen phosphorylase is composed of two identical ∼90 kDa sub-units. The catalytic capability of the enzyme is regulated by phosphorylation/dephosphorylation of a specific serine residue by a glycogen phosphorylase kinase/phosphatase and, non-covalently, by allosteric effectors such as AMP. The enzyme can exist in two states, an unphosphorylated, inactive form and a phosphorylated, active form. The covalent attachment of a phosphate group at serine 14 elicits conformational changes at remote sites mediated through subunit-subunit interactions. These changes lead to opening of a gate, created by a loop that closes the catalytic site in the inactive form, and by rearrangement of certain amino acids, creates a high affinity substrate-binding site. The unphosphorylated form can be activated by binding AMP at a site close to serine 14 which causes conformational changes leading to correct orientation of amino acids involved in subunit-subunit interactions [104,105,106]. By contrast, in the *E. coli* maltodextrin phosphorylase, the relative orientation of the two subunits is quite different and there are significant changes in subunit-subunit interactions. Sequence changes result in loss of the allosteric control sites resulting in the loop, which acts as a gate to control access to the catalytic site in the muscle phosphorylase, being held in the open conformation. In addition to binding (1→4)-α-glucan substrates in the active site, muscle phosphorylase has a non-catalytic, (1→4)-α-glucan-binding site, or 'glycogen storage site,' through which the enzyme may


His 377
Asp 339
-1
+1
Thr 676
Pyridoxal
Arg 569
Lys 568
Lys 574

a


His 377
Asp 339
Thr 676
Pyridoxal
Arg 569
Lys 568
Lys 574

b


His 377
Asp 339
Thr 676
Pyridoxal
Arg 569
Lys 568
Lys 574

c

Figure 12
Proposed catalytic mechanism for a phosphorylase based on the maltose phosphorylase GSG4-P ternary enzyme-substrate complex. (a) The oligosaccharide substrate is bound across the catalytic site in subsites −1 and +1. The phosphate substrate is within hydrogen-bonding distance of the glycosidic oxygen and is stabilized in this position by contacts to Arg569 and Lys574. The phosphate substrate is within hydrogen-bonding distance of the pyridoxal phosphate 5′-phosphate group whose position is stabilized by contacts to water (shown as *asterisks*) and polar amino acid residues. The co-factor 5′-phosphate acts as a general acid to promote attack of inorganic phosphate on the glycosidic oxygen. The oxygen is protonated leading to the cleavage of the C1–O bond and formation of the oxonium-carbonium ion intermediate. (b) The carbonium ion is promoted and stabilized by the inorganic phosphate dianion, the carbonyl oxygen of His377 and the lone pair of electrons on the 3′-OH of the saccharide in subsite +1. The 2′-OH group of the saccharide at subsite −1 is polarized by contact with the inorganic phosphate. (c) The cleaved oligosaccharide is free to diffuse away. Nucleophilic attack by the inorganic phosphate at the C1 position of the carbonium ion leads to the formation of ⟨-glucose 1-phosphate (from [52])

be attached to glycogen particles in vivo [107]. The site is on the enzyme surface some 30 Å from the catalytic site. The affinity for oligosaccharides at this site is higher than for the active site itself [108]. Such a 'storage' site is also present on plant phosphorylases but is lacking in the bacterial enzymes.

Mechanistic studies on phosphorylases suggest a double displacement mechanism is operating. The reaction proceeds through a rapid equilibrium bi-bi kinetic mechanism. There is random addition of the α-glucan and inorganic phosphate substrates, but the binding of one substrate increases the enzyme's affinity for the other substrate. Phosphorylases are able to catalyze the arsenolysis of Glc-1-P and the exchange between Glc-1-P and Pi, but only in the presence of an acceptor [109,110]. The rate-limiting step in the reaction is the interconversion of the ternary phosphorylase-phosphate-glycosyl complex to the ternary phosphorylase-Glc-1-P-glycosyl-complex [111]. By analogy with other retaining enzymes it may be suggested that a glycosyl-enzyme intermediate is involved, but such an intermediate has not been isolated. Positional isotope exchange results with the potato enzyme are consistent with such an intermediate. At present the identity of the presumptive nucleophile is unknown as there is some uncertainty due to the conformational flexibility at the catalytic site, and other mechanisms have been suggested [90]. The role of the essential pyridoxal phosphate bound at Lys 680 by a Schiff base close to the substrate phosphate in glycogen phosphorylase has not been resolved [4,99,112,113,114]. However, studies with the closely related *E. coli* maltodextrin phosphorylase [53] (*Fig. 12*), show that no acidic group participates directly in catalysis, but the 5′-phosphate of the pyridoxal cofactor promotes general acid attack by the inorganic phosphate, resulting in protonation of the glycosidic oxygen of the bond to be cleaved.

Substrate analogue inhibitors of phosphorylase include the ground state analogues, glucose; glucose-1-methylenephosphate and glucose-1,2-cyclic phosphate [107,114,115,116,117]; the dead-end product, heptulose 2-phosphate [99,118] and glucose derivatives with an sp^2 hybridized anomeric center; heptenitol [118,119]; D-gluconohydroximo-1,5-lactone *N*-phenylurethane [120,121]; acarbose [122]; and isofagomine [123,124,125]; nojirimycin tetrazole [113]; glucose spirohydantoins [126,127,128]; alkyl-dihydropyridine-dicarboxylic acid [129] and *N*-acetyl-β-D glucopyranosylamine [130]. 4-Deoxymaltohexaose/pentaose are good competitive inhibitors of phosphorylase with K_m values some 40-fold and 10-fold lower, respectively, than the underivatized oligosaccharides [52]. The increased affinity may relate to a stabilizing effect of the hydroxyl group on the ground state.

2.7 Hydrolase Inhibitors

2.7.1 Reagents Reacting with Amino Acid Residues Required for Activity or Binding

Inhibition or abolition of enzyme activity by group specific reagents that modify amino acid residues with a function in catalysis, or in binding the substrate can provide a preliminary indication of the identity of these important residues. Thus, various carbodiimides and Woodward's reagent K inhibit by reacting with carboxyl groups, other examples include *N*-bromosuccinimide reacting with tryptophans, tetranitromethane reacting with tyrosines, diethyl pyrocarbonate reacting with histidines, a variety of acylating reagents that modify lysines, alkylating reagents that modify thiol groups, and phenylglyoxal that modifies arginines [25,131]. The interpretation of the results needs to be made cautiously since many of the reagents are not absolutely specific and without corroborating evidence it is not possible to assign a catalytic or binding role to the amino acids. Differential labeling with unlabeled reagent in the presence of a substrate or substrate analogue, dissociation of the ligand and reaction with labeled reagent can provide better evidence for the role of specific amino acids in catalytic activity. Thus, acarbose (see *Fig. 17g*), a pseudo-tetrasaccharide substrate analogue and inhibitor of glucoamylase, protected the essential Trp120 from reaction with *N*-bromosuccinimide [132]. Although such reagents can be useful in exploring structure/function relationships, particularly when coupled with sequencing and differential labeling, most often site-directed mutagenesis is preferred in structure/function investigations [16,17]. Formation of hybrid or chimeric enzymes is another powerful tool in structure/function and structure/stability investigations and for protein engineering [133,134,135].

2.7.2 Irreversible Active-Site-Directed Inhibitors

Affinity Labeling These reagents [1,136,137] combine a glycosyl substrate moiety that targets the compound to the catalytic site and a functional group capable of reacting to form a stable, covalent linkage with a group in the active site, sometimes a catalytic amino acid, leading to inactivation of the enzyme. One of the first of this class of inhibitors was the 2,3-epoxypropyl-β-chitobiose used to inactivate lysozyme [138,139,140,141]. These substrate analogues require an acidic group for the protonation of the oxirane ring and a nucleophile for the formation of the covalent bond. The reaction is shown in *Fig. 13*.

Reduction of the glycosyl ester so formed converts the carboxylic acid-bearing residue to the corresponding alcohol and can be used to identify the catalytic amino acid at the site of reaction [139,140]. A panel of these reagents [142] has been used to probe the active sites of a series of β-glucan hydrolases [143] and to 'specifically' label and identify their catalytic nucleophiles [144,145,146,147,148,149,150]. The 2,3-epoxypropyl-α-glucoside reacts with a carboxylic acid in β-amylase [151]. The length of the epoxyalkyl chain influences the effectiveness of the inhibition, which depends on whether the oxirane ring is within reach of the catalytic group [145]. Because the alkyl chains show flexibility, active-site carboxyl groups in the vicinity of the catalytic nucleophile may be labeled, and it is necessary to verify the identity of the labeled group by independent methods such as X-ray crystallography of the enzyme complexed with the inhibitor. This has been achieved for hen egg white lysozyme [152] and

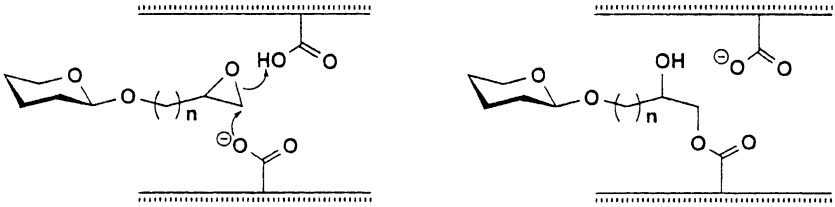


Figure 13
Reaction pathway for epoxyalkyl β-glycoside substrate analogue with a β-glucan hydrolase involving protonation of the oxirane ring by an acid group and the formation of a covalent bond with a nucleophile at the active site

the hybrid *B. macerans/B. amyloliquefaciens* (1→3; 1→4)-*endo*-β-glucanase [150] or with a bona fide mechanism-based inhibitor.

Significant differences in the stereospecificity of the interaction of the epoxyalkyl group between the (1→3; 1→4)-β-glucan hydrolases from *Hordeum vulgare* (Family 17) and *Bacillus subtilis* (Family 6), with indistinguishable substrate specificities, were shown by comparing the series of (*RS*)-epoxalkyl-β-cellobiosides with the pure (3*R*)- and (3*S*)-diastereoisomers [153]. The 3*S*-isomer was most effective for the barley enzyme whereas the 3*R*-isomer was almost without effect. The opposite was true of the *B. subtilis* enzyme. The structure of the *Fusarium oxysporium* (1→4)-β-glucan endohydrolase [154] affinity-labeled with a mixture of *R*- and *S*-forms of epoxybutyl-β-cellobioside shows the butyl moiety lying over the −1 subsite. Both isomers appear to be present, but modeling studies suggest that the electron density is much stronger for the *R*-isomer. Examples of other reagents with functional groups that have been used in affinity labeling of glycoside hydrolases are the oxiranes, conduritol B epoxide (*Fig. 11a*), and the naturally occurring cyclophellitol (*Fig. 11b*) [155], and various glycosyl epimines (aziridines) [156,157], the *N*-bromoacetyl-glycosylamines (*Fig. 11c*) [158], the glucosyl isothiocyanates (*Fig. 11d*) [159], and diazoketones [160]. As is the case with the epoxyalkyl affinity reagents the labeled amino acid residue may not necessarily be the catalytic nucleophile.

Bromoketone *C*-cellobioside and *N*-bromoacetyl-β-D-cellobiosylamine [161] differentially inactivated *Cellulomonas fimi* β-glucanases. The bromoketone derivatives inactivated the (1→4)-β-glucan endohydrolases, CenD and CenA, with stoichiometric labeling of the enzymes, but not the (1→4)-β-glucan exohydrolase Cex, whereas the *N*-bromoacetyl derivatives inactivated CenA and Cex, but not CenD. This is attributed to the different modes of hydrogen bonding of the amide moiety. The amino acid residues labeled were not identified. However, stoichiometric labeling of the putative acid/base, Glu276, in *Saccharomyces cerevisiae* α-glucosidase was obtained using the 1′-bromo-3′-(α-D-mannopyranosyl)-2′-propanone [162]. This Glu is completely conserved in other Family 13 α-glucosidases and α-amylases whose three-dimensional structures have implicated this residue as the acid/base catalyst, i. e., the proton donor. Thus, the bromoketone derivatives are good candidate reagents for labeling the general acid/base catalyst.

Mechanism-Based Inhibitors Another class of active-site-directed inhibitors are glycoside derivatives that require mechanism-based activation in order to react covalently with

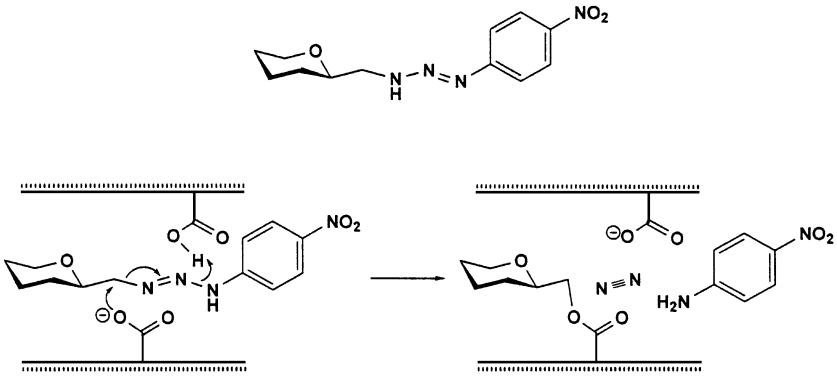


Figure 14
Reaction pathway in glycosyl methyltriazene activation [163]

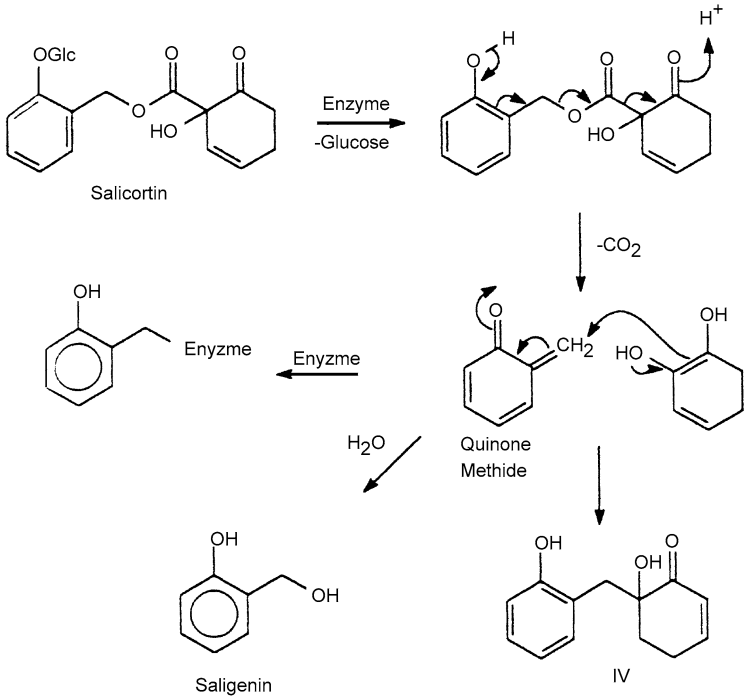


Figure 15
Reaction pathway for salicortin inhibition (after [165])

the enzyme. One class of such nucleophile-specific reagents are the glycosylmethyl-triazenes [163] that require protonation after binding, and decompose to yield a very reactive glycosylmethyl-carbenium ion (*Fig. 14*).

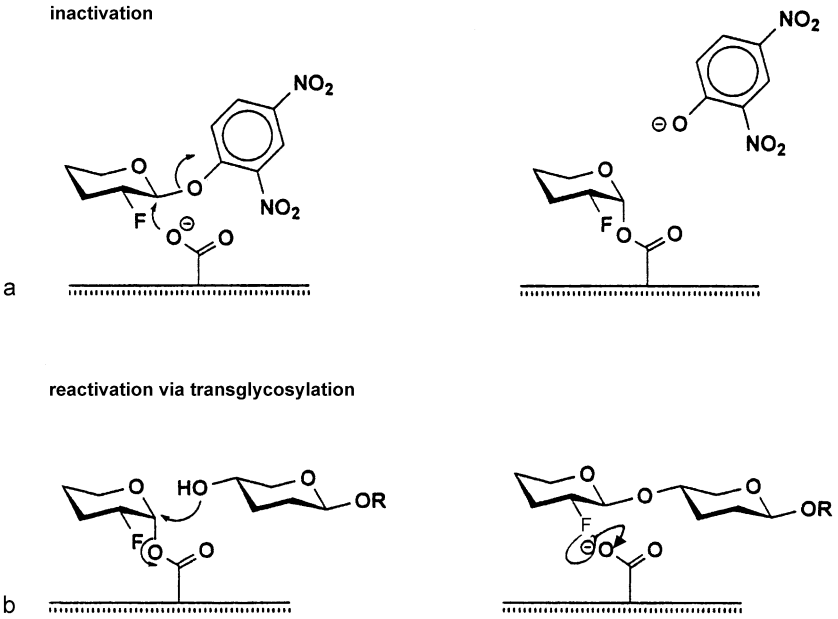


Figure 16
Reaction pathway and mode of activation of a β-glucosidase by 2,4-dinitrophenol 2 deoxy-2-fluoro-D-glycoside (after [137,167])

A second class of mechanism-based inhibitors are those that in the course of their cleavage by the enzyme release an aglycone that rearranges to give a reactive species. Two examples may be cited. Salicortin [164], a natural product, is hydrolyzed by β-glucosidases to produce, by rearrangement, a reactive quinone methide as shown in ❯ *Fig. 15* [165].
The *o*- and *p*-difluoromethylaryl β-glucosides (❯ *Fig. 11e*) [166] are further examples of this class of mechanism-based inhibitors. They are cleaved by β-glucosidases to yield a fluorohydrin that decomposes to hydrogen fluoride and an acyl fluoride that reacts at the active site. A third class of mechanism-based inhibitors are the 2-deoxy-2-fluoro-D-glucosides with good leaving groups, e. g., 2,4-dinitrophenol [167] and the 5-fluoro-D-glucosyl fluoride [168,169]. These reagents form reactive species only after cleavage of the glucosyl moiety and hence are highly specific inhibitors of β-glucosidases and have been used successfully to label a number of $e \rightarrow e$ enzymes. In the reaction (❯ *Fig. 16*) "the covalent intermediate formed is stabilized against hydrolysis, on the one hand, by the inductive destabilization of the positively charged oxacarbenium ion-like transition states by the electronegative fluorine and, on the other, by loss of crucial transition state hydrogen-bonding interactions" [4]. When a good leaving group is incorporated, the first step in the reaction proceeds at a reasonable rate. In some $e \rightarrow e$ enzymes the turnover rate of the intermediate is measurable, but the intermediate is stable enough to allow identification of the nucleophile. The identity of the amino acid residue labeled by affinity- or mechanism-based inhibitors has been determined using sensitive, mass spectrometric methods without the need for radiolabeled reagents [137,170,171].

Many of the affinity- and mechanism-based inhibitors were designed for use with glycoside hydrolases, however, they have potential applications with glycan hydrolases. Their effectiveness may be improved by extending the glycosyl portion to provide tighter binding in the active site. This was demonstrated to be the case in a family of epoxyalkyl β-glucosides and oligoglucosides of various chain lengths [170]. Those with longer oligosaccharide chains were more reactive and the inhibition more specific (see [1,5,137] for a discussion of the operation of these mechanism-based inhibitors).

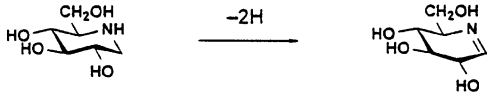
2.7.3 Transition State Inhibitors

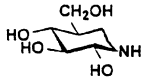
Transition-state analogues represent a class of inhibitors distinct from the active site-directed inhibitors. They are reversibly bound in the active site, and compete with the substrate by mimicking the shape and/or the charge of the transition state. A large number of transition-state inhibitors is available for glycosidases, some of which were developed for therapeutic or other applications, some are effective with polysaccharide depolymerizing enzymes. Transition-state inhibitors are useful tools for elucidating important residues in the enzyme-substrate interaction by protein crystallography. For glycoside hydrolases, an important transition state is believed to be a planar glycosyl-oxacarbenium ion [4,69]. Examples of transition state-analogue inhibitors are aldonolactones such as D-glucono-1,5-lactone (*Fig. 17a*) [1], basic sugar analogues such as the glycosylamines [1], glycosylimidazoles [1], cyclic aminoalditols, such as 1-deoxynojirimycin, (*Fig. 17b*) and isofagomine (*Fig. 17c*) [123,124]; glycotetrazoles, e. g., glucotetrazole (*Fig. 17d*) [69,172,173], pseudosubstrates such as glycals, e. g., D-glucal (*Fig. 17e*) [1], and the aminocyclitols, e. g., valienamine [1], carbacyclic pseudosaccharides incorporating valienamine or an analogue into an oligosaccharide chain, e. g., methylacarviosin (*Fig. 17f*), acarbose (*Fig. 17g*) [35,36,45,101,174,175,176], adiposin (D-*gluco*-dihydroacarbose) (*Fig. 17h*), methyl β-acarviosin (*Fig. 17i*), and methyl β-acarboside (*Fig. 17j*) [177,178,179] and the oxazoline-like pseudotrisaccharide, allosamidine (*Fig. 17k*) [72,74].

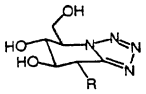
From a structural point of view some of these compounds are not perfect transition state analogues, but have closely related structures to, and behave kinetically as, transition state inhibitors [180]. One way to test the validity of a compound as a transition state analogue is by determining the relationship between K_i and k_{cat}/K_m for a series of mutated enzymes such as, e. g., for the acarbose-glucoamylase system [181] where these parameters gave a linear relationship in a double logarithmic plot. The pre-steady state kinetic analysis [177] indicated, however, that the interaction with acarbose elicited a unique conformational change in the enzyme as the final step in its binding mechanism [21]. Thus, thorough analyses of specific inhibitor-enzyme pairs may reveal aspects that are not evident from conventional kinetics. For Family 18 chitinases that follow a substrate-assisted catalytic mechanism involving a neighboring substrate C2′ acetamido group stabilization, molecular mechanical calculations have shown that oxazolines are potential inhibitors [73]. This is supported by the structure of the complex with the natural oxazoline, allosamidin (*Fig. 17k*) [74].

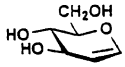
Figure 17
Structures of some transition-state inhibitors of glycosyl and glycan hydrolases

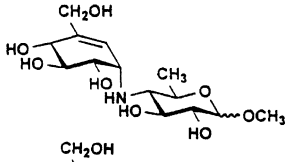
a 

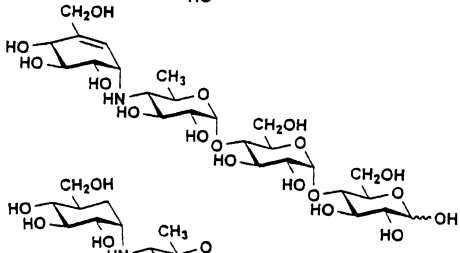
b 

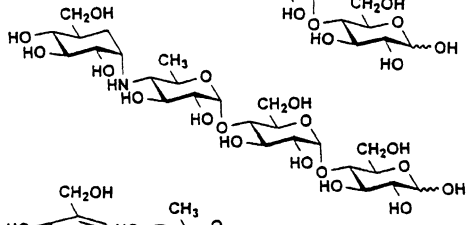
c 

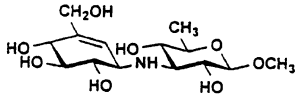
d 

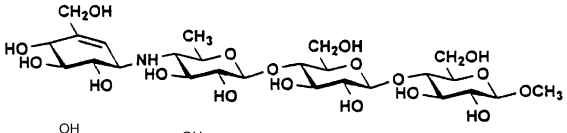
e 

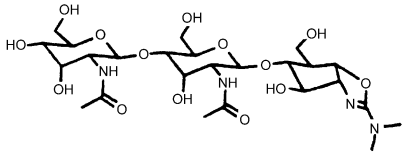
f 

g 

h 

i 

j 

k 

Recently, a series of tetrazoles, including *gluco*tetrazole (◗ *Fig. 17d*), was designed as neutral, stable lactone analogues and correlations were found between k_{cat}/K_m of various glycosidase-substrate pairs [172]. The tetrazole transition-state inhibitors have been used to classify hydrolases according to the *syn* or *anti* in-plane-protonation of the oxygen of the glycosidic bond cleaved [69,173].

2.7.4 Proteinaceous Inhibitors

A variety of polysaccharide hydrolases is inhibited by endogenous or exogenous proteins. These include α-amylases and limit dextrinase [182,183,182,185,186,187,188,189], fungal xylanases [190,191] and polygalacturonases [192]. Most of these inhibitors are from plants [183,185,186,187,192] but some are of microbial origin [182]. Structures of the enzyme-inhibitor complexes are available in some cases, in others only the structure of the free inhibitor is known. Five distinctly different α-amylase-inhibitor complexes have been reported including two for the pancreatic enzyme, one for the barley and two for insect α-amylases. Structural mimicry of the substrate has been observed in the bean (*Phaseolus aureus*) inhibitor-pancreatic α-amylase complex and in the complex of the small amaranth (*Amaranthus hypochondriacus*) inhibitor of the knottin family and the yellow mealworm (*Tenebrio mollitor*) α-amylase [183,187]. For the barley enzyme-inhibitor complex, contacts between the side chains and the carboxylic acid residues at the catalytic site of the enzyme involve a hydrogen bond network mediated through structural water molecules and a fully hydrated calcium ion at the protein-protein interface [186]. The role of the individual amino acid side chains in complex stabilization has been investigated by mutational analysis [135].

2.8 Structural Organization of Glycan Hydrolases

Many microbial polysaccharide hydrolases are secreted in a soluble form into the extracellular environment where their target substrates are found, whereas some bacterial cellulases are components of high molecular mass ($2.0–2.5 \times 10^6$ Da) cellulosome complexes that are anchored to the outer cell surface [193]. The α-amylases found in mammalian saliva and intestinal tract are soluble components of parotid gland and pancreatic secretions, respectively. The aleurone cells of germinating cereal seeds secrete a range of soluble glycanases concerned with mobilization of endosperm cell wall polysaccharides and starch. Some plant cellulases of unknown function are membrane-anchored [194]. In mammalian tissues polysaccharide hydrolases (e. g., lysozyme) may be compartmentalized in membrane-bound organelles (lysosomes), where they are concerned with turnover of cellular constituents and degradation of foreign material. Delivery to, and maintenance of the enzymes in, these various cell, tissue, and extracellular locations, involve specific structural features beyond those participating in substrate-binding and catalysis. Structural analysis of polysaccharide hydrolases (and glycosyl transferases) has shown that most are globular proteins carrying a catalytic domain. Additional non-catalytic domains involved, for example, in substrate binding, association with structural proteins, with membranes or with other catalytic monomers are known. In many instances the presence of a non-catalytic domain has been recognized but the function has not been identified.

2.8.1 Catalytic Domains

The various folding patterns of the polypeptide chain constituting the catalytic domains of glycoside hydrolase families have been reviewed [4]. At the time of writing three-dimensional structures are available for members of over 31 hydrolase families and one unclassified hydrolase and three families of glycosyl transferases and one unclassified (Bernard Henrissat, personal communication).

Topology of Active Sites Most often the active site is confined to one domain but there are examples where the active site is contributed by amino acid residues from two different domains [195,196]. The topology of the active sites of glycan hydrolases may be described in one of three general shapes: a pocket, a cleft or groove, or a tunnel, irrespective of whether the enzyme is inverting or retaining. Examples of the three types are shown in ❯ *Fig. 18*.
In some enzymes of the first group, the pocket may be quite shallow, but in others the pocket is quite deep [45,196,197]. The cleft or groove topology is shown by many polysaccharide hydrolases [46,101,174,176,198,199,201]. The length of the cleft will define the number of glycosyl binding sub-sites and hence the length of the polysaccharide chain that can be accommodated. In cellobiohydrolase II (Cel6A) from *Humicola insolens* [173] and cellulase Cel6A from *Trichoderma reesei* [202] the groove-type topology has been modified to form a tunnel by covering the open cleft with a polypeptide loop [203]. A tunnel-shaped active site would appear to be conducive to multiple attack. It is envisaged that the substrate is threaded through the tunnel, and after a productive binding and cleavage, one product is released and the other remains bound and processes in the active site to give a new productive association allowing repetitive (multiple) attack. How the enzyme-substrate complex is formed is not understood, but may involve opening the 'lid' over the cleft to allow binding. In *T. reesei* Cel 6A, Trp272 at the entrance of the tunnel is critical for the degradation of crystalline cellulose and has been proposed to have a role in threading the glucan chain into the tunnel [204]. In bacterial cyclodextrinase (Family 13) the shape of the active site pocket is altered by homodimer formation in which the *N*-terminal domain of one subunit covers the entrance to the catalytic site of the other, thus causing a change in substrate specificity [205,206].

Substrate Recognition at the Active Site The specificity of glycan hydrolases is determined by the ability of monosaccharide residues in the polysaccharide substrate to be recognized by, and be bound to, contact amino acid residues in the binding site. Generally, the binding of monosaccharide residues is promoted by the presence of planar hydrophobic groups on the enzyme surface, contributed by tyrosine, phenylalanine, or tryptophan side chains, which form a platform on which the hydrophobic face of the glycosyl ring sits. This appears to be a general strategy for protein-monosaccharide ring associations as was first shown for the bacterial arabinose transporting protein [19,207]. However, specificity is determined by substituents on the ring that make hydrogen bonds, either directly or through water molecules and polar interactions, with an array of specific amino acids lining the active site, as illustrated in ❯ *Fig. 19* [208].
There is absolute specificity with respect to the anomeric configuration of the monosaccharide residues in the substrate and the regioselectivity for linkage position is usually absolute for the glycone, but the 'aglycone' binding interactions may be much less stringent. Quite subtle dif-

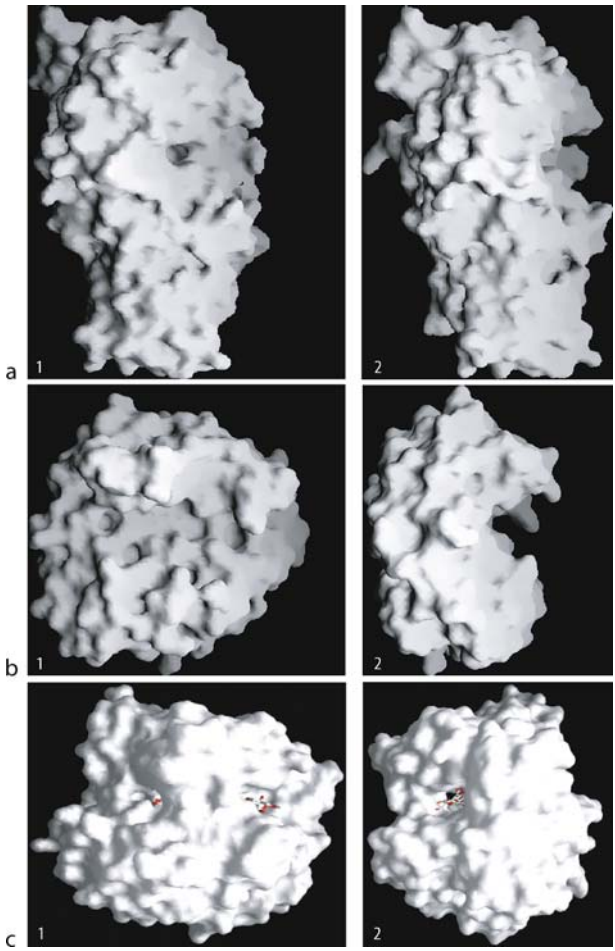


Figure 18

Three types of active sites found in glycan hydrolases. (a) *Hordeum vulgare* β-glucan exo-hydrolase [196]; 1. Viewed from above the catalytic face. 2. Rotated 90° along the long axis. The site appears as a compact hole in the middle of the molecule. (b) *Hordeum vulgare* (1→3;1→4)-β-glucan endo-hydrolase [199]; 1. Viewed from above the catalytic groove, that runs east-west. 2. Rotated 90° perpendicular to the groove axis. (c) *Trichoderma reesei* Cel6A (1→4)-β-glucan cellobiohydrolase (formerly cellobiohydrolase II) [200]; 1. Side-on view of the tunnel containing a substrate analogue. 2. End-on view of the same complex as in 1

ferences in the contact amino acids lead to changes in substrate specificity. This is illustrated by the (1→3)-β- and (1→3;1→4)-β-glucanases from barley, that both belong to Family 17, and have an almost identical folding pattern, with a binding site in the form of a deep cleft (*Fig. 18b*). Both require a laminaribiosyl residue in the −2 and −1 subsites, but the (1→3)-β-glucan hydrolase accepts only a 3-linked glucosyl in the +1 site, whereas the (1→3;1→4)-β-glucan hydrolase accepts only a 4-linked glucosyl residue in this site. Modeling studies suggest that the determinants of substrate specificity are identifiable amino acids in the active

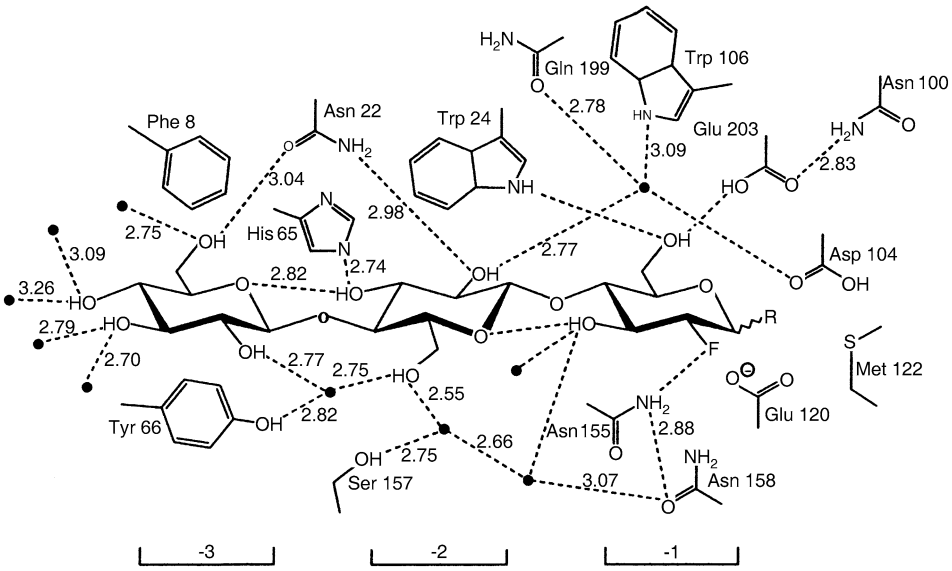


Figure 19
Active site of the *Streptomyces lividans* endoglucanase CelB2 showing the non-bonded interactions between the covalent 2-deoxy-2-fluorocellotriosyl intermediate complex with the nucleophile Glu120 and adjacent amino acids. The F-cellotriosyl portion sits in subsites −3, −2, and −1 (from [208])

site near the conserved catalytic amino acids [199]. The *Rhizopus arrhizus* β-glucan hydrolase has the same glycosyl requirement as the barley and *Bacillus* enzymes but will accept either a 3- or a 4-linked glucosyl residue in the +1 site suggesting a less stringent "aglycone"-binding interaction [209]. The homomorphous (1→4)-β-glucans and (1→4)-β-xylans are sometimes discriminated by cellulases and in other cases they are not [210]. Some chitinases that catalyze the hydrolysis of chitin, a β tGlcNAc homopolymer, also cleave the (1→4)-linkages in murein the homomorphous polysaccharide backbone of the bacterial cell wall peptidoglycan. In murein every other residue is muramic acid, the lactyl ether of GlcNAc, and this is the substrate for animal and bacteriophage lysozymes [211]. Lysozymes, e. g., from hen egg white may also hydrolyze chitin [212]. By an approach involving systematic removal of hydroxy groups [213], on disacccharide substrate it has been possible to determine which OH groups on maltose and isomaltose are needed for binding and hydrolysis by glucoamylase [214,215], but the difficulties of synthesis makes this an impractical strategy for most polysaccharide hydrolyzing enzymes. There are several examples where conformational changes of the protein are seen to accompany substrate binding but it is beyond the scope of the present chapter to include a detailed description of these events. Some well-established cases include: cyclodextrin glycanotransferase where the recognition of the substrate at the active site has been shown to be achieved by concomitant alteration of the conformation of both the substrate and the protein [216]. Structural evidence for a conformational change on substrate binding has been provided for porcine pancreatic α-amylase [36], soybean β-amylase (binding maltose) [197], and the cellobiohydrolase II, Cel6A, from *Humicola insolens* [173] and from *Trichoderma reesei* [217].

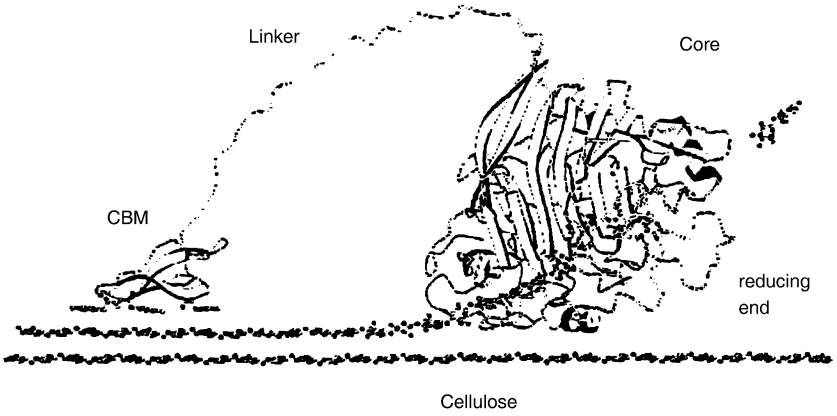


Figure 20
Model of intact cellobiohydrolase 1 (CBH1) from *Trichoderma reesei* bound to a cellulose microfibril surface. The model is based on the crystal structure of CBH1 in complex with cellohexaose, the NMR structure of the cellulose-binding module (CBM) and a modeled linker region. CBH1 is believed by some to act repetitively from the reducing end towards the non-reducing end of the cellulose chain (from [231])

Non-Catalytic, Carbohydrate-Binding Modules Hydrolases acting on large, and sometimes insoluble, substrates [10] such as starch [218,219,220,221], cellulose [222,223,224], xylans [225] and (1→3)-β-glucans [226]. The (1→3)-β-glucan binding modules are conserved in non-catalytic (1→3)-β-glucan-binding proteins from invertebrates [227]. An arabinofuranosidase has been reported with a (1→4)-xylanase-binding module [228] and a xylanase with a cellulose-binding module [229]. A xylanase from a thermophilic bacterium has a binding module for soluble xylans and (1→3;1→4)-β-glucans [230]. The binding module confers thermostability on the enzyme. In cellulases that depolymerize crystalline cellulose, the catalytic domain is connected by a linker peptide to a functionally independent, cellulose-binding module. This module associates very tightly with cellulose microfibril surfaces either in crystalline or non-crystalline areas, depending on the enzyme source, and anchors the associated catalytic domain to the solid substrate thus bringing and maintaining a high local concentration of the enzyme on the cellulose microfibril surface (*Fig. 20* [231]). Removal of the CBM from the catalytic domain results in a significant decrease in the ability to hydrolyze solid substrates.

The smallest CBM, in Family 1, consists of fewer than 40 amino acids and its binding is governed by a planar strip of three tyrosine residues which stack on every second residue of the cellulose chain [232]. The binding is quite tight and not very sensitive to changes in pH but, after replacing two of the tyrosines with histidine by site-directed mutagenesis, binding is still tight but is now pH-sensitive [233], supporting the involvement of the aromatic residues in the binding. A planar binding strip is also a feature of several other CBM families [222]. Some CBMs, in addition to binding the enzyme to the cellulosic surface, appear to disrupt non-covalent interactions between cellulose chains in the crystalline substrate and may potentiate the action of the cellulase [234]. A Family II CBM has been shown to move laterally on the cellulose surface [235] and this may aid in the translocation of the substrate

molecules in the active site in a repetitive action. The CBM of the *Thermomonospora fusca* E4–68 cellulase does not exhibit a planar binding strip topology, but binds to individual chains and is postulated [57] to act in feeding the cellulose substrate into the active site allowing the repetitive (multiple) attack (*Fig. 20*). Non-catalytic, chitin-binding-proteins that adhere to crystalline α-chitin, but not to β-chitin (chitosan), with high affinity, are produced by certain bacteria [236] and the binding involves specific Trp residues. The three-dimensional structure of an isolated starch-binding domain (SBD) from *Aspergillus niger* glucoamylase has been described [221,237,238] and a homologous SBD is seen in the three-dimensional structure of the *Bacillus circulans* CGTase [47]. The SBD is considered to disrupt polysaccharide chain interactions on the starch surface [239] since the release of glucose during solid starch degradation by a proteolyzed form of glucoamylase lacking its SBD is increased by addition of isolated SBD. A mutated SBD from *A. niger* glucoamylase was used to demonstrate that the two binding sites for β-cyclodextrin depend on aromatic residues for function and that they have slightly different binding affinities [238]. Replacement of a CGTase SBD with the one from *A. niger* glucoamylase caused a drastic decrease in activity suggesting that efficient interplay between the binding and catalytic domains is structure-dependent. CGTase has no flexible linker region, as in glucoamylase, but two globular domains connecting the binding and catalytic domains [214,215]. The crystal structure of the SBD of *Bacillus cereus* β-amylase complexed with maltose has been determined [240]. It is homologous to the *A. niger* SBD.
In many cellulose and starch hydrolases with multidomain architecture, highly *O*-glycosylated linker regions connect the catalytic domain and binding and module are considered to provide flexibility in the relative position of the domains on the substrate. The three-dimensional structures of the whole molecules are not known but dynamic light scattering studies of glucoamylase in solution using ligands having a catalytic site ligand, acarbose, connected by spacers of varying length to an SBD ligand, β-cyclodextrin, show that the two domains can come into close contact in solution [241]. Isothermal titration calorimetry indicates that strain is required to accommodate the two target moieties on the two domains in a single molecule of glucoamylase [242].

Adhesion Domains The core of the cellulosome in *Clostridium thermocellum* and *C. cellulolyticum*, is a large (210 kDa in *C. thermocellum*) multidomain, non-enzymic, cellulosome-integrating protein (Cip) (or scaffolding protein, 'scaffoldin'), that binds multiple *endo*-cellulases through high affinity ($K_D = 2.5 \times 10^{-10}$ M) interactions with multiple receptor ('cohesin') domains on the scaffold protein and 'dockerin' domains on the cellulases [243]. The Cip also has a single CBD [222,244]. There is significant synergism in the attack on crystalline cellulose so that the depolymerization of individual cellulose crystals appears to occur completely once attack has been initiated [245].

Other Non-Catalytic Domains. Many polysaccharide depolymerizing enzymes are secreted by bacteria and fungi from their intracellular site of synthesis across the periplasmic space and cell wall into the surroundings. These exported enzymes reach their destination via the general secretory pathway [246] which involves membrane protein complexes such as the bacterial ABC transporters [247] that recognize domains on the secreted proteins. There are several examples of non-catalytic domains on microbial carbohydrases that may play a role in transport from the cells. These include a repeat domain on a chitinase involved

in cell wall/membrane association [248] and cell membrane-associated pullulanase [249]. Fibronectin type III-like repeats found in several microbial carbohydrases may play a role in secretion [250].

Structure-Stability Relationships The thermostability and pH stability of polysaccharide hydrolases and glycosyl transferases vary according to the source. Generally, the enzymes originating from animal, plant, and mesophilic organisms are irreversibly inactivated above about 50 °C. Thermophilic bacteria and archaeans are a source of thermostable polysaccharide hydrolases with stabilities up to 80 °C and exceptionally up to 100 °C; and some of these enzymes are also halotolerant [251,252]. Polysaccharide hydrolases are normally stable in the pH range 4–8 but again there are acidophilic and alkalophilic organisms whose enzymes are stable at lower or higher pH values, respectively. The range of catalytic rate constants and half-lives of the extremophiles makes them candidates for application in industrial processes. Although metal ions do not participate directly in the catalysis by polysaccharide hydrolases, some e. g. a-amylases require Ca^{2+} and Na^{+} for activity [253,254]. One or more Ca^{2+} ions bind to and stabilize certain folding elements that are necessary for maintaining the active conformation of the enzyme. The requirement is usually detected by abstraction of the metal ion with a chelator such as EDTA or EGTA [255].

3 Application of Polysaccharide Depolymerases to Manufacture Bioactive Oligosaccharides

3.1 Manufacture of Nutritional Oligosaccharides

Recent years have seen an increase in interest in oligosaccharides that can play a role in optimizing health (See Chap. 5.3). These include the prebiotic oligosaccharides [256,257] and also oligosaccharides designed to inhibit adhesion of bacteria and toxins to human cells [258,259,260,261]. Whilst synthetic chemistry can manufacture almost any carbohydrate structure, large scale manufacture will probably not be based on traditional chemistry techniques. Most manufacturing technologies for bulk carbohydrates rely on enzymes, mostly polysaccharide depolymerases [262]. Polysaccharides represent a huge resource of chemical structures that can be liberated by enzymes.

A rapidly developing class of functional carbohydrates are the prebiotics (Chap. 5.3). Currently these selectively stimulate health-positive bacteria in the colon [256]. There are, however, possibilities to enhance their functionality. One such functional enhancement would be persistence of the prebiotic fermentation to the distal colon [263]. One way to approach this is to produce oligosaccharide mixtures with controlled molecular weight distributions.

3.1.1 Hydrolysis of Inulin

Inulin is a recognized prebiotic carbohydrate in its own right but is also used as a source of fructo-oligosaccharides by hydrolysis. Inulin, a fructan polymer is widely distributed in plants [264] and chicory (*Cichorium intybus*) is extensively used as a source of inulin for commercial purposes [265]. The roots of chicory contain 15 to 20% inulin and 5 to 10%

oligofructose, low amounts of glucose, fructose and sucrose have also been reported [266]. Inulin derived from chicory has a degree of polymerization of 2 to 60 with an average DP of 12, whereas the oligofructose component has a DP between 2 (F2) to 5 (GF4). A polydisperse carbohydrate, inulin exists as a linear β-(2→1) linked D-fructofuranose terminating in a glucose residue at the reducing end. The demand for fructan type compounds such as inulin and oligofructose is increasing due to the reported physiological benefits conferred on the host when these products are consumed. Such benefits include; modulation of the colonic microflora as inulin has purported prebiotic effects in vivo [256,267], enhances mineral absorption [268,269], modulates either the digestion and/or absorption or the metabolism of lipids [270] and exerts a positive influence on the body's defense mechanisms [265]. Although inulin occurs naturally in many fruits and vegetables, its quantity may be too low to beneficially influence the host, additionally cooking of these foods may influence the amount of inulin subsequently available, thus an attractive approach is the large scale manufacture of inulinoligosaccharides (IOS), a term which can be used interchangeably with fructooligosaccharides (FOS). It must be mentioned that FOS can be produced by one of two methods; (1) the hydrolysis of inulin by endoinulinases and (2) by enzymatic synthesis using sucrose as the substrate and a β-fructosidase possessing transfructosylation activity.

Using inulin as a substrate can yield two very different products depending on the enzyme used. An exo-inulinase (EC 3.2.1.26 β-D-fructofuranoside) is primarily used for the complete hydrolysis of inulin to fructose to produce an ultra high fructose syrup, whereas an endo-inulinase (EC 3.2.1.7 2→1-β-fructan-fructanohydrolase) attacks the internal linkages of inulin to produce oligosaccharides of varying DP. While extensive research is available on the production of fructose syrups there is less literature available on the manufacture of IOS, however this trend is changing due to the realization of the health benefits associated with IOS. Yazawa and Tamura [271] produced IOS through the acid treatment of inulin, although not an enzymatic process, IOS of different DP was produced. Results from a subsequent in vitro study suggested these oligomers of DP4 to 25 were selectively metabolized by *Bifidobacterium infantis*, a beneficial bacteria associated with host well being. Yeast, filamentous fungi and bacteria are all sources of inulinases their properties and applications have been reviewed elsewhere [272,273,274] and it is the action of the endo-inulinases which are of particular interest here. Such enzymes have been isolated from *Aspergillus*, *Pseudomonas* and *Xanthanomonas* sp.

Free enzyme versus immobilized enzyme can influence the yield of IOS, additionally an immobilized system would be favorable economically as the biocatalyst can be reused, enables continuous production and the end product is free of contamination. Kim et al. [275] intended to make a comparison between the reaction kinetics of free and immobilized endo-inulinases in a batch reactor however significant differences were observed in the reaction behavior and product composition due to the form of enzyme used and the initial concentration of substrate. Yun et al. [276] investigated the effect of inulin concentration on the production of IOS by free and immobilized endo-inulinase from *Pseudomonas* sp. Their findings corroborate those of Kim et al. [275] whereby different products are formed depending on the form of enzyme; a soluble enzyme yielded inulobiose and DP3 products, whereby the immobilized form predominantly produced inulobiose. As the concentration of inulin increased the yield of IOS did not increase in the soluble system and in the immobilized the yield remained the same. Although the enzyme was derived from *Pseudomonas* the immobilized form required a differ-

ent optimal pH to the soluble, pH 3.5 as opposed to pH 6, the lower pH resulted in a faster rate of hydrolysis of inulin, in addition a higher quantity of enzyme was used as the effectiveness factor is significantly decreased following immobilization. The yields of IOS produced by the free and immobilized endo-inulinase were 72 and 83%, respectively.

Form of enzyme and concentration of substrate influence the yield and type of IOS produced, however another study by Yun et al. [277] used whole cells of *Escherichia coli* HB101 whereby the gene for an endo-inulinase from *Pseudomonas* sp was expressed. In a batch reactor the production of IOS using the recombinant enzyme was compared to that of the native and there were differences in the DP distribution of the products formed although the yield of IOS was comparable. The native form produced less inulobiose and more DP4 and >DP4 products compared to the recombinant. Additionally Yun et al. [277] studied the effect of different reactor conditions; batch processing versus continuous using immobilized recombinant cells but no differences in product composition were observed.

Nakamura et al. [278] used immobilized endo-inulinase from *Aspergillus niger* in a continuous process to produce IOS. This particular enzyme was isolated from an endo-inulinase hyperproducing mutant 817 from the wild type strain *A. niger* 12 [279] which had previously been used in the production of fructose syrups by complete hydrolysis of inulin. Although the extracellular inulinase was partially purified, exo-inulinase activity remained, albeit at lower levels to the endo type. Various I/S ratios (I = endo-inulinase activity, S = exo-inulinase activity) were investigated (I/S = 4.5, I/S = 25 and I/S = 50) and the higher I/S ratio of 50 at the maximal flow rate of 6.0 ml min^{-1} produced a mixture of oligomers with a DP of 3–5.

Purified and partially purified endo-inulinases from *Aspergillus ficuum* for production of IOS were studied by Zhengyu et al. [280]. The yield of IOS was markedly different depending on the biocatalyst purity, the partially pure endo-inulinase produced an IOS yield of 50%, whereas the purified form produced an IOS yield of 70% in the same operating conditions. The product composition was also different; inulobiose – triose and tetraose where the predominant products of the partially purified form and the DP distribution for the purified enzyme ranged form DP2 to DP8 with relatively high levels of DP3 and DP4. The optimal substrate for the purified endo-inulinase was also determined. Pure inulin, Jerusalem artichoke powder and Jerusalem artichoke juice were used and the latter substrate produced an IOS yield of 80% during a 72-h reaction with a wide DP range (2–8), however no DP 6 was produced. A mixed enzyme system using endo-inulinases from *Xanthanomonos* and *Pseudomonas* sp is an interesting approach in the production of IOS [281].

3.1.2 Partial Hydrolysis of Dextran

Isomalto-oligosaccharides are commercial prebiotics in Japan [257] (Chap. 5.3). These are (1→6)-α-gluco-oligosaccharides with a degree of polymerization of two to five. The low DP means that the commercial product is partially metabolized by humans. Larger oligodextrans would be expected to have lower digestibility and increased colonic persistence. A promising manufacturing route to controlled molecular weight oligodextrans has been proposed using enzyme membrane reactors [282]. Such reactor systems use ultrafiltration membranes to retain enzymes in the system and to control the molecular weight of the product oligosaccharides (*Fig. 21*).

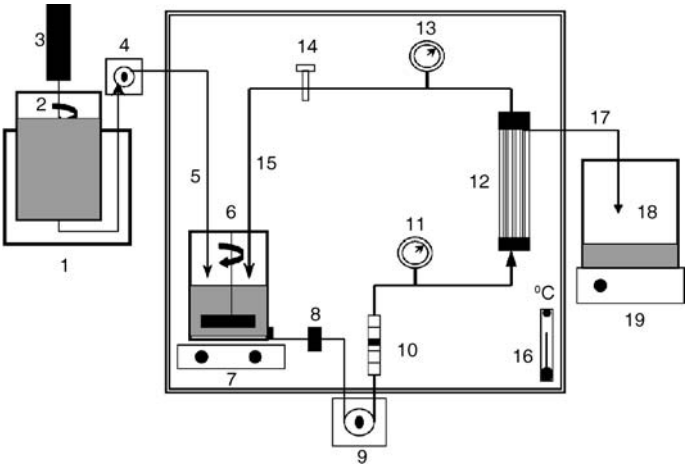


Figure 21
The continuous enzyme membrane reactor (CMR). (1): Temperature-controlled water-bath; (2): Feed tank; (3): Stirrer motor for feed tank; (4): Feed pump; (5): Feed inlet line to the reaction vessel; (6): Reaction vessel; (7): Magnetic stirring table; (8): Prefilter; (9): Recycle pump; (10): Flowmeter; (11): Membrane inlet pressure gauge; (12): Hollow fiber membrane cartridge; (13): Membrane outlet pressure gauge; (14): Pressure adjustment valve; (15): Retentate recycle line; (16): Air bath environment; (17): Permeate (product) line; (18): Permeate collection vessel; (19): Electronic balance

Table 1
Carbohydrate composition (%) of commercially available IMO product [285]

DP*	Components	IMO+
DP1	Glucose	3.8
DP2	Maltose	4.5
	Isomaltose	22.8
	Others**	13.1
DP3	Maltotriose	0.9
	Panose	11.6
	Isomaltotriose	16.7
DP4	Isomaltotetraose and others	17.7
DP5	Isomaltopentaose and others	7.2
≥DP6	Isomaltohexaose and others	1.7

+Isomalt 900® from Showa Sangyo Co. Ltd, Tokyo, Japan. *Denotes degree of polymerization. **Nigerose and kojibiose

In conventional production of IMO, starch is first hydrolyzed to maltooligosaccharides by α-amylase (EC 3.2.1.1) and pullulanase (EC 3.2.1.41). Then, α-glucosidase (EC 3.2.1.20) is added to convert the $\alpha 1 \rightarrow 4$ linked maltooligosaccharide into $\alpha 1 \rightarrow 6$ linked isomaltooligosaccharides. Finally, glucose is removed, giving a maximum IMO concentration of around 40%. About 40% glucose remains in the mixture [283,284]. ❯ *Table 1* [285] lists the composition of a commercial IMO product.

Oligodextran is a collective term used to indicate carbohydrate mixtures derived from dextran. It consists of an IMO part and a dextran part (< 70 kDa) [282], because of the resulting higher DP, oligodextrans are less digestible than IMO [286]. Commercially, they are either produced by acid hydrolysis of the native dextran from *Leuconostoc mesenteroides*, or by direct in vitro synthesis catalyzed by dextransucrase (EC 2.4.1.5) [287,288,289]. The dextran produced by *L. mesenteroides* NRRL B512F consists of 95% $\alpha 1 \rightarrow 6$ bonds, constituting the backbone of the molecule, and 5% side chains attached to the backbone through $\alpha 1 \rightarrow 3$ bonds [289]. Dextranases (1,6-α-D-glucanhydrolases EC 3.2.1.11) belong to the general class of hydrolytic enzymes and their name is derived from their catalytic function of hydrolyzing dextrans. These enzymes are used for the production of isomaltose or higher IMO's.

Dextranases from different sources exhibit diverse action patterns, but fundamentally they specifically hydrolyze the $\alpha(1 \rightarrow 6)$ glucosidic linkages of dextrans to produce primary isomlatose and some D-glucose as final products [290]. They are further divided into two classes, namely endodextranases and exodextranases. Endodextranases catalyze the rapid and random hydrolysis of the dextan $\alpha(1 \rightarrow 6)$ glucosidic linkages, irrespective of the position of these linkages in the dextran molecule (some selectivity exists), whereas endodextranases attack the $\alpha(1 \rightarrow 6)$ glucosidic bonds endwise liberating glucose [291]. The selectivity and limitation of the hydrolytic action of dextranases is attributed to their inability in hydrolyzing the branch linkages of dextrans ($\alpha(1 \rightarrow 3)$ bonds) [292]. In the case of endodextranases presence of excess branch linkages to a dextran molecule constitutes the polysaccharide highly resistant to degradation since the enzyme binding necessary for the hydrolytic action is inhibited [293]. The number of binding subsites of the endodextranases vary between 5–12 and a similar number of continuously $\alpha(1 \rightarrow 6)$ linked glucose residues, without any branch linkages in between are necessary for the attachment to take place [290]. Moreover the $\alpha(1 \rightarrow 6)$ glucosidic linkages in the vicinity of the branch linkages possess a greater degree of resistance towards hydrolysis compared to more distant ones [294]. Endodextranases have been isolated from extracellular fluids or cell extracts of fungi (e. g. *Penicillium funiculosum*, *P. lilacinum*, *Aspergillus carneous* and some of their mutants) and bacterial sources (*Lactobacillus bifidus*, *Cytophaga*, *Pseudomonas)* [292,293,295,296,297,298]. The most commonly used sources of endodextranases are from the Penicillium species in particular *P. funiculosum*, *P. lilacinum*, and *P. notatum* since they exert high enzyme yields and are easily purified [293,295,297].

Endodextranase type enzymes due to their rapid and random hydrolysis of dextrans are of interest for oligosaccharide production. Their inability to hydrolyze the branch linkages of dextrans determines the degree of hydrolysis and consequently gives rise to a variety of products depending on the frequency and the length of these branch points [292]. Bourne et al. [295,299] reported that *P. lilacinum* and *P. funiculosum* dextranses act upon a virtually unbranched dextran and the final products are glucose isomaltose and isomaltotriose with isomaltose being the main product and with the *P. funiculosum* enzyme producing slightly more glucose. Branched dextran hydrolysis results in a greater variety of products, which in the case of *L. mesenteroides* dextran are in the range between tri- to octasaccharide, linear or containing a branch linkage [294,299]. These products of the first hydrolysis can act as substrate for further hydrolysis but as the molecular weight decreases the rate of hydrolysis decreases [294]. In more recent studies [300] it was reported (for dextranse degradation of dextran hydro-gels) that the action of the enzyme could be divided into two steps. In the first step long unsubstituted chains (18 or

more) of glucose residues are hydrolyzed and in the second the hydrolysis products become the substrate for further breakdown. In studies concerning the rate of hydrolysis it was reported that the hydrolysis increases at relatively low substrate concentrations due to decreased viscosity and reduced product inhibition. High enzyme concentrations decreased the hydrolysis efficiency due to increased transglycosylation reactions, and presence of other low molecular weight sugars in the digests did not affect the enzyme activity or the total yield of hydrolysis products [301].

Mountzouris et al. [302] investigated the types of oligodextrans that would be generated via controlled enzymatic depolymerization of dextran using an endodextranase (Dextranase 50 L from *Penicillium lilacinum*). The reaction was carried out in an ultrafiltration stirred-cell membrane reactor fitted with a 10,000 MWCO membrane. It was found that the DE (dextrose equivalent) increased with increasing enzyme concentration and decreasing substrate concentration. At high enzyme concentrations, product DE values ranged from 18–38 and the oligosaccharide content of the products ranged from 27–82% (w/w). The oligosaccharides consisted mainly of isomaltose and isomaltotriose. At medium and low enzyme concentrations, the oligosaccharide yield was lower (30–40%) and the DE values of the oligosaccharides produced were also lower (in the range of 18–22). Anaerobic batch culture fermenters showed that dextran and oligodextrans supported growth of bifidobacteria with high levels of persistence for 48 h [303].

In a later study, Mountzouris et al. [282] investigated the continuous production of oligodextrans in a continuous stirred tank (CSTR) membrane reactor system. Substrate (industrial grade dextran from *Leuconostoc mesenteroides* B 512F) and enzyme were the same as those used previously. Oligodextran yields ranged between 84.4 and 98.7%. DE values ranged from 22–41 and the oligosaccharide content ranged from 55.9 to 93.4% (w/w), which was higher than in the batch membrane reactor. Three types of oligosaccharides with respect to their polysaccharide content were produced: a. < 15%, b. 15–30% and c. > 30% and < 44 %. The oligosaccharide content was found to be affected, not only by substrate and enzyme concentrations, but also by the residence time, higher residence time resulting in products with higher oligosaccharide content. The oligosaccharide content was fairly constant at 21.7–24.8% in all products, DP 4, 7, and 8 being the predominant ones. Isomaltotriose ranged from 21.6 to 35.2% (w/w). The novel oligodextrans generated in this work are expected to have lower digestibility than the already commercially available IMO and to give different physicochemical properties. Tanriseven and Dogan [304] produced IMO by using *L. mesenteroides* B 512 FM dextransucrase immobilized in alginate fibers. The product profile was the following: fructose (20.4%), glucose (21.2%), leucrose (3.8%), DP2 (4.7%), DP3 (5.1%), DP4 (6.6%), DP5 (6.4%), DP6 (5.7%), DP7 (4.4%), DP8 (2.2%), DP9 (2.6%), DP10 and bigger (16.9%). The authors concluded that the use of alginate fibers rather than beads leads to better performance and repetitive use.

3.1.3 Hydrolysis of Xylan to Produce Xylo-Oligosaccharides

XOS are manufactured by enzymatic hydrolysis of xylan from corn cobs, oat spelt xylan or wheat arabinoxylan [257]. Xylan, a principle type of hemicellulose exists as a linear polymer of β-D-xylopyranosyl units linked by 1→4 glycosidic bonds. Depending on the origin, different substituents such as arabinfuranosyl, 4-*O*-methylglucuronosyl and acetyl groups are present. Typical raw materials for XOS production include hardwoods, corn cobs, straws,

bagasses, hulls, malt cakes and bran [305]. Three different approaches exist for production of XOS: 1. enzyme treatment of lignicellulosic material (LCM), 2. chemical fractionation of LCM, with subsequent enzymatic hydrolysis and 3. hydrolytic degradation of xylan by steam, water or dilute acid solutions [305].

In the enzymatic production of XOS, low enzymatic activities are preferable in order to avoid degradation down to monosaccharides, di-, tri- and tetrasaccharides being the required products [305]. The enzyme used for XOS production is an endo-1-4-β-xylanase (1,4-β-D-xylan

Table 2
Summary of xylanase characteristics

Enzyme source	MW	Opt. pH	Opt. T (°C)	Best yield conditions	Products	Yield XOS
P. chyrosporium, G. trabeum, H. sheweinitzii, T. viride	Nd	4.5–5.0	40	Substrate: 1 g of hylglucuronoxylan	Aldotetraouronic acid from *T. viride**	75.7–91.3%
Trichoderma longibrachiatum	37.7 kDa	5.0–6.0	45	Oat spelt	Xylobiose, xylotetraose and higher DP XOSs	Nd
Robillarda sp.	23,400 Da	4.5–5.5	55		Xylobiose	
Irpex lacteus				Xyloglycans and aryl-xylooligosaccharides	1,4-β Glycosidic linkages, or 1,3 and 1,4 β xylosidic linkages	
T. viride				23 g reduced hardwood xylan at 40 °C for 24 h	4-*O*-Methyl-α-D-glucopyranosyl-(1-2)-D-xylotriose and 4-*O*-methyl-α-D-glucopyranosyl-(1-2)-D-xylotetraose	Neutral and acidic products were 13.0 g and 2.1 g, respectively
A. oryzae	Nd	Nd	Nd	Xylan (3 g)	Xylotriose, xylobiose, and xylose	1 g/L each
T. reesei	Nd	Nd	Nd	Unbleached birch kraft pulp, with endoxylanase 40 °C for 24 h at pH 5.5, with β-xylosidase at 40 °C for 24 h at pH 5.5 and then α-glucuronidase for 48 h at 40 °C	Acidic : (4ΔUA)-β-D-xylotetraose, (4ΔUA)-β-D-xylopentaose, (4-*O*-methyl-α-D-glucurono)-β-D-xylotetraose and (4-*O*-methyl-α-D-glucurono)-β-D-xylopentaose Neutral: were D- xylose, β-1,4-D-xylobiose and β-1,4-D-xylotriose	Nd

xylanohydrolase EC 3.2.1.8). This endo- acting enzyme cleaves the glycosidic bonds in the xylan backbone subsequently reducing the degree of polymerization of the substrate. Factors influencing the extent of hydrolysis include chain length, the degree of branching and presence of substituents. Additionally, purity of the enzyme is critical to reduce or prevent complete hydrolysis to xylose by contaminating β-xylosidases. *T. reesei*, *T. harzianum*, *T. viride* and *T. koningii* are known to produce xylanases [306]. *Table 2* lists other reported sources and extensive reviews on xylanases are available [307,308]. Fungal endo xylanases have an activity between 40–80 °C and between pH4–6.5. Extracellular enzymes from *Thermobifida fusca* were successfully used to produce XOS from various lignocellulolytic agricultural waste, however physical pretreatment of the substrate was a prerequisite to expose the xylosidic linkages to enable enzymatic hydrolysis [309]. Although various treatments are used (alkaline extraction, acidic treatment and cooking), for the production of XOS the acidic approach is undesirable as high quantities of xylose may be produced. An endo xylanase from *Aspergillus versicolor* has also been reported for production of XOS using xylan as substrate, interestingly no xylobiose or xylose were detected [310]. Techniques used to characterize the XOS include capillary zone electrophoresis and MALDI-TOF-MS [311].

3.1.4 Partial Hydrolysis of Pectins

Pectins are classified as dietary fiber and are reported to exert a beneficial effect on the gastrointestinal tract of the host (Chap. 5.3), thus the oligosaccharides derived from these may have similar or additional functionalities. Also, the oligosaccharides would be of a lower molecular weight, possibly more soluble than the native pectin as well as less viscous making them more amenable to be incorporated into foodstuffs. The approach of controlled partial hydrolysis applied to dextrans above has also been used to manufacture pectic oligosaccharides [312]. Enzyme membrane reactors were established to hydrolyze high methoxy citrus pectin (60–68% methylation) and low methoxy apple pectin (8% methylation) to low molecular weight oligosaccharides. A commercial endo-polygalacturonase (EC 3.2.1.15) from *Aspergillus pulverulentus* (Pectinas DL "Amano") chosen after screening several available preparations for the absence of pectin methylesterase (EC 3.1.1.11). Conversion to product was dependent on the degree of methylation where higher product conversions were achieved when the apple pectin was used as substrate compared to the citrus pectin. Yamaguchi et al. [313] studied the partial hydrolysis of citrus pectin (75% degree of esterification) using a crude enzyme mix from *Kluyveromyces fragilis*. Although a low molecular weight pectin was formed, the extent of hydrolysis was not comparable to that achieved by Olano-Martin et al. [312] as the molecular weight was 66,000 Daltons compared to the 3500 Dalton product from the continuous hydrolysis of a 60–60% methylated citrus pectin. It has been suggested that the pectins with a high degree of esterification are less susceptible to hydrolysis by endo-polygalacturonases than the pectin esterases, however other parameters such as temperature, enzyme concentration, origin and purity of enzyme, substrate concentration, pH, time and reactor will no doubt also influence the hydrolysis. The functionalities of the products produced by Yamaguchi et al. [313] and Olano-Martin et al. [312] have been investigated in vitro and although results are promising further in vitro studies would be required to substantiate a health claim before clinical trials are performed. The product oligosaccharides were found to have prebiotic potential using in vitro model fermentation systems [314]

and were characterized for their ability to inhibit the adhesion of *E. coli* verocytotoxins to human cells [315]. Lipid accumulation by the liver of rats was repressed, however serum cholesterol was not suggesting the molecular weight of the oligosaccharides could influence this.

3.1.5 Degradation of Bacterial Exopolysaccharide (EPS)

Bacterial EPS are a diverse group of polysaccharides comprised of a wide range of monosaccharide residues such as glucose, galactose and mannose. Historically most attention has been given to the EPS of Gram negative bacteria, particularly those of the various serotypes of *Klebsiella* spp. and *E. coli*, where the EPS carries the K antigenic specificity. Many of these EPS sequences carry potentially bioactive oligosaccharide sequences cryptically encoded in their primary structures. Many recognized receptor-active sequences can be found in these EPS molecules (*Table 3*). Economic methods of hydrolysis and isolation of these sequences would facilitate large scale production of receptor active oligosaccharides for subsequent development into more potent therapeutic derivatives [259,260].

Bacterial EPS tend to be rather refractory to most polysaccharide hydrolases [79]. A promising approach to preparation of oligosaccharides, however, is the use of bacteriophage. Many bacteriophage produce capsular depolymerase enzymes to degrade the EPS capsule and gain access to the bacterial cell surface. These enzymes are capable of reducing the EPS of their host organism to the repeating oligosaccharide unit. Many of these enzymes have been characterized [79,316] and they frequently display a very high degree of specificity for one particular linkage in the EPS sequence. They can either be hydrolases or lyases [79] although deacetylases (carboxylic ester or amide hydrolases) can also be found [317]. These cleave the acetyl substituents from the EPS.

No pure depolymerase enzymes are available commercially, and they have to be isolated from the bacteriophage particles [318]. Bacteriophage may be found in the natural habitat of their host bacteria, for example bacteriophages of the *Enterobacteriaceae* can be isolated from human or animal feces or in sewage [317,319]. It is in the base plate of the bacteriophage where the depolymerase is thought to be located [317]. However, it has also been discovered that in many of the systems studied the enzyme also exists as a soluble protein, found in the cell lysates following viral maturation. Yurewicz [320] has investigated the catalytic and molecular properties of enzyme from each source, finding that both are similar.

The exopolysaccharide can be broken down rapidly or slowly by endo- or exo- acting enzymes respectively. The majority of enzymes investigated to date are endoglycanases and endo- acting polysaccharide lyases [79]. Exceptions to this are the findings by Mishra et al. [321] who identified a polysaccharide which required both endo- and exo- acting enzymes to be degraded. Although phage enzymes are highly specific, it has been observed that some enzymes can act on more than one polysaccharide. Depolymerases from *E. coli* systems have been found to be active against *Aerobacter cloacae* polysaccharide and vice versa [322].

Various parameters need to be considered before performing EPS hydrolysis, namely substrate specificity, temperature, reaction medium, duration, and pH. Carboxyl reduction of EPS [323,324] prior to hydrolysis renders the bacteriophage enzyme inactive against the EPS, even though the polymer structure remains intact. A preliminary indication of the presence of depolymerase activity can be seen by the presence of a translucent halo surrounding the phage

Table 3
List of oligosaccharide repeating units resulting from bacteriophage hydrolysis (modified from Sutherland, 1999). Highlighted linkages represent anti-adhesive sequences

Oligosaccharide	Source
Glc[1]β1→4GlcAβ1→2Rhaα1→3Rhaα1→2Rhaα1→3Gal	Klebsiella K36 EPS
Gal[1]β1→4GlcAα1→3**Manα1→2Man**α1→3Gal	Klebsiella K74 EPS
Glcβ1→2Rhaα1→4GlcAα1→3Rha 3 ↑ Rhaα1	Klebsiella K17 EPS
Glcβ1→3Man[1]β1→4GlcAα1→3Manα1→3Galα1→3Gal	Klebsiella K46 EPS
Glcα1→4GlcAβ1→3Galα1→3Manα1→3Glc ↑ ↑ 2 2 Glcβ1 Glcβ1	Klebsiella K60 EPS
Galα1→3GalA[2]α1→2Fuc	Klebsiella K63 EPS
Man[1]β1→4GlcAα1→3Fucα1→3Glc	Klebsiella K6 EPS
Man[1]α1→4GalAα1→3**Manα1→2Man**α1→3Gal	Klebsiella K3 EPS
Gal[1]β1→4Glcβ1→6Glcα1→4GlcAα1→3**Manα1→2Man**α1→3Gal	Klebsiella K26 EPS
Manβ1→4GlcAβ1→2**Manα1→2Man**α1→3Gal	Klebsiella K43 EPS
Gal[1]α1→3Rhaα1→3**Galα1→4Gal**[1]α1→3Rhaα1→3Gal	E. coli K103 EPS
Manβ1→4GlcAβ1→3**Manα1→2Man**α1→3Glc	Klebsiella K24 EPS
Glcβ1→4GlcAβ1→2Manα1→4GlcAβ1→2Manα1→3Glc	Klebsiella K39 EPS
GlcAβ1→2Rhaα1→3Rhaα1→3Glcβ1→4Glc	Klebsiella K44 EPS
Glcα1→4GlcAβ1→2Manα1→3Glc 3 ↑ Galα1	***E. coli*** K39 EPS
GlcAβ1→2Rhaα1→4GlcNAcα1→6GalNAc	***E. coli*** K44 EPS
Gal[1]β1→4GlcAα1→3Manβ1→4Glcα1→3Glc	Klebsiella K13 EPS
GlcAα1→3Manβ1→4Glcα1→3Glc	Klebsiella K2 EPS
GlcA[1]β1→4Fucα1→3Glc	Klebsiella K1 EPS
Manα1→4GalAα→2Manα1→3Gal	Klebsiella K57 EPS

[1]Pyruvylated, [2]acetylated and formylated

plaques on confluent agar cultures of the host organism. In some studies the phage induced enzymes were recovered from cell lysates [323,324] for hydrolysis of the EPS and in others, crude preparations of phage were used [325,326,327]. Different temperatures have been used for EPS hydrolysis, 35 °C [328], 37 °C [323,325,326,329], and 40 °C [318]. The duration of the reaction has also varied. Dutton et al. [326] investigates hydrolysis at 24, 67, and 72 h.
Oligosaccharides prepared using this manufacturing technology can be used in in vitro assays to determine their anti-adhesive properties and might be suitable for anti-adhesion therapy as

they are likely to be non-immunogenic, are of low molecular weight and are soluble [258]. It is hoped that these compounds would be specific for the pathogen to be eliminated, unlike antibiotics which are not as specific, however extensive trials in vitro and in vivo would be required to confirm this.

References

1. Legler G (1990) Adv Carbohydr Chem Biochem 48:319
2. Sinnott ML (1990) Chem Rev 90:1171
3. Davies G, Henrissat B (1995) Structure 3:853
4. Davies GJ, Sinnott ML, Withers SG (1998) In: Sinnott ML (ed) Comprehensive biological catalysis: a mechanistic reference, vol 1. Academic Press, San Diego, p 119
5. Zechel DL, Withers SG (2000) Acc Chem Res 33:11
6. Hehre EJ (2000) Adv Carbohydr Chem Biochem 55:265
7. Henrissat B, Bairoch A (1993) Biochem J 293:781
8. Hehre EJ, Okada G, Genghof DS (1973) Advan Chem Ser 117:309
9. NC-IUBMB, Webb EC (1992) Enzyme nomenclature 1992: Recommendations of the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology on the nomenclature and classification of enzymes. Academic Press, San Diego
10. Wilson DB, Spezio M, Irwin D, Karplus A, Taylor J (1995) In: Sadler JN, Penner MH (eds) Enzymatic degradation of insoluble carbohydrates (ACS Symposium Series), vol 618. American Chemical Society, Washington, DC, p 1
11. Henrissat B (1991) Biochem J 280:309
12. Henrissat B, Bairoch A (1996) Biochem J 316:695
13. Henrissat B, Davies G (1997) Curr Opin Biotech 7:637
14. Henrissat B, Callebaut I, Fabrega S, Lehn P, Mornon JP, Davies G (1995) Proc Natl Acad Sci USA 92:7090
15. Mayans O, Scott M, Connerton I, Gravesen T, Benen J, Visser J, Pickersgill R, Jenkins J (1997) Structure 5:677
16. Svensson B, Søgaard M (1993) J Biotechnol 29:1
17. Ly HD, Withers SG (1999) Annu Rev Biochem 68:487
18. Davies G, Wilson KS, Henrissat B (1997) Biochem J 321:557
19. Quiocho FA (1993) Biochem Soc Trans 21:442
20. Johnson LN, Cheetham J, McLaughlin PJ, Acharya KR, Barfod D, Phillips DC (1988) Curr Topics Microbiol Immunol 139:81
21. Olsen K, Christensen U, Sierks MR, Svensson B (1993) Biochemistry 32:9686
22. Bundle DR, Sigurskjold BW (1994) Methods Enzymol 247:288
23. Peters T, Pinto BM (1996) Curr Opin Struct Biol 6:710
24. André G, Buléon A, Haser R, Tran V (1999) Biopolymers 50:751
25. Weimar T, Stoffer B, Svensson B, Pinto BM (2000) Biochemistry 39:300
26. Mayer C, Zechel DL, Reid SP, Warren RAJ, Withers SG (2000) FEBS Lett 466:40
27. Hiromi K (1970) Biochem Biophys Res Commun 40:1
28. Allen JD, Thoma JA (1976) Biochem J 159:121
29. Ajandouz EH, Abe J, Svensson B, Marchis-Mouren G (1992) Biochim Biophys Acta 1159:193
30. Suganuma T, Ohnishi M, Hiromi K, Nagahama T (1996) Carbohydr Res 282:171
31. Robyt JF, French D (1970) J Biol Chem 245:3917
32. Allen JD (1980) Methods Enzymol 64:248
33. Ajandouz EH, Marchis-Mouren G (1995) Carbohydr Res 268:267
34. Alkazaz M, Desseaux V, Marchis-Mouren G, Payan F, Forest E, Santimone M (1996) Eur J Biochem 241:787
35. Qian M, Haser R, Payan F (1995) Protein Sci 4:747
36. Qian M, Haser R, Buisson G, Duée E, Payan F (1994) Biochemistry 33:6284
37. Robyt JF, French D (1967) Arch Biochem Biophys 122:8
38. Srisodsuk M, Kleman-Leyer K, Keränen S, Kirk TK, Teeri TT (1998) Eur J Biochem 251:885
39. Väljämäe P, Sild V, Pettersson G, Johansson G (1998) Eur J Biochem 253:469
40. Kramhøft B, Svensson B (1998) In: Ballesteros A, Plou FJ, Iborra JL, Halling PJ (eds) Stability

and stabilization of biocatalysts. Elsevier, Amsterdam, p 343
41. Thoma JA, Spradlin J, Dygert S (1971) In: Boyer PD (ed) The Enzymes, vol 5. Academic Press, New York, p. 115
42. Nelson TE (1975) Biochim Biophys Acta 377:139
43. Nelson TE, Johnson Jr J, Jantzen E, Kirkwood S (1969) J Biol Chem 244:5972
44. Barras DR, Stone BA (1969) Biochim Biophys Acta 191:329
45. Aleshin AE, Stoffer B, Firsov LM, Svensson B, Honzatko RB (1996) Biochemistry 35:8319
46. Strokopytov B, Knegtel RMA, Penninga D, Rozeboom HJ, Kalk KH, Dijkhuizen L, Dijkstra BW (1996) Biochemistry 35:4241
47. Yoshioka Y, Hasegawa K, Matsuura Y, Katsube Y, Kubota M (1997) J Mol Biol 271:619
48. Hrmova M, MacGregor EA, Biely P, Stewart RJ, Fincher GB (1998) J Biol Chem 273:11134
49. Wind RD, Buitelaar RM, Dijkhuizen L (1998) Eur J Biochem 253:598
50. Mikami B, Yoon H-J, Yoshigi N (1999) J Mol Biol 285:1235
51. Yu SK, Bojsen K, Svensson B, Marcussen J (1999) Biochim Biophys Acta 1433:1
52. Mosi R, Withers SG (1999) Biochem J 338:251
53. Watson KA, McClerverty C, Geremia S, Cottaz S, Driguez H, Johnson LN (1999) EMBO J 18:4619
54. Bailey JM, French D (1957) J Biol Chem 226:1
55. Adachi M, Mikami B, Katsube T, Utsumi S (1998) J Biol Chem 273:19859
56. Barr BK, Hsieh Y-L, Ganem B, Wilson DB (1996) Biochemistry 35:586
57. Sakon J, Irwin D, Wilson DB, Karplus PA (1997) Nat Struct Biol 4:810
58. Armand S, Drouillard S, Schülein M, Henrissat B, Driguez H (1997) J Biol Chem 272:2709
59. Vernon CA (1967) Proc R Soc London Ser B 167:389
60. Koshland DE (1953) Biol Rev 28:416
61. McIntosh LP, Hand G, Johnson PE, Joshi MD, Korner M, Plesniak LA, Ziser L, Wakarchuk WW, Withers SG (1996) Biochemistry 35: 9958
62. McCarter JD, Withers SG (1994) Curr Opin Struc Biol 4:885
63. Tanaka Y, Tao W, Blanchard JS, Hehre EJ (1994) J Biol Chem 269:32306
64. Davies GJ, Dauter M, Brzozowski M, Bjørnvad ME, Andersen KV, Schülein M (1998) Biochemistry 37:1926
65. Damude HG, Ferro V, Withers SG, Warren RAJ (1996) Biochem J 315:467
66. Fang T-Y, Ford C (1998) Protein Eng 11:383
67. Nielsen JE, Beier L, Outzen D, Borchert TV, Frantzen HB, Andersen KV, Svendsen A (1999) Eur J Biochem 264:816
68. Ford C (1999) Curr Opin Biotechnol 10:353
69. Heightman TD, Vasella AT (1999) Angew Chem Int Ed 38:750
70. Varrot A, Schülein M, Pipelier M, Vasella A, Davies GJ (1999) J Am Chem Soc 121:2621
71. Sulzenbacher G, Driguez H, Henrissat B, Schülein M, Davies GJ (1996) Biochemistry 35:15280
72. Tews I, Terwisscha van Scheltinga AC, Perrakis A, Wilson KS, Dijkstra BW (1997) J Am Chem Soc 119:7954
73. Brameld KA, Shrader WD, Imperiali B, Goddard WA, III (1998) J Mol Biol 280:913
74. Terwisscha van Scheltinga AC, Armand S, Kalk KH, Isogai A, Henrissat B, Dijkstra BW (1995) Biochemistry 34:15619
75. Thunnissen A, Rozeboom HJ, Kalk KH, Dijkstra BW (1995) Biochemistry 34:12729
76. van Asselt EJ, Dijkstra AJ, Kalk KH, Takacs B, Keck W, Dijkstra BW (1999) Structure 7:1167
77. Weaver LH, Grutter MG, Matthews BW (1995) J Mol Biol 245:54
78. Matsumura I, Kirsch JF (1996) Biochemistry 35:1881
79. Sutherland IW (1995) FEMS Microbiol Rev 16:323
80. Pickersgill R, Jenkins J, Harris G, Nasser W, Robertbaudouy J (1994) Nat Struct Biol 1:717
81. Preston JF, III, Rice JD, Ingram LO, Keen NT (1992) J Bacteriol 174:203
82. Jurnak F, Kita N, Garrett M, Heffron SE, Scavetta R, Boyd C, Keen N (1996) In: Visser J, Voragen AGJ (eds) Pectin and Pectinases. Elsevier, Amsterdam, p 295
83. Scavetta RD, Herron SR, Hotchkiss AT, Kita N, Keen NT, Benen JAE, Kester HCM, Visser J, Jurnak F (1999) Plant Cell 11:1081
84. Henrissat B, Heffron SE, Yoder MD, Lietzke SE, Jurnak F (1995) Plant Physiol 107:963
85. Lehmann J, Schroeter E (1972) Carbohyd Res 23:359
86. Hehre EJ, Genghof DS, Sternlicht H, Brewer CF (1977) Biochemistry 16:1780
87. Yu SK, Ahmad T, Kenne L, Pedersen M (1995) Biochim Biophys Acta 1244:1
88. Reichenbecher M, Lottspeich F, Bronnenmeier K (1997) Eur J Biochem 247:262

89. Mieyal JJ, Abeles RH (1972) The Enzymes, vol 7. Academic Press, New York, p 515
90. Palm D, Klein HW, Schinzel R, Buehner M, Helmreich EJM (1990) Biochemistry 29:1099
91. Weinhausel A, Griessler R, Krebs A, Zipper P, Haltrich D, Kulbe KD, Nidetzky B (1997) Biochem J 326:773
92. Watson KA, Schinzel R, Palm D, Johnson LN (1997) EMBO J 16:1
93. Johnson LN, Weber IT, Wild DL, Wilson KS, Yeates DGR (1977) Proceedings of regulatory mechanisms of carbohydrate metabolism: Phosphorylase, vol 42, p 183
94. Fletterick RJ, Sygusch J, Murray N, Madsen NB, Johnson LN (1976) J Mol Biol 103:1
95. Lin K, Hwang PK, Fletterick RJ (1995) J Biol Chem 270:26833
96. Lin K, Rath VL, Dai SC, Fletterick RJ, Hwang PK (1996) Science 273:1539
97. Rath VL, Lin K, Hwang PK, Fletterick RJ (1996) Structure 4:463
98. Weber IT, Johnson LN, Wilson KS, Yeates DGR, Wild DL, Jenkins JA (1978) Nature 274:433
99. Johnson LN, Acharya KR, Jordan MD, McLaughlin PJ (1990) J Mol Biol 211:645
100. O'Reilly M, Watson KA, Schinzel R, Palm D, Johnson LN (1997) Nat Struct Mol Biol 4:405
101. Kadziola A, Søgaard M, Svensson B, Haser R (1998) J Mol Biol 278:205
102. Nakano K, Fukui T (1986) J Biol Chem 261:8230
103. Mori H, Tanizawa K, Fukui T (1993) Protein Sci 2:1621
104. Johnson LN (1992) FASEB J 6:2274
105. Sprang S, Goldsmith E, Fletterick R (1987) Science 237:1012
106. Sprang SR, Withers SG, Goldsmith EJ, Fletterick RJ, Madsen NB (1991) Science 254:1367
107. Johnson LN, Hajdu J, Acharya KR, Stuart DI, McLaughlin PJ, Oikonomakos NG, Barford D (1989) In: Herve G (ed) Allosteric Enzymes. CRC Press, Boca Raton, FL, p 81
108. Kasvinsky PJ, Madsen NB, Fletterick RJ, Sygusch J (1978) J Biol Chem 253:1290
109. Goldemberg SH, Marechal LR, De Souza BC (1966) J Biol Chem 241:45
110. Marechal LR (1967) Biochim Biophys Acta 146:417
111. Graves DJ, Wang JH (1972) In: Boyer PD (ed) The Enzymes. vol 7. Academic Press, New York, p 435
112. Newgard CB, Hwang PK, Fletterick RJ (1989) Crit Rev Biochem Mol Biol 24:69
113. Mitchell EP, Withers SG, Ermert P, Vasella AT, Garman EF, Oikonomakos NG, Johnson LN (1996) Biochemistry 35:7341
114. Martin JL, Johnson LN, Withers SG (1990) Biochemistry 29:10745
115. Withers SG, Madsen NB, Sprang SR, Fletterick RJ (1982) Biochemistry 21:5372
116. Goldsmith EJ, Sprang SR, Hamlin R, Xuong NH, Fletterick RJ (1989) Science 245:528
117. Johnson LN (1992) Protein Sci 1:1237
118. Duke EMH, Wakatsuki S, Hadfield A, Johnson LN (1994) Protein Sci 3:1178
119. McLaughlin PJ, Stuart DI, Klein HW, Oikonomakos NG, Johnson LN (1984) Biochemistry 23:5862
120. Barford D, Schwabe JWR, Oikonomakos NG, Acharya KR, Hajdu J, Papageorgiou AC, Martin JL, Knott JCA, Vasella A, Johnson LN (1988) Biochemistry 27:6733
121. Papageorgiou AC, Oikonomakos NG, Leonidas DD, Bernet B, Beer D, Vasella A (1991) Biochem J 274:329
122. O'Reilly M, Watson KA, Johnson LN (1999) Biochemistry 38:5337
123. Dong W, Jespersen T, Bols M, Skrydstrup T, Sierks MR (1996) Biochemistry 35:2788
124. Williams SJ, Hoos R, Withers SG (2000) J Am Chem Soc 122:2223
125. Waagepetersen HS, Westergaard N, Schousboe A (2000) Neurochem Int 36:435
126. Krulle TM, Watson KA, Gregoriou M, Johnson LN, Crook S, Watkin DJ, Griffiths RC, Nash RJ, Tsitsanou KE et al (1995) Tetrahedron Lett 36:8291
127. De la Fuente C, Krulle TM, Watson KA, Gregoriou M, Johnson LN, Tsitsanou KE, Zographos SE, Oikonomakos NG, Fleet GWJ (1997) Synlett 485
128. Gregoriou M, Noble MEM, Watson KA, Garman EF, Krulle TM, De La Fuente C, Fleet GWJ, Oikonomakos NG, Johnson LN (1998) Protein Sci 7:915
129. Zographos SE, Oikonomakos NG, Tsitsanou KE, Leonidas DD, Chrysina ED, Skamnaki VT, Bischoff H, Goldmann S, Watson KA, Johnson LN (1997) Structure 5:1413
130. Oikonomakos NG, Kontou M, Zographos SE, Watson KA, Johnson LN, Bichard CJF, Fleet GWJ, Acharya KR (1995) Protein Sci 4: 2469
131. Lundblad RL, Noyes CM (1984) Chemical reagents for protein modification. CRC Press, Boca Raton, FL

132. Clarke AJ, Svensson B (1984) Carlsberg Res Commun 49:559
133. Juge N, Rodenburg KW, Guo X-J, Chax J-C, Svensson B (1995) FEBS Lett 363:299
134. Kuriki T, Stewart Douglas C, Preiss J (1997) J Biol Chem 272:28999
135. Rodenburg KW, Vallée F, Juge N, Aghajari N, Guo X-J, Haser R, Svensson B (2000) Eur J Biochem 267:1019
136. Lalegerie P, Legler G, Yon JM (1982) Biochimie 64:977
137. Withers SG, Aebersold R (1995) Protein Sci 4:361
138. Thomas EW, McKelvy JF, Sharon N (1969) Nature 222:485
139. Eshdat Y, McKelvy JF, Sharon N (1973) J Biol Chem 248:5892
140. Eshdat Y, Dunn A, Sharon N (1974) Proc Nat Acad Sci USA 71:1658
141. Legler G (1977) Methods Enzymol 46:368
142. Rodriguez EB, Stick RV (1990) Aust J Chem 43:665
143. Høj PB, Rodriguez EB, Stick RV, Stone BA (1989) J Biol Chem 264:4939
144. Høj PB, Fincher GB (1995) Plant 7:367
145. Chen L, Fincher GB, Høj PB (1993) J Biol Chem 268:13318
146. Macarron R, van Beeumen J, Henrissat B, dela Mata I, Claeyssens M (1993) FEBS Lett 316:137
147. Klarskov K, Piens K, Stahlberg J, Høj PB, Vanbeeumen J, Claeyssens M (1997) Carbohydr Res 304:143
148. Legler G, Bause E (1973) Carbohyd Res 28:45
149. Clarke AJ, Strating H (1989) Carbohydr Res 188:245
150. Keitel T, Simon O, Borriss R, Heinemann U (1993) Proc Natl Acad Sci USA 90:5287
151. Nitta Y, Isoda Y, Toda H, Sakiyama F (1989) J Biochem (Tokyo) 105:573
152. Moult J, Eshdat Y, Sharon N (1973) J Mol Biol 75:1
153. Høj PB, Rodriguez EB, Iser JR, Stick RV, Stone BA (1991) J Biol Chem 266:11628
154. Sulzenbacher G, Schülein M, Davies GJ (1997) Biochemistry 36:5902
155. Withers SG, Umezawa K (1991) Biochem Biophys Res Commun 177:432
156. Tong MK, Ganem B (1988) J Am Chem Soc 110:312
157. Caron G, Withers SG (1989) Biochem Biophys Res Commun 163:495
158. Black TS, Kiss L, Tull D, Withers SG (1993) Carbohydr Res 250:195
159. Shul'man ML, Shiyan SD, Khorlin AY (1976) Biochim Biophys Acta 445:169
160. Sinnott ML (1982) CRC Crit Rev Biochem 12:327
161. Howard S, Withers SG (1998) J Am Chem Soc 120:10326
162. Howard S, Withers SG (1998) Biochemistry 37:3858
163. Marshall PJ, Sinnott ML, Smith PJ, Widdows D (1981) J Chem Soc Perkin Trans 1 366
164. Clausen TP, Keller JW, Reichardt PB (1990) Tetrahedron Lett 31:4537
165. Zhu J, Withers SG, Reichardt PB, Treadwell E, Clausen TP (1998) Biochem J 332:367
166. Halazy S, Berges V, Ehrhard A, Danzin C (1990) Bioorg Chem 18:330
167. Withers SG, Rupitz K, Street IP (1988) J Biol Chem 263:7929
168. McCarter JD, Withers SG (1996) J Biol Chem 271:6889
169. McCarter JD, Withers SG (1996) J Am Chem Soc 118:241
170. Høj PB, Condron R, Traeger JC, McAuliffe JC, Stone BA (1992) J Biol Chem 267:25059
171. Tull D, Burgoyne DL, Chow DT, Withers SG, Aebersold R (1996) Anal Biochem 234: 119
172. Ermert P, Vasella A, Weber M, Rupitz K, Withers SG (1993) Carbohydr Res 250:113
173. Varrot A, Schülein M, Davies GJ (1999) Biochemistry 38:8884
174. Dauter Z, Dauter M, Brzozowski AM, Christensen S, Borchert TV, Beier L, Wilson KS, Davies GJ (1999) Biochemistry 38:8385
175. Aleshin AE, Firsov LM, Honzatko RB (1994) J Biol Chem 269:15631
176. Brzozowski AM, Davies GJ (1997) Biochemistry 36:10837
177. McAuliffe JC, Stick RV, Tilbrook DMG, Watts AG (1998) Aust J Chem 51:91
178. Fairweather JK, Stick RV, Tilbrook DMG, Driguez H (1999) Tetrahedron 55:3695
179. Fairweather JK, Stick RV (1999) 10th European Carbohydrate Symposium. Conference Centre, National University of Ireland, Galway, Ireland, Poster Abstract 062
180. Søgaard M, Kadziola A, Haser R, Svensson B (1993) J Biol Chem 268:22480
181. Berland CR, Sigurskjold BW, Stoffer B, Frandsen TP, Svensson B (1995) Biochemistry 34: 10153
182. Wiegand G, Epp O, Huber R (1995) J Mol Biol 247:99

183. Bompard-Gilles C, Rousseau P, Rouge P, Payan F (1996) Structure 4:1441
184. Ishimoto M, Chrispeels MJ (1996) Plant Physiol 111:393
185. Strobl S, Maskos K, Betz M, Wiegand G, Huber R, Gomis-Rüth FX, Glockshuber R (1998) J Mol Biol 278:617
186. Vallée F, Kadziola A, Bourne Y, Juy M, Rodenburg KW, Svensson B, Haser R (1998) Structure 6:649
187. Pereira PJB, Lozanov V, Patthy A, Huber R, Bode W, Pongor S, Strobl S (1999) Structure 7:1079
188. Lu S, Deng P, Liu X, Luo J, Han R, Gu X, Liang S, Wang X, Li F, Lozanov V, Patthy A, Pongor S (1999) J Biol Chem 274:20473
189. MacGregor AW, Macri LJ, Schroeder SW, Bazin SL (1994) J Cereal Sci 20:33
190. Rouau X, Surget A (1998) J Cereal Sci 28:63
191. McLauchlan WR, Garcia-Conesa MT, Williamson G, Roza M, Ravenstein P, Maat J (1999) Biochem J 338:441
192. Leckie F, Mattei B, Capodicasa C, Hemmings A, Nuss L, Araci B, De Lorenzo G, Cervone F (1999) EMBO J 18:2352
193. Beguin P, Lemaire M (1996) Crit Rev Biochem Mol Biol 31:201
194. Brummell DA, Catala C, Lashbrook CC, Bennett AB (1997) Proc Natl Acad Sci USA 94:4794
195. Hrmova M, Harvey AJ, Wang J, Shirley NJ, Jones GP, Stone BA, Høj PB, Fincher GB (1996) J Biol Chem 271:5277
196. Varghese JN, Hrmova M, Fincher GB (1999) Structure 7:179
197. Mikami B, Degano M, Hehre EJ, Sacchattini JC (1994) Biochemistry 33:7779
198. Sabini E, Sulzenbacher G, Dauter M, Dauter Z, Jørgensen PL, Schülein M, Dupont C, Davies GJ, Wilson KS (1999) Chem Biol 6: 483
199. Varghese JN, Garrett TPJ, Colman PM, Chen L, Høj PB, Fincher GB (1994) Proc Natl Acad Sci USA 91:2785
200. Rouvinen J, Bergfors T, Teeri T, Knowles JK, Jones TA (1990) Dcience 249:380
201. Schmidt A, Gübitz GM, Kratky C (1999) Biochemistry 38:2403
202. Zou J, Kleywegt GJ, Ståhlberg J, Driguez H, Nerinckx W, Claeyssens M, Koivula A, Teeri TT, Jones TA (1999) Structure 7:1035
203. Machius M, Vértésy L, Huber R, Wiegand G (1996) J Mol Biol 260:409
204. Koivula A, Kinnari T, Harjunpaa V, Ruohonen L, Teleman A, Drakenberg T, Rouvinen J, Jones TA, Teeri TT (1998) FEBS Lett 429:341
205. Kim JS, Cha SS, Kim HJ, Ha NC, Oh ST, Cho HS, Cho MJ, Kim MJ, Lee HS, Kim JW, Choi KY, Park KH, Oh BH (1999) J Biol Chem 274:26279
206. Park KH, Kim TJ, Cheong TK, Kim JW, Oh BH, Svensson B (2000) Biochim Biophys Acta 1478:165
207. Spurlino JC, Rodseth LE, Quiocho FA (1992) J Mol Biol 226:15
208. Sulzenbacher G, Mackenzie LF, Wilson KS, Withers SG, Dupont C, Davies GJ (1999) Biochemistry 38:4826
209. Anderson MA, Stone BA (1975) FEBS Lett 52:202
210. Daniel AS, Martin J, Vanat I, Whitehead TR, Flint HJ (1995) J Appl Bacteriol 79:417
211. Iseli B, Armand S, Boller T, Neuhaus J-M, Henrissat B (1996) FEBS Lett 382:186
212. Rupley JA (1967) Proc R. Soc London Ser B 167:416
213. Hindsgaul O, Khare DP, Bach M, Lemieux RU (1985) Can J Chem 63:2653
214. Sierks MR, Svensson B (1992) Protein Eng 5:185
215. Frandsen TP, Stoffer BB, Palcic MM, Hof S, Svensson B (1996) J Mol Biol 263:79
216. Uitdehaag JCM, Mosi R, Kalk KH, van der Veen BA, Dijkhuizen L, Withers SG, Dijkstra BW (1999) NatStrucBiol 6:432
217. Zhou J (1999) Biotechnol Bioeng 62:618
218. Hayashida S (1975) Agric Biol Chem 39: 2093
219. Svensson B, Larsen K, Svendsen I, Boel E (1983) Carlsberg Res Commun 48:529
220. Svensson B, Jespersen H, Sierks MR, MacGregor EA (1989) Biochem J 264:309
221. Sorimachi K, Jacks AJ, Le Gal-Coëffet M-F, Williamson G, Archer DB, Williamson MP (1996) J Mol Biol 259:970
222. Bayer EA, Chanzy H, Lamed R, Shoham Y (1998) Curr Opin Struct Biol 8:548
223. Raghothama S, Simpson PJ, Szabo L, Nagy T, Gilbert HJ, Williamson MP (2000) Biochemistry 39:978
224. Ponyi T, Szabo L, Nagy T, Orosz L, Simpson PJ, Williamson MP, Gilbert HJ (2000) Biochemistry 39:985
225. Fernandes AC, Fontes C, Gilbert HJ, Hazlewood GP, Fernandes TH, Ferreira LMA (1999) Biochem J 342:105

226. Watanabe T, Kasahara N, Aida K, Tanaka H (1992) J Bacteriol 174:186
227. Ochiai M, Ashida M (2000) J Biol Chem 275:4995
228. Nogawa M, Yatsui K, Tomioka A, Okada H, Morikawa Y (1999) Appl Environ Microbiol 65:3964
229. Gill J, Rixon JE, Bolam DN, McQueen-Mason S, Simpson PJ, Williamson MP, Hazlewood GP, Gilbert HJ (1999) Biochem J 342:473
230. Meissner K, Wassenberg D, Liebel W (2000) Mol Microbiol 36:898
231. Munoz IG, Stahlberg J, Divne C, Kleywegt G, Johansson G, Pettersson G, Mowbray S, Jones TA (1999) 3rd Carbohydrate Bioengineering Conference, Newcastle-upon-Tyne, Poster Abstract 4, 7
232. Linder M, Mattinen M-L, Kontteli M, Lindeberg G, Staahlberg J, Drakenberg T, Reinikainen T, Pettersson G, Annila A (1995) Protein Sci 4:1056
233. Linder M, Nevanen T, Teeri TT (1999) FEBS Lett 447:13
234. Din N, Damude HG, Gilkes NR, Miller Jr RC, Warren RAJ (1994) Proc Natl Acad Sci USA 91:11383
235. Jervis EJ, Haynes CA, Kilburn DG (1997) J Biol Chem 272:24016
236. Zeltins A, Schrempf H (1997) Eur J Biochem 246:557
237. Sorimachi K, Le Gal-Coëffet MF, Williamson G, Archer DB, Williamson MP (1997) Structure 5:647
238. Williamson MP, Le Gal-Coëffet MF, Sorimachi K, Furniss CSM, Archer DB, Williamson G (1997) Biochemistry 36:7535
239. Southall SM, Simpson PJ, Gilbert HJ, Williamson G, Williamson MP (1999) FEBS Lett 447:58
240. Mikami B, Adachi M, Kage T, Sarikaya E, Nanmori T, Shinke R, Utsumi S (1999) Biochemistry 38:7050
241. Payre N, Cottaz S, Boisset C, Borsali R, Svensson B, Henrissat B, Driguez H (1999) Angew Chem Int Ed 38:974
242. Sigurskjold BW, Christensen T, Payre N, Cottaz S, Driguez H, Svensson B (1998) Biochemistry 37:10446
243. Fierobe HP, Pages S, Belaich A, Champ S, Lexa D, Belaich JP (1999) Biochemistry 38: 12822
244. Ciruela A, Gilbert HJ, Ali BRS, Hazlewood GP (1998) FEBS Lett 422:221
245. Boisset C, Chanzy H, Henrissat B, Lamed R, Shoham Y, Bayer EA (1999) Biochem J 340:829
246. Pugsley AP (1993) Microbiol Rev 57:50
247. Saier Jr MH (1994) Microbiol Rev 58:71
248. Walter S, Rohde M, Machner M, Schrempf H (1999) Appl Environ Microbiol 65:886
249. Possot OM, Gerard-Vincent M, Pugsley AP (1999) J Bacteriol 181:4004
250. Bork P, Doolittle RF (1992) Proc Natl Acad Sci USA 89:8990
251. Zeikus JG (1996) In: Park KH, Robyt JF, Choi Y-D (eds) Enzymes for carbohydrate engineering. Elsevier, Amsterdam, p 145
252. Jaenicke R, Bohm G (1998) Curr Opin Struct Biol 8:738
253. Bush DS, Sticher L, van Huystee R, Wagner D, Jones RL (1989) J Biol Chem 264:19392
254. Machius M, Declerck N, Huber R, Wiegand G (1998) Structure 6:281
255. Bertoft E, Andtfolk C, Kulp SE (1984) J Inst Brew 90:298
256. Gibson GR, Roberfroid MB (1995) J Nutr 125:1401
257. Playne MJ, Crittenden R (1996) Bull IDF 31:10
258. Zopf D, Roth S (1996) The Lancet 347:1017
259. Mulvey G, Kitov PI, Marcato P, Bundle DR, Armstrong GD (2001) Biochimie 83**:**841
260. Armstrong GD, Rowe PC, Goodyer P, Orrbine E, Klassen TP, Wells G, MacKenzie A, Lior H, Blanchard C, Auclair F, Thompson B, Rafter DJ, McLaine PN (1995) J Infect Dis 171:1042
261. Ofek I, Hasty DL, Doyle R (ed) (2003) Bacterial adhesion to animal cells and tissues. ASM Press, Washington, DC
262. Casci T, Rastall RA (2006) In: Gibson GR, Rastall RS (eds) Prebiotics: development and application. Wiley, Chichester, p 29
263. Rastall RA, Maitin V (2002) Curr Opin Biotechnol 13:490
264. Van Loo J, Coussement P, De Leenheer L, Hoebregs H, Smits G (1995) Critic Rev Food Sci Nutr 35:525
265. Roberfroid M (2005) Br J Nutr 93:S13
266. Niness KR (1999) J Nutr 129:1402S
267. Harmsen HJM, Raangs GC, Franks AH, Wildeboer-Veloo CM, Welling GW (2002) Microb Ecol Health Dis 14:211
268. Fairweather-Tait SJ, Johnson IT (1999) Bioavailability of minerals. In: Gibson GR, Roberfroid MB (eds) Colonic microbiota: nutrition and health. Kluwer Academic, Dordrecht, p 233
269. Cashman K (2002) Br J Nutr 87:S169

270. Delzenne NM, Daubioul C, Neyrinck M, Lasa M, Taper HS (2002) Br J Nutr 87:S255
271. Yazawa K, Tamura Z (1982) Bifidobacteria Nicroflora 1:39
272. Vandamme EJ, Derycke DG (1983) Adv Appl Microbiol 29:139
273. Pandey A, Soccol CR, Selvakumar P, Soccol VT, Krieger N, Fontana JD (1999) Appl Biochem Biotechnol 81:55
274. Singh P, Gill PK (2006) Food Technol Biotechnol 44:151
275. Kim DH, Cho YJ, Song SK, Yun JW (1997) Biotechnol Lett 19:369
276. Yun JW, Kim DH, Yoon HB, Song SK (1997) J Ferment Bioeng 84:365
277. Yun JW, Yong YJ, Song CH, Song SK (1999) J Biosci Bioeng 87:291
278. Nakamura T, Ogata Y, Kamo Y, Hirayama, Ohta K (2001) Food Sci Technol Res 7:145
279. Nakamura T, Nagatomo Y, Hamada S, Nishino Y, Ohta K (1994) J Ferment Bioeng 78:134
280. Zhengyu J, Jing W, Bo J, Xueming X (2005) Food Res Inter 38:301
281. Cho YJ, Sinha J, Park JP, Yun JW (2001) Enz Microbiol Tech 29:428
282. Mountzouris KC, Gilmour SG, Rastall RA (2001) J Food Sci 67:1767
283. Kohmoto T, Fukui F, Takaku H, Machida Y, Arai M, Mitsuoka T (1998) Bifidobact Microflora 7:61
284. Kuriki T, Yanase M, Takata H, Takesada Y, Imanaka T, Okada S (1993) Appl Environ Microbiol 59:953
285. Kaneko T, Yokoyama A, Suzuki M (1995) Biosci Biotechnol Biochem 59:1190
286. Mountzouris KC, Gilmour SG, Jay AJ, Rastall RA (1999) J Appl Microbiol 87:546
287. Robyt JF (1992) In: Alexander RJ, Zobel HF (eds) Developments in carbohydrate chemistry. American Association of Cereal Chemists, Minnesota, p 261
288. Alsop RM (1983) In: Bushell ME (ed) Progress in industrial microbiology. Elsevier, Amsterdam, p 1
289. Gasciolli V, Choplin L, Paul F, Monsan P (1991) J Biotechnol 19:193
290. Franssen O, van Ooijen RD, de Boer D, Maes RAA, Herron JN, Hennink WE (1997) Macromol 30:7408
291. Wheatley MA, Moo-Young M (1977) Biotechnol Bioeng 19:219
292. Sidebotham RL (1971) Dextrans. In: Tipson RS, Horton D (eds) Advances in carbohydrate chemistry and biochemistry. Academic Press, New York, p 371
293. Das DK, Dutton SK (1996) J Biochem Cell Biol 28:107
294. Walker GJ, Dewar MD (1975) Carbohydr Res 39:303
295. Bourne EJ, Hutson DH, Weigel H (1962) Biochem J 85:158
296. Bailey RW, Hutson DH, Weigel H (1961) Biochem J 80:514
297. Szczodrak J, Pleszczynska M, Fiedurek J (1994) J Industrial Mircobiol 13:315
298. Richards GN, Streamer M (1978) Carbohyr Res 62:191
299. Bourne EJ, Hutson DH, Weigel H (1963) Biochem J 86:555
300. Franssen O, van Oojen RD, de Boer D, Maes RAA, Hennink WE (1999) Macromol 32:2896
301. Pleszczynska M, Szczodrak J, Roglaski J (1997) Mycological Res 101:69
302. Mountzouris KC, Gilmour SG, Grandison, AS, Rastall EA (1999) Enzyme Microbiol Tech 24:75
303. Olano-Martin E, Mountzouris KC, Gibson GR, Rastall RA (2000) Br J Nutr 83:247
304. Tanriseven A, Dogan S (2002) Process Biochem 37:1111
305. Vazquez MJ, Alonso JL, Dominguez H, Parajo JC (2000) Trends Food Sci Technol 11:387
306. Chen C, Chen JL, Lin TY (1997) Enzyme Microbiol Technol 21:91
307. Polizeli MLTM, Rizzatti ACS, Monti R, Terenzi HF, Jorge JA, Amorium DS (2005) Appl Microbiol Biotechnol 67:577
308. Biely P, VrSanskh M, Tenkanen M, Kluepfel D (1997) J Biotechnol 57:151
309. Yang CG, Yang SF, Liu WH (2007) J Agric Food Chem 55:3955
310. Carmona EC, Brochette-Braga MR, Pizzirani-Kleiner AA, Jorge JA (1998) FEMS Microbiol Lett 166:311
311. Rydlund A, Dahlman O (1997) Carbohydr Res 300:95
312. Olano-Martin E, Mountzouris KC, Gibson GR, Rastall RA (2001) J Food Sci 66:966
313. Yamaguchi F, Shimizu N, Hatanaka C (1994) Biosci Biotech Biochem 58:679
314. Olano-Martin E, Gibson GR, Rastall RA (2002) J Appl Microbiol 93:505
315. Olano-Martin E, Williams MR, Gibson GR, Rastall RA (2003) FEMS Microbiol Lett 218:101

316. Rieger-Hug D, Stirm S (1981) Virology 113:363
317. Geyer H, Himmelspach K, Kwiatkowski B, Schlecht S, Stirm S (1983) Pure Appl Chem 55:637
318. Stirm S (1994) Examination of the repeating units of bacterial exopolysaccharides. In: BeMiller JN, Manners DJ, Sturgeon RJ (eds) Methods in carbohydrate chemistry, vol X. Wiley, London, p 143
319. Hughes KA, Sutherland IW, Clark J, Jones MV (1998) J Appl Micro 85:583
320. Yurewicz EC, Ghalambor MA, Duckworth DH, Heath EC (1971) J Biol Chem 246:5607
321. Mishra C, Robbins PW (1995) Glycobiology 5:645
322. Sutherland IW, Wilkinson JF (1965) J Gen Microbiol 39:373
323. Sutherland IW (1971) J Gen Microbiol 70:331
324. Sutherland IW (1971) European J Biochem 23:582
325. Niemann H, Beilharz H, Stirm S (1978) Carbohydr Res 60:353
326. Dutton GGS, Di Fabio JL, Leek DM, Merrifield EH, Nunn JR, Stephen AM (1981) Carbohydr Res 97:127
327. Elsässer-Beile U, Stirm S (1981) Carbohydr Res 88:315
328. Sutherland IW (1976) J Gen Microbiol 94:211
329. Thurow H, Niemann H, Stirm S (1975) Carbohydr Res 41:257

Part 12
Glycomedicine

12.1 Novel Approaches for Glycodrug Discovery

Hirosato Kondo
Discovery Research Laboratories, Shionogi & Co., Ltd., 12-4, Sagisu 5-chome, Fukushima-ku 553-0002, Osaka, Japan
hirosato.kondou@shionogi.co.jp

Abstract

This chapter will give you an overview of recent approaches for drug development, focusing on the function and roles of carbohydrates. ❯ *Section 2* will outline the biochemical functions of polysaccharide (neutral polysaccharide, chitin and chitosan, glycosaminoglycan, and synthetic polysaccharides) and glycopolymers and the perspectives of their medicinal/medical use. ❯ *Section 3* will outline recent biopharmaceutical research and development, utilizing

In: *Glycoscience*. Fraser-Reid B, Tatsuta K, Thiem J (eds)
Chapter-DOI 10-1007/978-3-540-30429-6_61:

the functions of carbohydrates. *Section 4* will outline drug discovery and development of new medicines targeting carbohydrates. In addition, recent advances in carbohydrate synthesis technology, needed for development of new medicines, will also be described.

Keywords

Glycodrug; Carbohydrate; Polysaccharides; Glycoprotein; Biopharmaceutical; Therapeutic antibody; Lysosomal disease; Lectin; Vaccine

Abbreviations

ADCC antibody-dependent cellular cytotoxicity
CHO chinese hamster ovary
DDS drug delivery systems
EPO erythropoietin
FGF fibroblast growth factors

1 Introduction

Technology for carbohydrate synthesis and analysis has been advanced remarkably in recent years. Studies have been actively performed for the purpose of elucidating the properties and importance of carbohydrate which is one of the three major macromolecules existing in the biological body (DNA, protein, and carbohydrate) for the purpose of creating new medicines through such studies. It is known that the carbohydrate not only serves as a source of energy for organisms but also is involved in intercellular signal transduction and immune system through forming complex carbohydrates (carbohydrate bound to protein or lipid). Furthermore, carbohydrate is known to be closely involved in sicknesses such as infection and cancer. For these reasons, research on drug development targeting carbohydrates has been actively performed. In addition, research on development of drug delivery systems (DDS) through attaching carbohydrate to various medicines has been carried out, making use of the unique tissue transfer and host protein-protective function of carbohydrate in vivo. This chapter will outline recent approaches for drug development, focusing on the function and roles of carbohydrates.

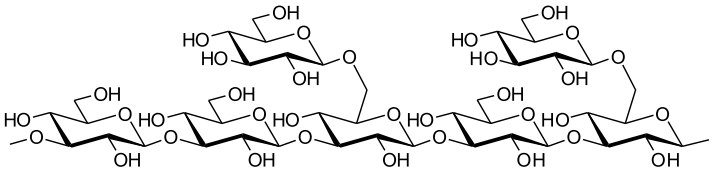
2 Polysaccharide-Based Glycodrugs

2.1 Introduction

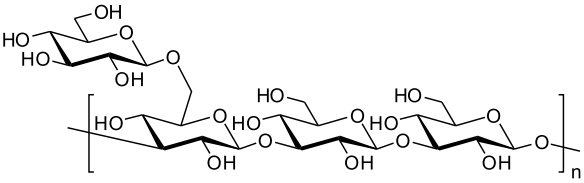
Polysaccharides commonly found in nature (starch, cellulose, mannan, xylan, laminaran, gum, etc.) are playing an important role not only in the food industry but also for use as materials. However, recent research tends to focus on applying these polysaccharides as substances with diverse biochemical functions or medicines. This section will outline the biochemical functions of polysaccharides and synthetic carbohydrate-bound polymers and the perspectives of their medicinal/medical use.

2.2 Neutral Polysaccharides

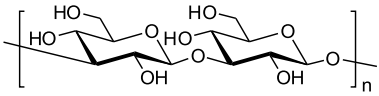
The most famous example of non-specific polysaccharide immunoadjuvants is lentinan isolated from *Lentinus edodes*. Since 1985, lentinan has been used as an antitumor agent. Lentinan is a (1→3)-β-glucopyranan bearing (1→6)-β-glucopyranoside branches (*Scheme 1*). This neutral polysaccharide suppresses the proliferation of sarcoma 180 tumors implanted subcutaneously in mice. Its antitumor activity was due to a host-mediated reaction with participation of thymus or thymus-dependent cells [1]. The lymphokines induced by this polysaccharide enhance the activity of lymphocytes (e. g., macrophages, helper T cells, and/or natural killer cells) in a dramatic manner [2,3,4,5]. Lentinan-like (1→3)-β-glucopyranan (schizophyllan) contained in the insoluble fraction of *Schizophyllum commune* has the structure shown in *Scheme 2* and is known to possess antitumor activity [6].



Scheme 1



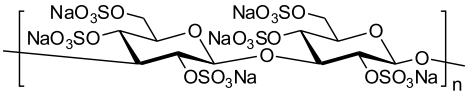
Scheme 2



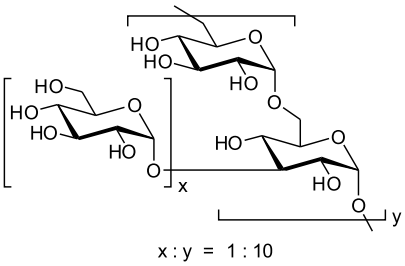
Scheme 3

Curdlan, which is a linear (1→3)-β-D-glucan (*Scheme 3*), also suppresses the proliferation of subcutaneously implanted sarcoma 180 markedly [7]. Curdlan sulfate with a high degree of substitution (*Scheme 4*) is reported to exert strong anti-human immunodeficiency virus (anti-HIV) activity [8,9].

Dextran (*Scheme 5*) solutions have been used as a plasma volume expander since 1947 owing to their non-immunogenic and well-tolerated nature as plasma substitutes [10]. Furthermore, since dextran suppresses erythrocyte aggregation and reduces blood viscosity, it has



Scheme 4

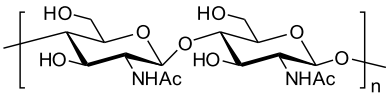


Scheme 5

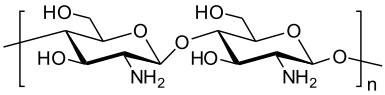
been reported to be useful for improving blood flow [11,12]. Dextran sulfates work as heparin-like anticoagulation agents and it has been reported that dextran sulfates strongly inhibited the replication of human immunodeficiency virus 1 (HIV-1) in vitro [13,14].

2.3 Chitin and Chitosan

Chitin is a cellulose-like (1→4)-β-D-glycan bearing an acetamide functional group at the C2 position of pyranose (*Scheme 6*), and is widely distributed in nature. Although chitin has been reported as a potentially biocompatible molecular material, it has a disadvantage of poor solubility in common organic solvents due to its extremely rigid crystalline structure [15]. Chitosan, derived from chitin by complete *N*-deacetylation under suitable alkaline conditions, has much improved solubility and has been used as a versatile intermediate to prepare a variety of functional derivatives (*Scheme 7*) [16]. Partially *N*-deacetylated chitin and some chitooligosaccharides exhibited significant immunoadjuvant activity in mice [17,18,19]. This finding accelerated the applications of chitin-related compounds in the biomedical and pharmaceutical fields. Immunological investigations on chitin and chitosan provide valuable information pertaining to activation of peritoneal macrophages, suppression of tumor growth and host protection against bacterial infection [20]. Minami et al. reported that when chitin or chitosan was used as a biological filling agent after surgical removal of tumor cells, it allowed rapid formation of granular tissue and remarkable improvement of surgical applications with no side effects [21,22,23].



Scheme 6

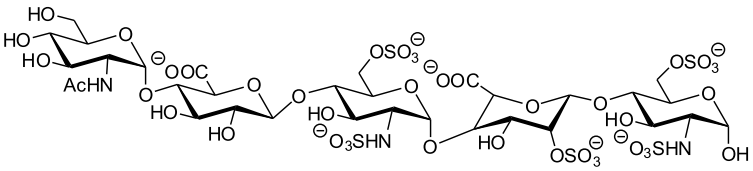


Scheme 7

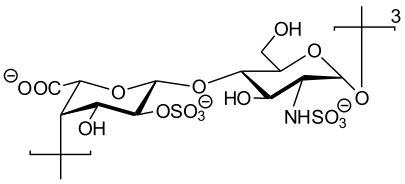
2.4 Glycosaminoglycans

Heparan sulfate proteoglycans, wich contain heparin, are ubiquitously distributed on the cell surface and in extracellular matrix. Interactions between heparan sulfate and various protein ligands are mediated by the specific oligosaccharide structures of polysaccharides (*Scheme 8*, *Scheme 9* and *Scheme 10*). The pentasaccharide sequence of heparin, bound to antithrombin III (AT-III), is a well-known example (*Scheme 8*). This sequence is indispensable for this kind of compound to serve as an anticoagulant clinically used for the prevention and treatment of thromboembolic disease [24]. Basic and acidic fibroblast growth factors (bFGF and aFGF) are known to bind to a dual receptor system composed of heparan sulfate proteoglycans and tyrosine kinase receptors on the cell surface. The binding structure of heparin sulfates with bFGF or aFGF are shown in *Scheme 9* and *Scheme 10*, respectively. Because basic and acidic FGF are additionally involved in the growth of tissue and blood vessels in fetuses, children and adults, and in the ectopic vascularization seen in patients with cancer or diabetic retinopathy, close attention has been paid to heparan sulfate that interacts with FGF. Development of drugs which can control these interactions is expected to cultivate a new dimension of treatment for these diseases [25].

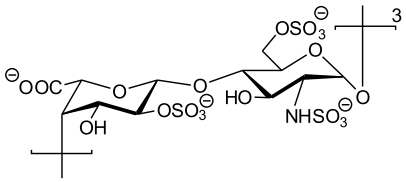
Hyaluronic acid is a non-sulfated and free glycosaminoglycan chain composed of D-glucuronic acid (GlcA) and *N*-acetyl-D-glucosamine (GlcNAc) (*Scheme 11*). Like collagen, fibronectin and heparan sulfate proteoglycan, hyaluronic acid is one of the most important macromolecules found in extracellular matrix. A highly viscous solution containing hyaluronic acid is suited for shock absorbing functions in cartilage. Recently, a high molecular mass



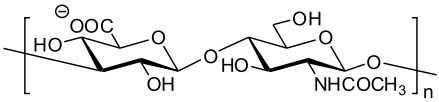
Scheme 8



Scheme 9



Scheme 10

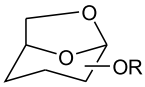


Scheme 11

hyaluronic acid named "Artz" (600,000–1,200,000) was developed as a drug effective for osteoarthritis of the knee and frozen shoulder syndrome. Artz has been shown to protect the articular cartilage, to inhibit unusual emigration of cartilage matrix and to alleviate the severe pain in arthrotic diseases.

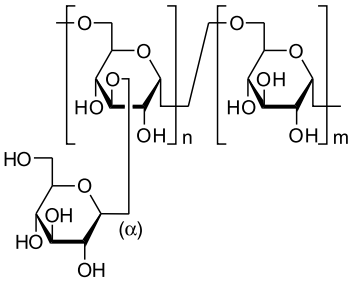
2.5 Synthetic Polysaccharides and Glycopolymers

Chemical synthesis is a promising strategy for creating bioactive polysaccharides and their analogues by utilizing polysaccharides found in nature because chemical synthesis is a powerful method not only for the construction of natural polysaccharides but also for the design of non-natural polysaccharide analogues having tailored biological functions. Polysaccharides and their derivatives can be synthesized if monosaccharides or their derivatives or their promoters are subjected directly to polycondensation reactions in the presence of appropriate condensation reagents. However, this type of synthesis seems to require much labor and time for regioselective protection and deprotection. Ring-opening polymerization of sugar orthoesters or anhydrosugar derivatives represents one of the best methods for the construction of stereoregular polysaccharides with high molecular masses. For example, branched type (1→6)-α-D-glucopyranan with remarkable hypoglycemic activity was synthesized by polymerization of 1,6-anhydrosugar derivative (*Scheme 12*) and subsequent chemoenzymatic manipulations (*Scheme 13*) [26,27]. The versatility of 1,6-anhydrosugar derivatives was demonstrated by the synthesis of a variety of stereoregular and bioactive polysaccharides with high molecular masses. In recent years, a new type of polymerizable glycoside has been often utilized for synthesis of glycopolymers with carbohydrate branches [28]. Glycopolymer obtained from acrylamide glycosides has excellent solubility in water and has been used as ligands of glycosylated proteins such as lectins and glycosyltransferases [29]. Glycopolymers containing sialic acid or sialo-oligosaccharide have been synthesized and used as potent inhibitors of influenza virus infection. Cross-linked glycopolymers and polystyrene-based glycopolymers are suitable for research on cell adhesion, cell culture and designing polymer films and fibers used as virus filters utilizing their specific carbohydrate-protein interaction. Among



R = protective group

Scheme 12



Scheme 13

numerous neoglycoconjugates, glycopolymers are the simplest in design, the most facile of the diverse types and the easiest to synthesize for use in biomedical research and practical utilization.

3 Development of Next-Generation Biopharmaceuticals by Glycoengineering

3.1 Introduction

Many of the proteins found in the biological body have complex carbohydrates. Modification with carbohydrates is increasingly acknowledged as one of the most important "post-translational modifications" determining the fate of the protein, as is the case with phosphorylation. Although the functions of carbohydrates bound to proteins in vivo have not yet been fully clarified, carbohydrates are not a mere decoration but are known to contribute greatly to stabilization of protein structure, improvement in water solubility, and elevation of the degree of diversity and selectivity of bioreactions. In recent years, biopharmaceutical research and development, utilizing the functions of carbohydrates, has been actively carried out. Examples of practical application of the findings from such research are presented briefly in this section.

3.2 Functionalization of Protein-Based Biopharmaceuticals by Glycoengineering

Various bioactive proteins produced by genetic engineering (e. g., recombinant cytokines, hormones, and so on) have already been clinically used. However, these bioactive proteins involve

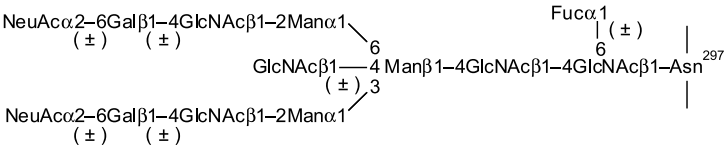
various problems. How to resolve the problems of instability in the biological body and short half-life in blood are important open issues. Carbohydrate modification has recently been attracting close attention as one means of resolving these issues [30].

The most noteworthy example of success in recent years is erythropoietin (EPO), a glycoprotein which stimulates erythrocyte formation. This protein has three *N*-linked carbohydrates and one *O*-linked carbohydrate on its surface. These carbohydrates are indispensable for the stable structure of EPO and expression of its activity in vivo [31]. It has been shown that when the number of *N*-linked carbohydrates was increased from 3 to 5 by recombinant gene technology with CHO cells, derivatives with three times prolonged half-life in blood were obtained [30,32]. The five *N*-linked carbohydrates cover the surface of the protein, making the EPO less likely to bind to receptors but to be retained in blood for longer periods of time, resulting in long-lasting efficacy. The sialic acid at the terminal decreases the clearance of this EPO in the liver and kidney, possibly leading to prolonged half-life in blood.

Attempts at changing the fate of compounds in vivo by means of carbohydrate modification are not confined to EPO. For example, it has been reported that new functional characteristics (e. g., prolonged half-life in blood, improved solubility in water, etc.) can be achieved by attaching carbohydrates to particular positions of compounds, e. g., thrombopoietin with only *O*-linked carbohydrate [32] and insulin without any carbohydrates [33,34]. Thus, carbohydrate modification is being acknowledged as a useful technique for modifying the functions of proteins.

3.3 Functionalization of Antibody-Based Biopharmaceuticals by Glycoengineering

Following recent advances in antibody engineering, it is now possible to synthesize humanized antibodies [35]. Because of clinical applicability, antibody-based medicines are one of the biopharmaceutical categories that have attracted close attention in recent years [36]. IgG has been most frequently used for the prevention and treatment of various diseases. The CH2 region of IgG has *N*-linked carbohydrates (usually complex type double stranded chains) in diverse manners (*Scheme 14*). It has recently been revealed that differences in the carbohydrate structure affect the effecter activity of antibodies such as antibody-dependent cellular cytotoxicity (ADCC) and their half-life in blood, inviting close attention from the viewpoint of improving the responses to antibody-based medicines [37,38,39].



Scheme 14

A noteworthy recent finding is that the ADCC activity of antibodies was elevated more than 100-fold by elimination of fucose from the *N*-linked carbohydrates in the CH2 region [38,39]. This is probably because elimination of fucose elevates the binding affinity of FcγRIIIa

expressed on NK cells serving as a trigger for expression of ADCC activity [40]. Another known approach to improve ADCC activity is modification of the amino acid sequence of the Fc region of antibody [41]. However, when applying this technique, the problem with antigenicity needs to be resolved. Fucose-free antibodies are expected to be highly safe since their structure resembles the one seen in the biological body [42].

There are many unresolved questions over the relationship of carbohydrate structure to antibody activity, kinetics and specificity. However, it is evident that the carbohydrate structure greatly affects the antibody functions. Attempts at improving the functions of antibodies by means of carbohydrate modification are now viewed as valuable from the standpoint of their potential for commercialization [43,44].

3.4 Drug Delivery Systems of Enzyme-Based Biopharmaceuticals Utilizing Carbohydrate Functions

The term "lysosomal disease" is used to indicate a group of genetic diseases involving reduced enzyme activity due to mutation of genes encoding lysosomal enzyme and its related factors and accumulation of the undegraded substrate which should normally be degraded by lysosomal enzyme [45,46]. Enzyme replacement therapy has been introduced clinically for the treatment of this disease [47,48,49]. With this therapy, uptake of lysosomal enzyme by cells is promoted by the binding of the specific carbohydrate (attached to the enzyme) to the carbohydrate receptor on the target cell surface and by the mechanism for transport of the enzyme into lysosomes.

For example, alglucerase (extracted from placenta) and imiglucerase (recombinant type) have been used for the treatment of Gaucher's disease [50,51]. These drugs are glucocerebrosidase which has carbohydrates involving mannose at the termini and targets macrophages [52]. In addition, agalsidase β was developed as an agent for the treatment of Fabry's disease, targeting lysosome [53,54]. Agalsidase β is recombinant α-galactosidase A produced in CHO cells. Its highly sialylated carbohydrates suppress uptake of the drug by the liver, and its uptake into lysosome is enabled in a reliable manner by keeping the mannose-6-phosphoric acid level for the efficient transport of the drug into the target lysosome [55].

As illustrated above, utilization of the targeting function of carbohydrates has enabled efficient enzyme replacement therapy and its clinical application. Efforts are currently being made to develop new technology to ensure adequate amounts of enzymes and to enable the application of this kind of drug to the treatment of lysosome disease complicated by disorders of the central nervous system.

3.5 Recent Advances in Glycoprotein Synthesis

Production systems based on animal cells are usually used for the production of glycoprotein-based biopharmaceuticals. Following advances in genetic recombinant technology, it is now possible to control the site and number of carbohydrate insertion in a highly sophisticated manner, and new-generations of glycoprotein-based biopharmaceuticals with novel functions are being developed one after another. However, production systems based on animal cells involve disadvantages such as difficulty in quality control and relatively high production costs.

These systems have the shortcoming of yielding glycoproteins in which carbohydrates not seen in humans are partially present. Furthermore, these carbohydrates are not homogeneous. To resolve these problems, attempts have been made to develop new glycoprotein production techniques making use of some other cell expression systems [56].

Among the research topics in this field, the most advanced pertains to development of a recombinant protein production system with yeast serving as the host [57]. The carbohydrates of the glycoprotein, produced with yeast serving as the host, are usually the high mannose type which cannot be utilized as medicines because of antigenicity. However, by modifying the yeast genetically to yield a recombinant yeast capable of expressing the enzyme needed for synthesis of human type carbohydrates, it is becoming possible to produce human type glycoprotein and also to introduce sialic acid to the *N*-glycan terminal [58].

Regarding research on topics other than those related to yeast expression systems, Betenbaugh and Lee et al., and Jarvis et al. reported synthesis of proteins with modified human type carbohydrates using insect cells [59,60,61,62]. Abei et al. and Shultz et al. reported innovative techniques allowing synthesis of glycoproteins with *Escherichia coli* serving as the host [63,64,65]. Further advances are expected in these fields of research.

Research involving chemical or enzymatic glycosylation has also been carried out actively [66,67]. For example, reports have been published concerning methods of site-specific glycosylation of proteins, using microbial transglutaminase [33,34,68]. Attempts have also been made to expand the line-up of recombinant glycosyltransferases which can be massively produced, including some examples of success in developing practical protein glycosylation techniques making use of GalNAc transferase or sialic acid transferase or selective PEGylation techniques [69,70,71].

Increasing attention has recently been paid to the chemical synthesis and semisynthesis of glycoproteins making use of the native chemical ligation [72,73] and the expressed protein ligation [81]. In this approach, glycopeptide fragments are first synthesized or biosynthesized as building blocks, then fragment peptides are conjugated by chemoselective reactions to make full-size glycoproteins [74,75,80,82]. One of the limitations of this method is that an *N*-terminal cysteine residue is required for the peptide conjugation, which restricts the target protein sequence. To overcome this limitation, some cysteine-free ligation techniques also have been developed [78,79]. These chemical-based approaches which make it possible to generate homogeneous glycoproteins would be a powerful tool in the research of biological functions of glycoproteins and their applications.

4 Drug Discovery Strategies Focusing on Disease-Related Carbohydrates

4.1 Introduction

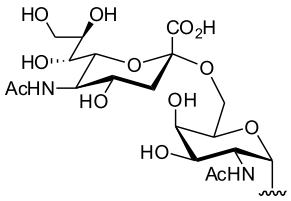
Carbohydrates play important roles in recognition of substances on the cell surface and in transduction of signals. Carbohydrates are involved closely in various diseases such as infection and cancer. For these reasons, increasing efforts have been made to conduct research for drug development, focusing on carbohydrates in the biological body. This section will outline development of new medicines targeting carbohydrates. In addition, recent advances in

carbohydrate synthesis technology, needed for development of new medicines, will also be presented.

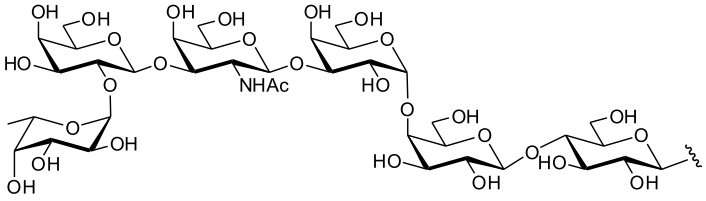
4.2 Drug Discovery Targeting Disease-Related Changes in Carbohydrates

Regarding carbohydrates in the biological body, it is known that the expression level of specific carbohydrates changes in association with the onset of various diseases. It is often unknown whether the change in carbohydrate structure and the onset of disease are both underlying factors or the change in carbohydrate structure induces the symptoms of the disease. However, if the causal relationship between such changes in carbohydrate structure and onset of diseases is clarified and if the condition of individual sick patients is assessed on the basis of changes in carbohydrate for subsequent specific therapy targeting particular carbohydrates, it will be possible to treat these diseases efficiently.

Following carcinogenesis, remarkable changes are noted in the carbohydrates of the superficial layers of cells [83]. Analysis of changes in cell surface carbohydrate during the course of progression and metastasis of tumors will be useful in the diagnosis and treatment of tumors. For example, it has been reported that sialyl Tn antigen (*Scheme 15*), composed of incomplete *O*-linked carbohydrates, showed high expression in various tumors [84]. Following such a finding, efforts have been made to develop a vaccine for cancer targeting the sialyl Tn antigen [85]. Among others, intensive efforts are now made to develop synthetic vaccines composed of sialyl Tn carbohydrates bound to carrier protein KLH for use in the management of metastatic breast cancer and metastatic uterine cancer [86,87]. In addition, various carbohydrate antigens such as glycolipids highly expressed in melanoma cells and globo-H (*Scheme 16*) (a protruding epitope seen in prostate cancer) have been discovered [88,89,90,91], and efforts to develop vaccines based on these antigens have been made.



Scheme 15



Scheme 16

Regarding the metastasizing potentials of tumor cells, close attention is now paid to GlcNAc transferase V (GnT-V) which acts on the branching structure of *N*-linked carbohydrates and form a characteristic carbohydrate structure (β-1,6-GlcNAc chain) [92,93]. GnT-V is highly expressed on cancer with high metastatic potentials. It has been reported that metastasis of cancer was suppressed in GnT-V knockout mice [94]. These findings suggest that GnT-V is associated with the malignancy level of various cancers. If GnT-V can be controlled, suppression of cancer metastasis may be possible. For example, it has been shown at the cellular level that the metastatic potential of cancer was reduced by increasing the expression of GnT-III which competes with GnT-V during carbohydrate synthesis [95,96,97]. Thus, GnT-V seems to serve as a key factor suppressing the metastasis of cancer. It has become a noteworthy target during efforts for drug development.

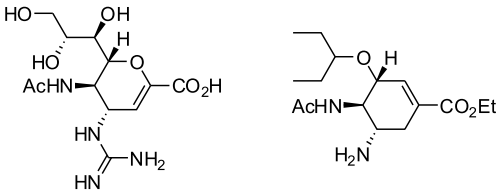
Other than changes in carbohydrates associated with cancers, various carbohydrate changes specific to some particular diseases have been reported, including changes in IgG carbohydrates associated with rheumatoid arthritis [98], abnormal α-dystroglycan carbohydrates associated with muscular dystrophy [99,100,101], and insufficiency of *O*-linked carbohydrates of the hinge region associated with IgG nephropathy [102,103]. Active research has been performed to investigate the causal relationship of these changes to sickness and to develop valid means of treatment. Following recent advances in carbohydrate structure analysis technology, it is now considered important to develop a database of changes in carbohydrates associated with each disease.

4.3 Drug Discovery Targeting Carbohydrate-Lectin Interactions

Numerous reports have been published concerning the attempts of developing drugs targeting interactions between carbohydrates and lectin in vivo. For example, active research has been performed on ligands of selectin (playing an important role in the onset of inflammation) to carbohydrates [104,105,106,107]. It has also been reported that interactions between carbohydrates and lectin are important when viruses and bacteria invade host cells [108]. This section will outline efforts of drug development targeting carbohydrate-lectin interactions, citing research on anti-influenza drugs as an example.

The influenza virus is one of the most extensively studied viruses at present. The influenza virus recognizes sialic acid on the host cell surface, followed by infection and transmission [109]. Two spike-formed proteins (hemagglutinin and neuraminidase) are found on the surface of the influenza virus [110,111]. Hemagglutinin binds to the host cells through recognizing the ligand containing sialic acid on the surface of infected cells, while neuraminidase eliminates sialic acid from the cell surface to avoid inhibition of the transmission of newly formed virus particles.

Because hemagglutinin has a high affinity for the infected cells by the binding of its multivalent binding site to multiple sialic acid molecules, it is not easy to develop low-molecular-weight hemagglutinin inhibitors which can inhibit interactions between hemagglutinin and receptors. To develop inhibitors with high inhibitory activity, research has been carried out to develop clusters of sialic acid derivatives that serve as ligand molecules. For example, reports have been published concerning research on efficiency inhibition of hemagglutinin functions using macromolecules with scaffolds made of synthetic polymers, cyclodextrin, dendrimers,



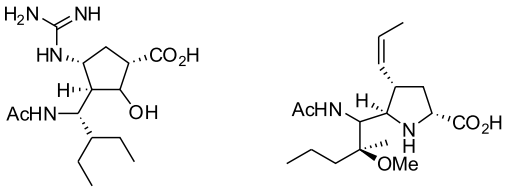
Zanamivir Oseltamivir

Scheme 17

peptides, etc. [112,113,114,115,116]. However, there are many open issues which need to be resolved before developing oral-dose preparations of large ligand molecules.

Neuraminidase inhibitors are a more attractive target of drug development. This is because it is possible to design an inhibitor with a high affinity for neuraminidase by mimicking the transition state seen during fragmentation of sialic acid with neuraminidase [117]. As examples of such neuraminidase inhibitors, zanamivir (a lung administration type drug derived from sialic acid) and oseltamivir (an oral-dose prodrug derived from shikimic acid) have been developed (*Scheme 17*) [118,119], and they have been introduced clinically as remarkable new medicines effective against both type A and B influenza virus. However, species of influenza virus resistant to each of these drugs have appeared. It is still an important issue to develop new anti-influenza drugs [120].

Compounds derived from a cyclopentane or pyrrolidine skeleton have recently been attracting close attention as next-generation neuraminidase inhibitors [121,122,123,124,125], and active efforts have been made for screening of such compounds (*Scheme 18*). Attempts have also been made to prevent virus infection by removing the tracheal sialic acid receptors (the starting point of viral infection) through direct inhalation of sialidase [126].

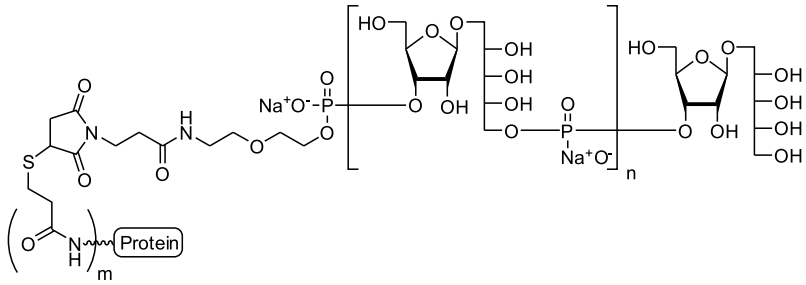


Scheme 18

4.4 Drug Discovery Targeting Carbohydrates of Pathogens

Microorganisms responsible for infection (various bacteria, viruses, protozoan, etc.) often have carbohydrates different from those seen in humans. For this reason, it will be possible to prevent the onset of associated diseases by using drugs targeted at these particular carbohydrates [127]. This section will outline these drugs.

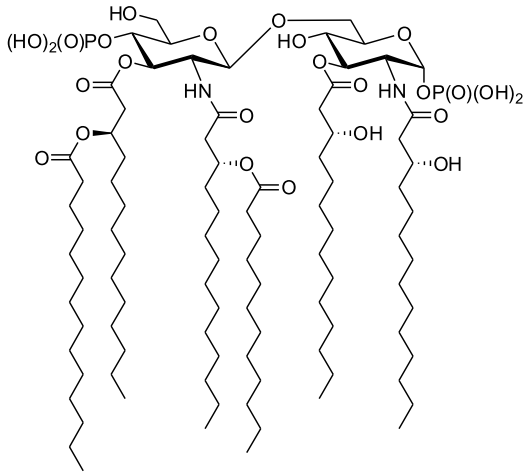
Capsular polysaccharide of bacteria is one of the significant targets for development of vaccines [128]. For example, *Haemophilus influenza* type B (Hib) vaccine, made of a polysaccharide-protein conjugate, is now used worldwide as a highly effective therapy [129]. Verez-



Scheme 19

Bencomo et al. demonstrated that a chemically synthesized Hib vaccine had an activity comparable to that of existing vaccines prepared from Hib polysaccharides and that vaccines based on synthetic complex carbohydrates would be useful (*Scheme 19*) [130]. Their report now serves as a base for establishing similar approaches to the control of other human pathogens.

Pathogenic Gram-negative bacteria have complex glycolipids called lipo-oligosaccharide and lipo-polysaccharide on the tunica external of the cells. These glycolipids also serve as the target for developing vaccines for use in the prevention of infection, similar to those stated above [131]. Meanwhile, lipo-polysaccharides, released as a result of collapse of bacteria can trigger serious inflammation. The skeleton needed for induction of such inflammation is called "Lipid A" (*Scheme 20*), which can induce fever and acute inflammatory reactions (so-called "septic shock") even at a very small amount [132]. In recent years, Lipid A analogues which may be used as antagonists have been explored, toward the goal of developing new drugs for the treatment of sepsis [133,134,135].



E. Coli Lipid A

Scheme 20

Other than the above-mentioned drugs, vaccines targeted at the glycoprotein on the surface of *Staphylococcus aureus* (a Gram-positive bacterium) are now under development [136]. Research on development of drugs targeted at the specific carbohydrates of protozoan parasites [137,138], and viruses such as HIV is also under way [139,140].

4.5 Recent Advances in Carbohydrate Synthesis for Glycodrug Discovery

Carbohydrates assume very complex structures in humans. To understand their diverse functions and apply them as medicines, it is essential to synthesize and evaluate various carbohydrates and complex carbohydrates in a compressive and high throughput manner. In practice, improvement of the synthetic throughput of diverse samples is contributing greatly to advancing genomics and proteomics. Methods of synthesis for establishment of the library of diverse carbohydrate compounds are also being developed.

For example, one-pot synthesis of carbohydrates is a valid means of immediately obtaining various carbohydrates [141]. Among others, the technique of programmable one-pot synthesis developed by Wong et al. allows efficient synthesis by means of quantitative evaluation and programmed control of the reactivity of each building block utilized for glycosylation [142]. The technique of solid-phase synthesis is a method of chemical synthesis with excellent throughputs [143]. Seeberger et al. succeeded in developing an automated synthesis of carbohydrates by utilizing solid-phase synthetic technique [144,145].

Remarkable advances have also been made in the techniques for synthesis making use of enzymes. Enzymatic methods have exerted their high efficacy when various glycoconjugates are synthesized [141]. For example, Thorson et al. developed glycorandomization, a technique by which a library of glycosylated compounds is established through enzymatic reactions from free monosaccharide libraries, using engineered enzymes [146,147]. They have carried out studies, applying this technique to various natural compounds. Nishimura et al. are developing carbohydrate synthesis engineering with glycosyltransferase, using unique substrates bound to water-soluble polymers. Their method is applicable to high throughput synthesis of not only carbohydrates but also glycoconjugates, such as glycolipids and glycopeptides [148,149,150]. The libraries associated with the carbohydrates, prepared in these ways, have been integrated into carbohydrate micro-arrays [151,152], which are now being utilized for screening of new specific carbohydrate antigens and for drug development research towards the goal of vaccine development.

References

1. Maeda YY, Hamuro J, Yamada YO, Ishimura K, Chihara G (1973) In: Immunopotentiation. Wolstenholme GEW, Knight J (eds) Elsevier, Excerpta Medica, North-Holland, p 2592
2. Maeda YY, Chihara G (1971) Nature 229:634
3. Hamuro J, Röllinghoff M, Wagner H (1978) Cancer Res 38:3080
4. Hamuro J, Wagner H, Röllinghoff M (1978) Cell Immunol 38:328
5. Hamuro J, Röllinghoff M, Wagner H (1980) Immunology 39:551
6. Yamashita T, Yamamoto T, Kagawa K, Higuchi Y, Tsubura E (1978) Gan to Kagakuryoho 5:103

7. Sasaki T, Abiko N, Sugino Y, Nitta K (1978) Cancer Res 38:379
8. Yoshida T, Hatanaka K, Uryu T, Kaneko Y, Suzuki E, Miyano H, Yoshida O, Yamamoto N (1990) Macromolecules 23:3717
9. Yamamoto I, Takayama K, Konma K, Gonda T, Yamazaki K, Matsuzaki K, Hatanaka K, Uryu T, Yoshida O, Nakashima H, Yamamoto N, KanekoY, Mimura T (1991) Carbohydr Polym 14:53
10. Grönwall A, Ingelman B (1984) Vox Sang 47:96
11. Richter W (1966) Acta Chir Scand 131:1
12. Gregersen M I, Peric B, Usami S, Chien S (1964) Bibliotheca Anat 4:58
13. Ito M, Baba M, Sato A, Pauwels R, De Clerq E (1988) Antiviral Res 7:361
14. Ueno R, Kuno S (1987) Lancet 1:1379
15. Muzzarelli RAA (1977) Chitin. Pergamon Press, Oxford
16. Muzzarelli RAA, Muzzarelli B (1998) In: Polysaccharides, Structural Diversity and Functional Versatility. Dumitriu S (ed) Marcel Dekker, New York, p 569
17. Nishimura K, Nishimura S-I, Nishi N, Saiki I, Tokura S, Azuma I (1984) Vaccine 2:93
18. Nishimura K, Nishimura S-I, Nishi N, Saiki I, Tokura S, Azuma I (1984) Vaccine 2:129
19. Suzuki S, Suzuki K, Tokoro A, Okawa Y, Suzuki M (1986) In: Chitin in Nature and Technology. Muzzarelli RAA, Jeuniaux C, Gooday GW (eds) Plenum, New York, p 485
20. Tokura S, Tamura H, Azuma I (1999) In: Chitin and Chitinases. Jollés P, Muzzarelli RAA (eds) Birkhäuser Verlag, Basel, p 279
21. Minami S, Okamoto Y, Matsuhashi A, Sashiwa H, Saimoto H, Shigemasa Y, Tangigawa T, Tanaka Y, Tokura S (1992) In: Advances in Chitin and Chitosan. Brine CJ, Sandford PA, Zikakis JP (eds) Elsevier Applied Science, London, p 61
22. Minami S, Okamoto Y, Matsuhashi A, Sashiwa H, Saimoto H, Shigemasa Y, Tangigawa T, Tanaka Y, Tokura S (1992) In: Advances in Chitin and Chitosan. Brine CJ, Sandford PA, Zikakis JP (eds) Elsevier Applied Science, London, p 70
23. Okamoto Y, Minami S, Matsuhashi A, Sashiwa H, Saimoto H, Shigemasa Y, Tanigawa T, Tanaka Y, Tokura S (1993) J Vet Med Sci 55:743
24. Lindahl U (1989) In: Heparin, Chemical and Biological Properties, Clinical Applications. Lane DA, Lindahl U (eds) Edward Arnold, London, p 159
25. Gallagher JT (1998) Trends Glycosci Glycotech 10:137
26. Hatanaka K, Song S-C, Maruyama A, Kobayashi A, Kuzuhara H, Akaike T (1992) Biochem Biophys Res Commun 188:16
27. Hatanaka K, Song S-C, Maruyama A, Kobayashi A, Kuzuhara H, Akaike T (1993) Polymer J 25:373
28. Nishimura S-I, Lee YC (1998) In: Polysaccharides, Structural Diversity and Functional Versatility. In: Dumitoriu S (ed) Marcel Dekker, New York, p 523
29. Roy R (1996) Polymer News 21:226
30. Sinclair AM, Elliott S (2005) J Pharm Sci 94:1626
31. Tsuda E, Kawanishi G, Ueda M, Masuda S, Sasaki R (1990) Eur J Biochem 188:405
32. Elliott S, Lorenzini T, Asher S, Aoki K, Brankow D, Buck L, Busse L, Chang D, Fuller J, Grant J, Hernday N, Hokum M, Hu S, Knudten A, Levin N, Komorowski R, Martin F, Navarro R, Osslund T, Rogers G, Rogers N, Trail G, Egrie J (2003) Nat Biotechnol 21:414
33. Sato M, Sadamoto R, Niikura K, Monde K, Kondo H, Nishimura S (2004) Angew Chem Int Ed 43:1516
34. Sato M, Furuike T, Sadamoto R, Fujitani N, Nakahara T, Niikura K, Monde K, Kondo H, Nishimura S (2004) J Am Chem Soc 126:14013
35. Chadd HE, Chamow SM (2001) Curr Opin Biotechnol 12:188
36. Carter PJ (2006) Nat Rev Immunol 6:343
37. Wright A, Morrison SL (1998) J Immunol 160:3393
38. Shields RL, Lai J, Keck R, O'Connell LY, Hong K, Meng YG, Weikert SH, Presta LG (2002) J Biol Chem 277:26733
39. Shinkawa T, Nakamura K, Yamane N, Shoji-Hosaka E, Kanda Y, Sakurada M, Uchida K, Anazawa H, Satoh M, Yamasaki M, Hanai N, Shitara K (2003) J Biol Chem 278:3466
40. Okazaki A, Shoji-Hosaka E, Nakamura K, Wakitani M, Uchida K, Kakita S, Tsumoto K, Kumagai I, Shitara K (2004) J Mol Biol 336:1239
41. Shields RL, Namenuk AK, Hong K, Meng YG, Rae J, Briggs J, Xie D, Lai J, Stadlen A, Li B, Fox JA, Presta LG (2001) J Biol Chem 276:6591
42. Harada H, Kamei M, Tokumoto Y (1987) Anal Biochem 164:374
43. Li H, Sethuraman N, Stadheim TA, Zha D, Prinz B, Ballew N, Bobrowicz P, Choi BK, Cook WJ, Cukan M, Houston-Cummings NR, Davidson R, Gong B, Hamilton SR, Hoopes JP, Jiang Y, Kim

N, Mansfield R, Nett JH, Rios S, Strawbridge R, Wildt S, Gerngross TU (2006) Nature Biotechnol 24:210
44. Hodoniczky J, Yuan ZZ, James DC (2005) Biotechnol Prog 21:1644
45. Vellodi A (2005) Brit J Haematol 128:413
46. Winchester B (2005) Glycobiology 15:1R
47. Neufeld EF (1991) Ann Rev Biochem 60:257
48. Vellard M (2003) Curr Opin Biotechnol 14: 444
49. Pastores GM, Barnett NL (2005) Expert Opin Emerg Drugs 10:891
50. Barton NW, Brady RO, Dambrosia JM, Di Bisceglie AM, Doppelt SH, Hill SC, Mankin HJ, Murray GJ, Parker RI, Argoff CE, Grewal RP, Yu KT, Graham OC, Holder CA, Howard KD, Kaneski CR, Oliver KL, Riesz S, Verderse CL, Zirzow GC (1991) New England J Med 324:1464
51. Grabowski GA, Barton NW, Pastores G, Dambrosia JM, Banerjee TK, McKee MA, Parker C, Schiffmann R, Hill SC, Brady RO (1995) Annals Intern Med 122:33
52. East L, Isacke CM (2002) Biochimica et Biophysica Acta – General Subjects 1572:364
53. Germain DP (2002) Exp Opin Invest Drugs 11:1467
54. Lee K, Jin X, Zhang K, Copertino L, Andrews L, Baker-Malcolm J, Geagan L, Qiu H, Seiger K, Barngrover D, McPherson JM, Edmunds T (2003) Glycobiology 13:305
55. Sly WS (2004) Missouri Med 101:100
56. Borman STU (2006) Chem Eng News 84:13
57. Wildt S, Gerngross TU (2005) Nat Rev Microbiol 3:119
58. Hamilton SR, Davidson RC, Sethuraman N, Nett JH, Jiang Y, Rios S, Bobrowicz P, Stadheim TA, Li H, Choi BK, Hopkins D, Wischnewski H, Roser J, Mitchell T, Strawbridge RR, Hoopes J, Wildt S, Gerngross TU (2006) Science 313:1441
59. Tomiya N, Betenbaugh MJ, Lee YC (2003) Accounts Chem Res 36:613
60. Tomiya N, Narang S, Lee YC, Betenbaugh MJ (2004) Glycoconj J 21:343
61. Jarvis DL, Finn EE (1996) Nat Biotechnol 14:1288
62. Aumiller JJ, Hollister JR, Jarvis DL (2003) Glycobiology 13:497
63. Wacker M, Linton D, Hitchen PG, Nita-Lazar M, Haslam SM, North SJ, Panico M, Morris HR, Dell A, Wren BW, Aebi M (2002) Science 298:1790
64. Feldman MF, Wacker M, Hernandez M, Hitchen PG, Marolda CL, Kowarik M, Morris HR, Dell A, Valvano MA, Aebi M (2005) Proc Natl Acad Sci USA 102:3016
65. Zhang Z, Gildersleeve J, Yang YY, Xu R, Loo JA, Uryu S, Wong CH, Schultz PG (2004) Science 303:371
66. Davis BG (2002) Chem Rev 102:579
67. Liu L, Bennett CS, Wong CH (2006) Chem Commun (Camb):21
68. Sato H, Hayashi E, Yamada N, Yatagai M, Takahara Y (2001) Bioconjug Chem 12:701
69. Johnson KF (1999) Glycoconjugate J 16:141
70. Warnock D, Bai X, Autote K, Gonzales J, Kinealy K, Yan B, Qian J, Stevenson T, Zopf D, Bayer RJ (2005) Biotechnol Bioeng 92:831
71. DeFrees S, Wang ZG, Xing R, Scott AE, Wang J, Zopf D, Gouty DL, Sjoberg ER, Panneerselvam K, Brinkman-Van der Linden EC, Bayer RJ, Tarp MA, Clausen H (2006) Glycobiology 16:833
72. Dawson PE, Muir TW, Clark-Lewis I, Kent SB (1994) Science 266:776
73. Dawson PE, Kent SBH (2000) Ann Rev Biochem 69:923
74. Marcaurelle LA, Mizoue LS, Wilken J, Oldham L, Kent SBH, Handel TM, Bertozzi CR (2001) Chem A Eur J 7:1129
75. Mezzato S, Schaffrath M, Unverzagt C (2005) Angew Chem Int Ed 44:1650
76. Kochendoerfer GG, Chen SY, Mao F, Cressman S, Traviglia S, Shao H, Hunter CL, Low DW, Cagle EN, Carnevali M, Gueriguian V, Keogh PJ, Porter H, Stratton SM, Wiedeke MC, Wilken J, Tang J, Levy JJ, Miranda LP, Crnogorac MM, Kalbag S, Botti P, Schindler-Horvat J, Savatski L, Adamson JW, Kung A, Kent SB, Bradburne JA (2003) Science 299:884
77. Chen SY, Cressman S, Mao F, Shao H, Low DW, Beilan HS, Cagle EN, Carnevali M, Gueriguian V, Keogh PJ, Porter H, Stratton SM, Wiedeke MC, Savatski L, Adamson JW, Bozzini CE, Kung A, Kent SB, Bradburne JA, Kochendoerfer GG (2005) Chem Biol 12:371
78. Brik A, Yang YY, Ficht S, Wong CH (2006) J Am Chem Soc 128:5626
79. Wu B, Chen J, Warren JD, Chen G, Hua Z, Danishefsky SJ (2006) Angew Chem Int Ed 45:4116
80. Tolbert TJ, Wong CH (2000) J Am Chem Soc 122:5421
81. Muir TW (2003) Ann Rev Biochem 72:249
82. Macmillan D, Bertozzi CR (2004) Angew Chem Int Ed 43:1355

83. Fuster MM, Esko JD (2005) Nat Rev Cancer 5:526
84. Brockhausen I (1999) Biochim Biophys Acta 1473:67
85. Koganty RR, Reddish MA, Longenecker BM (1996) Drug Discov Today 1:190
86. Sandmaier BM, Oparin DV, Holmberg LA, Reddish MA, MacLean GD, Longenecker BM (1999) J Immunother 22:54
87. Ibrahim NK, Murray JL (2003) Clin Breast Cancer 3(4):S139
88. Bitton RJ, Guthmann MD, Gabri MR, Carnero AJ, Alonso DF, Fainboim L, Gomez DE (2002) Oncol Rep 9:267
89. Carr A, Rodríguez E, Arango MC, Camacho R, Osorio M, Gabri M, Carrillo G, Valdés Z, Bebelagua Y, Pérez R, Fernández LE (2003) J Clin Oncol 21:1015
90. Ragupathi G, Livingston PO, Hood C, Gathuru J, Krown SE, Chapman PB, Wolchok JD, Williams LJ, Oldfield RC, Hwu WJ (2003) Clin Cancer Res 9:5214
91. Slovin SF, Ragupathi G, Adluri S, Ungers G, Terry K, Kim S, Spassova M, Bornmann WG, Fazzari M, Dantis L, Olkiewicz K, Lloyd KO, Livingston PO, Danishefsky SJ, Scher HI (1999) Proc National Academy of Sciences of the United States of America 96:5710
92. Dennis JW, Laferte S, Waghorne C (1987) Science 236:582
93. Dennis JW, Granovsky M, Warren CE (1999) Biochim Biophys Acta 1473:21
94. Granovsky M, Fata J, Pawling J, Muller WJ, Khokha R, Dennis JW (2000) Nat Med 6:306
95. Yoshimura M, Nishikawa A, Ihara Y, Taniguchi S, Taniguchi N (1995) Proc Natl Acad Sci USA 92:8754
96. Yoshimura M, Ihara Y, Matsuzawa Y, Taniguchi N (1996) J Biol Chem 271:13811
97. Zhao Y, Nakagawa T, Itoh S, Inamori K, Isaji T, Kariya Y, Kondo A, Miyoshi E, Miyazaki K, Kawasaki N, Taniguchi N, Gu J (2006) J Biol Chem 281:32122
98. Malhotra R, Wormald MR, Rudd PM, Fischer PB, Dwek RA, Sim RB (1995) Nature Medicine 1:237
99. Endo T (1999) Biochim Biophys Acta 1473: 237
100. Yoshida A, Kobayashi K, Manya H, Taniguchi K, Kano H, Mizuno M, Inazu T, Mitsuhashi H, Takahashi S, Takeuchi M, Herrmann R, Straub V, Talim B, Voit T, Topaloglu H, Toda T, Endo T (2001) Dev Cell 1:717
101. Grewal PK, Holzfeind PJ, Bittner RE, Hewitt JE (2001) Nat Genet 28:151
102. Odani H, Hiki Y, Takahashi M, Nishimoto A, Yasuda Y, Iwase H, Shinzato T, Maeda K (2000) Biochem Biophys Res Comm 271:268
103. Hiki Y, Odani H, Takahashi M, Yasuda Y, Nishimoto A, Iwase H, Shinzato T, Kobayashi Y, Maeda K (2001) Kidney Int 59:1077
104. Hiramatsu Y, Moriyama H, Kiyoi T, Tsukida T, Inoue Y, Kondo H (1998) J Med Chem 41:2302
105. Tsukida T, Moriyama H, Kurokawa K, Achiha T, Inoue Y, Kondo H (1998) J Med Chem 41:4279
106. Simanek EE, McGarvey GJ, Jablonowski JA, Wong CH (1998) Chem Rev 98:833
107. Magnani JL (2004) Arch Biochem Biophys 426:122
108. Karlsson KA (1995) Curr Opin Struct Biol 5:622
109. Suzuki Y (2005) Biol Pharm Bull 28:399
110. Meanwell NA, Krystal M (1996) Drug Discov Today 1:316
111. Meanwell NA, Krystal M (1996) Drug Discov Today 1:388
112. Kamitakahara H, Suzuki T, Nishigori N, Suzuki Y, Kanie O, Wong CH (1998) Angew Chem Int Ed 37:1524
113. Furuike T, Aiba S, Suzuki T, Takahashi T, Suzuki Y, Yamada K, Nishimura SI (2000) J Chem Soc, Perkin Transactions 1:3000
114. Fulton DA, Stoddart JF (2001) Bioconjug Chem 12:655
115. Reuter JD, Myc A, Hayes MM, Gan Z, Roy R, Qin D, Yin R, Piehler LT, Esfand R, Tomalia DA, Baker JR, Jr. (1999) Bioconjug Chem 10:271
116. Ohta T, Miura N, Fujitani N, Nakajima F, Niikura K, Sadamoto R, Guo CT, Suzuki T, Suzuki Y, Monde K, Nishimura SI (2003) Angew Chem Int Ed 42:5186
117. Moscona A (2005) New England J Med 353:1363
118. von Itzstein M, Wu WY, Kok GB, Pegg MS, Dyason JC, Jin B, Van Phan T, Smythe ML, White HF, Oliver SW, et al. (1993) Nature 363:418
119. Kim CU, Lew W, Williams MA, Liu H, Zhang L, Swaminathan S, Bischofberger N, Chen MS, Mendel DB, Tai CY, Laver WG, Stevens RC (1997) J Am Chem Soc 119:681
120. De Clercq E (2006) Nat Rev Drug Discov 5:1015
121. Smee DF, Huffman JH, Morrison AC, Barnard DL, Sidwell RW (2001) Antimicrob Agents Chemother 45:743

122. Sidwell RW, Smee DF, Huffman JH, Barnard DL, Bailey KW, Morrey JD, Babu YS (2001) Antimicrob Agents Chemother 45:749
123. Chand P, Bantia S, Kotian PL, El-Kattan Y, Lin TH, Babu YS (2005) Bioorg Med Chem 13:4071
124. Hanessian S, Bayrakdarian M, Luo X (2002) J Am Chem Soc 124:4716
125. Kati WM, Montgomery D, Carrick R, Gubareva L, Maring C, McDaniel K, Steffy K, Molla A, Hayden F, Kempf D, Kohlbrenner W (2002) Antimicrob Agents Chemother 46:1014
126. Malakhov MP, Aschenbrenner LM, Smee DF, Wandersee MK, Sidwell RW, Gubareva LV, Mishin VP, Hayden FG, Kim DH, Ing A, Campbell ER, Yu M, Fang F (2006) Antimicrob Agents Chemother 50:1470
127. Vliegenthart JF (2006) FEBS Lett 580:2945
128. Sood RK, Fattom A, Pavliak V, Naso RB (1996) Drug Discov Today 1:381
129. Schneerson R, Barrera O, Sutton A, Robbins JB (1980) J Exp Med 152:361
130. Verez-Bencomo V, Fernández-Santana V, Hardy E, Toledo ME, Rodríguez MC, Heynngnezz L, Rodriguez A, Baly A, Herrera L, Izquierdo M, Villar A, Valdés Y, Cosme K, Deler ML, Montane M, Garcia E, Ramos A, Aguilar A, Medina E, Toraño G, Sosa I, Hernandez I, Martínez R, Muzachio A, Carmenates A, Costa L, Cardoso F, Campa C, Diaz M, Roy R (2004) Science 305:522
131. Robbins JB, Schneerson R, Szu SC, Pozsgay V (1999) Pure Appl Chem 71:745
132. Raetz CRH (1990) Ann Rev Biochem 59:129
133. Holst O (1995) Angew Chem Int Ed Engl 34:2000
134. Christ WJ, Asano O, Robidoux AL, Perez M, Wang Y, Dubuc GR, Gavin WE, Hawkins LD, McGuinness PD, Mullarkey MA, et al. (1995) Science 268:80
135. Ingalls RR, Monks BG, Savedra Jr R, Christ WJ, Delude RL, Medvedev AE, Espevik T, Golenbock DT (1998) J Immunol 161:5413
136. García-Lara J, Masalha M, Foster SJ (2005) Drug Discov Today 10:643
137. Hewitt MC, Seeberger PH (2001) J Org Chem 66:4233
138. Schofield L, Hewittt MC, Evans K, Slomos MA, Seeberger PH (2002) Nature 418:785
139. Pantophlet R, Wilson IA, Burton DR (2003) J Virology 77:5889
140. Wang LX (2006) Curr Opin Drug Discov Develop 9:194
141. Koeller KM, Wong CH (2000) Chem Rev 100:4465
142. Zhang Z, Ollmann IR, Ye XS, Wischnat R, Baasov T, Wong CH (1999) J Am Chem Soc 121:734
143. Seeberger PH, Haase WC (2000) Chem Rev 100:4349
144. Plante OJ, Palmacci ER, Seeberger PH (2001) Science 291:1523
145. Seeberger PH, Werz DB (2005) Nat Rev Drug Discov 4:751
146. Fu X, Albermann C, Jiang J, Liao J, Zhang C, Thorson JS (2003) Nat Biotechnol 21:1467
147. Langenhan JM, Griffith BR, Thorson JS (2005) J Natural Prod 68:1696
148. Yamada K, Fujita E, Nishimura S (1997) Carbohydr Res 305:443
149. Fumoto M, Hinou H, Matsushita T, Kurogochi M, Ohta T, Ito T, Yamada K, Takimoto A, Kondo H, Inazu T, Nishimura SI (2005) Angew Chem Int Ed 44:2534
150. Fumoto M, Hinou H, Ohta T, Ito T, Yamada K, Takimoto A, Kondo H, Shimizu H, Inazu T, Nakahara Y, Nishimura SI (2005) J Am Chem Soc 127:11804
151. Feizi T, Chai W (2004) Nat Rev Mol Cell Biol 5:582
152. Borman S (2005) Chem Eng News 83:41

12.2 Biomedicine of Monosaccharides

Helen M. I. Osborn[1], *Philip G. Evans*[1], *Karel Bezouska*[2]
[1] School of Pharmacy, University of Reading, Whiteknights, Reading, Berkshire RG6 6AD, UK
[2] Department of Biochemistry, Faculty of Science, Charles University Prague, Hlavova 8, 12840 Praha 2, Czech Republic
h.m.i.osborn@reading.ac.uk, bezouska@biomed.cas.cz

In: *Glycoscience*. Fraser-Reid B, Tatsuta K, Thiem J (eds)
Chapter-DOI 10-1007/978-3-540-30429-6_62:

Abstract

The importance of monosaccharides in chemical biology is now well appreciated and as a result much attention has been placed on the development of new biomedical applications for both monosaccharides and their derivatives. Within this chapter an overview of the biochemistry and medical applications of monosaccharides is presented to introduce the reader to their key applications and highlight opportunities for future developments. Thus, a systematic review of the key monosaccharides that feature within chemical biology and biomedical applications is presented (e. g. pentoses, hexoses, heptoses, aminosaccharides, deoxysaccharides, alditols, uronic and aldulosonic acids, phosphorylated and sulfated monosaccharides) with an emphasis being placed on understanding, predicting, and influencing the interactions of the monosaccharides with biological systems.

Keywords

Monosaccharide complexes; Enzymes; Transporters; Receptors; Metabolic engineering; Monosaccharide mimetics; Diagnosis of malignant diseases; Cancer therapy

Abbreviations

ADP adenosine diphosphate
ATP adenosine triphosphate
cAMP cyclic adenosine monophosphate
cDNA complementary DNA
COG cluster of orthologous groups
Da Dalton
GLUT glucose transporter
K_d apparent dissociation constant
K_m Michaelis constant
NAD nicotinamide adenosine dinucleotide
NADP nicotinamide adenosine dinucleotide phosphate
UCP uncoupling protein
V_{max} maximal velocity of the enzymic reaction

1 Introduction

The aim of this chapter is to discuss the chemical biology and biomedical applications of monosaccharides that occur as important metabolic intermediates, as essential components of various body fluids, and as important biological recognition markers. In these capacities monosaccharides enter into numerous interactions with specific membrane transporters, intracellular enzymes, and carbohydrate receptors, which in turn defines their chemical biology and suggests avenues for important biomedical developments. In particular, the use of monosaccharides in diagnostics and therapies will be discussed herein with emphasis being clearly placed on newly emerging and recently established applications. In general the monosaccha-

rides discussed contain at least five carbon atoms. Although a detailed discussion of monosaccharide chemistry will not be provided herein it should be appreciated that carbohydrates are rich in functional groups and in particular that ring-closed monosaccharides can be considered to contain three principal chemical facets: the reactive hemiacetal hydroxy group, the outer chain of hydrophilic hydroxy side groups, and the inner hydrophobic hydrocarbon core. Additional functionality can be incorporated by introduction of other substituents onto the carbohydrate backbone, for example polar, charged groups such as carboxylic acids, phosphates, or sulfates are often incorporated within biologically significant monosaccharides. A recent review by Kremer and Gallo-Rodriguez that has highlighted the properties and synthesis of naturally occurring monosaccharides is recommended for additional background reading [1]. The roles of carbohydrates as therapeutics and their use in the design of drugs have also been reviewed [2].

2 Ketoses and Aldoses

2.1 Pentoses

The structures of pentoses which occur in nature and enter into interactions 2 with the sensor, transport, enzymatic, and receptor systems are presented in *Fig. 1*.

While both D- and L-arabinose as well as D-xylose occur in Nature mostly as pyranoses in a C1 chair configuration, D-ribose and D-lyxose are found in the furanose forms.

Pentoses are of prime importance for contemporary organisms as structural components of nucleic acids. The existence of the primordial RNA world, in which RNA is suggested to act as a catalyst as well as an informational macromolecule, assumes a large prebiotic source of ribose. Alternatively, the possible existence of pre-RNA molecules with backbones different from ribose phosphate has been considered [3].

Pentoses are important metabolites for microorganisms, many of which can use them as a sole source of carbon and energy. In order to metabolize pentoses efficiently, microorganisms possess specific periplasmic sensory/binding systems for these sugars, which cooperate

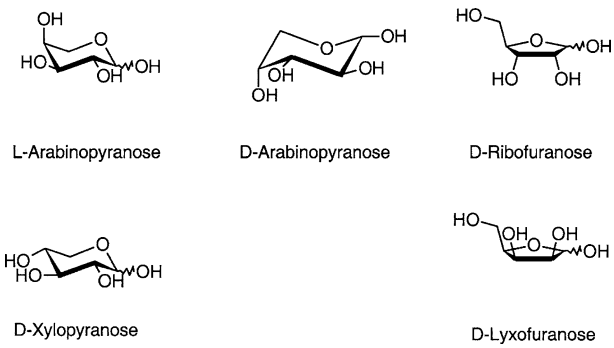


Figure 1
Structures of pentoses in which these monosaccharides interact with the biomolecules of living matter

with pentose/H^+ membrane symporters. Intracellularly, pentoses are converted enzymatically into their phosphate derivatives, and then introduced into the common pool of intracellular metabolites through a series of enzymatic conversions leading to sugars that may eventually enter the pentose phosphate and Entner–Doudoroff pathways. Thus, L-arabinose is converted into L-ribulose by L-arabinose isomerase, and then to L-ribulose 5-phosphate and D-xylulose 5-phosphate by the concerted action of ribulokinase and ribulose 5-phosphate-4-epimerase in order to enter into the intermediary metabolism through reactions of the pentose phosphate cycle. D-Ribose is similarly converted to D-ribose 5-phosphate by ribokinase, which is then isomerized into D-ribulose 5-phosphate by ribose phosphate isomerase, and finally to D-xylulose 5-phosphate by ribulose 5-phosphate-3-epimerase in order to link to the above pathways. The reaction scheme for D-xylose is simple: inside the cell it is converted to D-xylulose by xylose isomerase, and phosphorylated to D-xylulose 5-phosphate by xylulokinase to join once again the above pathway. Microbial metabolism of certain pentoses is more complex; for instance, L-xylulose is intracellularly reduced to xylitol by L-xylulose reductase and then reoxidized to D-xylulose by D-xylulose oxidase. Alternatively, this sugar may also be first phosphorylated by L-xylulokinase to L-xylulose 5-phosphate, which is subsequently epimerized into L-ribulose 5-phosphate by ribulose 5-phosphate-3-epimerase.

In plants and some other eukaryotes, pentoses are components of cell wall polysaccharides such as xylans and arabinogalactans. More interestingly, all "green" forms of life, i. e., those that perform photosynthesis, contain ribulose 1,5-bisphosphate carboxylase/oxygenase (often abbreviated as Rubisco) as the central enzyme involved in carbon dioxide fixation. Consequently, this enzyme has become of interest in numerous genetic engineering projects aimed at the improvement of photosynthesis in agriculturally important plants [4].

In vertebrates, pentoses are poorly metabolized in vivo, and often even decrease the efficiency of utilization when used in combination with other monosaccharides, such as glucose [5,6]. Small amounts of pentoses can apparently be introduced into the metabolism through the activity of intestinal microflora. Thus, interesting isolates of *Clostridium methylpentosum* that ferment only methylpentoses (L-rhamnose and L-fucose) and pentoses (L-lyxose, D-arabinose) have been obtained from the human intestine [7]. However, these reactivities represent quantitatively a negligible capacity when compared to the turnover of other sugars in the intestine. An explanation for this small ability of the vertebrates to utilize pentoses may be found not only in the lack of the specific enzymes of pentose metabolism, but also in the inhibitory effects that pentoses exert upon intestinal sucrase and other hexoses-metabolizing enzymes [8]. The same findings also seem to apply under in vitro conditions, where it has been shown that isolated organs such as the working heart metabolize certain pentoses extremely slowly, and this leads to rapid glycogen depletion and decreases in oxygen consumption [9]. In transformed cells, pentose phosphate pathways are considered to be important in the biochemistry of the tumor proliferation processes because of their role in supplying tumor cells with NADP for intracellular anabolic reactions [10]. Interestingly, the metabolism of pentoses is interlinked with that of glycuronic acids, and ascorbic acid (vitamin C) biosynthesis. Ascorbic acid is readily oxidized into dehydroascorbic acid, and this is dehydrated and decarboxylated into L-xylose. This carbohydrate is subsequently turned into L-xylulose, xylitol, and D-xylulose, which can enter the pentose phosphate cycle after phosphorylation.

2.1.1 Genetics of Pentose Enzymes and Transporters

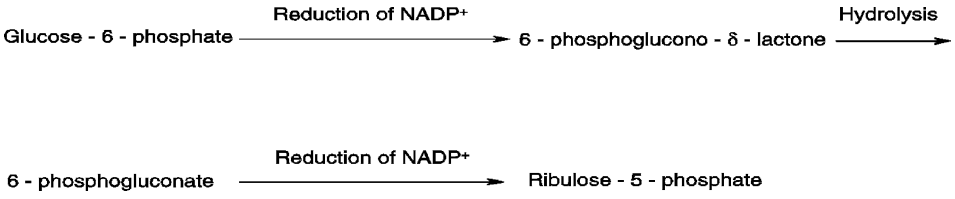
The genetics of certain enzymatic systems responsible for the metabolism of pentoses have been well studied, and the corresponding genes cloned, sequenced, and their position within the genome confirmed during the recent whole genome sequencing initiatives in several microorganisms. Moreover, the clustering of these genes and their protein products into clusters of orthologous groups (COG) of proteins has been performed by the National Center for Biotechnological Information of the NIH for 21 complete genomes representing 17 major phylogenetic lineages. The availability of the complete DNA sequence for several microbial genomes and of the COG assignments has become extremely useful for considerations on the evolutionary perspectives of the most important pentose-specific binding proteins, membrane transporters, and enzymes in prokaryotes and some eukaryotes. A detailed review on the current state of genome sequencing with regard to monosaccharides and glycomics has been published [11].

The metabolism of D-ribose is relatively simple because of its easy link to reactions of the pentose phosphate cycle. D-Ribose is phosphorylated by ribulokinase, an enzyme that is again widely distributed among various organisms except those that are parasites depending on the metabolism of their host: a total of 40 proteins are listed in COG0524.

A wide range of organisms must synthesize pentoses from common hexoses, for which they utilize the pentose phosphate pathway that occurs widely among many prokaryotes and eukaryotes. In an oxidative reaction, glucose 6-phosphate is converted into ribulose 5-phosphate according to (*Eq. 1*).

The pentose phosphate pathway also catalyzes the interconversion of three-, four-, five-, six-, and seven-carbon sugars in a series of non-oxidative reactions. All these reactions occur in the cytosol, and in plants part of the pentose phosphate pathway also participates in the formation of hexoses from CO_2 in photosynthesis. Thus, D-ribulose 5-phosphate can be directly converted into D-ribose 5-phosphate by phosphopentose isomerase, or to D-xylulose 5-phosphate by phosphopentose epimerase. D-Xylulose 5-phosphate can then be combined with D-ribose 5-phosphate to give rise to sedoheptulose 7-phosphate and glyceraldehyde-3-phosphate. This reaction is a transfer of a two-carbon unit catalyzed by transketolase. Both products of this reaction can be further converted into erythrose 4-phosphate and fructose 6-phosphate. The four-carbon sugar phosphate erythrose 4-phosphate can then enter into another transketolase-catalyzed reaction with the D-xylulose 5-phosphate to form glyceraldehyde 3-phosphate and fructose 6-phosphate, both of which can finally enter glycolysis.

Some eukaryotes are unable to assimilate pentoses efficiently. For instance, the yeast *Saccharomyces cerevisiae* cannot metabolize D-xylose, although four strains were found to utilize



Equation 1

small amounts of xylose aerobically in the presence of a mixture of other facilitating substrates [12]. When [^{14}C]-xylose was utilized in the presence of ribose under anaerobic conditions, the radioactive label was detected mainly in xylitol, but no label in ethanol was found. Under aerobic conditions the radioactive label was distributed between xylitol (91.3%), CO_2 (2.6%), and biomass (1.7%); no other metabolic products were detected [12].

In those organisms that perform photosynthesis, further phosphorylation of D-ribulose 5-phosphate into D-ribulose 1,5-diphosphate by phosphoribulokinase represents an important prerequisite in CO_2 fixation. As has already been mentioned, the key and extensively studied enzyme involved in this reaction is D-ribulose 1,5-bisphosphate carboxylase/oxygenase. The subfamily divergence in this multigene family has been studied in certain plants of Triticeae and other families [13]. The occurrence of this enzyme in anoxic Archaea is very interesting, since it had to evolve in the absence of molecular oxygen [14].

In vertebrates, pentoses and their derivatives are known to cause hepatic ATP catabolism by trapping the inorganic phosphate in the form of glycerol 3-phosphate as a consequence of an increase in the NADH/NAD$^+$ ratio. In isolated rat hepatocytes, 5 mM D-xylulose decreased ATP by 80% and the intracellular phosphate by 70% within 5 min [15]. D-Xylulose increased the concentration of D-xylulose 5-phosphate approximately 25-fold in 5 min. However, the main cause of the depletion of ATP and inorganic phosphate in these cells was the dramatic increase (approximately 50-fold) in the concentration of seduheptulose 7-phosphate in 10 min. The D-ribose 5-phosphate concentration increased approximately 15-fold [15]. Xylulose 5-phosphate was found to be an important regulator of several metabolic pathways in rats, including the hepatic expression of glucose 6-phosphatase and phosphoenol pyruvate carboxykinase [16], regulation of L-type pyruvate kinase through the glucose response element [17], and inhibition of glucose utilization in liver by fatty acids [18].

2.1.2 Structural Investigations of Pentose Enzymes, Transporters and Receptors

Molecular details beyond those found by classical enzymology are rapidly becoming available from structural studies, and these should allow a better understanding of the nature of binding and catalytic activities of the various enzymes, membrane transporters, and receptor proteins. Along these lines of research, periplasmic carbohydrate-binding proteins have been studied in particular detail. The periplasmic arabinose-binding protein (*araF* gene product) of *Escherichia coli* has been extensively characterized by X-ray crystallography and directed mutagenesis by Quiocho's group. The structure of this protein complexed with L-arabinose molecule (in the $^4C_1$ pyranose chair conformation) has been solved by high-resolution protein crystallography [19]. The sugar molecule is buried in the cleft between the two lobes of this bilobate protein. All sugar hydroxy groups are hydrogen-bonded to side-chain residues: beta-OH to Lys_{10} and Asp_{90}, 2-OH to Lys_{10}, 3-OH to Glu_{14} and Asn_{205}, and 4-OH to Asn_{232}. Interestingly, Lys_{10}, Glu_{14}, and Asp_{90} are associated with one domain of the protein, while Asn_{205} and Asn_{232} are in the second lobe (*Fig. 2*).

Structural changes in the protein accompany binding, as indicated by the inaccessibility of the bound L-arabinose to the aqueous environment [19]. Interestingly, unlike many other carbohydrate binding proteins that display a strict anomeric selectivity, α-L-arabinose can be equally well accommodated in the carbohydrate-binding site of this protein. In this case, however, only

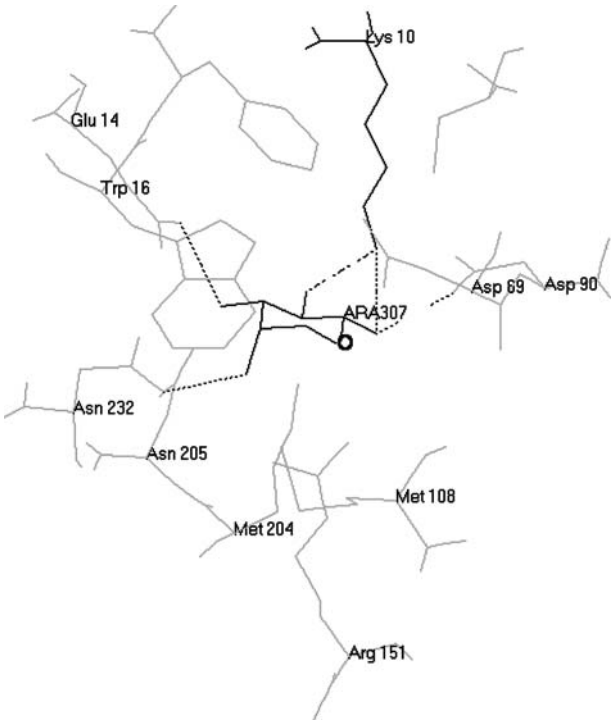


Figure 2
Arabinose binding protein in complex with β-L-arabinose (ARA)

a single hydrogen bond between the anomeric hydroxyl group and the aspartic acid D90 of the protein may be formed (*Fig. 3*).

L-Arabinose isomerase has been purified from *Escherichia coli*. The enzyme catalyzes the isomerization of L-arabinose to L-ribulose by a proton transfer mechanism in contrast to xylose isomerase which uses a hydride transfer mechanism to perform a similar isomerization [20]. This enzyme has high substrate specificity for L-arabinose, and its catalytic activity is dependent on manganese(II). It has been shown, however, that the enzyme could catalyze the exchange of the proton attached to C2 of arabinose with the solvent even in the absence of metal ion [20].

The periplasmic binding proteins for D-ribose and L-glucose/L-galactose of Gram-negative bacteria compete for a common inner membrane receptor in bacterial chemotaxis, and are the essential primary receptors for their respective membrane transport systems. The comparison of the high resolution structures of the periplasmic receptors for ribose (from *E. coli*) and glucose or galactose (from both *E. coli* and *Salmonella typhimurium*) highlights the features of these proteins important in their dual functions. Many structural features including some hydrophobic core interactions, a buried aspartate residue, and several unusual turns are conserved between the two proteins. The ribose receptor adopts the most closed structure, while the *Salmonella* glucose-galactose receptor has the most open structure [21]. The X-ray struc-

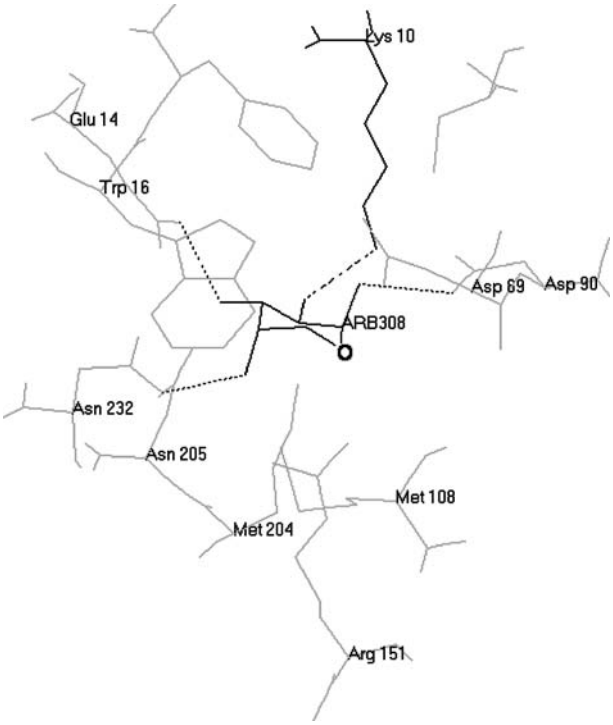


Figure 3
Arabinose binding protein with α-L-arabinose (ARB)

ture of the periplasmic ribose-binding protein from *E. coli* was solved at 3 Å resolution by the method of multiple isomorphous replacement. The protein consists of two highly similar structural domains, each of which is composed of a core of parallel β-sheets flanked on both sides by α-helices. The two domains are related to each other by an almost perfect two-fold axis of rotation, with the C termini of the β-strands of each sheet pointing toward the center of the molecule. Three short stretches of amino acid chain link these two domains, and presumably act as a hinge to allow relative movement of the domains in functionally important conformational changes. The ligand β-D-ribose is bound between the domains in its pyranose form, held by interactions with side-chains of the interior loops. The binding site is precisely tailored, with a combination of hydrogen bonding, hydrophobic and steric effects giving rise to tight binding (0.1 µM for ribose) and high specificity. Four out of seven binding-site residues are charged (2 each of aspartate and arginine) and contribute two hydrogen bonds each. The remaining hydrogen bonds are contributed by asparagine and glutamine residues. Three phenylalanine residues supply the hydrophobic component, packing against both faces of the sugar molecule [21].

The kinetic intermediates in the folding of ribose-binding protein have been examined by fluorescence spectroscopy in protein possessing of the engineered tyrosine residues [22]. The conformational changes of the ribose-binding protein upon the binding of sugar were inves-

tigated using small-angle X-ray scattering experiments. The monomeric form of the protein decreased in radius of gyration upon ligand binding. While the closed form of the protein is stabilized by the protein-sugar interaction, the open conformation is stable due to close contacts between the two domains [22].

2.1.3 Biomedicine

The use of pentoses and other monosaccharides in biomedicine should be based on utilizing the chemical biology of individual compounds and their derivatives as described in the preceding section. Although this is certainly true in a good number of examples, many other instances of the practical use of pentoses are still largely based on empirical observations.

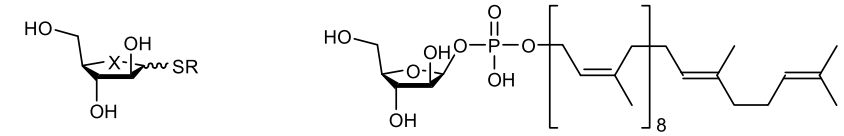
Some important biomedical implications of pentoses are based on the "non-specific" reactivity of the reducing hemiacetal group with proteins and other macromolecular components of living systems. The identification of pentosidine (imidazo[4,5-*b*]pyridinium) molecules synthesized through cross-linking of lysines and arginines in aged proteins by some pentoses (arabinose, lyxose, ribose, xylose) raised new questions concerning the role of pentoses in the ageing process [23]. Pentoses have been shown to participate in the formation of the advanced glycation end-products in collagen [24] and in rat liver histone octamers [25].

Biotypes of *Burkholderia pseudomallei*, a free-living organism causing the potentially lethal tropical infection melioidosis endemic in many parts of eastern Asia and northern Australia, can be distinguished based on their ability to assimilate L-arabinose. Whereas some soil isolated specimens can utilize this substrate, all clinical isolates tested so far could not. In addition to a cultivation test, the arabinose-negative clinical isolates can be diagnosed using a monoclonal antibody reacting with a high-molecular mass component expressed specifically only in L-arabinose minus strains [26]. A marked increase in the urinary secretion of arabinose and analogues of Kreb's cycle metabolites was observed in two brothers with autistic features (autism is a psychological disorder characterized by daydreaming hallucinations and disregard of external reality) [27].

Synthetic arabinose derivatives have recently shown promise as antimycobacterial, antiviral, and antithrombotic agents. Thus, a series of hydrolytically stable aza analogues of arabinofuranose was prepared, and evaluated as agents against *Mycobacterium tuberculosis* and *Mycobacterium avis.* The compounds were designed to mimic the putative arabinose donor (1-β-D-arabinofuranosyl decaprenol-1-phosphate) involved in the biogenesis of the essential cell wall polysaccharide, arabinogalactan (● *Fig. 4*). One of the synthesized compounds showed significant antimycobacterial activity in infected macrophage models [28].

1-β-D-Arabinofuranosylthymine and its 2,2′-anhydro derivative are potent anti-herpes simplex drugs. The three-dimensional crystal structure of 2,2′-anhydro-1-β-D-arabinofuranosylthymine revealed a rigid structure with the arabinose ring in the unusual O1′-*endo*-pucker conformation [29]. 4-Cyanophenyl- and 4-nitrophenyl-1,5-dithio-L- and -D-arabinopyranoside derivatives showed oral antithrombic effects in rats higher than beciparcil used as the reference compound ● *Fig. 5* [30].

Oligonucleotides containing 2′-deoxy-2′-fluoro-D-arabinose or -D-ribose form stable, triple-helical complexes stabilized by the intermolecular 2′-OH phosphate contacts and sugar puckering: these derivatives have promoted interest in relation to compounds synthesized for DNA targeting in vivo ● *Fig. 6* [31].



arabinofuranose mimic
X = O or NH,
R = aryl or alkyl side-chains

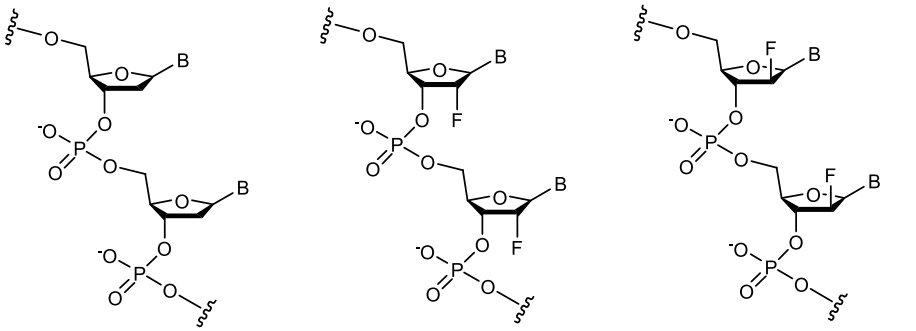
1-β-D-arabinofuranosyl decaprenol-1-phosphate

Figure 4
Antibacterial arabinose donor mimic

Beciparcil

α-L-arabinopyranoside

Figure 5
Beciparcil and analogues



DNA

2'F-RNA
2'-deoxy-2'-fluororibonucleic acid

2'F-ANA
2'-deoxy-2'-fluoroarabinonucleic acid

Figure 6
DNA and fluorinated version for DNA targeting

D-Lyxose, an epimer of D-arabinose rarely found in Nature, has been identified as an important component of the glycolipid adhesion receptors for mycobacteriophages, and the derivatives of this monosaccharide are currently used to evaluate the precise structural requirements for the attachment of the mycobacteriophage to the cell wall of its host [32]. Derivatives of α-nucleotides of D- and L-lyxofuranosylbenzimidazoles and ribofuranosylbenzimidazoles have been evaluated as antiviral agents *Fig. 7* [33].

In relation to the participation of ribose in ribonucleotides, experiments with oral or intravenous administration of this monosaccharide to patients with myoadenylate deaminase

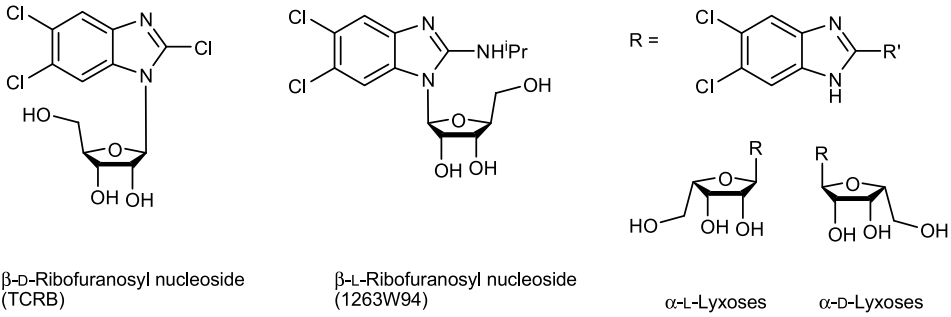


Figure 7
Antiviral agents

deficiency were performed: it appeared that the sugar may serve both as an energy source and as the stimulant of de novo synthesis of purine nucleotides [34]. Oral administration of 10 mmol/kg/day of D-ribose to patients with an inherited deficit of adenylosuccinase displaying severe psychomotor retardation and epilepsy resulted in improved behavior and a progressive reduction of seizure frequency [35].

D-Ribose has been added to the solutions used for the preservation of heart transplants since it maintains high levels of ATP and other high energy phosphates in these organs [36]. Infusion of ribose facilitates the redistribution of ^{201}Tl used in the diagnosis of coronary artery disease [37]. This effect is especially dramatic during the early imaging as compared with the late, 24 h imaging [37].

A xylose breath test has found usage for the diagnosis of small bowel bacterial overgrowth. These are performed with ^{13}C or ^{14}C techniques, and have been found to predict reliably small-bowel bacterial overgrowth in susceptible children [38,39].

Essential pentosuria is a metabolic condition characterized by deficiency in the major isoform of L-xylulose reductases in human tissues; homozygote individuals have only a residual activity of the enzyme [40], and thus excrete large concentrations of L-xylulose into urine. A minimum estimate of the frequency of the pentosuria allele in an Ashkenazi-Jewish population is 0.0127 [41]. This metabolic disease is absolutely harmless, especially because the specific and selective analytical techniques in use in modern diabetology should prevent a false positive diagnosis in essential pentosuria patients.

2.2 Hexoses

The predominant structures in which hexoses exert their biological activities are highlighted in *Fig. 8*. Hexoses most often found in biological systems, such as D-glucose, D-galactose, and D-mannose, are mostly observed in their cyclic pyranosyl forms (chair conformations).

Hexoses are monosaccharides of central importance in Nature since they are often the sole or the most important single source of easily available energy and feature as sources of carbon building blocks in a wide range of metabolic pathways. The origin of elementary hexoses during the chemical evolution has been proposed to occur according to a five-stage scenario [42].

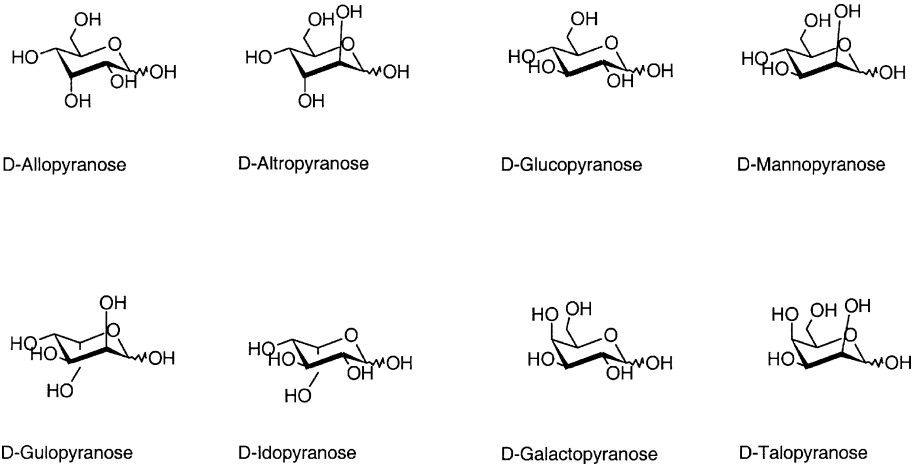


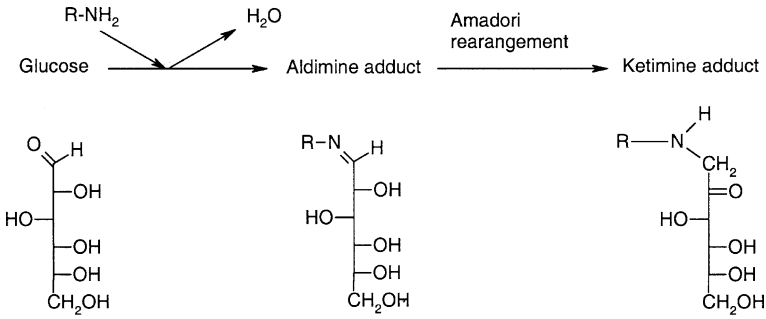
Figure 8
Structures of hexoses in which these monosaccharides interact with the biomolecules of living matter

The utilization of galactose as a key recognition molecule (based on its distinctive axial C4 hydroxy group [43], and its outermost location in glycoconjugates) may have evolved rather late, concomitantly with the evolution of multicellular organisms [44]. A possible scenario on the origin of saccharides and an investigation of saccharide structures, classic glycochemical reactions and biosynthetic features of monosaccharides, has been recently reviewed. In this hypothesis, fructose, glucose and mannose are the saccharides having been generated before the birth of life, whereas the others are characterized as "late-comer" saccharides as a result of the development of metabolism [45].

Glucose is clearly the most important single hexose and the most abundant of all monosaccharides and indeed the most abundant organic molecule on the planet. This unique feature of D-glucose is probably related to its chemical reactivity with proteins, which is thought to be the lowest of all the hexoses. Hence, it shows a minimal tendency to form glycation products with the reactive amino acid side-chains of proteins. On the other hand, glucose derivatives are formed under non-physiologically high concentrations of glucose by a mechanism involving the formation of Schiff-base adducts composed primarily of the ε-amino groups of lysyl residues in proteins and glucose, followed by Amadori rearrangement leading to more stable ketimine adducts shown in (*Eq. 2*).

Recently, the rapid preparation of carbohydrates has been facilitated by a synthetic route based on aldol coupling of three aldehydes used for the de novo production of polyol differentiated hexoses in only two chemical steps. The dimerization of alpha-oxyaldehydes, catalyzed by L-proline, is followed by a tandem Mukaiyama aldol addition-cyclization step catalyzed by a Lewis acid. Differentially protected glucose, allose, and mannose stereoisomers can each be selected, in high yield [46]. Microwave irradiation is becoming an increasingly popular method of carbohydrate synthesis and has been the subject of a recent review [47].

Hexoses play a central role in the carbohydrate metabolism in a large number of organisms. *Scheme 1* provides a basic summary allowing one to understand the central position of hexoses, and D-glucose in particular, in the intermediary metabolism.



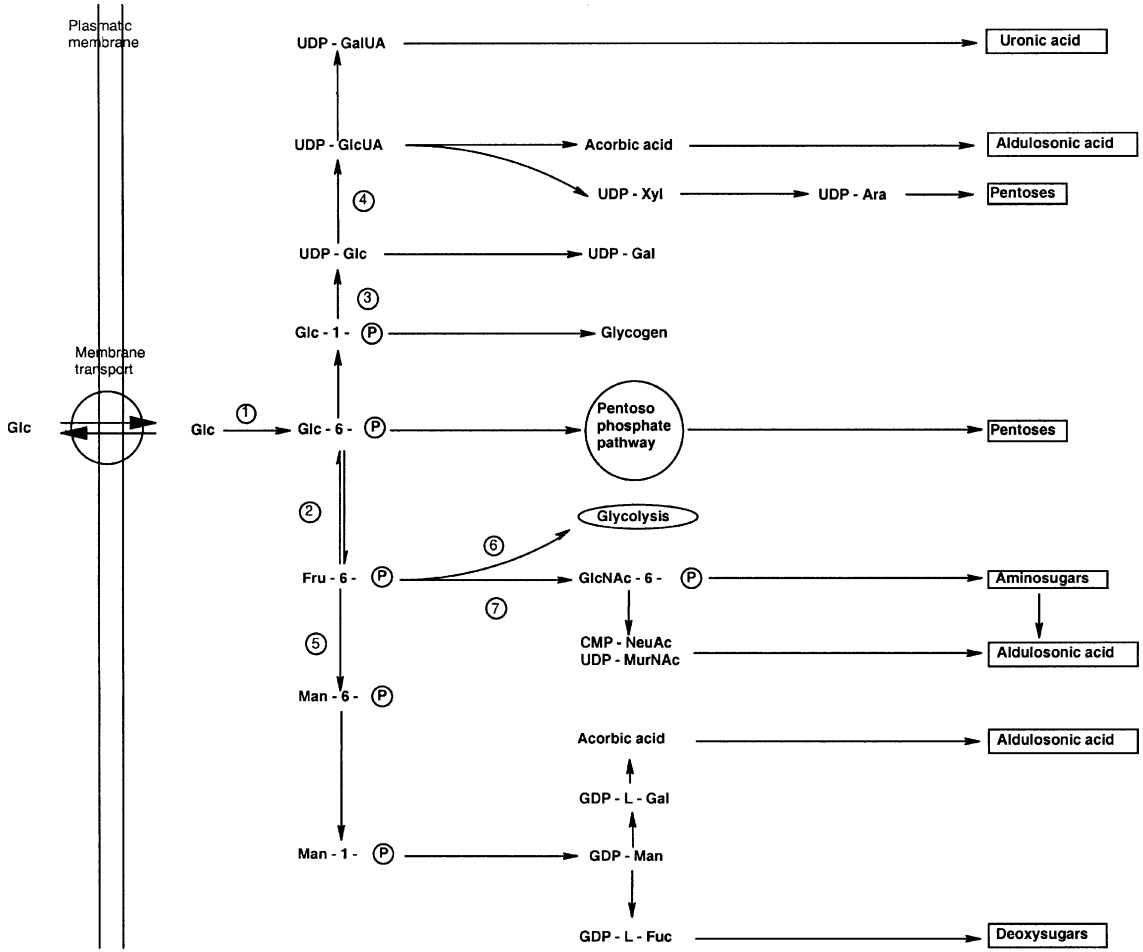
Equation 2

In many higher eukaryotes, including humans, glucose represents an important source of readily available energy for many cells, notable for cells of the nervous system. The concentration of glucose in the blood and other body fluids is thus under a tight control of hormonal and nervous regulatory circuits. The blood glucose level in a healthy person after an overnight fast is around 4.4 mmol L^{-1}, and during the day usually ranges from 4.4 to 6.6 mmol L^{-1}. The blood glucose level is controlled primarily by the liver, which can take up or release large amounts of glucose in response to hormonal signals and the level of glucose itself. Glucose is the sole fuel for the brain, except during prolonged starvation when it is replaced by ketone substances. In the resting state, the brain accounts for up to 60% of the utilization of glucose by the whole body. Non-invasive ^{13}C-NMR measurements have shown that, due to the existence of the well known blood-brain barrier, the concentration of glucose in the brain is about 1 mmol L^{-1} when the plasma level is 4.7 mmol L^{-1}. However, when the plasma level of glucose drops down below 2.2 mmol L^{-1} during hypoglycemia, the glucose concentration in the brain can approach the K_m of hexokinase (approximately 0.05 mmol L^{-1}) leading to a significant slow down in the rate of glycolysis.

Unlike the liver, muscles cannot produce glucose due to the lack of glucose 6-phosphatase. Therefore, they retain glucose from the blood supply as fuel. Since the rate of glycolysis in the actively contracting muscle far exceeds that of the citric acid cycle, part of the pyruvate is converted into lactate and alanine; these compounds are then rebuilt into glucose in liver. Since the adipose tissue is not able to phosphorylate endogenous glycerol because of the lack of a specific kinase, it needs glucose for the synthesis of triacylglycerols.

As opposed to D-hexoses, the role of L-hexoses in the metabolism of vertebrates has been the subject of much dispute. It became obvious from recent studies that only some L-sugars such as L-fructose or L-gulose can make an appreciable contribution to the metabolism, while some others such as L-glucose cannot [48]. Similarly to the situation with pentoses, microorganisms residing in the large intestine are thought to be mainly responsible for the uptake and utilization of these carbohydrates.

Intracellular hexose-specific receptors have been tightly linked with the processes of refolding and intracellular targeting of glycoproteins. Proteins are glycosylated cotranslationally immediately after their passage into the cisternae of the endoplasmic reticulum by *en bloc* transfer of the entire oligosaccharide precursor containing two moles of *N*-acetyl-D-glucosamine,



Scheme 1
Central role of hexoses in the intermediary metabolism of the monosaccharides. Individual reactions are catalyzed by the following enzymes: 1, glucokinase; 2, glucose 6-phosphate isomerase; 3, UTP:glucose 1-phosphate uridyltransferase; 4, UDP-glucose dehydrogenase; 5, mannose 6-phosphate isomerase; 6, phospofructokinase; 7, glutamine:fructose 6-phosphate amidotransferase

nine moles of mannose, and three moles of glucose. Two of these three terminally positioned glucoses are cleaved off soon after glycosylation, but the third glucose is used by the endoplasmic reticulum chaperones calreticulin (soluble) and calnexin (membrane-bound) to bind the partially folded glycoprotein, and assist in its correct folding [49,50]. Completely folded glycoproteins are then fully deglucosylated, and leave the endoplasmic reticulum. Folded glycoproteins seem to interact with the two chaperones to different extents, depending on the nature of the particular (glyco)protein [51].

The role of hexoses in eukaryotic organisms goes beyond their metabolic and structural functions, because they are seminal in the biological recognition phenomena as well. Hexose-specific receptors are crucial for the adherence of microbes to eukaryotic cells, uptake of aged serum glycoproteins, antigen phagocytosis, presentation to the cells of the immune system, and complement activation by bacteria. Mannose is the hexose that often features prominently in these reactions of the innate immune systems, since it is abundantly expressed at the surface of microbes, but only rarely occurs in structures of any of the various eukaryotic glycoconjugates. Antibodies against certain hexoses such as galactose are the most abundant natural auto-antibodies, and thus represent major concerns in transplantation immunology, especially in relation to the use of xenotransplants.

2.2.1 Structural Investigations of Hexose Enzymes, Receptors and Transporters

The structure of the D-allose binding protein from *Escherichia coli* (AlsB) bound to D-allose has been solved at 1.8 Å resolution using the molecular replacement method. As in other members of this family, especially the ribose-binding protein, with which the allose-binding protein shares 35% sequence identity, the structure consists of two similar domains joined by a three-stranded hinge region [52]. The protein exists in solution in a dynamic equilibrium of closed and open conformations which is an important part of its function. In the closed, ligand-bound form, D-allose is buried at the domain interface. Only the β-anomer of allopyranose is seen in the crystal structure, although the α-anomer can potentially bind with a similar affinity. Extensive hydrogen bonding as well as hydrophobic interactions are found to be important for the specific binding; altogether 10 residues from both the domains form 14 hydrogen bonds with the sugar. In addition, three aromatic rings, one from each domain with faces parallel to the plane of the sugar ring and a third perpendicular to this plane make up hydrophobic stacking surfaces for the ring hydrogen atoms. These results indicate that the aromatic rings forming the sugar binding cleft can sterically block the binding of any hexose epimer except D-allose, 6-deoxy-D-allose, or 3-deoxy-D-glucose; the latter two are expected to bind with reduced affinity due to the loss of some hydrogen bonds. However, the pyranose form of the pentose D-ribose can also fit into the allose-binding protein cleft, allowing this protein to function as a low affinity transporter for D-ribose as well [52].

The D-glucose/D-galactose-binding protein of 33 kDa found in the periplasm of bacterial cells serves as the primary high-affinity receptor for active transport, and as the receptor mediating chemotaxis towards both sugar epimers. The protein from *Escherichia coli* bindsD-glucose with $K_d = 2 \times 10^{-7}$ M, and binding of D-glucose is about two-fold tighter than binding of D-galactose. The crystal structures of both complexes refined to 2.0 Å resolution provide an understanding of the chemistry of the recognition of both monosaccharides in atomic detail [53,54]. In the two complex structures, the sugar ring is positioned identically in the

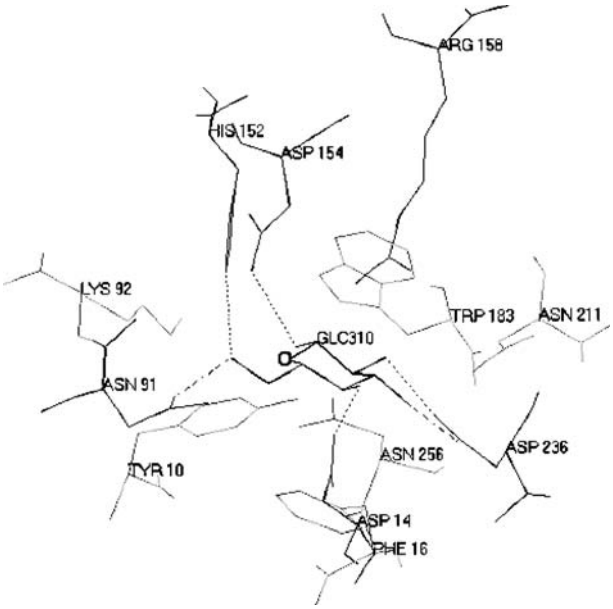


Figure 9
D-Galactose/D-glucose binding protein complexed with β-D-glucose (Glc)

binding site, and each hydroxy group common to both monosaccharides is involved in very similar cooperative hydrogen-bonding interaction with protein residues and ordered water molecules. Only the β-anomer of both monosaccharides is bound, and aspartic acid D154 is primarily responsible for accepting a hydrogen bond from the anomeric hydroxy group (*Fig. 9*).

Recognition of both sugar epimers is accomplished primarily by hydrogen bonding of D14 through an OD1 interaction with the equatorial OH-4 of D-glucose, or through the OD2 interaction with the axial OH-4 of D-galactose (see *Fig. 6*). Bacterial periplasmic binding proteins for L-arabinose, D-ribose, and D-glucose/D-galactose show noticeable similarities at the level of their three-dimensional structures, despite the lack of the obvious similarities in their primary structures [55]. All three binding proteins, which may have diverged from a common ancestor, serve as the primary receptors for bacterial high-affinity sugar transport systems. Many of the conserved residues reside in the binding site located in the cleft between the two lobes of these proteins. Moreover, D-ribose-binding protein and D-glucose/D-galactose-binding proteins also serve as receptors for chemotaxis. An exposed site located in one domain and including G_{74} is responsible for interaction with the trg transmembrane signal transducer thus triggering chemotaxis through the two respective monosaccharide-binding proteins. This site, together with the one proven to exist in the other domain of the protein, could be used by the signal transducer with which both binding proteins interact in such a way that it allows discrimination between the substrate-loaded "closed cleft" structure and the unligated "open cleft" conformation structures [55].

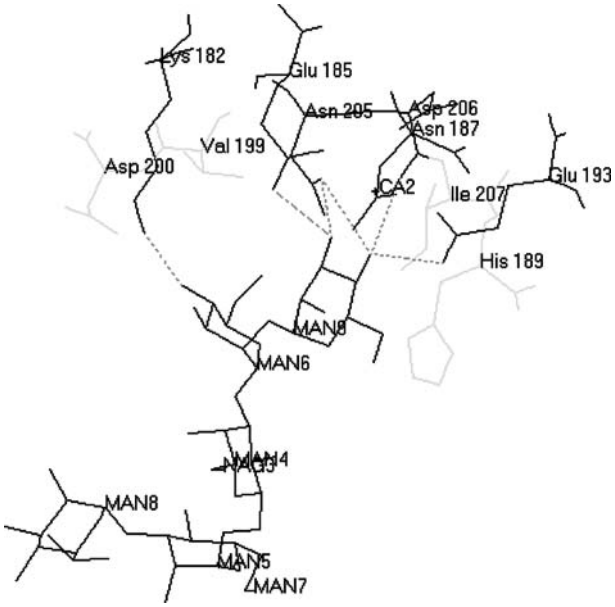


Figure 10
Mannose-binding protein A (lectin domain) complexed with Ca^{2+}(CA) and glucopeptide (MAN: mannose, NAG: *N*-acetylgalactosamine)

The crystal structure of human hexokinase in complex with the ATP analogue AMP-PNP has been solved at 2.25 Å resolution. One molecule of the ATP analogue is bound to each *N*-terminal domain of the dimeric enzyme in a surface cleft showing specific interaction with the nucleotides and a localized positive electrostatic effect. The molecular symmetry brings the two ATP analogue molecules to the center of two extended surface regions [56].

A group of *C*-type lectin receptors specific for D-mannose and related monosaccharides is important in the reactions of innate immunity based on the recognition of mannose-rich bacterial cell wall polysaccharides generally not occurring at the surface of eukaryotic cells. They involve both soluble mannose-binding proteins produced in the liver and secreted into serum as important acute-phase inflammation proteins, and the cell-bound macrophage mannose receptor that is an important pattern recognition receptor involved in host defense [57,58]. The rat mannose-binding protein became a prototype molecule for the structure of the protein fold of the entire *C*-type lectin superfamily [59]. The crystal structure of a C-type mannose-binding protein complexed with an oligomannose asparaginyl-oligosaccharide revealed that Ca^{2+} forms coordination bonds with the carbohydrate ligand. The carbohydrate specificity was determined by a network of coordination and hydrogen bonds that stabilizes the ternary complex of protein, Ca^{2+}, and sugar. The primary determinant of the monosaccharide specificity of this protein is the presence of hydroxy groups in positions equivalent to the 3- and 4-OH groups in D-mannose (*Fig. 10*).

This arrangement allows the recognition of some other hexoses (D-glucose), hexosamines (*N*-acetyl-D-glucosamine), and deoxyhexoses (L-fucose), while preventing some other mono-

saccharides from binding (e. g., D-galactose or *N*-acetyl-D-galactosamine both having axial 4-OH groups) [60]. Moreover, the structure of a mannose-binding carbohydrate-recognition domain in complex with a saccharide ligand suggests that two glutamic acid-asparagine pairs are essential determinants of ligand binding by this domain [60]. In *C*-type lectins that bind galactose with higher affinity than mannose one of these pairs is replaced by glutamine-aspartic acid. On the basis of this knowledge it was possible to engineer galactose-binding activity into a C-type mannose-binding protein by a simple switch in a position of a single amide group [61].

The natural forms of mannose-binding proteins are highly oligomeric molecules composed of trimers forming higher order complexes. The structure of the trimeric unit of the mannose-binding protein has been solved by protein crystallography to 1.8 Å resolution. The carbohydrate-binding sites in this trimeric protein are too far apart for a single trimer to bind multivalently with a typical mammalian high-mannose oligosaccharide. Thus, this protein can recognize pathogens selectively by binding avidly only to the widely spaced, repetitive sugar arrays on the pathogenic cell surface [62]. Molecular determinants of oligomer formation and complement fixation by these proteins have been investigated for proteins overexpressed in eukaryotic (Chinese hamster ovary) cells. Recombinant eukaryotic protein was posttranslationally modified in the same way as the native protein. Analysis of the isolated oligomers of mannose-binding protein revealed that the larger oligomers are more efficient in the activation of complement [63].

Macrophage mannose receptors are unique in their possession of eight linearly arrayed carbohydrate-recognition domains as a part of a single polypeptide. Although additional binding contributions coming from domains 5 and 7 may occur in complex saccharide ligand recognition, carbohydrate domain 4 of this receptor can alone account for the specificity of the entire receptor towards various monosaccharides [64]. For this reason, the greatest attention in binding and modeling studies has been paid to this particular protein segment of the entire receptor [65]. Recently, additional molecules belonging to the macrophage mannose receptor family have been cloned, sequenced and functionally characterized; of these, the dendritic cell receptor decalectin bearing a linear array of ten modules related to C-type lectin carbohydrate-recognition domains is worth mentioning [66].

The D-galactose-binding protein essential for transport and chemotaxis in *Escherichia coli* has been analyzed structurally by X-ray diffraction from the 3 Å resolution electron-density map obtained from multiple isomorphous replacement phases. Despite the lack of significant sequence homology, the overall course of the polypeptide backbone of the D-galactose-binding protein is remarkably similar to that of the L-arabinose-binding protein, the first of this family of proteins to be solved. Both structures are elongated (axial ratios 2:1) and composed of two globular domains. For both proteins, the arrangements of the elements of the secondary structure in both domains are identical: both lobes contain a core of β-plated sheet with a pair of helices on either side of the plane of the sheet. The four major hydrophobic clusters that stabilize the structure of the L-arabinose-binding protein are also present in the D-galactose-binding protein [67].

Extracellular or cell surface membrane receptors for D-galactose are widespread among higher eukaryotes. Isothermal titration calorimetry of the *C*-type lectin receptors from the tunicate *Polyandrocarpa misakiensis* revealed the presence of a single calcium atom per monomer with a dissociation constant of 2.6 µM, and confirmed the monosaccharide speci-

ficity of this receptor [68]. The X-ray crystal structure of the lectin complexed with D-galactose was solved at 2.2 Å resolution. Analytical ultracentrifugation revealed that the receptor behaved as a dimer in solution, and this was reflected by the presence of two subunits in the asymmetric unit with the dimeric interface formed by antiparallel pairing of the two *N*-terminal β-strands and hydrophobic interactions. The receptor adopts a typical *C*-type lectin fold with the main differences in structure from other *C*-type lectins being seen in the diverse loop region, and in the second α-helix, which is involved in the formation of the dimeric interface. D-Galactose is bound through coordination of the 3- and 4-hydroxy oxygen atoms with a bound calcium atom. Additional hydrogen bonds are formed directly between serine, aspartate, and glutamate side-chains of the protein and the sugar 3- and 4-hydroxy groups. Comparison of D-galactose binding by this receptor with D-mannose binding by the rat mannose-binding protein reveals how monosaccharide specificity is achieved in this receptor. A tryptophan side chain close to the monosaccharide-binding site, and the distribution of hydrogen-bond acceptors and donors around the 3- and 4-hydroxy groups of the sugar are essential determinants of the specificity. These elements are, however, arranged in a very different way than those in an engineered galactose-specific mutant of the mannose-binding protein, which is the only other D-galactose-specific receptor studied by rigorous structural procedures [61,68].

The uptake of monosaccharides across the plasma membrane of most mammalian cells is mediated by a family of facilitative transporters known as GLUTs. There are six known GLUT isoforms (GLUT1–GLUT7, GLUT6 is a pseudogene with no protein product) that are expressed in a tissue-specific manner [69]. Of the different GLUT transporters, GLUT5 is thought to function specifically as a fructose transporter with little ability to mediate the uptake of glucose. The intestine is a major site of GLUT5 expression consistent with the idea that this carrier plays a major role in the absorption of dietary fructose. Expression of this transporter has been studied in colorectal carcinoma cell line Caco-2, and found to be controlled by the carbohydrate content in the culture medium, and by the metabolic status of the cells [70]. More recently, it has been shown that the expression is co-regulated by thyroid hormone and D-glucose [70]. GLUT5 is also expressed in a number of other human tissues including skeletal muscle, adipose tissue, brain, and sperm, however, the precise physiological role in these tissues remains somewhat less well defined. It has been found lately that D-fructose transport facilitated by GLUT5 in skeletal muscle is saturable, is not inhibited by cytochalasin B, and is not acutely regulated by exercise. In contrast, the uptake of glucose is suppressed by cytochalasin B and regulated by exercise [71].

The domains responsible for D-fructose specificity of GLUT5 were investigated by creating chimeras of GLUT5 with the selective glucose transporter GLUT3, which were expressed in *Xenopus* oocytes. From the detailed analysis of these chimeras, it appeared that the amino terminal through the third transmembrane domain together with the polypeptide portion between the fifth and the eleventh transmembrane stretches seemed to be necessary for fructose uptake [72].

The crystal structure of the entire 6-phosphofructo-2-kinase/fructose 2,6-bisphosphatase, a key bifunctional regulator of both glycolysis and gluconeogenesis, has been solved at 2.0 Å resolution. The entire enzyme is a homodimer of 55 kDa subunits arranged in a head-to-head fashion, with each monomer consisting of an independent kinase and phosphatase domain. The location of γ-*S*-ATP and inorganic phosphate in the kinase and phosphatase domains,

respectively, allowed one to locate and describe the active sites of both domains. While the similarity of the phosphatase domain to a family of proteins containing the cofactor-independent phosphoglycerate mutases and acid phosphatases has already been mentioned, the kinase domain is clearly related to the superfamily of mononucleotide binding proteins, particularly with adenylate cyclase and the nucleotide-binding portion of the G proteins [73].

2.2.2 Regulation of Hexose Metabolism in Multicellular Eukaryotic Organisms

Glucose is the primary fuel for many eukaryotic cells. In higher, multicellular eukaryotes, the glucose levels in cells and body fluids are under a tight hormonal control assuring the homeostasis of the inner environment. During normal daily living, two peptide hormones – insulin and glucagon – regulate blood glucose levels in these organisms. Insulin acts on many cells in the body to produce both immediate and long-term effects. Within seconds insulin binds to its cell surface receptor, induces receptor autophosphorylation, and activation of the receptor protein tyrosine kinase. In subsequent minutes a 10–20 fold increase in the rate of glucose transport into many tissues (notably muscle cells and adipocytes of the fat tissues) occurs primarily due to an increase in the number of plasma membrane glucose transporters. This dramatic increase in the capacity of intracellular glucose uptake is not dependent on gene induction or new protein synthesis, but proceeds through an interesting mechanism involving rapid externalization on the cell surface of an abundant GLUT4 glucose transporter (an isoform found only in muscle cells and adipocytes) from intracellular endosome-like vesicles. Such an externalization results in an up to ten-fold increase in the number of cell surface transporters, which, together with some activation of these transporters accounts for the above-mentioned profound effects in glucose internalization. Continued exposure to higher concentrations (approximately 10^{-8} M) of insulin will eventually result in increased expression of liver enzymes that synthesize glycogen, and of adipocyte enzymes that synthesize triacylglycerols. Insulin also functions as a growth factor for many cells, but its growth-promoting actions may be the result of its binding to the insulin-like growth factor 1 receptor, rather then to the insulin receptor itself, especially since the latter receptor is significantly downregulated by prolonged exposure to insulin action.

The concentration of glucose in the blood of higher eukaryotes including humans has been thought to represent an important signal that regulates food intake and energy homeostasis. In the well-known Mayers's glucostatic hypothesis it has been postulated that small declines in glucose concentrations or utilization trigger meal initiation [74]. Both the liver and the brain have been hypothesized to monitor and respond to changes of immediately available energy, and control food intake [75,76]. Although elevated glucose concentrations have been shown to trigger insulin biosynthesis at the level of transcription and translation, the molecular mechanisms underlying the immediate transcriptional control have only recently been revealed. Evidence has been obtained that physiologically stimulated insulin secretion from the pancreatic β-cells promote insulin biosynthesis by enhancing insulin gene transcription in an autocrine manner [77]. As the most recent investigations point to the central role of the hexosamine biosynthetic pathway as a cellular "sensor" of energy and glucose availability. Nevertheless, current investigations support the hypothesis that glucose inhibits glutaminolysis in pancreatic β-cells in a concentration-dependent manner. In the basal interprandial state, glutaminolysis is partially turned on in β-cells because of glutamate dehydrogenase activation [78].

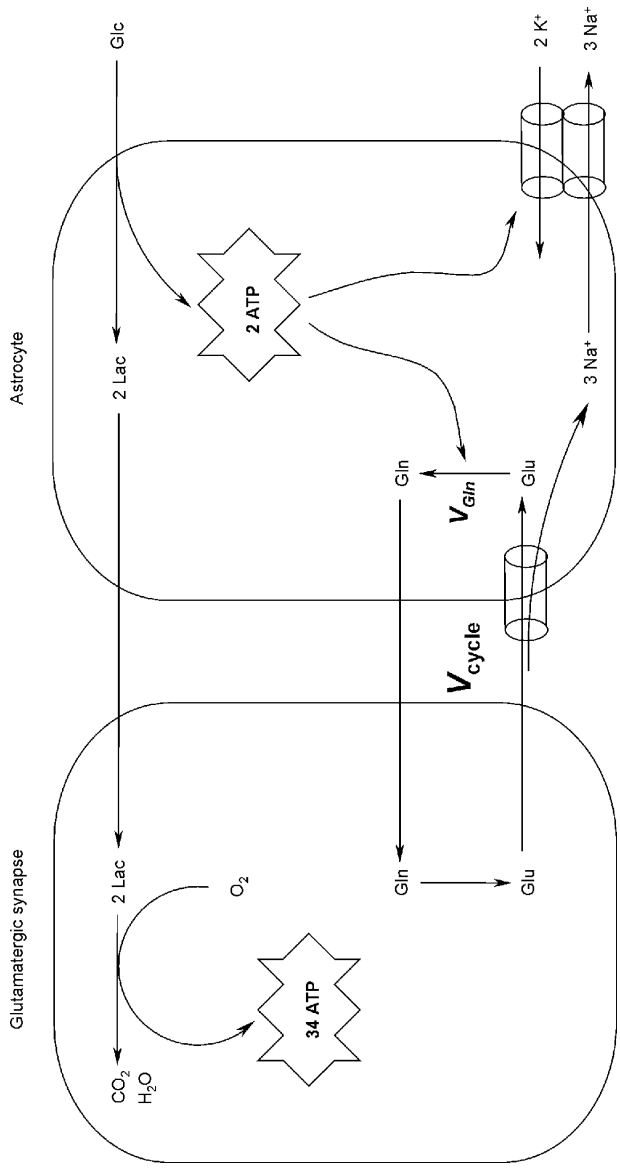
The brain has evolved as the central organ responsible for glucose sensing and glucose metabolism regulation. It receives neural inputs from glucosensors in the periphery but also contains neurons that directly sense changes in the blood glucose levels by using glucose as a signal to alter their transmission activities. Both obesity and diabetes are associated with several alterations in brain glucose sensing [79]. This role of glucose in the metabolism and coordinated activity of neurons has been clarified to considerable extent only recently [80]. As schematically depicted in ❯ *Scheme 2*, both the release of high concentrations of glutamate, and high energy production are required for glutamatergic neuronal synapses in order to mediate efficient activation of specific brain areas, processes that can be directly visualized by positron emission tomography or functional magnetic resonance imaging. Specialized glial cells in the brain, astrocytes, are assisting neurons in performing these tasks by accumulating glucose through their end-feet that cover virtually all capillary walls in the brain, and are enriched in glucose transporters. Inside the astrocyte, glucose is metabolized through glycolysis giving rise to both 2 mols of ATP necessary to drive the glutamate/glutamine cycle as well as transmembrane Na^+/K^+-ATPase (see ❯ *Scheme 2*).

Two moles of lactate are transported to neurons, and metabolized oxidatively to supply ATP for the energy demands of the latter cell [80]. Continuous (4 days) intracerebroventricular leptin infusion enhances insulin-stimulated glucose metabolism and favors the expression of the uncoupling proteins UCP1, UCP2 and UCP3 [81].

In other tissues, numerous hepatic and adipocytic genes are transcriptionally controlled by glucose and insulin. This is the case, for example, for the pyruvate kinase L gene in the liver and of the spot 14 gene in adipocytes [82]. At the hepatic level, the role of insulin is mainly to stimulate the synthesis of glucokinase, needed for phosphorylation of D-glucose to D-glucose 6-phosphate. However, an efficient regulation of the pyruvate kinase L gene by glucose also requires the synthesis of the glucose transporter GLUT2: in its absence the transcription of the gene is independent of the concentration of glucose in the medium. The role of GLUT2 seems to be in the depletion of the gluconeogenic cells in D-glucose 6-phosphate when cultivated without glucose.D-Glucose 6-phosphate seems to act through one of its pentose phosphate pathway metabolites, most probably through D-xylulose 5-phosphate controlling the activity of the cellular protein kinase/protein phosphatase cascades [82].

Ethanol impairs monosaccharide uptake and affects the synthesis, intracellular transport, subcellular distribution and secretion of these glycoproteins, suggesting alterations in glycosylation. Ethanol has been shown to increase the uptake of monosaccharides and the protein levels of GLUT1 but decreased those of mannosidase II. It alters the carbohydrate moiety of proteins and increases cell surface glycoproteins containing terminal mannose [83].

Membrane translocation of the glucose transporters in eukaryotic cells requires a complicated membrane architecture involving many accessory proteins as parts of the intracellular transport machinery. Thus, insulin-stimulated glucose transport inside the cells and translocation of the GLUT4 transporter require regulated interactions between the v-SNARE adaptor, the VAMP2 protein, and syntaxin 4. A novel syntaxin 4-binding protein, Synip, has also been isolated. Insulin induces dissociation of the Synip:syntaxin 4 complex due to an apparent decrease in the binding affinity of Synip for syntaxin 4. In contrast, the carboxy-terminal domain of Synip does not dissociate from syntaxin 4 in response to insulin stimulation, but inhibits glucose transport and GLUT4 translocation [84].



Scheme 2
Stoichiometry of glutamate (Glu)-mediated synaptic transmission and glucose usage. V_{cycle} is the reaction of the tricarboxylic acid cycle, V_{Gln} is the rate of the neurotransmitter cycle converting glutamate to glutamine (Gln)

Glucose deprivation has been shown to cause activation of Lyn kinase, c-Jun *N*-terminal kinase 1 and to increase the expression of basic fibroblast growth factor and c-Myc in adriamycin-resistant human breast cancer cells. Thus, glucose oxidation may be linked to cellular oxidative stress as evidenced from increase in the steady state of oxidized glutathione and intracellular peroxides [85]. A unique glucose-dependent apoptotic pathway is induced by the cellular oncogen c-Myc in transformed fibroblasts [86].

Noninsulin-dependent diabetes mellitus (NIDDM) is an increasingly common disease, and involves an increase in the capacity of the intestine to absorb monosaccharides. This has been linked to a combination of intestinal structural change with a specific increase in the expression of the monosaccharide transporters SGLT1, GLUT5, and GLUT2 [87].

Some hexoses are not common components of the metabolic pathways, but can have potential pharmacological effects. A notable example here is the all-*cis*-hexose D-allose that causes a striking down-regulation of hexose transport [88], most probably through the action of D-allose phosphate as the active down-regulator [89]. The fact that treatment of cells with tunicamycin, a well documented inhibitor of complex *N*-linked glycan biosynthesis, tightens this blockage of hexose transport has been correlated to impairment of glycosylation of specific glucose transporters [90]. More recent data indicate that D-allose-mediated depression of hexose transport in mammalian cells is mediated through the modulation of the intrinsic activity of the glucose transporter GLUT4 [91].

The ability of all eight D-aldohexoses to stimulate insulin release and biosynthesis was compared with their ability to serve as a metabolic substrate for the isolated islets of Langerhans. Insulin release and synthesis were stimulated by glucose or mannose, but not by allose, altrose, gulose, idose, galactose, or talose. Lactate formation was increased above values found in the absence of added substrate by 20 mmol/L glucose or mannose, but not by allose, altrose, gulose, galactose, or talose. These results support the substrate-site hypothesis for the recognition of sugars as the stimulus of insulin release and synthesis [92].

2.2.3 Generation of New Hexose-Utilizing Organisms by Metabolic Engineering

It has been shown that the diversion of lactose carbon through the tagatose pathway reduces the intracellular fructose 1,6-bisphosphate and growth rate in 20 strains of *Streptococcus bovis*. In these strains, glucose and the glucose moiety of lactose were metabolized by the Embden–Meyerhoff–Parnas pathway, but the galactose moiety of lactose was catabolized by the tagatose pathway through a scheme that bypasses D-fructose 1,6-bisphosphate [93]. Expression of the *pmi* gene from *Escherichia coli* encoding phosphomannose isomerase in *Zymomonas mobilis* leads to utilization of mannose as a novel growth substrate [94].

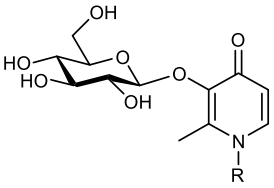
2.2.4 Biomedicine

Hexoses have found many biomedical applications, based mostly on their specific biological functions, but also on the unspecific (mostly protective) effects they have on tissues and organs. Considering the latter effects, hexoses have been found to exert a profound influence on the vitrification properties of ethylene glycol-based solutions used in cryobiology for the preservation of oocytes and other cells [95]. Glucose, mannose, and fructose have been shown to preserve physiological activities of various tissues, and have thus become important components

of the tissue and organ preserving solutions [96]. Solutions of both D-hexoses and L-hexoses, with the exception of D-idose, have been shown to protect cells significantly against thermal damage, increasing their survival factor up to one hundred times [97]. By direct determination of some hexoses such as D-talose, D-idose, and L-galactose it has been found that the polyhydroxy compounds must accumulate intracellularly for cellular heat protection. Many effects of hexoses on the cells of the vertebrate immune system have been recorded over the last period, but many of these effects were shown to be non-specific on later examination. As an example, monosaccharides were found to inhibit antigen-specific T-cell activation, but this inhibition was later interpreted as a chemical reactivity of these compounds dependent on the presence of an aldehyde function in their molecule [98].

The chemical reactivity of glucose leading to the glycation of abundant blood proteins such as albumin or hemoglobin is of eminent medical importance, and the determination of the fraction of glycated albumin or glycated hemoglobin has become a routine and reliable tool for the assessment of the long-term management of the diabetes patients [99]. Increased levels of the hemoglobin A1c serve as the definitive indicators of a long-term decompensation in these patients, and can even be used for postmortem diagnosis in cases of unsuspected diabetes mellitus leading to otherwise unexplained deaths [100]. Moreover, the advanced glycation end-products resulting from the reaction of hexoses with tissue proteins represent an important factor in diabetic complications [101] and Alzheimer's disease [102]. Aminoguanidine, a hydrazine-like molecule, has been studied extensively both in vitro and in vivo as an inhibitor of the formation of advanced glycation end-products [101]. More recently, a new class of inhibitors has been evaluated based on aryl or heterocyclic carboxamidophenoxy-isobutyric acids and related molecules [103].

The use of a series of small carbohydrates to stabilize proteins/peptides that misfold in Alzheimer's amyloid-beta peptides (Abeta) has been reported. The carbohydrates vary only in their distribution of potential H-bonding partners thus promoting various structural changes in Abeta. The utilization of a combinatorial strategy may represent a novel drug design strategy [104]. Two of the biochemical features of Alzheimer's disease that contribute to neurodegeneration are intracellular oxidative stress and elevated levels of trace metal irons including Fe^{III}, Cu^{II} and Zn^{II} [105]. Recently, a glucose prodrug which contains metal ion chelating ligands has been reported for use as a potential new therapy ❯ *Fig. 11*. The compound contains a functionalized bidentate hydroxypyridinone moiety and is designed to cross the blood brain barrier. The glucose-prodrug strategy solves the potential problem of premature metal binding by using carbohydrates as both the masking and directing substituent [106].



Glycosylated prodrug
where R = Me, Ph, 4-hydroxyphenyl
4-[I^{125}]-phenyl

Figure 11
Metal chelating compounds

Metabolic flux through the hexosamine biosynthetic pathway (HBP) is increased in the presence of high glucose and potentially stimulates the expression of genes associated with the development of diabetic nephropathy. A number of synthetic processes are coupled to the HBP, including enzymatic intracellular *O*-glycosylation, the addition of single *O*-linked *N*-acetylglucosamine monosaccharides to serine or threonine residues. In glomerular mesangial cells, HG- stimulated plasminogen activator inhibitor-1 (PAI-1) gene expression requires the HBP and the transcription factor, Sp1. It has been shown that among the pathways served by the HBP, *O*-GlcNAcylation, is obligatory for HG- induced PAI-1 gene expression and Sp1 transcriptional activation in mesangial cells [107].

The identification of leptin as an important hormonal factor involved in glucose utilization, diabetes, and the regulation of body weight lead to the current evaluation of this protein as a drug for the treatment of diabetes and obesity. In this respect, it has been found that leptin can inhibit glucose-induced insulin secretion from pancreatic islets [108], enhance insulin-stimulated glucose metabolism and expression of white adipose tissue uncoupling protein UCP2 and muscle uncoupling protein UCP3 [109], and contributes to the reduction of whole body adiposity by enhancing energy consumption in brown adipose tissue and muscle, and by attenuating the energy storage in the white adipose tissue [110].

Specifically designed boronic acids containing fluorescent sensors have been prepared, by a simple one-step synthetic procedure, in the progression toward non-invasive and continuous glucose monitoring. The sensors contain a boronic acid moiety as an effective glucose chelating group with the 6-methoxyquinolinium nucleus as a fluorescent indicator and are readily water soluble ❯ *Fig. 12* [111].

Similarly, 2-acrylamidophenylboronate has been synthesized and its ability to bind with glucose investigated both in solution and when integrated into a holographic sensor [112]. A glucose sensor comprising a reflection hologram incorporated into a thin, acrylamide hydrogel film bearing the *cis*-diol binding ligand, 3-actylamidophenylboronic acid, has been described. The potential for using this holographic sensor to detect real-time changes in bacterial cell metabolism was demonstrated by monitoring the germination and subsequent vegetative growth of Bacillus subtilis spores [113]. The boronic acid group has also been linked to an azo dye showing a color change from orange to purple in the presence of sugar at neutral pH [114].

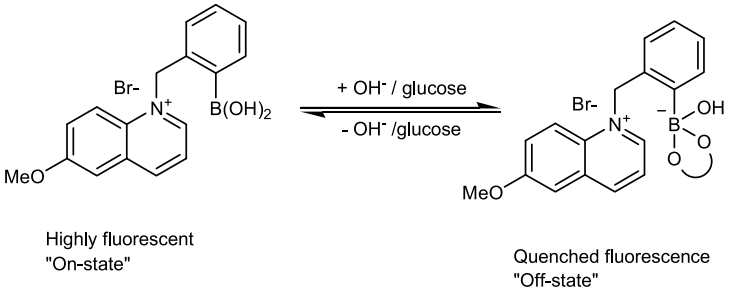


Figure 12
Glucose chelating boronic acid compounds used as sensors

Most cancer cells exhibit increased glycolysis and use this metabolic pathway for generation of ATP as a main source of their energy supply. This event is known as the Warburg effect and is considered as one of the most fundamental metabolic alterations during malignant transformation. Developing new therapies to inhibit glycolysis will provide a new class of anticancer agent that is likely to have broad therapeutic applications. These strategies have been recently reviewed [115].

Oral or intravenous administration of glucose has been evaluated as a potential possibility for sensitization before thermoradiotherapy. It has been concluded that intravenous and intravenous plus oral glucose administration are equally effective in inducing tumor acute acidification [116].

Recently, complexes of vanadyl(IV) with four monosaccharides were tested in two osteoblast-like cell lines (MC3T3E1 and UMR106). Many complexes caused stimulation of UMR106 proliferation (120% basal) in the range of 2.5 to 25 µmol/L. In the nontransformed osteoblasts, some vanadyl-saccharide complexes stimulated the mitogenesis (115% basal) in the same range of concentration. The glucose and sucrose complexes were the most efficient inhibitory agents (65 and 88% of inhibition vs. basal, respectively) for tumoral cells at 100 µmol/L. The galactose and turanose complexes exerted a similar effect in the nontransformed osteoblasts [117]. These vanadium derivatives are potential antitumor drugs.

A library of glycoconjugate benzenesulfonamides that contain carbohydrate-triazole units has been evaluated for their ability to inhibit three human transmembrane carbonic anhydrase isozymes hCA IX, hCA XII and hCA XIV which have potential as drug targets ❯ *Fig. 13* [118].

The glucose transporter GLUT1 has been identified as a highly sensitive marker of malignancy in body cavity effusions [119] suggesting that GLUT1 immunostaining may become useful in diagnostic cytopathology. Similarly, the expression of GLUT1 in colorectal carcinoma is a marker of poor prognosis [120]. Malignantly transformed cells transport glucose more efficiently than normal cells, both due to the increased expression of glucose transporters, and due to the expression of higher affinity transporters, such as GLUT2. Impaired regulation of GLUT2 seems to be the molecular cause of the hereditary maturity-onset diabetes of the young type I [121,122]. GLUT1 is expressed in complex hyperplasia with atypia and in adenocarcinoma. It has been recently shown that GLUT-1 immunostaining is useful in distinguishing benign hyperplasia from hyperplasia strongly associated with malignancy and that GLUT-1-

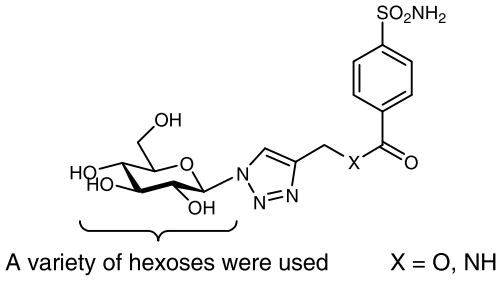


Figure 13
Benzenesulfonamide carbohydrates

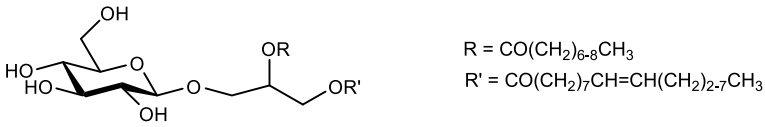


Figure 14
Antimicrobial diglyceride carbohydrates

mediated glucose transport may allow hypoxic tumor cells distant from stromal blood vessels to survive through glycolysis. This would suggest that the expression of GLUT-1 transporter may be closely related to the malignant transformation of epithelial endometrial tumors by supporting their increased need for glucose metabolism. GLUT-1 is preferentially expressed in atypical endometrial hyperplasia and endometrial adenocarcinoma [123].

The NO-mediated iron mobilization from cells has been linked to the presence of glucose. Due to its high affinity for iron, nitrogen monoxide (NO) affects cellular iron metabolism. It has been demonstrated that NO generators increase Fe-59 efflux from cells pre-labeled with Fe-59-transferrin by adding D-glucose to the reincubation medium this mechanism was potentiated. It has been shown that the metabolizable D-glucose and D-mannose stimulated NO-mediated Fe-59 mobilization. Other sugars, not easily metabolized by fibroblasts, had no effect [124]. This may have important implications for intracellular signaling by NO and also the NO-mediated cytotoxicity of activated macrophage that is partly due to iron release from tumor target cells.

The antimicrobial activity of a number of monoglucosyl diglycerides has been evaluated against gram-positive and gram-negative bacteria and fungi. It was shown that the 1,2-disubstition of the glyceride and the presence of an octanoyl side chain were essential structural features for activity ➤ *Fig. 14* [125].

Novel monosaccharide derivatives of D-allose have been evaluated as potent inhibitors of cellular proliferation in thymidine incorporation assays. These compounds show promise for the development of selective inhibitors of the growth of tumor cells as opposed to their normal counterparts [126]. It has also been shown that D-allose has a significant inhibitory effect on cancer cell proliferation in a dose-dependent manner, although the exact mechanism remains unclear [127]. This suggests that D-allose may be an effective adjuvant therapeutic agent against cancer in the future and may represent a new class of compounds with possible therapeutic potential. More specifically, D-allose has been shown to have a significant inhibitory effect on ovarian cancer cell proliferation in a dose-dependent manner, causing moderate G2/M arrest in the cell cycle, up-regulation of Cdk inhibitors p21 and p27 levels, and the induction of apoptosis in OVCAR-3 cells [128].

D-Mannose is an interesting hexose from the biomedical point of view, since the dietary inclusion of somewhat high doses of this monosaccharide was shown to provide a specific protection against various microorganisms in a wide variety of biological models. Thus, such treatment has been shown to be associated with the reduction in the numbers of *Salmonella enteretidis* in the caecal contact of chicken challenged by the food [129]. Similarly, mannose and its derivatives are able to prevent infection of macrophages and lymphocytes by HIV-1 [130], and adherence of type 1-piliated *Escherichia coli* to carbohydrate structures of vaginal mucosa that play a major role in the pathogenesis of ascending urinary tract infec-

tions in women [131]. Obviously, to make these treatments more useful for the above-mentioned applications, the proper clustering of D-mannose residues will be an important prerequisite.

The effect of several hexoses on wound healing has been investigated in cylindrical hollow sponge implants used as an inductive matrix for the growth of granulation tissues. While the infusion of the implant with D-mannose solutions inhibited the inflammation reaction in wound healing, D-galactose infusions enhanced wound healing [121].D-Galactose is an important hexose used clinically for the evaluation of hepatic functions in liver cirrhosis [122]. Clinical studies have been initiated aiming at the use of the D-galactose-specific liver receptors. It has been shown that treatment with galactose can prevent hepatic metastases in colorectal carcinoma patients [132], and that galactosylated cationic liposomes may be used as vehicles for gene transfer into the liver because of the interaction with the above asialoglycoprotein receptor [133].

D-Fructose has found numerous applications in biomedicine as an important component of tissue infusion and preservation solutions. It has been reported that fructose is the monosaccharide that best maintains the ATP and 2,3-diphosphoglycerate levels in red blood cells during the fourth to sixth week of whole blood storage [134]. Similarly, D-fructose protects rat hepatocytes against injury during the process of isolation and microencapsulation [135], and protects cells against oxidative injury by iron chelation [136]. Malabsorption of D-fructose is important in medical diagnosis since it has been shown to be associated with early signs of mental depression, although further studies will be needed to clarify the nature of this association [137]. Expression of the fructose transporter GLUT5 as an early diagnostic marker of human breast cancer opens opportunities for the development of novel strategies for the early diagnosis and treatment of this form of cancer [138].

The extensive expression of GLUT2 aswell as GLUT5 (glucose/fructose and fructose transporters, respectively) in malignant human tissues indicates that fructose may be a good energy substrate in tumor cells. This suggests that fructose uptake could be used for positron emission tomography imaging and, may possibly represent a novel target for the development of therapeutic agents in different human cancers [139]. Positron therapy with F-18-2-deoxy-2-fluoro-D-glucose (F-18-FDG) in a breast cancer animal model has been used to affect tumor growth rate and survival (positherapy) [140,141,142]. The molecular and cellular regulation of glucose transporter proteins in cancer has been recently reviewed [143] and [144].

Hydroxylated derivatives of topiramate, a novel antiepileptic drug based on D-fructose, are being evaluated in ongoing investigations [145,146].

D-Tagatose is a hexose of a great promise for biomedicine. This monosaccharide has been evaluated as an antioxidant and cytoprotective agent against chemically induced cell injury [147]. It has also become a center of attention for nutritionists and diabetologists as a new candidate sweetener that is as sweet as sucrose yet provides zero available energy [148]. Its application as a sweetener would certainly be supported by very low genotoxicity when tested in many assays including the Ames *Salmonella typhimurium* reverse mutation assay, the *Escherichia coli*/mammalian microsome assays, and chromosomal aberration assays in Chinese hamster ovary cells [149], and its low general influence on several clinically followed biochemical parameters [150].

Tagatose is currently being developed for the potential treatment of obesity and type two diabetes and is also under investigation for the potential treatment of anemia, hemophilia and

medical problems related to infertility, birth weight and excessive maternal food intake. Phase I and II clinical trials have been completed [151].
In Escherichia coli, L-arabinose isomerase (ECAI) catalyzes the isomerization of L-arabinose to L-ribulose. This enzyme is also of commercial interest as it catalyzes the conversion of D-galactose to D-tagatose in vitro. The crystal structure of ECAI was solved and refined at 2.6 angstrom resolution [152].

2.3 Heptoses

Despite the fact that many heptoses are by far less prominent in Nature than hexoses these monosaccharides are found both as metabolic intermediates, and as structural carbohydrates of bacterial cell walls.D-Sedoheptulose 7-phosphate is an important intermediate of the pentose cycle, and D-sedoheptulose 1,7-bisphosphate is present in plants as an intermediate of the dark phase of photosynthetic reactions. L-Glycero-D-*manno*-heptose was isolated from the oligosaccharides obtained by partial acid hydrolysis of the lipopolysaccharide from *Escherichia coli* K-12 strain W3100 [153] and *Haemophilus influenzae* [154]. Both L-glycero-D-*manno*-heptose and D-glycero-D-*manno*-heptose were isolated from the lipopolysaccharide of *Vibrio parahaemolyticus* [155].
Despite their quite recent history of investigation, two key enzymes of heptose metabolism have been purified and characterized biochemically, which open opportunities for structural studies [156].

2.3.1 Biomedicine

D-Mannoheptulose is currently used as a tool for competitive inhibition of D-glucose phosphorylation, metabolism, and functional effects in pancreatic islet β-cells, and other tissues [157]. An important diagnostic marker linking the early diabetes to obesity is the loss of D-mannoheptulose inhibition of glucose-induced insulin secretion in obese individuals, indicating the loss of sensitivity to competitive inhibition at the level of glucokinase [158]. On the other hand, the glucokinase activity in tumor cells is characterized by a high K_m for D-glucose, lack of feedback inhibition by D-glucose 6-phosphate, and potent inhibition by the specific glucokinase inhibitor, D-mannoheptulose. Moreover, mannoheptulose can cause efficient inhibition of glucose uptake by tumor cells, and inhibits growth of tumor cells in culture [159]. Rates of growth of human tumors in experimental animals are also dramatically reduced by as much as 65–80% during the administration of a dose of 1.7 mg of D-mannoheptulose per g body weight per day for 5 days. Thus, a naturally occurring D-mannoheptulose, which can be purified from avocados and is assumed to be of low toxicity, may become a potent natural anticarcinogen [159].

3 Aminosaccharides

Figure 15 presents the structures of the three most important and abundant hexosamines found in nature: *N*-acetyl-D-glucosamine, *N*-acetyl-D-galactosamine, and *N*-acetyl-D-mannosamine.

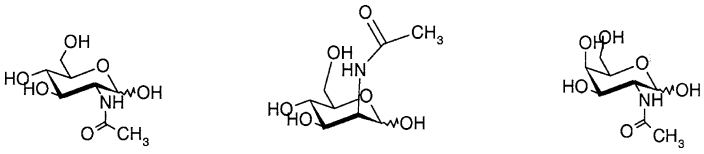
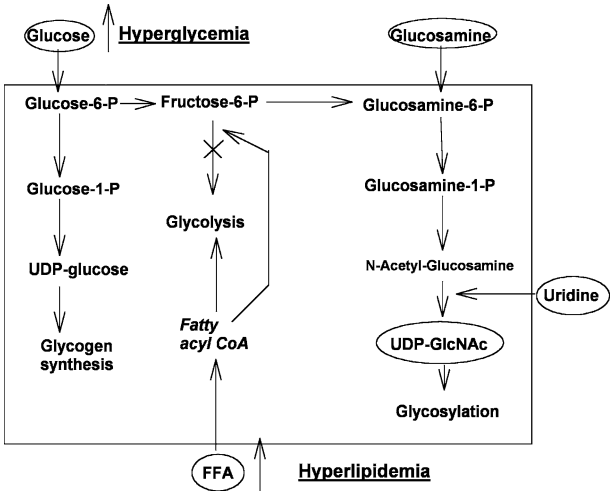


Figure 15
Structures of some aminohexoses in which these monosaccharides interact with the biomolecules of living matter

Aminosaccharides represent a specific class of sugars in that they bear both hydrophilic and ionizable, basic functionalities. However, in most naturally occurring aminosaccharides the amino group is acylated (most often acetylated), and this additional modification provides the *N*-acetylaminosaccharides with an extra group that is often critical in their biological interactions, despite its only moderately hydrophobic character (a similar aspect is also seen for the deoxysaccharides). In their acetylated form, aminosaccharides are among the most abundant constituents of cell wall polysaccharides, with chitin, a polysaccharide containing mostly linear β-1→4 linked *N*-acetyl-D-glucosamine chains, being the second most widely synthesized polysaccharide in nature. From this point of view, the biosynthesis of hexosamines catalyzed by glutamine: fructose 6-phosphate amidotransferase as the rate-limiting enzyme proceeds actively in a very wide range of organisms. In several vertebrates hexosamine biosynthesis has been identified as the critical cellular "sensor" of energy availability mediating the effects of glucose on the expression of several gene products [160]. Recently, evidence has been provided for rapid activation of the *obese* (*ob*) gene producing the "leaning" protein leptin in skeletal muscle by glucosamine [161]. Increased tissue concentrations of the end-product of the hexosamine biosynthesis pathways, UDP-*N*-acetyl-D-glucosamine (see *Scheme 3*), result in rapid and marked increases in leptin messenger RNA and protein levels.
Plasma leptin levels and leptin mRNA and protein levels in adipose tissues also increase. Most importantly, data indicate that leptin synthesis is also stimulated by either hyperglycemia or hyperlipidemia, which also increases tissue levels of UDP-*N*-acetyl-D-glucosamine in conscious rodents. These findings unveil an important biochemical link between the increased availability of certain nutrients and leptin expression [161].
The widespread availability of the acetylated hexosamines at the microbial surfaces [162], [163] certainly makes these monosaccharides favorite targets for both the innate and the adoptive immune system responses. The corresponding carbohydrate receptors are found among the animal lectins, anticarbohydrate antibodies as well as enzymes of the carbohydrate metabolism. While some of these recognition systems are related to the receptors already described in preceding sections due to their broad carbohydrate reactivities, a few systems seem to be unique for the recognition of the acetylated hexosamines. Notable among these proteins is the rat NKR-P1A antigen that has been characterized as an important activation receptor of rat natural killer cells, a specialized subset of lymphocytes involved in antitumor and antiviral defense [164,165]. The recombinant soluble form of this receptor binds specifically to both *N*-acetyl-D-glucosamine and *N*-acetyl-D-galactosamine linked to protein or chemically synthesized linkers [166,167]. Although some complex oligosaccharide sequences



Scheme 3

Sites of entry of glucose, glucosamine and uridine into the hexosamine biosynthesis pathway and the role of free fatty acids (FFA). Glucose is phosphorylated to produce glucose 6-phosphate, which is used mainly in glycogen synthesis and glycolysis pathways. The hexosamine pathway receives approximately 1 to 3% of the incoming glucose through the conversion of fructose 6-phosphate into glucosamine 6-phosphate by the rate-limiting enzyme glutamine:fructose 6-phosphate amidotransferase. Alternatively, glucosamine can enter the cell directly, and is phosphorylated to produce glucosamine 6-phosphate. Further metabolites are formed by the subsequent acylation and uridylation of the glucosamine 6-phosphate. Increased FFA availability generates increased levels of fatty-acyl-CoA, leading to inhibition (crossed reaction arrow) of glycolysis. An increased accumulation of fructose 6-phosphate leads to the increase in the concentration of the corresponding substrate. Modified according to ref. [161]

can possibly represent physiological ligands for this protein [168], and despite the evidence that linear oligomers of *N*-acetyl-D-glucosamine can be good ligands for this receptor [169], the carbohydrate-binding modules of this receptor have, even in their monomeric form, measurable reactivities towards simple *N*-acetyl-D-hexosamines, with *N*-acetyl-D-mannosamine being the best ligand in some oligosaccharide sequences [170].

3.1 Biomedicine

The central role of hexosamine biosynthesis intermediates as the sensors of the general metabolic status led to attempts to use the corresponding monosaccharides for feeding modulation in experimental animals.D-Glucosamine, *N*-acetyl-D-glucosamine, and 2,5-anhydro-D-mannitol were infused into the third cerebroventricle of the rat to compare their effects on food intake. D-Glucosamine at 24 μmol L^{-1} accelerated eating, and concomitantly increased plasma glucose, free fatty acids, and glycerol without affecting plasma insulin concentrations while *N*-acetyl-D-glucosamine induced feeding, but only upon oral administration [171].
The measurements of serum fructosamine may provide an alternative diagnostic tool to the more established monitoring methods, and become a clinically useful procedure for monitoring diabetes mellitus in veterinary medicine [172].

N-Acetyl-D-glucosamine has been evaluated as a clinically useful osmotic solute in peritoneal dialysis solutions [173]. This usage of the sugar would be well in accord with its good biological compatibility and beneficial influence on wound healing [174].

Synthetically engineered polyvalent dendrimer glucosamine conjugates have been used to prevent scar tissue formation. Water-soluble conjugates of D(+)-glucosamine and D(+)-glucosamine 6-sulfate have been shown to block fibroblast growth factor-2 mediated endothelial cell proliferation and neoangiogenesis in human Matrigel and placental angiogenesis assays [175].

Glucosamine represents one of the most commonly used drugs to treat osteoarthritis although the mechanism of its antiarthritic activity is still poorly understood. It has been shown that both glucosamine and *N*-acetylglucosamine inhibit IL-1 beta- and TNF-a-induced NO production in normal human articular chondrocytes. A novel mechanism for the inhibition of inflammatory processes has been suggested. The suppression of IL-1 beta -induced NO production is associated with inhibition of inducible NO synthase mRNA and protein expression. In addition, *N*-acetylglucosamine also suppresses the production of IL-1 beta -induced cyclooxygenase-2 and IL-6 [176].

In Alzheimer's disease, the major pathological features are diffuse and senile plaques that are primarily composed of the amyloid-beta (A beta) peptide. It has been proposed that proteoglycans and glycosaminoglycans (GAG) facilitate amyloid fibril formation and help stabilize the plaque aggregates, hence the use of therapeutics based on A beta -GAG interactions. The development of potential compounds based on the model of A beta -chondroitin sulfate binding can lead to effective inhibitors of the GAG-induced amyloid formation. It has been shown, that the sulfated D-GalNAc-4S, D-GalNAc-6S, and D-GalNAc-4S,6S, but not D-GalNAc, D-GlcNAc, or Delta UA-GalNAc, induce the fibrillar features of A beta -GAG interactions and compete with the intact chondroitin sulfate and heparin GAGs for A beta binding, as illustrated by competitive inhibition ELISAs [177].

Some unusual aminosaccharides can have antibacterial and antibiotic effects: this is the case of 3-amino-3-deoxy-D-glucose isolated from deep-sea bacteria (a *Bacillus* strain) collected at a depth of 4310 m [178]. More importantly, cellular receptors responsible for the adherence of microbes to eukaryotic cells are often lectin-type adhesins [179], and aminosaccharide treatment can result in the protection against microbial infections [180].

Finally, a notable antitumor effect of free or clustered aminosaccharides should be mentioned. Some aminosaccharides such as D-galactosamine have direct cytotoxic effects on certain experimental tumors [181], while D-mannosamine is selectively toxic for human malignant T-lymphoid cell lines [182]. It has been shown that aminosugars, especially D-mannosamine, produce hydrogen peroxide to cause DNA damage, which mediates apoptosis resulting in tumor growth inhibition [183].

4 Deoxysaccharides

By removing the hydrophilic hydroxy group from the structure of the respective monosaccharides in deoxy sugars, the hydrophobic character of certain areas in the molecule is increased. This fact may be considered to be rather minor, but in reality can often have a dramatic effect on the way how these molecules behave in solution, and most importantly, how they are rec-

ognized by their specific recognition proteins. In most instances, increasing the hydrophobic character increases the possibility for a deoxysaccharide to interact with various receptors or transporters more specifically. This is the reason why so many important antibiotics, drugs, and ligand mimetics fall into this class of monosaccharides. The above-mentioned effect seems to be most evident when the hydroxy group has been removed from the exocyclic C6 substituent. Classical example of this specific effect are provided by the 6-deoxymonosaccharides such as L-fucose (6-deoxy-L-galactose). In complex oligosaccharide sequences this monosaccharide is often localized in the terminal position, such as in blood group oligosaccharide sequences. This fact together with their relatively higher hydrophobic character, allows them to play such a dominant role in many important intercellular adhesive reactions [184], activation of cellular immunity, and in the immune antibody response [185].

Microorganisms can synthesize a wide range of deoxysaccharides and deoxyaminosaccharides that are not commonly contained in eukaryotic organisms, many of which occur as sugar components of important antibiotics. The nomenclature used for these aminosaccharides is still mostly trivial, and the biosynthetic pathways for some of these compounds are now being explored. *N*-Acetylviosamine is a component of bacterial lipopolysaccharide sequences, and the gene coding for an enzyme involved in dTDP-*N*-acetylviosamine (viaminose is 4-amino-4,6-dideoxy-D-glucose) synthesis has been identified in the gene cluster involved in the O7-specific lipopolysaccharide biosynthesis [186]. Similarly, the identification of the perosamine (perosamine is 4-amino-4,6-dideoxy-D-mannose) synthase gene of *Brucella melitensis* 16M may eventually allow us to understand the biosynthesis of the smooth lipopolysaccharides thought to be important virulence factors of these organisms [187].

The process of guanosine 5′-diphosphate-L-fucose biosynthesis is conserved throughout evolution from prokaryotes to man. In animals, GDP-L-fucose is the substrate for fucosyltransferases that participate in the biosynthesis and remodeling of the glycoconjugates including ABH blood group and Lewis-system antigens. The de novo pathway of GDP-L-fucose biosynthesis starts from GDP-D-mannose and proceeds through GDP-D-mannose 4,6 dehydratase and GDP-4-keto-6-deoxy-D-mannose epimerase/reductase. The genetic syndrome named severe leukocyte adhesion deficiency type II is known to result from a deficiency in this de novo pathway, resulting in the severe hypofucosylation of all glycoconjugates bearing fucose in various linkages. The crystal structures of the apo- and holoepimerase have been solved at 2.1 and 2.2 Å resolution, respectively. Each subunit of the homodimeric enzyme having a 34 kDa subunit is composed of two domains. The *N*-terminal domain binds $NADP^+$, and the *C*-terminal domain of 100 residues has α/β topology. The nicotinamide of the nucleotide and the connected ribose ring are located close to residues Ser_{107}, Tyr_{136}, and Lys_{140}. The enzyme is reminiscent of the fold observed in UDP-galactose epimerase [188].

4.1 Biomedicine

Deoxysaccharides express a range of extremely interesting biological activities that include antibiotic, antitumor, and antimalaria activities, as well as profound effects on cell–cell adhesion and cellular activation. Thus, 3-fluoro-3-deoxy-D-galactose displays high affinity for aldose reductase allowing the use of this compound as an excellent probe for investigations related to aldose reductase activity involved in cataract formation (see also the following

section) [189,190]. 2-Deoxy-D-ribose induces programmed cell death (apoptosis) in HL-60 leukemic cells [191], and has radiosensitizing effects on cancer cells [192,193]. Similarly, the apoptotic effects of the steroidal alkaloid solamargine have been shown to be dependent on the L-rhamnose moiety found in this compound [194]. 2-Deoxy-D-glucose enhances the cytotoxicity of topoisomerase inhibitors in human tumor cell lines [195].
1-Deoxy-D-xylulose 5-phosphate reductoisomerase inhibitors are potent antimalarial drugs [196]. The identical compound causes activation of T cells bearing Vγ9/Vδ2 receptors [197]. A peptide derivative of L-fucopyranose enhanced the activity of natural killer lymphocytes more than interleukin-2 [198]. The measurements of the serum levels of fucose are a important tumor diagnosis marker [199]. Interesting sugar mimetics have been recently constructed as E-selectin inhibitors in order to substitute an expensive oligosaccharide ligand, sialyl Lewis^x; in these derivatives, a five-membered fucose ring occurs which can still bind to calcium in the carbohydrate-recognition domain of E-selectin [200].

5 Alditols

Alditols are often observed as the end-products of monosaccharide metabolism that are stored in various cellular and tissue compartments with low redox potentials. However, there are also examples of alditols as important metabolic intermediates allowing the interconversion of rare forms of certain monosaccharides. In enteric bacteria such as *Escherichia coli* the hexitol galactitol is taken up through enzyme II of the phosphoenol pyruvate-dependent phosphotransferase system and accumulated inside the cell as galactitol 1-phosphate. The genes involved in galactitol metabolism have been cloned on a 7.8 kb DNA fragment [201].
The formation of alditols in eye lenses is one of the major complications in diabetic patients. Interestingly, the detailed molecular pattern of the lens alditols allows us to differentiate between the diabetic lens having increased concentrations of sorbitol 3-phosphate and fructose 3-phosphate, and the galactosemic lens, where the major products are galactitol 2-phosphate and galactitol 3-phosphate [202]. 2,5-Anhydro-D-mannitol is an analogue of fructose that has become popular in cellular metabolic studies using perfused rat liver [203].

5.1 Biomedicine

Alditols have been finding an increasing number of biomedical applications. Xylitol is a sweetener with important technological properties, an anticariogenic compound with low caloric value and negative dissolution heat. It has been used successfully in food formulations and in the pharmaceutical industry [204]. Xylitol has been shown to reduce plaque acid, and its anti-caries efficiency in combination with sodium fluoride has been confirmed by numerous clinical studies [205]. Xylitol is the only commercially used sugar substitute proven to have an antimicrobial effect on pneumococci [206]. Moreover, xylitol has been shown to display anti-adhesive effect on otopathogenic bacteria [207]. Determination of D-arabinitol/L-arabinitol ratios in urine may serve as an important diagnostic marker of invasive candidiasis [208]. Stable butyrate derivatives of D,L-xylitol have been shown to represent important cellular activation factors that can induce maturation and apoptosis in human acute myeloid leukemia cells [209].

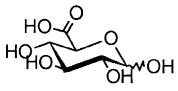
6 Uronic and Aldulosonic Acids

This group of monosaccharides includes structurally somewhat heterogeneous groups of compounds bearing the carboxylic group in their molecule (see ❯ *Fig. 16* for structures of some common compounds). The chemical effects of this group are primarily in conferring charge, and thus further increasing both the solubility of these derivatives, and their ability to interact with enzymes, receptors, and transporters through an ionic interaction. Ascorbic acid (vitamin C) is often arbitrarily included into this group of compounds that otherwise contain mostly glucuronate, galacturonate, mannuronate, sialic acids, *N*-acetylmuramic acid, 3-deoxy-D-*manno*-2-octulosonic acid, and other monosaccharides. However, one should keep in mind that ascorbic acid is ***not*** a monosaccharide bearing a carboxyl function, being chemically L-*threo*-hexulosono-1,4-lactone-2,3-enediol (see structures in ❯ *Fig. 16*).

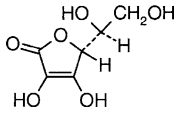
It has been suggested that the evolvement of sialic acids in microorganisms and higher animals may have stimulated evolution and rendered organisms less vulnerable to environmental attacks from viruses [210].

Glucuronic acid and sialic acid are normally present in conjugated forms. After degradation of these components in lysosomes, the free monosaccharides are released by a specific membrane transport system. The lysosomal sialic acid transporter from rat liver has been purified to apparent homogeneity in a reconstitutively active form. The transporter recognized structurally different types of acidic monosaccharides such as sialic acid, glucuronic acid, and iduronic acid. The transport was proton gradient dependent, and saturable with a K_m of approximately 0.4 mM [211].

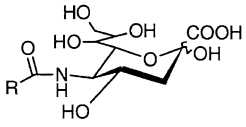
It has been suggested that de-*N*-acetyled sialic acid occurs on native gangliosides of certain tumors and cell lines. Although claims for their natural existence were based upon monoclonal antibodies and pulse-chase experiments, there have been no reports of their chemical detection until the report by Varki providing the first chemical proof for naturally occurring de-*N*-acetyl-



D-glucuronic acid



L-ascorbic acid



R = CH_3 N-acetyl-neuraminic acid

R = CH_2OH N-glycolyl-neuraminic acid

Figure 16
Structures of some charged monosaccharides in which these compounds interact with the biomolecules of living matter

gangliosides [212]. Synthetic de-*N*-acetyl-gangliosides have been shown to possess potent biological properties in vitro.

The interesting chemistry of sialic acid has been discussed in several reviews [213,214,215]. Sialic acid is an important component of complex oligosaccharides, when it is placed in the terminal position, and often masks the penultimate saccharide sequences. In this position, it interacts with numerous receptors including the influenza virus neuraminidase, and receptors mediating intercellular interactions such as the selectins (members of the *C*-type animal lectin family) and siglecs (or sialoadhesins, members of the immunoglobulin superfamily). The structure of influenza B neuraminidase in complex with sialic acid has been solved by X-ray crystallography at 2.2 Å resolution (*Fig. 17*) [216].

The good structural knowledge of this complex led to the development of structural mimetics of the sialic acid ligands with the advent of the development of inhibitors of viral binding and infection. One such approach has been to prepare derivatives of sialic acid with extended *N*-acyl groups, such as *N*-propanoyl, *N*-butanoyl, and *N*-pentanoyl; an elongation of *N*-acyl group resulted in inhibition of influenza A virus binding and infection [217]. The structure, function and evolution of viral sialate-*O*-acetylesterases has been recently reviewed with specific focus on the hemagglutinin-esterases of nidoviruses [218]. These enzymes help the virus achieve cell selectivity via protein-carbohydrate interactions of specific Sia subtypes.

The use of monosaccharide-based arrays to detect bacterial toxins, viruses and bacteria has been reported [219]. Arrays of *N*-acetyl galactosamine and *N*-acetylneuraminic acid derivatives were immobilized and were probed with fluorescently labeled bacterial cells and protein

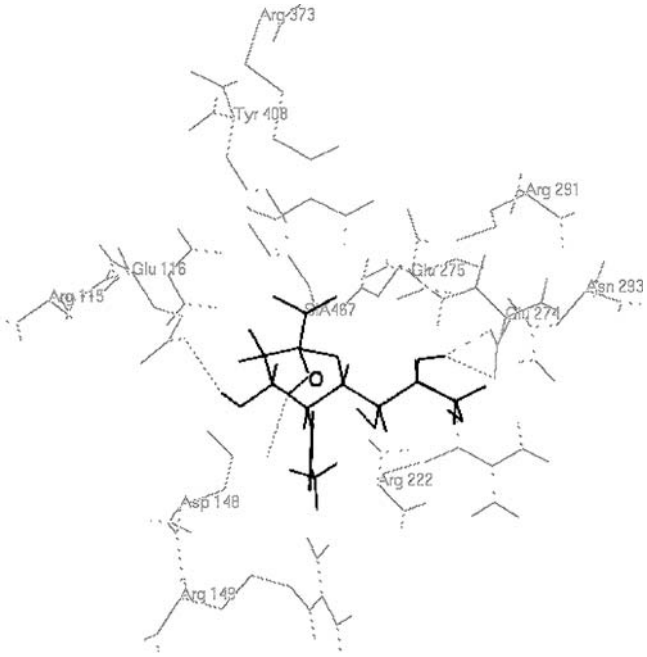


Figure 17
Influenza B neuraminidase and its complex with sialic acid (SIA)

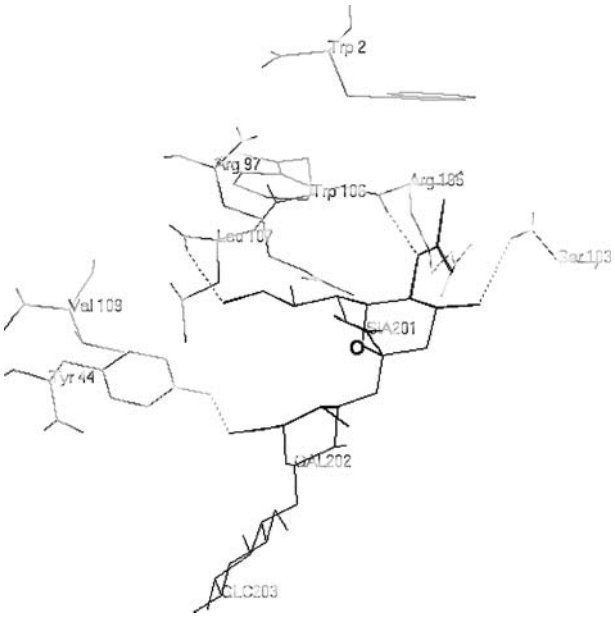


Figure 18
The structure of the *N*-terminal domain of sialoadhesin in complex with 3′siallylactose (SIA: sialic acid, GAL: galactose, GLC: glucose)

toxins. Salmonella typhimurium, Listeria monocytogenes, Escherichia coli and staphylococcal enterotoxin B (SEB) did not bind to either of the monosaccharides however, cholera toxin and tetanus toxin bound to GaINAc and Neu5Ac. This shows that the binding of the toxins to the carbohydrates is density dependent and semi-selective.

The involvement of sialic acid in the interaction of leukocyte adhesion receptors of the selectins with their ligands has been a subject of numerous reviews. It is obvious from these studies that sialic acid is involved in this interaction because of the charge of its carboxylic group that can be substituted with sulfate [220]. More recently, selectin/mannose-binding protein chimeras have been constructed to understand the mechanism involved in the binding of sialylated $Lexis^x$ oligosaccharide by this receptor [221].

The crystal structure of the *N*-terminal domain of sialoadhesin in complex with sialyllactose as ligand has been solved at 1.85 Å resolution [222]. The structure of the immunoglobulin domain contained in this receptor conforms to the V-set immunoglobulin-like fold but contains several distinctive features. The carboxylate of sialic acid forms a salt bridge with a side chain of Arg_{97} (◉ *Fig. 18*).

Hydrophobic interactions are made between the acetamido methyl group and the side chain of Trp_2, C9 of the glycerol side chain and the side chain of Trp_{106}, and C6 of galactose and the side chain of Leu_{107}. Hydrogen bonds are made between sialic acid and the main chain atoms of Ser_{103}, Leu_{107}, and Arg_{105}.

Although plant-derived ascorbate remains the major source of vitamin C in the human diet, its biosynthetic pathway is still unknown. However, recently it has been found that both

D-mannose and L-galactose are efficient precursors for ascorbate synthesis, that proceeds through a pathway involving GDP-D-mannose, GDP-L-galactose, L-galactose, and L-galactono-1,4-lactone [223].

6.1 Biomedicine

Monitoring of the levels of total sialic acid and lipid-associated sialic acid in serum is an important diagnostic tool for the detection of precancerous lesions [224]. Transition-state analogues derived from sialic acid are potent inhibitors of sialidases from various influenza virus strains, and these have been evaluated clinically as influenza remedies [225]. A detailed review of our current knowledge regarding the occurrence, specificity and function of sialic acid-specific lectins, particularly those that occur in viruses, bacteria and non-vertebrate eukaryotes has been recently published [226].

Several diseases are caused by defective sialic acid metabolism. Sialic acids cleaved off from degraded sialoglyco-conjugates are exported from lysosomes by a membrane transporter, named sialin, which is defective in two allelic inherited diseases: infantile sialic acid storage disease (ISSD) and Salla disease [227]. The function of sialin is basically known. Sialin's biosynthesis and intracellular trafficking as well as functional consequences of disease mutations have been characterized as well as the expression, localization, and targeting of the wild-type sialin [228]. A functional assay of human sialin has been developed and used to determine how pathogenic mutations affect transport [229].

"Sialic acid engineering" refers to the strategy whereby cell surface carbohydrates are modified by the bio-synthetic incorporation of metabolic intermediates, such as non-natural *N*-acetylmannosamine (ManNAc) analogues, into cellular glycoconjugates. Acetylated ManNAc analogues for use in sialic acid engineering applications have been developed and have been shown to be metabolized up to 900-fold more efficiently than their natural counterparts [230].

The protective effects of ascorbate as both a general antioxidant and a specific cofactor in the biosynthesis of collagen and other tissue proteins has been widely discussed in both the scientific and lay community, but its protective effects against the formation of cataract [231] and apoptosis of T cells [232] are probably worth mentioning. The estimated average requirement of ascorbate for the human diet is approximately 100 mg per day, and is usually adequately met by consumption of fruits and vegetables [233].

7 Phosphorylated and Sulfated Monosaccharides

Phosphorylated monosaccharides are common metabolic intermediates of all classes of monosaccharides, and as such they have already been mentioned in the preceding sections. Here, mostly the role of these compounds related to the cellular recognition events, and cell-cell adhesion events will be discussed. Two important receptors specific for D-mannose 6-phosphate are responsible for the proper delivery of the lysosomal enzymes into their target compartment. It appears from functional studies that there is certain redundancy in the biological function of these receptors; but knock-out mouse studies indicate that the lack of the large (205 kDa) cation-independent mannose 6-phosphate receptor appears to perturb lysosome function to a greater degree than the lack of the smaller (46 kDa) cation-dependent receptor [234]. The molecular basis for lysozyme recognition has been recently revealed

from X-ray analysis of the extracytoplasmic domain of the cation dependent receptor in complex with mannose 6-phosphate at 1.8 Å resolution [235]. The monomeric fold contained a nine-stranded, flattened β-barrel which bears a striking resemblance to avidin.
The second class of proteins that seem to bind mannose 6-phosphate and sulfated oligosaccharides derived from heparin are the seminal plasma spermadhesins, a group of small, structurally well defined proteins with possible biological function in sperm capacitation and sperm-egg interaction [236,237].
The sulfated monosaccharides have been gaining prominence recently as they seem to play a well documented role as the critical components of the oligosaccharide ligands for the selectins. The structural and biochemical progress in this area is very fast with both the contribution from chemical synthesis and enzymology of the specific sulfotransferases [238].

7.1 Biomedicine

Phosphorylated and sulfated monosaccharides promise a great potential from the point of view of their biomedical applications. Administration of certain phosphorylated monosaccharides such as fructose 1,6-bisphosphate during early reperfusion significantly improves the function of the post-ischemic heart [239]. The mannose 6-phosphate/insulin-like growth factor 2 receptor has been identified as the putative breast tumor suppressor factor, and it seems to be involved in multiple aspects of tumor pathophysiology including the deregulated growth, apoptosis, angiogenesis, and invasion [240]. The structure, distribution and function in the central nervous system of mannose 6-phosphate/insulin-like growth factor 2 has been reviewed [241].
The use of heparin-related sulfated oligosaccharides attached to long-circulating liposomes has been reported for cancer immunomodulation by the carbohydrate ligands for the lymphocyte receptor NKR-P1 [242].

8 Other Monosaccharides

Additional monosaccharide derivatives have been involved in biological studies that do not fall into the previously listed chemical categories. Of these, the acetylated monosaccharide derivatives have attracted the highest consideration. In particular, the insulinotropic action of L-glucose pentaacetate has been evaluated in animal and clinical studies [243]. Similarly, the tetraacetate derivative of 2-deoxy-D-glucose has been evaluated because of its cytotoxic effects upon lymphocytes, fibroblasts, and melanoma cells [244]. 2-Deoxy-D-glucose tetraacetate has been shown to display cytostatic and cytotoxic activity in various lines of tumoral cells. It was found to inhibit cell growth and confer chemosensitivity to cisplatin in two lines of human melanoma cells, poorly responsive to cisplatin [245].

9 Conclusions and Future Perspectives

In recent years much progress has been made in elucidating the roles of monosaccharides in chemical biology and biomedicine. The impact of monosaccharides within biological and medicinal applications has been further extended by using synthetic chemistry to prepare

novel derivatives. Of central importance is the incorporation of monosaccharides, as well as larger saccharides, within microarray technologies, to facilitate the identification of new carbohydrate-binding proteins as well as new ligands for already known carbohydrate-binding proteins. It is envisaged that both naturally occurring and synthetic carbohydrates will be required for the manufacture of microarrays that are sufficiently comprehensive and representative of entire glycomes. New leads to biological pathways that involve carbohydrate-protein interactions and new therapeutic targets are amongst the biomedically important outcomes that are anticipated from such applications [246,247]. Carbohydrate arrays fabricated on gold films have already been used to study carbohydrate-protein interactions with surface plasmon resonance (SPR) imaging. The binding of the carbohydrate-binding proteins concanavalin A (ConA) and jacalin to arrays composed of the monosaccharides mannose and galactose was monitored and adsorption coefficients were obtained [248]. Moreover, the interactions of single-walled carbon nanotubes with monosaccharides have been studied [249] and the ability of self-assembled receptors to stereoselectively recognize a saccharide has been described [250]. Another area of recent interest has been the use of monosaccharides in the design and synthesis of novel glycopeptides, nonpeptidal peptidomimetics for peptides, and abiotic proteins and this has necessitated the de novo synthesis of carbohydrate amino acids and carbohydrate scaffolds and carboproteins. This area has been recently reviewed [251,252]. Monosaccharides have also been used within combinatorial chemistry strategies, as scaffolds for the production of solid phase supported libraries. Such approaches have exploited the presence of multiple hydroxyl functionalities within the sugar scaffold, by derivatization of each hydroxyl group in turn to incorporate a range of different pharmacophores. This has even allowed entry to a variety of bicyclic and polycyclic polyfunctionalized structures of interest [253,254].

To conclude, considering our still rather immature understanding of many cellular receptors for the monosaccharides, it can be predicted that the collaboration between carbohydrate chemists and biologists in this fascinating field of research will continue to thrive in the coming years offering the potential for many new and significant biomedical advances.

References

1. Kremer RMD, Gallo-Rodriguez C (2004) Adv Carbohyd Chem Bi 59:9-67
2. Evans PG, Gemmell N, Osborn HMI, Osborne SD (2004) J Pharm Pharmacol 56:691
3. Larralde R, Robertson MP, Miller SL (1995) Proc Natl Acad Sci USA 29:8158
4. Mann CC (1999) Science 283:314
5. Schutte JB, de Jong J, van Weerden EJ, van Baak MJ (1992) Br Poult Sci 33:89
6. Schutte JB, de Jong J, van Weerden EJ, Tamminga S (1992) Br J Nutr 68:195
7. Himelbloom BH, Canale-Parola E (1989) Arch Microbiol 151:287
8. Seri K, Sanai K, Matsuo N, Kawakubo K, Xue C, Inoue S (1996) Metabolism 45:1368
9. Mahoney JR, Sako EY, Seymour KM, Marquardt CA, Foker JE (1989) J Surg Res 47:530
10. Boros LG, Lee PW, Brandes JL, Cascante M, Muscarella P, Schirmer WJ, Melvin WS, Ellison EC (1998) Med Hypotheses 50:55
11. Gabius HJ (2000) Naturwissenschaften 87(3): 108
12. van Zyl C, Prior BA, Kilian SG, Kock JL (1989) J Gen Microbiol 135:2791
13. Sasanuma T, Miyashita NT (1998) Genes Genet Syst 73:297
14. Watson GM, Yu JP, Tabita FR (1999) J Bacteriol 180:1569
15. Vincent MF, van den Berghe G, Hers HG (1989) FASEB J 3:1855
16. Massillon D, Chen W, Barzilai N, Prus-Wertheimer D, Hawkins M, Liu R, Taub R, Rossetti L (1998) J Biol Chem 273:228

17. Doiron B, Cuif MH, Chen R, Kahn A (1996) J Biol Chem 271:5321
18. Liu YQ, Uyeda K (1996) J Biol Chem 271:8824
19. Newcomer ME, Gilliland GL, Quiocho FA (1981) J Biol Chem 256:1321
20. Banerjee S, Anderson F, Farber GK (1995) Protein Eng 8:1189
21. Mowbray SL (1992) J Mol Biol 227:418
22. Kim D, Kim C, Park C (1994) J Mol Biol 240:385
23. Sell DR, Monnier VM (1989) J Biol Chem 264:21597
24. Paul RG, Avery NC, Slatter DA, Sims TJ, Bailey AJ (1998) Biochem J 330:1241
25. Gugliucci A (1994) Biochem Biophys Res Commun 203:588
26. Sirisinha S, Anuntagool N, Intachote P, Wuthiekanun V, Puthucheary SD, Vadivelu J, White NJ (1998) Microbiol Immunol 42:731
27. Shaw W, Kassen E, Chaves E (1995) Clin Chem 41:1094
28. Maddry JA, Bansal N, Bermudez LE, Comber RN, Orme IM, Suling WJ, Wilson LN, Reynolds RC (1998) Bioorg Med Chem Lett 8:237
29. Harrison DH, Schinazi RF, Rubin BH (1982) J Med Chem 25:1507
30. Bozo E, Boros S, Kuszmann J (1998) Carbohydr Res 311:191
31. Wilds CJ, Damha MJ (1999) Bioconjug Chem 10:299
32. Khoo KH, Suzuki R, Dell A, Morris HR, McNeil MR, Brennan PJ, Besra GS (1996) Biochemistry 35:11812
33. Migawa MT, Girardet JL, Walker JA, Koszalka GW, Chamberlain SD, Drach JC, Townsend LB (1998) J Med Chem 41:1242
34. Gross M, Reiter S, Zollner N (1989) Klin Wochenschr 67:1205
35. Salerno C, D'Eufemia P, Finocchiaro R, Celli M, Spalice A, Iannetti P, Crifo C, Giardini O (1998) Biochim Biophys Acta 1453:135
36. Muller C, Zimmer H, Gross M, Gresser U, Brotsack I, Wehling M, Pliml W (1998) Eur J Med Res 3:554
37. Perlmutter NS, Wilson RA, Angello DA, Palac RT, Lin J, Brown BG (1991) J Nucl Med 32:193
38. Dellert SF, Nowicki MJ, Farell MK, Delente J, Heubi JE (1997) J Pediatr Gastroenterol Nutr 25:153
39. Nogueira I, Ramos SS (1997) J Bacteriol 179:7705
40. Lane AB (1985) Biochem Genet 23:61
41. Lane AB, Jenkins T (1985) Ann Hum Genet 49:227
42. Hirabayashi J (1996) Q Rev Biol 71:365
43. Drickamer K (1997) Structure 5:465
44. Drickamer K, Taylor ME (1998) Trends Biochem Sci 23:321
45. Hirabayashi J (2004) Trends Glycosci Glyc 16(88):63
46. Northrup AB, MacMillan DWC (2004) Science 305(5691):1752
47. Corsaro A, Chiacchio U, Pistara V, Romeo G (2004) Curr Org Chem 8(6):511
48. Livesey G, Brown JC (1995) J Nutr 125:3020
49. Keller SH, Lindstrom J, Taylor P (1998) J Biol Chem 273:17064
50. Cannon KS, Helenius A (1999) J Biol Chem 274:7537
51. Zhang Q, Salter RD (1998) J Immunol 160:831
52. Chaudhuri BN, Ko J, Park C, Jones TA, Mowbray SL (1999) J Mol Biol 286:1519
53. Vyas NK, Vyas MN, Quiocho FA (1988) Science 242:1290
54. Vyas MN, Vyas NK, Quiocho FA (1994) Biochemistry 33:4762
55. Vyas NK, Vyas MN, Quiocho FA (1991) J Biol Chem 266:5226
56. Rosano C, Sabini E, Rizzi M, Deriu D, Murshudov G, Bianchi M, Serafini G, Magnani M, Bolognesi M (1999) Structure 7:1427
57. Weis WI, Taylor ME, Drickamer K (1998) Immunol Rev 163:19
58. Stahl PD, Ezekowitz RA (1998) Curr Opin Immunol 10:50
59. Weis WI, Kahn R, Fourme R, Drickamer K, Hendrickson WA (1991) Science 254:1608
60. Weis WI, Drickamer K, Hendrickson WA (1992) Nature 360:127
61. Drickamer K (1992) Nature 360:183
62. Weis WI, Drickamer K (1994) Structure 2:1227
63. Wallis R, Drickamer K (1999) J Biol Chem 274:3580
64. Taylor ME, Bezouška K, Drickamer K (1992) J Biol Chem 267:1719
65. Hitchen PG, Mullin NP, Taylor ME (1998) Biochem J 333:601
66. Steinman RM (1996) Exp Hematol 24:859
67. Vyas NK, Vyas MN, Quiocho FA (1983) Proc Natl Acad Sci USA 80:1792
68. Poget SF, Legge GB, Proctor MR, Butler PJ, Bycroft M, Williams RL (1999) J Mol Biol 290:867
69. Stephens JM, Pilch PF (1995) Endocr Rev 16:529

70. Matosin-Matekalo M, Mesonero JE, Laroche TJ, Lacasa M, Brot-Laroche E (1999) Biochem J 339:233
71. Darakhshan F, Kristiansen S, Richter E, Hundal HS (1997) Biochem Soc Trans 25:473S
72. Buchs AE, Sasson S, Joost HG, Cerasi E (1998) Endocrinology 139:827
73. Hasemann CA, Istvan ES, Uyeda K, Diesenhofer J (1996) Structure 4:1017
74. Mayer J, Thomas DW (1967) Science 156:328
75. Friedman MI (1997) Proc Nutr Soc 56:41
76. Campfield LA, Smith FJ (1990) Int J Obes 14:15
77. Leibiger IB, Leibiger B, Moede T, Berggren PO (1998) Mol Cell 1:993
78. Gao ZY, Li G, Najafi H, Wolf BA, Matschinsky FM (1999) Diabetes 48:1535
79. Levin BE, Dunn-Meynell AA, Routh VH (1999) Am J Physiol 276:R1223
80. Magistretti PJ, Pellerin L, Rothman DL, Shulman RG (1999) Science 283:496
81. Cusin I, Zakrzewska KF, Boss O, Muzzin P, Giacobino JP, Ricquier D, Jeanrenaud B, Rohner-Jeanrenaud F (1998) Diabetes 47:1014
82. Kahn A (1997) Biochimie 79:113
83. Tomas M, Fornas E, Megias L, Duran JM, Portoles M, Guerri C, Egea G, Renau-Piqueras J (2002) J Neurochem 83(3):601
84. Min J, Okada S, Kanzaki M, Elmendorf JS, Coker KJ, Syu LJ, Noda Y, Saltiel AR, Pessin JE (1999) Mol Cell 3:751
85. Blackburn RV, Spitz DR, Liu X, Galoforo SS, Sim JF, Ridnour LA, Chen JC, Davis BH, Corry PM, Lee YJ (1999) Free Radic Biol Med 26:419
86. Shim H, Chun YS, Lewis BC, Dang CV (1998) Proc Natl Acad Sci USA 95:1511
87. Dyer J, Wood IS, Palejwala A, Ellis A, Shirazi-Beechey SP (2002) Am J Physiol-Gastr L 282(2):G241
88. Ullrey DB, Kalckar HM (1986) Proc Natl Acad Sci USA 83:5858
89. Kalckar HM (1991) Annu Rev Biochem 60:1
90. Ullrey DB, Kalckar HM (1987) Proc Natl Acad Sci USA 84:3678
91. Pratt SE, Colby-Germinario S, Manuel S, Germinarion RJ (1994) J Cell Physiol 161:580
92. Ashcroft SJ, Lowry M (1979) Diabetologia 17:165
93. Bond DR, Tsai BM, Russell JB (1998) Appl Microbiol Biotechnol 49:600
94. Weisser P, Kramer R, Sprenger GA (1996) Appl Environ Microbiol 62:4155
95. Kuleshova LL, MacFarlane DR, Trounson AO, Shaw JM (1999) Cryobiology 38:119
96. Wada H, Okada Y, Uzuo T, Nakamura H (1998) Brain Res 788:144
97. Henle KJ, Monson TP, Nagle WA, Moss AJ (1985) Int J Hyperthermia 1:371
98. Rhodes J, Zheng B, Lifely MR (1992) Immunology 75:626
99. Turner AP, Chen B, Piletsky SA (1999) Clin Chem 45:1596
100. Khuu HM, Robinson CA, Brissie RM, Konrad RJ (1999) J Forensic Sci 44:643
101. Rahbar S, Nadler JL (1999) Clin Chim Acta 287:123
102. Munch G, Schinzel R, Loske C, Wong A, Durany N, Li JJ, Vlassara H, Smith MA, Perry G, Riederer P (1998) J Neural Transm 105: 439
103. Rahbar S, Kumar Yernini K, Scott S, Gonzales N, Lalezari I (1999) Biochem Biophys Res Commun 262:651
104. Fung J, Darabie AA, McLaurin (2005) J Biochem Bioph Res Comm 328(4):1067
105. Gaggelli E, Kozlowski H, Valensin G (2006) Chem Rev 106:1995
106. Schugar H, Green DE, Bowen ML, Scott LE, Storr T, Bohmerle K, Thomas F, Allen DD, Lockman PR, Merkel M, Thompson KH, Orvig C (2007) Angew Chem Int Ed 46:1716
107. Goldberg HJ, Whiteside CI, Hart GW, Fantus IG (2006) Endocrinology 147(1):222
108. Roduit R, Thorens B (1997) FEBS Lett 415:179
109. Cusin I, Zakrzewska KF, Boss O, Muzzin P, Giacobino JP, Ricquier D, Jeanrenaud B, Rohner-Jeanrenaud F (1998) Diabetes 47:1014
110. Wang JL, Chinookoswong N, Scully S, Qi M, Shi ZQ (1999) Endocrinology 140:2117
111. Badugu R, Lakowicz JR, Geddes CD (2005) Bioorg Med Chem 13(1):113
112. Yang XP, Lee MC, Sartain F, Pan XH, Lowe CR (2006) Chem-A Eur J 12(33):8491
113. Lee MC, Kabilan S, Hussain A, Yang XP, Blyth J, Lowe CR (2004) Analyt Chem 76(19):5748
114. DiCesare N, Lakowicz JR (2001) Org Lett 3(24):3891–3893
115. Pelicano H Martin DS, Xu RH, Huang P (2006) Oncogene 25(34):4633
116. Leeper DB, Engin K, Wang JH, Cater JR, Li DJ (1998) Int J Hyperthermia 14:257
117. Barrio DA, Cattaneo ER, Apezteguia MC, Etcheverry SB (2006) Can J Physiol Pharmacol 84(7):765
118. Wilkinson BL, Bornaghi LF, Houston TA, Innocenti A, Vullo, Supuran CT, Poulson S-A (2007) Bioorg Med Chem Lett 17:987

119. Burstein DE, Reder I, Weiser K, Tong T, Pritsker A, Haber RS (1998) Mod Pathol 11:392
120. Haber RS, Rathan A, Weiser KR, Pritsker A, Itzkowitz SH, Bodian C, Slater C, Weiss A, Burstein ED (1998) Cancer 83:34
121. Kossi J, Peltonen J, Ekfors T, Niinikoshi J, Laato M (1999) Eur Surg Res 31:74
122. Garello E, Battista S, Bar F, Niro GA, Capello N, Rizzetto M, Molino G (1999) Dig Dis Sci 44:782
123. Ashton-Sager A, Paulino AFG, Afify AM (2006) Appl Immunohist Mol Morph 14(2):187
124. Watts RN, Richardson DR (2001) J Biol Chem 276(7):4724
125. Cateni F, Bonivento P, Procida G, Zacchigna M, Favretto LG, Scialino G, Banfi E (2007) Bioorg Med Chem 15:815
126. Colquhoun A, Alaluf S, Brandley A, Gemmel N, Gibbs G, Osborn HM, Harwood LM, Newsholme EA (1997) Cell Biochem Funct 15:243
127. Sui L, Dong YY, Watanabe Y, Yamaguchi F, Hatano N, Tsukamoto I, Izumori K, Tokuda M International (2005) J Oncol 27(4):907
128. Sui L, Dong YY, Watanabe Y, Yamaguchi F, Hatano N, Izumori K, Tokuda M (2005) Anticancer Res 25(4):2639
129. Allen VM, Fernandez F, Hinton MH (1997) Br Poult Sci 38:485
130. Seddiki N, Rabehi L, Benjouad A, Saffar L, Ferriere F, Gluckman JC, Gattegno L (1997) Glycobiology 7:1229
131. Rajan N, Cao Q, Anderson BE, Pruden DI, Sensibar J, Duncan JL, Schaeffer AJ (1999) Infect Immun 67:5027
132. Isenberg J, Stoffel B, Stutzer H, Otte K, Beuth J (1997) Anticancer Res 17:3767
133. Kawakami S, Yamashita F, Nishikawa M, Takakura Y, Hashida M (1998) Biochem Biophys Res Commun 252:78
134. Dawson RB, Hershey RT, Myers CS, Zuck TF (1980) Transfusion 20:110
135. Yu CH, Leng XS, Peng JR, Wei YH, Liu JC, Du RY (1999) Transplant Proc 31:1080
136. Valeri F, Boess F, Wolf A, Goldlin C, Boelsterli UA (1997) Free Radic Biol Med 22:257
137. Ledochowski M, Sperner-Unterweger B, Widner B, Fuchs D (1998) Eur J Med Res 17:295
138. Zamora-Leon SP, Golde DW, Concha II, Rivas CI, Delgado-Lopez F, Baselga J, Nualart F, Vera JC (1996) Proc Natl Acad Sci USA 93: 1847
139. Godoy A, Ulloa V, Rodriguez F, Reinicke K, Yanez AJ, Garcia MD, Medina RA, Carrasco M, Barberis S, Castro T, Martinez F, Koch X, Vera JC, Poblete MT, Figueroa CD, Peruzzo B, Perez F, Nualart F (2006) J Cell Physiol 207(3):614
140. Moadel RM, Weldon RH, Katz EB, Lu P, Mani J, Stahl M, Blaufox MD, Pestell RG, Charron MJ, Dadachova E (2005) Cancer Res 65(3): 698
141. Pauwels EKJ (2001) Drugs of The Future 26(7):659
142. Pauwels EKJ, Sturm EJC, Bombardieri E, Cleton FJ, Stokkel MPM (2000) J Cancer Res Clin Oncol 126(10):549
143. Macheda ML, Rogers S, Best JD (2005) J Cell Physiol 202(3):654
144. Medina RA, Owen GI (2002) Biol Res 35(1):9
145. Nortey SO, Wu WN, Maryanoff BE (1997) Carbohydr Res 304:29
146. Shank RP, Gardocki JF, Streeter AJ, Maryanoff BE (2000) Epilepsia 41:S3 Suppl. 1
147. Paterna JC, Boess F, Staubli A, Boelsterli UA (1998) Toxicol Appl Pharmacol 148:117
148. Levin GV, Zehner LR, Saunders JP, Beadle JR (1995) Am J Clin Nutr 62:1161S
149. Kruger LC, Whittaker MH, Frankos VH (1999) Regul Toxicol Pharmacol 29:S36
150. Saunders JP, Donner TW, Sadler JH, Levin GV, Makris NG (1999) Regul Toxicol Pharmacol 29:S57
151. Moore MC (2006) Current Opin Investig D 7(10):924
152. Manjasetty BA, Chance MR (2006) J Mol Biol 360(2):297
153. Holst O, Zahringer U, Brade H, Zamojski A (1991) Carbohydr Res 215:323
154. Bernlind C, Oscarson S (1997) Carbohydr Res 297:251
155. Kondo S, Watabe T, Haishima Y, Hisatsune K (1993) Carbohydr Res 245:353
156. Ding L, Seto BI, Ahmed SA, Coleman WG (1994) J Biol Chem 269:24384
157. Scruel O, Vanhoutte C, Sener A, Malaisse WJ (1998) Mol Cell Biochem 187:113
158. Kibenge MT, Chan CB (1995) Obes Res 3:171
159. Board M, Colquhoun A, Newsholme EA (1995) Cancer Res 55:3278
160. McClain D, Paterson A, Roos M, Wei X, Kudlow J (1992) Proc Natl Acad Sci USA 89:8150
161. Wang J, Liu R, Hawkins M, Barzilai N, Rossetti L (1998) Nature 393:684
162. Pickering BT (1965) Nature 206:400
163. Yoneyama T, Koike Y, Arakawa H, Yokoyama K, Sasaki Y, Kawamura T, Araki Y, Ito E, Takao S (1982) J Bacteriol 149:15

164. Giorda R, Rudert WA, Vavassori C, Chambers WH, Hiserodt JC, Trucco M (1990) Science 249:1298
165. Ryan JC, Seaman WE (1997) Immunol Rev 155:79
166. Bezouska K, Vlahas G, Horvarth O, Jinochova G, Fiserova A, Giorda R, Chambers WH, Feizi T, Pospisil M (1994) J Biol Chem 269: 16945
167. Bezouska K, Kren V, Kieburg C, Lindhorst TK (1998) FEBS Lett 426:243
168. Bezouska K, Yuen C-T, O'Brien J, Childs RA, Chai W, Lawson AM, Drbal K, Fiserova A, Pospisil M, Feizi T (1994) Nature 372:150
169. Bezouska K, Sklenar J, Dvorakova J, Havlicek V, Pospisil M, Thiem J, Kren V (1997) Biochem Biophys Res Commun 238:149
170. Sedmera P, Prikrylova V, Bezouska K, Rajnochova E, Thiem J, Kren V (1998) J Carbohydr Chem 17:1351
171. Sakata T, Kurokawa M (1992) Am J Clin Nutr 55:272S
172. Elliott DA, Nelson RW, Reusch CE, Feldman EC, Neal LA (1999) J Am Vet Med Assoc 214:1794
173. Breborowicz A, Wieczorowska-Tobis K, Kuzlan M, Kupsz J, Korybalska K, Polubinska A, Krzystof P, French I, Tam P, Wu G (1997) Perit Dial Int 17:S80
174. Schmidt RJ, Spyratou O, Turner TD (1989) J Pharm Pharmacol 41:781
175. Shaunak S, Thomas S, Gianasi E, Godwin A, Jones E, Teo I, Mireskandari K, Luthert P, Duncan R, Patterson S, Khaw P, Brocchini S (2004) Nat Biotechnol 22(8):977
176. Shikhman AR, Kuhn K, Alaaeddine N, Lotz M (2001) J Immunol 166(8):5155
177. Fraser PE, Darabie AA, McLaurin J (2001) J Biol Chem 276(9):6412
178. Fusetani N, Ejima D, Matsunaga S, Hashimoto K, Itagaki K, Akagi Y, Taga N, Suzuki K (1987) Experientia 43:464
179. Wendlinger G, Loessner MJ, Scherer S (1996) Microbiology 142:985
180. (a) Ghannoum MA, Abu-Elteen K, Ibrahim A, Stretton R (1991) Microbios 67:95; (b) Haidar M, Gluckman JC, Gattegno L (1992) Glycobiology 2:429
181. St Arneault G, Walter L, Bekesi JG (1971) Int J Cancer 7:483
182. Onoda T, Morikawa S, Harada T, Suzuki Y, Inoue K, Nishigami K (1982) Cancer Res 42:2867
183. Hiraku Y, Kawanishi S (1999) Free Rad Res 31(5):389
184. Feizi T (1992) Biochem Soc Trans 20:274
185. Feizi T (1985) Nature 314:53
186. Marolda CL, Feldman MF, Valvano MA (1999) Microbiology 145:2485
187. Godfroid F, Taminiau B, Danese I, Denoel I, Tibor A, Weynants V, Cloeckaert A, Godfroid J, Letesson JJ (1998) Infect Immun 66: 5485
188. Rizzi M, Tonetti M, Vigevani P, Sturla L, Bisso A, Flora AD, Bordo D, Bolognesi M (1998) Structure 6:1453
189. Secchi EF, Lizak MJ, Sato S, Kador PF (1999) Curr Eye Res 18:277
190. Murata M, Ohta N, Sakurai S, Alam S, Tsai JY, Kador PF, Sato S (2001) Chem-Biol Interact 130(1-3):617
191. Giuseppina Monti M, Ghiaroni S, Barbieri D, Franchesci C, Marverti G, Moruzzi MS (1999) Biochem Biophys Res Commun 257:460
192. Dwarakanath BS, Adhikari JS, Jain V (1999) Ind J Radiat Oncol Biol Phys 43:1125
193. Khaitan D, Chandna S, Arya MB, Dwarakanath BS (2006) Cancer Biol Ther 5(9):1142
194. Chang LC, Tsai T, Wang JJ, Lin CN, Kuo KW (1998) Biochem Biophys Res Commun 242:21
195. Dwarukanath BS, Khaitan D, Ravindranath T (2004) Cancer Biol Ther 3(9):864
196. Jomaa H, Wiesner J, Sanderbrand S, Altincicek B, Weidemeyer C, Hintz M, Turbachova I, Eberl M, Zeidler J, Lichtenthaler HK, Soldati D, Beck E (1999) Science 285:1573
197. Jomaa H, Feurle J, Luhs K, Kunzmann V, Tony HP, Herderich M, Wilhelm M (1999) FEMS Immunol Med Microbiol 25:371
198. Abbott SD, Gagnon L, Lagraoui M, Kadhim S, Attardo G, Zacharie B, Penney CL (1998) J Med Chem 41:1909
199. Shashikanth MC, Rao BB (1994) Indian J Dent Res 5:119
200. Hiramatsu Y, Moriyama H, Kiyoi T, Tsukida T, Inoue Y, Kondo H (1998) J Med Chem 41: 2302
201. Nobelmann B, Lengeler JW (1996) J Bacteriol 178:6790
202. Kappler F, Su B, Szwergold BS, Randall WC, Brown TR (1995) Metabolism 44:1527
203. Bruynseels K, Bergans N, Gillis N, van Dorpen F, van Hecke P, Stalmans W, Vanstapel F (1999) NMR Biomed 12:145
204. Silva SS, Felipe MG, Mancilha IM (1998) Appl Biochem Biotechnol 70:331

205. Gaffar A, Blake-Haskins JC, Sullivan R, Simone A, Schmidt R, Saunders F (1998) Int Dent J 48:22
206. Tapiainen T, Kontiokari T, Sammalkivi L, Ikaheimo I, Koskela M, Uhari M (2001) Antimicrob Agents Ch 45(1):166
207. Kontiokari T, Uhari M, Koskela M (1998) J Antimicrob Chemother 41:563
208. Christensson B, Wiebe T, Pehrson C, Larsson L (1997) J Clin Microbiol 35:636
209. Santini V, Scappini B, Gozzini A, Grossi A, Villa P, Ronco G, Douillet O, Pouillart P, Bernabei PA, Rossi Ferrini P (1998) Br J Haematol 101: 529
210. Schauer R (2004) Zoology 107(1):49
211. Havellar AC, Mancini GM, Beerens CE, Souren RM, Verheijen FW (1998) J Biol Chem 273:34568
212. Sonnenburg JL, van Halbeek H, Varki A (2002) J Biol Chem 277(20):17502
213. Furuhata K (2004) Trends Glycosci Glyc 16(89):143
214. (a) Boons G-J (2000) Chem Rev 100:4539 (b) Ress DK, Linhardt RJ (2004) Curr Org Chem 1:31
215. Okamoto K, Goto T (1990) Tetrahedron 46, 17, 5835
216. Burmeister WP, Ruigrok RWH, Cusack S (1992) EMBO J 11:49
217. Keppler OT, Herrmann M, von der Lieth CW, Stehling P, Reutter H, Pawlita M (1998) Biochem Biophys Res Commun 253:437
218. de Groot RJ (2006) Glycoconjugate J 23:(1–2):59
219. Ngundi MM, Taitt CR, McMurry SA, Kahne D, Ligler FS (2006) Biosens Bioelectron 21(7):1195
220. Yuen C-T, Bezouska K, O'Brien J, Stoll M, Lemoine R, Lubineau A, Kiso M, Hasegawa A, Bockovich NJ, Nicolaou KC, Feizi T (1994) J Biol Chem 269:1595
221. Torgersen D, Mullin NP, Drickamer K (1998) J Biol Chem 273:6254
222. May AP, Robinson RC, Vinson M, Crocker PR, Jones EY (1998) Mol Cell 1:719
223. Wheeler GL, Jones MA, Smirnoff N (1998) Nature 393:365
224. Rao VR, Krishnamoorty L, Kumaraswamy SV, Ramaswamy G (1998) Cancer Detect Prev 22:237
225. Driguez PA, Barrere B, Quash G, Doutheau A (1994) Carbohydr Res 262:297
226. Lehmann F, Tiralongo E, Tiralongo J (2006) Cell Mol Life Sci 63(12):1331
227. Strehle EM (2003) Genetic Test 7(2):113
228. Aula N, Jalanko A, Aula P, Peltonen L (2002) Mol Genet Metab 77(1–2):99
229. Morin P, Sagne C, Gasnier B (2004) EMBO J 23(23):4560
230. Jones MB, Teng H, Rhee JK, Lahar N, Baskaran G, Yarema KJ (2004) Biotech Bioeng 85(4): 394
231. Yokoyama T, Sasaki H, Giblin FJ, Reddy VN (1994) Exp Eye Res 58:207
232. Campbell JD, Cole M, Bunditrutavorn B, Vella AT (1999) Cell Immunol 194:1
233. Levine M, Rumsey SC, Daruwala R, Park JB, Wang Y (1999) JAMA 281:1415
234. Sohar I, Sleat D, Gong Liu C, Ludwig T, Lobel P (1998) Biochem J 330:903
235. Roberts DL, Weix DJ, Dahms NM, Kim JJ (1998) Cell 93:639
236. Solis D, Romero A, Jimenez M, Diaz-Maurino T, Calvete JJ (1998) FEBS Lett 431:273
237. Bezouska K, Sklenar J, Novak P, Halada P, Havlicek V, Kraus M, Ticha M, Jonakova V (1999) Protein Sci 8:1551
238. Bowman KG, Bertozzi CR (1999) Chem Biol 6:R9
239. Takeuchi K, Cao-Danh H, Friehs I, Glynn P, D'Agostino D, Simplaceaunu F, McGowan FX, del Nido PJ (1998) J Thorac Cardiovasc Surg 116:335
240. Oates AJ, Schumaker LM, Jenkins SB, Pearce AA, DaCosta SA, Arun B, Ellis MJ (1998) Breast Cancer Res Treat 47:269
241. Hawkes C, Kar S (2004) Brain Res Review 44(2–3):117
242. Pospisil M, Vannucci L, Horvath O, Fiserova A, Krausova K, Bezouska K, Mosca F (2000) Int J Oncol 16:267
243. Malaisse WJ, Kadiata MM (1998) Int J Mol Med 2:331
244. Reinhold U, Malaisse WJ (1998) Int J Mol Med 1:427
245. Reinhold U, Ugurel S, Tilgen W, Kadiata MM, Olivares E, Nadi AB, Malaisse WJ (2000) Oncol Rep 7(5):1093
246. Feizi T, Fazio F, Chai WC, Wong CH (2003) Curr Opin Struct Biol 13(5):637
247. Houseman BT, Mrksich M (2002) Chem Biol 9(4):443
248. Smith EA, Thomas WD, Kiessling LL, Corn RM (2003) J Am Chem Soc 125(20):6140

249. Polaczek E, Tomasik PJ, Mazurkiewicz J, Wrzalik R, Stobinski L, Tomasik P, Lin HM (2005) J Nanosci Nanotechnol 5(3):479
250. Ishi-i T, Mateos-Timoneda MA, Timmerman P, Crego-Calama M, Reinhoudt DN, Shinkai S (2003) Angew Chem Int Ed 42(20):2300
251. Jensen KJ, Brask J (2005) Biopolymers 80(6):747
252. Le GT, Abbenante G, Becker B, Grathwohl M, Halliday J, Tometzki G, Zuegg J, Meutermans W (2003) Drug Discov Today 8(15):701
253. Cipolla L, Peri F, La Ferla B, Redaelli C, Nicotra F (2005) Curr Org Synth 2(2):153
254. Opatz T, Kallus C, Wunberg T, Schmidt W, Henke S, Kunz H (2002) Carbohydr Res 337(21–23):2089

12.3 Structure and Function of Mammalian Carbohydrate-Lectin Interactions

Kevin Anderson, *David Evers*, *Kevin G. Rice*
Division of Medicinal and Natural Products Chemistry,
College of Pharmacy, University of Iowa,
115 South Grand Avenue, Iowa City, IA 52242-1112, USA
kevin-rice@uiowa.edu

In: *Glycoscience*. Fraser-Reid B, Tatsuta K, Thiem J (eds)
Chapter-DOI 10-1007/978-3-540-30429-6_63:

Abstract

Over the past three decades the field of glycobiology has expanded beyond a basic understanding of the structure and biosynthesis of glycoprotein, proteoglycans, and glycolipids toward a more detailed picture of how these molecules afford communication through binding to mammalian lectins. Although the number of different mammalian lectin domains appears to be finite and even much smaller than early estimates predicated based on the diversity of glycan structures, nature appears capable of using these in numerous combinations to fine tune specificity. The following provides an overview of the major classes of mammalian lectins and discusses their glycan binding specificity. The review provides a snapshot of the field of glycobiology that continues to grow providing an increasing number of examples of biological processes that rely upon glycan-lectin binding.

Keywords

Lectin; Mannose binding protein; Selectin; Asialoglycoprotein receptor; Macrophage mannose receptor; Mannose-6-phosphate receptor; Galectin; Pentraxin; Siglec; Oligomerization

Abbreviations

CTLDs	C-type lectin-like domains
CRD	C-type carbohydrate recognition domains
MBP-A	mannose-binding protein-A
MBP	mannose-binding protein
MDCC	MBP-dependent cell-mediated cytotoxicity
TNF-α	tumor necrosis factor-α
IL-1	interleukin-1
LAD-II	leukocyte adhesion deficiency type II
EGF	epidermal growth factor-like
CR	consensus repeat
HUVECs	human umbilical cord vein endothelial cells
ASGPR	asialoglycoprotein receptor
HL	hepatic lectin
RTG-r	rat testis Gal receptor
CHL	chicken hepatic lectin
AHL	alligator hepatic lectin
MMR	macrophage mannose receptor
LH	lutropin
DC-SIGN	dentritic cell specific ICAM-3 grabbing non-integrin
ICAM-3	integrin cell adhesion molecule 3
M6P	Man-6-phosphate
IGF-II/CI-MPR	insulin-like growth factor II/cation-independent M6P receptor
CLC	Charcot–Leydon crystal protein
HL	Hodgkin's lymphoma
CRP	C reactive protein

SAP	serum amyloid protein
PE	phosphoethanolamine
PC	phosphocholine
Ig	immunoglobulin
SIGLECS	sialic acid binding **i**mmuno**g**lobulin-like lectins
MAG	myelin-associated glycoprotein

1 Introduction

Lectins are proteins which bind to carbohydrate and must be neither of immune origin (e. g. not antibodies) nor catalytic (e. g. not enzymes) [1]. Carbohydrate recognition systems have been described as a functional triad of lectin, carrier (protein, peptide, or lipid), and carbohydrate [2]. Glycorecognition systems are important in the mediation of inflammation and immune response, protein trafficking, developmental regulation, cell adhesion and migration, cancer, nerve communication, cytotoxicity, fertilization, and the regulation of serum half-life of proteins and hormones, among other proposed or documented functions [2,3]. A number of unrelated microorganisms utilize lectin-carrier-carbohydrate interactions to infect hosts and prolong their survival [3]. Lectins are often found in the seeds of legume plants where they exhibit a wide variety of carbohydrate specificities, often for saccharides which are not native to the legume where they are found [4,5]. Legume lectins are nearly always oligomeric and are useful reagents for the agglutination of cells and precipitation of multivalent carbohydrates. Although there is little known regarding the native functions of legume lectins, they have proved themselves valuable as *in vitro* tools for immunology and glycobiology [4].

While many of the subtleties of glycorecognition have not been fully elucidated, two admittedly unique examples of the functional roles for carbohydrates may serve to convince the reader of their general importance in biology. First, ABO and Lewis blood groups are carbohydrate antigens present on *N*- and *O*-linked oligosaccharides and glycolipids. The fate of the recipient of a human blood transfusion who has not been adequately crossmatched for blood typing highlights the biological importance of carbohydrate recognition [6]. Secondly, ricin is a 60 kDa heterodimeric protein composed of catalytic (A subunit) and Gal-specific lectin (B subunit) domains and has been described as the most toxic substance known to man [5]. Cell adhesion is mediated by sugar binding to the B subunit, followed by endocytosis of the A and B subunits during non-specific homeostatic receptor recycling. In the reducing environment of the cell, the subunits dissociate and the A subunit is freed to catalyze depurination of the 28S ribosomal RNA subunit [5].

In the following, we present the structure and function of carbohydrate-lectin interactions with particular emphasis on the five known families of animal lectins [7,8]. These groups of structurally (and often functionally) distinct lectins can be divided into *C*-type (calcium-requiring) lectins, P-type (Man-6-phosphate binding) lectins, *S*-type (galectins) lectins, pentraxins, and I-type (immunoglobulin-like) lectins. Within these groups of lectins a wide (e. g. within *C*-type lectins) or a very narrow range of carbohydrate specificities (e. g. within galectins) have been observed. A number of recent X-ray crystal structures of lectins bound to ligand (see discussion below) have served to further our understanding of how binding specificity is achieved.

The number of identical or different subunits which oligomerize to constitute a biologically relevant receptor is considered in detail. Additionally, we present the sometimes well-documented but often speculative biological roles for the members of these lectin families.

2 C-Type Lectins

2.1 C-Type Lectin Subgroups

Of the 65 human proteins identified as containing *C*-type lectin-like domains (CTLDs), approximately half function as lectins. The *C*-type carbohydrate recognition domains (CRD) can be classified broadly into two groups, based on the relative orientation of the 3- and 4-hydroxyl groups of monosaccharides which interact with the primary ligand-binding site, with preferred ligands possessing equatorial geometry at these positions [9]. Approximately two-thirds of lectins are Man-type; containing CRD with affinity for mannose and *N*-acetylglucosamine, the others are referred to as Gal-type, and bind galactose and *N*-acetylgalactosamine [10]. The CRDs are characterized by a common overall fold and sequence homology, including 14 invariant and 23 conserved residues along their 115 to 130 amino acid length [11]. While the CRDs of *C*-type lectins are conserved, the overall topology and organization of the non-CRD regions of these lectins are diverse as reviewed elsewhere [11,12,13,14]. *C*-type lectins in group I are monomeric proteoglycans with domains flanking the *C*-terminal CRD (*Fig. 1*). Group II (type II receptors, indicating that they are integral membrane proteins with an intracellular *N*-terminus and an extracellular *C*-terminus), III (soluble, e. g. not membrane-associated proteins, collectins), and V (type II lymphocyte antigens) *C*-type lectins have trimeric structures (see discussion below) by virtue of the association of non-lectin *N*-terminal regions [15]. Group III collectins can form higher-ordered structures, such as pulmonary surfactant protein which forms 12-mer cruciform structures, as evidenced by electron microscopy, where the *N*-terminal regions of four trimers associate into a macromolecular cross [16]. It is unclear whether functional group IV selectin receptors (see discussion below) oligomerize.

Group VI consists of receptors with multiple lectin domains in a single polypeptide and is composed of the macrophage mannose, phospholipase A_2, DEC 205, and S4GGnM receptors [18]. Consistent with the observation that group VI lectins serve to endocytose ligands, they are referred to as the endocytic *C*-type lectins, although it may be noted that a number of group II *C*-type lectins, such as asialoglycoprotein receptors, are also endocytic. Group VII *C*-type lectins are the simplest structures consisting of a CRD without flanking domains [11]. Thirteen human proteins containing CTLDs have been identified by genomic screens and have been classified into seven additional groups [10].

2.2 Mannose Binding Protein

The conserved sequences and available X-ray crystal structures of *C*-type lectin CRDs correlate with a common structural motif identified in rat serum mannose-binding protein-A (MBP-A) as represented in *Fig. 2* [19,20]. There are two α-helices, 6 β-strands, and 2 disulfide bonds common to this motif. Divalent cation binding sites are located in the loop regions.

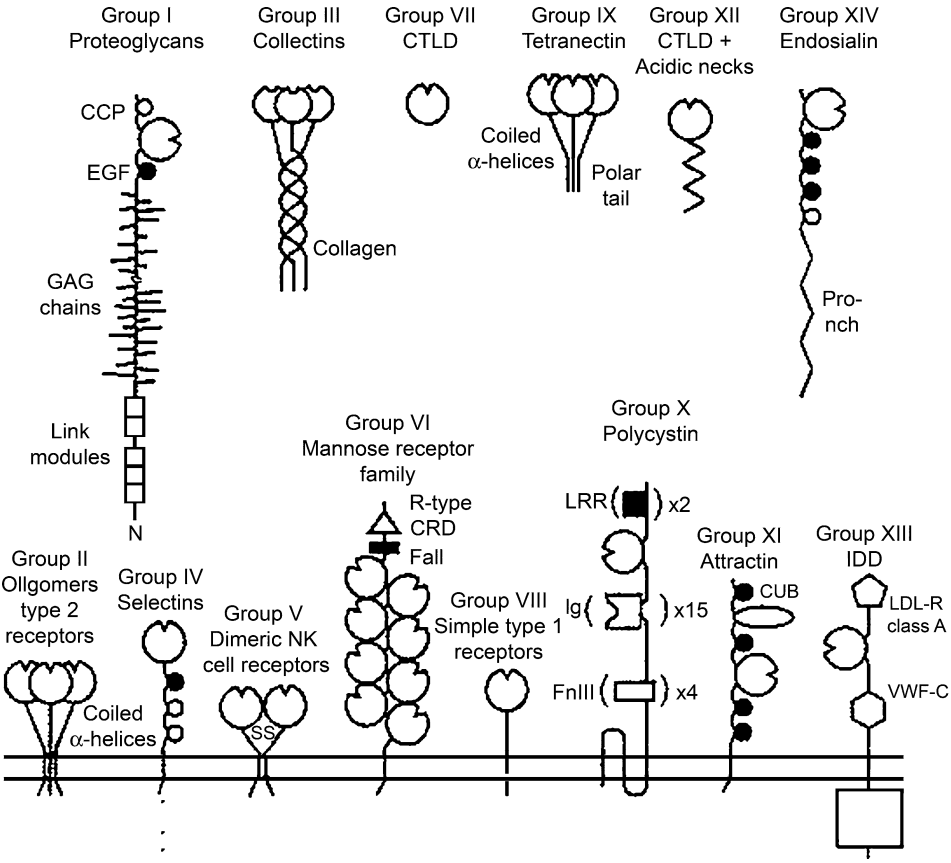


Figure 1

Domain organization of proteins containing CTLDs. CCP, complement control protein; CUB, complement components Clr/Cls, Uegf and bone morphogenetic protein-I; EGF, epidermal growth factor; Fn, fibronectin; LDL-R, low-density lipoprotein receptor; LLR, leucine-rich repeat; SS, disulfide bond; VWF-C, von Willebrand factor. Groups I, III, VII, IX, XII, and XIV are found in solution, while groups II, IV–VI, VII, X–XI, and XIII are membrane-associated. (Reproduced with permission from [17])

Calcium one is considered to be involved in maintaining lectin structure while calcium two coordinates directly with carbohydrate, and calcium three is believed to be an artifact of crystallization [20].

Chemical cross-linking experiments have determined, and X-ray crystallography has confirmed, that the oligomeric arrangement of MBP-A is a homotrimer with three axes of symmetry held together by the *N*-terminal regions interacting in a triple coiled-coil of α-helices [15]. The binding sites are 53 Å apart, significantly farther apart than that determined for the rat hepatic lectin (see discussion below). This indicates that a trimer of MBPs is unlikely to show cooperativity between subunits to interact with high avidity to a single high-mannose oligosaccharide, while foreign carbohydrate polymers such as yeast mannan would be expected to bind

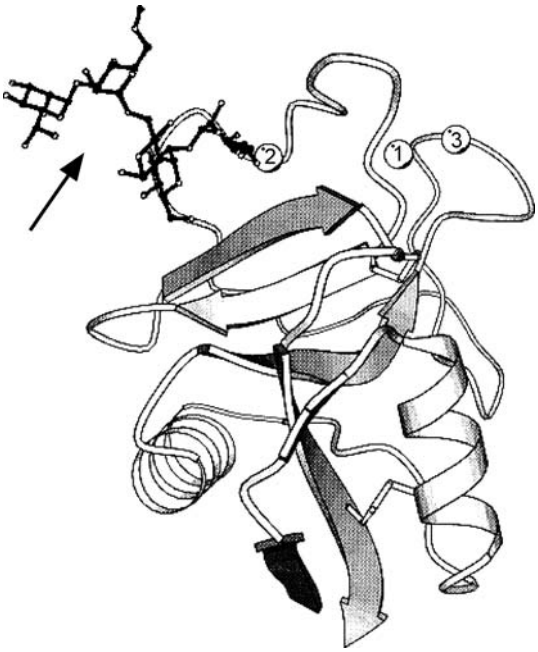


Figure 2
The structure of mannose binding protein-A. Bound $Man_6GlcNAc_2Asn$ asparaginyl-oligosaccharide is indicated by top left *arrow*. *Numbered circles* represent calcium ions. See text for details. (Reproduced with permission from [20])

more tightly by interacting with all three subunits of a MBP-trimer [15]. The three-dimensional X-ray crystal structure of MBP-C [21] demonstrates a topology nearly identical to that of MBP-A.

The mode of specific recognition of different sugars by highly homologous *C*-type lectin CRDs have been advanced by Drickamer and colleagues. It may be noted that while MBPs recognize Man, other hexoses with vicinal, equatorial hydroxyls at positions 3 and 4 have been shown to specifically bind these lectins, suggesting that MBPs may bind to a variety of saccharides [22]. Ligand discrimination between GlcNAc, Gal/GalNAc, Man, and Fuc residues have been delineated to the orientation of similar binding site residues which interact with or exclude individual hexoses based on equatorial and/or axial substituents [23]. Subtle mutations of the binding site residues of MBP-A have been shown to alter its binding specificity from Man to mimic the binding properties of asialoglycoprotein receptors, that bind to Gal [24,25], GalNAc [26,27], and GlcNAc [28]. MBP has also been altered to possess E-selectin-like specificity for the Fuc-containing sialyl $Lewis^x$ tetrasaccharide [29,30,31].

MBPs have been proposed to act as a primitive complement-dependent arm of the immune system [32]. A functional role has been shown in the clearance of yeast [33], complement fixation of bacteria [34], and the blocking of viral infectivity (human immunodeficiency and influenza viruses) [35,36]. Two serine proteases (MASP-1 and MASP-2) have been shown to be associated with the MBP receptor and are required for complement activation by MBP [37]. A recent

study suggests that there is an additional complement-independent pathway mediating MBP cytotoxicity [38]. MBP-dependent cell-mediated cytotoxicity (MDCC) has demonstrated anti-cancer activity in mice injected with human colorectal carcinoma cells [38]. Patients with a deficiency in the MBP gene demonstrate a life-long increased susceptibility to infection [39]. This is evident for cystic fibrosis patients in which MBL-deficiency influence the severity of chronic infection by *Pseudomonas aeruginosa* and *Burkholderia cepacia* [40]. An affinity purified form of MBL has been developed and used in a phase 1 clinical to treat patients with opsinization defect [41].

2.3 Selectins

The selectins are a family of lectins that mediate the movement of leukocytes from the bloodstream into sites of inflammation. This multistep cascade proceeds by initial attachment leading to the rolling of leukocytes along endothelial vasculature via selectin-carbohydrate interactions. Subsequent firm adhesion requires chemoattractant activation of leukocyte β_2 integrins which mediate arrest through interaction with endothelial receptors. Leukocytes then extravasate between endothelial cells in the direction of chemoattractants as indicated in *Fig. 3* [42].

There are three selectin molecules: E (endothelial), P (platelet), and L (leukocyte), (reviewed in [43]) named for their cellular localization. P-selectin is released to the cell surface from storage vesicles in endothelial cells and platelets minutes after stimulation by a number of activators, such as thrombin or histamine [42]. E-selectin expression on the endothelial cell surface

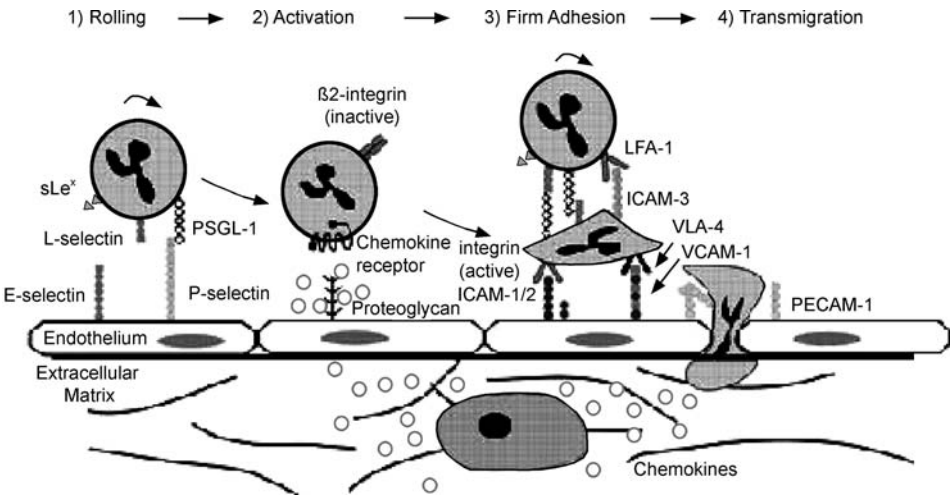


Figure 3

Leukocyte rolling and adhesion. The migration of leukocytes to sites of inflammation begins with attachment (1a.) followed by rolling (1b.), activation (2), cessation of movement and flattening against the vessel wall (3), and trans-endothelial migration (4). Steps 1a–2 are mediated by selectins, 1b–4 by chemoattractants, and 3–4 by β_2 integrins. (Adapted with permission from [42])

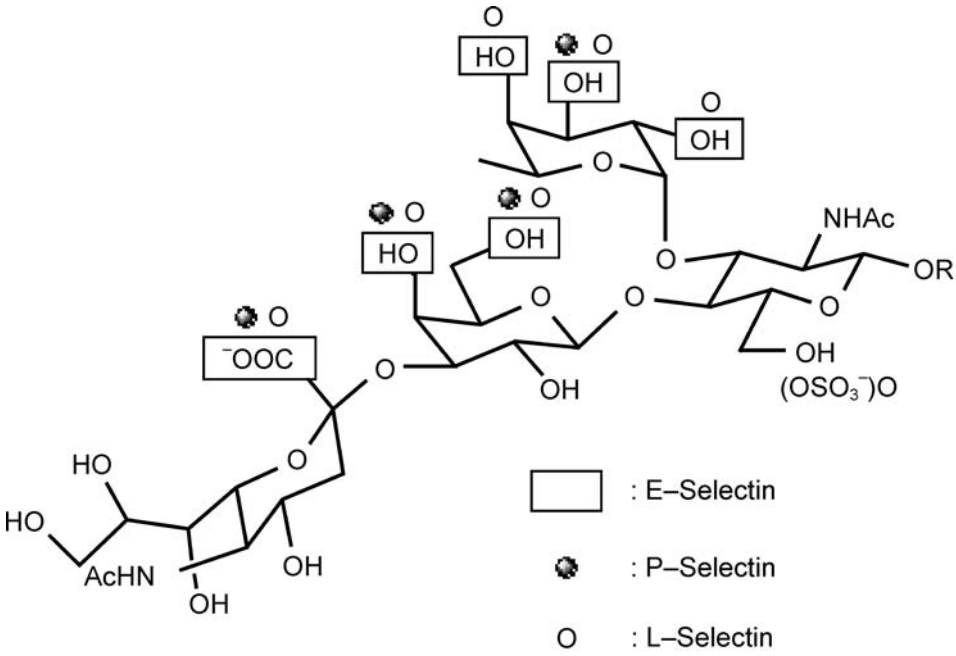


Figure 4
Selectins bind to different functional groups of the sialyl Lewis^x tetrasaccharide. Functional groups considered to be important for binding to each of the selectins are indicated. (Reprinted with permission from [49])

requires de novo mRNA transcription, peaking 3–6 hours after stimulation with activators such as tumor necrosis factor-α (TNF-α), interleukin-1 (IL-1), or bacterial lipopolysaccharide [44]. These cytokines also upregulate the transcription of P-selectin, resulting in increased expression 2–4 hours following stimulation [42].

All three selectins bind sialyl Lewis^x (SLe^x, NeuAcα2-3Galβ1-4[Fucα1-3]GlcNAc) which is proposed to be the minimal carbohydrate ligand for all three selectins, although sialyl Lewis^a (NeuAcα2-3Galβ1-3[Fucα1-4]GlcNAc), 3′SO_4-Le^a and 3′SO_4-Le^x bind with affinities similar to SLe^x [45,46,47,48].

Pharmacophore studies utilizing ligands with addition or deletion of hydroxyl groups and modifications to sialic acids have produced a foundation for understanding the nature of the interactions between SLe^x and each selectin as shown in *Fig. 4* [49]. Selectins do not tolerate the deletion of Fuc hydroxyls, while the only moiety of NeuAc necessary for binding is the *C*-1 carboxylate. Gal C-4 and C-6 hydroxyls are important for binding, while the remainder of Galβ1-4GlcNAc is considered to function primarily as a scaffold on which Fuc and NeuAc (or sulfate) are positioned for receptor binding [49].

A variety of pathological consequences are the result of malfunctions related to the selectins. Patients with a rare disease known as leukocyte adhesion deficiency type II (LAD-II, whereas LAD-I is a defect in the common β_2 integrin subunit) lack the appropriate de novo fucosylation pathway to biosynthesize SLe^x, and consequently suffer from repeated severe bacterial

infections in their early youth [50,51]. Selectin "knockout" mice have been genetically engineered and demonstrate abnormal leukocyte rolling, as well as an increased susceptibility to infection [52]. It has also been proposed that the selectins mediate tumor metastases via tumor cell expression of SLe^x [48].

The SLe^x epitope is expressed on inflammation-induced acute phase proteins which serve as endogenous competitors for selectins and mediate a dampening of the immune response, returning inflammation to homeostasis [53]. Inhibitors of the selectins have been proposed to be useful therapies for treating inflammatory disorders including respiratory distress [54], hypersensitivity responses [55], and surgically induced myocardial ischemic reperfusion injury [56]. Animal studies have shown that the inhibition of selectin-carbohydrate interactions can alleviate these inflammatory responses.

E-, P-, and L-selectin each consist of an *N*-terminal *C*-type lectin carbohydrate recognition domain (CRD), an epidermal growth factor-like (EGF) subunit, a number of short consensus repeat (CR) units, a membrane spanning region, and a *C*-terminal cytoplasmic tail. There is approximately 72% homology for analogous selectin CRDs across species, and ~ 52% homology between different selectins within a species [43]. Notably, the selectins are unique among *C*-type lectins as they have only one calcium binding site [57]. While the 3-dimensional structure of E-selectin CRD and EGF domains without ligand bound has been elucidated by X-ray crystallography [57], there is not yet an unambiguous understanding of the binding interactions between selectins and their ligands. The interactions of a mutant MBP with its binding specificity subtly altered to mimic E-selectin (as described above) have been elucidated by X-ray crystallography as shown in ❯ *Fig. 5* [58].

As expected, this structure shows that calcium complexes directly to Fuc hydroxyls rather than serving a purely structural role. Gal C-4 and C-6 hydroxyls are hydrogen bonded to lectin side chains, while the acetyl and backbone groups of GlcNAc participate in hydrophobic bonding. Interestingly, this structure fails to demonstrate interactions between the lectin and the C-1 carboxylate of NeuAc [58]. Lys 213 [57,59] (in MBP numbering), or alternatively Arg 97 [60] have been independently identified as likely electrostatic binding partners with the negatively charged carboxylate. The NMR structure of SLe^x bound to E-selectin suggests a binding pocket where Lys 111 is positioned to hydrogen bond the C-7 hydroxyl of NeuAc, and Lys 113 is an electrostatic binding partner for the C-1 carboxylic acid of NeuAc [61]. E-selectin mutants Lys113Ala (residue 213 in MBP numbering) and Arg97Ala demonstrate a complete loss of affinity for SLe^x [57,59]. Conservative mutation of these residues (e. g. Lys113Arg [60] and Arg97Lys [57]) failed to restore SLe^x binding activity, while mutant Lys113Glu bound as well as wild-type [60].

Although NeuAc (or an appropriately spaced negatively charged carboxylate isostere) is required for binding to the selectins [62], the model of an electrostatic interaction between the C-1 carboxylate and binding site lysine(s) has not been confirmed. A short sequence of lysine and arginine residues shown to be conserved in six bacterial sialic acid binding lectins [*H. pylori* hpaA, *E. coli* fimbral adhesion (SfaS), *E. coli* K99 fibrillar adhesion protein (K99), *E. coli* colonization factor antigen I (CFA/1), *V. cholerae* heat-labile toxin B subunit (CT-B), and heat-labile *E. coli* enterotoxin (LTI-B)] was proposed to comprise a sialic acid binding consensus sequence, and the positive charge of these residues in the binding pocket prompted the theory that they were electrostatic binding partners for the C-1 carboxylate of sialic acids [63]. The "basic consensus sequence" residues of the K99 and SfaS lectins were

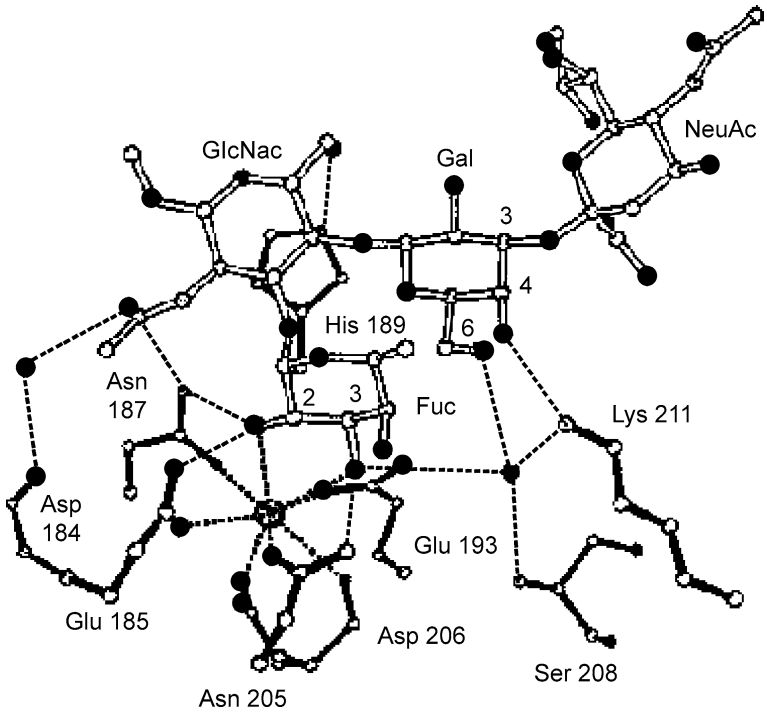


Figure 5
X-ray crystal structure of the binding site of a mutant mannose binding protein with specificity for sialyl Lewis X with ligand bound. See text for details. (Reprinted with permission from [30])

mutated to neutral residues, and resulted in a loss of binding [64,65]. It has been shown that selective chemical modification of lysine and arginine residues of CT-B blocks binding to sialic acids [66]. It was not until the X-ray crystal structures of the CT-B and LTI-B bound to ligands ganglioside G_{M1} and lactose [67,68] were published that this theory was questioned. The conserved lysine and arginine residues are not part of the binding site, and their roles are structural rather than important for binding.

There is evidence that biologically relevant selectin receptors are multimeric. Gel permeation chromatography of P-selectin isolated from human platelets showed an apparent molecular weight of 2000 kDa which was shifted to an apparent molecular weight of 470 kDa (the apparent molecular weight of P-selectin by SDS-PAGE is 120 kDa) when eluted with detergent [69]. Analytical ultracentrifugation of P-selectin has indicated that it forms higher ordered species from dimers to hexamers [70]. Likewise, E-selectin demonstrates the formation of high molecular weight species by size-exclusion chromatography [71].

Another line of evidence for multimeric organization of selectins comes from increased binding of multivalent ligands. Selectin-ligand interactions are noted to be weak (K_D = mM) consistent with the interactions of many lectins with monovalent ligands [72]. Bivalent SLe^x ligands show a five-fold increase over monovalent ligands in blocking neutrophil binding to immo-

bilized E-selectin [73,74]. Bovine serum albumin (BSA) bearing 7, 11, or 16 SLe^x moieties displayed a 20, 900, or 1000-fold increase in avidity, respectively, in an assay measuring leukocyte binding to immobilized E-selectin [75]. Monovalent, divalent, and tetravalent SLe^x moieties on linear and branched carbohydrate scaffolds inhibited L-selectin dependent adhesion of lymphocytes with potencies increasing in order of valency to an IC_{50} of 1 nM for tetravalent structures [76,77]. Liposomes bearing SLe^x have shown enhanced inhibition of human umbilical cord vein endothelial cells (HUVECs) binding to immobilized E-selectin by seven orders of magnitude compared to monovalent ligand [78]. One explanation for the increased potency of multivalent ligands is that a cluster of appropriately spaced ligands is in a unique position to interact with the multiple binding sites of a cluster of lectins (the "multivalency effect" as discussed below). Another recently offered explanation suggests that the apparent avidity of lectin-carbohydrate interactions depends upon whether the lectin or the carbohydrate is immobilized [79]. When the carbohydrate is immobilized, the lectin binding as measured by surface plasmon resonance is five orders of magnitude greater than when the lectin is immobilized. It has been proposed that the "multivalency effect" of lectin-carbohydrate interactions often attributed to lectin oligomerization can sometimes be attributed to an artifact of a slower k_{off} for ligands which are either larger than monosaccharides or bear multiple ligands [79].

Although SLe^x, SLe^a or their sulfated derivatives are widely considered to be "the ligands" for the selectins, appropriate sialylation/sulfation and fucosylation for biologically relevant binding to selectins could potentially occur on any number of *N*- or *O*-linked oligosaccharides, carbohydrate polymers, or glycolipids. A number of candidate selectin ligands have been identified (reviewed in [43,80,81]). The best characterized ligand for P-selectin is dimeric PSGL-1, presenting SLe^x–type determinants on the non-reducing end of *O*-linked oligosaccharides [82,83,84]. Three candidate ligands for L-selectin have been isolated. They are GlyCAM-1, CD34, and MAdCAM-1 presenting SLe^x on the non-reducing end of *O*-linked oligosaccharides [85,86]. Ligands for E-selectin have been advanced as either glycolipids [87,88] or the glycoprotein ESL-1, the latter presenting relevant carbohydrate determinants on *N*-linked oligosaccharides [89,90,91]. Dissociation constants for the binding of selectins to multivalent ligands and potential ligand-carriers are typically in the range of nM [70,92,93,94], while binding to monovalent sialylated or sulfated Lewis antigens are in the range of mM [95].

2.4 Asialoglycoprotein Receptors

The nomenclature of "asialoglycoprotein receptor" was first used to describe the apparent function of a mammalian liver cell surface lectin. This lectin was shown to bind and endocytose glycoproteins by virtue of the non-reducing terminal Gal residues present on oligosaccharides exposed by non-specific or enzymatic removal of sialic acid residues, and was consistent with the rapid clearance of serum glycoproteins following desialylation [96]. It has become apparent that a number of endocytic receptors rapidly purge appropriately glycosylated proteins of differing carbohydrate specificities from serum [97].

The canonical asialoglycoprotein receptor (ASGPR) is the hepatic lectin (HL) common to mammals, first reported in the rat in 1974 by Ashwell and co-workers [98] {reviewed in [99,100,101]}. Rat HL is a complex of two homologous *C*-type lectins (46 kDa HL-1 and

50 kDa HL-2/3) specific for Gal and GalNAc [102]. The "multivalency effect" encompasses the idea that the binding of appropriately spaced single ligands on a common backbone may achieve an avidity equal to the sum of individual affinities limited only by the loss of translational and rotational entropy intrinsic to the spatial limitations of the backbone. In perhaps the clearest demonstration of this effect, it has been determined that the binding of rat HL is inhibited by monovalent Gal (IC_{50} 1 mM), biantennary non-reducing *N*-linked oligosaccharide terminating in Gal (IC_{50} 1 μM), triantennary Gal-terminating oligosaccharide (IC_{50} 10 nM for one isoform, the others being less potent), and tetraantennary Gal-terminating oligosaccharide (IC_{50} 1 nM). This indicated that a trimeric lectin binds to a trimeric oligosaccharide with high affinity [103,104]. It has been estimated that the distance between binding sites is from 17–43 Å [105]. Ligand photo-affinity labeling of the subunits of the rat HL has defined the geometry of subunits as a trimeric receptor composed of one HL-2/3 subunit and two HL-1 subunits [106]. Capitalizing on the high avidity binding of triantennary Gal/GalNAc-terminating oligosaccharides for endocytic HL, agents have been prepared as potential carriers of therapeutic DNA for receptor-mediated gene delivery to parenchymal hepatocytes [107,108,109]. Although a biological role has been established for HL, transgenic mice which lack the minor subunit of HL (HL-2 "knockout mice") demonstrate only a mildly abnormal phenotype, and while they fail to clear appropriately glycosylated glycoproteins, there is no serum accumulation of these glycoproteins [110,111].

A 32–34 kDa Gal/GalNAc specific *C*-type lectin (M-ASGP-BP) has also been detected in mammalian macrophages [112]. It is 59% identical to RHL-1 and 45% identical to RHL-2/3 at the amino acid level, with a shorter cytoplasmic tail and an insert between the membrane spanning region and CRD. Enhanced avidity of clustered saccharide ligands suggests that in contrast to HL recognition of a distinct triantennary oligosaccharide, oligomeric M-ASGP-BP binds to triantennary oligosaccharides with potencies equal to a random cluster of Gal residues [113]. Additionally, while HL prefers ligands terminating in GalNAc, M-ASGP-BP is selective for Gal [113].

Fuc/GlcNAc specific receptors have been identified in mammalian tissues. An endocytic *C*-type lectin recognizing Fucα1-3Galβ1-4GlcNAc found in Kupffer cells (macrophages specific to the liver), is present as two subunits of 77 and 88 kDa [114,115,116,117,118]. A 180 kDa Fuc/GlcNAc specific lectin has also been localized to Kupffer cells and rat alveolar macrophages [119,120].

A Gal/GalNAc specific *C*-type lectin which is antigenically similar to RHL-2/3 has been detected in mammalian testes. The rat testis Gal receptor (RTG-r) is expressed throughout testicular development, and in spermatids and sperm it is restricted to the plasma membrane overlaying the dorsal portion of the head. Rather than functioning in the clearance of serum glycoproteins, this lectin has been proposed to be important in spermatogenesis and events leading to sperm-egg recognition [121].

The avian homolog of the mammalian liver asialoglycoprotein receptor has been reported. Isolated as a 26 kDa glycoprotein, the chicken hepatic lectin (CHL) exhibits divalent cation-dependent lectin activity requiring a pH greater than 6.5 [122]. It has been shown to mediate endocytosis of glycoproteins bearing non-reducing terminal GlcNAc residues [123]. Avian glycoproteins possess Gal-terminating oligosaccharides, which could be trimmed to bear GlcNAc residues for clearance by CHL [124,125]. Chemical cross-linking of CHL has demonstrated oligomeric species ranging from monomer to hexamer [126]. However, radiation inac-

tivation and sedimentation equilibrium analysis has shown a distribution of oligomeric species, suggesting that while diffuse higher ordered species may be present, the functional state of the receptor is likely a trimer [127].

The major liver lectin of the alligator has been isolated and characterized as a calcium-dependent D-Man/L-Fuc specific lectin, designated as the alligator hepatic lectin (AHL). AHL is present as two bands of equal intensity at 21 and 23 kDa by SDS-PAGE. Binding to ligand is inhibited below pH 6.5, consistent with dissociation in acidic endosomal compartments [128]. Chemical cross-linking of AHL showed dimeric, trimeric, and diffuse higher ordered species of lectin, with the formation of oligomers being favored in the presence of multimeric ligand. This result was consistent with size-exclusion chromatography which indicated species present as monomers, dimers, and trimers. Unfortunately, this lectin has yet to be sequenced or cloned making its designation as a *C*-type lectin unclear. It has been speculated that a hierarchy of hepatic lectins in different organisms from mammals (recognizing Gal/GalNAc) to birds (recognizing GlcNAc) to reptiles (recognizing Man) may coincide with the terminal monosaccharides present on complex-type *N*-glycans in these classes of vertebrates, and function in their clearance from serum [128].

2.5 Endocytic C-Type Lectins

The macrophage mannose receptor (MMR) is a type I membrane-spanning protein with an intracellular *C*-terminus and an extracellular *N*-terminus possessing eight extracellular *C*-type lectin domains (CRD 1–8). CRD 4 is the only lectin domain that requires divalent cation to bind Man when expressed alone [129]. Only CRDs 4 and 5 resist digestion by a protease specific for a sequence that is inaccessible when bound to calcium [130], indicating that the remaining CRDs are not functional. However, the expression of truncated polypeptides has shown that CRDs 4–8 are required for binding to mannosylated glycoproteins, mannan, invertase, and Man-BSA neoglycoprotein, as well as to demonstrate ligand endocytosis equivalent to wild-type MMR [130]. Site-directed mutagenesis has established that the second ("structural") calcium binding site in MMR is located in a position unique to this *C*-type lectin [131]. Similar to other Man-specific lectins, this receptor also binds weakly to hexoses with equatorial vicinal C-3 and C-4 hydroxyls such as Fuc and Glc/NAc. The biologically relevant ligands for MMR were originally described (see main discussion) as polyvalent mannosylated ligands such as microbial carbohydrate polymers which could interact with the clustered CRDs of a single polypeptide. thus MMR would appear to function in host defense, clearing foreign molecules by endocytosis [131]. The orientation of Man bound to MMR as determined by NMR [132] is notably different from that of MBP and may reflect subtle differences in the binding specificity for these lectins.

The 180 kDa S4GGnM receptor is structurally similar *C*-type lectin located in liver Kupffer cells and hepatocytes and is specific for 4-sulfo-GalNAcβ1-4GlcNAcβ1-2Manα-bearing glycopeptides such as lutropin (LH) and thyrotropin. this highly regulated carbohydrate determinant is responsible for the rapid clearance of these hormones from the serum following their release from the pituitary gland [133]. This receptor and the pituitary synthesis of its unique carbohydrate ligand are conserved in vertebrates ranging from fish to mammals [134].

Since MMR and S4GGnM-R possess significant sequence homology, and similar topology and antigenicity, the cross-reactivity of their binding specificity has been examined. Surprisingly, both MMR and S4GGnM-R contain eight *C*-type lectin CRDs specific form Man as well as a single *N*-terminal cysteine-rich region specific for S4GGnM [135]. Since both lectins have dual specificity for both ligands and are highly related, the above described limited roles for each lectin have called into question [135]. It has been suggested that a more appropriate nomenclature for these lectins would be the liver and lung Man/S4GGnM-R. A potential explanation for the biological differences between these lectins utilizes the receptor oligomerization theory. Perhaps a cluster of *N*-terminal S4GGnM-recognizing regions of a dimeric/multimeric liver Man/S4GGnM-R could bind ligands such as lutropin tightly, while a monomeric lung Man/S4GGnM-R could not [135].

DC-SIGN and DC-SIGNR are homologous *C*-type lectins belonging to subgroup II. The DC-SIGN gene is located on the human chromosome 19p13 in close proximity to the DC-SIGNR (dendritic cell-specific ICAM-3 grabbing non-integrin-related) gene. These type II transmembrane receptors mediate cell-cell interactions within the immune system by binding to intracellular adhesion molecule-3 independent of the conventional ICAM-3 ligand, lymphocyte function-associated antigen-1 [9].

DC-SIGN (CD209) binds high mannose oligosaccharides, such as the $Man_9GlcNAc_2$ oligosaccharide found on the gp120 envelope glycoprotein of HIV-1, and in this way facilitates viral infection in *trans* of CD4+ T-cells [136]. DC-SIGN also functions as a receptor for *Mycobacterium tuberculosis* [137], hepatitis C virus [138], Ebola virus [139], and human cytomegalovirus [140].

3 P-Type Lectins

3.1 Architecture of P-Type Lectins

The two lectins that bind specifically to Man-6-phosphate (M6P) and compose the family of P-type lectins are the 43–46 kDa cation-dependent M6P receptor (CD-MPR) and the 275–300 kDa insulin-like growth factor II/cation-independent M6P receptor (IGF-II/CI-MPR). The structure and function of these receptors have been reviewed in detail elsewhere [141,142,143,144]. The IGF-II/CI-MPR is composed of a short *N*-terminal region, followed by a large extracellular domain of 15 repeating subunits, followed by a single transmembrane domain, and a *C*-terminal cytoplasmic tail. The CD-MPR is composed of a short *N*-terminal region, followed by a single extracellular domain homologous to each of the 15 repeating domains of the IGF-II/CI-MPR, followed by a single transmembrane domain, and a *C*-terminal cytoplasmic tail [141]. A schematic representation of the two M6PR lectins is shown in ❯ *Fig. 6* [143].

The binding sites for M6P have been localized to the extracellular domain of CD-MPR and to extracellular repeats 3 and 9 of IGF-II/CI-MPR [145]. Site-directed mutagenesis has implicated Arg 435 (in repeat 3) and Arg 1334 (in repeat 9) in bovine IGF-II/CI-MPR as important for binding [146]. Scanning Ala mutants of bovine CD-MPR in Gln 66, His 105, Arg 111, Glu 133, Arg 135, and Tyr 143 among others, demonstrated a loss in binding affinity and predicted that these residues may be important for binding [147]. The 3-dimensional structure of the

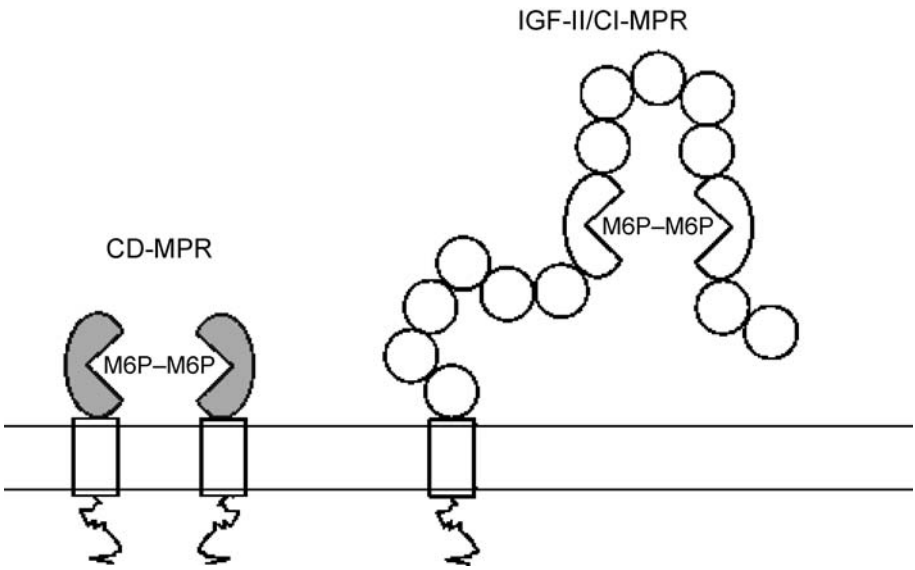


Figure 6
Model for receptor interactions with bivalent mannose-6-phosphate ligands. Dimeric CD-MPR subunits each bind one molecule of M6P, while monomeric IGF-II/CI-MPR has two lectin domains, each binding one molecule of M6P. (Adapted with permission from [143])

extracellular region of bovine CD-MPR complexed with M6P has recently been determined by X-ray crystallography [148]. The composition of this domain includes nine β-strands and one α-helix with three disulfide bonds oriented such that strands 5–9 and 1–4 form perpendicular β-sheets in a β-barrel arrangement as shown in *Fig. 7*. All six cysteines participate in inter-sheet stabilization by three disulfide bonds [148].

Amino acids involved in binding M6P are shown in *Fig. 8*. The phosphate moiety is coordinated by His 105, Asn 104, Asp 103 and divalent cation. Each hydroxyl of Man is hydrogen-bonded to binding site residues, while no interactions between the hydrophobic carbohydrate backbone and aromatic side chains are observed [148]. The multiple contacts between lectin and M6P, in contrast to that observed for many lectins which mediate cell surface interactions, are reflected in μM binding constants, reminiscent of bacterial arabinose-binding protein which functions as a sugar transport protein [149].

In contrast to CD-MPR, IGF-II/CI-MPR lacks residues corresponding to Glu 101 and Arg 111 which coordinate divalent cation, and may help explain why this lectin does not require metals for high-affinity binding [148]. In contrast to the *C*-type lectins which show no binding in the presence of EDTA, the absence of divalent cation only moderately raises the binding constant of CD-MPR to ligands such as pentamannose phosphate from 2×10^{-6} M to 2.5×10^{-5} M [150].

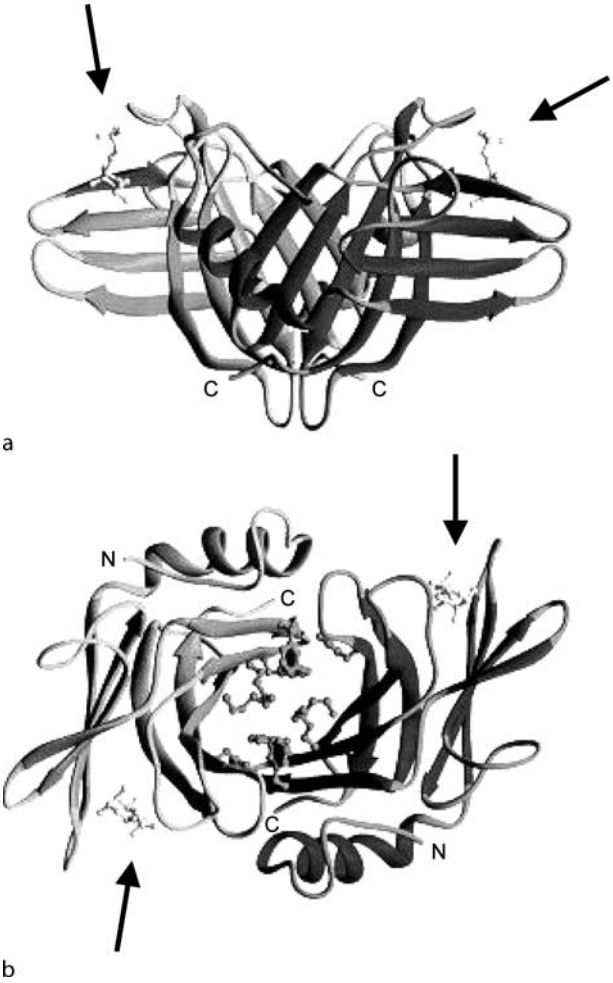


Figure 7
X-ray crystal structure of dimeric bovine cation-dependent mannose-6-phosphate receptor bound to ligand. (a) "side" view, (b) "top view." (Reproduced with permission from [148])

3.2 Oligomerization of P-Type Lectins

Equilibrium binding experiments with monomeric M6P and diphosphorylated high-mannose oligosaccharides have established that CD-MPR has one binding site per monomer which can be occupied by a single monovalent or a single divalent ligand in solution [150], while similar studies on the IGF-II/CI-MPR revealed two binding sites per monomer which can be occupied by two monovalent or a single divalent ligand in solution [151]. The relative avidities as measured by equilibrium dialysis of CD-MPR for M6P or a diphosphorylated high-mannose oligosaccharide were 8×10^{-6} M and 2×10^{-7} M, respectively [150], as compared

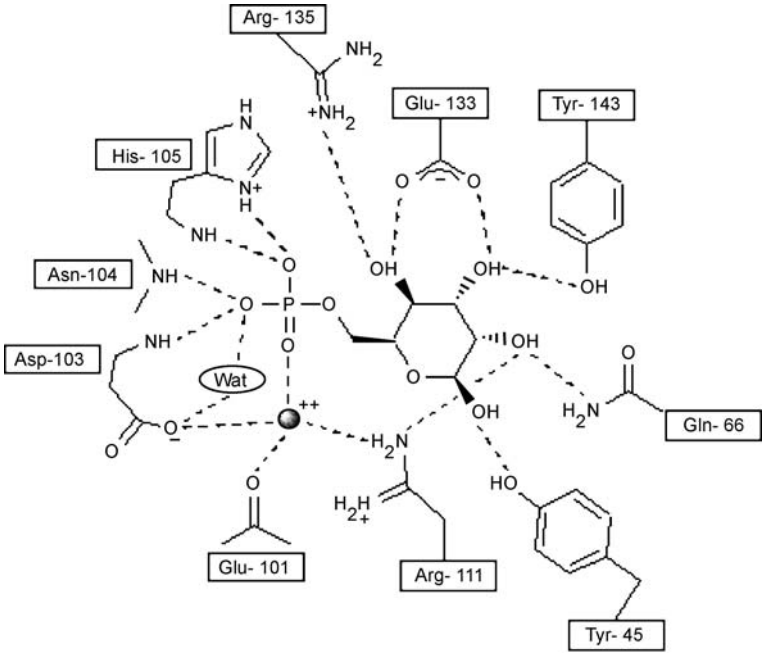


Figure 8
Representation of the binding site for the bovine cation-dependent mannose-6-phosphate receptor bound to ligand. *Dashed lines* indicate hydrogen bonds while the *bottom left circle* indicates divalent cation. (Adapted with permission from [148])

to 7×10^6 M and 2×10^{-9} M, respectively, for IGF-II/CI-MPR [151]. On the basis of this data, it has been proposed that a dimeric MPR binds with high affinity to glycoproteins decorated with diphosphorylated oligosaccharides, while no such dimerization may be required for ligand interactions with IGF-II/CI-MPR [141]. The two binding sites of IGF-II/CI-MPR are not functionally equivalent, as evidenced by the failure of mutants of critical Arg residues in either domain 3 or 9 to show compensation for each other [152].

The CD-MPR crystallizes as a dimer with relevant binding sites spaced 40 Å apart [148]. The distance between antennae for *N*-linked oligosaccharides is generally considered to be 20 Å [72], although X-ray crystallography has shown the non-reducing ends of a biantennary oligosaccharides as far as 24–27 Å apart [153], which indicates that a diphosphorylated high-mannose oligosaccharide is unlikely to interact with both binding sites of a dimeric receptor. The authors propose that a tetrameric receptor may be capable of mediating this interaction, although this theory is at odds with above mentioned data noting a binding ratio of one diphosphorylated oligosaccharide to two CD-MPR [148].

Both M6P receptors have been noted to form oligomers. Cross-linking of soluble CD-MPR and IGF-II/CI-MPR have demonstrated that receptor oligomerization depends upon pH, receptor concentration, and the presence or absence of ligand [154,155,156]. The on-off rates of association are relatively fast, indicating that oligomerization is a process in dynamic equilibrium.

Dimeric and tetrameric forms of CD-MPR are favored by a pH optimum of 6.3–6.4, dropping sharply above and below, and by the presence of ligand [156]. There is evidence that upon binding glycoproteins bearing multiple M6P moieties, monovalent IGF-II/CI-MPR dimerizes [157]. An attractive theory which explains the formation of oligomeric M6PRs has been advanced. High-affinity binding of M6PRs may be due to self-association in the golgi near neutral pH at high receptor concentration. Upon vesicular transport to endosomes, the dilution of receptor and lowering of pH induces self-dissociation, causing release of ligand, providing a mechanism for receptor recycling between golgi and endosome [141]. However, the experimental conditions of isolated receptors and membrane preparations under which this hypothesis has been tested may or may not reflect the actual cellular environment. Recently, cross-linking of CD-MPR in intact cells has been performed with membrane-permeable reagents. In this in situ system, CD-MPR remained dimeric upon raising endosomal pH, metabolic blockage of M6P ligand, or localization to the cell surface [158]. These findings suggest that tuning of receptor affinity for ligand is unlikely to be regulated by oligomerization. Additionally, the authors suggest that an alternative mechanism of regulating receptor affinity may be due to conformational changes in the lectin or ligand [158].

3.3 Biological Roles of P-Type Lectins

In few cases have the biological roles of lectins been more firmly established than for CD-MPR and IGF-II/CI-MPR. In the trans-golgi both receptors bind to proteins which have been decorated with M6P, such as acid hydrolases. The receptor *C*-terminal cytoplasmic targeting amino acid sequence serves as a recognition signal for transport to the endosomal compartment, where dissociation of lectin and M6P-protein takes place [141,142,143]. Lysosomal storage diseases (reviewed in [159]) are characterized by the failure to deliver acid hydrolases (or the delivery of ineffective acid hydrolysases) to the endosome, *en route* to the lysosomal compartment. This causes the accumulation of undegraded macromolecules in these compartments. Although the vast majority of lysosomal storage diseases are traced to biological defects other than M6P receptors, these two lectins are the only known proteins which serve to guide endosomal trafficking. CD-MPR and IGF-II/CI-MPR have overlapping but not identical roles. There is evidence that while IGF-II/CI-MPR is capable of endocytosis of ligands from the cell surface, CD-MPR is not. IGF-II/CI-MPR but not CD-MPR serves to turnover IGF by endocytosis [160,161]. Ligands for the M6P receptors are overlapping but not redundant, such that a high affinity ligand for CD-MPR is not necessarily a high affinity ligand for IGF-II/CI-MPR [162].

A strain of genetically engineered homo- and heterozygous “knockout” mice lacking CD-MPR has been shown to be indistinguishable from normal mice, including levels and activities of lysosomal enzymes [163]. In contrast, the IGF-II/CI-MPR knockout mutation is lethal [164]. However, IGF-II/CI-MPR null mice can be “rescued” to viable but abnormal phenotypes by genomic alteration to provide a background lacking in either IGF or IGF1r (IGF-receptor) [165]. The implication is that death is a consequence of overstimulation by IGF, while a milder phenotype is likely mediated by the CI-MPR function in IGF-II/CI-MPR knockout mice [165].

The IGF-II /CI-MPR has also been found in the circulation as a result of proteolytic release from the cell surface. The soluble receptor has been postulated to sequester IGF-II, thereby controlling the proliferative effects of this mitogen [166]. Consequently, the ability of the IGF-II/CI-MPR to target lysosomal enzymes, to bind TGF-b, and to bind IGF-II has implicated it in a role in tumor suppression. The discovery that there are a variety of mutations in the IGF-II/CI-MPR in human cancers have provided further evidence of this association [167,168,169,170].

4 *S*-Type Lectins

4.1 Architecture of the Galectins

The *S*-type (referring to conserved cysteine residues) lectin family has been designated the galectins [171], reflecting their common specificity for Gal presented as Galβ1-4Glc/NAc or Galβ1-3Glc/NAc. Thus far, there are fourteen numbered galectins (galectins 1–10, and 12–14) and five galectin like proteins (GRIFIN, HSPC159, PP13, PPL13, and OvGal11). Galectins do not contain a signal peptide and are secreted by non-classical mechanisms. They are not integral membrane proteins and binding does not require metals [172]. A diagrammatic representation of the known mammalian galectins is shown in *Fig. 9*. With the possible exceptions of galectins 3, 5, and 10, all known mammalian galectins are either dimeric or contain two CRDs in a single compact structure.

The structure and function of galectins have been reviewed in detail elsewhere [171,172,173,174,175,176]. Three-dimensional structures for galectins 1–3, 7, and 10 have been determined

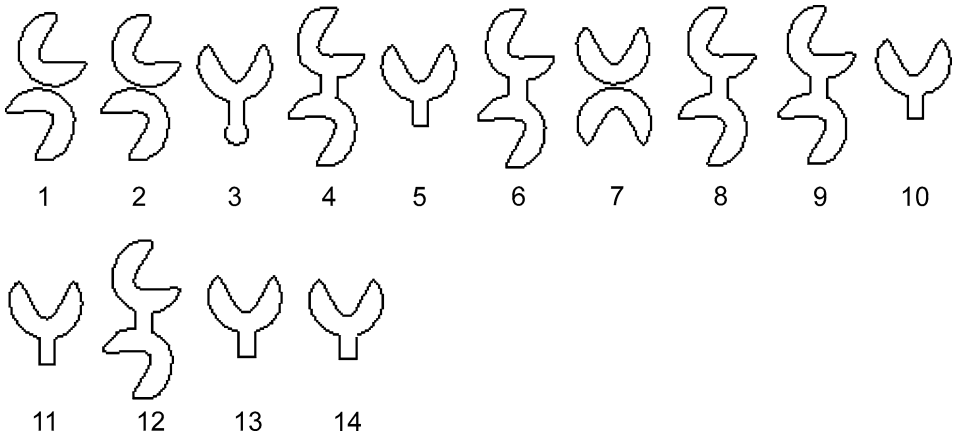


Figure 9
Representation of the structures of galectins 1–14. Galectins 1, 2, and 7 are shown as dimers, while monomeric galectins 4, 6, 8, 9, and 12 each contain two lectin domains. Galectins 3, 5, 10, 11, 13, and 14 are depicted as monomers each with a single lectin domain. (Adapted with permission from [175])

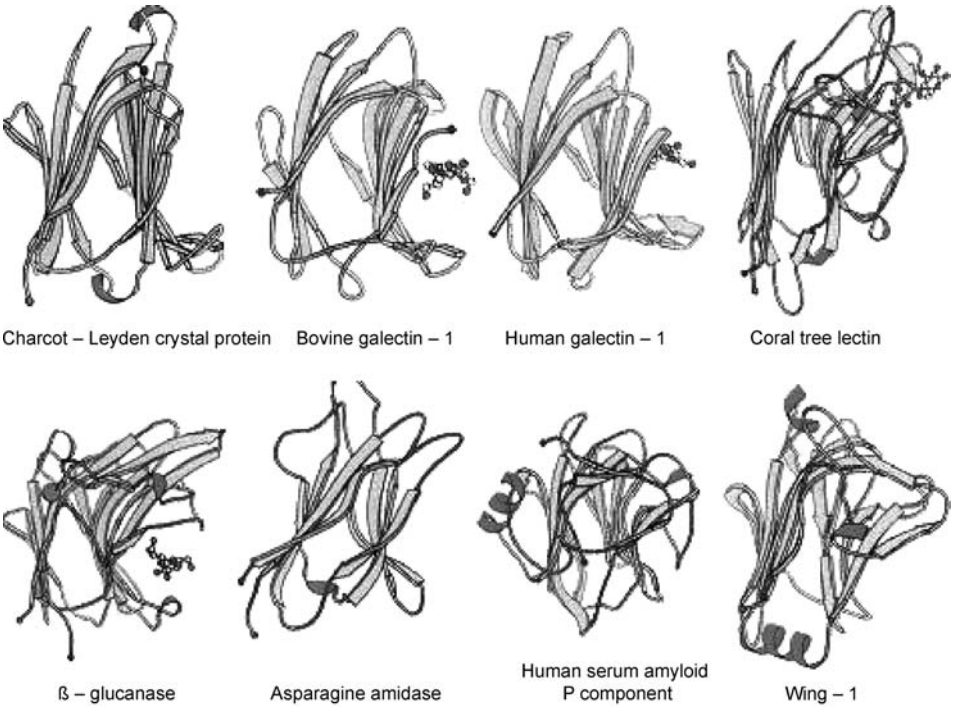


Figure 10
X-ray crystal structures of the lectin domains of human and bovine galectin 1, galectin 10 and similar domains from other proteins. (Reproduced with permission from [181])

by X-ray crystallography [177,178,179,180,181]. A common motif determined for galectins consists of 5 conserved β-strands forming the first β-sheet, opposed by 6 β-strands forming a second β-sheet oriented in a β-sandwich which has become known as the "jelly-roll" motif [182], named for its resemblance to a pastry. This motif is common to all known legume lectins (reviewed in [4,5]), being first identified in concanavalin A, and likely is the result of convergent evolution from different ancestors since there is little primary amino acid or nucleotide homology between legume lectins and galectins. *Figure 10* depicts the lectin domain of three galectins, and shows that the motif is common to many lectins and otherwise unrelated proteins [181].

4.2 Oligomerization of Galectins

The dimeric nature of galectins 1 and 2 was first demonstrated by size-exclusion chromatography and crosslinking experiments [183]. Their 3-dimensional structures have more recently been characterized as dimers of individual 14 kDa CRDs, demonstrating two-fold axes of symmetry [153,177,178]. The CRDs of a single dimer are oriented away from each other such that binding sites cannot each interact with antennae of a single biantennary oligosaccha-

ride [153], but can serve as a cross-linker for β-galactosides. Non-covalent dimeric association is observed in galectins 1, 2, 5, 7, 10, 13, and 14 [184]. Galectins 4, 6, 8, and 9 and 12 contain two CRDs in a single compact protein, oriented apart as seen in galectins 1 and 2 [185, 186,187].

The crystal structure of galectins 1, 2, 3, 7, and 10 have been solved providing a detailed understanding of their dimeric nature [153,178,179,180]. Residues corresponding to the *N*-terminal region of galectins 1 and 2 which have been implicated as important for dimerization are absent in galectin 3 [179]. However, it has been shown that full-length galectin 3 forms oligomers which can be trapped by chemical cross-linking, while galectin 3 expressed as only the *C*-terminal lectin domain is primarily monomeric [188]. Additionally, full-length galectin 3 demonstrates hemagglutination activity while the *C*-terminal fragment is devoid of this activity [188]. Galectin 3 also shows positive cooperativity in binding to laminin, giving rise to speculation that both subunits of this possibly dimeric lectin can simultaneously bind Gal on the antennae of a biantennary *N*-glycans [189]. It is unclear if galectin 5 is functionally a monomer or an oligomer. While it does not show the presence of higher ordered species by size exclusion chromatography, its ability to hemagglutinate erythrocytes suggests that its functional state is oligomeric [184].

Charcot–Leydon crystal protein (CLC) has been designated galectin 10 based on its 15–30% sequence identity and structural similarity to other galectins [181]. Additionally, seven of twelve residues conserved in galectin carbohydrate binding sites are present [181]. CLC autocrystallizes in a hexameric bipyramidal arrangement [190]. CLC weakly binds to GlcNAc or LacNAc and its lectin activity has only been shown by binding to these monosaccharides immobilized on affinity columns [185]. CLC demonstrates enzymatic activity as a lysolecithin acylhydrolase and is a major constituent of total protein in eosinophils and basophils [191].

The amino acids considered to be required for ligand binding by galectins are shown for the representative galectin 2 in complex with lactose in ❯ *Fig. 11* [178]. The hydrogen bonding interactions are dominated by the C-4 hydroxyl group of Gal which binds to the side chains of His 45, Asn 47, and Arg 49, and the C-3 hydroxyl group of Glc which is in contact with Arg 49, Asn 58, Glu 68, and Arg 70. Also seen are hydrogen bonds from Asn 58 and Glu 68 to the C-6 hydroxyl of Gal. The multiple hydrogen bonds to the C-4 hydroxyl and the C-6 hydroxyl suggests that ligand discrimination is accomplished through recognition of the axial C-4 hydroxyl. Hydrophobic interactions of the indole side chain of Trp 65 with the pyranose ring of Gal suggest a reason why the binding site cannot accommodate a monosaccharide with an equatorial C-4 hydroxyl [178]. X-ray crystal structures of galectins bound to lactose, LacNAc or biantennary *N*-acetyllactosamine terminating oligosaccharide show a conserved hydrogen bonding network consistent with that determined for galectin 2 [164,181] (❯ *Table 1*).

While the terminal Gal residue is in intimate contact with the lectin binding pocket, the Glc/GlcNAc residue is oriented away from this region, presumably making a smaller contribution to binding specificity and affinity [177]. The lactose/LacNAc binding sites of galectins are located within a wide protein cleft, suggesting that the internal sugar residues of carbohydrate polymers, such as polylactosaminoglycans, can be accommodated by galectins [180]. No thiol residues participate in ligand binding and the cysteines can be either oxidized or reduced without affecting binding [177]. Additionally, the complex of galectin 1 with biantennary oligosaccharide shows the ability of galectins to crosslink molecules bearing greater than one ligand in solution [153].

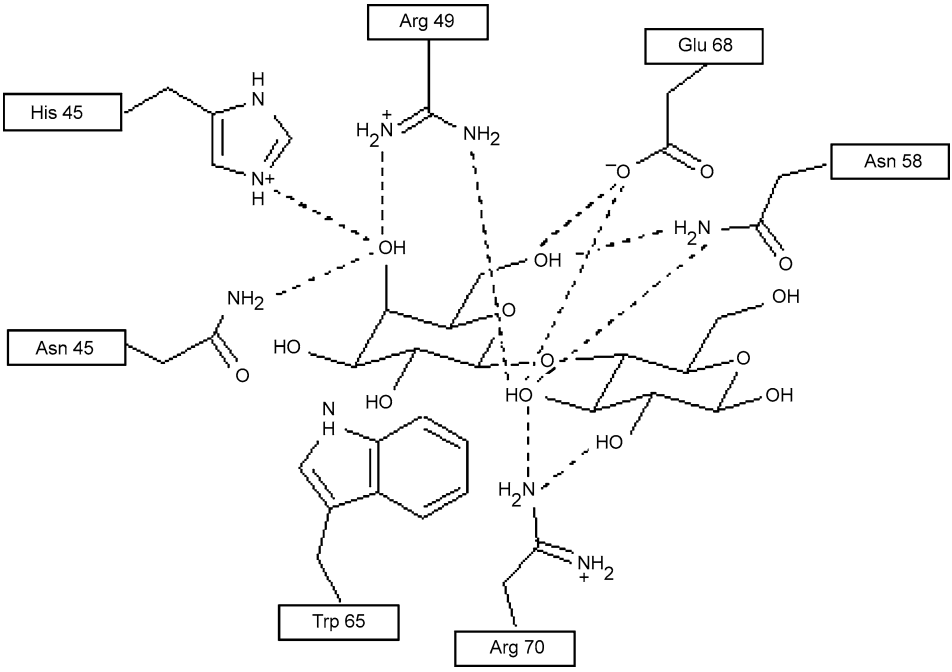


Figure 11
X-ray crystal structure of lactose bound to galectin 2. *Dashed lines* indicate hydrogen bonds. (Adapted with permission from [178])

4.3 Proposed Biological Roles of Galectins

Galectins are conserved in a number of unrelated organisms such as marine sponges, nematodes, frogs, birds, and mammals, suggesting that β-galactoside binding lectins may be biologically important [175,185,187,192]. The tissue-specific localization of several galectins is the strongest evidence for their functions. While it is clear from the examples given below that a number of galectins are involved in growth regulation, cell adhesion, cell migration, cancer, and immune responses, it is not clear that any of these processes are meditated by galectins in non-redundant pathways. Genetically engineered "knockout" mice lacking galectins 1 and 3 show no phenotypic abnormalities, indicating that these galectins under normal conditions are not necessary for survival [193]. However, in thioglycolate-induced peritonitis, galectin 3 null mice demonstrate reduced recruitment of leukocytes to sites of inflammation [194].
Galectin 1 is proposed to mediate cell adhesion, regulate cellular proliferation, and mediate cellular apoptosis [192]. Galectin 3 has been implicated in cell adhesion, regulation of inflammation, pre-mRNA splicing, as well as protecting against induced apoptosis [192]. Cell-cell and/or cell-extracellular matrix cross-linking functions have been proposed for galectins 4 and 6 [185]. The selective expression of galectin 5 in erythrocytes suggests roles in maturation and/or cell adhesion of erythrocytes [185]. Likewise, galectin 7 is a constitutive marker for

Table 1
Conservation of the hydrogen bonding pattern of galectins binding to Galβ(1-4_Glc/NAc

Sugar Atom	Bovine galectin 1 [175]	Human galectin 2 [177]	Human galectin 3 [176]	Human galectin 7 [178]	Human galectin 10 [180]
Gal O4	His 44	His 45	His 158	His 49	His 53
	2.8	2.7	2.7	2.7–2.9	2.9
Gal O4	Arg 48	Arg 49	Arg 162	Arg 53	–
	3.0	3.3	3.1	2.6–3.2	
Gal O4	Asn 46	Asn 47	Asn 160	Asn 51	Gln 55
	3.4	3.1	3.2–3.3	3.2–3.3	2.7
Gal O5	Arg 48	–	Arg 162	Arg 53	–
	2.9		3.0	2.8–3.1	
Gal O5	–	–	–	Glu 72	–
				3.0	
Gal O6	Asn 61	Asn 58	Asn 174	Asn 62	Asn 65
	2.7	2.7	2.9	2.4–2.9	3.4
Gal O6	Glu 71	Glu 68	Glu 184	Glu 72	Gln 75
	2.8	2.6	3.1	2.5–3.1	2.9
Glc/GlcNAc O3	Arg 73	Arg 70	Arg 162	–	–
	3.3	3.1	2.7–2.8		
Glc/GlcNAc O3	Arg 48	Arg 49	–	Arg 53	Arg 61
	2.8	2.6		3.1	2.9
Glc/GlcNAc O3	Glu 71	Glu 68	Glu 184	Glu 72	Gln 75
	2.4	2.8	2.6	2.6	2.8
Glc/GlcNAc O3	Arg 48	–	Arg 162	–	-
	3.2		3.0		

Summary of the hydrogen bonding pattern of galectins bound to ligand(s). Beneath the identified binding amino acid, the interatomic hydrogen bond distance is given in Å. A range of interatomic distances indicates that the X-ray structure has been independently solved bound to different ligands. Interactions with Glc O2 have been omitted for the purpose of clarity. Adapted from Leonidas et al. [180] with permission

all keratinocyte subtypes but its specific function is unclear at this time [195]. Galectin 9 may mediate clonal deletion of thymocytes, as evidenced by its role in lactose-dependent apoptosis of these cells in vitro [196]. An isoform of galectin 9 shows chemoattractant activity which is selective for eosinophils [197].

Several galectins have altered expression in tumor cells relative to normal cells. The high expression of galectin 1 correlates with thyroid tumors whereas high and low expression of galectin 3 has been proposed as a marker for a number of other malignancies [192]. The absence of galectin 7 in squamous carcinoma cell lines may serve a diagnostic role [195]. Likewise, the overexpression of galectin 8 may serve as a marker for prostate cancer [187]. The expression of galectin 9 is increased in Hodgkin's lymphoma (HL) patients where 50% of clinical populations possess antibodies against galectin 9 [185].

5 Pentraxins

The non-covalent association of identical subunits of C reactive protein (CRP) and serum amyloid protein (SAP) into flat pentameric disc-like arrangements identifies the pentraxin family of lectins [198]. These calcium-dependent plasma-soluble lectins containing a conserved octapeptide "signature sequence" of HXCXS/TWXS are categorized into the classical short pentraxins CRP and SAP (reviewed in [199,200]), and the more recently described long pentraxins (reviewed in [201]). The *C*-terminal region of long pentraxins bears significant homology to short pentraxins [202].

5.1 Short Pentraxins

The short pentraxins were first described in the 1930s [203] and are the major components of a host of cytokine-induced liver proteins which mediate clearance of pathogens, known as acute phase proteins [200]. CRP shows 67–76% and SAP demonstrates 63–70% identical amino acids for the same protein in different species, while CRP and SAP share 42–54% identity in the same species [200]. Short pentraxins are conserved in vertebrates and are also present in unrelated organisms such as the horseshoe crab [202]. A number of biological functions have been proposed for CRP and SAP including complement activation, binding to sites of injury, activation of leukocytes, and association with amyloid fibrils [200].
The 3-dimensional structure of SAP as determined by X-ray crystallography is considered to be representative of pentraxin structures [204,205]. The overall topography is consistent with earlier studies observing a decameric stacking of pentamers of disc-like identical monomer subunits (*Fig. 12*). However, CRP oligomerizes into pentameric, but not decameric multimers [202]. Pentraxin monomers display a jelly-roll motif, similar to galectins and legume lectins, with two anti-parallel β-sheets of seven and eight β-strands respectively, around a hydrophobic interior where one strand is flanked on either end by α-helices [204,205]. Notably, pentraxins and legume lectin domains show approximately 11% sequence identity [201]. SAP contains two bound calcium ions, one of which occupies a coordination site

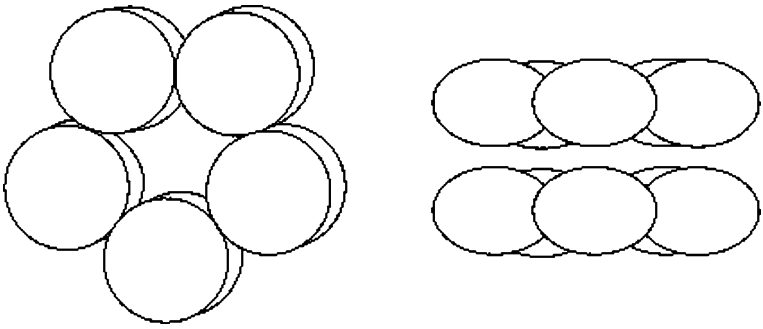


Figure 12
Representation of the decameric arrangement of pentraxin serum amyloid P component subunits. (a) "top" view, (b) "side" view

between β-strands and serves to orient binding site amino acid residues [204,205]. Although the highest affinity ligands for SAP are phosphoethanolamine (PE) and phosphocholine (PC), the lectin selectively binds to terminal β-galactosides and also binds to polyanions such as heparin [206,207]. The complex of SAP with 4,6-*O*-(1-carboxyethylidene)-β-D-Gal showed only one hydrogen bond to Gal while the majority of binding interactions were specific to the aglycone, explaining the failure of monosaccharides to show specific binding. An adjacent and overlapping but not identical site has been identified as responsible for binding to PE and PC [204,205].

CRP binds terminal β-galactosides at a pH optimum of 5–6 and lectin binding assays require that both lectin and ligand be immobilized. Thus it has been suggested that serum CRP does not function as a lectin under normal physiological conditions and binds carbohydrate only under conditions where tissues are inflamed or hypoxic [208].

5.2 Long Pentraxins

Although more than seven highly homologous long pentraxins have been biologically characterized, in few cases they have been shown to bind carbohydrate. For example, lectin activity was satisfied for 46 kDa rat Narp (**n**euronal **a**ctivity-**r**egulated **p**entraxin) by binding to agar in the absence of EDTA [209]. Alternatively, the multimeric organization of pentraxins has been well documented. The formation of higher molecular weight oligomers of apexin [210,211], PTX3 [212], and horseshoe crab limulin [213,214,215] has been shown to require non-reducing conditions, which converge to monomeric species in the presence of DTT.

Horseshoe crab limulin binds to mammalian sialic acids and bacterial KDO [213] and is a sialic acid-specific hemagglutinin as well as cytolytic to erythrocytes in vitro [214]. In contrast to other pentraxins, limulin subunits have been proposed to associate as twelve 45 kDa AB complexes in doubly stacked hexameric disc-like oligomers [202,214].

6 I-Type Lectins

6.1 The Immunoglobulin Fold

Proteins belonging to the immunoglobulin (Ig) superfamily possess domains homologous to the constant (C) and variable (V) regions of antibodies. C regions of antibodies are composed of 5–10 amino acid long anti-parallel β-strands with strands ABDE and GFC (*Fig. 13*) organized into sandwiches of two opposing β-sheets around a hydrophobic interior stabilized by disulfide bonds between strands B and F. V regions bear similar topology to C regions yet contain two additional β-strands making GFCC′C″ β-sheets [216] as depicted in *Fig. 13*. This structural topology was originally described for antibodies [217], and subsequent structural determinations have shown that a large and diverse superfamily of proteins, some of which are lectins, contain domains with homology to the immunoglobulins [216].

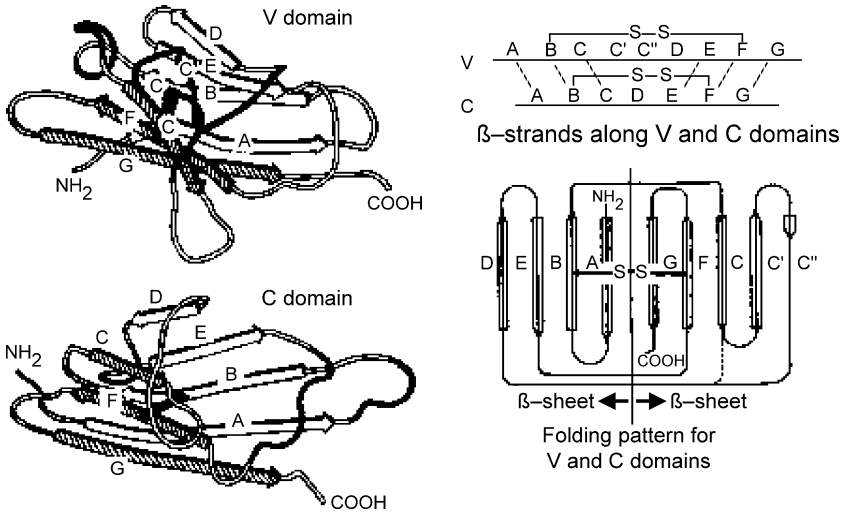
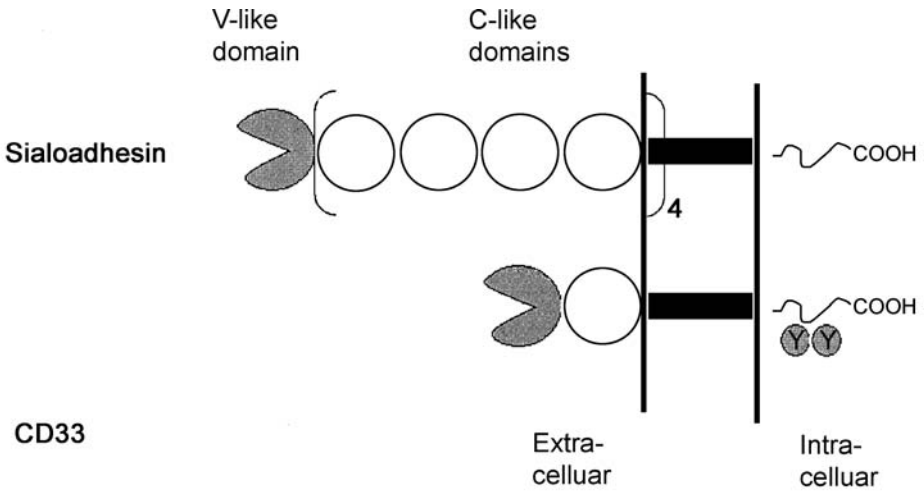


Figure 13
The immunoglobulin fold shown for an antibody. β-strands associate into β-sheets which compose the V and C regions. See text for details. (Reproduced with permission from [216])

6.2 The Architecture of Siglecs

A subset of lectins belonging to the I-type (Immunoglobulin-like) family formerly known as sialo adhesions [218] has been designated the siglecs (**s**ialic acid binding **i**mmuno**g**lobulin-like **lec**tins) [219]. The single *N*-terminal domain of the siglecs is homologous to the V region of antibodies and is followed by a varied number of repeat domains homologous to the C region of antibodies (*Fig. 14*). V-like and C2-like domains are followed by a transmembrane segment, and a *C*-terminal cytoplasmic tail which typically contains tyrosine motifs. The currently known members of the siglec family are sialoadhesion (Sn, hSiglec-1), CD22 (hSiglec-2), CD33 (hSiglec-3), myelin-associated glycoprotein (MAG, hSiglec-4), and hSiglec-5–hSiglec-11. Homologies between siglecs are realized in their *N*-terminal V-like carbohydrate-binding domains and in their cytoplasmic tyrosine motifs, and variation is seen in the tissue and cell types they are expressed in [220]. The I-type lectins which do not bind sialic acid and have C2 or V1-like domains include integrins ICAM-1 (CD54), PECAM-1 (CD31), and N-CAM comprised of repeats of C2-like domains, while P_o glycoprotein contains a single V-like domain [7].

Antibody sequence homology predictions and the X-ray crystal structure of hSiglec-1 in complex with NeuAcα2-3Galβ1-4Glc [221] have shown that although the V-like *N*-terminal domain adheres to the canonical Ig fold, four major structural differences between antibodies and siglecs exist. The siglecs have an intersheet disulfide bond between strands B and E, rather than between B and F in immunoglobulins. Likewise, the G strand is split into two smaller strands, the B-C loop is larger and extended, and there is a lack of β structure in the C$''$ region of siglecs [221].



Human siglecs	Number of Ig domains	Number of cytoplasmic tyrosine motifs	Sialic-acid-binding specificity	Expression
Sialoadhesin/hSiglec-1	17	0	2,3>2,6>2,8	Mac
CD22/hSiglec-2	7	6	2,6	B
CD33/hSiglec-3	2	2	2,6>2,3	My-pro, Mon
MAG/hSiglec-4	5	1	2,3>>2,6	Oligo, Schwann
hSiglec-5	4	2	2,3=2,6>2,8	Mon, N, B
hSiglec-6	3	2	STn	B, Plac
hSiglec-7	3	2	2,8>>2,6>2,3	Mon, NK
hSiglec-8	3	2	2,3>2,6	Eos, Baso, Mast
hSiglec-9	3	2	2,3=2,6	Mon, N, NK, B
hSiglec-10	5	3	2,3=2,6	Mon, NK
hSiglec-11	5	2	2,8	Mac

Figure 14
Representation of siglec domains. *N*-terminal lectin domains are shown with a binding pocket, extracellular domains are depicted as *circles*, transmembrane units as *black rectangles*, *C*-terminal cytoplasmic tails by *lines*, and tyrosine-based motifs as Y. 2,3, 2,6 and 2,8 refer to the α2,3- α2,6- and α2,8-glycosidic linkages of sialic acid to galactose (2,3 and 2,6) or sialic acid itself (2,8). B, Bcells; Baso, basophils; Eos, eosiniphils; Mac, macrophages; Mast, mast cells; Mon, monocytes; My-pro, myeloid progenitors; N, neutrophils; NK, natural killer cells; Oligo, oligodendrocytes; Plac, placentral trophoblasts; Schwann, Schwann cells

The region of antibodies responsible for binding antigen is on the ends of the β-sandwich, at the hypervariable regions [216]. However, the carbohydrate binding sites of hSiglec-1 and CD22 were predicted by site-directed mutagenesis to be along the face of the β-sandwich, with key residues along strands G, F, C, and C′ implicated as important for binding [222,223]. The 3-dimensional structure of hSiglec-1 has confirmed these predictions, and demonstrated the role of a basic arginine residue as an electrostatic partner for the C-1 carboxylate of

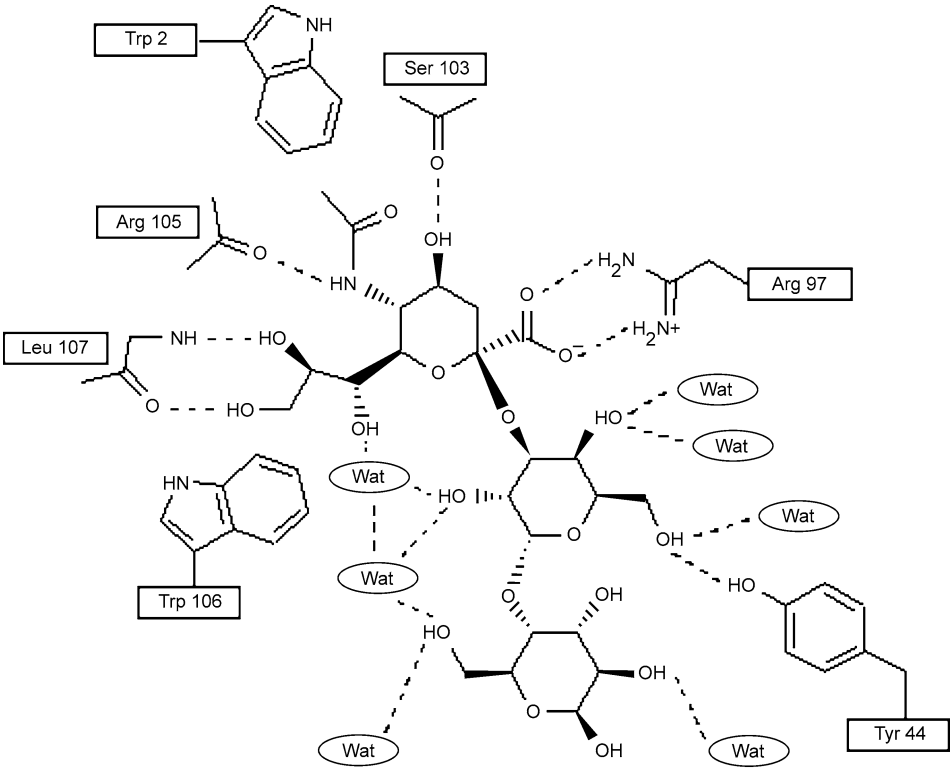


Figure 15
X-ray crystal structure of the binding site of sialoadhesion with NeuAcα2-3Galβ1-4Glc bound. *Dashed lines* indicate hydrogen bonds. (Reproduced with permission from [221])

NeuAc [221]. The complex of sialoadhesin with NeuAcα2-3Galβ1-4Glc (*Fig. 15*) shows seven hydrogen bonds and two hydrophobic interactions with sialic acid, while Gal and Glc make fewer contacts [221], consistent with the proposal that lactose is a scaffold upon which ligand discrimination centers on NeuAc.

6.3 Specificities of Siglecs

The ligands for all 11 known siglecs are sialylated carbohydrates. Although the most common sialic acid is NeuAc, it may be noted that there are over 40 biologically occuring sialic acids [224]. The most prevalent modifications to NeuAc are glycolyl acylation on the 5 position (a modification common in mammals but absent in humans [225]) and acetylation on the glycerol side chain [226]. The ligand specificities of siglecs were first determined by evaluating the binding of siglec-Fc chimeras to erythrocytes which had been desialylated, and then resialylated with sialyltransferases to prepare NeuAcα2-3Galβ1-3GalNAc (3O), NeuAcα2-3Galβ1-4GlcNAc (3N), or NeuAcα2-6Galβ1-4GlcNAc (6N) structures. CD22 bound only to

RBCs expressing 6N, MAG bound only to 3O, while CD33 and Sn bound to either 3O or 3N [218]. Subsequent studies showed that while CD22 binds only to α2-6 sialylated structures and siglec-5 binds to either α2-6 or α2-3 sialylated structures, all other siglecs bind only to α2-3 sialylated structures [227]. Chemical modification of the C-1 carboxylate, *N*-acetyl group, and glycerol side chain of NeuAc blocks the binding of MAG, Sn, and SMP to relevant glycolipids, while MAG but not Sn requires the C-7 and C-4 hydroxyls [228,229]. Modifications to the *N*-acyl group of NeuAc showed distinct binding specificities between siglecs, which has been proposed to function as a tuning mechanism for ligand discrimination [230].

Figure 16 shows truncated sialic structures highlighting the relevant interactions between sialic acid and siglec in comparison to other sialic acid specific lectins [8,231]. It is apparent that sialic acids are recognized by siglecs as the non-reducing terminal moiety of saccharides. The minimal structure required by human MAG was shown to be NeuAcα2-3Galβ1-4Glc [228], while Sn requires NeuAcα2-3Galβ1-4GlcNAc, human CD22 requires NeuAc(Gc)α2-6Galβ1-4Glc(NAc), and mouse CD22 requires NeuGcα2-6Galβ1-4Glc(NAc) [229]. The preference of murine CD22 for a C-5 *N*-glycolyl substituent is not mirrored by mouse MAG which prefers C-5 *N*-acetyl substituents [228,229].

6.4 Regulation of Siglec Binding

The sialylated structures which mediate siglec binding are relatively common epitopes, occurring on *N*- and *O*-linked oligosaccharides and glycolipids disseminated throughout the body. This observation leads to speculation that siglecs might bind non-specifically to any appropriately glycosylated structure. However, the binding of Fc-Siglec chimeras has been shown to be limited to certain tissue types [218] (*Fig. 17*). One explanation for this apparent discrepancy invokes the compartmentalization argument, noting that siglecs are restricted to certain cell lineages and physiological locations. An alternative explanation is that siglecs are masked or blocked until activated for binding. In order to demonstrate sialic acid-dependent binding for recombinant CD22, CD33, MAG, and SMP expressed in COS or CHO cells, it is necessary to enzymatically remove sialic acids from the surface of these cells [232,233,234,235]. The nomenclature adopted by the field has been to describe *cis* ligands as those expressed on the cell expressing siglecs, in contrast to *trans* ligands on the target cell [235]. Glycosylation on a single asparagine residue of CD22 and CD33 is required for the proper lectin folding, evident by failure to bind ligand, and lack of recognition by monoclonal antibodies when these lectins are either deglycosylated by treatment with *N*-glycosidase F or when this asparagine is mutated to no longer support glycosylation [233,234]. However, when the *N*-linked oligosaccharides at this site are sialylated, sialic acid recognition by the lectins is abrogated [234]. This indicates that either sialylated structures adjacent to siglecs, or the siglecs themselves block binding to exogenous sialylated ligands. This phenomenon is not observed for sialoadhesion, which can be rationalized by its 17 Ig-like domains extending the lectin domain from the cell surface, presumably out of reach of endogenous sialylated structures, or that Sn does not contain this *N*-glycosylation site [235,236]. However, this phenomenon is also absent for siglec-5, whose extracellular domains are a length intermediate between CD22 and CD33, and contain this *N*-glycosylation site [237]. It has been proposed that “masking” serves as a downregulatory mechanism where cell surface siglecs are blocked from binding cellular or secreted ligands

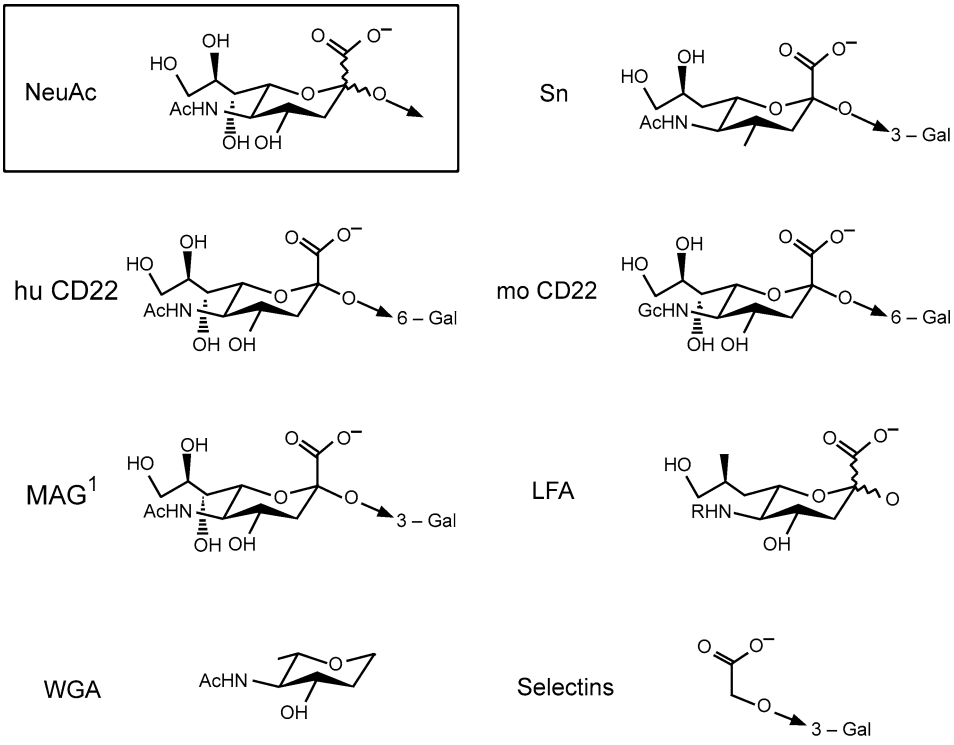


Figure 16
Truncated structures of sialic acids show the moieties required for binding to various sialic acid-specific lectins. LFA is *Limax flavus* agglutinin [231] and WGA is wheat germ agglutinin [5]. The binding importance of sialic acid C-4 and C-7 hydroxyls for MAG has not yet been determined. R indicates that LFA tolerates a variety of equatorial bulky substituents on the C-5 *N*-acyl group of sialic acids

until the cell upon which they are expressed is activated by desialylation [234]. This has been reproduced in biologically relevant B cells in the case of CD22 [234]. Desialylation could be accomplished by downregulation of sialyltransferases, or the action of neuraminidases.

6.5 Biological Roles of the Six Known Siglecs

Much is known about siglec structure and ligand-binding specificity, and in many cases tissue and ligand distribution have been elucidated (*Fig. 17*). However, it has been more difficult to determine the biological role of siglecs. Since Sn is only expressed in mature macrophages, and binds preferentially to cells of myeloblast lineage, it has been proposed to function in the regulation of myeloid cell interactions [218]. This is supported by the observation that the highest level of Sn expression under non-inflammatory conditions is in secondary lymphoid tissues and bone marrow macrophages where it is found at contact sites between macrophages and developing granulocytes [236].

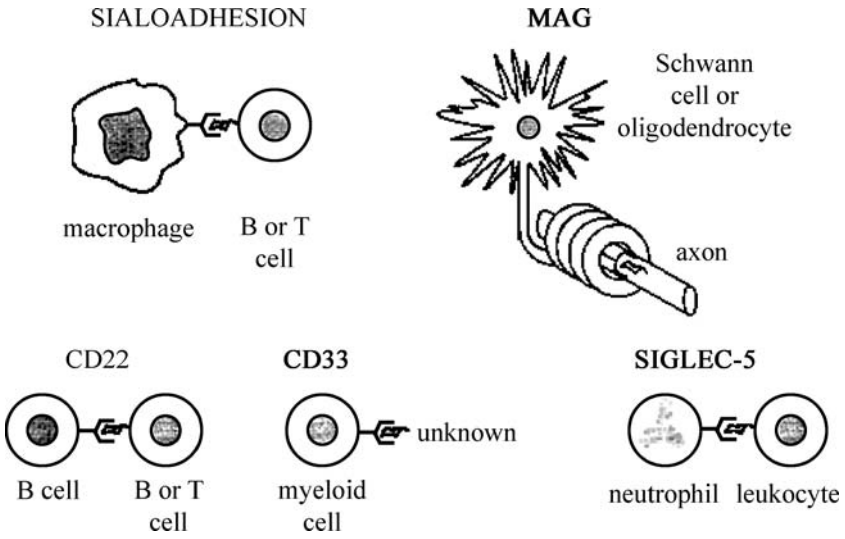


Figure 17
The tissue locations of siglecs and the tissues to which they have been shown to bind. Cells expressing siglecs are indicated by prongs. Cells expressing ligands are indicated by pyranose rings. The tissue CD33 binds have not yet been determined. Abbreviations: MAG (myelin-associated glycoprotein). (Adapted with permission from [218])

CD22 was first recognized as a useful marker for the stages of B cell development. Following the identification of CD22 as an Ig-superfamily member it was determined that cross-linking of CD22 with anti-CD22 IgG-conjugated beads induced an enhanced B cell immune response [238]. The cytoplasmic tail of CD22 contains a V/IXYXXL sequence which has been shown in other proteins to interact with an effector of the B cell antigen receptor SHP-1. CD22 has been proposed to interact with and dephosphorylate SHP-1, explaining its down-regulatory function in B cell antigen response [238]. The hyperimmune response of CD22 knockout mice is not as severe as that seen in SHP-1 knockout mice [239], suggesting a redundancy in this pathway. In addition to helping set the threshold for B cell receptor response, CD22 has also been proposed to have a role in B cell elimination, since it is a stage-specific marker [238]. In addition to the *C*-terminal cytoplasmic domain, the *N*-terminal lectin domain of CD22 could participate in regulation of antigen receptor signaling, since concomitant with the activation of B cells is the upregulation of βGal α2-6 sialyltransferase, which would cause self-association or association with molecules other than those in contact with the B cell receptor, drawing CD22 away from its function as a negative modulator [238]. This could also be accomplished through CD22 interactions with other α2-6 sialylated cells when they are in intimate contact with B-cells.

CD33 was first identified as a 67 kD differentiation marker for subsets of myeloid cells. It has also been useful as a marker for myeloid versus lymphoid leukemias. CD33 expression in cells of hematopoietic origin has led to the proposal that it participates in myeloid cell differentiation and mediates cellular interactions [240].

MAG is a 100 kD glycoprotein which has 8–9 potential *N*-linked glycosylation sites and is typically 30% carbohydrate by weight [241]. The location of MAG in myelinating oligodendroglia and myelin-forming Schwann cells and its intimate contact with axons has led to the proposal that it participates in the myelination of neurons [218]. MAG knockout mice have been shown to myelinate normally [242], or alternatively with subtly altered periaxonal architecture [243], indicating that degenerate pathways may operate in myelination or that the role of MAG in myelination is minimal. MAG has also been proposed to inhibit axon regeneration, as evidenced by in vitro and in vivo experiments. Cytoplasmic phosphorylation sites have been identified in MAG and a possible role has been elucidated in signal transduction pathways [241].

Siglec-5 has been proposed to participate in acute inflammatory responses, since it is localized to neutrophils and it does not require cell-surface desialylation to unmask its sialic acid binding. hSiglec-5-4L, a an hSiglec-5 isoform has been identified, and a soluble Fc chimaeric protein containing the hSiglec-5-4L extracellular domain binds glycophorin A in a sialic acid-dependent manner [244].

Other putative functions of Siglecs include binding sialic acid on the same cell surface, interacting with soluble sialylated glycoprotein ligands, or inhibiting cell-cell interactions involving Siglecs [10].

7 Conclusions

Great strides have been made with respect to the ligand binding specificities and molecular structures of the five families of mammalian lectins. A rational pattern of ligand discrimination emerges based upon the three-dimensional orientation of binding site amino acid residues, specific ligand tissue distribution, and lectin regulation to exhibit or lack binding activity under different cellular conditions. Lectins bind monosaccharides and monovalent ligands with low affinity while binding multimeric ligands and ligand-carrier complexes with high-avidity. Lectin oligomeric state and carbohydrate-binding domain orientation in multimeric association serves to discriminate between ligands (e. g. MBP is most likely to bind with high-avidity to foreign carbohydrate polymers while RHL is most likely to bind to triantennary Gal-terminating *N*-glycans on endogenous glycoproteins, both by virtue of the quaternary topology of their oligomeric subunits). Since the spacing of carbohydrate residues on oligosaccharides is important for biologically relevant lectin-carbohydrate interactions, the utilization of unnatural carbohydrate structures linked to synthetic polymer scaffolds as ligands for lectins may require careful consideration of spatial constraints.

Progress is being made to determine the biological functions of animal lectins. There is evidence for a myriad of roles for lectin-carbohydrate interactions, suggesting that lectins and sugars mediate their effects through non-redundant pathways. Lectin “knockout” experiments have demonstrated phenotypic abnormalities which can be traced to important biological roles. Since glycosylation is a highly regulated biological process conserved throughout evolution, it is obvious that these often subtle interactions serve to optimize and complement the function of organisms to provide an evolutionary advantage.

Acknowledgement

The authors thank V. Hayden Thomas for valuable suggestions and comments on this manuscript. This work was supported by NIH grants AI33189, GM48049 and NIH training grant GM07767.

References

1. Goldstein IJ, Hughes RC, Monsigny M, Osawa T, Sharon N (1980) Nature 285:66
2. Crocker PR, Feizi T (1996) Curr Opin Struct Biol 6:679
3. Varki A (1993) Glycobiology 3:97
4. Loris R, Hamelryck T, Bouckaert J, Wyns L (1998) BBA-Protein Struct M 1383:9
5. Rini JM (1995) Annu Rev Biophys Biom 24: 551
6. Hughes RC (1983) Glycoproteins. Chapman & Hall, New York
7. Gabius HJ (1997) Eur J Biochem 243:543
8. Weis WI, Drickamer K (1996) Annu Rev Biochem 65:441
9. Mitchell DA, Fadden AJ, Drickamer K (2001) J Biol Chem 276:28939
10. Varki A, Angata T (2006) Glycobiology 16:1R
11. Drickamer K (1993) Current Opin Struct Biol 3:393
12. Day AJ (1994) Biochem Soc Trans 22:83
13. Drickamer K (1988) J Biol Chem 263:9557
14. Bezouska K, Crichlow GV, Rose JM, Taylor ME, Drickamer K (1991) J Biol Chem 266:11604
15. Weis WI, Drickamer K (1994) Structure 2:1227
16. Holmskov U, Malhotra R, Sim RB, Jensenius JC (1994) Immunol Today 15:67
17. Drickamer K (2002) Biochem Soc Symp 69:59–72
18. Wu K, Yuan J, Lasky LA (1996) J Biol Chem 271:21323
19. Weis WI, Drickamer K, Hendrickson WA (1992) Nature 360:127
20. Weis WI, Kahn R, Fourme R, Drickamer K, Hendrickson WA (1991) Science 254:1608
21. Ng KKS, Iobst ST, Weis WI, Drickamer K (1996) J Biol Chem 271:663
22. Lee RT, Ichikawa Y, Fay M, Drickamer K, Shao MC, Lee YC (1991) J Biol Chem 266:4810
23. Drickamer K (1997) Structure 5:465
24. Drickamer K (1992) Nature 360:183
25. Iobst ST, Drickamer K (1996) J Biol Chem 271:6686
26. Kolatkar AR, Weis WI (1996) J Biol Chem 271:6679
27. Kolatkar AR, Leung AK, Isecke R, Brossmer R, Drickamer K, Weis WI (1998) J Biol Chem 273:19502
28. Burrows L, Iobst ST, Drickamer K (1997) Biochem J 324:673
29. Blanck O, Iobst ST, Gabel C, Drickamer K (1996) J Biol Chem 271:7289
30. Ng KKS, Weis WI (1997) Biochemistry 36:979
31. Torgersen D, Mullin NP (1998) J Biol Chem 273:6254
32. Turner MW (1996) Immunol Today 17:532
33. Tabona P, Mellor A, Summerfield JA (1995) Immunology 85:153
34. Kuhlman M, Joiner K, Ezekowitz RA (1989) J Exp Med 169:1733
35. Ezekowitz RAB, Kuhlman M, Groopman JE, Byrn RA (1989) J Exp Med 169:185
36. Hartshorn KL, Sastry K, White MR, Anders EM, Super M, Ezekowitz RA, Tauber AI (1993) J Clin Invest 91:1414
37. Thiel S, Vorup-Jensen T, Stover CM, Schwaeble W, Laursen SB, Poulsen K, Willis AC, Eggleton P, Hansen S, Holmskov U, Reid KBM, Jensenius JC (1997) Nature 386:506
38. Ma Y, Uemura K, Oka S, Kozutsumi Y, Kawasaki N, Kawasaki T (1999) PNAS 96:371
39. Summerfield JA, Ryder S, Sumiya M, Thursz M, Gorchein A, Monteil MA, Turner MW (1995) Lancet 345:886
40. Garred P, Pressler T, Madsen HO, Frederiksen B, Svejgaard A, Hoiby N, Schwartz M, Koch C (1999) J Clin Invest 104:431
41. Valdimarsson H, Vikingsdottir T, Bang P, Saevarsdottir S, Gudjonsson JE, Oskarsson O, Christiansen M, Blou L, Laursen I, Koch C (2004) Scand J Immunol 59:97
42. Springer TA (1994) Cell 76:301
43. Kansas GS (1996) Blood 88:3259
44. Larsen E, Celi A, Gilbert GE, Furie BC, Erban JK, Bonfanti R, Wagner DD, Furie B (1989) Cell 59:305

45. Phillips ML, Nudelman E, Gaeta FCA, Perez M, Singhal AK, Hakomori Si, Paulson JC (1990) Science 250:1130
46. Walz G, Aruffo A, Kolanus W, Bevilacqua M, Seed B (1990) Science 250:1132
47. Foxall C, Watson SR, Dowbenko D, Fennie C, Lasky LA, Kiso M, Hasegawa A, Asa D, Brandley BK (1992) J Cell Biology 117:895
48. Yuen CT, Lawson AM, Chai W, Larkin M, Stoll MS, Stuart AC, Sullivan FX, Ahern TJ, Feizi T (1992) Biochemistry 38:9126
49. Simanek EE, McGarvey GJ, Jablonowski JA, Wong CH (1998) Chem Rev 98:833
50. Etzioni A, Frydman M, Pollack S, Avidor I, Phillips ML, Paulson JC, Gershoni-Baruch R (1992) N Engl J Med 327:1789
51. Sturla L, Etzioni A, Bisso A, Zanardi D, De Flora G, Silengo L, De Flora A, Tonetti M (1998) FEBS Letters 429:274
52. Frenette PS, Wagner DD (1997) Thromb Haemostas 78:60
53. Brinkman-van der Linden ECM, de Haan PF, Havenaar EC, van Dijk W (1998) Glycoconjugate J 15:177
54. Mulligan MS, Paulson JC, DeFrees S, Zheng Z, Lowe JB, Ward PA (1993) Nature 364:149
55. Borges E, Tietz W, Steegmaier M, Moll T, Hallman R, Hamann A, Vestweber D (1997) J Exp Med 185:573
56. Buerke M, Weyrich AS, Zheng Z, Gaeta RFC, Forrest MJ, Lefer AM (1994) J Clin Invest 93:1140
57. Graves BJ, Crowther RL, Chandran C, Rumberger JM, Li S, Huang KS, Presky DH, Familletti PC, Wolitzky BA, Burns DK (1994) Nature 367:532
58. Weis WI (1997) Curr Opin Struct Biol 7:624
59. Erbe DV, Watson SR, Presta LG, Wolitzky BA, Foxall C, Brandley BK, Lasky LA (1993) J Cell Biology 120:1227
60. Kogan TP, Revelle BM, Tapp S, Scott D, Beck PJ (1995) J Biol Chem 270:14047
61. Cooke RM, Hale RS, Lister SG, Shah G, Weir MP (1994) Biochemistry 33:10591
62. Brandley B, Kiso M, Abbas S, Nikrad P, Srivasatava O, Foxall C, Oda Y, Hasegawa A (1993) Glycobiology 3:633
63. Evans DG, Karjalainen TK, Evans DJ, Jr., Graham DY, Lee CH (1993) J Bacteriol 175:674
64. Jacobs AA, Simons BH, de Graaf FK (1987) Embo J 6:1805
65. Morschhauser J, Hoschutzky H, Jann K, Hacker J (1990) Infect Immun 58:2133
66. Ludwig DS, Holmes RK, Schoolnik GK (1985) J Biol Chem 260:12528
67. Sixma TK, Pronk SE, Kalk KH, Wartna ES, van Zanten BAM, Witholt B, Hoi WGJ (1991) Nature 351:371
68. Merritt EA, Sarfaty S, Akker FVD, L'Hoir C, Martial JA, Hol WGJ (1994) Protein Sci 3:166
69. Skinner MP, Lucas CM, Burns GF, Chesterman CN, Berndt MC (1991) J Biol Chem 266:5371
70. Ushiyama S, Laue TM, Moore KL, Erickson HP, McEver RP (1993) J Biol Chem 268:15229
71. Lobb RR, Chi-Rosso G, Leone DR, Rosa MD, Bixler S, Newman BM, Luhowskyj S, Benjamin CD, Dougas IG, Goelz SE, Hession C, Chow EP (1991) J Immunol 147:124
72. Lee YC (1993) Biochem Soc Trans 21:460
73. DeFrees SA, Gaeta FCA (1993) J Am Chem Soc 115:7549
74. DeFrees SA, Kosch W, Way W, Paulson JC, Sabesan S, Halcomb RL, Huang DH, Ichikawa Y, Wong CH (1995) J Amer Chem Soc 117:66
75. Welply JK, Abbas SZ, Scudder P, Keene JL, Broschat K, Casnocha S, Gorka C, Steininger C, Howard SC, Schmuke JJ, Graneto M, Rotsaert JM, Manger ID, Jacob GS (1994) Glycobiology 4:259
76. Turunen JP, Majuri ML, Seppo A, Tiisala S, Paavonen T, Miyasaka M, Lemstrom K, Penttila L, Renkonen O, Renkonen R (1995) J Exp Med 182:1133
77. Renkonen O, Toppila S, Penttila L, Salminen H, Helin J, Maaheimo H, Costello CE, Turunen JP, Renkonen R (1997) Glycobiology 7:453
78. Stahn R, Schafer H, Kernchen F, Schreiber J (1998):311
79. Shinohara Y, Hasegawa Y, Kaku H, Shibuya N (1997) Glycobiology 7:1201
80. Rosen SD, Bertozzi CR (1994) Cur Op in Cell Biol 6:663
81. Lasky L (1992) Science 258:964
82. Moore KL, Stults NL, Diaz S, Smith DF, Cummings RD, Varki A, McEver RP (1992) J Cell Biology 118:445
83. Wilkins PP, McEver RP, Cummings RD (1996) J Biol Chem 271:18732
84. Norgard KE, Moore KL, Diaz S, Stults NL, Ushiyama S, McEver RP, Cummings RD, Varki A (1993) J Biol Chem 268:12764
85. Lasky LA, Singer MS, Dowbenko D, Imai Y, Henzel WJ, Grimley C, Fennie C, Gillett N, Watson SR, Rosen SD (1992) Cell 69:927
86. Imai Y, Singer MS, Fennie C, Lasky LA, Rosen SD (1991) J Cell Biology 113:1213

87. Stroud MR, Handa K, Salyan ME, Ito K, Levery SB, Hakomori S, Reinhold BB, Reinhold WN (1996) Biochemistry 35:758
88. Stroud MR, Handa K, Salyan ME, Ito K, Levery SB, Hakomori S, Reinhold BB, Reinhold VN (1996) Biochemistry 35:770
89. Levinovitz A, Muhlhoff J, Isenmann S, Vestweber D (1993) J Cell Biology 121:449
90. Lenter M, Levinovitz A, Isenmann S, Vestweber D (1994) J Cell Biology 125:471
91. Steegmaier M, Levinovitz A, Isenmann S, Borges E, Lenter M, Kocher HP, Kleuser B, Vestweber D (1995) Nature 373:615
92. Croce K, Freedman SJ, Furie BC, Furie B (1998) Biochemistry 37:16472
93. Mehta P, Cummings RD, McEver RP (1998) J Biol Chem 273:32506
94. Nicholson MW, Barclay AN, Singer MS, Rosen SD, van der Merwe PA (1998) J Biol Chem 273:763
95. Varki A (1997) J Clin Inves 99:158
96. Kawasaki T, Ashwell G (1976) J Biol Chem 251:1296
97. Ashwell G, Harford J (1982) Annu Rev Biochem 51:531
98. Hudgin RL, Pricer WE, Jr., Ashwell G, Stockert RJ, Morell AG (1974) J Biol Chem 249: 5536
99. Breitfeld PP, Simmons CFJ, Strous GJ, Geuze HJ, Schwartz AL (1985) Int Rev Cytol 97:47
100. Spiess M (1990) Biochemistry 29:10009
101. Lodish HF (1991) Tibs 16:374
102. Drickamer K, Mamon JF, Binns G, Leung JO (1984) J Biol Chem 259:770
103. Lee YC, Townsend RR, Hardy MR, Lonngren J, Arnarp J, Haraldsson M, Lonn H (1983) J Biol Chem 258:199
104. Townsend RR, Hardy MR, Wong TC, Lee YC (1986) Biochemistry 25:5716
105. Lee YC, Lee RT, Rice K, Ichikawa Y, Wong TC (1991) Pure & Appl Chem 63:499
106. Rice KG, Weisz OA, Barthel T, Lee RT, Lee YC (1990) J Biol Chem 265:18429
107. Chiu MH, Tamura T, Wadhwa MS, Rice KG (1994) J Biol Chem 269:16195
108. Chiu MH, Thomas HV, Stubbs HJ, Rice KG (1995) J Biol Chem 270:24024
109. Wadhwa MS, Rice KG (1995) J Drug Target 3:111
110. Ishibashi S, Hammer RE, Herz J (1994) J Biol Chem 269:27803
111. Braun JR, Willnow TE, Ishibashi S, Ashwell G, Herz J (1996) J Biol Chem 271:21160
112. Ii M, Kurata H, Itoh N, Yamashina I, Kawasaki T (1990) J Biol Chem 266:11295
113. Ozaki K, Lee RT, Lee YC, Kawasaki T (1995) Glycoconj J 12:268
114. Lehrman MA, Pizzo SV, Imber MJ, Hill RL (1986) J Biol Chem 261:7412
115. Lehrman MA, Hill RL (1986) J Biol Chem 261:7419
116. Lehrman MA, Haltiwanger RS, Hill RL (1986) J Biol Chem 261:7426
117. Hoyle GW, Hill RL (1988) J Biol Chem 263:7487
118. Hoyle GW, Hill RL (1991) J Biol Chem 266:1850
119. Haltiwanger RS, Lehrman MA, Eckhardt AE, Hill RL (1986) J Biol Chem 261:7433
120. Haltiwanger RS, Hill RL (1986) J Biol Chem 261:7440
121. Abdullah M, Kierszenbaum AL (1989) J Cell Biol 108:367
122. Kawasaki T, Ashwell G (1977) J Biol Chem 252:6536
123. Mellow TE, Halberg D, Drickamer K (1988) J Biol Chem 263:5468
124. Lee RT, Rice KG, Rao NBN, Ichikawa Y, Barthel T, Piskarev V, Lee YC (1989) Biochemistry 28:8351
125. Loeb JA, Drickamer K (1987) J Biol Chem 262:3022
126. Verrey F, Drickamer K (1993) Biochem J 292 (Pt 1):149
127. Steer CJ, Osborne JC, Jr., Kempner ES (1990) J Biol Chem 265:3744
128. Lee RT, Yang GC, Kiang J, Bingham JB, Golgher D, Lee YC (1994) J Biol Chem 269:19617
129. Taylor ME, Bezouska K, Drickamer K (1992) J Biol Chem 267:1719
130. Taylor ME, Drickamer K (1993) J Biol Chem 268:399
131. Mullin NP, Hitchen PG, Taylor ME (1997) J Biol Chem 272:5668
132. Hitchen PG, Mullin NP, Taylor ME (1998) Biochem J 333:601
133. Mengeling BJ, Manzella SM, Baenziger JU (1995) PNAS 92:502
134. Manzella S, Dharmesh S, Beranek M, Swanson P, Baenzinger J (1995) J Biol Chem 270:21665
135. Fiete DJ, Beranek MC, Baenziger JU (1998) Proc Natl Acad Sci U S A 95:2089
136. Engering A, Geijtenbeek TBH, van Vliet SJ, Wijers M, van Liempt E, Demaurex N, Lanzavecchia A, Fransen J, Figdor CG, Piguet V, van Kooyk Y (2002) J Immunol 168:2118

137. Geijtenbeek TBH, van Vliet SJ, Koppel EA, Sanchez-Hernandez M, Vandenbroucke-Grauls CMJE, Appelmelk B, van Kooyk Y (2003) The J Exp Med 197:7
138. Wang QC, Feng ZH, Nie QH, Zhou YX (2004) Chinese Med J 117:1395
139. Alvarez CP, Lasala F, Carrillo J, Muniz O, Corbi AL, Delgado R (2002) J of Virology 76:6841
140. Halary F, Amara A, Lortat-Jacob H, Messerle M, Delaunay T, Houles C, Fieschi F, Arenzana-Seisdedos F, Moreau J-F, Dechanet-Merville J (2002) Immunity 17:653
141. Kornfeld S (1992) Annu Rev Biochem 61:307
142. Hille-Rehfeld A (1995) BBA - Rev Biomembranes 1241:177
143. Dahms NM (1996) Biochem Soc Trans 24:136
144. Dahms NM, Hancock MK (2002) BBA - Gen Subjects 1572:317
145. Westlund B, Dahms NM, Kornfeld S (1991) J Biol Chem 266:23233
146. Dahms NM, Rose PA, Molkentin JD, Zhang Y, Brzycki MA (1993) J Biol Chem 268:5457
147. Wendland M, Waheed A, von Figura K, Pohlmann R (1991) J Biol Chem 266:2917
148. Roberts DL, Weix DJ, Dahms NM, Kim JJP (1998) Cell 93:639
149. Quiocho FA, Vyas NK (1984) Nature 310:381
150. Tong PY, Kornfeld S (1989) J Biol Chem 264:7970
151. Tong PY, Gregory W, Kornfeld S (1989) J Biol Chem 264:7962
152. Marron-Terada PG, Brzycki-Wessell MA, Dahms NM (1998) J Biol Chem 273:22358
153. Bourne Y, Bolgiano B, Liao DI, Strecker G, Cantau P, Herzberg O, Feizi T, Cambillau C (1994) Nat Struct Biol 1:863
154. Li M, Distler JJ, Jourdian GW (1990) Arch Biochem Biophys 283:150
155. Waheed A, Hille A, Junghans U, von Figura K (1990) Biochemistry 29:2449
156. Dahms N, Kornfeld S (1989) J Biol Chem 264:11458
157. York SJ, Arneson LS, Gregory WT, Dahms NM, Kornfeld S (1999) J Biol Chem 274:1164
158. Punnonen EL, Wilke T, Von Figura K, Hille-Rehfeld A (1996) Eur J Biochem 237:809
159. Neufeld EF (1991) Annu Rev Biochem 60:257
160. Pohlmann R, Boeker MWC, Von Figura K (1995) J Biol Chem 270:27311
161. Munier-Lehmann H, Mauxion F, Bauer U, Lobel P, Hoflack B (1996) J Biol Chem 271: 15166
162. Sleat DE, Lobel P (1997) J Biol Chem 272:731
163. Ludwig T, Ovitt CE, Bauer U, Hollinshead M, Remmler J, Lobel P, Ruther U, Hoflack B (1993) Embo J 12:5225
164. Wang ZQ, Fung MR, Barlow DP, Wagner EF (1994) Nature 372:464
165. Ludwig T, Eggenschwiler J, Fisher P, D'Ercole AJ, Davenport ML, Efstratiadis A (1996) Develop Biol 177:517
166. Zaina S, Squire S (1998) J Biol Chem 273:28610
167. Yamada T, De Souza AT, Finkelstein S, Jirtle RL (1997) Proc Natl Acad Sci 94:10351
168. Hankins GR, De Souza AT, Bentley RC, Patel MR, Marks JR, Iglehart JD, Jirtle RL (1996) Oncogene 12:2003
169. Kong FM, Anscher MS, Washington MK, Killian JK, Jirtle RL (2000) Oncogene 19:1572
170. Gemma A, Hosoya Y, Uematsu K, Seike M, Kurimoto F, Yoshimura A, Shibuya M, Kudoh S (2000) Lung Cancer 30:91
171. Barondes SH, Castronovo V, Cooper DNW, Cummings RD, Drickamer K, Felzi T, Gitt MA, Hirabayashi J, Hughes C, Kasai K-i (1994) Cell 76:597
172. Barondes SH, Cooper DNW, Gitt MA, Leffler H (1994) J Biol Chem 269:20807
173. Lobsanov YD, Rini JM (1997) Trends Glycosci Glycotechnol 9:145
174. Kasai K-i, Hirabayashi J (1996) J Biochem (Tokyo) 119:1
175. Leffler H (1997) Trends Glycosci Glycotechnol 9:9
176. Leffler H, Carlsson S, Hedlund M, Qian Y, Poirier F (2002) Glycoconjugate J V19:433
177. Liao D, Kapadia G, Ahmed H, Vasta GR, Herzberg O (1994) PNAS 91:1428
178. Lobsanov YD, Gitt MA, Leffler H, Barondes SH, Rini JM (1993) J Biol Chem 268:27034
179. Seetharaman J, Kanigsberg A, Slaaby R, Leffler H, Barondes SH, Rini JM (1998) J Biol Chem 273:13047
180. Leonidas DD, Vatzaki EH, Vorum H, Celis JE, Madsen P, Acharya KR (1998) Biochemistry 37:13930
181. Leonidas DD, Elbert BL, Zhou Z, Leffler H, Ackerman SJ, Acharya KR (1995) Structure 3:1379
182. Branden C, Tooze J (1991) Introduction to Protein Structure. Garland Publishing, New York
183. Brewer CF (1997) Trends Glycosci Glycotechnol 9:155
184. Gitt MA, Wiser MF, Leffler H, Herrmann J, Xia YR, Massa SM, Cooper DN, Lusis AJ, Barondes SH (1995) J Biol Chem 270:5032

185. Gitt MA, Jordan ET, Leffler H (1997) Trends Glycosci Glycotechnol 9:87
186. Gitt MA, Colnot C, Poirier F, Nani KJ, Barondes SH, Leffler H (1998) J Biol Chem 273:2954
187. Hadari YR, Eisenstein M, Zakut R, Zick Y (1997) Trends Glycosci Glycotechnol 9:103
188. Hsu DK, Zuberi RI, Liu FT (1992) J Biol Chem 267:14167
189. Massa SM, Cooper DN, Leffler H, Barondes SH (1993) Biochemistry 32:260
190. Charcot JM, Robin C (1853) C R Mem Soc Biol 5:44
191. Zhou Z, Tenen DG, Dvorak AM, Ackerman SJ (1992) J Leukoc Biol 52:588
192. Perillo NL, Marcus ME, Baum LG (1998) J Mol Med 76:402
193. Colnot C, Ripoche MA, Scaerou F, Foulis D, Poirier F (1996) Biochem Soc Trans 24:141
194. Colnot C, Ripoche MA, Milon G, Montagutelli X, Crocker PR, Poirier F (1998) Immunology 94:290
195. Magnaldo T, Fowlis D, Darmon M (1998) Differentiation 63:159
196. Wada J, Kanwar YS (1997) J Biol Chem 272:6078
197. Matsumoto R, Matsumoto H, Seki M, Hata M, Asano Y, Kanegasaki S, Stevens RL, Hirashima M (1998) J Biol Chem 273:16976
198. Osmand AP, Friedenson B, Gewurz H, Painter RH, Hofmann T, Shelton E (1977) Proc Natl Acad Sci USA 74:739
199. Steel DM, Whitehead AS (1994) Immunol Today 15:81
200. Gewurz H, Zhang XH, Lint TF (1995) Curr Opin Immunol 7:54
201. Goodman AR, Cardozo T, Abagyan R, Altmeyer A, Wisniewski HG, Vilcek J (1996) Cytokine Growth F R 7:191
202. Srinivasan N, White HE, Emsley J, Wood SP, Pepys MB, Blundell TL (1994) Structure 2:1017
203. Tillett WS, Francis T, Jr. (1930) J Exp Med 52:561
204. Emsley J, White HE, O'Hara BP, Oliva G, Srinivasan N, Tickle IJ, Blundell TL, Pepys MB, Wood SP (1994) Nature 367:338
205. Hohenester E, Hutchinson WL, Pepys MB, Wood SA (1997) J Med Biol 269:570
206. Tennent GA, Pepys MB (1994) Biochem Soc Trans 22:74
207. Zahedi K (1997) J Biol Chem 272:2143
208. Kottgen E, Hell B, Kage A, Tauber R (1992) J Immunol 149:445
209. Tsui CC, Copeland NG, Gilbert DJ, Jenkins NA, Barnes C, Worley PF (1996) J Neuroscience 16:2463
210. Noland TD, Friday BB, Maulit MT, Gerton GL (1994) J Biol Chem 269:32607
211. Reid MS, Blobel CP (1994) J Biol Chem 269:32615
212. Introna M, Alles VV, Castellano M, Picardi G, De Gioia L, Bottazzi B, Peri G, Breviario F, Salmona M, De Gregorio L, Dragani TA, Srinivasan N, Blundell TL, Hamilton TA, Mantovani A (1996) Blood 87:1862
213. Armstrong PB, Swarnakar S, Srimal S, Misquith S, Hahn EA, Aimes RT, Quigley JP (1996) J Biol Chem 271:14717
214. Tennent GA, Butler PJG, Hutton T, Woolfitt AR, Harvey DJ, Rademacher TW, Pepys MB (1993) Eur J Biochem 214:91
215. Robey FA, Liu TY (1981) J Biol Chem 256:969
216. Williams AF, Barclay AN (1988) Ann Rev Immunol 6:381
217. Schiffer M, Girling RL, Ely KR, Edmundson AB (1973) Biochemistry 12:4620
218. Kelm S, Pelz A, Schauer R, Filbin MT, Tang S, de Bellard ME, Schnaar RL, Mahoney JA, Hartnell A, Bradfield P, Crocker PR (1994) Curr Biol 4:965
219. Crocker P, Clark E, Filbin M, Gordon S, Jones Y, Kehrl J, Kelm S, Le Douarin N, Powell L, Roder J, Schnaar R, Sgroi D, Stamenkovic K, Schauer R, Schachner M, van den Berg T, van der Merwe P, Watt S (1998) Glycobiology 8:0v
220. Crocker PR (2002) Curr Opin Struct Biol 12:609
221. May AP, Robinson RC, Vinson M, Crocker PR, Jones EY (1998) Molecular Cell 1:719
222. van der Merwe PA, Crocker PR, Vinson M, Barclay AN, Schauer R, Kelm S (1996) J Biol Chem 271:9273
223. Vinson M, Kelm S, May AP, Jones EY, Crocker PR (1996) J Biol Chem 271:9267
224. Schauer R, Kamerling JP (1997) In: Montreuil J, Vliegenhart J, Schachter H (eds) Glycoproteins II (Glycoproteins), vol 29b. Elsevier, Amsterdam, p 243
225. Irie A, Koyama S, Kozutsumi Y, Kawasaki T, Suzuki A (1998) J Biol Chem 273: 15866
226. Varki A (1992) Glycobiology 2:25
227. Powell LD, Varki A (1994) J Biol Chem 269:10628
228. Collins BE, Yang LJS, Mukhopadhyay G, Filbin MT, Kiso M, Hasegawa A, Schnaar RL (1997) J Biol Chem 272:1248

229. Collins BE, Kiso M, Hasegawa A, Tropak MB, Roder JC, Crocker PR, Schnaar RL (1997) J Biol Chem 272:16889
230. Kelm S, Brossmer R, Gross HJ, Strenge K, Schauer R (1998) Eur J Biochem 255:663
231. Knibbs RN, Osborne SE, Glick GD, Goldstein IJ (1993) J Biol Chem 268:18524
232. Dulac C, Tropak MB, Cameron-Curry P, Rossier J, Marshak DR, Roder J, Le Douarin NM (1992) Neuron 8
233. Sgroi D, Nocks A, Stamenkovic I (1996) J Biol Chem 271:18803
234. Razi N, Varki A (1998) Proc Natl Acad Sci USA 95:7469
235. Tropak MB, Roder JC (1997) J Neurochem 68:1753
236. Crocker PR, Mucklow S, Bouckson V, McWilliam A, Willis AC, Gordon S, Milon G, Kelm S, Bradfield P (1994) EMBO J 13:4490
237. Cornish AL, Freeman SD, Forbes G, Ni J, Zhang M, Cepeda M, Gentz R, Augustus M, Carter KC, Crocker PR (1998) Blood 92:2123
238. Cyster JG, Goodnow CC (1997) Immunity 6:509
239. O'Keefe TL, Williams GT, Davies SL, Neuberger MS (1996) Science 274:798
240. Freeman SD, Kelm S, Barber EK, Crocker PR (1995) Blood 85:2005
241. Quarles RH (1997) J Mol Neurosci 8:1
242. Li C, Tropak MB, Gerial R, Clapoff S, Abramow-Newerly W, Trapp B, Peterson A, Roder J (1994) Nature 369:747
243. Montag D, Giese KP, Bartsch U, Martini R, Lang Y, Bluthmann H, Karthigasan J, Kirschner DA, Wintergerst ES, Nave KA, Zielasek J, Toyka KV, Lipp HP, Schachner M (1994) Neuron 13:229
244. Connolly NP, Jones M, Watt SM (2002) Brit J Haematol 119:221

12.4 Multivalency in Protein–Carbohydrate Recognition

Laura L. Kiessling, *Travis Young*, *Todd D. Gruber*, *Kathleen H. Mortell*
Department of Chemistry, University of Wisconsin, 1101 University Avenue, Madison, WI 53706, USA
kiessling@chem.wisc.edu

In: *Glycoscience*. Fraser-Reid B, Tatsuta K, Thiem J (eds)
Chapter-DOI 10-1007/978-3-540-30429-6_64:

Abstract

The binding of proteins to individual glycoconjugates is often weak. In physiological settings, carbohydrate–protein interactions often benefit from multivalency—multiple copies of the carbohydrate epitope engage with multiple copies of a carbohydrate–binding protein. This mode of molecular recognition is especially important in cell adhesion because the individual interactions need to be kinetically labile yet highly specific. This review discusses why protein–carbohydrate interactions are especially well-suited for multivalent binding, the binding modes and energetic forces that underlie multivalent protein–carbohydrate complexes, and the utility of multivalent carbohydrate derivatives for exploring multivalent binding and inhibiting physiologically and medically important protein–carbohydrate interactions.

Keywords

Carbohydrates; Polyvalency; Molecular recognition; Lectins; Cell adhesion

Abbreviations

ASGR	asialoglycoprotein receptor
ATRP	atom-transfer radical polymerization
bFGF	basic fibroblast growth factor
BSA	bovine serum albumin
ConA	concanavalin A
ELLA	enzyme-linked lectin inhibition assay
EcorL	*Erythrina corallodendrum*
FRET	Förster resonance energy transfer
LOL I	*Lathyrus ochrus* lectin I
LT	labile enterotoxin
LFA	*Limax flavus*
MBPs	mannose-binding proteins
PAMAM	polyamidoamine
ROMP	ring-opening metathesis polymerization
SLT	Shiga-like toxins
SPR	surface plasmon resonance

1 Introduction

Protein–carbohydrate recognition events mediate numerous physiological processes including fertilization, pathogen-cell adhesion, and the inflammatory response [1,2,3,4,5]. Given the

specificity of these biological recognition events, it is surprising that monovalent protein–carbohydrate interactions are so weak. A general perception is that high affinity complexation is required for physiological recognition processes ($K_D \sim 10^{-9}$ M), yet monovalent protein–carbohydrate interactions typically occur with low affinity ($K_D \sim 10^{-3}$ M). Despite the physiological importance of such low affinity interactions, most studies of biologically relevant molecular recognition events focus on high affinity interactions. Here, we will focus on low affinity but multivalent protein–carbohydrate interactions. Multivalency (multiple binding epitopes interact with multiple receptors) is a key mode of recognition that operates in essential physiological recognition processes. Multivalent binding is prevalent in protein–carbohydrate interactions.

1.1 Characteristics of Protein–Carbohydrate Interactions

To understand the importance and functional roles of multivalent interactions, background in the interactions involved in monovalent protein–carbohydrate binding is valuable. X-ray crystallographic studies have served to illuminate the molecular interactions that govern carbohydrate binding by proteins [6,7,8,9,10]. Extracellular carbohydrate binding sites are generally shallow and solvent-exposed. Thus, the proteins make only a handful of direct contacts to the ligand. A brief summary of the types of interactions seen in carbohydrate–protein complexes will be presented as a prelude to an in-depth discussion of multivalent binding.

1.1.1 Hydrogen Bonds

With the high density of hydroxy groups displayed by carbohydrates, hydrogen bonding is expected to play an important role in their recognition. Although the energetic contributions of hydrogen bonds to protein–carbohydrate binding have been debated, it is clear that the directionality of hydrogen bonds is critical to the specificity of carbohydrate recognition [11]. One example of hydrogen bond-mediated specificity control is found in the complex of galectin-1, a mammalian lectin that specifically binds lactose derivatives, with *N*-acetyllactosamine [12]. The protein uses arginine, asparagine, and histidine residues to form direct hydrogen bonds with the axial 4-position hydroxyl group of the galactose residue (❯ *Fig. 1*). Thus, the protein exploits several contacts to a single hydroxyl group to differentiate between galactose and related carbohydrates (e. g., glucose or mannose) that have an equatorial hydroxyl group at the 4-position. Carbohydrate functional groups that are not critical for discriminating between related carbohydrate epitopes are often solvent-exposed [6].

The recognition of carbohydrates by lectins is also faciliated by the high density of potential hydrogen-bonding partners in the former. Vicinal hydroxyl groups in both the equatorial–axial and equatorial–equatorial configurations are spaced about 2.8 Å apart, such that bidentate interactions with two atoms of the same amino acid residue are favorable. In several complexes determined by X-ray crystallography, galactose 3- and 4-hydroxyl groups are observed to hydrogen bond to a single aspartic acid side chain carboxylate group (❯ *Fig. 2*). In addition to direct interactions between amino acid residues and saccharide hydroxyl groups, indirect, water-mediated hydrogen bonds are common (❯ *Sect. 1.1.5*) [13,14].

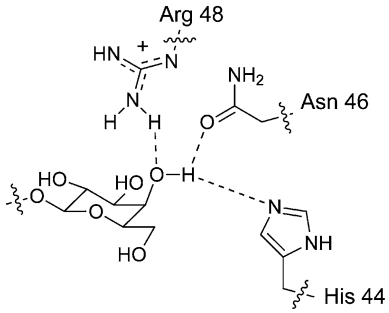


Figure 1
Putative recognition of 4-position hydroxyl of *N*-acetyl-lactosamine through hydrogen-bonding interactions observed in the X-ray structure of galectin-1 from bovine spleen. (PDB entry 1slt [12]). Other protein–ligand contacts are omitted for clarity

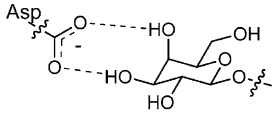


Figure 2
Bidentate hydrogen-bonding interaction with an aspartate sidechain and the 3 and 4 position hydroxyl groups of galactose

1.1.2 Hydrophobic Effects

Although carbohydrates have many polar functional groups, many mono- and oligosaccharides are amphiphilic. Thus, they present significant nonpolar surface area that can interact with complementary hydrophobic amino acid residues. The directionality of hydrogen bonds formed in the binding site serves to orient the carbohydrate so that the complementary surfaces are aligned [15]. Indeed, aromatic amino acid side chains are observed to interact with bound sugars in numerous structures determined by X-ray crystallographic analysis. For example, the α-face of β-galactose presents a continuous nonpolar surface from carbons three through six, and the α-face is observed to pack against aromatic side chains in lectin–galactose complexes (*Fig. 3*).

As noted by Weis and Drickamer the aliphatic protons of the sugar ring bear a small partial positive charge, which could lead to weak, favorable interactions with the π-cloud of aromatic residues in an analogous fashion to edge-face stacking of aromatic residues in protein structures [6,16,17]. In accord with the hydrophobicity of lectin binding sites, carbohydrates bearing aromatic aglycones are often bound with significantly greater affinity than the corresponding glycosides lacking such hydrophobic substituents [18]. Indeed, a screen for inhibitors of the dendritic cell lectin DC-SIGN resulted in the identification of aromatic small molecules that do not resemble its carbohydrate ligands [19].

An additional role for aromatic groups in carbohydrate binding may be to stabilize the *exo*-anomeric effect [20]. The *exo*-cyclic O1 lone pair of a pyranose (e. g., glucose) can donate electron density into the C1–O5 σ^* orbital while decreasing electron density in the C1–O5

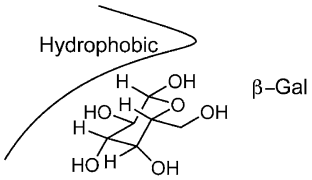


Figure 3
The α-face of galactose is nonpolar

bond (see *Fig. 7*); this preference for the gauche glycosidic conformation, is referred to as the *exo*-anomeric effect [20,21]. As indicated by the resonance structure, the partial positive character on C1 would be stabilized by cation-π interactions with aromatic amino acids in the binding site [20]. It is an intriguing possibility that this stabilization is the source of the prevalence of aromatic amino acid side chains in the carbohydrate-binding sites of lectins.

1.1.3 The Role of Calcium Ions

Calcium ions are required for activity of many families of carbohydrate-binding proteins. The most common function for calcium ions in lectins is structural; the calcium ion can serve to orient important protein functional groups for ligand coordination. For example, in legume lectins it is common for conserved amino acid residues (e. g., asparagine and aspartic acid) to simultaneously coordinate a calcium ion and participate in hydrogen-bonding interactions with the bound saccharide [22,23]. The mammalian *C*-type lectins employ direct contacts between calcium and hydroxyl groups of the bound saccharide. In the structure of rat serum mannose binding protein, a C-type lectin, the binding site calcium ion is coordinated by the equatorial hydroxy groups at C3 and C4 of mannose [24] (*Fig. 4*). Ligand exchange for calcium(II) complexes is rapid [25], suggesting saccharide binding events that occur with direct ligand–calcium coordination could benefit from fast association rates.

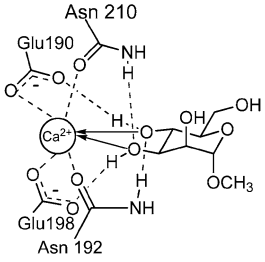


Figure 4
Mannose forms direct contents to a protein coordinated calcium ion in complex with MBP-C

1.1.4 Coulombic Interactions

Many carbohydrate-binding proteins discriminate between potential ligands through Coulombic interactions. Certain biologically important saccharides, such as sialyl Lewis^x, other

sialic acid derivatives, and heparin, contain anionic groups that are recognized by their respective receptors. Crystallographic studies of a heparin-derived saccharide bound to basic fibroblast growth factor (bFGF) suggest that several anionic sulfate substituents of the heparin fragment make ion pairing or direct hydrogen-bonding contacts with protein side chains [26,27,28,29,30]. The anionic substituents of heparin are also partially solvent exposed, interacting both with bFGF residues and solvent. Another illustrative example of the contributions that Coulombic interactions can make to the association of carbohydrates and proteins is found in the structure of sialyl Lewis^x bound to P-selectin [22]. A salt bridge is formed between the C1 carboxylate of sialic acid and a protein arginine, enabling binding and recognition. Structure–function studies performed in other systems also highlight the importance of anionic carbohydrate substituents in binding their protein targets. One system, involving L-selectin-mediated recruitment of leukocytes to sites of tissue damage, will be discussed in detail in *Sect. 2*.

1.1.5 The Role of Water

The ability of water to act as both a hydrogen bond donor and acceptor makes it ideal for interacting with the dense display of hydroxyl groups furnished by carbohydrates. Moreover, numerous X-ray crystallographic studies demonstrate that water plays an important role in the recognition of carbohydrates. Indeed, 20 water molecules are involved in the binding of *Lathyrus ochrus* lectin I (LOL I) to the linear trisaccharide Man(α1-3)Man(β1-4)GlcNAc [31]. Comparison of complexes containing different sugars bound to the same lectin often shows the same water-mediated hydrogen bonds between protein and carbohydrate [6].

Though valuable, crystallographic studies do not reveal the energetic contributions of water-mediated hydrogen bonds to protein–carbohydrate complexes. The energetic impetus for complex formation has been debated, and Lemieux has argued that the driving force for carbohydrate complexation is the reorganization of water molecules upon binding [14]. In this model, solvation of the amphiphilic carbohydrate and the protein binding pocket perturbs the low energy organization of the bulk water structure. Release of some of these higher energy water molecules to the bulk upon binding accounts for the favorable enthalpy of binding [32,33]. This rationalization was stimulated by extensive calorimetric studies of the thermodynamics of saccharide epitope binding by lectins which revealed enthalpic- entropic compensation as the hydrophobicity of the carbohydrate ligands was varied. The data suggest that hydrophobic patches on the carbohydrate surface raise the energy of water at the surface and release of this disordered water to bulk is enthalpically favorable but results in a decrease in entropy. The disruption of the more ordered water structure about polar hydroxyl groups on ligand binding is enthalpically unfavorable but entropically favorable. According to this model it is the reorganization of water molecules which provides the driving force for protein–carbohydrate complexation. A detailed understanding of the role of solvation in protein-saccharide interactions remains elusive. Studies on carbohydrate ligands binding to the lectin concanavalin A (ConA) have added additional insight [34], although the complex results of a further concanavalin A study led to the suggestion that the specific details of the binding site dictated energetic outcomes, and that empirical rules regarding water displacement were not predictive [35].

1.1.6 Conformational Flexibility

Oligosaccharides possess varying degrees of conformational flexibility. The conformational flexibility of the constituent monosaccharides is typically low, with the common pyranose residues residing primarily in a single conformation. Pyranoses with more than one axial substituent, such as iduronic acid, tend to be more flexible due to steric congestion present in all conformers. The conformational flexibility of iduronic acid has been suggested to be important for the recognition of heparin by heparin-binding proteins (*Fig. 5*) [36]. For most oligosaccharides, however, the source of conformational flexibility is associated with rotation about the glycosidic linkages [37,38].

Linear oligosaccharides often exhibit the characteristics of multiple conformers. The conformational flexibility of oligosaccharides suggests that there is a loss in conformational entropy upon binding; therefore, conformationally constrained analogs may be bound with higher affinity. It is difficult to dissect the loss of conformational entropy on ligand binding from other thermodynamic parameters; however, the conformational restriction of inter-saccharide bond rotamers upon binding has been estimated to carry a conformational entropy penalty as large as 1–2 kcal/mol [39].

Despite predictions that restricting flexibility will be beneficial, it has proved difficult to generate high affinity inhibitors by restricting conformational entropy. Indeed, the installation of conformational constraints can have complex effects on oligosaccharide binding. The challenges are highlighted in a study in which the thermodynamics of conformationally constrained trisaccharides binding to a well-characterized, monoclonal antibody Se155-4, were determined calorimetrically. X-ray crystallographic analysis of the parent trisaccharide bound to the antibody suggested that tethers could be used to constrain glycosidic bond rotamers without disrupting protein–carbohydrate contacts (*Fig. 6*). Nevertheless, the oligomethylene ether-tethered compounds **4–6** (*Fig. 6*) show weaker overall binding relative to the parent trisaccharide **1** with an unanticipated loss in entropy and a favorable change in enthalpy. The aryl dithioether tethered trisaccharide, **7**, functions as designed, exhibiting tighter antibody binding due to a favorable entropy term. The magnitude of the affinity increase, however, is diminished by an unfavorable change in the enthalpy of binding relative to the parent ligand. The conformations of the acetal-tethered trisaccharides **2**, **3** (*Fig. 6*) are significantly distorted from the native trisaccharide and are not recognized by the antibody. The conformations of the remaining four constrained trisaccharides appear to be similar to that of the parent structure by both NMR experiments and molecular dynamics simulations. Thus, the origins of the ther-

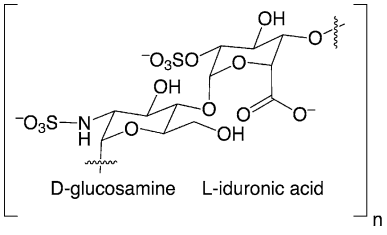


Figure 5
Structure of a disaccharide unit found in heparin and heparan sulfate proteoglycans

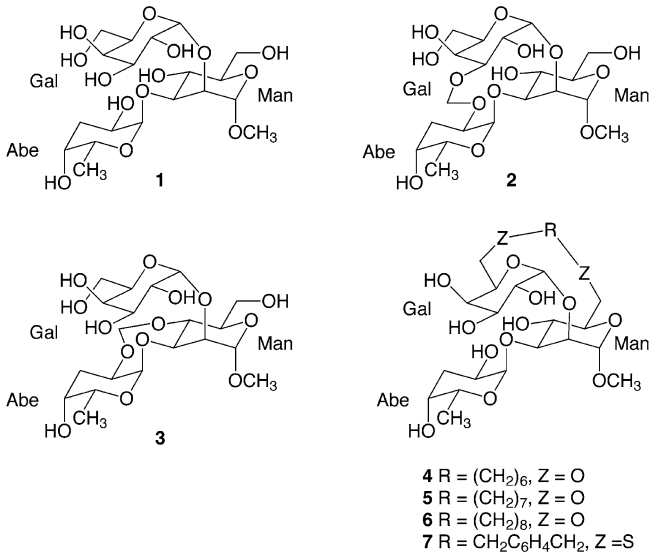


Figure 6
Conformationally constrained trisaccharide derivatives synthesized and studied by Bundle and coworkers

modynamic differences in binding were attributed to the structures of the tethers. These results can be explained by Lemieux's model: Higher energy water structure about the hydrophobic oligomethylene linker in compounds **4–6** could contribute to the favorable enthalpy of binding as these higher energy molecules are released to the bulk water. This study suggests that only small affinity enhancements are to be gained by the preorganization of glycosidic bond rotamers in the design of carbohydrate-based inhibitors [40,41].

Conformational controls are naturally present in oligosaccharides. For example, the *exo*-anomeric effect appears to preorganize oligosaccharides. The *exo*-anomeric effect describes the observed conformational preference of the exocyclic C1–O (glycosidic) bond for a rotamer in which a lone pair of electrons (*Fig. 7*) of the exocyclic oxygen is antiperiplanar to the polar, endocyclic C1–O bond [21,42]. The extent to which the *exo*-anomeric effect is operative and serves to rigidify glycosidic linkages in aqueous solvent has been debated.

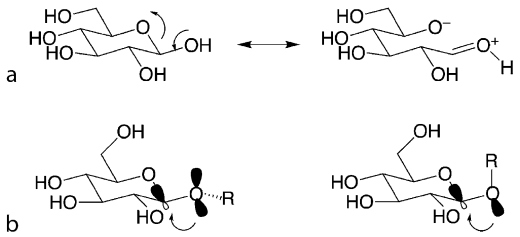


Figure 7
The *exo*-anomeric effect favors conformations in which the O1 oxygen is oriented such that its lone pair is an antiperiplanar orientation to the C1–O bond

However, molecular modeling studies in which the force field parameters are adjusted for the *exo*- anomeric effect tend to yield conformational preferences consistent with those observed in solution. Further, two recent structures of protein–carbohydrate complexes reveal that the bound carbohydrate ligands adopt conformations consistent with the *exo*-anomeric effect—remarkably a lack of electron density between C1 and O5 could be observed [20,43]. Thus, likely due to a combination of steric and electronic factors, many branched oligosaccharides appear to occupy a single low energy conformation in solution. Still, the potential energy wells are predicted to be relatively shallow and broad, suggesting that these oligosaccharides may alter their conformations without significant energetic cost [38,39,44].

1.2 Mechanisms of Multivalent Ligand Binding

Multivalent ligands can engage membrane-bound receptors through multiple different binding modes (*Fig. 8*). Indeed, many naturally occuring lectins and their ligands are multivalent. For example, membrane-bound glycoproteins can serve as scaffolds that present multiple oligosaccharide determinants. Additionally, saccharide-binding receptors often form oligomeric quaternary structures. Thus, an understanding of the various mechanisms that mediate monovalent protein–carbohydrate interactions is necessary but not sufficient. The understanding of complex physiological systems is complicated by the multiplicity of individual ligand-receptor interactions [45,46,47,48,49,50,51], which afford energetic contributions by different mechanisms. There are several distinct mechanisms that contribute to the increases in binding affinity commonly attributed to "multivalency" or "cluster" effects. Briefly, the variables that can be added to the intrinsic affinity of the relevant monovalent protein–saccharide interaction are the sum of individual interactions, the distance and orientation of recognition elements, and the contributions of the scaffold on which these recognition elements are displayed. The sum of these factors can result in an increased *functional affinity* of the multivalent ligand for its cognate receptor. The term *functional affinity* is useful as it describes a measurable quantity, the apparent association constant, for multivalent ligands. An understanding of the contributions that individual mechanisms can make to protein–carbohydrate recognition is required to effectively design multivalent ligands.

1.2.1 Entropy

Consideration of entropic effects is vital to an understanding of multivalent interactions. For multivalent ligands, loss of translational entropy is complete with the binding of the first recognition element; subsequent binding interactions can proceed without paying an additional translational entropy penalty [52,53]. This favorable entropic contribution inherent to multipoint binding is counteracted by the conformational energy cost upon binding of the multidentate ligand. The conformational entropy term is affected by structural features of the multivalent ligand such as the conformational rigidity of the scaffold on which the binding determinants are displayed and the favored orientation of each recognition subunit [48,54,55,56]. It should be noted that the translational entropy penalty for binding membrane bound receptors is not as great as that incurred for a receptor–ligand interaction in solution, because membrane bound molecules diffuse in two dimensions rather than three.

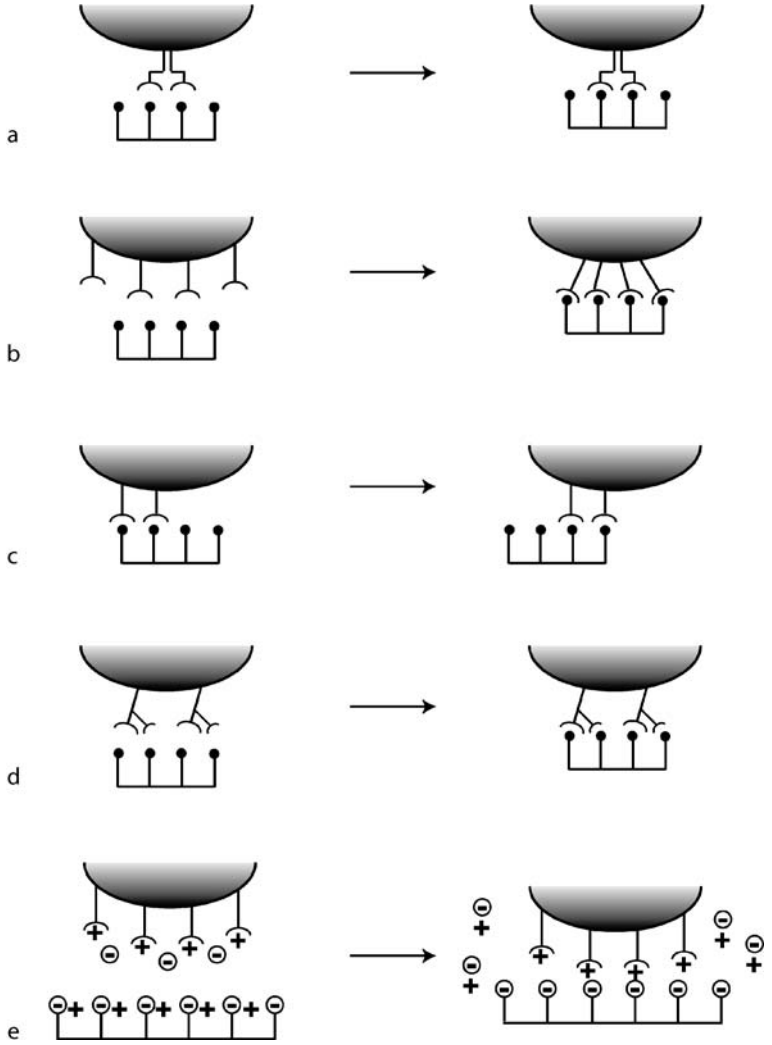


Figure 8
Modes of multivalent ligand binding. Possible molecular mechanisms for increases in functional affinities for multivalent ligands. (a) Interaction of a multivalent ligand with a multivalent receptor (chelate effect). (b) Clustering of receptors by a multivalent ligand. (c) Increased local ligand concentration (slow off-rates due to statistical effects). (d) Subsite binding (occupation of secondary binding sites on a single receptor). (e) The polyelectrolyte effect

1.2.2 The Chelate Effect

The simultaneous interaction of multiple saccharide epitopes of polyvalent ligands with oligomeric receptors can lead to interactions of high functional affinity when the orientation of the saccharide recognition elements corresponds to that of the receptor (*Fig. 8*). The mechanism of affinity increase is often described as the chelate effect [52,53]. This idea is

consistent with models developed for antibody–antigen interactions [57]. Historically, the chelate effect describes the observation that electrophilic metal centers are often coordinated with higher affinity by divalent, covalently linked, electron-donating groups. Thus, the chelate effect describes multiple interactions which, after formation of the first contact, are effectively intramolecular [53]. The observed free energy of binding will be related to the sum of the free energies of binding for each individual receptor–ligand interaction such that the free energy of interaction (ΔG_{obs}) for an oligomeric receptor $R_1 - R_2 \cdots - R_n$ with an oligomeric ligand, $L_1 - L_2 \cdots - L_n$ can be estimated by the equation:

$$\Delta G_{obs} = \Delta G(R_1 - L_1) + \Delta G(R_2 - L_2) + \cdots + \Delta G(R_n - L_n)$$

This model predicts that the formation of additional contacts should give exponential affinity increases, but fails to take into account an entropy cost related to the conformational flexibility of the tethers linking the binding epitopes [52]. Loss of conformational entropy can diminish the functional affinity enhancement derived from chelation. In a system in which the saccharide epitopes are favorably oriented, high functional affinities for multivalent receptor–ligand interactions can be achieved for relatively small complexes with only a few individual protein–carbohydrate interactions. The potential of chelate effect-mediated affinity enhancement has been demonstrated by the binding of a non-natural trimer of the glycopeptide antibiotic vancomycin to a trimer of its dipeptide ligand D-Ala-D-Ala; this complex has an affinity similar to that of the avidin–biotin interaction [58]. The design of this artificial multivalent interaction was facilitated by the fact that both binding epitopes are relatively small molecules and the spacing between recognition elements could be controlled.

1.2.3 Receptor Clustering

The chelate effect is distinct from multivalent binding modes that require receptor clustering [59]. Since the cell membrane is fluid, receptors are free to diffuse in two dimensions; this mobility can lead to clustering of receptors. In such a system, the unliganded state of the saccharide-binding protein may not have the same distribution on the cell surface as the bound species. This clustering occurs with an entropy cost, but it does not require that the multidentate saccharide display match precisely the display of receptors on the cell surface.

1.2.4 Concentration Effects

A dense array of receptors and/or ligands can lead to an increase in observed affinity through mechanisms other than the chelate effect. For example, a high effective concentration of recognition epitopes and/or receptors can result in high functional affinity. This concept can be demonstrated by modeling the binding event as a two-step process. The first event is the formation of a transitory encounter complex that can either proceed to form the final, stable ligand receptor complex or dissociate in an unproductive event. In the case of a multivalent ligand, the probability that the encounter complex will arrive at a productive receptor–ligand interaction is higher due to an increased local concentration of ligand. The structural requirements for multidentate saccharide derivatives to bind to oligomeric lectins with increased affinity due to concentration effects are not as stringent as those for classical chelate effect-mediated affinity enhancement [60].

1.2.5 Extended and Secondary Binding Sites

Lectins can bind multidentate or large saccharides with enhanced affinity due to interactions of the ligand with recognition elements outside the primary saccharide binding site [61]. Ligand substituents can interact with a second thermodynamically independent binding site or through extension of the primary binding site. In the first case, distinct secondary binding sites may offer interactions of a different character to those present in the primary interactions. In the second case, groups that are not recognized by the lectin can independently form van der Waals or water-mediated contacts to the protein. The distinction between the carbohydrate binding site and recognition elements outside it is not always clearly defined because binding sites are shallow [6,9]. However, structural studies of certain carbohydrate binding proteins have revealed distinct, secondary saccharide binding sites [61,62]. Although the physiological relevance of such sites remains obscure, multivalent ligands with the ability to occupy both sites simultaneously will be more effective inhibitors. The secondary binding sites may endow lectins not only with higher binding affinities for selected oligosaccharides, but with the ability to discriminate between multivalent ligands by binding site or saccharide orientation.

1.2.6 Polyelectrolyte Effect

As previously discussed for monovalent binding, some naturally occurring, multivalent saccharide derivatives possess multiple anionic groups. Although specific anionic groups are important in certain systems, the charge density associated with such ligands can influence their binding behavior. As in the case of protein–DNA association, the polyelectrolyte effect plays an important role [63]; the entropically favorable release of counterions into solution as the protein binds provides a driving force for association of the two macromolecules. Physiological mucin-like ligands, which display multiple copies of sulfated saccharide epitopes, may rely on similar effects for their interaction with the selectins. The polyelectrolyte effect is distinct from the specific Coulombic interactions discussed in *Sect. 1.1.4*; however, ligands that display the multiple charged groups required for specific interaction with a receptor will also likely benefit from this effect.

1.2.7 Evolutionary Considerations

Monosaccharides are bound with high affinity by some proteins [10]. In contrast, extracellular proteins typically bind monovalent carbohydrate epitopes with low affinity; they exhibit marked preferences for binding to multidentate saccharide ligands. What advantages are conferred by multivalent binding?

Because cell–cell recognition events mediated by protein–carbohydrate interactions occur on the cell surface, it is necessary to consider the effect of an array of ligand receptor interactions. Multivalent interactions allow mechanisms other than the intrinsic affinity of a recognition epitope for a receptor to control recognition events. If a tight interaction is observed between a monomeric receptor–ligand pair, multiple copies of this interaction will also bind tightly, leading to a lack of specificity. Multivalency can facilitate discrimination between structurally similar recognition epitopes when monovalent binding events are of low affinity [54,64,65]. Consider a monovalent ligand that is bound by a given receptor with some affinity: A small change in the ligand structure might diminish the affinity of the interaction

by one order of magnitude. In contrast, multivalent arrays of the two ligands can be much more selective: A trivalent receptor could, in principle, exhibit a 10^3 difference in functional affinity for trivalent displays of the two ligands. Thus, multivalent binding can be highly selective [51,64,65,66].

Multivalent systems allow for high functional affinity but exhibit unique kinetics. For example, they exhibit greater reversibility of binding in the presence of competing ligands [58]. In contrast, for monovalent binding events, the dissociation is determined solely by the rate constant for dissociation and is not affected by the addition of competing ligands. Their kinetic lability confers additional flexibility on processes mediated by multivalency. For example binding events composed of multiple weak interactions are expected to be more resistant to shear stress than a single tight interaction [67]; this feature is likely to be significant for some cell–cell recognition processes.

Finally, in multivalent systems, changes in the spacing or orientation of recognition elements could alter the nature of an interaction, without changes in the recognition elements themselves [54,55]. Indeed, the epitope density of a multivalent display has been shown to affect receptor cluster stoichiometry and rate of formation [54], and the architecture of the multivalent scaffold has been shown to determine whether a multivalent display acts as an inhibitor or effector [55]. Thus, different types of glycoconjugates might bind the same lectin yet elicit very different physiological responses.

2 Inhibitors of Multivalent Interactions

We discuss various ligands and their functions as well as provide some interpretation of the mechanisms underlying their function. A variety of different scaffolds has been used to present multiple copies of recognition epitopes. The structure of the scaffold can have a marked effect on its mode of multivalent binding [55], so the features of a ligand can be chosen for a specific purpose (e. g., clustering the target receptor or occupying a subsite). We have grouped these into the following classes: low molecular weight ligands, dendrimers, large spherical displays, and linear polymeric scaffolds. This review is not intended to be comprehensive, rather, specific examples have been chosen to represent a larger body of work.

2.1 Low-Molecular-Weight Ligands

Small templates that present 2–10 saccharide epitopes can exhibit enhanced affinity for target proteins. The small size of such multivalent ligands allows for characterization of the spatial relationships between relevant epitopes. Studies on recognition events with ligands of this type have provided interesting results. However, templates of this type present recognition epitopes in a manner that differs significantly from common, natural systems in valency and geometry. It can be useful to think of these small molecules as mimicking single, multiantennary saccharide epitopes (◉ *Fig. 9a*), but not the larger displays furnished by glycoproteins (◉ *Fig. 9b*). Thus, to mimic glycoporteins, different types of multivalent ligands are required. Intriguingly, there are have been exciting examples in which relatively small molecules can participate in multivalent binding.

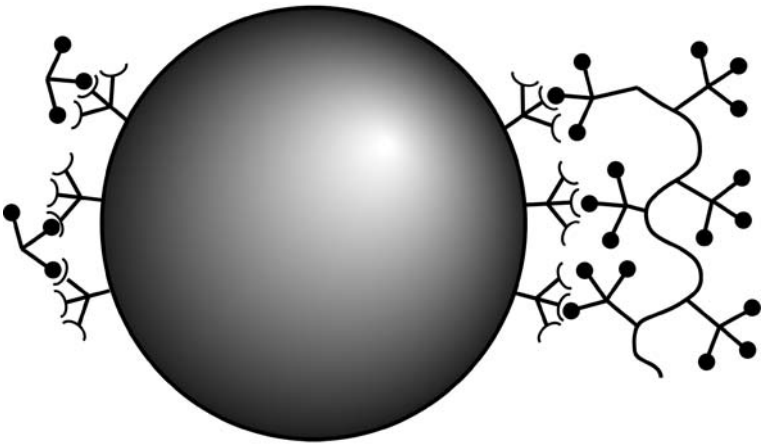


Figure 9
Two distinct types of multivalent binding used by natural systems. (a) Multiantennary saccharide derivatives bind multivalently. (b) Glycoproteins display several epitopes and can span large distances

2.1.1 Ligands for the Hepatic Lectin

Studies on saccharide recognition by the asialoglycoprotein receptor (ASGR), also known as the hepatic lectin, have yielded insight into the features of multivalent ligands that give rise to enhanced functional affinities. Although the physiological significance of this receptor is not completely clear, the ASGR is an endocytotic receptor found on hepatocytes and is believed to be responsible for the recycling of desialylated proteins that carry triantennary oligosaccharide chains [68,69]. In groundbreaking investigations, Lee and coworkers evaluated the ability of synthetic, peptide-based multivalent carbohydrate displays to bind to the asialoglycoprotein receptor, using templates based on aspartic acid, glutamic acid and aminotris(hydroxymethyl)methane (*Fig. 10*) [65,70].

Variations in the spacing of recognition epitopes, through the use of different linker groups as well as the valency of the template, were studied by monitoring inhibition of a radiolabelled ligand for ASGR, asialoorosomucoid, binding to either isolated receptors or hepatocytes. The synthetic clusters displaying three *N*-acetylgalactosamine residues, in which 6-aminohexylgalactosides were linked to glutamic acid residues through amide bonds, YEE(ah-GalNAc)$_3$, compound **12** (*Fig. 11*), exhibited the best inhibitory potency for rat hepatic lectin. Structurally related divalent derivatives were less potent and the monovalent derivatives showed the poorest activity. Trivalent saccharide displays based on the aminotris(hydroxymethyl)methane template were less effective than those derived from the glutamate or aspartate scaffolds (*Fig. 10*). The galactose bearing template **9** (sug=β-galactose) was only two-fold more potent than a monovalent galactose derivative when binding to solubilized lectin was assayed and approximately 70-fold more active when binding to hepatocytes was evaluated. These modest increases suggest that the saccharide residues on the aminotris(hydroxymethyl)methane are too closely spaced (9 Å maximum) to span more

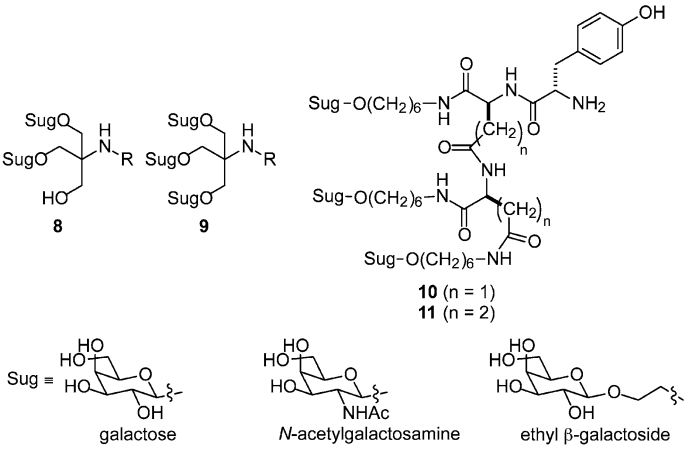


Figure 10
Lee's cluster galactoside inhibitors of the hepatic lectin

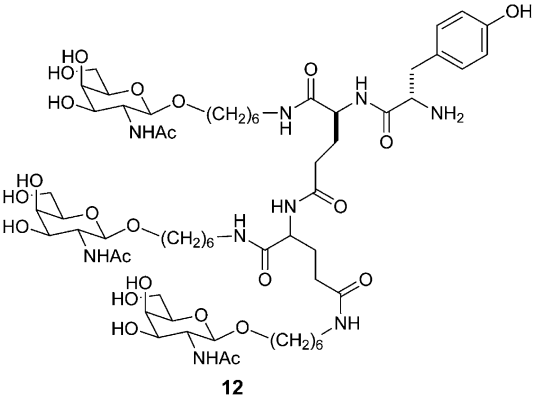


Figure 11
Compound 12, displaying three GalNAc residues, was found to bind tightly to the asialoglycoprotein receptor (ASGR)

than one ASGR carbohydrate binding site. This result is consistent with investigations with triantennary saccharides that suggest that the optimum spacing between galactose residues is 15–30 Å [71,72].

The binding properties of isolated ASGR and the cell surface receptor are different, highlighting the interdependence of receptor organization in the membrane and ligand display. The trivalent aspartic and glutamic acid-based inhibitors **10** and **11** (*Fig. 10*) were significantly more active (1000–10,000-fold) against hepatocytes than were the corresponding monovalent analogs, yet smaller differences (approximately 100-fold) were observed with purified lectin. There are two potential reasons for this difference: the lectin may be organized differently on the cell surface than in the isolated form, or the ligands may bind to more than one lectin on the

cell surface. Because receptor-mediated endocytosis is believed to involve aggregation of cell surface receptors, the triantennary galactose derivatives may trigger endocytosis by oligomerizing the cytoplasmic domains of ASGR subunits. Thus, the role of multivalency, in this case, may be to facilitate receptor-mediated endocytosis [73]. Though how these multivalent ligands function is not known, it is clear from these studies that the presentation of one or two additional carbohydrate epitopes on an appropriate template can have a dramatic effect on the functional affinity of small molecule ligands for membrane-bound lectins.

2.1.2 Ligands for the Mannose Binding Proteins

The substantial binding enhancement observed with small multivalent ligands in the case of the ASGR stimulated efforts to apply similar templates to other lectins. The mannose-binding proteins (MBPs) play a significant role in the innate immune system of human infants [74]. However, the mannose derivative of ligand **11** (*Fig. 10*) showed no appreciable affinity enhancement for rat MBP-A, a homotrimeric protein with a mannose binding site on each subunit. This observation is consistent with the X-ray structural data which show the crystallized trimeric rat MBP-A fragment to have carbohydrate binding sites separated by 53 Å, larger than the maximum distance between ligands displayed on the tripeptide template [75]. The ligands, in this case, cannot exploit inter-oligomer binding, as the target protein is not membrane-bound. For intra-oligomer binding, a template with the required spacing would have to overcome a considerable entropy penalty to simultaneously place recognition epitopes in binding sites separated by such large distances. The ASGR recognizes branching within an oligosaccharide epitope, but the MBPs appear to recognize widely spaced mannosylated epitopes on a glycoprotein or cell surface.

Thus, low molecular weight templates appear to be suitable for only a subset of systems. Finding the right template for binding epitope display is difficult. An additional challenge is that the syntheses of such defined multivalent displays can be rather complex. Some general strategies that address these issues are emerging. Some elegant illustrations of applying modular design principles to generate potent multivalent ligands for the AB_5 are described in the next section.

2.1.3 Ligands for AB_5 Toxins

The AB_5 family of bacterial toxins, which includes cholera, pertussis, Shiga, and Shiga-like toxins (SLT), consists of a single A subunit and a pentagonal arrangement of 5 B subunits [76]. These toxins invade cells through the multivalent interaction of the B subunits with carbohydrate residues of gangliosides. The functional toxin receptor on mammalian cells for SLTs is the trisaccharide of the glycolipid Gb_3 [77]. A dimeric Gb_3 derivative was generated and found to occupy both the primary and secondary binding sites of SLT [78]. The affinity of this dimeric ligand for SLT is 40-fold greater than the monomer [79]. To generate an even more potent ligand, a pentamer was designed, in which the dimeric Gb_3 derivative was appended to a glucose core. The resulting compound, dubbed STARFISH, displays five copies of the trisaccharide dimer (*Fig. 12*). Remarkably, STARFISH is an inhibitor that is > 1,000,000-fold more potent in inhibiting SLT-mediated cell death over the monovalent trisaccharide [79].

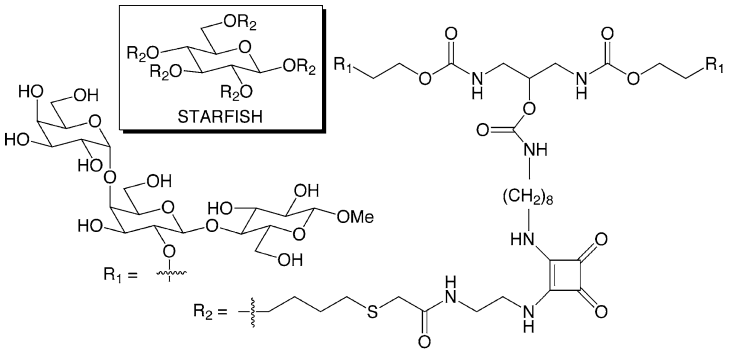


Figure 12
Potent decameric ligands for the AB_5 toxin Shiga-like toxin

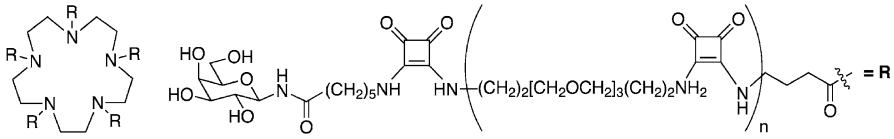


Figure 13
Modular pentavalent ligands for the AB_5 toxin enterotoxin

Although the pentavalent compound was designed to interact one to one with the pentameric SLT, structural studies showed one STARFISH inhibitor binding to two SLT pentamers. These studies again highlight the ability of multivalent ligands to mediate protein assembly [50,55,80,81]. Though rational design was not completely successful at predicting the ultimate mode of action, the resulting multivalent display exhibited tremendous efficacy.

A different study focused on the heat-labile enterotoxin (LT) [82]. A series of pentavalent saccharide derivatives were generated using a modular synthetic approach. These compounds were designed to examine the effect of linker length and overall ligand dimensions on LT binding (*Fig. 13*). The best pentavalent inhibitor is 100,000-fold more active than the monovalent ligand. Importantly, the activity of the compound depends upon the effective linker length—the most effective inhibitor displays sugars separated by distances similar to those of the binding sites within the LT pentamer. The results suggest that the inhibitor saccharides bind simultaneously to all five of the B subunits of the toxin. Further work on similar systems by Fan and coworkers generated additional inhibitors and observed 1:1 binding in the structure of the complex [83,84].

Pieters and coworkers [85] designed cholera toxin inhibitors based on three factors: (1) usage of the authentic receptor oligosaccharide as the monovalent ligand, (2) use of multivalent dendritic scaffolds, and (3) the used of elongated spacer arms of optimized length. This strategy enabled the synthesis of multivalent dendrimers with 47,500-fold tighter binding than monovalent ligand. In a similar system, recent work by Kiick and coworkers [86] demonstrated, using galactose-bearing glycopolypeptides, that correct saccharide spacing, as analyzed from crystallographic structures, was a primary determinant for controlling binding to cholera toxin.

2.1.4 Bivalent Inhibitors of the Selectins

E-Selectin, a lectin involved in mediating the inflammatory response [87,88,89,90,91], has also been targeted with low molecular weight, multidentate ligands [92,93]. Unlike ASGR or influenza virus hemagglutinin, E-selectin is not known to have an oligomeric quaternary structure. Nevertheless, the protein is anchored in the membrane and therefore could participate in multivalent receptor-ligand interactions. To probe this possibility, Wong, Paulson and coworkers explored the dimerization of sialyl Lewis^x (sLe^x) recognition epitopes in the study of multivalent ligand recognition by E-selectin [94]. They prepared dimeric sLe^x derivatives linked by different spacer groups (● *Fig. 14*) and evaluated their activities.

Some of the bivalent ligands were more active than the corresponding monomeric sLe^x derivative **15**. The most potent ligand, **18**, was 5-fold more potent. Bivalent ligands linked by 1,4-butanediol, **13**, or 1,5-pentanediol, **14**, exhibited no increased potency over the pentasaccharide standard. These results emphasize the importance of entropy concerns in the binding of multivalent ligands. Related compounds, for example those with increased hydrophobicity, generated by Miyauchi and coworkers are more effective inhibitors (● *Fig. 15*) [95].

By appending sLe^x epitopes to an aromatic template and functionalizing the amino group of sLe^x with a large, hydrophobic residue **21** (● *Fig. 15*), a compound more effective (8-fold) than the aforementioned dimers was generated. The small increases in functional affinities obtained with these bivalent compounds suggest that they may make favorable contacts outside of the

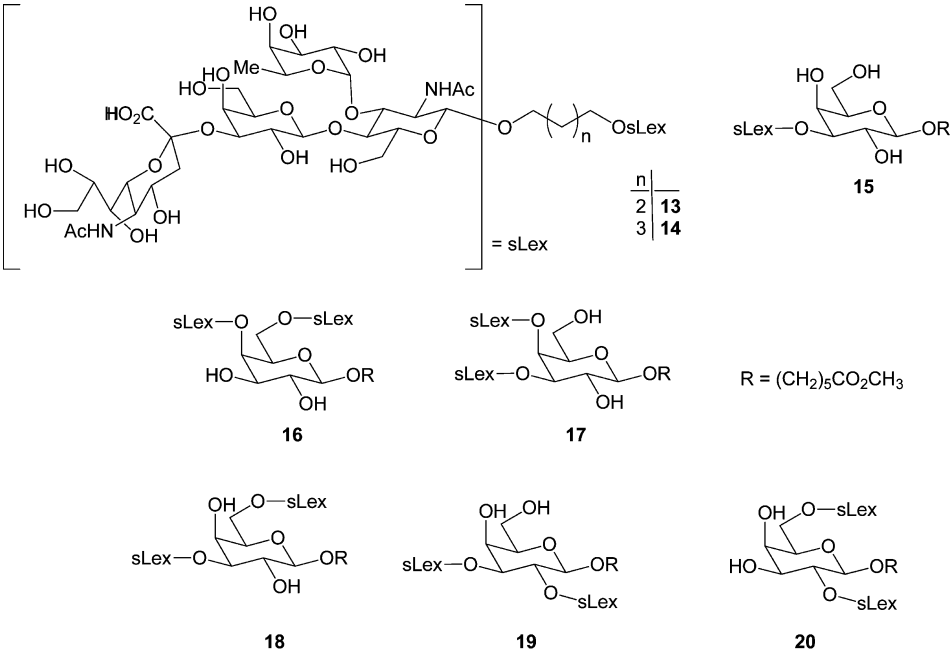
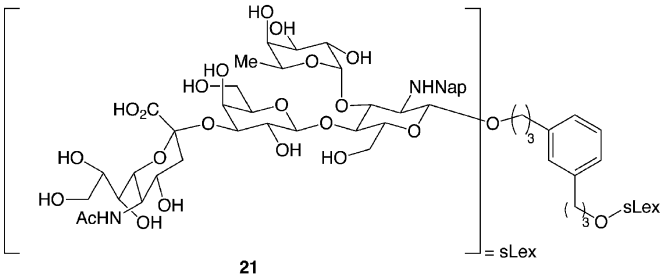


Figure 14
Dimeric sLe^x derivatives with varying linkers



21

Figure 15
Hydrophobic sLe^x dimer with rigid linker generated by Miyauchi and coworkers

binding site rather than simultaneously interacting with two different proteins. These results are consistent with studies suggesting that the physiological ligands are more complex than sLe^x [88,96].

2.1.5 Low-Molecular-Weight Multivalent Displays Can Cluster Proteins

Low molecular weight multivalent displays of ligands can have increased activities relative to monovalent ligands, even when it is clear that the multivalent ligands are too small to span the multiple binding sites of their target lectin [97]. This phenomenon was investigated with using a low molecular weight macrocycle displaying three mannose residues to test binding to the mannose-binding lectin concanavalin A [80]. A trivalent macrocycle (*Fig. 16*) was more potent than a monovalent mannose derivative and, contrasting previous reports of lectin-ligand precipitation [98], was found to function by favorably clustering lectins in solution. Surface plasmon resonance (SPR) was used to demonstrate increased interaction with a liganded surface and Förster resonance energy transfer (FRET) to show clustering in solution. Together, these results demonstrate that low molecular weight ligands, despite their small size, can favorably interact with multiple binding partners.

2.2 Dendrimers as Inhibitors of Protein–Carbohydrate Interactions

Recent developments in the synthesis of dendrimers have stimulated interest in the use of these novel macromolecules. Although many applications have been proposed for dendritic molecules, perhaps one of the most promising is as scaffolds for the study of molecular recognition events [99,100,101,102,103,104]. Dendrimers are oligomeric molecules; however they offer some of the advantages of both polymers and small molecules. Thus, they are capable of displaying a large number of functional groups, yet are monodisperse. The iterative fashion in which dendrimers are synthesized, combined with their relatively small size, results in defined structures which can be characterized readily by using conventional analytical techniques, such as mass spectrometry. It is also possible to rapidly generate a series of dendrimers based on a common, backbone to systematically characterize small differences in recognition of the displayed ligands. The dense display of functional groups by dendrimers has prompted their derivatization with diverse, biologically relevant molecules. The potential of dendrimers

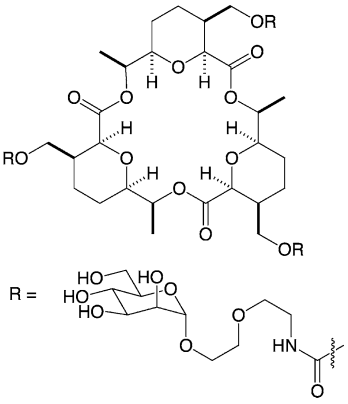


Figure 16
A trivalent mannose derivative that promotes lectin clustering

to act as scaffolds for the multivalent presentation of ligands has sparked the interest of several research groups studying protein–carbohydrate interactions [47,105,106,107,108].

2.2.1 Dendrimers Targeted at Influenza Virus Hemagglutinin

Roy and coworkers have designed inhibitors of influenza virus hemagglutinin based on several different dendrimer backbones [109,110]. Using solid phase syntheses, they generated sialic acid-displaying dendrimers based on a poly-L-lysine core structure (*Fig. 17*) as ligands for influenza virus hemagglutinin. These dendrimers, which bear 2, 4, 8, or 16 pendant sialic

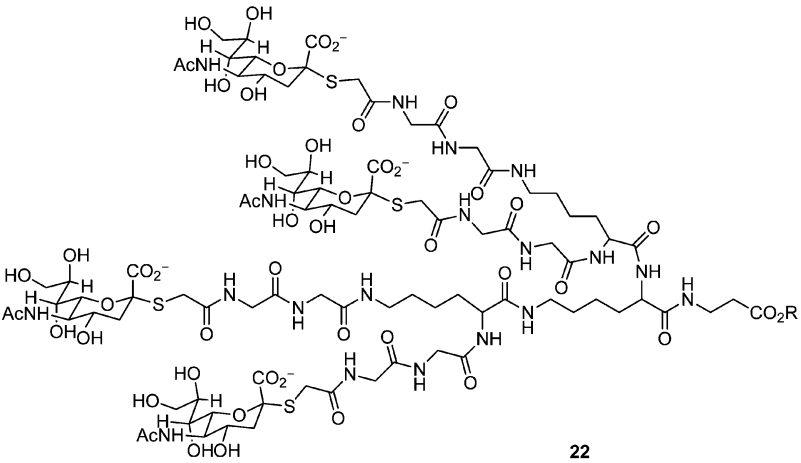


Figure 17
Sialosides displayed on a poly-L-lysine core structure

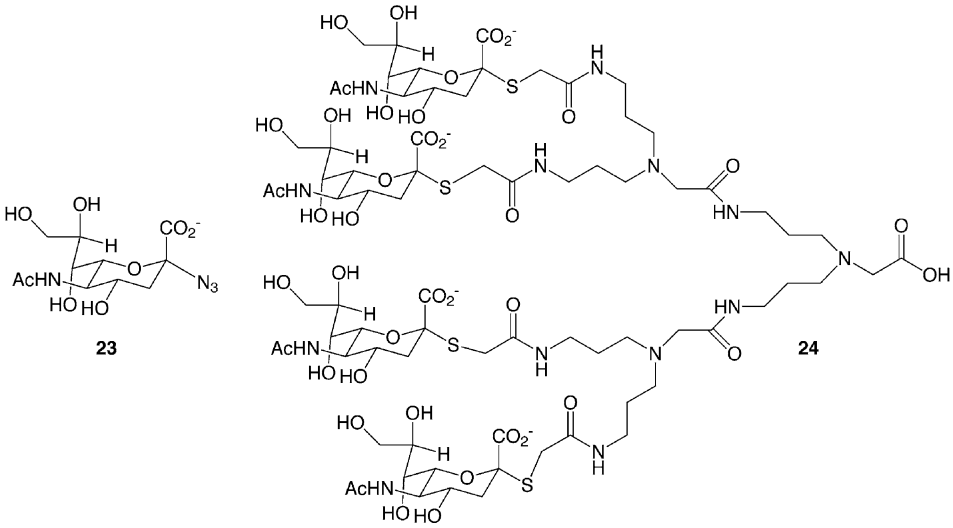


Figure 18
Representative divergent dendrimer, bearing four sialosides

acid residues inhibit viral hemagglutinin-mediated agglutination of erythrocytes at dendrimer concentrations of 625, 312, 156, and 19 μM, respectively. Still, the enhancements in inhibitory potency on a saccharide residue basis (compared to a monosialoside) range from 2-fold to 10-fold. Thus, the dendrimers display only modest enhancements in potency.

Using a competitive enzyme-linked lectin inhibition assay (ELLA), Roy and coworkers compared the activities of structurally related dendrimer displays. These displays varied in the point of branching on the dendritic scaffold. The test assessed the ability of each dendrimer to inhibit the binding of a dimeric lectin from *Limax flavus* (LFA) to the highly sialylated A_1-acid glycoprotein, (10–12.5% sialic acid by weight). Several α-thiosialoside derived from 2 different dendrimer scaffolds (*Fig. 18* and *Fig. 19*) were prepared and evaluated using the aforementioned assay. The divergent dendrimers were prepared in valencies of 2, 4, 8, and 16 (*Fig. 18*) and the radially substituted dendrimers in valencies of 4, 8, and 12 (*Fig. 19*). The dendrimer displaying four thiosialosides was found to be the most efficient inhibitor from either series, with an IC_{50} value of 12 nM. The potency of the dendrimers on a per sialoside basis was 127-fold greater than that of the standard, NeuAcαN_3, **23**. Though this > 100-fold enhancement is significant, not all the dendrimers were more potent. Those of the divergent series displaying 8 and 16 thiosialosides were less efficient on a per sialoside basis than the standard. Interestingly, insoluble complexes were formed when the dendrimers were incubated with LFA, presumably due to crosslinking of the lectin by the inhibitors.

Although some structural features can influence the activity of the dendrimers [111], Roy's studies identified few major differences between dendrimers of different structure. If the dendrimers were acting via the chelate effect, changes in the distance between epitopes would be expected to perturb their functional affinity significantly. As stated above, however, the dendrimers crosslink their target proteins. Because small variations in their structure are unlikely

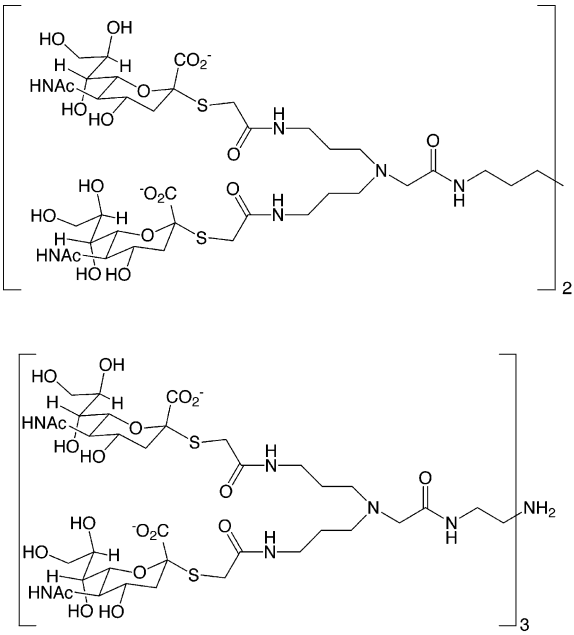


Figure 19
Roy's radially distributed sialoside-bearing dendrimers of valency 4 and 6

to have large effects on their ability to mediate lectin clustering, all of the dendrimers have similar potencies.

In a study comparing dendrimers to other architectures for multivalent display [including low molecular weight displays, globular proteins, and ring-opening metathesis polymerization (ROMP)-derived polymers], dendrimers were found to cluster a different soluble lectin (concanvalin A) [55]. These results suggest that a major mode by which dendrimers function is to cluster lectins. The benefits and drawbacks of clustering by PAMAM dendrimers and other scaffolds have been discussed [50].

2.3 Spherical Displays

2.3.1 Liposomes

Liposomes provide spherical displays with intriguing potential for the polyvalent presentation of carbohydrates (*Fig. 20*). Liposomes, assemblies in which a lipid bilayer surrounds an aqueous core, are particularly interesting due to their potential use as drug delivery agents. The therapeutic applications of lipid vesicles have been studied; it is thought that incorporation of specific recognition elements in the lipid bilayer could allow tissue-targeted drug delivery [112,113].

Studies by the Whitesides group highlighted the ability of saccharide-bearing liposomes to function as efficient inhibitors of protein–carbohydrate recognition events [114]. The lipo-

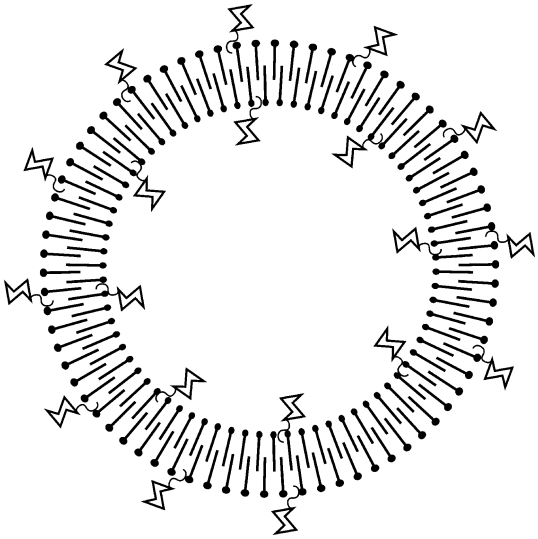


Figure 20
A schematic representation of carbohydrate-bearing liposomes

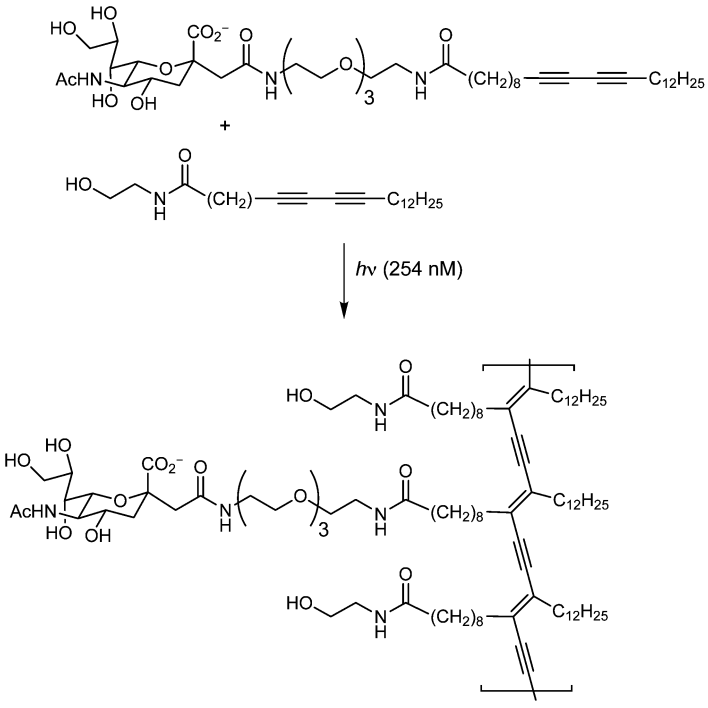


Figure 21
Covalent crosslinking of sialic acid-displaying liposomes

somes are much more potent than those of the low molecular weight multidentate ligands. Within the liposome assemblies, which are held together by noncovalent interactions, the saccharide-substituted lipids can reorganize to bind optimally to the target protein. Covalently crosslinked liposomes also can function as highly effective multivalent displays [115]. Presumably they present ligands at a range of distances and geometries, some of which are complementary to the distribution of the protein receptors on the virus surface. The potencies of the best liposomal inhibitors suggest that these displays may be able to mimic glycoproteins rather than just saccharide epitopes.

In studies by Spevak et al. [115], lipids carrying sialic acid residues attached through *C*-glycosidic linkages were synthesized and crosslinked with varying amounts of unsubstituted lipid to give covalently crosslinked liposomes displaying different concentrations of sialic acid residues (*Fig. 21*) [115]. As in the case of sialic acid-substituted acrylamide polymers (*Sect. 2.4.1*), inhibitory potency was found to increase with the amount of sialic acid displayed on the liposome surface up to a certain concentration; no inhibition of hemagglutination was observed for liposomes carrying 30% or 60% of the sialic acid derivative. This finding emphasizes the importance of the display of charged ligands for multivalent binding and shows that higher density presentations do not necessarily exhibit higher potency. On a saccharide residue basis, liposomes can be highly effective, with minimum inhibitory concentrations of sialic acid of approximately 20 nM. Liposomes present a polyvalent surface array that is quite large; consequently, some of the enhanced inhibitory potency may derive from steric stabilization (*Sect. 2.4.1*).

Liposomal inhibitors of the selectins have been generated that are highly potent inhibitors of these proteins. Liposomes that present multiple copies of the tetrasaccharide sLe^x exhibit affinity enhancements for E-selectin of approximately 500-fold relative to the monomeric sLe^x derivatives [116]. Spevak et al. have incorporated sLe^x derivatives in covalently crosslinked liposomes (*Fig. 22*) [117]. These highly anionic liposomes, prepared with lipid **25**, containing 5% of sLe^x derivatized lipid, inhibit at nanomolar concentrations, the interaction of P-selectin with sLe^x. In contrast, the glycoliposomes prepared with the neutral matrix lipid, **26**, showed no inhibitory activity. Maltose and lactose functionalized lipids, used as controls, were inactive as monomers. However, cross-linked liposomes displaying lactose or maltose prepared with anionic lipid, **25**, were only 10-fold less active than the sLe^x display. These results are consistent with the preference, and apparent broad specificity, of P-selectin for binding highly anionic ligands [118,119,120,121,122,123]. Moreover, they highlight the difficulties of using liposomes as specific inhibitors.

In another example of liposome usage, a multivalent display of a C-3-modified sialoside on liposomes has been shown to inhibit influenza virus replication [124]. Hemagglutinin and sialidase are both involved in both glycoproteins, are involved in attachment and detachment of viral particles on the host cell through recognition of sialyl residues. In this case, the C-3-modified sialoside liposome seems to be such a successful inhibitor of viral replication because it successfully targets two different enzymes that act on multivalent carbohydrate displays.

2.3.2 Resin Bound Ligands

Another venue for saccharide display is on solid support. Resins that have been employed include polystyrene beads, polystyrene modified with polyethylene glycol (TentaGel), and

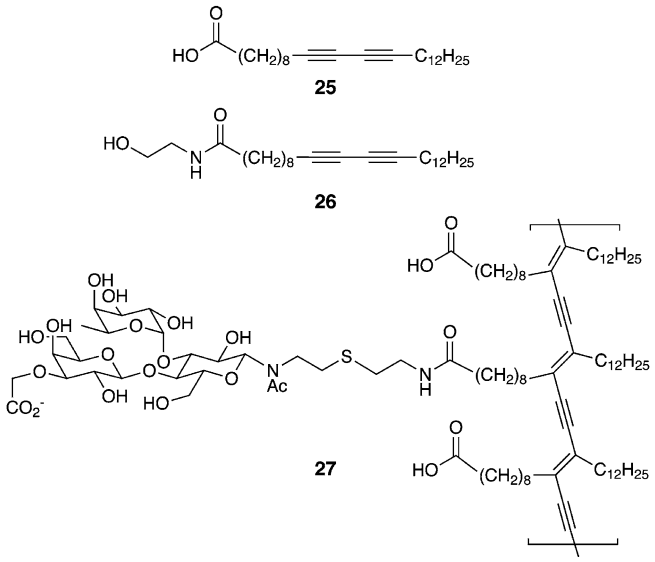
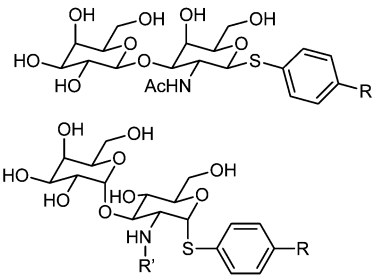


Figure 22
Covalently crosslinked inhibitors of P-selectin

Synsorb. Because several of these supports are also used in solid-phase synthesis, an advantage of using carbohydrate-displaying beads as model systems for polyvalent recognition events is that libraries of saccharides can be assembled. Therefore, structure–function relationships can be determined rapidly.

Kahne and coworkers exploited the opportunity to probe multivalent recognition using libraries of carbohydrate derivatives on TentaGel beads [125,126]. They synthesized an encoded library designed to contain approximately 1300 disaccharide derivatives, including a derivative of the known ligand, Galβ(1-3)GalNAc. The functionalized beads were then assayed by treatment with biotinylated lectin from *Bauhinia purpurea*, incubation with streptavidin- conjugated alkaline phosphatase, and washing and staining of the beads gave a colorimetric readout of the amount of bound lectin. Decoding of the library indicated that compounds other than the known ligand bind to the lectin with high affinity when displayed on TentaGel. To evaluate the potency of the new compounds on a monovalent basis, compounds **28**, **30**, and **31** were synthesized and assayed for their ability to inhibit lectin binding to TentaGel beads derivatized with Galβ1,3GalNAc (◉ *Fig. 23*).

Surprisingly, the binding affinities of the corresponding soluble, monovalent compounds and those displayed on the insoluble resin did not correlate. Compound **28** was the best monovalent inhibitor, while the polyvalent derivative **29** was stained only at high lectin concentrations. Kahne and coworkers suggest the intriguing possibility that biological displays may have unique properties that do not reflect those of the parent monovalent ligand. Although this hypothesis is consistent with the results, other interpretations of the data are possible. For example, it can be difficult to evaluate the efficiency of solid-phase syntheses; thus, the beads could present varying densities of ligand to the lectin. Other studies reveal that differences in the amount of ligand displayed can have dramatic effects on functional affinities [127].



28 R=H, 25 μM
29 R = $OCH_2C(O)NH$-TentaGel

30 R = H, R' = Nitrobenzoyl: 35 μM
31 R = H, R' = Isovaleryl: 46 μM

Figure 23
Disaccharide derivatives identified from a library of compounds, which Kahne and coworkers synthesized on TentaGel beads

Moreover, the resin-bound ligands may interact not only with the targeted binding site on the protein, but also with a secondary binding site. For the library used, several of the functional groups incorporated into the synthesis of the saccharide-like epitopes are not found in naturally occurring saccharides. The disaccharide derivatives studied bear aromatic groups at the anomeric position, as well as hydrophobic acyl groups at the two position, which considerably distort the hydrophobicity of these molecules with respect to the known carbohydrate ligands. Some lectins have aromatic binding sites which are distinct from the saccharide binding regions [18]. Therefore, a secondary site could have a very different character, complicating the interpretation of structure–activity relationships. Furthermore, the resin displaying the ligands may alter the solvation sphere of the ligands, thereby changing the energetics of ligand receptor binding events. Elucidation of the origin of the observed alteration in specificity and its relevance to biological recognition await further experiments.

In addition to resin-bound carbohydrate derivatives, surface-bound carbohydrates or glycoconjugates are also proving valuable in the form of carbohydrate microarrays. Array technology, adapted from cDNA microarray printing techniques, allows the coating of various surfaces with carbohydrate moieties. Thus, the specificity of a lectin for different oligosaccharides or glycoconjugates can be determined in a single of experiment. Indeed, glycan arrays have begun to illuminate the specificities of carbohydrate binding partners with previously unattainable rapidity. This topic has been the subject of several reviews [128,129,130,131,132,133], which provide additional details. Because the surface presentation of the oligosacharide epitopes is multivalent, the glycoconjugate displays can exhibit increased avidity (and perhaps specificity) for the target receptor [134,135]. Moreover, properties of multivalency (such as influence of density, orientation, or pattern) can be studied.

2.4 Linear Polymeric Scaffolds

With the exception of trivalent ligands designed for the ASGR receptor and those that block the AB_5 toxins, efforts to design small multidentate ligands typically result in relatively modest increases in functional affinities. This may be because most small ligand clusters mimic naturally occurring, multi-antennary oligosaccharides, yet the physiological recognition events

occur over large surface areas. In physiological settings, saccharide epitopes can be separated by large distances. For example, multiple copies of carbohydrate ligands can be displayed on glycoprotein backbones. Additionally, multiple copies of a glycoprotein can be presented on the cell surface. Similarly, many copies of any given glycolipid are distributed throughout a membrane. To study the importance of multivalency in a given system, macromolecular displays can be used as scaffolds for the presentation of carbohydrate epitopes. For example, proteins have been used as vehicles for displaying multiple carbohydrate epitopes [136]. The display of carbohydrate ligands on inert carrier proteins such as bovine serum albumin (BSA) can afford useful reagents for the study of multivalent protein–carbohydrate binding in a given system. Still, proteins have fixed dimensions and therefore can span a limited distances. In contrast, polymers can span large distances, and their features, sizes and valencies can be altered systematically. Thus, they serve as excellent probes of multivalent binding events.

2.4.1 Acrylamide Polymers

An efficient method for the synthesis of carbohydrate bearing polymers is the radical polymerization of acrylic acid derivatives [137]. The mild conditions of the radical polymerization and the convenient preparation of the polymerization substrates, make acrylamide polymers suitable for the study of multivalent binding [47,138,139,140,141,142,143]. Several research groups have synthesized acrylamide-based polymers and demonstrated their efficacy in inhibition of influenza virus binding to erythrocytes [144,145,146,147,148]. The demonstration that sialic acid-displaying acrylamide polymers are effective inhibitors of hemagglutinin-mediated viral adhesion has stimulated considerable effort to optimize the application of this strategy.
Polymers incorporating varying mole fractions of sialic acid on a polyacrylamide or polyacrylate backbone have been synthesized by two general strategies (*Fig. 24*). Free radical copolymerization of a sialic acid-bearing acrylamide with acrylamide itself or with acrylamide derivatives affords the target materials. Alternatively, an amine-bearing sialic acid derivative

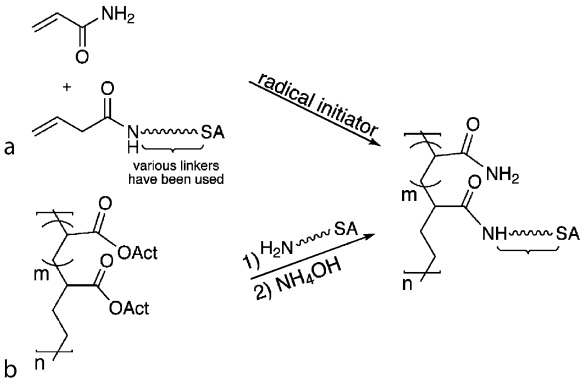


Figure 24
Two strategies for the production of polyacrylamides bearing sugar residues. (a) Copolymerization of acrylamide and a sugar-bearing monomer. (b) Coupling to activated esters on a preformed polymeric backbone. SA = α-*N*-acetyl-neuraminic acid (sialic acid), Act = activating group

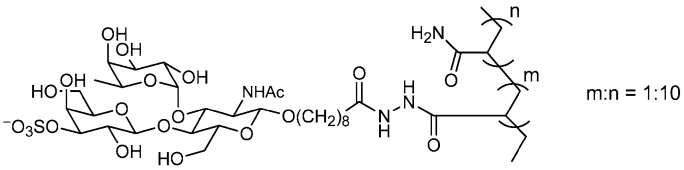
can be coupled to activated esters on a preformed acrylamide backbone [147,148]. Varying monomer reaction rates probably complicate copolymerization by the former strategy, while the latter strategy is expected to result in a random distribution of ligands. A key benefit of the latter strategy is that structure–function relationships in which the ligand structure is varied can be more readily addressed. The polymers prepared by the radical polymerization strategy are large, composed of approximately 2000 monomer units.

Sialic acid-bearing acrylamide polymers prepared by both methods were tested in two in vitro assays of multivalent binding affinity and were found to function as effective multivalent ligands [145,149]. In both cases, the polymers showed large enhancements of inhibitory potency relative to monovalent sialic acid derivatives: Enhancements of 10^3–10^4 were attained for the polymeric sialosides. The inhibitory potency was found to depend significantly on the density of the sialic acid groups displayed on the polymer backbone, with optimal activity at relatively low ratios of sialoside to acrylamide. The optimal proportion varies somewhat and is likely dependent on the sialic acid derivative employed, the method of synthesis, and the assay system used to evaluate the polymers; but sialic acid densities of 10% and 20% have been reported to give the most effective inhibition. These studies contributed to establishing the applicability of synthetic polymeric materials for the study of multivalent molecular recognition events.

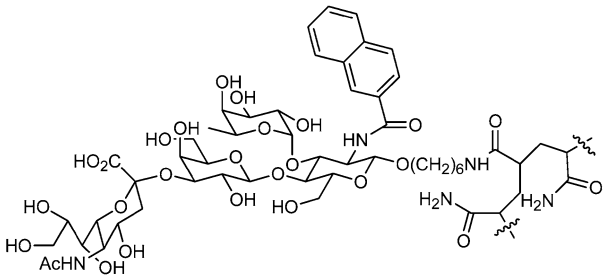
These promising results encouraged efforts to optimize the activity of acrylamide-based displays by modifying polymer structure. The modifications include varying the backbone charge, changing the length of the linker to the sialoside, and copolymerization with monomers bearing aromatic groups. Interestingly, polymers derivatized with *C*-glycosidic analogs of sialic acid exhibited inhibitory potencies equal to their *O*-linked counterparts in a hemagglutination inhibition assay; these results are relevant to the development of carbohydrate-containing therapeutic agents because *C*-glycosides are resistant to in vivo degradation. A highly active of this class is a copolymer resulting from three-components, which bears 20% [145] of a sialic acid *C*-glycoside and 10% benzylamine on a polyacrylamide backbone [149]. The incorporation of benzylamine was inspired by the observation that the most effective monomeric sialosides reported bear hydrophobic, aromatic substituents. Thus, there is a consistency between features that control monovalent binding and those that contribute to multivalent recognition.

Acrylamide polymers derivatized with sLe^x or sLe^x mimetics have been synthesized to target the selectins (◉ *Fig. 25*) [144,150]. Inhibition of glycoprotein binding to immobilized E-selectin by polymers displaying sLe^x derivatives is enhanced approximately 10–100-fold over the potency of monomeric ligands [144]. The effective enhancement of ligands **32** and **33** over their respective monomers is small compared to the increased potencies observed for polymers targeted at influenza virus hemagglutinin. The differences may result from the assays employed. For example, immobilized E-selectin may not mimic the cell surface display of this protein. Alternatively, the differences may result from the organization of the protein receptors. Specifically, viral hemagglutinin exists as a trimer in the membrane, but E-selectin appears to be a monomer.

Syntheses of sialic acid-bearing acrylamide based polymers have established general procedures for generating saccharide-displaying polymers. These methods can be readily applied to multivalent molecular recognition events in other systems. Carbohydrate-substituted polyacrylamides have been shown to be good substrates for glycosidases and glycosyl transferases [151,152], and these enzymes can be used to generate multivalent derivatives of complex carbohydrates to be prepared. Advances in the synthesis of complex oligosaccharide-deriva-



32



33

Figure 25
Acrylamide copolymers, bearing derivatives of sLe^x

tized acrylamide polymers should facilitate future studies of multivalent binding in different receptor systems [153,154].

A disadvantage of traditional acrylamide polymerization reactions is the heterogeneity of the products that result. A radical polymerization method that produces polymers of similar structure but that are much more homogeneous is atom-transfer radical polymerization (ATRP) [155,156]. ATRP has been used to synthesize carbohydrate-substituted polymers with low polydispersities [157,158,159,160,161]. Materials that display sugar residues such as glucofuranose [160], glucopyranose [161], and *N*-acetyl-D-glucosamines [159].

Multivalent Binding and Steric Effects Several research groups have found that polymeric materials are more efficient at inhibiting hemagglutinin-mediated viral adhesion. However, the molecular basis for the observed increase in inhibitory potency is not known. Indeed, antibodies targeted to viral receptors unrelated to hemagglutinin were blocked from binding to their targets, indicating that sites other than the sialic acid receptor were involved [162]. The increased inhibitory efficiency could be attributed to the sum of the effects detailed in ❯ *Sect. 1.2*. Alternatively, the large size of these high molecular weight polymers and their accompanying solvation spheres could, once bound to the erythrocyte surface, sterically obstruct the approach of virus to the cell surface.

To dissect the contributions of classical multivalent effects from the steric mechanism, Whitesides and coworkers demonstrated that when monomeric sialic acid derivatives were added along with the acrylamide polymer, more efficient (on a molar sialic acid basis) inhibition of agglutination was observed [149,162,163]. Addition of the monomeric inhibitors is proposed to have two effects. First, displacement of a polymer-displayed ligand by monomer would lower the affinity of the polymer for the surface of the virus. Second, the size of the polymer not bound to the surface is necessarily increased, and thereby increases the ability of the polymer

to inhibit the approach to the erythrocyte surface. This mechanism is termed "*steric stabilization*". Thus, for the most efficient inhibition, complete binding of polymer-displayed sialic acid epitopes is not necessary or desirable. This steric stabilization mechanism results from a combination of specific (ligand–receptor) interactions and nonspecific steric effects. On the basis of this mechanism, a novel strategy for the inhibition of adhesion events was proposed: Receptors which are not directly involved in adhesion may be used to bind a large polymer to the cell surface which then may occlude, nonspecifically, the approach and adhesion of cells. This mechanism highlights the complexity of the systems involved in multivalency and the difficulty in ascribing affinity enhancements to specific molecular mechanisms without detailed analysis of binding data.

2.4.2 Polyamino Acid Backbones

Several groups have employed polyamino acid backbones as templates for the display of multiple carbohydrate residues [164,165]. For example, potent E-selectin antagonists have been generated from polylysine [166,167] These materials are highly effective inhibitors of leukocyte rolling. Multivalent polylysine displays have also been used to inhibit Shiga-like toxin binding [168].

Highly potent inhibitors of influenza virus have been generated using a polyglutamic acid backbone. Specifically, Kanie, Wong and coworkers modified polyglutamate with lysoganglioside GM_3 (*Fig. 26*) [169]. The polyglutamic acid backbone was chosen for its low toxicity, low immunogenicity and biodegradability. It was anticipated that presenting the saccharide epitope in conjunction with the sphingosine group would afford potent inhibitors of influenza virus-mediated hemagglutination that could function by more than one mechanism. Specifically, the lipid groups could interact with hydrophobic amino acid side-chains of hemagglutinin or insert into the viral membrane. To facilitate binding studies, the polymer was tagged with a fluorescent reporter. They found that the resulting compounds could inhibit the target process at picomolar concentrations.

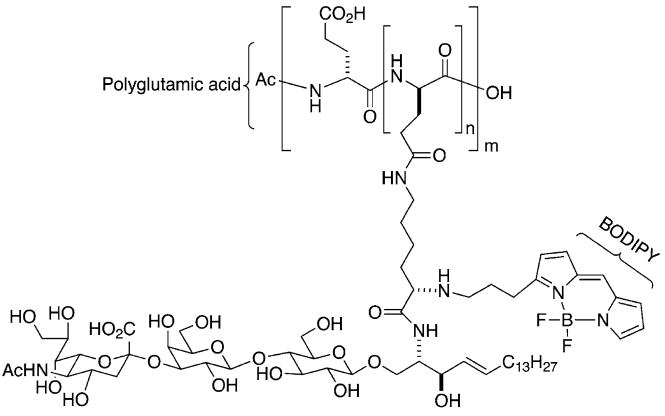


Figure 26
Lysoganglioside GM3-polyglutamic acid conjugate

2.4.3 Polymers Generated by Ring-Opening Metathesis

As described above, acrylamide polymerization reactions are commonly used to generate multivalent carbohydrate derivatives. Still, these reactions produce highly polydisperse mixtures [154], which complicates the interpretation of binding data. The disadvantages posed by standard acrylamide polymerizations have spurred the application of alternative polymerization strategies that can be used to generate more homogenous materials. One such polymerization is the ring-opening metathesis polymerization.

Ruthenium-catalyzed ROMP tolerates highly functionalized, water-soluble monomers. ROMP initiators, such as molybdenum alkylidenes, are effective for oligomerization of protected saccharides [170]. Moreover, ROMP performed under aqueous conditions can be used to produce polymers bearing unprotected carbohydrate substituents [171,172]. Indeed, the ruthenium carbenes are so tolerant that they can oligomerize monomers bearing sulfate groups [121,173]. In addition, initiation can be fast relative to polymer chain propagation, and termination events are generally rare; therefore, this strategy can be used to produce polymers of low polydispersity and defined length [59,174,175,176].

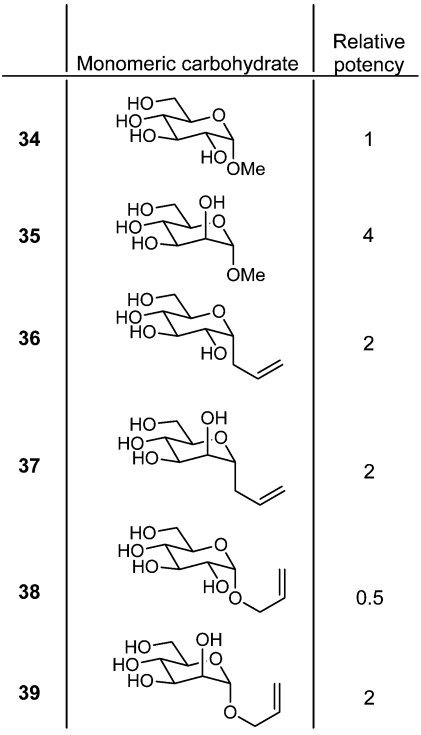
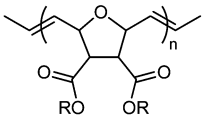


Figure 27
Different monovalent carbohydrate derivatives show similar low affinity for ConA. Relative potency = IC_{50} compound / IC_{50} α-Me-glucoside

Like the dendrimers described in the previous section, polymers generated by ROMP can vary dramatically in size and valency – their size can correspond to the small decavalent ligands discussed in *Sect. 2.1* as well as the large linear acrylamide polymers described in *Sect. 2.3.1* Accordingly, ROMP can give rise to multidentate ligands with saccharide residues that can occupy widely spaced binding sites (e. g., 10–300 Å). The favorable characteristics of ROMP for the synthesis of polymers bearing biologically functional ligands has been exploited [5,56,59,176,177], Mortell et al. first showed that saccharide epitopes appended from the polymer backbone could interact specifically with carbohydrate binding proteins, and that such materials are effective multidentate ligands [171]. In addition they demonstrated the discriminatory potential of multivalent displays of different binding epitopes for their protein targets [64]. The ligand specificity was revealed by assaying mannose- and glucose-derived polymers for the ability to inhibit agglutination of red blood cells by the mannose- and glucose-binding plant lectin, concanavalin A. Carbohydrate-bearing polymers were found to inhibit agglutination at concentrations of up to 50,000-fold lower than monomeric ligands on a saccharide residue basis (*Fig. 27*).

To test the specificity of ligand interaction, the activities of a monovalent glucoside and mannoside with those of their multivalent counterparts were compared (*Fig. 27* and *Fig. 28*).



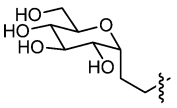
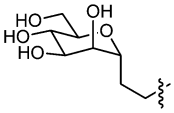
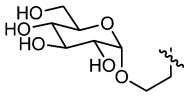
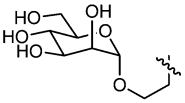
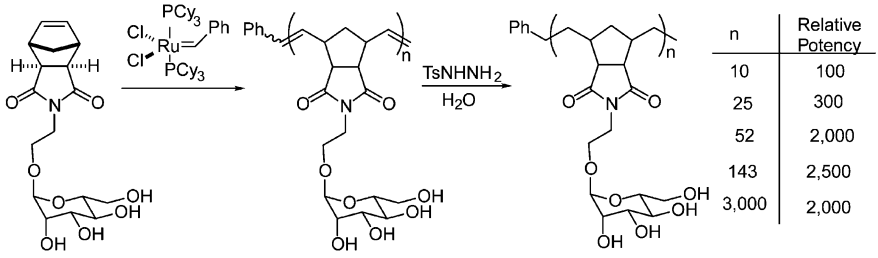
Carbohydrate Polymer Substituents		Relative inhibitory potency	
R =		10^3	**40**
R =		10^5	**41**
R =		250	**42**
R =		4 x 10^3	**43**

Figure 28
Multivalent arrays of carbohydrate derivatives show range of affinities for ConA. Relative potency IC_{50} polymer/(n)(IC_{50} α-Me-glucoside)



n	Relative Potency
10	100
25	300
52	2,000
143	2,500
3,000	2,000

Figure 29
Synthesis of defined-length, mannose-bearing neoglycopolymers using Grubbs' ruthenium initiator

ConA binds the monovalent *C*-mannoside, **37**, with a slightly (0.2 kcal mol^{-1}) [178] more favorable free energy than the corresponding *C*-glucoside, **36**. Given this small difference, it is not surprising that ConA cannot discriminate between the monovalent mannose and glucose ligands in a hemagglutination assay. In contrast, inhibitory potency of the mannose-substituted **41** versus the glucose-substituted polymer **40** was 100-fold greater [64]. Thus, the discrimination between the polymers was remarkable.

This enhancement of specificity through multipoint contacts may contribute to the exquisite specificity of cell–cell recognition events in nature. Although the observed enhancements in affinity and specificity, with which these neoglycopolymers bind to ConA, are consistent with the chelate effect, polymers of varying length were required for further elucidation of the mechanism by which the binding of these neoglycopolymers to ConA is enhanced.

A key advantage of the ROMP strategy is that defined metallocarbene initiators can be used to generate materials of different lengths, simply by varying the ratio of initiator to monomer. The Grubbs ruthenium catalyst gives initiation rates with reactive, strained, cyclic olefins that are significantly greater (approximately eight-fold) than the rate of propagation [174,179]. These relative rates allow for low product dispersity, if the polymerization is living. Thus, the average polymer length can be controlled by the reaction stoichiometry (monomer: initiator ratio). Accordingly, a defined ruthenium-carbene was chosen for the synthesis of a series of oligomers differing in valency (*Fig. 29*) [180]. By varying the reaction conditions, mannose-displaying oligomers ranging in average length from 10 to 150 monomer units were synthesized. In addition a series was generated in which the backbone was reduced to eliminate alkene isomers and to evaluate the role of backbone flexibility on activity. The resulting two sets of oligomers were then tested in the ConA-mediated hemagglutination assay described above.

The inhibitory potency of the mannose displaying polymers increases dramatically with polymer length, with materials of 50 residues displaying activity approximately 2000-fold higher activity than α-methyl mannoside on a saccharide residue basis [180]. Inspection of the ConA tetramer structure reveals that the saccharide binding sites within the homotetramer are approximately 65 Å apart. The neoglycopolymers produced in this study are a mixture of alkene backbone isomers. Molecular modeling of the most compact, limiting case, containing all *cis* backbone alkene isomers would require approximately 35 residues to span two binding sites within the ConA tetramer. Polymers composed of all *trans*-backbone alkenes would bridge two sites with approximately 17 monomer units. Moreover, it should be anticipated

that polymer potency would increase dramatically up to average length 50, as this length is adequate for all neoglycopolymer backbone alkene isomers to place saccharides in two binding sites simultaneously. The dependence of inhibitory potential on the average length of the polymers is consistent with these materials operating via the chelate effect. Polymers that can place saccharide residues in two binding sites on the lectin simultaneously apparently have the highest activity [180,181].

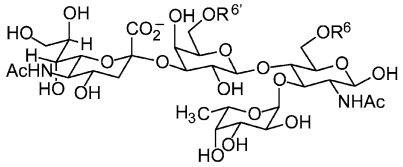
Although the chelate effect can account for much of the observed affinity enhancements, it cannot completely explain the dependence of affinity on polymer length. If the chelate effect was the only mechanism for increasing functional affinities, polymers composed of more residues than required to span the binding sites should display decreased activities, on a saccharide residue basis, relative to the shorter polymers, but they do not. The increased potency of longer polymers and those too short to bridge distinct saccharide binding sites, can be attributed to statistical effects, which lead to slower off-rates. That is, the high local concentration of mannose derivatives favors rebinding of individual epitopes and disfavors dissociation of the polymer.

The role of statistical effects can be investigated using multivalent ligands to explore a protein–carbohydrate interaction that cannot benefit from the chelate effect. The lectin of *Erythrina corallodendrum*, (EcorL) is a dimer with saccharide binding sites oriented in opposite directions; therefore, a multidentate ligand is expected to occupy only one saccharide binding site. Lactose-substituted polymers were synthesized and tested for inhibition of EcorL-mediated hemagglutination [60]. The lactose-bearing neoglycopolymers exhibited only a 5-fold increase in inhibitory potency on a saccharide residue basis. This increase in inhibitory potency is similar in magnitude to those observed with small molecules and dendrimers.

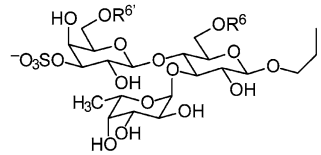
It is also conceivable that steric stabilization plays a role in the inhibition of hemagglutination by the longer polymers although it should be noted that the polymers generated via ROMP are considerably smaller than the acrylamide- based polymers that benefit from the steric mechanism. Moreover, the assays for the ROMP-derived ligands do not involve direct binding to a surface-immobilized receptor. It is also possible that the ability of ligands to cluster multiple copies of ConA is a key factor in their potency.

Selectins appear to be another example in which multivalency provides high functional affinity and specificity, yet these proteins are not oligomeric. The selectins (E-, P-, and L-selectin), a family of three saccharide binding proteins that mediate the recruitment of leukocytes from the bloodstream to sites of tissue damage, have been the subject of intense study since their discovery [88,96,182,183,184,185,186]. The three proteins all bind sLe^x weakly as a monomer; however, it appears that the endogenous ligands are more complex and bind with higher affinity. Thus, it seems likely that it is the features of the glycoproteins, including the multivalent presentation of saccharide epitopes may be important for binding [187,188,189,190].

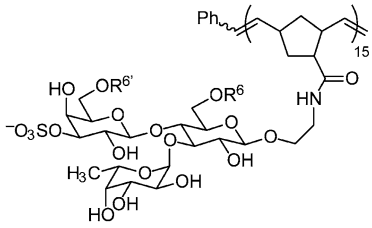
L-Selectin is one family member that may participate in multivalent interactions. Interestingly, endogenous L-selectin ligands, such as GlyCAM-1 [191,192], are mucins:. They are higly *O*-glycosylated, extended polypeptides from which multiple saccharide recognition elements extend [193,194]. GlyCAM-1 has been shown to primarily present the saccharide capping groups 6-sulfo-sialyl Lewis^x (6-sulfo sLe^x) and 6′-sulfo-sialyl Lewis^x (6′-sulfo sLe^x). The identification of these tetrasaccharide capping groups, coupled with the requirement for GlyCAM-1 sulfation for L-selectin binding, led to the hypothesis that the specificity of GlyCAM-1 binding to L-selectin might be controlled by multivalent recognition of the displayed, sulfated saccharides [173,195,196,197,198,199].



44 sLe^x $R^6 = R^{6'} = H$
45 6-sulfo sLe^x $R^6 = SO_3^-$, $R^{6'} = H$
46 6'-sulfo sLe^x $R^6 = H$, $R^{6'} = SO_3^-$



47 $R^{3'} = R^{6'} = SO_3^-$, $R^6 = H$
48 $R^{3'} = R^6 = SO_3^-$, $R^{6'} = H$



49 $R^{3'} = R^{6'} = SO_3^-$, $R^6 = H$
50 $R^{3'} = R^6 = SO_3^-$, $R^{6'} = H$

Figure 30
sLe^x derivatives and the sulfated Le^x derivatives generated by Kiessling and coworkers for the study of inhibition of leukocyte rolling

To study the importance of sulfation in the context of multivalent presentation, Sanders et al. sought to mimic the critical features of the physiological L-selectin ligand using the ROMP-generated neoglycopolymers [173]. The neoglycopolymers, like mucins, display multiple saccharide epitopes on an extended backbone. Monovalent compounds in which the sialic acid residue of sLe^x has been replaced by a sulfate were synthesized along with their multivalent counterparts (*Fig. 30*) [67,200,201]. Examination of the monovalent compounds in an ELISA using immobilized GlyCAM-1 as the L-selectin ligand, the trisaccharide 3′,6′-disulfo Le^x, **47** showed inhibitory potential similar to sLe^x, but 3′,6-disulfo Le^x, **48**, exhibited higher potency. This finding indicates that 6-sulfo sLe^x is an important recognition epitope, which binds to L-selectin with higher affinity than sLe^x [197,198,202].

Investigation of the inhibitory behavior of the sulfated neoglycopolymers, **49** and **50**, revealed that the multivalent arrays display similar activity in a static assay; both are approximately 80-fold (per saccharide residue) more effective at inhibiting L-selectin binding to heparin than is monovalent sLe^x. The ability of both polymeric ligands to block binding is consistent with numerous reports in which a variety of highly sulfated, multivalent ligands show comparable inhibition of L-selectin binding in static assays [203]. In contrast, under the more physiologically relevant conditions of shear flow [204], there is a dramatic difference in the inhibitory behavior of the two related ligands [173,199,205]. Neoglycopolymer **49** shows no ability to diminish L-selectin mediated cell rolling, while the 6-sulfated material, **50**, shows potent inhibitory capacity. One hypothesis the authors suggest is that compound **50** is poised to form multivalent interactions with cell surface L-selectin. Intriguingly, compound **50** is the most

potent inhibitor of L-selectin known. These studies reveal that both the recognition epitope and its multivalent display are important for specific recognition.

The challenges of synthesizing highly sulfated polymers led to the development of an alternative route to multivalent ligands through ROMP. Because *N*-hydroxysuccinimide esters are compatible with metathesis, ROMP-derived polymers can be generated that bear these activated esters, and nucleophilic functional groups bearing carbohydrate ligands can readily be appended [206]. Additionally, unique functional groups can be installed through selective end-capping of polymers generated by ROMP [207]. Fluorescent reporter groups have installed, and these can be used to test for multivalent binding [208,209]. These techniques have enabled ROMP-derived polymers to aggregate Jurkat T cell leukemia cells by noncovalently crosslinking [81], inhibit L-selectin [207] [208], induce protein oligomerization [156], and bind to B cells by engaging the lectin CD22 [210].

The development of new polymer chemistry will continue to provide access to diverse and novel multivalent ligands. These can be used to ask new types of mechanistic quesitons and to generate highly effective ligands for carbohydrate-binding proteins. As the importance of protein – carbohydrate interactions continues to grow, access to ligands that can be used to illuminate and control these processes will become even more valuable.

3 Applications

Researchers are beginning to exploit empirical evidence that synthetic, multivalent, saccharide-bearing materials can interact effectively with carbohydrate-binding receptors. While the study of multivalent protein–carbohydrate interactions is an emerging field, there is great potential for intervention in therapeutically relevant biological processes ranging from cell–cell recognition in inflammation to the attachment of pathogens such as influenza virus to regulating the immune system. Thus, synthetic materials with the ability to modulate these interactions are of considerable interest.

The applications of synthetic, multivalent carbohydrate derivatives range from their use as inhibitors to biosensors [211] to cell type-specific targeting agents [81,212,213]. Further investigations into the fundamental issues of protein–carbohydrate recognition will provide additional tools and insights for the modulation of physiological processes.

4 Conclusions

Complex networks of variables influence the interactions of proteins with carbohydrates on cell surfaces. The affinity and specificity of cell–cell recognition processes mediated by protein–carbohydrate interactions depends not only on the inherent specificity of one receptor for a specific saccharide epitope, but factors which are not operative in monovalent interactions. These factors include the location of the protein in vivo (on the cell surface or in circulation), the oligomerization state of the protein, its display of saccharide binding sites, the display of saccharide recognition epitopes, and the conformational preferences of the scaffold for ligand display. Given this range of variables, the factors that contribute to favorable multivalent protein–carbohydrate interactions are difficult to dissect. Moreover, cryptic alterations in the environment surrounding recognition elements may have dramatic effects on recogni-

tion [214,215]. Even with these challenges, considerable progress has been made towards illuminating the molecular mechanisms and physiological implications of multivalent processes. These advances are critical because an understanding of multivalent recognition events can be applied to the design of synthetic molecules that can modulate these processes. Progress in synthetic chemistry is providing powerful tools for the study of these fundamental, biological recognition events. This interplay, in which developments in chemical synthesis are used to advance our knowledge of biological systems, provides chemistry and biology with new challenges and opportunities.

References

1. Varki A (1993) Glycobiology 3:97
2. Sharon N (2007) J Biol Chem 282:2753
3. Gabius HJ (2006) Crit Rev Immunol 26:43
4. Nilsson CL (2003) Anal Chem 75:348A
5. Bertozzi CR, Kiessling LL (2001) Science 291: 2357
6. Weis WI, Drickamer K (1996) Annu Rev Biochem 65:441
7. Zelensky AN, Gready JE (2005) FEBS J 272: 6179
8. Barre A, Bourne Y, Van Damme EJ, Peumans WJ, Rouge P (2001) Biochimie 83:645
9. Rini JM (1995) Annu Rev Biophys Biomol Struct 24:551
10. Quiocho FA (1989) Pure & Appl Chem 61: 1293
11. Williams BA, Chervenak MC, Toone EJ (1992) J Biol Chem 267:22907
12. Liao D-I, Kapadia G, Ahmed H, Vasta GR, Herzberg O (1994) Proc Natl Acad Sci USA 91:1428
13. Loris R, Stas PPG, Wyns L (1994) J Biol Chem 269:26722
14. Lemieux RU (1996) Acc Chem Res 29:373
15. Quiocho RA, Wilson DK, Vyas NK (1989) Nature 340:404
16. Burley SKP (1985) Science 229:23
17. Butterfield SM, Sweeney MM, Waters ML (2005) J Org Chem 70:1105
18. Lis H, Sharon N (1986) Annu Rev Biochem 55:35
19. Borrok MJ, Kiessling LL (2007) J Am Chem Soc 129:12780
20. Borrok MJ, Kiessling LL, Forest KT (2007) Protein Sci 16:1032
21. Lemieux RU, Pavia AA, Martin JC, Watanabe KA (1969) Can J Chem 47:4427
22. Somers WS, Tang J, Shaw GD, Camphausen RT (2000) Cell 103:467
23. Feinberg H, Mitchell DA, Drickamer K, Weis WI (2001) Science 294:2163
24. Weis WI, Kahn R, Fourme R, Drickamer K, Hendrickson WA (1991) Science 254:1608
25. Forseen SKK (1994) Calcium in biological systems. In: Bertini IG, Lippard HB, Valentine JS (eds) Bioinorganic Chemistry. University Science Books, Mill Valley, pp 107–166
26. Faham S, Hileman RE, Fromm JR, Linhardt RJ, Rees DC (1996) Science 271:1116
27. DiGabriele AD, Lax I, Chen DI, Svahn CM, Jaye M, Schlessinger J, Hendrickson WA (1998) Nature 393:812
28. Pellegrini L, Burke DF, von Delft F, Mulloy B, Blundell TL (2000) Nature 407:1029
29. Pellegrini L (2001) Curr Opin Struct Biol 11:629
30. Mohammadi M, Olsen SK, Goetz R (2005) Curr Opin Struct Biol 15:506
31. Bourne Y, Rouge P, Cambillau C (1990) J Biol Chem 265:18161
32. Chervenak MC, Toone EJ (1994) J Am Chem Soc 116:10533
33. Chervenak MC, Toone EJ (1995) Biochemistry 34:5685
34. Clarke C, Woods RJ, Gluska J, Cooper A, Nutley MA, Boons GJ (2001) J Am Chem Soc 123:12238
35. Li Z, Lazardis T (2005) J Phys Chem B 109:662
36. Casu B, Petitou M, Provasoli M, Sinay P (1988) TIBS 13:221
37. Xu QW, Bush CA (1996) Glycobiology 6:1704
38. Woods RJ (1995) Curr Opin Struct Biol 5:591
39. Carver JP (1993) Pure & Appl Chem 65:763
40. Bundle DR, Alibes R, Nilar S, Otter A, Warwas M, Zhang P (1998) J Am Chem Soc 120:5317
41. Lindh I, Hindsgaul O (1991) J Am Chem Soc 113:216
42. Praly J-P, Lemieux RU (1987) Can J Chem 65:213

43. Johnson GP, Stevens ED, French AD (2007) Carbohydr Res 342:1210
44. Kurutz JK, Kiessling LL (1997) Glycobiology 7:337
45. Lee YC, Lee RT (1995) Acc Chem Res 28:321
46. Kiessling LL, Pohl NL (1996) Chem Biol 3:71
47. Roy R (1996) Curr Opin Struct Biol 6:692
48. Mammen M, Choi SK, Whitesides GM (1998) Angew Chem, Int Ed Eng 37:2754
49. Kitov PI, Bundle DR (2003) J Am Chem Soc 125:16271
50. Lundquist JJ, Toone EJ (2002) Chem Rev 102:555
51. Wolfenden ML, Cloninger MJ (2005) J Am Chem Soc 127:12168
52. Jencks WP (1981) Proc Natl Acad Sci 78:4046
53. Page MI, Jencks WP (1971) Proc Natl Acad Sci 68:1678
54. Cairo CW, Gestwicki JE, Kanai M, Kiessling LL (2002) J Am Chem Soc 124:1615
55. Gestwicki JE, Cairo CW, Strong LE, Oetjen KA, Kiessling LL (2002) J Am Chem Soc 124:14922
56. Kiessling LL, Strong LE, Gestwicki JE (2000) Annu Rep Med Chem 35:321
57. Crothers DM, Metzger H (1972) Immunochemistry 9:341
58. Rao JH, Lahiri J, Isaacs L, Weis RM, Whitesides GM (1998) Science 280:708
59. Kiessling LL, Gestwicki JE, Strong LE (2006) Angew Chem-Int Edit 45:2348
60. Pohl NL, Kiessling LL (1999) Synthesis SI:1515
61. Wright CS (1984) J Mol Biol 178:91
62. Hester G, Kaky H, Goldstein IJ, Wright CS (1995) Nat Struct Biol 2:472
63. Anderson CF, Record MT (1990) Annu Rev Biophys Biophys Chem 19:423
64. Mortell KH, Weatherman RV, Kiessling LL (1996) J Am Chem Soc 118:2297
65. Lee RT, Lee YC (1987) Glycoconjugate J 4:317
66. Carlson CB, Mowery P, Owen RM, Dykhuizen EC, Kiessling LL (2007) ACS Chem Biol 2:119
67. Alon R, Hammer DA, Springer TA (1995) Nature 374:539
68. Ashwell G, Harford J (1982) Ann Rev Biochem 51:531
69. Lodish HF (1991) TIBS 16:374
70. Connolly DT, Townsend RR, Kawaguchi K, Bell WR, Lee YC (1982) J Biol Chem 257:939
71. Lee YC, Townsend RR, Hardy MR, Lönngren J, Arnarp J, Haraldsson M, Lönn H (1983) J Biol Chem 258:199
72. Baenziger JU, Maynard Y (1980) J Biol Chem 255:4607
73. Spiess M (1990) Biochemistry 29:10009
74. Ikeda K, Sannoh T, Kawasaki N, Kawasaki T, Yamashina I (1987) J Biol Chem 262:7451
75. Weis WI (1994) Structure 2:147
76. Fan E, Merritt EA, Verlinde CLMJ, Hol WGJ (2000) Curr Opin Struct Biol 10:680
77. Karmali MA, Petric M, Lim C, Fleming PC, Arbus GS, Lior H (1985) J Infect Dis 151:775
78. Ling H, Boodhoo A, Hazes B, Cummings MD, Armstrong GD, Brunton JL, Read RJ (1998) Biochemistry 37:1777
79. Kitov PI, Sadowska JM, Mulvey G, Armstrong GD, Ling H, Pannu NS, Read RJ, Bundle DR (2000) Nature 403:669
80. Burke SD, Zhao Q, Schuster MC, Kiessling LL (2000) J Am Chem Soc 122:4518
81. Gestwicki JE, Strong LE, Cairo CW, Boehm FJ, Kiessling LL (2002) Chem Biol 9:163
82. Fan EK, Zhang ZS, Minke WE, Hou Z, Verlinde C, Hol WGJ (2000) J Am Chem Soc 122:2663
83. Zhang Z, Merritt EA, Ahn M, Roach C, Hou Z, Verlinde CL, Hol WG, Fan EK (2002) J Am Chem Soc 124:12991
84. Merritt EA, Zhang Z, Pickens JC, Ahn M, Hol WG, Fan EK (2002) J Am Chem Soc 124:8818
85. Pukin AV, Branderhorst HM, Sisu C, Weijers CA, Gilbert M, Liskamp RM, Visser GM, Zuilhof H, Pieters RJ (2007) ChemBioChem 8:1500
86. Polizzotti BD, Maheshwari R, Vinkenborg J, Kiick KL (2007) Macromolecules 40:7103
87. Bevilacqua MP, Stengelin S, Gimbrone MA, Seed B (1989) Science 243:1160
88. Lasky LA (1995) Annu Rev Biochem 64:113
89. Luster AD, Alon R, von Andrian UH (2005) Nat Immunol 6:1182
90. Ley K (2003) Trends Mol Med 9:263
91. Gonzalez-Amaro R, Sanchez-Madrid F (1999) Crit Rev Immunol 19:389
92. Simanek EE, McGarvey GJ, Jablonowski JA, Wong C-H (1998) Chem Rev 98:833
93. Kaila N, Thomas BE (2003) Expert Opinion on Therapeutic Patents 13:305
94. DeFrees SA, Kosch W, Way W, Paulson JC, Sabesan S, Halcomb RL, Huang DH, Ichikawa Y, Wong C-H (1995) J Am Chem Soc 117:66
95. Miyauchi H, Yuri M, Tanaka M, Kawamura N, Hayashi M (1997) Bioorg Med Chem Lett 7:989
96. Rosen SD, Bertozzi CR (1994) Curr Opin Cell Biol 6:663
97. Weatherman RV, Mortell KH, Chervenak M, Kiessling LL, Toone EJ (1996) Biochemistry 35:3619

98. Dimick SM, Powell SC, McMahon SA, Moothoo DN, Naismith JH, Toone EJ (1999) J Am Chem Soc 121:10286
99. Aoi K, Itoh K, Okada M (1995) Macromolecules 28:5391
100. Smith DK, Diederich F (1998) Chem-Eur J 4: 1353
101. Seebach D, Abele S, Gademann K, Guichard G, Hintermann T, Juan B, L. MJ, Schreiber JV (1998) Helv Chim Acta 81:932
102. Kim Y, Zimmerman SC (1998) Curr Opin Chem Biol 2:733
103. Fréchet JM (1994) Science 263:1710
104. Cloninger MJ (2002) Curr Opin Chem Biol 6:742
105. Dubber M, Lindhorst TK (1998) J Chem Soc Chem Commun 1265:
106. Ashton PR, Hounsell EF, Jayaraman N, Nilsen TM, Spencer N, Stoddart JF, Young M (1998) J Org Chem 63:3429
107. Lindhorst TK, in Host-Guest Chemistry. (2002), vol. 218, pp 201
108. Wolfenden ML, Cloninger MJ (2006) Bioconj Chem 17:958
109. Roy R, Zanini D, Meunier SJ, Romanowska A (1993) J Chem Soc, Chem Commun 1869
110. Reuter JD, Myc A, Hayes MM, Gan Z, Roy R, Qin D, Yin R, Piehler LT, Esfand R, Tomalia DA, Baker Jr. JR (1999) Bioconjugate Chem 10: 271
111. Vrasidas I, Andre S, Valentini P, Bock C, Lensch M, Kaltner H, Liskamp RMJ, Gabius HJ, Pieters RJ (2003) Organic & Biomolecular Chemistry 1:803
112. Yamazaki N, Kojima S, Gabius S, Gabius H (1992) Int J Biochem 24:99
113. Matsuo H, Funato K, Harashima H, Kiwada H (1994) J Drug Targeting 2:141
114. Kingery-Wood JE, Williams KW, Sigal GB, M. WG (1992) J Am Chem Soc 114:7303
115. Spevak W, Nagy JO, Charych DH, Schaefer ME, Gilbert JH, Bednarski MD (1993) J Am Chem Soc 115:1146
116. DeFrees SA, Phillips L, Guo L, Zalipsky S (1996) J Am Chem Soc 118:6101
117. Spevak W, Foxall C, Charych DH, Dasgupta F, Nagy JO (1996) J Med Chem 39:1018
118. Preobrazhenskaya ME, Berman AE, Mikhailov VI, Ushakova NA, Mazurov AV, Semenov AV, Usov AI, Nifant'ev NE, Bovin NV (1997) Biochem & Molecular Biology International 43:443
119. Larsen GR, Sako D, Ahern TJ, Shaffer M, Erban J, Sajer SA, Gibson RM, Wagner DD, Furie BC, Furie B (1992) J Biol Chem 267:11104
120. Liu W-j, Ramachandran V, Kang J, Kishimoto TK, Cummings RD, McEver RP (1998) J Biol Chem 273:7078
121. Manning DD, Hu X, Beck PJ, Kiessling LL (1997) J Am Chem Soc 119:3161
122. Sako D, Comess KM, Barone KM, Camphausen RT, Cumming DA, Shaw GD (1995) Cell 83:323
123. Pouyani T, Seed B (1995) Cell 63:333
124. Guo C-T, Sun X-L, Kanie O, Shortridge KF, Suzuki T, Miyamoto D, Hidari K, Wong C-H, Suzuki Y (2002) Glycobiology 12:183
125. Liang R, Loebach J, Horan N, Ge M, Thompson C, Yan L, Kahne D (1997) Proc Natl Acad Sci USA 94:10554
126. Liang R, Yan L, Loebach J, Ge M, Uozumi Y, Sekanina K, Horan N, Gildersleeve J, Thompson C, Smith A, Biswas K, Still WC, Kahne D (1996) Science 274:1520
127. Rheinnecker M, Hardt C, Ilag LL, Kufer P, Gruber R, Hoess A, Lupas A, Rottenberger C, Plückthun A, P. P (1996) J Immunol 157:2989
128. Love KR, Seeberger PH (2002) Angew Chem-Int Edit 41:3583
129. Timmer MSM, Stocker BL, Seeberger PH (2007) Curr Opin Chem Biol 11:59
130. Stevens J, Blixt O, Paulson JC, Wilson IA (2006) Nature Reviews Microbiology 4:857
131. Turnbull JE, Linhardt RJ (2006) Nat Chem Biol 2:449
132. Paulson JC, Blixt O, Collins BE (2006) Nat Chem Biol 2:238
133. Fukui S, Feizi T, Galustian C, Lawson AM, Chai WG (2002) Nat Biotechnol 20:1011
134. Smith EA, Thomas WD, Kiessling LL, Corn RM (2003) J Am Chem Soc 125:6140
135. Kiessling LL, Cairo CW (2002) Nat Biotechnol 20:234
136. Weply J, Abbes S, Scudder P, Keene J, Broschel K, Casnocha S, Gorka C, Steinenger C, Howard S, Schmuke J, Granelo M, Roissert J, Manger I, Jacob G (1994) Glycobiology 4:256
137. Horejsi V, Smolek P, Kocourek J (1978) Biochim Biophys Acta 538:293
138. Kallin E (1994) Methods Enzymol 242:221
139. Kobayashi K, Akaike T, Usui T (1994) Methods Enzymol 242:226
140. Nishimura S, Furuike T, Matsuoka K (1994) Methods Enzymol 242:235
141. Roy R (1996) Trends Glycosci Glycotechnol 8:79

142. Bovin NV, Gabius H-J (1995) Chem Soc Rev 24:413
143. Bovin NV, Korchagina EY, Zemlyanukhina TV, Byramova NE, Galanina OE, Zemlyakov AE, Ivanov AE, Zubov VP, Mochalova LV (1993) Glycoconjugate J 10:142
144. Miyauchi H, Tanaka M, Koike H, Kawamura N, Hayashi M (1997) Bioorg Med Chem Lett 7:985
145. Spaltenstein A, Whitesides GM (1991) J Am Chem Soc 113:686
146. Gamian A, Chomik M, Laferriere CA, Roy R (1991) Can J Microbiol 37:233
147. Matrosovich MN, Mochalova LS, Marinina VP, Byramova NE, Bovin NV (1990) FEBS Lett 272:209
148. Roy R, Tropper F (1988) J Chem Soc Chem Commun 1058:
149. Mammen M, Dahmann G, Whitesides GM (1995) J Med Chem 38:4179
150. Roy R, Park WKC, Srivastava OP, Foxall C (1996) Biorg Med Chem Lett 6:1399
151. Fan J-Q, Quesenberry MS, Takegawa K, Iwahara S, Kondo A, Kato I, Lee YC (1995) J Biol Chem 270:17730
152. Nishimura SI, Kee KB, Matsuoka K, Lee YC (1994) Biochem Biophys Res Commun 199:249
153. Ohno K, Tsujii Y, Miyamoto T, Fukuda T, Goto M, Kobayashi K, Akaike T (1998) Macromolecules 31:1064
154. Hawker CJ (1998) Acc Chem Res 30:373
155. Matyjaszewski K, Xia JH (2001) Chem Rev 101:2921
156. Griffith BR, Allen BL, Rapraeger AC, Kiessling LL (2004) J Am Chem Soc 126:1608
157. Spain SG, Gibson MI, Cameron NR (2007) J Polym Sci Pol Chem 45:2059
158. Narain R (2006) Reactive & Functional Polymers 66:1589
159. Vazquez-Dorbatt V, Maynard HD (2006) Biomacromolecules 7:2297
160. Muthukrishnan S, Zhang MF, Burkhardt M, Drechsler M, Mori H, Muller AHE (2005) Macromolecules 38:7926
161. Sen Gupta S, Raja KS, Kaltgrad E, Strable E, Finn MG (2005) Chem Commun 4315
162. Sigal GB, Mammen M, Dahmann G, Whitesides GM (1996) J Am Chem Soc 118:3789
163. Choi S-K, Mammen M, Whitesides GM (1996) Chem Biol 3:97
164. Thoma G, Magnani JL, Ohrlein R, Ernst B, Schwarzenbach F, Duthaler RO (1997) J Am Chem Soc 119:7414
165. Thoma G, Ernst B, Schwarzenbach F, Duthaler RO (1997) Bioorg Med Chem Lett 7:1705
166. Ali M, Hicks AE, Hellewell PG, Thoma G, Norman KE (2004) FASEB J 18:152
167. Thoma G, Duthaler RO, Magnani JL, Patton JT (2001) J Am Chem Soc 123:10113
168. Imai Y, Matsuura Y, Ono Y, Ishikawa T, Ito Y (2001) J Biochem (Tokyo) 130:665
169. Kamitakahara H, Suzuki T, Nishigori N, Suzuki Y, Kanie O, Wong C-H (1999) Angew Chem Int Ed Eng 37:1524
170. Nomura K, Schrock RR (1996) Macromolecules 29:540
171. Mortell KH, Gingras M, Kiessling LL (1994) J Am Chem Soc 116:10253
172. Fraser C, Grubbs RH (1995) Macromolecules 28:7248
173. Sanders WJ, Gordon EJ, Beck PJ, Alon R, Kiessling LL (1999) J Biol Chem 274: 5271
174. Lynn DM, Kanaoka S, Grubbs RH (1996) J Am Chem Soc 118:784
175. Trnka TM, Grubbs RH (2001) Acc Chem Res 34:18
176. Sampson NS, Lee Y (2006) Curr Opin Struct Biol 16:544
177. Kiessling LL, Gestwicki JE, Strong LE (2000) Curr Opin Chem Biol 4:696
178. Weatherman RV, Kiessling LL (1996) J Org Chem 61:534
179. Dias EL, Nguyen ST, Grubbs RH (1997) J Am Chem Soc 119:3887
180. Kanai M, Mortell KH, Kiessling LL (1997) J Am Chem Soc 119:9931
181. Mann DA, Kanai M, Maly DJ, Kiessling LL (1998) J Am Chem Soc 120:10575
182. Kansas GS (1996) Blood 88:3259
183. McEver RP, Moore KL, Cummings RD (1995) J Biol Chem 270:11025
184. Tedder TF, Steeber DA, Chen A, Engel P (1995) FASEB J 9:866
185. Rosen SD, Bertozzi CR (1996) Curr Biol 6:261
186. Varki A (1994) Proc Natl Acad Sci 91: 7390
187. Hemmerich S, Leffler H, Rosen SD (1995) J Biol Chem 270:12035
188. Wilkins PP, McEver RP, Cummings RD (1996) J Biol Chem 271:18732
189. Shailubhai K, Streeter PR, Smith EE, Jacob GS (1997) Glycobiology 7:305
190. Moore KL, Eaton SF, Lyons DE, Lichenstein HS, Cummings RD, McEver RP (1994) J Biol Chem 269:23318

191. Hemmerich S, Rosen SD (1994) Biochemistry 33:4830
192. Hemmerich S, Bertozzi CR, Leffler H, Rosen SD (1994) Biochemistry 33:4820
193. Uchimura K, Rosen SD (2006) Trends Immunol 27:559
194. Uchimura K, Gauguet JM, Singer MS, Tsay D, Kannagi R, Muramatsu T, von Andrian UH, Rosen SD (2005) Nat Immunol 6:1105
195. Sanders WJ, Katsumoto TR, Bertozzi CR, Rosen SD, Kiessling LL (1996) Biochemistry 35:14862
196. Yuen C-T, Lawson AM, Chai W, Larkin M, Stoll MS, Stuart AC, Sullivan FX, Ahern TJ, Feizi T (1992) Biochemistry 31:9126
197. Galustian C, Lawson AM, Komba S, Ishida H, Kiso M, Feizi T (1998) Biochem Biophys Res Commun 240:748
198. Mitsuoka C, Kasugai-Sawada M, Ando-Furui K, Izawa M, Nakanishi H, Nakamura S, Ishida H, Kiso M, Kannagi R (1998) J Biol Chem 273:11225
199. Gordon EJ, Sanders WJ, Kiessling LL (1998) Nature 392:30
200. Green PJ, Tamatani T, Watanabe T, Miyasaka M, Hasegawa A, Kiso M, Yuen C-T, Stoll MS, Feizi T (1992) Biochem Biophys Res Commun 188:244
201. Green PJ, Yuen C-T, Childs RA, CHai W, Miyasaka M, Lemoine R, Lubineau A, Smith B, Ueno H, Nicolaou KC, Feizi T (1995) Glycobiology 5:29
202. Scudder PR, Shailubhai D, Duffin KL, Streeter PR, Jacob GS (1994) Glycobiology 4:929
203. Crottet P, Kim YJ, Varki A (1996) Glycobiology 6:191
204. Finger EB, Puri KD, Alon R, Lawrence MG, von Andrian UH, Springer TA (1996) Nature 379:266
205. Mowery P, Yang ZQ, Gordon EJ, Dwir O, Spencer AG, Alon R, Kiessling LL (2004) Chem Biol 11:725
206. Strong LE, Kiessling LL (1999) J Am Chem Soc 121:6193
207. Gordon EJ, Gestwicki JE, Strong LE, Kiessling LL (2000) Chem Biol 7:9
208. Owen RM, Gestwicki JE, Young T, Kiessling LL (2002) Org Lett 4:2293
209. Roberts KS, Sampson NS (2004) Org Lett 6:3253
210. Yang ZQ, Puffer EB, Pontrello JK, Kiessling LL (2002) Carbohydr Res 337:1605
211. Charych D, Cheng Q, Reichert A, Kuziemko G, Stroh M, Nagy JO, Spevak W, Stevens RC (1996) Chem Biol 3:113
212. Gonsho A, Irie K, Susaki H, Iwasawa H, Okuno S, Sugawara T (1994) Biol Pharm Bull 17:275
213. Merwin JR, Noell GS, Thomas WL, Chiou HC, Derome ME, McKee TD, Spitalny GL, Findeis MA (1994) Bioconj Chem 5:612
214. Crocker PR, Feizi T (1996) Curr Opin Struct Biol 6:679
215. Lind SS, Parker W, Everett ML, Platt JL (1998) Glycobiology 8:433

12.5 Biomedicine of Enkephalin-Derived Glycopeptide Analgesics

Robin Polt
University of Arizona, Tucson, AZ 85721-0041, USA
polt@u.arizona.edu

Abstract

The incorporation of glycosides into peptide neurotransmitters imparts drug-like character to the neurotransmitter "message" via "membrane hopping". The importance of the glycopeptide-membrane interaction is emphasized, and the *biousian* theory is briefly explained. Application of this approach to enkephalins, the endogenous opioid peptides, leads to potent analgesic compounds capable of systemic delivery. The clinical applications of these compounds are advocated by the author.

Keywords

Glycoproteins; Blood-brain barrier; Hormones; *O*-Linked glycopeptides; Opiates; Transport

Abbreviations

ACTH	adrenocorticotropic hormone
BBB	blood-brain barrier
CHO	Chinese hamster ovary
CNS	central nervous system
CFS	cerebrospinal fluid
DAMGO	Tyr-*D*Ala-Gly-*N*-Me-Phe-glycinol
DPDPE	cyclo-*S*-*S*-[Tyr-D-Pen-Gly-Phe-D-Pen] or D-Pen-D-enkephalin
GPCR	G protein-coupled receptor
GMP	good manufacturing practice

In: *Glycoscience*. Fraser-Reid B, Tatsuta K, Thiem J (eds)
Chapter-DOI 10-1007/978-3-540-30429-6_65:

POMC pro-opiomelanorcortin
MSH melanocyte-stimulating hormone
M6G morphine-6-glucuronide
SAR structure-activity relationship

1 Endogenous Opioid Peptides

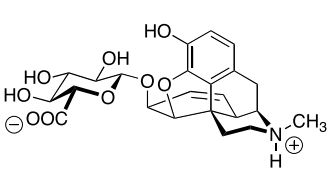
Enkephalins are the endogenous peptide ligands for the opioid receptors that were discovered in the 1970s [1]. The opioid receptors are members of the G protein-coupled receptor (GPCR) family, a class of membrane-bound receptors that exhibit a conserved topology and represent 80% of all known receptors [2]. Analysis of the deduced amino acid sequences of the many GPCRs that have been cloned indicates that they have a seven transmembrane motif with three external and three internal loops. The domains with the greatest similarity are the transmembrane regions and the intracellular loops, and the most divergent are the extracellular loops and the amino- and carboxyl-terminals. This well-populated class of receptors that includes numerous neurotransmitter and hormonal receptors represents a very attractive target for drug design. In neuronal cells, activation of opioid receptors is coupled, via pertussis-toxin-sensitive G proteins, to various effectors including adenylate cyclase and K^+ and Ca^{2+} channels [3].
The endogenous ligands of these opiate receptors have been identified as a family of more than 20 opioid peptides that originate from the three precursor proteins: pro-opiomelanorcortin (POMC), the common protein precursor for β-endorphin, ACTH, and additional MSH-containing peptides; pro-enkephalin A, the precursor for met- and leu-enkephalin and several larger enkephalin-containing peptides (e. g. peptides E and F) and pro-enkephalin B (prodynorphin), another precursor for leu-enkephalin and for larger opioid peptides (e. g. the dynorphins and neo-endorphins). The opioid peptides are generated from their respective precursors by selective proteolysis, mainly at basic and paired basic residues. Collectively, these peptides are referred to as "endorphins", and all possess the Tyr-Gly-Gly-Phe opioid "message". The activation of the opioid receptors produces analgesia as well as many other effects. Endogenous opioid peptides typically have short half-lives in vivo due to rapid enzymatic biodegradation, and poor bioavailability in tissues and organs. The challenge facing chemists and pharmacologists is to improve metabolic stability and extend their bioavailability while maintaining potency and selectivity [4,5].
The isolation of various endogenous opioid peptides and their classification into families from common precursor proteins has led to attempts to pair particular opioid peptides with opioid receptor sub-types, μ (mu, MOP or MOR), δ (delta, DOP or DOR), or κ (kappa, KOP or KOR). All of the endogenous peptide ligands are agonists at the opioid receptors. The pro-enkephalin-related peptides all show a preference for the δ receptor. All dynorphins and neo-endorphins show κ agonism, but dynorphin A_{1-8} retains a preference for the δ receptor, whereas dynorphin A_{1-13} is potent at both μ and κ receptors. The pro-opiomelanocortin family gives rise to only one opioid peptide, β-endorphin, which is a very potent agonist at both μ and δ receptors with a slight preference for δ. None of the peptide ligands are particularly selective for any opioid receptor sub-type. The precise biological significance of this ligand-receptor "overlap" is unknown [6].

Table 1
Representative endogenous opioids

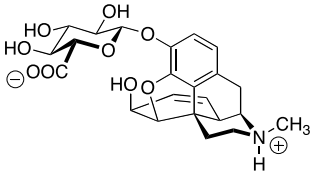
POMC (porcine)	
β-Endorphin:	**Tyr-Gly-Gly-Phe-Met**-Thr-Ser-Glu-Lys-Ser-Gln-Thr-Pro-Leu-Val-Thr-Leu-~Phe-Lys-Asn-Ala-Ile-Val-Lys-Asn-Ala-His-Lys-Lys-Gly-Gln
Pro-Enkephalin (human)	
Met-Enkephalin:	**Tyr-Gly-Gly-Phe-Met**
Leu-Enkephalin:	**Tyr-Gly-Gly-Phe-Leu**
Peptide E:	**Try-Gly-Gly-Phe-Met**-Arg-Arg-Val-Gly-Arg-Pro-Glu-Trp-Trp-Met-Asp-~Tyr-Gln-Lys-Arg-**Tyr-Gly-Gly-Phe-Leu**
Pro-Dynorphin (porcine)	
α-Neo-Endorphin:	**Tyr-Gly-Gly-Phe-Leu**-Arg-Lys-Tyr-Pro-Lys
β-Neo-Endorphin:	**Tyr-Gly-Gly-Phe-Leu**-Arg-Lys-Tyr-Pro
Dynorphin A_{1-8}:	**Tyr-Gly-Gly-Phe-Leu**-Arg-Arg-Ile
Dynorphin A_{1-17}:	**Tyr-Gly-Gly-Phe-Leu**-Arg-Arg-Ile-Arg-Pro-Lys-Leu-Trp-Asp-Asp-Gln
Dynorphin B_{1-13}:	**Tyr-Gly-Gly-Phe-Leu**-Arg-Arg-Gln-Phe-Lys-Val-Val-Thr

Currently, almost all clinically available opioid drugs target the μ receptor for relief of moderate to severe pain, reflecting a common mechanism of action similar to morphine. Morphine shows preference for μ-opioid receptors, but also binds to δ and κ receptors with low affinity. These "narcotics" remain the most widely used drugs for severe pain, and also produce a suite of side effects, including pruritus (itching), miosis (contraction of the pupils), sedation, constipation, and respiratory depression. All currently available μ-agonists display tolerance in vivo and have a potential to produce both physical and psychological addiction, and statistical data reveal that opiate drug abuse and addiction is increasing [7].

Results from Bilsky [8], Chang [9], Negus [10], Dykstra [11], and others show that compounds with mixed μ and δ opioid receptor agonist activity ($\mu + \delta$ approach) may have utility in achieving a similar, or even a greater degree of analgesia and fewer adverse effects than μ-selective agonists such as fentanyl or morphine. Activation of δ-opioid receptors was initially thought to have *no* side effects, but this does not seem likely. Using other criteria, it has been shown that mixed μ-agonist and δ-antagonist activity (μ–δ approach) produces analgesia



Morphine-6-β-D-Glucuronide



Morphine-3-β-D-Glucuronide

Figure 1
Glycosylated metabolites of morphine [16,22,23]

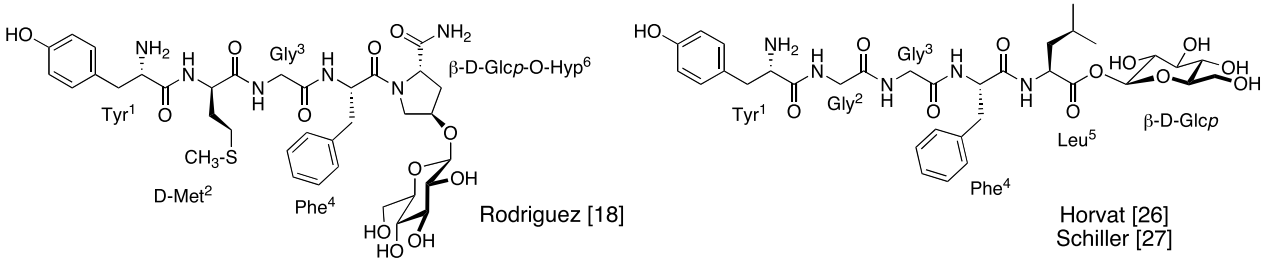


Figure 2
Glycosylated enkephalin derivatives [25,26,27]

with lowered tolerance [12,13]. Clearly, opioids with altered agonist profiles should be studied as a treatment for various types of pain, including severe pain, chronic pain, and neuropathic pain. The use of glycosylated "*address segments*" to promote "*membrane hopping*" and penetration of the blood-brain barrier (BBB) should allow such studies to proceed without more invasive means to deliver peptide "*messages*" to the CNS [14,15].

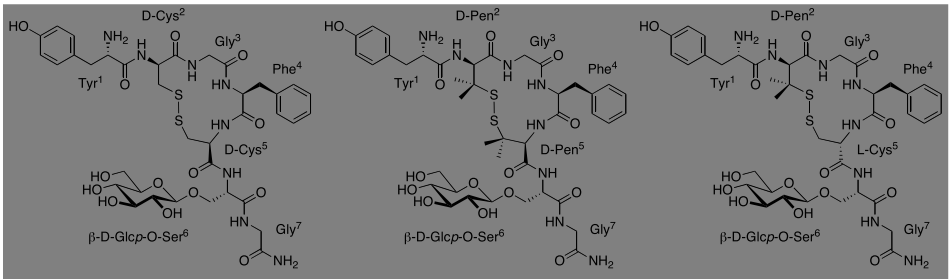
Initially, we were intrigued by reports that morphine-6-glucuronide (M6G) (❯ *Fig. 1*) apparently penetrated the BBB and produced potent analgesia [16], despite the fact that the resulting zwitterionic metabolite could not possibly be diffusing through the lipophilic BBB [17]. This fact, coupled with a report that an enkephalin galactoside analogue was much more potent than the corresponding glucoside [18], (❯ *Fig. 2*) led us to hypothesize (*incorrectly*) that the glucose transport system GLUT-1 might be involved [19]. Early work in the Davis lab [20] and more recent work by others [21] confirmed that GLUT-1 was *not* involved in the transport of glycopeptides across the BBB [24]. Several research groups began to explore the attachment of carbohydrate groups to opioid peptides in order to influence receptor binding (μ vs. δ), improve serum stability, and enhance peptide delivery to the brain or spinal column.

2 Design Principles for Glycosylated Enkephalins with Improved Bioavailability

Enkephalin analogues bearing glycosyl residues were first synthesized and evaluated pharmacologically by Rodriguez [18], Schiller [27], and Horvat [26] (❯ *Fig. 2*). The hydroxyproline glycoside linkage in the Rodriguez compound is found in nature and may be properly referred to as a *glycopeptide*. The acyl glycoside on the right, and other glycosidic linkages *not* found in nature are more properly referred to as *neo-glycopeptide structures*. Many glycosylated enkephalin analogues have now been prepared and tested in vitro and in vivo. It is safe to generalize that the glycoside or carbohydrate moiety should not be located near the *N*-terminus to avoid interference with binding to the opioid receptors. Likewise, it may be safe to say that selectivity for the opioid subtype ($\mu/\delta/\kappa$) is largely, but not exclusively determined by the amino acid sequence. For earlier reviews of many synthetic glycopeptide derivatives of enkephalin, and some other glycopeptides, along with more complete discussions of opioid SAR, see Polt [25] and Horvat [26]. Many glycosylated enkephalins show greatly enhanced binding activity to μ- and/or δ-receptors and remarkable central analgesic activity when administered peripherally. The potential of glycoconjugates to modify potency, receptor selectivity, and bioavailability has now been clearly and abundantly demonstrated. What remains to be demonstrated are the clinical benefits of glycosylated enkephalin derivatives.

Our own work in this area began with glycosylated DPDPE analogues depicted in ❯ *Table 2*. The synthesis of these cyclic disulfides was not always straightforward, and the –S–S– bond proved to be somewhat labile. While most peptide chemists were focusing on improving $\mu/\delta/\kappa$ selectivity for the generation of better pharmacological *tools*, we chose to focus our efforts on the development of better *drugs*, by selecting for enhanced potency, bioavailability, and penetration of the BBB [4,5,28].

Table 2
Disulfide-linked glycopeptides related to DPDPE



	Glyco-DCDCE	Glyco-DPDPE	Glyco-DPLCE
δ-binding	46.4 ($\pm$ 53) nM	85 nM	<1 nM
μ-binding	45.4 ($\pm$ 26) nM	48000 nM	>400 nM

Building on the extensive opioid SAR of the enkephalins, it was relatively straightforward to choose opioid peptide agonists with the desired μ/δ selectivity, and convert them into glycopeptides with similar selectivity. We tried to avoid, or at least ignore κ-agonists, since this receptor has been associated with dysphoria. Enkephalins possess highly amphipathic conformations that promote strong interactions with membranes [29,30,31], and it is likely that most if not all peptide neurotransmitters share this feature. Thus, when a peptide neurotransmitter is released from a pre-synaptic vesicle, it needs to travel only several hundred angstroms at most (the distance between the pre- and post-synaptic membranes), where it can adsorb strongly to the post-synaptic neuronal membrane. A rapid *2-dimensional* "search" for the receptor in turn facilitates rapid binding of the neurotransmitter to a membrane-bound receptor. Surface-assisted "reduction-of-dimensionality" calculations were examined by Max Delbrück in which he quantitatively demonstrated the viability of this theory [32]. The probability of a membrane-bound neuropeptide finding its corresponding receptor approaches 100%, assuming the absence of convection and that the membrane is fluid (*Fig. 3*). This model has also been endorsed by Robert Schwyzer, who invokes the membrane as a "catalyst" or "matrix" for receptor-ligand interactions [33].

In a recent study [34], we chose to minimize the confounding effects of $\mu/\delta/\kappa$-selectivity by choosing DAMGO (D-**A**lanine-*N*-**M**ethylphenylalanine-**G**lycin**o**l-Enkephalin), a highly μ-selective ligand, for our "message segment". The attachment of a glycoside (or other water-soluble moieties) near the *C*-terminus of the peptide led to increased stability of the aqueous state without perturbation of the membrane-bound conformation of the peptide message [35]. Thus, instead of simply binding to a biologic membrane, the glycopeptide can "*hop off*" the membrane, free to travel some distance before it encounters another membrane that will permit the glycopeptide to "*hop on*" again. By balancing the free energy of the two states (e. g. membrane-bound state versus aqueous state, *Fig. 4*), the glycopeptide will spend the optimal amount of time in the aqueous phase in order to impart drug-like characteristics to the largely membrane-bound peptide neurotransmitter. This can be accomplished by the addition

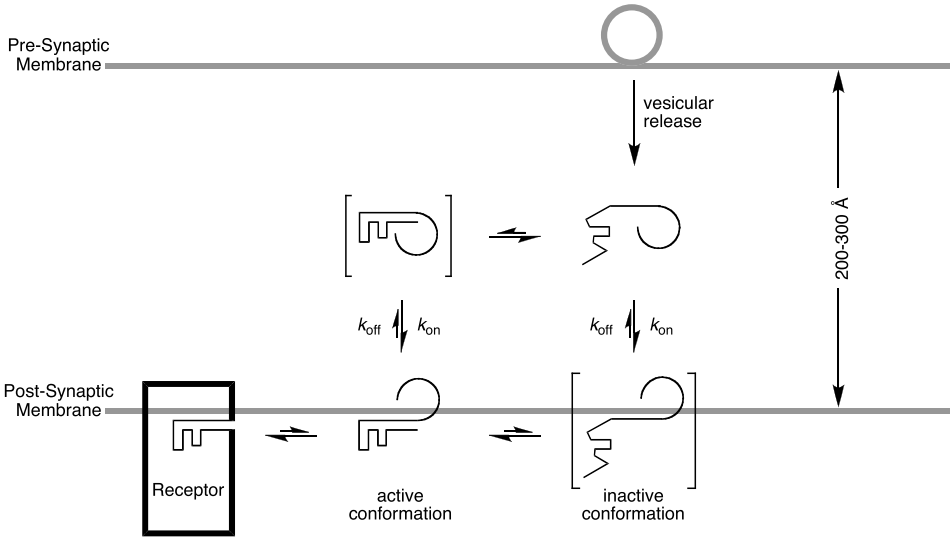


Figure 3
The membrane plays a dominant role in "the search" for the receptor and subsequent binding [33]

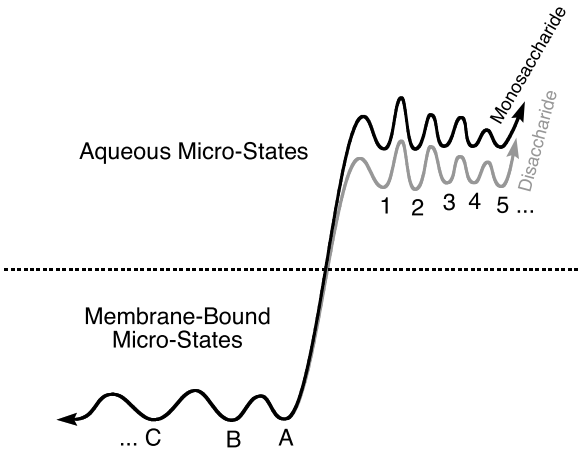
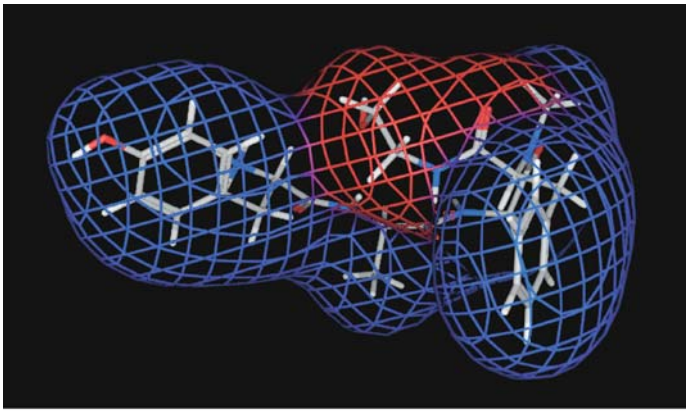
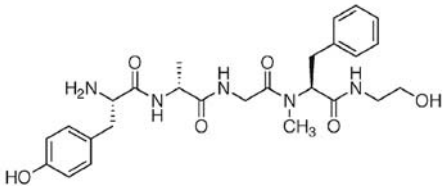


Figure 4
Membrane bound conformations (microstates) vs. aqueous conformations

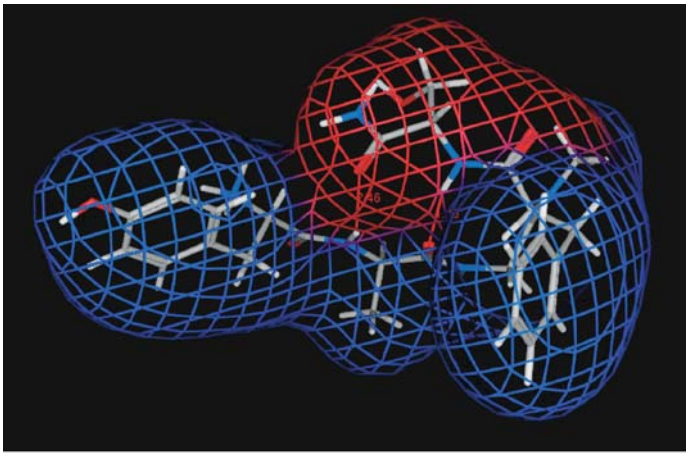
of larger saccharides with increased hydrodynamic volumes (e. g. disaccharide vs. monosaccharide) [36].

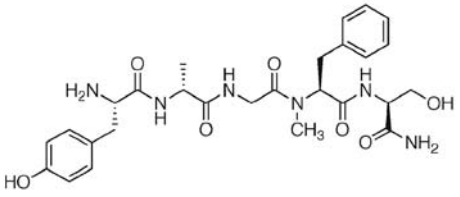
If the glycopeptide spends too much time in the aqueous compartment, then the interaction with its receptors will be reduced. Thus, the glycopeptide needs to spend sufficient time in the membrane to bind to the receptor and undergo endocytosis (vide infra). We have attempted to quantify this process, or at least to rationalize it, by calculating the solvent-accessible sur-



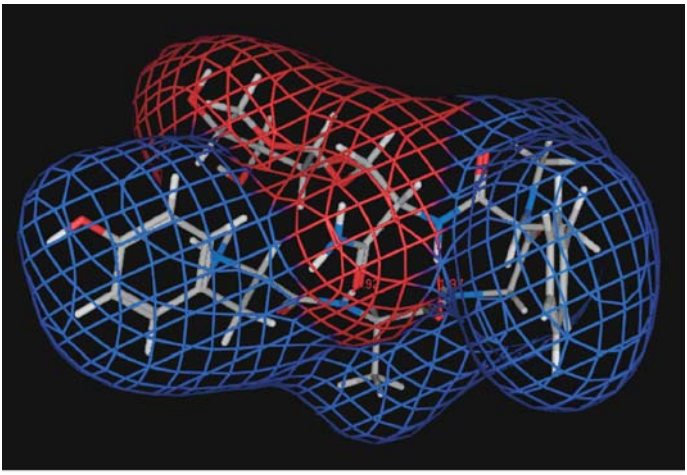


a

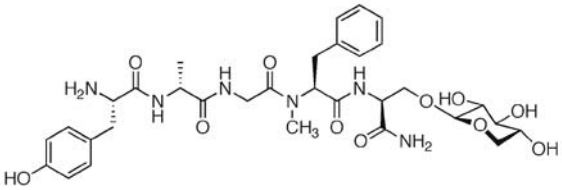


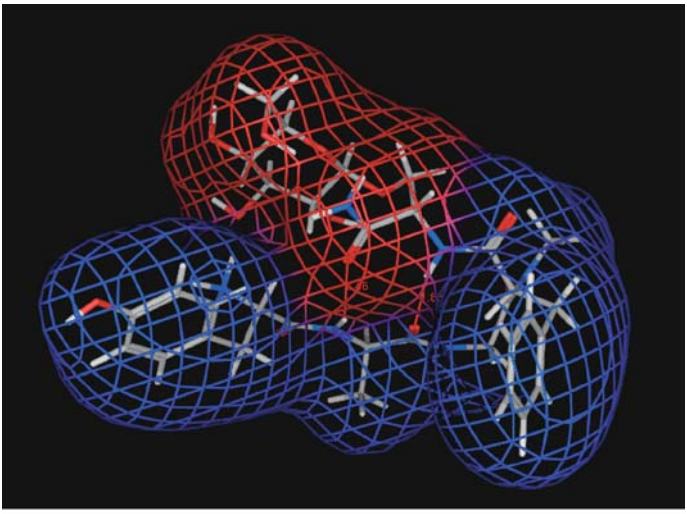


b

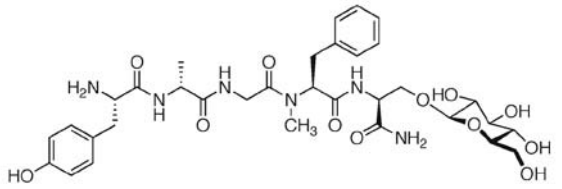


c


H2N
O
N
H
O
N
H
O
N
CH3
O
N
H
O
HO
O
OH
OH
O
NH2
HO



d


H2N
O
N
H
O
N
H
O
N
CH3
O
N
H
O
HO
O
OH
OH
O
NH2
HO

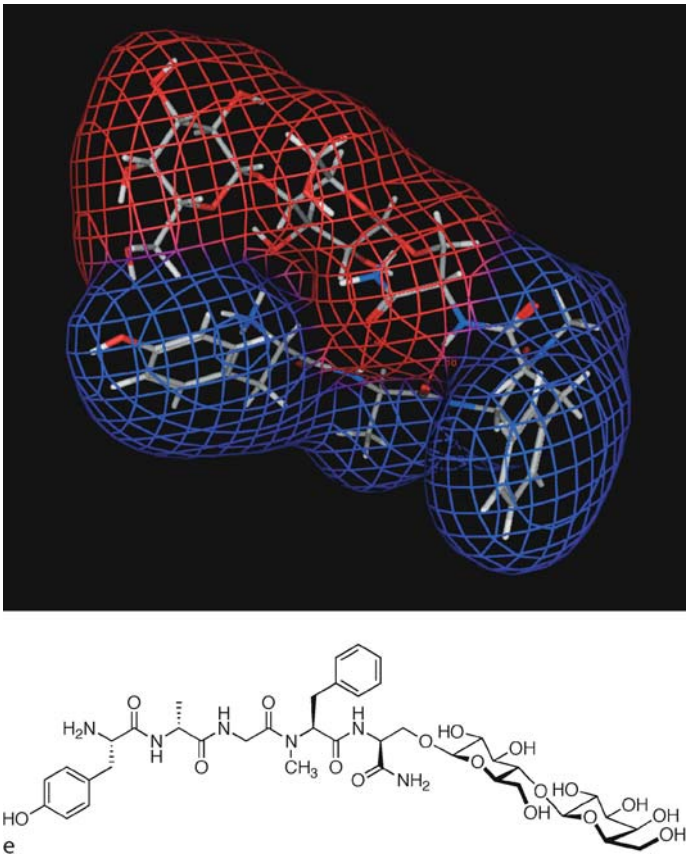


Figure 5
Glycosylated DAMGO derivatives

face area (Connolly surface, *Fig. 5*), and dividing that area into a hydrophilic surface and a lipophilic surface [34].

The amphipathic nature of the DAMGO derivatives was visualized (*Fig. 5*) using the molecular mechanics package of MOE® (Chemical Computing Group, Montreal, Canada) and by labelling the surface blue to indicate the lipophilic portion of the surface (A_{lipid}) and red to indicate the hydrophilic portion of the surface (A_{water}). The ratio of the two types of surface areas was used to create an expression of the amphipathicity in the form of an equilibrium constant using the formula $\mathbf{A} = e^{Awater/Alipid}$. The precise amphipathicity of each molecule is fluxional as each molecule is flexible, existing as an ensemble of conformations (A, B, C… or 1,2,3,4,5… c.f. *Fig. 4*), even when bound to a membrane or micelle [35,37].

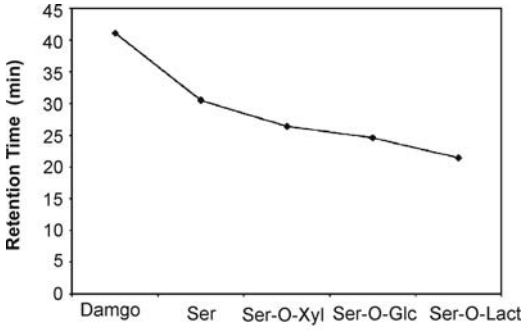
The DAMGO derivatives (*Fig. 5*), including DAMGO itself were administered to mice using intravenous (i.v.) injection, as well as after intracerebroventricular (i.c.v.) injection, and the in vivo responses measured using the 55 °C tail-flick assay by our collaborators in the Bilsky lab at the University of New England. Full dose-response curves were obtained, and the

antinociceptive potency of each drug was recorded as the amount of drug required to obtain 50% activity, or the A_{50} value. Interestingly, when the i.v. potency was plotted versus several measures of amphipathicity, similar U-shaped or V-shaped curves were obtained (*Fig. 6*). The binding affinities of each compound for the $\mu/\delta/\kappa$-opiate receptors were measured by our collaborators in the Bidlack lab at the University of Rochester Medical Center. Binding was measured using radioimmunoassay methods with human opiate receptors expressed in CHO cells. All of the DAMGO derivatives were extremely potent and selective for the μ-receptor. The antinociception and binding values are summarized in *Table 3*, along with values for morphine and ACR-150 (vide infra).

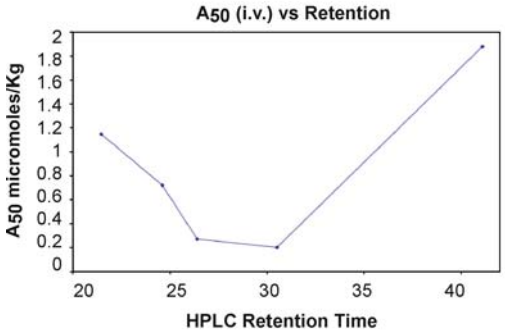
Several of these μ-selective DAMGO derivatives are sufficiently bioavailable and potent enough to be clinically useful, but still show side effect profiles that are similar to morphine or fentanyl. Use of the opioid agonist bearing an *O*-serinyl β-lactoside ACR-150 (*Table 3*) based on the Roques sequence Tyr-D-Thr-Gly-Phe-Leu-Ser (DTLES) [38] is expected to show reduced side effects, and has been subjected to more intense study. Fractional-kilogram amounts of GMP-grade ACR-150 have been produced by two independent contractors, and bioanalytical methods have been developed in order to monitor serum levels of the glycopeptide. This *lactoside* glycopeptide not only shows increased penetration of the BBB and potent antinociception in mice, but also shows diminished side effects normally associated with μ-agonists [4] (*Fig. 7*). Studies of the corresponding *glucoside* have been published [39].

By now it is clear that glycosylation of an enkephalin-derived peptide increases penetration of the peptide across the rat BBB in situ and increases systemic bioavailability of the compound in vivo. A significant feature of these studies is the use of reliable methods for the synthesis of multi-gram quantities of ACR-150 [40,41], which has allowed for further study in vivo. This has allowed us to observe the increased potency of the opioid glycopeptide following i.v. or s.c. (subcutaneous) administration, and assessment of the side-effect profile of ACR-150 following systemic administration. By administering the compound systemically, the glycopeptide presumably accesses opioid receptors in both the brain and spinal cord, an important consideration in the development of physical dependence in mice [42]. The results indicate that ACR-150 produced less physical dependence in mice than equivalent doses of morphine using an acute model (4 day), and additional studies are needed to assess the development of physical dependence using chronic models. Unlike μ-agonists that show strong locomotor effects in mice, locomotor stimulation by ACR-150 is absent. It is hoped that the pre-clinical studies done with this glycopeptide will result in the commercial application of glycosylated enkephalins for the treatment of pain, and that these studies can be published in the open literature very soon.

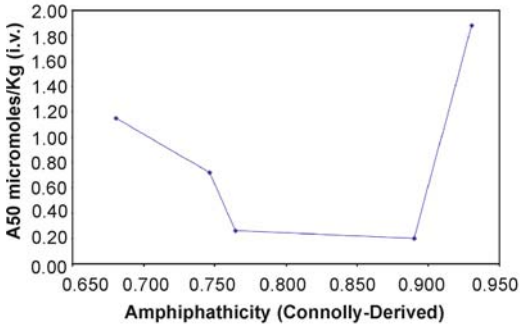
It is too far early to assess the actual *clinical* relevance of ACR-150 (*Fig. 7*) and related glycopeptides, but it is not too early to state with some assurance that this compound or a similar glycopeptide should be subjected to the necessary pre-clinical studies with the intention of testing glycopeptide-based analgesia in human subjects. Other glycosylated opioid peptides would seem to have similar promise, most notably those prepared in the laboratories of Rocchi [43,44], and Negri [45]; dermorphins and related glycopeptoids; the laboratory of Tomatis who has been working with deltorphins in addition to dermorphins [46]; Schiller [47]; Horvat [48]; and Toth, who has been working with glycolipoyl enkephalin analogues [49] (*Fig. 8*). All of these compounds bear some promise for the future, but ACR-150 has been manufactured in fractional-kilogram amounts under the conditions of good manufacturing



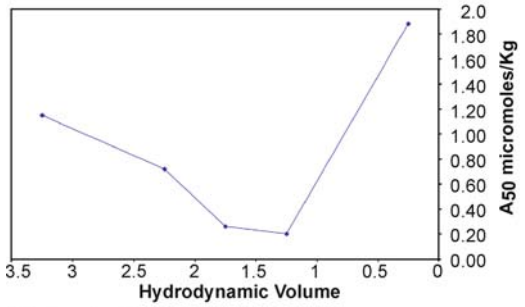
a RP-HPLC Retention Time vs Glycosylation



b Antinociception vs RP-HPLC Retention Time



c Antinociception vs Amphiphathicity



d Antinociception vs Hydrodynamic Volume

Figure 6

Centrally mediated antinociception in mice shows a U-shaped or V-shaped curve when correlated with three different measures of amphipathicity. All three analyses produce a U- or V-shape, as predicted by the biousian hypothesis [15]. Amphipathicity values A, were calculated using the formula $A = e^{Awater/Alipid}$. (a) RP-HPLC retention time vs. glycosylation. Retention times on a C_{18} column decrease monotonically as the degree of glycosylation is increased. (b) Antinociception vs. RP-HPLC retention time. Antinociception, or A_{50} values, as measured by 55 °C mouse tail-flick assay after i.v. administration of each glycopeptide [34] is recorded on the Y-axis as a function of HPLC retention times. (c) Antinociception vs. amphipathicity (A values). A_{50} values as a function of amphipathicity. Connolly-derived amphipathicity values are plotted along the X-axes, and A_{50} values derived from mouse intravenous tail-flick data are plotted on the Y-axis. (d) Antinociception vs. hydrodynamic volume. A_{50} values as a function of hydrodynamic values (glucose units) for the hydrophilic address segment

Table 3

Peak antinociception (A_{50} values) determined following i.c.v. and i.v. administration and receptor binding (K_i values) of the μ-selective DAMGO derivatives. Values for the balanced μ/δ-agonist ACR-150 (YtGFLS*, where S* is the β-lactoside of L-serine amide) are also listed, along with values for morphine sulfate included for comparison. A_{50} values were determined after intracerebral ventricular (i.c.v.) administration or intravenous (i.v.) administration using the mouse 55 °C tail-flick assay. Binding was determined in Chinese hamster ovary (CHO) cell membranes, expressing either the human μ, δ, or κ opioid receptor. Each radiolabelled reference drug (indicated above) was incubated with 12 concentrations of the glycopeptide. Nonspecific binding was measured by the inclusion of 10 mM naloxone. Binding data are the mean K_i values from three experiments performed in triplicate [34].

Opioid drug	A_{50} i.c.v. (pmol/mouse)	A_{50} i.v. (μmol/kg)	[^{3}H]DAMGO (μ)
Morphine	2384	7.84	0.79
ACR-150	19	4.07	5.00
DAMGO	30	1.88	0.56
DAMGO-Ser	2.0	0.20	0.68
DAMGO-Ser-Xyl	2.0	0.27	1.30
DAMGO-Ser-Glc	19	0.72	1.30
DAMGO-Ser-Lact	2.0	1.15	0.66

Opioid drug	[^{3}H]Naltrindole (δ)	[^{3}H]U69,593 (κ)	μ:δ:κ
Morphine	290	12	1:370:15
ACR-150	9.20	42	1:2:8
DAMGO	990	270	1:1900:510
DAMGO-Ser	600	190	1:880:280
DAMGO-Ser-Xyl	730	160	1:560:120
DAMGO-Ser-Glc	54% at 10 IM	270	1:>5000:210
DAMGO-Ser-Lact	1600	350	1:2400:530

process (GMP), a fact which has greatly facilitated consideration of this class of glycosylated enkephalins as pharmaceutical entities.

The two circled compounds bear further comment. Terasaki's peptide, TAPA [50], is known to penetrate the BBB *without* the benefit of glycosylation. Presumably, this is because the

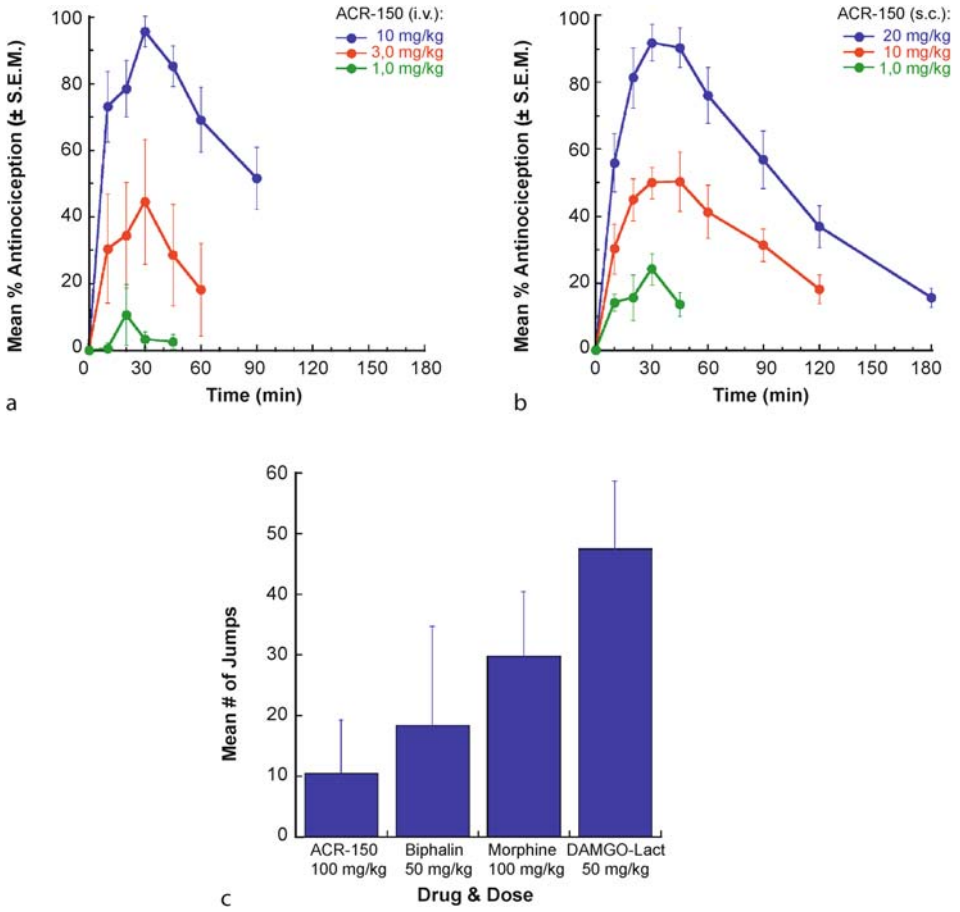


Figure 7
Effects of ACR-150 in mice. a Potency i.v. b Potency s.c. c Addiction potential of ACR-150, compared with biphalin, morphine, and the DAMGO-lactoside, in order of increasing μ-selectivity

charged group is highly solvated and the peptide can "hop" along the membrane in the same way that is postulated for the glycopeptides [34]. It is possible that Rocchi's Lys^7 compounds may also benefit from the additional charge in the molecule.

3 Putative Mechanism for Blood-Brain Barrier Penetration

The BBB of the brain has two major components. An endothelial layer lies between the arterial blood in the brain capillaries and the interstitial fluid of the brain. In humans, the surface area of the endothelial layer in the brain is approximately 21 square meters [51]. An epithelial layer lies between venous blood and the cerebrospinal fluid (CSF) in the choroid plexus, and has a surface area of only 0.021 square meters in humans [52]. At the spinal cord, the BBB

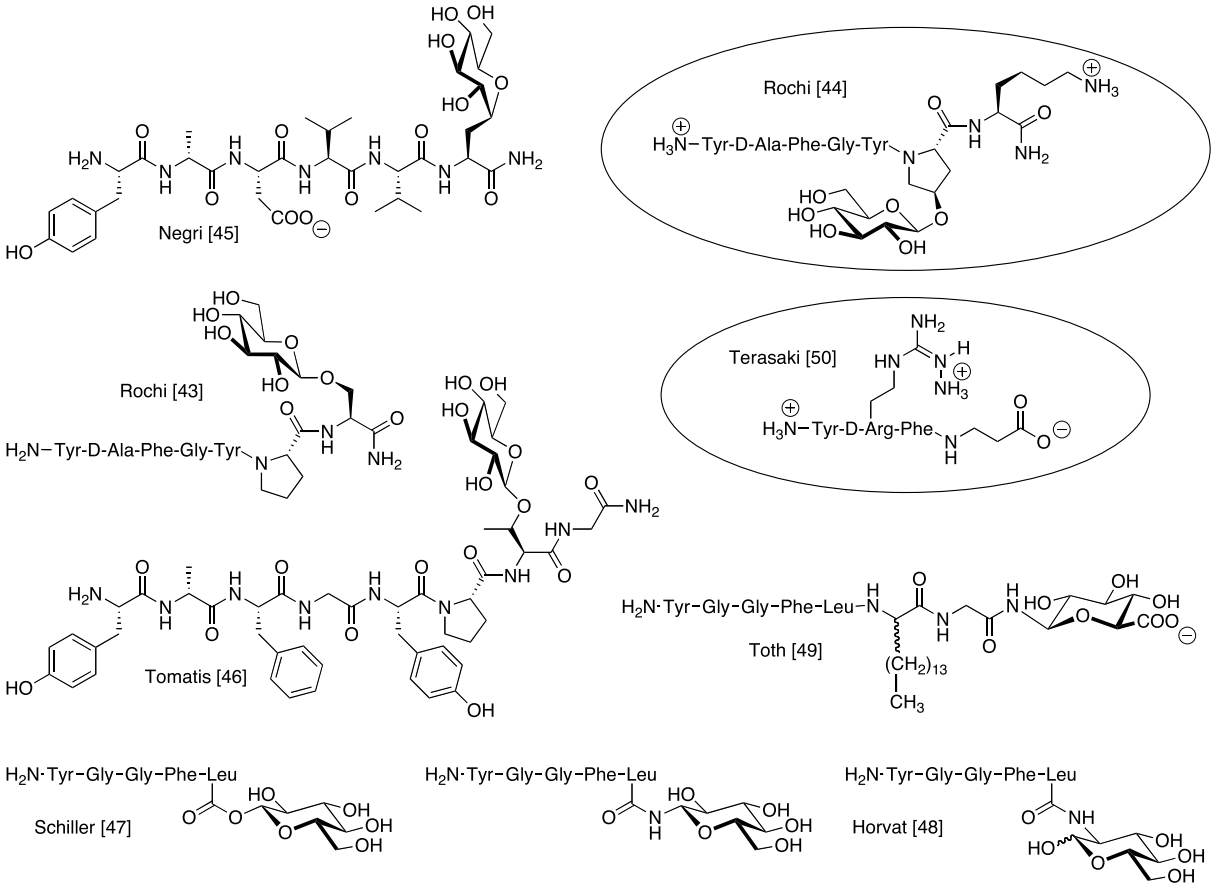


Figure 8
Other peptide-based opioids

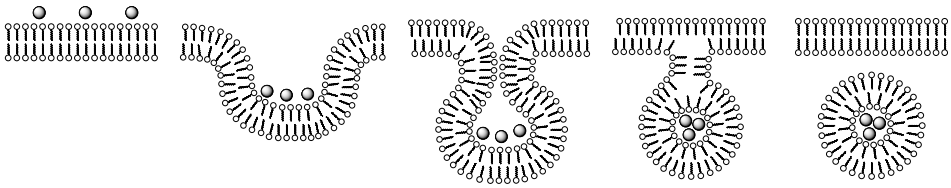


Figure 9
Endocytosis of glycopeptide drugs

only consists of the endothelial barrier. Thus, CSF is created in the choroid plexus and flows around the brain and down to the spinal column. Complete turnover of CSF takes about 5 h in humans. The BBB presents a physical obstacle, and also a metabolic one as well, possessing both oxidative enzymes and peptidases, including enkephalinase. Thus, peptides are normally rapidly degraded before they reach the CSF. This enzymatic gauntlet may be as important as the lipophilic barrier in excluding peptide pharmaceuticals from the CNS. It should also be noted that entry to the CSF does not guarantee that a drug will enter the brain, as many peptides are rapidly exported back to the bloodstream. For a complete discussion of the BBB, or more properly, the neurovascular unit, see the review by Hawkins and Davis [53].

We postulate that BBB penetration of the glycopeptides is initiated by an endocytotic event, (*Fig. 9*) and can probably best be described as adsorptive transcytosis [4,5,15]. The peptide E-2078, an analog of Dynorphin1-8, was shown to cross the BBB via adsorptive mediated endocytosis and transcytosis [54]. Current studies in our lab and our collaborators' labs are designed to more fully elucidate the mechanism of BBB transport of our drug candidates. Studies with much larger dynorphin or endorphin analogues with amphipathic helix address regions indicate that much larger glycopeptides (>2000 M.W.) can penetrate the BBB, as well [55].

4 Drugs for the Brain from the Brain

Outlook. Pain is ubiquitous in medical practice, and analgesics are essential tools for any clinician. While there has been much research into new applications and new formulations of long-standing opioid drugs, it is surprising that very few new drug entities have come into the marketplace for the treatment of pain in decades. The past three decades of research with enkephalin peptide analogues have generated many potent and selective opiate ligands, which have proved invaluable for research in the mechanisms of pain transmission. Use of these ligands as *pharmaceuticals* and not only as biochemical tools has been advanced by an understanding of the factors that govern their stability in serum and in the brain, and how they can penetrate the BBB. Our prediction is that as the greater pharmacological community becomes aware of the hard-won advances in this field, and begins to gain an appreciation of the potential benefits offered by peptide-based analgesics, that the problems associated with the use of peptide and glycopeptide pharmaceuticals will be rapidly solved, and that glycopeptides will find widespread clinical application, particularly in the CNS.

Key to the solution of these problems is the realization of the role of the cell-surface membrane in altering the conformation of even short glycopeptides. It seems certain that a better

understanding of the interaction of glycopeptide amphipaths with membranes could lead to a new pharmacology based on endogenous neuropeptides that have already evolved to agonize or antagonize particular receptors (nature's inherent design). Over 250 endogenous peptides have already been identified that bind to some 50 different receptors in the CNS. What is suggested here are the first necessary steps toward understanding the interaction of glycopeptides with membranes, which is key to both receptor binding and BBB transport. Thus, drugs *for* the brain can be derived *from* the brain.

Acknowledgement

We would like to thank the National Science Foundation, the Office of Naval Research, and the National Institutes of Health (NINDS) for support of our work with glycopeptides. We would also like to thank numerous collaborators in both industry and academia, but especially Edward J. Bilsky, Richard D. Egleton, Peg Davis, Frank Porreca, Henry I. Yamamura, Jean M. Bidlack, S. Steven Negus, and Victor J. Hruby for their invaluable assistance.

References

1. Hughes J, Smith TW, Kosterlitz HW, Fothergill LA, Morgan BA, Morris HR (1975) Identification of two related pentapeptides from the brain with potent agonist activity. Nature (Lond) 258:577–579
2. Fanelli F, De Benedetti PG (2005) Computational Modeling Approaches to Structure-Function Analysis of G Protein-Coupled Receptors. Chem Rev 105:3297–3351
3. Kane BE, Svensson B, Ferguson DM (2006) Molecular Recognition of Opioid Receptor Ligands. The AAPS J 8:E126–E137 http://www.aapsj.org
4. Muthu D, Polt R (2005) New Prospects for Glycopeptide Based Analgesia: Glycoside-Induced Penetration of the Blood-Brain Barrier. Curr Drug Deliv 2:59–73
5. Muthu D, Keyari CM, Polt R (2005) Glycosylated Neuropeptides: A New Vista for Neuropsychopharmacology. Med Res Rev 25:557–585
6. Gutstein HB, Akil H (2001) Opioid Analgesics. In: Hardman JG, Limbird LE, Gilman AG (eds) Goodman & Gilman's The Pharmacological Basis of Therapeutics, 10th ed. McGraw-Hill, New York, pp. 569–619
7. Evans CJ (2004) Secrets of the opium poppy revealed. Neuropharmacol 47:203–299
8. Elmgabari NO, Egleton RD, Palian MM, Lowey JJ, Schmid WR, Davis P, Navratilova E, Dhnasekaran M, Keyari CM, Yamamura HI, Porreca F, Hruby VJ, Polt R, Bilsky EJ (2004) Antinociceptive Structure-Activity Studies with Enkephalin-Based opioid Glycopeptides. J Pharmacol Exp Ther 311:290–297
9. Gengo PJ, Chang K-J (2004) Mixed Opioid Receptor Agonists as a New Class of agents for the Treatment of Moderate to Severe Pain. In: Chang K-J, Porreca F, and Woods JH (eds) The Delta Receptor. Marcel Dekker, New York, pp. 231–244
10. Stevenson GW, Folk JE, Linsenmayer DC, Kenner C, Rice KC, Negus SS (2003) Opioid Interactions in Rhesus Monkeys: Effects of $\delta + \mu$ and $\delta + \kappa$ Agonists on Schedule-Controlled Responding and Thermal Nociception. J Pharm Exp Therap 307:1054–1064
11. Dykstra LA, Granger AL, Allen RM, Zhang X, Rice KC (2002) Antinociceptive effects of the selective delta opioid agonist SNC80 alone and in combination with mu opioids in the squirrel monkey titration procedure. Psychopharmacology 163:420–429
12. Weltrowska G, Lemieux C, Chung NN, Schiller PW (2004) A chimeric opioid peptide with mixed μ agonist / δ antagonist properties. J Peptide Res 63:63–68
13. Daniels DJ, Lenard NR, Etienne CL, Law PL, Roerig SC, Portoghese PS (2005) Opioid-induced

tolerance and dependence in mice is modulated by the distance between pharmacophores in a bivalent ligand series. Proc Natl Acad Sci USA 102:19208–19213
14. Muthu D, Keyari CM, Polt R (2005) Glycosylated Neuropeptides: A New Vista for Neuropsychopharmacology. Med Res Rev 25:557–585
15. Egleton RD, Bilsky EJ, Tollin G, Muthu D, Lowery J, Alves I, Davis P, Porreca F, Yamamura HI, Yeomans L, Keyari CM, Polt R (2005) Biousian glycopeptides penetrate the blood-brain barrier. Tetrahedron Asymm 16:65–75
16. Osborne R, Joel S, Trew D, Slevin M (1988) Analgesic Activity of Morphine-6-Glucuronide. Lancet 1:828
17. Bourasset F, Cisternino S, Temsamani J, Scherrmann JM (2003) Evidence for an active transport of morphine-6-beta-D-glucuronide but not P-glycoprotein-mediated at the blood-brain barrier. J Neurochemistry 86:1564–1567
18. Rodriguez RE, Rodriguez FD, Sacristan MP, Torres JL, Valencia G, Garcia-Anton JM (1989) New glycosylpeptides with high antinociceptive activity. Neurosci Lett 101:89–94
19. Polt R, Porreca F, Szabó LZ, Bilsky EJ, Davis P, Abbruscato TJ, Davis TP, Horvath R, Yamamura HI, Hruby VJ. (1994) Glycopeptide enkephalin analogues produce analgesia in mice: Evidence for penetration of the blood-brain barrier. Proc Natl Acad Sci 91:7114–7118
20. Williams SA, Abbruscato TJ, Szabo L, Polt R, Hruby VP, Davis TP (1996) The Effect of Glycosylation on the Uptake of an Enkephalin Analogue into the Central Nervous System. In: Couraud & Scherman (eds) Biology and Physiology of the Blood-Brain Barrier. Plenum Press, New York, pp. 69–77
21. Garcia-Alvarez I, Garrido L, Fernandez-Mayoralas A (2007) Studies on the uptake of glucose derivatives by red blood cells. Chem Med Chem 2:496–504
22. Smith MT, South SM (2001) The role of morphine-6-glucuronide (M6G) in pain control. Pain Rev 8:171–191
23. Romberg R, van Dorp E, Hollander J, Kruit M, Binning A, Smith T, Dahan A (2007) A randomized, double-blind, placebo-controlled pilot study of IV morphine-6-glucuronide for postoperative pain relief after knee replacement surgery. Clin J Pain 23:197–203
24. Guo XL, Meiyu G, Guanhua D (2005) Glucose transporter 1, distribution in the brain and in neural disorders: Its relationship with transport of neuroactive drugs through the blood-brain barrier. Biochem Gen 43:175–187
25. Polt R, Mitchell SA (2000) Chemical Biology and Biomedicine: Enkelphalin-Derived Glycopeptide Analgesics. In: Fraser-Reid B, Tastuta K, Thiem J (eds) Glycoscience-Chemistry and Chemical Biology. Springer, Berlin Heidelberg New York, Vol. 3, Chapter 8.4
26. Horvat S (2001) Opioid Peptides and their Glycoconjugates: Structure-Activity Relationships. Curr Med Chem 1:133–154
27. Horvat J, Horvat S, Lemieux C, Schiller PW (1988) Synthesis and Biological Activity of [Leu^5] enkephalin Derivatives Containing D-Glucose. Int J Peptide Protein Res 31:499–507
28. Palian MM, Polt R (2001) Glycopeptide Analgesics. Drugs of the Future 26:561–576
29. Deaton KR, Feyen EA, Nkulabi HJ, Morris KF (2001) Pulsed-field gradient NMR study of sodium dodecyl sulfate micelle–peptide association. Magn Reson Chem 39:276–282
30. Chatterjee C, Majumder B, Mukhopadhyay C (2004) Pulsed-field gradient and saturation transfer difference NMR study of enkephalins in the ganglioside GM_1 micelle. J Phys Chem B 108:7430–7436
31. Watts CR, Tessmer MR, Kallick DA (1995) Structure of Leu^5-Enkephalin Bound to a Model Membrane as Determined by High-Resolution NMR. Lett Peptide Sci 2:59–70
32. Adam G, Delbruck M (1968) In: Rich R, Davidson N (eds) Structural Chemistry and Molecular Biology. Freeman & Co., San Francisco, p. 198
33. Sargent DF, Schwyzer R (1986) Membrane lipid phase as catalyst for peptide receptor interactions. Proc Natl Acad Sci USA 83:5774–5778
34. Lowery JJ, Yeomans L, Keyari CM, Davis P, Porreca F, Knapp BI, Bidlack JM, Bilsky EJ, Polt R (2007) Glycosylation Improves the Central Effects of DAMGO. Chem Biol Drug Discov 69:41–47
35. Palian MM, Boguslavsky VI, O'Brien DF, Polt R (2003) Glycopeptide-Membrane Interactions: Glycosyl Enkephalin Analogues Adopt Turn Conformations by NMR and CD in Amphipathic Media. J Am Chem Soc 125:5823–5831
36. Guile GR, Rudd PM, Wing DR, Prime SB, Dwek RA (1996) A rapid high-resolution high-performance liquid chromatographic method for separating glycan mixtures and analyzing oligosaccharide profiles. Anal Biochem 240:210–226

37. Kriss CT, Lou BS, Szabo LZ, Mitchell SA, Hruby VJ, Polt R (2000) Enkephalin-Based Drug Design: Conformational Analysis of *O*-Linked Glycopeptides by NMR and Molecular Modeling. Tetrahedron Asymm 11:9–25
38. Gacel G, Daugé V, Breuzé P, Delay-Goyet P, Roques BP (1988) Development of conformationally constrained linear peptides exhibiting a high affinity and pronounced selectivity for δ opioid receptors. J Med Chem 31:1891–1897
39. Bilsky EJ, Egleton RD, Mitchell SA, Palian MM, Davis P, Huber JD, Jones H, Yamamura HI, Janders J, Davis TP, Porreca F, Hruby VJ, Polt R (2000) Enkephalin glycopeptide analogues produce analgesia with reduced dependence liability. J Med Chem 43:2586–2590
40. Pratt MR, Bertozzi CR (2005) Synthetic glycopeptides and glycoproteins as tools for biology. Chem Soc Rev 34:58–68
41. Mitchell SA, Pratt MR, Hruby VJ, Polt R (2001) Solid-phase synthesis of *O*-linked glycopeptide analogues of enkephalin. J Org Chem 66:2327–2342
42. Bilsky EJ, Bernstein RN, Wang ZJ, Sadee W, Porreca F (1996) Effects of neutral and negative antagonists and protein kinase inhibitors on acute morphine dependence and antinociceptive tolerance in mice. J Pharmacol Exp Ther 277:484–490
43. Negri L, Lattanzi R, Tabacco F, Scolaro B, Rocchi R (1998) Glycodermorphins: opioid peptides with potent and prolonged analgesic activity and enhanced blood-brain barrier penetration. Brit J Pharmacol 124:1516–1522
44. Biondi L, Filira F, Giannini E, Gobbo M, Lattanzi R, Negri L, Rocchi R (2007) Novel glycosylated [Lys7]-dermorphin analogues: synthesis, biological activity and conformational investigations. J Pept Sci 13:179–189
45. Negri L, Lattanzi R, Tabacco F, Orru L, Severini C, Scolaro B, Rocchi R (1999) Dermorphin and deltorphin glycosylated analogues: Synthesis and antinociceptive activity after systemic administration. J Med Chem 42:400–404
46. Tomatis R, Marastoni M, Balboni G, Guerrini R, Capasso A, Sorrentino L, Santagada V, Caliendo G, Lazarus LH, Salvadori S (1997) Synthesis and pharmacological activity of deltorphin and dermorphin-related glycopeptides. J Med Chem 40:2948–2952
47. Horvat S, Horvat J, Varga-Defterdarovic L, Pavelic K, Chung NN, Schiller PW (1993) Methionine-Enkephalin Related Glycoconjugates. Int J Peptide Protein Res 41:399–404
48. Roscic M, Horvat S (2006) Transformations of bioactive peptides in the presence of sugars – Characterization and stability studies of the adducts generated via the Maillard reaction. Bioorg Med Chem 14:4933–4943 (vide penetus)
49. Kellam B, Drouillat B, Dekany G, Starr MS, Toth I (1998) Synthesis and in vitro evaluation of lipoamino acid and carbohydrate-modified enkephalins as potential antinoceptive agents. Int J Pharm 161:55–64
50. Deguchi Y, Miyakawa Y, Sakurada S, Naito Y, Motimoto K, Ohtsuki S, Hosoya K, Terasaki T (2003) Blood-brain barrier transport of a novel μ_1 specific opioid peptide, H-Tyr-D-Arg-Phe-β-Ala-OH (TAPA). J Neurochem 84:1154–1161
51. Crone C (1963) The permeability of capillaries in various organs as determined by the use of the "indicator diffusion" method. Acta Physiol Scand 58:292–305
52. Dohrmann GJ (1970) The choroid plexus: a historical review. Brain Res 18:197–218
53. Hawkins BT, Davis TP (2005) The Blood-Brain Barrier/Neurovascular Unit in Health and Disease. Pharmacol Rev 57:173–185
54. Terasaki T, Hirai K, Sato H, Kang Y-S, Tsuji A (1989) Absorptive–mediated endocytosis of a dynorphin-like analgesic peptide, E2078, into blood-brain barrier. J Pharmacol Exp Ther 251:351–357
55. Muthu D, Palian MM, Alves I, Yeomans L, Keyari CM, Davis P, Bilsky EJ, Egleton RD, Yamamura HI, Jacobsen NE, Tollin G, Hruby VJ, Porreca F, Polt R (2005) Glycopeptides Related to β-Endorphin Adopt Helical Amphipathic Conformations in the Presence of Lipid Bilayers. J Am Chem Soc 127:5435–5448

12.6 Antitumor and Antimicrobial Glycoconjugates

Thisbe K. Lindhorst
Otto Diels-Institut für Organische Chemie, Christian-Albrechts-Universität zu Kiel, 24098 Kiel, Germany
tklind@oc.uni-kiel.de

Abstract

Carbohydrates are found in nature in manyfold conjugated ways and in addition, they are conjugated to numerous carrier molecules by synthetic means. This chapter outlines the structures and most important biological and organic aspects of glycosylated natural products which serve as antimicrobial drugs as well as antitumor therapeutics. In addition, modern concepts of synthetic glycosciences leading to unusual glycoconjugates are surveyed in the last two sections, covering glycoconjugates derived from sugar amino acids as well as functional glycoconjugates, such as multivalent derivatives or those equipped with a functional tether.

In: *Glycoscience*. Fraser-Reid B, Tatsuta K, Thiem J (eds)
Chapter-DOI 10-1007/978-3-540-30429-6_66:

Keywords

Glycoconjugates; Antitumor and antimicrobial drugs; Deoxysugars; Molecular recognition; Antibiotic resistance; Antiadhesives; Sugar amino acids; Functional glycoconjugates

Abbreviations

AFGPs antifreeze glycoproteins
CPO chorion peroxidase
SAA sugar amino acid

1 Introduction

Molecules, in which carbohydrates are covalently bound to proteins or lipids, have been termed glycoconjugates. They are a fascinating group of molecules for chemists and biochemists because of the enormous number and variability of the structures in which they occur as well as the wide array of their biological functions. Consequently, it has been difficult to classify these molecules in a comprehensive way, the more so since new carbohydrate-containing compounds with new linkage types are still being found. For example, to the rich world of *O*- and *N*-linked glycoproteins other types of carbohydrate-peptide junctions have to be added such as an α-*C*-glycosidic linkage of D-mannose bound to the C2 of a tryptophan (Trp) amino acid residue (*Fig. 1*) [1,2], which has been discovered on human ribonuclease (RNase) 2. NMR studies have shown that the mannose moiety in native RNase 2 adopts different orientations around its *C*-linkage compared to the denatured protein. Thus the three-dimensional structure of the protein apparently induces a specific orientation of the mannose residue.

In RNase 2 the main structural roles of the *C*-linked mannose seem to be to stabilize the *N*-terminal loop of the protein and to keep Trp7 in a specific orientation, compared to the non-*C*-glycosylated form [3]. Site-directed mutagenesis has revealed that the sequence Trp-x-x-Trp is required for the *C*-mannosylation of the first tryptophan [4]. This *C*-glycosidic linkage has been proved to occur also in a number of other mammalian proteins, such as interleukin-12, complements, properdin, thrombospondin, erythropoietin receptor, mucins, and a bovine lens protein [5,6,7,8,9,10,11]. It has been demonstrated that the *C*-mannosylation is an enzyme-catalyzed event [12].

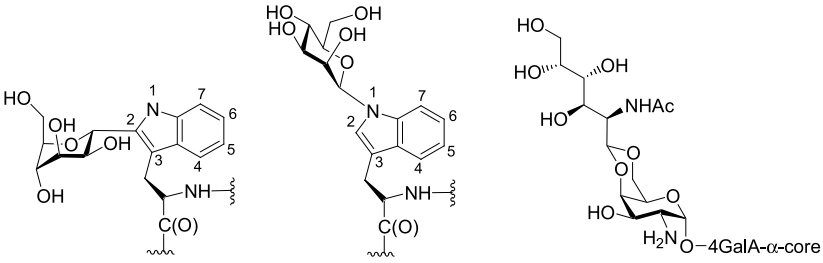


Figure 1
Unusual glycoconjugates

Recently an *N*-mannosylated tryptophan residue has been described for *Aedes aegypti* chorion peroxidase (CPO) [13] which plays a crucial role in chorion hardening by catalyzing chorion protein cross-linking through dityrosine formation. This enzyme is extremely resistant to denaturing conditions, which seems intimately related to its posttranslational modifications, including glycosylation. Other than in RNAse 2 mannose is linked to the N1 atom of the indole ring of Trp residue (*Fig. 1*). Three of the seven available Trp residues in mature CPO are partially or completely mannosylated.

Also in the field of glycolipids unusual types of glycosidic linkage have been discovered [14], such as an open-chain acetal-linked *N*-acetylgalactosamine in the core part of the lipopolysaccharides from *Proteus* microorganisms (*Fig. 1*).

A growing number of special glycoconjugate structures with unusual functions are known such as the antifreeze glycoproteins [15] found in antarctic and arctic fish families. These compounds have the ability to depress the freezing point of the organism such that it can survive the subzero temperatures encountered in its environment. Their antifreeze activity is lowering the freezing temperature in organs around 500 times more than predicted from solution colligative properties. Antifreeze glycoproteins (AFGPs) can be isolated as a heterogenous mixture of compounds from different marine fish, which makes a good characterization difficult. They have a simple, repeating unit glycoprotein structure, which can be expressed as Ala-[Ala-(Galβ1→3GalNAcα1)Thr-Ala]-Ala. AFGP 1 is the largest fraction (33.7 kDa) and AFGP 8 (2.6 kDa, *Fig. 2*) is the lowest molecular weight fraction. Molecular cloning does not allow us to produce the glycosylated antifreeze proteins, while their chemical synthesis has been successfully carried out [16].

The physical properties of AFGPs are very attractive for the cryopreservation of cells, tissues, and organs. Recently, a functional *C*-linked analogue of native AFGP 8 has been synthesized and tested in vitro [17]. These studies have revealed that AFGP 8 is cytotoxic to human embryonic liver and human embryonic kidney cells at concentrations higher than 2 and 0.63 mg/mL, respectively, whereas in contrast, the synthetic *C*-linked AFGP analogue displayed no in vitro cytotoxicity even at high concentrations (*Fig. 2*).

A large amount of natural products, however, contain carbohydrate moieties attached to other molecules than proteins and lipids in a variety of ways [18]. All of these compounds have in common that they are composed of an aglycon and one or more appended saccharide units. They are classified according to the nature of their aglycon moiety or are divided into categories depending on their biological activity. As such glycosylated natural products very often

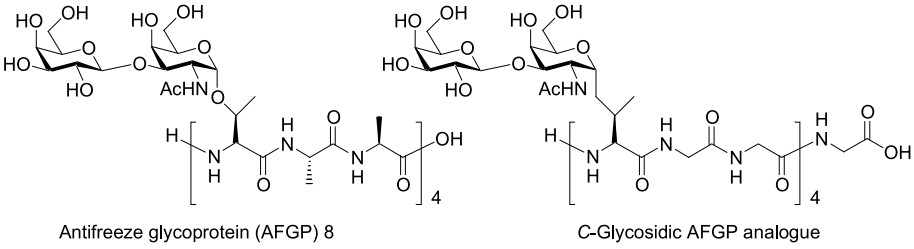


Figure 2
Glycoconjugates with antifreezing properties

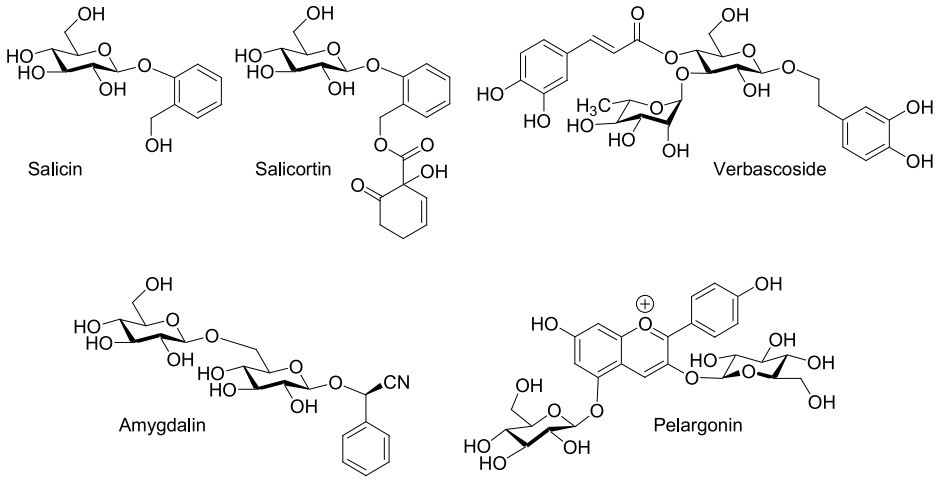


Figure 3
'Historical' glycoconjugates

exhibit activities against a variety of microbes and human tumors, they have attracted significant attention in medicine and pharmaceutical industry [19]. Some of them are of low molecular weight having structures and activities which have been identified much earlier than when the importance of cell membrane glycoproteins and glycolipids was elucidated. Salicin for example (*Fig. 3*), is the β-glucoside of salicylic alcohol and has been found in the middle of the 19th century to be responsible for the pain relieving effects of the bark of the willow- and the poplar-tree. These properties have been known since the times of Hippocrates and include antipyretic, analgesic, antiphlogistic, and antirheumatic activities. Salicortin A (*Fig. 3*) is a derivative of salicin, which also occurs as an ingredient of willow-tree bark and protects these plants against insects. Phenylethyl glycosides represent another example of relatively simple molecules, in which a 2-phenylethyl β-D-glucopyranoside is functionalized with an aromatic acid, e. g. cinnamic acid, caffeic acid, or ferulic acid [20]. They are an interesting group of natural products which are widely distributed in the plant kingdom. Verbascoside (*Fig. 3*) is a typical representative, belonging to the largest class of phenylethyl glycosides, those with a rhamnose residue attached to the 3-position of the central glucose unit. It has been established that verbascoside is a potent inhibitor of protein kinase C and of aldose reductase. Furthermore, it possesses antibacterial, antiviral, and antitumor activity, as well as cytotoxic and immunomodulatory properties. A total synthesis of verbascoside has been reported [21].

A large variety of relatively simple glycosylated cyanohydrins are known as cyanogenics, that is, they liberate hydrogen cyanide. The poisonous properties of some HCN-liberating plants were known already in ancient Egypt and liberated HCN was first detected and reported in 1802 by the pharmacist Bohm in Berlin. Amygdalin (*Fig. 3*) is a famous example of this class of cyanogenic glycosides, which was first synthesized in 1924 [22].

Anthocyanidin glycosides, on the other hand are interesting color pigments in plants. All glycosylated anthocyanidins are called anthocyanins and named according to the exact substitution pattern of the flavylium aglycon. Pelargonin (*Fig. 3*) has an orange color in acidic

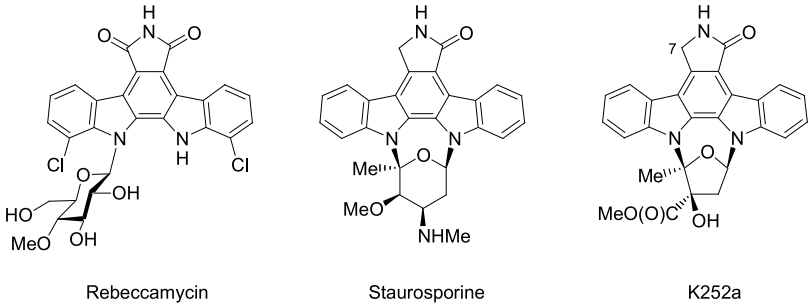


Figure 4
Secondary metabolites based on carbohydrates

solution. Color stabilization and variation of anthocyanins seems to originate from the hydrophobic interaction of the aromatic moieties in aqueous solution as well as from slight pH changes [23].

The majority of bioactive glycoconjugates have been found to be biosynthesized by bacteria and other microorganisms. Over the last 50 years, many molecules which appear to be dispensable for the primary metabolic activities of microorganisms such as growth, have been isolated and their structures have been elucidated; they have been called "secondary metabolites". Secondary metabolites from microorganisms are of substantial interest for the field of drug discovery and for pharmaceutical enterprises. Compared to plant derived metabolites, bioactive compounds from microorganisms are accessible in unlimited quantities by in vitro-bioprocessing technologies. Rebeccamycin for instance (*Fig. 4*) [24,25,26,27] was isolated from *Nocardia aerocolonigenes*, a member of the ascomycetales family. It consists of a symmetrical indolocarbazole chromophore which is glycosylated with a 4-*O*-methyl-glucoside and shows strong antitumor activity. It is structurally related to staurosporine (*Fig. 4*), an indole alkaloid produced by *Streptomyces staurosporeus*, which is effective as an antimycotic and inhibits blood coagulation. Furthermore, staurosporin is a potent, however unspecific inhibitor of protein kinases. Protein kinases are vital to the signal transduction pathways which carry external signals to the nucleus of the cell. The disruption of cellular signal transduction via protein kinase malfunction has been related to the onset of several disease states, including rheumatoid arthritis, systemic lupus, erythematosis, diabetes melitus, Alzheimer's disease and cancer. Highly specific inhibitors of protein kinases are sought as valuable tools to study the cellular roles of individual protein kinases and also seem to be of value for chemotherapeutic intervention.

Searching for more specific inhibitors of protein kinases, numerous derivatives of staurosporine have been isolated as well as synthesized. Among those, a furanosylated congener, K252a (*Fig. 4*), has been isolated and found to inhibit protein kinase C with an IC_{50} value of 32 nanomolar. The synthesis of C(7)-methyl derivatives of K252a has been reported [28].

Synthetic means have often been successfully used to improve specificities and activities of naturally occurring drugs which are taken as lead structures. Also molecular hybrids [29], combining structural elements of different classes of natural products have offered advantages, such as synthetic 2-phenylchinolin-carbohydrate hybrids, which were shown to cleave DNA on photoinduction [30].

Many other glycoconjugates have much more complex structures than the molecules mentioned so far in this chapter. They have been classified according to the chemical nature of their aglycon moieties and a collection of these compounds will be discussed in the following sections. The elucidation of their structures, biosynthesis and biological properties has been equally fascinating and difficult. Moreover, the interest in the discovery of even new classes of glycosylated natural products and their synthesis [31,32] has become increasingly important in view of the increasing antibiotic resistances of bacteria [33]. While the era of antibiotics, which has been started in the 1930s with the introduction of penicillin and the sulfonamides, suggested a pharmacological victory over formerly lethal bacterial infections, this vision is fading away nowadays. Since the first clinical use of antibiotics in the Second World War, a variety of novel antibacterial drugs from microorganisms have been introduced and basically against all of them bacteria have developed resistances. This experience has not stopped researchers from seeking new antimicrobial drugs [34] and indeed some of the most recently investigated compounds form the last line of defense against the most dangerous multi-resistant bacterial strains. In addition to their antimicrobial activities these compounds are highly attractive because of their potent activity against a variety of human tumors.

The bacterial antibiotics, herein discussed, contain a rich spectrum of deoxysugars [35] and aminosugars, often having branched chains. These exhibit a great range of activities; target binding and specificity functions have been ascribed to the deoxysugar components of numerous bacterial antibiotics such as in the case of aureolic acids, anthracyclines, or enediynes. The deoxysugar components of these compounds add a great deal to structure-activity relationships.

Presumably, bacterial infections cannot be treated by solely an armada of drugs and even further new, active natural products. A better understanding of their biology and of biosynthesis of the involved molecules, including the biosynthesis of the deoxysugar portions [36], and gaining more knowledge about control of biosynthesis, the related gene clusters and related biochemical effects, will eventually lead to alternatives in antimicrobial treatment. Microbiological modifications of microorganisms have led to modified secondary metabolites with an altered spectrum of activities. This has been especially successful with a group of Gram-positive bacteria belonging to the family actinomycetales, particularly the streptomycetes, which are especially prolific producers of secondary metabolites, many of which show antibiotic activity. Various strategies of "antibiotic pharmacobiotechnology", including mutagenesis, biotransformation, mutational biosynthesis and precursor feeding have been applied to actinomycetes to alter the spectrum of metabolites formed with the aim to produce compounds with high potency or a broad activity spectrum. Among these efforts, the selective mutagenesis of the producer of the avermectins, represents one of the successful examples [37].

Thus, methods of microbial chemistry have complemented sophisticated organic chemistry and led to a variety of half-synthetic products against which no resistances have been developed as yet. Also numerous mimetics of active carbohydrate structures and glycoconjugates have been synthesized, among which antiadhesive multivalent glycoconjugates and functionalized glycoconjugates form promising groups of molecules in drug research. Therefore, approaches in which carbohydrates have been modified and tested as molecular tools as vaccines or to inhibit microbial adhesion to their host cells, for example, will be mentioned in the last section.

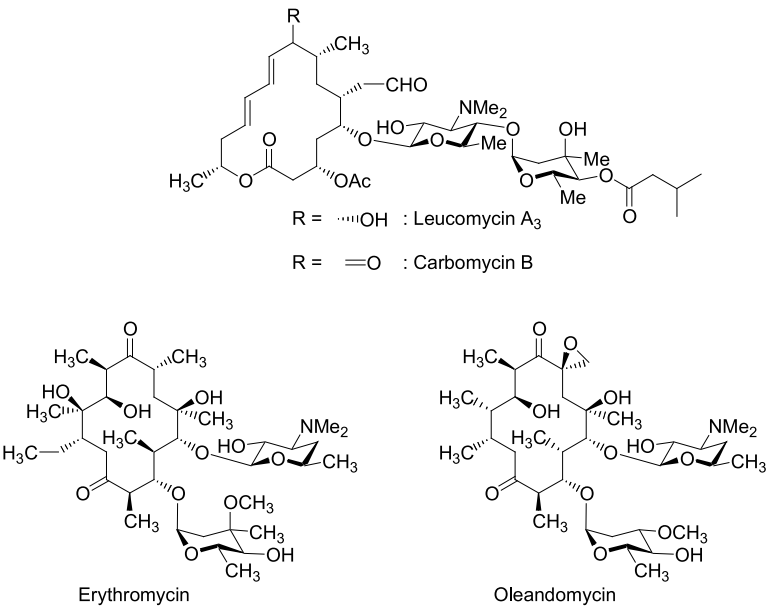
2 Structures, Occurrence and Properties of Antitumor and Antimicrobial Glycoconjugates

2.1 Macrolide Antibiotics

Macrolide antibiotics are glycosides with a macrocyclic lactone aglycon, which is formed in the polyketide biosynthetic pathway [38,39]. The lactone ring is 12-, 14-, 16-, or 18-membered. The polyoxo-macrolides such as the classical antibiotic erythromycin are produced by *streptomyces* microorganisms and form a clinically important group of polyketide antibiotics. Erythromycin (● *Fig. 5*) is the major component out of a mixture of macrolide antibiotics which is formed by *Saccharopolyspora erythraea*. It contains two deoxysugars attached to the aglycon, L-cladinose and D-desosamine.

More than 500 different representatives of the macrolide antibiotics are known, most of which are biologically active against Gram-positive bacteria, displaying a relatively low toxicity. Clinically used are erythromycin, oleandomycin, carbomycin and leucomycin (● *Fig. 5*). They act as inhibitors of the bacterial protein biosynthesis by binding to the 50S-ribosomal subunit. The synthesis of the two clinically important 16-membered ring macrolide antibiotics leucomycin A3 and carbomycin B could be started from D-glucose, which was chosen because it contained three of the required stereocenters [40].

Macrolide antibiotics contain a wide spectrum of deoxysugars which are biosynthesized on different pathways [41]. For chlorothricin (● *Fig. 6*) for example, the stereochemical course



● Figure 5
Glycosylated macrolide antibiotics

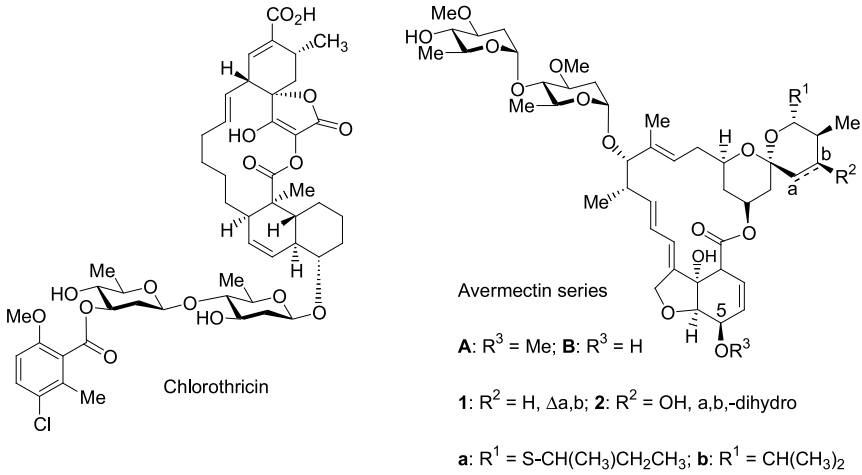


Figure 6
Macrolide glycoconjugates with polycyclic aglycon

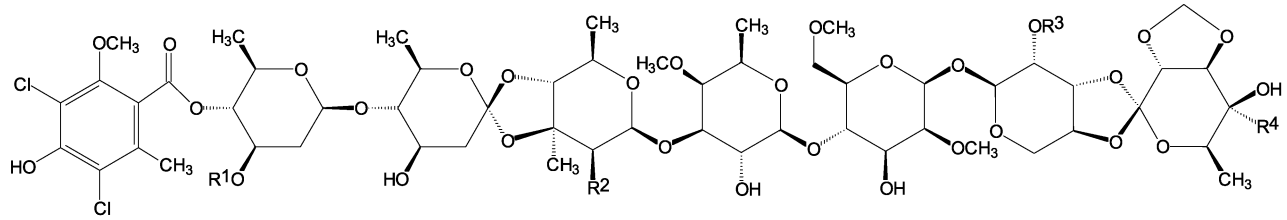
of the replacement of the 2-hydroxy group by hydrogen in the biosynthesis of the 2,6-dideoxy-sugar moieties (β-D-olivose) was shown to proceed with inversion of configuration at C-2 by feeding deuterated glucose to chlorothricin-producing *Streptomyces antibioticus* Tü99 [42]. The chemical synthesis of macrolide antibiotics has been reviewed [43].

The avermectins [44,45] represent a special class of macrolide antibiotics, comprising a complex of 16-membered pentacyclic lactones containing two L-oleandrose units attached to the aglycon. The avermectins are produced and isolated from *Streptomyces avermitilis* and exhibit broad insecticidal, ascaricidal, and antithelmintic activity even at very low concentrations without showing any antibacterial or antifungal activity. They interact in a unique manner with the GABA (γ-aminobutyric acid)-regulated chloride channels. The complex is composed of eight related components (*Fig. 6*), four major (**A1a**, **A2a**, **B1a** and **B2a**) and four minor ones (**A1b**, **A2b**, **B1b** and **B2b**).

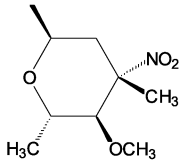
It has been shown for avermectin B that the hydroxyl group at C-5 and a disaccharide moiety are essential for antiparasitic, mitocidal and selective insecticidal activities. Chemical modification of the macrolide-type avermectins resulted in ivermectin which was commercialized to be used in agriculture and animal health care as well as in human medicine.

2.2 Orthosomycins

The orthosomycin antibiotics [46] are structurally characterized by one or more interglycosidic spiro-ortholactone linkages, replacing the traditional acetal junctions [47]. The natural occurrence of orthoesters is rare, but within recent years several examples of antibiotics possessing this common structural feature have been described. The orthosomycins include flambamycin, the everninomicins-B, -C, -D and –2, hygromycin, the destomycins-A, -B and -C and the antibiotics SS-56-C and A-396-I (*Fig. 7*). The orthosomycins have been divided into two



sug =



Flambaymycin: $R^1 = H, R^2 = OH, R^3 = COCH(CH_3)_2, R^4 = COCH_3$

Everninomicin B: $R^1 = sug, R^2 = OH, R^3 = CH_3, R^4 = CH(OCH_3)CH_3$

Everninomicin C: $R^1 = sug, R^2 = H, R^3 = CH_3, R^4 = CH(OCH_3)CH_3$

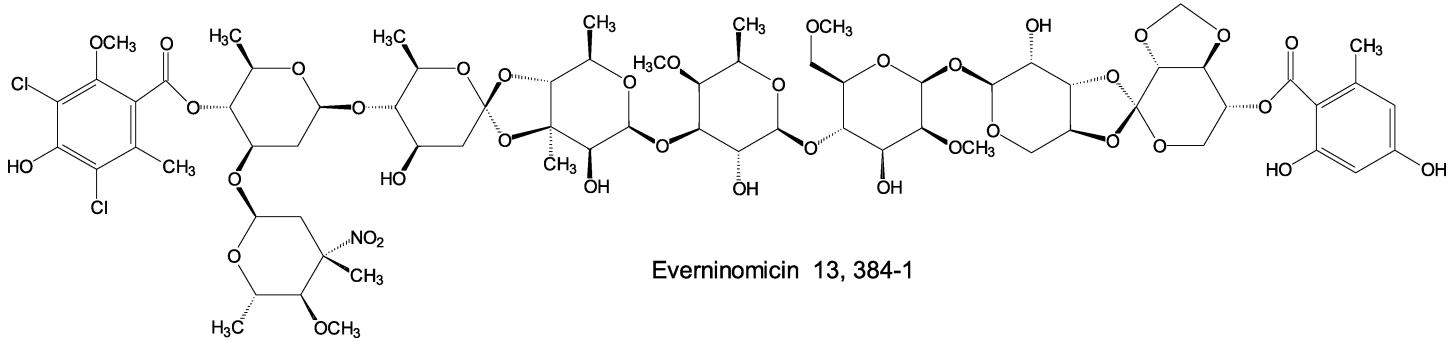
Everninomicin D: $R^1 = sug, R^2 = H, R^3 = CH_3, R^4 = (S)\text{-}CH(OCH_3)CH_3$

Everninomicin 2: $R^1 = H, R^2 = H, R^3 = CH_3, R^4 = (S)\text{-}CH(OCH_3)CH_3$

Avilamycin A: $R^1 = H, R^2 = H, R^3 = COCH(CH_3)_2, R^4 = COCH_3$

Avilamycin C: $R^1 = H, R^2 = H, R^3 = COCH(CH_3)_2, R^4 = (S)\text{-}CH(OCH_3)CH_3$

Curamycin A: $R^1 = H, R^2 = H, R^3 = COCH_3, R^4 = COCH_3$



Everninomicin 13, 384-1

Figure 7
Structures of the orthosomycins

distinct series on the basis of additional structural features, namely those which are esters of dichloroisoeverninic acid, for example flambamycin and the everninomicins and those which contain an aminocyclitol residue, for example hygromycin B and the destomycins. Orthosomycin antibiotics have been isolated for example from several streptomyces species.

The avilamycin antibiotic complex is comprised of 16 components, synthesized from *Streptomyces viridochromogenes*. Avilamycins as many other orthosomycins are active against Gram-positive bacteria, however advantageously show no toxicity or mutagenicity, as it was found for a number of other orthosomycins. The antibacterial effect was shown to operate via protein synthesis perturbation. By blocking of the binding site of aminoacyl-tRNA, a ribosomal attachment is inhibited. This results in an interruption of the translation step and thus a defect in peptide processing.

Everninomicin 13,384–1 (◉ *Fig. 7*) has recently become an important synthetic target as it has been shown to be effective against methicillin-resistant staphylococci as well as against vancomycin-resistant streptococci and enterococci and therefore, represents a promising new drug [48,49,50]. It was first isolated from *Micromonospora carbonacea var. africana*, found in Kenya and carries a novel oligosaccharide structure with two sensitive orthoester moieties. The oligosaccharide chain is terminated at both ends by two highly substituted aromatic esters; furthermore, it contains a 1,1′-interglycosidic linkage, a nitrosugar (evernitrose), 13 ring systems and 35 stereocenters.

2.3 Saccharomicin Antibiotics

Rapid emergence of bacterial resistance to antimicrobial agents has been recognized as an epidemic of global proportions. Particularly in hospital-acquired infections, many pathogenic bacteria are multiply resistant to several classes of antibiotics, effectively narrowing therapeutic options. For multiply resistant *Staphylococcus aureus*, vancomycin currently represents the last line of defense, whereas for certain resistant strains of *Mycobacterium tuberculosis* and vancomycin-resistant enterococci, no viable alternatives exist. Consequently there is renewed emphasis on the discovery of novel antibiotics. Lately, from an antibiotic complex designated LL-C19004, produced by the rare actinomycete *Saccharothrix espanaensis*, two novel heptadecaglycoside antibiotics have been isolated and named saccharomicin A and B (◉ *Fig. 8*) [51].

Saccharomicins represent a rather new class of antibiotics in which the absolute stereochemistry of the sugar residues still awaits clarification. This family is unrelated to all other known families of antibiotics and they show activity both in vivo and in vitro against multiply resistant strains of *Staphylococcus aureus* as well as vancomycin-resistant enterococci. The aglycon *N*-(*m*, *p*-dihydroxycinnamoyl)taurine was the first reported example of an amide bond linkage between cinnamic acid and taurine. In addition, the oligosaccharide portion of the molecules also exhibits unique features. Two of the four 4-*epi*-vancosamine residues are protonated, and considering the sulfate groups, the molecule apparently exists as a quadruple zwitterion. The antimicrobial activity of the saccharomicins may be attributed to their zwitterionic property that presumably folds the molecule into two loops. The macrocyclic backbone conformation may interact with bacterial cell walls and thus cause damage.

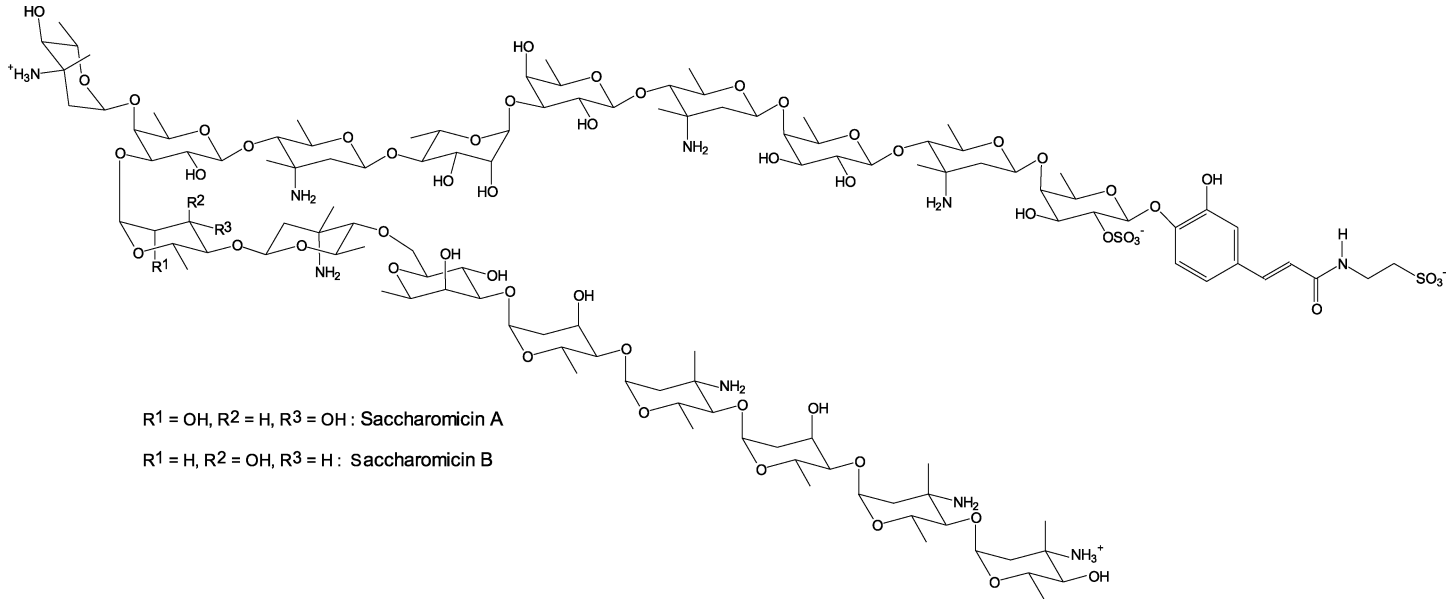


Figure 8
Structures of the saccharomicin antibiotics

2.4 Toxins of Marine Origin

Marine sources have only begun to be utilized for natural product search in the last 50–60 years [52]. Many of the compounds isolated form marine origins have been shown to be extremely effective toxins, like the famous tetrodotoxin. Diarrhetic shellfish poisoning is associated with eating bivalves such as mussels, scallops, or clams which have accumulated dinoflagellate toxins. The causative organisms have been identified as several dinoflagellates in the genus *Dinophysis*. Appearance of the dinoflagellates leads to toxification of the shellfish. The toxins responsible for most human diarrhetic shellfish poisoning, connected with symptoms such as diarrhea, nausea, vomiting and abdominal pain are okadaic acid and pectenotoxin 1 (❯ *Fig. 9*).

Okadaic acid was first isolated from the sponges *Halichondria okadai* and *Halichondria melanodocia* and was shown to be ultimately produced by dinoflagellates such as *Prorocentrum lima*. Okadaic acid contributes substantially to the analysis of phosphorylation and

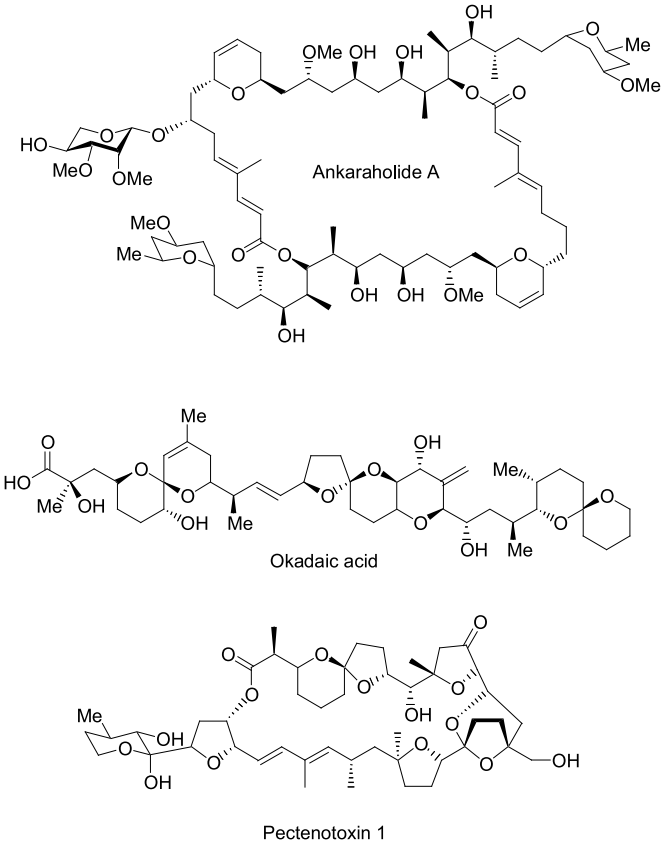


Figure 9
Carbohydrate-conjugated marine toxins

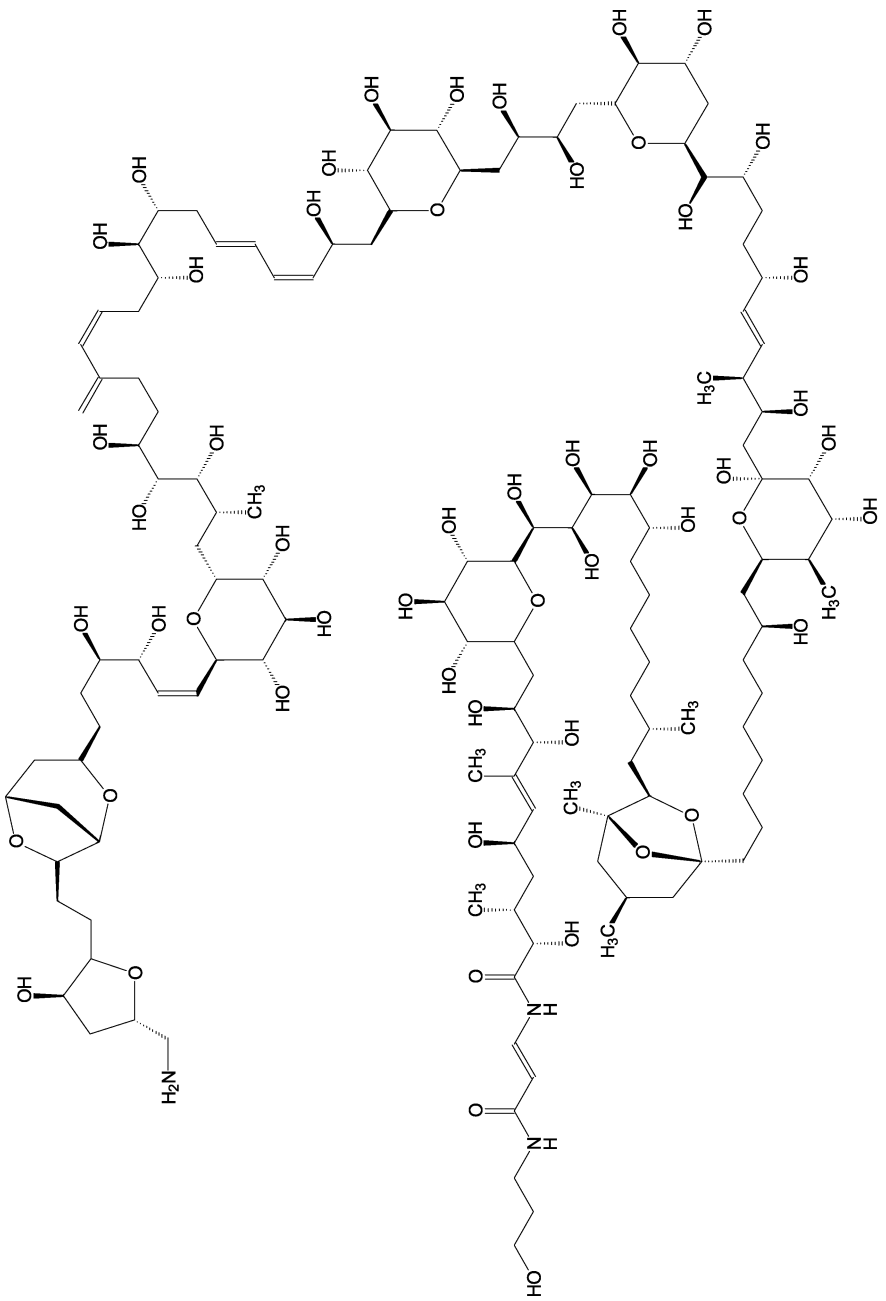


Figure 10
Structure of the famous marine toxin palytoxin

dephosphorylation processes in eukaryotic cell metabolism leading to the understanding of tumor growth processes with kinases as cell proliferation stimulators (tumor promotors) and phosphatases as cell proliferation inhibitors (tumor suppressors) [53].

Palytoxin (Fig. 10) is one of the most poisonous non-peptidic substances known, being even 25 times more toxic than tetrodotoxin. It was first discovered as the toxic principle of the Hawaiian legendary "limu-make-O-Hana (deadly seaweed of Hana)", which is actually the soft coral *Palythoa toxica*. Palytoxin has been isolated from soft corals of the genus *Palythoa* representing one of the largest marine-derived natural products yet isolated. In the meantime, palytoxin and its analogs have also been found in a wide variety of other organisms, in a seaweed *Chondria armata*, in crabs belonging to the genera *Demania* and *Lophozozymus*, in a triggerfish *Melichtys vidua* and in a file-fish *Alutera scripta*.

Palytoxin [54] has 64 stereocenters and 7 carbon-carbon double bonds that could exhibit geometrical isomerism; thus palytoxin could theoretically exist as 2^{71} different stereoisomeric forms, nevertheless, it is produced in nature as a single stereoisomer. Palytoxin acts by forming pores in the cell membrane by which membrane permeability for Na^+ and K^+ ions—at higher palytoxin concentrations also for ATP—is elevated. This results in induction of contraction of smooth muscle in every contracting organ.

Search for new biologically active secondary metabolites from marine cyanobacteria has led to the isolation of swinholide A from Fijian cyanobacterium [55]. This compound was formerly isolated from marine sponges. The bioassay-guided investigation of another cyanobacterium from Madagascar resulted in the identification of two other glycosylated swinholide derivatives, ankaraholide A and B, which differ only in one methyl group (Fig. 11). As cyanobacteria are symbiotic microorganisms of sponges, these bacteria might be the true producers of the swinholide metabolites.

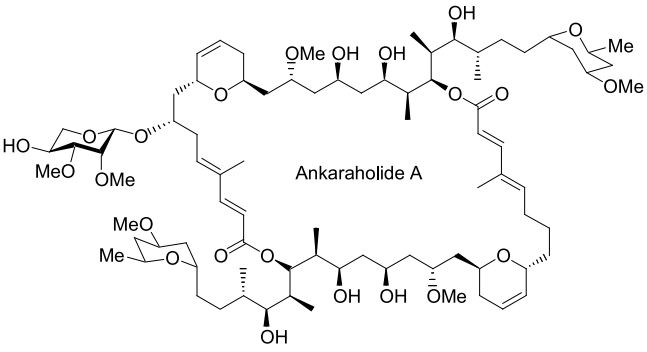


Figure 11
Structure of the glycosylated swinholide derivative ankaraholide

2.5 Spiroketal Glycosides and Cardiac Glycosides

Spiroketal glycosides are widely distributed in nature [56] displaying a broad spectrum of biological activities. Typical examples are the papulacandines, tricyclic spiroketal glycosides

which are produced by *Papularia sphaerosperma* (● *Fig. 12*). They have been found to have antibiotic activity and to inhibit β-glucan biosynthesis in yeast [57]. For papulacandin D [58] antifungal activity was reported.

Some members of a group of natural products, called the saponins also contain spiroglycosidic moieties, such as digitonin (● *Fig. 13*). The saponins resemble a family of amphiphilic glycosides forming colloidal detergent-type solutions in water. According to the chemical nature of their aglycon (sapogenine), they are grouped into steroid saponins and triterpene saponins. There are only very few saponins which originate from animals, such as from asteroids which produce the so-called holuthuroids and asterosaponins, respectively, which are toxic steroidal saponins an used for self-defense. They possess a wide variety of pharmacological activities, they are cytotoxic, haemolytic, and microbicidal. Cyclic saponins such as sepositosid A (● *Fig. 13*) are typical for the asteroid species *Echinaster*.

Osladin (● *Fig. 14*), a steroidal trisaccharide saponin, is the sweet principle of a fern, *Polypodium vulgare*, being more than 500 times sweeter than sucrose. These compounds are of interest as society shows a continuous and growing demand for highly sweet non-caloric and non-cariogenic sucrose substitutes, otherwise having the pleasant properties of sucrose. Some naturally occurring terpene glycosides, such as the sesquiterpenoid glycosides from *Sapindus rarak*, are especially interesting in this regard [59]. Structure-sweetness relationships among the natural terpenoid glycoside sweeteners have been a matter of investigation [60].

A steroid moiety is also typical for cardiac glycosides [61], for which the synonyms digitalis and cardiac steroids are often used. Cardiac glycosides act on the heart by direct as well as indirect mechanisms to enhance the force and velocity of contraction. In addition, they slow down the rate of atrioventricular conduction. Higher doses cause toxic and sometimes lethal effects. Plants and preparations containing cardiac glycosides have been used as poisons and drugs in herbal medicine since ancient times. The cardiac glycosides are characterized by a lactone ring at the 17β-position of the steroid aglycon and a sugar moiety at its 3β-position. They are found in a variety of plants distributed over a broad geographic area, and in some animals as well. Many natural products of the cardiac glycoside type have also been isolated from marine sources [62]. The best known source of cardiac glycosides is *Digitalis purpurea* (purple foxglove) found in Europe and Asia. The prototype of digitalis glycosides is digitoxin, a steroidal aglycon called digitoxigenin equipped with a trisaccharide of three D-digitoxose moieties (● *Fig. 14*).

2.6 Aureolic Acids

Aureolic acids are a group of tetrahydroanthracene oligosaccharide antibiotics, the name of which is related to their characteristic golden appearance. Members of the aureolic acid group include chromomycin A_3, olivomycin A and mithramycin, as the most prominent examples (● *Fig. 15*).

The aglycon of chromomycin A_3 is identical to that of mithramycin, but differs from that of olivomycin A by a methyl group. Conversely, the sugar components for chromomycin A_3 and olivomycin A are nearly identical, while those in mithramycin are different. Research has established that these drugs exhibit their activity by intercalating into host DNA and thus inhibiting normal cellular functions. Aureolic acids require a divalent metal ion, preferably

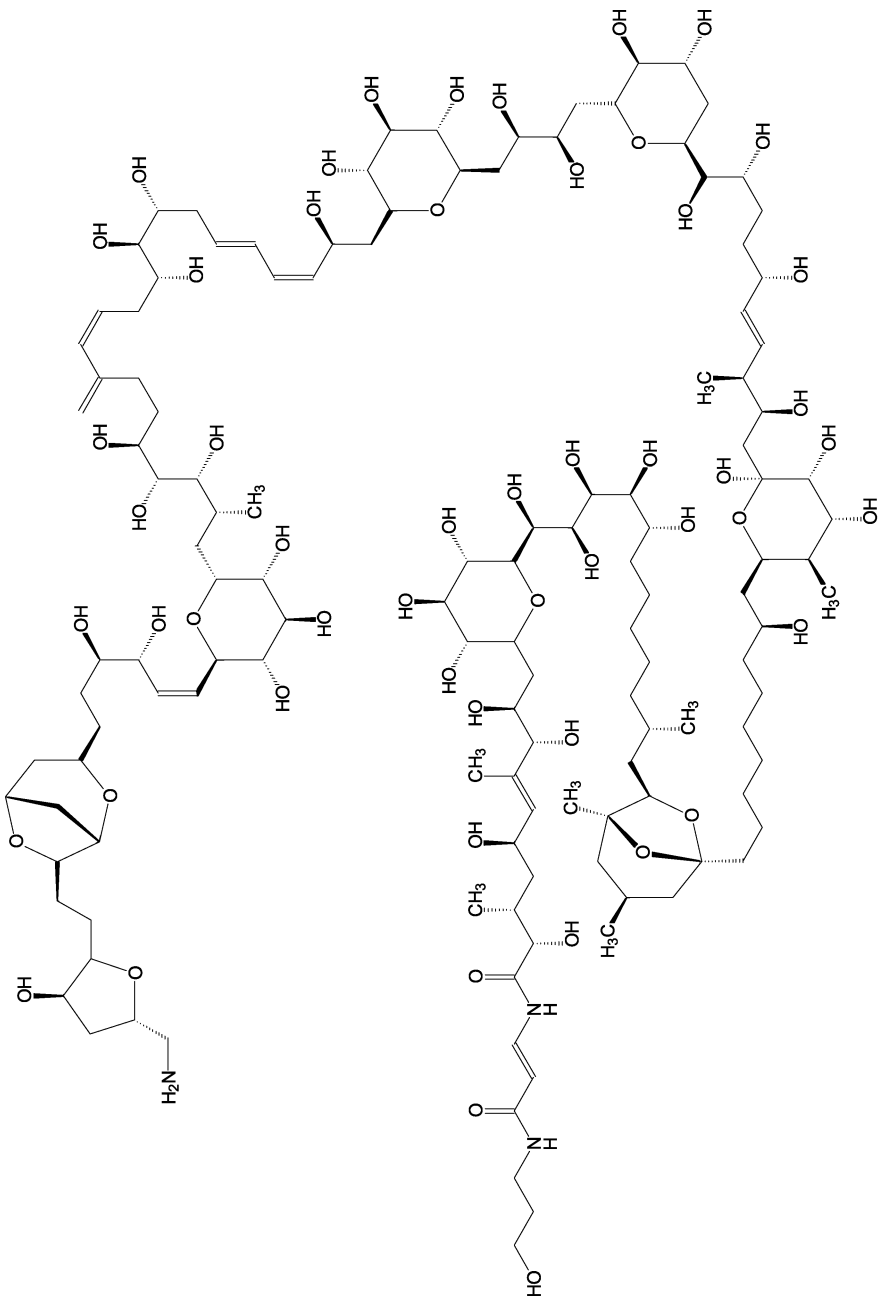
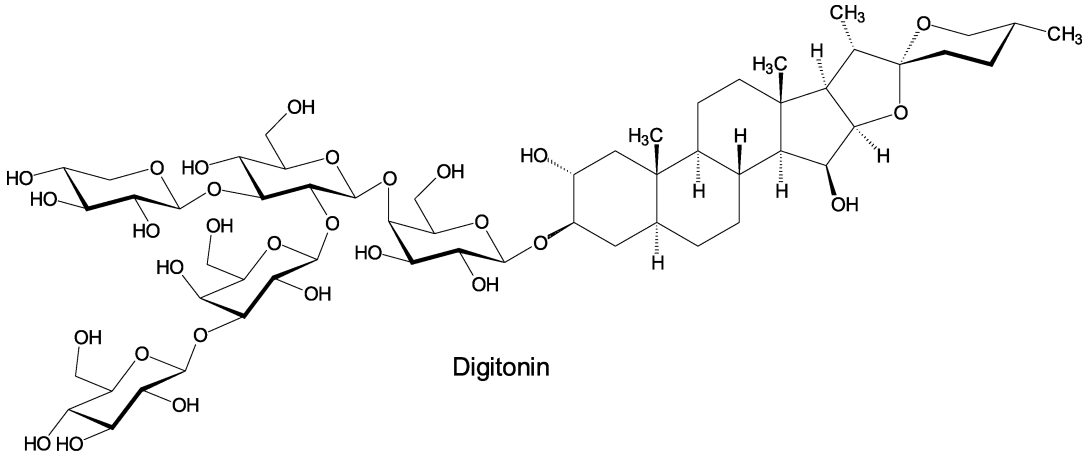


Figure 12
Structure of glycosidic spiroketals



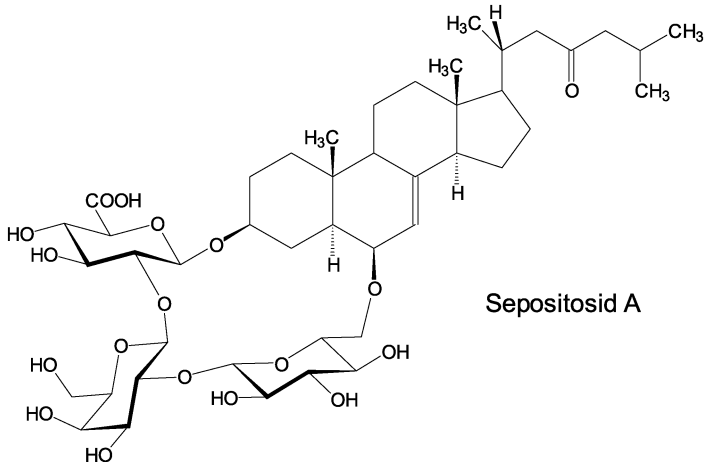


Figure 13
Glycosylated steroid derivatives

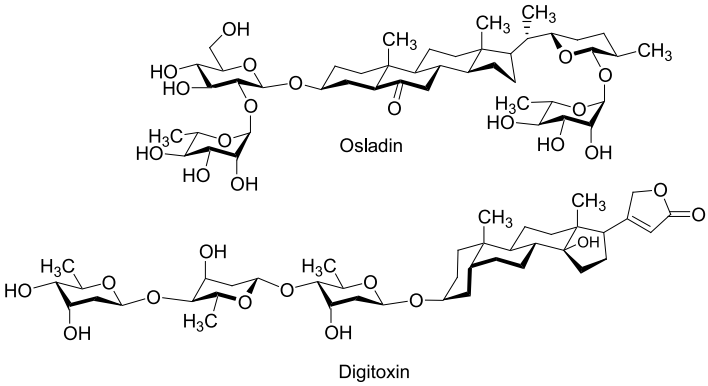


Figure 14
Cardiac glycosides depicted with their conformational characteristics

a Mg^{2+} ion [63] and a guanine-containing target for activity [64]. They form symmetrical dimer cation complexes and bind in the minor groove of DNA thus inhibiting RNA synthesis. Comparing the effects of aureolic acids with different sugar portions supports the idea that the carbohydrate moiety plays a role in determining the sequence specificity of the drug. The structures of the 2:1 complexes of chromomycin A_3 [65] and mithramycin [66] with DNA have been resolved by NMR spectroscopy. In these experiments, the hydrogen-bonding interactions of the deoxysugars with both strands of the DNA double helix were unveiled. The trisaccharide moieties in these aureolic acids were found to align with the minor groove, and they are also involved in the stabilization of the 2:1 drug-metal complex before DNA binding occurs. Removal of the trisaccharide inactivates the drug.

2.7 Anthracycline Antibiotics and Angucyclines

Anthracycline antibiotics are effective antitumor agents and are widely prescribed in chemotherapeutic treatments for a number of human tumors, with doxorubicin (adriamycin) and daunorubicin (daunomycin) being the most popular members (*Fig. 16*) [67]. These compounds contain a tetracyclic chromophore to which one or more deoxysugars are attached with or without amino groups. Anthracyclines may also contain deoxyuloses as in aclacinomycin A and even the aglycon structure can be further complicated as seen in arugomycin (*Fig. 16*). Removal of these sugars results in at least a partial loss of biological activity.
Several X-ray crystal structures and solution NMR structures have shown that the tetracyclic chromophore of these drugs intercalates DNA and alters the local topology [68]. In addition the glycosidic moiety binds in the minor groove – in case of arugomycin also to the major groove [69]. Another mode of activity for anthracyclines is direct inhibition of topoisomerase II [70]. In the case of anthracycline antibiotics, it appears that the oligosaccharide chains do not play a significant role in the DNA sequence selectivity as can be seen with calicheamicins. Intercalation of the chromophore into DNA seems to be the dominant interaction which both determines preferred binding sequences and directs the binding of the saccha-

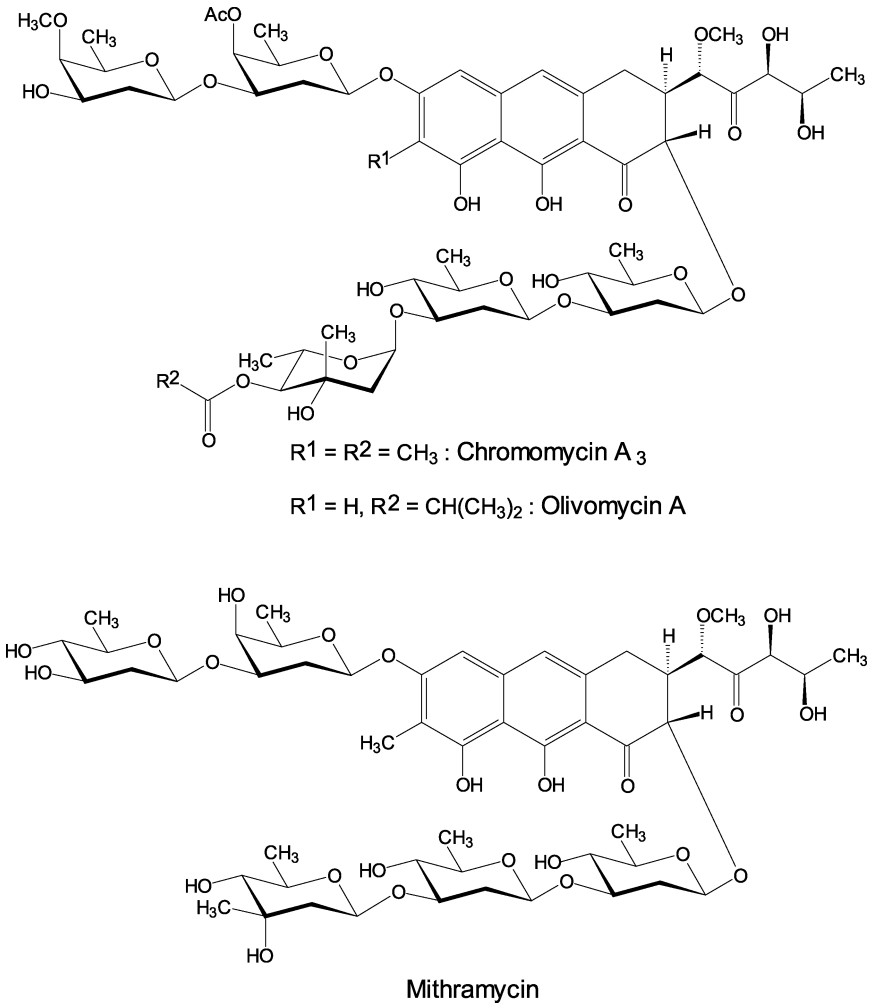


Figure 15
Structures of aureolic acids

rides into the minor groove. Nevertheless, the saccharide side chains have been shown to dramatically enhance the binding affinity of the drugs for DNA. Additionally, in aclacinomycin A the trisaccharide binding forced the DNA to kink toward the major groove with opening of the minor groove [71].

An emerging new class of antibiotics are the quinoide angucyclines [72] which in contrast to the related linearly condensed anthracyclines, do not only show anticancer activity but also a multitude of other interesting biological activities such as antibacterial or antiviral properties and enzyme inhibition. The angucyclinones have been classified based on the degree of oxygenation and *C*-glycoside formation [73]. Representative examples are aquayamycin

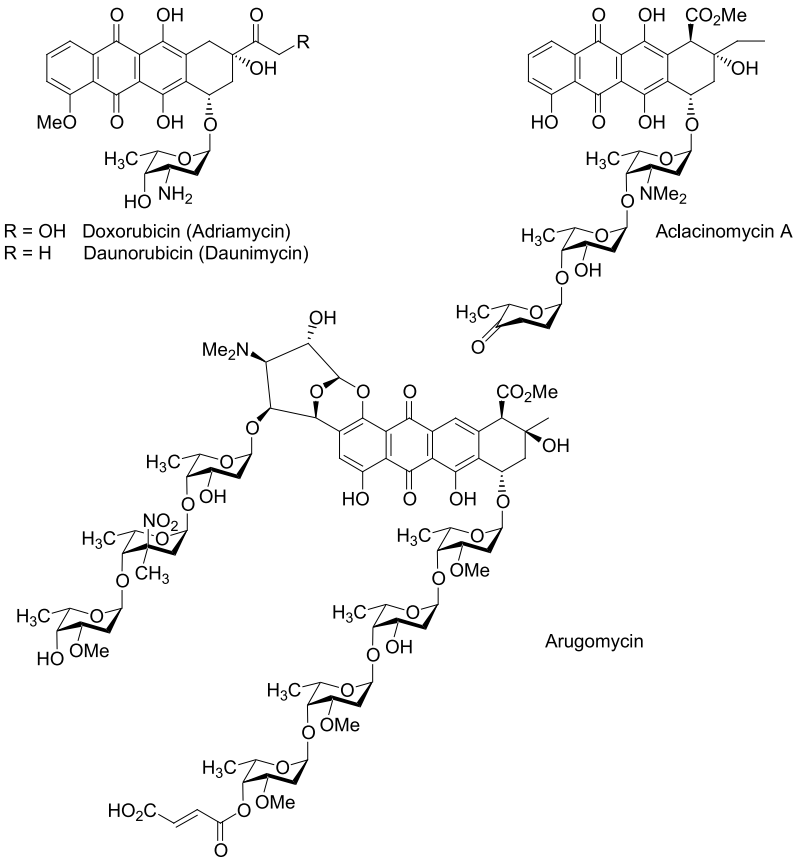


Figure 16
Structures of anthracycline antibiotics

in which an olivose is *C*-glycosidically linked and the saquayamycins, such as saquayamycin A (*Fig. 17*).

Many investigations have been dedicated toward the elucidation of the biosynthetic pathways leading to the deoxysugars found in anthracyclines and angucyclines such as in case of the angucycline antibiotic landomycin A (*Fig. 18*), produced by the actinomycete strain *Streptomyces cyanogenus* [74]. Landomycin A has a unique structure consisting of a decaketide aglycon and a hexadeoxysaccharide chain.

2.8 Enediyne Antibiotics

The enediyne antibiotics [75,76] have arrived on the scene in the 1980s with calicheamicin γ_1^I isolated from *Micromonospora echinospora calichensis*. They have a novel molecular structure which makes the enediyne antibiotics natural prodrugs; that is, the enediyne antibiotics

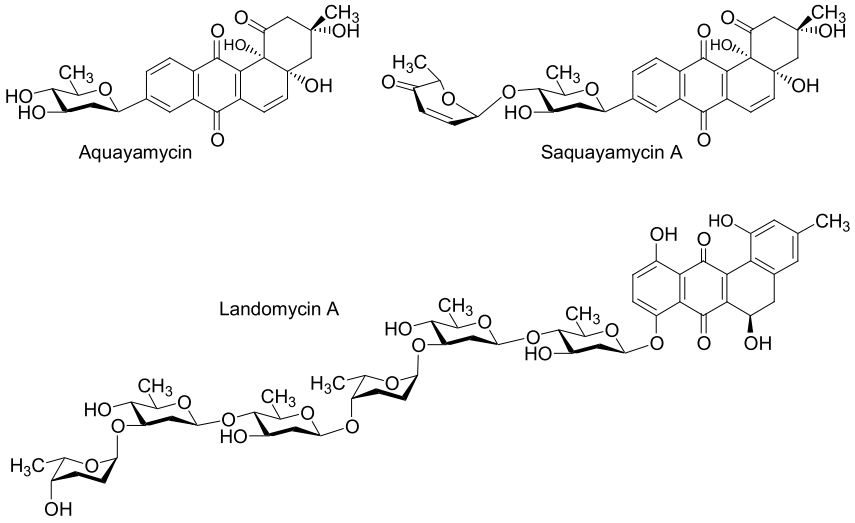
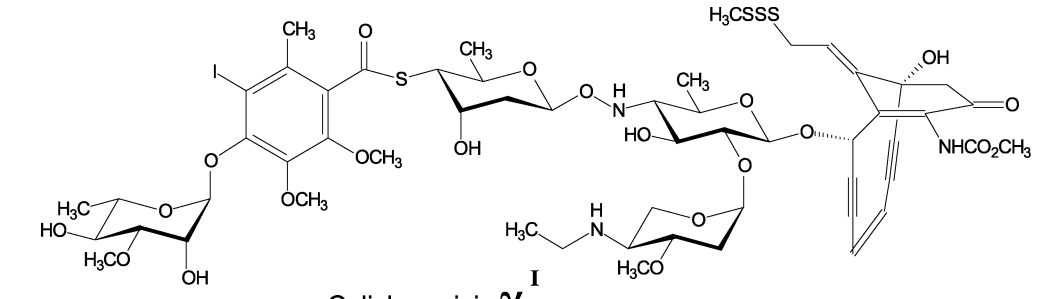
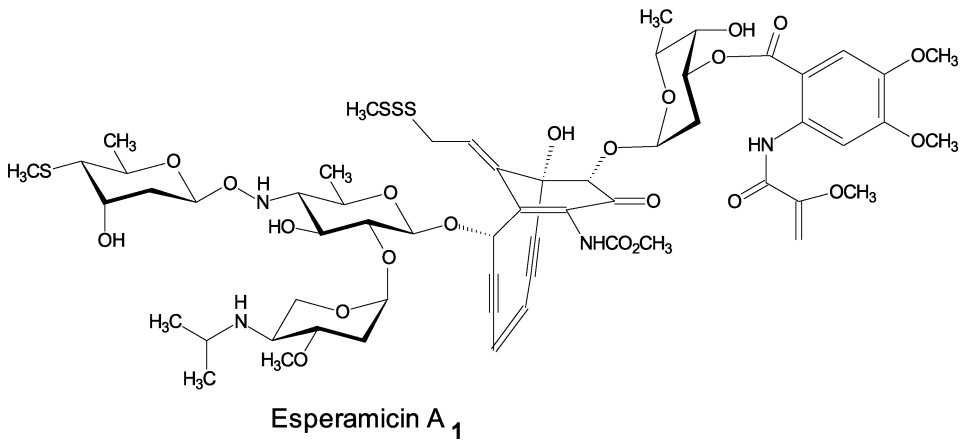


Figure 17
Structures of quinoide angucyclines

have to undergo an activation step before they can unfold their biological activity. In brief, this process includes strong binding to duplex DNA, double-strand cleavage of the genetic material by formation of a benzenoid diradical, and as a consequence potent antitumor and antibiotic activity. The antitumor activity of calicheamicin γ_1^I is over 1000 times higher than that of adriamycin, which is widely used in chemotherapy, when tested in murine tumor models [77]. The structure of calicheamicin γ_1^I (*Fig. 18*) is comprised of an unusual carbohydrate domain, a hexasubstituted benzene ring and a cyclic enediyne chromophore carrying a trisulfide moiety that acts as a triggering device for the cascade of events which leads, via a Bergman cycloaromatization [78], to the diradical species and DNA rupture. First the (Z)-1,5-diyn-3-ene unit is sequence-specifically [79,80] delivered to the minor groove of the DNA helix, whereupon reduction of the chromophore leads to cycloaromatization of the enediyne and 1,4-benzenoid diradical formation. The diradical is positioned to abstract hydrogen atoms from the sugar phosphate backbone of the nearby DNA, thus causing single- or double-strand scission of the helix and forming the molecular basis for the cytotoxic activity of these drugs. The oligosaccharide domain of calicheamicin γ_1^I is endowed with high affinity for certain DNA sequences, and acts as the delivery system of the molecule to its biological target. Thus, the sugar portions of the enediyne antibiotics are vital to the biological activity; although the aglycon can bind to DNA, it does so with far less affinity and no sequence selectivity [81]. Additionally, the calicheamicinone aglycon is not efficient at double-strand scission. On the other hand, the carbohydrate portion without the aglycon is able to bind DNA with nearly the same interactions as those of the carbohydrates in the whole drug [82] and it can block transcription. These observations are in contrast to the anthracycline antibiotics, whose sequence selectivity is determined primarily by the aglycon and whose carbohydrate moieties are unable to bind DNA in the absence of the aglycon.





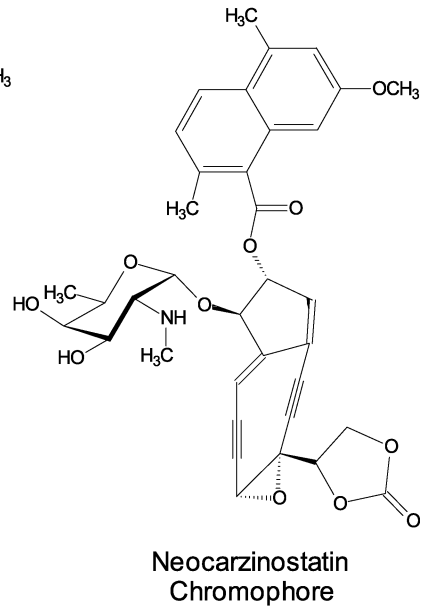
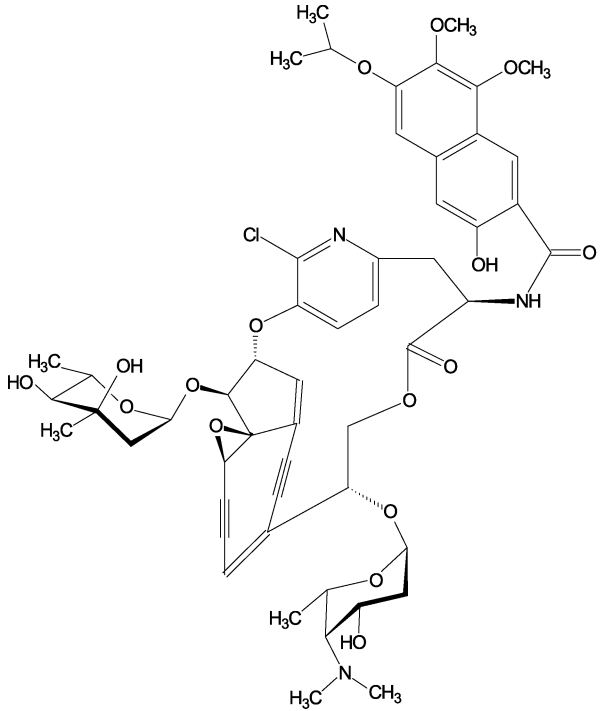
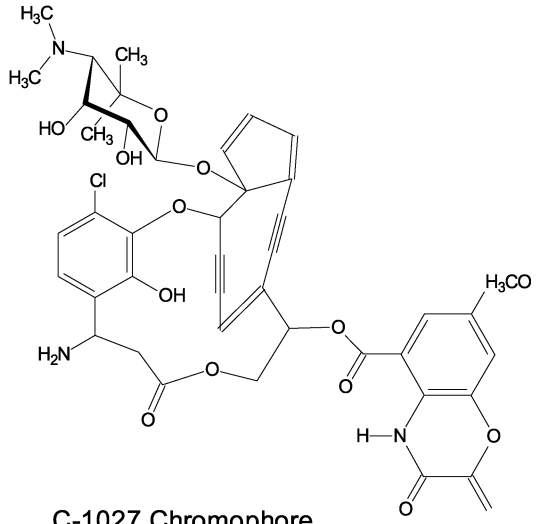


Figure 18
Structures of important enediyne antibiotics



Kedarcidin Chromophore



C-1027 Chromophore

Figure 19
Enediyne antibiotics with unusual chromophore

The family of enediynes includes the calicheamicins, esperamycins, dynemicin A, which is not glycosylated, the C-1027 chromophore and kedarcidin (◗ *Fig. 18* and ◗ *Fig. 19*). The neocarzinostatin chromophore, which does not contain an (Z)-enediyne, is also included in this family of compounds because of a similar mode of action.

The calicheamicins have been isolated from *Micromonospora echinospora calichensis*. They are highly active against Gram-positive as well as Gram-negative bacteria and especially effective against tumors. The esperamycins are a subclass of the naturally occurring enediynes, being effective against a broad spectrum of bacteria and effective as cytostatics.

Neocarzinostatin (NCS) is a naturally occurring 1:1-complex of a protein moiety and the NCS-chromophore, which can be isolated from *Streptomyces carzinostaticus* [83]. The chromophore portion is responsible for the biological activity, whereas the protein part is important for stabilization and transportation of the drug, which shows strong cytostatic and antibacterial activity, based on its ability to cleave DNA.

How would neocarzinostatin-containing species protect themselves against its DNA-disrupting properties? The answer lies in the protein part of neocarzinostatin, which is responsible for selecting thiol groups with which to react. It was shown that the neocarzinostatin apoprotein is responsible for shielding the labile chromophore from attack by the most abundant eukaryote cellular thiol groups before the drug reaches the target DNA. For example, while 2-mercaptoethanol can inactivate the protein-bound chromophore, glutathione is rejected by the protein. This selective inactivation may be the basis of the self-protective functionality of the microorganisms that produce neocarzinostatin.

Highly complex synthetic schemes have been devised for a number of members of the enediyne class of potent anticancer agents. However, despite their potency, the enediynes' reactivity and toxicity have limited their usefulness as drugs. Therefore, attempts have been made to prepare semi-synthetic enediyne derivatives to reduce toxic side effects of the enediyne antibiotics [84] and to simplify their synthesis including the preparation of enediyne mimetics [85], as for example to increase the reactivity of enediynes by metal-ion coordination. Furthermore, hybrid molecules have been prepared. As in the anthracycline daunorubicin the non-glycosylated aglyone is inactive and also does not show DNA intercalation, molecular hybrids of calicheamicin γ_1^I and the anthracycline aglycon have been synthesized for which the name "chalichearubinine" was coined, in order to combine DNA sequence specificity of calicheamicin γ_1^I with the activity spectrum of daunorubicin [86].

Relatively little has been known about how these bacterial natural products are synthesized in nature. Recently, it has been shown that the enediyne cores of both C-1027 and calicheamicin are synthesized via a common polyketide pathway, suggesting that all enediynes are biosynthesized in the same manner [87]. The stretch of the *Streptomyces globisporus* genome necessary for biosynthesis of the nine-membered enediyne C-1027 was cloned and characterized and it was found that disruption of the single polyketide synthase (PKS) gene in this region stops C-1027 production. By disrupting a hydroxylase gene in the C-1027 cluster, a novel C-1027 analog with improved chemical stability could be created and thus new drug candidates may arise by genetic manipulation of the biosynthetic pathways.

2.9 Aminoglycoside Antibiotics

This group of antibiotics consists entirely of moieties derived from carbohydrate metabolism, i. e., the molecules can be rather characterized as "pseudooligosaccharides" than as "glycoconjugates". Aminoglycoside antibiotics have long been used as very efficient drugs against Gram-positive and Gram-negative bacteria as well as mycobacterial infections. Streptomycin was the first oligosaccharide analog with pharmaceutical relevance, discovered in 1944 [88]. Streptomycin and its relatives were found in culture filtrates of several streptomyces strains. Since the discovery of streptomycin, a great number of aminoglycoside antibiotics have been isolated worldwide from the cultures of actinomycetes and bacteria such as *Streptomyces*, *Nocardia*, *Micromonospora*, and *Bacillus*. Early works of the structure elucidation were reviewed [89].

The terms "aminoglycoside" as well as "aminocyclitol antibiotic" which are used for this class of natural products refer to the structural aspect. Most of the aminoglycoside antibiotics contain aminosugars and an aminocyclitol or a cyclitol moiety. The middle furanose part in streptomycin is called streptose and was the first branched-chain sugar obtained from microorganisms. The other rings in streptomycin are a *N*-methyl-L-glucosamine moiety and an aminocyclitol called streptidine (❯ *Fig. 20*).

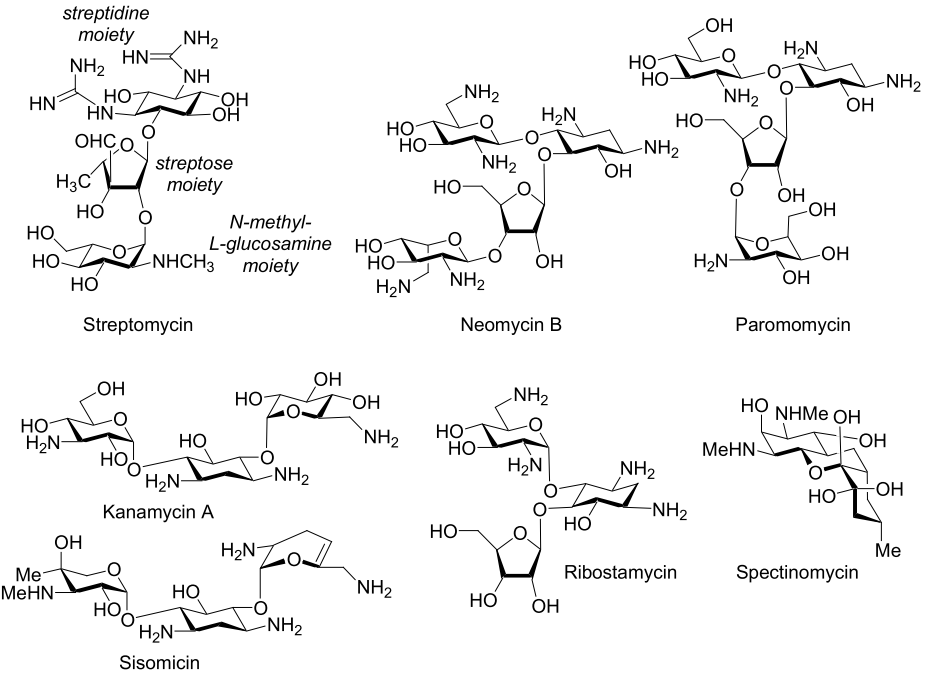


Figure 20
Structures of important aminoglycosides

The structural diversity of aminoglycosides is large. They consist of two to five monomers and contain almost a one-to-one ratio between amino and hydroxy groups. The antibiotics of the neomycin group have a 4,5-disubstituted 2-deoxy-streptamine. Neomycins B and C were isolated from the culture broth of *Streptomyces fradiae* and contain four rings [90]. They are closely related to paromomycin which has been an important aminoglycoside for biological studies owing to its excellent affinity for the A site region of rRNA of bacterial ribosomes [91]. The kanamycins possess a 4,6-disubstituted 2-deoxystreptamine; they are produced by *Streptomyces kanamyceticus*, and have been clinically used against a variety of bacteria, especially against streptomycin-resistant tuberculosis. Sisomicin was isolated from the culture broth of *Micromonospora inyoensis*; it contains an unsaturated aminosugar at the 4-position of 2-deoxystreptamine [92]. *Micromonospora inyoensis* produced a dimeric compound, named 66–40C as a minor component of sisomicin [93]. This compound suggested that the 6′-position of sisomicin is biosynthetically aminated via an aldehyde intermediate. Sisomicin has been used as a chemotherapeutic agent.

The spectinomycins (❯ *Fig. 20*) have a fused three-ring system which consists of a 4,5-disubstituted actinamine and an ulose. Spectinomycin is produced by *Streptomyces spectabilis* and *Streptomyces flavopersius*. It has been clinically used for treatment of the gonorrhea infection. A related structure has been discovered with spenolimycin which was isolated from the culture filtrate of *Streptomyces gilvospiralis*. Efforts have been directed toward semisynthetic aminoglycoside antibiotics to treat multi-drug resistant bacteria [94].

Aminoglycoside antibiotics interact with a great variety of RNA molecules [95] including sequences that are important for retroviral replication. The potential of RNA as a new drug target has recently come up, with the recognition that RNA molecules can adopt complex three-dimensional structures that, as with proteins, enable the design of specific ligands. Another reason for the present interest comes from the fact that many pathogenic agents, such as retroviruses, encode their genetic information in RNA strands. Thus, any biological function involving RNA is a potential target.

The interactions between aminoglycosides and RNA are influenced by the number and basicity of amino groups present [96] and thus the strength of this interaction is dominated by electrostatics, with the positively charged aminoglycosides displacing metal ions. As the metal ions which are associated with the highly charged RNA molecules participate in RNA three-dimensional folding and provide active centers in catalytic RNA molecules, aminoglycosides can promote conformational changes, prevent folding of the RNA into an active tertiary architecture or displace catalytically active ions [97]. To better understand the function of aminoglycosides at a molecular level, their interactions with RNA are considered a good model system [98]. The elucidation of the molecular details of this interaction would eventually facilitate the development of new antibiotics.

Unfortunately, as with many other antibiotics, the efficiency of aminoglycosides has been compromised by the emergence of resistant bacterial strains [99]. Bacteria produce enzymes such as phosphotransferases and acetyltransferases to inactivate aminoglycosides by enzymatic modification of specific hydroxyl groups or amino groups of the aminoglycosides. *Escherichia coli*, for example, can inactivate streptomycin by using an enzyme which transfers AMP from ATP onto the 3″-OH of the antibiotic to produce 3″-O-adenylylstreptomycin (❯ *Fig. 21*) [100]. Chemically modified aminoglycosides have been used to investigate their interactions with nucleic acids. A neomycin-acridine conjugate (❯ *Fig. 22*), for example, has been used to

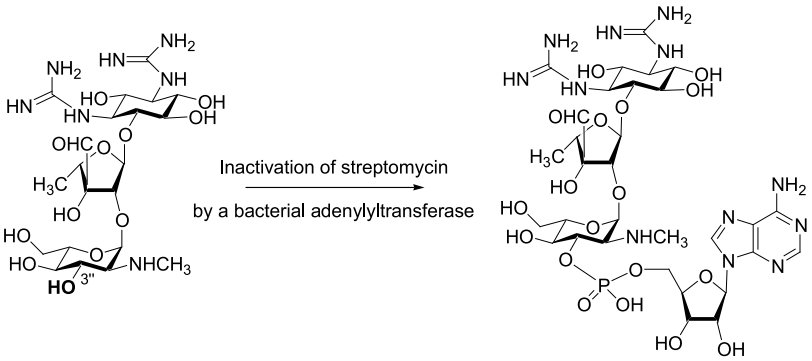


Figure 21
Enzymatic inactivation of streptomycin

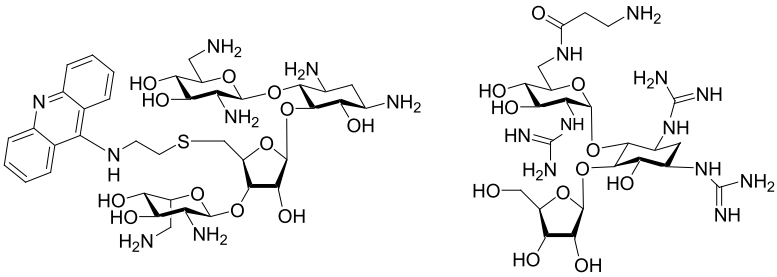


Figure 22
Functional aminoglycoside conjugates

show preference of neomycin binding to A-form nucleic acids [101]. A β-alanyl-modified ribostamycin derivative (*Fig. 22*) has been synthesized to allow its immobilization and to facilitate functional assays [102]. Additional guanidinium moieties improve RNA binding.

2.10 Moenomycins

The moenomycin antibiotics possess sophisticated glycolipid-type structures, as shown for moenomycin A (*Fig. 23*), which are difficult to synthesize [103]. The first total synthesis of moenomycin A has only recently been accomplished [104].

An explanation for the antibiotic activity of this class of bioactive molecules has been suggested [105,106]. It is believed that moenomycins interact in the transglycosylation step of the biosynthesis of the bacterial cell wall peptidoglycan. Two groups of antibiotics are known to interfere in this step, the glycopeptide antibiotics such as vancomycin, preventing the approach of the transglycosylase to its substrate by binding to mucopeptide precursor molecules, and the moenomycin antibiotics, which function by an entirely different mechanism: Moenomycins

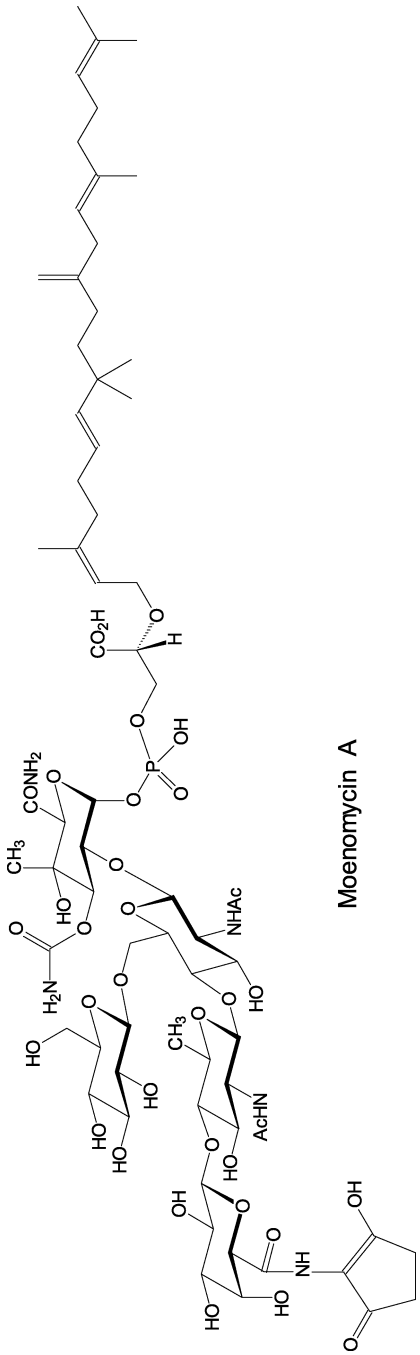


Figure 23
Structure of moenomycin

prevent transglycosylation of the peptide chain by reversibly binding to the necessary enzyme transglycosylase [107,108], thus acting as enzyme inhibitors.

Transglycosylases are specific oligosaccharyltransferases which transfer a precursor-oligosaccharide onto a peptide chain [109,110,111]. They use lipid-bound sugar pyrophosphates as their substrates. The lipid unit in bacteria is a C_{55} isoprenoid (bacterioprenol) whereas in the eukaryotic dolichol cycle long-chain polyprenol alcohols (dolichols) of varying chain lengths (C_{40}–C_{120}) are used. Bacterial transglycosylation is believed to proceed in such a way that the growing peptidoglycan chain linked to bacterioprenol via a pyrophosphate bridge acts as the glycosyl donor whereas a disaccharide intermediate, so-called lipid II, acts as glycosyl acceptor.

In *E. coli* the transglycosylation reaction is catalyzed by the high-molecular weight penicillin-binding proteins. These are bifunctional enzymes which catalyze the transglycosylation reaction as well as the transpeptidation reaction, in which the sugar strands are cross-linked via short peptide chains [112]. As peptidoglycan glycosyltransferases, which catalyze the polymerization step of cell-wall biosynthesis are highly conserved across all bacteria, they form an important drug target for antibiotic-resistant bacteria. An X-ray structure of the bifunctional cell-wall cross-linking enzyme, including its transpeptidase and glycosyltransferase domains has recently be determined [113].

The moenomycins are believed to compete for the binding sites of the enzyme with its natural substrates. Its oligosaccharide moiety is recognized by transglycosylase in a highly specific interaction. A second prerequisite for the transglycosylase-inhibiting activity of moenomycin has been shown to be intercalation of the antibiotic into the cell membrane mediated by its lipid moiety [114]. However, the lipophilic isoprenoid chain of moenomycin also accounts for its poor pharmacokinetic properties, preventing its clinical use in humans. As removal of this unit completely abolishes biological activity, current research is dedicated to the question, whether this portion of the molecule can be replaced by a shorter lipid [115].

2.11 Glycopeptide Antibiotics – the Vancomycin Group

The increasing problem of bacterial resistances against the classical antibiotics has led to a growing importance of the glycopeptide antibiotics as new potent drugs. Molecules of this group have a peptide backbone of 7 amino acids in common with over 100 examples being structurally elucidated. Glycopeptide antibiotics [116,117,118] are effective against a broad spectrum of Gram-positive and Gram-negative bacteria. The most well-known representative is vancomycin (❯ *Fig. 24*) which is used to treat infections caused by the penicillin-resistant *Staphylococcus aureus*, which has also developed resistances against methicillin and almost all other common antibiotics such as the cephalosporins, tetracyclins, aminoglycosides, erythromycin and the sulfonamides.

Vancomycin was discovered in the 1950s in a soil sample collected in the jungle of Borneo and isolated by Eli Lilly from culture broth of the actinomycete *Streptomyces orientalis*, which was later named *Nocardia orientalis* and finally *Amycolatopsis orientalis*. It was first used in clinics in 1959, and today also teicoplanin (❯ *Fig. 25*) is being used clinically. The other glycopeptides of the vancomycin group such as ristocetin A (❯ *Fig. 26*) have too many toxic side effects to be used as a therapeutic. Vancomycin and teicoplanin are among the last antibiotics

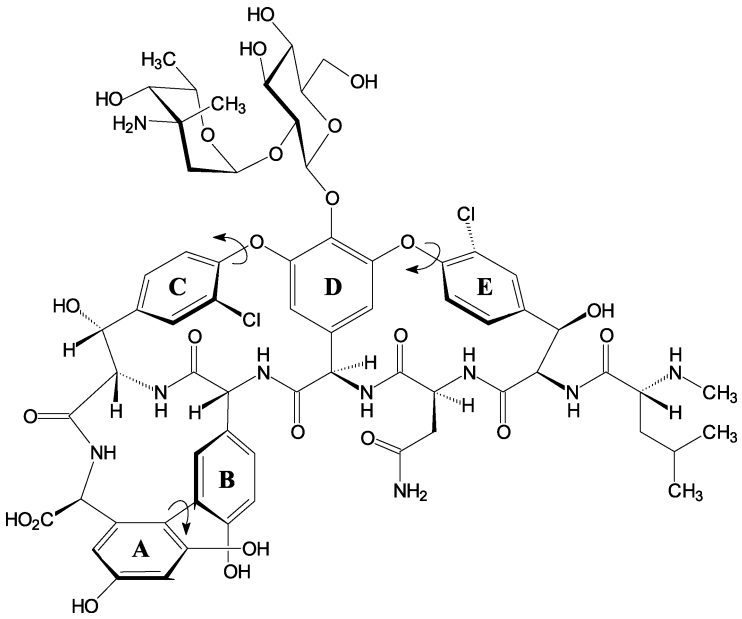


Figure 24
Structure of vancomycin. Vancomycin has three elements of atropisomerism as indicated by *three arrows*

forming a line of defense against resistant bacteria, however, also vancomycin-resistant strains have already developed [119,120].

Glycopeptide antibiotics are normally produced as a mixture of compounds by a single bacterial strain. They are structurally varied especially with regard to their glycosylation. Until today, ristocetin A is the only glycopeptide antibiotic which has been found to have a tetrasaccharide moiety bound to the fourth amino acid residue of the peptide backbone. Teicoplanin has a glucosamine moiety at this position which is *N*-acylated with a C_{11}-side chain. This side chain is responsible for the advantageous pharmacological features of this compound. The sugar moieties in glycopeptide antibiotics improve the solubility of the respective compounds thus facilitating transport of the drugs to their targets.

Most of the monosaccharide moieties found in glycopeptide antibiotics fall into two groups, hexoses and 6-deoxy-hexoses on one hand and aminotrideoxyhexoses on the other(◘ *Fig. 27*). A number of rare sugars such as 4-oxovancosamine and ureidovancosamine are also found.

Most of the difficult syntheses leading to these monosaccharide have recently been reviewed [121]. Vancosamine, which can be regarded as the *C*-3-methyl analog of daunosamine, is one of the most widely distributed monosaccharide constituents in glycopeptide antibiotics [122]. De novo syntheses of vancosamine derivatives have recently been reported [123,124].

The structure of vancosamine was elucidated in 1982 [125] and its total synthesis was finally accomplished in 1999. By 1998 two groups, had reported independent total syntheses of the vancomycin aglycon [126,127,128,129,130] and by early 1999 the first total synthesis [131,132,133,134,135] of vancomycin itself appeared in the literature followed by another

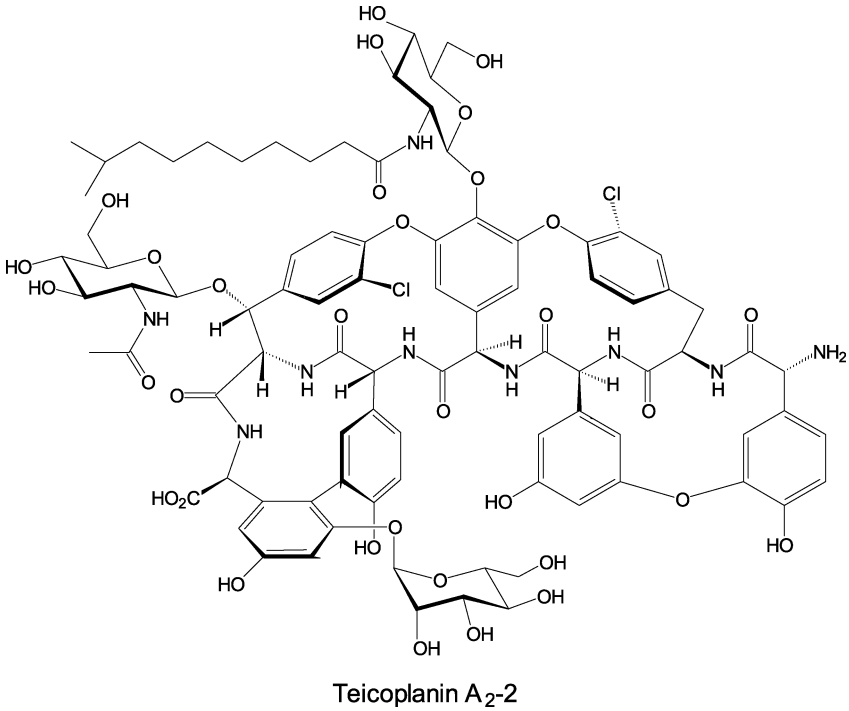


Figure 25
Structure of teicoplanin

report of the aglycon synthesis [136]. Besides the difficult carbohydrate chemistry, for the total synthesis of vancomycin a total of 18 stereocenters had to be correctly assembled. Furthermore three elements of atropisomerism in the aglycon had to be established, given by the biarylic system AB as well as the ether-linked biaryl moieties C-O-D and D-O-E (*Fig. 24*).

The vancomycin antibiotics act by an intriguing mechanism [137,138]. They interfere with the transglycosylation step of the bacterial peptidoglycan biosynthesis. Vancomycin binds to mucopeptide precursor molecules of the bacterial cell wall, terminating in Lys-D-Ala-D-Ala [139] and thereby preventing the approach of the transglycosylase.

Bacterial cell wall biosynthesis proceeds in two steps at the exterior of the cell. First disaccharide moieties with peptides attached are transported from the cytoplasm to the outside of the cell, where they are linked together catalyzed by transglycosylase. Subsequently the resulting long glycopeptide chains are, catalyzed by a transpeptidase, cross-linked through their peptide moieties to improve their mechanical stability. The enzyme recognizes the sequence D-Ala-D-Ala at the end of a peptide chain, cleaves the terminal alanin residue and then links the remainder to the peptide chain of an adjacent glycopeptide. How does vancomycin interact in this process?

The antibiotic forms a number of hydrogen bonds to the mucopeptide precursor molecule with which it interacts. This alignment is strengthened by hydrophobic interactions between

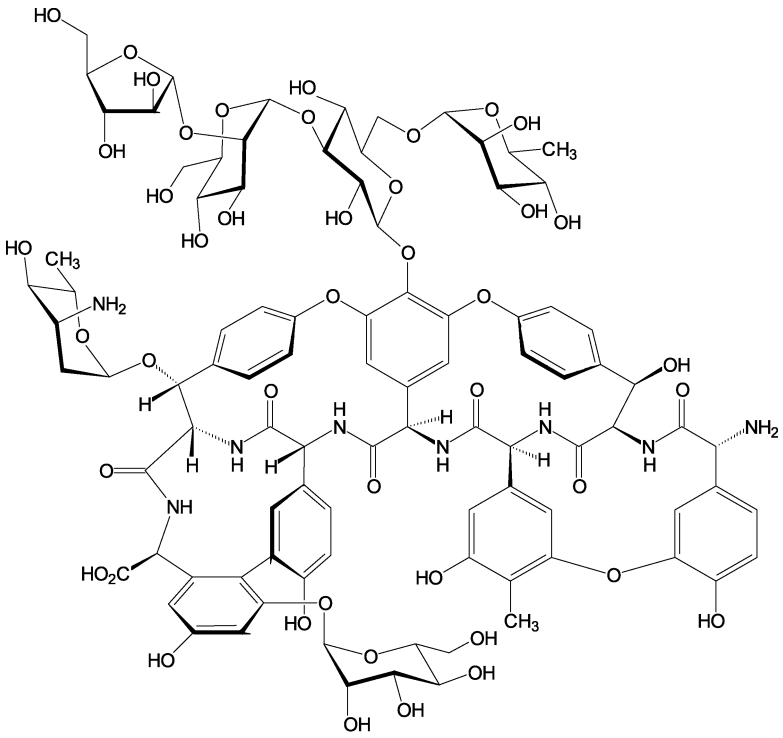
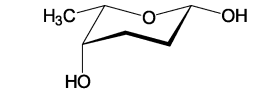


Figure 26
Structure of ristocetin

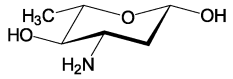
the alanyl methyl groups of the ligand and the antibiotic's hydrocarbon moieties (*Fig. 28*, hydrogen bonds are indicated by dashed lines).

As shown with a synthetic cell wall analog, effective binding is critically dependent on firm complexation of the carboxylate at the *C*-terminus of the mucopeptide. This interaction adds the largest contribution to the exothermic term of the interaction (20–30 kJ mol^{-1}). To guarantee this, binding of the carboxylate group in a pocket of three NH-groups of the antibiotic has to be stronger than its association in water. This is most likely accomplished by forcing these three CO-NH-dipoles into an especially close neighborhood, for which the necessary energy has been paid during the biosynthesis of the drug. This allows a better solvatization of the carboxylate than water would be able to provide it.

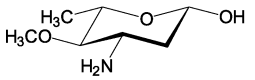
The two amide-amide-hydrogen bonds, also indicated in *Fig. 28*, add almost no direct contribution to improve ligand binding. Indirectly, however, they play an important role as they help to orient the ligand in such a way that the hydrophobic interactions with the alanyl-methyl groups can optimize binding, finally increasing binding energy by a factor of 10^3. Moreover, they restrict the flexibility of the ligand in the binding pocket of the antibiotic, thereby improving binding of the terminal carboxylate.



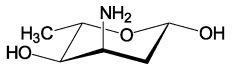
L-Rhodinose
(2,3,6-Trideoxy-L-threo-hexopyranose)



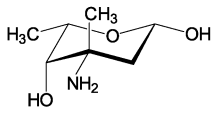
L-Acosamine (3-Amino-2,3-6-trideoxy-
L-arabino-hexopyranose)



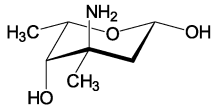
L-Actinosamine (3-Amino-2,3,6-trideoxy-
4-*O*-methyl-L-arabino-hexopyranose)



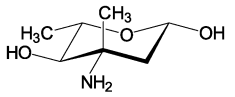
L-Ristosamine (3-Amino-2,3,6-trideoxy-
L-ribo-hexopyranose)



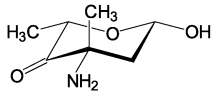
L-Vancosamine (3-Amino-2,3,6-trideoxy-
3-*C*-methyl-L-lyxo-hexopyranose)



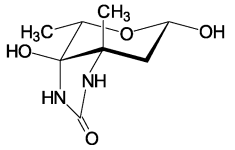
3-*epi*-L-Vancosamine (3-Amino-2,3,6-trideoxy-
3-*C*-methyl-L-xylo-hexopyranose)



L-Eremosamine (3-Amino-2,3,6-tridesxy-3-*C*-methyl-
L-arabino-hexopyranose, 4-*epi*-L-vancosamine)



4-Oxovancosamine (3-Amino-2,3,6-trideoxy-3-*C*-methyl-
L-threo-hexopyranos-4-ulose)



Ureido-4-oxovancosamine (3aR, 4S, 6R, 7aS)-
Octahydro-3a-hydroxy-4,7a-dimethyl-2-oxo-
pyrano[3,4-d]imidazol-6-yl]

Figure 27
Unusual monosaccharide components found in glycopeptide antibiotics

This concert of interactions may reflect a general principle, according to which weak interactions can lead to strong molecular recognition. Thus, a total of weak interactions provided by closely positioned functional groups can provide a significant increase in binding energy by restricting the flexibility of a receptor-bound ligand. This implies that the cooperativity of the binding process is not determined by entropy such as in case of the classical chelate effect,

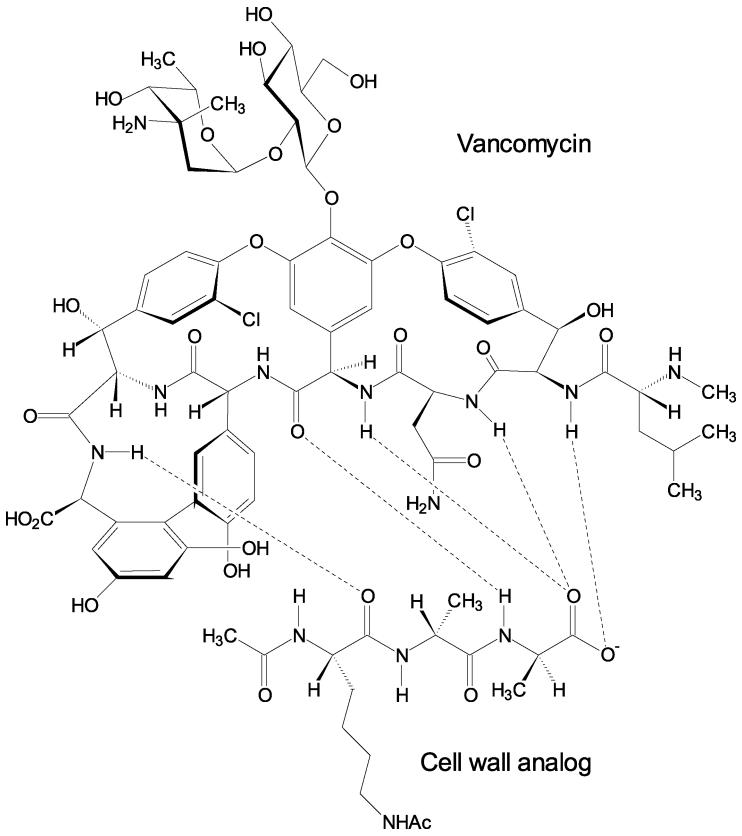


Figure 28
Strong binding of vancomycin to a cell wall analog results from a concert of weak interactions

where many interactions to the same molecule lead to anchoring the ligand to the receptor, but rather by an enthalpic effect, cooperatively increasing the strength of one particular interaction. This is an intriguing observation, which might be of general importance when molecular recognition processes are determined by weak interactions.

2.12 Sugar Amino Acids

Conjugation of carbohydrates with amino acids has become an important field of modern glycosciences. This type of compounds has been named 'sugar amino acids', abbreviated SAAs, and has found many interesting applications in drug design [140]. However, also among the naturally occurring compounds many SAAs are found. One prominent family of sugar amino acids consists of neuraminic acid derivatives which are *N*- or *O*-acylated. The main substituents on nitrogen are acetyl and glycosyl. Derivatives of glucosaminuronic acid are detected in the cancomycin family of antibiotics, similar to vancomycin [141] and natural SAAs

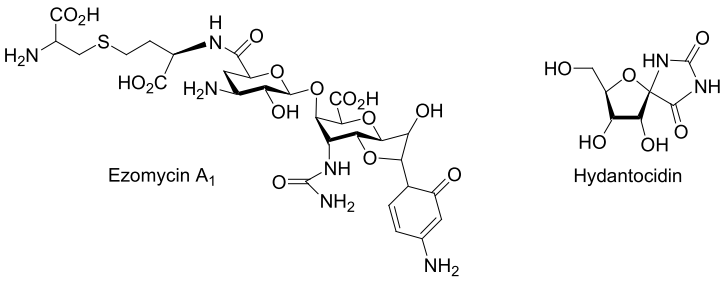


Figure 29
Naturally occuring glyco amino acids

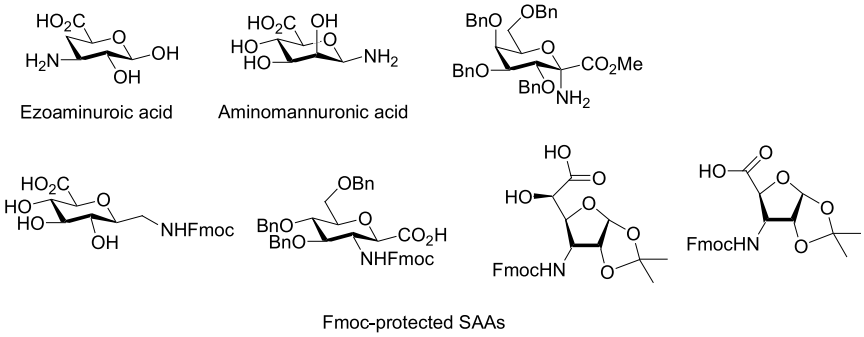


Figure 30
Synthetic glyco amino acids

can also be found in nucleoside antibiotics [142]. Two examples of SAA-containing natural products are depicted in *Fig. 29*: Ezomycin A_1 containing two different 3-amino-3-deoxy uronic acids [143] and the furanoid SAA (+)-hydantocidin, which represents a spiro hydantoin derivative and exhibits herbicidal activity [144].

SAAs are valuable building blocks in drug design and drug research, allowing for the preparation of modified analogs of biologically active peptides or glycoconjugates as well as oligomerization and cyclization to conformationally restricted and designed glycopeptides. Sugar amino acids have been made of furanoid or pyranoid α- β-, γ- or δ-amino acids. Some examples can be seen in *Fig. 30*.

Synthesis of individual SAAs and their incorporation into peptidic or saccharidic structures has been successful in medicinal chemistry [145]. A number of carbohydrate-based structures have been designed as mimetics of the potent peptidic antagonist of integrins, the RGD-motif. Integrins are proteins which are located at the cell surface of a number of different cell types. They play an important role in cell-matrix interactions as well as in tumorgenesis. This invoked a pharmaceutical interest in antagonists of integrins to block angiogenesis. The recognition sequence of integrins such as $\alpha_V\beta_3$ has been shown to be the RGD(Arg-Gly-Asp)-motif (*Fig. 31*). Sugar amino acids were used as turn mimetics [146] in the RGD sequence

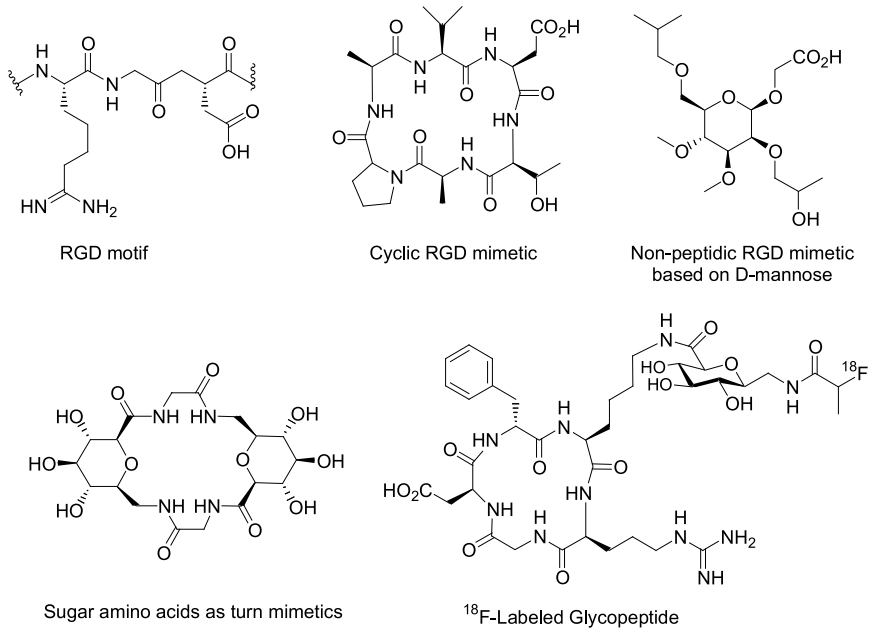


Figure 31
Use of glyco amino acids in biological chemistry

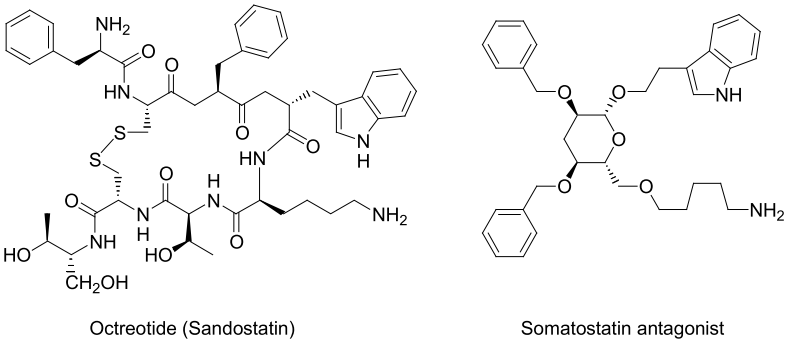


Figure 32
Carbohydrate scaffolds can serve to mimic more complex biologically active compounds

to improve the pharmacokinetic properties of these compounds and to introduce radio-nuclides such as ^{18}F for tumor imaging (*Fig. 31*) [147]. Carbohydrate derivatives such as of D-mannose have also been used to mimic peptidic RGD mimetics (*Fig. 31*).

Furthermore, as carbohydrates provide a source of multifunctional molecules with well-defined stereochemistry, they have been used as scaffolds to present pharmacophoric groups in a distinct arrangement. The first so-called carbohydrate privileged structure has been a somatostatin agonist which contained a deoxyglucose nucleus [148] (*Fig. 32*). Somato-

statin (also known as growth hormone inhibiting hormone (GHIH) or somatotropin release-inhibiting hormone (SRIF) is a peptide hormone that regulates the endocrine system and affects neurotransmission and cell proliferation via interaction with G-protein-coupled somatostatin receptors and inhibition of the release of numerous secondary hormones. Somatostatin analogs are of great interest in medical therapy such as for treatment of acromegaly. Octreotide (sandostatin) is a cyclic α-octapeptide which serves as analogue of the tetradecapeptide somatostatin.

2.13 Functional Glycoconjugates

As glycoconjugates on cell surfaces are involved in numerous processes of cellular biology such as inflammation, cell-cell interactions, cell recognition, immunological response, metastasis and fertilization, they form interesting targets for medicinal chemistry and pharmaceutical research as well as industry. Moreover, the glycocalyx of a cell has been found to be specific for a particular species, cell type and even its developmental or health status. To gain insight into the biological processes in which glycocalyx carbohydrates are involved, synthetic glycoconjugates have been utilized to probe protein-carbohydrate, carbohydrate-carbohydrate and carbohydrate-nucleic acid interactions [149].

Many different approaches to functionalized glycoconjugates have been investigated and published. Multivalent glycoconjugates [150,151,152,153] have been synthesized in the form of glycodendrimers [154] for example, with respect to the multivalency effects [155] which are of importance in glycobiology [156]. In addition, multivalent assemblies and arrays of carbohydrates have been realized, such as in the case of glyco-decorated self assembled monolayers (SAMs) [157,158], as gold glyconanoparticles [159] or as carbohydrate microarrays in so-called glyco-chips [160]. Moreover, carbohydrates have been attached to functional probes such as fluorescent markers, biotin or magnetic beads [161]. *Figure 33* shows an example of a glucose-based cluster mannoside which has been conjugated to fluoresceine [162] together with an carbohydrate-functionalized fluorescent, poly(*p*-phenylene ethylene) (PPE) which can be used for the detection of *Escherichia coli* [163]. A fluorescence resonance energy transfer (FRET) experiment allows to detect *E. coli* binding to the mannose-decorated polymer.

As bacteria as well as many other microbes bind to carbohydrates which are displayed on the surface of their host and target cells to allow firm adhesion to the cell surface, inhibition of this carbohydrate-dependent adhesion process has been the goal of a concept named anti-adhesion therapy [164]. For the approach to treating microbial infection target-designed multivalent glycomimetics have been investigated in vitro [165,166,167,168,169,170,171].

In addition, carbohydrate conjugates can be utilized to develop carbohydrate-based vaccines for treatment of diseases in which carbohydrate toxins or antigens, respectively, are involved. Thus, fully synthetic glycoconjugates might be developed into vaccines for malaria, and other parasites such as different bacteria, tuberculosis, tropical diseases, leishmaniasis and the *Influenzae* virus.

The last two decades have witnessed the production and clinical testing of polysaccharide–protein conjugates specific for at least four different bacteria which normally cause considerable mortality and morbidity, especially in young children. In some cases, immunizing children from 4 months of age, with a booster early in the second year, has resulted in remarkably high success rates in protecting them from disease. For one pathogen, *Haemophilus*

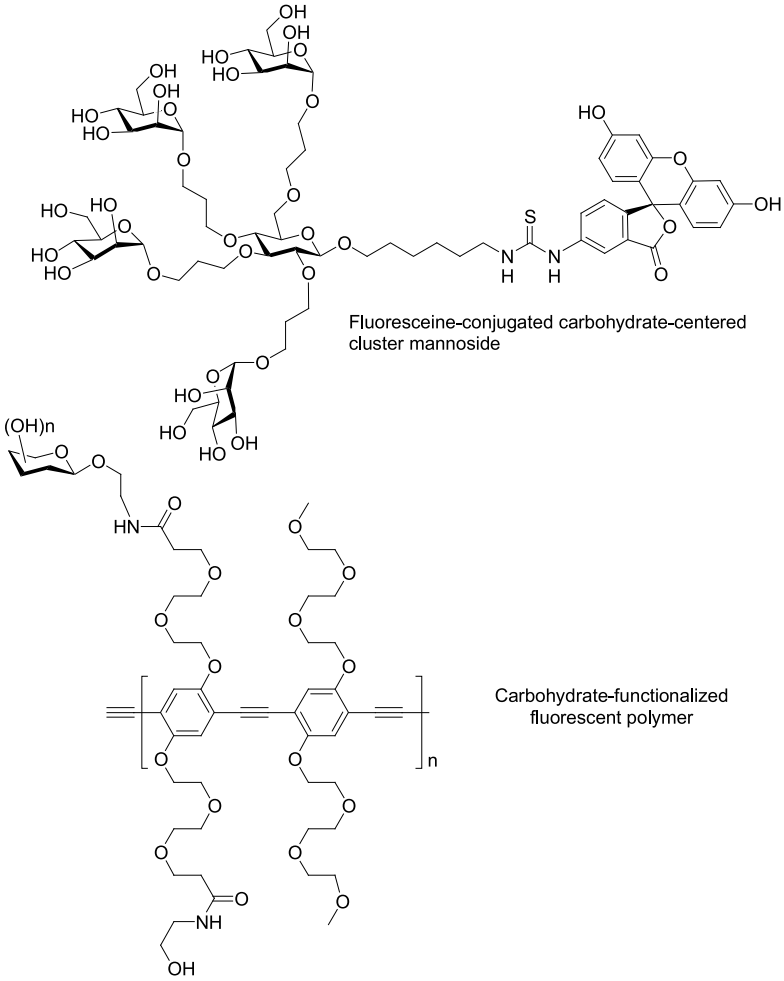


Figure 33
Fluorescent glycoconjugates

influenza type B, the success rate has been sufficiently high (> 95%). The results of immunization with conjugate vaccines to *Streptococcus pneumoniae*, *Neisseria meningiditis*, and *Salmonella typhi* are also very encouraging [172].

Some examples of research under development include a synthetic conjugated polysaccharide vaccine against *Haemophilus influenzae* type B, which has been described in 2004 (*Fig. 34*). The vaccine prototype was obtained by conjugation to human serum albumin (HSA) [173].

Furthermore, very recently a malaria vaccine has been suggested [174] which is based on the total synthesis of the malaria toxin responsible for the morbidity and mortality associated with malaria. The malaria parasite expresses a large amount of glycosylphosphatidylinositol

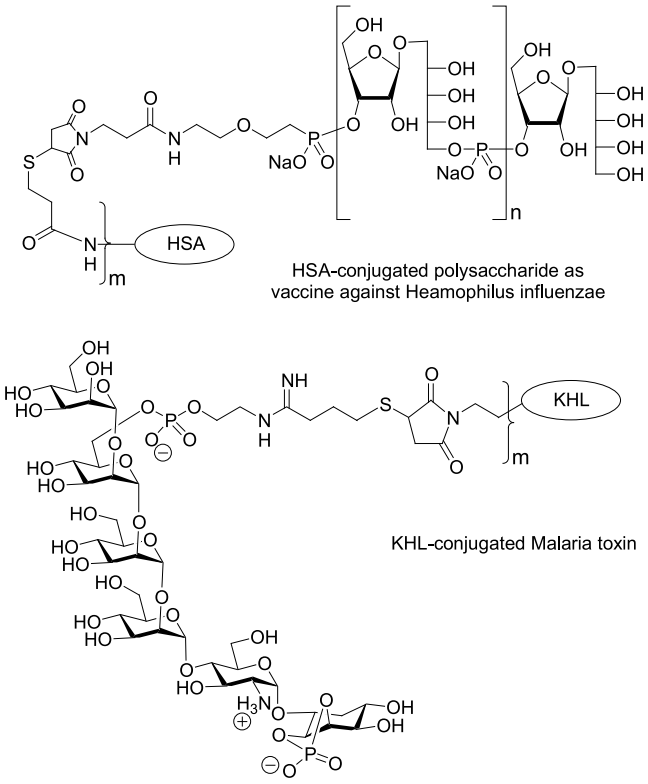


Figure 34
Protein-conjugated carbohydrates as vaccine candidates

(GPI) on the cell surface and there is evidence that the inflammatory cascade is triggered by this malaria toxin. A malaria vaccine could serve as an alternative to current malaria treatment which suffers from growing drug resistance. Fully synthetic material toxin has been conjugated to the carrier protein KHL (*Fig. 34*) and this conjugate proved highly immunogenic in mice and was able to prevent infection.

References

1. Löffler A, Doucey MA, Jansson AM, Müller DR, de Beer T, Hess D, Meldal M, Richter WJ, Vliegenthart JFG, Hofsteenge J (1996) Biochemistry 35:12005
2. Vliegenthart JFG, Casset F (1998) Curr Opin Struct Biol 8:565
3. Ihara Y, Manabe S, Kanda M, Kawano H, Nakayama T, Sekine I, Kondo T, Ito Y (2005) Glycobiology 15:383
4. Krieg J, Hartmann S, Vicentini A, Gläsner W, Hess D, Hofsteenge J (1998) Mol Biol Cell 9:301

5. Doucey MA, Hess D, Blommers MJ, Hofsteenge J (1999) Glycobiology 9:435
6. Hofsteenge J, Blommers M, Hess D, Furmanek A, Miroshnichenko O (1999) J Biol Chem 274:32786
7. Hartmann S, Hofsteenge J (2000) J Biol Chem 275:28569
8. Gonzalez de Peredo A, Klein D, Macek B, Hess D, Peter-Katalinic J, Hofsteenge J (2002) Mol Cell Proteomics 1:11
9. Hofsteenge J, Huwiler KG, Macek B, Hess D, Lawler J, Mosher DF, Peter-Katalinic J (2001) J Biol Chem 276:6485
10. Perez-Vilar J, Randell SH, Boucher RC (2004) Glycobiology 14:325
11. Ervin LA, Ball LE, Crouch RK, Schey KL (2005) Investig Ophthalmol Vis Sci 46:627
12. Doucey MA, Hess D, Cacan R, Hofsteenge J (1998) Mol Biol Cell 9:291
13. Li JS, Cui L, Rock DL, Li J (2005) J Biol Chem 280:38513
14. Vinogradov E, Bock K (1999) Angew Chem 111:712; Angew Chem Int Ed 38:671
15. Feeney RE, Osuga DT, Yeh Y (1994) J Carbohydr Chem 13:347
16. Jiang W-T, Hsiao K-F, Chen S-T, Wang K-T (1999) Synthesis 1687
17. Liu S, Wang W, von Moos E, Jackman J, Mealing G, Monette R, Ben RN (2007) Biomacromolecules 8: DOI: 10.1021/bm061044o
18. Nicolaou KC, Mitchell HJ (2001) Angew Chem 113:1624; Angew Chem Int Ed 40:1576
19. Křen V, Martínková L (2001) Curr Medicin Chem 8:1313
20. Jiménez C, Riguera R (1994) Nat Prod Rep 11:591
21. Duynstee HI, de Koning MC, Ovaa H, van der Marel GA, van Boom JH (1999) Eur J Org Chem 2623
22. Campbell R, Haworth T (1924) J Chem Soc 125:1337
23. Yoshida K, Kondo T, Okazaki Y, Katou K (1995) Nature 373:291
24. Gallant M, Link JT, Danishefsky SJ (1993) J Org Chem 58:343
25. Link JT, Raghavan S, Danishefsky SJ (1995) J Am Chem Soc 117:552
26. Link JT, Raghavan S, Gallant M, Danishefsky SJ, Chou TC, Ballas LM (1996) J Am Chem Soc 118:2825
27. Wood JL, Stoltz BM, Goodman SN (1996) J Am Chem Soc 118:10656
28. Wood JL, Petsch DT, Stoltz BM, Hawkins EM, Elbaum D, Stover DR (1999) Synthesis 1529
29. Dräger G, Garming A, Maul C, Noltemeyer M, Thiericke R, Zerlin M, Kirschning A (1998) Chem Eur J 4:1324
30. Toshima K, Takano R, Maeda Y, Suzuki M, Asai A, Matsumura S (1999) Angew Chem 111:3953; Angew Chem Int Ed 38:3733
31. Nicolaou KC, Vourloumis D, Winssinger N, Baran PS (2000) Angew Chem 112:46; Angew Chem Int Ed 39:44
32. Nicolaou KC, Sorensen EJ, Winssinger N (1998) J Chem Edu 75:1226
33. Neu HC (1992) Science 257:1064
34. Brown ED, Vright GD (2005) Chem Rev 105:759
35. Johnson DA, Liu H-W (1999) In: Pinto BM (ed) Comprehensive Natural Products Chemistry, vol 3. Elsevier, Amsterdam, p 311
36. Kirschning A, Oelkers C, Ries M, Schönberger A, Wohlert S-E, Rohr J (1999) In: Diederichsen U, Lindhorst TK, Westermann B, Wessjohann LA (eds) Bioorganic Chemistry. Wiley-VCH, Weinheim, p 298
37. August PR, Yu T-W, Floss HG (1999) In: Grabley S, Thiericke R (eds) Drug Discovery from Nature. Springer, Berlin Heidelberg New York, p 215
38. Carreras CW, Pieper R, Khosla C (1997) Top Curr Chem 188:85
39. Zerlin M, Thiericke R (1999) In: Diederichsen U, Lindhorst TK, Westermann B, Wessjohann LA (eds) Bioorganic Chemistry. Wiley-VCH, Weinheim, p 28
40. Nicolaou KC, Pavia MR, Seitz SP (1982) J Am Chem Soc 104:2027
41. Kirschning A, Bechthold AF-W, Rohr J (1997) Top Curr Chem 188:1
42. Lee JJ, Lee JP, Keller PJ, Cottrell CE, Chang C-J, Zähner H, Floss HG (1986) J Antibiot 39:1123
43. Patterson I, Mansuri MM (1985) Tetrahedron 41:3569
44. Davies HG, Green RH (1991) Chem Soc Rev 20:211
45. White JD, Bolton GL, Dantanarayana AP, Fox CMJ, Hiner RN, Jackson RW, Sakuma K, Warrier US (1995) J Am Chem Soc 117:1908
46. Wright DE (1979) Tetrahedron 35:1207
47. Beau J-M, Jaurand G, Esnault J, Sinaÿ P (1987) Tetrahedron Lett 28:1105
48. Nicolaou KC, Mitchell HJ, Suzuki H, Rodríguez RM, Baudoin O, Fylaktakidou KC (1999)

Angew Chem 111:3523; Angew Chem Int Ed 38:3334
49. Nicolaou KC, Rodríguez RM, Fylaktakidou KC, Suzuki H, Mitchell HJ (1999) Angew Chem 111:3529; Angew Chem Int Ed 38:3340
50. Nicolaou KC, Mitchell HJ, Rodríguez RM, Fylaktakidou KC, Suzuki H (1999) Angew Chem 111:3535; Angew Chem Int Ed 38:3345
51. Kong F, Zhao N, Siegel MM, Janota K, Ashcroft JS, Koehn FE, Borders DB, Carter GT (1998) J Am Chem Soc 120:13301
52. Yasumoto T, Murata M (1993) Chem Rev 93:1897
53. Haystead TAJ, Sim ATR, Carling D, Honnor RC, Tsukitani Y, Cohen P, Hardie DG (1989) Nature 337:78
54. Suh EM, Kishi Y (1994) J Am Chem Soc 116:11205
55. Andrianasolo EH, Gross H, Goeger D, Musafija-Girt M, McPhail K, Leal RM, Mooberry SL, Gerwick WH (2005) Org Lett 7:1375
56. Perron F, Albizati KF (1989) Chem Rev 89:1617
57. Barrett AGM, Peña M, Willardsen JM (1996) J Org Chem 61:1982
58. Barrett AGM, Pena M, Wilardsen A (1995) J Chem Soc Chem Commun 1147
59. Kinghorn AD, Kim N-C, Kim DSHL (1999) In: Ikan R (ed) Naturally Occurring Glycosides. Wiley, Chichester, p 399
60. Choi Y-H, Hussain RA, Pezzuto JM, Kinghorn AD, Morton JF (1989) J Nat Prod 53:118
61. Albrecht HP (1999) In: Ikan R (ed) Naturally Occurring Glycosides. Wiley, Chichester, p 83
62. D'Auria MV, Minale L, Riccio R (1993) Chem Rev 93:1839
63. Gao X, Patel DJ (1990) Biochemistry 29:10940
64. Ward DC, Reich E, Goldberg ICH (1965) Science 149:1259
65. Gao X, Patel DJ (1989) Biochemistry 28:571
66. Banville DL, Keniry MA, Shafer RH (1990) Biochemistry 29:9294
67. Lown JW (1993) Chem Soc Rev 165
68. Searle MS, Maynard AJ, Williams HEL (2003) Org Biomol Chem 1:605
69. Searle MS, Bicknell W, Wakelin LPG, Denny WA (1991) Nucleic Acids Res 19:2897
70. Johnson DA, Liu H-W (1999) In: Pinto BM (ed) Comprehensive Natural Products Chemistry vol 3. Elsevier, Amsterdam, p 311
71. Yang D, Wang AH-J (1994) Biochemistry 33:6595
72. Krohn K, Rohr J (1997) Top Curr Chem 188:127
73. Rohr J, Thiericke R (1992) Nat Prod Rep 9:103
74. Wohlert S-E, Bechthold A, Beninga C, Henkel T, Holzenkämpfer M, Kirschning A, Oelkers C, Ries M, Weber S, Weißbach U, Westrich L, Rohr J (1999) In: Diederichsen U, Lindhorst TK, Westermann B, Wessjohann LA (eds) Bioorganic Chemistry. Wiley-VCH, Weinheim, p 305
75. Nicolaou KC, Dai W-M (1991) Angew Chem 103:1453; Angew Chem Int Ed 30:1387
76. Smith AL, Nicolaou KC (1996) J Med Chem 39:2103
77. Lee MD, Ellestad GA, Borders DB (1991) Acc Chem Res 24:235
78. Lockhart TP, Bergman RG (1981) J Am Chem Soc 103:4091
79. Smith AL, Nicolaou KC (1996) J Med Chem 39:2103
80. Ikemoto N, Kumar RA, Ling T-T, Ellestad GA, Danishefsky SJ, Patel DJ (1994) J Am Chem Soc 116:9387
81. Drak J, Iwasawa N, Danishefsky S, Crothers DM (1991) Proc Natl Acad Sci USA 88:7464
82. Gomez-Paloma L, Smith JA, Chazin WJ, Nicolaou KC (1994) J Am Chem Soc 116:3697
83. Cin D-H (1999) Chem Eur J 5:1084
84. Maier ME (1995) Synlett 13
85. König B (2000) Eur J Org Chem 381
86. Depew KM, Zeman SM, Boyer SH, Denhart DJ, Ikemoto N, Danishefsky S, Crothers DM (1996) Angew Chem 108:2972; Angew Chem Int Ed 35:2797
87. Liu W, Christenson SD, Standage S, Shen B (2002) Science 297:1170
88. Schatz A, Bugie E, Waksman SA (1944) Proc Soc Exp Biol Med 55:66
89. Lemieux RU, Wolfrom ML (1948) Adv Carbohydr Chem 3:337
90. Waksman SA, Lechevalier HA (1949) Science 109:305
91. Francois B, Szychowski J, Adhikari SS, Pachamuthu K, Swayze EE, Griffey RH, Migawa MT, Westhof E, Hanessian S (2004) Angew Chem 116:6903; Angew Chem Int Ed 46:2971
92. Reimann H, Cooper DJ, Mallams AK, Jaret RS, Yehaskel A, Kugelman M, Vernay HF, Schumacher D (1974) J Org Chem 39:1451
93. Davies DH, Greeves D, Mallams AK, Morton JB, Tkach RW (1977) J Chem Soc Perkin Trans 1:1407
94. Kondo S, Ikeda Y, Ikeda D, Takeuchi T, Usui T, Ishii M, Kudo T, Gomi S, Shibahara S (1994) J Antibiotics 47:821
95. Michael K, Torr Y (1998) Chem Eur J 4:2091

96. Wang H, Torr Y (1998) Angew Chem 110:117; Angew Chem Int Ed 37:109
97. Hermann T, Westhof E (1998) J Mol Biol 276:903
98. Walter F, Vicens Q, Westhof E (1999) Curr Opin Chem Biol 3:694
99. Wright GD (2003) Curr Opin Chem Biol 7:563
100. Shaw KJ, Rather PN, Hare RS, Miller GH (1993) Microbiol Rev 138:138
101. Arya DP, Xue L, Willis B (2003) J Am Chem Soc 125:10148
102. Werz DB, Seeberger PH (2005) Chem Eur J 11:3104
103. Welzel P (1993) In: Krohn K, Krist H, Maas H (eds) Antibiotics and Antiviral Compounds-Chemical Synthesis and Modifications. VCH, Weinheim, p 373
104. Taylor JG, Li X, Oberthür M, Zhu W, Kahne DE (2006) J Am Chem Soc 128:15084
105. Ritzeler O, Hennig L, Fineisen M, Welzel P, Müller D, Markus A, Lemoine G, Lampilas M, van Heijenoort J (1997) Tetrahedron 53:1675
106. Kosmol R, Hennig L, Welzel P, Findeisen M, Müller D, Markus A, van Heijenoort J (1997) J Prakt Chem 339:340
107. El-Abadla N, Lapilas M, Hennig L, Findeisen M, Welzel P, Müller D, Markus A, van Heijenoort J (1999) Tetrahedron 55:699
108. van Heijenoort Y, Leduc M, Singer H, van Heijenoort J (1987) J Gen Microbiol 133:667
109. Wang R, Steensma DH, Takaoka Y, Yun JW, Kajimoto T, Wong C-H (1997) Bioorg Med Chem 5:661
110. Hashimoto H, Endo T, Kajihara Y (1997) J Org Chem 62:1914
111. Xu T, Werner RM, Lee K-C, Fettinger JC, Davis JT, Coward JK (1998) J Org Chem 63:4767
112. Ghuysen J-M (1991) Ann Rev Microbiol 45:37
113. Lovering AL, de Castro LH, Lim D, Strynadka NCJ (2007) Science 315:1402
114. Anikin A, Buchynsky A, Kempin U, Stembera K, Welzel P, Lantzsch G (1999) Angew Chem 111:3931; Angew Chem Int Ed 38:3703
115. Adachi M, Zhang Y, Leimkuhler C, Sun B, LaTour JV, Kahne DE (2006) J Am Chem Soc 128:14012
116. Nicolaou KC, Boddy CNC, Bräse S, Winssinger N (1999) Angew Chem 111:2230; Angew Chem Int Ed 38:2096
117. Williams DH, Bardsley B (1999) Angew Chem 111:1264; Angew Chem Int Ed 38:1172
118. Kahne DE, Leimkuhler C, Lu W, Walsh C (2005) Chem Rev 105:425
119. Mayville P, Novick RP, Muir TW (1999) Proc Nat Acad Sci 96:1218
120. Sieradzki K, Tomasz A (1999) New England J Med 340:5179
121. Sztaricskai F, Pelyvás-Ferenczik I (1994) In: Nagarajan R (ed) Glycopeptide Antibiotics. Dekker, New York, p 105
122. Hauser FM, Ellenberger SR (1986) Chem Rev 86:35
123. Nicolaou KC, Mitchell HJ, van Delft FL, Rübsam F, Rodríguez RM (1998) Angew Chem 110:1972; Angew Chem Int Ed 37:1871
124. Parker KA, Chang W (2003) Org Lett 5:3891
125. Harris CM, Harris TM (1982) J Am Chem Soc 104:4293
126. Evans DA, Wood MR, Trotter BW, Richardson TL, Barrow JC, Katz JK (1998) Angew Chem 110:2864; Angew Chem Int Ed 37:2700
127. Evans DA, Dinsmore CJ, Watson PS, Wood MR, Richardson TL, Trotter BW, Katz JL (1998) Angew Chem 110:2868; Angew Chem Int Ed 37:2704
128. Nicolaou KC, Natarajan S, Li H, Jain NF, Hughes R, Solomon ME, Ramanjulu JM, Boddy CNC, Takayanagi M (1998) Angew Chem 110:2872; Angew Chem Int Ed 37:2708
129. Nicolaou KC, Jain NF, Natarajan S, Hughes R, Solomon ME, Li H, Ramanjulu JM, Takayanagi M, Koumbis AE, Bando T (1998) Angew Chem 110:2879; Angew Chem Int Ed 37:2714
130. Nicolaou KC, Takayanagi M, Jain NF, Natarajan S, Koumbis AE, Bando T, Ramanjulu JM (1998) Angew Chem 110:2881; Angew Chem Int Ed 37:2717
131. Nicolaou KC, Mitchell HJ, Jain NF, Winssinger N, Hughes R, Bando T (1998) Angew Chem 111:253; Angew Chem Int Ed 38:240
132. Nicolaou KC, Li H, Boddy CNC, Ramanjulu JM, Yue TY, Natarajan S, Chu X-J, Bräse S, Rübsam R (1999) Chem Eur J 5:2584
133. Nicolaou KC, Boddy CNC, Li H, Koumbis AE, Hughes R, Natarajan S, Jain NF, Ramanjulu JM, Bräse S (1999) Chem Eur J 5:2602
134. Nicolaou KC, Koumbis AE, Takayanagi M, Natarajan S, Jain NF, Bando T, Li H, Hughes R (1999) Chem Eur J 5:2622
135. Nicolaou KC, Mitchell HJ, Jain NF, Hughes R, Winssinger N, Natarajan S, Koumbis AE (1999) Chem Eur J 5:2648
136. Boger DL, Miyazaki S, Kim SH, Wu JH, Loiseleur O, Castle SL (1999) J Am Chem Soc 121:3226

137. Williams DHR, Bardsley B (1999) Angew Chem 111:1264; Angew Chem Int Ed 38:1172
138. Ge M, Chen Z, Onishi HR, Kohler J, Silver LL, Kerns R, Fukuzawa S, Thompson C, Kahne D (1999) Science 284:507
139. Perkins HR (1969) Biochem J 111:195
140. Gruner SAW, Locardi E, Lohof E, Kessler H (2002) Chem Rev 102:491
141. Waltho JP, Williams DH, Selva E, Ferrari P (1987) J Chem Soc Perkin Trans 1 92:2103
142. Knapp S (1995) Chem Rev 95:1859
143. Knapp S, Jaramillo C, Freeman BJ (1994) Org Chem 59:4800
144. Haruyama H, Takayama T, Kinoshita T, Kondo M, Nakajima T, Haneishi TJ (1991) Chem Soc, Perkin Trans 1, 7:1637
145. Haubner R, Finsinger D, Kessler H (1997) Angew Chem 109:1440; Angew Chem Int Ed Engl 36:1347
146. Stockle M, Voll G, Gunther R, Lohof E, Locardi E, Gruner S, Kessler H (2002) Org Lett 4:2501
147. Haubner R, Wester H-J, Weber WA, Mang C, Ziegler SI, Goodman SL, Senekowitsch-Schmidtke R, Kessler H, Schwaiger M (2001) Cancer Res 61:1781
148. Hirschmann R, Nicolaou KC, Pietranico S, Salvino J, Leah EM, Sprengler PA, Furst G, Smith AB, II (1992) J Am Chem Soc 114:9218
149. Werz DB, Seeberger PH (2005) Chemistry Eur J 11:3194
150. Lee RT, Lee YC (1987) Glycoconjugate J 4:317
151. Lindhorst TK (2002) Top Curr Chem 218:201
152. Bertozzi CR, Kiessling LL (2001) Science 291:2357
153. Bovin NV (1998) Glycoconj J 15:1573
154. Röckendorf N, Lindhorst TK (2001) Top Curr Chem 217:201
155. Mammen M, Choi S-K, Whitesides GM (1998) Angew Chem 110:2908; Angew Chem Int Ed 37:2754
156. Kiessling LL, Gestwicki JE, Strong LE (2006) Angew Chem Int Ed 45:2348
157. Kleinert M, Röckendorf N, Lindhorst TK (2004) Eur J Org Chem 3931
158. Mrkisch M (2000) Chem Soc Rev 29:267
159. de la Fuente JM, Barrientos AG, Rojas TC, Rojo J, Cañada J, Fernández A, Penadés S (2001) Angew Chem 113:2317; Angew Chem Int Ed 40:2257
160. Brun MA, Disney MD, Seeberger PH (2006) ChemBioChem 7:421
161. Pankhurst QA, Connolly J, Jones SK, Dobson J (2003) J Phys D Appl Phys 36:R167
162. Dubber M, Lindhorst TK (2000) J Org Chem 65:5275
163. Disney MD, Zheng J, Swager TM, Seeberger PH (2004) J Am Chem Soc 126:13343
164. Bavington C, Page C (2005) Antiadhesion Respiration 72:335
165. Dubber M, Sperling O, Lindhorst TK (2006) Org Biomol Chem 4:3901
166. Sperling O, Fuchs A, Lindhorst TK (2006) Org Biomol Chem 4:3913
167. Bouckaert J, Mackenzie J, de Paz JL, Chipwaza B, Choudhury D, Zavialov A, Mannerstedt K, Anderson J, Piérard D, Wyns L, Seeberger PH, Oscarson S, De Greve H, Knight SD (2006) Molec Microbiol 61:1556
168. Nagahori N, Lee RT, Nishimura S-I, Pagé D, Roy R, Lee YC (2002) ChemBioChem 3:836
169. Kahane I, Ofek I (eds) (1996) Toward anti-adhesion therapy of microbial diseases. Adv Exp Med Biol, vol 408
170. Ofek I, Hasty DL, Sharon N (2003) FEMS Immunol Medic Microbiol 38:181
171. Hakomori S (1996) Cancer Res 56:5309
172. Ada G, Isaacs D (2003) Clinic Microbiol Inf 9:79
173. Verez-Bencomo V, Fernández-Santana V, Hardy E, Toledo ME, Rodríguez MC, Heynngnezz L, Rodriguez A, Baly A, Herrera L, Izquierdo M, Villar A, Valdés Y, Cosme K, Deler ML, Montane M, Garcia E, Ramos A, Aguilar A, Medina E, Toraño G, Sosa I, Hernandez I, Martínez R, Muzachio A, Carmenates A, Costa L, Cardoso F, Campa C, Diaz M, Roy R (2004) Science 305:522
174. Schofield L, Hewitt MC, Evans K, Siomos MA, Seeberger PH (2002) Nature 418:785

12.7 Glycoside vs. Aglycon: The Role of Glycosidic Residue in Biological Activity

Vladimír Křen
Institute of Microbiology, Academy of Sciences of the Czech Republic, Centre of Biocatalysis and Biotransformation, Vídeňská 1083, 142 20 Prague, Czech Republic
kren@biomed.cas.cz

In: *Glycoscience*. Fraser-Reid B, Tatsuta K, Thiem J (eds)
Chapter-DOI 10-1007/978-3-540-30429-6_67:

Abstract

A large number of biologically active compounds are glycosides. Sometimes the glycosidic residue is crucial for their activity, in other cases glycosylation only improves pharmacokinetic parameters. Recent developments in molecular glycobiology brought better understanding of aglycon vs. glycoside activities, and made possible the development of new, more active or more effective glycodrugs based on these findings – a very illustrative recent example is vancomycin. The new enzymatic methodology "glycorandomization" enabled preparation of glycoside libraries and opened up paths to the preparation of optimized or entirely novel glycoside antibiotics. This chapter deals with an array of glycosidic compounds currently used in medicine but also covers the biological activity of some glycosidic metabolites of known drugs. The chapter discusses glycosides of vitamins, polyphenolic glycosides (flavonoids), alkaloid glycosides, glycosides of antibiotics, glycopeptides, cardiac glycosides, steroid and terpenoid glycosides etc. The physiological role of the glycosyl moiety and structure-activity relations (SAR) in the glycosidic moiety (-ies) are also discussed.

Keywords

Natural products; Glycosides; Structure-activity relations; Glycorandomization; Flavonoid glycosides; Vitamin glycosides; Antibiotics; Alkaloid glycosides; Steroid glycosides; DNA interactions

Abbreviations

ATP adenosine triphosphate
CNS central nervous system
DNA deoxyribonucleic acid
GABA gamma amino butyric acid
HBA heavenly blue anthocyanidin
i.v. intra venous
PBMC peripheral blood mononuclear cells
p.o. per oral
MRSA methicilline-resistant *Staphylococcus aureus*
NK natural killer
RA retinoic acid
RAG retinoic acid glucuronide
RNA ribonucleic acid

1 Introduction

Many biologically active compounds are glycosides. Glycosides comprise several important classes of compounds such as antibiotics, hormones, sweeteners, alkaloids, flavonoids, etc. Sometimes the glycosidic residue is crucial for their activity; in other cases it only improves pharmacokinetic parameters. Owing to recent developments in molecular glycobiology a greater understanding has been gained of aglycon vs. glycoside activity and – based on these findings – it has become possible to develop new, more active or more effective glycodrugs.

Nevertheless, it is nearly impossible to define a general pattern of biological activities for the glycosides compared to the respective aglycons. Some well-selected examples can illustrate general trends and show the effects and potential applications of the compounds carrying the glycosidic moiety in relation to the respective aglycons.

It is generally accepted that glycosides are more water-soluble than the respective aglycons. Attaching the glycosidic moiety to a molecule increases its hydrophilicity. This effect influences pharmacokinetic properties of the respective compounds, e. g., circulation, elimination and concentration in the body fluids.

Modified hydrophilicity, however, influences mainly the membrane transport. Some compounds enter the cells just because of their "solubility" in the membrane components. Glycosylation can, in some cases, restrict or inhibit cell uptake of the particular compounds. Glycosylation can strongly influence transport through such important barriers as the hemato-encephalitic barrier and block the entrance of many compounds into brain tissue. Contrary to this, some glucosides can be transported actively into the brain tissue using the glucose-transport system. Another important barrier, in which glycosylation plays a crucial role is the placental barrier. Here entry of many glucuronides to fetal tissue is blocked, thus preventing intoxication by metabolites of xenobiotics. On the other hand, some glycosidic moieties can interact with receptors or lectins on the cell surfaces followed by their active uptake. A good example is the high affinity of β-galactosides to hepatocytes due to galectin-C occurring in high concentration on their surface.

Carbohydrates and glycosides recently emerged as a novel class of nucleic acid-binding compounds. Detailed study of the factors affecting the site-selectivity of some recently discovered antitumor antibiotics has shed new light on the role that oligosaccharides may play in nucleic acid recognition.

An important aspect for prediction of the respective glycoconjugate activities is also their susceptibility towards glycosidic cleavage at various sites of application. In the stomach and in the intestine most of the glycosides are hydrolyzed, either by the action of the acidic environment (stomach) or by the action of glycosidases (small intestine). There are, however, glycosides that are not hydrolyzed easily – e. g. α-galactosides – and such compounds are either unable to pass the hemato-intestinal barrier or they penetrate unhydrolyzed. Nonresorbed glycoconjugates can be cleaved later in the colon or metabolized by the action of the intestinal microflora. Glycosidases are present also in other body fluids, e. g. blood serum contains lysozyme that cleaves effectively β-*N*-acetylhexosaminides.

There exist, however, glycosides with specific individual biological activity that cannot be simply derived from the respective activity of the aglycon. The final activity is then given by the overall molecular structure.

Comparison of biological activities of aglycon and its respective glycoside can indicate some structure-effect correlations and also demonstrate the advantage (or uselessness) of introducing glycosyl moieties into pharmacologically interesting molecules.

Exploitation of the SAR data of the glycosides (most often glycosidic antibiotics) enabled synthetic modification and optimization of their glycon part. Recently, however, the new brilliant methodology "glycorandomization", based on the enzymatic modifications of glycosides, enabled the preparation of glycoside libraries. This method employing natural and mutant glycosyltransferases with wobbling specificities opened up new paths to the preparation of optimized or entirely novel glycoside antibiotics.

This chapter deals with "small molecules" – glycosides – aiming at the description of the function of the glycon part in the biological activities of the respective compounds. New, recently developed methods, leading to the bio-functional optimization of this moiety are also described.

2 Polyphenol Glycosides

Most of the polyphenols are produced by plants. Phenolic OH groups are generally good targets for biological glycosylations and many phenolic compounds occur virtually only in their glycosylated forms.

Glycosides of polyphenolic compounds, e. g. those of flavonoids, constitute a large group and novel findings increase their number extremely quickly. The list of flavone and flavonol glycosides known up to 1986 contains about 900 entries [1]. In the survey of Harborne and Williams [2] covering discoveries in anthocyanins and other flavonoids from January 1995 to December 1997 over 160 new glycosides of flavonoids are reported.

Many of them served as important components of traditional medicines. It is very interesting from the point-of-view of structure-effect correlations that minute alterations in their structures, e. g. positional changes of OH groups, bring about dramatic changes in their biological effects. Naturally, in the view of such minor changes, glycosylation of these compounds

sometimes completely changes their activity. For instance, quercetin (3,3′,4′,5,7-pentahydroxyflavon, 16) is now considered to be an extremely dangerous toxic and mutagenic compound (even though it was used previously as a colorant). However, its various glycosides [e. g., quercitrin ([17]), hyperin, spiraeosid, rutin (see also ❯ *Sect. 6.1.4* in this chapter)] occur frequently in plant material sometimes used as food, and many are considered to be beneficial to human health.

2.1 Anthocyanins

Anthocyanins are the most important group of water-soluble plant pigments visible to human eyes. Glycosylation of anthocyanins has many effects – physico-chemical effects such as antioxidant stabilization and also color stabilization or "deepening" caused by "sandwich stacking" controlled by the glycosyl substitution. This effect is well-known in "heavenly blue anthocyanidin" (HBA, **1**) (❯ *Scheme 1*) from *Ipomoea tricolor* (Convolvulaceae). HBA carries complex glycosyl-ester moieties attached to C-3 of peonidin (aglycon) containing ester-bound residues of caffeyl acid (❯ *Fig. 1*).

The whole structure then folds into a sandwich structure, in which aromatic residues of caffeic acid interact with the aromatic system of the aglycon peonidin [3]. This stacking results in the bathochromic shift of the complex. Moreover, this and similar glycosyl-acyl conjugates of the anthocyanidines stabilize these pigments at various pH within the plant vacuoles in the range of 4–6 against hydration and tautomerization resulting in profound changes of electron spectral properties (loss of color). Stability of color is often crucial for the biology of the plants, e. g. pollination by insects.

Anthocyanidins may also be important factors – with other flavonoids – in the resistance of plants to insect attack. Thus, the complete cyanidin 3-β-glucoside (**2**)(❯ *Scheme 2*) and not only the aglycon itself was shown to protect cotton leaves against the feeding of tobacco budworm [1].

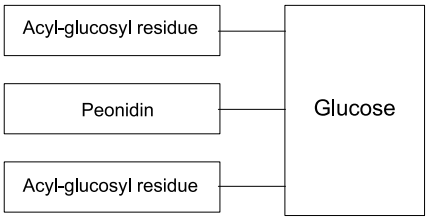
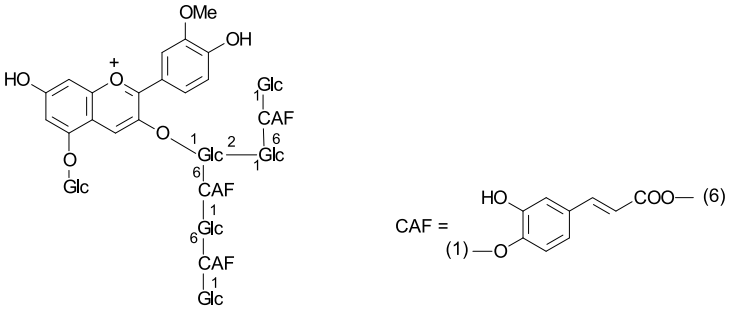


Figure 1
Model of the sandwich stacking

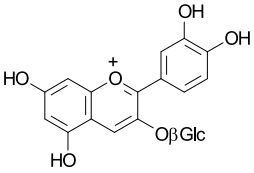
2.2 Flavonoid and Isoflavonoid Glycosides

The raison d'être for the proliferation of flavone and flavonol glycosides in nature continues to intrigue plant scientists. The ability of UV-B radiation to damage DNA, RNA and proteins as



1 (HBA) peonidin acyl-glycoside

Scheme 1



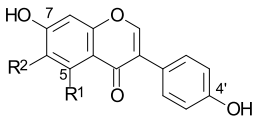
2

Scheme 2

well as to impair processes like photosynthesis is well known. Constituents of leaf epidermis such as polyphenols and their glucosides provide a means of absorbing the damaging radiation. An example of this has been demonstrated in the case of the needles of *Pinus silvestris* (pine tree). Two main compounds, isoquercitrin 3′,6′-di-*p*-coumarate and kaempferol 3-β-glucoside, have been found to be induced in seedlings under simulated global radiation. The concentration of the acylated kaempferol 3-β-glucoside reached 2.4 µmol/g fresh wt. [4].

In the case of plant pollens, it has long been recognized that flavonol glycosides are widely present and apparently contribute to the yellow color in the pollen. It is interesting that the two glycosides most frequently encountered are the 3-sophorosides of kaempferol and quercetin, and, furthermore, if these are not present, closely related 2′-*O*-glycosides of a flavonol 3-β-glucoside are normally encountered. Further examples are two methylated herbacetin 3-sophorosides reported in *Ranunculus*, *Raphanus* and *Klea* pollens [1,5]. The function of flavonoid glycosides in pollen is still uncertain in most plants, but in the case of *Petunia hybrida*, there is rather good evidence that it has a critical role in subsequent pollen tube growth, once the pollination has occurred. Structure-activity relationships have been explored in the case of *Petunia*, and a kaempferol diglycoside does appear to be the most active constituent [6].

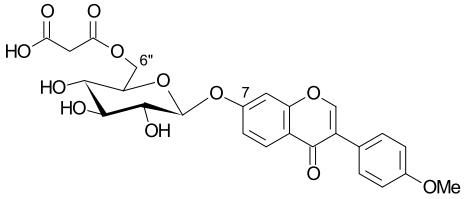
Glycosides and malonyl glycosides of isoflavones, and some other flavonoids often serve as their storage and transport forms. The most-studied systems for isoflavonoid glycoside metabolism are white lupine (*Lupinus albus*) root and cell cultures containing a range of mono- and diglucosides of genistein (**4**), 2′-hydroxygenistein and their 6- or 3′-prenyl derivatives [7], and soybean seeds and seedlings, and chickpea cell suspension cultures.



3 R^1 = H, R^2 = H

4 R^1 = OH, R^2 = H

5 R^1 = H, R^2 = OMe



6

Scheme 3

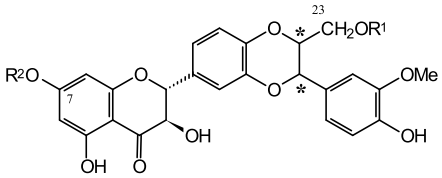
Features of the accumulation and metabolism of these compounds differ somewhat in the different species. In soybean seed hypocotyls, the 7-*O*-*β*-glucosides, 7-*O*-*β*-glucoside-6″-*O*-malonates, and 7-*O*-*β*-glucoside-6″-*O*-acetates of the isoflavones daidzenin (**3**), genistenin (**4**), and glycitein (**5**) occur (*Scheme 3*). All of them have been shown to increase during seed development in the pod to maximum levels between 45 and 60 days after flowering [8]. Three days after germination, the metabolism of the young leaf shifts from isoflavonoid to flavonoid accumulation [9], although low levels of isoflavone glycosides remain [10].

A similar pattern of isoflavonoids and their glycosides is observed in alfalfa where formonetin 7-*O*-*β*-glucoside-6″-*O*-malonate (**6**) accumulates in roots along with other isoflavonoid constituents [11].

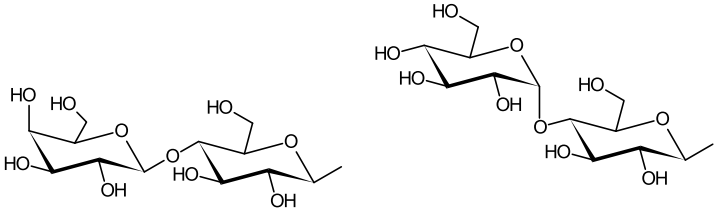
Infection of soybean with pathogenic fungus *Phytophthora sojae* leads to dramatic changes in isoflavone glycoside profiles and distribution. In the leaves, the nonglycosylated pterocarpan glyceolin accumulates to high levels only in the hypersensitive lesion formed in a resistant interaction, whereas the glucosides and malonyl glucosides of **3**, **4** and **5** accumulate in a broad area around the lesion [10]. In cotyledons the already constitutive pools of isoflavone glycosides are rapidly mobilized and, in the case of **3** the aglycon can be used for phytoalexin synthesis [12].

Some flavonoid glycosides are prepared synthetically or by biotransformations, usually for pharmaceutical purposes. Silybin (**7**) is a flavonolignan that is extracted from seeds of milk thistle (*Silybum marianum*) and it is used extensively as a potent hepatoprotectant and an antidote in mushroom poisoning. The major drawback of this compound is its low water solubility (about 0.43 g/L). Therefore, its glycosylation was accomplished by biological [13] and chemical methods [14].

Biotransformation using plant cell cultures afforded silybin 7-*β*-glucoside (**8**) [13] and chemical glycosylations gave silybin glycosides at C-23 (*β*-glucoside, *β*-galactoside, *β*-maltoside and *β*-lactoside **9**–**12**). The solubility of silybin glycosides was improved considerably compared to the aglycon (**9**, 13.0; **10**, 1.7; **11**, 3.8 and **12**, 5.6 g/L)(*Scheme 4*). Biological tests



7 R^1 = H, R^2 = H (silybin)
8 R^1 = H, R^2 = βGlc
9 R^1 = βGlc, R^2 = H
10 R^1 = βGal, R^2 = H
11 R^1 = βLac, R^2 = H
12 R^1 = βMalt, R^2 = H

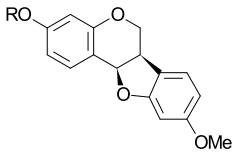


β-Lactosyl (β-Lac) β-Maltosyl (β-Malt)

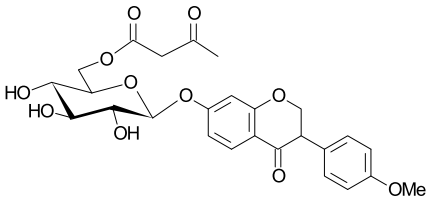
Scheme 4

showed that the above silybin glycosides have considerably higher antioxidative activity, the β-glucoside being the best that is ca. 10-times better than aglycon itself. Silybin monoglycosides show also better hepatoprotective activity than the aglycon. Similar effects were observed also in the tests with membrane lipoperoxidation of the mitochondrial membranes, in which silybin glycosides, mainly β-glucoside and β-galactoside, were found to be ca. 50–70% better antilipoperoxidants. The cytotoxicity at higher concentrations of silybin towards hepatocytes at conc. over 0.1 mM can be lowered by its glycosylation – in this case especially the β-maltoside and β-lactoside display substantially lower toxicity at high concentrations. Silybin β-galactoside was found to have better hepatoprotective activities in vivo presumably due to β-galactosyl residue causing higher affinity towards hepatocytes.

Flavonoids occurring virtually in all plant-derived food have a strong effect on the mammalian biology, mainly on immunity, inflammation and cancer [15]. A well described case of an immunomodulative effect for flavonoid glycoside is plantagoside (5,7,3′,4′,5′-pentahydroxyflavanone 3′-β-*O*-glucoside) that inhibits proliferative response of Balb/c spleen cells to the T-cell mitogen concanavalin A. On the contrary it has no effect on the mitogenic activity of lipopolysaccharides or phytohaemaglutinin thus demonstrating that the later two mitogens probably utilize activation pathways that are insensitive to this particular glyco-flavonoid. The fact that plantagoside is an α-mannosidase inhibitor [16] is of interest as well. Glycosides of isoflavonoids also have important functions in the communication between leguminous plants and nitrogen-fixing bacteria. Some of them are able to trigger *Nod* genes responsible for for-

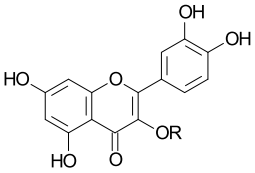


13 (-)-medicarpin R = H
14 R = βGlc



15

Scheme 5



16 quercetin R = H
17 quercitrin R = βGlc

Scheme 6

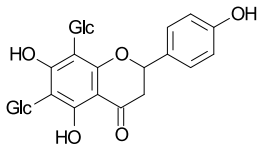
mation of lipochitooligosaccharide signal molecules, which in turn induce root hair curling and the cortical cell division that characterize the early development of the nodule.

Root exudates from alfalfa inoculated with *Rhizobium melioti* contain the pterocarpan medicarpin (**13**) and its β-glucoside (**14**), as well as formonetin 7-*O*-β-glucoside-6″-*O*-malonate (**15**) (Scheme 5). Levels of **14** are increased when plants are grown under low nitrogen conditions [17]. Formonetin and its 7-*O*-glucoside do not possess *Nod* gene-inducing activity for the alfalfa symbiont. Surprisingly, however, **15** can induce *Rhizobium nod* genes via interactions with both the NodD1 and NodD1-recognition proteins [18]. The Nod factors synthesized as a result of *Nod* gene activation are active on alfalfa roots at concentrations of around 10^{-9} M.

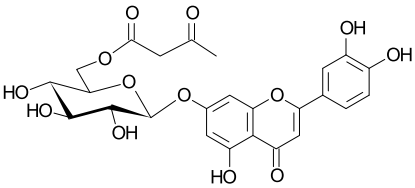
Flavonoids and their glycosides mediate also communication between plants and insects, serving like "pollination factors". This effect is caused mainly by the final color of the anthocyanin complex and is not related to the extent of glycosylation. Flavonoid glycosides also serve as feeding stimulants, e. g. isoquercitrin (quercetin 3-*O*-β-glucoside, **17**)(Scheme 6) that serves as a "biting factor" in feeding of *Bombyx mori* (silkworm) on the *Morus alba* leaves [19].

There are further examples of feeding stimulants from the group of flavonoid glycosides such as, e. g. kaempferol 3-*O*-β-xyloside and some others.

These compounds can be also effective antifeedants and here the glycosylation pattern is of crucial importance. Rutin [91] is a feeding stimulant to the tobacco hornworm, *Manduca sexta*,



18



19

Scheme 7

and quercetin 3-rhamnoside (less one glucose in comparison to rutin) acts as an antifeedant to the same species [20].

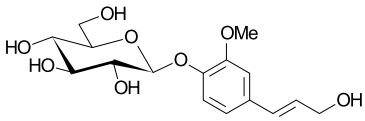
The larvae of most butterfly and moth species are monophagous or oligophagous, i. e. they feed from only one or a few (usually closely related) host plants. Caterpillars are much less mobile than winged adults and are unlikely to find the right food if the eggs, from which they hatch, have been laid on the wrong plant species. Thus, the oviposition choice of an adult female is crucial to larval survival. Most butterflies use visual and/or olfactory cues for the oviposition. The female will lay eggs only when she detects specific oviposition stimulants on the plant. As oviposition stimulants usually a complex cocktail is involved and flavonoid glycosides are an important part of this cocktail.

The first flavonoid glycoside that was found to induce ovipositional response in a butterfly was vicenin-2 (apigenin 6,8-di-*C*-glucoside, **18**) in the *Citrus*-feeding swallowtail *Papilio xuthus* [21]. Later more flavonoid glycosides were identified to be oviposition stimulants (naringenin 7-*O*-β-rutinoside, hesperetin 7-*O*-β-rutinoside and rutin) [22]. Another documented oviposition inducing flavonoid glycoside is 7-(6″-malonyl-β-D-glucopyranosyl)-luteolin (**19**) present in carrot (*Daucus carota*) that attracts *Papilio polyxenes* (Black swallowtail). The malonyl group in the flavonoid glycoside seems to be crucial, since the unmalonylated luteolin 7-glucoside and the corresponding glucuronide (also present in the carrot leaves) has little or no effect [23] (*Scheme 7*).

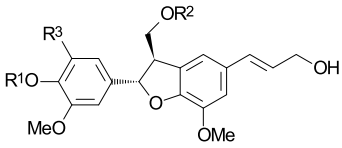
Flavonoid glycosides after intake by the insects are also sequestrated without hydrolysis and can be found in the insect tissues.

2.3 Lignins and Lignans

Coniferin (**20**) is one of the most abundant lignan glycosides occurring mostly in coniferous trees and some other plants. It is the dominant transport and storage form of coniferyl alcohol, which in turn constitutes the main building block of lignins.



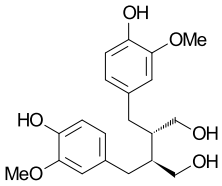
20



21 R^1 = βGlc R^2 = R^3 = H

22 R^1 = H R^2 = βGlc R^3 = H

23 R^1 = H R^2 = βGlc R^3 = OMe



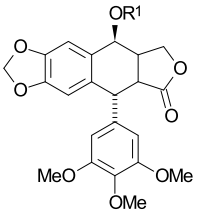
24

Scheme 8

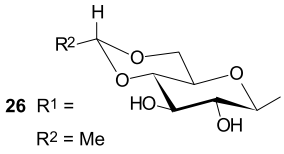
β-Glucosides of dehydroconiferyl alcohols (**21**–**23**) (*Scheme 8*) and the respective 2,3-epimers are growth factors isolated from transformed *Vinca rosea* tumor cells [24]. These compounds are not fragments of cell walls, but are produced directly from coniferyl alcohol [25]. Nonglycosylated derivatives do not have the cytokinine activity.

One of the most abundant lignans, for example in flaxseed, is di-β-glucoside of (+)-secoisolariciresinol (**24**) (both phenolic OH glycosylated). This compound is in vivo deglycosylated by intestinal microflora and by further demethylation and oxidation converted into enterolactone. These compounds called "mammalian lactones" enter the enterohepatic circulation, and there exists good evidence that their presence lovers incidence of the hormone-related cnacers [26]. They also lower the incidence of colon tumors. The above-mentioned diglycoside of **24** was also shown to act as an effective radical scavenger [27].

One of the most prominent lignans in terms of pharmacological application, in which glycosylation plays a crucial role, is podophylotoxin (**25**). This compound is the main active principle of podophyllin, a resin obtained from rhizomes of *Podophyllum peltatum*. Early pharmacological results with *Podophylum* lignans were disappointing. In spite of its antitumor promise direct administration of **25** was compromised by severe gastrointestinal toxicity. However, derivatized glycosides showed very promising activity in vitro and in vivo [28] and two of



25 podophylotoxin R^1 = H



26 R^1 =
R^2 = Me

27 R^2 =

Scheme 9

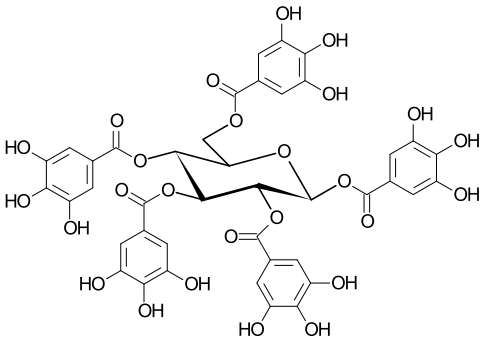
them, the ethylidene derivative etoposide (**26**) and the 4,6-alkylidene derivative of thiophencarbaldehyde, teniposide (**27**), were developed as anticancer drugs. In contrast to a classical spindle poison **25**, which causes arrest of the cells in the metaphase, attaching sugar moieties prevented tubulin interaction and thus microtubular assembly was not inhibited [29,30]. Etoposide (**26**) has proven to be useful in the treatment of patients with small-cell lung cancer, Kaposi's sarcoma, lymphoma and leukemia while teniposide (**27**) is mainly used to treat acute lymphatic leukemia, neuroblastoma and brain tumors in children (*Scheme 9*).

2.4 Gallotannins

Gallotannins (e. g., 1,2,3,4,6-penta-*O*-galloyl-β-D-glucopyranose, **28** (*Scheme 10*)) are abundant compounds of glycosidic character in plant material. As plant preparations these compounds had been used from ancient times as traditional medicine in wound healing, as an astringent and as an antidote in many intoxications (precipitation of the noxa) and the compounds also act as an effective antioxidant. These effects, however, cannot be attributed to the glycosidic moiety in the center of the molecule but mostly to the polyphenolic nature of the galloyl residues.

3 Glycosidic Antibiotics

Antibiotics form a large group of compounds of a vast structural diversity that are able to suppress or inhibit growth of one type of cells (microbial, fungal, tumor, genetically altered, etc.) and do not or to a lesser extent affect the growth of host cells. This definition is only very broad as it cannot embrace all the subtle effects of the various types of antibiotics.



28

Scheme 10

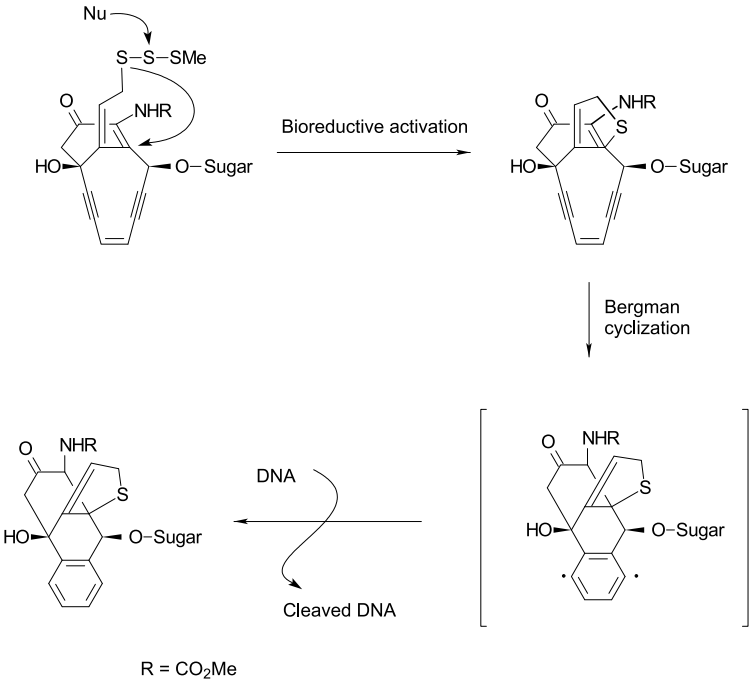
As in virtually all other groups of biologically active compounds, also in the group of antibiotics sugars play a very important role [31]. Occurrence of quite uncommon deoxy- and aminosugars is another typical feature of a large group of glycosidic antibiotics.
In this section we would like to focus on recent findings or typical cases.

3.1 Enediyne Antibiotics

The enediyne antibiotics [32] are extremely potent antitumor agents with a unique bicyclo[7.3.1]enediyne substructure. This group includes mainly calicheamycins [33], esperamycins [34], and dynemycins. Neocarzinostatin [35] and others (e. g., maduropeptin and kedarcidin) are also classified as belonging to this family. At room temperature in the presence of DNA, the core system of an enedyine antibiotic undergoes an interesting reaction yielding diradicals on sp^2 carbon that causes DNA strand breakage (*Scheme 11*). During studies of the structures of calicheamycin γ^1 and esperamycin A_1, it became apparent that enedyines could be triggered to aromatize via cleavage of the trisulfide with formation of a diradical intermediate as the biologically active species.

3.1.1 Calicheamycin

Calicheamycin γ^1 (**29**) was isolated from the cultivation broth of *Micromonospora echinospora* enriched with iodide. Compound **29** binds to the minor groove of DNA showing a high degree of base pair sequence-specificity, and a high degree of double strand to single strand cleavage of DNA. This occurs mostly at the sites 5′-TCCT/AGGA and 5′-TCTC/GAGA. The double stranded scissions suggest that **29** binds in the minor groove with the diradical positioned in such a way as to allow hydrogen abstraction from each strand [36]. The degree of specificity has been attributed to the carbohydrate moiety of the molecule, with the oligosaccharidic moiety aligned towards the 3′-end of the DNA. Studies have shown that the aglycon of **29** binds to DNA with lower affinity and in a nonsequence-specific manner [37]. Carbohydrates of calicheamycin γ^1 are lipophilic, which makes them favorable to interact with DNA.

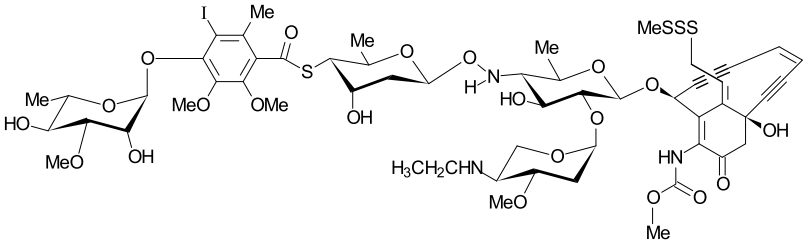


Scheme 11

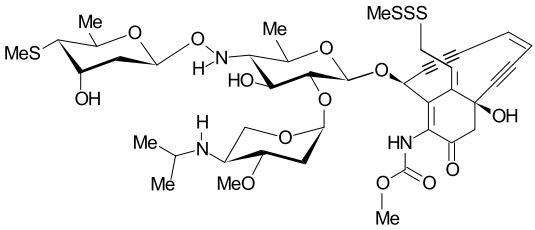
There is evidence that the iodine atom of the aromatic ring interacts with the C-2 amino groups of one or both guanines in TCCT:AGGA duplex DNA [38]. The binding energy of the oligosaccharide domain of **29** to the above DNA sequence within the duplex DNA is 31.6 kJ/mol. The sulfur atom of the third glycosidic moiety forms a hydrogen bond with an exposed amino proton from 3′-guanine in a drug-DNA complex [39]. This sugar was shown to be positioned edgewise in the minor groove allowing the aromatic ring to be placed between the minor groove with its iodine and methyl group positioned deep inside the minor groove.

Esperamycin C (**30**) can be considered to be an analogue of calicheamycine void of the rhamnosyl moiety and the aromatic part. Compound **30** causes double strand DNA cleavage but with considerably lower sequence selectivity. The aromatic and rhamnose moieties of **29** seem to determine the sequence specificity and the monodirectional mode of binding but they have little or no effect on orientating the aglycon within the DNA. NMR studies and molecular modeling demonstrated that also other sugar moieties in **29** specifically interact with DNA, e. g. the second carbohydrate moiety within the sugar-phosphate of DNA. The hydroxylamine glycosidic linkage causes two linked sugars to adopt an unusual eclipsed conformation giving the oligosaccharides the correct shape and rigidity to allow selective binding in the minor groove of DNA [38] (*Scheme 12*).

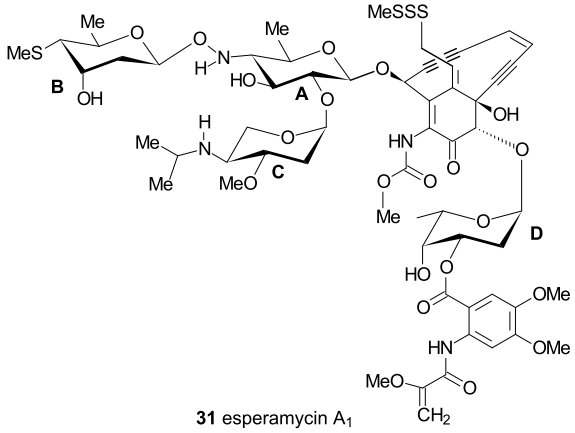
The monomer and the "head-to-head" dimer of the oligosaccharide segment of **29** have been shown to inhibit DNA transcription factors [39]. The dimeric glycoside was a more effective inhibitor than the corresponding respective monomer. The inhibitory effect has been attribut-



29 calicheamycin γ^1



30 esperamycin C



31 esperamycin A_1

Scheme 12

ed to the conformational changes in DNA caused by interactions with the sugar and this in turn interferes with the transcription factor DNA recognition. The base sequence selectivity of the calicheamycin carbohydrate to DNA may open up new possibilities for the chemical control of genetic information. It can, however, be stated that obviously deoxy- and substituted saccharides should be employed to enhance hydrophobic interactions.

3.1.2 Esperamycin

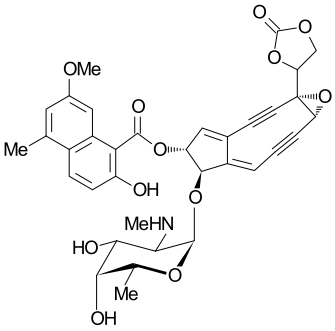
Esperamycins (esperamycin C **30**, esperamycin A_1 **31** and others), isolated from *Actinomadura verrucosospora*, display very strong antitumor activity due to DNA cleavage effected by the aglycon [40]. Esperamycin A_1 causes mostly single strand DNA cleavage and is far less sequence-specific than **29** having a similar aglycon.
Structure analysis of the esperamycin A_1–d(CGGATCCG) duplex showed that the drug bound in the minor groove with the methoxyanthranilate moiety intercalates at the (G2-G3):(6′-C7′) segment [41]. It was demonstrated that the isolated deoxyfucose anthranilate group did not interact with DNA, however, this moiety was shown to contribute 4–8 kJ/mol to the binding energy of **31**.
Esperamycin C causes double strand lesions involving deoxyribose hydrogen abstraction from the 4′- and 5′-positions, whereas esperamycin A_1 causes lesions from the 1′- and 5′-positions. This change to the 1′-position for esperamycin A_1 is believed to be due to deoxyfucose anthranilate, which intercalates in the DNA. The A-B sugars of the A-B-C trisaccharide are positioned deep into the minor groove. The OH groups of the trisaccharide A-B-C are positioned close to potential hydrogen bond acceptors, and favorable van der Waals interactions thus help to stabilize the complex [42].

3.1.3 Neocarzinostatin

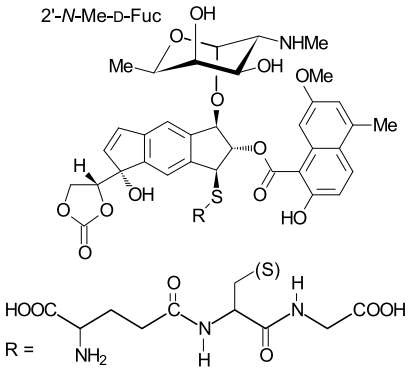
Neocarzinostatin, isolated from *Streptomyces carzinostaticus*, consists of a mixture of chromophore (**32**) and a protein [43]. This compound having potent antitumor and antibacterial activities induces single strand and double strand scissions in a ration 5:1, respectively, with a preference for T and A residues. The sequence specificity differs from other enedyines partly because of the different glycosylation pattern of the molecule.
The 2′-*N*-methyl-D-fucosamine, naphthoate and tetrahydroindacene of the postactivated-neocarzinostatin (rearranged complex with oligopeptide, **33**) (*Scheme 13*) were shown to have the correct structure to bind to specific sequences of DNA such as AGC sites. Specific sites are suggested for chemical modification that may alter selectivity of binding/cleavage, one of these sites being the N-2 of the 2′-*N*-methyl-D-fucosamine [44].

3.2 Anthracyclines

Anthracyclines represent a relatively large group of natural, semisynthetic and synthetic compounds [45]. Some compounds of this type are used in cancer treatment, several promising candidates are in clinical testing. Anthracyclines have also good antibacterial activity, however, due to their toxicity they are not used clinically in the treatment of infectious diseases. At present daunomycin (**34**) is used for the treatment of leukemia and adriamycin (**35**) is used for treatment of some solid tumors. These two compounds were studied in great detail. Modifications on both the aglycon and saccharidic parts have been performed. Changes of the saccharidic moiety were achieved by substitution of existing groups, e. g. *N*-alkylation, *N*-acylation, *O*-alkylation, formation of tetrahydropyranyl derivatives etc.



32 neocarzinostatin



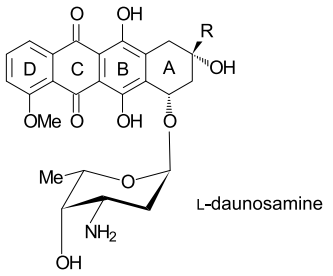
33 post-activated neocarzinostatin

Scheme 13

3.2.1 Daunomycin

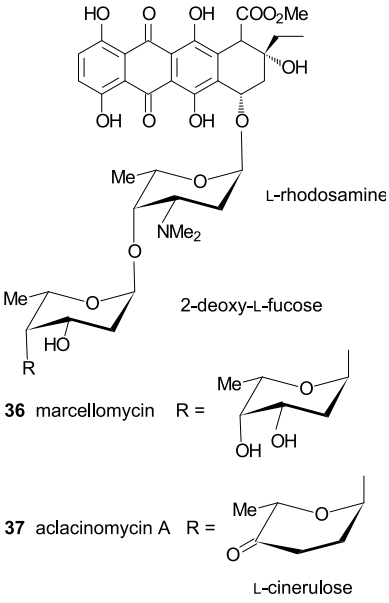
The anticancer activity of daunomycin (**34**) (*Scheme 14*) is attributed to the intercalation of the drug with DNA. This pattern of effect is common to all anthracyclines [46]. Two daunomycin molecules intercalate at each of the two C:G sites at either site of the duplex d(CGTACG). This binding results in an increase in base pair-separation from 3.4 to 6.8 Å as the molecule associates, which in turn leads to an unwinding of the helix and the formation of a noncovalent complex with the DNA. This results in the inhibition of DNA replication and RNA transcription [47].

Rings B-D of the aglycon are intercalated with the DNA, leaving ring A and the aminosugar as anchoring units. The cationic aminosugar and the hydroxy group of ring A fill the minor groove, which displaces water and ions from the groove. The aminosugar does not show hydrogen bonds to adjacent bases. The amino and hydroxy groups of the sugar face out of the minor groove, and it has been suggested that they may interact with polymerases, which could prevent or retard the action of these enzymes.



34 daunomycin R = COMe

35 adriamycin R = $COCH_2OH$



Scheme 14

It was observed that anthracyclines containing several glycosyl moieties have, in general, less side effects. For example, marcellomycin (**36**) and aclacinomycin A (**37**) show lower cardiotoxicity than daunomycin.

This is also the reason for the many attempts to introduce one or more additional carbohydrate moieties. Reaction of daunomycinone with diols providing its 7-*O*-hydroxyalkyl derivatives enabled attachment of the sugars in a different manner leading to derivatives with lower toxicity [48].

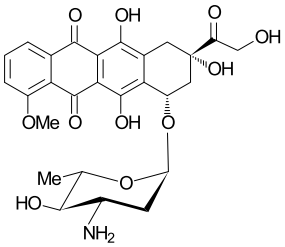
Nogalamycine, which contains two sugars (one attached to the D-ring and the second attached at C-7 of the A-ring), binds to DNA in a slightly different manner [31].

3.2.2 Glycosyl Derivatives of Anthracyclines

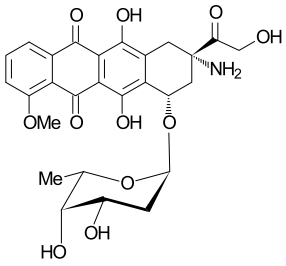
In natural anthracyclines the sugars are linked by α-glycosidic linkages, the corresponding β-glycosides were found to be inactive [49]. L-arabino analogues of daunomycin and doxorubicin, 4′-epidaunomycin and 4′-epidoxorubicin (**38**), were shown to have similar activity to the natural anthracyclines, the latter (**38**) also has a lower cardiotoxicity and is now used clinically [50] (*Scheme 15*).

Neutral synthetic sugar derivatives, for example SM-5887 (**39**), display good antitumor activity with reduction in local tissue toxicity and cardiotoxicity compared to, for example, adriamycin (**35**) [51].

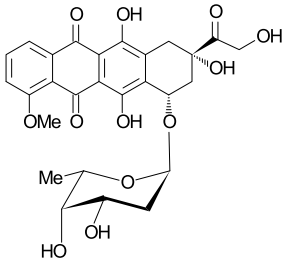
Drug resistance in the anthracycline cancerostatics has been addressed by the study of the influence of the amino group of daunosamine. It seems that this group is recognized by the P-gp multidrug transporter (P-glycoprotein is an ATP-dependent transmembrane drug exporter). Comparative studies of doxorubicin and hydroxyrubicin (**40**) in drug-resistant tumor cells demonstrated that **40** partially or completely circumvents P-gp-mediated drug resistance due to its decreased transport by P-gp compared with doxorubicin [52].



38

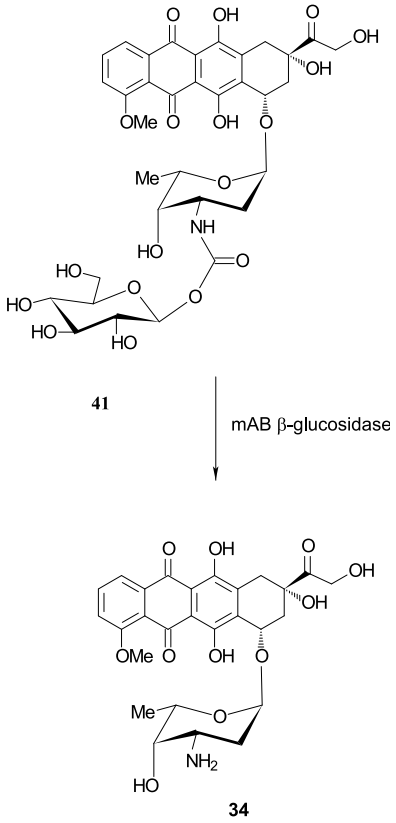


39



40

Scheme 15



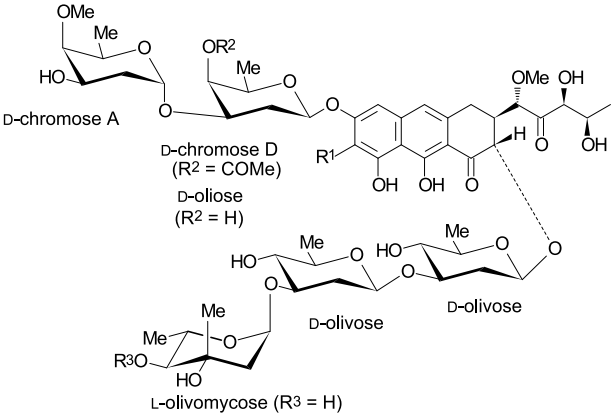
Scheme 16

There have been also attempts to prepare targeted drugs based on anthracyclines employing the recognition of some specific sugars by receptors on the tumor cells. Thus, a β-glucoside derivative of adriamycin (**41**) was used in an antibody-directed enzyme-prodrug therapy (*Scheme 16*) [53].

3.2.3 Aureolic Acids

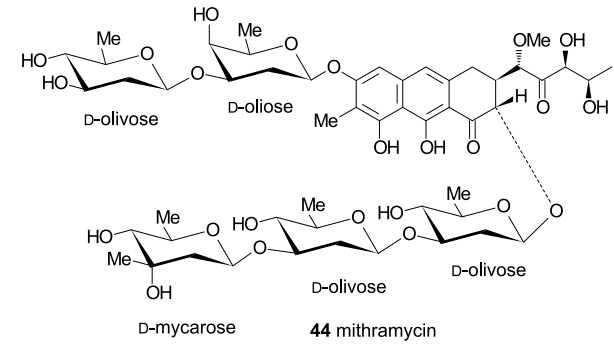
Members of the aureolic acid group include chromomycin A_3 (**42**), olivomycin A (**43**) and mithramycin (plicamycin) (**44**). The aglycon of chromomycin A_3 is identical to that of mithramycin, but differs from that of olivomycin A by a methyl group. Conversely, the sugar components for chromomycin A_3 and olivomycin are nearly identical, while those in mithramycin are different (*Scheme 17*).

These drugs require a divalent metal ion, preferably an Mg^{2+} ion and a guanine-containing target for activity. They were identified to be inhibitors of DNA-dependent RNA polymerase. It was found that **42** lacking some of the sugar moieties was less active [54] and the aglycon does not bind to DNA at all [55].



42 chromomycin A_3 R^1 = Me, R^2 = R^3 = COMe

43 olivomycin A R^1 = H, R^2 = COMe, R^3 = $COCHMe_2$



44 mithramycin

Scheme 17

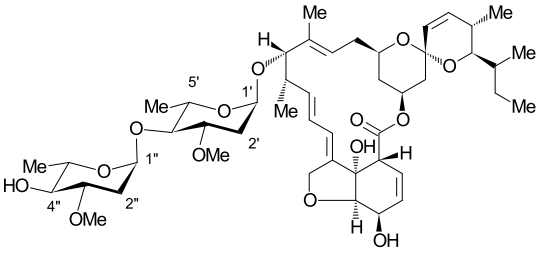
The mode of action of these antibiotics is different from other anthracyclines despite their aglycon similarity. Aureolic acids do not intercalate into DNA but they bind to CG rich regions. NMR studies demonstrated that **42** forms a symmetrical dimer with self complementary d(TTGGCCAA) with the hydrophobic edge of the chromophore located next to the GC:CC site [56]. The sugars in the olivose-olivose-olivomycose trisaccharide extend towards the 3′-direction of the octanucleotide in the minor groove. The chromose disaccharide and the hydrophilic side chain are directed to the phosphate backbone. The trisaccharide part olivose-olivose-olivomycose is essential for stabilization of the drug-Mg^{2+} complex that binds effectively to DNA, removal of chromose A has no effect on the complex formation [57].

The action of mithramycin (**44**) is virtually the same, however, it is less specific for the respective DNA sequence, which can be explained by the additional hydrogen-binding potential of D-mycarose [58].

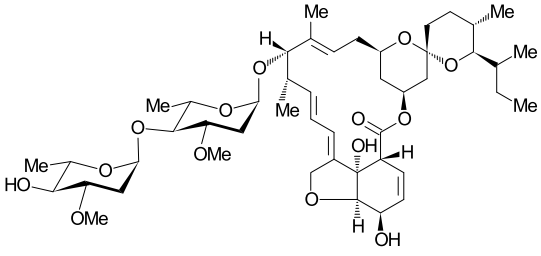
3.3 Avermectins

This class of macrolide antibiotic has mostly antiparasitic activity. Avermectin B_{1a} (**45**) and ivermectin (**46**) (Scheme 18) are used mostly in veterinary medicine, however, some semisynthetic derivatives are also used for treatment of onchocerciasis in humans [59]. The action of avermectin is believed to stimulate specific chloride ion transport systems increasing the membrane permeability to Cl^- ions via GABA (γ-butyrate) receptors and non-GABA receptors [60].

The aglycon of avermectin has poor antiparasitic activity. Positions 4′ and 4″ of the oleandrose moieties were modified partly because of feasibility. The synthesis of the 4″-amino-4″-deoxyoleandrose derivative of avermectin B_1 and ivermectin was based on the observation that most macrolide antibiotics contained a basic amino group [61]. One of the derivatives, 4″-*epi*-acetamido-4″-deoxyavermectin B_1, is currently under development as a novel avermectin endectoside [61]. Besides, also 2″-α-fluoro (*ax.*) and 2″-β-fluoro (*eq.*) derivatives of avermectins were prepared to strengthen the glycosidic bond. These derivatives have interesting activities in some bioassays compared to the parental compounds [62].



45 avermectin B_{1a}



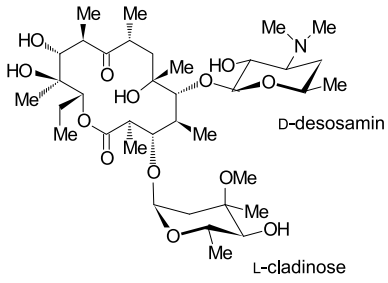
46 ivermectin

Scheme 18

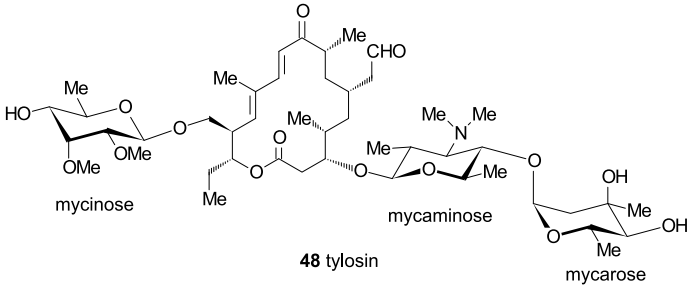
3.4 Macrolide Antibiotics

Macrolide antibiotics containing a large lactone ring (hence the title macrolide) consisting of 14 atoms are biosynthetically related to avermectins being also polyketides. The first member of this group erythromycin A (**47**) was isolated in 1952 [63] and it soon became an important antibiotic widely used in clinical medicine against Gram-positive bacteria. It constitutes the main treatment for many pulmonary infections such as Legionnaire's disease. Other families of macrolides are known containing 12- or 16-membered macrolide rings. A representative of the latter family is tylosin (**48**) (*Scheme 19*), a commercially important antibacterial agent used in veterinary medicine. The 12-membered ring macrolides have not been used clinically. Most of the macrolides contain deoxy- and aminosugars. Although it was established [64,65] that the sugar moieties are absolutely essential for the microbial activity of these compounds it is difficult to judge their role in the activity as the exact mode of action is unclear.

The mode of action of macrolide antibiotics involves the inhibition of protein synthesis of specific binding to the 50S ribosomal subunit but without a specific target at the 23S ribosomal subunit and various proteins [66]. Nevertheless, the exact interaction of the macrolide and the ribosome unit is still not fully understood. In principle, the macrolide antibiotic should inhibit also mammalian mitochondrial protein synthesis but they are unable to penetrate the mitochondrial membrane.



47 erythromycin A



48 tylosin

Scheme 19

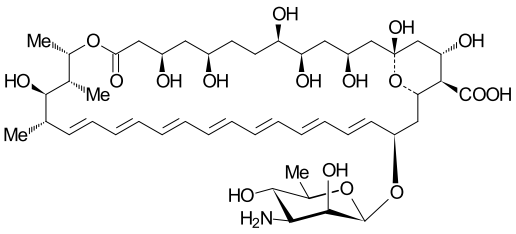
Sugar moieties have been linked with many pharmacological properties displayed by the macrolide antibiotics [67]. Basicity of the nitrogen in aminosugars improves the active transport of the compounds into the cell, however, its protonation/deprotonation does not affect the uptake. It was shown [68] that the more basic macrolides are the more effective ones. Esterification of the glycosidic position 2′ improves pharmacological and pharmacokinetic properties of the resulting drug lowering unwanted gastrointestinal side effects [69].

3.5 Polyenes

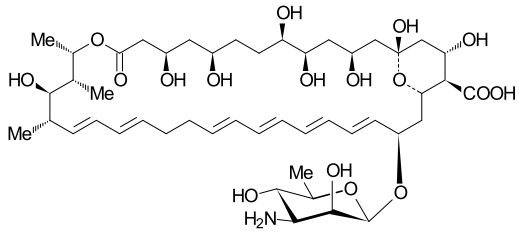
Polyene macrolides consist of a macrocyclic lactone, a polyene chromophore, and a polyhydroxylated chain which often bears glycosidic residues. These antibiotics, e. g., amphotericin B (**49**) and nystatin A_1 (**50**) (› *Scheme 20*), are used as potent antifungal antibiotics although they cause side effects such as nephrotoxicity, hemolytic anemia and electrolyte abnormalities [70].

These compounds interact with ergosterol contained in cell membranes of lower eukaryotes (fungi, yeasts) forming pores in their membranes. These channels cause a shunt in the membrane potential and leach the cellular ions [71].

Structure-activity studies on amphotericin B (**49**) have shown the importance of the basic nitrogen of D-mycosamine for the activity [72]. The polyene is orientated with the polar head at the



49 amphotericin B



50 nystatin A_1

Scheme 20

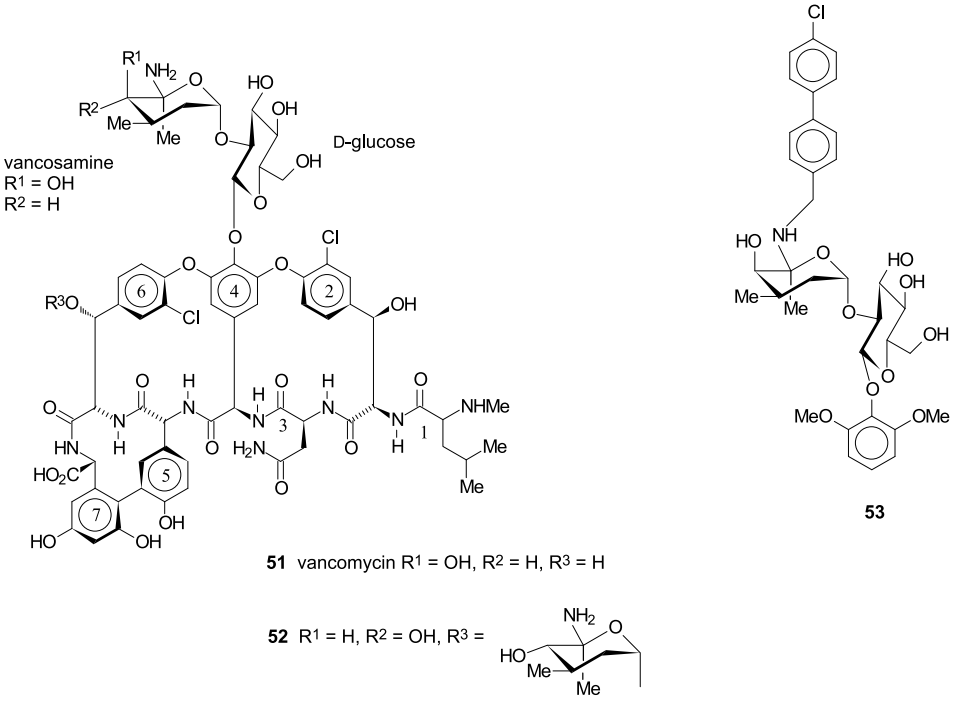
membrane–water interface and it has been suggested that the basic amino group of the sugar and the carbonyl group form a "cage-like" hydrogen-bonded structure with sterol containing a 3-β-hydroxy group and water.
N-Glycosylation of D-mycosamine in the polyenes leads to antibiotics with reduced toxicity and similar activity to the parent compound but increased water solubility and, therefore, better bioavailability [73].

3.6 Glycopeptides

This group of antibiotics is made up of complex polypeptide aglycones, which are abundantly glycosylated with mono-, di- and tetrasaccharides [74]. To date about 100 members of this group are known, such as vancomycin, teicoplanin, bleomycin, ristocetin etc. Some of them are very important, as they are often *ultimum refugium* – the last possible treatment for multiply resistant bacterial infections. Recent discoveries demonstrated that glycosidic residues play crucial roles in their activity and this opens up a large area for developing new glycosidic antibiotics, as shown in the case of vancomycin.
The story of the involvement of glycosyl moieties in vancomycin activity is probably one of the most illustrative and also very topical as glycosyl derivatives of vancomycin quite recently proved to be very effective against multiply resistant (even vancomycin-resistant) bacteria.

3.6.1 Vancomycin

Vancomycin (**51**) (❯ *Scheme 21*) [75] is a glycopeptide antibiotic that is widely used in the treatment of Gram-positive bacterial infections. It was discovered in a soil sample from the jungles of Borneo during a research program carried out by the pharmaceutical company Eli Lilly in the mid-1950s. At present its importance has increased considerably as it is one of the few antibiotics effective against nosokomial infections, e. g., multiply resistant bacteria – a typical example is methiciline-resistant *Staphylococcus aureus* (MRSA), and together with gentamycin it is an antibiotic of last resort.
Vancomycin consists of a core heptapeptide with attached saccharide moieties, one of which is the deoxyaminosugar vancosamine. Vancomycin exhibits its antibacterial activity by binding bacterial cell wall mucopeptide precursors terminating in the sequence L-Lys-D-Ala-D-Ala [76], thereby impeding further processing of these intermediates into peptidoglycans. Five hydrogen bonds account for this binding specificity, and the disruption of one of these hydrogen bonds by the replacement of the terminal alanine with lactate (D-Ala-D-Lac) in the mucopeptide precursor is the molecular basis for the resistance to **51**.
Because of the ultimate importance of vancomycin in the treatment of often fatal infections caused by multiply resistant bacteria, extensive efforts have been directed toward the discovery of vancomycin derivatives with activity against the drug-resistant bacteria [77]. As a result of these efforts, several derivatives such as **52** (LY264826) have been found to be up to 500 times more effective than vancomycin itself. The most notable difference is the presence of another aminosugar (R^3) attached to the amino acid 6. This extra sugar possibly facilitates the dimerization of the antibiotic and/or anchors the antibiotic to the cell membrane, both of which have been shown to be important for vancomycin activity [75]. It was also demonstrated that the conformations of vancomycin and its aglycon differ in their alignment of the amide protons



Scheme 21

that participate in the hydrogen-binding network with cell-wall precursors [78]. Alkylation of the 3-amino group on the disaccharide at amino acid residue 4 further enhances the activity, probably by serving as a hydrophobic anchor to the cell membrane [79].

Another target for modification of vancomycin is the vancosamine moiety. Recently, it was found that *N*-alkylation with *n*-decyl or 4-chlorobiphenylyl groups results in an antibiotic acting by a different mechanism than vancomycin itself [80] and, therefore, exhibits activity against vancomycin-resistant microorganisms. Sugar-altered vancomycin interferes with the polymerization of a glycopeptide monomer and disrupts its ability to form cell-wall material in the first place. Thus, altered vancomycin may disrupt cell-wall formation at multiple targets of the glycopeptide monomer, straight-chain polymer, and cross-link reinforcement.

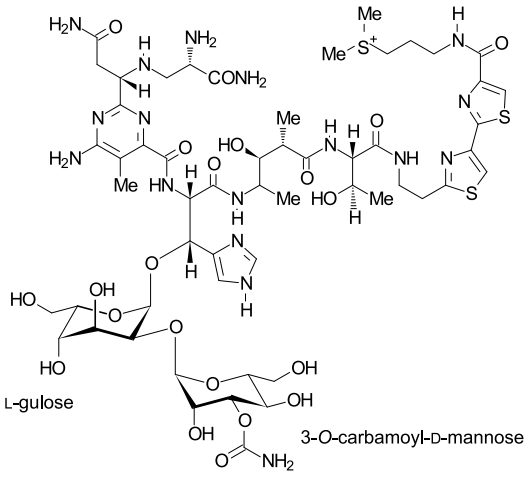
In fact, the altered disaccharide derivative itself (**53**) has strong antibacterial activity, even if not attached to the rest of the large vancomycin molecule [80]. This compound, however, acts by a different mechanism than vancomycin. It inhibits the transglycosylation step that precedes the transpeptidase step blocked by vancomycin itself. This finding explains a complex mechanism of action of the derivative **52**.

These findings offer possibilities to use sugar chemistry to make glycopeptide antibiotics such as vancomycin more potent against bacteria, for example, if vancosamine was replaced by daunosamine (3-desmethyl-vancosamine). Although this derivative differs from **51** only in one methyl group at the terminal sugar, it shows some notable differences in activity [81]. A general methodology for selective glycosylation of the vancomycin aglycon has been developed

by Kahne's group [82]. It is quite likely that a wide-ranging investigation of different sugars will lead to more significant improvements across a range of bacterial strains. This case and analogous approaches are discussed in detail in ❯ *Sect. 8* of this chapter.

3.6.2 Bleomycin

Bleomycin A_2 (**54**) (❯ *Scheme 22*) has antimicrobial activity against both Gram-positive and Gram-negative bacteria but it is clinically often used as an antitumor agent. The mode of action of bleomycin is DNA strand scission occurring in the presence of Fe^{2+} and molecular oxygen [74]. It was demonstrated that the sugars do not contribute to the binding affinity or the DNA cleavage selectivity, although the presence of the sugar enhances the DNA cleavage efficiency by 2–5 times [78]. The sugar moieties are suggested to contribute to the binding of O_2 and activation and protection of the reactive iron-oxo or perferryl intermediate to activated bleomycin [83].

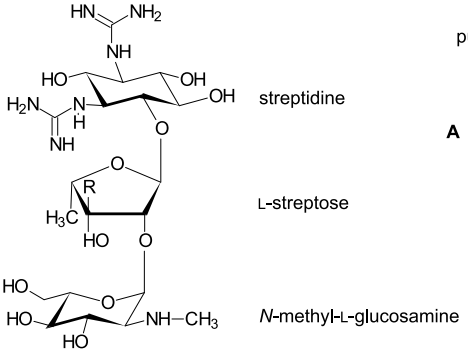


54 bleomycin A_2

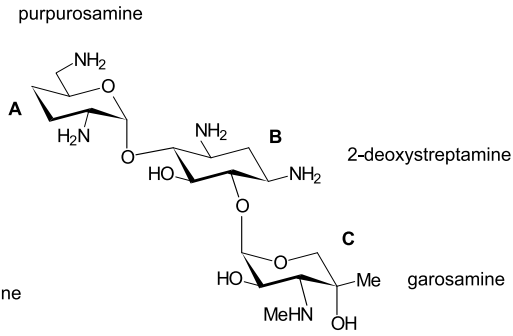
Scheme 22

3.7 Aminocyclitol Antibiotics

The term "aminoglycoside" refers to the structural aspects. Most of the aminoglycoside antibiotics contain aminosugar(s) and an aminocyclitol or a cyclitol moiety; therefore, this group is also called "aminocyclitol antibiotics." A review on structural aspects, synthesis and chemical modifications of these compounds was published by Ikeda and Umezawa in 1999 [210].
Streptomycin (**55**), dihydrostreptomycin (**56**) and gentamycin (**57**) (❯ *Scheme 23*) are typical representatives of a large group of the aminocyclitol antibiotics. They interact with ribosomal RNA causing a decrease in translational accuracy and inhibit translocation of the ribosome [84]. These antibiotics bind to a conserved sequence of rRNA that is near the site of codon-anticodon recognition in the aminoacyl-tRNA site of the 30S subunit.



55 streptomycin R = CHO
56 dihydrostreptomycin R = CH_2OH



57 gentamycin A

Scheme 23

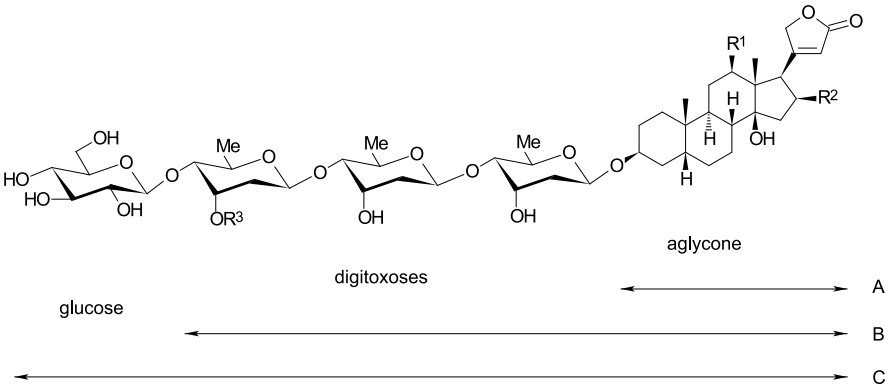
A comparison of various aminoglycosidic antibiotics suggests chemical groups that are essential for their function [85]. Both amino groups of the B ring (see e. g. in **57**), which are positively charged at physiological pH, contribute to RNA binding as hydrogen donors.
Recent investigation of the binding mechanism of gentamycin C1A (**57**) has shown that it binds in the major groove of the RNA. Rings A and B of gentamycin direct specific RNA-drug interactions. Ring C of gentamycin, which distinguishes this subclass of aminoglycosides, also directs specific RNA interactions with conserved base pairs [86].
Resistance to this type of antibiotic is usually mediated by acetylation, adenylation and phosphorylation mostly in rings A and B causing thus lowered or abolished antibiotic affinity to the RNA [87].

4 Steroid and Terpenoid Glycosides

This group involves a considerable number of physiologically active compounds whose activity is largely dependent on the complete structure, including the glycosidic moiety. A number of these compounds have detergent properties, e. g., saponins, and such physicochemical properties are partly responsible for their toxicity, e. g. the hemolytic activity of some saponins is caused by the damage of the erythrocyte membrane.

4.1 Cardiac Glycosides

Cardiac glycosides consist of a five-ring cardenolide aglycon, called a genin, with a number of attached monosaccharides, which often include deoxysugars. These steroidal compounds are usually isolated from plant material (digitoxin, strophantidine), but they have been dis-



Structure part	Description	R^1	R^2	R^3
Part A (aglycon)	Digitoxigenin	H	H	–
	Gitoxigenin	H	H	–
	Digoxigenin	OH	H	–
Part B (trioside)	Digitoxin	H	H	H
	Gitoxin	H	OH	H
	Digoxin	OH	H	H
Part C (tetraoside)	Purpurea glycoside A	H	H	H
	Purpurea glycoside B	H	OH	H
		OH	H	H
Acetylated tetraosides	Lanatosid A	H	H	CH_3CO
	Lanatosid B	H	OH	CH_3CO
	Lanatosid C	OH	H	CH_3CO

Figure 2

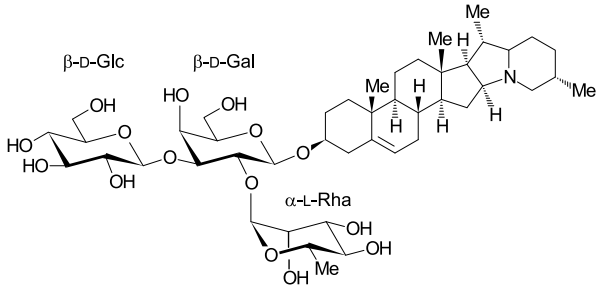
covered in higher mammals as adrenal cortical hormones (ouabain) [88]. Possibly using the same binding site as the natural hormone, cardiac glycosides inhibit Na^+/K^+-ATPases [89,90], resulting in an ionotropic activity that has proven to be useful in the treatment of various heart conditions, e. g., atrial tachyarrythmias, and it is used to produce a positive ionotropic effect in congestive heart failure [91]. This activity is enhanced several-fold due to the presence of the sugars in these compounds [92,93]. The sugar moiety is thought to be important in uptake, distribution and binding stability [94]. Cleavage of the sugar moieties forms the genin (*Fig. 2*) which retains a diminished affinity while the sugar moiety alone has no activity [94,95]. Altering the structure of the sugar(s) has a dramatic influence on the activity of the cardiac glycoside. Cardiac glycosides having $\beta(1\rightarrow4)$ sugar linkages have stronger activities than those with $\beta(1\rightarrow6)$ and $\beta(1\rightarrow2)$ linkages [96,97], α-linkages further diminish the activity [98]. Sugars having 4′-OH axial groups (e. g., galactose) diminish the binding, 3′-OH axial and 2′-OH equatorial groups appear to contribute to the binding [99]. Also, the addition of a 6′-OH group slightly diminishes the activity [99].

Therefore, existing natural digitoxose/glucose containing cardenolides seem to be optimal despite extensive structure-effect studies targeted to its optimization.

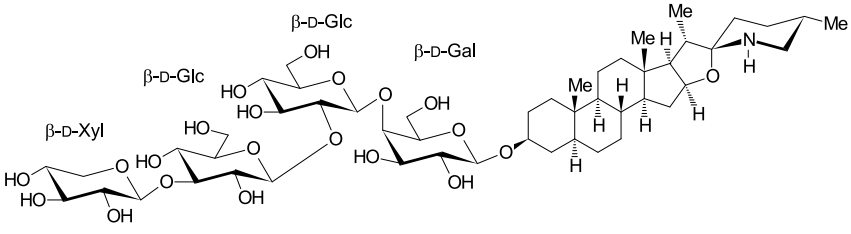
Some studies have revealed that the saccharides lengthen the half-life of the activity by preventing the host from modifying the genin and neutralizing its inhibitory activity [100]. The sugar proximate to the genin has the strangest effect on binding and activity of the drug [101]. Even though the structure of the cardiac glycoside binding site has been extensively studied [102], its detailed nature remains unknown, namely the aminoacid residues which interact with the sugar moiety. Nevertheless, a three-step model has been proposed to account for the interaction between a cardiac glycoside and its receptor. The binding event can proceed via reversible binding of the steroid core with the receptor, followed by a conformational change that exposes a sugar-binding motif, which then binds to the saccharides and hinders dissociation [103].

4.2 Saponins

Originally, the name "saponin" implied a group of soap-like natural surfactants that form long-lasting bubbles on shaking an aqueous solution. An aglycone of saponins is designated as sapogenin or sapogenol (triterpene sapogenins and steroidal sapogenins). Saponins having one sugar moiety are called monodesmosides, and those having two sugar moieties at different positions are called bisdesmosides. Saponins often occur in higher plants, which are frequently used for human and animal consumption, therefore, their physiological activities are of utmost



58 α-solanine



59 α-tomatine

Scheme 24

importance [104]. They can be found in larger amounts in, e. g., soybeans, peas, beans, oats, potatoes, tomatoes, tea, liquorice, ginseng, and in forage as, e. g., alfalfa, lupine, sunflowers, guar etc. Some of the saponins contain nitrogen and these are sometimes included in the alkaloids, e. g., solasodine, α-solanine (**58**) or α-tomatine (**59**) (› *Scheme 24*) and they are rather toxic acting mostly as acetylcholine esterase inhibitors. Hydrolysis (cleavage of glycoside) of, e. g., α-solanin results in overall loss of its toxicity.

Recently, a series of 20 steroidal glycosides from *Solanum* (e. g., glycosides of diosgenin, solasodine, solasonine, etc.) was examined for their anticarcinogenic activities and the following conclusions were drawn: As regarding the sugar linkage, the glycosides possessing a terminal α-L-rhamnosyl moiety (β-chacotriose) are the most effective; as for the aglycon, the glycosides carrying a spirostanol aglycon showed the strongest activity, and even the aglycon without the sugar had strong activity [105]. Generally, saponins due to their large structural diversity have many biological effects. As a raison d'être of saponins in plants its defensive role against microbial pathogens and some predators is often reported. Many saponins were also tested for their antibiotic and anticancer activities, however, no detailed data for the function of their glycosidic moieties were given, therefore, it is beyond of the scope of this review.

4.2.1 Ginsengosides

Ginseng saponins, e. g., dammarane-type saponins are extracted from the root of *Panax ginseng*. They have been used in various formulations in oriental countries for more than 5000 years especially as a tonic [106]. Ginsengosides are highly glycosylated and their activity often differs depending on the number of glycosyl units attached. They can be interconverted by trimming by glycosidases and this is a way by which some more scarce ginsengosides are produced [107].

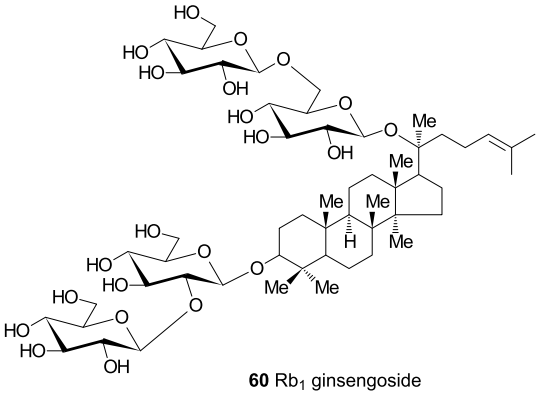
Ginsengoside Rb_1 (**60**) is a representative of the saponins derived from 20(*S*)-protopanaxadiol. It exhibits central nervous system-depressant and antipsychotic activity, protection against stress ulcer, increase of gastrointestinal motility and weak anti-inflammatory action. Rg_1 (**61**) – the major saponin of 20(*S*)-protopanaxatriol – shows weak CNS-stimulant action, antifatigue activity and blood pressure activity [108,109] (› *Scheme 25*).

Some other saponins have been reported to have potentiation activity on the nerve growth factor [110], stimulation of the pituitary-adrenocortical system [111], etc.

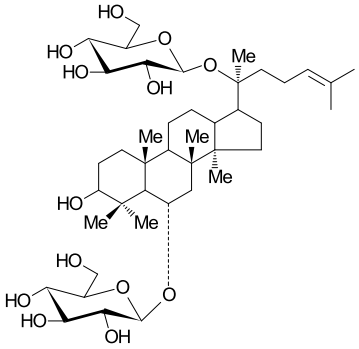
4.3 "Sweet Glycosides"

4.3.1 Osladin

Osladin (**62**) is a steroidal glycoside that is about 500-times sweeter than sucrose. It was isolated by the Czech chemists Jizba and Herout in 1967 [112] from the rhizomes of European fern *Polypodium vulgare* known for its very sweet taste. Its structure has been recently revised [113] by total synthesis. During the synthesis it was shown that minute changes in the structure result in total loss of the sweet taste. Thus, this is a typical glycoside whose overall structure – including the glycosidic part – is crucial for the respective activity.



60 Rb_1 ginsengoside



61 Rg_1 ginsengoside

Scheme 25

4.3.2 Steviol Glycosides

Leaves of *Stevia rebaudiana* (Compositae) are a source of several sweet glycosides of steviol (**63**) [114]. The major glycoside, stevioside (**64**), is used in oriental countries as a food sweetener and the second major glycoside named rebaudioside (**65**), which is sweeter and more delicious than stevioside, is utilized in beverages.

Stevioside (**64**) and rubusoside (**66**) taste somehow bitter, and show aftertaste. To improve the sweetness and the taste, modifications of sugar moieties of both the glycosides were performed by enzymatic glycosylations and/or enzymatic trimming. Cyclomaltodextrin-glucanotransferase (CGTase) efficiently catalyzes transfer of the α-glucosyl moiety (one or more) from starch onto the 4-OH of a glucosyl moiety. Stevioside was treated with this system and it resulted in a complex mixture of mono-, di-, tri-, and polyglucosylated derivatives on both existing glucose moieties of the original compound. Significant improvements of the quality of

taste were observed for most of the glucosylated products, especially for S1a and S2a, which were mono- and di- glucosylated at the 13-sophorosyl moiety of stevioside [115]. A remarkable enhancement of the intensity of sweetness was also observed for both these products, while glucosylation at the 19-*O*-glucosyl moiety (ester bound) resulted in a decrease of the intensity of sweetness (◉ *Table 1*).

Rubusoside was also transglucosylated by the same enzyme system and a large number of products were obtained. Strong enhancement of the sweetness intensity was observed for the products, which were di- or tri-glucosylated at the 13-*O*-glucosyl moiety. Tetraglucosylation at the 13-*O*-glucosyl moiety as well as glucosylation at the 19-*O*-glucosyl moiety led to a decrease in sweetness [116]. Therefore, after transglucosylation the products are "trimmed" enzymatically using β-amylase which releases maltose from the nonreducing end of the α-glucoside. By this treatment maltotriose- and higher α1-4-glucosyl chains were trimmed into monoglucosides and maltosides, which resulted in further improvement of sweetness. The commercial product of transglucosylated stevioside always undergoes this treatment [117].

Replacement of the 19-*O*-glucosyl group by a β-galactosyl group led to worsening of the taste [118]. Significant improvement of the taste quality was achieved by enzymatic transfructosylation at the 19-*O*-Glc moiety of both stevioside and rubusoside [119]. The β-fructofuranosyl moiety is, however, unstable and prone to hydrolysis.

This case is a practical demonstration of the advantage of a glycosidic moiety modification for optimization of physiological properties of the compound(s).

4.3.3 Glycosides of Glycyrrhetic Acid

The major sweet principle of licorice root (*Glycyrrhiza glabra*), glycyrrhizin (**67**) (◉ *Scheme 28*) (content ~ 4%), has been used as a sweetener and flavor enhancer in foods, tobacco products and also in medicine as an anti-inflammatory agent.

A structure-sweetness relationship study has shown that the monoglucuronide of glycyrrhetic acid is about 5-times sweeter than **67** and ca. 1000-times sweeter than sucrose [120]. Replacement of the second glucuronic acid by, e. g., xylose or glucose resulted also in an improvement of the taste and enhancement of sweetness [120].

The monoglucuronide of **67** is now commercially produced using a β-glucuronidase from yeast *Cryptococcus magnus* MG-27 that cleaves selectively only one GlcA moiety [121]. This compound has also better pharmacological properties such as, e. g., an inhibitory effect against skin carcinogenesis and pulmonary tumorigenesis (in mice) [122].

4.4 Glycosides as Aroma Precursors

Glycosides serve often as aroma precursors. It is known that many plants after crushing develop strong aroma. Many of these cases can be attributed to the hydrolysis of glycosides as released aglycons are more volatile than the respective glycosides.

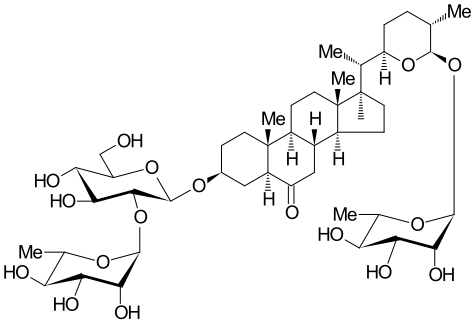
Saliva contains α-glucosidase (ptyalin) and, therefore, experiments with α-glucosides of some aromatic alcohols were performed to enable slow liberation of the fragrances while chewing the food or a chewing gum. However, these approaches have not been very successful up to now.

Table 1

Sweetness quality and intensity of various glucosyl derivatives of stevioside

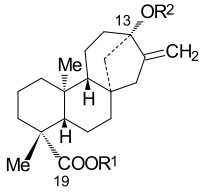
Compound	19-*O*-glycosyl (R1)	13-*O*-glycosyl (R2)	RS	QT
Stevioside (**64**)	-Glc	-Glc-Glc	160	0
Rebaudioside (**65**)	-Glc	-Glc-Glc Glc	210	+2
Rubusoside (**66**)	-Glc	-Glc	134	-2
S1a	-Glc	-Glc-Glc-α-Glc	180	+4
S2a	-Glc	-Glc-Glc-α-Glc-α-Glc	205	+4
S3a	-Glc	-Glc-Glc-α-Glc-α-Glc-α-Glc	117	+3
S1b	-Glc-α-Glc	-Glc-Glc	133	+2
S2b	-Glc-α-Glc	-Glc-Glc-α-Glc	136	+1
S3b	-Glc-α-Glc	-Glc-Glc-α-Glc-α-Glc	146	0
S2c	-Glc-α-Glc-α-Glc	-Glc-Glc	136	0
S3c	-Glc-α-Glc-α-Glc	-Glc-Glc-α-Glc	150	+1
S3d	-Glc-α-Glc-α-Glc-α-Glc	-Glc-Glc	121	+3

RS – relative sweetness to sucrose; QT – quality of taste, stevioside: 0, + better, - worse



62 osladin

Scheme 26



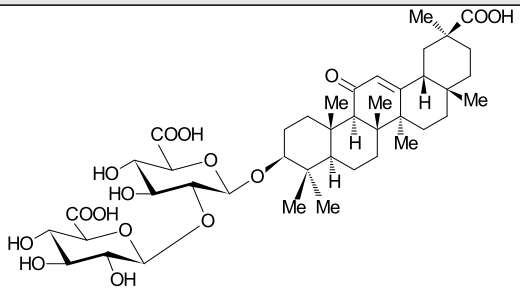
63 steviol R^1 = H, R^2 = H

64 stevioside R^1 = β-D-Glcp, R^2 = Glc(β1→2)Glcβ

65 rebaudioside-A R^1 = β-D-Glcp, R^2 = Glc(β1→2)Glcβ
3
↑
Glcβ1

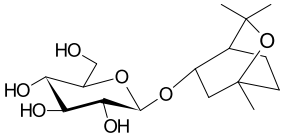
66 Rubusoside R^1 = β-D-Glcp, R^2 = β-D-Glcp

Scheme 27

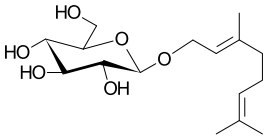


67 glycyrrhizin

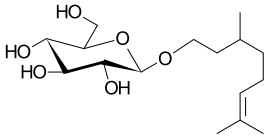
Scheme 28



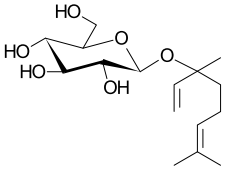
68 1,8-epoxy-*p*-menthan-yl β-D-glucopyranoside



69 geranyl β-D-glucopyranoside



70 citronellyl β-D-glucopyranoside



71 linalyl β-D-glucopyranoside

Scheme 29

Another possibility is to use glycosides of some fragrances to be liberated from tobacco products (by pyrolysis). Since the respective glycosides are not volatile this would limit loss of aroma during storage.

4.4.1 Terpenoid Glycosides from Ginger

A typical example for glycosidic aroma precursors are the terpenoid glycosides discovered recently in fresh young ginger [123].

Linalool and geraniol contribute the most to the fresh ginger aroma. From fresh ginger β-glucopyranosides of 5-hydroxyborneol, 1,8-epoxy-*p*-menthan-3-ol (**68**), 2-heptanol, geraniol (**69**), nerol, citronellol (**70**), (*R*)-linalool (**71**), and some others were isolated. Enzymatic hydrolysis of these compounds with acetone powder prepared from ginger (cold acetone protein precipitate that usually contains intact enzymes from the respective tissue) liberated from the above glycosides the respective fragrance aglycons. This demonstrates that some fragrance components are stored in their glycosidic form and they are released during processing of the ginger.

5 Alkaloid Glycosides

Although alkaloids are mostly produced by plants where glycosylation occurs quite often there are only a few examples of natural alkaloid glycosides, and their biological activities have not been studied to a large extent. There are, however, many examples of alkaloid glycosides prepared artificially for a specific (pharmacological) reason.

5.1 Glycosides of Ergot Alkaloids

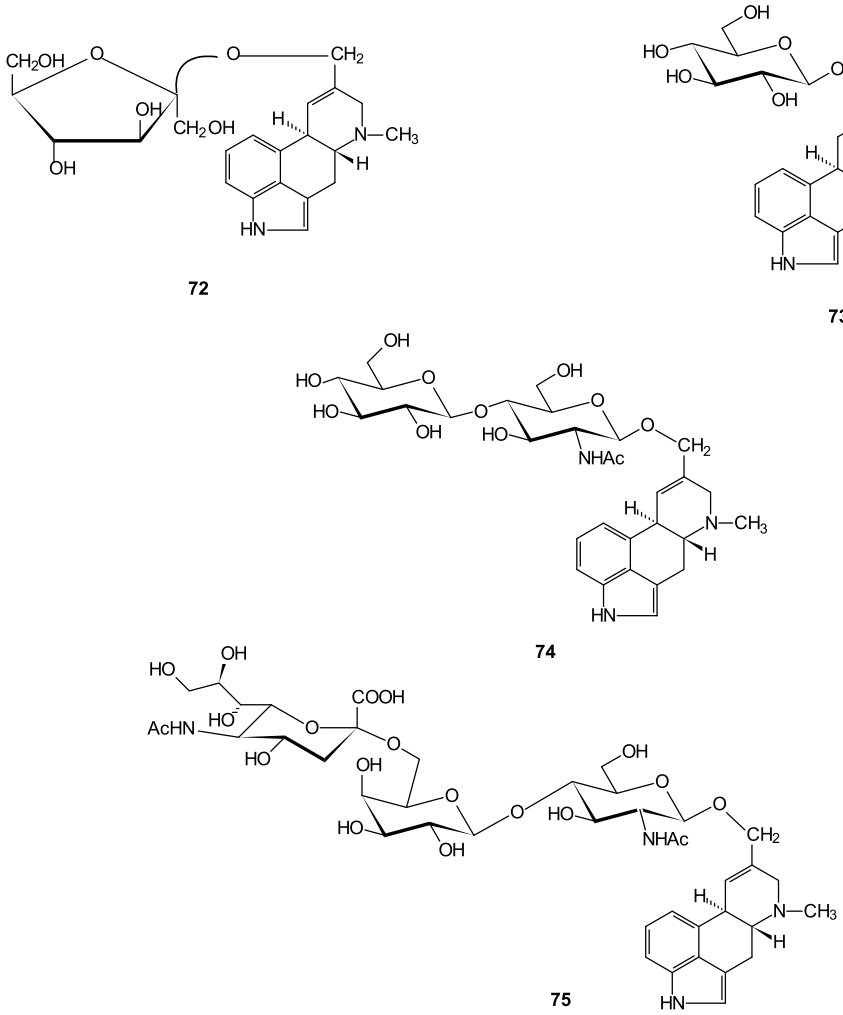
Ergot alkaloids that are produced by the parasitic fungus *Claviceps purpurea* (Ergot) [124] cover a large field of therapeutic uses as drugs of high potency in the treatment of uterine atonia, postpartum bleeding, migraine, orthostatic circulatory disturbances, senile cerebral insufficiency, hypertension, hyperprolactinemia, acromegaly, and parkinsonism [125]. Recently, new therapeutic uses have emerged, such as, for example, against schizophrenia, applications based on newly discovered antibacterial and cytostatic effects, immunomodulatory and hypolipemic activity [126,127,128,129]. Glycosides of EA were isolated as naturally occurring products, and recently a number were prepared by chemical and enzymatic methods. Their promising physiological effects stimulate future research in this field. The first natural EA glycoside, elymoclavine-*O*-β-D-fructofuranoside (**72**) (*Scheme 30*), was isolated from a saprophytic culture of *Claviceps* sp. strain SD-58 by Floss et al. [130]. This glycoside was formed from elymoclavine produced by the microorganism by the action of enzyme invertase present in the fungal mycelium.

Recently, a large series of glycosides of ergot alkaloids was prepared using both chemical and enzymatic methods [131,132,133,134]. Preliminary results obtained indicate that some of these derivatives could have very interesting activities compared to their aglycons.

Some glycosides of 9,10-dihydrolysergol and elymoclavine were tested for their inhibitory activity to prolactin secretion [135,136]. Only 9,10-dihydrolysegol-*O*-β-D-glucopyranoside (**73**) exhibited significantly ($p < 0.001$) higher inhibitory activity.

More systematic studies were performed on the immunomodulatory activity of these new alkaloid glycosides. A large panel of the glycosides, e. g., of elymoclavine and 9,10-dihydrolysergol was tested for their stimulatory activity on cytotoxic lymphocytes. These lymphocytes form the effector arm of cell-mediated immune responses to infection and tumors.

The effect of the alkaloid glycosides was tested on natural killer (NK) cell mediated cytotoxicity of the human resting and activated peripheral blood mononuclear cells (PBMC) against MOLT4 T lymphoma cells (resistant to lysis by fresh PBMC cells and sensitive to activated cells) [128,134,137]. All compounds were tested in a concentration range from 10^{-6} to 10^{-15} M. The maximum immunomodulatory effect was obtained at 10^{-10} M. These effects are mediated by cell surface receptors. After addition of free saccharides to the effector-target cell mixture, the cytotoxic activity of resting fresh lymphocytes has been enhanced in all cases. The most potent stimulation was observed by addition of glucose. The glycosylation of elymoclavine does not influence its cytotoxic activity, however in the case of DH-lysergol its cytotoxic-potentiating activity strongly increased especially in the case of the β-glucoside

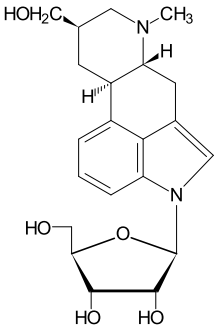


Scheme 30

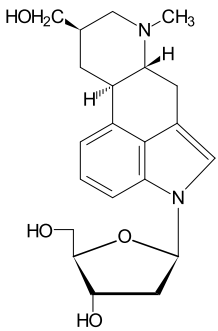
(**73**) and the β-lactoside (**74**). The lytic capacity of activated killer cells was not influenced by any preparation tested.

The effects of elymoclavine and DH-lysergol glycosides were tested also with the cytotoxicity of resting PBMC cells against the NK-sensitive cell tumor line K562 and against the NK-resistant RAJI tumor cell line [128,134,137]. Stimulation of NK cells against the sensitive K 562 was best with the DH-lysergol itself, its glycosylation lowered the stimulatory effects. Galactosylation of elymoclavine potentiated stimulation activity compared to the aglycon.

Interesting results were obtained in stimulation of NK cells against the resistant tumor cell line RAJI. The attachment of β-Glc and mainly Neu5Acα(2→6)Galβ(1→4)GlcNAcβ-*O*-



76



77

Scheme 31

glycosidic moieties to elymoclavine (**75**) had a strong stimulatory effect [134,137]. β-Galactosylation of DH-lysergol also potentiated its effects.

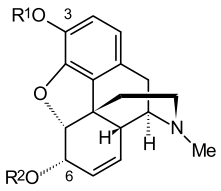
These and some other glycosides were further tested for their immunomodulatory activity on mouse splenocyte models (Balb/c and athymic nude Nu/Nu mice). Here mostly elymoclavine β-galactoside and lactoside had the highest activity [128].

A large panel of ergot alkaloid glycosides and their aglycons was tested for antiviral activities including anti-HIV activity against the replication of HIV-1(III_B) and HIV-2(ROD) in acutely infected MT-4 cells, for their activity in persistently infected HUT-78/III_B cells and for broad spectrum antiviral activity in E_6SM cells cultures against *Herpes simplex* virus-1 (KOS), *Herpes simplex* virus-2 (G), *Vaccinia* virus, *Vesicular stomatitis* virus, *Herpes simplex* virus-1 TK^- B2006 and *Herpes simplex* virus-1 TK^- VMW1837, in HeLa cell cultures against *Vesicular stomatitis* virus, *Coxsackie* virus B4 and *Respiratory syncytial* virus and in Vero cell cultures against *Parainfluenza*-3 virus, *Reovirus*, *Sindbis* virus, *Coxsackie* virus B4 and *Punta Toro* virus [134,138,139]. It was found, however, that virtually all alkaloids and their glycosides tested have cytotoxic concentration below the threshold of their antiviral activity. In some alkaloids, e. g., dihydrolysergol, their cytotoxity was considerably lowered by *N*-ribosylation (**76**) and deoxyribosylation (**77**) (*Scheme 31*). Testing of cytostatic activity of mainly *N*-glycosides of ergot alkaloid (ribosides and deoxyribosides) will be necessary.

Some selected β-galactosides of ergot alkaloids and respective aglycons (elymoclavine and chanoclavine) were tested for their capability to influence basal and forskoline stimulated adenylate-cyclase in ciliary protrusion (*recessus cilliaris*) of the rabbit eye. No significant effects were found. However, both ergot alkaloidgalactosides strongly diminished intraocular pressure in rabbits and their effect was considerably higher than in the respective aglycons [134]. This suggests their potential anti-glaucoma activity.

5.2 Morphine Glucuronides

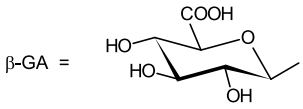
Morphine (**78**) is an important analgesic with a long history of usage. This drug has, however, unwanted side effects. Recently it was found that one glycosidic metabolite – morphine-β-6-



78 (morphine) $R^1 = H$, $R^2 = H$

79 $R^1 = H$, $R^2 = \beta$-GA

80 $R^1 = \beta$-GA, $R^2 = H$



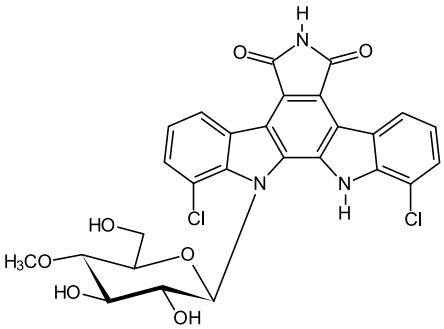
Scheme 32

glucuronide (**79**) – has quite interesting biological effects that are rather different from those of the parental drug.

Morphine has two nucleophilic sites that may be glucuronylated, a 3-hydroxy group on an aromatic ring and an alcoholic 6-hydroxy group.

Conjugation of the 3-hydroxy group occurs in many mammalian species, while glucuronylation at the 6 position appears to be unique to man [140]. This metabolite accumulates in blood of humans with chronic dosing to values greater than morphine [141]. Although the analgesic potency of **79** has been acknowledged for more than two decades, the potential clinical use of this glycoside has only recently been recognized [142]. Morphine β-6-glucuronide (**79**), but not the morphine β-3-glucuronide (**80**) (*Scheme 32*), binds to μ_1 and μ_2 receptors with affinities similar to morphine in mouse brain [140,141]. On the basis of the pharmacodynamic action the 6-glucuronide (**79**) is more than three-times more active than an equimolar subcutaneous dose of morphine in mice [140]. Recent clinical studies in cancer patients given **79** indicated that useful analgesic effects are achieved with an absence of nausea and vomiting [143] that is often caused by morphine itself. Interestingly, **79** present in plasma is distributed into the cerebral spinal fluid, but only one tenth compared to morphine [142]. The potency of morphine β-6-glucuronide is in part explained by an unexpectedly high lipophilicity of the folded form of the molecule. Studies [143] using force-field and quantum mechanical calculations indicate that the glucuronide conjugates of morphine can exist in conformational equilibrium between the extended and folded form of the molecule. The extended conformer, which predominates in aqueous media, is highly hydrophilic because it efficiently exposes polar groups of the molecule to the surrounding aqueous milieu. On the other hand, the folded conformers mask part of these polar groups and are much more lipophilic. The folded forms likely predominate in media of low polarity such as biological membranes [143].

Contrary to its 6-isomer (**79**) morphine-β-3-glucuronide is not an analgesic but it is a potent μ-opioid receptor antagonist. The 3-glucuronide also resembles morphine in that it cau-



81 rebeccamycin

Scheme 33

ses allodynia, a condition in which an ordinarily innocuous stimulus is perceived as painful [144].

5.3 Rebeccamycin

Rebeccamycin **81** (*Scheme 33*) is an interesting alkaloid-type antibiotic that is produced by *Streptomycetes* [145]. Rebeccamycin inhibits the growth of human lung adenocarcinoma cells and produces single strand breaks in the DNA of these cells. A related antibiotic without chlorine, staurosporin, produced by *Streptomyces staurosporeus*, was reported to have antifungal and hypotensive activity [143]. In both structures the 4-*O*-methyl-β-glucopyranosyl moiety forms an *N*-glycosidic bond. A certain analogy with the nucleosides can be traced which may be responsible for its antineoplastic activity.

6 Glycosides of Vitamins

Both hydrophilic and lipophilic glycosides of vitamins often occur in nature and some of them were prepared by chemical and biochemical methods. Recently, an outbreak of papers [146,147,148,149,150,151] was observed dealing with their chemistry and biology, and a comprehensive review on the biological significance of various vitamin glycosides was published [152].

Glycosylated vitamins have an advantage over the respective aglycons in their better solubility in water (especially the lipophilic ones), stability against UV-light, heat and oxidation, reduction in bitter taste and odor (e. g., thiamin), and resistance to enzymatic actions. Some of the vitamin glycoconjugates have altered or improved pharmacokinetic properties. For the synthesis and biosynthesis of vitamin glycosides the presence of a free OH group is mandatory.

We will discuss the vitamin glycosides at length not only because of the quite active research in this area but also for the reason that many vitamin glycosides occur in natural sources as vitamers or they are formed as important and physiologically active metabolites in organism.

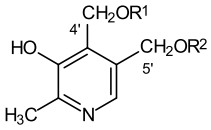
6.1 Water-Soluble Vitamins

6.1.1 Pyridoxine

Extensive research on biological activities has been performed on 5′-*O*-(*β*-D-glucopyranosyl)-pyridoxine **83** (*Scheme 34*), because it is a major form of pyridoxine (**82**, vitamin B_6) in plant-derived foods. Experiments in vivo and in vitro have indicated the lowering of uptake of vitamin B_6 caused by its *β*-glucoside **83** [153,154,155,156]. The presence of pyridoxine *β*-glucoside reduces the bioavailability of pyridoxine usually from plant material by 75–80% [157]. This compound (**83**) is cleaved partly only in the small intestine, but not in the liver [148]. A highly selective pyridoxine-*β*-D-glucosidase was found in the jejunal pig mucosa [149].
Less work was, however, devoted to *α*-glucosides, *β*-glucosides, and *β*-fructosides of pyridoxine, which occur less commonly in nature as a result of transglycosylations.
Recently, it was demonstrated that pyridoxine *α*-glucosides (**84**, **85**) serve nutritionally as well as pyridoxine in terms of transport across everted intestinal sacs, metabolic conversion to active form by liver or kidney homogenate, and the appearance of B_6-derivatives (pyridoxine, pyridoxal, and pyridoxal phosphate) in the blood of B_6-deficient rats after oral administration, while **83** serves very poorly as a B_6 nutrient [157]. Moreover, it has been reported that the 5′-*O*-(*α*-D-glucopyranosyl)-pyridoxine (**84**) does not inhibit the uptake of pyridoxine and is readily converted to pyridoxine in freshly isolated liver cells, while **83** competitively inhibits the uptake of pyridoxine into liver cells [158]. It was found that the 5′-glucoside (**84**) is cleaved in the liver about 5–6-times faster than its regioisomer 4′-*O*-(*α*-D-glucopyranosyl)-pyridoxine (**85**). These facts support the excellent nutritional efficiency of pyridoxine-*α*-glucosides, especially the 5′-isomer. Pyridoxine *α*-glucoside (**84**) has been observed to be much more stable against UV light irradiation and heating than pyridoxine itself [159].
The case of pyridoxine glycosides clearly demonstrates that type of glycosylation strongly modulates the biological activity. In addition, the type of glycosidic linkage (*α*- vs. *β*-) and the site of attachment at quasi-identical carbons also has a strong influence on the final biological activity of the drug.

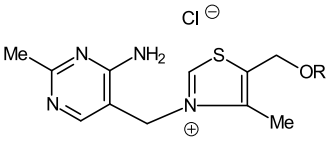
6.1.2 Thiamin

Thiamin (**86**, vitamin B_1) (*Scheme 35*) occurs in nature free and in phosphorylated form. Its glycosides have not been identified in natural material. The artificial glycosylation of



82 (pyridoxine) $R^1 = R^2 = H$
83 $R^1 = H$, $R^2 = \beta$-Glc*p*
84 $R^1 = H$, $R^2 = \alpha$-Glc*p*
85 $R^1 = \alpha$-Glc*p*, $R^2 = H$

Scheme 34



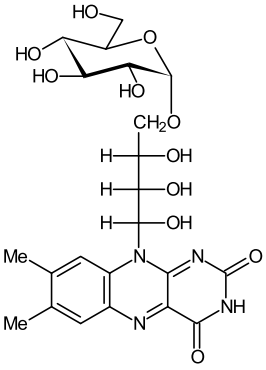
86 (thiamin) R = H
87 R = β-Gal*p*
88 R = α-Glc*p*
89 R^1 = β-Glc*p*NAc

Scheme 35

thiamin was motivated mostly by needs to remove unpleasant taste and odor of the compound and to increase its stability against UV light. *O*-*β*-Galactoside (**87**) [160], *O*-*α*-glucoside (**88**) [161], and *0*-*β*-*N*-acetylglucosaminide (**89**) [162] of thiamin were prepared by enzymatic synthesis using glycosidases. Biological activities were tested only with the *α*-glucoside. This compound, similarly to other glycosides does not have the specific thiamin odor and strong tongue-pricking taste, in contrast it is mildly sweet. *O*-*α*-Glucosylthiamine has 73% activity of a molar equivalent of thiamin hydrochloride, when its use by thiamin-deficient male rats on semisynthetic diet was examined in terms of its effect on the growth curve, food intake, liver weight, and hepatic thiamin content [161]. Pig liver *α*-glucosidase completely hydrolyzed **88** to glucose and thiamin thus some of the thiamin glycosides can be useful as food additives or in medicine.

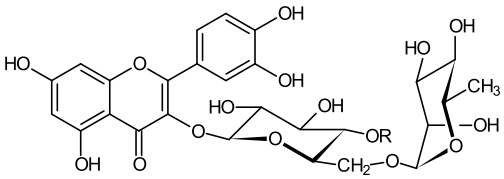
6.1.3 Riboflavin

Riboflavin (vitamin B_2) occurs in nature also in the form of the glycoside, e. g., as riboflavin 5′-*α*-D-glucoside in the liver (**90**) (*Scheme 36*). It is formed from free riboflavin by liver *α*-transglucosylase employing maltose as a donor for an *α*-glucosyl unit [163]. This glycoside is nearly bio-equivalent to riboflavin [164]; it enters the isolated hepatocytes more slowly than free riboflavin and it is hydrolyzed to the free vitamin upon entry.



90

Scheme 36



91 (rutin) R = H
92 R = α-Glcp

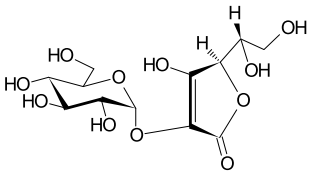
Scheme 37

6.1.4 Rutin

Rutin (**91**, vitamin P) (*Scheme 37*) is one of the most widespread quercetin glycosides (quercetin-3-*O*-rutinoside). Formally, it should be classified as a flavonoid but because of its activities (essential as a capillary protectant) it is included in the vitamin group. Rutin is a glycoside itself, composed of quercetin and the disaccharide rutinose [6-*O*-(α-L-rhamnopyranosyl)-D-glucose]. Its water solubility is rather low (ca. 0.1 g/l) and, therefore, by enzymatic glycosylation 4^{G}-α-D-glucopyranosyl-rutin (**92**) was prepared [165]. Its solubility was considerably increased (30000-times higher than rutin). Its improved biological activity has been predicted based on the finding that this glucoside is hydrolyzed by pig liver α-glucosidase [165].

6.1.5 Ascorbic Acid

Ascorbic acid (vitamin C) is involved in many biological processes, such as collagen synthesis, antioxidation, intestinal absorption of iron, and metabolism of some aminoacids. Search for and preparation of its glycosides was motivated mostly by the need for more stable compounds (resistant to oxidation) having the same or better bioavailability. The first glycoside of ascorbic acid was prepared by Yamamoto [166] and its structure was presumed to be 6-*O*-α-glucopyranosyl-L-ascorbic acid, and it was postulated that this glycoconjugate is formed in vivo by α-glucosidases. Its sensitivity to oxidation was the same as ascorbic acid and its biological activity was inferior to it. Later, another glycoside, 2-*O*-α-glucopyranosyl-L-ascorbic acid (**93**) (*Scheme 38*), was synthesized by an enzymatic regioselective transglucosylation [167]. This glycoside was found to show similar bioavailability as the aglycon in vivo [168] and in vitro [169]. This compound was, however, considerably more stable towards oxidative stress and UV irradiation. In fact, the glycoside **93** itself has no reducing power because of the substitution at the 2-OH group that is involved in redox reactions of ascorbic acid. Oral administration of **93** (guinea pigs) resulted in a remarkable increase of ascorbate in various tissues as well as in plasma [170]. Ascorbic acid was released from **93** at the mucosal side, and it was actively taken up across the intestinal membrane into the serosal side, whereas **93** itself permeated but poorly. Hydrolysis of **93** was mediated by maltase that could be inhibited by castanospermine. Ascorbate was transported by an active uptake system.
Ascorbate, because of its in vivo inhibitory action on melanin synthesis is also used as a skin-whitening agent in cosmetics (e. g., in Japan) [171]. Glycoside **93** was tested in vitro and in



93

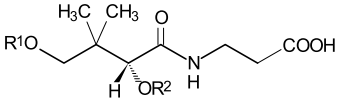
Scheme 38

vivo in humans and compared with ascorbate and ascorbic acid 2-phosphate that is conventionally used for these purposes. It was found that after percutaneous application of **93** the level of ascorbate sustained for a longer period than in the case of phosphate. Also the melanin synthesis (in B18 melanoma cells) was better inhibited by **93** than by ascorbic acid 2-phosphate. From these in vivo and in vitro results it was concluded that by using **93** the level of ascorbate sustains for a longer period. The compound has also UV protective activity against UV-induced damage of human skin keratocytes and fibroblasts.

6.1.6 Pantothenate

Pantothenic acid (**94**) (*Scheme 39*) (denoted as vitamin B_3 or B_5 – sometimes confused with nicotinamide) is widely distributed in animals, plants, and microorganisms as a component of CoA and thus plays an important role in many metabolic pathways. 4′-*O*-(*β*-glucopyranosyl)-D(*R*)-pantothenic acid (**95**) was isolated from tomato juice [172,173] as a growth factor of malo-lactic fermentation bacteria, responsible for the fermentative formation of malic and lactic acids in various food processes, such as for example, wine making. The structure of compound **95** was confirmed by the synthesis and also other glycosides, e. g., 2′-*O*-*β*-glucopyranoside (**96**) and respective glycosides from L and DL forms of **94** were prepared [173]. From comparative studies it follows that the 4′-glucoside was about 50-times more effective than the respective 2′-glucoside or the aglycon (all in D form), the L-form being ineffective. It was suggested that the activity of (**95**) was due to a specific membrane transport of malo-lactic fermentative bacteria. This case also demonstrates the significance of regioisomers of glycosides for their physiological activities.

Following these studies, D-pantothenic acid *β*-glucopyranoside (**95**) was prepared by action of various *β*-glucosidases [174]. Also *α*-glucoside of pantotheic acid was isolated, however, the



94 (D-pantothenic acid) $R^1 = H,\ R^2 = H$
95 $R^1 = \beta\text{-Glc}p,\ R^2 = H$
96 $R^1 = H,\ R^2 = \beta\text{-Glc}p$

Scheme 39

site of the glucosyl attachment was not exactly determined (proposed 4′- position) [175]. Studies of pantothenate metabolism in dogs have demonstrated in vivo glucosylation [176]. Following administration of an oral dose of pantothenate (3 mg/kg), about 60% of pantothenate was excreted in the form of its glucoside **95**.

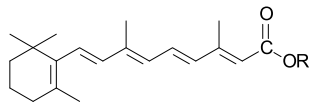
6.1.7 Nicotinamide

Glycosides of this vitamin (sometimes denoted as vitamin B_3) were not described, however, there exists an evidence that bound niacin from wheat bran (termed niacytin) has a single nicotinic acid moiety at least partially linked to an aromatic amine with glucose, xylose, and arabinose in a 6:3:1 molar ratio per molecule, with approximately three cinnamic acid esters [152,177,178]. It seems that these glycosidic complexes limit the bioavailability of the nicotinamide and for its liberation they must be treated, e. g., by soaking corn in a lime solution, traditionally performed in Central America during production of tortillas [179].

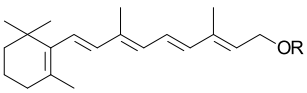
6.2 Lipophilic Vitamins

6.2.1 Retinol

Vitamin A has long been recognized as an essential nutrient in mammals for growth, vision, reproduction, cell differentiation, and the integrity of the immune system [180]. Different metabolic forms of vitamin A show different activities: 11-*cis* retinal is the major ligand for the opsins in vision, 9-*cis* and all-*trans* retinoic acid (**97**, RA) (*Scheme 40*) are active in cell differentiation and in embryogenesis, and retinol and its esters primarily serve as transport and storage forms of the vitamin [180]. In addition to fulfilling their physiological functions, vitamin A and its derivatives and analogs show therapeutic utility in several types of cancer and skin disorders [180]. After the important observation that retinoic acid reduced the onset and number of chemically induced papillomas in mice [181], many analogs, now termed retinoids,



97 (all-*trans*-retinoic acid) R = H
100 (retinoyl β-glucuronide) R = β-glucuronyl



98 (all-*trans*-retinol) R = H
99 (retinol β-glucuronide) R = β-glucuronyl

Scheme 40

were synthesized and tested for their efficacy [182]. The major problem that arose in using retinoids therapeutically was their toxicity [180].

Besides synthetic derivatives useful in therapy there exists a natural derivative of retinol (**98**), and retinoic acid(s) e. g., all-*trans*-retinyl β-glucuronide (**99**) [183] and all-*trans*-retinoyl β-glucuronide (**100**) [184]. These vitamin A glycosides were first identified as billiary metabolites of vitamin A, and they are now prepared also chemically for their favorable biological effects [185,186].

After administration of all-*trans*-, 13-*cis*- or 9-*cis*-retinoic acid, the respective retinoyl glucuronides (RAG) have been identified as the major metabolite in most mammal models including the human system. Like RA, RAG is biologically active in promoting the growth of vitamin A-deficient rats [187] and in the induction of differentiation of, e. g., HL-60 cells [188], but unlike RA, RAG is less cytotoxic and teratogenic. Like RA, RAG is also effective in the topical treatment of human acne [189], but unlike RA, it does not produce any side effects associated with RA therapy [189].

All-*trans*-RAG was investigated (in mice) for its teratogenicity because of its lower placental transfer compared to free RA [190]. Surprisingly, it was found to be a more potent teratogen than RA itself because intravenous or subcutaneous application of RAG was followed by its fast hydrolysis causing high levels of RA. Pharmacokinetic studies nevertheless confirmed lower transplacental transfer of RAG compared to RA [190].

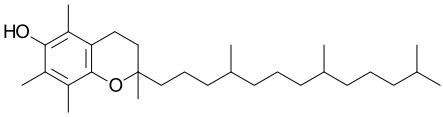
All-*trans*-retinyl β-glucuronide (**99**), another vitamin A glucuronide metabolite was prepared synthetically [185] and it was found to be equally effective in its growth-promoting activity as its aglycon [185,191]. The glucuronide, although converted to retinol in vivo, might also act physiologically in an intact form.

The mechanism of action for retinoid glucuronides is not established [192]. A novel action of the conjugates is suggested, because the glucuronides do not bind to cellular retinoic acid-binding proteins or to nuclear receptors for retinoic acid, although they are active prior to their hydrolytic cleavage.

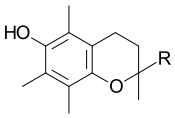
There exist other glycoconjugates of retinoic acid that have potential physiological application. Retinyl phosphate is mannosylated by rat liver microsomes using GDP-mannose in a way analogous to formation of mannophosphoryldolichol, however, the latter reaction is reversible compared to the former one [193]. Probably due to this irreversible reaction retinol is no longer available for its physiological functions in the organism. Besides retinoyl glucuronide, the corresponding galacturonide is formed in the organism at about 10–30% of the rate of the glucuronide formation [194]. Other mostly synthetic glycoconjugates such as retinoyl β-glucose, retinoyl adenosine, and retinoyl sucrose [195] show usually lower activities than RAG [196].

6.2.2 Tocopherol

Tocopherol (vitamin E) (**101**) (*Scheme 41*) is considered to be one of the most prominent antioxidants and radical scavengers in the organism, and it plays also an important role in fertility. It is generally accepted that it is incorporated into biological membranes and is especially effective as a lipid-peroxyl radical scavenger. Glycosylation of tocopherol is motivated mostly by the needs to increase its solubility. Therefore, an attachment of glucosyl unit(s) would be especially useful. Attachment of β-glucosyl, β-maltosyl, and β-oligomaltosyl units via the 6-OH group of tocopherol was achieved [197]. The biological applicability of such a compound



101 α-tocopherol



102 R = COOH
103 R = CH_2OH
104 R = CH_2O-α-Glc*p*

Scheme 41

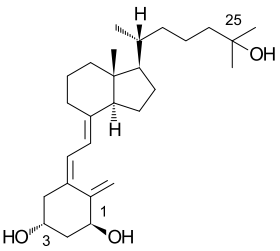
is, however, dependent on its deglycosylation in vivo because the free 6-OH group is essential for the redox properties of tocopherol.

Another probably more feasible strategy is based on the partly water soluble tocopherol derivative Trolox (**102**). The solubility of this compound in water is quite low, therefore enzymatic glycosylation of the corresponding alcohol (**103**) using α-glucosidase from *Saccharomyces* sp. was attempted [198]. The glycosylated product having the structure **104** is well water-soluble (> 1 g/mL) and its radical scavenging activity measured using the 1,1-diphenyl-2-picrylhydrazyl radical is nearly the same as that of α-tocopherol, Trolox, and ascorbic acid. Kinetic studies of the inhibition of the radical chain reaction of methyl linoleate in solution demonstrated that the peroxyl radical scavenging activity was not changed by the replacement of the phytyl side chain of tocopherol to the glucosyl group [199]. Its effectiveness was even higher than that of ascorbic acid when a liposomal suspension was exposed to a lipid-soluble radical generator 2,2′-azobis(2,4-dimethyl-valeronitrile) (AMVN). Also, formation of cholesteryl ester hydroperoxides in human plasma exposed to radical generators was considerably retarded by using compound **104**.

This is probably a good example for a positive influence of glycosylation by transforming a compound active almost solely in the lipid pool into the water pool of the organism. There is, however, one concern – glycosides bearing highly lipophilic moieties can act as nonionic detergents thus causing negative effects like hemolysis, etc. These potential problems must be addressed by thorough in vivo tests.

6.2.3 Calcitriol

Calcitriol (1α,25-dihydroxycholecalciferol, vitamin D_3, **105**) (*Scheme 42*) is involved in calcium metabolism. Glycosides of this compound occur in plants [200,201] and also as metabolites in mammal organisms. β-Glucopyranosides of **105** linked to the C-1, C-3, or C-25 hydroxyfunctions were tested in vivo (rats) and compared with the aglycon. These glycosides occur as potential sources of vitamin D in plants, e. g., *Solanum malacoxylon* [200] where



105

Scheme 42

they cause pathogenic calcifications in cattle grazing on this pasture. It was concluded that all three glycosides were equipotent but they were less active than the parental compound. The least active was the 3-β-glucopyranoside [150]. Their activity is probably a result of their hydrolysis [150,202] after administration *i.v.* or *p.o.*

Biological activity of the three β-glucopyranosides of **105** was studied in chicks and Japanese quails. While the 1- and the 3-glucoside showed no or only little effect on serum calcium, bone weight, calcium binding protein or calcium deposition in the egg shell, the 25-glucoside was found to be more than half as active as the aglycon **105**. The bioactivity of this glucoside parallels a higher binding constant to the intestinal calcitriol receptor compared to those of the two glucosides [203]. β-Glucuronides are the most prominent mammal metabolites of **105** [204].

7 Carbohydrate – Nucleic Acid Interactions

Traditionally, the glycan chains of the DNA-binding antibiotics have been viewed as molecular elements that control the pharmacokinetics of a drug, such as absorption, distribution, metabolism, and excretion. This notion changed a few years ago with the finding that the carbohydrate residues present in the calichenamycin antibiotics partially determine the selectivity of the process and even the calichenamycin aryltetrasaccharide (without possessing the enediyne ring) [39] binds well to the DNA.

Carbohydrate minor groove DNA [205,206] binders are neutral neoglycoconjugates, whose sugar residues are deoxygenated, thus showing a delicate balance between hydrophylic and hydrophobic domains (*Fig. 3*). There is rarely more than a single positively charged ammonium group in the respective carbohydrate and many of the sugars are of the 2,6-deoxyhexopyranose type. Often, the remaining hydroxy groups are further alkylated or acylated. Thus, the recognition process relies on just a few H-bonding interactions and is mostly driven by the hydrophobicity of the DNA-binding saccharides. This suggest that rather the complementary shapes, and hence specific van der Waals contacts between the carbohydrates and minor groove, may determine the site-selectivity displayed by many carbohydrate-containing minor groove binders [207]. Additionally, the relatively rigid carbohydrates may take advantage of the sequence-dependent flexibility of double stranded DNA to bind their target

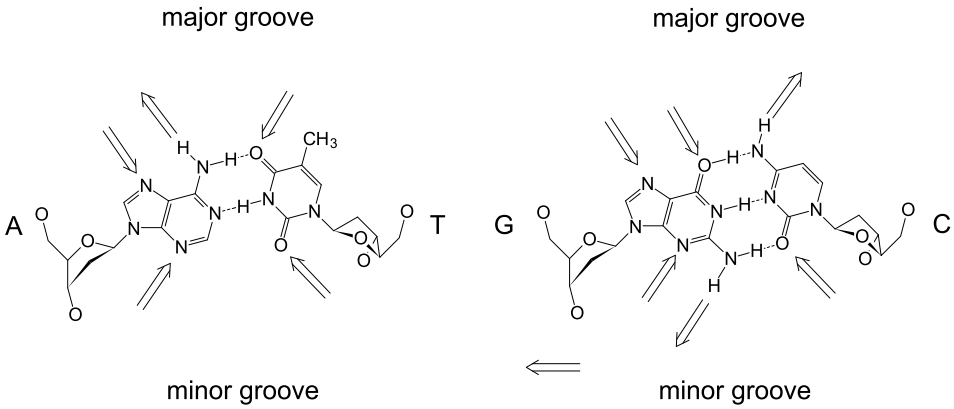
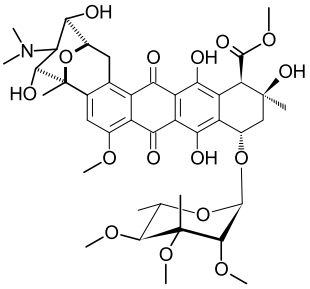


Figure 3
Substitution pattern of H-bonding donor and acceptor groups for adenine-thymine and guanine-cytosine base pairs in the major and minor groove

sites [208]. Furthermore, the hydrophobicity of DNA-binding carbohydrates may differentiate these from the hydrophilic and often negatively charged cell-surface saccharides, which are involved in protein recognition and may enhance the cellular uptake of the respective drugs.
Quite recently evidence has accumulated that some carbohydrates can be also effective DNA major groove binders. So far, there is only a limited number of examples of glycosides that have been found to interact with the DNA major groove [209]: neocarzinostatin (**32**, DNA cleavage agent); altromycin B (alkylating agent; nogalamycin **106** (*Scheme 43*), respinomycin (tandem-intercalative-groove binding ligands); neomycin-Hoechst33258 conjugate (dual-groove binding ligand).
Dual groove binding natural products such as nogalamycin have shown enhanced binding affinity over parent single carbohydrate (single groove binding) compounds (e. g., daunomycin **34**). Therefore, the application of synthetic chemistry and/or glycorandomization can result in novel ligands capable of dual groove recognition and increased binding affinity.



106 nogalamycin

Scheme 43

8 Glycorandomization

The isolation of several sugar biosynthesis gene clusters and glycosyltransferases from different antibiotic-producing organisms, and the increasing knowledge about these biosynthetic pathways opened up the possibility of generating novel bioactive compounds through combinatorial biosynthesis. Recent advances in this area indicate that antibiotic glycosyltransferases show some substrate flexibility that can allow one to alter the types of sugar transferred to the different aglycons or, less frequently, to change the position of its attachment [211,212].

Recently, two complementary glycorandomization strategies have been described, namely, neoglycorandomization, a chemical approach based on a one-step sugar ligation reaction that does not require any prior sugar protection or activation, and chemoenzymatic glycorandomization, a biocatalytic approach that relies on the substrate promiscuity of enzymes to activate and attach sugars to natural products.

These "glycorandomization" approaches (occasionally referred to as "glycodiversification") are expected to foster our understanding of the role of sugars in a variety of glycoconjugates and the exploitation of these critical attachments. Neoglycorandomization that was developed by J. S. Thorson and coworkers relies on a broad variety of reducing monosaccharides without protection/deprotection methodology [212]. Neoglycorandomization is based on the selective formation of glycosidic bonds between reducing sugars and secondary alkoxyamine-containing aglycons [213] to form a library of "neoglycosides".

Sugars and acceptors containing secondary alkoxyamines are reacted with reducing sugars to generate oligosaccharide and glycopeptide mimics. Unlike primary alkoxyamines, which provide open-chain oxime isomers [112] secondary alkoxyamines react to form closed-ring neoglycosides. Presumably, such secondary alkoxyamines react with reducing sugars to form an intermediate oxo-imminium species, which then undergoes ring closure [113].

Chemoenzymatic glycorandomization [215,216] employs the inherent or engineered substrate promiscuity of anomeric kinases and nucleotidylyltransferases, and it attempts to provide activation pathways for the synthesis of nucleotide diphosphosugar donor libraries. In chemoenzymatic glycorandomization these activated sugar libraries, in turn, serve as substrates for inherently promiscuous natural product glycosyltransferases, thereby providing rapid chemoenzymatic means to glycodiversity. Such multienzyme, one-pot reactions offer an attractive alternative to the extensive synthetic manipulation typically required for chemical glycosylation strategies.

In the last three years a variety of examples has appeared supporting this robust methodology: Four glycosyltransferases from two distinct natural product biosynthetic pathways – calicheamicin and vancomycin – readily catalyzed reversible reactions, allowing sugars and aglycons to be exchanged. More than 70 differentially glycosylated calicheamicin and vancomycin variants were prepared. This study [217] suggests the reversibility of glycosyltransferase-catalyzed reactions and may be general and useful for generating exotic nucleotide sugars, establishing in vitro glycosyltransferase activity in complex systems. A monoglycosylated vancomycin library was prepared by the use of flexible glycosyltransferases with nucleotide diphosphosugar libraries [218].

High-level expression of three macrolide glycosyltransferases established a synthetic "tool kit" with such plasticity that 12 modified macrolide antibiotics (oleandromycin and erythromycin) have been readily created [219].

It is quite obvious that more interesting examples will follow bringing this method close to practical exploitation. Here, the basic knowledge of the role of carbohydrates in the biological activity of glycosides can be capitalized upon.

Acknowledgement

The work in the author's laboratory was supported by grants from the Czech Ministry of Education LC06010, OC 170, Czech National Science Foundation 205/05/0172 and the Grant Agency of the Czech Academy of Sciences IAA400200503.

References

1. Williams CA, Harborne JB (1993) Flavone and flavonol glycosides. In: Harborn JB (ed) The flavonoids. Chapman & Hall, London, pp 337–385
2. Harborne JB, Williams CA (1998) Nat Prod Rep 15:631
3. Steglich W, Fugmann B, Lang-Fugmann S (eds) (1997) Naturstoffe - Römpp Lexikon. Georg Thieme Verlag, Stuttgart, p 43
4. Schnitzler JP, Jungblut TP, Heller W, Kofferlein M, Hutzler P, Heinzmann U, Schmelzer E, Ernst D, Langebartels C, Sandermann H (1996) New Phytol 132:247
5. Markham KR, Campos M (1996) Phytochemistry 43:763
6. Vogt T, Wollenweber E, Taylor LP (1995) Phytochemistry 38:580
7. Hallard D, Bleichert A, Gagnon H, Tahara S, Ibrahim R (1992) Z Naturforsch, Teil C 46:725
8. Kudou S, Fleury Y, Welti D, Magnolato D, Uchida T, Kitamura K, Okubo K (1991) Agric Biol Chem 55:2227
9. Graham TL (1991) Plant Physiol 95:594
10. Morris PF, Savard ME, Ward EWB (1991) Physiol Mol Plant Pathol 39:229
11. Tiller SA, Parry AD, Edwards R (1994) Physiol Plant 91:27
12. Graham TL, Kim JE, Graham MY (1990) Mol Plant-Microbe Interact 3:157
13. Křen V, Minghetti A, Sedmera P, Havlíček V, Přikrylová V, Crespi-Perellino N (1998) Phytochemistry 47:217
14. Křen V, Sedmera P, Kubisch J, Halada P, Přikrylová V, Jegorov A, Cvak L, Gebhardt R, Ulrichová J, Šimánek V (1997) J Chem Soc, Perkin Trans 1 1997:2467
15. Middleton E, Kandaswami C (1993) The impact of plant flavonoids on mammalian biology: implications for immunity, inflammation and cancer. In: Harborn JB (ed) The flavonoids. Chapman & Hall, London, pp 337–385
16. Yamada H, Nagai T, Takemoto N, Endoh H, Kiyohara H, Kawamura H, Otsuka Y (1989) Biochem Biophys Res Commun 165: 1292
17. Coronado C, Zuanazzi JAS, Sallaud C, Quirion J-C, Esnault R, Husson H-P, Kondorosi A, Ratet P (1995) Plant Physiol 108:533
18. Dakora F, Joseph C, Phillips D (1993) Plant Physiol 101:819
19. Hamamura Y, Hayashiya K, Naito K-I, Matsuura K, Nishida J (1962) Nature 194:754
20. deBoer G, Hanson FE (1987) Entomol Exp Appl 45:123
21. Ohsugi T, Nishida R, Fukami H (1985) Agric Biol Chem 49:1897
22. Nishida R, Ohsugi T, Kobuko S, Fukami H (1987) Experientia 43:342
23. Feeny P, Sachdev K, Rosenberry L, Carter M (1988) Phytochemistry 27:3439
24. Lynn DG, Chen RH, Manning KS, Wood HN (1987) Proc Natl Acad Sci 84:615
25. Orr JD, Lynn DG (1992) Plant Physiol 98:343
26. Setahell KDR, Lawson AM, Borriello SP, Adlercreutz H, Axelson M (1982) Falk Symp 31 (Colonic carcinogenesis): 93
27. Prasad K (1997) Moll Cell Biochem 168:117
28. Keller-Juslén C, Kuhn M, von Wartburg A, Stähelin H (1971) J Med Chem 14:963
29. Lockie JD, Horwitz SB (1976) Biochemistry 15:5435
30. Stähelin H (1972) Planta Med 22:336
31. Weymouth-Wilson AC (1997) Nat Prod Rep 14:99
32. Nicolaou KC, Dai W-M (1991) Angew Chem Int Ed Engl 30:1387

33. Lee MD, Dunne TS, Chang CC, Ellestad GA, Siegel MM, Morton GO, McGahren WJ, Borders DB (1987) J Am Chem Soc 109:3466
34. Golik J, Clardy J, Dubay G, Groenewold G, Kawaguchi H, Konishi M, Krishnan B, Ohkuma H, Saitoh K, Doyle TW (1987) J Am Chem Soc 109:3461
35. Edo K, Mizugaki M, Koide Y, Seto H, Furihata K, Ohtake N, Ishida N (1985) Tetrahedron Lett 26:331
36. Walker SL, Andreotti AH, Kahne DE (1994) Tetrahedron 50:1351
37. Drak J, Iwasawa N, Danishefsky S, Crothers DM (1991) Proc Natl Acad Sci USA 88:7464
38. Nicolaou KC, Tsay SC, Suzuki T, Joyce GF (1992) J Am Chem Soc 114:7555
39. Ho SN, Boyer SH, Schreiber SL, Danishefsky SJ, Crabtree GR (1994) Proc Natl Acad Sci USA 91:9203
40. Long BH, Golik J, Forenza S, Ward B, Rehfuss R, Onbeowiak JC, Catino JJ, Musial ST, Brookshire KW, Doyle T (1989) Proc Natl Acad Sci USA 86:2
41. Kumar RA, Ikemoto N, Patel DJ (1997) J Mol Biol 265:173
42. Lu L, Golik J, Harrison R, Dedon P (1994) J Am Chem Soc116:9733
43. Stassinopoulos A, Goldberg IH (1995) Bioorg Med Chem 3:713
44. Gao X, Stassinopoulos A, Rice JW, Goldberg IH (1995) Biochemistry 34:40
45. Arcamone F (1984) Med Res Rev 4:153
46. Wang H-J, Ughetto G, Quinley GJ, Rich A (1987) Biochemistry 26:1152
47. Moore MH, Hunter WN, d'Estaintot BL, Kennard O (1989) J Mol Biol 206:693
48. Jizba JV, Sedmera P, Přikrylová V, Vokoun J, Mikulík K, Vanìk Z (1898) Collect Czech Chem Commun 54:1104
49. Arcamone F (ed) (1980) Anticancer agents based on natural product models. Academic Press, New York
50. Casazza MA, De Marco A, Bertazzoli C, Formeli F, Giuliani F, Pratesi F (1978) Curr Chemother, Proc Int Cong Chemother, 10th, p 502
51. Igarashi J, Sunagawa M (1995) Bioorg Med Chem Lett 5:2923
52. Lothstein L, Sweatman TW, Priebe W (1995) Bioorg Med Chem Lett 5:1807
53. Leenders RGG, Schreeren HW, Houba PHJ, Boven E, Haishma HJ (1995) Bioorg Med Chem Lett 5:2975
54. Haysaka T, Inoue Y (1969) Biochemistry 8:2342
55. Kaziro Y, Kamiyama MJ (1967) J Biochem (Tokyo) 62:424
56. Gao X, Patel DJ (1989) Biochemistry 28:751
57. Silva DJ, Kahne DE (1993) J Am Chem Soc 115:7962
58. Sastry M, Patel DJ (1993) Biochemistry 32:6588
59. Greene BM, Brown KR, Taylor HR (1989) Use of Ivermectins in humans. In: Ivermectin and Abamectin, Campbell WC (ed) Springer, Berlin Heidelberg New York, p 311
60. Martin RJ, Pennington AJ (1989) Br J Pharmacol 98:747
61. Mrozik H, Eskola P, Arison BH, Linn BO, Lusi A, Matzuk A, Shih TL, Tischler M, Waksmunski FS, Wyvratt MJ, Blizzard TA, Margiatto GM, Fischer MH, Shoop WL, Egerton JR (1995) Bioorg Med Chem Lett 5:2435
62. Bliard C, Escribano FC, Lukacs G, Olesker A, Sarda P (1987) J Chem Soc Chem Commun 1987:368
63. McGuire JM, Bunch RL, Anderson RC, Boaz HE, Flynn EH, Powell M, Smith JW (1952) Antibiot Chemoter 2:281
64. Flynn EH, Murphy HW, McMahon RE (1955) J Am Chem Soc 77:3104
65. Jones PH, Rowley EK (1968) J Org Chem 33:665
66. Boger D, Colleti SL, Teramoto S, Ramsey TM, Zhou J (1995) Bioorg Med Chem 3:1281
67. Williams JD, Sefton AM (1993) J Antimicrob Chemother 31: Suppl C 11
68. Glaude RP, Bright GM, Isaacson RE, Newborg MF (1989) J Antimicrob Chemother 33:277
69. Kirst HA, Sides GD (1989) Antimicrob Agents Chemother 33:1413
70. Brajtburg J, Powderly WG, Kobayashi GS, Medoff G (1990) Antimicrob Agents Chemother 34:183
71. Hartsel SC, Benz SK, Peterson RP, Whyte BS (1991) Biochemistry 30:77
72. Herve M, Dedouzy JC, Borowski E, Cybulska B, Gary-Bobo CM (1989) Biochim Biophys Acta 980:261
73. Cheron M, Cybulska B, Mazerski J, Grzybowska J, Czerwinski A, Borowski E (1988) Biochem Pharmacol 37:827
74. Stubbe J, Kozarich JW (1987) Chem Rev 87:1107
75. Williams DH (1996) Nat Prod Rep 13:469
76. Perkins HR (1969) Biochem J 111:195
77. Nagarajan R (1993) J Antibiot 46:1181

78. Boger DL, Honda T, Menezes RF, Colletti SL (1994) J Am Chem Soc 116:5631
79. Nicas TI, Mullen DL, Flokowitsch JE, Preston DA, Snyder NJ, Zweifel MJ, Wilkie SC, Rodriquez MJ, Thompson RC, Cooper RDG (1996) Antimicrob Agents Chemother 40: 2194
80. Ge M, Chen Z, Onishi HR, Kohler J, Silver LL, Kerns R, Fukuzawa S, Thompson C, Kahne D (1999) Science 284:507
81. Ge M, Thopson C, Kahne D (1998) J Am Chem Soc 120:11014
82. Thopson C, Ge M, Kahne D (1999) J Am Chem Soc 121:1237
83. Grdadolnik SG, Pristovsek P, Mierke DF (1998) J Med Chem 41:2090
84. Davies J, Davis BD (1968) J Biol Chem 243:3312
85. Shaw KJ, Rather PN, Hare RS, Miller GH (1993) Microbiol Rev 57:138
86. Yoshizawa S, Fourmy D, Puglisi JD (1998) EMBO J 17:6437
87. Davies J (1994) Science 264:375
88. Blaustein MP (1993) Am J Physiol 264:C1367
89. Ahmed K, Rohrer DC, Fullerton DS, Deffo T, Kitatsuji E, From AHL (1983) J Biol Chem 258:8092
90. Fullerton DS, Kihara M, Deffo T, Kitatsuji E, Ahmed K, Simat B, From AHL, Rohrer DC (1984) J Med Chem 27:256
91. Doi SA, Landless PN (1995) Br J Clin Pract 49:257
92. Forbush III B (1983) Top Membr Transp 19:113
93. Rohrer DC, Kihara M, Deffo T, Rathore H, Ahmed K, From AHL, Fullerton DS (1984) J Am Chem Soc 106:8269
94. Thomas RC, Boutagy J, Gelbert A (1974) J Pharm Sci 63:1649
95. Wallick ET, Pitts BJR, Lane LK, Schwartz A (1980) Arch Biochem Biophys 202:442
96. Takechi M, Uno C, Tanaka Y (1998) Biol Pharm Bull 21:1234
97. Takechi M, Uno C, Tanaka Y (1997) Phytochemistry 44:299
98. Takechi M, Tanaka Y (1994) Phytochemistry 37:1421
99. Rathore H, From AHL, Ahmed K, Fullerton DS (1986) J Med Chem 29:1945
100. Repke KRH, Schönfeld W, Weiland J, Megges R, Hache A (1989) In: Sandler M, Smith HJ (eds) Design of enzyme inhibitors as drugs. Oxford University Press, Oxford, p 435
101. Beer J, Kunze R, Herrman I, Portius HJ, Mirsalichova NM, Abubakirov NK, Repke KRH (1988) Biochim Biophys Acta 937:335
102. Kasturi R, Yuan J, McLean LR, Ball WJ Jr (1998) Biochemistry 37:6658
103. Adams RJ, Schwartz A, Grupp G, Grupp I, Lee SW, Wallick ET, Powell T, Twist VW, Gathiram P (1982) Nature 296:167
104. Oakenfull D, Sidhu GS (1986) In: Cheebo PR (ed) Toxicants of plant origin. CRC Press, Boca Raton, FL, pp 98–116
105. Nakamura T, Komori C, Lee Y-Y, Hashimoto F, Yahara S, Nohara T, Ejima A (1996) Biol Pharm Bull 19:546
106. Fulder S (1980) The root of being. Ginseng and the pharmacology of harmony. Hutchinson and Co., Publ. Ltd., London
107. Kohda H, Tanaka O (1975) Yakugaku Zasshi 95:246
108. Takagi K, Saito H, Nabata H (1972) Japan J Pharmacol 22:245
109. Nabata H, Saito H, Takagi K (1973) Japan J Pharmacol 23:29
110. Saito H, Suda K, Schwab M, Thoenen H (1977) Japan J Pharmacol 27:445
111. Hiai S, Yokozawa H, Oura H, Yano S (1979) Endocrinol Japan 26:661
112. Jizba J, Herout V (1967) Collect Czech Chem Commun 32:2867
113. Yamada H, Nishizawa M (1995) J Org Chem 60:386
114. Tanaka O (1997) Pure Appl Chem 69:675
115. Fukunaga Y, Miyata T, Nakayasu N, Mizutani K, Kasai R, Tanaka O (1991) Agric Biol Chem 53:1603
116. Ohtani K, Aikawa Y, Ishikawa H, Kasai R, Kitahata S, Mizutani K, Doi S, Nakamura M, Tanaka O (1991) Agric Biol Chem 55:449
117. Kasai R, Kaneda N, Tanaka O, Yamasaki K, Sakamoto I, Morimoto K, Okada S, Kitahata S, Furukawa H (1981) Nippon Kagakukaishi 1981:726; Chem Abstr (1981) 95:169682
118. Esaki S, Tanaka R, Kamiya S (1984) Agric Biol Chem 48:1831
119. Ishikawa H, Kitahata S, Ohtami K, Ikuhara C, Tanaka O (1990) Agric Biol Chem 54:2043
120. Mizutani K, Kuramoto K, Tamura Y, Ohtake N, Doi S, Nakamura M, Tanaka O (1994) Biosci Biotech Biochem 58:554
121. Kuramoto T, Ito Y, Oda M (1994) Biosci Biotech Biochem 58:455
122. Mizutani K (1994) In: Ho C, Osawa M, Huangand M, Rosen RT (eds) Food phytochemicals

for cancer prevention II, ACS Symposium Series 547. Am Chem Soc, Washington, DC, pp 322–328
123. Sekiwa Y, Mizuno Y, Yamamoto Y, Kubota K, Kobayashi A, Koshino H (1999) Biosci Biotechnol Biochem 63:384
124. Křen V, Cvak L (eds) (1999) Ergot - Genus *Claviceps*, Medicinal & Aromatic Plants - Industrial Profiles. Harwood Publ Ltd, London, pp 1–515
125. Eich E, Pertz H (1994) Pharmazie 42:867
126. Markstein R, Seiler MP, Jaton A, Briner U (1992) Neurochem Int 20:211S
127. Eich E, Eichberg D, Schwarz G, Clas F, Loos M (1985) Arzneim-Forsch/Drug Res 35 (II):1760
128. Fišerová A, Kovářů H, Hajduová Z, Mareš V, Křen V, Flieger M, Pospíšil M (1997) Physiol Res 46:119
129. Cincotta AH, Meier AH (1989) Life Sci 45:2247
130. Floss HG, Günther H, Mothes U, Becker I (1966) Z Naturforsch 22b:399
131. Křen V, Sedmera P, Havlíček V, Fišerová A, Šíma P (1992) Tetrahedron Lett 33:7233
132. Křen V, Ščigelová M, Přikrylová V, Havlíček V, Sedmera P (1994) Biocatalysis 10:181
133. Křen V, Augé C, Sedmera P, Havlíček V (1994) J Chem Soc Perkin Trans 1 1994:2481
134. Křen V (1997) Top Curr Chem 186:45
135. Seifert K, Johne S (1984) Arch Pharm (Weinheim) 317:577
136. Cassady JM, Li GS, Spitzner EB, Floss HG (1974) J Med Chem 17:300
137. Křen V, Fišerová A, Augé C, Sedmera P, Havlíček V, Šíma P (1996) Bioorg Med Chem 4:869
138. Křen V, Pískala A, Sedmera P, Havlíček V, Přikrylová V, Witvrouw M, De Clercq E (1997) Nucleosides Nucleotides 16:97
139. Křen V, Olšovský P, Havlíček V, Sedmera P, Witvrouw M, De Clercq E (1997) Tetrahedron 53:4503
140. Mulder GJ (1992) Trends Pharmacol Sci 13:302
141. Paul D, Standifer KM, Inturrisi CE, Pasternak GW (1989) J Pharmacol Exp Ther 251:477
142. Portenoy RK, Khan E, Layman N, Lapin J, Malkin MG, Foley KM, Thaler HT (1991) Neurology 41:1457
143. Carrupt P-A, Testa B, Bechalany A, El Tayar N, Descas P, Perrissoud D (1991) J Med Chem 34:1272
144. Yaksh TL, Harty GJ (1988) J Pharmacol Exp Ther 244:501
145. Bush JA, Long BH, Catino JJ, Bradner WT (1987) J Antibiotics 40:668
146. Trumbo PR, Banks MA, Gregory III JF (1990) Proc Soc Exp Biol Med 195:240
147. Nakano H, McMahon LG, Gregory III JF (1997) J Nutr 127:1508
148. Nakano H, Gregory III JF (1995) J Nutr 125:2751
149. McMahon LG, Nakano H, Levy M-D, Gregory III JF (1997) J Biol Chem 272:32025
150. Londowski JM, Kost SB, Gross M, Labler L, Meier W, Kumar R (1985) J Pharmacol Exp Ther 234:25
151. Joseph T, McCormick DB (1995) J Nutr 125:2194
152. Gregory III JF (1998) Annu Rev Nutr 18:277
153. Gilbert JA, Gregory III JF (1992) J Nutr 122:1029
154. Kabir H, Leklem J, Miller LT (1983) J Food Sci 1983:1422
155. Nakano H, Gregory III JF (1995) J Nutr 125:926
156. Zhang Z, Gregory III JF, McCormick DB (1993) J Nutr 123:85
157. Tsuge H, Maeno M, Hayakawa T, Suzuki Y (1996) J Nutr Sci Vitaminol 42:377
158. Joseph T, Tsuge H, Suzuki Y, McCormick DB (1996) J Nutr 126:2899
159. Kawai F, Yamada H, Ogata K (1971) J Vitaminol 17:121
160. Suzuki Y, Uchida K (1994) Biosci Biotech Biochem 58:1273
161. Suzuki Y, Uchida K (1998) Biosci Biotech Biochem 62:221
162. Křen V, Huňková Z, Halada P, Suzuki Y (1998) Biosci Biotech Biochem 62:2415
163. Suzuki Y, Uchida K (1980) Meth Enzymol 66:327
164. Joseph T, McCormick DB (1995) J Nutr 125:2194
165. Suzuki Y, Suzuki K (1991) Agric Biol Chem 55:181
166. Yamamoto I, Muto N, Nagata E, Nakamura T, Suzuki Y (1990) Biochim Biophys Acta 1035:44
167. Aga H, Yoneyama M, Sakai S, Yamamoto I (1990) Agric Biol Chem 55:1751
168. Yamamoto I, Suga S, Mitoh Y, Tanaka M (1990) J Pharmacobio-Dyn 13:688
169. Yamamoto I, Muto N, Murakami K, Akyiama J (1992) J Nutr 122:871
170. Muto N, Terasawa K, Yamamoto I (1992) Internat J Vit Nutr Res 62:318
171. Kumano Y, Sakamoto T, Egawa M, Iwai I, Tanaka M, Yamamoto I (1998) J Nutr Sci Vitaminol 44:345

172. Amachi T, Imamoto S, Yoshizumi H, Senoh S (1970) Tetrahedron Lett 1970:4871
173. Amachi T, Imamoto S, Yoshizumi H (1971) Agric Biol Chem 35:1222
174. Kawai F, Yamada H, Ogata K (1974) Agric Biol Chem 38:831
175. Kawai F, Maezato K, Yamada H, Ogata K (1972) Biochim Biophys Acta 286:91
176. Nakano K, Sugawara Y, Ohashi M, Harigaya S (1986) Biochem Pharmacol 35:3745
177. Mason JB, Gibson M, Kodicek E (1973) Br J Nutr 30:297
178. Wall JS, Carpenter KJ (1988) Food Technol 42:198
179. van den Berg H (1997) Eur J Clin Nutr 51(Suppl):S64
180. Sporn MB, Roberts AB, Goodman DS (eds) (1994) The retinoids: Biology, chemistry, and biochemistry, 2nd ed. Raven Press, New York, pp 1–679
181. Bollag W, Matter A (1981) Ann N Y Acad Sci 359:9
182. Davson MI, Hobbs PD (1994) The synthetic chemistry of retinoids. In: Sporn MB, Roberts AB, Goodman DS (eds) The retinoids: Biology, chemistry, and biochemistry, 2nd ed. Raven Press, New York, pp 5–178
183. Zachman RD, Dunagin PE, Olson JA (1966) J Lipid Res 7:3
184. Unagin PE, Meadows EH, Olson JA (1965) Science 148:86
185. Barua AB, Olson JA (1987) Biochem J 244:231
186. Becker B, Barua AB, Olson JA (1996) Biochem J 314:249
187. Nath K, Olson JA (1967) J Nutr 93:461
188. Gallup JM, Barua AB, Furr HC, Olson JA (1987) Proc Soc Exp Biol Med 186:269
189. Gunning DB, Barua AB, Lloyd RA, Olson JA (1994) J Dermatol Treat 5:181
190. Nau H, Elmazar MM, Ruhl R, Thiel R, Sass JO (1996) Teratology 54:150
191. Barua AB, Batres RO, Olson JA (1988) Biochem J 252:415
192. Kauffman FC, Zaleski J, Thurman RG, Kwei GY (1994) Biologically active conjugates of drugs and toxic chemicals. In: Kauffman FC (ed) Conjugation-deconjugation reactions in drug metabolism and toxicity. Springer, Berlin Heidelberg New York
193. Frot-Coutaz J, Letoublon R, Degiuli A, Fayet Y, Audigier-Petit C, Got R (1985) Biochim Biophys Acta 841:299
194. Genchi G, Wang W, Barua AB, Bidlack WR, Olson JA (1996) Biochim Biophys Acta 1289:284
195. Planer WS, Olson JA (1994) Retinol and retinoic acid metabolism. In: Sporn MB, Roberts AB, Goodman DS (eds) The Retinoids: Biology, Chemistry, and Biochemistry, 2nd ed. Raven Press, New York, pp 229–255
196. Formelli F, Barua AB, Olson JA (1996) FASEB J 10:1014
197. Lahmann M, Thiem J (1997) Carbohydr Res 299:23
198. Murase H, Yamauchi R, Kato K, Kunieda T, Terao J (1997) Lipids 32:73
199. Murase H, Moon JH, Yamauchi R, Kato K, Kunieda T, Yoshikawa T, Terao J (1998) Free Radic Biol Med 24:217
200. Wasserman RH, Henion JD, Haussler MR, McCain TA (1976) Science 194:853
201. Hughes MR, McCain TA, Chang SY, Haussler MR, Villareale M, Wasserman RH (1977) Nature 268:347
202. Londowski JM, Kost SB, Meier W, Labler L, Kumar R (1986) J Pharmacol Exp Ther 237: 837
203. Rambeck WA, Weiser H, Meier W, Labler L, Zucker H (1985) Int J Vitam Nutr Res 55:263
204. Shimada K, Nakatani I, Saito K, Mitamura K (1996) Biol Pharm Bull 19:491
205. Shimada K, Kamezawa Y, Mitamura K (1997) Biol Pharm Bull 20:596
206. Dervan PB (2001) Bioorg Med Chem 9:2215
207. Hunziker J (1996) Chimia 50:248
208. Marin JN, Muñoz EM, Schwergold C, Souard F, Asensio JL, Jiménez-Barbero J, Cañada J, Vicent C (2005) J Am Chem Soc 127: 9518
209. Willis B, Arya DP (2006) Curr Org Chem 10:663
210. Ikeda D, Umezawa S (1999) In: Ikan R (ed) Naturally occurring glycosides. Wiley, New York, pp 1–42
211. Mendez C, Salas JA (2001) Trends Biotech 19:449
212. Langenhan JM, Griffith BR, Thorson JS (2005) J Nat Prod 68:1696
213. Van Vranken D L, Chisolm JD (2000) J Org Chem 65:7541
214. Peri F, Nicotra F (2004) Chem Commun 2004:623
215. Griffith BR, Langenhan JM, Thorson JS (2005) Curr Opin Biotechnol 16:622

216. Daines AM, Maltman BA, Flitsch SL (2004) Curr Opin Chem Biol 8:106
217. Zhang C, Griffith BR, Fu Q, Albermann C, Fu X, Lee I-K, Li L, Thorson JS (2006) Science 313, 1291
218. Fu X, Albermann C, Jiang J, Liao J, Zhang C, Thorson JS (2003) Nature Biotechnol 21:1467
219. Yang M, Proctor MR, Bolam DN, Errey JC, Field RA, Gilbert HJ, Davis BG (2005) J Am Chem Soc 127:9336

12.8 Mucin-Based Vaccines

Jonathan P. Richardson, Derek Macmillan
Department of Chemistry, University College London, 20 Gordon Street, WC1H 0AJ London, UK
d.macmillan@ucl.ac.uk

Abstract

Mucins are heavily *O*-glycosylated cell surface and secreted glycoproteins. In addition to orchestrating cell-extracellular matrix and cell-cell interactions in healthy organisms mucins are also the major carriers of altered glycosylation in carcinomas. Tumor-associated antigens displayed by cancer cells comprise oligosaccharide and glycopeptide motifs not encountered in the same locale or at the same frequency in healthy cells, and potentially confer a selective advantage to the tumor. Frequently tumor-associated antigens are under-glycosylated and prematurely sialylated, and it is these relatively simple saccharide and glycopeptide structures

In: *Glycoscience*. Fraser-Reid B, Tatsuta K, Thiem J (eds)
Chapter-DOI 10-1007/978-3-540-30429-6_68:

that have been targeted to serve as drug candidates in most cases. A major goal is to assemble glycopeptide vaccine candidates based on partial mucin sequences and displaying tumor-associated antigens that can mount a potent immunological tumor-specific response when, in reality, the tumor has already coerced the immune system into a state of co-existence.

Keywords

Mucins; Vaccines; Cancer; Carbohydrates; Antigens; Glycosylation; Glycopeptides

Abbreviations and Symbols

Ac	acetyl
ADCC	antibody dependent cell-mediated cytotoxicity
APC	antigen presenting cell
Bn	benzyl
BCR	B-cell receptor
BSA	bovine serum albumin
CAN	cerium (IV) ammonium nitrate
CDC	complement dependent cytotoxicity
CTL	$CD8^+$ cytotoxic T-lymphocytes
DBU	1,8-diazabicyclo[5.4.0]undec-7-ene
DC	dendritic cell
DNA	deoxyribonucleic acid
EGF	epidermal growth factor
ELISA	enzyme-linked immunosorbent assay
ER	endoplasmic reticulum
Et	ethyl
Fmoc	9-fluorenylmethoxycarbonyl
GalNAc	*N*-acetyl galactosamine
GlcNAc	*N*-acetyl glucosamine
HPLC	high pressure liquid chromatography
IFN	interferon
Ig	immunoglobulin
IL	interleukin
KLH	keyhole limpet hemocyanin
MALDI-TOF MS	matrix-assisted laser desorption-time of flight mass spectrometry
MAPK	mitogen-activated protein kinase
MDa	MegaDalton
Me	methyl
MHC	major histocompatibility complex
MUC	mucin
NeuAc	*N*-acetyl neuraminic acid (sialic acid)
NMR	nuclear magnetic resonance
NOE	nuclear Overhauser effect
Ph	phenyl
ppGalNAcT	UDP-GalNAc-polypeptide-α-*O*-GalNAc transferase

PSA	prostate-specific antigen
RNA	ribonucleic acid
Su	succinimide
TAA	tumor-associated antigen
***t*-Bu**	tertiary butyl
TCR	T-cell receptor
TGF	transforming growth factor
$\mathbf{T}_H$	$CD4^+$ T-helper cell
TIPS	triisopropylsilyl
TSE	2-trimethylsilyl ethyl
UDP	uridine diphosphate
VNTR	variable number of tandem repeats

1 Introduction

Mucins are cell surface and secreted glycoproteins characterized by high levels of *O*-linked glycosylation, often over 50% carbohydrate by weight, which is generally densely clustered in repeating units of 20 or so amino acid residues (termed tandem repeats) of protein sequence rich in serine, threonine and proline [1]. The occurrence of tandem repeats is characteristic of "true" mucins and sets them apart from other heavily *O*-glycosylated proteins involved in, for example, lymphocyte homing and inflammation. The number of tandem repeats in any particular mucin expressed by the cells from different individuals is highly variable (a phenomenon termed polymorphism) hence the abbreviation VNTR: variable number of tandem repeats, that is often used to describe these repeating domains. Mucin molecules can also be cross-linked via intermolecular disulfide linkages between cysteine residues which results in a highly heterogeneous and hydrophilic matrix with molecular weights in excess of 1 MDa. Mucin molecules tend to have highly extended structures resulting from both intrinsic conformational preferences and intramolecular (hydrogen bonding) interactions typical of the α-*O*-glycopeptide linkage, and repulsive interactions between negative charges on densely clustered sialic acid residues. The highly charged and highly hydrated proteins form a viscous gel-like protective barrier between epithelial cells (other than skin cells that are water impermeable) and the external environment. Impressively, mucins which line the gastrointestinal tract allow endothelial cells to withstand the strongly acidic (pH 2) environment of the stomach and maintain a pH of 6–7 near the cell surface by binding bicarbonate secreted by gastric cells. In addition to the gastrointestinal tract, mucins are also secreted by epithelial cells of the tracheobronchial, respiratory, ocular and reproductive tracts, and serve to maintain tissue homeostasis, to trap and export invading pathogenic organisms as in nasal secretions as well as being implicated in numerous adhesion and signaling processes since, when bound to the cell surface, mucins can also mediate cell-cell interactions [2,3]. In this short review we aim to introduce the structure and occurrence of mucins and explain how their molecular characteristics influence their macroscopic physical properties and proposed biological activities. We will go on to describe their relevance to tumor development and metastasis and explain why mucins in particular have been targeted for vaccine development, primarily in the context of cancer treatment. In order to put the biological results in context we will describe the basic aspects of immunology

relevant to the design of mucin related anti-cancer vaccines and the general considerations for vaccine development. We will finally describe how anti-cancer vaccine candidates are prepared in the laboratory and summarize biological findings with such molecules.

2 Mucins: Structure, Function and Localization

Mucins are heavily *O*-glycosylated proteins. As the discovery of the number of ways in which proteins are modified by carbohydrates has increased, as a result of detailed glycobiological investigations, *O*-linked glycosylation has come to mean many things [4] but in the context of mucins it is specifically characterized by an α-*O*-glycosidic linkage between *N*-acetylgalactosamine (GalNAc) at the reducing end of the saccharide moiety and the hydroxyl groups of threonine and serine amino acid residues within the protein sequence [5]. The carbohydrate-rich VNTR domains in the wild-type mucin often contain densely clustered, short, branched oligosaccharide chains and the occurrence and role of glycosylation will be discussed in more detail in *Sect. 3* Mucins, for which 20 genes have been identified thus far [6], are often divided into two major classes: those that are membrane-associated (e. g. MUC1, MUC3, MUC4, MUC12) by virtue of a hydrophobic *C*-terminal transmembrane domain, and those which are constitutively secreted, for example MUC2, by specialized cells. The membrane associated mucins, though different in size, polymorphisms and sequence share several features (*Fig. 1*) [3,7].

Most membrane-bound mucins are proteolytically processed shortly after translation and MUC1 is translocated to the cell surface as a stable heterodimer (originating from a single encoded polypeptide) resulting in an *N*-terminal domain comprised mainly of the VNTR domain. The *C*-terminal domain possesses a hydrophobic transmembrane domain and a highly conserved short *C*-terminal cytosolic "tail" that can be phosphorylated on serine and tyrosine and is known to interact with elements of the cytoskeleton and β-catenin which is a mediator of the MAPK signaling pathway. Indeed activation of mucins, stimulated by ligand binding or a change in extracellular environment, with ensuing proteolysis of the MUC1 backbone and release of soluble mucin is accompanied with *C*-terminus internalization in complex with β-catenin suggesting transcription factor activity [2,8,9]. While it is feasible that the "normal" beneficial role for this function (association with intracellular signaling molecules) may be to protect epithelial cells from stress-induced apoptosis, this process has also been implicated in the hosts tolerance to cellular transformation (and chemotherapeutics) by over-expression of "oncoprotein" MUC1, allowing it to sequester various signaling proteins (such as c-Abl, c-Src, and EGF1 that interact with the cytosolic tail of MUC1 via their SH2 domains) at the cell membrane when they might otherwise initiate apoptosis [10]. In addition, membrane-bound mucins have extracellular EGF-like domains that are also implicated in signaling [11,12]. In this case association of the proteolytically released intramembrane domain of e. g. MUC4 (one of the best studied examples) exposes the EGF domain facilitating association with the integral membrane protein ErbB2 and initiates phosphorylation of ErbB2 which is involved in intracellular signaling through the MAPK pathway [2,11,13]. The roles for mucin-Erb interactions are largely speculative and poorly understood at the molecular level [2,3]. Secreted mucins, on the other hand, do not contain the *C*-terminal tail, or transmembrane domain and contain fewer EGF domains but instead (with the exception of MUC7) possess C and

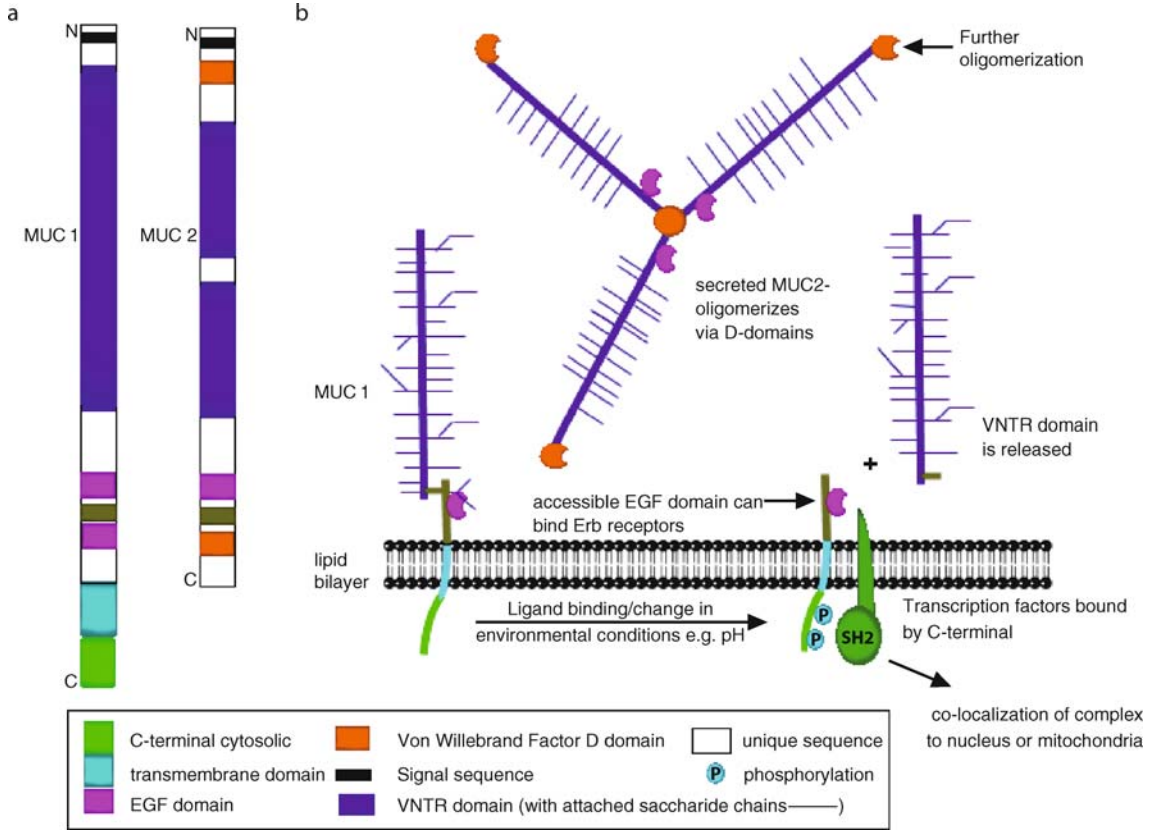


Figure 1

(a) Domain structures (*left*) of membrane-associated (MUC1) and secreted (MUC2) polypeptide sequences (before posttranslational modification and processing). (b) After processing (*right*), suggesting some possible functions and key interactions

N-terminal Von Willebrand factor D domains that are cysteine rich and facilitate intermolecular covalent cross-linking via disulfide bond formation. Apomucins are glycosylated as they traverse the Golgi apparatus and are localized exclusively at the apical cell surface by virtue of their transmembrane domain. Membrane associated mucins can also be secreted as a result of proteolytic processing of the cell surface protein or RNA splice variation that encodes the secreted protein.

3 Mucin Glycosylation

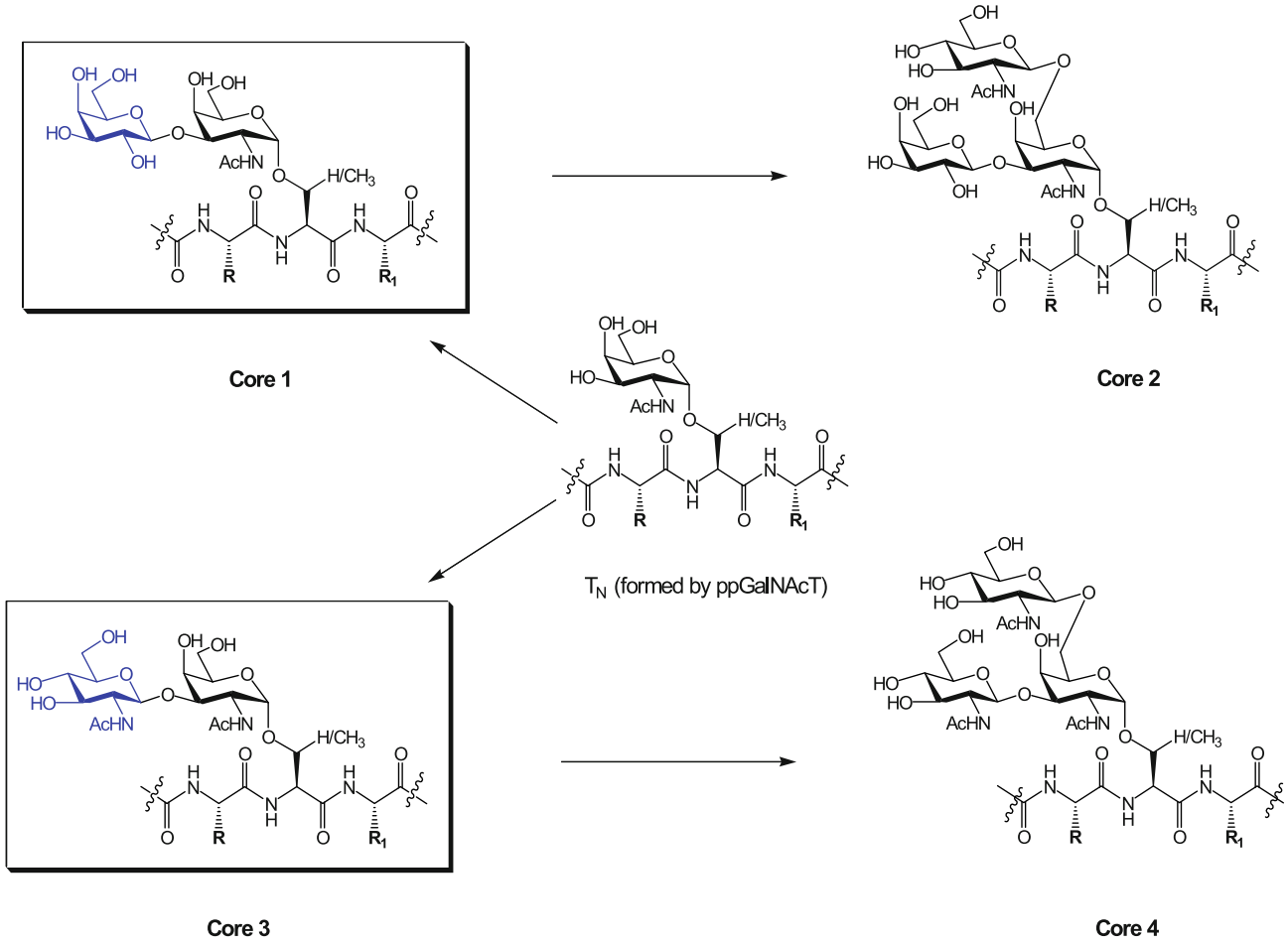
The repeating Ser/Thr/Pro rich sequence which is, in part, characteristic of each mucin is repeated from 10 to over 100 times (between 20 and 120 for MUC1) and although the num-

Table 1
Occurrence of some mucin species and their VNTR domains [1]

Mucin species	(s)ecreted, (m)embrane bound	D-domains	Number of tandem repeats in VNTR domain	Tandem repeats Tandem repeat sequence (number of residues in tandem repeat)
MUC1	m, s	–	pm	AHGVTSA**PDTRP**APGSTAPP (20)
MUC2	s	+	pm 16	PTTTPITTTTTVTPTPPTPTGTQT (23) PPTTTPSPPTTTTTTP (16)
MUC3	m, s	–	pm	HSTPSFTSSITTTETTS (17)
MUC4	m, s	–	pm	TSSASTGHATPLPVTD (16)
MUC5AC	s	+	pm	TTSAPTTS (8)
MUC5B	s	+	11	Irregular repeats of 29 aa
MUC6	s	+	pm	169 aa repeat
MUC7	s	– (non-gel forming)	6	TTAAPPTPSATTPAPPSSSAPPE (23)
MUC8	?	?	?	TSCPRPLQEGTRV (13) and TSCPRPLQEGTPGSRAAHALSR-RGHRVHELPTSSPGGDTGF (41)
MUC9	?	?	9	GEKTLTPVGHQSVTP (15)
MUC11	?	?	?	SGLSEESTTSHSSPGSTHTTLSPASTTT (28)

Pm = polymorphic

Figure 2
Modification of a peptide by ppGalNAcT glycosylates the peptide via an α-*O*-glycosidic linkage to serine and threonine (forming what is known as the T_N antigen). In healthy cells this structure is most commonly modified by core 1–core 4 forming glycosyltransferases. The core structure which predominates in any particular mucin depends on substrate and glycosyltransferase expression patterns but is generally core 2 in mucins with core 3 and 4 predominating in the gastrointestinal tract


Core 1
Core 2
T_N (formed by ppGalNAcT)
Core 3
Core 4

ber of repetitions varies from one individual to another (i. e. is polymorphic), the identity of the amino acids in tandem repeating sequences (in the VNTR domain) in typical mucins is generally known (◉ *Table 1*) [1].

The VNTR domain of a mucin can comprise 50–80% of the whole encoded protein sequence. The bulk of *O*-linked glycosylation occurs in these peptide sequences rich in serine, threonine and proline amino acid residues. The presence of proline appears important for efficient glycosylation by the enzyme UDP-GalNAc-polypeptide-α-*O*-GalNAc transferase (ppGalNAcT). The factors that influence whether a specific serine or threonine residue is glycosylated are poorly understood when compared to *N*-linked glycosylation but the recent disclosure of the crystal structure of ppGalNAcT, that adds the first *N*-acetylgalactosamine (GalNAc) residue of *O*-linked saccharides may yet provide insight into the origin of the apparent sequence context rules [14]. Peptides which are decorated with only this first sugar (GalNAc) give rise to what is known as the T_N antigen. Understanding the requirements for *O*-linked glycosylation is further complicated by the fact that at least fifteen ppGalNAcT enzymes exist, expression of which can be restricted to specific cell types and have been shown to exhibit different substrate specificities. ppGalNAcT I and ppGalNAcT II glycosylate "naked" (apomucin) peptide sequences whereas ppGalNAcT IV requires that the peptide is already glycosylated by ppGalNAcT I [15,16]. Despite the lack of a clear consensus sequence for *O*-linked glycosylation, the probability of serine and threonine residues being glycosylated can be predicted with some accuracy and Ser/Thr residues with proline located at the −1 and/or +3 positions are particularly favorable [17]. The T_N antigen is normally elongated to form one of the eight distinct "core structures" [5]. Core 1 and particularly core 2 are by far the most abundant core structures on mucins (◉ *Fig. 2*) in healthy cells though the prevailing core type is essentially the overall outcome of a competition between several glycosyltransferases for the same substrate. Core 3 and core 4 structures are also observed, though far less frequently, on mucins except in the intestinal tract [18]. Depending on the tissue type, the cores are often terminated by the addition of only a few saccharides, the blood group determinants for example, but usually with *N*-acetylglucsoamine, galactose, sialic acid, fucose and additionally with sulfate. The core structures from the breast epithelium are additionally elongated with polylactosamine repeating units and may comprise more than 15 monosaccharide residues. Typical mucin oligosaccharide structures are depicted in ◉ *Fig. 3*.

Normal mucin maturation is an iterative and dynamic process and shortly after expression the mucin that is displayed at the cell surface is only partially sialylated and a significant proportion of the oligosaccharides are neutral (uncharged). Maturation occurs as the cell surface membranes turnover and are recycled back to the *trans* Golgi where they undergo further sialylation before they are again transported back to the cell surface.

Once a peptide is glycosylated the attached *O*-linked saccharides make specific interactions with the peptide backbone that disrupt the formation of regular, or expected (based on protein sequence), secondary structure and promotes the formation of β-turns (◉ *Fig. 4*). The presence of a stable, defined, extended structure has been demonstrated by detailed NMR investigations of mucin and mucin-type glycopeptides which, in general display many more NOE's in the NMR spectrum than would normally be expected from a "randomly" glycosylated and unstructured short peptide sequence. The formation of α-helical conformations are particularly disrupted by *O*-glycosylated peptides. Specific to α-*O*-glycopeptides (the corresponding β-*O*-linked GalNAc does not exert the same influence) [19], the observed conformational pref-

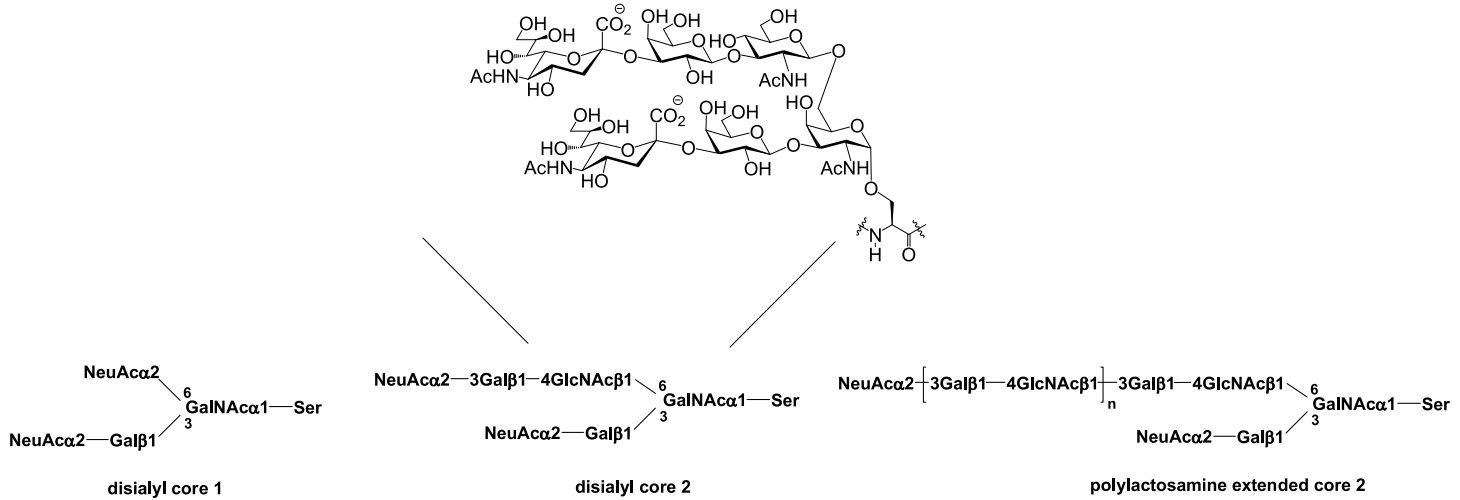
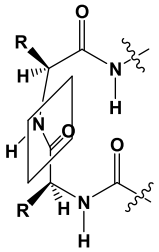
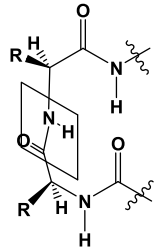


Figure 3

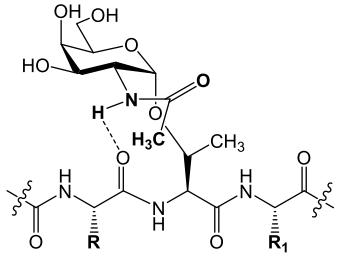
Typical *O*-linked mucin oligosaccharides. The initial GalNAc is modified by core 1 and core 2-forming glycosyltransferases and further with sialic acid, galactose, *N*-acetyl-glucosamine, (additionally also with fucose, and sulfate culminating in the formation of selectin ligands such as sLe^x (see *Fig. 5b*))



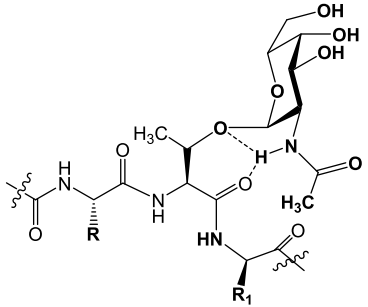
Type I beta turn



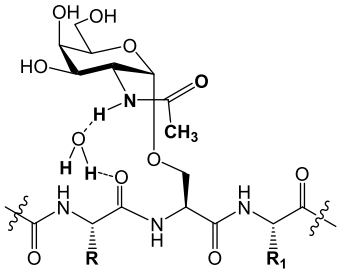
Type II beta turn



Hydrogen bond to backbone NH



Alternative hydrogen bond to backbone NH



Bridging water

Figure 4

Common β-turn motifs found in protein structures and possible hydrogen-bonding interactions between the carbohydrate and the peptide backbone that are believed to influence glycopeptide structure. Note that the turn-type structure is most clearly depicted in the central structure

erence has been attributed to stereoelectronic effects (the exo-anomeric effect in particular) and hydrogen bonding interactions between the N–H bond of the *N*-acetyl group of GalNAc and the amide carbonyl group of the proximal amino acid residue (◉ *Fig. 4*) in the underlying peptide sequence [20], and alternatively to hydrogen bonding interactions between the N–H bond of the *N*-acetyl group of GalNAc and the carbonyl group of the same (threonine) amino

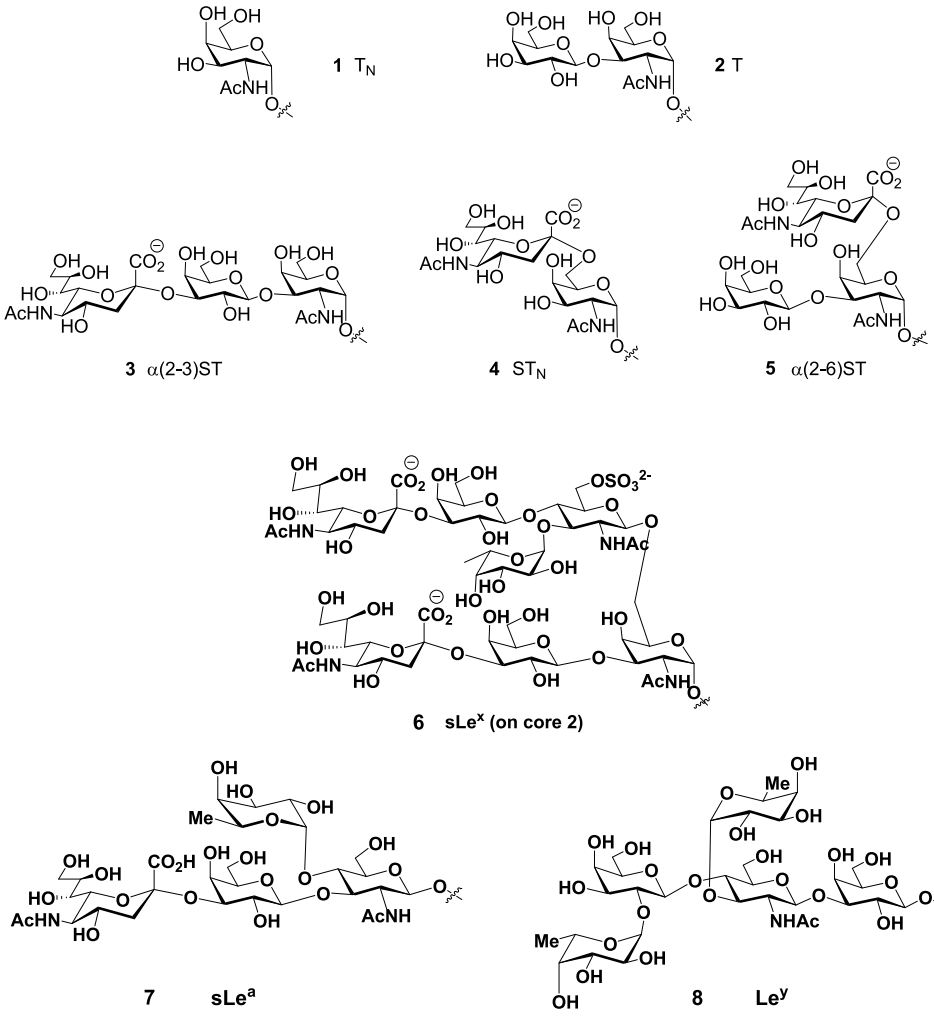


Figure 5

(a) Tumor-associated antigens are often under-glycosylated e. g. T_N (1), T (2), and α2,3-sialylT (3), and prematurely sialylated e. g. α2,6-sialyl-T_N (4), and α2,3-α2,6-disialyl-T_N (5). The premature sialylation on the 6-hydroxyl group of the *N*-acetylgalactosamine sugar precludes the formation of core 2 type structures. (b) Some Lewis antigens associated with cancer: (see ◉ *Table 2*) sLe^x (sialylLewis x, 6), sLe^a, (7) and Le^y (8) sLe^x (sialylLewis x, 6), sLe^a, (7) and Le^y (8)

acid residue [19,21]. Recent investigations suggest that H-bonding interactions are perhaps too weak for α-*O*-glycosylserine containing peptides and are not observed in conformational studies of model glycoamino acid diamides by NMR spectroscopy. This report suggests that inclusion of bridging water molecules may stabilize the extended conformation [22].

A detailed conformational analysis of glycopeptides corresponding to fragments of LI-cadherin (sequence: LAALD**S**QGAIV) containing native and cancer associated glycopeptide antigens was also described by Kunz and co-workers showing clear evidence for turn formation in peptides glycosylated on serine (bold) with the T_N (**1**), T (**2**), α2,3-sialylT (**3**), α2,6-sialyl-T_N (**4**), α2,3-α2,6-disialyl-T_N (**5**), and glycophorin structures (● *Fig. 5a*). Furthermore the report demonstrated dramatic and significantly different conformational effects depending on the identity of the appended saccharide. Peptides modified with T_N, T, and their mono and disialylated analogues all showed a more extended peptide structure relative to the unglycosylated peptide, and the larger (sialylated) saccharides orchestrating the formation of a more profound γ-turn motif. The T_N and T antigens, which are associated with tumors, imposed β type-III' and β type-II turn conformations on the peptide respectively [23]. In addition to the clear conformational preference of the appended saccharides it is additionally proposed that the charge repulsion originating from the large number of sialic acid residues in close proximity further dictates that mucins adopt an extended rod-like structure. Consequently the rod-like peptide sequence is often described as merely a scaffold on which saccharides are displayed. As mentioned above, the hydrophilic and charged nature of mucins dictates their physical properties, sequestering water molecules and ionic species they form a gel-like protective layer that coats the surface and are secreted from endothelial cells which are specialized for mucin production.

4 Mucin Glycosylation and Cancer

Mucins are the major carriers of altered glycosylation in carcinomas. The description of "abnormal" mucin expression and aberrant glycosylation in tumor cell lines is often attributed to the following key observations [3,24,25,26]. First, mucin expression is up-regulated in tumor cells or indeed present in tumor cells when the healthy cells are devoid of such mucin expression. This is exemplified by over-expression of MUC1 in many cancers, and with MUC4 expression in pancreatic cancer since the pancreatic epithelial cells do not normally express MUC4 [7]. Second, the carbohydrates present in the mucin are abnormal, often truncated. This is believed to stem from premature sialylation of the T_N antigen or T (also known as the T, Thomsen–Friedensreich) antigen which precludes further biosynthesis (i. e. prevents action of core 2 GlcNAc transferase) [27]. Additionally the T_N (**1**) or T (**2**) antigens themselves (or even the "naked" apomucin) may be present (● *Fig. 5a*). It has been proposed that the appearance of such antigens might result from mucin over-expression, the abundance of acceptor sites saturating the posttranslational apparatus. Indeed it has been demonstrated that over-expression of MUC1 (which is over-expressed in 90% of human breast carcinomas) is correlated with mammary gland transformation in vivo [12] and the truncated sialylT_N antigen (**4**, ● *Fig. 5a*) is a hallmark of this type of cancer and vaccines based on this epitope are in clinical trials for cancer therapy [25]. This hypoglycosylation of mucins can additionally arise from transformed cells which have lost core 2-forming β-1,6-GlcNAc transferase activity [28] or

Table 2
Common expression patterns of cancer glycans found on mucins in malignant tissues. Adapted from Dube and Bertozzi [25]

Tumor-associated antigen	Malignant tissue					
	Ovary	Pancreas	Breast	Colon	Prostate	Lung
sLe^x		X	X	X		X
sLe^a		X	X	X		X
sT_N	X	X	X	X	X	X
T	X		X	X	X	
Le^y	X	X	X	X	X	X

over-express sialytransferases [27]. Depending on the physiological conditions, the peripheral sugars may vary and give rise to further tumor-associated antigens (◉ *Fig. 5b*). Expression of tumor-associated antigens is often a result of deregulation of glycosyltransferase expression or neoexpression of glycosyltransferases that are normally only observed during development giving rise to carbohydrate epitopes such as Le^a (Lewis a), Le^b, Le^x, Le^y (**8**), sLe^a (sialyl Lewis a), sLe^c, and sLe^x (**6**) which, though not unique to tumor cells, are the hallmark of various tumors as a result of their "wrong place-wrong time" appearance (◉ *Table 2*). Why specific sets of tumor-associated antigens predominate in particular carcinomas is unclear, however, it is likely that they have been selected from mixed populations to confer some advantageous properties to the tumor cell such as evasion from immunosurveillance machinery and increased adhesion through expression of selectin ligands (e. g. sLe^x) [29,30] that additionally facilitate tissue invasion and colonization of secondary sites [24].

Finally, in normal epithelial cells mucin expression is restricted to the apical domain (oriented towards the lumen) and consequently once proteolyzed or secreted the mucin is delivered to the lumen. However in many cancer cells the regulated apical expression of mucins appears corrupted with mucin expression (and secretion) also into intercellular spaces and into tissues and plasma. Indeed, soluble mucins are often observed in the blood of cancer patients. This seems to result in abnormal cell topology which inhibits normal cell-cell adhesive interactions (through increased negative charge and increased steric-bulk) facilitating tumor cell metastasis and evasion from the normal immune response. Pre-incubation of tumor cells with α-benzyl GalNAc which inhibits *O*-linked protein glycosylation (by diverting the biosynthetic machinery to the small molecule) reduces the metastatic potential of various tumor cell lines, directly linking glycosylation with malignancy [31]. A similar affect is observed when mucin expression is attenuated in cancer cell lines. For example, suppression of MUC1 over-expression in human pancreatic cancer cell line S2-013 using siRNA technology showed slightly increased adhesion in vitro (reduced over glycosylation) and reduced metastasis to secondary sites in vivo when re-administered to the pancreas of athymic mice [32].

Not all mucin over-expression is associated with carcinogenesis and much of what we know about mucins and their role in tumor development has been proposed as a result of knock-out studies in mice, where the expression of genes encoding various (apo)mucins has been completely abrogated. In contrast to the demonstrated relationship between MUC1 over-expression and tumor progression, in a recent study knock-out mice lacking the most abundant secreted

mucin in the gastrointestinal tract MUC2 were shown to have abnormal organ morphology and behavior [33]. Sixty-five percent of *muc2* null mice had developed tumors (that had originated from epithelial cells) in the colon, and the small and large intestine by the age of 1 year suggesting tumor suppressor activity for MUC2.

Given the importance of e. g. MUC1, and other cell-surface associated mucins in tumor development and metastasis [7] it is reasonable to presume that if altered mucin glycosylation accompanies tumor progression then any tumor-associated glycopeptide motif may be cancer cell specific and hence an immune response raised against aberrantly glycosylated MUC1, may serve to counter the spread of tumor cells. The replacement of regular core 1 and core 2 structures with their truncated T_N, T and sialylated T_N/T precursors is particularly prolific and it is these "unnatural" structures that have been the focus for much vaccine development since, potentially, they also allow antigenic elements of the peptide backbone, that would otherwise not be accessible in healthy cells to be targeted. The value of searching for vaccines against these epitopes is further validated by the clear demonstration of tumor-associated antigen (TAA)-specific glycopeptide conformations [23].

Because of the overwhelming evidence that aberrant mucin glycosylation is the hallmark of various cancers and may play a causative role in immune evasion and metastasis it is unsurprising that mucin based glycopeptides have been targeted for anti-cancer vaccine development. It should be noted that mucins are also *N*-glycosylated but targeting *N*-linked oligosaccharides on mucins has, to date, not received so much attention. However, over-expression of the glycosyltransferase GlcNAcT V which allows modification of *N*-linked glycans with *N*-acetyllactosamine repeats is associated with poor prognosis for cancer patients and is the focus of considerable investigation [25,34]. There are considerable advantages to pursuing mucin based vaccines. First, the vaccine itself may be otherwise non-toxic to healthy cells unlike many standard chemotherapy agents and is additionally non-radioactive so the potential for unpleasant side-effects may be reduced. However the use of cancer cell derived mucins as immunogens has been largely unsuccessful since they are heterogeneous. Generally containing a mixture of normal and tumor associated glycosylation, isolated mucins are very weakly immunogenic and recognized as "self." The remainder of this review will focus on how chemistry and biology have combined in an attempt to break this immunotolerance and to initiate a high quality cancer cell-specific response. The type of immune response generated has ramifications for the efficacy of the treatment. For example if B-cells are activated and antibodies are produced then they may only be beneficial in the clearance of non-adherent (e. g. metastasizing) tumor cells and circulating tumor cells after surgery since antibodies have limited capacity to penetrate the interior of solid tumors. Solid tumors also often shed their VNTR domains from the cell surface which neutralizes the effect of antibodies, though secreted mucins may still serve to activate B-lymphocytes (and hence T_H-cells) that express membrane-bound antibodies. Additionally, antibodies raised against tumor-associated antigens displayed by mucins (MUC1 in particular) have been shown to fail to recruit effector functions that kill cancer cells [35]. For reasons like this the use of antibodies raised against TAAs are often employed more widely as diagnostic agents or may possibly be utilized in antibody directed enzyme pro-drug therapy (ADEPT). A T-cell response, on the other hand, has the potential to invade cancerous tissue and is generally considered more selective than an antibody response and is consequently more desirable. Furthermore there is no requirement for additional effector molecules since activated cytotoxic $CD8^+$ T-cells can kill tumor cells directly through granule mediated exo-

cytosis. Carbohydrates are generally incapable of eliciting a cytotoxic $CD8^+$ T-cell response since, by themselves, carbohydrates are not presented by major histocompatibility complex (MHC) molecules though there is strong evidence that cells specialized for presenting tumor-associated antigens to T-cells are capable of presenting glycopeptides [36]. However, cancer cells have been shown to subvert the efficacy of a tumor specific T-cell response by attenuating the T-cells ability to transduce sufficient co-stimulatory signals [37]. The factors that need to be considered when initiating T-cell and B-cell responses to tumor-associated antigens are discussed briefly in the following section.

5 Generation of a Basic Immune Response

In order to discuss the current state of mucin-based vaccine technology within the context of generating an immune response, it is first necessary to understand a little about the way in which a basic immune response is initiated. For more detailed information regarding the induction and regulation of immune responses the reader is referred elsewhere [38,39]. The immune system is composed of two main branches: the innate and adaptive systems, that function together to protect the host against invading pathogens and the generation of potentially harmful cells within the body. This functionality is dependent upon the ability of the immune system to differentiate between healthy cells and those that require elimination. All nucleated cells in the body identify themselves to the immune system using antigens: short peptide fragments that act as "immunological identity cards." The adaptive immune system can be split into the humoral and cellular branches, and the implementation of either branch depends upon the successful recognition of antigens by lymphocytes. Antigens are continually exposed to scrutiny by both T- and B-lymphocytes, and those identified as foreign induce an immune response resulting in clearance. The adaptive immune response to a specific antigen becomes increasingly potent following each repeated encounter with the same specific antigen, and culminates in the development of long-lived antigen-specific immunological memory.

Cancer cells express many altered cell surface antigens that can potentially serve as targets for an anti-tumor immune response. However, as all cancers arise from the transformation of an otherwise healthy cell, many characteristics of the healthy cell are maintained, and consequently, the boundaries that enable the immune system to differentiate between "self" and "altered self" are often blurred. Immunological tolerance to "self" antigens is required to maintain tissue homeostasis and prevent the development of autoimmune disease. However, tolerance to "altered self" antigens facilitates disease progression. Hence, selectively breaking tolerance to promote tumor clearance is an area of intense research. The immune system often fails to clear tumors due to immunological tolerance towards altered self-antigens and cancer-mediated suppression of the immune response. Thus, whilst the immune system acts in a concerted manner to protect against invaders, the removal of well-established tumors is generally achieved through surgical intervention followed by chemotherapy or radiotherapy. However, during the early stages of cancer development, numerous encounters occur between T- and B-lymphocytes and developing cancer cells, and in some instances, successful elimination of these cells may occur. The activation of T- and B-lymphocytes during a basic immune response is summarized below.

6 Antigen Presentation

A T-lymphocyte (T-cell) response is triggered when antigens are acquired by professional antigen presenting cells (APCs). The APC family consists of B-lymphocytes (B-cells), monocytes, macrophages and dendritic cells (DCs), which are by far the most potent type of APC. Antigens are derived from endogenous and exogenous sources and both are assembled onto molecules of the MHC through "classically" segregated presentation pathways. Endogenous proteins synthesized in the cytosol are degraded in the proteasome into peptide antigens and transported to the lumen of the endoplasmic reticulum (ER), where they are assembled onto MHC class I molecules. The antigen-MHC complex is then transported to the surface of the APC (*Fig. 6*, *A–D*).

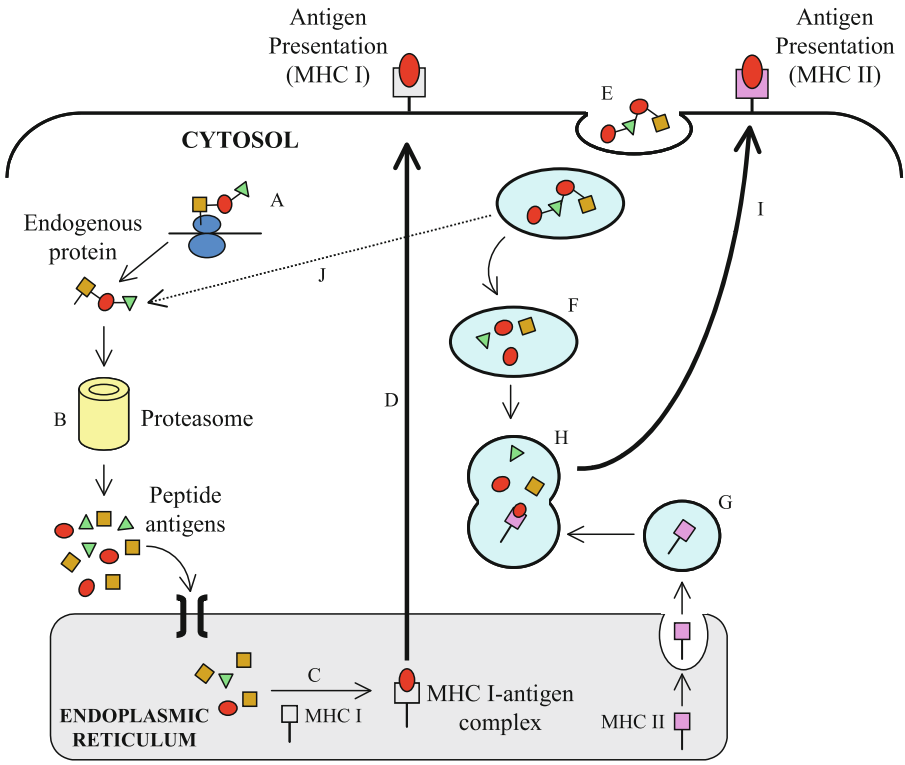


Figure 6

Pathways of antigen presentation by MHC class I/II molecules. *A–D*. Translation of cytosolic mRNA produces endogenous proteins that are degraded into antigen peptides by the proteasome. The antigens are imported into the endoplasmic reticulum and assembled onto MHC class I molecules that are subsequently transported to the cell surface. *E–I*. Pathogens and antigens are internalized by phagocytosis. Acidification of the intracellular vesicles triggers proteolysis generating peptide antigens that are assembled onto MHC class II molecules that are presented at the cell surface. *J*. Proteins derived from the phagocytosis of exogenous material can be presented on MHC class I molecules through the cross-presentation pathway. (Modified from Heath and Carbone [58])

Exogenous antigens are derived from intracellular vesicles following the phagocytosis of invading pathogens by macrophages or internalization of antigens by DCs and B-cells. Vesicular acidification triggers proteolytic processing of the captured proteins into antigenic peptide fragments (Fig. 6, *E–F*). Like class I, MHC class II molecules also reside in the ER lumen, although in contrast to class I, class II molecules are transported from the ER to the cytoplasm in vesicles which subsequently fuse with the acidified vesicles containing the processed antigen fragments. The antigens are then loaded onto the MHC class II molecule and the resulting MHC-antigen complex is presented on the APC surface (Fig. 6, *G–I*). Accordingly, MHC class I molecules are expressed on all nucleated cells that synthesize endogenous proteins with expression of class II molecules restricted to APCs that can phagocytose both pathogens and exogenous antigens. Antigen presentation by either class of MHC molecule can be crudely likened to a "Hot Dog in a bun," where the antigen is presented in an extended conformation (the "Hot Dog") bound by the vice-like structure of the MHC antigen binding groove ("the bun"). Antigens presented by MHC class I molecules are typically between 8 and 10 amino acid residues in length whilst those presented by class II antigens can be larger. DCs that have acquired antigen receive maturation signals from the site of infection and migrate to the lymphoid tissues where the antigen is presented to circulating T-cells. The T-cells interact with antigen-MHC complexes by virtue of antigen-specific T-cell receptors (TCR) expressed on their surface. Random variations in the amino acid sequence (and hence structure) of the antigen binding regions of the TCR generate specificity to a potentially limitless pool of antigens. Binding of TCRs to either MHC class I or II molecules is determined by the expression of co-receptors (CD8 or CD4) that define T-cell effector function. $CD8^+$ T-cells recognize antigens bound to MHC class I molecules and $CD4^+$ T-helper (T_H) cells interact with antigens bound to MHC class II. Binding of CD8 and CD4 receptors to MHC class I and II respectively increases the affinity of TCR binding to the MHC-antigen complex. Critically, activation of either $CD8^+$ T-cells or $CD4^+$ T_H cells is dependent upon at least two stimulatory signals which must be provided by the same APC.

7 Activation of $CD8^+$ T-Cells

At least two signaling events must occur to successfully initiate a $CD8^+$ T-cell response (Fig. 7*a*). First, the TCR must recognize a cognate antigen bound to MHC class I molecules presented by APCs. Several TCR-antigen interactions are required to stimulate efficient activation, with additional co-stimulatory signaling between CD28 on the T-cell and B7 molecules on the APC completing activation.

Signaling between CD28 and B7 acts to "validate" the authenticity of the antigen-TCR interaction. Indeed, T-cells that do not receive co-stimulation fail to initiate an immune response against a presented foreign antigen. This process is termed anergy, and once imposed, results in immunological unresponsiveness to the same antigen in subsequent encounters, even in the presence of co-stimulatory signaling. In the presence of co-stimulation, proliferation and differentiation of activated $CD8^+$ T-cells is driven by secretion of the cytokine interleukin-2 (IL-2). Binding of IL-2 to a cognate IL-2 receptor expressed on the activated cell stimulates clonal expansion into large populations of antigen specific cytotoxic T-lymphocytes (CTLs) (Fig. 7*b*). CTLs do not require co-stimulatory signals to initiate an immune attack, and

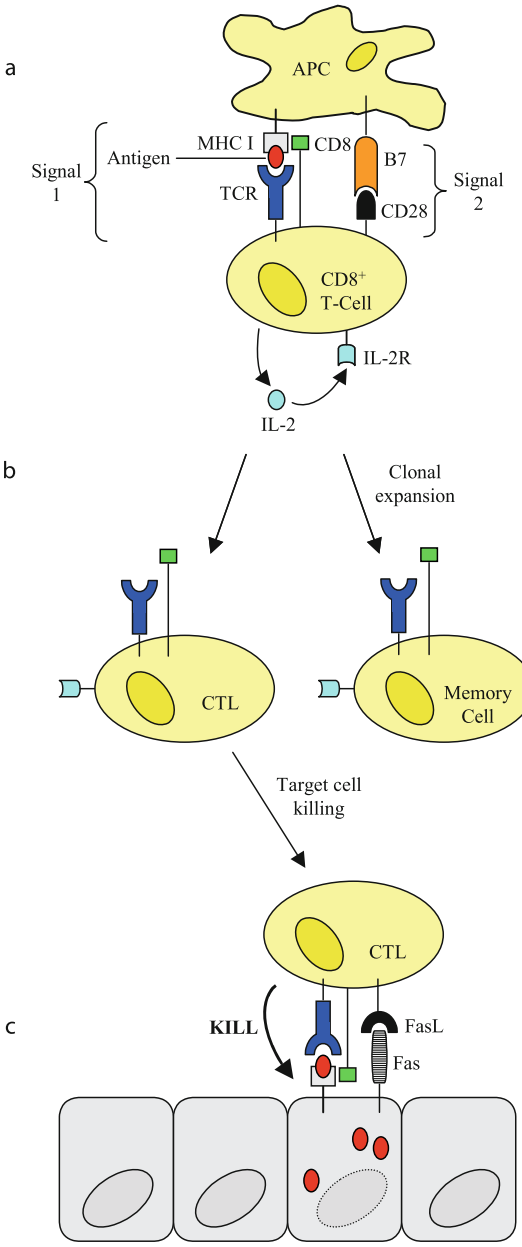


Figure 7

Activation of CD8+ T-cells. (a) APCs present antigens on MHC class I molecules that are recognized by the TCR and CD8 co-receptor respectively. Co-stimulation between B7 and CD28 is also required for activation and autocrine signaling via IL-2 drives clonal expansion. (b) Activated CD8+ T-cells develop into either CTLs or memory T-cells. (c) Killing of target cells by antigen-specific CTLs proceeds by either granule exocytosis or FasL-Fas interaction. (Modified from Janeway et al. [38] Figure 8.22)

are released into the body where they migrate to the peripheral tissues to search for target cells. A proportion of the expanded T-cell population may proliferate to form antigen-specific memory T-cells. The development of immunological memory (humoral or cellular) following a primary challenge against an encountered antigen enables a rapid and aggressive immune response to be re-generated against the same antigen should it be encountered in a subsequent infection, and thus forms the basis upon which the administration of prophylactic vaccines is founded.

8 CTL-Mediated Killing of Target Cells

CTLs can eliminate target cells by inducing programmed cell death (apoptosis) by two contact dependent mechanisms, the first of which is granule exocytosis. When a CTL encounters a target cell, an interface is formed between the TCR and the MHC class I-antigen complex. Formation of this interface stimulates the CTL to synthesize cytotoxic granules: specialized lysosomes containing factors that promote apoptosis. Once synthesized, the granules migrate to the junction between the effector and target cell where their contents are released. The target cell is breached by perforin, a small monomeric protein that inserts into the plasma membrane and polymerizes to form a pore through which the pro-apoptotic factors are delivered. Apoptosis is triggered by the release of powerful serine proteases called granzymes that degrade multiple substrates resulting in death of the target cell. Once apoptosis is initiated, the CTL can detach and engage new target cells. CTLs can eliminate target cells using a second contact dependent mechanism mediated between Fas ligands (FasL) on the CTL and Fas molecules expressed on the target cell. Once a FasL-Fas interaction is established, a proteolytic cascade is initiated that culminates in the activation of caspase enzymes, thereby promoting apoptosis (◉ *Fig. 7c*).

9 Activation of CD4$^+$ T_H Cells

Unlike CD8$^+$ T-cells, CD4$^+$ T_H cells do not exhibit cytotoxic effector functions, and are thus unable to destroy infected cells or those that display altered self-antigens. However, CD4$^+$ T_H cells are crucial in activating and directing the effector functions of other immune cells. Activation of CD4$^+$ T_H cells occurs in a similar manner to CD8$^+$ T-cells, the critical difference being the class of MHC molecule to which the antigen is bound. Recognition of an MHC class II-antigen complex by the CD4$^+$ TCR together with co-stimulatory signaling through B7 and CD28 drives activation, and autocrine signaling via IL-2 stimulates clonal expansion (◉ *Fig. 8a*).

The expanded population of T_H cells proliferate into either T_{H1} or T_{H2} subsets (◉ *Fig. 8a*), which are distinguished by the different cytokines they produce (◉ *Fig. 8b*). T_{H1} cells secrete numerous cytokines including IL-2 and interferon gamma (IFN-γ) that activate macrophages and CD8$^+$ T-cells. Cytokines secreted by T_{H2} cells include IL-4 and IL-5, which are involved in the stimulation of antibody responses (◉ *Fig. 8c*). In general, the cytokines secreted by one subtype of T_H cells are inhibitory to the development of the other, resulting in an effector response driven by either T_{H1} or T_{H2} but not both simultaneously. Thus, the secreted cytokine milieu determines if the overall CD4$^+$ T_H cell response is cellular or humoral which, as we will see, has important ramifications for vaccine development.

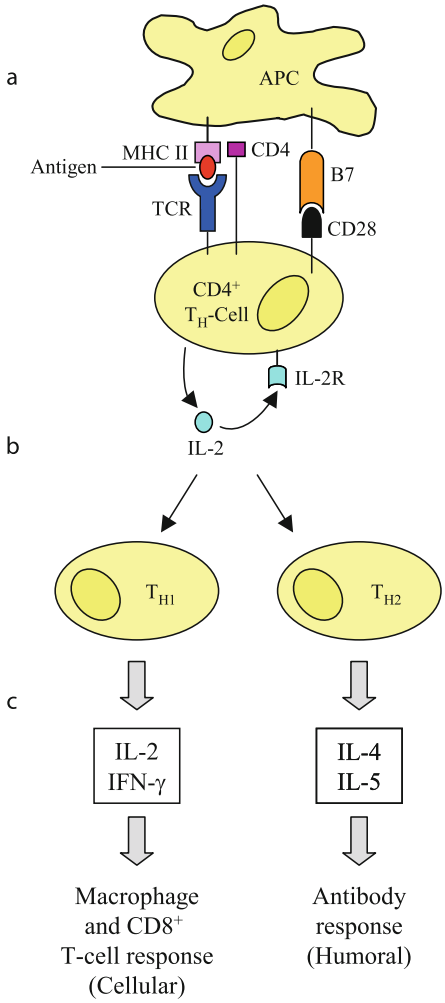
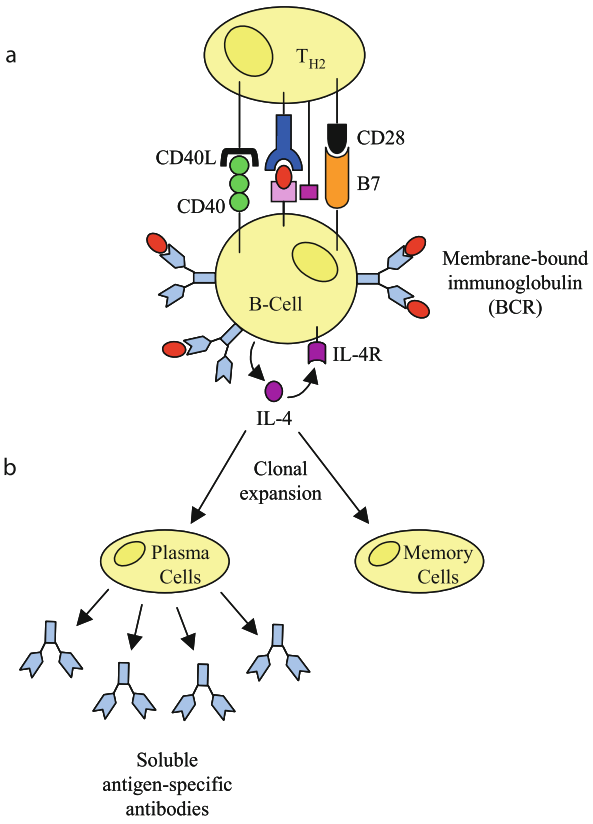


Figure 8
Activation of CD4⁺ T_H cells. (a) CD4⁺ T_H cells recognize antigen presented on MHC class II molecules. Co-stimulation and autocrine signaling are required to complete activation and drive clonal expansion. (b) Activated CD4⁺ T_H cells may proliferate into either T_{H1} or T_{H2} cells. (c) T_{H1} and T_{H2} cells secrete a variety of cytokines that stimulate the generation of either cellular or humoral immune responses. (Modified from Janeway et al. [38] Figure 8.10)

10 Activation of B-Cells

As previously mentioned, B-cells are APCs that must first acquire antigen to become activated. In contrast to the TCR expressed by T-cells, the B-cell receptor (BCR) consists of antigen specific cell surface immunoglobulins (IgM and IgD). Each B-cell expresses a BCR that is specific for a different antigen. Antigen specificity in the B-cell population within the body

derives from variations in the amino acid sequence of the antigen-binding region of the BCR molecule, which allows a huge number of potential antigen structures to be recognized. In contrast to $CD8^+$ T-cell and $CD4^+$ T_H cell activation, which proceed after each receives stimulatory signals from the same APC, the stimulatory inputs required for B-cell activation are provided from two different sources. The first is provided by antigen binding to a cognate BCR, which is subsequently internalized by receptor-mediated endocytosis. The antigen is processed and bound to MHC class II molecules that are transported to the B-cell surface. An activated $CD4^+$ T-cell provides the second signal upon recognition of the antigen-MHC class II complex. Co-stimulatory signaling between CD40L on the T-cell and CD40 on the B-cell completes activation, with secretion of IL-4 driving clonal expansion (◉ *Fig. 9a*).



◘ Figure 9
Activation of B-cells. (a) Binding of soluble antigen to a BCR is followed by receptor-mediated endocytosis and presentation by MHC class II molecules. An antigen-specific T_{H2} cell binds the antigen with its TCR and provides additional stimulation through CD28 and CD40L. Clonal expansion of activated B-cells occurs after autocrine signaling through IL-4. (b) Activated B-cells may proliferate into memory B-cells or plasma cells that secrete a large amount of antigen-specific antibodies

Activated B-cells undergo proliferation and may differentiate into antibody secreting plasma cells that facilitate antigen clearance, or become committed to a pool of long-lived memory B-cells (◉ *Fig. 9b*). Whilst the specificity of secreted antibodies derives from the complementarity determining regions of the molecule, the constant region of the molecule triggers different effector responses depending upon the antibody isotype. Binding of secreted antibody to a cognate epitope may target a cell for destruction either by antibody dependent cell-mediated cytotoxicity (ADCC), or by complement-dependent cytotoxicity (CDC).

11 Mucins, Immune System Evasion and Cancer Development

Cancers avoid immunological detection in a variety of ways including down-regulating the expression of immuno-stimulatory antigens, MHC molecules and co-stimulatory receptors required for immune cell activation [40]. These strategies are usually accompanied by a concomitant over-expression of immuno-modulatory cytokines [40] such as transforming growth factor-beta (TGF-β), which imposes numerous immuno-suppressive effects [41]. TGF-β-specific impairment of murine DC migration and antigen presentation has been observed [42] whilst expression of perforin, granzyme A, granzyme B, FasL and IFN-γ by CTLs is also inhibited [43]. A TGF-β-dependent increase in MUC4 expression is observed in pancreatic adenocarcinoma [11], an effect which is likely to contribute to cancer development, as MUC4 silencing with antisense MUC4 RNA was observed to suppress the growth of pancreatic tumor cells [13]. Premature termination of glycosylation is a hallmark of cancer-associated mucins, and the presence of such structures facilitates immune evasion, thereby contributing to disease progression. Addition of recombinant MUC1 displaying *O*-linked sialylated core 1 (sialyl-T) oligosaccharides to human DC impaired maturation and activation capacity in vitro [44]. Moreover, secretion of sialyl-T_N-MUC1 by numerous pancreatic tumor cell lines abrogated DC-mediated T-cell activation [45]. Secretion of MUC1 by ovarian, pancreatic and breast cancer cells blocked T-cell activation directly [46], and induces T-cell anergy that was reversible by addition of IL-2 [47].

One of the ways in which the immune system eliminates unwanted cells from the body is by CDC. CDC proceeds following the secretion and binding of specific classes and isotypes of immunoglobulins (IgM, IgG1 and IgG3 in humans) from B-cells to target cells. Blood serum from patients vaccinated with QS-21 and keyhole limpet hemocyanin (KLH) conjugated to either clustered T, sT_N, a 32 amino acid MUC1 peptide displaying *O*-linked T_N or the unglycosylated MUC1 peptide alone produced antigen-specific antibodies of the appropriate classes and isotypes to mediate complement fixation [35]. These antibodies bound to MCF-7 cancer cells expressing sT_N, T, MUC1, GM2, globoH, and Le^Y in vitro but failed to induce lysis by CDC. In contrast, parallel experiments revealed that antibodies generated against the expressed glycolipid antigens directed CDC efficiently [35]. From these observations it was proposed that the extended rod-like structure of glycosylated mucins projecting away from the cell surface may hinder the recruitment of certain immune system components towards cells that express them [35]. However, alternative immunoglobulin-dependent mechanisms of target cell clearance are not inhibited by the molecular architecture of cancer-associated mucins, as antibodies from breast cancer patients vaccinated with a KLH-MUC1 peptide conjugate promoted ADCC on MUC1 expressing cancer cells in vitro [48]. In contrast to CDC however,

the observed ADCC occurred by recruitment of natural killer cells to eliminate targeted cancer cells [48] rather than the components of the complement cascade.

The spread of cancer occurs following metastasis from the site of transformation and requires that cancer cells disengage from the primary tumor and adhere to new cells and target tissues elsewhere in the body. Galectin-3 is a galactoside-binding lectin expressed on the surface of endothelial cells. Addition of recombinant galectin-3 facilitated the adhesion of MUC1 positive, but not negative breast and colon cancer cells to human umbilical vein endothelial cells in vitro [49]. The observed adhesion was largely dependent upon an interaction between galectin-3 and the MUC1 cancer-associated T antigen, as removal of T from the surface of cancer cells reduced binding. Furthermore, the distribution of MUC1 on the surface of the cancer cell was altered following galectin-3 addition, and the observed adhesion to endothelial cells markedly reduced by co-incubation with antibodies specific for E-Selectin and CD44H, suggesting that the galectin-3-T interaction facilitates the adhesion of cancer cells to epithelial targets by unveiling adhesion molecules previously masked by cancer-associated MUC1 [49].

However, an oncogenic capacity of tumor-associated MUC1 is not restricted to the hypo-glycosylated extracellular domain. In response to damage of cellular DNA, the tyrosine kinase c-Abl is targeted to the nucleus of the cell where it promotes apoptosis. The cytoplasmic domain of MUC1 confers resistance to apoptosis in response to DNA damage by binding to c-Abl, thus preventing its import into the nucleus [10]. The MUC1-c-Abl interaction therefore allows DNA damage to accumulate, which may contribute to cancer development and/or progression. The induction of cellular apoptosis acts as a safeguard whereby potentially deleterious cells undergo programmed destruction before they can cause harm to the body. Activation of the anti-apoptotic Bcl-$_{XL}$ and phosphoinositide 3-kinase pathways was observed in MUC1 transfected rat 3Y1 fibroblast cells in vitro [50]. Moreover, MUC1 circumvents apoptosis in response to treatment with the chemotherapeutic agent cisplatin [51], pro-apoptotic drugs 1-β-D-arabinofuranosylcytosine and gemcitabine [50] and in response to hypoxic stress [52]. Collectively, these observations demonstrate the involvement of cancer-associated mucins not only in the process of immune evasion, but also in the process of cellular transformation itself.

12 Considerations for the Design of Mucin-Based Vaccines

The paramount feature of any synthetic vaccine molecule is safety. Vaccines must exhibit a negligible level of toxicity and have very few side effects. The molecule should be inexpensive to produce, soluble and chemically stable. Since the number of antigen-TCR interactions influences the efficiency of T-cell activation, the in vivo stability of the vaccine and rate of biological clearance must also be considered. Synthetic vaccines against cancers that express aberrantly glycosylated mucins are generally comprised of one or more TAAs that may be covalently conjugated to a carrier molecule. These immunogenic molecules serve to modify the chemical properties of synthetic antigens (e. g. improved solubility) and facilitate transport across membranes within the body. For a molecule to be an effective carrier, several criteria must be satisfied. Firstly, it must be safe to use with minimal toxicity and side effects, it must be soluble and easily produced with minimal molecular heterogeneity. The molecule must be amenable to chemical modification and conjugation to target antigens and lastly, it must be immunogenic to seed cross-reactivity to the conjugated antigen.

On healthy cells, the MUC1 peptide backbone is shielded from immune detection by an elaborate network of carbohydrates displayed upon the molecule. Hypo-glycosylation of mucins on cancer cells exposes cryptic peptide motifs and generates mucin-based glyco-antigens (e. g. T, T_N, sT_N) that can serve as TAAs in vaccine design. Consequently, synthesis of mono and multi-valent peptide and glycopeptide vaccines is an area of intense research. The synthesis of ever more complex multi-valent vaccines may pave the way forward in a multi faceted approach to cancer therapy, exploiting an array of peptide and glycopeptide TAAs that generate a simultaneous immune response against multiple targets. Multi-valent vaccines aim to provide the immune system with molecules that represent the antigenic architecture expressed on the surface of transformed cells. However, as the total number of aberrantly glycosylated mucins expressed on cancer cells may not be known, the number of mucin-based TAAs identified thus far may act as a limitation to vaccine efficacy. Furthermore, to facilitate the induction of a potent immune response, the design of synthetic vaccine molecules must proceed in a way that allows completely intact TAAs to be presented by APCs following internalization and proteolytic processing.

Tumors can be considered as a mixed population of cells, each of which is genetically unstable, resulting in the generation of a large number of heterogeneous cell surface antigens. It is postulated [53] that cancer progression stems directly from this antigenic heterogeneity and is, paradoxically, facilitated by an active immune system that destroys highly immunogenic cells whilst leaving those that do not invoke an immune response to proliferate. This process is termed immuno-editing and may inadvertently provide the growing tumor with a means of escape from clearance. Critically, antigenic heterogeneity may facilitate tumor adaptation against therapeutic agents that do not promote rapid clearance. The process of immuno-editing has profound implications in relation to the design of mucin-based vaccines. In many cases, structural characterization of mucin-type antigens prior to chemical synthesis requires that an established tumor be surgically removed. However, it is likely that many of the aberrant structures selected for synthesis and incorporation into vaccines are those that, due to immuno-editing, exhibit the lowest intrinsic levels of immunogenicity. Poor immunogenicity notwithstanding, expression of these "edited" antigens are nonetheless likely to be restricted to cancer cells alone, and as such should not be overlooked in vaccine design. Indeed, attempts to break tolerance using synthetic vaccines that display poorly immunogenic altered "self" antigens often rely upon the use of immunological adjuvants, molecules that, when mixed and co-administered with a vaccine boost the immune response against the associated antigen. Due to the poor immunogenicity of some TAAs, the choice of adjuvant can be critical in determining the effectiveness of the induced immune response. In many cases however, the use of vaccines in an adjuvant setting directs a sub-optimal induction of immune responses. Characterization of the adjuvant-mediated augmentation of immune responses to synthetic MUC1 [54], and glycosylated MUC2 [55] conjugate vaccines revealed high antigen-specific antibody titers with no direct evidence of CTL activation.

The generation of an immune response is most efficient when antigens are presented to T- and B-cells by APCs in a wholly biological context. Direct presentation of antigens by APCs negates the requirement for immunological adjuvants and provides an environment rich in essential co-stimulatory molecules. DCs can be isolated from the blood of cancer patients and provided with mucin-based antigens in vitro that allow processing and presentation to occur in the absence of immuno-subversion by established tumors. Thus, a pool of antigen presenting

DCs are generated that can be re-introduced to the patient (adoptive transfer). Adoptive transfer of DCs pulsed with lysate obtained from tumors that express cancer-associated mucins may potentially provide a means of exploiting a large number of hitherto unidentified mucin-based TAAs. As discussed above (see ❯ *Sect. 4*), premature termination of mucin glycosylation is a cancer-specific phenomenon that produces a pool of altered self TAAs. An interesting question that arises from this observation is "are these TAAs derived exclusively from the exposure of mucin core peptide sequences following the premature termination of glycosylation?" Murine DCs were observed to internalize, process and present MUC1 glycopeptides on MHC class II molecules in vitro [36], demonstrating that *O*-linked carbohydrate groups are not removed during antigen processing and suggesting that the poor immunogenicity of MUC1 stems not from the inability of MHC molecules to bind glycopeptides, but rather, from a reduction in the interaction between TCRs and MHC-glycopeptide complexes caused by carbohydrate presentation [36].

Furthermore, lymph node lymphocytes isolated from colorectal carcinoma patients and stimulated with MUC1 transfected B-cells in vitro produced CTLs that lysed MUC1 expressing targets cells [56]. Significantly, target cell lysis was enhanced by inhibiting *O*-linked glycosylation with phenyl-*N*-acetyl-α-D-galactosaminide, and was blocked completely by co-incubation with anti-T monoclonal antibodies [56]. Lysis was similarly titrated by addition of a Galβ1-3GalNAc disaccharide [56], demonstrating in vitro at least, that CTL recognition of mucin-based carbohydrate antigens can contribute to the clearance of MUC1 expressing cells.

13 Processing of Mucin-Based Vaccines

The apparent dichotomy in antigen presentation pathways (❯ *Fig. 6*) has important implications for mucin-based vaccines and serves to highlight the challenges faced in vaccine design. As vaccines are essentially exogenous antigenic molecules injected into the bloodstream, any resulting immune response is likely to proceed through the "classical" pathway of presentation by MHC class II molecules to $CD4^+$ T_H cells. This pathway excludes the presentation of vaccine-derived antigens by MHC class I molecules and hence, the production of antigen-specific immunoglobulins is frequently observed with activated CTLs being generally excluded from the immune response.

However, certain subsets of DCs can prime $CD8^+$ T-cell responses against exogenous antigens presented on MHC class I molecules in a process called cross-presentation [57] ❯ *Fig. 6*, *J*). Cross-presentation of exogenous antigens to $CD8^+$ T-cells is well documented in viral immunity (reviewed in [58]), and a more detailed understanding of the molecular events that transpire during antigen processing may enable specific features of synthetic mucin-based vaccines to be modified to facilitate the processing of exogenous mucin-based antigens by this pathway, thus increasing the chances of generating cancer associated mucin-specific CTLs to sites of cancer development in vivo.

Interestingly, several reports demonstrate that direct recognition of MUC1 peptide sequences by cognate TCRs can occur without a requirement for proteolytic processing of antigen or presentation by MHC molecules [59,60,61,62]. It is generally believed that the highly ordered, repetitive rod-like structure of the MUC1 VNTR domain enables direct TCR recognition

by virtue of multiple antigenic motifs that allow numerous TCR interactions to occur, thus bypassing the requirement for MHC-based stabilization of a limited number of TCR-antigen complexes. Generation of so-called "MHC-unrestricted" MUC1-specific CTLs that eliminate MUC1 expressing myeloma, breast and pancreatic cancer cells has been demonstrated in vitro [59,60,61,62]. Significantly, MHC-unrestricted lysis of MUC1 expressing cancer cells was abrogated in vitro by the addition of anti-MUC1 peptide antibody [60], suggesting that aberrant glycosylation arising from the cancerous state unmasked previously shielded antigenic peptide sequences that subsequently allowed this distinct and unusual form of direct recognition to occur. Although encouraging, it must be noted that MHC-unrestricted recognition of cancer-associated mucins by CTLs is uncommon, and has yet to be observed in response to the administration of a mucin-based vaccine.

Although the immune system functions in part to destroy deleterious cells as they arise in the body, cancers still develop in immuno-competent hosts. Cancers that over-express hypo-glycosylated mucins generally fail to induce potent immune responses due to a lack of intrinsic immunogenicity, the emergence of tolerance and the avoidance of immunological detection as discussed above. However, despite the prevalence of poorly immunogenic, aberrantly glycosylated mucins in many cancers, mucin-based immune responses in cancer patients have been observed [63,64], albeit infrequently. Thus, there is currently a great demand for mucin-based vaccine technologies that augment immune responses against hypo-glycosylated TAAs to facilitate tumor clearance. The synthetic methodologies commonly used to produce such vaccines are summarized below.

14 Chemical and Enzymatic Synthesis of Mucin-Based Vaccines

As discussed previously the heterogeneous nature of cancer cell-derived mucins (containing both normal and tumor-associated saccharides) generally makes them poor vaccine candidates and as a result chemists, for the most part, have provided pure samples of defined glycopeptides for structural and biological evaluation. Since the pioneering syntheses of Kunz [65], Lemieux [66], Meldal [67,68,69,70], Paulsen [71], and Schmidt [72], the synthesis of glycopeptides displaying "simple" *O*-linked saccharides such as the T_N and T antigens has become largely routine and is discussed in detail in various recent reviews [73,74,75,76,77,78]. Indeed simply glycosylated amino acids (e. g. GalNAc Ser/Thr) in suitably protected form for use in solid phase peptide synthesis (SPPS) are commercially available, and incorporation of simple saccharides into large proteins such as mucins [79,80] and mucin-like proteins [81] has been described. Consequently, rather than focusing on the construction of specific glycopeptide TAA's, this section highlights notable syntheses of these antigens (*Fig. 5*) and their union with mucin-derived peptide fragments resulting in homogeneous glycopeptide and glycopeptide-conjugate vaccine candidates poised to break immune tolerance. Specific examples will only be employed to emphasize an overall strategy.

While the assembly of oligopeptides and oligonucleotides is routine and has become highly automated, the automated synthesis of oligosaccharides is still in its infancy despite some notable advances [82,83,84]. The situation becomes even more complex when the protecting group requirements for both the oligosaccharide assembly, and peptide assembly combine in the same molecule to stretch this aspect of chemical synthesis to its limits. Conse-

quently most glycopeptides are prepared using a combination of SPPS and solution-based techniques and often not entirely in automated fashion because close monitoring of coupling reactions is usually desired in order to reduce the quantity of precious glycosylated building blocks required. The oligosaccharides and glycosylated building blocks are often prepared using standard solution-based techniques. In the past researchers have attempted to simplify the synthetic requirement by utilizing glycopeptide mimics that may be linked to a peptide

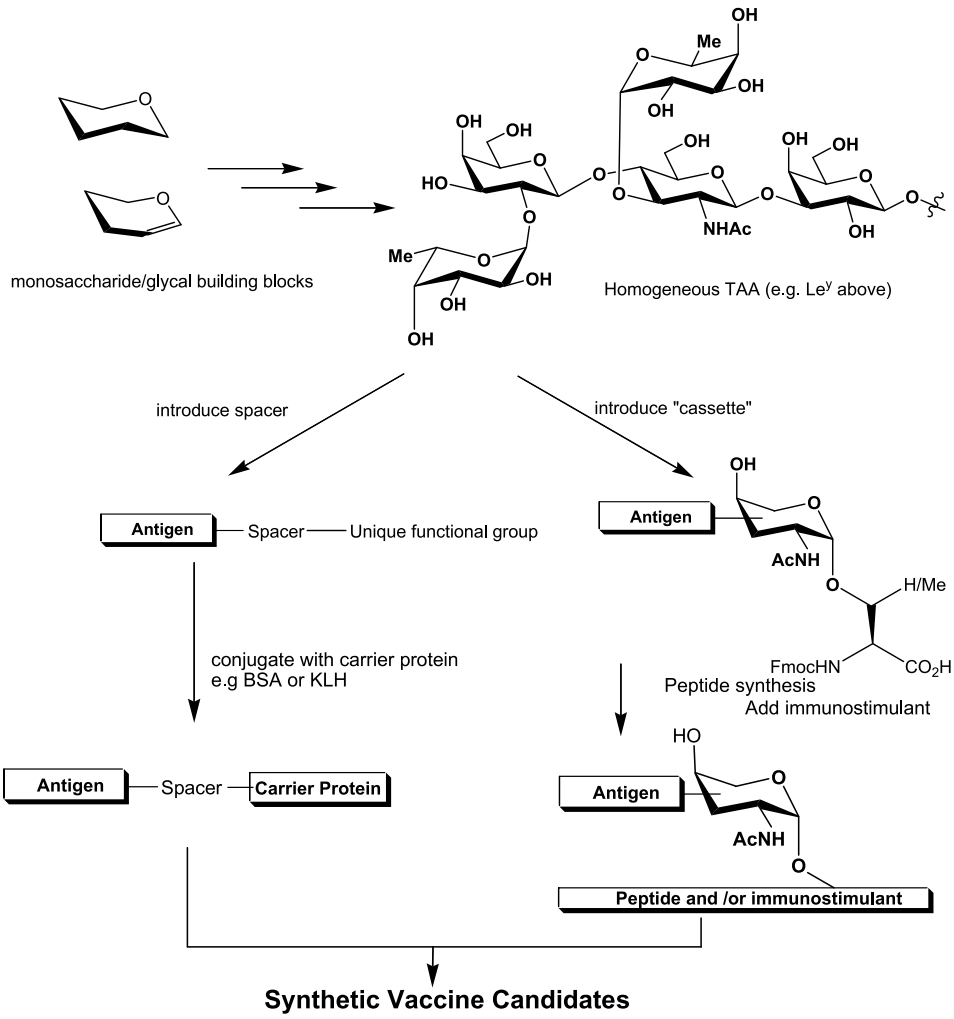


Figure 10

Common strategies for the assembly of tumor-associated antigen and glycopeptide vaccines. Synthetic TAAs and glycopeptide mimics displaying TAAs are either linked to large immunogenic carrier proteins that facilitate the display of hundreds of copies of antigen simultaneously, or alternatively glycopeptides may be adorned with additional functionality to include immunostimulants such as T-cell epitopes or inflammatory lipids

backbone of a carrier protein via an, often unnatural, linker or spacer motif (◉ *Fig. 10*). This can be particularly advantageous where the display of multiple copies of antigenic saccharides is required to arouse the strongest immunological response [85,86,87,88,89,90,91], and is also believed to more appropriately represent the multi-valent, clustered, environment in which tumor associated saccharides are displayed at the cell surface [92,93]. Additionally, it is considered to be particularly useful when the carrier protein is also considerably antigenic, such as Bovine Serum Albumin (BSA) or KLH which, in itself, stimulates the immune system, possibly through proteolytic processing and display on MHC class II, facilitating the recruitment of $CD4^+$ T-cell help. KLH can display hundreds of copies of oligosaccharide or glycopeptide antigen per protein molecule. Vaccines of this type will be discussed in more detail in ◉ *Sect. 15* The caveat in this approach is that elements of any unnatural linker motif, acting alone or in concert with the saccharide appendage, could give rise to the "incorrect" immunological response [89,94] and although widely practiced (with considerable success [95]) there is a desire to prepare homogeneous samples of oligosaccharides that are linked to native peptide sequences e. g. of MUC1 via regular α-*O*-glycosidic linkages to serine and threonine such that antigenic motifs which comprise elements of the saccharide and the underlying peptide backbone can be investigated. This makes sense because, often, TAAs are truncated saccharides arising from premature "capping" of core structures with sialic acid (◉ *Fig. 5a*) and the underlying peptide is believed to be more exposed. Obviously the use of unnatural spacers and linkers between the saccharide and peptide backbone precludes this type of investigation though are still useful for investigations of "saccharide-only" antigens, especially where the attached oligosaccharide is sufficiently sterically demanding so as to conceal any underlying non-natural structural elements. There are two general modes by which the synthesis of naturally occurring mucin-type glycopeptides might be envisaged. The first is the convergent union of pre-assembled saccharides with the appropriate peptides. Whilst this is commonplace in the assembly of *N*-linked glycopeptides it is far less common in *O*-linked glycopeptide synthesis due to the difficulties in controlling the stereochemistry of the new glycosidic linkage and the compatibility between glycosidation reactions (requiring strictly anhydrous conditions) and the solubility of peptides. Instead a "building block" or "cassette" approach is adopted where fully protected glycoamino acids are prepared and then introduced into regular SPPS. The added bulk of the glycoamino acid has ramifications for glycopeptide synthesis and the most sterically demanding glycoamino acids can give rise to low coupling yields in solid-phase glycopeptide assembly. However, it is possible that simple glycoamino acids can be introduced and then modified by using appropriate enzymes, provided that the appropriate enzymes are available [96]. These methods are shown schematically in ◉ *Fig. 10* and specific examples will be discussed below in more detail.

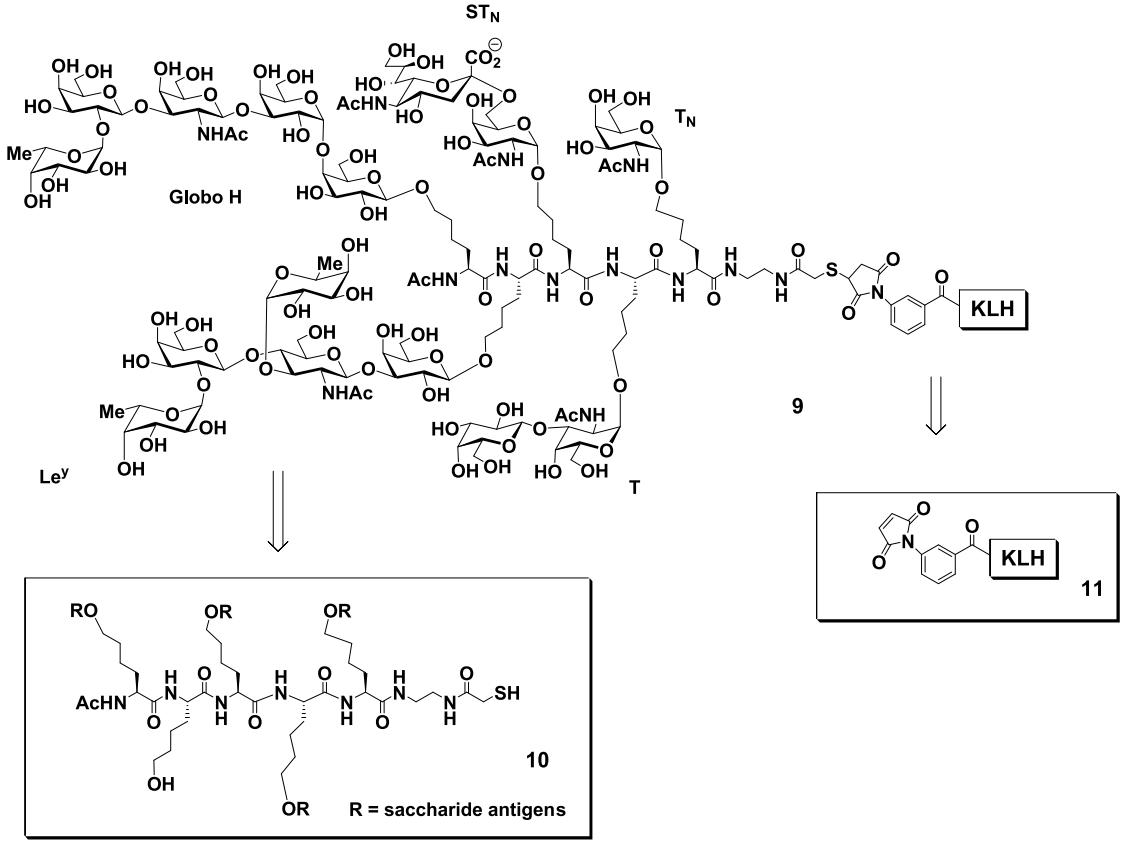
15 Multivalent Antigen Display on Immuno-Stimulatory Carrier Proteins: Eliciting an Antibody Response

Oligosaccharides often interact strongly with their receptors in a multi-valent fashion hence the term avidity, (arising from the sum of many individual interactions) rather than affinity (which

is more commonly employed to describe the interaction between a specific substrate-ligand pair) that is used to describe saccharide-protein interactions [92,93]. Following on from this it is known that numerous copies of specific cell surface oligosaccharides (or "glycans") direct fundamental processes such as cell-cell adhesion in this fashion since interactions between single synthetic or isolated oligosaccharides and their ligands are relatively weak (mM–μM range). Therefore in the development of therapeutic anti-cancer vaccines it has been essential to develop multi-valently displayed samples to more accurately replicate the situation at the surface of a cancer cell. Immuno-stimulatory proteins and short non-immunogenic linker motifs have been used for this purpose for decades since the chemistry is relatively well developed [97] and most often employs simple, non-specific, reductive amination type-reactions between antigen-displayed aldehyde functional groups and the ε-amino groups of lysine residues contained within, for example, KLH. In this way many problems associated with glycopeptide synthesis (such as the increased demand for protecting groups) are circumvented allowing the work to be conducted in non-specialist laboratories. Historically, the latent aldehyde functionality at the reducing end of isolated and synthetic oligosaccharides has been utilized in this way for conjugation to proteins but is less frequently employed nowadays as it irreversibly destroys the cyclic structure of the terminal sugar at the reducing end.

In a recent and impressive study the Danishefsky laboratory prepared a unimolecular pentavalent vaccine, displaying five different tumor-associated antigens (GloboH, Le^y, sT_N, T_N and T antigens) concomitantly, on a short peptide scaffold comprised mainly of an unnatural amino acid (ε-hydroxynorleucine) and subsequently conjugated to KLH (**9**, ➤ *Fig. 11a*) [98].

This approach is designed to elicit a multi-faceted response to several tumor-associated antigens simultaneously. Interestingly the choice of unnatural amino acid arises more from the shortcomings in chemical synthesis encountered along the way, rather than entirely by design. However its increased flexibility compared to the native α-*O*-glycopeptide linkage to serine and threonine undoubtedly facilitates chemical synthesis by relieving steric crowding in the clustered glycopeptide sequence. Alternative synthetic vaccines prepared using dendrimeric non-immunogenic lysine core-based multiple antigenic glycopeptides (MAG's) have also benefited from the use of similarly flexible homoserine-based glycoamino acid building blocks [87]. Initially the Danishefsky group had investigated the construction of allyl glycosides of antigens since they could be converted to aldehydes upon treatment with ozone. However, since the allyl glycoside was not fully compatible with the global reductive debenzylation of the pre-assembled saccharide at the end of the synthesis [99] an n-pentenyl glycoside (a well-known glycosyl donor [100]) was employed instead. Although more stable to the reductive debenzylation, the n-pentenyl glycosides, when employed as glycosyl donors, failed to provide access to the natural glycoamino acid (GalNAc-Ser/Thr) cassettes directly so an ozonolysis/Horner–Emmons/asymmetric hydrogenation sequence was ultimately preferred for the preparation of antigens containing the native GalNAc-α-*O*-glycosidic linkage. These neoglycoamino acid building blocks could be readily employed in SPPS [85]. A key step in the vaccine assembly was the conjugation between the neoglycopeptide (**10**) and maleimide modified KLH (**11**). To facilitate this, the neoglycopeptide was equipped with a diaminopropyl spacer that was subsequently linked to a 2-mercoptoacetate moiety employing the pentafluorophenyl (pfp) active ester. Unmasking of the thiol functionality (with NaOH) allowed reaction with maleimide-activated-KLH (**11**). This approach appears generally useful with GalNAc-α-*O*-ε-hydroxynorleucine derivatives, containing the native α-*O*-glycosidic linkage such as the



a

Figure 11

(a) A unimolecular penta-valent glycopeptide vaccine 9assembled by ligation between neoglycopeptide-thiol 10 and maleimide modified KLH 11. The component parts are readily prepared fromglycal precursors

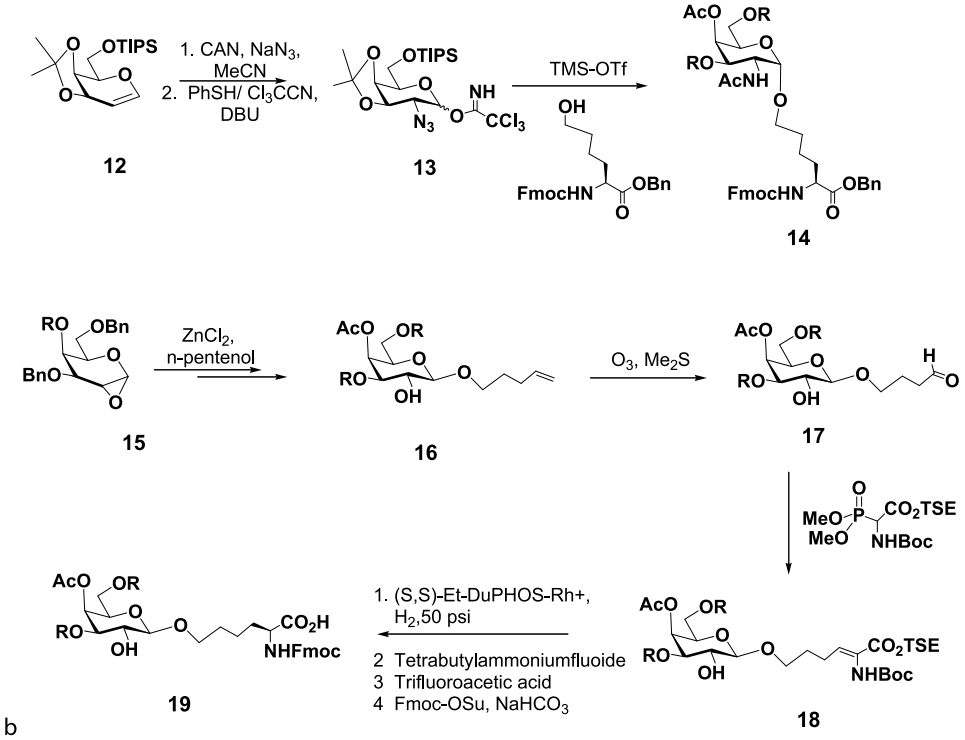


Figure 11 (continued)
(b) A Glycal 12 can be converted to a suitable glycosyl donor (13, X = leaving group) following Lemieuxs' azidonitration. Glycosidation with suitably protected ε-hydroxynorleucine, facilitated by the 2-azido group on the sugar, affords the α-*O*-linked tumor-associated antigens. For β-*O*-linked products the terminal glycal epoxide (e. g. 15) is opened with-n-pentenol. The n-pentenyl glycoside 16 provides access to the glycoamino acid cassette 19 following ozone-mediated double bond cleavage, Horner–Emmons reaction with a suitable phosphonate reagent and Rhodium-catalyzed asymmetric reduction of the carbon–carbon double bond

T_N and T antigens (**1** and **2**, ❯ *Fig. 5a*) and sialylated analogues being readily prepared by glycosylation of the activated 2-azido-2-deoxy sugars (**13**, ❯ *Fig. 11b*). Antigens containing β-*O*-glycosidic linkages are also readily prepared through solvolysis of the appropriate glycal epoxides [101]. All carbohydrate components are prepared from readily available glycals, emphasizing the power of glycal assembly methodology [102]. More recently [103] preliminary immunological findings with this class of KLH-linked antigen has been disclosed and the findings are discussed in more detail in the following section.

In another recent example Kunz and co-workers assembled glycopeptide vaccine **23** corresponding to fragments of the VNTR domain of MUC1 (❯ *Fig. 12*) modified with the sT_N antigen (**4**, ❯ *Fig. 5a*). Assembled glycopeptide **20** was displayed on BSA using diethylsquarate (3,4-diethoxy-3-cyclobutene-1,2-dione, **21**). The preformed sqauric monoamide **22** was conjugated to BSA in approximately 60% yield after 24 h [104].

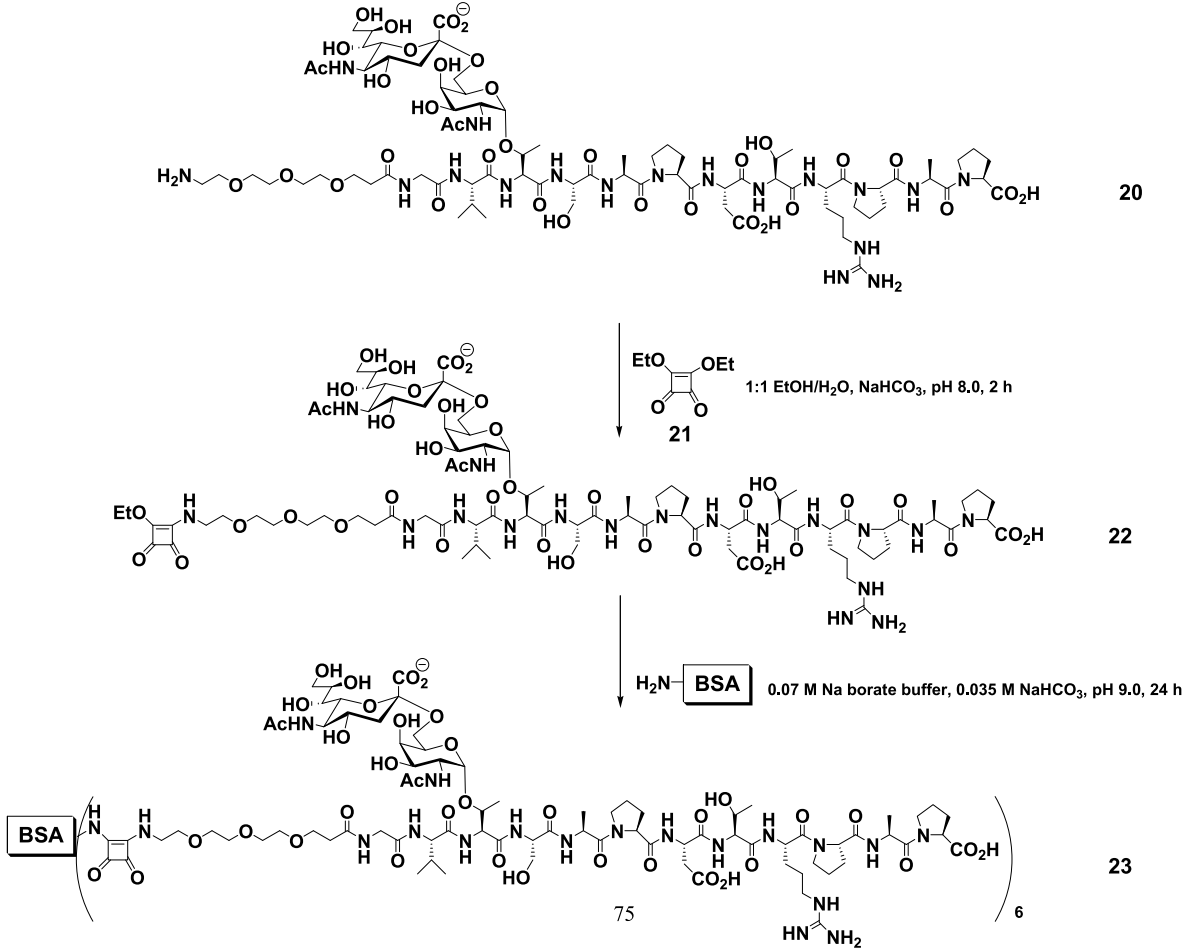


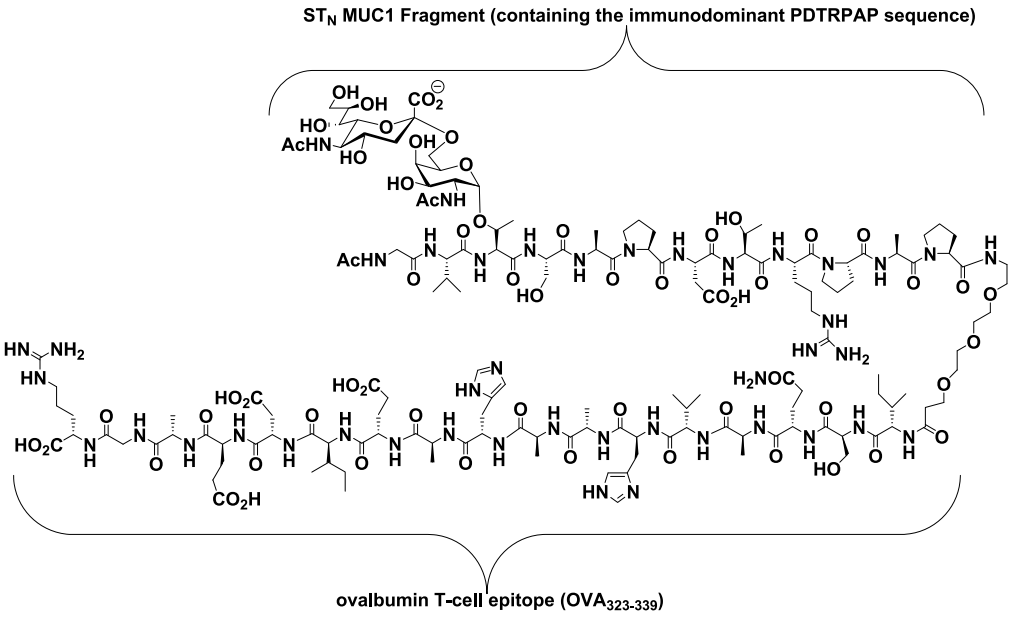
Figure 12
Synthesis of sT_N containing MUC1 fragment. sT_N was introduced as the fully protected glycoamino acid

16 Synthetic Glycopeptides Designed to Elicit a CD4$^+$ T-Helper (T_H) Response

A major goal in vaccine design is to enlist strong T-cell help from CD4$^+$ T_H-cells to activate naïve B-lymphocytes and establish immunological memory such that tumor cells can be cleared if they attempt to re-establish themselves in a host after vaccination. In an attempt to invoke a T_H-cell response Kunz and co-workers describe a strategy where various TAAs are displayed on fragments (usually containing the immunodominant peptide motif: PDTRPAP, from the VNTR domain) of MUC1. The MUC1 fragment is linked via a non-immunogenic spacer to a known T-cell epitope (i. e. is known to be presented by MHC molecules). In this example (*Fig. 13*) the epitope is a fragment of ovalbumin ($OVA_{323-229}$) [105] though T-cell epitopes from Tetanus toxin have also been employed in the past [106].

It can be envisaged that exposing B-cells (in peripheral blood lymphocytes, or whole organisms!) to these glycopeptides may result in internalization of the immunogenic saccharide antigen after recognition by cell-surface bound immunoglobulins. After proteolytic processing, the T-cell epitope can then be displayed on MHC class II to T_H-cells, facilitating the release of sufficient co-immunostimulatory cytokines to allow B-cell maturation and antibody secretion. In this case the vaccines were used to immunize transgenic mice whose T-cells over-express a receptor specific for the ovalbumin sequence. The synthesis of such vaccines follows a similar route regardless of the antigen being presented and begins with the assembly of a suitably protected T_N displaying glycoamino acid (*Fig. 14a*) [69,70]. This protected glycoamino acid is commercially available though can be prepared on gram scale in few steps from readily availably starting materials. In the literature various glycosyl donors (glycosyl halides, thioglycosides, trichloroacetimidates) have all been effective in reactions with suitably protected serine and threonine residues. The versatility of this building block has been demonstrated in many glycopeptide syntheses and can be readily elaborated to various tumor associated glycopeptide carbohydrate antigens. *Figure 14b* highlights the synthesis of cassettes **27**, **28** and **31** from **24** by Kunz and co-workers, corresponding to the sT_N (**4**), ST (**5**), T (**2**) and T_N (**1**) antigens respectively [23]. Following assembly of the sT_N cassette it was incorporated into a MUC1 fragment. The peptide was cleaved from the resin employing a fluoride cleavable linker which allows the glycopeptide *C*-terminal acid, to be obtained in an otherwise fully protected form. This peptide was coupled to the pre-assembled ovalbumin sequence via a triethyleneglycol spacer and the fully deprotected vaccine was obtained in 42% yield, based on loading of the first amino acid of the ovalbumin fragment, after preparative HPLC purification.

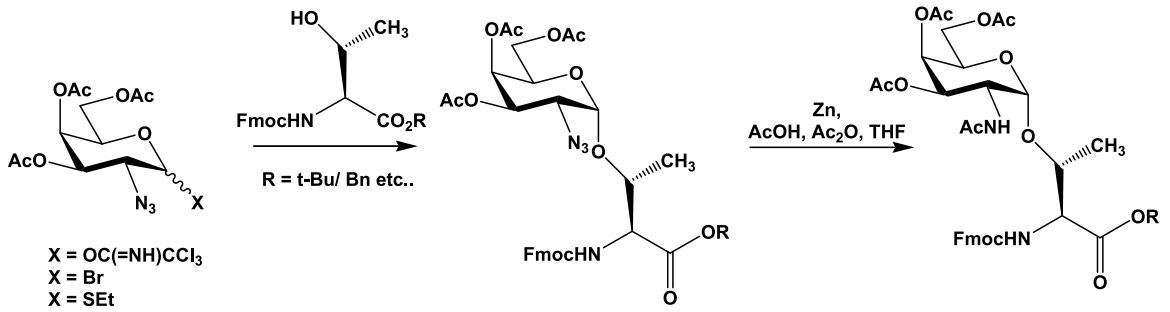
Boons and co-workers envisaged a similar strategy for the construction of a vaccine containing the T_N antigen (**32**, *Fig. 15*) [107]. Their vaccine contained the threonine-linked T_N antigen, conjugated to a T-cell epitope (comprised of a 20mer "YAF" peptide derived from the outer membrane protein of *Neisseria meningitides*) via a short 1,3-diaminopropyl linker. Furthermore the vaccine construct contained an S-[(R)-2,3-dipalmitoyloxy-propyl]-*N*-palmitoyl-(R)-cysteine (Pam_3Cys) motif that is known to function as an adjuvant, stimulating the secretion of pro-inflammatory cytokines that stimulates antigen presenting cells and the development of T_H-cells. This vaccine was prepared by coupling the three pre-assembled components. First the T-cell epitope was prepared using standard SPPS and the lipid modified cysteine residue **33**



76

Figure 13

Glycopeptide vaccine recently described by Kunz and co-workers comprising the immunodominant (PDTRP) sequence of the VNTR domain of MUC1, adorned with the sT$_N$ antigen and connected to a T-cell epitope (OVA$_{323-339}$) known to be presented to MHC class II molecules



a

Figure 14

(a) Glycoamino acid building blocksor "cassettes" such as 24 are readily prepared via glycosidation of suitably protected serine and threonine residues with common glycosylhalide trichloroacetimidate or thioglycoside donors. Following reduction of the azido group (which helps separation of α and β anomers) andcarboxyl deprotection the building blocks are ready for solid phase peptide synthesis or the T_N antigen 24 can be elaborated tovarious glycoamino acids

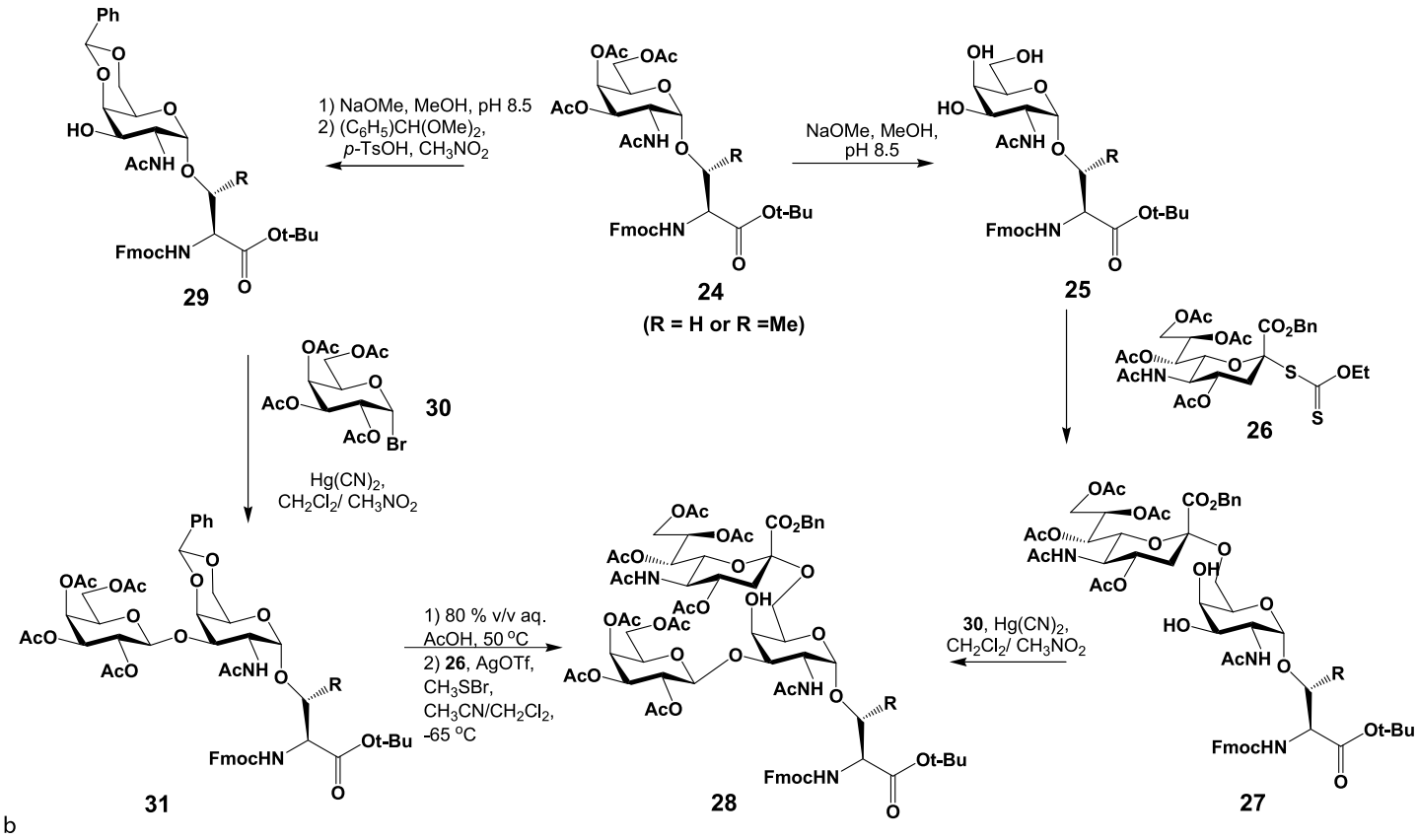


Figure 14 (continued)
(b) Synthesis of cassettes 27, 28 and 31 from 24 by Kunz and co-workers, corresponding to the sT_N (4), ST (5), T (2) and T_N (1) antigens, respectively [23]

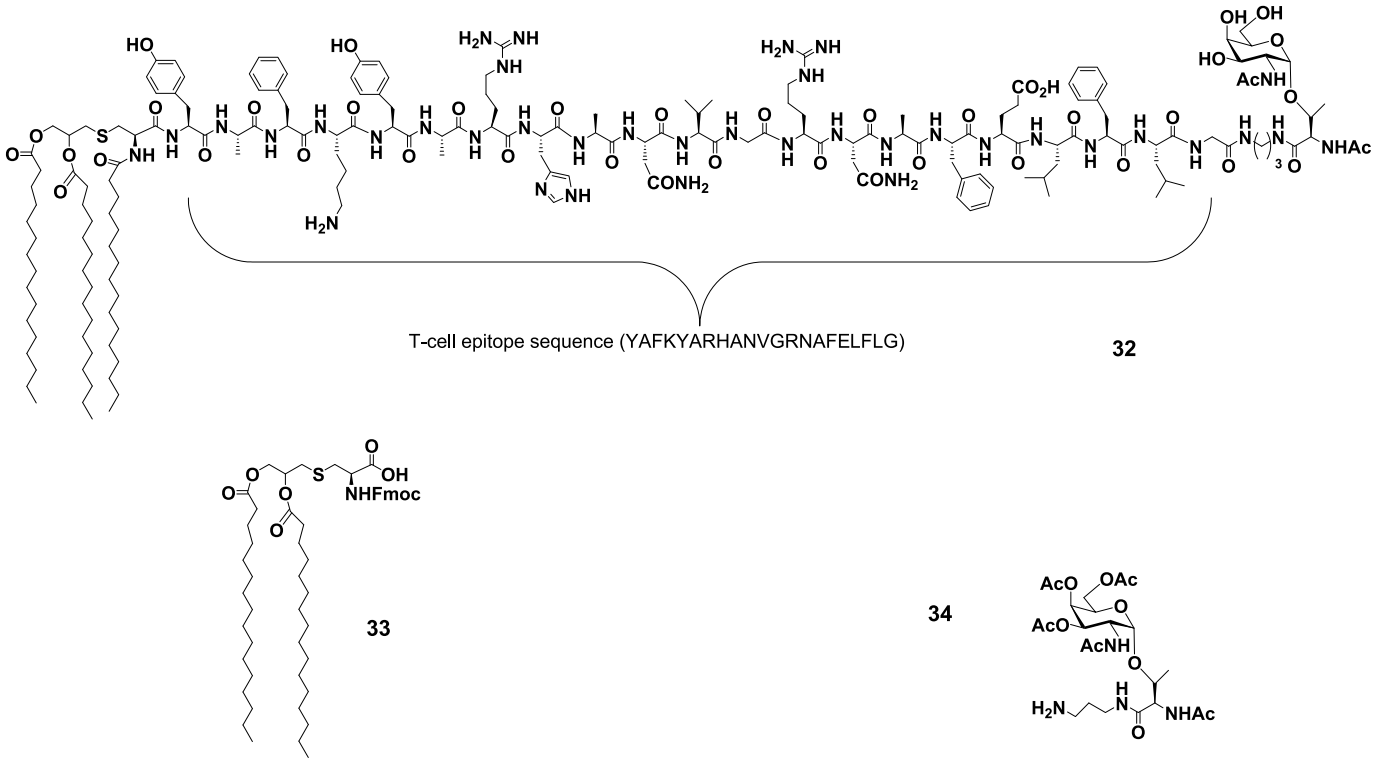


Figure 15

Synthetic glycopeptide vaccine (32) described by Boons and co-workers is assembled by first coupling the pre-assembled T-cell epitope to the lipidated cysteine residue 33 on solid phase. Following addition of the final palmitoyl group the lipopeptide is cleaved from solid support to present only the *C*-terminal acid to the T_N cassette 34. Glycolipopeptide 32 is finally obtained after protecting group removal and purification [107]

was finally coupled. Following adjustment of the cysteine protecting group and degree of lipidation the vaccine precursor was cleaved from the solid phase, as before to free only the *C*-terminal acid, though employing a different linker (HMPB-MBHA resin). Finally, the T_N cassette **34** was coupled and the vaccine was obtained after removal of protecting groups, size exclusion and reverse-phase HPLC. The Danishefsky group have also employed the use of the Pam_3Cys motif in various vaccine constructs (though not in concert with a T-cell epitope) though have favored more the KLH-conjugated constructs as a result of their higher observed IgG titers [74].

17 Enzymatic and Chemo-Enzymatic Synthesis of Mucin-Based Vaccine Candidates

Enzymatic assembly of tumor associated glycopeptide antigens is highly desirable since the obstacles that are characteristic of carbohydrate chemistry, the requirement for many protecting group manipulations and stereoselective glycosidation, are overcome due to the exquisite regio- and stereoselectivity of enzymes. Additionally the transformations can be conducted in aqueous solution. The drawbacks are not so much the cost of the required glycosylnucleotide donors (UDP-Gal, UDP-GalNAc, CMP-NeuAc, UDP-GlcNAc) since substrate recycling protocols are well developed in many instances, but more the general availability of suitable glycosyltransferases. Many glycosyltransferases are not commercially available. However recent reports have shown that ppGalNAcT enzymes, that transfer the first GalNAc residues to "naked" mucins, when working either alone or in concert with additional transferases can produce mucin associated antigens suitable for immunological study [108,109]. In one study, Freire et al. showed that MUC6 fragments from the tandem repeat domain could be over-expressed in *E. coli*. Interestingly MUC6 is over-expressed in MCF-7 breast cancer cell lines but its expression is normally restricted to gastric tissues of healthy individuals. The purified, unglycosylated, fragments obtained were then glycosylated with recombinant bovine ppGalNAc T I and the products, adorned with the T_N antigen, could be prepared in quantities of up to 3 mg. Peptides were also glycosylated with MCF-7 breast cancer cell extracts and could be similarly glycosylated (i. e. with T_N). The degree of glycosylation of MUC6 peptides was greater when using the breast cancer cell extract (T_N occupying 54–58% of the potential *O*-glycosylation sites) than with the recombinant ppGalNAcT I (T_N occupying 42% of the potential *O*-glycosylation sites). However, the degree of peptide glycosylation by bovine ppGalNAcT I depended significantly on the size of the MUC6 peptide fragment, 42% (as stated above) for glycosylation of a peptide corresponding to half a tandem repeat yet up to 64% glycosylation site occupancy for the whole tandem repeat. These peptide fragments were used to immunize mice and the results are described in more detail in the following section. In a further study [108], a synthetic peptide comprising 60 amino acids (three MUC1 tandem repeats) was fully glycosylated (five GalNAc residues per tandem repeat) employing recombinant ppGalNAcT enzymes (*Fig. 16*). Furthermore the T_N antigen was elaborated to T and sT_N antigens by the action of recombinant β-1,3-GalT and α-2,6-NeuAcT (a.k.a. ST6GalNAc I) respectively. The reactions were simply conducted in aqueous buffered solutions and monitored by reversed-phase HPLC and MALDI-TOF mass spectrometry. As enzymes become more avail-

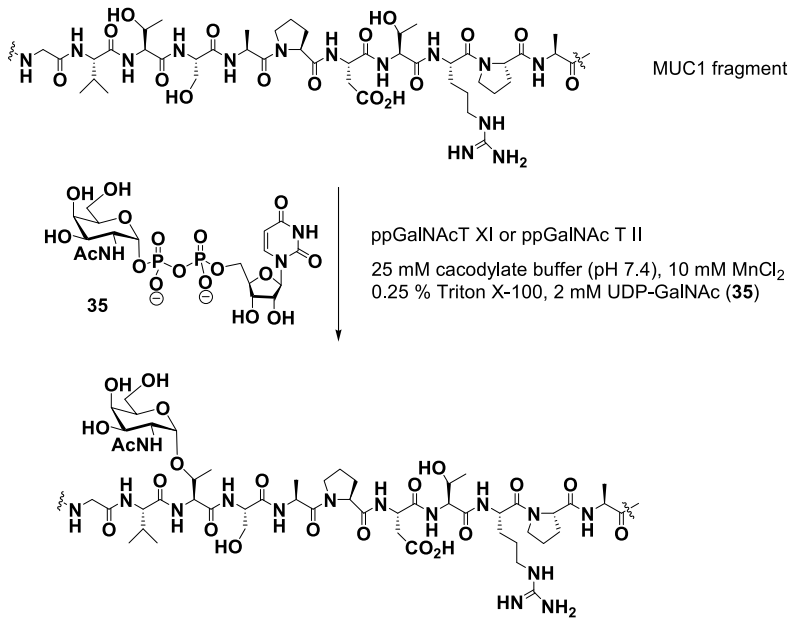


Figure 16
Peptide *O*-glycosylation employing glycosyltransferases utilizes the activated glycosyl nucleotide donor UDP-*N*-acetylgalactosamine (35) to modify mucin fragments with complete regio and stereocontrol in the absence of protecting groups

able the assembly of mucin based vaccines employing this approach, possibly in concert with chemical protein synthesis, is sure to become more commonplace.

For over a decade mucin-type glycopeptides have been amenable to automated synthesis employing the cassettes approach [110], and more recently many tumor associated glycopeptide antigens have been prepared using semi-synthetic methods in combination with glycopeptide substrate immobilization on water soluble polymer supports [111]. Along with technological advances in glycopeptide synthesis e. g. microwave assisted peptide synthesis [112] it might also be possible to decrease the time taken (increase the yield of oligosaccharide and glycopeptide antigens) to discover new tumor-associated antigenic motifs, and to prepare pure homogeneous samples for immunological investigations. Some immunological findings employing the glycopeptides highlighted above are summarized in the final section of this review.

18 Immunological Evaluation of Mucin-Based Vaccines

Aberrantly glycosylated mucins over-expressed on the surface of cancer cells reveal peptide sequences and prematurely sialylated carbohydrate moieties that would otherwise be inaccessible or absent in the healthy uncancerous state. Consequently, this provides a source of peptide and glycopeptide-based, mucin-associated antigens from which synthetic vaccines

can be constructed. The majority of synthesized vaccines are accordingly based on the solvent exposed peptide sequences and cancer-specific mucin-based antigens within the VNTR region, although additional sequences outwith the VNTR region have also yielded encouraging results. The remainder of this section seeks to evaluate the impact of such vaccines in preclinical murine studies and in clinical trials. A summary of the experimental results obtained from these investigations is also presented (● *Table 3*).

19 Unglycosylated Mucin Peptide Vaccines

The process of cellular transformation adversely affects the molecular machinery involved in the post-translational modification of mucin peptides, such that glycosylation is prematurely terminated by the addition of sialic acid to the developing carbohydrate chains. Consequently the mucin peptide backbone is often under-glycosylated, and core peptide sequences that would otherwise be inaccessible are exposed in the disease state. Many unglycosylated, mucin-based peptide vaccines are therefore based on these exposed sequences, as it is the accessibility to these motifs that differentiates between healthy and cancerous cells in vivo. Preclinical evaluation of these peptide vaccine formulations relies heavily upon the use of murine models and immortalized cancer cell lines. However, human MUC1 is highly immunogenic in wild-type mice. Introduction of human MUC1 into wild type murine models is therefore not immunologically representative of the human situation, where an absence of an immune response is frequently observed due to immunological tolerance to cancer-associated MUC1 antigens. For this reason, transgenic (Tg) mouse models were developed, where genetic engineering allows the human MUC1 glycoprotein to be expressed as a "self" molecule in mice from birth. In such models, immunological tolerance to "self" (the human MUC1 glycoprotein) and "altered-self" (human MUC1 decorated with prematurely sialylated cancer-associated antigens) is established from the outset, and thus provides a more representative immunological model in which to assess the efficacy of vaccines designed to break tolerance to human, cancer-associated mucin antigens.

The production of CTL precursors and extremely low titers of MUC1-specific antibodies was observed following vaccination of mice transgenic for human MUC1 (MUC1.Tg), with an unglycosylated MUC1 peptide (5 × PAHGVTSAPDTRPAPGSTAP: see [113] ● *Table 3*)-glutathione-*S*-transferase fusion protein (MUC1-GST) or unglycosylated MUC1-GST conjugated to oxidized mannan [113]. The use of oxidized mannan in these vaccine formulations acts as an immunological adjuvant, and is observed to favor the induction of cell-mediated, rather than antibody-based immune responses against the conjugated molecule [114]. Using an alternate strategy, MUC1.Tg mice were vaccinated with *Vaccinia* virus engineered to express both human MUC1 and the cytokine IL-2. In this system, MUC1 expression occurs through a virus-mediated introduction of mucin DNA into the recipient cell. As a result, production of human MUC1 is orchestrated by the host (murine) translational machinery, and is therefore endogenous. Consequently, the processed MUC1 antigenic fragments are presented on the surface of murine cells by MHC class I molecules. The generation of CTL precursor cells was also observed using this system [113]. Importantly, whilst the production of a fully blown anti-MUC1 immunological response was not observed, these results demonstrate that the stim-

Table 3
Summary of peptide and glycopeptide sequences used in mucin-based vaccine formulations

Vaccine Type	Peptide Sequence(s)[a]	Glyco-Antigen(s)	Carrier Molecule	Immunological Adjuvant	Model System	Cancer	Antibody Response	CTL Response	Reference
PV	5 × PAHGVTSAPD TRPAPGSTAP-GST	None	None	IL-2	Tg Mice	None	Minimal	↑ CTLp	[113]
PV	5 × PAHGVTSAPD TRPAPGSTAP-GST	None	Ox. Man	IL-2	Tg Mice	None	Minimal	↑ CTLp	[113]
PV	1 × VTSAPDTRPA PGSTAPPAHG VTSAPDTRPA	None	KLH	QS-21	Human Clinical	Breast	IgM, IgG1, IgG2, IgG3	ND	[115]
PV	7 × GVTSAPDTRP APGSTAPPAH	None	None	GM-CSF	WT Mice	T-Cell lymphoma	IgG1, IgG2b	ND	[123]
PV	7 × GVTSAPDTRP APGSTAPPAH	None	None	SB-AS2	WT Mice	T-Cell lymphoma	IgG1, IgG2b, IgG3	ND	[123]
PV	7 × GVTSAPDTRP APGSTAPPAH	None	None	GM-CSF	Tg Mice	T-Cell lymphoma	IgM	ND	[123]
PV	7 × GVTSAPDTRP APGSTAPPAH	None	None	SB-AS2	Tg Mice	T-Cell lymphoma	IgG1, IgG2b, IgG3	ND	[123]
PV	1 × ATWGQDVTSV	None	None	None	Human in vitro	Pancreatic	NT	T-Cell activation	[118]
PV	1 × ALWGQDVTSV	None	None	None	Human in vitro	Pancreatic	NT	T-Cell activation	[118]
PV	5 × GVTSAPDTRP APGSTAPPAH	None	None	SB-AS2	Phase I Clinical	Pancreatic	IgM, IgG	YES	[116]
PV	3 × PDTRPAPGST APPAHGVTSA	None	Ox. Man	None	Phase III Clinical	Breast	IgM, IgG	T-Cell activation	[117]

Table 3 (continued)

Vaccine Type	Peptide Sequence(s)[a]	Glyco-Antigen(s)	Carrier Molecule	Immunological Adjuvant	Model System	Cancer	Antibody Response	CTL Response	Reference
GPV	None	T_N (c)	KLH	QS-21	Phase I Clinical	Prostate	IgM, IgG1, IgG3	NT	[119]
GPV	None	T_N (c)	Palmitic acid	QS-21	Phase I Clinical	Prostate	IgM, IgG	NT	[119]
GPV	1 × YAFKYARHAN VGRNAFELFL-(spacer)-[T]	T_N	None	Pam_3Cys	WT Mice	None	IgM, IgG	NT	[107]
GPV	1 × YAFKYARHAN VGRNAFELFL-(spacer)-[T]	T_N	None	Pam_3Cys + QS-21	WT Mice	None	IgM, IgG	NT	[107]
GPV	1 × GV[T]SAPDTRP AP	ST_N	$OVA_{323-339}$	CFA	OVA Tg Mice	None	IgG	NT	[105]
GPV	1 × P[TTT]PI[TTTTT]V [T]P[T]P[T]P[T]G[T]Q[T]P[TTT]PI[STT]C	$T_{2/4}$ GalNAc transferase	KLH	QS-21	Human	Prostate	IgM, IgG1, IgG3	NT	[55]
GPV	1 × P[TTT]PI[TTTTT]V [T]P[T]P[T]P[T]G[T]Q[T]P[TTT]PI[STT]C	$T_{2/4}$ GalNAc transferase	KLH	GPI-0100	Human	Prostate	IgM, IgG1, IgG3	NT	[55]
GPV	5 × hydroxy norleucine	T T_N ST_N Le^Y globoH	KLH	QS-21	Mice	None	IgM, IgG IgM IgG None IgM, IgG	NT	[103]
GPV	5 × hydroxy norleucine	T T_N ST_N Le^Y globoH	Pam_3Cys	QS-21	Mice	None	IgG IgM None IgM IgM	NT	[103]
GPV	3 × V[TS]APD[T]RPA PG[ST]APPAHG	T_N	KLH	CFA	Tg Mice	None	IgM, IgG1	NT	[108]
GPV	3 × V[TS]APD[T]RPA PG[ST]APPAHG	ST_N	KLH	CFA	Tg Mice	None	IgM, IgG1, IgG2A, IgG2B	NT	[108]

Table 3 (continued)

Vaccine Type	Peptide Sequence(s)[a]	Glyco-Antigen(s)	Carrier Molecule	Immunological Adjuvant	Model System	Cancer	Antibody Response	CTL Response	Reference
AT	7 × GVTSAPDTRP APGSTAPPAH	None	None	None	WT Mice	T-Cell lymphoma	ND	YES	[123]
AT	7 × GVTSAPDTRP APGSTAPPAH	None	None	None	Tg Mice	T-Cell lymphoma	ND	YES	[123]
AT	1 × TRPAPGSTAP PAHGVTSAPD TRPAPGSTAP	None	None	IL-2	Human Clinical	Breast + Lung	NT	YES	[127]
AT	1 × AHGVTSA[3 × PDTRPAPGSTA PPAHGVTSA]P ESRPAPGSTAPA AHGVTASPESR PAPGSTAPPAV	None	Ox. Man	None	Phase II Clinical	Adeno-carcinoma	IgM, IgG	YES	[128]
AT	1 × STPPVHNV and LLLLTVLTV with PADRE	None	None	IL-2	Phase I Clinical	Renal Carcinoma	NT	T-Cell activation	[130]

PV = Peptide vaccine. GPV = Glycopeptide vaccine. AT = Adoptive transfer.

[a] The position of attatched glyco-antigens is indicated by highlighting the relevant amino acid residue in square brackets.

GST = Glutathione-S-transferase. Ox. Man = Oxidized mannan. CTL = Cytotoxic T-lymphocyte. CTLp = Cytotoxic T-lymphocyte precursors. ND = Not detected. NT = Not tested. $OVA_{323-339}$ = Ovalbumin peptide fragment 323-339. KLH = Keyhole limpet hemocyanin. T_N (c) = T_N antigen in trimeric cluster. IL-2 = Interleukin-2. GM-CSF = Granulocyte-macrophage colony-stimulating factor. SB-AS2 = An oil-in-water emulsion of 3-deacylated-monophosphoryl lipid A and QS-21. Pam_3Cys = S-[(R)-2,3-dipalmitoyloxy-propyl]-N-palmitoyl-(R)-cysteine. CFA = complete Freund's adjuvant. QS-21 = a water-soluble triterpene glucoside plant extract from the Soap Bark tree *Qillaja saponaria*. GPI-0100 = A semi-synthetic saponin based on QS-21. Ig = Immunoglobulin

ulation of an immune response against human MUC1 can proceed without the induction of autoimmunity against "self" antigens expressed in the transgenic model.

In a small clinical study, nine high-risk breast cancer patients were immunized with 1.5 unglycosylated synthetic MUC1 VNTR domains (VTSAPDTRPAPGSTAPPAHGVTSAPDTRPA: see [115] ❯ *Table 3*) conjugated to KLH and administered with the adjuvant QS-21 [115]. Analysis of peripheral blood after vaccination revealed elevated levels of MUC1-specific IgM and IgG that were sustained for up to 137 weeks following a single vaccination [115]. The observed switch in immunoglobulin isotype from antigen-specific IgM to IgG is importantly, only observed following the stimulation of B-cells by T_H-cells. The presence of antigen-specific IgG in post-vaccination blood plasma is therefore indicative of an activated $CD4^+$ T_H-cell response in vivo. IgM produced by seven of the nine patients in the study bound to MCF-7 tumor cells (expressing sT_N, T, GM2, globoH and Le^Y) in vitro, although none of the patients presented evidence of CTL activation [115]. Low but detectable increases in MUC-1 specific IgG responses were observed in a phase I clinical trial on 16 patients with resected or advanced pancreatic cancer that received a synthetic unglycosylated vaccine comprised of 100 amino acids from the MUC1 VNTR region (5 × GVTSAPDTRPAPGSTAPPAH: see [116] ❯ *Table 3*) and the immunological adjuvant SB-AS2 [116].

An unglycosylated peptide comprised of 60 amino acids from the VNTR region of human MUC1 (3 × PDTRPAPGSTAPPAHGVTSA: see [117] Table 3) was expressed and purified from bacterial culture and conjugated to oxidized mannan. The resulting vaccine was evaluated in a phase III clinical trial on 31 patients with early (stage II) breast cancer [117]. Patients received either placebo or vaccine with an observed cancer recurrence rate of 27% and 0% respectively after 5.5 years [117]. Of the patients treated with vaccine, 69% generated anti-MUC1 antibody responses and 40% generated a MUC1-specific T-cell response as evidenced by an increase in secretion of interferon-gamma (IFN-γ) in vitro [117].

In another study, a computational analysis of the human MUC1 amino acid sequence was performed to identify those regions most likely to interact with a polymorphic variant of MHC class I molecule [118]. Interestingly, several peptide sequences located outside of the VNTR region were identified as capable binding partners ($ATWGQDVTSV_{92-101}$, $ALLVLVCVLV_{1162-1171}$, $TISDVSVSDV_{1135-1144}$, $ALAIVYLIAL_{1172-1181}$, $VLVALAIVYL_{1169-1178}$ and $YLIALAVCQC_{1177-1186}$). The identified peptide sequences were constructed synthetically and all were observed to stimulate an increase in secretion of IFN-γ from T-cells, indicative of activation in vitro [118]. Furthermore, the most potent MHC binding peptide (ATWGQDVTSV) was modified to produce an agonist epitope (ALWGQDVTSV: see [118] ❯ *Table 3*) that exhibited improved characteristics as a MUC1 antigen when compared to the natural MUC1 peptide sequence identified in the computational screen [118]. The cytolytic efficiency of generated CTLs was quantified using an in vitro ^{111}In release assay, in which the amount of ^{111}In released from labeled target cells expressing MUC1-based antigens is proportional to lysis mediated by CTLs. Use of this assay importantly demonstrated that pulsing the agonist peptide into autologous DCs from patients with pancreatic cancer produced CTLs that lysed MUC1-expressing cancer cells more efficiently than CTLs generated from the natural MUC1 peptide sequence [118]. These results not only demonstrate that the immunogenicity of MUC1 is not restricted to the VNTR region of the molecule, but also that amino acid substitutions in the mucin peptide backbone can generate an immune response of increased potency against a naturally expressed cancer-associated mucin.

The results presented here demonstrate that immunoglobulin and cellular immune responses are observed with unglycosylated mucin-based peptide-only vaccines. Whether antibodies or T-cells become the predominant species in the immune response is most likely dependent upon the choice of immunological adjuvant and/or carrier protein used in the vaccine formulation.

20 Mucin-Based Glycopeptide Vaccines

In addition to exposed peptide motifs, the generation of aberrantly sialylated carbohydrate moieties provides yet another means of differentiating between cells expressing normal or cancer-associated mucins in vivo (see *Sect. 4*). Chemical synthesis allows a potentially unlimited number of vaccine formulations to be constructed incorporating one, or a combination of antigenic mucin-based carbohydrate structures.

In a pre-clinical study, mice were vaccinated with phospholipid-based liposomes containing a synthetic lipidated glycopeptide vaccine (see [107] *Table 3*) comprised of the T_N antigen (as a B-cell epitope), the MHC class II restricted outer membrane peptide fragment from *Neisseria meningitidis* (YAFKYARHANVGRNAFELFL) as a T-cell epitope, and the lipopeptide adjuvant Pam_3Cys with and without QS-21 [107]. Use of this multi-component vaccine aimed to achieve an immunoglobulin-dependent (and therefore) B-cell mediated recognition of the T_N antigen followed by T_H-cell recruitment through the T-epitope. In this situation, secretion of cytokines from the T_H-cell to the B-cell would then be envisaged to stimulate B-cell clonal expansion and differentiation into plasma cells that would subsequently produce soluble T_N antigen-specific antibodies. Encouragingly, analysis of post-vaccination sera by enzyme-linked immunosorbent assay (ELISA) using a T_N-BSA conjugate as bait revealed the production of T_N-specific IgM and IgG, with the highest titers observed from those mice that received QS-21 [107], suggesting that a T_H cell-mediated stimulation of B-cells had indeed occurred.

A similar immunological scenario was envisaged by Kunz and co-workers, who used a polar, non-immunogenic spacer molecule to couple a synthetic glycopeptide vaccine composed of the MUC1 glycopeptide (GV[T]SAPDTRPAP) displaying sialylated T_N attached to threonine (bold), and an ovalbumin peptide sequence (residues 323-339: ISQAVHAAHAEINEAGR [105] *Table 3*). This vaccine was administered to mice expressing TCRs transgenic for $OVA_{323-339}$ together with complete Freund's adjuvant. Analysis of blood serum by ELISA using MUC1-sT_N-BSA as bait revealed that sT_N-specific antibodies had been produced. Most importantly, co-incubation with either un-glycosylated MUC1 peptide or MUC4 peptide displaying sT_N neutralized antibody binding [105], suggesting that the immunoglobulins generated were only specific for MUC1 peptides decorated with a cancer-specific mucin-based antigen (sT_N).

Multi-valent vaccines aim to recreate the antigenic architecture that predominates on the surface of cancer cells following premature termination of glycosylation. A synthetic bivalent vaccine comprised of a glycosylated MUC2 peptide with a C-terminal cysteine residue (PTTTPITTTTTVTPTPTPTGTQTPTTTPISTTC: see [55] *Table 3*) conjugated to KLH and a globoH-KLH conjugate was evaluated in a clinical trial with biochemically relapsed prostate cancer patients, when administered with either QS-21 or GPI-0100 as immunological adjuvants [55]. The MUC2 peptide component was constructed synthetically and glycosylated

by treatment with ppGalNAcT II and ppGalNAcT IV glycosyltransferases. Analysis of post vaccination blood sera by ELISA using glycosylated MUC2 peptides conjugated to human serum albumin identified MUC2-specific IgM, IgG1 and IgG3 [55]. The highest antibody titers were derived from the administration of vaccine in combination with the QS-21 adjuvant. Of the nine patients who received the glycosylated MUC2-globoH-KLH vaccine, only one produced IgG1 and IgG3 capable of directing CDC against MCF-7 cells in vitro [55]. Similarly poor induction of CDC has been reported elsewhere [35,48,119].

In a recent study, the heterogeneous nature of cancer-associated antigens was addressed by Danishefsky and co-workers, in which a fully synthetic prostate cancer-specific vaccine incorporated multiple antigens (T, T_N, sT_N, Le^Y and globoH) displayed on a single polypeptide backbone (See [103] ❯ *Table 3*). This uni-molecular penta-valent vaccine was conjugated to either KLH or Pam_3Cys and administered to mice with QS-21. ELISA was used to compare antibody titers with those obtained from a control model that received each individual mono-valent antigen conjugated to KLH (T-KLH, T_N-KLH, sT_N-KLH, Le^Y-KLH and globoH-KLH). Mice that received the individually pooled mono-valent formulations induced high titer IgM responses against T_N, Le^Y and globoH. However, immunoglobulin class switching was not observed in response to any antigen using this formulation. In contrast, notable IgG titers were raised against T, T_N and sT_N following treatment with the uni-molecular penta-valent formulation coupled to KLH [103], whilst the Pam_3Cys conjugate generated IgG against T only, and at a significantly lower titer. Thus, the simultaneous administration of multiple carbohydrate antigens in a KLH-coupled, uni-molecular format served to recruit T_H-cells as part of the immune response against multiple antigens expressed by a specific (prostate) cancer. Importantly, IgG from sera generated against the penta-valent-KLH conjugate bound to MCF-7 cells (expressing globoH, Le^Y, sT_N, T and T_N), LSC cells (expressing Le^Y, sT_N and T_N) and DU-145 cells (expressing Le^Y and T) in vitro. In addition, the same study [103] describes the construction of additional uni-molecular penta-valent vaccine formulations (incorporating T, sT_N, T_N, GM2 and globoH), and increasingly elaborate hexa-valent molecules (displaying T, sT_N, T_N, Le^Y, GM2 and globoH) designed to specifically target breast and prostate cancer. Pre-clinical studies with these multivalent formulations are presently ongoing.

The efficacy of a fully synthetic mucin-based glycopeptide vaccine: α-*N*-acetylgalactosamine-*O*-serine/threonine (T_N) conjugated to either KLH or the B-cell activator palmitic acid (See [119] ❯ *Table 3*), was analyzed in a phase I clinical trial conducted with 15 patients presenting biochemically relapsed prostate cancer [119]. Patients presenting consecutive elevations in prostate-specific antigen (PSA), a biomarker that can be used to indirectly monitor disease progression, were eligible for trial. The vaccine was administered in a trimeric "clustered" formulation in combination with the adjuvant QS-21. Post vaccination blood sera were isolated and analyzed for immunoglobulin production and PSA levels. The highest antibody titers (IgM, IgG1 and IgG3) were observed with QS-21 as the adjuvant source, but presented a limited capacity to induce CDC in vitro (1 out of 15 patients). In contrast, patients that received $(T_N)_3$-palmitic acid produced IgM and very low titers of IgG. Notably, following treatment with $(T_N)_3$-KLH/QS-21, 33% of patients exhibited a reduction in PSA levels greater than 50%, whilst 46% presented PSA levels that remained stable after vaccination [119]. Importantly, 6/15 patients produced T_N-specific antibodies capable of binding to T_N expressing LSC colon carcinoma cells in vitro, whilst five patients enrolled in the trial remained disease free for 55 months [119].

In an alternative approach, a chemo-enzymatic strategy was used in the production of a multivalent mucin-based vaccine formulation. A MUC1 synthetic peptide comprised of 3 VNTR domains (3 × VTSAPDTRPAPGSTAPPAHG) was synthesized and glycosylated in vitro with purified recombinant human glycosyltransferase enzymes (GalNAc-T II, GalNAc-T IV and GalNAc-T XI: see [108] ❯ *Table 3*), such that 5 *O*-linked glycan groups were coupled per VNTR repeat. The extent of *O*-linked glycosylation with T_N was determined by MALDI-TOF analysis. The GalNAc-substituted MUC1 glycopeptide was then additionally sialylated (to form sT_N) with recombinant mouse ST6GalNAc-I, and coupled to KLH [108]. Wild type mice and mice transgenic for human MUC1 were vaccinated with these chemo-enzymatically derived formulations ($MUC1_{60}T_{N15}$ or $MUC1_{60}sT_{N15}$) with Freund's adjuvant to stimulate an immune response, and blood sera was analyzed by ELISA for the generation of antigen-specific immunoglobulins. Significantly, the transgenic model yielded MUC1 specific IgM and IgG1 with the $MUC1_{60}T_{N15}$ vaccine, and IgM, IgG1, IgG2A and IgG2B in response to administration of $MUC1_{60}sT_{N15}$, demonstrating immunoglobulin class switching in response to the administration of either vaccine. Whilst high MUC1-specific immunoglobulin titers were observed in both wild type and transgenic models, the highest values were observed when human MUC1 was expressed as self in the transgenic model, thus indicating that immunological tolerance can be broken. Furthermore, immunocytochemical analysis of sera produced following vaccination with $MUC1_{60}T_{N15}$ demonstrated positive staining of T_N-expressing T47D human breast cancer cells [108].

Interestingly, the immunogenicity of glycopeptides corresponding to the MUC1 VNTR domain was positively enhanced by substitution of amino acid residues involved in MHC binding [120]. Wholly synthetic, multi-component glycopeptide libraries modeled on human MUC1 were generated and administered with complete Freund's adjuvant to BALB/c mice. Lymph node T-cells were isolated from mice vaccinated with one such library containing the "MHC-anchor modified" glycopeptide APP(A)**MHG(V)LTSAP(D)ITRPA** (where the native residue in brackets is substituted for the residue in bold, and where GalNAc is attached to the boldface threonine). Characterization of the murine lymph nodes revealed the presence of T_H-cells that readily proliferated in response to the native, unmodified MUC1 glycopeptide in vitro [120]. Significantly, an equivalent response was not observed using the native unmodified MUC1 glycopeptide, clearly demonstrating that glycopeptide antigenicity was positively augmented by the engineered amino acid substitutions. Furthermore, use of the same glycopeptide library produced T_H-cells that proliferated following exposure to human breast cancer cell lysate in vitro [120], suggesting that the substitution of MHC anchoring amino acid residues does not adversely affect the presentation (and recognition) of the associated carbohydrate epitope. Hence, modification of peptide sequences lying within and outside of the VNTR domain with single or multiple amino acid substitutions can augment the immunogenicity of the corresponding peptide and glycopeptide antigens [118,120].

The majority of experimental results presented here demonstrate that a predominant antibody response is induced following the administration of mucin-based glycopeptide vaccines. The involvement of T-cells is observed however (as evidenced by immunoglobulin class switching from IgM to IgG mediated through T_H-cell stimulation of B-cells), although the production of antigen-specific $CD8^+$ CTLs is not a predominant feature of the observed immune responses. However, it must be noted that many of these vaccine formulations are designed specifically to induce an immunoglobulin response, where it is hoped that these formulations may serve by

eliminating the small number of remaining cells and circulating micrometastases that sometimes remain following the surgical removal of established tumors. However, when considering the structural nature of cancer-associated mucins and the ability of antibodies to mediate clearance by mechanisms such as CDC, perhaps a cautionary note of warning may be prudent where the administration of such vaccines is concerned.

21 Adoptive Transfer of Mucin-Based Antigens

Dendritic cells (DCs) are the most potent of antigen presenting cells, and function solely to present antigens to $CD8^+$ and $CD4^+$ T-cells via MHC class I and class II molecules respectively. Generating pools of activated, antigen-specific T-cells is often hampered in vivo by a cancer-driven subversion of the immune system. Adoptive transfer therapy has the potential to provide a way around this problem. DCs are removed from the cancer patient and cultured ex vivo with a source of cancer-specific antigen(s). During this time, internalization and presentation of antigen occurs on the DC cell surface, after which the cells can be re-administered to the donor patient to stimulate the production of cancer-specific T-cells in vivo. Alternatively, they can be used ex vivo to stimulate proliferation of antigen-specific T-cells which are cultured and re-introduced in large numbers back into the donor. The adoptive transfer of autologous (patient donated) DCs pulsed with cell lysates from MUC1-positive tumors or with nucleic acids encoding mucins or mucin-based peptides therefore aims to provide the recipient with a source of mucin-based antigens or antigen-specific T-cells from which potent, antigen-specific immunity can be derived. A number of in vitro human cell culture studies and pre-clinical in vivo murine studies demonstrate the feasibility of such an approach.

A murine model transgenic for human MUC1 that developed spontaneous pancreatic tumors expressing hypo-glycosylated MUC1 produced MUC1-specific $CD8^+$ CTLs in a natural immune response that failed to halt disease progression [121], an observation most likely explained by an inability of the immune system to distinguish "self" (healthy) from "altered self" (cancerous). Encouragingly however, the adoptive transfer of these unresponsive cells into healthy MUC1.Tg mice provided immunity against injected MUC1-positive tumor cells [121]. The significance of these results is twofold. First, they demonstrate that MUC1-specific CTLs can be produced naturally in response to the presence of MUC1-expressing tumors, and secondly, they serve to highlight that disease progression frequently occurs by cancer-mediated immune subversion despite the induction of an antigen-specific cellular immune response.

In another study, wild type mice were vaccinated with murine DCs transfected with MUC1 RNA from human breast cancer cells. In this system, human cancer-associated MUC1 RNA is translated by the murine DC cellular apparatus and presented on MHC class I molecules. Using this strategy, anti-MUC1 CTL-based immunity against a challenge with MC38 MUC1-positive, but not negative tumor cells was observed in vivo [122]. Critically, complete tumor clearance was observed in wild type mice with pre-established MUC1-expressing tumors following vaccination. However, introduction of these DCs into mice transgenic for human MUC1 did not provide immunity against MUC1-positive tumor cells due to pre-established immunological tolerance to mucins in the transgenic system. Importantly, a cellular anti-MUC1 immune response was observed in the transgenic model, but only when interleukin-12

(IL-12) was included in the treatment [122]. IL-12 is a multifunctional cytokine that (amongst other functions) is known to promote the expression of co-stimulatory molecules on DCs. The use of IL-12 with the adoptive transfer of murine DCs in this experimental system thus demonstrates that immunological tolerance to human MUC1 can be broken in vivo.

Vaccination of murine DCs pulsed with a synthetic peptide of 7 VNTR domains (7 × GVTSA-PDTRPAPGSTAPPAH: see [123] ❯ *Table 3*) into both wild type and MUC1.Tg mice induced only T-cell based immunity with a 90% rejection rate of injected murine T-cell lymphomas expressing human MUC1 [123]. The incorporation of 7 VNTR domains (rather than 1 or 2) was considered important for efficacy as previously, a synthetic MUC1 peptide composed of 7 VNTR domains was observed to be efficiently internalized and processed by DCs [124]. Using this formulation, MUC1-specific $CD8^+$ and $CD4^+$ T-cells were produced by the wild type mice (where human MUC1 is immunogenic) whereas significantly, only MUC1-specific $CD8^+$ T-cells were produced in the transgenic model [123]. These results demonstrate that whilst $CD4^+$-based immunological tolerance against human MUC1 was not broken in the transgenic system, tumor clearance can still proceed in the face of established immunological tolerance through a $CD8^+$ CTL-based mechanism. Furthermore and most importantly, these observations highlight that tumor clearance can proceed in the absence of an immunoglobulin or $CD4^+$ T-cell response in vivo.

In an in vitro study, DCs were pulsed with lysate from breast carcinoma cell lines MCF-7 (expressing sT_N, T, GM2, globoH and Le^Y) or MDA-MB-231 (expressing MUC1 TAAs as evidenced by binding of monoclonal antibody SM3 that recognizes the exposed peptide sequence PDTRP), and with tumor necrosis factor alpha and prostaglandin-E_2 as immunological adjuvants [125]. The pulsed DCs were then used to stimulate peripheral blood mononuclear cells in vitro. As a result, both activated $CD4^+$ T_{H1} cells and importantly, $CD8^+$ T-cells were obtained, indicating that presentation of MUC1 antigen to MHC class II molecules and cross-presentation (see also ❯ *Sect. 13* and ❯ *Fig. 6 J*) to MHC class I molecules had occurred [125].

Moreover, fusion of isolated MUC1-positive metastatic human colorectal carcinoma cells with autologous DCs provided a source of aberrantly glycosylated mucin together with the appropriate co-stimulation required to activate cellular immune responses in vitro [126]. Stimulation of patient-derived peripheral blood mononuclear cells with carcinoma-DC fusions generated antigen-specific CTLs that destroyed autologous MUC1-positive carcinoma cells in vitro [126].

A trial of 14 patients presenting advanced or metastatic cancer of the breast or lung (of which nine patients were MUC1-positive and five MUC1-negative) was conducted with DCs pulsed with either synthetic MUC1 peptide (TRPAPGSTAPPAHGVTSAPDTRPAPGSTAP: see [127] ❯ *Table 3*) or lysate from autologous cancer cells cultured in vitro [127]. Of the nine MUC1-positive patients, seven responded to vaccination and acquired antigen-specific immunity that caused a reduction in tumor size and prolonged survival in comparison with the MUC1-negative patients. However, following treatment, long-term immunity was found to be unsustainable in a number of patients. All of the MUC1-positive patients (four out of the seven that responded to vaccination) that received DCs pulsed with autologous tumor cell lysate subsequently relapsed. The three patients immunized with DCs pulsed with MUC1 peptide were alive at 3, 19 and 42 months post-vaccination at the time of trial [127].

A phase I clinical trial was conducted with a MUC1 peptide containing 3 VNTR domains and additional MUC1 flanking sequence (AHGVTSA[3 × PDTRPAPGSTAPPAHGVTSA]PESR-

PAPGSTAPAAHGVTSAPESRPAPGSTAPPAV: see [128] ❯ *Table 3*). The peptide was conjugated to oxidized mannan as an adjuvant and pulsed to autologous DCs that were administered to ten patients with advanced MUC1-positive adenocarcinomas [128]. Following treatment, vaccine-specific $CD4^+$ and $CD8^+$ T-cell responses were observed in all patients that were sustained for 1 year, whilst continual treatment led to disease stabilization in two of the ten patients [128].

The use of peptides to pulse DCs in adoptive transfer experiments is not restricted to the use of a single amino acid sequence. In one study, a computational screen of MUC1 was performed to identify peptide sequences that would bind to MHC class II (HLA-A2) molecules. Two peptide sequences were identified: STAPPVHNV from the VNTR domain and LLLLTVLTV from the leader sequence of the molecule [129]. Both of these peptide sequences were synthesized and used to pulse DCs, which were then further pulsed with the T-helper epitope PADRE to stimulate a $CD4^+$ T-cell response (See [130] ❯ *Table 3*). The efficacy of these "triple-pulsed" DCs was evaluated in a clinical trial on 20 patients presenting MUC1 positive metastatic renal cell carcinoma [130]. Following five vaccinations, patients additionally received treatment with IL-2. Post-treatment, MUC1 antigen-specific CTLs were produced that lysed target cells in ^{51}Cr release assays and caused disease stabilization in four patients and regression in six [130].

In spite of the immunological tolerance observed in murine studies, consideration of the above results demonstrate that humoral and cellular immune responses can be generated from the adoptive transfer of cancer-associated mucins in vitro and in vivo. In the clinical setting, the most promising results appear to be those obtained with DCs pulsed with synthetic MUC1 peptides rather than tumor cell lysate. One possible explanation for this observation may lie in the purity of the antigen preparation. The presence of a heterogeneous pool of TAAs derived from whole-cell preparations may result in the generation of large numbers of DCs, of which only a small proportion actually present the required antigen.

22 Concluding Remarks

The central challenge of mucin-based vaccine technology remains to break immunological tolerance to one or more cancer-specific antigens, such that immunoglobulins and cytotoxic T-lymphocytes can combine to rapidly clear established tumors from cancer patients. The combined use of both chemistry and biology in addressing this challenge has thus far provided encouraging data in preclinical trials, where in the majority of experiments, mutually exclusive immune responses producing either immunoglobulins or cytotoxic T-lymphocytes against specific antigenic targets is observed. Translation of these formulations to the clinical setting highlights that whilst positive progress is undoubtedly being made, a consistently effective vaccine formulation has yet to be produced. However, encouragement must be drawn from these initial clinical observations, which will undoubtedly pave the way forward towards the generation of more effective treatments. Re-addressing the "either-or" nature of vaccines to harness the full capacity of the immune system, so that both cancer-cell specific immunoglobulins and cytotoxic T-lymphocytes can be brought to bear against established tumors in the clinical setting remains the ultimate goal.

References

1. Hanisch FG, Muller S (2000) Glycobiology 10:439
2. Singh PK, Hollingsworth MA (2006) Trends Cell Biol 16:467
3. Hollingsworth MA, Swanson BJ (2004) Nature Reviews Cancer 4:45
4. Spiro RG (2002) Glycobiology 12:43R
5. Wopereis S, Lefeber DJ, Morava E, Wevers RA (2006) Clin Chem 52:574
6. Moniaux N, Junker WM, Singh AP, Jones AM, Batra SK (2006) J Biol Chem 281:23676
7. Moniaux N, Andrianifahanana M, Brand RE, Batra SK (2004) Brit J Cancer:1633
8. Huang L, Chen D, Liu D, Yin L, Kharbanda S, Kufe D (2005) Cancer Res 65:10413
9. Ren J, Raina D, Chen W, Li G, Huang L, Kufe D (2006) Mol Cancer Res 4:873
10. Raina D, Ahmad R, Kumar S, Ren J, Yoshida K, Kharbanda S, Kufe D (2006) EMBO J 25:3774–3783
11. Jonckheere N, Perrais M, Mariette C, Batra SK, Aubert J-P, Pigny P, Seuningen IV (2004) Oncogene 23:5729
12. Schroeder JA, Masri AA, Adriance MC, Tessier JC, Kotlarczyk KL, Thompson MC, Gendler SJ (2004) Oncogene 23:5739
13. Singh AP, Moniaux N, Chauhan SC, Meza JL, Batra SK (2004) Cancer Res 64:622
14. Fritz TA, Hurley JH, Trinh L-B, Shiloach J, Tabak LA (2004) Proc Natl Acad Sci USA 101:15307
15. Pratt MR, Hang HC, Ten Hagen KG, Rarick J, Gerken TA, Tabak LA, Bertozzi CR (2004) Chem Biol 11:1009
16. Hang HC, Bertozzi CR (2005) Bioorg Med Chem 13:5021
17. VandenSteen P, Rudd PM, Dwek RA, Opdenakker G (1998) Crit Rev Biochem Mol Biol 33:151
18. Brockhausen I (2006) EMBO Reports 7:599–604
19. Coltart DM, Royyuru AK, Williams LJ, Glunz PW, Sames D, Kuduk SD, Schwarz JB, Chen XT, Danishefsky SJ, Live DH (2002) J Am Chem Soc 124:9833
20. Schuman J, Campbell AP, Koganty RR, Longenecker BM (2003) J Pep Res 61:91
21. Dziadek S, Griesinger C, Kunz H, Reinscheid UM (2006) Chem Eur J 12:4981
22. Corzana F, Busto JH, Jimenez-Oses G, Asensio JL, Jimenez-Barbero J, Peregrina JM, Avenoza A (2006) J Am Chem Soc 128:14640
23. Kuhn A, Kunz H (2007) Angew Chem Int Ed 46:454
24. Kim YJ, Varki A (1997) Glycoconjugate J V14:569
25. Dube DH, Bertozzi CR (2005) Nature Rev Drug Discov 4:477
26. Varki A (1999) In: Varki A, Cummings R, Esko J, Freeze H, Hart G, Marth J (eds) Essentials of Glycobiology. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York
27. Sewell R, Backstrom M, Dalziel M, Gschmeissner S, Karlsson H, Noll T, Gatgens J, Clausen H, Hansson GC, Burchell J, Taylor-Papadimitriou J (2006) J Biol Chem 281:3586
28. Huang MC, Chen HY, Huang HC, Huang J, Liang JT, Shen TL, Lin NY, Ho CC, Cho IM, Hsu SM (2006) Oncogene 25:3267
29. Fukuda M (2002) Biochim Biophys Acta 1573:394
30. Somers WS, Tang J, Shaw GD, Camphausen RT (2000) Cell 103:467
31. Brown JR, Fuster MM, Sarkar AK, Esko JD (1999) Glycobiology 9:68
32. Tsutsumida H, Swanson BJ, Singh PK, Caffrey TC, Kitajima S, Goto M, Yonezawa S, Hollingsworth MA (2006) Clin Cancer Res 12:2976
33. Velcich A, Yang W, Heyer J, Fragale A, Nicholas C, Viani S, Kucherlapati R, Lipkin M, Yang K, Augenlicht L (2002) Science 295:1726
34. Dennis JW, Granovsky M, Warren CE (1999) Biochim Biophys Acta-Gen Subjects 1473:21
35. Ragupathi G, Liu NX, Musselli C, Powell S, Lloyd K, Livingston PO (2005) J Immunol 174:5706
36. Vlad AM, Muller S, Cudic M, Paulsen H, Otvos L, Jr., Hanisch FG, Finn OJ (2002) J Exp Med 196:1435
37. Carlos CA, Dong HF, Howard OMZ, Oppenheim JJ, Hanisch F-G, Finn OJ (2005) J Immunol 175:1628
38. Janeway C, Travers P, Walport M, Shlomchik M (2005) Immuno Biology: The Immune System in Health and Disease. Garland Science Publishing, New York
39. Male D, Brostoff J, Roth D, Roitt I (2006) Immunology. Mosby, Elsevier, Canada

40. Zitvogel L, Tesniere A, Kroemer G (2006) Nat Rev Immunol 6:715
41. Li MO, Wan YY, Sanjabi S, Robertson A-KL, Flavell RA (2006) Annu Rev Immunol 24:99
42. Kobie JJ, Wu RS, Kurt RA, Lou S, Adelman MK, Whitesell LJ, Ramanathapuram LV, Arteaga CL, Akporiaye ET (2003) Cancer Res 63:1860
43. Thomas DA, Massague J (2005) Cancer Cell 8:369
44. Rughetti A, Pellicciotta I, Biffoni M, Backstrom M, Link T, Bennet EP, Clausen H, Noll T, Hansson GC, Burchell JM, Frati L, Taylor-Papadimitriou J, Nuti M (2005) J Immunol 174:7764
45. Monti P, Leone BE, Zerbi A, Balzano G, Cainarca S, Sordi V, Pontillo M, Mercalli A, Di Carlo V, Allavena P, Piemonti L (2004) J Immunol 172:7341
46. Chan A, Bernstorff DCLWv, Spanjaard RA, Joo H-G, Eberlein TJ, Goedegebuure PS (1999) Int J Cancer 82:721
47. Agrawal B, Krantz MJ, Reddish MA, Longenecker BM (1998) Nature Med 4:43
48. Snijdewint F, Mensdorff-Pouilly Sv, Karuntu-Wanamarta AH, Verstraeten AA, Livingston PO, Hilgers J, Kenemans P (2001) Int J Cancer 93:97
49. Yu L-G, Andrews N, Zhao Q, McKean D, Williams JF, Connor LJ, Gerasimenko OV, Hilkens J, Hirabayashi J, Kasai K, Rhodes JM (2007) J Biol Chem 282:773
50. Raina D, Kharbanda S, Kufe D (2004) J Biol Chem 279:20607
51. Ren J, Agata N, Chen D, Li Y, Yu W-h, Huang L, Raina D, Chen W, Kharbanda S, Kufe D (2004) Cancer Cell 5:163
52. Yin L, Kharbanda S, Kufe D (2007) J Biol Chem 282:257
53. Dunn GP, Old LJ, Schreiber RD (2004) Annu Rev Immunol 22:329
54. Kim SK, Ragupathi G, Musselli C, Choi S-J, Park YS, Livingston PO (1999) Vaccine 18:597
55. Slovin SF, Ragupathi G, Fernandez C, Jefferson MP, Diani M, Wilton AS, Powell S, Spassova M, Reis C, Clausen H, Danishefsky S, Livingston P, Scher HI (2005) Vaccine 23:3114
56. Bohm C, Mulder MC, Zennadi R, Notter M, Schmitt-Graff A, Finn OJ, Taylor-Papadimitriou J, Stein H, Clausen H, Riecken EO, Hanski C (1997) Scand J Immunol 46:27
57. Shen L, Rock KL (2006) Current Opin Immunol 18:85
58. Heath WR, Carbone FR (2001) Nature Rev Immunol 1:126
59. Barnd DL, Lan MS, Metzgar RS, Finn OJ (1989) Proc Natl Acad Sci USA 86:7159
60. Jerome KR, Domenech N, Finn OJ (1993) J Immunol 151:1654
61. Noto H, Takahashi T, Makiguchi Y, Hayashi T, Hinoda Y, Imai K (1997) Int Immunol 9:791
62. Magarian-Blander J, Ciborowski P, Hsia S, Watkins SC, Finn OJ (1998) J Immunol 160:3111
63. Dittmann J, Karin K-M, Toni W, Thomas K, Tobias H, Horst-Dieter B, Stefan S, Hans-Georg R, Cécile G (2005) Cancer Immunol Immunother V54:750
64. Choi C, Witzens M, Bucur M, Feuerer M, Sommerfeldt N, Trojan A, Ho A, Schirrmacher V, Goldschmidt H, Beckhove P (2005) Blood 105:2132
65. Kunz H, Birnbach S (1986) Angew Chem Int Ed Engl 25:360
66. Lemieux RU, Ratcliffe RM (1979) Can J Chem 57:1244
67. Komba S, Meldal M, Werdelin O, Jensen T, Bock K (1999) J Chem Soc Perkin Trans 1:415
68. Hilaire PMS, Cipolla L, Franco A, Tedebark U, Tilly DA, Meldal M (1999) J Chem Soc Perkin Trans 1:3559
69. Mathieux N, Paulsen H, Meldal M, Bock K (1997) J Chem Soc Perkin Trans 1:2359
70. Meinjohanns E, Meldal M, Schleyer A, Paulsen H, Bock K (1996) J Chem Soc Perkin Trans 1:985
71. Paulsen H (1990) Angew Chem 102:851
72. Grundler G, Schmidt RR (1984) Liebigs Ann 1826
73. Herzner H, Reipen T, Schultz M, Kunz H (2000) Chem Rev 100:4495
74. Danishefsky SJ, Allen JR (2000) Angew Chem Int Ed 39:836
75. Marcaurelle LA, Bertozzi CR (2002) Glycobiology 12:69R
76. Brocke C, Kunz H (2002) Bioorg Med Chem 10:3085
77. Pellissier H (2005) Tetrahedron 61:2947
78. Buskas T, Ingale S, Boons GJ (2006) Glycobiology 16:113R
79. Karsten U, Serttas N, Paulsen H, Danielczyk A, Goletz S (2004) Glycobiology 14:681
80. Hojo H, Matsumoto Y, Nakahara Y, Ito E, Suzuki Y, Suzuki M, Suzuki A, Nakahara Y (2005) J Am Chem Soc 127:13720
81. Macmillan D, Bertozzi CR (2004) Angew Chem Int Ed 43:1355

82. Plante OJ, Palmacci ER, Seeberger PH (2001) Science 291:1523
83. Werz DB, Seeberger PH (2005) Chem Eur J 11:3194
84. Seeberger PH (2003) Chem Commun: 1115
85. Allen JR, Harris CR, Danishefsky SJ (2001) J Am Chem Soc 123:1890
86. Cipolla L, Rescigno M, Leone A, Peri F, La Ferla B, Nicotra F (2002) Bioorg Med Chem 10:1639
87. Vichier-Guerre S, Lo-Man R, Huteau V, Deriaud E, Leclerc C, Bay S (2004) Bioorg Med Chem Lett 14:3567
88. Buskas T, Li YH, Boons GJ (2005) Chem Eur J 11:5457
89. Buskas T, Li YH, Boons GJ (2004) Chem Eur J 10:3517
90. Grigalevicius S, Chierici S, Renaudet O, Lo-Man R, Deriaud E, Leclerc C, Dumy P (2005) Bioconjugate Chem 16:1149
91. Wan Q, Chen JH, Chen G, Danishefsky SJ (2006) J Org Chem 71:8244
92. Mammen M, Choi SK, Whitesides GM (1998) Angew Chem Int Ed 37:2755
93. Lundquist JJ, Toone EJ (2002) Chem Rev 102:555
94. Lo-Man R, Vichier-Guerre S, Perraut R, Deriaud E, Huteau V, BenMohamed L, Diop OM, Livingston PO, Bay S, Leclerc C (2004) Cancer Res 64:4987
95. Slovin SF, Keding SJ, Ragupathi G (2005) Immunol Cell Biol 83:418
96. Leppanen A, White SP, Helin J, McEver RP, Cummings RD (2000) J Biol Chem 275:39569
97. Kochendoerfer GG (2005) Current Opin Chem Biol 9:555
98. Keding SJ, Danishefsky SJ (2004) Proc Natl Acad Sci USA 101:11937
99. Allen JR, Danishefsky SJ (1999) J Am Chem Soc 121:10875
100. Udodong UE, Madsen R, Roberts C, Fraser-Reid B (1993) J Am Chem Soc 115:7886
101. Keding SJ, Endo A, Danishefsky SJ (2003) Tetrahedron 59:7023
102. Danishefsky SJ, Bilodeau MT (1996) Angew Chem Int Ed Engl 35:1380
103. Ragupathi G, Koide F, Livingston PO, Cho YS, Endo A, Wan Q, Spassova MK, Keding SJ, Allen J, Ouerfelli O, Wilson RM, Danishefsky SJ (2006) J Am Chem Soc 128:2715
104. Dziadek S, Kowalczyk D, Kunz H (2005) Angew Chem Int Ed 44:7624
105. Dziadek S, Hobel A, Schmitt E, Kunz H (2005) Angew Chem Int Ed 44:7630
106. Keil S, Claus C, Dippold W, Kunz H (2001) Angew Chem Int Ed 40:366
107. Buskas T, Ingale S, Boons GJ (2005) Angew Chem Int Ed 44:5985
108. Sorensen AL, Reis CA, Tarp MA, Mandel U, Ramachandran K, Sankaranarayanan V, Schwientek T, Graham R, Taylor-Papadimitriou J, Hollingsworth MA, Burchell J, Clausen H (2006) Glycobiology 16:96
109. Freire T, Lo-Man R, Piller F, Piller V, Leclerc C, Bay S (2006) Glycobiology 16:390
110. St Hilaire PM, Meldal M (2000) Angew Chem Int Ed 39:1163
111. Fumoto M, Hinou H, Ohta T, Ito T, Yamada K, Takimoto A, Kondo H, Shimizu H, Inazu T, Nakahara Y, Nishimura SI (2005) J Am Chem Soc 127:11804
112. Matsushita T, Hinou H, Fumoto M, Kurogochi M, Fujitani N, Shimizu H, Nishimura SI (2006) J Org Chem 71:3051
113. Acres B, Apostolopoulos V, Balloul JM, Wreschner D, Xing PX, Ali-Hadji D, Bizouarne N, Kieny MP, McKenzie IFC (2000) Cancer Immunol Immunother 48:588
114. Apostolopoulos V, Pietersz GA, McKenzie IFC (1996) Vaccine 14:930
115. Gilewski T, Adluri S, Ragupathi G, Zhang S, Yao T-J, Panageas K, Moynahan M, Houghton A, Norton L, Livingston PO (2000) Clin Cancer Res 6:1693
116. Ramanathan RK, Lee KM, McKolanis J, Hitbold E, Schraut W, Moser AJ, Warnick E, Whiteside T, Osborne, J, Kim H, Day R, Troetschel M, Finn OJ (2005) Cancer Immunol Immunother V54:254
117. Apostolopoulos V, Pietersz G, Tsibanis A, Tsikkinis A, Drakaki H, Loveland B, Piddlesden S, Plebanski M, Pouniotis D, Alexis M, McKenzie I, Vassilaros S (2006) Breast Cancer Res 8:R27
118. Tsang K-Y, Palena C, Gulley J, Arlen P, Schlom J (2004) Clin Cancer Res 10:2139
119. Slovin SF, Ragupathi G, Musselli C, Olkiewicz K, Verbel D, Kuduk SD, Schwarz JB, Sames D, Danishefsky S, Livingston PO, Scher HI (2003) J Clin Oncol 21:4292
120. Gad M, Jensen T, Gagne R, Komba S, Daugaard S, Kroman N, Meldal M, Werdelin O (2003) Eur J Immunol 33:1624

121. Mukherjee P, Ginardi AR, Madsen CS, Sterner CJ, Adriance MC, Tevethia MJ, Gendler SJ (2000) J Immunol 165:3451
122. Koido S, Kashiwaba M, Chen D, Gendler S, Kufe D, Gong J (2000) J Immunol 165:5713
123. Soares MM, Mehta V, Finn OJ (2001) J Immunol 166:6555
124. Hiltbold EM, Vlad AM, Ciborowski P, Watkins SC, Finn OJ (2000) J Immunol 165:3730
125. Bohnenkamp HR, Coleman J, Burchell JM, Taylor-Papadimitriou J, Noll T (2004) Cellular Immunol 231:112
126. Koido S, Hara E, Torii A, Homma S, Toyama Y, Kawahara H, Ogawa M, Watanabe M, Yanaga K, Fujise K, Gong J, Toda G (2005) Int J Cancer 117:587
127. Kontani K, Taguchi O, Ozaki Y, Hanaoka J, Sawai S, Inoue S, Abe H, Hanasawa K, Fujino S (2003) Int J Mol Med 12:493
128. Loveland BE, Zhao A, White S, Gan H, Hamilton K, Xing P-X, Pietersz GA, Apostolopoulos V, Vaughan H, Karanikas V, Kyriakou P, McKenzie IFC, Mitchell PLR (2006) Clin Cancer Res 12:869
129. Brossart P, Heinrich KS, Stuhler G, Behnke L, Reichardt VL, Stevanovic S, Muhm A, Rammensee HG, Kanz L, Brugger W (1999) Blood 93:4309
130. Wierecky J, Muller MR, Wirths S, Halder-Oehler E, Dorfel D, Schmidt SM, Hantschel M, Brugger W, Schroder S, Horger MS, Kanz L, Brossart P (2006) Cancer Res 66:5910

12.9 Polysaccharide-Based Vaccines

Violeta Fernández Santana, *Yury Valdés Balbin*, *Janoi Chang Calderón*, *Luis Peña Icart*, *Vicente Verez-Bencomo*
Center for Synthetic Antigens, Faculty of Chemistry,
University of Havana, Ciudad Habana, Cuba 10400
violeta@fq.uh.cu, yury@fq.uh.cu, janoi@fq.uh.cu, luisp@fq.uh.cu,
vicente@fq.uh.cu

Abstract

Capsular polysaccharides (CPS) and lipopolysaccharides from bacteria are employed for the production of vaccines against human diseases. Initial development of CPS as a vaccine was followed by the development and introduction of conjugate polysaccharide-protein vaccines. The principles leading to both developments are reviewed.

Keywords

Polysaccharide vaccine; Conjugate vaccine; Meningitis; Pneumonia; *Haemophilus influenzae* type b

Abbreviations

ADH adipic acid dihydrazide
CPS capsular polysaccharides
CRM_{197} cross reacting mutant from diphtheria toxin
DT diphtheria toxoid

In: *Glycoscience*. Fraser-Reid B, Tatsuta K, Thiem J (eds)
Chapter-DOI 10-1007/978-3-540-30429-6_69:

Hib *Haemophilus influenzae* type b
Im imidazol
KLH keyhole limpet hemocyanine
LPS lipopolysaccharide
Nm *Neisseria meningitidis*
Sp *Streptococcus pneumoniae*
TT tetanus toxoid

1 Introduction

A variety of glycans with important structural and functional roles are produced as part of bacteria cell walls. Capsular polysaccharides (CPS), teichoic acids, and lipopolysaccharides (LPS) are the three most important and represent the bacterial first line of defense against complement and bacteriophages. The carbohydrate portion of this glycoconjugate contains the major antigenic determinants that distinguish various serotypes of bacteria [1].
It has long been known that the presence or absence of CPS correlates in many cases with disease and that bactericidal and opsonic antibodies directed against them protect the host from invasive disease caused by encapsulated bacteria [2]. Antibodies directed against the LPS play also a major role in protection against infection by avoiding bacterial colonization of specific mucosa or organ [3].
The interest in polysaccharide-based vaccines was originated already in the 1930s and associated mainly with bacterial meningitis and pneumonia. In 1945, a Sp-vaccine made from CPS of four serotypes was studied [4] but the advent of chemotherapeutics and antibiotics diminished the interest in this area of study. A resurgence of interest in preventative medicine occurred only after several decades due to the increase in bacterial antibiotic resistance and the occurrence of meningitis epidemics.

2 Vaccines Based on Capsular Polysaccharide

Haemophilus influenzae, *Neisseria meningitidis*, and *Streptococcus pneumoniae* are the leading cause of bacterial meningitis, pneumoniae, and sepsis. Although six types of *Haemophilus influenzae* [5], 13 groups of *Neisseria meningitidis*, and more than 90 serotypes of *Streptococcus pneumoniae* [6] are known, only a limited number are clinically important. A single type b of *Haemophilus influenzae* is associated with more than 90% of all Hi pneumonia and meningitis. This is a typical pathogen of children under five years with a peak of incidence occurring mainly in the first year of age [7]. The Hib-Meningitis is associated with a high rate of sequelae among survivors.
Serogroups A, B, C, W135, and Y are responsible for more than 90% of meningococcal meningitis. The disease affects the early infancy but also occurs in older children and particularly adolescents [8]. In the so-called "meningitis belt" of Africa, a well-characterized epidemic pattern is observed for group A with some occurrence of group C.
Pneumococcal infections can occur at all ages, but they have two peaks of incidence in the first years of life and in the elderly [9]. At least 23 serotypes are associated with the disease,

Table 1
Commercially available polysaccharide vaccines

Bacteria	Vaccine	Company	Valency	CPS
Streptococcus pneumoniae	PneumovaxII	Aventis-MSD	23	1, 2, 3, 4, 5, 6B, 7F, 8, 9N, 9V, 10A, 11A, 12F, 14, 15B, 17F, 18C, 19F, 19A, 20, 22F, 23F, 33F
	Pnu-immune	Wyeth	23	
Neiseria meningitidis	AC Vax	GSK	2	A, C
	Mengivax (A + C)	Aventis	2	
	Menomune	Aventis	4	A, C, W-135, Y
	Mencevax	GSK	4	
Salmonella typhi	Typhim Vi	Aventis	1	Vi

but its distribution is associated not only with age but also with geographical regions [10]. An important international collaborative effort is actually been made for establishing the burden of disease worldwide and the serotypes associated with them [11].

The potential of CPS as human vaccines for prevention of meningitis and pneumonia was ascertained by the fact that protective immune response is serotype-specific [4,12] and is directed mainly against the capsular polysaccharides; other cellular components are recognized by protecting antibodies to a significantly lower extent [13].

For Hib, Nm, and Sp the CPS expression correlates with increasing virulence [14,15]. The capsules themselves are highly hydrated shells surrounding the bacterium and mediating the adherence of the bacterium to biological and inorganic surfaces. The capsule inhibits or interferes with recognition of the bacteria cell wall by antibodies, adherence of complement, and ingestion of bacteria by host cells and killing of the bacterium after phagocytosis [16]. The polyanionic properties of the vast majorities of CPS help them in modulating the passage of molecules and especially ions to the bacterial cell envelope.

Capsular polysaccharides purified from bacteria are almost non-toxic and considerably less toxic than the whole-cell and elicited anti-polysaccharide antibodies at a single dose of 10–50 μg [17]. The antibodies provide type-specific immunity lasting for several years in healthy humans [18]. The simplicity of polysaccharides as vaccines and its low toxicity prompted development of the first polysaccharidic pneumococcal vaccine (tetravalent) marketed as early as 1946 [19]. Polysaccharide vaccines were only really introduced during the 1970s to counteract the emergence of bacterial resistance to antibiotics.

In 1974 vaccines against *Neisseria meningitidis* [20] followed by *Streptococcus pneumoniae* [21,22] in 1977 and later *Haemophilus influenzae* type b [23] were licensed (see *Table 1*).

2.1 Anti-CPS Response and the Structure of the CPS

The anti-CPS antibody response recognizes linear epitopes on the polysaccharides [24]. The response is strong, but of relative low avidity, with a production mostly of IgM and IgG2

antibodies, which are not good activators of complement. The IgM is rapidly cleared but the IgG2 response lasts for a few years. Typically the antibody response drops to about a quarter of its initial value after four years in adults, and more rapidly in children [25].

Polysaccharides are T-independent antigens [26] and as a general rule the response to polysaccharides occurs 3–18 months after birth in humans, but it depends on the CPS structure; generally, children less than 2 years of age respond poorly. Immunological memory is not induced and as a result no boosting of the antibody response takes place on reimmunization. The affinity maturation and isotype switching of the antibody repertoire is absent.

Considering their relative simplicity of production and the low toxicity, this pattern of immunological response was generally acceptable in "the absence of better vaccines" and is acceptable now only for vaccines used to protect the elderly from pneumonia or high risk groups over short periods, for example during epidemics.

Some of the capsular polysaccharides are good immunogens and others are poor immunogens although there is no general rule associating structure with immune response. For example, the capsular polysaccharide of group A Meningococci [27,28] is a good immunogen in infants starting at the age of 3 to 6 months and produces protective levels of IgG antibodies. One or two-dose regimes of group A or bivalent A + C vaccine may be useful to protect infants during epidemics, although without offering a long-term protection [29]. Pneumococcal capsular polysaccharides of serotypes 3 and 18C induce also a good antibody response in infants as young as 6 months, but only older children respond to the CPS from serotypes 19F and 23F [30]. The capsular polysaccharide of serotype 6 is a poor immunogen at all ages [31]. *Haemophilus influenzae* type b capsular polysaccharide elicits long-lived and complement-mediated bactericidal antibodies in adults [13].

The structure of the repeating unit from the CPS used in polysaccharide vaccine is provided in ❯ *Table 2*. Some of them, particularly those from *Neisseria meningitidis* with a simpler repeating unit, were solved primarily using one-dimensional ^{13}C NMR spectroscopy and were amongst the first complex polysaccharide structures elucidated using these techniques.

Capsular polysaccharides are characterized by the presence of a strict repeating unit as for example in *Haemophilus influenzae* type b where a phosphodiester bond is a bridge between ribosyl-ribitol without any additional substituent. In other CPS the structure becomes heterogeneous by incomplete substitution or multiple *O*-acetylation sites. The repeating unit in *Neisseria meningitidis* is linear and composed only of a mono or disaccharide. All have a small repeating unit, are of relatively low-molecular weight and contain either phosphate or Neu5Ac as the anionic group. The repeating unit of *Streptococcus pneumoniae* capsular polysaccharides have a more complex structure (between 2–7 monosaccharide residues), often contain uronic acids but not sialic acid and are substituted with a wide range of groups, including *O*-acetyl, pyruvate acetal, and glycerol phosphate groups.

O-Acetyl groups are immunologically relevant in several CPS. For example, de-*O*-acetylation reduces or completely abolishes the recognition by anti-CPS sera in Nm serogroup A [32] and also in Sp capsular polysaccharides from serotype 1 [33], 11A [34], 15F [35], and 34 [36]. In a few cases the *O*-acetyl group can be removed without significant loss of activity toward serotype-specific serum. The most remarkable is Nm serogroup C. About 15% of meningococcal group C isolates produce a non-acetylated polysaccharide but antibodies induced by this CPS are still protective against infections caused by strains expressing *O*-acetylated CPS [37]. The fine specificity of antibodies reveals that both polysaccharides could be used in the produc-

Table 2
Structure of CPS included in commercially available vaccines

Haemophilus influenzae		
Type	**Primary structure**	**Ref.**
b	→3)-β-D-Ribf-(1→1)-D-Ribitol-(5-OPO_3→	[42]

Neisseria meningitidis		
Group	**Primary structure**	**Ref.**
A	→6)-α-D-ManNAc-(1-OPO_3→ (3Ac, 85%)	[43]
C	→9)-α-Neu5Ac-(2→ (7/8 Ac, 85%)	[44]
W-135	→6)-α-D-Gal-(1→4)-α-Neu5Ac-(2→	[45]
Y	→6)-α-D-Glc-(1→4)-α-Neu5Ac-(2→	[45]

Streptococcus pneumoniae		
Serotype	**Primary structure**	**Ref.**
1	→3)AAT-α-D-Gal-(1→4)-α-D-GalA-(1→3)-α-D-GalA-(1→ (Ac, 30%)	[46]
2	α-D-GluA-(1→6)-α-D-Glc 1 ↓ 2 →4)-β-D-Glc-(1→3)-α-L-Rha-(1→3)- α-L-Rha-(1→3)-β-L-Rha-(1→	[47]
3	→3)-β-D-GlcA-(1→4)-β-D-Glc-(1→	[48]
4	→3)-β-D-ManNAc-(1→3)-α-L-FucNAc-(1→3)- α-D-GalNAc-(1→4)-α-D-Gal2,3(*S*)Pyr-(1→	[49]
5	→4)-β-D-Glc-(1→4)-α-L-FucNAc-(1→3)-β-D-Sug-(1→ 3 ↑ 1 α-L-PneNAc-(1→2)-β-D-GlcA	[50]
6B	→2)-α-D-Gal-(1→3)-α-D-Glc-(1→3)-α-L-Rha-(1→4)- D-Rib-ol-(5→OPO_3→	[51]
7F	β-D-Gal 1 ↓ 2 →6)-α-D-Gal-(1→3)-β-L-Rha2Ac-(1→4)-β-D-Glc-(1→3)- β-D-Gal*NAc*-(1→ 4 ↑ 1 α-D-Glc-(1→3)-α–L-Rha	[52]

Table 2 (continued)

Streptococcus pneumoniae		
Serotype	**Primary structure**	**Ref.**
8	→4)-β-D-GlcA-(1→4)-β-D-Glc-(1→4)-α-D-Glc-(1→4)- α-D-Gal-(1→	[53]
9N	→4)-α-D-GlcA-(1→3)-α-D-Glc-(1→3)-β-D-ManNAc-(1→4)-β-D-Glc-(1→4)-α-D-GlcNAc- (1→	[54]
9V	(2Ac, 17%) (4Ac, 6%) (3Ac, 25%) (6Ac, 55%) →4)-α-D-GlcA-(1→3)-α-D-Gal-(1→3)-β-D-ManNAc-(1→4)-β-D-Glc-(1→4)-α-D-Glc- (1→ (2Ac, 3%) (3Ac, 4%)	[55]
10A	β-D-Gal 1 ↓ 6 5)-β-D-Gal*f*-(1→3)-β-D-Gal-(1→4)-β-D-GalNAc-(1→3)-α-D-Gal-(1→2)-D-Rib-ol- (5→OPO_3→ 3 ↑ 1 β-D-Gal*f*	[16]
11A	Gro*l*-(1→OPO_3 ↓ (+ OAc) 4 →6)-α-D-Glc2/3Ac-(1→4)-α-D-Gal-(1→3)-β-D-Gal-(1→4)-β-D-Glc-(1→	[16]
12F	α-D-GalNAc 1 ↓ 3 →4)-α-L-FucNAc-(1→3)-β-D-GalNAc-(1→4)- β-D-ManNAcA-(1→ 3 ↑ 1 α-D-Glc-(1→2)-α-D-Glc	[56]
14	→6)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-Glc-(1→ 4 ↓ 1 β-D-Gal	[57]
15B	→6)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)- 4 β-D-Glc-(1→ ↑ 1 α-D-Gal-(1→2)-β-D-Gal 3 (+ OAc, 70%) ↑ $Cho_{0.2}$→P	[58]

Table 2 (continued)

Streptococcus pneumoniae		
Serotype	**Primary structure**	**Ref.**
17F	→3)-β-L-Rha-(1→4)-β-D-Glc-(1→3)-α-D-Gal-(1→3)-β-L-Rha2Ac-(1→4)-α-L-Rha-(1→2)-D-Ara-ol-(1→OPO_3→ 4 ↑ 1 α-D-Gal	[59]
18C	Gro-(1→OPO_3 ↓ 3 →4)-β-D-Glc-(1→4)-β-D-Gal-(1→4)-α-D-Glc-(1→3)-β-L-Rha-(1→ 2 ↑ 1 α-D-Glc (Ac, 30%)	[60]
19F	→ 4)-β-D-ManNAc-(1→4)-α-D-Glc-(1→2)-α-L-Rha-(1→OPO_3→	[61]
19A	→4)-β-D-ManNAc-(1→4)-α-D-Glc-(1→3)-α-L-Rha-(1→OPO_3→	[62]
20	→6)-α-D-Glc-(1→6)-β-D-Glc-(1→3)-β-5,6-OAc-D-Gal*f*-(1→3)-β-D-Glc-(1→3)-α-D-GlcNAc-(1→OPO_3→ 4 ↑ 1 β-D-Gal*f*	[63]
22F	α-D-Glc 1 ↓ (β-l-Rha, 2Ac, 80%) 3 →4)-β-D-GlcA-(1→4)-β-L-Rha-(1→4)-α-D-Glc-(1→3)-α-D-Gal*f*-(1→2)-α-L-Rha-(1→	[64]
23F	Gro-(2→OPO_3 ↓ 3 →4)-β-D-Glc-(1→4)-β-D-Gal-(1→4)-β-L-Rha-(1→ 2 ↑ 1 α-L-Rha	[65]
33F	(2Ac, 40%) →3)-β-D-Gal-(1→3)-α-D-Gal-(1→3)-β-D-Gal*f*-(1→3)-β-D-Glc*f*-(1→5)-β-D-Gal*f*-(1→ 2 ↑ 1 (+ OAc, 20%) α-D-Gal	[41]

tion of glycoconjugate vaccine with similar efficiency [38]. Sp serotype 20 is also recognized by anti-Sp20 serum after de-*O*-acetylation [39].
Other substituents in CPS are also immunodominant in most cases. For example, the pyruvate groups are immunodominant in serogroup 4. Its removal clearly changes immunological properties of the CPS [40]. Phosphates are very frequent in Sp CPS mainly as a phosphodiester bridge, but it could also be part of a backbone as free phosphate, glycerol 1-phosphate, glycerol 2-phosphate, ribitol phosphate. As a backbone the substituents become immunodominant and their loss abolishes the recognition, for example after dephosphorylation of Sp23F [65].
Lipopolysaccharides are also considered T-independent antigens [66]. LPS share some common features with CPS as for example direct stimulation of B-cells in the absence of T-cells, weak immunogenicity and protective capacity of the anti-LPS antibody response. While CPS initiate the immune response binding directly to the IgM receptor on B-cells and crosslinking them through their numerous repeating epitopes and are nontoxic, LPS are potent B-cell mitogens and are highly toxic [67], a fact that precludes its direct use as human vaccine.
Vaccination of infants and young children with CPS failed in producing serotype-specific antibody and general vaccination of humans of all-age with LPS is unacceptable due to its toxicity [68]. This failure is the main driving force behind the development of conjugate vaccines.

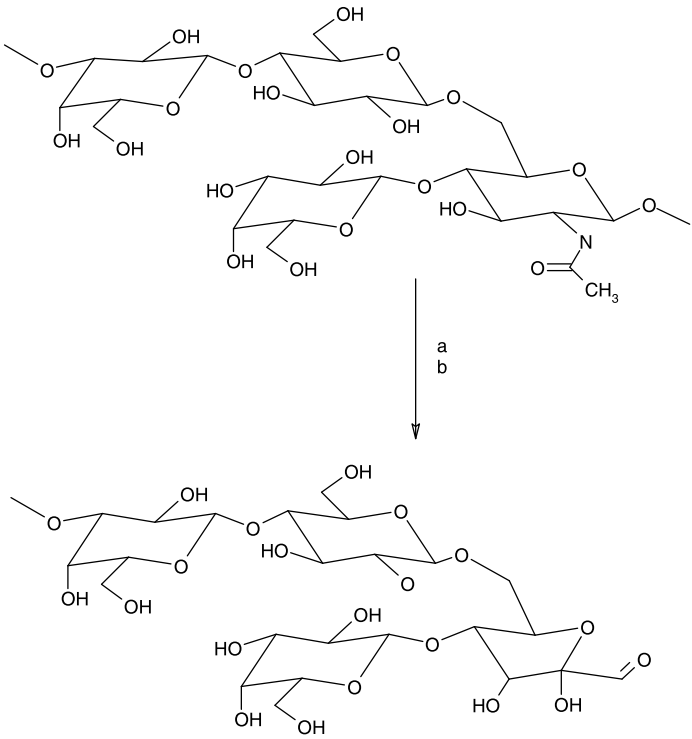
3 Glycoconjugate Vaccines

Polysaccharides are ideally located on the surface of bacteria and are a source of the most highly conserved protective epitopes. One of the challenges in vaccine development during the final quarter of the 20th century was to develop a better anti-carbohydrate immune response for CPS and LPS. Conjugation of polysaccharides to proteins provided a solution to some of the limitations found for natural compounds.
The concept of hapten and hapten-protein conjugate was introduced in the 1920s by Karl Landsteiner to provide a chemically defined system for studying the binding of and individual antibody to a unique epitope on a complex protein antigen [69]. According to this concept small organic molecules (haptens) were chemically coupled to a large protein (carrier) giving a hapten-carrier conjugate. Animals immunized with such a conjugate produce specific antibodies not only against the carrier protein (unaltered epitopes) but also for the hapten determinant.
Avery and Goebel applied this concept to cellobiouronic acid [70], and also to the Sp serotype 3 CPS. Both conjugates to horse serum globulin were able to induce polysaccharide specific antibodies and remarkably they conferred immunity to challenge by live *Streptococcus pneumoniae* serotype 3.
The development of conjugate vaccine as a pharmaceutical was initially undertaken for *Haemophilus influenzae* type b [71] and later, after the impressive success was extended to other CPS from *Streptococcus pneumoniae* [72] and *Neisseria meningitidis* [73]. The development of glycoconjugate vaccines from group B *Streptococcus* [74,75], *Salmonella typhi* [76], *Staphylococcus aureus* [77] CPS are also under different investigation stages. The development of conjugate vaccine from the carbohydrate portion of lipopolysaccharide is less straightforward and although the concept was proved at experimental level for *Vibrio cholerae* O1 [78] and O139 [79], *Shigella sonnei* and *flexneri* 2a [80], the vaccines are in earlier stages

of development-application and the concept is not fully proven in clinics. It is impossible to cover all aspects of the development of glycoconjugate vaccine in this chapter. For a detailed inventory of research the reader can consult excellent reviews on this matter [81,82].
Basically, the success in development of carbohydrate-protein conjugate vaccines as pharmaceuticals [83] is defined by the adequate selection of: (a) polysaccharide modification strategy, (b) the carrier protein and its modification, (c) the conjugation methods and by the establishment of analytical methods for the evaluation of the final glycoconjugate.

3.1 Polysaccharide Modification Strategy

Molecular weight is an important property for CPS vaccines as the immune response is induced by interaction of repeating epitopes with IgM B-cell receptors and their crosslinking. In glycoconjugate vaccines, the immune response proceeds by a different mechanism involving T-cells. Therefore, the length of the carbohydrate chain in a CPS-protein conjugate is not as important and restricted as for the CPS vaccines. The epitope recognized by antibodies to CPS is generally located in a linear small fragment of the capsular polysaccharide and consequently polysaccharide-protein conjugates prepared from CPS with a broad range of molecular weight have been shown to be excellent immunogens experimentally and in clinical practice [59]. T-cell dependency is more pronounced for conjugates with small-sized fragment polysaccharides. For example, small dextran molecules (4 KDa) were better immunogens in mice than for large ones ($\leq$ 40 KDa) [84,85]. Glycoconjugates from high molecular weight polysaccharides are immunogenic in humans at lower dose [86].
Very short oligosaccharides of one or two repeating units may not always raise antibodies specific for the polysaccharide [87,88] and usually are excluded from the preparation of glycoconjugate vaccines. Apart from that, fractions selected with a broad spectrum of size could be considered and for instance were employed successfully in the glycoconjugate vaccines developed so far. In the design of a glycoconjugate vaccine, the selection of the polysaccharide size is actually dictated by the conjugation technology and the removal of free CPS from the final product rather than by immunological consideration. Molecular sizing of the polysaccharides can be seen as a simple and effective means to ensure consistency of the final conjugate [89]. Therefore, partial depolymerization is usually performed before activation even if the intention is to use high molecular weight polysaccharides for conjugation.
For high and low molecular weight polysaccharides the strategy for conjugation differs markedly, specially for the type of activation: In large polysaccharides the activation modified randomly internal sites situated along the chain. Small oligo- or polysaccharide fragments could be selectively modified by terminal activation at one or both ends of the carbohydrate chain [90].
The methods for depolymerization more frequently employed are acid [91] or enzymatic hydrolysis [92] and ultrasonic irradiation [93,94]. The mixture of fragments is separated into fractions of selected molecular sizes by methods such as ion exchange chromatography [95] or gel filtration or ultrafiltration [96]. In some particular cases, endoglycosidases could provide selective cleavage of a specific bond. For example, Sp serotype 14 could be fragmented by the catalytic hydrolysis of endo-β-galactosidase from *Flavobacterium keratolyticus* [97].



Scheme 1
(a) NaOH, (b) $NaNO_2$

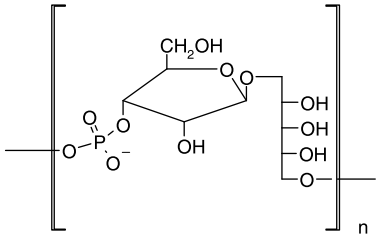
Defragmentation and activation could be performed simultaneously. For example, partial de-*N*-acetylation and nitrous acid deamination of the Sp serotype 14 CPS [98] proceeds with a rearrangement leading to a terminally activated oligo or polysaccharide according to the *Scheme 1*.

For some polysaccharides as for example Nm group C [99] and *H. Influenzae* type b [100] the periodic oxidation occurs selectively at a position leading to fragmentation of the CPS.

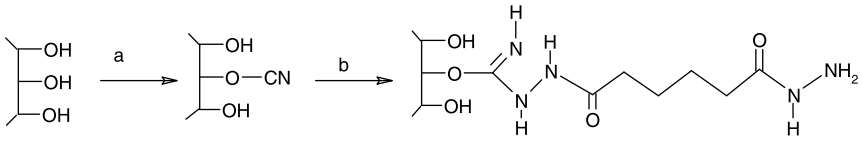
Depolymerized CPSs are further modified for activation. The following examples of *Haemophilus influenzae* type b CPS modification illustrate very well the broad spectrum of present technology [101] (see *Scheme 2*).

Method A. A high molecular weight polysaccharide was activated by reaction with cyanogen bromide at basic pH, in a similar fashion Sepharose is activated for affinity chromatography. The reaction takes place at all possible hydroxy groups to produce O–CN randomly

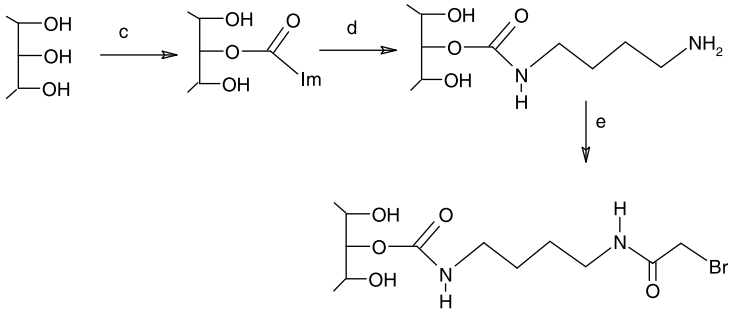
Scheme 2
Modification of Hib capsular polysaccharide for conjugation. (a) CNBr/OH–; (b) Adipic acid dihydrazide; (c) Carbonyldimidazole/OH–; (d) Butanediamine; (e) p-nitrophenylbromoacetate; (f) Sodium periodate; (g) Acetic acid; (h) Ethylendiamine, sodium cyanoborohydride



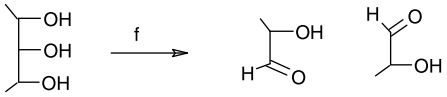
METHOD A



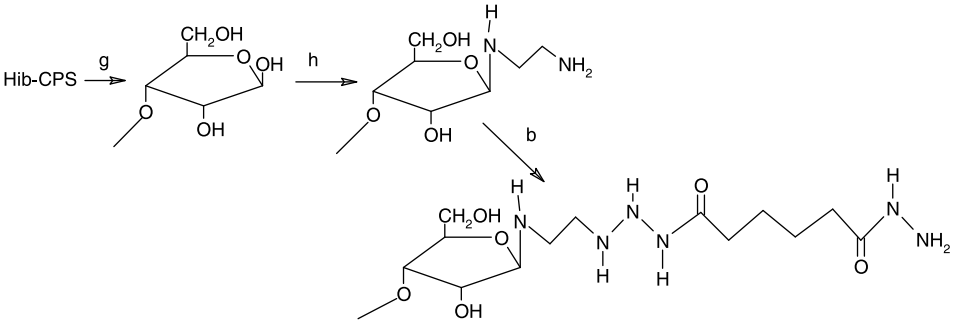
METHOD B



METHOD C



METHOD D



distributed all along the chain [102]. A spacer was introduced with adipic acid dihydrazide (ADH). The reactive group for conjugation is a very reactive hydrazide.

Method B. A high molecular weight polysaccharide was size-reduced and further transformed into a tetrabutylamonium salt for solubilization in dimethylformamide or dimethylsulfoxide [103]. The reaction sequence for activation is rather complex and includes activation of hydroxy groups with carbonylimidazol followed by introduction of a spacer with a terminal bromoacetate function.

Method C. The polysaccharide was directly oxidized by sodium periodate. The oxidation is restricted to the ribitol moiety leading to simultaneous fragmentation and terminal activation. As a result a fragment was obtained having aldehyde functions at both ends [100].

Method D. The polysaccharide was hydrolyzed with acetic acid to fragments having between 5 and 15 repeating units. After purification, ethylendiamine reacts at the terminal reducing end introducing a terminal amino function. Similarly to method A, the reaction with ADH introduced a spacer having terminally a very reactive hydrazide function [104].

The modification of the polysaccharide depends strongly on the structure. For example the meningococcal group C polysaccharide is partially acetylated at positions 7-8. Acetylation preclude the periodic oxidation, therefore the oxidation occurs only at those units with a free diol. The oxidation similarly to Hib polysaccharide breaks the structure producing size-reduced and terminal activated polysaccharide. The same process for Sp serotype 19F could take place at several internal bonds without size-reduction generating aldehydes randomly distributed along the chain (see ❯ *Scheme 3*). The control of the processes is essential in both cases to achieve the introduction of an ideal number of active groups. Over-activation will affect the structure of the CPS and reduce their recognition; under-activation will reduce the yield on the conjugation step or will produce less stable conjugates since the carbohydrate-protein attachment is effected only through a low number of sites.

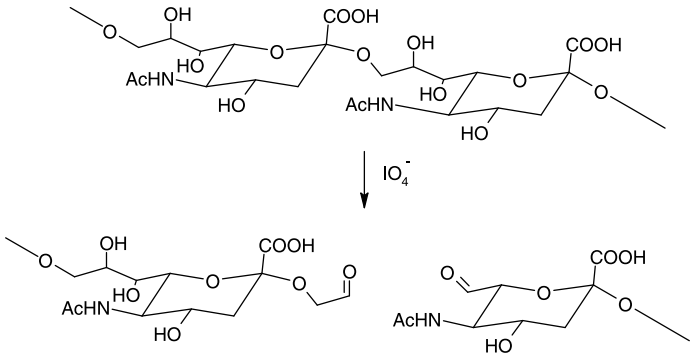
3.2 Selection of the Protein Carrier and its Modification

Several highly immunogenic bacterial proteins have been proposed as carriers for conjugate vaccines. The list includes bacterial pili, outer membrane proteins (OMPs), and excreted toxins of pathogenic bacteria [105], preferably in toxoid form. Tetanus and diphteria toxoids, which are readily available and accepted for human use, are the preferred carrier despite the lot-to-lot variations during the detoxification procedure.

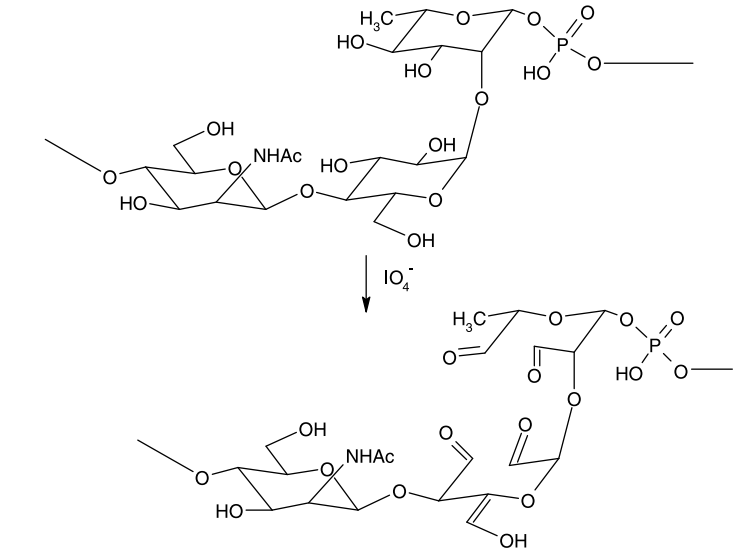
Alternative recombinant carrier proteins have been developed, for example CRM_{197} a non-toxic analogue of diphtheria toxoid [106,107] and a 42 KDa outer membrane protein from *Haemophilus influenzae* [108].

The lysine amino function is usually employed for conjugation but in some cases the amino group was modified to a thiol, for example, 25% of free amines in Nm Outer Membrane Protein were modified with *N*-acetylhomocysteine tiolactone [103].

The development of tetravalent meningococcal and 13-valent pneumococcal conjugate vaccine pointed to the future need of new carrier proteins. Experimentally or in very early clinical development are several proteins. Examples are *Bortedella pertussis* fimbriae [109], recombinant 64 KDa Nm OMP [110], Nm porin [111], and Sp pneumolysin [112].



a



b

Scheme 3
Sodium periodate oxidation of (a) *Neisseria meningitidis* group C; (b) *Streptococcus pneumoniae* serogroup 19F

3.3 Conjugation Methods

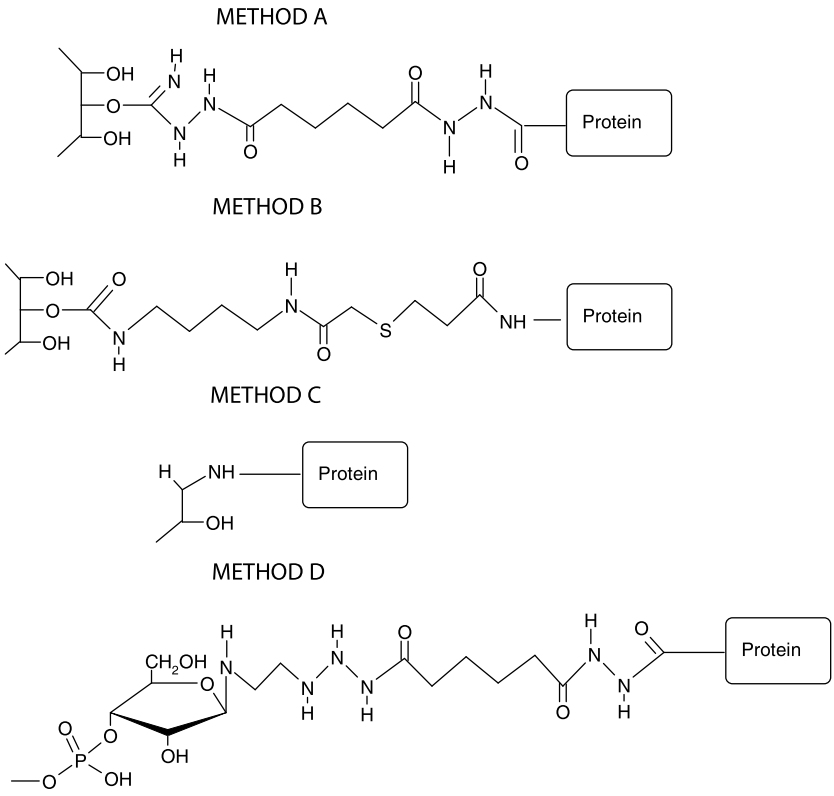
The polysaccharide modification strategy, the selected protein carrier and the conjugation methods are interdependent. The structure of the carbohydrate-protein conjugate could be very diverse ranging from a neoglycoprotein type made from smaller fragments of the CPS linked

to a protein through a single point to a multicomponent lattice-type which contains several molecules of both components each linked at a multiple point.

A wide variety of glycoconjugation methods were explored and are covered in the numerous reviews on the subject [113]. Carbohydrate-conjugates to bovine or human serum albumine are usually obtained as the very first step in conjugation research. Such conjugates serve not only as a model but also for studying the antigenicity of the conjugate polysaccharide, i. e. the ability of being recognized by antibodies of the unmodified CPS.

For the development of polysaccharide-protein conjugates as pharmaceuticals it is necessary not only to perform covalent coupling of both molecules but also to fulfill [83] some the following important requirements:

- The linkages formed during the conjugation process should be stable and also immunologically and pharmacologically acceptable.
- The integrity of both components has to be maintained.
- The residual reagent should be eliminated. Any residual reagent left should be below acceptable levels related to toxicity.



Scheme 4
Structure of conjugates made by methods (a)–(d)

- Neoepitopes formed either as a result of the conjugation or after active group capping could not be immunologically relevant.
- The product must be as defined as possible to facilitate their control.
- The conjugates have to be reproducible and water soluble.
- The immunogenicity of the conjugate must be optimized in terms of polysaccharide size, loading of the protein carrier, and CPS–protein ratio.

The direct introduction of active groups on the polysaccharide and the protein could seriously affect the efficiency of the coupling by steric hindrance [114]. By direct attachment, the saccharide epitopes could be shielded by the secondary structure of the protein. To avoid this problem, a spacer arm is frequently introduced either on the modified CPS or on the protein carrier during the design of a glycoconjugate. The conjugates represented in *Scheme 4* have been shown to be highly stable. They contained aliphatic amines [115], amides [116,117], or thioethers [118].

3.4 Glycoconjugate Vaccines from Lipopolysaccharides

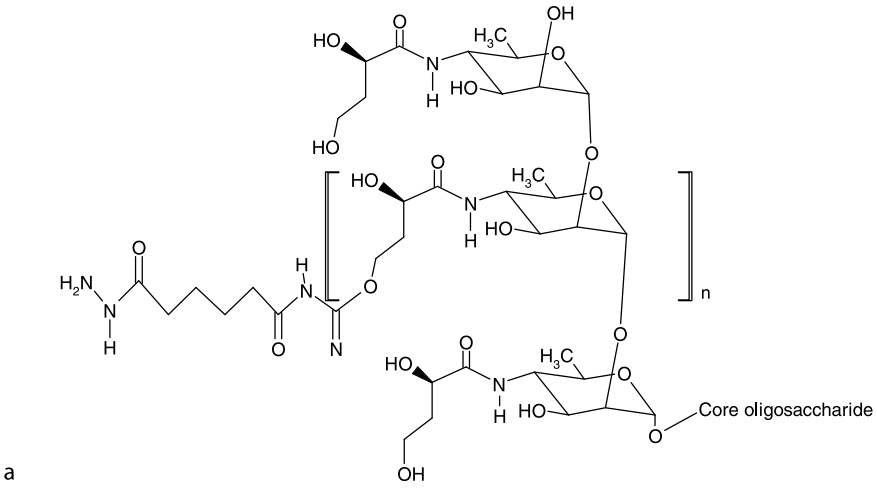
Anti-lipopolysaccharide response plays an important role in protection against disease [3] by interfering with colonization. The use of intact LPS as vaccines however, is precluded by their toxicity, which must be reduced as a critical first step before any further manipulation is done to improve the anti-LPS immune response. If detoxified and the LPS retains the epitopes it could be conjugated to a protein carrier to develop a better anti-carbohydrate response in a similar fashion to CPS.

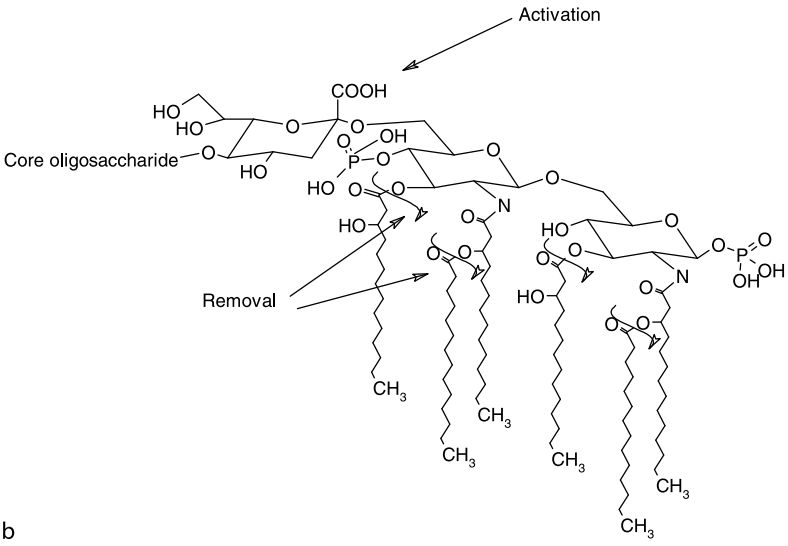
The general simplified structure of the LPS contains three domains; (a) *O*-polysaccharide chain, responsible for serotype specificity, (b) core polysaccharide, a conserved, usually short oligosaccharide chain containing KDO and (c) lipid A, which is responsible for the toxicity and is composed of GlcNAc-disaccharide substituted by several fatty acids linked as esters or amides.

The modification and activation of LPS for conjugation to a protein is more complex than for CPS. The first step is detoxification by fragmentation of the lipid A moiety either by the cleavage of the acid-sensitive KDO-lipid A linkage or by saponification of fatty acids. Acid hydrolysis completely eliminates lipid A and could be used in the absence of acid-sensitive monosaccharides for isolation of the *O*-polysaccharide. Saponification of fatty acids by base hydrolysis or hydrazinolysis also reduces the toxicity to acceptable levels.

LPS with a long *O*-polysaccharide chain could be regarded as CPS for conjugation and is usually activated by random modification of the polysaccharide chain. Examples illustrating the complexity and potential of LPS for glycoconjugate vaccines are shown in *Scheme 5*. Detoxification by hydrazinolysis conserved the structure better than NaOH saponification for *Vibrio cholerae* O1 LPS [119,120]. The remaining detoxified LPS is then activated by Method A (*Scheme 2*) and conjugated to cholera toxin B-subunit, a pathogen related carrier. Immunization of 38 human volunteers with two doses of vaccine [78] induces anti-LPS IgM, IgG responses, and vibriocidal titers comparable to those immunized with whole-cell vaccine.

A number of bacteria display LPS with short *O*-polysaccharide chains. These short *O*-polysaccharide chains are more conserved and could potentially afford a broader protection. Once detoxified, the strategy for conjugation differs from that previously used for LPS O-polysac-





Scheme 5

(a) Structure of the *O*-polysaccharide from *V. cholerae* O1 Inaba after detoxification followed by modification using Method A. (b) Structure of LPS from non-typable *Haemophilus influenzae*. The sites and type of modification are shown

charide. Random activation of such a small carbohydrate is impossible without affecting the integrity of the molecule and its recognition. Terminal activation is more suitable and is used for example in the modification of non-typable Hi-LPS [121]. The detoxified LPS are activated via the residual KDO-COOH by ADH and carbodiimide. Finally, the conjugation is performed in a similar fashion as described in Method D (*Scheme 2*).

3.5 Properties of Conjugates

Carbohydrate-protein conjugates stimulate the immune response through a very complex process. In a simplified model it involves the binding of the glycoconjugate to the surface immunoglobulin of the follicular B cell population with appropriate specificity for the saccharide hapten [122]. The glycoconjugate is endocyted and the carrier is fragmented into peptides. Selected peptides that bind to major histocompatibility complex class II molecules are displayed on the cell surface and recognized by peptide-specific T-cell receptors on the T-cell [123]. The cognate interaction between PS-specific B-cell and T-cell is modulated by interaction not only of MHC II-TCR, but also by CD40-CD40L and CD28-B7-1/-2 [124]. The interaction induces activation of the B-cell with differentiation into antibody secreting cells, germinal-center formation, class switch recombination, and somatic hypermutation. Memory B cells generated during this T cell-dependent process can be activated by subsequent stimulation with either PS-protein conjugates or by the CPS alone or by the whole-bacteria.

Glycoconjugate vaccines are considered biological pharmaceuticals. With the upsurge of recombinant biological pharmaceuticals, an evolution from traditional bioassay to physicochemical characterization is taking place. This evolution influences strongly the development of more and more complex glycoconjugate formulations as one of the limiting steps was the inability to fully characterize the products or their combination. Actually, modern physicochemical methods allow glycoconjugates to be characterized with a degree of precision previously not possible [125,126]. A combination of methods such as circular dichroism, size exclusion chromatography, high performance anion exchange chromatography, mass spectrometry, and NMR spectroscopy [127,128] are used.

The first glycoconjugate vaccine introduced into clinical practice was *Haemophilus influenzae* type b in the 1990s. After initial introduction in industrialized countries, the vaccination coverage was extended slowly to cover 26% of children worldwide. The World Health Organization recommended the inclusion of anti-Hib in the extended program of immunization [139].

Glycoconjugate vaccines induce a strong IgG immune response in infants with memory and also prevent nasopharyngeal Hib colonization. Few vaccines in history have induced such a dramatic decline in incidence over such a short period as have the Hib conjugates [140]. At present, six different glycoconjugate vaccines are commercially available (see *Table 3*). The formulations contain the vaccine either alone or in a preferred pentavalent combination containing conjugate anti-Hib, anti-hepatitis B and the classical Diphtheria, Pertussis, and Tetanus combination [141].

The spectacular success of anti-Hib conjugate vaccine prompted the development and introduction of other glycoconjugate vaccines. Several meningococcal group C-tetanus toxoid conjugates based on periodate oxidize polysaccharides were licensed. Glycoconjugates for other meningococcal groups were also developed and are now available in bivalent AC or tetravalent A, C, Y, and W135 combination. The development of multivalent glycoconjugate vaccine introduces further complexity. A challenge for modern technology was the development of multivalent glycoconjugate vaccine against *S. pneumoniae.* Prevenar, a licensed heptavalent vaccine containing serotypes 4, 6B, 9V, 14, 18C, 19F, and 23F conjugate to recombinant CRM_{197} protein induced a dramatic decline in bacterial pneumoniae in the US and other countries [134]. However, this vaccine could not be considered universal because serotype distribution of Sp depends strongly on geographical regions [142]. Vaccines containing more than ten

Table 3
Commercially available conjugate vaccines

Vaccine	CPS	Carrier	Company	Ref.
HibTITER	Hib	CRM_{197}	Wyeth Vaccines	[129]
PedvaxHIB	Hib	OMP Nm	Merck	[130]
ActHIB	Hib	TT	Sanofi Pasteur	[131]
ProHIBIT	Hib	DT	Connaught Laboratories	[132]
Quimi-Hib	Hib (synthetic)	TT	Heber-Biotec	[133]
Prevnar	Sp 4, 6B, 9V, 14, 18C, 19F and 23F	CRM_{197}	Wyeth Lederle	[134]
MENINGITEC	Nm C	CRM_{197}	Wyeth Lederle	[135]
MENJUGATE	Nm C	TT	Baxter	[136]
NEISVAC-C	Nm C	CRM_{197}	Chiron	[137]
Menactra	Nm A,C, Y and W_{135}	DT	Sanofi Pasteur	[138]

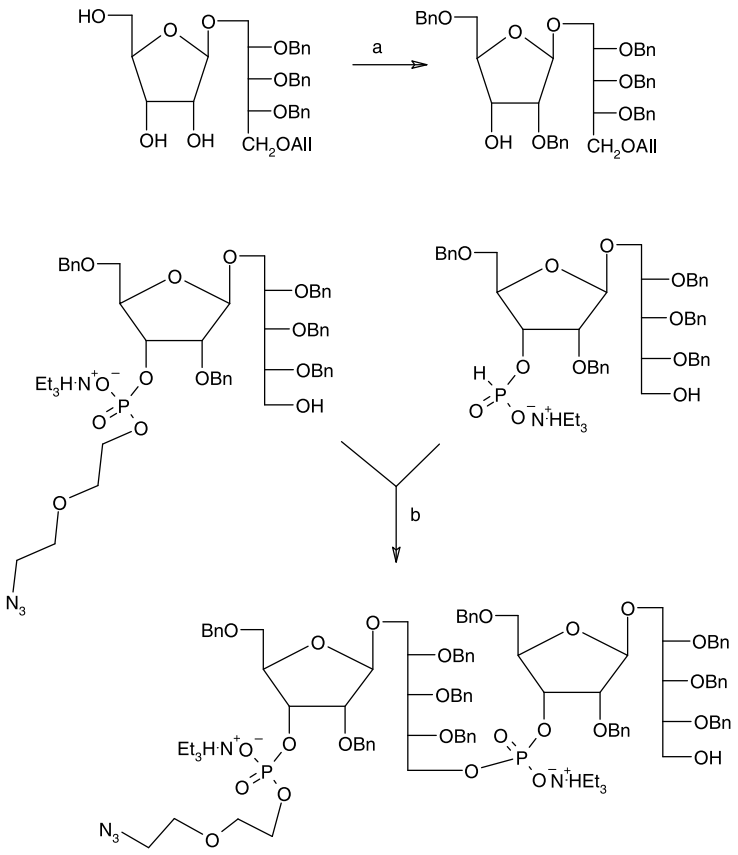
serotypes are in an advanced stage of development [108,143] and this shows the level attained by modern technology.
Vaccination with Prevenar reduced seven serotypes of Sp included in the conjugate vaccines. An increase in circulation of serotype 19A, not included in a vaccine, warns of the possibility of a replacement phenomenon [144]. Capsular polysaccharide serogroup seems not to be a fixed attribute of a strain for Nm [145] and Sp [146], as the exchange of single gene cassettes involved in CPS biosynthesis is a frequent process that could affect the speed of a replacement phenomenon.
Glycoconjugate vaccines from other capsular polysaccharides are in preclinical or clinical evaluation steps for Group B *Streptococcus*, *Salmonella typhi*, Shigellosis, Cholera, and *Staphylococcus aureus*.

3.6 Synthetic Carbohydrate-Protein Conjugate Vaccines

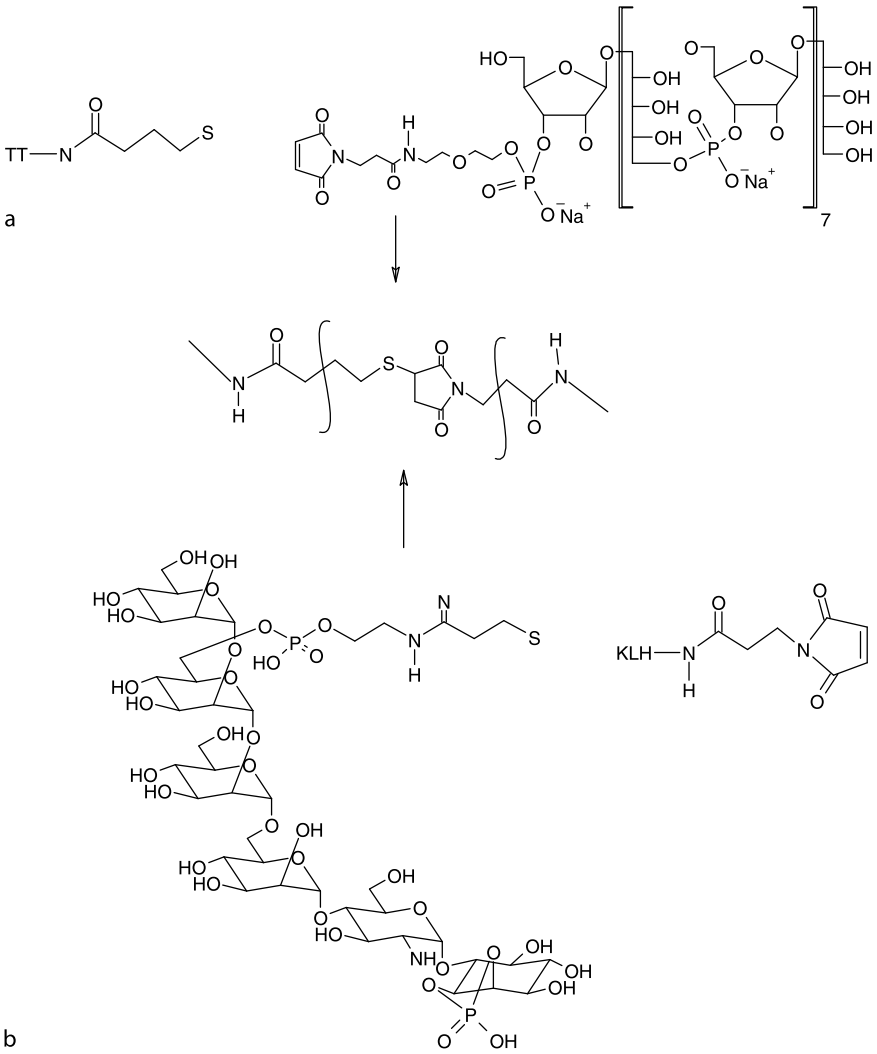
The production of CPS by bacteria fermentation-extraction-purification procedures presently reaches a very efficient level. Purified CPS or LPS should be chemically modified for conjugation to a protein by a very complex procedure and give a final product that needs to be as defined as possible. If the production or purification of the carbohydrate material from the natural source is difficult and one bears in mind the complexity of producing a consistent glycoconjugate from a natural polymer then opportunities for modern synthetic carbohydrate chemistry are at hand.
The field is actually in a good position to contribute to the development of bacterial polysaccharide glycoconjugate vaccines. Numerous fragments of capsular polysaccharides, lipopolysaccharides, or other bacterial carbohydrates have been synthesized and their protein conjugates obtained [87].
They supply valuable information on the influence of saccharide size and structure on immunogenicity. The use of synthetic oligosaccharides in the development of glycoconjugate vaccines for human use is, however, a very complex task far beyond a laboratory synthesis of a given oligosaccharide and oligosaccharide-protein conjugate.

Anti-Hib vaccine was the first glycoconjugate to be developed up to clinical use in humans. The structure of the Hib-CPS was an attractive target for synthetic carbohydrate chemistry in the 1990s. Accordingly, several syntheses of Hib-oligosaccharides and their protein conjugates were developed. An important step was the demonstration that a tetrameric oligosaccharide-tetanus toxoid conjugate induced anti CPS responses with an increasing IgG/IgM ratio in adult mice and monkeys [147] and that the synthesis of an oligosaccharide fragment having three repeating units coupled to T helper synthetic peptides raised protective antibodies against Hib infection in an infant rat model [148].

For anti-Hib vaccine, synthetic chemistry was compelled to reach the efficiency level attained during the production of the glycoconjugate vaccine from natural CPS. The contribution of several groups [149] to the synthesis of Hib-oligosaccharides paved the way for further improvement. A simplified synthesis of the repeating disaccharide unit and chain elongation process in one step through a polycondensation reaction were essential achievements before further developments and clinical evaluation were possible (see ❯ *Scheme 6*). The synthetic



❑ Scheme 6
Reagents (a) BnCl, Bu_2SnO, Bu_4NI, NaH, Tol; (b) PivCl; I_2; Gel filtration



Scheme 7
Simplified structure of synthetic oligosaccharide fragment-protein conjugates using thiol-maleimido chemistry. (a) Conjugation of synthetic Hib-CPS fragment to thiolated Tetanus Toxoid. (b) Conjugation of *Plasmodium falciparum* synthetic pentasaccharide to KLH

oligosaccharide fragment with a terminal spacer arm activated through a maleimido function could be coupled to thiolated-TT with good yields and reproducibility [150] (see *Scheme 7*). The clinical evaluation in infants showed a comparable high efficiency for the vaccine containing the synthetic antigen against Hib infection [133] allowing licensing and use of the vaccine in infants since 2004.

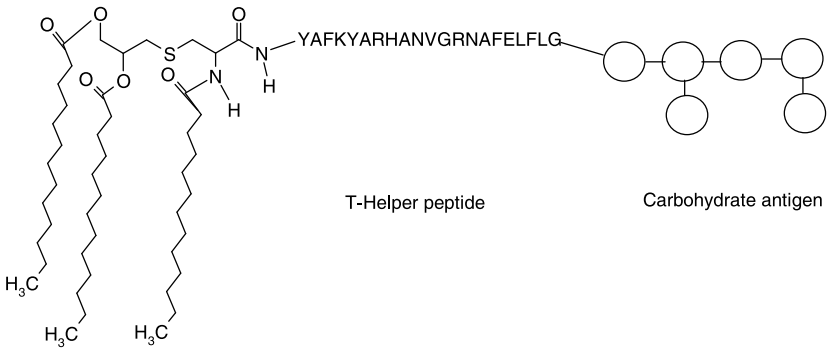
Activity in the field has increased in the last few years and other projects targeting the synthesis of CPS or LPS fragments are in development. Citing only some of them, vaccines against

shigellosis [151,152], cholera [153], and Sp [154] are in preclinical or early clinical evaluation. While these projects afford valuable information on the antigen-antibody recognition, their feasibility and competing advantages over natural sources of antigens for the production of glycoconjugate vaccines remains to be demonstrated. Other projects as for example malaria [155] which identifies a pentasaccharide as an important carbohydrate antigen associated with the toxic effect of Plasmodium have additional opportunities (see *Scheme 7*). In this particular case the improved syntheses as well as the difficulties in isolating the carbohydrate from natural sources are positive circumstances.

4 Future Prospects

The impact on health care of polysaccharide-based vaccines for bacterial diseases in the last 15 years attracted much attention to this research field [156]. The evolution is not only toward the development of new vaccines for unsolved diseases but also for increasing the efficiency of existing vaccines. Noteworthy, a fully designed synthetic molecule induces in laboratory animals an improved anti-carbohydrate response [157]. Although there are several explored and potential designs for this new generation anti-carbohydrate vaccine, the selected example (see *Scheme 8*) contains three separate parts in a fully synthetic molecule: (a) carbohydrate antigen, (b) universal T-helper peptide, (c) lipopeptide with a dual function; anchored to a liposome and support as a ligand for toll-like receptors. Constructions like that will certainly impact the future technology.

Diverging with the previous tendency but with similar promise is the production of carbohydrate-protein glycoconjugates directly in a single bacterium. *Escherichia coli* was engineered to do three operations that usual technology performs separately. The recombinant bacteria produce the polysaccharide and the protein carrier and also transfer the polysaccharide to a protein carrier with the help of oligosaccharyltransferase in a way similar to the *N*-linked glycosylation in higher organisms [158].



Scheme 8

References

1. Heidelberger M, Avery OT (1923) J Exp Med 38:73
2. Robbins JB, Schneerson R, Szu SC (1997) In: Levine MM, Woodrow GC, Kaper JB Cobon GS (eds) New Generation Vaccines. Marcel Dekker, New York, p 803
3. Zähringer U, Lindner B, Rietschel ETh (1994) Adv Carbohydr Chem Biochem 50:211
4. MacLeod CM, Hodges RG, Heidelberger M, Bernhard WG (1945) J Exp Med 82:445
5. Jennings H (1984) Adv Carbohydr Chem Biochem 42:155
6. Henrichsen J (1995) J Clin Microbiol 33:2759
7. Cadoz M, Prince-DM, Diop MI, Denis F (1983) Pathol Biol 31:128
8. Girard MP, Preziosi MP, Aguado MT, Kieny MP (2006) Vaccine 24:4692
9. Hausdorff W, Bryant J, Paradiso P, Siber G (2000) Clin Infect Dis 30:100
10. Hausdorff W, Siber G, Paradiso P(2001) Lancet 357:950
11. Cherian T(2007) PneumoADIP and Hib initiative surveillance Networks Investigators Meeting, Bangkok, Thailand
12. Anderson P, Peter G, Johnston RB, Wetterlow H, Smith DH (1972) J Clin Invest 51:39
13. Heidelberger M, MacLeod CM, Di Lapi MM (1949) J Exp Med 88:369
14. Pittman M (1931) J Exp Med 53:471
15. Moxon ER, Kroll JS (1990) Curr Topics Microbiol Immun 150:65
16. Jones C (1998) Carbohydrates in Europe 21:10
17. Francis TJr, Tillet WS (1930) J Exp Med 52:573
18. Goldschneider I, Gotschlich EC, Artenstein MS (1969) J Exp Med 129:1307
19. Cadoz M (1998) Vaccine 16:1391
20. Gotschlich EC, Goldschneider I, Artenstein MS (1969) J Exp Med 129:1385
21. Austrian R (1981) Rev Infect Dis 3 (suppl):S1
22. Robbins JB, Austrian R, Lee CJ, Rastogi SC, Schiffman G, Henrichsen J, Mäkelä PH, Broome CV, Facklam RR, Tiesjema RH, Parke JC (1983) J Infect Dis 148:1136
23. Jennings HJ (1998) Carbohydrates in Europe 21:17
24. Rautonen N, Pelkonen J, Sipinen S, Käyhty H, Mäkelä O (1986) J Immunol 137:2670
25. Lucas AH, Granoff DM (2001) Pediatr Infec Dis J 20:235
26. Lucas AH, Reason DC (1999) Immunol Rev 171:89
27. Mäkelä PH, Peltola H, Käyhty H (1977) J Infect Dis 136:543
28. Robbins JB, Schneerson R, Gotschlich (1997) Lancet 350:1709
29. Reingold AI, Broome CV, Hightower AW, Ajello GW, Bolan GA, Adamsbaum C, Jones EE, Phillips C, Tiendrebeogo H, Yada A (1985) Lancet 2:114
30. Douglas RM, Paton JC, Duncan SJ, Hansman DJ (1983) J Inf Dis 148:131
31. Shelly MA, Jacoby H, Riley GJ, Draves BT, Pichichero M, Treanor JJ (1997) Infect Immun 65:242
32. Berry DS, Lynn F, Lee CH, Frasch CE, Bash MC (2002) Infect Immun 70:3707
33. Guy RCE, How MJ, Stacey M, Heidelberger M (1967) J Biol Chem 242:5106
34. Kennedy DA, Buchanan JG, Baddiley J (1969) Biochem J 115:37
35. Perry MB, Bundle DR, Daoust V, Carlo DJ (1982) Mol Immun 19:235
36. Chitenden GJF, Roberts WK, Buchanan JG, Badiley J (1968) Biochem J 109:597
37. Peltola H, Safary A, Käyhty H, Karankoo V, Andre TE (1985) Pediatrics 76:91
38. Arakere G, Frasch CE (1991) Infect Immun 59:4349
39. Richards JC, Perry MB, Carlo DJ (1983) J Biochem Cell Biol 61:178
40. Bennett LG, Bishop CT (1977) Immunochemistry 14:693
41. Crisel RM, Baker RS, Dorman DE (1975) J Biol Chem 250:4926
42. Bundle DR, Smith IC, Jennings HJ (1974) J Biol Chem 249:2275
43. Bhattacharjee AK, Jenning HJ, Kenny CP, Martins A, Bundle DR (1975) J Biol Chem 250:1926
44. Bhattacharjee AK, Jenning HJ, Kenny CP, Martins A, Smith ICP (1976) Can Biochem 54:1
45. Lindberg B, Lindqvist J, Lönngren J, Powell DA (1980) Carbohydr Res 78:111
46. Jansson PE, Lindberg B, Andersson M, Lindquist U, Henrichsen J (1988) Carbohydr Res 182:111
47. Reeves RE, Goebel WF (1941) J Biol Chem 139:511
48. Jones C (1990) Carbohydr Res 198:353

49. Jansson PE, Lindberg B, Lindquist U (1985) Carbohydr Res 140:101
50. Kenne L, Lindberg B, Madden JK (1979) Carbohydr Res 73:175
51. Moreau M, Richards JC, Perry MB, Kniskern PJ (1988) Carbohydr Res 182:79
52. Jones JKN, Perry MB (1957) J Am Chem Soc 79:2787
53. Jones C, Mulloy B, Wilson A, Dell A, Oates JE (1985) J Chem Soc, Perkin Trans 1:1665
54. Rutherford TJ, Jones C, Davies DB, Elliott AC (1994) Carbohydr Res 218:175
55. Leontein K, Lindberg B, Lönngren J, Carlo DJ (1983) Carbohydr Res 114:257
56. Lindberg B, Lönngren J, Powell DA (1977) Carbohydr Res 58:111
57. Jansson PE, Lindberg B, Lindquist U, Ljungberg J (1987) Carbohydr Res 162:111
58. Jennings HH, Pon RA (1996) Polysaccharides and glycoconjugates as human vaccines. In: Dumitriu S (ed) Polysaccharides in Medicinal Applications. Marcel Dekker, New York, p 443
59. Lindberg J (1990) PhD thesis, University of Stockholm, Sweden
60. Ohno N, Yadomae T, Miyazaki (1980) Carbohydr Res 80:297
61. Katzenellenbogen E, Jennings HJ (1983) Carbohydr Res 124:35
62. Beynon LM, Richards JC, Perry MB (1997) Eur J Biochem 250:163
63. Richards JC, Perry MB, Kniskern PJ (1989) Can J Chem 67:638
64. Richards JC, Perry MB (1988) Biochem Cell Biol 66:758
65. Richards JC, Perry MB, Kniskern PJ (1984) Can J Biochem Cell Biol 62:666
66. Stein KE (1992) J Infect Dis 165(suppl 1):S49
67. Roantree RJ (1967) Annu Rev Microbiol 21:443
68. Peltola H, Käyhty H, Sivonen A, Mäkelä PH (1977) Pediatrics 60:730
69. Kuby J (1997) Immunology, 3rd edn. Freeman WH and company, New York
70. Avery OT, Goebel WF (1931) 54:437
71. Schneerson R, Barrera O, Sutton A, Robbins JB (1980) J Exp Med 152:361
72. Klein DL, Ellis RW (1997) In: Levine MM, Woodrow GC, Kaper JB and Cobon GS (eds) New Generation Vaccines. Marcel Dekker, New York, p 503
73. Zollinger WD (1997) In: Levine MM, Woodrow GC, Kaper JB and Cobon GS (eds) New Generation Vaccines. Marcel Dekker, New York, p 469
74. Lagergard T, Shiloach J, Robbins JB, Schneerson R (1990) Infect Immun 58:687
75. Wessels MR, Paoletti LC, Kasper DL, DiFabio JL, Michon F, Holme K, Jennings HJ (1990) J Clin Invest 86:1428
76. Szu S, Stone AL, Robbins JD, Robbins JB, Schneerson R (1987) J Exp Med 166:1510
77. Fattom A, Naso RB (1997) In: Levine MM, Woodrow GC, Kaper JB and Cobon GS (eds) New Generation Vaccines. Marcel Dekker, New York, p 979
78. Gupta RK, Taylor RK, Bryla DA, Robbins JB, Szu SC (1998) Infect Immun 66:3095
79. Boutonnier A, Villeneuve S, Nato F, Dassy B, Fournier JM (2001) Infect Immun 69:3488
80. Ashkenazi S, Passwell JH, Harlev E, Mirror D, Dagan R, Farzan N, Ramon R, Majadly F, Bryla DA, Karpass AB, Robbins JB, Schneerson R (1999) J Infect Dis 179:1565
81. Robbins JB (1999) Pure Appl Chem 71:745
82. Kamerling JP (2000) In: Tomasz A (eds) *Streptococcus pneumoniae* – Molecular Biology & Mechanisms of Disease. Mary Ann Liebert, Inc., Larchmont, NY, USA, p 81
83. Jennings HJ, Sood RK (1994) Synthetic Glycoconjugates as humans vaccines. In: Lee YC and Lee RT (eds) Neoglycoconjugates: Preparation and Applications. Academic Press, San Diego, p 325
84. Sepala I and Mäkelä O (1989) J Immunol 143:1259
85. Fernández C, Sverremark E (1994) Cell Immunol 153:67
86. Romero-Steiner S, Fernández J, Biltoft Ch, Wohl ME, Sanchez J, Feris J, Balter S, Levine O, Carlone GM (2001) 8:1115
87. Pozsgay V (2000) Adv Carbohyd Chem Biochem 56:153
88. Lifely MR, Mareno C, Lindon JC (1987) Vaccine 5:11
89. Parisi L, von Hunolstein C (1999) J Chromatog A 847:209
90. Jennings HJ (1990) Curr Top Microbiol Immunol 150:97
91. Anderson P (1983) Infect Immun 39:233
92. Finne J, Makela PH (1985) J Biol Chem 260:1265
93. Szu SC, Zon G, Schneerson R, Robbins JB (1986) Carbohydr Res 155:7
94. Pawlowski A, Svenson SB (1999) FEMS Microbiol Lett 174:255
95. Constantino P, NOrelli F, Giannozzi A, D'Ascenzi S, Bartolini A, Kaur S, Tang D,

Seid R, Viti S, Paffeti R, Bigio M, Pennantini C, Averani G, Guamieri V, Gallo E, Ravenscroft N, Lazzeroni C, Rappuoli R, Ceccarini C (1999) Vaccine 17:1251
96. Laferriére CA, Sood RK, de Muys JM, Michon F, Jennings HJ (1997) Vaccine 15:179
97. Wessels MR, Poszgay V, Kasper DL, Jennings HJ (1987) J Biol Chem 262:8262
98. Laferriere CA, Sood RK, De Muys JM, Michon F, Jennings HJ (1998) Infect Immun 66:2441
99. Beuvery EC, Roy R, Kanhai V, Jennings HJ (1986) Dev Biol Standard 65:197
100. Anderson PW, Pichichero ME, Insel RA, Betts R, Eby R, Smith DH (1986) J Immunol 137:1181
101. Fernández Santana V, Pena Icart L, Beurret M, Costa L, Verez Bencomo V (2006) Methods Enzymol 415:153
102. Chu CY, Schneerson R, Robbins JB, Rastogi SC (1983) Infect Immun 40:245
103. Marburg S, Jorn D, Tolman RL, Arison B, McCanley J, Kniskern PJ, Hagopian A, Vella PP (1986) J Amer Chem Soc 108:5282
104. Porro M, Constantino S, Viti S, Vannozzi F, Naggi A, G Torri (1985) Mol Immun 22:907
105. Peeters CCAM, Lagerman PR, Weers O, Oomen LA, Hoogerhout P, Beurret M, Poolman JT (2003) Preparation of polysaccharide-conjugate vaccines. In: Robinson A, Hudson MJ, Cranage MP (eds) Vaccine Protocols, 2nd ed. Humana Press, Totowa, New Jersey, p 153
106. Uchida T, Pappenheimer AM Jr, Harper AA (1972) Science 175:901
107. Barington T, Skettrup M, Juul L, Heilmann C (1993) Infect Immun 61:432
108. Prymula R, Peeters P, Chrobok V, Kriz P, Novakova E, Kaliskova E, Kohl I, Lommel P, Poolman J, Prieels J, Schueman L (2006) Lancet 367:740
109. Crowley LA, Reddin KM, Gorringe AR, Hudson MJ, Robinson A (2001) Vaccine 19:3399
110. Pérez AE, Dickinson FO, Banderas F, Serrano T, Llanes R, Guzmán D, Díaz P, Alvarez A, Guirola M, Caballero E, Canaan-Haden L, Guillén G (2006) FEMS Immunol Med Microbial 46:386
111. Fusco PC, Michon F, Laude-Sharp M, Minetti CA, Huang CH, Heron I, Blake MS (1998) Vaccine 16:1842
112. Michon F, Fusco PC, Minetti CA, Laude-Sharp M, UitzC, Huang CH, D'Ambra AJ, Moore S, Remeta DP, Heron I, Blake MS (1998) Vaccine 16:1732
113. Ravenscroft N, Jones C (2000) Curr Opin Drug Develop 3:222
114. Beuvery EC, van de Kaaden A, Kanhai V, Leussink AB (1983) Vaccine 1:31
115. Jennings HJ, Lugowski C (1981) J Immunol 127:1011
116. Beuvery EC, Miedema F, van Delft R, Haverkamp J (1983) Infect Immun 40:39
117. Costantino P, Viti S, Vannozzi F, Serafín G Marsili I, Valeri A (1986) Ann Sclavo Collana Monogr 3:359
118. Shafer DE, Toll B, Schuman RF, Nelson BL, Mond JJ, Lees A (2000) Vaccine 18:1273
119. Taylor RK, Kim TJ, Bose N, Stonehouse E, Tripathi SA, Kovac P, Wade WF (2004) Chem Biodiversity 1:1036
120. Gupta RK, Szu SC, Finkelstein RA, Robbins JB (1992) Infect Immun 60:3201
121. Wu TH, Gu XX (1999) Infect Immun 67:5508
122. Siber GR (1994) Science 265:1385
123. Guttormsen HK, Sharpe AH, Chandraker AK, Brigtsen AK, Sayegh MH, Kasper DL (1999) Infect Immun 67:6375
124. Lucas AH, Apicella MA, Taylor CE (2005) Clin Infect Dis 41:705
125. Egan W, Frasch CE, Anthony BF (1995) JAMA 273:888
126. Holliday MR, Jones C (1999) Biologicals 27:51
127. Jones C, Lemercinier X, Crane DT, Gee CK, Austin S (2000) Dev Biol (Basel) 103:121
128. Ravenscroft N, D'Ascenzi S, Proietti D, Norelli F, Costantino P (2000) Dev Biol (Basel) 103:35
129. Black SB, Shinefield HR, Fireman B, Hiatt R, Polen M, Vittinghoff E (1991) Pediatr Infect Dis J 10:97
130. Santosham M, Wolff M, Reld R (1991) N Engl J Med 324:1767
131. Booy R, Hodgson S, Carpenter L (1994) Lancet 344:362
132. Eskola J, Kahty H, Takala AK (1990) N Engl J Med 323:1381
133. Verez-Bencomo V, Fernández-Santana V, Hardy E, Toledo ME, Rodríguez MC, Heynngnezz L, Rodríguez A, Baly A, Herrera L, Izquierdo M, Villar A, Valdés Y, Cosme K, Deler ML, Montane M, García E, Ramos A, Aguilar A, Medina E, Toraño G, Sosa I, Hernández I, Martínez R, Muzachio A, Carmenates A, Costa L, Cardoso F, Campa C, Díaz M, Roy R (2004) Science 305:522

134. O'Brien KL, Santosham M (2004) Am J Epidemiol 159:634
135. Richmond P, Borrow R, Miller E, Clark S, Sadler F, Fox A (1999) J Infect Dis 179:1569
136. Richmond P, Borrow R, Goldblatt D, Findlow J, Martin S, Morris R (2001) J Infect Dis 183:160
137. Richmond P, Goldblatt D, Fusco PC, Fusco JD, Heron I, Clark S, Borrow R, Michon F (1999) Vaccine 18:641
138. Lagos R, Papa T, Muñoz A, Ryall R, Piña M, Bassily E (2005) Hum Vaccin 1(6):228
139. WHO/UNICEF immunization coverage estimates, data as of September 2006
140. Peltola H (2000) Clin Rev Microbiol 13:302
141. Tregnaghi M, López P, Rocha C, Rivera L, David MP, Rüttimann R, Schuerman L (2006) Pan Amer J Public Health 19:179
142. Meeting report (2007) Vaccine 25:6557
143. Puumalainen T, Zeta-Capeding M, Kayhty H (2002) Pediatr Infect Dis J 21:309
144. Singleton RJ, Hennessy TW, Bulkow LR, Hammitt LL, Zulz T, Hurlburt DA, Butler JC, Rudolph K, Parkinson A (2007) JAMA 297:1784
145. Swartley JS, Marfin AA, Edupuganti S, Liu LJ, Cieslak P, Perkins B, Wengler JD, Stephens DS (1997) Proc Natl Acad Sci USA 94:271
146. Coffey TJ, Enright ME, Daniels M, Morona JK, Hryniewicz W, Paton JC, Spratt BG (1998) Mol Microbiol 27:73
147. Peeters CC, Evenberg D, Hoogerhout P, Kayhty H, Saarinen L, Van Boeckel CA, Van der Marel GA, Van Boom JH and Poolman JT (1992) Infect Immun 60:1826
148. Chong P, Chan N, Kandil A, Triplet B, James O, Yang YP, Shi SP and Klein M (1997) Infect Immun 65:4918
149. Verez Bencomo V, Roy R, Rodríguez MC, Villar A, Fernandez-Santana V, García E, Valdes Y, Heynngnezz L, Sosa S, Medina E (2007) Carbohydrate Vaccines in ACS Symposium series, p 40
150. Fernández-Santana V, González-Lio R, Sarracent-Perez J, Verez-Bencomo V (1998) Glycoconj J 306:163
151. Pozsgay V(1998) J Org Chem 63:5983
152. Bélot F, Wright K, Costachel C, Phalipon A, Murlard LA (2004) J Org Chem 69:1060
153. Ma X, Saksena R, Chernyak A, Kovac P (2003) Org Biomol Chem 1:775
154. Benaissa-Trouw B, Lefeber DJ, Kamerling JP, Vliegenthart JFG, Kraaijeveld K, Snipe H (2001) Infect Immun 69:4698
155. Schofield L, Hewitt MC, Evans K, Siomos A, Seeberger PH (2002) Nature 418:785
156. Roy R (2004) Drug Discov Today 327
157. Buskas T, Ingale S, Boons GJ (2005) Angew Chem Int Ed 44:5985
158. Kowarik M, Numao S, Feldman MF, Schulz BL, Callewaert N, Kiermaier E, Catrein I, Aebi M (2006) Science 314:1148

Appendix

Nomenclature of Carbohydrates

(Recommendations 1996)

International Union of Pure and Applied Chemistry
and
International Union of Biochemistry and Molecular Biology

Joint Commission on Biochemical Nomenclature*

Prepared for publication by
Alan D. McNaught

The Royal Society of Chemistry, Thomas Graham House, Science Park, Milton Road, Cambridge CB4 4WF, UK

*Members of the Commission (JCBN) at various times during the work on this document (1983–1996) were as follows:

Chairmen: H.B.F. Dixon (UK), J.F.G. Vliegenthart (Netherlands), A. Cornish-Bowden (France); ***Secretaries:*** A. Cornish-Bowden (France), M.A. Chester (Sweden), A.J. Barrett (UK), J.C. Rigg (Netherlands); ***Members:*** J.R. Bull (RSA), R. Cammack (UK), D. Coucouvanis (USA), D. Horton (USA), M.A.C. Kaplan (Brazil), P. Karlson (Germany), C. Lièbecq (Belgium), K.L. Loening (USA), G.P. Moss (UK), J. Reedijk (Netherlands), K.F. Tipton (Ireland), S. Velick (USA), P. Venetianer (Hungary).

Additional contributors to the formulation of these recommendations:

Expert Panel: D. Horton (USA) (*Convener*), O. Achmatowicz (Poland), L. Anderson (USA), S.J. Angyal (Australia), R. Gigg (UK), B. Lindberg (Sweden), D.J. Manners (UK), H. Paulsen (Germany), R. Schauer (Germany).

Nomenclature Committee of IUBMB (NC-IUBMB) (those additional to JCBN): A. Bairoch (Switzerland), H. Berman (USA), H. Bielka (Germany), C.R. Cantor (USA), W. Saenger (Germany), N. Sharon (Israel), E. van Lenten (USA), E.C. Webb (Australia).

American Chemical Society Committee for Carbohydrate Nomenclature: D. Horton (*Chairman*), L. Anderson, D.C. Baker, H.H. Baer, J.N. BeMiller, B. Bossenbroek, R.W. Jeanloz, K.L. Loening, W.A. Szarek, R.S. Tipson, W.J. Whelan, R.L. Whistler.

In: *Glycoscience*. Fraser-Reid B, Tatsuta K, Thiem J (eds)
Chapter-DOI 10-1007/978-3-540-30429-6_70 © IUPAC (1996) *Pure and Applied Chemistry* **68**, 1919–2008

Corresponding Members of the ACS Committee for Carbohydrate Nomenclature (other than JCBN and the expert panel): R.F. Brady (USA), J.S. Brimacombe (UK), J.G. Buchanan (UK), B. Coxon (USA), J. Defaye (France), N.K. Kochetkov (Russia), R.U. Lemieux (Canada), R.H. Marchessault (Canada), J.M. Webber (UK).

Correspondence on these recommendations should be addressed to Dr Alan D. McNaught at the above address or to any member of the Commission.

Preamble

These Recommendations expand and replace the Tentative Rules for Carbohydrate Nomenclature [1] issued in 1969 jointly by the IUPAC Commission on the Nomenclature of Organic Chemistry and the JUB-IUPAC Commission on Biochemical Nomenclature (CBN) and reprinted in [2]. They also replace other published JCBN Recommendations [3,4,5,6,7] that deal with specialized areas of carbohydrate terminology; however, these documents can be consulted for further examples. Of relevance to the field, though not incorporated into the present document, are the following recommendations:

- Nomenclature of cyclitols, 1973 [8]
- Numbering of atoms in *myo*-inositol, 1988 [9]
- Symbols for specifying the conformation of polysaccharide chains, 1981 [10]
- Nomenclature of glycoproteins, glycopeptides and peptidoglycans, 1985 [11]
- Nomenclature of glycolipids, in preparation [12]

The present Recommendations deal with the acyclic and cyclic forms of monosaccharides and their simple derivatives, as well as with the nomenclature of oligosaccharides and polysaccharides. They are additional to the Definitive Rules for the Nomenclature of Organic Chemistry [13,14] and are intended to govern those aspects of the nomenclature of carbohydrates not covered by those rules.

2-Carb-0 Historical Development of Carbohydrate Nomenclature [15]

2-Carb-0.1 Early Approaches

In the early nineteenth century, individual sugars were often named after their source, e. g. grape sugar (Traubenzucker) for glucose, cane sugar (Rohrzucker) for saccharose (the name sucrose was coined much later). The name glucose was coined in 1838; Kekulé in 1866 proposed the name ‘dextrose’ because glucose is dextrorotatory, and the laevorotatory ‘fruit sugar’ (Fruchtzucker, fructose) was for some time named ‘laevulose’ (American spelling ‘levulose’).

Very early, consensus was reached that sugars should be named with the ending '-ose', and by combination with the French word 'cellule' for cell the term cellulose was coined, long before the structure was known. The term 'carbohydrate' (French 'hydrate de carbone') was applied originally to monosaccharides, in recognition of the fact that their empirical composition can be expressed as $C_n(H_2O)_n$. However the term is now used generically in a wider sense (see ❯ *2-Carb-1.1*).

2-Carb-0.2 The Contribution of Emil Fischer

Emil Fischer [16] began his fundamental studies on carbohydrates in 1880. Within ten years, he could assign the relative configurations of most known sugars and had also synthesized many sugars. This led to the necessity to name the various compounds. Fischer and others laid the foundations of a terminology still in use, based on the terms triose, tetrose, pentose, and hexose. He endorsed Armstrong's proposal to classify sugars into aldoses and ketoses, and proposed the name fructose for laevulose, because he found that the sign of optical rotation was not a suitable criterion for grouping sugars into families.

The concept of stereochemistry, developed since 1874 by van't Hoff and Le Bel, had a great impact on carbohydrate chemistry because it could easily explain isomerism. Emil Fischer introduced the classical projection formulae for sugars, with a standard orientation (carbon chain vertical, carbonyl group at the top); since he used models with flexible bonds between the atoms, he could easily 'stretch' his sugar models into a position suitable for projection. He assigned to the dextrorotatory glucose (via the derived glucaric acid) the projection with the OH group at C-5 pointing to the right, well knowing that there was a 50% chance that this was wrong. Much later (Bijvoet, 1951), it was proved correct in the absolute sense.

Rosanoff in 1906 selected the enantiomeric glyceraldehydes as the point of reference; any sugar derivable by chain lengthening from what is now known as D-glyceraldehyde belongs to the D series, a convention still in use.

2-Carb-0.3 Cyclic Forms

Towards the end of the nineteenth century it was realized that the free sugars (not only the glycosides) existed as cyclic hemiacetals or hemiketals. Mutarotation, discovered in 1846 by Dubrunfaut, was now interpreted as being due to a change in the configuration of the glycosidic (anomeric) carbon atom. Emil Fischer assumed the cyclic form to be a five-membered ring, which Tollens designated by the symbol <1,4>, while the six-membered ring received the symbol <1,5>.

In the 1920s, Haworth and his school proposed the terms 'furanose' and 'pyranose' for the two forms. Haworth also introduced the 'Haworth depiction' for writing structural formulae, a convention that was soon widely followed.

2-Carb-0.4 Nomenclature Commissions

Up to the 1940s, nomenclature proposals were made by individuals; in some cases they were followed by the scientific community and in some cases not. Official bodies like the Interna-

tional Union of Chemistry, though developing and expanding the Geneva nomenclature for organic compounds, made little progress with carbohydrate nomenclature. The IUPAC Commission on Nomenclature of Biological Chemistry put forward a classification scheme for carbohydrates, but the new terms proposed have not survived. However in 1939 the American Chemical Society (ACS) formed a committee to look into this matter, since rapid progress in the field had led to various misnomers arising from the lack of guidelines. Within this committee, the foundations of modem systematic nomenclature for carbohydrates and derivatives were laid: numbering of the sugar chain, the use of D and L and α and β, and the designation of stereochemistry by italicized prefixes (multiple prefixes for longer chains). Some preliminary communications appeared, and the final report, prepared by M.L. Wolfrom, was approved by the ACS Council and published in 1948 [17].

Not all problems were solved, however, and different usages were encountered on the two sides of the Atlantic. A joint British-American committee was therefore set up, and in 1952 it published 'Rules for Carbohydrate Nomenclature' [18]. This work was continued, and a revised version was endorsed in 1963 by the American Chemical Society and by the Chemical Society in Britain and published [19]. The publication of this report led the IUPAC Commission on Nomenclature of Organic Chemistry to consider the preparation of a set of IUPAC Rules for Carbohydrate Nomenclature. This was done jointly with the IUPAC-JUB Commission on Biochemical Nomenclature, and resulted in the 'Tentative Rules for Carbohydrate Nomenclature, Part I, 1969', published in 1971/72 in several journals [1]. It is a revision of this 1971 document that is presented here. In the present document, recommendations are designated 2-Carb-*n*, to distinguish them from the Carb-*n* recommendations in the previous publication.

2-Carb-1 Definitions and Conventions

2-Carb-1.1 Carbohydrates

The generic term 'carbohydrate' includes monosaccharides, oligosaccharides and polysaccharides as well as substances derived from monosaccharides by reduction of the carbonyl group (alditols), by oxidation of one or more terminal groups to carboxylic acids, or by replacement of one or more hydroxy group(s) by a hydrogen atom, an amino group, a thiol group or similar heteroatomic groups. It also includes derivatives of these compounds. The term 'sugar' is frequently applied to monosaccharides and lower oligosaccharides. It is noteworthy that about 3% of the compounds listed by Chemical Abstracts Service (i. e. more than 360 000) are named by the methods of carbohydrate nomenclature.

Note. Cyclitols are generally not regarded as carbohydrates. Their nomenclature is dealt with in other recommendations [8,9].

2-Carb-1.2 Monosaccharides

Parent monosaccharides are polyhydroxy aldehydes H-[CHOH]$_n$-CHO or polyhydroxy ketones H-[CHOH]$_n$-CO-[CHOH]$_m$-H with three or more carbon atoms.

The generic term 'monosaccharide' (as opposed to oligosaccharide or polysaccharide) denotes a single unit, without glycosidic connection to other such units. It includes aldoses, dialdos-

es, aldoketoses, ketoses and diketoses, as well as deoxy sugars and amino sugars, and their derivatives, provided that the parent compound has a (potential) carbonyl group.

1.2.1 Aldoses and Ketoses

Monosaccharides with an aldehydic carbonyl or potential aldehydic carbonyl group are called aldoses; those with a ketonic carbonyl or potential ketonic carbonyl group, ketoses.

Note. The term 'potential aldehydic carbonyl group' refers to the hemiacetal group arising from ring closure. Likewise, the term 'potential ketonic carbonyl group' refers to the hemiketal structure (see ❯ *2-Carb-5*).

1.2.2 Cyclic Forms

Cyclic hemiacetals or hemiketals of sugars with a five-membered (tetrahydrofuran) ring are called furanoses, those with a six-membered (tetrahydropyran) ring pyranoses. For sugars with other ring sizes see ❯ *2-Carb-5*.

2-Carb-1.3 Dialdoses

Monosaccharides containing two (potential) aldehydic carbonyl groups are called dialdoses (see ❯ *2-Carb-9*).

2-Carb-1.4 Diketoses

Monosaccharides containing two (potential) ketonic carbonyl groups are termed diketoses (see ❯ *2-Carb-11*).

2-Carb-1.5 Ketoaldoses (Aldoketoses, Aldosuloses)

Monosaccharides containing a (potential) aldehydic group and a (potential) ketonic group are called ketoaldoses (see ❯ *2-Carb-12*); this term is preferred to the alternatives on the basis of ❯ *2-Carb-2.1.1* (aldose preferred to ketose).

2-Carb-1.6 Deoxy Sugars

Monosaccharides in which an alcoholic hydroxy group has been replaced by a hydrogen atom are called deoxy sugars (see ❯ *2-Carb-13*).

2-Carb-1.7 Amino Sugars

Monosaccharides in which an alcoholic hydroxy group has been replaced by an amino group are called amino sugars (see ❯ *2-Carb-14*). When the hemiacetal hydroxy group is replaced, the compounds are called glycosylamines.

2-Carb-1.8 Alditols

The polyhydric alcohols arising formally from the replacement of a carbonyl group in a monosaccharide with a CHOH group are termed alditols (see ❯ *2-Carb-19*).

2-Carb-1.9 Aldonic Acids

Monocarboxylic acids formally derived from aldoses by replacement of the aldehydic group by a carboxy group are termed aldonic acids (see ❯ *2-Carb-20*).

2-Carb-1.10 Ketoaldonic Acids

Oxo carboxylic acids formally derived from aldonic acids by replacement of a secondary CHOH group by a carbonyl group are called ketoaldonic acids (see ❯ *2-Carb-21*).

2-Carb-1.11 Uronic Acids

Monocarboxylic acids formally derived from aldoses by replacement of the CH2OH group with a carboxy group are termed uronic acids (see ❯ *2-Carb-22*).

2-Carb-1.12 Aldaric Acids

The dicarboxylic acids formed from aldoses by replacement of both terminal groups (CHO and CH_20H) by carboxy groups are called aldaric acids (see ❯ *2-Carb-23*).

2-Carb-1.13 Glycosides

Glycosides are mixed acetals formally arising by elimination of water between the hemiacetal or hemiketal hydroxy group of a sugar and a hydroxy group of a second compound. The bond between the two components is called a glycosidic bond.
For an extension of this definition, see ❯ *2-Carb-33*.

2-Carb-1.14 Oligosaccharides

Oligosaccharides are compounds in which monosaccharide units are joined by glycosidic linkages. According to the number of units, they are called disaccharides, trisaccharides, tetrasaccharides, pentasaccharides etc. The borderline with polysaccharides cannot be drawn strictly; however the term ‘oligosaccharide’ is commonly used to refer to a defined structure as opposed to a polymer of unspecified length or a homologous mixture. When the linkages are of other types, the compounds are regarded as oligosaccharide analogues. (See ❯ *2-Carb-37*.)

Note. This definition is broader than that given in [6], to reflect current usage.

2-Carb-1.15 Polysaccharides

'Polysaccharide' (glycan) is the name given to a macromolecule consisting of a large number of monosaccharide (glycose) residues joined to each other by glycosidic linkages. The term poly(g1ycose) is not a full synonym for polysaccharide (glycan) (cf. [20]), because it includes macromolecules composed of glycose residues joined to each other by non-glycosidic linkages.

For polysaccharides containing a substantial proportion of amino sugar residues, the term polysaccharide is adequate, although the term glycosaminoglycan may be used where particular emphasis is desired.

Polysaccharides composed of only one lund of monosaccharide are described as homopolysaccharides (homoglycans). Similarly, if two or more different kinds of monomeric unit are present, the class name heteropolysaccharide (heteroglycan) may be used. (See 2-Carb-39.)

The term 'glycan' has also been used for the saccharide component of a glycoprotein, even though the chain length may not be large.

The term polysaccharide has also been widely used for macromolecules containing glycose or alditol residues in which both glycosidic and phosphate diester linkages are present.

2-Carb-1.16 Conventions for Examples

1.16.1

Names of examples are given with an initial capital letter (e. g. 'L-*glycero*-β-D-gluco-Heptopyranose') to clarify the usage in headings and to show which letter controls the ordering in an alphabetical index.

1.16.2

The following abbreviations are commonly used for substituent groups in structural formulae: Ac (acetyl), Bn or $PhCH_2$ (benzyl), Bz or PhCO (benzoyl), Et (ethyl), Me (methyl), Me_3Si (not TMS) (trimethylsilyl), Bu^tMe_2Si (not TBDMS) (*tert*-butyldimethylsilyl), Ph (phenyl), Tf (triflyl = trifluoromethanesulfonyl), Ts (tosyl = toluene-*p*-sulfonyl), Tr (trityl).

2-Carb-2 Parent Monosaccharides

2-Carb-2.1 Choice of Parent Structure

In cases where more than one monosaccharide structure is embedded in a larger molecule, a parent structure is chosen on the basis of the following criteria, applied in the order given until a decision is reached:

2.1.1

The parent that includes the functional group most preferred by general principles of organic nomenclature [13,14]. If there is a choice, it is made on the basis of the greatest number of

occurrences of the most preferred functional group. Thus aldaric acid > uronic acidketoaldonic acid/aldonic acid > dialdose > ketoaldosdaldose > diketose > ketose.

2.1.2

The parent with the greatest number of carbon atoms in the chain, e. g. a heptose rather than a hexose.

2.1.3

The parent with the name that comes first in an alphabetical listing based on:

2.1.3.1. the trivial name or the configurational prefix(es) of the systematic name, e. g. allose rather than glucose, a *gluco* rather than a *gulo* derivative;

2.1.3.2. the configurational symbol D rather than L;

2.1.3.3. the anomeric symbol α rather than β.

2.1.4

The parent with the most substituents cited as prefixes (bridging substitution, e. g. 2,3-*O*-methylene, is regarded as multiple substitution for this purpose).

2.1.5

The parent with the lowest locants (see [14], p. 17) for substituent prefixes.

2.1.6

The parent with the lowest locant for the first-cited substituent.

The implications of these recommendations for branched-chain structures are exemplified in ❯ *2-Carb-18*.

Note 1. To maintain homomorphic relationships between classes of sugars, the (potential) aldehyde group of a uronic acid is regarded as the principal function for numbering and naming (see ❯ *2-Carb-2.2.1* and ❯ *2-Carb-22*).

Note 2. To maintain integrity of carbohydrate names, it is sometimes helpful to overstep the strict order of principal group preference specified in general organic nomenclature [13,14]. For example, a carboxymethyl-substituted sugar can be named as such, rather than as an acetic acid derivative (see ❯ *2-Carb-31.2*).

2-Carb-2.2 Numbering and Naming the Parent Structure

The basis for the name is the structure of the parent monosaccharide in the acyclic form. ❯ *Charts I* and ❯ *IV* (❯ *2-Carb-10*) give trivial names for parent aldoses and ketoses with up to six carbon atoms. ❯ *2-Carb-8.2* and ❯ *2-Carb-10.3* describe systematic naming procedures.

2.2.1 Numbering

The carbon atoms of a monosaccharide are numbered consecutively in such a way that:

2.2.1.1. A (potential) aldehyde group receives the locant 1 (even if a senior function is present, as in uronic acids; see ❯ *2-Carb-2.1*, note 1);

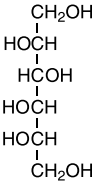
Chart I
Trivial names (with recommended three-letter abbreviations in parentheses) and structures (in the aldehydic, acyclic form) of the aldoses with three to six carbon atoms. Only the D-forms are shown; the L-forms are the mirror images. The chains of chiral atoms delineated in bold face correspond to the configurational prefixes given in italics below the names

2.2.1.2. The most senior of other functional groups expressed in the suffix receives the lowest possible locant, i. e. carboxylic acid (derivatives) > (potential) ketonic carbonyl groups.

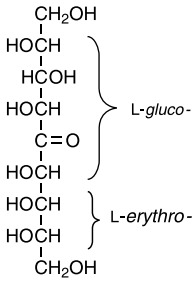
2.2.2 Choice of Parent Name

The name selected is that which comes first in the alphabet (configurational prefixes included). Trivial names are preferred for parent monosaccharides and for those derivatives where all stereocentres are stereochemically unmodified.

Examples:



L-Glucitol
not D-gulitol



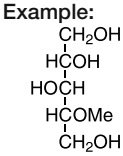
L-*erythro*-L-*gluco*-Non-5-ulose
not D-*threo*-D-*allo*-non-5-ulose

2.2.3 Choice Between Alternative Names for Substituted Derivatives

When the parent structure is symmetrical, preference between alternative names for derivatives should be given according to the following criteria, taken in order:

2.2.3.1. The name including the configurational symbol D rather than L.

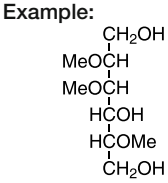
Example:



4-*O*-Methyl-D-xylitol
not 2-*O*-methyl-L-xylitol

2.2.3.2. The name that gives the lowest set of locants (see [14], p. 17) to the substituents present.

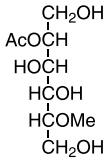
Example:



2,3,5-Tri-*O*-methyl-D-mannitol
not 2,4,5-tri-*O*-methyl-D-mannitol

2.2.3.3. The name that, when the substituents have been placed in alphabetical order, possesses the lowest locant for the first-cited substituent.

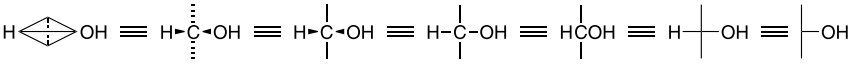
Example:



2-*O*-Acetyl-5-*O*-methyl-D-mannitol
not 5-*O*-Acetyl-2-*O*-methyl-D-mannitol

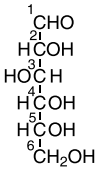
2-Carb-3 The Fischer Projection of the Acyclic Form

In this representation of a monosaccharide, the carbon chain is written vertically, with the lowest numbered carbon atom at the top. To define the stereochemistry, each carbon atom is considered in turn and placed in the plane of the paper. Neighbouring carbon atoms are below, and the H and OH groups above the plane of the paper (see below).



Conventional representation of a carbon atom (e.g. C-2 of D-Glucose) in the Fischer projection. Representation (e) will be used in general in the present document

The formula below is the Fischer projection for the acyclic form of D-glucose. The Fischer projections of the other aldoses (in the acyclic form) are given in *Chart I* (*2-Carb-2.2*).



D-Glucose

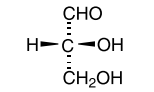
Note. The Fischer projection is not intended to be a true representation of conformation.

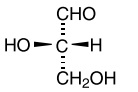
2-Carb-4 Configurational Symbols and Prefixes

2-Carb-4.1 Use of D and L

The simplest aldose is glyceraldehyde (occasionally called glyceral [21]). It contains one centre of chirality (asymmetric carbon atom) and occurs therefore in two enantiomeric forms, called D-glyceraldehyde and L-glyceraldehyde; these are represented by the projection formu-

lae given below. It is known that these projections correspond to the absolute configurations. The configurational symbols D and L should appear in print in small-capital roman letters (indicated in typescript by double underlining) and are linked by a hyphen to the name of the sugar.

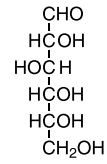

D-Glyceraldehyde

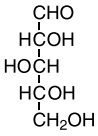

L-Glyceraldehyde

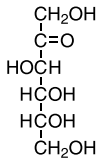
2-Carb-4.2 The Configurational Atom

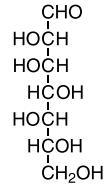
A monosaccharide is assigned to the D or the L series according to the configuration at the highest-numbered centre of chirality. This asymmetrically substituted carbon atom is called the 'configurational atom'. Thus if the hydroxy group (or the oxygen bridge of the ring form; see ❯ *2-Carb-6*) projects to the right in the Fischer projection, the sugar belongs to the D series and receives the prefix D-.

Examples:

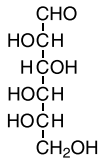

D-Glucose

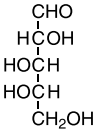

D-Xylose

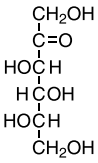

D-*arabino*-Hex-2-ulose
(D-Fructose)

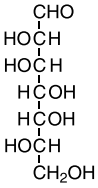

D-*glycero*-L-*gulo*-Heptose

D-Monosaccharides


L-Glucose


L-Arabinose


L-*xlyo*-Hex-2-ulose
(L-Sorbose)


L-*glycero*-D-*manno*-Heptose

L-Monosaccharides

2-Carb-4.3 Configurational Prefixes in Systematic Names

In the systematic names of sugars or their derivatives, it is necessary to specify not only the configuration of the configurational atom but also the configurations of all CHOH groups.

This is done by the appropriate configurational prefix. These prefixes are derived from the trivial names of the aldoses in ❯ *Chart I* (relevant portions of the structures are delineated in bold face). In monosaccharides with more than four asymmetrically substituted carbon atoms, where more than one configurational prefix is employed (see ❯ *2-Carb-8.3*), each group of asymmetrically substituted atoms represented by a particular prefix has its own configurational symbol, specifying the configuration (D or L) of the highest numbered atom of the group.
The configurational prefixes are printed in lower-case italic (indicated in typescript by underlining), and are preceded by either D- or L-, as appropriate. For examples see ❯ *2-Carb-4.2* and ❯ *2-Carb-6.2*

Note. In cyclic forms of sugars, the configuration at the anomeric chiral centre is defined in relation to the 'anomeric reference atom' (see ❯ *2-Carb-6.2*).

2-Carb-4.4 Racemates and *meso* Forms

Racemates may be indicated by the prefix DL-. Structures that have a plane of symmetry and are therefore optically inactive (e. g. erythritol, galactitol) are called *meso* forms and may be given the prefix '*meso*-'.

2-Carb-4.5 Optical Rotation

If the sign of the optical rotation under specified conditions is to be indicated, this is done by adding (+)- or (−)- before the configurational prefix. Racemic forms are indicated by (±)-.

Examples:

- **D-Glucose or (+)-D-glucose**
- **D-Fructose or (−)-o-fructose**
- **DL-Glucose or (±)-glucose**

2-Carb-5 Cyclic Forms and Their Representation

2-Carb-5.1 Ring Size

Most monosaccharides exist as cyclic hemiacetals or hemiketals. Cyclic forms with a three-membered ring are called oxiroses, those with a four-membered ring oxetoses, those with a five-membered ring furanoses, with a six-membered ring pyranoses, with a seven-membered ring septanoses, with an eight-membered ring octanoses, and so on. To avoid ambiguities, the locants of the positions of ring closure may be given; the locant of the carbonyl group is always cited first, that of the hydroxy group second (for relevant examples of this see ❯ *2-Carb-6.4*). Lack of ring size specification has no particular implication.

Note. The 'o' of oxirose, oxetose, and octanose is not elided after a prefix ending in 'o'.

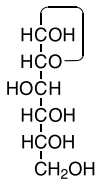
Example:

- **Nonooctanose, not nonoctanose.**

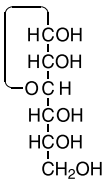
If it is to be stressed that an open-chain form of an aldose is under consideration, the prefix '*aldehydo*-' may be used. For ketoses, the prefix is '*keto*-'.

2-Carb-5.2 The Fischer Projection

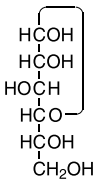
If a cyclic form of a sugar is to be represented in the Fischer projection, a long bond can be drawn between the oxygen involved in ring formation and the (anomeric) carbon atom to which it is linked, as shown in the following formulae for cyclic forms of α-D-Glucose (see 2-Carb-6 for the meaning of α and β):



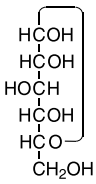
α-D-Glucooxirose



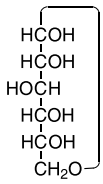
α-D-Glucooxetose



α-D-Glucofuranose



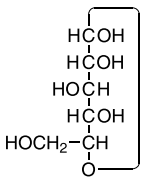
α-D-Glucopyranose



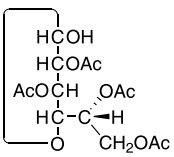
α-D-Glucoseptanose

2-Carb-5.3 Modified Fischer Projection

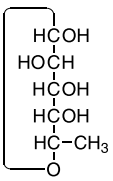
To clarify steric relationships in cyclic forms, a modified Fischer projection may be used. The carbon atom bearing the ring-forming hydroxy group, C-*n* (C-5 for glucopyranose) is rotated about its bond to C-(n − 1) ((C-4 for glucopyranose) in order to bring all ring atoms (including the oxygen) into the same vertical line. The oxygen bridge is then represented by a long bond; it is imagined as being behind the plane of the paper.
Examples are given below.



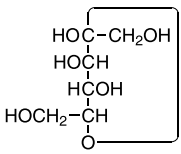
α-D-Glucopyranose



2,3,5,6-Tetra-*O*-acetyl-α-D-galactofuranose



β-L-Fucopyranose

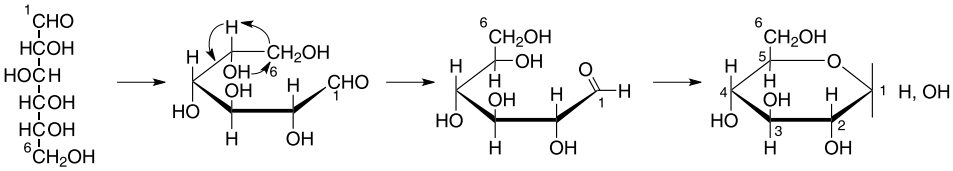


β-L-Fructofuranose

Thus the *trans* relationship between the hydroxymethyl group and the C-1 hydroxy group in α-D-glucopyranose, and the *cis* relationship between the methyl group and the C-1 hydroxy group in β-L-fucopyranose, are clearly shown. Note that representation of ketoses may require a different modification of the Fischer projection, as shown in the fructofuranose example above. Here C-2 is rotated about the bond with C-3 to accommodate the long bond to C-2 from the oxygen at C-5.

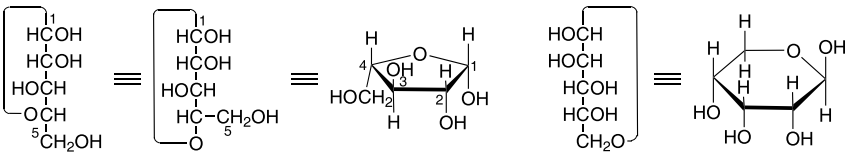
2-Carb-5.4 The Haworth Representation

This is a perspective drawing of a simplified model. The ring is orientated almost perpendicular to the plane of the paper, but viewed from slightly above so that the edge closer to the viewer is drawn below the more distant edge, with the oxygen behind and C-1 at the right-hand end. To define the perspective, the ring bonds closer to the viewer are often thickened.
The following schematic representation of pyranose ring closure in D-glucose shows the reorientation at C-5 necessary to allow ring formation; this process corresponds to the change from Fischer to modified Fischer projection.



Haworth representation of D-glucopyranose

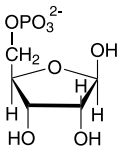
The orientation of the model described above results in a clockwise numbering of the ring atoms. Groups that appear to the right of the modified Fischer projection appear below the plane of the ring; those on the left appear above. In the common Haworth representation of the pyranose form of D-aldohexoses, C-6 is above the plane.
Generally, the configuration at the centre that yields the ring oxygen determines whether the rest of the carbon chain is below or above the plane of the ring.
Examples (for the use of α and β see ❯ *2-Carb-6*):



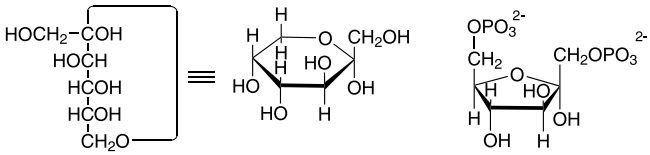
Fischer modified Fischer Haworth

β-L-Arabinofuranose representations

β-D-Ribopyranose

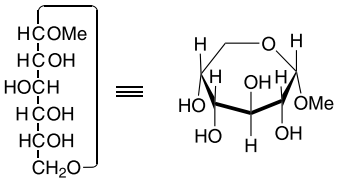


β-D-Ribopyranose 5-phosphate

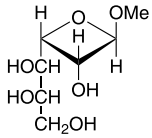


α-D-Fructopyranose

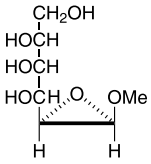
α-D-Fructofuranose 1,6-bisphosphate



Methyl α-D-glucoseptanoside



Methyl α-L-altrooxetoside



Methyl β-D-altrooxiroside

Note. In writing Haworth formulae, the H atoms bound to the carbon atoms of the ring are often omitted to avoid crowding ofthe lettering in the ring. For the sake of clarity, the form with H atoms included is preferred in this document.

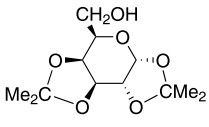
2-Carb-5.5 Unconventional Haworth Representations

It is sometimes desirable to draw Haworth formulae with the ring in other orientations (see ❯ *Chart II*), when there are bulky substituents to be represented, or when linkages in oligo- or poly-saccharides are to be shown. If the ring is inverted [as in (g)-(l)], the numbering runs counterclockwise.

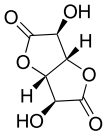
2-Carb-5.6 The Mills Depiction

In some cases, particularly where additional rings are present, structural formulae can be clarified by use of the Mills depiction. Here the main hemiacetal ring is drawn in the plane of the paper; dashed bonds denote substituents below this plane, and thickened bonds those above.

Examples:



1,2:3,4-Di-*O*-isopropylidene-α-D-galactopyranose



D-Glucaro-1,4:6,3-dilactone

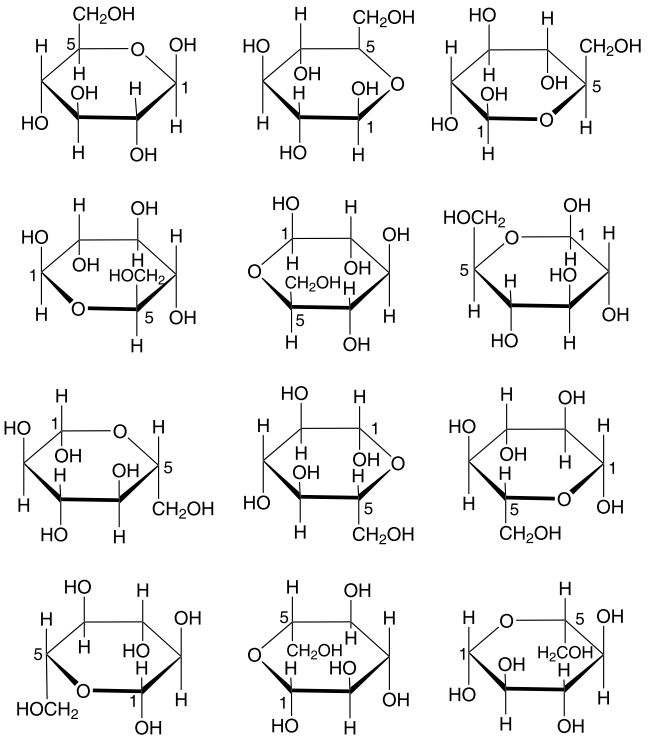


Chart II

***β*-D-Glucopyranose in the twelve possible Haworth representations (the hydrogen atoms are frequently omitted)**

2-Carb-5.7 Depiction of Conformation

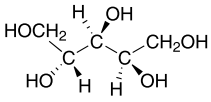
The Haworth representation implies a planar ring. However, monosaccharides assume conformations that are not planar: these may be represented by Haworth conformational formulae. The nomenclature of conformations is described in 2-Carb-7. For example, *β*-D-glucopyranose assumes a chair conformation:

β-D-Glucopyranose in a chair conformation

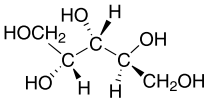
Note. The hydrogen atoms bonded to carbon are frequently omitted, but their inclusion may be necessary to make a stereochemical point.

2-Carb-5.8 Conformations of Acyclic Chains

Conformational depictions of acyclic sugar chains are conveniently expressed by locating certain atoms in the plane of the paper and orientating the remaining atoms or groups appropriately above and below that plane, as shown for D-arabinitol and xylitol (it should be recognized that the favoured conformation does not necessarily have all the carbon atoms in the same plane):



D-Arabinitol



Xylitol

2-Carb-6 Anomeric Forms; Use of α and β

2-Carb-6.1 The Anomeric Centre

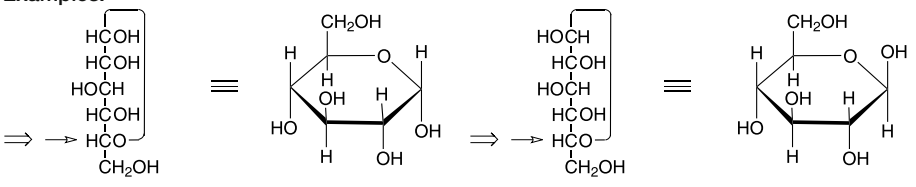
The new centre of chirality generated by hemiacetal ring closure is called the anomeric centre. The two stereoisomers are referred to as anomers, designated α or β according to the configurational relationship between the anomeric centre and a specified anomeric reference atom.

2-Carb-6.2 The Anomeric Reference Atom and the Anomeric Configurational Symbol (α or β)

The anomeric reference atom is the configurational atom (see 2-Carb-4.2 and 2-Carb-4.3) of the parent, unless multiple configurational prefixes (see 2-Carb-8.3) are used. If multiple configurational prefixes are used, the anomeric reference atom is the highest-numbered atom of the group of chiral centres next to the anomeric centre that is involved in the heterocyclic ring and specified by a single configurational prefix. In the α anomer, the exocyclic oxygen atom at the anomeric centre is formally *cis*, in the Fischer projection, to the oxygen attached to the anomeric reference atom; in the β anomer these oxygen atoms are formally *trans*.

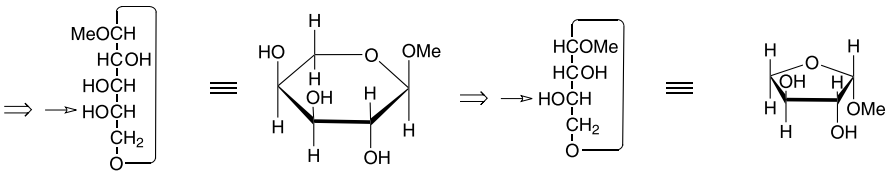
The anomeric symbol α or β, followed by a hyphen, is placed immediately before the configurational symbol D or L of the trivial name or of the configurational prefix denoting the group of chiral carbon atoms that includes the anomeric reference atom.

Examples:



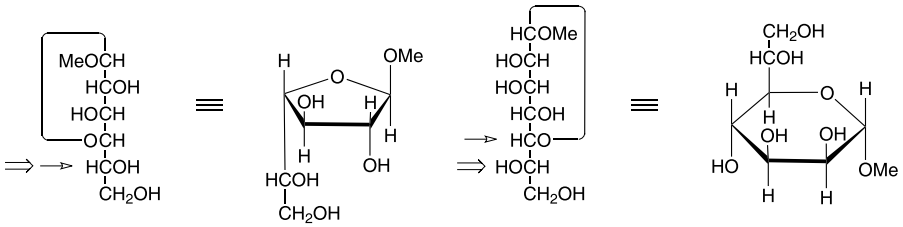
α-D-*gluco*
α-D-Glucopyranose

β-D-*gluco*
β-D-Glucopyranose



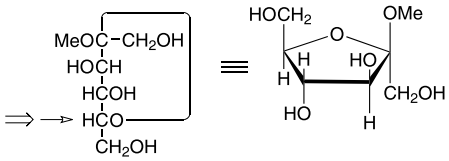
α-L-*arabino*
Methyl α-L-arabinopyranoside

β-L-*threo*
Methyl β-L-threofuranoside

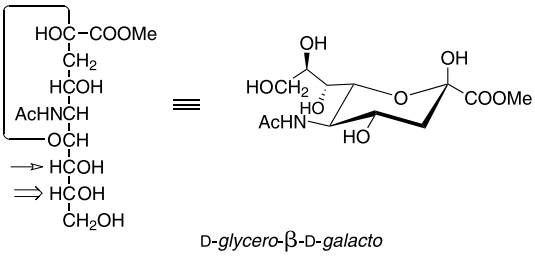


β-D-*galacto*
Methyl β-D-galactofuranoside

L-*glycero*-α-D-*manno*
Methyl L-*glycero*-α-D-*manno*-heptopyranoside



β-D-*arabino*
Methyl β-D-fructofuranoside



D-*glycero*-β-D-*galacto*
Methyl 5-acetoamido-3,5-dideoxy-D-*glycero*-β-D-*galacto*-non-2-ulopyranosonate (see 2-carb-14.2)

→ denotes the anomeric reference atom; ⟹ denotes the configurational atom.

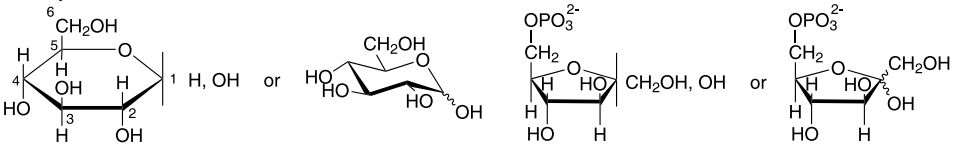
Note. For simple aldoses up to aldohexoses, and ketoses up to hept-zuloses, the anomeric reference atom and the configurational atom are the same.

2-Carb-6.3 Mixtures of Anomers

In solution, most simple sugars and many of their derivatives occur as equilibrium mixtures of tautomers. The presence of a mixture of two anomers of the same ring size may be indicated

in the name by the notation α, β-, e. g. α, β-D-glucose. In formulae, the same situation can be expressed by separating the representation of the ligands at the anomeric centre from the α and β bonds [see examples (a) and (c)], or by use of a wavy line [(b) and (d)] (particularly if hydrogen atoms are omitted).

Examples:



a b c d

α,β-D-Glucopyranose

α,β-D-Fructofuranose 6-phosphate

2-Carb-6.4 Use of α and β

The Greek letters α and β are applicable only when the anomeric carbon atom has a lower locant than the anomeric reference atom. In the case of dialdoses (cf. ❯ *2-Carb-9*), some diketoses (cf. ❯ *2-Carb-11*) and aldoketoses (cf. ❯ *2-Carb-12*), ring closure is also possible in the other direction, i. e. of a carbonyl group with a higher locant than the reference carbon atom with a hydroxy group having a lower locant. In this case, the configuration of the anomeric carbon atom is indicated by the appropriate symbol *R* or *S* according to the sequence rule (cf. Section E in [13]).

Examples:

(6*R*)-D-*gluco*-Hexadialdo-6,2-pyranose

(6*S*)-1,2-*O*-Isopropylidene-α-D-*gluco*-hexadialdo-1,4:6,3-difuranose

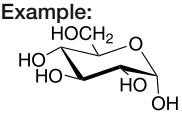
Note that locant numerals (potential carbonyl first) may be needed before the ring-size suffix in such cases.

2-Carb-7 Conformation of Cyclic Forms[1]

2-Carb-7.1 The Conformational Descriptor

The conformation, i. e. the (approximate) spatial arrangement of the ring atoms of a monosaccharide in the cyclic form, may be indicated by an italic capital letter designating the type of ring shape, and numerals, distinguishing the variants. The conformational descriptor is joined to the end of the name of the monosaccharide by a hyphen.

Example:



α-D-Glucopyranose-$^4C_1$

2-Carb-7.2 Notation of Ring Shape

The appropriate letters are as follows. Five-membered rings: *E* for envelope and *T* for twist; six-membered rings: *C* for chair, *B* for boat, *S* for skew, *H* for half-chair, and *E* for envelope. Examples are given in ❯ *Chart III*.

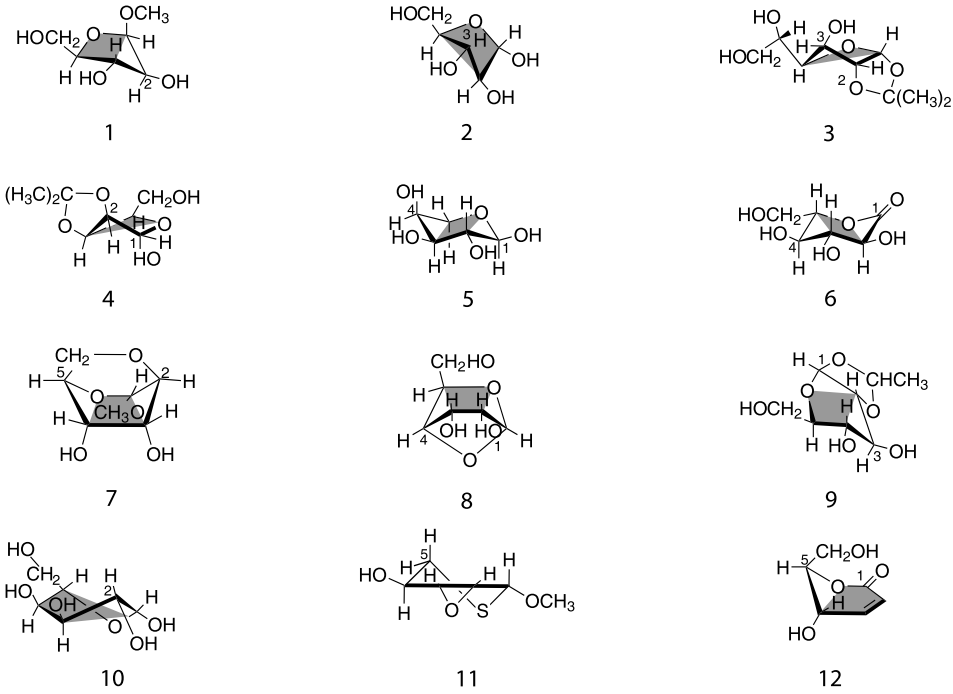
2-Carb-7.3 Notation of Variants

The variants are distinguished by the locants of those ring atoms that lie outside a reference plane (defined below) and are listed for some examples in ❯ *Table 1*. The locants of ring atoms that lie on the side of the reference plane from which numbering appears clockwise (i. e. the upper side in the normal Haworth representation of furanoses and pyranoses) are written as superscripts and precede the letter; those that lie on the other side are written as subscripts and follow the letter. Heteroatoms (e. g. O, S are indicated by their subscript or superscript atomic symbols. ❯ *Table 1* gives the notations and ❯ *Chart III* some examples.

Six-membered rings

Chairs. The reference plane is defined by two parallel ring sides, so chosen that the lowest-numbered carbon atom in the ring is exoplanar (examples **5** and **6**).
Boats. The reference plane is defined by the two parallel 'sides' of the boat (examples **7** and **8**).
Skews. Each skew form has two potential reference planes, containing three adjacent atoms and the remaining non-adjacent atom. The reference plane is so chosen that the lowest-numbered carbon atom in the ring, or the atom numbered next above it, is exoplanar, in that order of preference (examples **9** and **10**).
Half-chairs. The reference plane is defined by the four adjacent coplanar atoms (example **11**).
Envelopes. The reference plane is defined by the five adjacent coplanar atoms (example **12**).

[1] This is an abridged version of the document 'Conformational Nomenclature for Five- and Six-membered Ring Forms of Monosaccharides and their Derivatives. Recommendations 1980' [3].



1 Methyl β-D-arabinofuranoside-E_2

2 α-D-Arabinofuranose-^3E

3 1,2-*O*-Isopropylidene-β-L-idofuranose-$^3T_2$

4 2,3-*O*-Isopropylidene-α-D-luxofuranose-$^2T_1$

5 α-L-Arabinopyranose-$^4C_1$

6 L-Glucono-1,5-lactone-$^1C_4$

7 Methyl 2,6-anhydro-α-D-altropyranoside-$^{2,5}B$

8 1,4-Anhydro-α-D-allopyranose-$B_{1,4}$

9 1,2-*O*-Ethylidene-α-D-glucopyranose-$^1S_3$

10 β-L-Altropyranose-$^2S_0$

11 Methyl 2,3-anhydro-5-thio-β-L-lyxoyranoside-^5H_S

12 2,3-Dideoxy-D-*erythro*-hex-2-enono-1,5-lactone-^5E

Chart III
Drawings of the conformations listed in Table 1. The reference plane is stippled

Five-membered rings

Envelopes. The reference plane is defined by the four adjacent coplanar atoms (examples 1 and 2).

Twists. The reference plane is defined by three adjacent ring-atoms, so chosen that the exoplanar atoms lie on opposite sides of the plane (examples 3 and 4).

Note 1. Many of the possible conformations are not likely to contribute significantly to the chemistry of a particular monosaccharide, but must be stabilized by formation of additional rings, as in anhydrides or other derivatives. Some others may occur as transition-state intermediates.

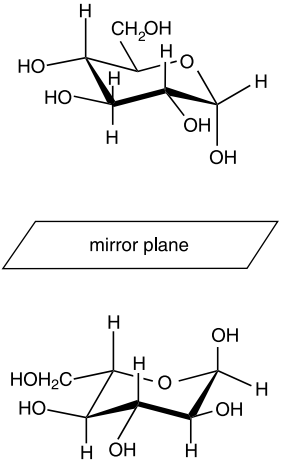
Note 2. A more precise specification of conformation can be achieved by use of the Cremer-Pople puckering parameters [22].

Table 1
Conformationsand their notations; some examples are shown in Chart III

Type of sugar	Conformation	Atoms of reference plane	Above plane	Below plane	Notation	Example
Aldofuranose	envelope	O-4,C-I,C-3,C-4	–	C-2	E_2	**1**
Aldofuranose	envelope	C-I,C-2,C-4,O-4	C-3	–	${}^3E$	**2**
Aldofuranose	twist	C-1,O-4,C-4	C-3	C-2	${}^3T_2$	**3**
Aldofuranose	twist	C-3,C-4,O-4	C-2	C-1	${}^2T_1$	**4**
Aldopyranose	chair	C-2,C-3,C-5,O-5	C-4	C-1	${}^4C_1$	**5**
Pyranoid lactone	chair	C-2,C-3,C-5,O-5	C-1	C-4	${}^1C_4$	**6**
Aldopyranose	boat	O-5,C-1,C-3,C-4	C-2,C-5	–	${}^{2,5}B$	**7**
Aldopyranose	boat	C-2,C-3,C-5,O-5	–	C-1,C-4	$B_{1,4}$	**8**
Aldopyranose	skew	C-2,C-4,C-5,O-5	C-1	C-3	${}^1S_3$	**9**
Aldopyranose	skew	C-1,C-3,C-4,C-5	C-2	O-5	${}^2S_0$	**10**
Aldopyranose	half-chair	C-1,C-2,C-3,C-4	C-5	S-5	${}^5H_s$	**11**
Pyranoid lactone	envelope	C-1,C-2,C-3,C-4,O-5	C-5	–	${}^5E$	**12**

2-Carb-7.4 Enantiomers

The conformational symbols for enantiomers are different. It is therefore important to state in the context whether the D or the L form is under consideration. Enantiomers have the same reference plane (see 2-Carb-7.3), and it should be noted that the mirror image of α-D-glucose-${}^4C_1$ is α-L-glucose-${}^1C_4$.



Mirror images: α-D-glucopyranose-${}^4C_1$ (upper) and α-L-glucopyranose-${}^1C_4$ (lower)

2-Carb-8 Aldoses

2-Carb-8.1 Trivial Names

The aldoses with three to six carbon atoms have trivial names which are given, together with the formulae in the Fischer projection, in ❯ *Chart I* (❯ *2-Carb-2.2*). (See also the alphabetical listing of trivial names in the Appendix.)
The trivial names form the basis of the configurational prefixes (see ❯ *2-Carb-4.3*).

2-Carb-8.2 Systematic Names

Systematic names are formed from a stem name and a configurational prefix or prefixes. The stem names for the aldoses with three to ten carbon atoms are triose, tetrose, pentose, hexose, heptose, octose, nonose, decose. The chain is numbered so that the carbonyl group is at position 1.
The configuration of the CHOH groups of the sugar is designated by the appropriate configurational prefix(es) from ❯ *Chart I*. When used in systematic names, these prefixes are always to be in lower case letters (with no initial capital), and italicized in print. Each prefix is qualified by D or L (❯ *Chart I* shows only the D structures).

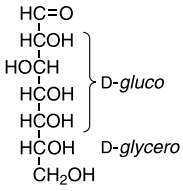
Examples:
D-*ribo*-Pentose for D-ribose
D-*manno*-Hexose for D-mannose

The trivial names are preferred for the parent sugars and for those derivatives where all stereocentres are unmodified.

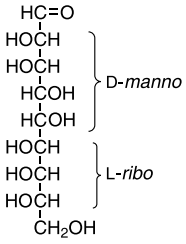
2-Carb-8.3 Multiple Configurational Prefixes

An aldose containing more than four chiral centres is named by adding two or more configurational prefixes to the stem name. Prefixes are assigned in order to the chiral centres in groups of four, beginning with the group proximal to C-1. The prefix relating to the group of carbon atom(s) farthest from C-1 (which may contain less than four atoms) is cited first.

Examples:



D-*glycero*-D-*gluco*-Heptose
not D-*gluco*-D-*glycero*-heptose

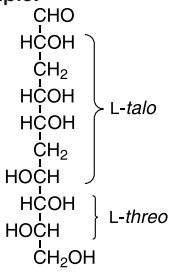


L-*ribo*-D-*manno*-Nonose
not D-*manno*-L-*ribo*-nonose

2-Carb-8.4 Multiple Sets of Chiral Centres

If sequences of chiral centres are separated by non-chiral centres, the non-chiral centres are ignored, and the remaining set of chiral centres is assigned the appropriate configurational prefix (for four centres or less) or prefixes (for more than four centres).

Example:



3,6-Dideoxy-L-*threo*-L-*talo*-decose

Note 1. This convention is not needed for parent aldoses, only for deoxy aldoses, ketoses and similar compounds (see ❯ *2-Carb-10.4* and ❯ *2-Carb-11.2*).

Note 2. Since all aldoses up to the hexoses have trivial names that are preferred, the systematic names apply only to the higher aldoses. However, the configurational prefixes are also used to name ketoses (see below) and other monosaccharides.

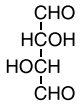
2-Carb-8.5 Anomeric Configuration in Cyclic Forms

For the specification of α and β in cyclic forms see ❯ *2-Carb-6*.

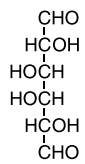
2-Carb-9 Dialdoses

Systematic names for individual dialdoses are formed from the systematic stem name for the corresponding aldose (see ❯ *2-Carb-8.2*), but with the ending 'odialdose' instead of 'ose', and the appropriate configurational prefix (❯ *Chart I*). A choice between the two possible aldose parent names is made on the basis of ❯ *2-Carb-2.2.2*.

Examples:



L-*threo*-Tetrodialdose

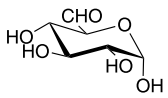


galacto-Hexodialdose

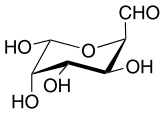
Note. The prefix '*meso*-' could be included in the latter case, but it is not needed to define the structure.

If a cyclic form is to be named, the locants of the anomeric centre and of the carbon atom bearing the ring oxygen atom must be given (in that order) (cf. ❯ *2-Carb-6.4*). If there is more than one ring size designator, they are placed in alphabetical order (e. g. furanose before pyranose).

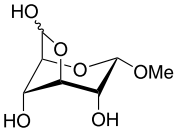
Examples:



α-D-*gluco*-Hexodialdo-1,5-pyranose



(6*R*)-D-*gluco*-Hexodialdo-6,2-pyranose



Methyl α-D-*gluco*-hexodialdo-6,3-furanose-1,5-pyranoside

2-Carb-10 Ketoses

2-Carb-10.1 Classification

Ketoses are classified as 2-ketoses, 3-ketoses, etc., according to the position of the (potential) carbonyl group. The locant 2 may be omitted if no ambiguity can arise, especially in a biochemical context.

2-Carb-10.2 Trivial Names

Ketoses with three to six carbon atoms are shown in ❯ *Chart IV*, with trivial names (and three-letter abbreviations) in parentheses. (See also the alphabetical listing of trivial names in the Appendix.)

The trivial names 'D-erythrulose' for D-*glycero*-tetrulose, 'D-ribulose' for D-*erythro*-pent-2-ulose, and 'D-xylulose' for D-*rhreo*-pent-2-ulose contain stereochemical redundancy and should not be used for naming derivatives. Sedoheptulose is the accepted trivial name for D-*altro*-hept-2-ulose.

2-Carb-10.3 Systematic Names

The systematic names are formed from the stem name and the appropriate configurational prefix.

The stem names are formed from the corresponding aldose stem names (❯ *2-Carb-8.2*) by replacing the ending '-ose' with '-ulose', preceded by the locant of the carbonyl group, e. g.

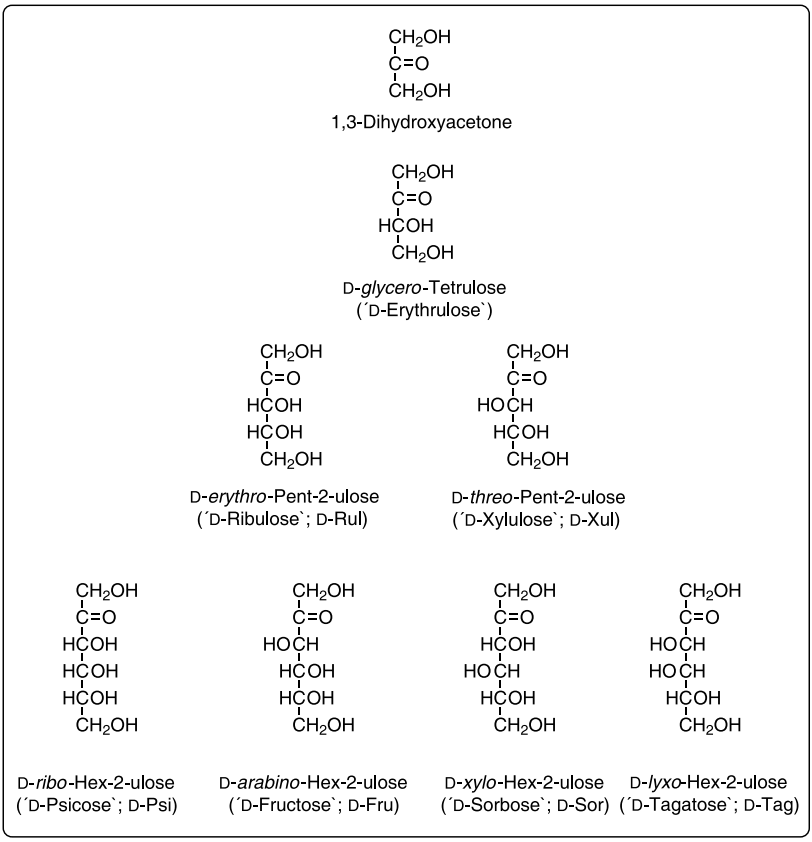


Chart IV
Structures, with systematic and trivial names, of the 2-ketoses with three to six carbon atoms

hex-3-ulose. The chain is numbered so that the carbonyl group receives the lowest possible locant. If the carbonyl group is in the middle of a chain with an odd number of carbon atoms, a choice between alternative names is made according to ❯ *2-Carb-2.2.2*.

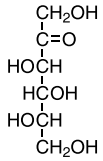
Note. In Chemical Abstracts Service (CAS) usage the locant for the carbonyl group precedes the stem name, e. g. 3-hexulose.

For examples see ❯ *2-Carb-10.4.*

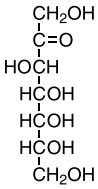
2-Carb-10.4 Configurational Prefixes

For 2-ketoses, configurational prefixes are given in the same way as for aldoses (see ❯ *2-Carb-8.2* and ❯ *2-Carb-8.3*).

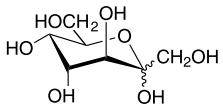
Examples:



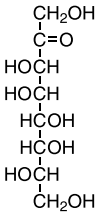
L-*xylo*-Hex-2-ulose
(L-Sorbose)



D-*altro*-Hept-2-ulose
(D-Sedoheptulose)



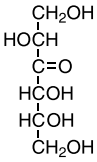
D-*altro*-Hept-2-ulopyranose



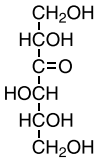
L-*glycero*- D-*manno*-Oct-2-ulose

For ketoses with the carbonyl group at C-3 or a higher-numbered carbon atom, the carbonyl group is ignored and the set of chiral centres is given the appropriate prefix or prefixes according to ❯ *Chart I* (cf. ❯ *2-Carb-8.4*).

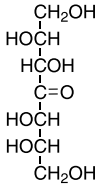
Examples:



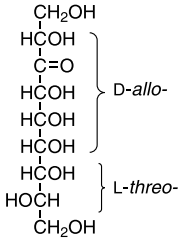
D-*arabino*-Hex-3-ulose



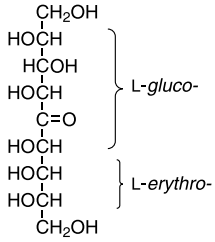
D-*xylo*-Hex-3-ulose
not L-*xylo*-hex-4-ulose



L-*gluco*-Hept-4-ulose
not D-*gulo*-hept-4-ulose



L-*threo*-D-*allo*-Non-3-ulose



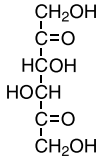
L-*erythro*-L-*gluco*-Non-5-ulose
not D-*threo*-D-*allo*-non-5-ulose

2-Carb-11 Diketoses

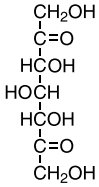
2-Carb-11.1 Systematic Names

The systematic name of a diketose is formed by replacing the terminal '-se' of the stem name by '-diulose'. The locants of the (potential) carbonyl groups must be the lowest possible and appear before the ending. The stem name is preceded by the appropriate configurational prefix. If there is a choice of names, a decision is made on the basis of 2-Carb-2.2.2. In cyclic forms, locants may be needed for the positions of ring closure; that of the (potential) carbonyl group is cited first.

Examples:



L-*threo*-Hexo-2,5-diulose

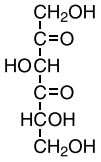


meso-*xylo*-Hepto-2,6-diulose

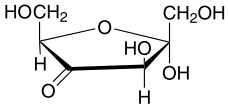
2-Carb-11.2 Multiple Sets of Chiral Centres

If the carbonyl group(s) divides the sequence of chiral centres, the configurational prefixes are assigned in the normal manner (see 2-Carb-8.4) for all chiral centres; the non-chiral centres are ignored.

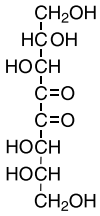
Examples:



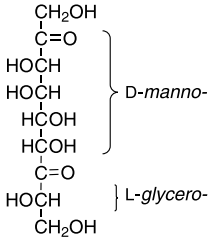
D-*threo*-Hexo-2,4-diulose



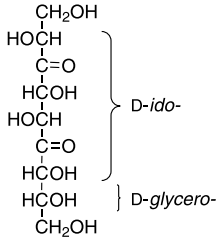
α-D-*threo*-Hexo-2,4-diulo-2,5-furanose



L-*altro*-Octo-4,5-diulose
not L-*talo*-octo-4,5-diulose



L-*glycero*-D-*manno*-Nono-2,7-diulose



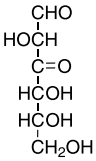
D-*glycero*-D-*ido*-Nono-3,6-diulose

2-Carb-12 Ketoaldoses (Aldoketoses, Aldosuloses)

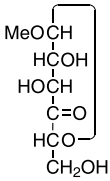
2-Carb-12.1 Systematic Names

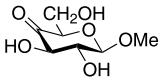
Names of ketoaldoses are formed in the same way as those of diketoses, but with use of the termination '-ulose' in place of the terminal '-e' of the corresponding aldose name (2-Carb-8.2). The carbon atom of the (potential) aldehydic carbonyl group is numbered 1, and this locant is not cited in the name. The locant of the (potential) ketonic carbonyl group is given (as an infix before '-ulose') unless it is 2; it may then be omitted (in this text, this locant is always retained for the sake of clarity). In cyclic forms, locants may be needed for the positions of ring closure; that of the (potential) carbonyl group is cited first. The position of the ring-size designator (e. g. pyrano) depends upon which carbonyl group is involved in ring formation (see examples).

Examples:

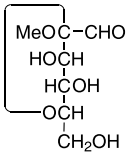


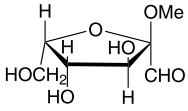
D-*arabino*-Hexos-3-ulose





Methyl β-D-*xylo*-hexopyranosid-4-ulose





Methyl α-L-*xylo*-Hexos-2-ulo-2,5-furanoside

2-Carb-12.2 'Dehydro' Names

In a biochemical context, the naming of aldoketoses as 'dehydro' aldoses is widespread. Thus D-*xylo*-hexopyranos-4-ulose would be termed 4-dehydro-D-glucose. This usage of 'dehydro' can give rise to names which are stereochemically redundant, and should not be employed for naming derivatives.

Note. In Enzyme Nomenclature [23] dehydro names are used in the context of enzymic reactions. The substrate is regarded as the parent compound, but the name of the product is chosen according to the priority given in 2-Carb-2.2.

Examples:

D-Glucose + O_2 = 2-dehydro-D-glucose + H_2O_2 (EC 1.1.3.10)

Sucrose + acceptor = β-D-fructofuranosyl 3-dehydro-α-D-allopyranoside + reduced acceptor (EC 1.1.99.13)

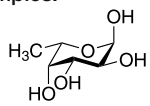
L-Sorbose + $NADP^+$ = 5-dehydro-D-fructose + NADPH (reaction of sorbose dehydrogenase, EC 1.1.1.123)

2-Carb-13 Deoxy Sugars

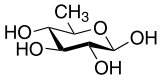
2-Carb-13.1 Trivial Names

Several deoxy sugars have trivial names established by long usage, e. g. fucose (Fuc), quinovose (Qui) and rhamnose (Rha). They are illustrated here in the pyranose form. These names are retained for the unmodified sugars, but systematic names are usually preferred for the formation of names of derivatives, especially where deoxygenation is at a chiral centre of the parent sugar. (See also the alphabetical listing of trivial names in the Appendix.)

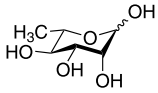
Examples:



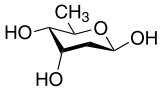
α-L-Fucopyranose
6-Deoxy-α-L-galactopyranose



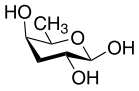
β-D-Quinovopyranose
6-Deoxy-β-D-glucopyranose



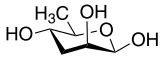
L-Rhamnopyranose
6-Deoxy-L-mannopyranose



2,6-Dideoxy-β-D-*ribo*-hexopyranose
(β-Digitoxopyranose)



3,6-Dideoxy-β-D-*xylo*-hexopyranose
(β-Abequopyranose)



3,6-Dideoxy-β-D-*arabino*-hexopyranose
(β-Tyvelopyranose)

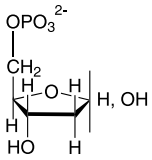
Other trivial names that have been used include ascarylose for 3,6-dideoxy-L-*arabino*-hexose, colitose for 3,6-dideoxy-L-*xylo*-hexose and paratose for 3,6-dideoxy-D-*ribo*-hexose.

Note. Sugars with a terminal CH_3 group should be named as *o*-deoxy sugars, as shown above, not *C*-methyl derivatives.

2-Carb-13.2 Names Derived from Trivial Names of Sugars

Use of 'deoxy-' in combination with an established trivial name (see ❯ *Charts I* and ❯ *II*) is straightforward if the formal deoxygenation does not affect the configuration at any asymmetric centre. However if 'deoxy' removes a centre of chirality, the resulting names contain stereochemical redundancy. In such cases, systematic names are preferred, especially for the naming of derivatives.

Note. The names 2-deoxyribose (for 2-deoxy-D-*erythro*-pentose) and 2-deoxyglucose (for 2-deoxy-D-*arabino*-hexose) are often used.



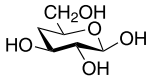
2-Deoxy-D-*erythro*-pentofuranose 5-phosphate

2-Carb-13.3 Systematic Names

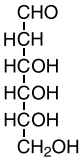
The systematic name consists of the prefix 'deoxy-', preceded by the locant and followed by the stem name with such configurational prefixes as necessary to describe the configuration(s) at the asymmetric centres present in the deoxy compound. Configurational prefixes are cited in order commencing at the end farthest from C-1. 'Deoxy' is regarded as a detachable prefix, i. e. it is placed in alphabetical order with any substituent prefixes.

Note. The treatment of 'anhydro' (see 2-Carb-26), 'dehydro' (see 2-Carb-17.3) and 'deoxy' as detachable prefixes follows long-standing practice in carbohydrate chemistry, but is in conflict with [14] (p. 12).

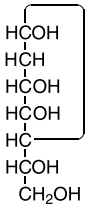
Examples:



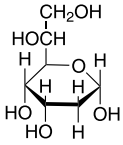
4-Deoxy-β-D-*xylo*-hexopyranose
not 4-deoxy-β-D-galactopyranose

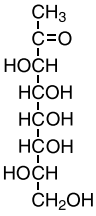


2-Deoxy-D-*ribo*-hexose not 2-deoxy-D-allose



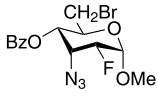
a





2-Deoxy-α-D-*allo*-heptopyranose

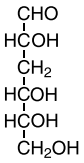
1-Deoxy-L-*glycero*-D-*altro*-oct-2-ulose



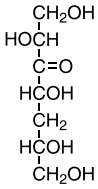
Methyl 3-azido-4-*O*-benzoyl-6-bromo-2,3,6-trideoxy-2-fluoro-α-D-allopyranoside

If the CH_2 group divides the chiral centres into two sets, it is ignored for the purpose of assigning a configurational prefix; the prefix(es) assigned should cover the entire sequence of chiral centres (see ❯ *2-Carb-8.4*).

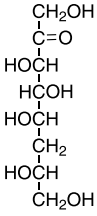
Examples:



3-Deoxy-D-*ribo*-hexose
not 3-deoxy-D-*erythro*-D-*glycero*-hexose



5-Deoxy-D-*arabino*-hept-3-ulose
not 5-deoxy-D-*glycero*-D-*glycero*-L-*glycero*-hept-3-ulose



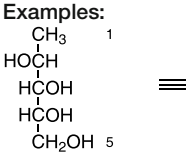
6-Deoxy-L-*gluco*-oct-2-ulose
not 6-deoxy-L-*glycero*-L-*xylo*-oct-2-ulose

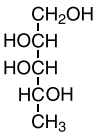
If the anomeric hydroxy group is replaced by a hydrogen atom, the compound is named as an anhydro alditol (❯ *2-Carb-26*).

2-Carb-13.4 Deoxy Alditols

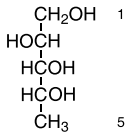
The name of an aldose derivative in which the aldehyde group has been replaced by a terminal CH_3 group is derived from that of the appropriate alditol (see ❯ *2-Carb-19*) by use of the prefix 'deoxy-'.

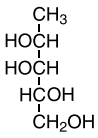
Examples:





1-Deoxy-D-arabinitol
not 5-deoxy-D-lyxitol





5-Deoxy-D-arabinitol
not 1-deoxy-D-lyxitol

The alditols from fucose and rhamnose are frequently termed fucitol and rhamnitol (see ❯ *2-Carb-19.1*).

2-Carb-14 Amino Sugars

2-Carb-14.1 General Principles

The replacement of an alcoholic hydroxy group of a monosaccharide or monosaccharide derivative by an amino group is envisaged as substitution of the appropriate hydrogen atom of the corresponding deoxy monosaccharide by the amino group. The stereochemistry at the carbon atom carrying the amino group is expressed according to 2-Carb-8.2, with the amino group regarded as equivalent to OH.

Some examples of *N*-substituted derivatives are given here; for a detailed treatment see 2-Carb-25.

2-Carb-14.2 Trivial Names

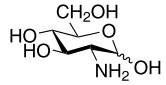
Accepted trivial names are as follows.

- D-Galactosamine for 2-amino-2-deoxy-D-galactose
- D-Glucosamine for 2-amino-2-deoxy-D-glucose
- D-Mannosamine for 2-amino-2-deoxy-D-mannose
- D-Fucosamine for 2-amino-2,6-dideoxy-D-galactose
- D-Quinovosamine for 2-amino-2,6-dideoxy-D-glucose
- Neuraminic acid for 5-amino-3,5-dideoxy-D-*glycero*-D-*galacto*-non-2-ulosonic acid
- Muramic acid for 2-amho-3-*O*-[(*R*)-1-carboxyethyl]-2-deoxy-D-glucose.

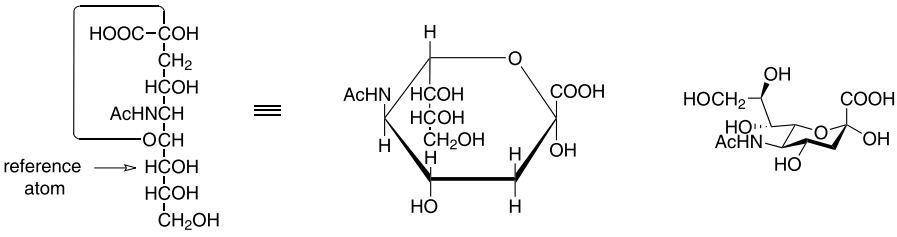
In the last two cases the trivial name refers specifically to the D enantiomer. (See also the alphabetical listing of trivial names in the Appendix.)

Such names as 'bacillosamine' for 2,4-diamino-2,4,6-trideoxy-D-glucose and 'garosamine' for 3-deoxy-4-C-methyl-3-methylamino-L-arabinose are not recommended, as they imply replacement of OH by NH_2 in a nonexistent parent sugar.

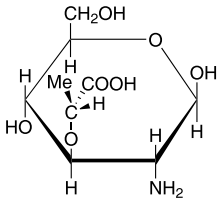
Examples:



2-Amino-2-deoxy-D-glucopyranose (D-glucosamine).



5-Acetamido-3,5-dideoxy-D-*glycero*-α-D-*galacto*-non-2-ulopyranosonic acid (*N*-acetyl-α-neuramic acid, α-Neu5Ac), drawn in three ways (note that C-7 is the anomeric reference atom



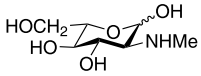
2-Amino-3-*O*-[(R)-1-carboxyethyl]-2-deoxy-β-D-glucopyranose (β-muramic acid)

For examples with nitrogen in the ring, see ❯ *2-Carb-34.1*.

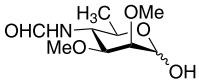
2-Carb-14.3 Systematic Names

The compounds are named by use of a combination of 'deoxy-' and 'amino-' prefixes. When the complete name of the derivative includes other prefixes, 'deoxy-' takes its place in the alphabetical order of detachable prefixes.

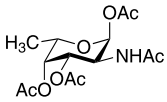
Examples:



2-Deoxy-2-methylamino-L-glucopyranose



4,6-Dideoxy-4-formamido-2,3-di-*O*-methyl-D-mannopyranose



2-Acetamido-1,3,4-tri-*O*-acethyl-2,6-dideoxy-α-L-galactopyranose

When the amino group is at the anomeric position, the compound is normally named as a glycosylamine (see ❯ *2-Carb-33.6*).

2-Carb-15 Thio Sugars and Other Chalcogen Analogues

Replacement of a hydroxy oxygen atom of an aldose or ketose, or of the oxygen atom of the carbonyl group of the acyclic form of an aldose or ketose, by sulfur is indicated by placing the prefix 'thio', preceded by the appropriate locant, before the systematic or trivial name of the aldose or ketose.

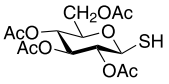
Replacement of the ring oxygen atom of the cyclic form of an aldose or ketose by sulfur is indicated in the same way, the number of the non-anomeric adjacent carbon atom of the ring being used as locant.

Selenium and tellurium compounds are named likewise, by use of the prefix 'seleno' or 'telluro'.

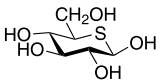
Sulfoxides (or selenoxides or telluroxides) and sulfones (or selenones or tellurones) may be named by functional class nomenclature [13].

Note. The appropriate prefix is thio, not thia; the latter is used in systematic organic chemical nomenclature to indicate replacement of CH_2 by S.

Examples:

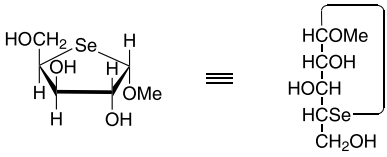


2,3,4,6-Tetra-*O*-acetyl-1-thio-β-D-glucopyranose

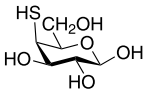


5-Thio-β-D-glucopyranose

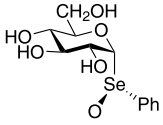
Methyl 2,3,4-tri-*O*-acetyl-1-thio-6-*O*-trityl-α-D-glucopyranoside



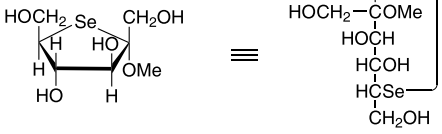
Methyl 4-seleno-α-D-xylofuranoside



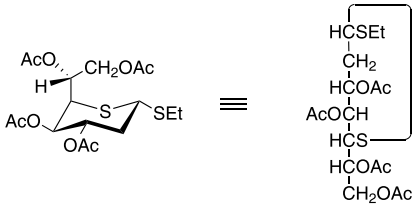
4-Thio-β-D-galactopyranose



α-D-Glucopyranosyl phenyl (*R*)-selenoxide



Methyl 5-seleno-α-D-fructofuranoside



Ethyl 3,4,6,7-tetra-*O*-acetyl-2-deoxy-1,5-dithio-α-D-*gluco*-heptopyranoside

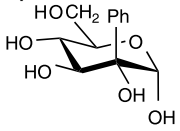
Note. It is common practice in carbohydrate names to regard 'thio' as detachable, and therefore alphabetized with any other prefixes.

2-Carb-16 Other Substituted Monosaccharides

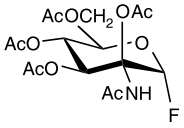
2-Carb-16.1 Replacement of Hydrogen at a Non-terminal Carbon Atom

The compound is named as a *C*-substituted monosaccharide. The group having priority according to the Sequence Rule ([13], Section E) is regarded as equivalent to OH for assignment of configuration. Any potential ambiguity (particularly when substitution is at the carbon atom where ring formation occurs) should be avoided by use of the *R*, *S* system to specify the modified stereocentre.

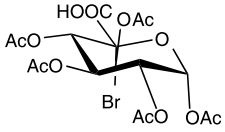
Examples:



2-*C*-Phenyl-α-D-glucopyranose



2-*C*-Acetamido-2,3,4,6-tetra-*O*-acetyl-α-D-mannopyranosyl fluoride

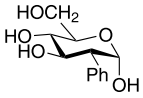


(5*R*)-1,2,3,4-Tetra-*O*-acetyl-5-bromo-α-D-*xylo*-hexopyranuronic acid
or 1,2,3,4-tetra-*O*-acetyl-5-bromo-β-L-idopyranuronic acid

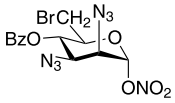
2-Carb-16.2 Replacement of OH at a Non-terminal, Non-anomeric Carbon Atom

The compound is named as a substituted derivative of a deoxy sugar. The group replacing OH determines the configurational description. Any potential ambiguity should be dealt with by the alternative use of the *R*, *S* system to specify the modified stereocentre.

Examples:



2-Deoxy-2-phenyl-α-D-glucopyranose or 2-deoxy-2-*C*-phenyl-α-D-glucopyranose
or (2*R*)-2-deoxy-2-phenyl-α-D-*arabino*-hexopyranose



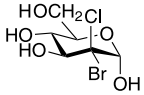
2,3-Diazido-4-*O*-benzoyl-6-bromo-2,3,6-trideoxy-α-D-mannopyranosyl nitrate

Note. Use of the symbol *C*- is essential only in cases of potential ambiguity, to make clear that substitution is at carbon rather than at a heteroatom (cf. 2-Carb-18.2); however, it may also be used for emphasis.

2-Carb-16.3 Unequal Substitution at a Non-terminal Carbon Atom

The compound is named as a disubstituted deoxy sugar. Configuration is determined by regarding the substituent having priority according to the Sequence Rule ([13], Section E), as equivalent to OH. Any potential ambiguity should be dealt with by the alternative use of the *R*, *S* system to specify the modified stereocentre.

Example:

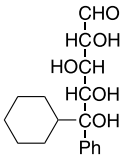


(2*R*)-2-Bromo-2-chloro-2-deoxy-α-D-*arabino*-hexose
or 2-bromo-2-chloro-2-deoxy-α-D-glucopyranose
(Br has priority over Cl)

2-Carb-16.4 Terminal Substitution

If substitution at the terminal carbon atom of the carbohydrate chain creates achiral centre, the stereochemistry is indicated by the *R*, *S* system.

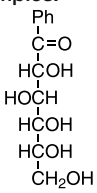
Example:



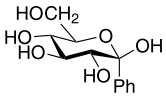
(5*R*)-5-*C*-Cyclohexyl-5-*C*-phenyl-D-xylose

Note. A monosaccharide with a terminal methyl group is named as a deoxy sugar, not as a *C*-methyl derivative. Substitution of aldehydic H by a ring or ring system is indicated simply with a *C*-substituent prefix.

Examples:



1-*C*-Phenyl-D-glucose
not 1-*C*-phenyl-D-*gluco*-hex-1-ulose

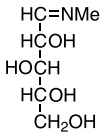


1-*C*-Phenyl-β-D-glucopyranose

2-Carb-16.5 Replacement of Carbonyl Oxygen by Nitrogen (Imines, Oximes, Hydrazones, Osazones etc.)

The imino analogue of a monosaccharide may be named as an imino-substituted deoxy alditol.

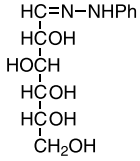
Example:



1-Deoxy-1-(methylimino)-D-xylitol

Oximes, hydrazones and analogues are named directly as oxime or hydrazone derivatives etc.

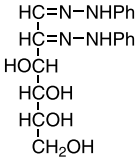
Example:



D-Glucose phenylhydrazone

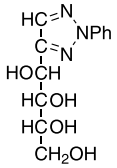
The vicinal dihydrazones formed from monosaccharides with arylhydrazines have been called arylosazones, but are preferably named as ketoaldose bis(phenylhydrazone)s

Example:



D-*arabino*-Hexos-2-ulose bis (phenylhydrazone)
or D-*arabino*-hex-2-ulose phenylosazone

The triazoles formed on oxidising arylosazones (commonly called osotriazoles) triazolylalditols.



D-*arabino*-Hexos-2-ulose phenylosotriazole
or (1*R*)-1-(2-phenyl-2*H*-1,2,3-triazol-4-yl))-D-erythritol
or 2-phenyl-4-(D-*arabino*-1,2,3,4-tetrahydroxybutyl)-2*H*-1,2,3-triazole

2-Carb-16.6 Isotopic Substitution and Isotopic Labelling

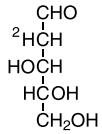
Rules for designating isotopic substitution and labelling are given in [13] (Section H). Parentheses indicate substitution; square brackets indicate labelling. The locant U indicates uniform labelling.

Examples:

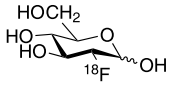
- D-(1-^{13}C)Glucose (substitution)
- D-(2-^{2}H)Mannose (substitution)
- L-[U-^{14}C]Arabinose (labelling)
- D-[1-^{3}H]Galactose (labelling)

When isotopic substitution creates a centre of chirality, configuration is defined as for other types of substitution (see ❯ *2-Carb-16.1* to ❯ *2-Carb-16.4*).

Example:



2-Deoxy-D-(2-^{2}H)lyxose (substituted)
or (2S)-2-deoxy-D-*threo*-(2-^{2}H)pentose



2-Deoxy-2-[^{18}F]fluoro-D-glucose (labelled)
or (2*R*)-2-deoxy-2-[^{18}F]fluoro-D-*arabino*-hexose

2-Carb-17 Unsaturated Monosaccharides[2]

2-Carb-17.1 General Principles

This section relates to the introduction of a double or triple bond between two contiguous carbon atoms of the backbone chain of a monosaccharide derivative. A double bond between a carbon atom of the backbone chain and an atom outside that chain, or a double or triple bond between two carbon atoms outside the backbone chain, will be treated according to the normal rules of organic nomenclature [13,14].

2-Carb-17.2 Double Bonds

Monosaccharide derivatives having a double bond between two contiguous carbon atoms of the backbone chain are named by inserting, into the name for the corresponding fully saturated derivative, the infix '*x*-en'. The infix is placed directly after the stem name that designates the chain length of the sugar. The locant *x* is the lower-numbered carbon atom involved in the double bond. Steric relations at a double bond are designated, if necessary, by the standard stereosymbols '(*Z*)-' and '(*E*)-' preceding the whole name ([13], Section E). For multiple double bonds, infixes such as '*x*, *y*-dien' are used (preceded by an inserted 'a' for euphony).

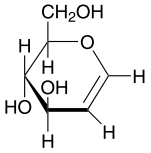
[2]This is based on the 1980 recommendations [5] Some examples have been omitted.

Note 1. The term 'glycal' is a non-preferred, trivial name for cyclic enol ether derivatives of sugars having a double bond between carbon atoms 1 and 2 of the ring. It should not be used or modified as a class name for monosaccharide derivatives having a double bond in any other position.

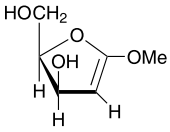
Note 2. Following the principle of first naming the saturated derivative, compounds having a C=CR-0-group as part of a ring system are named as unsaturated derivatives of anhydro alditols if R is hydrogen or carbon; if R is a halogen, chalcogen, or nitrogen-family element, the resulting name is that of a glycenose or glycenosyl derivative.

Note 3. The symbols (*Z*)- and (*E*)- may be omitted when the double bond is located within a ring system of six atoms or less, as steric constraints in such systems normally permit only one form.

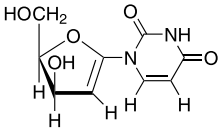
Examples:



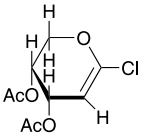
1,5-Anhydro-2-deoxy-D-*arabino*-hex-1-enitol (non-preferred trivial name D-glucal)



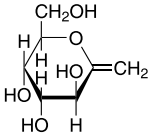
Methyl 2-deoxy-D-*threo*-pent-1-enofuranoside



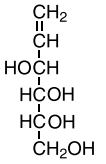
1-(2-Deoxy-D-*threo*-pent-1-enofuranosyl) uracil



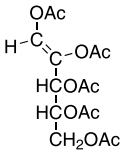
3,4-Di-*O*-acetyl-2-deoxy-D-*erythro*-pent-1-enopyranosyl chloride



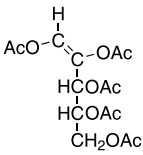
2,6-Anhydro-1-deoxy-D-*altro*-hept-1-enitol
(alphabetic preference over 2,6-anhydro-7-deoxy-D-*talo*-hept-6-enitol)



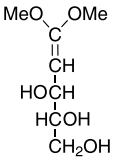
1,2-Dideoxy-D-*arabino*-hex-1-enitol



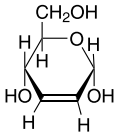
(*Z*)-1,2,3,4,5-Penta-*O*-acetyl-D-*erythbro*-pent-1-enitol



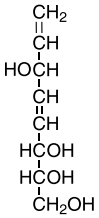
(*E*)-1,2,3,4,5-Penta-*O*-acetyl-D-*erythro*-pent-1-enitol



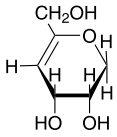
2-Deoxy-D-*threo*-pent-1-enose dimethyl acetal



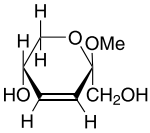
2,3-Dideoxy-α-D-*erythro*-hex-2-enopyranose



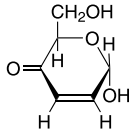
1,2,4,5-Tetradeoxy-D-*arabino*-octa-1,4-dienitol



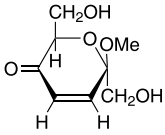
1,5-Anhydro-4-deoxy-D-*erythro*-hex-4-enitol
(enantiomeric precedence over 2,6-anhydro-3-deoxy-L-*erythro*-hex-2-enitol)



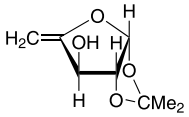
Methyl 3,4-dideoxy-β-D-*glycero*-hex-3-en-2-ulopyranoside



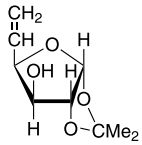
2,3-Dideoxy-α-D-*glycero*-hex-2-enopyranos-4-ulose



Methyl 3,4-dideoxy-β-D-*glycero*-hept-3-en-2-ulopyranosid-5-ulose



5-Deoxy-1,2-*O*-isopropylidene-β-L-*threo*-pent-4-enofuranose



5,6-Dideoxy-1,2-*O*-isopropylidene-α-D-*xylo*-hex-5-enofuranose

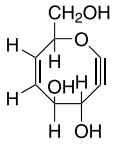
2-Carb-17.3 Triple Bonds and Cumulative Double Bonds

Monosaccharide derivatives having a triple bond or cumulative double bonds in the backbone chain are named by the methods of ❯ *2-Carb-17.2*, with the infix '*n*-yn' for a triple bond and infixes such as '*i*,*j*-dien' for cumulative double bonds.

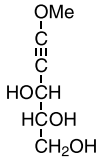
Note. This approach was not included in [5].

Alternatively they can be named on the basis of the corresponding fully saturated sugar by using the appropriate number of dehydro and deoxy prefixes (deoxy operations are regarded as formally preceding dehydro operations). The prefixes are placed in alphabetical order before the stem name.

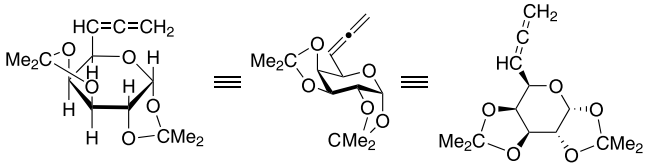
Examples:



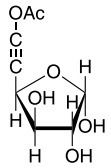
(*Z*)-1,7-Anhydro-2,5,6-trideoxy-D-*xylo*-oct-5-en-1-ynitol
or (*Z*)-1,7-anhydro-1,1,2,2-tetradehydro-2,5,6-trideoxy-D-*xylo*-oct-5-enitol



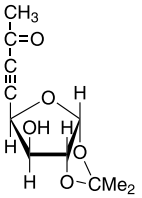
2-Deoxy-1-*O*-methyl-D-*threo*-pent-1-ynitol
or 1,1,2,2-tetradehydro-2-deoxy-1-*O*-methyl-D-*threo*-pentitol



6,7,8-Trideoxy-1,2:3,4-di-*O*-isopropylidene-α-D-*galacto*-octa-6,7-dienopyranose
or 6,7,7,8-tetradehydro-6,7,8-trideoxy-1,2:3,4-di-*O*-isopropylidene-α-D-*galacto*-octopyranose



6-*O*-Acetyl-5-deoxy-α-D-*xylo*-hex-5-ynofuranose
or 6-*O*-acetyl-5,5,6,6-tetradehydro-5-deoxy-α-D-*xylo*-hexofuranose



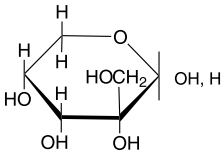
5,6,8-Trideoxy-1,2-*O*-isopropylidene-α-D-*xylo*-oct-5-ynofuranos-7-ulose
or 5,5,6,6-tetradehydro-5,6,8-trideoxy-1,2-*O*-isopropylidene-α-D-*xylo*-octofuranos-7-ulose

2-Carb-18 Branched-Chain Sugars[3]

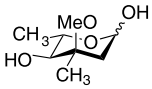
2-Carb-18.1 Trivial Names

Several branched monosaccharides have trivial names, some established by long usage. Examples are given below, together with systematic names for the (cyclic or acyclic) forms illustrated. (See also the alphabetical listing of trivial names in the Appendix.) Enantiomers of the sugars listed should be named systematically.

Examples:

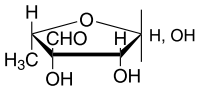


Hamamelose
2-*C*-(Hydroxymethyl)-D-ribopyranose

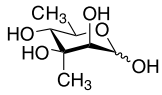


Cladinose
2,6-Dideoxy-3-*C*-methyl-3-*O*-methyl-L-*ribo*-hexopyranose

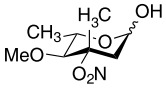
[3]This is a modified form of the 1980 recommendations [4]. Priority is now given to naming cyclic forms, since in most cases branched-chain monosaccharides will form cyclic hemiacetals or hemiketals.



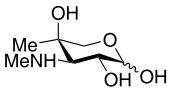
Streptose
5-Deoxy-3-*C*-formyl-L-lyxofuranose



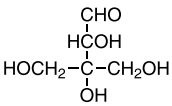
6-Deoxy-3-*C*-methyl-D-mannopyranose
(Evalose)



2,3,6-Trideoxy-3-*C*-methyl-4-*O*-methyl-3-nitro-L-*arabino*-hexopyranose
(Evernitrose)

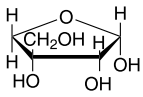


3-Deoxy-4-*C*-methyl-3-methylamino-L-arabinopyranose
(Garosamine)



D-Apiose (Api)
3-*C*-(Hydroxymethyl)-D-*glycero*-tetrose

Note. For the cyclic forms of apiose, systematic names are preferred, e. g.



3-*C*-(Hydroxymethyl)-α-D-erythrofuranose
[The name α-D-*erythro*-apiofuranose is ambiguous; Chemical Abstracts Service (CAS) uses the trivial name D-apio-α-D-furanose; Beilstein gives (3*R*)-α-D-apiofuranose]

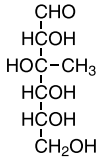
2-Carb-18.2 Systematic Names

A branched-chain monosaccharide is named as a substituted parent unbraxhed monosaccharide, as outlined in ❯ *2-Carb-16.1* to ❯ *2-Carb-16.4*.

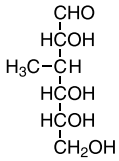
Note. *C*-Locants are essential only where there is potential ambiguity, to make clear whether substitution is at carbon or at a heteroatom (cf. ❯ *2-Carb-16*); however, they may also be used for emphasis.

2-Carb-18.3 Choice of Parent

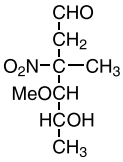
If the branched monosaccharide forms a cyclic hemiacetal or hemiketal, the chain which includes the ring atoms rather than any alternative open chain must be the basis of the name. Otherwise the parent is chosen according to the principles given in ❯ *2-Carb-2.1*.
Examples (see also ❯ *Chart V*):



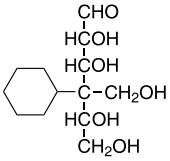
3-*C*-Methyl-D-glucose
(configuration determined by OH)



3-Deoxy-3-methyl-D-glucose
(configuration determined by CH_3)



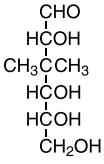
2,3,6-Trideoxy-3-*C*-methyl-4-*O*-methyl-3-nitro-D-*lyxo*-hexopyranose
(nitrogen has priority over carbon for determining configuration)



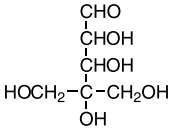
4-Cyclohexyl-4-deoxy-4-(hydroxymethyl)-D-allose
[oxygen (in CH_2OH) has priority over carbon (in cyclohexyl) at C-4]
or (4*R*)-4-cyclohexyl-4-deoxy-4-(hydroxymethyl)-D-*ribo*-hexose

If the two substituents at the branch point are identical, so that this centre has become achiral, the stereochemistry is specified as described in ❯ *2-Carb-8.4*.

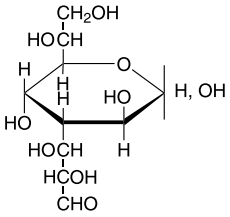
Examples:



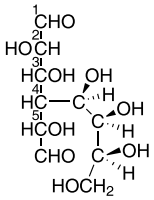
3-Deoxy-3,3-dimethyl-D-*ribo*-hexose



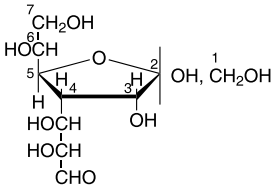
4-*C*-(Hydroxymethyl)-D-*erythro*-pentose



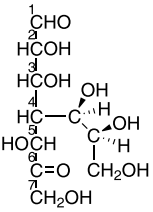
1



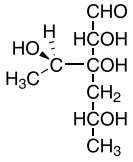
2



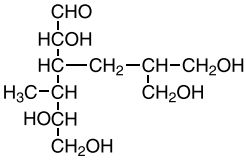
3



4



5



6

1 3-Deoxy-3-[(1*R*,2*S*)-1,2-dihydroxy-3-oxopropyl]-D-*glycero*-D-*altro*-heptopyranose
or 3-deoxy-3-(D-*threo*-1,2-dihydroxy-3-oxopropyl)-D-*glycero*-D-*altro*-heptopyranose
(not the alternative open-chain six-carbon dialdose or eight-carbon aldose, cf. **2**)

2 4-Deoxy-4-(L-*ribo*-1,2,3,4-tetrahydroxybutyl)-D-*altro*-hexodialdose

3 4-Deoxy-4-[(1*R*,2*R*)-1,2-dihydroxy-3-oxopropyl]-D-*allo*-heptulo-2,5-furanose
or 4-deoxy-4-(D-*erythro*-1,2-dihydroxy-3-oxopropyl)-D-*allo*-heptulo-2,5-furanose
(not the alternative ketoaldose, cf. **4**)

4 4-Deoxy-4-[(1*R*,2*S)*-1,2,3-trihydroxypropyl]-L-*talo*-heptos-6-ulose
or 4-deoxy-4-(L-*erytro*-1,2,3-trihydroxypropyl)-L-*talo*-heptos-6-ulose

5 4,6-Dideoxy-3-*C*-(D-*glycero*-1-hydroxyethyl)-D-*ribo*-hexose
(not the alternative pentose)

6 3,4-Dideoxy-3-[3-hydroxy-2-(hydroxymethyl)propyl]-4-*C*-methyl-L-mannose
(not the alternative *threo*-hexose)

◘ Chart V

Choice of parent in branchedchain monosaccharides. In the first names given for examples 1, 3 and 4, side-chain configuration is specified by use of *R* and *S*. This approach is generally preferred in all but the simplest cases, as less open to misinterpretation

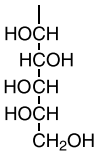
Note. Cyclization of the second example between C-1 and a CH_2OH group would necessitate a three-centre configurational prefix for the ring form.

Note. These recommendations may give rise to very different names for cyclic and acyclic forms of the same basic structures, resulting from different priorities. Thus, in ❯ *Chart V*, structures **1** and **2** are virtually identical, differing only by cyclization. The same holds for structures **3** and **4**.

2-Carb-18.4 Naming the Branches

Each branch will be named as an alkyl or substituted alkyl group replacing a hydrogen atom at the branch point of the parent chain. Within the branches, configurations around asymmetric centres can either be indicated using the *R*, *S*-system or, if they are carbohydrate-like and assignment is straightforward, by the use of the configurational prefix. For this purpose, and in the absence of a carbonyl group (or a terminal COOH or its equivalent) in the branch, the point of attachment of the branch (on the main chain) is regarded as to an aldehyde group.

Example:



L-*gluco*-1,2,3,4,5-Pentahydroxypentyl

If there is a carbonyl group in the branch (or a terminal COOH or its equivalent), its position (assigned lowest number when stereochemistry is being considered) is used to define the configurational prefix (see examples **1** and **3** in ❯ *Chart V*). Use of the *R*, *S* system is generally preferred, as less open to misinterpretation.
For an alternative approach to naming carbohydrate residues as substituents see ❯ *2-Carb-31.2* [which would give the name (1*R*)- or (1*S*)-L-arabinitol-1-*C*-yl for the above example, depending on the ligands at the branch point].

2-Carb-18.5 Numbering

The carbon atoms of the parent chain are numbered according to ❯ *2-Carb-2.2.1*. If a unique numbering is required for the branch(es) (e. g. for X-ray or NMR work), the carbon atoms may be given the locant of the appropriate branch point, with the internal branch locant as superscript, e. g. 4^2 for position 2 of the branch at position 4 of the main chain. This style of branch numbering is not to be used for naming purposes: e. g. the side-chain-methylated derivative of compound **5** is named 4,6-dideoxy-3-*C*-[(*R*)-1-methoxyethyl]-D-*ribo*-hexose, and not as a 3^1-*O*-methyl derivative.

2-Carb-18.6 Terminal Substitution

See ❯ *2-Carb-16.4*.

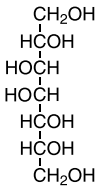
2-Carb-19 Alditols

2-Carb-19.1 Naming

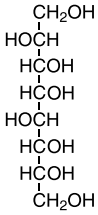
Alditols are named by changing the suffix '-ose' in the name of the corresponding aldose into '-itol'.

If the same alditol can be derived from either of two different aldoses, or from an aldose or a ketose, preference is ruled by ❯ *2-Carb-2.1* or ❯ *2.2.2* as appropriate.

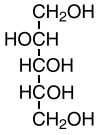
Examples:



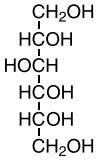
D-*glycero*-D-*galacto*-Heptitol
not L-*glycero*-D-*manno*-heptitol



D-*erythro*-L-*galacto*-Octitol
not D-*threo*-L-*gulo*-octitol

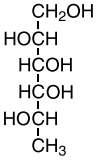


D-Arabinitol (Ara-ol) not D-lyxitol

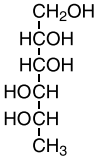


D-Glucitol (Glc-ol) not L-gulitol
(the trivial name sorbitol is not recommended)

The trivial names fucitol and rhamnitol are allowed for the alditols corresponding to the 6-deoxy sugars fucose and rhamnose.



L-Fucitol (L-Fuc-ol) or 1-deoxy-D-galactitol
not 6-deoxy-L-galactitol (cf. 2-Carb-2.2.3.1)



L-Rhamnitol (L-Rha-ol) or 1-deoxy-L-mannitol

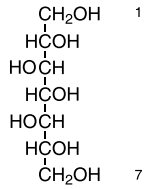
2-Carb-19.2 *meso* Forms

Alditols that are symmetric and therefore optically inactive – the *meso* forms – can be designated by the prefix *meso*-.

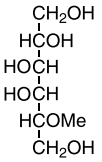
Examples:
***meso*-Erythriol**
***meso*-Ribitol**
***meso*-Galactitol**

The prefix D or L must be given when a derivative of a *meso* form has become asymmetric by substitution. It is also necessary to define the configurational prefixes as D or L in the case where there are more than four contiguous asymmetric carbon atoms.

Examples:



meso-D-*glycero*-L-*ido*-Heptitol
not L-*glycero*-D-*ido*-heptitol; cf. 2-Carb-2.2.3



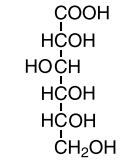
5-*O*-Methyl-D-galactitol
not 2-*O*-methyl-L-galactitol

2-Carb-20 Aldonic Acids

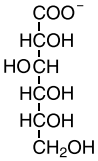
2-Carb-20.1 Naming

Aldonic acids are divided into aldotrionic acid, aldotetronic acids, aldopentonic acids, aldohexonic acids, etc., according to the number of carbon atoms in the chain. The names of individual compounds of this type are formed by replacing the ending '-ose' of the systematic or trivial name of the aldose by '-onic acid'.

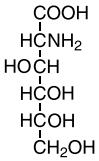
Examples:



D-Gluconic acid



D-Gluconate



2-Amino-2-deoxy-D-gluconic acid

2-Carb-20.2 Derivatives

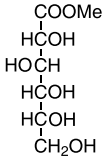
Salts are named by changing the ending '-onic acid' to '-onate', denoting the anion. If the counter ion is known, it is given before the aldonate name.

Example:
Sodium D-gluconate

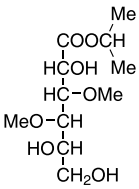
Esters derived from the acid function are also named using the ending '-onate'. The name of the alkyl (aryl, etc.) group is given before the aldonate name. Alternative periphrase names like 'aldonic acid alkyl (aryl, etc.) ester' may be suitable for an index.

Amides are designated by the ending '-onamide', and nitriles by the ending '-ononitrile'.

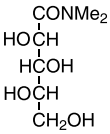
Examples:



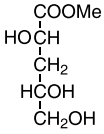
Methyl D-gluconate



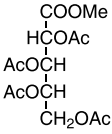
Isopropyl 3,4-di-*O*-methyl-L-mannonate



N,N-Dimethyl-L-xylonamide



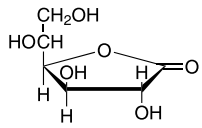
Methyl 3-deoxy-D-*threo*-pentonate



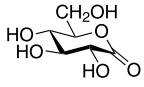
Methyl tetra-*O*-acetyl-L-arabinonate

Lactones are named with the ending '-onolactone'. The locants must be given (in the form '-ono-*i*, *j*-lactone): that of the carbonyl group (*i*) is cited first, and that of the oxygen (*j*) second (see examples below). Periphrase names (see alternatives in parentheses) appear widely in the literature but are not recommended. *Lactams* are named similarly by use of the ending '-onolactam'.

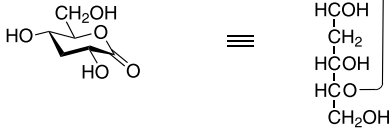
Examples:



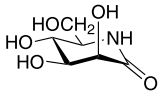
D-Glucono-1,4-lactone
(D-Gluconic acid γ-lactone)



D-Glucono-1,5-lactone
(D-Gluconic acid δ-lactone)



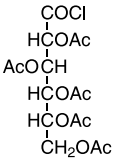
3-Deoxy-D-*ribo*-hexono-1,5-lactone



5-Amino-5-deoxy-D-mannono-1,5-lactam

Acyl halides are named by changing the ending '-onic acid' to '-onoyl halide;'.

Example:



Penta-*O*-acetyl-D-gluconoyl chloride

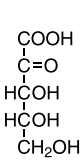
More complicated examples of general principles for naming acid derivatives can be found elsewhere [13,14].

2-Carb-21 Ketoaldonic Acids

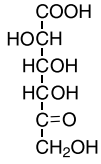
2-Carb-21.1 Naming

Names of individual ketoaldonic acids are formed by replacing the ending '-ulose' of the corresponding ketose by '-ulosonic acid', preceded by the locant of the ketonic carbonyl group. The anion takes the ending '-ulosonate'. The numbering starts at the carboxy group.
In glycosides derived from ketoaldonic acids, the ending is '-ulosidonic acid', with appropriate ring-size infix, e. g. '-ulopyranosidonic acid'.

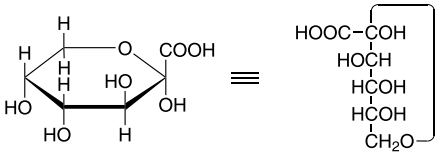
Examples:



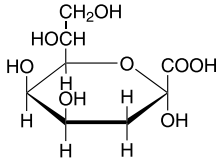
D-*erythro*-Pent-2-ulosonic acid



D-*arabino*-Hex-5-ulosonic acid



α-D-*arabino*-Hex-2-ulopyranosonic acid



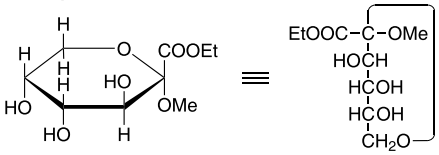
3-Deoxy-α-D-*manno*-oct-2-ulopyranosonic acid

Note. The last of the above examples is.one of the possible forms of the compound referred to by the three-letter symbol Kdo (formerly the abbreviation KDO, from the previously allowed trivial name ketodeoxyoctonic acid). Similarly the symbol Kdn for the C_9 sugar 3-deoxy-D-*glycero*-D-*galacto*-non-2-ulopyranosonic acid is widely used.

2-Carb-21.2 Derivatives

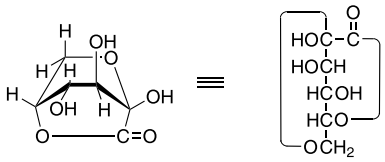
Esters, lactones, lactams, acyl halides etc. are named by modifying the ending '-ic acid' as described for aldonic acids (➤ *2-Carb-20.2*).

Examples:

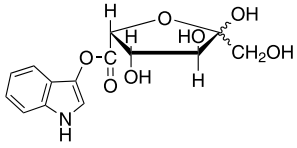


Ethyl (methyl α-D-*arabino*-hex-2-ulopyranosid)onate

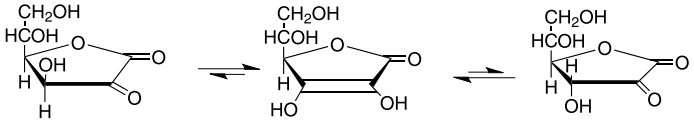
Note. The parentheses are inserted to distinguish between the ester alkyl group (cited first) and the glycosidic *O*-alkyl group.



β-D-*arabino*-Hex-2-ulopyranosono-1,5-lactone



Indol-3-yl D-*xylo*-hex-5-ulofuranosonate; trivial name isatan B



L-*xylo*-Hex-2-ulosono-1,4-lactone L-*threo*-Hex-2-enono-1,4-lactone L-*lyxo*-Hex-2-ulosono-1,4-lactone

(L-Ascorbic acid is the equilibrium mixture of all three isomers)

2-Carb-22 Uronic Acids

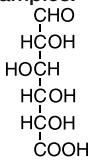
2-Carb-22.1 Naming and Numbering

The names of the individual compounds of this type are formed by replacing (a) the '-ose' of the systematic or trivial name of the aldose by '-uronic acid', (b) the '-oside' of the name of the glycoside by '-osiduronic acid' or (c) the '-osyl' of the name of the glycosyl group by '-osyluronic acid'. The carbon atom of the (potential) aldehydic carbonyl group (*not* that of the carboxy group as in normal systematic nomenclature [13,14]) is numbered 1 (see ➤ *2-Carb-2.1*, note 1).

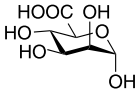
2-Carb-22.2 Derivatives

Derivatives of these acids formed by change in the carboxy group (salts, esters, lactones, acyl halides, amides, nitriles, etc.) are named according to ❯ *2-Carb-20.2*. The anion takes the ending '-uronate'. Esters are also named using the ending '-uronate'.

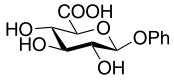
Examples:



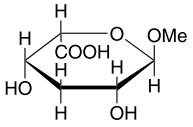
D-Glucuronic acid



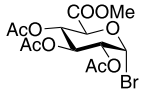
α-D-Mannopyranuronic acid



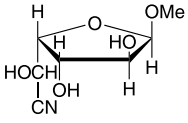
Phenyl β-D-glucopyranosiduronic acid
not phenyl β-D-glucuronoside or phenyl glucuronide



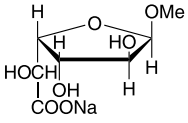
Methyl α-L-idopyranosiduronic acid



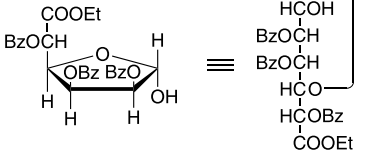
Methyl 2,3,4-tri-*O*-acetyl-α-D-glucopyranosyluronate bromide



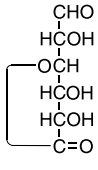
Methyl α-L-glucofuranosidurononitrile



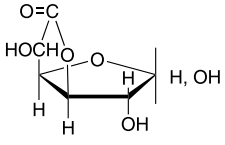
Sodium (methyl α-L-glucofuranosid)uronate



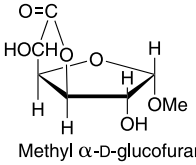
Ethyl 2,3,5-tri-*O*-benzoyl-α-D-mannofuranuronate



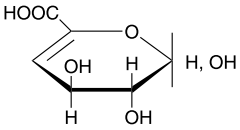
D-Glucurono-6,3-lactone



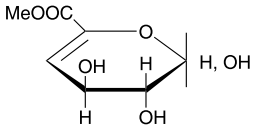
D-Glucofuranurono-6,3-lactone



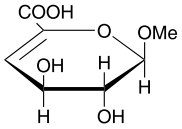
Methyl α-D-glucofuranosidurono-6,3-lactone



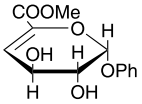
4-Deoxy-L-*threo*-hex-4-enopyranuronic acid



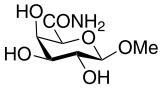
Methyl 4-deoxy-L-*threo*-hex-4-enopyranuronate



Methyl 4-deoxy-α-L-*threo*-hex-4-enopyranosiduronic acid



Methyl (phenyl-4-deoxy-β-L-*threo*-hex-4-enopyranosid)uronate



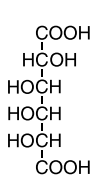
Methyl β-D-galactopyranosiduronamide

2-Carb-23 Aldaric Acids

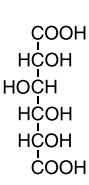
2-Carb-23.1 Naming

Names of individual aldaric acids are fonned by replacing the ending '-ose' of the systematic or trivial name of the parent aldose by '-aric acid'. Choice between possible names is based on 2-Carb-2.2.2.

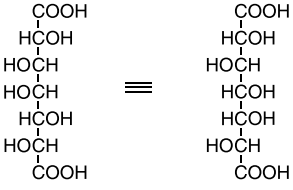
Examples:



L-Altraric acid
not L-talaric acid



D-Glucaric acid
not L-gularic acid

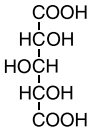


L-*glycero*-D-*galacto*-Heptaric acid
not L-*glycero*-D-*gluco*-heptaric acid

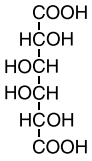
2-Carb-23.2 *meso* Forms

To the names of aldaric acids that are symmetrical, which therefore have no D- or L- prefix, the prefix '*meso*-' may be added for the sake of clarity. Examples: *meso*-erythraric acid, *meso*-ribaric acid, *meso*-xylaric acid, *meso*-allaric acid, *meso*-galactaric acid.
The D or L prefix must however be used when a *meso*-aldaric acid has become asymmetric as a result of substitution.

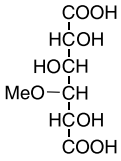
Examples:



meso-Xylaric acid



meso-Galactaric acid

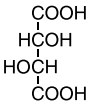


4-*O*-Methyl-D-galactaric acid
not 3-*O*-methyl-L-galactaric acid

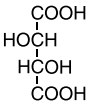
2-Carb-23.3 Trivial Names

For the tetraric acids, the trivial name tartaric acid remains in use, with the stereochemistry given using the *R*, *S* system. Esters are referred to as 'tartrates' (the second 'a' is elided).

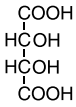
Examples:



(2*R*, 3*R*) - or (+)-Tartaric acid
or L-threaric acid



(2*S*, 3*S*) - or (–)-Tartaric acid
or D-threaric acid



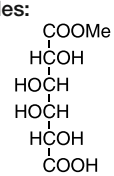
(2*R*, 3*R*) - or *meso*-Tartaric acid
or erythraric acid

Note. In the older literature, there is confusion about the use of D and L in the case of tartaric acids. It is therefore recommended to use the *R*, *S* system in this case.

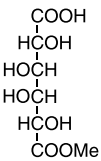
2-Carb-23.4 Derivatives

Derivatives formed by modifying the carboxy group (salts, esters, lactones, lactams, acyl halides, amides, nitriles etc.) are named by the methods of 2-Carb-20.2. Dilactones, half-esters, amic acids etc. are named by the methods of [13,14]. In cases of ambiguity, locants should be specified.

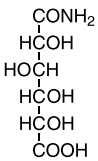
Examples:



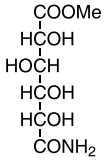
1-Methyl hydrogen D-galactarate



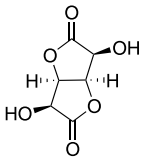
6-Methyl hydrogen D-galactarate



D-Glucar-1-amic acid



Methyl D-glucar-6-armate



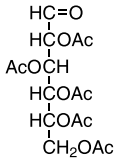
L-Mannaro-1,4:6,3-dilactone

2-Carb-24 *O*-Substitution

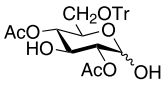
2-Carb-24.1 Acyl (Alkyl) Names

Substituents replacing the hydrogen atom of an alcoholic hydroxy group of a saccharide or saccharide derivative are denoted as *O*-substituents. The '*O*-' locant is not repeated for multiple replacements by the same atom or group. Number locants are used as necessary to specify the positions of substituents; they are not required for compounds fully substituted by identical groups. Alternative periphrase names for esters, ethers, etc. may be useful for indexing purposes. For cyclic acetals see 2-Carb-28.

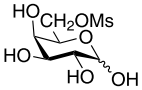
Examples:



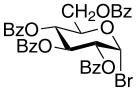
Penta-*O*-acetyl-*aldehydo*-D-glucose
or *aldehydo*-D-glucose pentaacetate



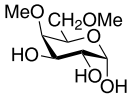
2,4-Di-*O*-acethyl-6-*O*-trityl-D-galactopyranose



6-*O*-Methanesulfonyl-D-galactopyranose
or D-galactopyranose 6-methanesulfonate



Tetra-*O*-benzoyl-α-D-glucopyranosyl bromide



4,6-Di-*O*-methyl-α-D-galactopyranose

Note. Acyl substituents on anomeric OH are designated (as above) by *O*-acyl prefixes. However, anomeric *O*-alkyl derivatives are named as glycosides (see 2-Carb-33).

2-Carb-24.2 Phosphorus Oxoacid Esters

24.2.1 Phosphates

Of special biochemical importance are the esters of monosaccharides with phosphoric acid. They are generally termed phosphates (e. g. glucose 6-phosphate). In biochemical use, the term 'phosphate' indicates the phosphate residue regardless of the state of ionization or the counter ions.

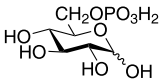
The prefix terms used for phosphate esters in organic nomenclature ([14], p. 65) are '*O*-phosphono-' and '*O*-phosphoato-' for the groups $(HO)_2P(O)$- and $(0^-)_2P(O)$- respectively, bonded to oxygen.

The term 'phospho-' is used for $(HO)_2P(O)$- or ionized forms in a biochemical context (see recommendations for the nomenclature of phosphorus-containing compounds [24]).

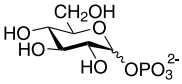
If a sugar is esterified with two or more phosphate groups, the compound is termed bisphosphate, trisphosphate etc. (e. g. fructofuranose 1,6-bisphosphate). The term diphosphate denotes an ester with diphosphoric acid, e. g. adenosine 5′-diphosphate.

Note. In abbreviations. a capital P is used to indicate a terminal $-PO_3H_2$ group or a non-terminal $-PO_2H-$ group (or dehydronated forms).

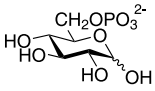
Examples:



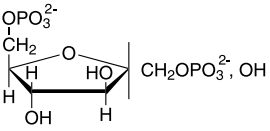
D-Glucopyranose 6-(dihydrogen phosphate)
or 6-*O*-phosphono-D-glucopyranose



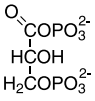
α-D-Glucopyranosyl phosphate
(biochemical usage: glucose 1-phosphate) (Glcl1P)



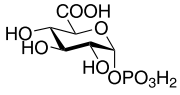
D-Glucopyranose 6-phosphate
(often shortened to glucose 6-phosphate)
or 6-*O*-phosphonato-D-glucopyranose
or 6-phospho-D-glucose (Glc6P)
(in a biochemical context)



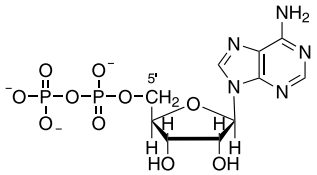
D-Fructofuranose 1,6-bisphosphate
(often shortened to fructose 1,6-bisphosphate)
or 1,6-di-*O*-phosphonato-D-fructofuranose
or 1,6-bisphospho-D-fructofuranose



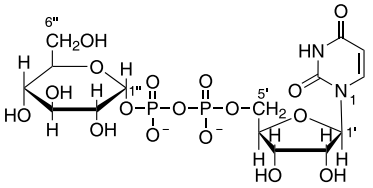
3-*O*-Phosphonato-D-glyceroyl phosphate
or 3-phospho-D-glyceroyl phosphate
or 1,3-bisphospho-D-glycerate (for biochemical usage)



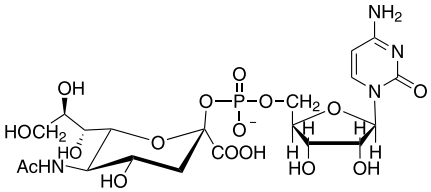
α-D-Glucopyranuronic acid
1-(dihydrogen phosphate)
(biochemical usage: glucuronate 1-phosphate) (GlcA1P)



Adenosine 5′-diphosphate (ADP) or 5′-diphosphoadenosine



Uridine 5′-(α-D-glucopyranosyl diphosphate)
(trivial name uridinediphosphoglucose) (UDP-Glc)

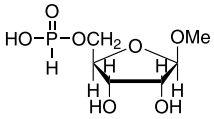


Cytidine 5′-(5-acetamido-3,5-dideoxy-D-*glycero*-β-D-galacto-non-2-ulopyranosylonic acid monophosphate) (CMP-Neu5Ac)

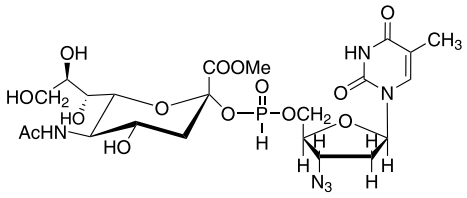
24.2.2 Phosphonates

The following examples illustrate the use of phosphonate terminology for esters of phosphonic acid, $HP(O)(OH)_2$. For formation of the alternative (substitutive) names, see 2-Carb-31.2.

Examples:



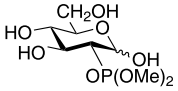
Methyl β-D-ribofuranoside 5-(hydrogen phosphonate)
or methyl 5-deoxy-β-D-ribofuranosid-5-yl hydrogen phosphonate



3'-Azido-3'-deoxythymidine 5'-[(methyl 5-acetamido-3,5-dideoxy-D-*glycerco*-α-D-*galacto*-non-2-ulopyranosylonate) phosphonate]

Derivatives substituted on phosphorus are named by standard procedures [13,14]; e. g. *P*-methyl derivatives are named as methylphosphonates.
Compounds with a phosphonate group linked by a P–C bond to a carbohydrate residue may be named as glycos-*n*-ylphosphonates (cf. ❯ *2-Carb-31.2*) or as *C*-substituted carbohydrates (cf. amino sugars, ❯ *2-Carb-14*).

Example:



2-Deoxy-2-dimethoxyphosphoryl-D-glucopyranose
(this usage of 'phosphoryl' is given in [13], Section D, Rule 5.68, and [14], p. 65)
or dimethyl 2-deoxy-D-glucopyranos-2-ylphosphonate

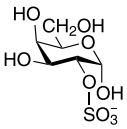
24.2.3 Phosphinates

Esters of phosphinic acid, $H_2P(O)(OH)$, are named by the same methods as used for phosphonates. For examples with two P–C bonds see ❯ *2-Carb-31.3*.

2-Carb-24.3 Sulfates

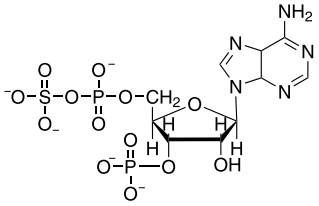
The prefix terms used for sulfuric esters are '*O*-sulfo-' and '*O*-sulfonato-', for the groups $(HO)S(O)_2$- and $(O^-)S(O)_2$- respectively, bonded to oxygen. Sulfates may also be named by citing the word 'sulfate', preceded by the appropriate locant, after the carbohydrate name.

Example:



α-D-Galactopyranose 2-sulfate
or 2-*O*-sulfonato-α-D-galactopyranose

The mixed sulfuric phosphoric anhydride (PAdoPS or PAPS) of 3′-phospho-5′-adenylic acid is named as an acyl sulfate:



3'-Phospho-5'-adenylyl sulfate (PAPS)

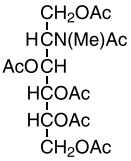
2-Carb-25 *N*-Substitution

Substitution, e. g. acylation, at the NH_2 group of an amino sugar can be dealt with in two different ways:

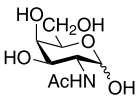
(a) The whole substituted amino group can be designated as a prefix, e. g. 2-acetamido- (or 2-butylamino-) 2-deoxy-D-glucose. For the purpose of the configurational prefix, the group is considered to take the place of the former OH group.
(b) If the amino sugar has a trivial name, the substitution is indicated by a prefix preceded by an italic capital *N*.

Note. In carbohydrate nomenclature, substitution at a heteroatom is normally indicated by citing the locant of the attached carbon atom, followed by a hyphen, and then the italicized heteroatom element symbol, e. g. 2-*O*-methyl, 5-*N*-acetyl. Substituents on the same kind of heteroatom are grouped (e. g. 2,3,4-tri-*O*-methyl), and substituents of the same kind are cited in alphabetical order of heteroatoms (e. g. 5-*N*-acetyl-4,8,9-tri-*O*-acetyl). The alternative format with superscript numerical locants (e. g. N^5, O^4, O^8, O^9-tetraacetyl), used in some other areas of natural product chemistry, is unusual in carbohydrate names.

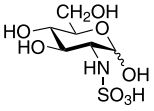
Examples:



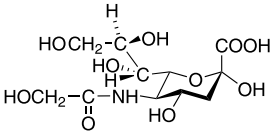
1,3,4,5,6-Penta-*O*-acetyl-2-deoxy-2-(*N*-methylacetamido)-D-glucitol



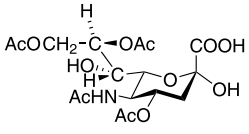
2-Acetamido-2-deoxy-D-galactopyranose
or *N*-acetyl-D-galactosamine



2-Deoxy-2-sulfoamino-D-glucopyranose
or *N*-sulfo-D-glucosamine



N-Glycoloyl-α-neuraminic acid (α-Neu5Gc)
(D is implied in the trivial name)



5-*N*-Acetyl-4,8,9-tri-*O*-acetyl-α-neuraminic acid
(α-Neu4,5,8,9Ac$_4$)

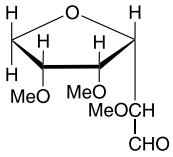
2-Carb-26 Intramolecular Anhydrides

An intramolecular ether (commonly called an intramolecular anhydride), formally arising by elimination of water from two hydroxy groups of a single molecule of a monosaccharide

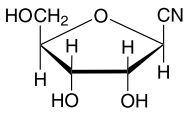
(aldose or ketose) or monosaccharide derivative, is named by attaching the (detachable) prefix 'anhydro-' preceded by a pair of locants identifying the two hydroxy groups involved.

Note. Detachable prefixes are cited in alphabetical order along with any substituent prefixes.

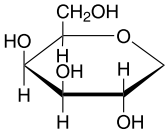
Examples:



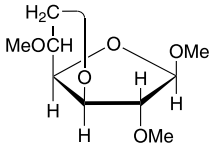
3,6-Anhydro-2,4,5-tri-*O*-methyl-D-glucose



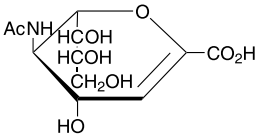
2,5-Anhydro-D-allononitrile



1,5-Anhydro-D-galactitol



Methyl 3,6-anhydro-2,5-di-*O*-methyl-β-D-glucofuranoside

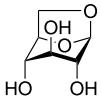


5-Acetamido-2,6-anhydro-3,5-dideoxy-D-*glycero*-D-*galacto*-non-2-enonic acid (Neu2en5Ac)

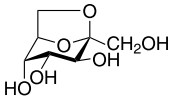
The compounds usually known as monosaccharide anhydrides or glycose anhydrides (earlier 'glycosans'), formation of which involves the anomeric hydroxy group, are named by the same procedure. In these cases the order of preference of ring size designators is pyranose > furanose > septanose. However, three- or four-membered rings should normally be cited as 'anhydro' if there is a choice.

Trivial names for anhydro monosaccharides, though established by usage, are not recommended because of possible confusion with polysaccharide names based on the use of the termination '-an'.

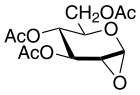
Examples:



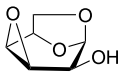
1,6-Anhydro-β-D-glucopyranose
not 1,5-anhydro-α-D-glucoseptanose
(older trivial name: levoglucosan)



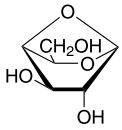
2.7-Anhydro-β-D-*altro*-hept-2-ulopyranose
(older trivial name: sedoheptulosan)



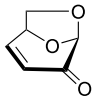
3,4,6-Tri-*O*-acetyl-1,2-anhydro-α-D-glucopyranose
not 3,4,6-tri-*O*-acetyl-1,5-anhydro-β-D-glucooxirose



1,6:3,4-Dianhydro-β-D-talopyranose



1,4-Anhydro-β-D-galactopyranose

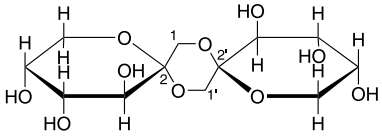


1,6-Anhydro-3,4-dideoxy-β-D-*glycero*-hex-3-enopyranos-2-ulose
(trivial name: levoglucosenone)

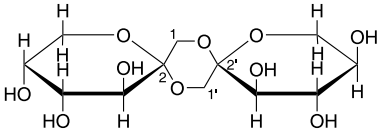
2-Carb-27 Intermolecular Anhydrides

The cyclic product of condensation of two monosaccharide molecules with the elimination of two molecules of water (commonly called an intermolecular anhydride), is named by placing the word 'dianhydride' after the names of the two parent monosaccharides. When the two parent monosaccharides are different, the one preferred according to the order of preference given in 2-Carb-2.1 is cited first. The position of each anhydride link is indicated by a pair of locants showing the positions of the two hydroxy groups involved; the locants relating to one monosaccharide (in a mixed dianhydride, the second monosaccharide named) are primed. Both pairs of locants immediately precede the word 'dianhydride'.

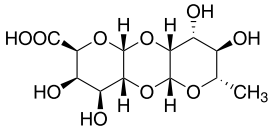
Examples:



α-D-Fructopyranose β-D-fructopyranose 1,2′:1′,2-dianhydride



α-D-Fructopyranose α-D-sorbopyranose 1,2′:1′,2-dianhydride

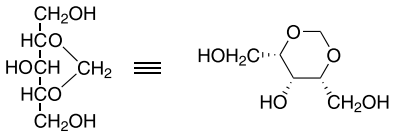


(α-D-Galactopyranuronic acid) β-L-rhamnopyranose 1,2′:1′,2-dianhydride

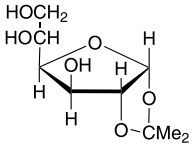
2-Carb-28 Cyclic Acetals

Cyclic acetals formed by the reaction of saccharides or saccharide derivatives with aldehydes or ketones are named in accordance with 2-Carb-24.1, bivalent substituent names (formed by general organic nomenclature principles) being used as prefixes. In indicating more than one cyclic acetal grouping of the same kind, the appropriate pairs of locants are separated typographically when the exact placement of the acetal groups is known.

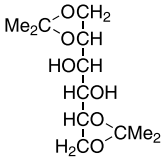
Examples:



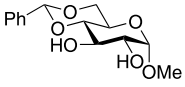
2,4-*O*-Methylenexylitol



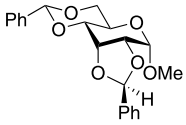
1,2-*O*-Isopropylidene-α-D-glucofuranose



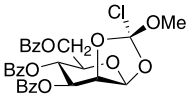
1,2:5,6-Di-*O*-isopropylidene-D-mannitol



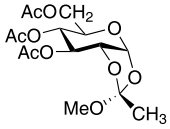
Methyl (*R*)-4,6-*O*-benzylidene-α-D-glucopyranoside



Methyl (*S*)-2,3:(*R*)-4,6-di-*O*-benzylidene-α-D-allopyranoside



3,4,6-Tri-*O*-benzoyl-[(*S*)-1,2-*O*-chloro (methoxy)methylene]-β-D-mannopyranose



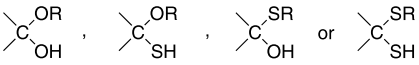
3,4,6-Tri-*O*-acetyl-α-D-glucopyranose (*R*)-1,2-(methyl orthoacetate)
or 3,4,6-tri-*O*-acetyl-[(*R*)-1,2-*O*-(1-methoxyethylidene)]-α-D-glucopyranose

Note 1. The last two examples contain cyclic ortho ester structures. These compounds are conveniently named as cyclic acetals.

Note 2. In the last four examples, new asymmetric centres have been introduced at the carbonyl carbon atom of the aldehyde or ketone that has reacted with the saccharide. When known, the stereochemistry at such a new centre is indicated by use of the appropriate *R* or *S* symbol ([13], Section E) placed in parentheses, immediately before the locants of the relevant prefix.

2-Carb-29 Hemiacetals, Hemiketals and Their Thio Analogues

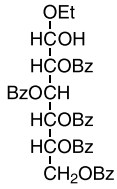
The compounds obtained by transforming the carbonyl group of the acyclic form of a saccharide, or saccharide derivative, into the grouping:



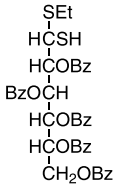
(R = alkyl or aryl)

are named as indicated in ❯ *2-Carb-30*, by using the terms 'hemiacetal', 'monothiohemiacetal', or 'dithiohemiacetal'(or the corresponding 'hemiketal' terms for ketone derivatives), as appropriate. The two isomers of a monothiohemiacetal are differentiated by use of *O* and *S* prefixes.

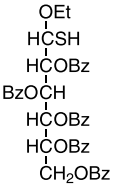
Examples:



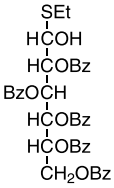
(1 *S*)-2,3,4,5,6-Penta-*O*-benzoyl-D-glucose ethyl hemiacetal



(1 *S*)-2,3,4,5,6-Penta-*O*-benzoyl-D-glucose ethyl dithiohemiacetal



(1 *R*)-2,3,4,5,6-Penta-*O*-benzoyl-D-glucose *O*-ethyl monothiohemiacetal

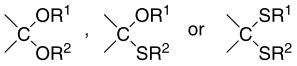


(1 *S*)-2,3,4,5,6-Penta-*O*-benzoyl-D-glucose *S*-ethyl monothiohemiacetal

Note. In these compounds carbon atom number 1 has become chiral. When known, the stereochemistry at this new chiral centre is indicated using the *R*, *S* system ([13], Section E).

2-Carb-30 Acetals, Ketals and Their Thio Analogues

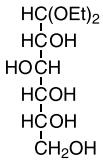
The compounds obtained by transfoming the carbonyl group of a saccharide or saccharide derivative into the grouping:



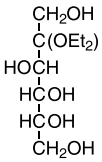
(R = alkyl or aryl)

are named by placing after the name of the saccharide or saccharide derivative the term 'acetal', 'monothioacetal' or 'dithioacetal' (or the corresponding 'ketal' terms for ketone derivatives) as appropriate, preceded by the names of the groups R^1 and R^2. With monothioacetals, the mode of bonding of two different groups R^1 and R^2 is indicated by the use of the prefixes *O* and *S*.

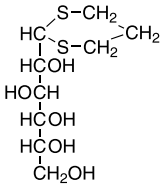
Examples:



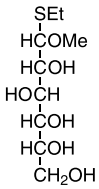
D-Glucose diethyl acetal



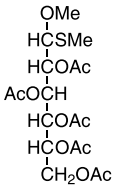
D-Fructose diethyl ketal



D-Glucose propane-1,3-diyl dithioacetal



(1 *S*)-D-Glucose *S*-ethyl *O*-methyl monothioacetal



(1 *R*)-2,3,4,5,6-Penta-*O*-acetyl-D-glucose dimethyl monothioacetal

Note. In the last two examples, carbon atom 1 has become chiral. When known, the stereochemistry at this new chiral centre is indicated by the *R*, *S* system, as specified in 2-Carb-29.

2-Carb-31 Names for Monosaccharide Residues

2-Carb-31.1 Glycosyl Residues

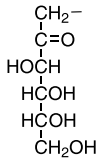
The residue formed by detaching the anomeric hydroxy group from a monosaccharide is named by replacing the terminal '-e' of the monosaccharide name by '-yl'. The general name

is 'glycosyl' residue. Terms of this type are widely used in naming glycosides and oligosaccharides. For examples (including glycosyl residues from uronic acids), see ➤ *2-Carb-33.2*. The term 'glycosyl' is also used in radicofunctional names, e. g. for halides such as the glucopyranosyl bromide in ➤ *2-Carb-24.1* and the mannopyranosyl fluoride in ➤ *2-Carb-16.1*, and esters such as the glucopyranosyl phosphate in ➤ *2-Carb-24.2.1* and the mannopyranosyl nitrate in ➤ *2-Carb-16.2*.

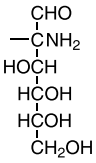
2-Carb-31.2 Monosaccharides as Substituent Groups

In order to produce names for structures in which it may be desirable for a non-carbohydrate portion to be cited as parent, prefix terms are required for carbohydrate residues linked through carbon or oxygen at any position on the main chain. These prefixes can be formed by replacing the final 'e' of the systematic or trivial name of a monosaccharide by '-*n*-C-yl', '-*n*-*O*-yl' or '-*n*-yl' (if there is no ambiguity). In each case the term '-yl' signifies loss of H from position *n*. At a secondary position (e. g. in 2-deoxy-D-glucos-2-y1, below) the free valency is regarded as equivalent to OH for assignment of configuration.

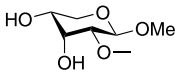
Examples:



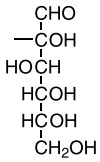
1-Deoxy-D-fructos-1-yl



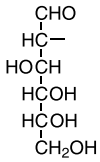
2-Amino-2-deoxy-D-glucos-2-yl



(Methyl β-D-ribopyranosid-2-*O*-yl)



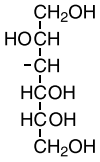
D-Glucose-2-*C*-yl



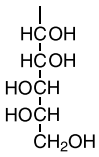
2-Deoxy-D-glucos-2-yl

The same endings can be used to form substituent prefixes for alditol residues.

Examples:



3-Deoxy-D-mannitol-3-yl

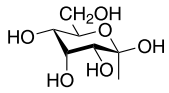


L-Arabinitol-1-*C*-yl (cf. 2-Carb-18.4)
(*R*/*S* to be specified at C-1)

The ending '-yl' without locants signifies loss of OH from the anomeric position (see 2-Carb-31.1). Loss of H from the anomeric OH is indicated by the ending '-yloxy', without locant. For examples see 2-Carb-33.

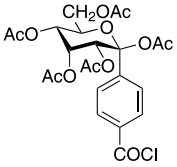
The situation in which the anomeric OH is retained but H is lost from the anomeric carbon atom is indicated by use of the ending '-yl' without locants in conjunction with the prefix '1-hydroxy-' (not by the ending '- 1-*C*-yl'). **N.B.** In this case, the anomeric prefix α or β refers to the free valency, not the OH group.

Example:

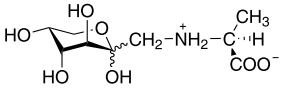


1-Hydroxy-α-D-allopyranosyl

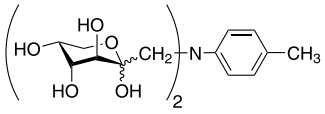
Examples of the use of substituent prefixes for carbohydrate residues:



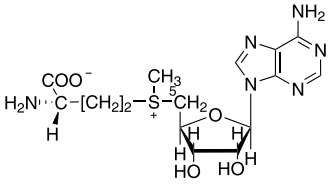
4-(1-Acetoxy-2,3,4,6-tetra-*O*-acetyl-α-D-allopyranosyl) benzoyl chloride



N-(1-Deoxy-D-fructopyranos-1-yl)-L-alanine

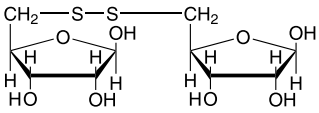


N,N-Bis-(1-deoxy-D-fructopyranos-1-yl)-*p*-toluidine

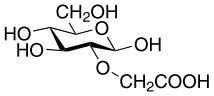


S-(5'-Deoxyadenosin-5'-yl)-L-methionine (AdoMet)

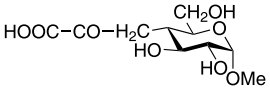
S-[(1-Adenin-9-yl)-1,5-dideoxy-β-D-ribofuranos-5-yl]-L-methionine
[trivial name *S*-adenosylmethionine (SAM)]



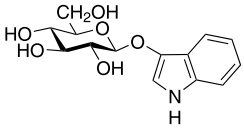
Bis(5-deoxy-β-D-ribofuranos-5-yl) disulfide
or bis(5-deoxy-β-D-ribofuranos-5-yl)disulfane



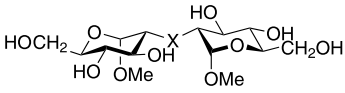
(β-D-Glucopyranos-2-*O*-yl)acetic acid
(more commonly named 2-*O*-carboxymethyl-β-D-glucopyranose; see 2-Carb-2.1, note 2)



(Methyl α-D-glucopyranosid-4-*O*-yl)pyruvic acid [or methyl 4-*O*-(oxalomethyl)-α-D-glucopyranoside]



3-(β-D-Glucopyranosyloxy) indole (or indol-3-yl β-D-glucopyranoside); trivial name indican



X=O

Methyl 2-*O*-(methyl 2-deoxy-α-D-glucopyranosid-2-yl)-α-D-glucopyranoside
or bis(methyl 2-deoxy-α-D-glucopyranosid-2-yl) ether

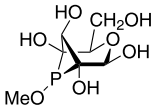
X=NH

Bis(methyl 2-deoxy-α-D-glucopyranosid-2-yl)amine

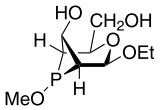
2-Carb-31.3 Bivalent and Tervalent Groups

The group formed by detaching one hydrogen atom from each of two (or three) carbon atoms of a monosaccharide is named by replacing the terminal '-e' of the monosaccharide name by '-diyl' (or '-triyl'), preceded by the appropriate locants.

Examples:



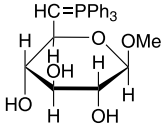
Methyl β-D-talopyranose-2-*C*,4-*C*-diylphosphinite
or 2-*C*,4-*C*-(methoxyphosphanediyl)-β-D-glucopyranose
or (2*R*,4*S*)-2-*C*,4-*C*-(methoxyphosphanediyl)-β-D-*threo*-hexopyrano



Methyl (ethyl 2,4-dideoxy-β-D-glucopyranoside-2,4-diyl)phosphinite
or ethyl 2,4-dideoxy-2,4-(methoxyphosphanediyl)-β-D-glucopyranoside

Residues formed by detaching two (or three) hydrogen atoms from the same carbon atom may be named similarly.

Example:



(Methyl 6-deoxy-β-D-glucopyranosid-6-ylidene)triphenyl-λ^5-phosphane
or methyl 6-deoxy-6-triphenyl-λ^5-phosphanylidene-β-D-glucopyranoside

Note. Names based on phosphane, rather than phosphine or phosphorane, are used in this document, as recommended in [14].

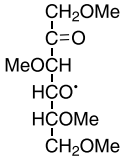
2-Carb-32 Radicals, Cations and Anions

Naming procedures described in this section follow the recommendations given in [25].

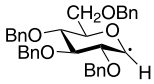
2-Carb-32.1 Radicals

Names for-radicals are formed in the same way as those for the corresponding substituent groups (see ❯ *2-Carb-31.2*).

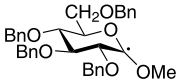
Examples:



1,3,5,6-Tetra-*O*-methyl-D-fructos-4-*O*-yl



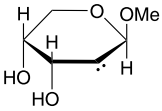
Tetra-*O*-benzyl-D-glucopyranosyl



2,3,4,6-Tetra-*O*-benzyl-1-methoxy-D-glucopyranosyl

Carbenes are named analogously by use of the suffix '-ylidene'.

Example:

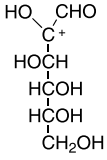


Methyl 2-deoxy-β-D-*erythro*-pentopyranosid-2-ylidene

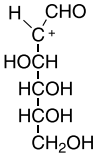
2-Carb-32.2 Cations

Cations produced by formal loss of H^- from a carbon atom are denoted by replacing terminal 'e' with the suffix '-ylium', in conjunction with appropriate locants and a 'deoxy-' prefix if necessary (cf. ❯ *2-Carb-31.2*).

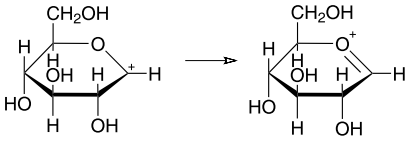
Examples:



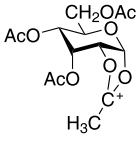
D-*arabino*-Hexos-2-*C*-ylium



2-Deoxy-D-*arabino*-hexos-2-ylium



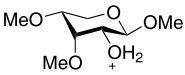
D-Glucopyranosylium



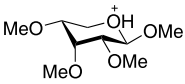
3,4,6-Tri-*O*-acetyl-1,2-*O*-ethyliumdiyl-α-D-allopyranose

Cations formed by hydronation of an OH group or at the hemiacetal ring oxygen are denoted by the suffix '-*O*-ium', with numerical locant.

Examples:



Methyl 3,4-di-*O*-methyl-β-D-ribopyranosid-2-*O*-ium

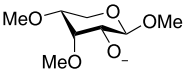


Methyl 2,3,4-tri-*O*-methyl-β-D-ribopyranosid-5-*O*-ium

2-Carb-32.3 Anions

Anions produced by formal loss of H^+ from an OH group are denoted by the suffix '-*O*-ate', with numerical locant.

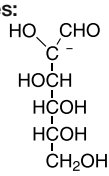
Example:



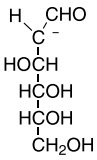
Methyl 3,4-di-*O*-methyl-β-D-ribopyranosid-2-*O*-ate

Anions produced by formal loss of H^+ from a carbon atom are denoted by the suffix '-ide', with appropriate locants and a 'deoxy-' prefix if necessary (cf. 2-Carb-31.2).

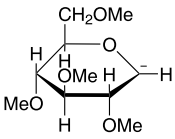
Examples:



D-*arabino*-Hexos-2-*C*-ide



2-Deoxy-D-*arabino*-hexos-2-ide

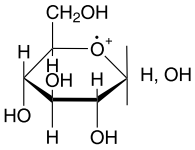


1,5-Anhydro-2,3,4,6-tetra-*O*-methyl-D-glucitol-1-ide

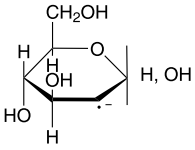
2-Carb-32.4 Radical Ions

Radical ions can be named by adding the suffix 'yl' to ion names. Alternatively, the words 'radical cation' or 'radical anion' may be added after the name of the parent with the same molecular formula, especially when the location of the radical ion centre is not to be specified.

Examples:



D-Glucopyranosiumyl, or D-glucopyranose radical cation



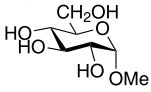
2-Deoxy-D-*arabino*-hexos-2-id-2-yl, or 2-deoxy-D-*arabino*-hexopyranos-2-ylidene radical anion

2-Carb-33 Glycosides and Glycosyl Compounds

2-Carb-33.1 Definitions

Glycosides were originally defined as mixed acetals (ketals) derived from cyclic forms of monosaccharides.

Example:



Methyl α-D-glucopyranoside

However, the term 'glycoside' was later extended to cover not only compounds in which, as above, the anomeric hydroxy group is replaced by a group -OR, but also those in which the replacing group is -SR (thioglycosides), -SeR (selenoglycosides), $-NR^1R^2$ (*N*-glycosides), or even $-CR^1R^2R^3$ (*C*-glycosides). 'Thioglycoside' and 'selenoglycoside' are legitimate generic terms; however the use of '*N*-glycoside' , although widespread in biochemical literature, is improper and not recommended here ('glycosylamine' is a perfectly acceptable term). '*C*-Glycoside' is even less acceptable (see Note to *2-Carb-33.7*). A glossary of terms based on 'glycose' is given in the Appendix.

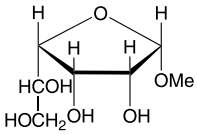
Particularly in naturally occurring glycosides, the compound ROH from which the carbohydrate residue has been removed is often termed the aglycone, and the carbohydrate residue itself is sometimes referred to as the 'glycone'.

Note. The spelling 'aglycon' is often encountered.

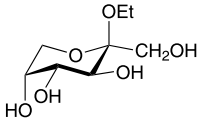
2-Carb-33.2 Glycosides

Glycosides can be named in three different ways:
(a) By replacing the terminal '-e' of the name of the corresponding cyclic form of the monosaccharide by '-ide' and preceding this, as a separate word (the intervening space is significant), the name of the group R (see examples below).

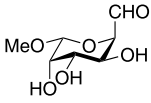
Examples:



Methyl α-D-gulofuranoside
not methyl-α-D-gulofuranosid



Ethyl β-D-fructopyranoside



Methyl (6*R*)-D-*gluco*-hexodialdo-6,2-pyranoside

Note. This is the 'classical' way of naming glycosides. It is used mainly when the group R is relatively simple (e. g. methyl, ethyl, phenyl).

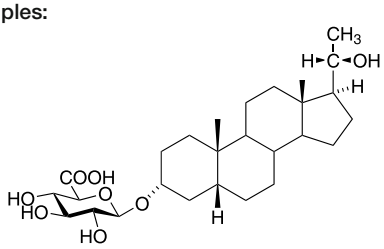
(b) By using the term 'glycosyloxy-', in the appropriate form for the monosaccharide, as prefix, for the name of the compound.

Note. This prefix includes the oxygen of the glycosidic bond. An example is given in 2-Carb-31.2; more are given below.

(c) By using the term '*O*-glycosyl-' as prefix to the name of the hydroxy compound.

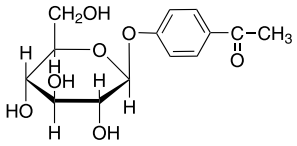
Note. This prefix does not include the oxygen of the glycosidic group. This is the appropriate method for naming natural products if the trivial name includes the OH group. The system is also used to name oligosaccharides (see 2-Carb-37).

Examples:

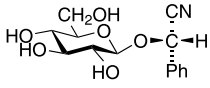


(20*S*)-20-Hydroxy-5β-pregnan-3α-yl β-D-glucopyranosiduronic acid
or (20*S*)-3α-(β-D-glucopyranosyloxyuronic acid)-5β-pregnan-20-ol;
for biochemical usage, pregnanediol 3-glucuronide

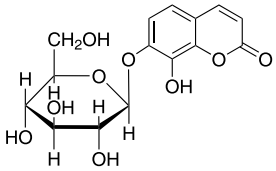
Note. A common biochemical practice would give the name (20*S*)-3α-(β-D-glucopyranuronosyloxy)-5β-pregnan-20-ol. This practice of naming glycosyl residues from uronic acids as 'glycuronosyl' is unsatisfactory because it implies the acceptance of the parent name 'glycuronose'. However the use of a two-word substituent prefix (glycosyloxyuronic acid), ending with a functional class name, remains inherently problematic, since it contravenes general organic nomenclature principles [13,14]. The latter practice has the advantage of retaining homomorphic relationships between glycoses and glycuronic acids.



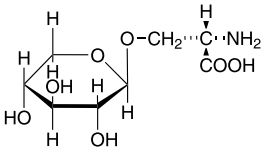
4-Acetylphenyl β-D-glucopyranoside
or 4′-(β-D-glucopyranosyloxy)acetophenone;
trivial name picein



(*S*)-*O*-β-D-Glucopyranosylmandelonitrile
or (*S*)-(β-D-glucopyranosyloxy)(phenyl)acetonitrile;
trivial name sambunigrin



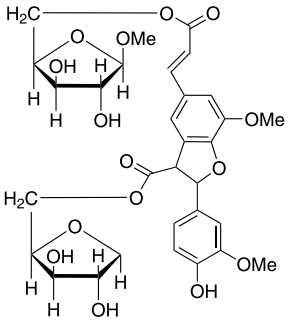
7-(β-D-Glucopyranosyloxy)-8-hydroxycoumarin;
trivial name daphnin



O-β-D-Xylopyranosyl-L-serine [(Xyl-)Ser]

Glycosides can be named as substituents by the methods of ❯ *2-Carb-31*.

Example:



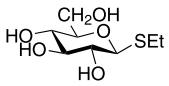
(Methyl 5-deoxy-β-D-xylofuranosid-5-yl) 2-(4-hydroxy-3-methoxyphenyl)-7-methoxy-
5-[2-(methyl β-D-xylofuranosid-5-*O*-ylcarbonyl)vinyl]-2,3-dihydrobenzofuran-3-carboxyl

2-Carb-33.3 Thioglycosides

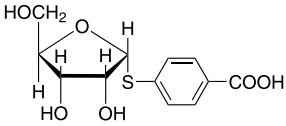
Names for individual compounds can be formed, like those for glycosides, in three ways, as follows.

(a) By using the term thioglycoside, preceded by the name of the group R.
(b) With the prefix 'glycosylthio-', followed by the name of the compound RH; this prefix includes the sulfur atom.
(c) With the prefix '*S*-glycosyl-' (not including the S atom), followed by the name of the thio compound. Sulfoxides and sulfones can also be named by functional class nomenclature [13,14].

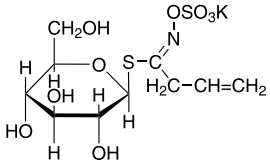
Examples:



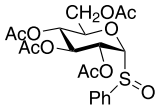
Ethyl 1-thio-β-D-glucopyranoside



4-(α-D-Ribofuranosylthio)benzoic acid
or 4-carboxyphenyl 1-thio-α-D-ribofuranoside



S-β-D-Glucopyranosyl (*Z*)-*O*-(potassium sulfonato)but-3-enehydroximothioate
(trivial name sinigrin)

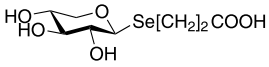


Phenyl tetra-*O*-acetyl-α-D-glucopyranosyl sulfoxide
or phenyl 2,3,4,6-tetra-*O*-acetyl-1-*thio*-α-D-glucopyranoside *S*-oxide

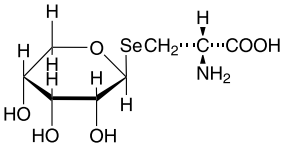
2-Carb-33.4 Selenoglycosides

Names are formed analogously to those for thioglycosides (2-Carb-33.3).

Examples:



2-Carboxyethyl 1-seleno-β-D-xylopyranoside
or 3-(β-D-xylopyranosylseleno)propanoic acid

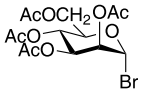


Se-β-D-Ribopyranosyl-D-selenocysteine
or (*S*)-2-amino-2-carboxyethyl 1-seleno-β-D-ribopyranoside
or 3-(β-D-ribopyranosylseleno)-D-alanine

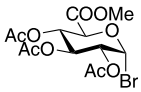
2-Carb-33.5 Glycosyl Halides

Compounds in which the anomeric hydroxy group is replaced by a halogen atom are named as glycosyl halides. Pseudohalides (azides, thiocyanates etc.) are named similarly.

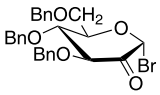
Examples:



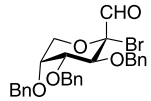
Tetra-*O*-acetyl-α-D-mannopyranosyl bromide



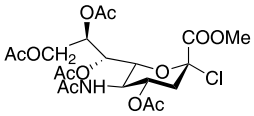
Methyl (2,3,4-tri-*O*-acetyl-α-D-glucopyranosyl)uronate bromide
not methyl 2,3,4-tri-*O*-acetyl-1-bromo-1-deoxy-α-D-glucopyranuronate



3,4,6-Tri-*O*-benzyl-α-D-*arabino*-hexopyranosyl-2-ulose bromide



3,4,5-Tri-*O*-benzyl-α-D-*arabino*-hexos-2-ulo-2,6-pyranosyl bromide
or 3,4,5-tri-*O*-benzyl-*aldehydo*-α-D-*arabino*-hexos-2-ulopyranosyl bromide

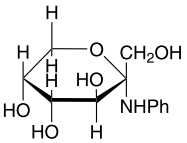


Methyl (5-acetamido-4,7,8,9-tetra-*O*-acetyl-3,5-dideoxy-D-*glycero*-α-D-*galacto*-non-2-ulopyranosyl)onate chloride

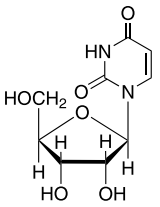
2-Carb-33.6 *N*-Glycosyl Compounds (Glycosylamines)

N-Glycosyl derivatives are conveniently named as glycosylamines. In the case of complex heterocyclic amines, such as nucleosides, the same approach is used.

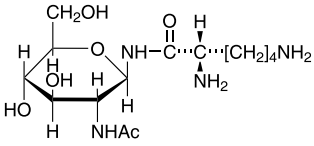
Examples:



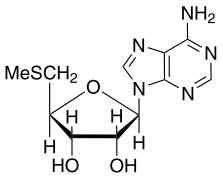
N-Phenyl-α-D-fructopyranosylamine
not aniline α-D-fructopyranoside



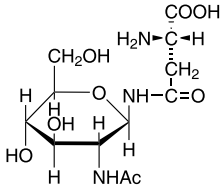
1-β-D-Ribofuranosyluracil (trivial name uridine)



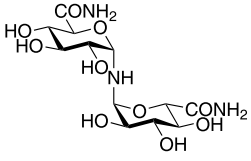
N^1-(2-Acetamido-2-deoxy-β-D-glucopyranosyl)-L-lysinamide (Lys-NH-GlcNAc)
[trivial name N^1-(*N*-acetylglucosaminyl)-L-lysinamide]



9-(5-*S*-Methyl-thio-β-D-ribofuranosyl)adenine



N^4-(2-Acetamido-2-deoxy-β-D-glucopyranosyl)-L-asparagine [(GlcNAc-)Asn]
or 2-acetamido-N^1-L-β-aspartyl-2-deoxy-β-D-glucopyranosylamine
(trivial name β-*N*-acetylglucosaminyl-L-asparagine)



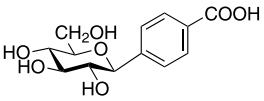
Bis(α-D-glucopyranosyluronamide)amine

2-Carb-33.7 *C*-Glycosyl Compounds

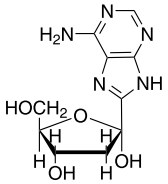
Compounds arising formally from the elimination of water from the glycosidic hydroxy group and an H atom bound to a carbon atom (thus creating a C-C bond) are named using the appropriate 'glycosyl-' prefixes (or other methods as appropriate, avoiding '*C*-glycoside' terminology).

Note. The term *C*-glycoside, introduced for naming pseudouridine (a nucleoside from transfer RNA), is a misnomer. All other glycosides are hydrolysable; the C-C bond of 'C-glycosides' is usually not. The use and propagation of names based on 'C-glycoside' terminology is therefore strongly discouraged.

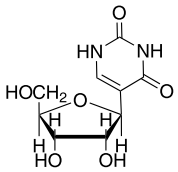
Example:



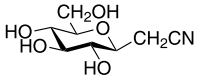
4-β-D-Glucopyranosylbenzoic acid
not 4-carboxyphenyl *C*-β-D-glucopyranosid



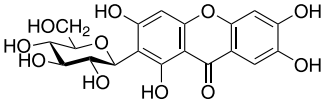
8-(2-Deoxy-β-D-*erythro*-pentofuranosyl)adenin
not adenine 8-(2-deoxyriboside)



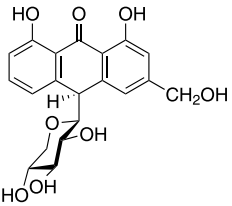
5-β-D-Ribofuranosyluracil; trivial name pseudouridine



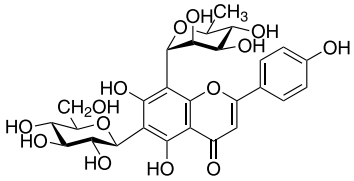
3,7-Anhydro-2-deoxy-D-*glycero*-D-*gulo*-octononitrile
or 2-*C*-(β-D-glucopyranosyl) acetonitrile
not cyanomethyl *C*-β-D-glucopyranoside



2-β-D-Glucopyranosyl-1,3,6,7-tetrahydroxyxanthen-9-one; trivial name mangiferin



(10*S*)-10-β-D-Glucopyranosyl-1,8-dihydroxy-3-(hydroxymethyl)anthracen-9(10*H*)-one; trivial names aloin A. (10*S*)-barbaloin



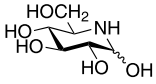
6-β-D-Glucopyranosyl-4',5,7-trihydroxy-8-α-L-rhamnopyranosylflavone; trivial name violanthin

2-Carb-34 Replacement of Ring Oxygen by Other Elements

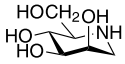
2-Carb-34.1 Replacement by Nitrogen or Phosphorus

Names should be based on those of the amino sugars (see ❯ *2-Carb-14*) (or the analogous phosphanyl sugars) with the amino or phosphanyl group at the non-anomeric position. Ring-size designators (furano, pyrano etc.) are the same as for the oxygen analogues.

Examples:

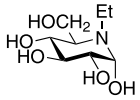


5-Amino-5-deoxy-D-glucopyranose; trivial name nojirimycin

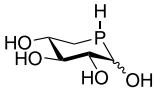


1-Amino-1,5-anhydro-1-deoxy-D-mannitol or 1,5-dideoxy-1,5-imino-D-mannitol; trivial name deoxymannojirimycin

Note the extension of the use of 'anhydro' in the above example to include the elimination of water between -NH_2 and -OH (Cf. ❯ *2-Carb-26*).

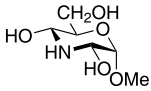


5-Deoxy-5-ethylamino-α-D-glucopyranose



5-Deoxy-5-phosphanyl-D-xylopyranose

Note. Use of the terms 'aza sugar', 'phospha sugar' etc. should be restricted to structures where carbon, not oxygen, is replaced by a heteroatom. Thus the structure below is a true aza sugar. The term 'imino sugar' may be used as a class name for cyclic sugar derivatives in which the ring oxygen atom has been replaced by nitrogen.



Methyl 3-deoxy-3-aza-α-D-*ribo*-hexopyranoside

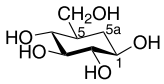
2-Carb-34.2 Replacement by Carbon

The (non-detachable) prefix 'carba-' signifies replacement of a heteroatom by carbon in general natural product nomenclature [26], and may be applied to replacement of the hemiacetal ring oxygen in carbohydrates if there is a desire to stress homomorphic relationships. If the original heteroatom is unnumbered, the new carbon atom is assigned the locant of the non-anomeric adjacent skeletal atom, with suffix 'a'.

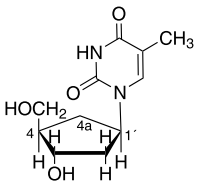
Note. The draft natural product rules [26] recommend that the new carbon atom takes the locant of the lower-numbered proximal atom. However, carbohydrate chemists regard the ring oxygen as formally originating from the non-anomeric (usually higher-numbered) position.

Additional stereochemistry (if any) at the new carbon centre is specified by use of the *R*/*S* system ([13], Section E).
Structures of this type can also be named as cyclitols [8].

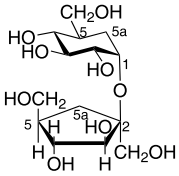
Examples:



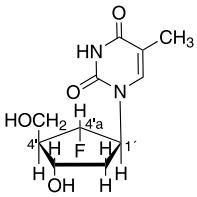
5a-Carba-β-D-glucopyranose



1-(2-Deoxy-4a-carba-β-D-*erythro*-pentofuranosyl)thymine
or 4'a-carbathymidine



5a-Carba-β-D-fructouranosyl 5a-carba-α-D-glucopyranoside



1-[(4a*S*)-2-Deoxy-4a-fluoro-4a-carba-β-D-*erythro*-pentofuranosyl]thymine
or (4'a*S*)-4'a-fluoro-4'a-carbathymidine

2-Carb-35 Carbohydrates Containing Additional Rings

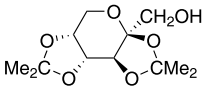
Internal bridging of carbohydrate structures by bivalent substituent groups creates additional rings, which can be named either by use of a substituent prefix representing the bridging group, or by fusion nomenclature. The following recommendations for the use of these two approaches are not thoroughly developed; they simply represent an attempt to rationalize ind codify current literature practice in the use of systems not in general well suited to carbohydrate applications. Bridging substituent prefix nomenclature (❯ *2-Carb-35.1*) is based on the system well established for simple cyclic acetals (❯ *2-Carb-28*), and fusion nomenclature (❯ *2-Carb-35.2*) on current literature usage and requirements for general natural product nomenclature [26].

2-Carb-35.1 Use of Bivalent Substituent Prefixes

Where the new bridge is attached to oxygen (or a replacement heteroatom, e. g. nitrogen in an amino sugar) already indicated in the name of the unbridged carbohydrate, the bivalent substituent prefix denotes substitution at two heteroatoms as outlined in ❯ *2-Carb-24.1* and ❯ *2-Carb-25* [method (b)]. Heteroatoms not directly bonded to the carbohydrate chain are regarded as part of the bridge.

Where the new bridge is attached through C-C bonds to the carbohydrate chain, the bridge prefix denotes a double *C*-substitution. Procedures are as outlined in ❯ *2-Carb-16*.

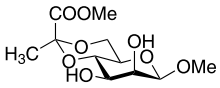
Examples:



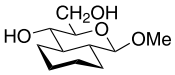
2,3:4,5-Di-*O*-isopropylidene-ß-D-fructopyranose

Note 1. The alternative fusion name (see ❯ *2-Carb-35.2*) is 2,2,2′,2′-tetramethyl-4,4′,5,5′-tetrahydro-(2,3,4,5-tetradeoxy-β-D-fructopyranoso)[2,3-*d*:4,5-*d*′]bis[1,3]dioxole; this is clearly less desirable on grounds of complexity.

Note 2. The use of prefixes ending in '-ylidene' for *gem*-bivalent substituent groups is traditional in the carbohydrate field, although no longer recommended in general organic nomenclature [14].

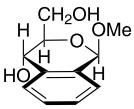


Methyl [(*S*)-4,6-*O*-(1-methoxycarbonylethylidene)]-β-D-mannopyranoside



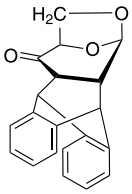
Methyl 2,3-(butane-1,4-diyl)-2,3-dideoxy-β-D-glucopyranoside

Note. The alternative fusion name (see 2-Carb-35.2) is hexahydro(methyl 2,3-dideoxy-*β*-D-glucopyranosido)[2,3]benzene

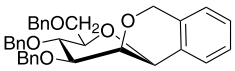


Methyl 2,3-(buta-1,3-diene-1,4-diyl)-2,3-dideoxy-β-D-*erythro*-hex-2-enopyranoside

Note. The alternative fusion name (see 2-Carb-35.2) is (methyl 2,3-dideoxy-*β*-D-*erythro*-hexopyranosido)[2,3]benzene



1,6-Anhydro-2,3-dideoxy-2,3-(9,10-dihydroanthracene-9,10-diyl)-β-D-*ribo*-hexopyranos-4-ulose

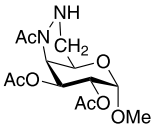


(1*S*)-1,5-Anhydro-3,4,6-tri-*O*-benzyl-1-*C*,2-*O*-(*o*-phenylenemethylene)-D-mannitol

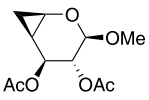
Note. The isomeric chromene would be named as a 2-*O*, 1-*C*-substituted system.

The prefix 'cyclo-' may be used for a single-bond bridge [14].

Examples:



Methyl 4-*N*-acetyl-2,3-di-*O*-acetyl-4,6-diamino-4,6-*N*-cyclo-4,6-dideoxy-α-D-galactopyranoside



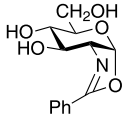
Methyl 2,3-di-*O*-acetyl-4,6-cyclo-4,6-dideoxy-β-D-galactopyranoside

2-Carb-35.2 Ring Fusion Methods

Fusion methods are employed as in general natural product nomenclature [26], except that the original carbohydrate ring is cited first, in parentheses (with terminal '-e', if present, replaced by '-o'). For designating stereochemistry, bonds in the new ring are considered as equivalent to OH, unless OH (or its equivalent) is still present at the ring junction. Substituents on the carbohydrate portion are included within the parentheses enclosing the fusion prefix. Substituents on the new ring (including 'hydro-' prefixes) precede the carbohydrate term(s). If there is a choice, the new ring is numbered in the direction used to define the fusion locants.

Note. General natural product fusion nomenclature [26] would require the carbohydrate portion to be cited last (e. g. oxazologlucopyranose), whereas it is cited first here and in the literature.

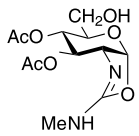
Examples:



2-Phenyl-4,5-dihydro-(1,2-dideoxy-α-D-glucopyranoso)[2,1-*d*]-1,3-oxazole

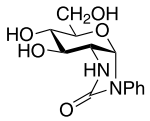
Note 1. The alternative name using a substituent prefix (see ❯ *2-Carb-35.1*) is 2-amino-1-*O*,2-*N*-(benzylylidene)-2-deoxy-α-D-glucopyranose.

Note 2. Literature fusion names for this type of compound use 'glucopyrano[2,l-d]oxazoline' terminology. However, names for partially hydrogenated heterocycles ending in 'oline' were abandoned by IUPAC in 1983 [27], in favour of 'dihydro.....ole'. Use of 'pyranoso' rather than 'pyrano' is recommended to avoid confusion with the normal fusion prefix from 'pyran' and to simplify rules for naming derivatives (e. g. glycosides).



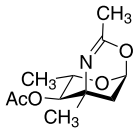
2-Methylamino-4,5-dihydro-(3,4,6-tri-*O*-acetyl-1,2-dideoxy-α-D-glucopyranoso)[2,1-*d*]-1,3-oxazole

Note. The alternative name using a substituent prefix (see 2-Carb-35.1) is 3,4,6-tri-*O*-acetyl-2-amino-2-deoxy-1-*O*,2-*N*-[(methylamino)methylylidene]-α-D-glucopyranose.



3-Phenyltetrahydro-(1,2-dideoxy-α-D-glucopyranoso)[1,2-*d*]imidazol-2-one

Note. The alternative name using a substituent prefix (see 2-Carb-35.1) is 2-amino-1,2-*N*-carbonyl-1,Z-dideoxy-1-*N*-phenyl-α-D-glucopyranosylamine



2-Methyl-5,6-dihydro-(4-*O*-acetyl-1,2,3,6-tetradeoxy-3-methyl-α-L-*ribo*-hexopyranoso)[3,2,1-de]-4*H*-1,3-oxazine

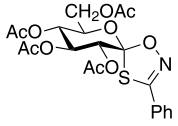
Note 1. The alternative name using a substituent prefix (see 2-Carb-35.1) is 4-*O*-acetyl-3-amino-2,3,6-trideoxy-1-*O*,3-*N*-(ethan-1-yl-1-ylidene)-3-C-methyl-α-L-ribo-hexopyranose

Note 2. This example would not normally be regarded as a fused system for nomenclature purposes, since it is not *ortho*- and *peri*-fused [13].

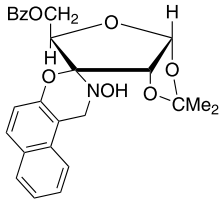
2-Carb-35.3 Spiro Systems

Spiro systems can be named by normal procedures [13]. For clarity, any anhydro or deoxy prefixes or chalcogen replacement prefixes (e. g. thio) referring to the spiro junction should appear next to the carbohydrate stem. The carbohydrate component is cited first. Configuration at the spiro junction is assigned by the *R*, *S* system.

Example:

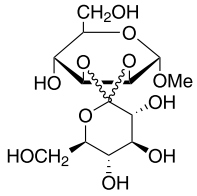


(1*R*)-2,3,4,6-Tetra-*O*-acetyl-3'-phenylspiro[1,5-anhydro-D-glucitol-1,5'-[1,4,2]oxathiazole]



(3*S*)-5-*O*-Benzoyl-1',2'-dihydro-1,2-*O*-isopropylidenespiro[3-deoxy-α-D-*elythro*-pentofuranose-3,3'-naphtho[1,2- e][1,3]oxazin]-2'-ol

The following spiro disaccharide example is best named by use of a *gem*-bivalent substituent prefix:



Methyl 2,3-*O*-D-glucopyranosylidene-α-D-mannopyranoside

Stereochemistry at C-1 of the glucose residue could be indicated as *R* or *S*, e. g. [(1*R*)-2,3-*O*-D-glucopyrano-sylidene]. . .

2-Carb-36 Disaccharides

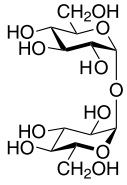
2-Carb-36.1 Definition

A disaccharide is a compound in which two monosaccharide units are joined by a glycosidic linkage.

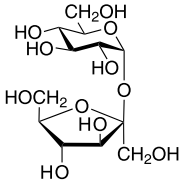
2-Carb-36.2 Disaccharides Without a Free Hemiacetal Group

Disaccharides which can be regarded as formed by reaction of the two glycosidic (anomeric) hydroxy groups with one another are named, systematically, as glycosyl glycosides. The parent (cited as the 'glycoside' component) is chosen according to ❯ *2-Carb-2.1*. Both anomeric descriptors must be included in the name.

Examples:



α-D-Glucopyranosyl α-D-glucopyranoside
[α-D-Glc*p*-(1↔1)-α-D-Glc*p*]
(trivial name α,α-trehalose)

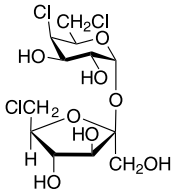


β-D-Fructofuranosyl α-D-glucopyranoside
[β-D-Fru*f*-(2↔1)-α-D-Glc*p*]
not α-D-glucopyranosyl β-D-fructofuranoside
(trivial names sucrose, saccharose)

Note. Such disaccharides are also known as non-reducing disaccharides.

If derivatives are named on the basis of the trivial name, the component cited first in the systematic name receives primed locants.

Example:



4,6,6'-Trichloro-4,6,6'-trideoxygalactosucrose
or 6-chloro-6-deoxy-β-D-fructofuranosyl 4,6-dichloro-4,6-dideoxy-α-D-galactopyranoside
('galactosucrose' is a trivial name for the 4-epimer of sucrose)

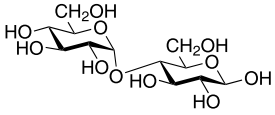
2-Carb-36.3 Disaccharides with a Free Hemiacetal Group

A disaccharide in which one glycosyl unit has replaced the hydrogen atom of an alcoholic hydroxy group of the other is named as a glycosylglycose. The locants of the glycosidic linkage and the anomeric descriptor(s) must be given in the full name.
There are two established methods in use for citing locants: either in parentheses between the glycosyl and glycose terms, or in front of the glycosyl prefix, as in the names of glycosides. The former (preferred) method is derived from that used to designate residues in oligosaccharides (see ❯ *2-Carb-37* and ❯ *2-Carb-38*).

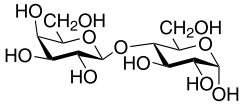
Note. The latter method is that used by Chemical Abstracts Service for disaccharides.

The *O*-locants used for the former method in the previous recommendations [1] are omitted here.

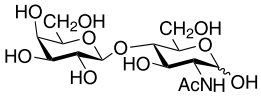
Examples:



α-D-Glucopyranosyl-(1→4)-β-D-glucopyranose
[α-D-Glc*p*-(1→4)-β-D-Glc*p*]
or 4-*O*-α-D-glucopyranosyl-β-D-glucopyrano
(trivial name β-maltose, not β-D-maltose)



β-D-Galactopyranosyl-(1→4)-α-D-glucopyranose
[β-D-Gal*p*-(1→4)-α-D-Glc*p*]
or 4-*O*-β-D-galactopyranosyl-α-D-glucopyranose
(trivial name α-lactose, not α-D-lactose)



β-D-Galactopyranosyl-(1→4)-*N*-acetyl-D-glucosamine (trivial name *N*-acetyllactosamine; LacNAc)

Note. Disaccharides with a free hemiacetal group are also known as reducing disaccharides.

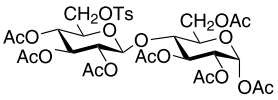
2-Carb-36.4 Trivial Names

Many of the naturally occurring disaccharides have well established trivial names. Some of these are listed below, together with the systematic names in both versions (see above).

Cellobiose β-D-Glucopyranosyl-(1→4)-D-glucose
4-*O*-β-D-Glucopyranosyl-D-glucose
Gentiobiose β-D-Glucopyranosyl-(1→6)-D-glucose
6-*O*-β-D-Glucopyranosyl-D-glucose
Isomaltose α-D-Glucopyranosyl-(1→6)-D-glucose
6-*O*-α-D-Glucopyranosyl-D-glucose
Melibiose α-D-Galactopyranosyl-(1→6)-D-glucose
6-*O*-α-D-Galactopyranosyl-D-glucose
Primeverose β-D-Xylopyranosyl-(1→6)-D-glucose
6-*O*-β-D-Xylopyranosyl-D-glucose
Rutinose α-L-Rhamnopyranosyl-(1→6)-D-glucose
6-*O*-α-L-Rhamnopyranosyl-D-glucose

The systematic names of trehalose, sucrose, maltose and lactose have been given already (with the formulae).
If derivatives are named on the basis of the trivial name, the component cited first in the systematic name receives primed locants.

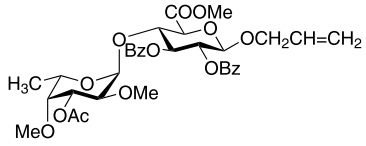
Example:



1,2,2',3,3',4',6-Hepta-*O*-acetyl-6'-*O*-tosyl-α-cellobiose
or 2,3,4-tri-*O*-acetyl-6-*O*-tosyl-β-D-glucopyranosyl-(1→4)-1,2,3,6-tetra-*O*-acetyl-α-D-glucopyranose

If the reducing terminal is a uronic ester glycoside, the ester alkyl group is cited at the beginning of the name, and the aglyconic alkyl group is cited with the name of the glycosidic residue.

Example:



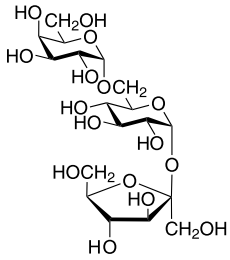
Methyl (3-*O*-acetyl-6-deoxy-2,4-di-*O*-methyl-α-L-galactopyranosyl)-(1→4)-(allyl 2,3-di-*O*-benzoyl-β-D-glucopyranosid)uronate

2-Carb-37 Higher Oligosaccharides

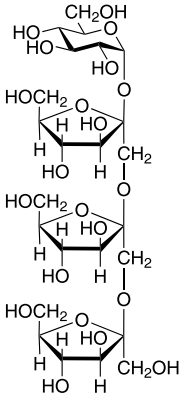
2-Carb-37.1 Oligosaccharides Without a Free Hemiacetal Group

Trisaccharides (for example) are named as glycosylglycosyl glycosides or glycosyl glycosy lglycosides as appropriate. A choice between the two residues linked through their anomeric positions for citation as the 'glycoside' portion can be made on the basis of ❯ *2-Carb-2.1*. Alternatively, a sequential (end-to-end) naming approach may be used, regardless of ❯ *2-Carb-2.1*. The names are formed by the preferred method of naming disaccharides (see ❯ *2-Carb-36.3*): the locant of the anomeric carbon atom, an arrow, and the locant of the connecting oxygen of the next monosaccharide unit are set in parentheses between the names of the residues concerned.

Examples:



β-D-Fructofuranosyl α-D-galactopyranosyl-(1→6)-α-D-glucopyranoside
(glucose preferred to fructose for citation as 'glycoside')
or α-D-galactopyranosyl-(1→6)-α-D-glucopyranosyl β-D-fructofuranoside
(sequential method)
[α-D-Gal*p*-(1→6)-α-D-Glc*p*-(1↔2)β-D-Fru*f*]
(trivial name raffinose)



β-D-Fructofuranosyl-(2→1)-β-D-fructofuranosyl-(2→1)-β-D-fructofuranosyl α-D-glucopyranoside
[β-D-Fru*f*-(2→1)-β-D-Fru*f*-(2→1)-β-D-Fru*f*-(2↔1) α-D-Glc*p*] (trivial name nystose)

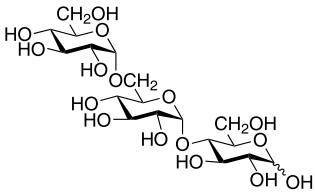
If derivatives are to be named on the basis of the trivial name, the component cited last in the systematic name receives locants with no primes, the preceding component singly-primed

locants, etc. However, naming of trisaccharide and higher oligosaccharide derivatives systematically is preferred, to avoid ambiguity.

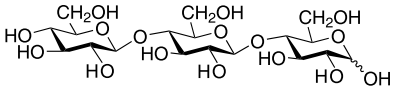
2-Carb-37.2 Oligosaccharides with a Free Hemiacetal Group

An oligosaccharide of this class is named as a glycosyl[glycosyl]$_n$glycose, i. e. the reducing sugar is the parent. Anomeric descriptors and locants are given as described in 2-Carb-37.1. The conventional depiction has the reducing sugar (glycose residue) on the right and the non-reducing end (glycosyl group) on the left. Internal sugar units are called glycosyl residues (the term 'anhydrosugar unit' is misleading and its use is discouraged). As the reducing end is often converted into the corresponding alditol, aldonic acid or glycoside derivative, the more general term 'downstream end' has been proposed for this end of the molecule.

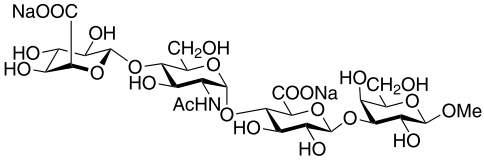
Examples:



α-D-Glucopyranosyl-(1→6)-α-D-glucopyranosyl-(1→4)-D-glucopyranose (trivial name panose)



β-D-Glucopyranosyl-(1→4)-β-D-glucopyranosyl-(1→4)-D-glucopyranose (trivial name cellotriose)

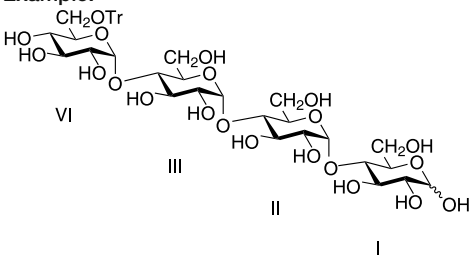


Methyl (sodium α-L-idopyranosyluronate)-(1→4)-(2-acetamido-2-deoxy-α-D-glucopyranosyl)-(1→4)-(sodium β-D-glucopyranosyluronate)-(1→3)-β-D-galactopyranoside
{Na_2[α-L-Ido*p*A-(1→4)-α-D-Glc*p*NAc-(1→4)-β-D-Glc*p*A-(1→3)-β-D-Gal*p*OMe]}

Higher oligosaccharides are named systematically in the same way. However, it is often preferable to give their structures by use of the symbolic approach outlined in 2-Carb-38.
Trivial names for linear oligosaccharides consisting only of 1→4 linked α-D-glucopyranosyl residues are maltotriose, maltotetraose etc. Similar names, based on the component sugar, are convenient for referring to other homo-oligosaccharides (e. g. xylobiose, galactotetraose), but such names should be used sparingly. Locants for naming substituted derivatives may be

obtained by assigning roman numerals to the residues in ascending order starting from the reducing end.

Example:



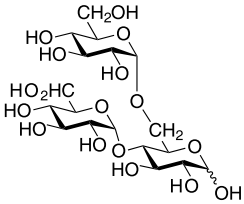
6^{VI}-*O*-Tritylmaltotetraose

Arabic numerals have also been used in this context, but confusion may result when component sugar residues have structural modifications (e. g. chain branches) requiring superscript locant numbers. The present recommendation follows long-established usage in glycolipids [21].

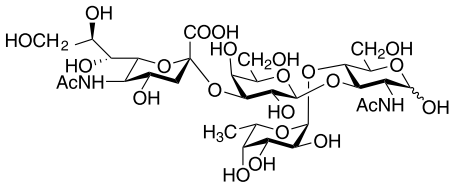
2-Carb-37.3 Branched Oligosaccharides

Terms designating branches should be enclosed in square brackets. In a branched chain, the longest chain is regarded as the parent. If two chains are of equal length the one with lower locants at the branch point is preferred, although some oligosaccharides are traditionally depicted otherwise, such as the blood group A trisaccharide exemplified below.

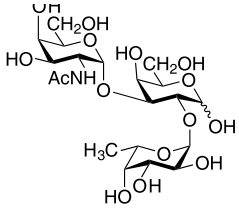
Examples:



α-D-Glucopyranosyl-(1→4)-[α-D-glucopyranosyl-(1→6)]-D-glucopyranose
[or 4,6-di-*O*-(α-D-glucopyranosyl)-D-glucopyranose]
(trivial name isopanose)



(5-Acetamido-3,5-dideoxy-D-*glycero*-α-D-*galacto*-non-2-ulopyranosylonic acid)-(2→3)-β-D-galactopyranosyl-(1→3)-[α-L-fucopyranosyl-(1→4)]-2-acetamido-2-deoxy-D-glucopyranose
or 5-*N*-acetyl-α-neuraminyl-(2→3)-β-D-galactopyranosyl-(1→3)-[α-L-fucopyranosyl-(1→4)]-2-acetamido-2-deoxy-D-glucopyranose
{α-Neu*p*5Ac-(2→3)-β-D-Gal*p*(1→3)-[α-L-Fuc*p*(1→4)]-D-Glc*p*NAc} (sialyl-Le^a trisaccharide)



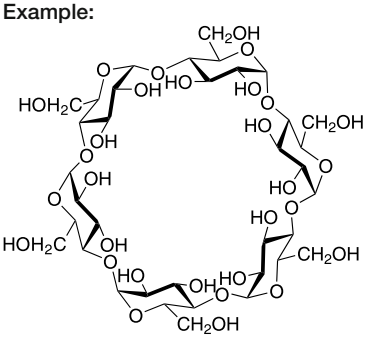
2-Acetamido-2-deoxy-α-D-galactopyranosyl-(1→3)-[α-L-fucopyranosyl-(1→2)]-D-galactopyranose
{α-D-Gal*p*NAc-(1→3)-[α-L-Fuc*p*(1→2)]-o-Gal*p*} (blood group A trisaccharide)

2-Carb-37.4 Cyclic Oligosaccharides

37.4.1 Semisystematic Names

Cyclic oligosaccharides composed of a single type of oligosaccharide unit may be named semisystematically by citing the prefix 'cyclo', followed by terms indicating the type of linkage [e. g. 'malto' for α-(1→4)-linked glucose units], the number of units (e. g. 'hexa' for six) and the termination '-ose'. The trivial names α-cyclodextrin (α-CD) for cyclomaltohexaose, β-cyclodextrin (β-CD) for cyclomaltoheptaose and γ-cyclodextrin (γ-CD) for cyclomaltooctaose are well established.

Example:



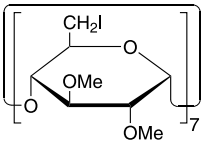
Cyclomaltohexaose (α-cyclodextrin, α-CD)

Structures with linkages other than (1→4) should be named systematically (see ❯ *2-Carb-37.4.2*).

Note. The cyclic oligosaccharides arising from enzymic transglycosylation of starch have been referred to as Schardinger dextrins. These names (and those of the cyclohexaamylose type) are not recommended, but the abbreviation CD is tolerated.

Derivatives with the same substitution pattern on each residue can be named semisystematically by assigning a single multiplicative prefix (e. g. hexakis, heptakis etc.) to the substituent prefixes as a group.

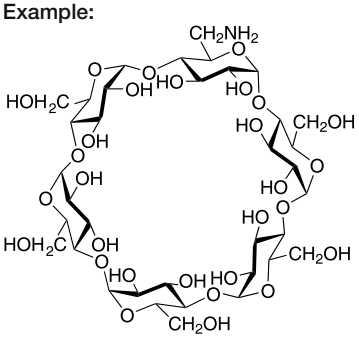
Example:



Heptakis(6-deoxy-6-iodo-2,3-di-*O*-methyl)cyclomaltoheptaose

Derivatives with different substitution patterns on the various residues can be named by the method of ❯ *2-Carb-37.2*, assigning a roman numeral to each residue.

Example:

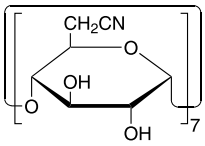


6'-Amino-6'-deoxycyclomaltohexaose

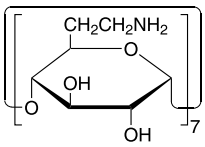
37.4.2 Systematic Names

Cyclic oligosaccharides composed of a single type of residue can be named by giving the systematic name of the glycosyl residue, preceded by the linkage type in parentheses, preceded in turn by 'cyclo-' with a multiplicative suffix (i. e. 'cyclohexakis-' etc.)

Examples:



Cycloheptakis-(1→4)-(6-deoxy-α-D-*gluco*-heptopyranosylurononitrile)



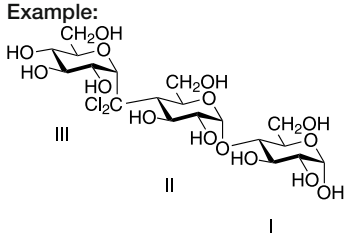
Cycloheptakis-(1→4)-(7-amino-6,7-dideoxy-α-D-*gluco*-heptopyranosyl)

The 1→6 isomer of cyclomaltohexaose should be named cyclohexakis-(1→6)-α-D-glucosyl, rather than cycloisomaltohexaose.

2-Carb-37.5 Oligosaccharide Analogues

Structures in which the linking glycosidic oxygen is replaced by -CH_2- may be named by use of the replacement prefix 'carba-' (cf. ❯ *2-Carb-34.2*) for emphasis of homomorphic relationships. The oxygen replaced is given the locant of the carbon atom to which it is attached in the residue with the lower roman numeral (cited as superscript) (ct ❯ *2-Carb-37.2*), with suffix 'a'.

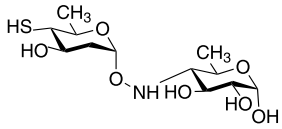
Example:



4'''a,4''a-Dichloro-4''a-carba-α-maltotriose

If the glycosidic oxygen link is replaced by -O-NH-, normal amino sugar nomenclature can be employed.

Example:



4-(2,6-Dideoxy-4-thio-α-D-*arabino*-hexopyranosyloxyamino)-4,6-dideoxy-α-D-glucopyranose

2-Carb-38 Use of Symbols for Defining Oligosaccharide Structures[4]

2-Carb-38.1 General Considerations

Oligosaccharide and polysaccharide structures occur not only in free form but often as parts of glycopeptides or glycoproteins [11] or of glycolipids [21]. It can be cumbersome to designate their structures by using the recommendations of ❯ *2-Carb-37*. The use of three-letter symbols for monosaccharide residues is therefore recommended. With appropriate locants and anomeric descriptors, long sequences can thus be adequately described in abbreviated form. Symbols for the common monosaccharide residues and derivatives are listed in ❯ *Table 2*. They are generally derived from the corresponding trivial names. Abbreviations for sub-

[4]The recommendations presented here are a modified version of the published 1980 recommendations [6]

Table 2
Symbols for monosaccharide residues and derivatives in oligosaccharide chains

Abequose	Abe	Iduronic acid	IdoA
Allose	All	Lyxose	Lyx
Altrose	Alt	Mannose	Man
Apiose	Api	Muramic acid	Mur
Arabinose	Ara	Neuraminic acid	Neu
Arabinitol	Ara-ol	*N*-Acetylneuraminic acid	Neu5Ac
2-Deoxyribose	dRib	*N*-Acetyl-2-deoxyneur-2-enaminicacid	Neu2en5Ac
Fructose	Fru	*N*-Glycoloylneuraminic acid	Neu5Gc
Fucose	Fuc	3-Deoxy-D-*manno*-oct-2-ulosonic acid	Kdo
Fucitol	Fuc-ol	Rhamnose	Rha
Galactose	Gal	3,4-Di-*O*-methylrhamnose	Rha3,4Mez
Galactosamine	GalN	Psicose	Psi
N-Acetylgalactosamine	GalNAc	Quinovose	Qui
β-D-GalaCtOpyranOSe 4-sulfate	β-D-Gal*p*4S	Ribose	Rib
Glucose	Glc	Ribose 5-phosphate	Rib5*P*
Glucosamine	GlcN	Ribulose	Ribulo (or Rul)
2,3-Diamino-2,3-dideoxy-D-glucose	GlcN3N	Sorbose	Sor
Glucitol	Glc-ol	Tagatose	Tag
N-Acetylglucosamine	GlcNAc	Talose	Tal
Glucuronic acid	GlcA	Xylose	Xyl
Ethyl glucopyranuronate	Glc*p*A6Et	Xylulose	Xylulo (or Xul)
Gulose	Gul	2-CMethylxylose	Xyl2*C*Me
Idose	Ido		

stituents (see 2-Carb-1.16.2), preceded by locants, follow the monosaccharide abbreviations directly.

2-Carb-38.2 Representations of Sugar Chains

For writing the structure of an oligo- or poly-saccharide chain, the glycose residue [the 'reducing group', i. e. the residue with the free hemiacetal group or modification thereof (e. g. alditol, aldonic acid, glycoside)] should be at the right-hand end. Also, when there is a glycosyl linkage to a non-carbohydrate moiety (e. g. protein, peptide or lipid) the glycosyl residue involved should appear at the right.
Numbering of monosaccharide units, if desired, should proceed from right to left.

2-Carb-38.3 The Extended Form

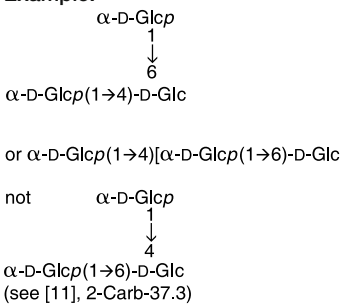
This is the form employed by the carbohydrate databank CarbBank, and is preferred for most purposes. Each symbol for a monosaccharide unit is preceded by the anomeric descriptor and

the configuration symbol. The ring size is indicated by an italic *f* for furanose or *p* for pyranose, etc. The locants of the linkage are given in parentheses between the symbols; a double-headed arrow indicates a linkage between two anomeric positions. In CarbBank, omission of α/β, D/L, or *f*/*p* means that this structural detail is not known.

Examples:
α-D-Gal*p*-(1→6)-α-D-Glc*p*-(1↔2)-β-D-Fruf for raffinose (see 2-Carb-37.1)
β-D-Glc*p*-(1→4)-β-D-Glc*p*-(1→4)-D-Glc for cellotriose (see 2-Carb-37.2)

Branches are written on a second line, or in brackets on the same line.

Example:



or α-D-Glc*p*(1→4)[α-D-Glc*p*(1→6)-D-Glc

(see [11], 2-Carb-37.3)

for 4,6-Di-*O*-α-D-glucopyranosyl-D-glucose

The hyphens may be omitted, except that separating the configurational symbol and the three-letter symbol for the monosaccharide.

2-Carb-38.4 The Condensed Form

In the condensed form, the configurational symbol and the letter denoting ring size are omitted. It is understood that the configuration is D (with the exception of fucose and iduronic acid which are usually L) and that the rings are in pyranose form unless otherwise specified. The anomeric descriptor is written in the parentheses with the locants.

Example:
Gal(α1-6)Glc(α1-2β)Fruf for raffinose

For most purposes, the short form (2-Carb-38.5) is preferred when abbreviation of the extended form is desirable.

2-Carb-38.5 The Short Form

For longer sequences, it is desirable to shorten the notation even further by omitting (i) locants of anomeric carbon atoms, (ii) the parentheses around the locants of the linkage and (iii) hyphens (if desired). Branches can be indicated on the same line by using appropriate enclosing marks (parentheses, square brackets etc.). Whenever necessary, configuration symbols and ring size designators etc. may be included, to make the notation more specific.

Example:
Galα-6Glcα-βFruf or Galα6Glcαβ*F*ruf for raffinose

The following examples show all three representations of the same structure:

a

extended form

β-D-Gal*p*-(1→4)-β-D-Glc*p*NAc-(1→2)-α-D-Man*p*-(1→
3
↑
1
α-L-Fuc*p*

b

condensed form in two lines

Gal(β1-4)GlcNAc(β1-2)Man(α1-
|
Fuc(α1-3)

or in one line

Gal(β1-4)[Fuc(α1-3)]GlcNAc(β1-2)Man(α1-

c

short form

Galβ-4(Fucα-3)GlcNAcβ-2Manα-
or Galβ4(Fucα3)GlcNAcβ2Manα-
or Galβ4GlcNAcβ2Manα-
|
Fucα-3

This last version is recommended as the most explicit representation of branching using the short form.

Note. These representations do not follow the recommendat ions for choice of main chain given in 2-Carb-37.3. Such deviations are common in depicting series of naturally occurring oligosaccharides where it is desirable to show homomorphic relationships.

2-Carb-39 Polysaccharides[5]

2-Carb-39.1 Names for Homopolysaccharides

A general term for a polysaccharide (glycan) composed of a single type of monosaccharide residue is obtained by replacing the ending '-ose' of the sugar name by '-an'.

Note. Examples of established usage of the '-an' ending are: xylan for polymers of xylose, mannan for polymers of mannose, and galactan for polymers of galactose. Cellulose and starch are both glucans, as they are composed of glucose residues.

2-Carb-39.2 Designation of Configuration of Residues

When the configurational series of the monomer residues is known, D- or L- may be included as a prefix to the name.

[5] This is a modified version of the 1980 recommendations on polysaccharide nomenclature [7]

Examples:

[4)-α-D-Glc*p*-(1→3)-α-D-Glc*p*-(1→]$_n$

A D-glucan (nigeran)

[5)-α-L-Ara*f*-(1→5)-α-L-Ara*f*-(1→5)-α-L-Ara*f*-(1→5)-α-L-Ara*f*-(1→]$_n$

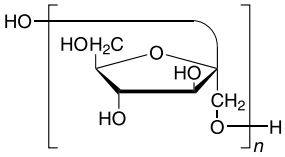
3 3
↑ ↑
1 1
α-L-Ara*f* α-L-Ara*f*

An L-arabinan (more specifically an α-L-arabinan)

2-Carb-39.3 Designation of Linkage

When the major linkage in a homopolysaccharide is known, it may be indicated in the name. The linkage designation shows the carbon atoms involved in the glycosidic bonds. When specific sugars are designated, notation for glycosidic linkages should precede the symbols designating the configuration of the sugar; thus, (1→4)-α-D-glucan.

Examples:

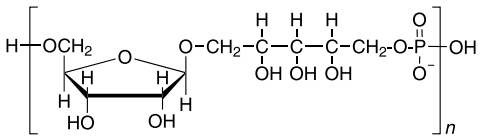


(2→1)-β-D-Fructofuranan (inulin has this structure, with a terminal α-D-glucopyranosyl group)

[4)-α-D-Glc*p*-(1→]$_n$

(1→4)-α-D-Glucopyranan (amylose)

Note. When the linkage between monosaccharide units is non-glycosidic (as in the phosphate derivative shown below), use of the glycan terminology is inappropriate; other methods of polymer nomenclature should be employed [20].



Poly[(β-D-ribofuranosyl-5-*O*-yl)(1-deoxy-D-ribitol-1-*C*,5-*O*-diyl)(oxidophosphoryl)]

[5)-β-D-Rib*f*-(1→1)-D-Rib-ol-5-*P*-(O→]$_n$

Such structures do not conform to the original strict definition of 'polysaccharide' but are generally classified as polysaccharides in current practice.

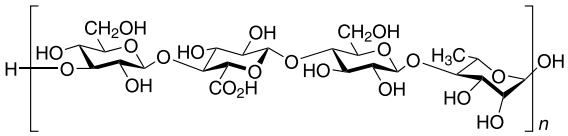
2-Carb-39.4 Naming of Newly Discovered Polysaccharides

Names assigned to newly discovered polysaccharides should end in '-an'.

Examples:

[6)-β-D-Glc*p*-(1→]$_n$

Pustulan (a glucan from the lichen *Umbilicaria pustulata*)



[3)-β-D-Glc*p*-(1→4)-β-D-Glc*p*A-(1→4)-β-D-Glc*p*-(1→4)-α-L-Rha*p*(1→]$_n$

Gellan (a bacterial polysaccharide originally designated S-60)

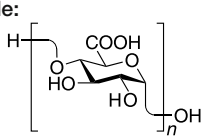
Note. The name ending in '-an' refers to the unsubstituted polysaccharide. Thus xylan occurs in nature in unacetylated and partially acetylated forms. Xylan designates the unacetylated material, and xylan acetate an acetylated derivative.

Well established names such as cellulose, starch, inulin, chitin, amylose and amylopectin are retained. 'Carrageenan' and 'laminaran' are now often used rather than the older names ending in '-in'.

2-Carb-39.5 Uronic Acid Derivatives

A polysaccharide (glycan) composed entirely of glycuronic acid residues is named by replacing '-ic acid' by '-an'. The generic name for this group of polysaccharides is 'glycuronan'.

Example:



(1→4)-α-D-galacturonan (pectin component)

Note. The term glycuronan is used instead of 'polyuronide'; the latter term is incorrect.

2-Carb-39.6 Amino Sugar Derivatives

A polysaccharide composed entirely of amino sugar residues is named by appropriate modification of the systematic amino sugar name.

Example:

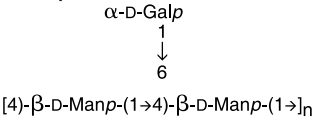
[4)-β-D-Glc*p*-NAc(1→]$_n$

(1→4)-2-acetamido-2-deoxy-β-D-glucan (chitin)

2-Carb-39.7 Polysaccharides Composed of more than one Kind of Residue

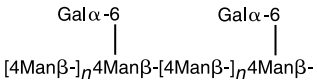
A heteropolysaccharide (heteroglycan) is a polymer containing two or more kinds of sugar (glycose) or modified sugar (e. g. aminodeoxyglycose or glycuronic acid) residue. When the polysaccharide has a principal chain ('backbone') composed of only one type of sugar residue, this residue should be cited last (as a 'glycan' term), and the other types of residue cited as 'glyco-' prefixes in alphabetical order. However, when no single type of sugar residue constitutes the principal chain, all sugar residues should be cited alphabetically as 'glyco-' prefixes, and the name should terminate with the suffix '-glycan'.

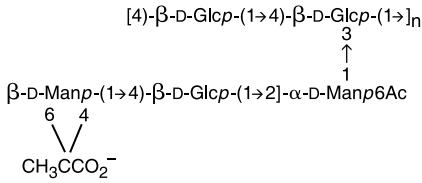
Examples:



A D-galacto-D-mannan (guaran)

Note. A less branched D-galacto-D-mannan could be shown in the short form as:



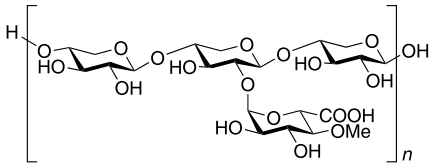


Xanthan

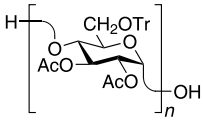
2-Carb-39.8 Substituted Residues

When substitution occurs in a polysaccharide (glycan), each type of substitutent is cited in the name at an appropriate position (in alphabetical order).

Examples:



(4-*O*-Methyl-α-D-glucurono)-D-xylan



2,3-Di-*O*-acetyl-6-*O*-tritylamylose

[4)-β-D-Glc*p*A-(1→3) -β-D-Glc*p*NAc-(1→]n

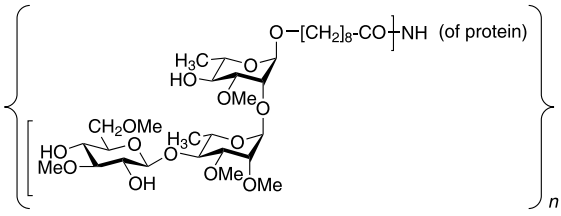
(2-Acetarnido-2-deoxy-D-gluco)-D-glucuronoglycan (hyaluronic acid or hyaluronan)

2-Carb-39.9 Glycoproteins, Glycopeptides and Peptidoglycans

Polymers containing covalently bound monosaccharide and amino-acid residues are termed glycoproteins, glycopeptides or peptidoglycans. It is not possible to give precise distinctions between these terms. In general, glycoproteins are conjugated proteins containing either oligosaccharide groups or polysaccharide groups having a fairly low relative molecular mass. Proteoglycans are proteins linked to polysaccharides of high molecular mass. Peptidoglycans consist of polysaccharide chains covalently linked to peptide chains. The nomenclature of these compounds is discussed in [11].

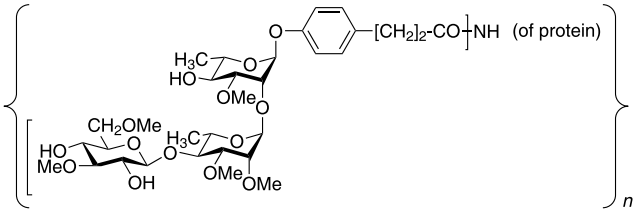
Synthetically produced or modified carbohydrate-protein conjugates are sometimes referred to as neoglycoproteins. The nomenclature for the carbohydrate-containing substituents in such structures is analogous to sequential oligosaccharide nomenclature (*2-Carb-37.2*)

Examples:



Poly-[(3,6-Di-*O*-methyl-β-D-glucopyranosyl)-(1→4)-(2,3-di-*O*-methyl-α-L-rhamnopyranosyl)-(1→2)-(3-*O*-methyl-α-L-rhamnopyranosyloxy)-(1-*O*→9)-nonanoyl-(1→N]-protein

{[β-D-Glc*p*3,6Me$_2$-(1→4) α-L-Rha*p*2,3Me$_2$-(1→2)-α-L-Rha*p*3Me-(1-*O*→9)-nonanoyl-(1→N]$_n$-protein}



Poly-[(3,6-Di-*O*-methyl-β-D-glucopyranosyl)-(1→4)-(2,3-di-*O*-methyl-α-L-rhamnopyranosyl)-(1→2)-(3-*O*-methyl-α-L-rhamnopyranosyloxy)-(1-*O*→4′)-(3-phenylpropanoyl)-(1→N^{Lys}]-protein

Appendix

Trivial Names for Carbohydrates and Derivatives with Their Systematic Equivalents and Symbols (Non-limiting List)

(a) Parent Monosaccharides

Allose (All)	*allo*-Hexose
Altrose (Alt)	*altro*-Hexose
Arabinose (Ara)	*arabino*-Pentose
Erythrose	*erythro*-Tetrose
Erythrulose	*glycero*-Tetrulose
Fructose (Fru)	*arabino*-Hex-2-ulose
D-Fucitol (D-Fuc-ol)	6-Deoxy-D-galactitol
L-Fucitol (L-Fuc-ol)	1-Deoxy-D-galactitol
Fucosamine (FucN)	2-Amino-2,6-dideoxygalactose
Fucose (Fuc)	6-Deoxygalactose
Galactosamine (GalN)	2-Amino-2-deoxygalactose
D-Galactosaminitol (GalN-ol)	2-Amino-2-deoxy-D-galactitol
Galactose (Gal)	*galacto*-Hexose
Glucosamine (GlcN)	2-Amino-2-deoxyglucose
Glucosaminitol (GlcN-ol)	2-Amino-2-deoxyglucitol
Glucose (Glc)	*gluco*-Hexose
Glyceraldehyde	2,3-Dihydroxypropanal
Glycerol (Gro)	Propane-1,2,3-triol
Glycerone (1,3-dihydroxyacetone)	1,3-Dihydroxypropanone
Gulose (Gul)	*gulo*-Hexose
Idose (Ido)	*ido*-Hexose
Lyxose (Lyx)	*lyxo*-Pentose
Mannosamine (ManN)	2-Amino-2-deoxymannose
Mannose (Man)	*manno*-Hexose
Psicose (Psi)	*ribo*-Hex-2-ulose

Quinovose (Qui)	6-Deoxyglucose
Quinovosamine	2-Amino-2,6-dideoxyglucose
Rhamnitol (Rha-ol)	1-Deoxymannitol
Rhamnosamine (RhaN)	2-Amino-2,6-dideoxymannose
Rhamnose (Rha)	6-Deoxymannose
Ribose (Rib)	*ribo*-Pentose
Ribulose (Rul)	*erythro*-Pent-2-ulose
Sorbose (Sor)	*xylo*-Hex-2-ulose
Tagatose (Tag)	*lyxo*-Hex-2-ulose
Talose (Tal)	*talo*-Hexose
Tartaric acid	Erythraric/Threaric acid
Threose	*threo*-Tetrose
Xylose (Xyl)	*xylc*-Pentose
Xylulose (Xul)	*threo*-Pent-2-ulose

(b) Common Trivial Names

Abequose (Abe)	3,6-Dideoxy-D-*xylo*-hexose
Amicetose	2,3,6-Trideoxy-D-*erythro*-hexose
Amylose	(1→4)-α-D-Glucopyranan
Apiose (Api)	3-*C*(Hydroxymethyl)-*glycero*-tetrose
Arcanose	2,6-Dideoxy-3-*C*-methyl-3-*O*-methyl-*xylo*-hexose
Ascarylose	3,6-Dideoxy-L-*arabino*-hexose
Ascorbic acid	L-*threo*-Hex-2-enono-1,4-lactone
Boivinose	2,6-Dideoxy-D-gulose
Cellobiose	β-D-Glucopyranosyl-(1→4)-D-glucose
Cellotriose	β-D-Glucopyranosyl-(1→4)-β-D-glucopyranosyl-(1→4)-D-glucose
Chacotnose	α-L-Rhamnopyranosyl-(1→2)-[α-L-rhamnopyranosyl-(1→4)]-D-glucose
Chalcose	4,6-Dideoxy-3-*O*-methyl-D-*xylo*-hexose
Cladinose	2,6-Dideoxy-3-*C*-methyl-3-*O*-methyl-L-*ribo*-hexose
Colitose	3,6-Dideoxy-L-*xylo*-hexose
Cymarose	6-Deoxy-3-*O*-methyl-*ribo*-hexose
2-Deoxyribose (dRib)	2-Deoxy-*erythro*-pentose
2-Deoxyglucose (2dGlc)	2-Deoxy-*arabino*-hexose
Diginose	2,6-Dideoxy-3-*O*-methyl-*lyxo*-hexose
Digitalose	6-Deoxy-3-*O*-methyl-D-galactose
Digitoxose	2,6-Dideoxy-D-*ribo*-hexose
Evalose	6-Deoxy-3-*C*-methyl-D-mannose
Evernitrose	2,3,6-Trideoxy-3-*C*-methyl-4-*O*-methyl-3-nitro-L-*arabino*-hexose
Gentianose	β-D-Fructofuranosyl β-D-gIucopyranosyl-(1→6)-α-D-glucopyranoside
Gentiobiose	β-D-Glucopyranosyl-(1→6)-D-glucose

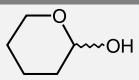
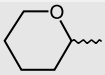
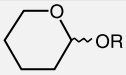
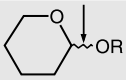
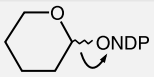
Hamamelose 2-*C*-(Hydroxymethyl)-D-ribose
Inulin (2→1)-β-D-Fructofuranan
Isolevoglucosenone 1,6-Anhydro-2,3-dideoxy-β-D-*glycero*-hex-2-enopyranos-4-ulose
Isomaltose α-D-Glucopyranosyl-(1→6)-D-glucose
Isomaltotriose α-D-Glucopyranosyl-(1→6)-α-D-glucopyranosyl-(1→6)-D-glucose
Isopanose α-D-Glucopyranosyl-(1→4)-[α-D-glucopyranosyl-(1→6)]D-glucose
Kojibiose α-D-Glucopyranosyl-(1→2)-D-glucose
Lactose (Lac) β-D-Galactopyranosyl-(1→4)-D-glucose
Lactosamine (LacN) β-D-Galactopyranosyl-(1→4)-D-glucosamine
Lactosediamine (LacdiN) 2-Amino-2-deoxy-β-D-galactopyranosyl-(1→4)-D-glucosamine
Laminarabiose β-D-Glucopyranosyl-(1→3)-D-glucose
Levoglucosan 1,6-Anhydro-β-D-glucopyranose
Levoglucosenone 1,6-Anhydro-3,4-dideoxy-β-D-*glycero*-hex-3-enopyranos-2-ulose
Maltose α-D-Glucopyranosyl-(1→4)-D-glucose
Manninotriose α-D-Galactopyranosyl-(1→6)-α-D-galactopyranosyl-(1→6)-D-glucose
Melezitose α-D-Glucopyranosyl-(1→3)-β-D-fructofuranosyl α-D-glucopyranoside
Melibiose α-D-Galactopyranosyl-(1→6)-D-glucose
Muramic acid (Mur) 2-Amino-3-*O*-[(*R*)-1-carboxyethyl]-2-deoxy-D-glucose
Mycarose 2,6-Dideoxy-3-*C*-methyl-L-*ribo*-hexose
Mycinose 6-Deoxy-2,3-di-*O*-methyl-D-allose
Neuraminic acid (Neu) 5-Amino-3,5-dideoxy-D-*glycero*-D-*galacto*-non-2-ulosonic acid
Nigerose α-D-Glucopyranosyl-(1→3)-D-glucose
Nojirimycin 5-Amino-5-deoxy-D-glucopyranose
Noviose 6-Deoxy-5-*C*-methyl-4-*O*-methyl-L-*lyxo*-hexose
Oleandrose 2,6-Dideoxy-3-*O*-methyl-L-*arabino*-hexose
Panose α-D-Glucopyranosyl-(1→6)-α-D-glucopyranosyl-(1→4)-D-glucose
Paratose 3,6-Dideoxy-D-*ribo*-hexose
Planteose α-D-Galactopyranosyl-(1→6)-β-D-fructofuranosyl α-D-glucopyranoside
Primeverose β-D-Xylopyranosyl-(1→6)-D-glucose
Raffinose β-D-Fructofuranosyl α-D-galactopyranosyl-(1→6)-α-D-glucopyranoside
Rhodinose 2,3,6-Trideoxy-L-*threo*-hexose
Rutinose α-L-Rharnnopyranosyl-(1→6)-D-glucose
Sarmentose 2,6-Dideoxy-3-*O*-methyl-D-*xylo*-hexose
Sedoheptulose D-*altro*-Hept-2-ulose
Sedoheptulosan 2,7-Anhydro-β-D-*altro*-hept-2-ulopyranose
Solatriose α-L-Rhamnopyranosyl-(1→2)-[β-D-glucopyranosyl-(1→3)]-D-galactose
Sophorose β-D-Glucopyranosyl-(1→2)-D-glucose

Stachyose	β-D-Fructofuranosyl α-D-galactopyranosyI-(1→6)-α-D-galactopyranosyl-(1→6)-α-D-glucopyranoside
Streptose	5-Deoxy-3-*C*-formyl-L-lyxose
Sucrose (saccharose)	β-D-Fructofuranosyl α-D-glucopyranoside
α,α-Trehalose	α-D-Glucopyranosyl α-D-glucopyranoside
Trehalosamine	2-Amino-2-deoxy-α-D-glucopyranosyl α-D-glucopyranoside
Turanose	α-D-GIucopyranosyl-(1→3)-D-fructose
Tyvelose (Tyv)	3,6-Dideoxy-D-*arabino*-hexose
Umbelliferose	β-D-Fructofuranosyl α-D-galactopyranosyl-(1→2)-α-D-galactopyranoside

(c) Trivial Names Formed by Modification of Non-standard Monosaccharide Parent Names

Acosamine	3-Amino-2,3,6-trideoxy-L-*xylo*-hexose
Bacillosamine	2,4-Diamino-2,4,6-trideoxy-D-glucose
Daunosamine	3-Amino-2,3,6-trideoxy-L-*lyxo*-hexose
Desosamine	3,4,6-Trideoxy-3-dimethylamino-D-*xylo*-hexose
Forosamine	2,3,4,6-Tetradeoxy-4-dimethylamino-D-*erythro*-hexose
Garosamine	3-Deoxy-4-*C*-methyl-3-methylamino-L-arabinose
Kanosamine	3-Amino-3-deoxy-D-glucose
Kansosamine	4,6-Dideoxy-3-*C*-methyl-2-D-methyl-L-mannose
Mycaminose	3,6-Dideoxy-3-dimethylamino-D-glucose
Mycosamine	3-Amino-3,6-dideoxy-D-mannose
Perosamine	4-Amino-4,6-dideoxy-D-mannose
Pneumosamine	2-Amino-2,6-dideoxy-D-talose
Purpurosamine C	2,6-Diamino-2,3,4,6-tetradeoxy-D-*erythro*-hexose
Rhodosamine	2,3,6-Trideoxy-3-dimethylamino-*lyxo*-hexose

Glossary of Glycose-Based Terms; Standard Forms [Common Biochemical Usage]

	Class	Amino sugar[a] {usually asacetamido (*N*-acetyl) derivative}	Uronic acid[b] *Uronate*	Ulosonic acid *Ulosonate*	The sialic acid family[b]
	Glycose	Aminodeoxy-glycose [Glycosamine]	Glycuronic acid *Glycuronate*	Glyculosonic acid *Glyculosonate* [Ketoglyconic acid]	Neuraminic acid Sialic acid
	Glycosyl	Aminodeoxy-glycosyl [Glycosaminyl]	Glycosyluronic acid *Glycosy-luronate* [Glycurono-syl]	Glyculosylonic acid *Glyculosylonate*	[Neuraminosyl][c] [Sialosyl][c]
	Glycoside	Aminodeoxy-glycoside [Glycosaminide]	Glycosiduronic acid *Gly-cosiduronate* [Glycuronide]	Glyculosidonic acid *Glyculosidonate*	[Neuraminoside] [Sialoside]
	Glycosidase (Glycoside hydrolase)	Aminodeoxy-glycosidase [Gly-cosaminidase]	Gly cosiduronase [Gly-curonidase]	Glyculosidonase	[Neuraminidase] [Sialidase]
	Glycosyl-trans-ferase	Aminodeoxy-glycosyltrans-ferase [Glycosaminyl-transferase]	Glycosylurona-tetransferas [Glycurono-syltrans-ferase]	Glyculosylonate-transferase	[Neuraminosyl-transferase] [Sialosyltrans-ferase]

[a]The biochemical usage is widely established in the literature. [b]The biochemical usage implies the parents 'glycuronose', 'sialose', and 'neuraminose'. [c]'Neuraminyl' and 'sialyl' have been used, but are likely to be interpreted as referring to acyl groups; the terms given are more consistent with the terms used for glycosides

References

1. IUPAC Commission on the Nomenclature of Organic Chemistty (CNOC) and IUPAC-IUB Commission on Biochemical Nomenclature (CBN) (1969) Tentative rules for carbohydrate nomenclature, Part I. Biochem J 125:673–695 (1971); Biochemistry 10:3983–4004 (1971); Biochim Biophys Acta 244:223–302 (1971); Eur J Biochem 21:455–477 (1971) and 25:4 (1972); J Biol Chem 247:613–635 (1972); ref.2, pp.127-148
2. International Union of Biochemistry and Molecular Biology (1992) Biochemical Nomenclature and Related Documents. Portland Press, London
3. IUPAC-IUB Joint Commission on Biochemical Nomenclature (JCBN), Conformational nomenclature for five- and six-membered ring forms of monosaccharides and their derivatives (Recommendations 1980). Eur J Biochem 111:295-298 (1980); Arch Biochem Biophys 207:469–472 (1981); Pure Appl Chem 53:1901–1905 (1981); ref. 2, pp. 158–161
4. IUPAC-IUB Joint Commission on Biochemical Nomenclature (JCBN), Nomenclature of branched-chain monosaccharides (Recommendations 1980), Eur J Biochem 119:5–8 (1981); corrections: Eur J Biochem 125:1 (1982); Pure Appl Chem 54:211–215 (1982); ref. 2, pp. 165–168
5. IUPAC-IUB Joint Commission on Biochemical Nomenclature (JCBN), Nomenclature of unsaturated monosaccharides (Recommendations 1980), Eur J Biochem 119:1–3 (1981); corrections: Eur J Biochem 125:1 (1982); Pure Appl Chem 54:207–210 (1982); ref. 2, pp. 162–164
6. IUPAC-IUB Joint Commission on Biochemical Nomenclature (JCBN), Abbreviated terminology of oligosaccharide chains (Recommendations 1980), J Biol Chem 257:3347–3351 (1982); Eur J Biochem 126:433–437 (1982); Pure Appl Chem 54:1517–1522 (1982); Arch Biochem Biophys 220:325–329 (1983); ref. 2, pp. 169–173
7. IUPAC-IUB Joint Commission on Biochemical Nomenclature (JCBN), Polysaccharide nomenclature (Recommendations 1980), Eur J Biochem 126:439–441 (1982); Pure Appl Chem 54:1523–1526 (1982); J Biol Chem 257:3352–3354 (1982); Arch Biochem Biophys 220:330–332 (1983); ref. 2, pp. 174–176
8. WAC-IUB Commission on Biochemical Nomenclature (CBN), Nomenclature of cyclitols (Recommendations 1973), Biochem J 153:23–31 (1976); Eur J Biochem 57:1–7 (1975); Pure Appl Chem 37:283–297 (1974); ref. 2, pp 149–155
9. Nomenclature Committee of IUB (NC-IUB), Numbering of atoms in myo-inositol (Recommendations 1988). Biochem J 258:1–2 (1989); Eur J Biochem 180:485–486 (1989); ref. 2, pp. 156–157
10. IUPAC-IUB Joint Commission on Biochemical Nomenclature (JCBN), Symbols for specifying the conformation of polysaccharide chains (Recommendations 1981). Eur J Biochem 131:5–7 (1983); Pure Appl Chem 55:1269–1272 (1983); ref. 2, pp. 177–179
11. IUPAC-IUB Joint Commission on Biochemical Nomenclature (JCBN), Nomenclature of glycoproteins, glycopeptides and peptidoglycans. Eur J Biochem 159:1–6 (1986); Glycoconjugute J 3:123–124 (1986); J Biol Chem 262:13–18 (1987); Pure Appl Chem 60:1389–1394 (1988); Royal Society of Chemistry Specialist Periodical Report, ‘Amino Acids and Peptides’, vol. 21, p. 329 (1990); ref. 2, pp. 84–89
12. IUPAC-IUB Joint Commission on Biochemical Nomenclature (JCBN), Nomenclature of glycolipids, in preparation
13. IUPAC Nomenclature of Organic Chemistry, Sections A, B, C, D, E. F and H, 1979 Edition. Pergamon Press, Oxford, U.K. Sections E and F are reprinted in ref. 2, pp. 1–18 and 19–26, respectively.
14. Guide to IUPAC Nomenclature of Organic Compounds, Recommendations 1993, Blackwell Scientific Publications, Oxford 1993
15. This text is largely based on an essay entitled ‘Development of Carbohydrate Nomenclature’ by D. Horton, included in ‘The Terminology of Biotechnology: A Multidisciplinary Problem’ (ed. K.L. Loening. Springer-Verlag, Berlin and Heidelberg, 1990)
16. E. Fischer (1890) Ber 23:2114
17. Rules of carbohydrate nomenclature (1948) Chem Eng News 26:1623
18. Rules of carbohydrate nomenclature (1952) J Chem Soc 5108 (1952); Chem Eng News, 31:1776
19. Rules of carbohydrate nomenclature (1963) J Org Chem 28:281
20. IUPAC Commission on Macromolecular Nomenclature, Nomenclature of regular single-strand

organic polymers (Recommendations 1975), Pure Appl Chem 48:373–385 (1976); 'Compendium of Macromolecular Nomenclature'. Blackwell Scientific Publications, Oxford, p. 91 (1991)
21. IUPAC-IUB Joint Commission on Biochemical Nomenclature (JCBN), The nomenclature of lipids (Recommendations 1976), Eur J Biochem 79:11–21 (1977); Hoppe-Seyfer's Z. Physiof. Chern. 358:617–631 (1977); Lipids 12:455–468 (1977); Mol Cell Biochem 17:157–171 (1977); Chem Phys Lipids 21:159–173 (1978); J Lipid Res 19:114–40728 (1978); Biochem J 171:21–35 (1978); ref. 2, pp. 180–191
22. Cremer D, Pople JA (1975) J Am Chem Soc 97:1354–1358; Boeyens JCA (1978) J Cryst Mol Struct 8:317–320; Koll P, John HG, Kopf J (1982) Liebigs Ann Chem 626–638
23. IUB Nomenclature Committee (1992) Enzyme Nomenclature. Academic Press, Orlando, Florida
24. IUPAC-IUB Commission on Biochemical Nomenclature (CBN), Nomenclature of phosphorus-containing compounds of biochemical importance (Recommendations 1976), Hoppe-Seylers Z. Physiol Chem 358:599–616 (1977); Eur J Biochem 79:1–9 (1977); Proc Natl Acud Sci USA 74:2222–2230 (1977); Biochem J 171:1–19 (1978) ref. 2, pp. 256–265
25. IUPAC Commission on Nomenclature of Organic Chemistry, Revised nomenclature for radicals, ions, radical ions and related species (Recommendations 1993). Pure Appl Chem 65:1357–1455 (1993)
26. IUPAC Nomenclature of Organic Chemistry, Section F, revised version in preparation.
27. IUPAC Commission on Nomenclature of Organic Chemistry, Revision of the extended Hantzsch-Widman system of nomenclature for heteromonocycles. Pure Appl Chem 55:409–416 (1983)

Index