Zbigniew J. Witczak
Roman Bielski *Editors*

Coupling and Decoupling of Diverse Molecular Units in Glycosciences

Springer

Coupling and Decoupling of Diverse Molecular
Units in Glycosciences

Zbigniew J. Witczak · Roman Bielski
Editors

Coupling and Decoupling of Diverse Molecular Units in Glycosciences

Springer

Editors
Zbigniew J. Witczak
Pharmaceutical Sciences Department
Wilkes University
Wilkes-Barre, PA
USA

Roman Bielski
Pharmaceutical Sciences Department,
  Nesbitt School of Pharmacy
Wilkes University
Wilkes-Barre, PA
USA

Foreword

Since both fields' inception, carbohydrate chemistry and glycobiology have enjoyed deep reciprocity, the biological activities of carbohydrates are critical for human health and have been harnessed in medicine for more than a century. Glycosylated natural products continue to play a central role as therapeutics for cancer and infectious disease, and microbial polysaccharides comprise some of the most effective vaccines and adjuvants. At the same time, carbohydrates offer an abundant natural source of highly functionalized, chiral starting materials that can be elaborated in the preparation of other bioactive materials. Thus, synthetic chemists have devoted effort to wielding sugar starting materials by selective transformations that retain their valued components. At the same time, the ability to chemically modify sugars has opened doors to structure–activity relationship studies of their biological modes of action. Work in recent decades has brought forth the ability to construct complex oligosaccharides and to conjugate them to carrier scaffolds, emulating glycoproteins and glycolipids found in nature. These methods have enabled construction of glycosylated biomolecules and their mimetics, reagents that have formed the crux of fundamental studies of biological mechanism and that are in development as therapeutic and diagnostic agents.

As underscored by the chapters in this book, chemical manipulation of glycosides has advanced in parallel with, while also being accelerated by, newly developed bioconjugation methods, such as next-generation thiol modifications and bioorthogonal click chemistries. Having advanced over the last two decades, these new transformations enable conjugation of highly functionalized molecules in aqueous environments and without need for protecting groups. Bioorthogonal click chemistries in particular have been a boon for carbohydrate chemists, as they are perfectly suited for conjugating sugars to peptides and other carriers and can also be deployed for in vivo transformation of glycoside prodrugs to active species.

These and other exciting developments are captured in the following pages, with topics as diverse as use of sugars as chiral auxiliaries, glycoconjugate synthesis and analysis, and vaccine optimization. We start with an overview by Brimble et al. of N-linked glycopeptide assembly using combinations of chemical and enzymatic methods. Inherent in such efforts is the deployment of methods for coupling sugar

building blocks via glycosidic linkages. One such method is presented in the chapter by Lim and Fairbanks, the use of DMC derivatives to activate the anomeric center for coupling reactions in water. In the contribution from Miura et al., we learn how synthetic carbohydrates reflecting minimal functional units of gly-cosaminoglycans can be assembled into multivalent glycopolymers that emulate the biological properties of their natural congeners. And synthetic carbohydrates can also be designed to inhibit carbohydrate processing enzymes, as reflected in Valeti and Sucheck's chapter on design of transition state analogs that block a *Mycobacterium tuberculosis* enzyme involved in cell wall biosynthesis.

Two contributions from Witczak and coworkers inform how chemoselective reactions can be harnessed for carbohydrate conjugation and library synthesis, one focusing on thiol exchange reactions and another on azide–alkyne cycloaddition—the canonical click chemistry. Rubio-Ruiz et al. report on another use of bioorthogonal chemistry—as a mechanism to unmask anticancer glycoside pro-drugs *in vivo*. Dias et al. summarized their work synthesizing glycosyl flavonoids, both natural and unnatural, as treatments for neurodegenerative disease. We also learn from Daskhan and coworkers that simple sugars can be manipulated by skeletal rearrangements to form new scaffolds for natural products synthesis and for non-natural molecules such as carbon-linked or "C-" disaccharides as well as structures with expanded ring systems. Beyond the goal of making bioactive sugar analogs, Patel and Andreana show that sugars, by virtue of their inherent chirality, can serve as auxiliaries for chiral induction during conversion of isocyanides to amino acids.

The impact that creative coupling-and-decoupling reactions can have in biomedical glycoscience is beautifully illustrated by four chapters in this collection. Three of these pertain glycoconjugate interactions with the immune system. Glaffig and Kunz summarize approaches to assembly of vaccine conjugates based on cancer-associated mucin glycopeptide structures. These synthetic materials hold promise as tumor vaccine components. Schoker et al. show how synthetic alpha-Gal-bearing conjugates can be used to map immunodominant epitopes from trypanosome cell surface structures, which has important implications for vaccine development. And Hayman et al. summarized their work conjugating innate immune cell-activating lipid antigens to glycopeptide vaccines in order to boost their efficacy.

Finally, while chemical modifications of carbohydrates and their conjugates populate most of these chapters, it should not be overlooked that knowing the most important structures to make is a critical prerequisite. Thus, methods for elucidating the structures of naturally occurring glycoconjugates are of paramount importance. The chapter by Gray and Flitsch summarizes how modern mass spectrometry and NMR methods combined with enzymatic digestions and lectin binding studies can provide high-resolution structure information that then informs one's choice of synthetic targets.

Collectively, the chapters in this book provide a broad view of the exciting and interdependent fields of carbohydrate chemistry and glycobiology, one that will appeal both to expert practitioners as well as those who are new to the area. Enjoy the read!

Carolyn Bertozzi

Preface

Connecting molecular units exhibiting specific physical, chemical, or biological properties is a pivotal area of chemical synthesis. For many years, only organic units were considered worthy of this noble chemistry. Not anymore. Constructs containing inorganic or biological units are more and more common. The number of papers on coupling various moieties has been growing rapidly for dozens of years. The introduction of click chemistry was a game changer. Today, besides the original click process—the Huisgen reaction forming cyclic triazole—we have several other good coupling reactions such as the addition of thiols to various olefins, acetylenes, and isocyanates, Diels-Alder click reactions, and the newest one sulfur(VI) fluoride exchange (SuFex) [1].

While we may consider our accomplishments in the area to be wonderful, we must humbly admit that nature performs coupling better than we do and that there is a tremendous room for improving the existing chemistries. Furthermore, there is an acute need for more and better decoupling methods.

The list of methods applicable to connecting molecular units is much longer than the list of disconnecting methods. This was one of the reasons that prompted us to look at possible coupling strategies applicable to situations when decoupling will be necessary later on. We published in 2013 a relevant review article "Strategies for Coupling Molecular Units if Subsequent Decoupling is Required" [2]. In the article, we argued that there was no reason not to employ click chemistry to connect units whenever possible even when disconnection is in the cards. We just must design our coupling strategy wisely. We proposed the use of sacrificial units. The concept was called CAD (coupling and decoupling). It is worth noting that carbohydrates and their derivatives have become one of the most often used molecular units employed in the CAD processes.

Recently, Blanco, Santoyo-Gonzales et al. [3] stated in a paper published in Advanced Synthesis and Catalysis:

> *The concept of "coupling-and-decoupling" (CAD) chemistry has been recently introduced by Bielski and Witczak as a strategy that aims at both the binding and subsequent disconnection of the target molecules, a desirable feature targeted release, the decoupling of a molecule from a solid support or a surface of interest after performing chemical*

trans-formations on it (e.g., solid phase synthesis), the modification of the surface of a
material or the quantification of the amount of compound bound to the surface or solid
support.

The interest in the topic prompted us to organize a symposium during the 2015 Pacifichem in Honolulu in December 2015. We hoped it would cover most important aspects of the CAD strategy. The title of the symposium was "*Strategies for Coupling and Decoupling Diverse Molecular Units in the Glycosciences*". It attracted several prominent speakers, had a relatively large attendance and was met with a lot of interest. The symposium led to the present book. Some of the chapters in this book are based on the presentations delivered at the symposium. Other contributions are also from the leading experts in the field of carbohydrate chemistry. Some of the chapters are literature reviews, and some describe recent experiments performed in the authors' laboratories. We believe that all articles are of very high standard and offer a novel perspective on the discussed and highly important subjects.

Since the symposium, a few wonderful papers in the area of CAD have been published. One paper by Santoyo-Gonzales et al. has been already mentioned. The authors report an exceptionally effective methodology for CAD employing vinyl sulfonates. The method is simple and yields are excellent. Another spectacular paper was published by Anslyn and coworkers [4] who developed a simple Meldrum's acid-based reactant. It can be easily coupled to amine and thiol. What is remarkable both reactants (primary amine and thiol) can be decoupled and recovered unchanged. The authors use the words click and declick to describe the process. Additionally, it is worth mentioning two relevant reviews that were published after our symposium [5–6].

The book chapters have been written by authors coming both from the industry and academia, representing very different educational, scientific and ethnic backgrounds and perspectives. While this is probably an asset of the book, it makes creating a logically satisfying table of content a very difficult task. We apologize for this but the sequence of chapters in the book is unavoidably rather accidental.

With a growing complexity of material science and biomedical disciplines, one can expect a dramatic increase in forming and disconnecting molecular units. Moreover, it is almost impossible to find undiscovered phenomena or applications in the main areas of various disciplines. Thus, researchers will more and more look for fertile research objects in the phenomena taking place at very exotic conditions or when two or more units (that may or may not be compatible) are connected. Furthermore, it can be expected that many novel drugs will consist of two or more components. Note that often a mixture of drugs (HIV, certain cancer therapies) is significantly more effective than a single API (active pharmaceutical ingredient). All these aspects require effective methods of coupling and decoupling molecular units.

We hope that the book you keep in your hands (or watch at the screen of your computer) is what you expected and that it fills an important void. In conclusion, we believe that the present book offers a thoughtful perspective on the important area of research and will steer new ideas and solutions.

Acknowledgements

We wish to thank Prof. Carolyn Bertozzi for finding time in her very busy schedule to write the forward and all the authors for their excellent contributions to this volume. We also wish to thank all reviewers of the chapters for their level of expertise and helpful efforts to improve the quality of the manuscripts.

 We dedicate this book to our wives Wanda and Barbara.

Wilkes-Barre, USA   Zbigniew J. Witczak
<div align="right">Roman Bielski</div>

References

1. Dong J, Krasnova L, Finn MG, et al (2014) Angew Chem Int 53: 9430–9448
2. Bielski R, Witczak ZJ (2013) Chem Rev 113: 2205–2243
3. Cruz CM, Ortega-Muñoz M, López-Jaramillo FJ et al (2016) Adv Synth Catal 358: 3394–3413
4. Diehl KL, Kolesnichenko IV, Robotham SA et al (2016) Nature Chem 8: 968–973
5. Pergolizzi G, Dedola S, Field RA (2017). Contemporary glycoconjugation chemistry, Carbohydr Chem 42: 1–46
6. Tiwari VK, Mishra BB, Mishra KB (2016). Cu-Catalyzed Click Reaction in Carbohydrate Chemistry Chem Rev 116: 3086–3240

Contents

Synthesis of N-Linked Glycopeptides Using Convergent Enzymatic Glycosylation Combined with SPPS

Renata Kowalczyk, Harveen Kaur, Antony J. Fairbanks and Margaret A. Brimble

Abstract Glycosylation of peptides and proteins has emerged as a promising strategy to improve the pharmacokinetic profile of peptide- and protein-based therapeutics. The synthesis of pure homogeneous N-linked glycopeptides and glycoproteins is a challenging task, and efficient routes to access them are in high demand. Endo-β-N-acetylglucosaminidise catalysed glycosylation of N-acetylglucosamine-tagged peptides, using activated oligosaccharide oxazolines as donors, has recently attracted attention due to the relative simplicity by which the process convergently affords glycoconjugates with complete control of stereo- and regioselectivity. Herein, a brief review of some examples of recent enzyme-mediated N-glycosylation used to synthesise glycopeptides with therapetic potential is provided.

1 Introduction

Protein glycosylation is the most complex and diverse form of post-translational modification leading to the formation of N-, O-, S- and C-glycosides, phosphoglycans and glypiated proteins (proteins that are covalently bonded to a glycosylphosphatidylinositol via their C-terminus) [1]. The most common glycan-protein linkages found in nature are formed using either the side chain amide nitrogen of asparagine (Asn) residues or the side chain of serine (Ser) and threonine (Thr) residues to afford N-linked or O-linked glycoproteins, respectively [2].

R. Kowalczyk · H. Kaur · M.A. Brimble (✉)
School of Chemical Sciences, The University of Auckland, 23 Symonds St,
Auckland, New Zealand
e-mail: m.brimble@auckland.ac.nz

A.J. Fairbanks · M.A. Brimble
Maurice Wilkins Centre for Molecular Biodiscovery, University of Auckland,
Private Bag 92019, Auckland 1010, New Zealand

A.J. Fairbanks
Department of Chemistry, University of Canterbury, Private Bag 4800,
Christchurch 8140, New Zealand

Examples of *C*-mannosylation via the indole C-2 carbon atom of tryptophan (Trp) and *S*-glycosylation using the thiol of cysteine (Cys) are rare and have also been described in the literature (Fig. 1) [3].

Glycosylation plays an important role in various biological processes including cell development [4], inflammation [5], cell–cell signalling, adhesion and immune responses [6]. The presence of sugar moieties affects glycoprotein tertiary structures [7, 8], facilitates folding [9] and improves proteolytic stability [9]. Altered glycosylation patterns affect the circulatory lifetime of glycoproteins [10] and are associated with numerous diseases including some cases of congenital disorders
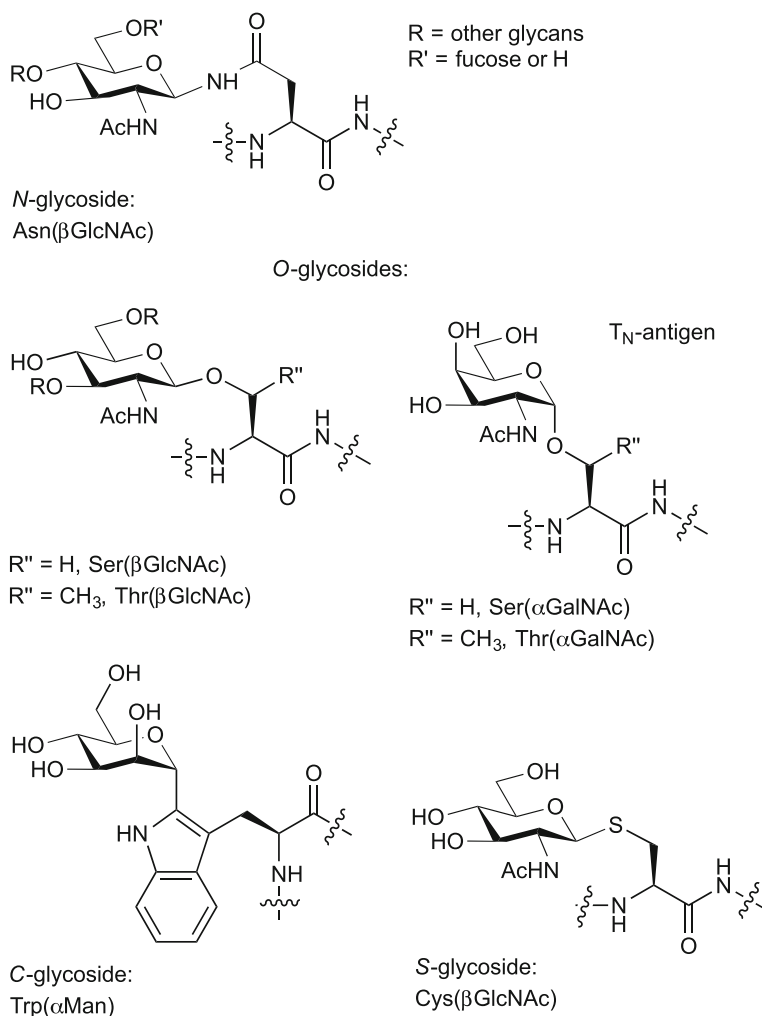

Fig. 1 Selected examples of *N*-, *O*-, *C*- and *S*-glycan-peptide linkages found in nature

[11, 12], leukocyte adhesion deficiency II [1], the aetiology of diabetes [13] and neurodegeneration [14, 15], cancer [16, 17] and Alzheimer's disease [18, 19].

Glycosylation of peptides and proteins with the intention of improving the pharmacokinetic profile of protein-based drugs has resulted in rapid expansion of the therapeutic peptide and protein market [20–26]. Increased proteolytic stability has been achieved by glycosylation of glucagon-like peptide-1 (GLP-1) [22], insulin [27], exendin-4 [28] and interferon-β [29]. The composition of protein-bound glycans can modulate the efficacy of protein therapeutics, for example, it has been demonstrated on numerous occasions that multiply sialylated versions of erythropoietin (EPO) [30] possess longer plasma half-lives as compared to their asialylated counterparts [10].

Robust and general strategies to prepare glycopeptides and glycoproteins in pure forms are highly sought after and have been investigated by many research groups [31–42]. N-Linked glycoproteins are prevalent and their glycans encompass diverse structures. They have gained significant scientific attention, not least due to their potential as therapeutic agents and thus are the main focus of this mini-review which describes the use of enzymatic glycosylation to access N-linked glycopeptides [24, 43].

2 Synthesis of Glycopeptides in Nature

The biosynthetic pathways of glycopeptide and glycoprotein synthesis in mammals are complex [2]. Briefly, the initial stage involves scavenging a simple monosaccharide unit, mostly glucose (Glc) but also galactose (Gal), mannose (Man) and glucosamine (GlcN) from the bloodstream by cells throughout the body using protein transporters (sodium-dependent co-transporters, SGLT and sodium independent facilitative transporters, GLUT) located in the plasma membrane of various tissues [2]. This is followed by intracellular de novo synthesis of additional sugar-based building blocks including fucose (Fuc), N-acetyl neuraminic acid (Neu, sialic acid) and N-acetylgalactosamine (GalNAc) by chemical processes including epimerisation, condensation and acetylation [2]. Subsequent phosphorylation of the monosaccharide units and pairing with corresponding nucleotides takes place in the cytosol to afford energy-rich nucleotide sugars required for further glycan synthesis [2]. Organelle-specific transporter proteins traffic nucleotide sugars from the cytosol into the endoplasmic reticulum (ER) and Golgi lumens where assembly of glycans of particular structures takes place. This complex process is mediated via the action of numerous transmembrane glycosyltransferases and glycosidases, the precise mechanisms of which are not yet fully understood [2, 24].

N-Linked glycosylation starts in ER where a dolichol phosphate oligosaccharide [$Glc_3Man_9(GlcNAc)_2$] is formed and then transferred *en block* onto an Asn residue on a nascently translated protein via the action of oligosaccharyltransferase (OST) [2]. The glycosylated asparagine residue is invariable within the conserved Asn-X-Thr/Ser peptide sequence (where X is any amino acid residue except

proline) of the unfolded protein. The oligosaccharide is then processed by the enzymes glucosidase I and II, affording a truncated GlcMan$_9$(GlcNAc)$_2$ glycoprotein interacting with calnexin and calreticulin and participating in the primary 'quality control' system distinguishing native from non-native protein conformations [44]. Subsequent removal of the Glc unit by glucosidase II releases Man$_9$(GlcNAc)$_2$-tagged glycoprotein from the chaperone, which if correctly folded can leave the ER. In case the protein exhibits non-native conformation, association with calnexin and calreticulin is renewed (via GlcMan$_9$(GlcNAc)$_2$ unit containing reattached Glc residue) and 'quality control' process is repeated until proper protein conformation is achieved. Glycoproteins permanently misfolded are eliminated from ER for degradation [44].

Removal of a terminal α(1-2)-linked mannose unit from either of the two arms of Man$_9$(GlcNAc)$_2$ subsequently takes place and is mediated by mannosidase I or II affording a Man$_8$(GlcNAc)$_2$-bound protein that is then transported to the *cis*-Golgi apparatus for further processing. Within the Golgi lumen, the common intermediate Man$_5$(GlcNAc)$_2$ is formed by the action of mannosidase IA and IB to remove α(1-2)-linked mannoses. Man$_5$(GlcNAc)$_2$ is then used to assemble, complex and hybrid subclasses of *N*-glycosides [2]. The partially processed glycans which were not trimmed to Man$_5$(GlcNAc)$_2$, or those which escape remodelling process from Man$_5$(GlcNAc)$_2$ to complex and hybrid *N*-glycoproteins, fall into high-mannose subclass of *N*-linked glycoproteins of the type Man$_{(5-9)}$GlcNAc$_2$ (Fig. 2) [45].

O-Linked glycosylation occurs in the Golgi apparatus and starts with the attachment of a monosaccharide unit (most often GalNAc) to a Ser or Thr residue present within the sequence of an already folded protein [2]. The subsequent formation of more complex oligosaccharides structures from the Ser/Thr(GalNAc) (Tn-antigen) core is then achieved via the sequential action of a variety of glycosyltransferases [2].

Protein glycosylation, unlike pure protein and oligonucleotide synthesis, is not template mediated and so depends on the activity and concentration of sugar substrates, the structural and conformational properties of glycosylation sites and the differential activities of numerous enzymes [41]. This means that glycoproteins are invariably produced as heterogeneous mixtures, termed glycoforms, where proteins possessing the same peptide chain vary in glycan structure [42, 47] and which may also vary in site occupancy. Access to well-defined homogeneous glycoproteins for subsequent structural and functional studies is, therefore, a challenging task.

Literature methods to prepare homogeneous glycopeptides and glycoproteins include the use of recombinant technology, fully synthetic techniques using chemical methods, enzymatic approaches and combinations of all the above. Each of these methodologies has its own advantages and limitations, and in depth discussion and progress to date has been summarised by a number of recent reviews [32, 33, 35, 36, 41, 42]. Our own endeavours towards the synthesis of glycosylated peptides and proteins with natural and non-natural glycan-peptide linkages began in

High-mannose:

Man-α(1,2)-Man $\diagdown$ α(1,6)
                        Man $\diagdown$ α(1,6)   Asn
Man-α(1,2)-Man $\diagup$ α(1,3)
                                        Man-β(1,4)-GlcNAc-β(1,4)-GlcNAc
                                            α(1,3)
Man-α(1,2)-Man-α(1,2) —— Man $\diagup$

Complex:

                                                                            Asn
NeuNAc-α(2,3)-Gal-β(1,4)-GlcNAc-β(1,2) —— Man $\diagdown$ α(1,6)
                                                    Man-β(1,4)-GlcNAc-β(1,4)-GlcNAc
NeuNAc-α(2,3)-Gal-β(1,4)-GlcNAc-β(1,2) —— Man $\diagup$ α(1,3)

                                                    β(1,4)   α(1,6)

                                                    GlcNAc   Fuc

Hybrid:

            Man-α(1,2)-Man $\diagdown$ α(1,6)   Asn
                                    Man $\diagdown$
            Man-α(1,2)-Man $\diagup$ α(1,3)
                                                    Man-β(1,4)-GlcNAc-β(1,4)-GlcNAc
NeuNAc-α(2,6)-Gal-β(1,4)-GlcNAc $\diagdown$ β(1,4)
                                    Man $\diagup$
NeuNAc-α(2,6)-Gal-β(1,4)-GlcNAc $\diagup$ β(1,2)

                                                    β(1,4)
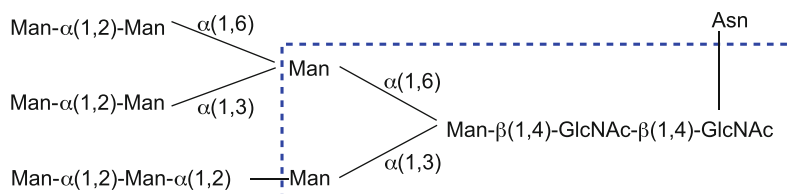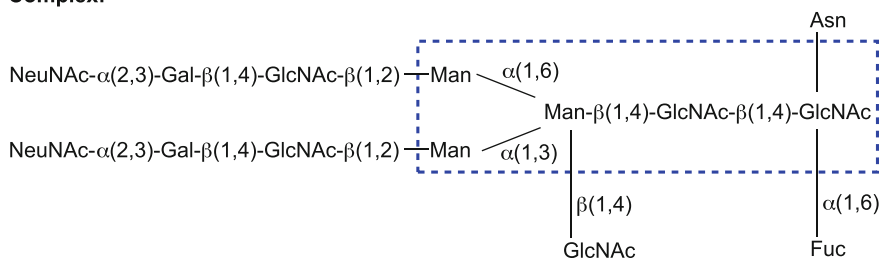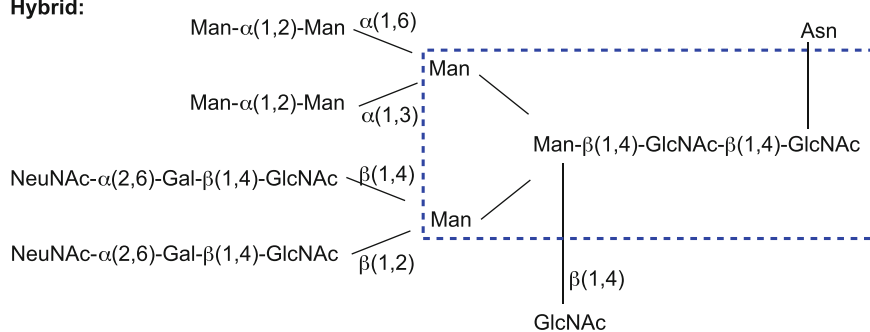
                                                    GlcNAc

Fig. 2 High-mannose, complex and hybrid subclasses of *N*-glycosides containing the core Man$_3$(GlcNAc)$_2$ pentasaccharide [46]

2008 and were mainly focused on using fully synthetic techniques [48–56]. Our current interest is now directed at combining synthetic techniques with chemoenzymatic methods to achieve the convergent synthesis of complex glycopeptides, [57, 58] and this is the primary focus of the current report. A comprehensive review on the topic, comprising elegant examples of chemoenzymatic approaches by other research groups, has been recently published by Wang and Amin [33].

3   Recombinant Approach to Access Glycopeptides and Glycoproteins

Recombinant methods using either fungal, plant or insect-based systems to produce *N*-glycoproteins seem promising but suffer from the limitation that heterogeneous products are invariably obtained, as well as the differences in glycosylation patterns between species, and batch-to-batch variability [24]. Expression systems based on bacteria are restricted to the synthesis of non-glycosylated proteins only (insulin, for example) owing to their inability to glycosylate due to the absence of glycosylation machinery [24]. Up to this point in time mammalian cell lines, for example, typically from Chinese hamster ovary (CHO), have been extensively used for the production of therapeutic glycoproteins due to their potential to produce certain human-like glycosylation patterns [24]. Other mammalian systems used for the synthesis of *N*-glycoproteins include baby hamster kidney (BHK-21) and murine myeloma (NS0 and Sp2/0) cells. Their use, however, is limited; BHK cell lines, similarly to CHO cells are not able to produce α(2,6)–linked terminal sialic acids present in human glycans, and immunogenity concerns are associated with the use of murine myeloma cells [24]. Nevertheless, recombinant methods using mammalian cell lines are routinely used to produce marketed therapeutic monoclonal antibodies [59]. Adalimumab (Humira®), the worlds best selling drug in 2015 [60] used for the treatment of rheumatoid arthritis [61], is expressed in CHO cell lines; Golimumab (Simponi®), used as an immunosuppressant [62], is expressed in Sp2/0 cells [59].

Human cell lines derived from embryonic kidney (HEK293), embryonic retinoblasts (PER.C6) or hybrid HKB11 cell lines composed of embryonic kidney cells (293S) and modified Burkitt's lymphoma cells (2B8) are attractive but expensive alternatives to CHO cells [63]. Significant scientific focus has therefore been directed into engineering effective and more straightforward yeast-based systems [64, 65]. It was found that 'humanized' *Pichia pastoris*-derived cell lines can be used to express homogeneous human *N*-linked glycoproteins bearing truncated complex [(GlcNAc)$_2$Man$_3$(GlcNAc)$_2$] units [65] and full-length complex sialylated glycans [64]. These studies have opened up further possibilities to access therapeutic glycoproteins via recombinant techniques using alternative yeast-derived expression systems [64, 65]. However, complete control of glycosylation, i.e. in order to obtain strictly homogeneous glycoproteins via recombinant methods is still challenging [66].

4   The Use of Chemical Synthesis to Access Homogeneous Glycopeptides and Glycoproteins

The use of synthetic techniques to counter challenges faced in the preparation of homogeneous glycopeptides and glycoproteins is under investigation by several research laboratories worldwide [37–40, 56, 67–71]. Solid phase peptide synthesis technique (SPPS) is the method of choice for the preparation of glycopeptides [32]. Despite the extensive development that SPPS has undergone [72] since its first discovery [73], it is still limited to the preparation of up to 30- to 50-residue long glycopeptides that carry relatively small oligosaccharide units. However, the combination of SPPS and ligation techniques, particularly native chemical ligation (NCL) [74] or expressed protein ligation (EPL) [75], has enabled synthetic access to more complex structures including large glycoproteins [35, 36]. The synthesis of 40- and 80 amino acid MUC1 glycoproteins bearing eight GalNAc units at corresponding Thr11 and Thr19 of tandem repeats was accomplished using a combination of 9-fluorenylmethoxycarbonyl (Fmoc) SPPS and a serine/threonine ligation technique [76–78]. The first total synthesis of glycocin F, an antimicrobial 43 amino acid glycopeptide with two βGlcNAc moieties at Ser18 and Cys43 was successfully undertaken by Brimble et al. [56]. The synthetic protocol involved initial Fmoc SPPS of three glycocin F fragments incorporating *O*- and *S*-linked GlcNAc unit using either Fmoc-Ser[GlcNAc(OAc)$_3$] or Fmoc-Cys[GlcNAc(OAc)$_3$] building blocks, respectively, accessed via total synthesis. Native chemical ligation [74] was then used to join these fragments which was followed by oxidative folding to effect the desired *C*-amidated glycocin F [56]. Many other examples of the synthesis of larger *N*-glycoproteins using ligation techniques have been reported in the literature [35, 36]. Notable examples include the synthesis of 166 amino acid interferon-β-1a [79], the 72 amino acid glycosylated analogue of interleukin-8 which was used in folding studies [80], the hydrophobic glycoprotein saposin C (80 amino acids) [81], and the 124 amino acid bovine ribonuclease (RNAse) C accessed via semisynthetic methods (EPL and NCL) [82] or total synthesis and NCL [83].

In general two approaches may be used to generate a linkage between an oligosaccharide and a peptide chain. The so-called 'linear approach' involves initial preparation of a glycosylated amino acid building block, which is then incorporated into a growing resin-bound peptide chain that is then typically extended using SPPS [32, 38, 84]. A major limitation of this technique is the significant effort required to prepare suitably protected carbohydrate-bearing building blocks in sufficient quantities for the subsequent coupling steps. Furthermore, the attachment of complex protected oligosaccharides to amino acid residues generates significant steric hindrance, which may diminish the effectiveness of the peptide coupling steps during SPPS, leading to by-product formation. Therefore, only short- to medium-sized glycopeptides, carrying relatively small (and typically *O*-linked) oligosaccharide units, can be accessed via the linear approach [33, 38, 42]. Nonetheless, this linear approach has been employed for the synthesis of

antitumour vaccine candidates based on mucin glycopeptide antigens [68, 85, 86], and for the synthesis of mannosylated peptides as components for synthetic vaccines [48]. Other literature examples that have adopted the linear strategy include the synthesis of fluorescent glycopeptides as biological probes [87, 88], and the synthesis of analogues of antifreeze glycoproteins [53] in addition to others reviewed elsewhere [84]. Kajihara et al. have employed the linear strategy to synthesise N-linked glycopeptides including a 79–85 fragment of EPO (ALLVNSS) bearing a complex biantennary sialyloligosaccharide [89], an EPO (85–95) fragment with two different glycans (asialo- and sialyloligosaccharides) attached [90], ligation partners for subsequent NCL to construct the full-length EPO mutants bearing one, two or three biantennary sialyloligosaccharides [91, 92] and a 38 amino acid cytotoxic T-lymphocyte-associated protein-4 (CTLA-4) fragment 113–150 with two complex-type undecadisialyloligosaccharides [93]. The glycopeptide thioesters bearing N-linked biantennary complex-type nonasaccharide unit required for subsequent ligation (EPL and NCL) to afford RNAse C were synthesised using the linear strategy by the group of Unverzagt [82, 83].

An alternative and convergent strategy involves the initial synthesis of a peptide chain followed by the direct attachment of a carbohydrate unit either on-resin or in-solution [32, 38, 40]. This technique is widely applicable for the preparation of N-linked glycopeptides where glycosylamines are attached to a pre-assembled peptide via the side chain carboxylic acid of embedded aspartic acid (Asp) residues using the Lansbury aspartylation [94, 95]. A notable example that employed this convergent strategy to access an N-linked glycoprotein was reported by Danishefsky et al. [69, 96, 97]. It involved the total synthesis of the 166 amino acid fully glycosylated homogeneous erythropoietin, containing N-linked branched dodecasaccharides at Asn24, Asn38 and Asn83, and an O-linked glycan at Ser126 [69, 96, 97]. A combination of Fmoc SPPS, NCL [74], O-mercaptoaryl ester rearrangement [96] and metal-free desulfurisation [98] was used to deliver the synthetic target. The highly complex N-linked oligosaccharides were prepared by total synthesis and introduced using a convergent selective amidation of the corresponding aspartic acid residue using a 'one flask' aspartylation approach [94, 95, 99, 100]. Further examples of the synthesis of glycopeptides have also been reported [32, 33, 36, 38, 42].

5 Enzymatic Approach to Access Homogeneous Glycopeptides and Glycoproteins

Enzymatic approaches to the glycosylation of peptides and proteins are increasing in popularity due to their simplicity, and excellent stereo- and regiochemical control [101]. This technique may employ a variety of enzymes including glycosidases, glycosyltransferases and glycosynthases to generate desired oligosaccharide structures and sugar-peptide linkages [101]. Glycosidases are responsible for

glycosidic bond hydrolysis and catalyse the breakdown of either terminal sugar units, from their non-reducing end, or internal glycosidic bonds (exo- and endo-glycosidases, respectively) [101]. Glycosyltransferases catalyse the formation of specific glycosidic linkages by transferring monosaccharide units from glycosyl donor substrates to corresponding acceptors [101]. Mutant glycosidases (commonly referred to known as glycosynthases) [102–104] may be used for glycopeptide and glycan synthesis. Glycosynthases have the ability to transfer activated oligosaccharides onto acceptors, and unlike glycosidases, possess little or no hydrolytic activity [101].

Endoglycosidases acting on glycan chains of glycoproteins can be further divided into two classes, those which hydrolyse the core region of *N*-linked oligosaccharides embedded within the glycoprotein chain, and those which recognise *O*-linkages between the sugar and the protein [105]; endo-β-*N*-acetylglucosaminidase (Endo-β-GlcNAc-ase, ENGase, endohexosaminidase) and endo-β-*N*-acetylgalactosaminidase are representative examples from each class of endoglycosidase, respectively [105]. To date, sequence analysis of the most synthetically useful ENGase enzymes employed for the synthesis of glycoproteins has lead to their classification in the carbohydrate-active enzymes (CAZy) database [106], as either members of family 18 or family 85 of the glycoside hydrolases (GH18 or GH85) [107]. The family GH18 enzymes are mostly derived from bacteria and fungi, while GH85 enzymes can be found in organisms ranging from bacteria to mammals [107].

In addition to their hydrolytic activity, endoglycosidases can also effectively transfer oligosaccharides onto corresponding hydroxyl-containing substrates by transglycosylation, or glycosylation. This dual capability makes endoglycosidases, especially the ENGases, valuable tools for the convergent synthesis of oligosaccharides and glycoconjugates [105, 108].

6 Wild-Type Enzymes and Peptide and Protein Glycosylation

A number of Endo-β-GlcNAc-ases have attracted scientific attention [107] due to their synthetic potential, and examples from family GH85 include Endo M (from *Mucor hiemalis*) [109–120], Endo A (from *Arthrobacter protophormiae*) [117, 121–128], Endo D (from *Streptococcus pneumoniae*) [129, 130], Endo OM (from *Ogataea minuta*) [107] and Endo-BH (from *Bacillus halodurans*) [131]. Selected examples of ENGases from family GH18 are Endo H (from *Streptomyces griseus*) [132], Endo S (from *Streptococcus pyogenes*) [133, 134] and Endo F1, F2 and F3 (from *Flavobacterium meningosepticum*) [135–137]. All ENGases cleave the *N,N*′-diacetylchitobiose unit [GlcNAcβ(1-4)GlcNAc], a common motif present within Asn-linked high-mannose, complex and hybrid *N*-glycans [105]. Enzyme hydrolysis is substrate specific; Endo H, Endo F1 and Endo A only act on
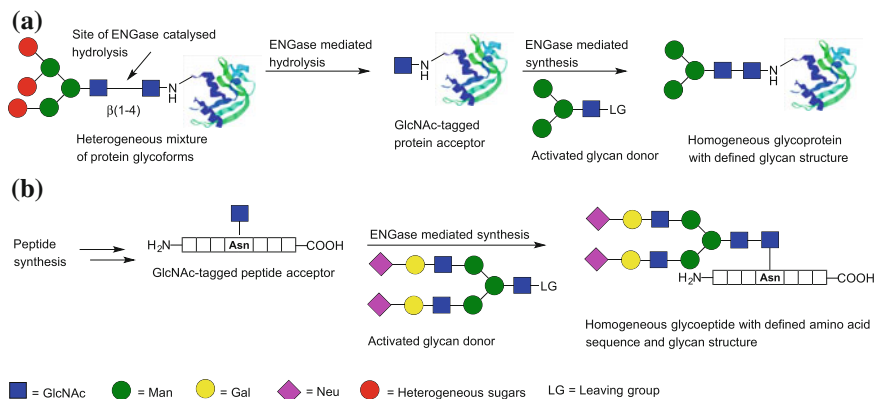
Scheme 1 **a** Glycoprotein remodelling approach and **b** convergent chemoenzymatic synthesis of glycopeptides using ENGases [43]

high-mannose and hybrid *N*-glycans, while Endo F3 also recognises bi- and tri-antennary complex glycans [43].

The synthetic ability of ENGases can, therefore, be used for the preparation of homogeneous glycoproteins in an enzyme-mediated glycoprotein remodelling approach (Scheme 1a) [33, 34, 43, 105, 138]. This approach involves initial enzyme-mediated hydrolysis of the β(1–4) bond of [GlcNAcβ(1-4)GlcNAc] unit, affording a GlcNAc-tagged protein. Subsequent reattachment of another glycan to this GlcNAc acceptor using the transglycosylation activity of an ENGase then gives a glycoprotein with a desired glycan structure [43, 105].

The protein remodelling approach can be extended to the convergent, chemoen-zymatic synthesis of glycopeptides. Herein, the synthesis of a peptide chain incor-porating a GlcNAc moiety as a sugar acceptor is performed first, and is then followed by enzymatic attachment of *N*-glycans using ENGases (Scheme 1b) [43]. However, the relatively low yields of products obtained during the process and hydrolytic activity of endoglycosidases often diminish their synthetic potential [43, 139].

7 Strategies to Improve Enzymatic Glycosylation

Approaches to effect improvements in ENGase catalysed glycosylation processes include the use of *N*-glycan oxazolines as activated sugar donors in combination with mutant variants of ENGases with altered activity towards product hydrolysis [43, 140–150].

The successful use of an oxazoline-activated disaccharide was first reported in 2001 by Shoda et al. [151]. The Manβ(1-4)GlcNAc-oxazoline (**1**) was successfully used as a glycosyl donor in an Endo M [109]- and Endo A [121]-mediated glycosylation reaction using GlcNAcOpNP (**2**) as the acceptor, affording the

Scheme 2 Oxazoline use for ENGase mediated glycosylation by Shoda et al. [43, 151]

3 : Endo A : 55% and Endo M : 44%
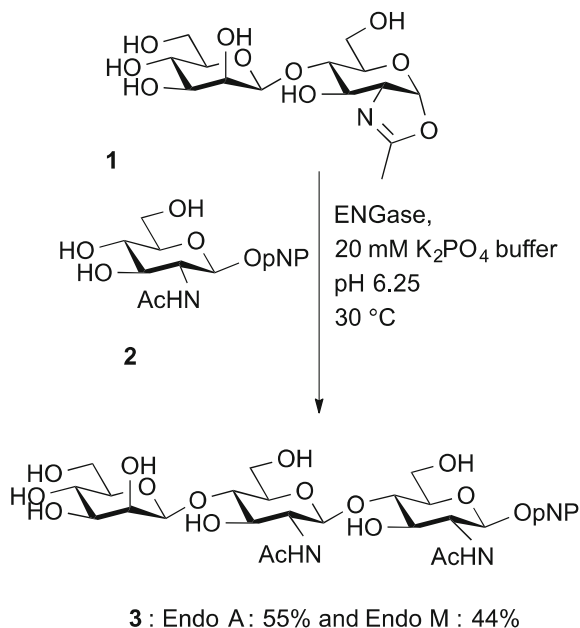
corresponding trisaccharide **3** (Scheme 2) [151]. The activity of oxazolines as donors for glycosylation processes catalysed by ENGases is related to the structural and functional similarities they share with oxazolinium ions, which are high energy intermediates in the enzymatic hydrolysis [151].

The use of sugar oxazolines as activated sugar donors in ENGase catalysed glycosylation has attracted the attention of many research laboratories worldwide [130, 133, 139–143, 152–165]. The power of ENGase mediated glycosylation using sugar oxazolines was recently demonstrated by Fairbanks et al. [166] who reported the first synthesis of phosphorylated glycoprotein bearing phosphorylated pentasaccharide unit attached via native *N*-linkage using the protein remodelling approach. The tetrasaccharide oxazoline in which two terminal mannose residues contained phosphate at the 6 position was accessed via chemical synthesis and attached to GlcNAc-tagged RNase B using Endo A [117, 121–128] glycosylation activity [166].

The extensive research on the utility of saccharide oxazolines for ENGase facilitated glycosylation of peptides and proteins [130, 133, 139–143, 152–165] revealed, however, that only smaller natural sugar oxazolines (di- and tri-saccharides) proved to be effective in glycosylation reactions using wild-type enzymes [139]. When larger, natural *N*-oligosaccharide oxazolines were used, significant reductions in glycosylation yields were observed due to competitive hydrolysis [139].

One potential solution to this problem might be the use of structurally modified sugar oxazolines. In principle, these highly activated glycan donors should be

readily processed by ENGases to afford glycosylation products that may not be substrates for hydrolytically active enzymes due to structural changes [140, 142].

Fairbanks et al. [140] showed that chemically synthesised di-, tri-, tetra- and hexasaccharide-oxazolines derived from the core sections of N-linked high-mannose glycans, containing a glucose moiety in place of a central mannose unit [Manβ(1-4)GlcNAc to Glcβ(1-4)GlcNAc] were substrates for Endo M [109] and Endo A [121], but not Endo H [132], and could be used to effect irreversible glycosylation. However, this approach only produces non-natural glycan structures, which may be considered a limiting factor.

Another strategy to curtail undesired product hydrolysis, and hence improve the yield of the glycosylation step, involves using mutant enzymes [43, 139]. Site-directed mutagenesis of wild-type enzymes is used to generate the requisite mutants [43, 139]. The idea originated from the use of glycosynthases [102–104] used for oligosaccharide synthesis, and was further developed to afford mutated versions of various ENGases, including Endo A [153, 167], Endo M [161, 168], Endo D [169] and Endo S [170, 171]. These mutated enzymes allowed the synthesis of glycoproteins [160, 161, 164] containing natural N-linked oligosaccharides. An important example is the commercially available N175Q mutant of Endo M, derived from family GH85 developed in the laboratories of Wang and Yamamoto [168]. Exchanging Asn175 for Gln in Endo M proved superior; the mutant exhibited greater glycosylation activity with significantly reduced hydrolytic activity as compared to wild-type Endo M and other mutants investigated [168]. Endo M N175Q became a valuable tool that may be used to access homogeneous N-liked glycopeptides and glycoproteins carrying natural high-mannose- and biantennary complex-type oligostructures [168].

The enzymatic remodelling of immunoglobulin G (IgG) [133, 170] using Endo S [133, 134] further expanded the synthetic potential of ENGases to access homogeneous antibodies (Abs) bearing well-defined sugar structures. Monoclonal antibodies (mAbs) are an important class of N-linked glycoprotein therapeutic which are produced using recombinant techniques as a mixture of multiple glycoforms of variable abundance and complexity depending on the expression system and cell line used [172]. The glycosylation pattern in the Fc region of mAbs is especially important and affects antibodies functions on immune cells via interaction with FcγR receptors [24]. The presence of biantennary N-linked oligosaccharide units with two terminal α2,6-linked sialic acids at the two Asn297 Fc glycosylation sites enhances the activity of immunoglobulin G against cancer and infectious and inflammatory diseases [173]. Straightforward access to pure glycoforms of mAbs is, therefore, the key to modulate their clinical effects and develop improved antibody-based therapeutics. Wong et al. [173] recently reported the synthesis of homogeneous Rituximab IgG1 (used for the treatment of rheumatoid arthritis and cancer) [174] using an enzymatic remodelling approach where the initial formation of a mono-GlcNAc-tagged antibody, achieved using Endo S [133, 134] and a fucosidase from *Bacteroides fragilis*, was followed by ligation of the well-defined synthetic glycan oxazolines using an Endo S D233Q mutant [170]. The synthetic utility of an enzymatic approach using mutated enzymes has been

also demonstrated by Davis et al. [171], who recently reported the synthesis of a homogeneous form of sialylated mAb Herceptin (Trastuzumab) [175], using a the same Endo S D233Q [170] and optimised reaction conditions (enzyme loading and oxazoline concentration). In addition to well-defined natural glycans, modified sugar oxazolines with handles or tags (such as azides or alkynes) incorporated via amidation of non-reducing terminal sialic acids of a decasaccharide unit were also incorporated onto GlcNAc-tagged Herceptin using the optimised protocol [171].

To broaden the scope and potential applications of enzymatic glycosylation, studies on the use of structurally modified GlcNAc or alternative sugar acceptors for ENGase mediated glycosylation have been undertaken [115, 120, 125, 135, 176, 177]. Fairbanks et al. [177] have recently reported the tolerance of various ENGases to transfer N-glycan oxazolines **4** and **5** to a structurally altered Asn (GlcNAc) acceptor in which the hydroxyl group of the glycan unit was protected with a benzyl ether at C-3 (**6**), C-4 (**7**) or C-6 (**8**) (Fig. 3). The OH-3 fucosylated Asn(GlcNAc) acceptor (**9**) was also tested but none of the enzymes studied (WT Endo M, N175Q Endo M, Endo A, Endo D) were able to effect this glycosylation [177]. The study revealed subtle structural preferences of each enzyme towards sugar acceptors, a factor which needs to be taken into consideration when choosing reaction partners [177].

8 Access to *N*-Linked Oligosaccharides

Despite the availability of various synthetic techniques, the synthesis of glycopeptides is still challenging [37, 38, 40, 67–71]. This is mainly due to the limited access to the oligosaccharide components that are generally obtained via multistep syntheses in specialised laboratories [35]. To date, although remarkable progress in carbohydrate synthesis has been made, reliable and general routes to prepare complex oligosaccharides are still needed [178]. Nevertheless, total synthesis gives access to a wide range of sugar constructs, either with natural or non-natural linkages.

A recent report by Shoda et al. [179] revealed a new and very convenient method to access GlcNAc-terminating oligosaccharide oxazolines. These oxazolines can be prepared from the corresponding 2-acetamido-2-deoxy reducing sugars (**10**) in water by activation using 2-chloro-1,3-dimethylimidazolinium chloride (DMC) as the dehydrating reagent to afford corresponding activated sugar donors (**11**) (Scheme 3) [179].

Although labour-intensive, approaches to access full-length *N*-oligosaccharides have been developed [43]. A general retrosynthetic overview for the synthesis of tetrasaccharide *N*-oxazolines is shown in Scheme 4 [43]. Due to the susceptibility of glycosyl oxazolines to acid and/or hydrogenation conditions, oxazoline formation must take place before the final base-catalysed removal of protecting groups is effected, or alternatively, the Shoda [179] approach described above can be used to synthesise the oxazoline in the last step. Successive glycan disconnections at C-6, and C-3 of the inner mannose unit allow for the installation of non-symmetrical
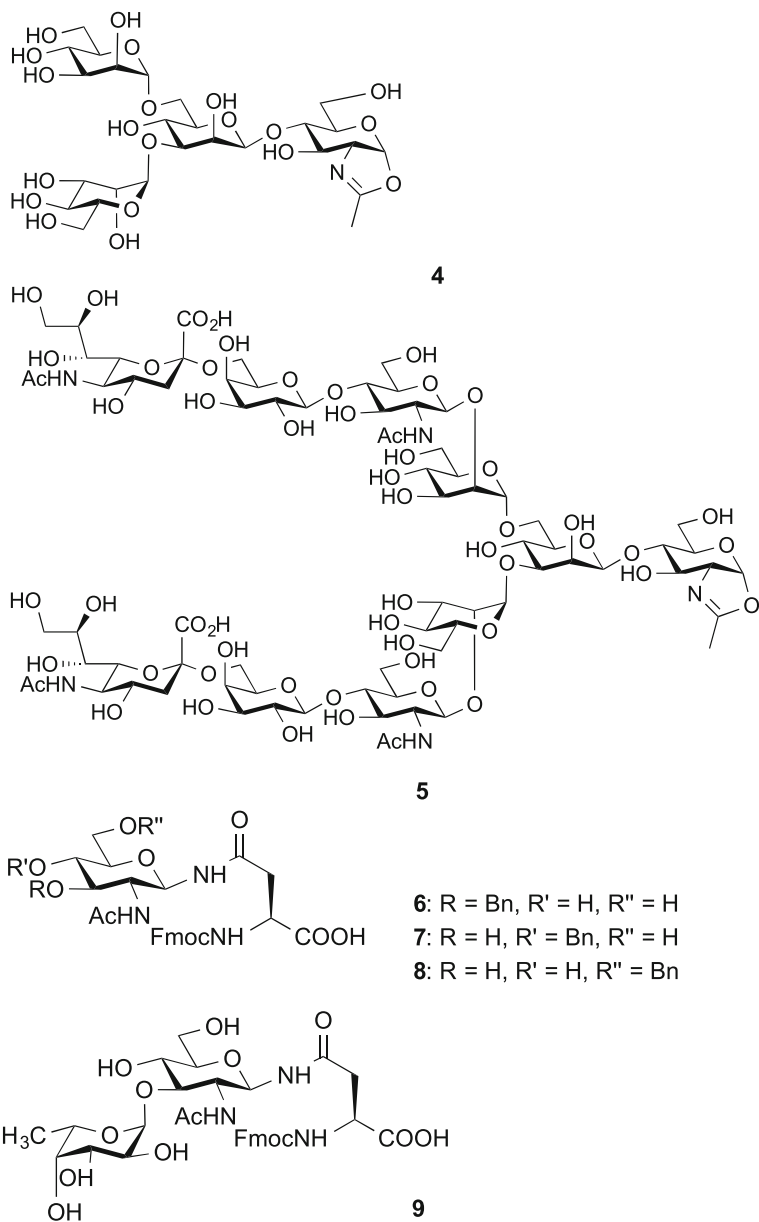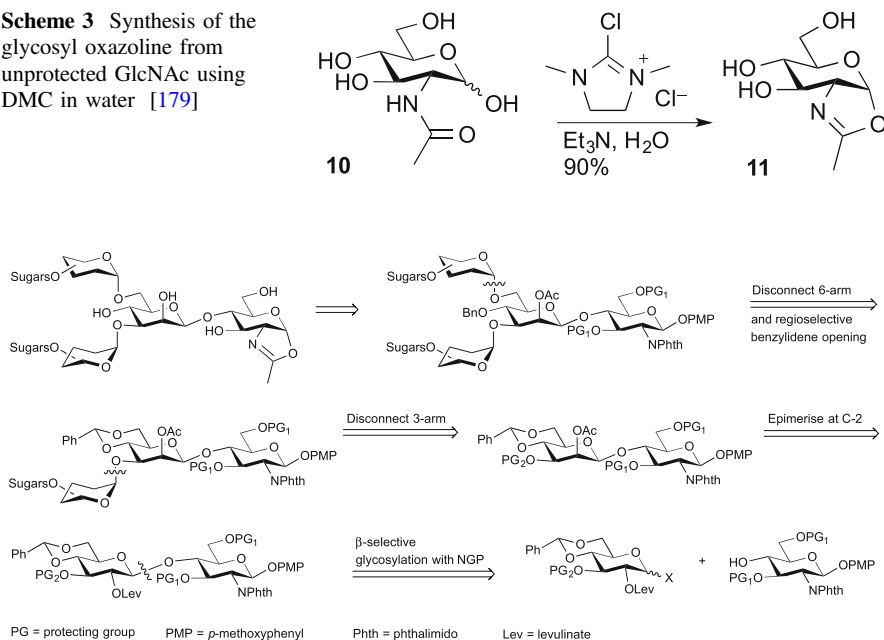
Fig. 3 Tetrasaccharide- and decasaccharide-oxazolines **4** and **5**, respectively, and modified glycosyl acceptors **6–9** targeted during the study [177]

Scheme 3 Synthesis of the glycosyl oxazoline from unprotected GlcNAc using DMC in water [179]

Scheme 4 Generalised retrosynthetic route to N-glycan oxazolines [43]

glycans, which can be achieved using 4,6-benzylidene acetal protection. Subsequent formation of a challenging β-mannosidic linkage takes place by inversion of configuration at C-2 of the selectively synthesised β-glucoside accessed using the neighbouring group participation (NGP) approach. The C-2 hydroxyl of the *gluco* donor is protected as a levulinate (Lev) ester, which can be removed selectively allowing installation of trifluoromethanesulfonate (triflate) leaving group that is subsequently displaced by acetate affording desired β-mannoside [43, 180].

Selected examples of oligosaccharide oxazolines accessed using these strategies are depicted in Fig. 4 [43].

The isolation of large oligosaccharides from natural sources may conveniently bypass laborious total synthesis routes. Some complex Asn-linked N-glycans, such as a sialic acid terminated complex biantennary unit [(NeuAcGalGlcNAcMan)$_2$Man (GlcNAc)$_2$] and a high-mannose glycan [(Man$_9$(GlcNAc)$_2$)] can be obtained in significant quantities from either egg yolks [181] or soy bean flour [182–184], respectively. Isolated and purified oligosaccharides can be chemically modified to afford suitably protected or activated building blocks for further incorporation into peptide or protein chains. It has also been shown that the biantennary glycan can be further modified using branch-specific exoglycosidases to access a broad variety of Asn-linked oligosaccharides, thus providing facile access to complex sugar structures [90].
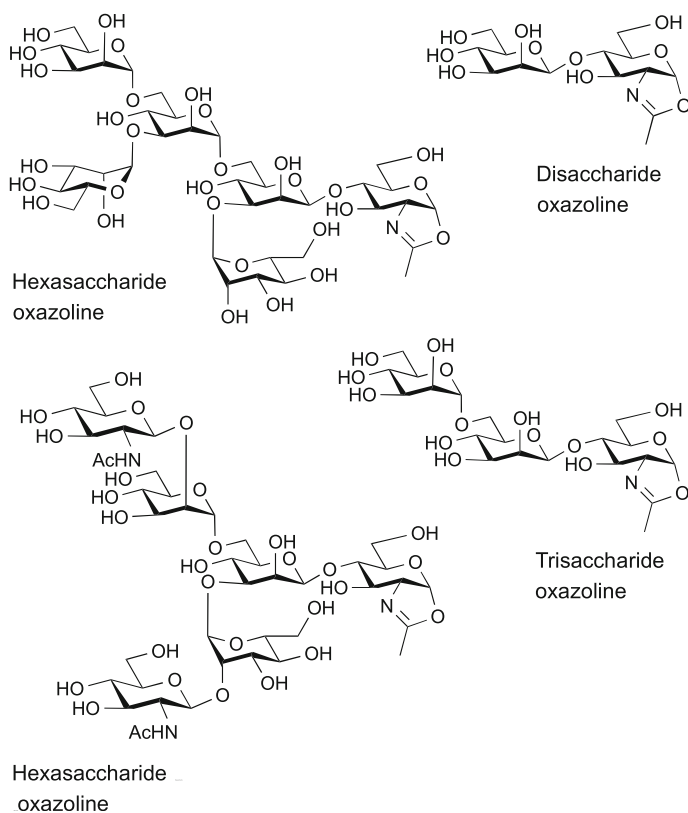
Fig. 4 Selected oligosaccharide oxazolines synthesised [43]

Kajihara et al. [92] reported the chemical synthesis of a mutated EPO variant, with alanine (Ala) residues replacing native glutamic acid (Glu) and glutamine (Gln) residues at position 21 and 78, respectively, using NCL technique [74]. The required asparaginyl sialyloligosaccharide, namely Asn[(NeuAcGalGlcNAcMan)$_2$Man(GlcNAc)$_2$]-(Asn83) was initially isolated from egg yolks. It was then suitably modified with phenacyl (Pac) protecting groups (acid-labile sialic acid residues) and the $N\alpha$-amino group of Asn was masked with a *tert*-butyloxycarbonyl (Boc) protecting group to allow the synthesis of sialylglycopeptide α-thioesters using Boc SPPS. Subsequent ligation of the corresponding peptide fragments using NCL [74] afforded the desired EPO mutant [92]. This synthetic protocol was recently extended by the same research group to the synthesis of full length EPO with one mutation site (Gln78 to Ala) bearing well-defined glycoforms (biantennary sialyloliosaccharide) at one (Asn83), two (Asn38 and Asn83, Asn24 and Asn83, Asn24 and Asn38) and three (Asn24, Asn38, Asn83) native EPO *N*-glycosylation sites [91].

9 Alternative Access to *N*-Linked Glycans and *N*-Linked Glycopeptides Using Glycosidic Bond Mimetics

The synthesis of glycoconjugate mimetics is an alternative approach to construct sugar structures or incorporate glycans onto peptides in a simplified way as compared to a total synthesis approach. The introduction of a glycosidic bond mimetic may improve the stability of the glycoconjugate towards chemical and enzymatic degradation, which is highly beneficial for pharmaceutical applications [185]. The use of the copper(I)-catalysed Huisgen 1,3-dipolar cycloaddition of alkynes and azides to afford a 1,2,3-triazole conjugate (CuAAC 'click chemistry') [186, 187] has increased in popularity in recent years, in the peptidomimetic field and as a bioconjugation strategy, due to its simplicity, the mild reaction conditions, its tolerance of various functional groups and its complete regioselectivity to form 1,4-disubstituted products [188].

The syntheses of 'click' mimetics of fish antifreeze glycopeptides [49, 189] and a 20 amino acid MUC 1 domain [54] were successfully undertaken in our laboratory [190]. We have also developed a powerful strategy where two ligation techniques, NCL [74] and CuAAC [186, 187], are carried out in a sequential manner to afford 'click' neoglycopeptides in a 'one pot' fashion [51]. Another highly attractive method which combines the CuAAC strategy [186, 187] and Shoda's [179] direct synthesis of sugar oxazolines from reducing sugars in water to afford 1,2,3-triazole-linked glycoconjugates in a 'one pot' reaction was recently reported [191]. This strategy allows the facile conjugation of reducing sugars with a diverse array of alkynes, including other sugars and peptides. This methodology can potentially be used as a simpler alternative to access homogeneous glycopeptides and possibly glycoproteins in cases where installation of the native *N*-linkage using an enzymatic approach fails or efficient access to glycosidic bond mimetics is required.

10 Applications of Convergent Enzymatic Glycosylation for the Synthesis of Glycopeptides with Therapeutic Potential

Our on-going interest in the synthesis of peptide-glycoconjugates [48, 52, 53, 56, 87] and glycopeptide mimetics that contain non-natural glycan-peptide linkages (neoglycopeptides) [190] using total synthesis prompted us to investigate the alternative enzymatic approach. Herein, a summary of our recent work on the convergent chemoenzymatic synthesis of *N*-linked glycopeptides with therapeutic potential is described.

10.1   Synthesis of a Library of Glycosylated Analogues of Pramlintide

Glycosylation of peptides and proteins is an important tool for producing therapeutic peptidomimetics with improved physicochemical and pharmacokinetic profiles [20, 21]. With this idea in mind, we investigated the effect of the *N*-linked glycosylation of the therapeutic peptide pramlintide (Symlin®), a 37-amino acid synthetic analogue of amylin that is currently used in conjunction with insulin for the treatment of type 1 and type 2 diabetes (Fig. 5) [192–194].

Based on the promising previous results obtained from in vitro and in vivo studies on the *N*-glycosylation of pramlintide at Asn3 and Asn21 with mono-, penta- and undecasaccharides [165], we undertook a systematic investigation into the effect of glycan structure and the position of the attachment of the *N*-glycan to the pramlintide peptide on the activity of glycosylated peptides to act as agonists of amylin receptors [58]. There are six possible sites for *N*-glycosylation of pramlintide; Asn3, Asn14, Asn21, Asn22, Asn31 and Asn35. The synthetic strategy was designed to accommodate the presence of GlcNAc units at defined Asn residues within the pramlintide sequence for subsequent enzymatic transfer of more complex sugar structures. In addition, the disulfide bond (Cys2/Cys7) and amidated *C*-terminus of the peptide had to be installed, as both features are required for biological activity of pramlintide [195].

First, we synthesised a non-glycosylated pramlintide as a control peptide, which was prepared using microwave-enhanced Fmoc SPPS to afford the reduced pramlintide precursor **12**. Disulfide bond formation of **12** was subsequently carried out upon the activation using 2,2′-dipyridyl disulfide (DPDS) in dimethyl sulfoxide (DMSO) [196] to afford the cyclic (Cys2/Cys7) and *C*-amidated pramlintide (**13**), Scheme 5.
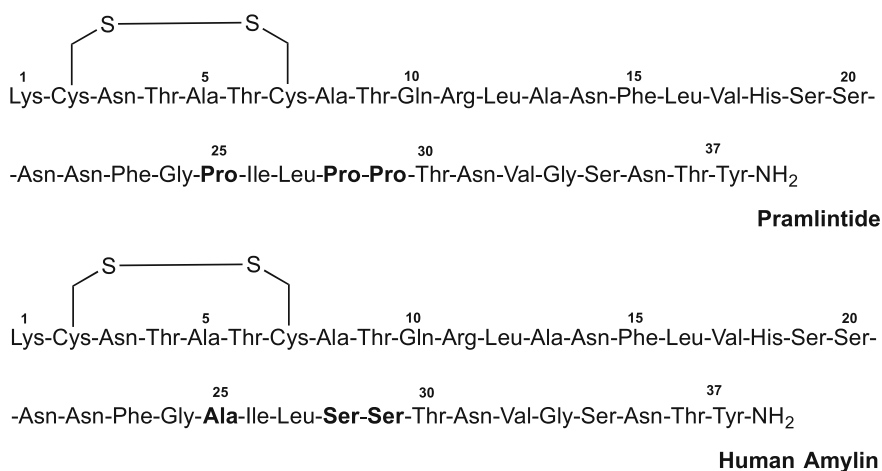
Fig. 5 Primary sequence of pramlintide and human amylin
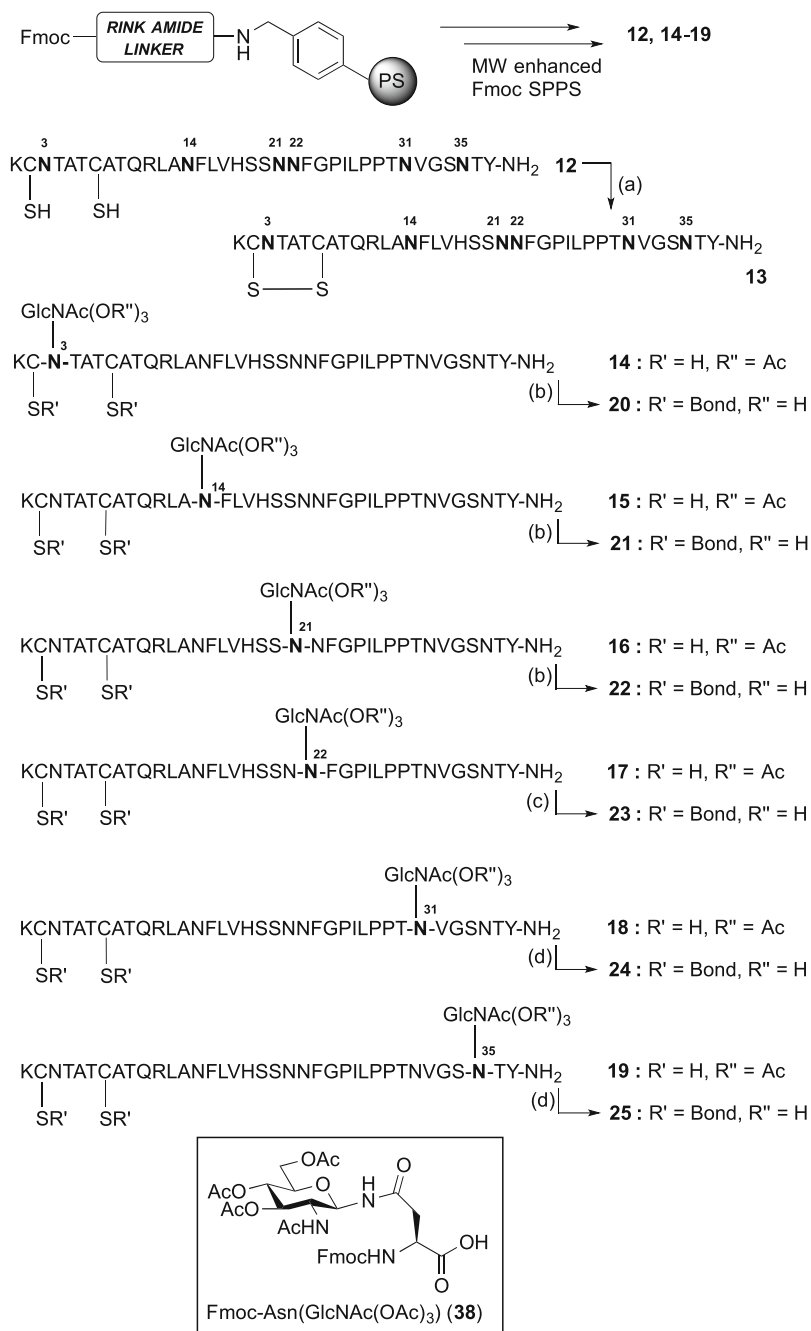
Scheme 5 Synthesis of pramlintide **13**, and pramlintide analogues **20–25** containing a GlcNAc residue at Asn3, Asn14, Asn21, Asn22, Asn31 or Asn35. Reagents and conditions: **a** DPDS, DMSO, rt; **b** DPDS, DMSO; then 13% NH$_2$NH$_2$·1.5 H$_2$O, rt; **c** 5% NH$_2$NH$_2$·1.5 H$_2$O, 10% DMSO, 85% 6 M Gu·HCl, 17 h, rt; **d** DPDS, DMSO; then 5% NH$_2$NH$_2$·1.5 H$_2$O, rt [58]

For the synthesis of monoglycosylated pramlintide analogues (**20**–**25**), comprising a GlcNAc unit at specific Asn residues, microwave-enhanced Fmoc SPPS was employed. The GlcNAc substitution was introduced using the per-*O*-acetylated Fmoc-Asn[GlcNAc(OAc)$_3$] building block (**38**) [197] to give linear, sugar hydroxyl protected pramlintide analogues **14**–**19**. Subsequent use of a 'one pot approach' developed by Hojo et al. [198] to form the Cys2/Cys7 disulfide bond with simultaneous removal of the glycan hydroxyl acetate protecting groups required long reaction times (17 h) to obtain the desired product **23**. We found that the reaction was significantly accelerated when both reactions were performed sequentially in the same vessel whereby the linear, acetate-protected glycopeptides **14**–**16**, and **18**–**19** were first treated with DPDS in DMSO to effect disulfide bond formation, then hydrazine hydrate was added to deprotect the sugar hydroxyls [58]. For the synthesis of **24** and **25,** the total reaction time was significantly reduced to 5 and 5.5 h, respectively. Faster acetate removal was achieved using a higher concentration of hydrazine hydrate (3 h in total, for analogues **20**–**22**), Scheme 5.

A library of pramlintide analogues **26**–**31** bearing the core *N*-glycan pentasaccharide [Man$_3$(GlcNAc)$_2$] was then synthesised using Endo A [121, 128] to transfer tetrasaccharide oxazoline **4** (accessed via total synthesis [140]) to the corresponding GlcNAc-tagged pramlintide **20**–**25** (Scheme 6).

This methodology was then extended to the preparation of pramlintide analogues **32**–**37** bearing a complex biantennary glycan [(NeuAcGalGlcNAcMan)$_2$Man (GlcNAc)$_2$]. In this case treatment of the decasaccharide-oxazoline **5**, synthesised from the corresponding reducing sugar that was isolated from egg yolks [181, 199], with the commercially available Endo M N175Q mutant [168, 177] in the presence of **20**–**25** enabled the preparation of pramlintide analogues **32**–**37**, (Scheme 7) [58].

A comprehensive series of 18 pramlintide analogues comprising mono-, penta- and undecasaccharides (**20**–**37**) were then tested as agonists of amylin receptors and their activity was compared to parent pramlintide (**13**). The parent pramlintide **13** and analogues **20**–**25** bearing GlcNAc unit were screened against the three best characterised amylin receptors (CT$_{(a)}$, AMY$_{1(a)}$, and AMY$_{3(a)}$, which contain the CT$_{(a)}$ splice variant of the calcitonin receptor) [200, 201] at which activity of pramlintide is equal to human or rat amylin [202]. Analogues **26**–**37** containing more complex glycans were only tested against AMY$_{1(a)}$ analogously to the previous study [165].

The study revealed that the presence of *N*-glycans was well tolerated at Asn21, Asn31 and Asn35 by the AMY$_{1(a)}$ receptor, and that the activity of analogues versus the amylin receptors decreases as the size of the glycan increases (GlcNAc > pentasaccharide > undecasaccharide). It was therefore established that *N*-glycosylation of pramlintide is a promising tool to afford analogues with improved therapeutic potential. In vivo studies to assess the biological importance of *N*-glycosylation of pramlintide are under investigation and results will be reported in due course [58].
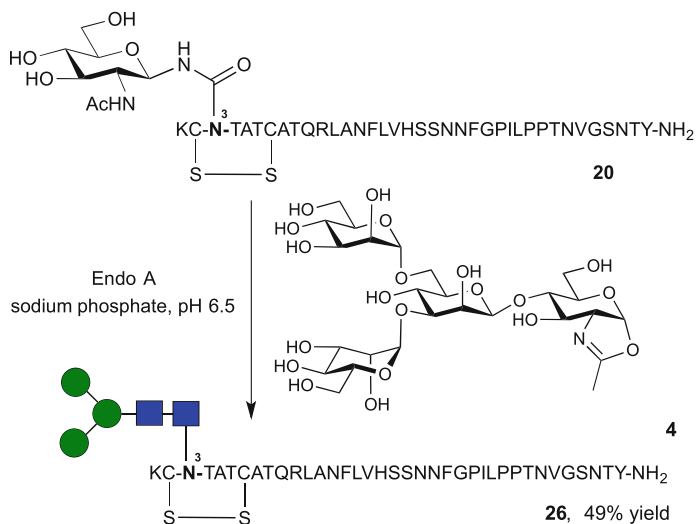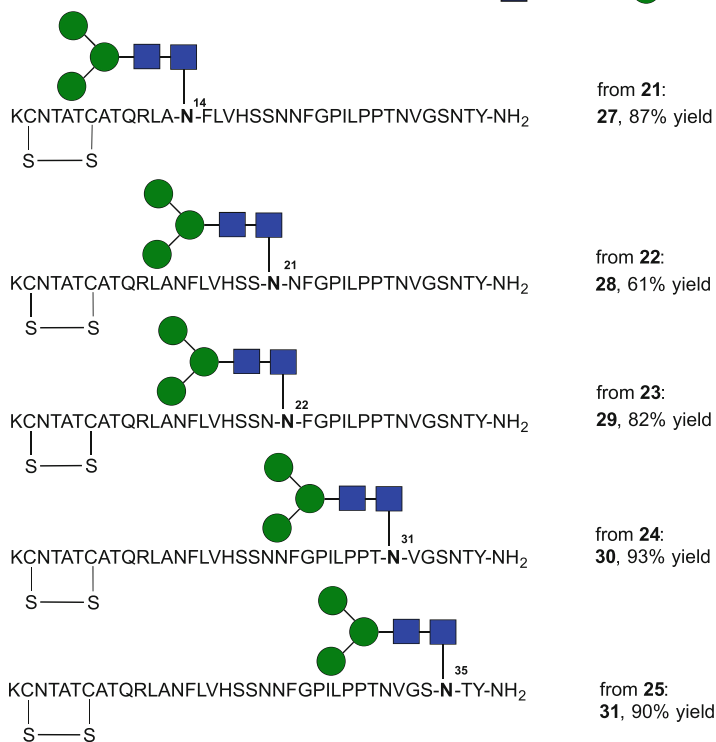
Scheme 6 Synthesis of pramlintide analogues **26–31** containing the core *N*-glycan pentasaccharide [Man$_3$(GlcNAc)$_2$] at position 3, 14, 21, 22, 31 or 35 [58]
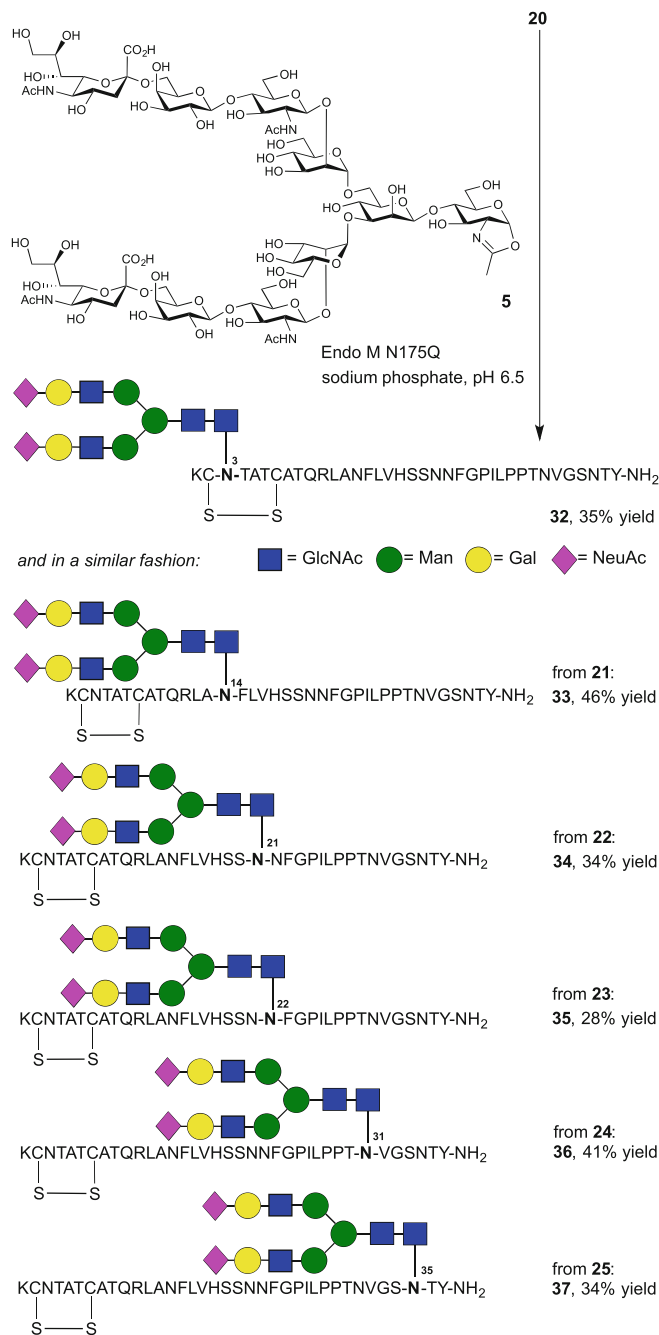
Scheme 7 Synthesis of pramlintide analogues 32–37 containing a complex biantennary glycan [(NeuAcGalGlcNAcMan)$_2$Man(GlcNAc)$_2$] at position 3, 14, 21, 22, 31 or 35 [58]

10.2 Synthesis of Mannosylated Glycopeptides

Our interest in the synthesis of glycopeptide-based vaccine candidates comprising mannose units to target antigen presenting cells (APCs) responsible for initiating an immune response via mannose receptor (MR), led us previously to prepare mono- and di-mannosylated and 5(6)-carboxyfluorescein (5(6)-CF) labelled glycopeptides by chemical synthesis [48, 87]. Subsequent progression of this work involved the synthesis of glycopeptide-based vaccine candidates comprising more complex high-mannose-type N-glycans and testing their ability to bind to APCs [57]. For this purpose, the pp65 protein fragment 491–509 from the cytomegalovirus (CMV) [ILARNLVPMVATVQGQNLK] incorporating peptide epitope pp65$_{495–503}$ (NLVPMVATV) recognised by human cytotoxic T-lymphocyte (CTL) was chosen as a synthetic target. The target peptide contains two asparagine residues Asn5 and Asn17 conveniently located within the pp65$_{491–509}$ sequence for potential attachment of sugar residues. To allow detection of the peptides using flow cytometry 5(6)-carboxyfluorescein was attached to the N-terminus of the glycopeptides.
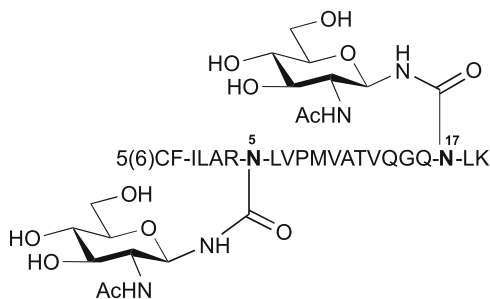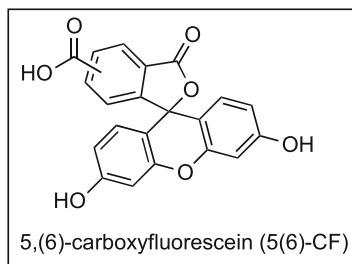
5(6)-CF-labelled control peptide **39**, and glycopeptides **40** and **41** comprising either one- or two-GlcNAc units, respectively, were synthesised using microwave-enhanced Fmoc SPPS wherein the Fmoc-Asn[GlcNAc(OAc)$_3$] building block **38** [197] replaced either Asn17 (for **40**) or Asn5 and Asn17 (for **41**) as required. Subsequent reduction of the methionine sulfoxide of the control peptide, used in place of methionine, was performed following literature procedures [203] and afforded the pp65 protein fragment 491–509 (**39**). Removal of sugar hydroxyl protecting groups of the per-O-acetylated pp65$_{491–509}$ precursors of **40** and **41** (sodium methoxide in methanol) was undertaken prior to reduction of the methionine sulfoxide [203] to give GlcNAc-tagged glycopeptides **40** and **41** ready for further enzymatic glycosylation (Fig. 6).

With GlcNAc-tagged glycopeptides **40** and **41** in hand, we next focused on the synthesis of N-glycopeptides bearing a Man$_3$-terminated pentasaccharide. This was successfully undertaken using oxazoline donor **4** [140], glycopeptide acceptors **40** and **41**, and the Endo A E173H mutant [153] to afford **42** and **44**, bearing either one- or two-pentasaccharide units, respectively, in good yield (68 and 89%, respectively), Scheme 8a.

To access full-length high-mannose N-linked glycopeptides **43** and **45** bearing nine mannose units at each glycosylation site, oxazoline donor **46** was used which was conveniently sourced from soy bean flour [161]. Somewhat surprisingly the Endo A E173H mutant [153] proved incapable of transferring the full-length high-mannose oxazoline **46** onto glycopeptide acceptors **40** and **41**, possibly due to an altered substrate tolerance of the mutated enzyme as compared to wild-type Endo A [57]. We, therefore, used commercially available Endo M N175Q mutant
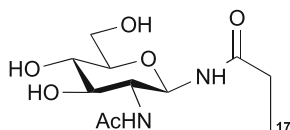
Fig. 6 CMV control peptide 39 and glycopeptides 40 and 41 containing a GlcNAc residue either at Asn17, or at both Asn5 and Asn17, respectively [57]

[168, 177] which produced the desired *N*-undecasaccharide-glycopeptides 43 and 45 with either one- or two-Man$_9$(GlcNAc)$_2$ units in 48% and 54% yield, respectively (Scheme 8b).

Subsequent analysis to assess glycopeptide binding levels to APCs indicated improved targeting of the peptide cargo to MR-expressing cells due to the presence of the high-mannose *N*-glycans. This effect was more pronounced for analogues glycosylated at both asparagines (44 and 45) as compared to counterparts bearing a single *N*-glycan at Asn17 (42 and 43). In addition, stronger binding was observed for glycopeptides bearing the high-mannose unit, Man$_9$(GlcNAc)$_2$ (43 and 45) than those with the truncated glycan, Man$_3$(GlcNAc)$_2$ (42 and 44). Importantly, it was also found that analogues in which the sugars were sited outside the epitope sequence (either Man$_3$(GlcNAc)$_2$ or Man$_9$(GlcNAc)$_2$ at Asn17, 42 and 43, respectively) were readily processed and presented by the APCs to human T cells.

These results provide important evidence that *N*-glycosylation of peptides using high-mannose glycans may produce superior compounds for vaccine development. Additionally, we have demonstrated the effectiveness of a convergent chemoenzymatic approach to readily obtain complex *N*-linked glycopeptides with therapeutic potential.

■ = GlcNAc ● = Man

(a)

5(6)CF-ILAR-**N**-LVPMVATVQGQ-**N**-LK

42, 68% yield

4

pp65$_{491-509}$ analogue (**40**)

(a)

(b)

46

5(6)CF-ILAR-**N**-LVPMVATVQGQ-**N**-LK

43, 48% yield

5(6)CF-ILAR-**N**-LVPMVATVQGQ-**N**-LK

44, 89% yield

4 | (a)

pp65$_{491-509}$ analogue (**41**)

46 | (b)

5(6)CF-ILAR-**N**-LVPMVATVQGQ-**N**-LK

45, 54% yield
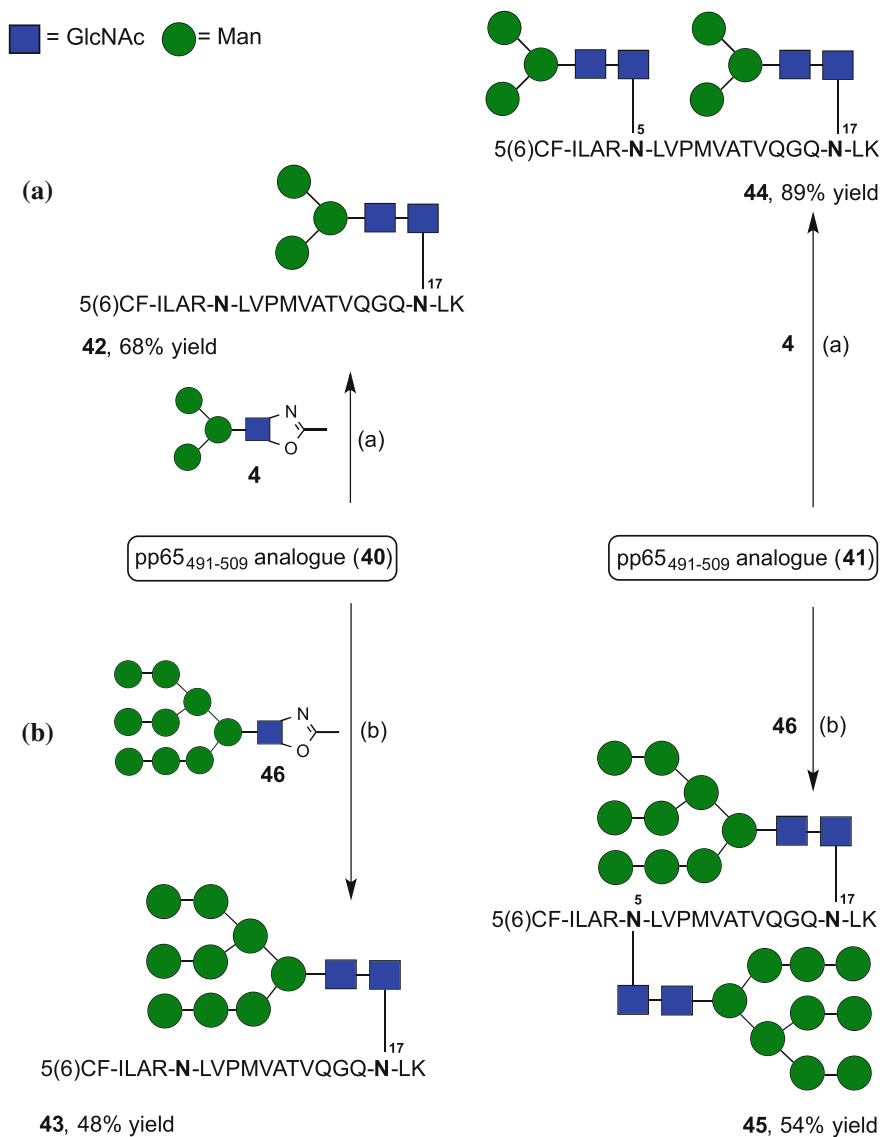
Scheme 8 Synthesis of **a** glycopeptides **42** and **44** containing a Man$_3$(GlcNAc)$_2$ residue, and **b** glycopeptides **43** and **45** containing a Man$_9$(GlcNAc)$_2$ residue, either at Asn17, or at both Asn5 and Asn17, respectively. Reagents and conditions: **a** Endo A E173H, sodium phosphate buffer pH 6.5; **b** Endo M N175Q, sodium phosphate buffer pH 6.5 [57]

11 Conclusions

The synthesis of homogeneous peptides and proteins is still a complex and onerous task. The synthesis of the sugar component is a limiting factor, especially when laborious total synthesis routes are employed. Fortunately, recent progress in the use of chemoenzymatic techniques using ENGase-mediated glycosylation has demonstrated significant potential. By careful design of reaction conditions and appropriate selection of partners for glycosylation, a wide range of peptide-oligosaccharide structures can be obtained. The use of enzyme-mediated synthesis in combination with chemical synthetic techniques provides a method to access complex, highly desirable glycoconjugates efficiently [101, 204].

References

1. Spiro RG (2002) Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds. Glycobiology 12:43R–56R
2. Meledeo MA, Yarema KJ, Begley TP (2007) Glycan biosynthesis in mammals, in wiley encyclopedia of chemical biology. Wiley, New Jersey
3. Lafite P, Daniellou R (2012) Rare and unusual glycosylation of peptides and proteins. Nat Prod Rep 29:729–738
4. Haltiwanger RS, Lowe JB (2004) Role of Glycosylation in Development. Annu Rev Biochem 73:491–537
5. Dube DH, Bertozzi CR (2005) Glycans in cancer and inflammation-potential for therapeutics and diagnostics. Nat Rev Drug Discov 4:477–488
6. Dwek RA (1996) Glycobiology: toward understanding the function of sugars. Chem Rev 96:683–720
7. Imperiali B, O'Connor SE (1999) Effect of N-linked glycosylation on glycopeptide and glycoprotein structure. Curr Opin Chem Biol 3:643–649
8. Wyss DF, Choi JS (1995) Conformation and function of the N-linked glycan in the adhesion domain of human CD2. Science 269:1270
9. Opdenakker G, Rudd PM, Ponting CP, Dwek RA (1993) Concepts and principles of glycobiology. FASEB J 7:1330–1337
10. Erbayraktar S, Grasso G, Sfacteria A, Xie Q-W, Coleman T, Kreilgaard M, Torup L, Sager T, Erbayraktar Z, Gokmen N, Yilmaz O, Ghezzi P, Villa P, Fratelli M, Casagrande S, Leist M, Helboe L, Gerwein J, Christensen S, Geist MA, Pedersen LØ, Cerami-Hand C, Wuerth J-P, Cerami A, Brines M (2003) Asialoerythropoietin is a nonerythropoietic cytokine with broad neuroprotective activity in vivo. Proc Nat Acad Sci USA 100:6741–6746
11. Schachter H, Freeze HH (2009) Glycosylation diseases: quo vadis? Biochim Biophys Acta Mol Basis Dis 1792:925–930
12. Murakami Y, Kinoshita T (2015) Congenital Disorders of Glycosylation: Glycosylphosphatidylinositol (GPI)-Related. In: Taniguchi N, Endo T, Hart GW, Seeberger PH, Wong C-H (eds) Glycoscience: biology and medicine. Springer, Japan, pp 1229–1236
13. Akimoto Y, Miura Y, Endo T, Kawakami H, Hart G (2015) Diabetes and O-GlcNAcylation. In: Taniguchi N, Endo T, Hart GW, Seeberger PH, Wong C-H (eds) Glycoscience: biology and medicine. Springer, Japan, pp 1207–1212
14. Hart GW, Housley MP, Slawson C (2007) Cycling of O-linked β-N-acetylglucosamine on nucleocytoplasmic proteins. Nature 446:1017–1022

15. Zeidan Q, Hart GW (2010) The intersections between O-GlcNAcylation and phosphory-lation: implications for multiple signaling pathways. J Cell Sci 123:13–22
16. Stowell SR, Ju T, Cummings RD (2015) Protein Glycosylation in Cancer. Annu Rev Pathol Mech Dis 10:473–510
17. Korekane H, Taniguchi N (2015) Glycosylation in cancer: enzymatic basis for alterations in N-glycan branching. In: Taniguchi N, Endo T, Hart GW, Seeberger PH, Wong C-H (eds) Glycoscience: biology and medicine. Springer, Japan, pp 1349–1356
18. Schedin-Weiss S, Winblad B, Tjernberg LO (2014) The role of protein glycosylation in Alzheimer disease. FEBS J 281:46–62
19. Gao C, Taniguchi N (2015) Chronic Obstructive Pulmonary Disease (COPD). In: Taniguchi N, Endo T, Hart GW, Seeberger PH, Wong C-H (eds) Glycoscience: Biology and Medicine. Springer, Japan, pp 1267–1274
20. Solá RJ, Griebenow K (2009) Effects of Glycosylation on the Stability of Protein Pharmaceuticals. J Pharm Sci 98:1223–1245
21. Solá RJ, Griebenow K (2010) Glycosylation of Therapeutic Proteins: an Effective Strategy to Optimize Efficacy. BioDrugs 24:9–21
22. Ueda T, Tomita K, Notsu Y, Ito T, Fumoto M, Takakura T, Nagatome H, Takimoto A, Mihara SI, Togame H, Kawamoto K, Iwasaki T, Asakura K, Oshima T, Hanasaki K, Nishimura SI, Kondo H (2009) Chemoenzymatic Synthesis of Glycosylated Glucagon-like Peptide 1: effect of glycosylation on proteolytic resistance and in vivo blood glucose-lowering activity. J Am Chem Soc 131:6237–6245
23. Sinclair AM, Elliott S (2005) Glycoengineering: the effect of glycosylation on the properties of therapeutic proteins. J Pharm Sci 94:1626–1635
24. Costa AR, Rodrigues ME, Henriques M, Oliveira R, Azeredo J (2014) Glycosylation: impact, control and improvement during therapeutic protein production. Crit Rev Biotechnol 34:281–299
25. Li HJ, d'Anjou M (2009) Pharmacological significance of glycosylation in therapeutic proteins. Curr Opin Biotechnol 20:678–684
26. Jefferis R (2009) Glycosylation as a strategy to improve antibody-based therapeutics. Nat Rev Drug Discov 8:226–234
27. Sato M, Furuike T, Sadamoto R, Fujitani N, Nakahara T, Niikura K, Monde K, Kondo H, Nishimura SI (2004) Glycoinsulins: Dendritic sialyloligosaccharide-displaying insulins showing a prolonged blood-sugar-lowering activity. J Am Chem Soc 126:14013–14022
28. Ueda T, Ito T, Tomita K, Togame H, Fumoto M, Asakura K, Oshima T, Nishimura SI, Hanasaki K (2010) Identification of glycosylated exendin-4 analogue with prolonged blood glucose-lowering activity through glycosylation scanning substitution. Bioorg Med Chem Lett 20:4631–4634
29. Runkel L, Meier W, Pepinsky RB, Karpusas M, Whitty A, Kimball K, Brickelmaier M, Muldowney C, Jones W, Goelz SE (1998) Structural and functional differences between glycosylated and non-glycosylated forms of human interferon-β (IFN-β). Pharm Res 15:641–649
30. Lappin TRJ, Maxwell AP (1989) Chemistry and assays of erythropoietin. In: Erythropoietin, W. Jelkmann and A. J. Gross, Editors. 1989, Springer Berlin Heidelberg: Berlin, Heidelberg. p. 7–18
31. Davis BG (2002) Synthesis of glycoproteins. Chem Rev 102:579–602
32. Gamblin DP, Scanlan EM, Davis BG (2009) Glycoprotein synthesis: an update. Chem Rev 109:131–163
33. Wang L-X, Amin Mohammed N (2014) Chemical and chemoenzymatic synthesis of glycoproteins for deciphering functions. Chem Biol 21:51–66
34. Wang L-X, Davis BG (2013) Realizing the promise of chemical glycobiology. Chem Sci 4:3381–3394
35. Unverzagt C, Kajihara Y (2013) Chemical assembly of N-glycoproteins: a refined toolbox to address a ubiquitous posttranslational modification. Chem Soc Rev 42:4408–4420

36. Izumi M, Okamoto R, Kajihara Y (2015) Chemical synthesis of homogeneous glycoproteins. In: Taniguchi N, Endo T, Hart GW, Seeberger PH, Wong C-H (eds) Glycoscience: biology and medicine. Springer, Japan, pp 313–321

37. Westerlind U (2012) Synthetic glycopeptides and glycoproteins with applications in biological research. Beilstein J Org Chem 8:804–818

38. Fernández-Tejada A, Brailsford J, Zhang Q, Shieh J-H, Moore MAS, Danishefsky SJ (2015) Total synthesis of glycosylated proteins. In: Liu L (ed) Protein ligation and total synthesis I. Springer International Publishing, Cham, pp 1–26

39. Kajihara Y, Yamamoto N, Okamoto R, Hirano K, Murase T (2010) Chemical synthesis of homogeneous glycopeptides and glycoproteins. Chem Rec 10:80–100

40. Payne RJ, Wong C-H (2010) Advances in chemical ligation strategies for the synthesis of glycopeptides and glycoproteins. Chem Commun 46:21–43

41. Rich JR, Withers SG (2009) Emerging methods for the production of homogeneous human glycoproteins. Nat Chem Biol 5:206–215

42. Xu C, Li X (2015) Glycopeptide/glycoprotein synthesis. In: Taniguchi N, Endo T, Hart GW, Seeberger PH, Wong C-H (eds) Glycoscience: biology and medicine. Springer, Japan, pp 323–330

43. Fairbanks AJ (2013) Endohexosaminidase-catalyzed synthesis of glycopeptides and proteins. Pure Appl Chem 85:1847–1863

44. Ellgaard L, Helenius A (2003) Quality control in the endoplasmic reticulum. Nat Rev Mol Cell Biol 4:181

45. Stanley P, Schachter H, Taniguchi N (2009) N-glycans. In: Varki A, Cummings RD, Esko JD, Freeze HH, Stanley P, Bertozzi CR, Hart GW, Etzler ME (eds) Essentials of glycobiology. Cold Spring Harbor Laboratory Press, New York

46. Herzner H, Reipen T, Schultz M, Kunz H (2000) Synthesis of glycopeptides containing carbohydrate and peptide recognition motifs. Chem Rev 100:4495–4538

47. Rudd PM, Dwek RA (1997) Glycosylation: heterogeneity and the 3D structure of proteins. Crit Rev Biochem Mol Biol 32:1–100

48. Brimble MA, Kowalczyk R, Harris PWR, Dunbar PR, Muir VJ (2008) Synthesis of fluorescein-labelled O-mannosylated peptides as components for synthetic vaccines: comparison of two synthetic strategies. Org Biomol Chem 6:112–121

49. Miller N, Williams GM, Brimble MA (2009) Synthesis of fish antifreeze neoglycopeptides using microwave-assisted "click chemistry". Org Lett 11:2409–2412

50. Kowalczyk R, Harris PWR, Dunbar PR, Brimble MA (2009) Stability of 5(6)-carboxyfluorescein in microwave-assisted synthesis of fluorescein-labelled O-dimannosylated peptides. Synthesis 2009:2210–2222

51. Lee DJ, Mandal K, Harris PWR, Brimble MA, Kent SBH (2009) A one-pot approach to neoglycopeptides using orthogonal native chemical ligation and click chemistry. Org Lett 11:5270–5273

52. Lee DJ, Harris PWR, Kowalczyk R, Dunbar PR, Brimble MA (2010) Microwave-assisted synthesis of fluorescein-labelled GalNAcα1-O-Ser/Thr (Tn) glycopeptides as immunological probes. Synthesis 2010:763–769

53. Peltier R, Evans CW, DeVries AL, Brimble MA, Dingley AJ, Williams DE (2010) Growth habit modification of ice crystals using antifreeze glycoprotein (AFGP) analogues. Cryst Growth Des 10:5066–5077

54. Lee DJ, Harris PWR, Brimble MA (2011) Synthesis of MUC1 neoglycopeptides using efficient microwave-enhanced chaotrope-assisted click chemistry. Org Biomol Chem 9:1621–1626

55. Lee DJ, Yang S-H, Williams GM, Brimble MA (2012) Synthesis of multivalent neoglyconjugates of MUC1 by the conjugation of carbohydrate-centered, triazole-linked gllycoclusters to MUC1 peptides using click chemistry. J Org Chem 77:7564–7571

56. Brimble MA, Edwards PJ, Harris PWR, Norris GE, Patchett ML, Wright TH, Yang S-H, Carley SE (2015) Synthesis of the antimicrobial S-linked glycopeptide, glycocin F. Chem Eur J 21:3556–3561

57. McIntosh JD, Brimble MA, Brooks AES, Dunbar PR, Kowalczyk R, Tomabechi Y, Fairbanks AJ (2015) Convergent chemo-enzymatic synthesis of mannosylated glycopeptides; targeting of putative vaccine candidates to antigen presenting cells. Chem Sci 6:4636–4642

58. Kowalczyk R, Brimble MA, Tomabechi Y, Fairbanks AJ, Fletcher M, Hay DL (2014) Convergent chemoenzymatic synthesis of a library of glycosylated analogues of pramlintide: structure-activity relationships foramyl in receptor agonism. Org Biomol Chem 12:8142–8151

59. Reichert JM (2012) Marketed therapeutic antibodies compendium. mAbs 4:413–415

60. http://top101news.com/2015-2016-2017-2018/news/health/best-selling-drugs-world/. Top 10 Best selling Drugs in the World

61. Bang LM, Keating GM (2004) Adalimumab a review of its use in rheumatoid arthritis. BioDrugs 18:121–139

62. Mazumdar S, Greenwald D (2009) Golimumab. mAbs 1:422–431

63. Swiech K, de Freitas M, Covas D, Picanço-Castro V (2015) Recombinant glycoprotein production in human cell lines. In: García-Fruitós E (ed) Insoluble proteins. Springer, New York, pp 223–240

64. Hamilton SR, Davidson RC, Sethuraman N, Nett JH, Jiang Y, Rios S, Bobrowicz P, Stadheim TA, Li H, Choi B-K, Hopkins D, Wischnewski H, Roser J, Mitchell T, Strawbridge RR, Hoopes J, Wildt S, Gerngross TU (2006) Humanization of yeast to produce complex terminally sialylated glycoproteins. Science 313:1441–1443

65. Hamilton SR, Bobrowicz P, Bobrowicz B, Davidson RC, Li H, Mitchell T, Nett JH, Rausch S, Stadheim TA, Wischnewski H, Wildt S, Gerngross TU (2003) Production of complex human glycoproteins in yeast. Science 301:1244–1246

66. Wang L-X, Lomino JV (2012) Emerging technologies for making glycan-defined glycoproteins. ACS Chem Biol 7:110–122

67. Fernandez-Tejada A, Danishefsky SJ (2014) Chapter 25 Development of cancer vaccines from fully synthetic mucin-based glycopeptide antigens. A vision on mucins from the bioorganic chemistry perspective. In: Carbohydrate Chemistry, vol 40. The Royal Society of Chemistry, pp 533–563

68. Gaidzik N, Westerlind U, Kunz H (2013) The development of synthetic antitumour vaccines from mucin glycopeptide antigens. Chem Soc Rev 42:4421–4442

69. Payne RJ (2013) Total synthesis of erythropoietin through the development and exploitation of enabling synthetic technologies. Angew Chem Int Ed 52:505–507

70. Kajihara Y, Okamoto R, Yamamoto N, Izumi M (2010) Chapter twenty-four—synthesis of glycopeptides. In: Minoru F (ed) Methods in enzymology. Academic Press, pp 503–519

71. Panda SS, Jones RA, Dennis Hall C, Katritzky AR (2015) Applications of chemical ligation in peptide synthesis via acyl transfer. In: Liu L (ed) Protein ligation and total synthesis I. Springer International Publishing, Cham, pp 229–265

72. Behrendt R, White P, Offer J (2016) Advances in Fmoc solid-phase peptide synthesis. J Pept Sci 22:4–27

73. Merrifield RB (1963) Solid phase peptide synthesis. I. The synthesis of a tetrapeptide. J Am Chem Soc 85:2149–2154

74. Dawson P, Muir T, Clark-Lewis I, Kent S (1994) Synthesis of proteins by native chemical ligation. Science 266:776–779

75. Muir TW (2003) Semisynthesis of proteins by expressed protein ligation. Annu Rev Biochem 72:249–289

76. Xu C, Lam HY, Zhang Y, Li X (2013) Convergent synthesis of MUC1 glycopeptides via serine ligation. Chem Commun 49:6200–6202

77. Zhang Y, Xu C, Lam HY, Lee CL, Li X (2013) Protein chemical synthesis by serine and threonine ligation. Proc Nat Acad Sci USA 110:6657–6662

78. Li X, Lam HY, Zhang Y, Chan CK (2010) Salicylaldehyde ester-induced chemoselective peptide ligations: enabling generation of natural peptidic linkages at the serine/threonine sites. Org Lett 12:1724–1727

79. Sakamoto I, Tezuka K, Fukae K, Ishii K, Taduru K, Maeda M, Ouchi M, Yoshida K, Nambu Y, Igarashi J, Hayashi N, Tsuji T, Kajihara Y (2012) Chemical synthesis of homogeneous human glycosyl-interferon-β that exhibits potent antitumor activity in vivo. J Am Chem Soc 134:5428–5431

80. Izumi M, Makimura Y, Dedola S, Seko A, Kanamori A, Sakono M, Ito Y, Kajihara Y (2012) Chemical synthesis of intentionally misfolded homogeneous glycoprotein: a unique approach for the study of glycoprotein quality Control. J Am Chem Soc 134:7238–7241

81. Hojo H, Tanaka H, Hagiwara M, Asahina Y, Ueki A, Katayama H, Nakahara Y, Yoneshige A, Matsuda J, Ito Y, Nakahara Y (2012) Chemoenzymatic Synthesis of hydrophobic glycoprotein: synthesis of saposin c carrying complex-type Carbohydrate. J Org Chem 77:9437–9446

82. Piontek C, Ring P, Harjes O, Heinlein C, Mezzato S, Lombana N, Pöhner C, Püttner M, Varón Silva D, Martin A, Schmid FX, Unverzagt C (2009) Semisynthesis of a homogeneous glycoprotein enzyme: Ribonuclease C: part 1. Angew Chem Int Ed 48:1936–1940

83. Piontek C, Varón Silva D, Heinlein C, Pöhner C, Mezzato S, Ring P, Martin A, Schmid FX, Unverzagt C (2009) Semisynthesis of a homogeneous glycoprotein enzyme: Ribonuclease C: part 2. Angew Chem Int Ed 48:1941–1945

84. Haase C, Seitz O (2007) Chemical synthesis of glycopeticles. In: Wittmann V (ed) Glycopeptides and glycoproteins: synthesis, structure, and application. Springer, Berlin, pp 1–36

85. Palitzsch B, Gaidzik N, Stergiou N, Stahn S, Hartmann S, Gerlitzki B, Teusch N, Flemming P, Schmitt E, Kunz H (2016) A synthetic glycopeptide vaccine for the induction of a monoclonal antibody that differentiates between normal and tumor mammary cells and enables the diagnosis of human pancreatic cancer. Angew Chem Int Ed 55:2894–2898

86. Hartmann S, Palitzsch B, Glaffig M, Kunz H (2014) Chapter 24 tumour-associated glycopeptide antigens and their modification in anticancer vaccines. In: Carbohydrate chemistry: volume 40. The Royal Society of Chemistry, pp 506–532

87. Kowalczyk R, Harris PWR, Dunbar RP, Brimble MA (2009) Stability of 5(6)-carboxyfluorescein in microwave-assisted synthesis of fluorescein-labelled o-dimannosylated peptides. Synthesis, 2210–2222

88. Lee DJ, Harris PWR, Kowalczyk R, Dunbar PR, Brimble MA (2010) Microwave-assisted synthesis of fluorescein-labelled GalNAc α 1-O-Ser/Thr (Tn) glycopeptides as immunological probes. Synthesis-Stuttgart, 763–769

89. Yamamoto N, Ohmori Y, Sakakibara T, Sasaki K, Juneja LR, Kajihara Y (2003) Solid-phase synthesis of sialylglycopeptides through selective Esterification of the sialic acid residues of an asn-linked complex-type sialyloligosaccharide. Angew Chem Int Ed 42:2537–2540

90. Kajihara Y, Suzuki Y, Yamamoto N, Sasaki K, Sakakibara T, Juneja LR (2004) Prompt chemoenzymatic synthesis of diversecomplex-type oligosaccharides and Its application to the solid-phase synthesisof a glycopeptide with asn-linked sialyl-undeca- and asialo-nonasaccharides. Chem Eur J 10:971–985

91. Murakami M, Kiuchi T, Nishihara M, Tezuka K, Okamoto R, Izumi M, Kajihara Y (2016) Chemical synthesis of erythropoietin glycoforms for insights into the relationship between glycosylation pattern and bioactivity. Sci Adv 2:1–12

92. Murakami M, Okamoto R, Izumi M, Kajihara Y (2012) Chemical synthesis of an erythropoietin glycoform containing a complex-type Disialyloligosaccharide. Angew Chem Int Ed 51:3567–3572

93. Yamamoto N, Takayanagi A, Yoshino A, Sakakibara T, Kajihara Y (2007) An approach for a synthesis of asparagine-linked sialylglycopeptides having intact and homogeneous complex-type undecadisialyloligosaccharides. Chem Eur J 13:613–625

94. Cohen-Anisfeld ST, Lansbury PT (1993) A practical, convergent method for glycopeptide synthesis. J Am Chem Soc 115:10531–10537

95. Anisfeld ST, Lansbury PT (1990) A convergent approach to the chemical synthesis of asparagine-linked glycopeptides. J Org Chem 55:5560–5562

96. Wang P, Dong S, Brailsford JA, Iyer K, Townsend SD, Zhang Q, Hendrickson RC, Shieh J, Moore MAS, Danishefsky SJ (2012) At Last: erythropoietin as a single Glycoform. Angew Chem Int Ed 51:11576–11584

97. Wang P, Dong S, Shieh J-H, Peguero E, Hendrickson R, Moore MAS, Danishefsky SJ (2013) Erythropoietin derived by chemical synthesis. Science 342:1357–1360

98. Yan LZ, Dawson PE (2001) Synthesis of peptides and proteins without cysteine Residues by native chemical ligation combined with desulfurization. J Am Chem Soc 123:526–533

99. Wang P, Aussedat B, Vohra Y, Danishefsky SJ (2012) An advance in the chemical synthesis of homogeneous N-linked glycopolypeptides by convergent aspartylation. Angew Chem Int Ed 51:11571–11575

100. Ullmann V, Rädisch M, Boos I, Freund J, Pöhner C, Schwarzinger S, Unverzagt C (2012) Convergent solid-phase synthesis of NGlycopeptides facilitated by pseudoprolines at consensus-sequence Ser/Thr residues. Angew Chem Int Ed 51:11566–11570

101. Katoh T, Yamamoto K (2015) Glycoenzymes in glycan analysis and synthesis. In: Taniguchi N, Endo T, Hart GW, Seeberger PH, Wong C-H (eds) Glycoscience: biology and medicine. Springer, Japan, pp 379–389

102. Malet C, Planas A (1998) From β-glucanase to β-glucansynthase: glycosyl transfer to α-glycosyl fluorides catalyzed by a mutant endoglucanase lacking its catalytic nucleophile. FEBS Lett 440:208–212

103. Mackenzie LF, Wang Q, Warren RAJ, Withers SG (1998) Glycosynthases: mutant glycosidases for oligosaccharide synthesis. J Am Chem Soc 120:5583–5584

104. Hancock SM, Vaughan MD, Withers SG (2006) Engineering of glycosidases and glycosyltransferases. Curr Opin Chem Biol 10:509–519

105. Yamamoto K (2015) Endo-enzymes. In: Taniguchi N, Endo T, Hart GW, Seeberger PH, Wong C-H (eds) Glycoscience: biology and medicine. Springer, Japan, pp 391–399

106. http://www.cazy.org. carbohydrate-active enZYmes database

107. Murakami S, Takaoka Y, Ashida H, Yamamoto K, Narimatsu H, Chiba Y (2013) Identification and characterization of endo-β-nacetylglucosaminidas from methylotrophic yeast Ogataea minuta. Glycobiology 23:736–744

108. Yamamoto K (2001) Chemo-Enzymatic synthesis of bioactive glycopeptide using microbial endoglycosidase. J Biosci Bioeng 92:493–501

109. Kadowaki S, Yamamoto K, Fujisaki M, Kumagai H, Tochikura T (1988) A novel endo-β-N-acetylglucosaminidase acting on complex oligosaccharides of glycoproteins in a fungus. Agric Biol Chem 52:2387–2389

110. Yamamoto KJ, Kadowaki S, Watanabe J, Kumagai H (1994) Transglycosylation activity of mucor hiemalis endo-β-N-acetylglucosaminidase which transfers complex oligosaccharides to the N-Acetylglucosamine moieties of peptides. Biochem Biophys Res Commun 203:244–252

111. Haneda K, Inazu T, Yamamoto K, Kumagai H, Nakahara Y, Kobata A (1996) Transglycosylation of intact sialo complex-type oligosaccharides to the N-acetylglucosamine moieties of glycopeptides by Mucor hiemalis endo-β-N-acetylglucosaminidase. Carbohydr Res 292:61–70

112. Yamamoto K, Fujimori K, Haneda K, Mizuno M, Inazu T, Kumagai H (1997) Chemoenzymatic synthesis of a novel glycopeptide using a microbial endoglycosidase. Carbohydr Res 305:415–422

113. Mizuno M, Haneda K, Iguchi R, Muramoto I, Kawakami T, Aimoto S, Yamamoto K, Inazu T (1999) Synthesis of a glycopeptide containing oligosaccharides: chemoenzymatic synthesis of eel calcitonin analogues having natural N-linked oligosaccharides. J Am Chem Soc 121:284–290

114. Haneda K, Inazu T, Mizuno M, Iguchi R, Yamamoto K, Kumagai H, Aimoto S, Suzuki H, Noda T (1998) Chemo-enzymatic synthesis of calcitonin derivatives containing N-linked oligosaccharides. Bioorg Med Chem Lett 8:1303–1306

115. Yamanoi T, Tsutsumida M, Oda Y, Akaike E, Osumi K, Yamamoto K, Fujita K (2004) Transglycosylation reaction of mucor hiemalis endo-β-N-acetylglucosaminidase using sugar

derivatives modified at C-1 or C-2 as oligosaccharide acceptors. Carbohydr Res 339:1403–1406

116. Osumi K, Makino Y, Akaike E, Yamanoi T, Mizuno M, Noguchi M, Inazu T, Yamamoto K, Fujita K (2004) Mucor hiemalis endo-β-N-acetylglucosaminidase can transglycosylate a bisecting hybrid-type oligosaccharide from an ovalbumin glycopeptide. Carbohydr Res 339:2633–2635

117. Haneda K, Takeuchi M, Tagashira M, Inazu T, Toma K, Isogai Y, Hori M, Kobayashi K, Takeuchi M, Takegawa K, Yamamoto K (2006) Chemo-enzymatic synthesis of eel calcitonin glycosylated at two sites with the same and different carbohydrate structures. Carbohydr Res 341:181–190

118. Makimura Y, Watanabe S, Suzuki T, Suzuki Y, Ishida H, Kiso M, Katayama T, Kumagai H, Yamamoto K (2006) Chemoenzymatic synthesis and application of a sialoglycopolymer with a chitosan backbone as a potent inhibitor of human influenza virus hemagglutination. Carbohydr Res 341:1803–1808

119. Yamanoi T, Yoshida N, Oda Y, Akaike E, Tsutsumida M, Kobayashi N, Osumi K, Yamamoto K, Fujita K, Takahashi K, Hattori K (2005) Synthesis of mono-glucose-branched cyclodextrins with a high inclusion ability for doxorubicin and their efficient glycosylation using Mucor hiemalis endo-β-N-acetylglucosaminidase. Bioorg Med Chem Lett 15:1009–1013

120. Tomabechi Y, Inazu T (2011) Preparation of pseudo glycoamino acid and its application to glycopeptide synthesis. Tetrahedron Lett 52:6504–6507

121. Takegawa K, Nakoshi M, Iwahara S, Yamamoto K, Tochikura T (1989) Induction and purification of endo-β-N-acetylglucosaminidase from arthrobacter protophormiae grown in ovalbumin. Appl Environ Microbiol 55:3107–3112

122. Fan J-Q, Huynh LH, Reinhold BB, Reinhold VN, Takegawa K, Iwahara S, Kondo A, Kato I, Lee YC (1996) Transfer of Man9GlcNAc tol-fucose by endo-β-N-acetylglucosaminidase from arthrobacter protophormiae. Glycoconjugate J 13:643–652

123. Takegawa K, Tabuchi M, Yamaguchi S, Kondo A, Kato I, Iwahara S (1995) Synthesis of neoglycoproteins using oligosaccharide-transfer activity with endo-β-N-acetylglucosaminidase. J Biol Chem 270:3094–3099

124. Fan J-Q, Quesenberry MS, Takegawa K, Iwahara S, Kondo A, Kato I, Lee YC (1995) Synthesis of Neoglycoconjugates by Transglycosylation with Arthrobacter protophormiae endo-β-N-acetylglucosaminidase: demonstration of a macro-cluster effect for mannose-binding proteins. J Biol Chem 270:17730–17735

125. Fan J-Q, Takegawa K, Iwahara S, Kondo A, Kato I, Abeygunawardana C, Lee YC (1995) Enhanced transglycosylation activity of arthrobacter protophormiae Endo-β-N-acetylglucosaminidase in media containing organic solvents. J Biol Chem 270:17723–17729

126. Takegawa K, Fujita K, Fan J-Q, Tabuchi M, Tanaka N, Kondo A, Iwamoto H, Kato I, Lee YC, Iwahara S (1998) Enzymatic synthesis of a neoglycoconjugate by transglycosylation with arthrobacter endo-β-N-acetylglucosaminidase: a substrate for colorimetric detection of endo-β-N-acetylglucosaminidase activity. Anal Biochem 257:218–223

127. Fujita K, Tanaka N, Sano M, Kato I, Asada Y, Takegawa K (2000) Synthesis of neoglycoenzymes with homogeneous Nlinked oligosaccharides using Immobilized endo-β-N-acetylglucosaminidase A. Biochem Biophys Res Commun 267:134–138

128. Takegawa K, Yamabe K, Fujita K, Tabuchi M, Mita M, Izu H, Watanabe A, Asada Y, Sano M, Kondo A, Kato I, Iwahara S (1997) Cloning, sequencing, and expression of arthrobacter protophormiae endo-β-N-acetylglucosaminidase in escherichia coli. Arch Biochem Biophys 338:22–28

129. Muramatsu H, Tachikui H, Ushida H, Song X-J, Qiu Y, Yamamoto S, Muramatsu T (2001) Molecular cloning and expression of endo-β-N-acetylglucosaminidase D, which acts on the core structure of complex type asparagine-linked oligosaccharides. J Biochem 129:923–928

130. Parsons TB, Patel MK, Boraston AB, Vocadlo DJ, Fairbanks AJ (2010) Streptococcus pneumoniae endohexosaminidase D; feasibility of using N-glycan oxazoline donors for synthetic glycosylation of a GlcNAc-asparagine acceptor. Org Biomol Chem 8:1861–1869
131. Fujita K, Takami H, Yamamoto K, Takegawa K (2004) Characterization of endo-β-N-acetylglucosaminidase from alkaliphilic bacillus halodurans C-125. Biosci Biotechnol Biochem 68:1059–1066
132. Tarentino AL, Maley F (1974) Purification and properties of an endo-β-N-acetylglucosaminidase from streptomyces griseus. J Biol Chem 249:811–817
133. Goodfellow JJ, Baruah K, Yamamoto K, Bonomelli C, Krishna B, Harvey DJ, Crispin M, Scanlan CN, Davis BG (2012) An endoglycosidase with alternative glycan specificity allows broadened glycoprotein remodelling. J Am Chem Soc 134:8030–8033
134. Collin M, Olsén A (2001) EndoS, a novel secreted protein from streptococcus pyogenes with endoglycosidase activity on human IgG. EMBO J 20:3046–3055
135. Huang W, Li J, Wang L-X (2011) Unusual transglycosylation activity of flavobacterium meningosepticum endoglycosidases enables convergent chemoenzymatic synthesis of core fucosylated complex N-glycopeptides. ChemBioChem 12:932–941
136. Trimble RB, Tarentino AL (1991) Identification of distinct endoglycosidase (endo) activities in flavobacterium meningosepticum: endo F1, endo F2, and endo F3. Endo F1 and endo H hydrolyze only high mannose and hybrid glycans. J Biol Chem 266:1646–1651
137. Tarentino AL, Plummer TH Jr (1994) [4] Enzymatic deglycosylation of asparagine-linked glycans: purification, properties, and specificity of oligosaccharide-cleaving enzymes from flavobacterium meningosepticum. In: Methods in enzymology, vol 230. Academic Press, Cambridge, pp 44–57
138. Zhao G, Liu Y, Wu Z, Zhu H, Yu Z, Fang J, Wang P (2015) Chemoenzymatic synthesis of glycoproteins. In: Taniguchi N, Endo T, Hart GW, Seeberger PH, Wong C-H (eds) Glycoscience: biology and medicine. Springer, Japan, pp 427–435
139. Wang L-X (2008) Chemoenzymatic synthesis of glycopeptides and glycoproteins through endoglycosidase-catalyzed transglycosylation. Carbohydr Res 343:1509–1522
140. Rising TWDE, Heidecke CD, Moir JWB, Ling ZL, Fairbanks AJ (2008) Endohexosaminidase-catalysed glycosylation with oxazoline donors: fine tuning of catalytic efficiency and reversibility. Chem Eur J 14:6444–6464
141. Rising TWDF, Claridge TDW, Davies N, Gamblin DP, Moir JWB, Fairbanks AJ (2006) Synthesis of N-glycan oxazolines: donors for endohexosaminidase catalysed glycosylation. Carbohydr Res 341:1574–1596
142. Rising TWDF, Claridge TDW, Moir JWB, Fairbanks AJ (2006) Endohexosaminidase M: exploring and exploiting enzyme substrate specificity. ChemBioChem 7:1177–1180
143. Parsons TB, Moir JWB, Fairbanks AJ (2009) Synthesis of a truncated bi-antennary complextype N-glycan oxazoline; glycosylation catalysed by the endohexosaminidases endo A and endo M. Org Biomol Chem 7:3128–3140
144. Huang W, Ochiai H, Zhang XY, Wang LX (2008) Introducing N-glycans into natural products through a chemoenzymatic approach. Carbohydr Res 343:2903–2913
145. Li B, Song HJ, Hauser S, Wang LX (2006) A highly efficient chemoenzymatic approach toward glycoprotein synthesis. Org Lett 8:3081–3084
146. Li B, Zeng Y, Hauser S, Song HJ, Wang LX (2005) Highly efficient endoglycosidase-catalyzed synthesis of glycopeptides using oligosaccharide oxazolines as donor substrates. J Am Chem Soc 127:9692–9693
147. Ochiai H, Huang W, Wang LX (2009) Endo-β-N-acetylglucosaminidase-catalyzed polymerization of β-Glcp-(1->4)-GlcpNAc oxazoline: a revisit to enzymatic transglycosylation. Carbohydr Res 344:592–598
148. Wang LX, Song HJ, Liu SW, Lu H, Jiang SB, Ni JH, Li HG (2005) Chemoenzymatic synthesis of HIV-1 gp41 glycopeptides: effects of glycosylation on the anti-HIV activity and alpha-helix bundle-forming ability of peptide C34. ChemBioChem 6:1068–1074

149. Wei YD, Li CS, Huang W, Li B, Strome S, Wang LX (2008) Glycoengineering of human IgG1-Fc through combined yeast expression and in vitro chemoenzymatic glycosylation. Biochemistry 47:10294–10304

150. Zeng Y, Wang JS, Li B, Hauser S, Li HG, Wang LX (2006) Glycopeptide synthesis through endo-glycosidasecatalyzed oligosaccharide transfer of sugar oxazolines: probing substrate structural requirement. Chem Eur J 12:3355–3364

151. Fujita M, Shoda S-i, Haneda K, Inazu T, Takegawa K, Yamamoto K (2001) A novel disaccharide substrate having 1,2-oxazoline moiety for detection of transglycosylating activity of endoglycosidases. Biochim Biophys Acta Gen Subj 1528:9–14

152. Fairbanks AJ (2011) Endohexosaminidase catalysed glycosylation with oxazoline donors: the development of robust biocatalytic methods for synthesis of defined homogeneous glycoconjugates. C R Chim 14:44–58

153. Heidecke CD, Ling Z, Bruce NC, Moir JWB, Parsons TB, Fairbanks AJ (2008) Enhanced glycosylation with mutants of endohexosaminidase A (endo A). ChemBioChem 9:2045–2051

154. Heidecke CD, Parsons TB, Fairbanks AJ (2009) Endohexosaminidase-catalysed glycosylation with oxazoline donors: effects of organic co-solvent and pH on reactions catalysed by endo A and endo M. Carbohydr Res 344:2433–2438

155. Li B, Zeng Y, Hauser S, Song H, Wang L-X (2005) Highly efficient endoglycosidase-catalyzed synthesis of glycopeptides using oligosaccharide oxazolines as donor substrates. J Am Chem Soc 127:9692–9693

156. Li H, Li B, Song H, Breydo L, Baskakov IV, Wang L-X (2005) Chemoenzymatic synthesis of HIV-1 V3 glycopeptides carrying two N-glycans and effects of glycosylation on the peptide domain. J Org Chem 70:9990–9996

157. Wang L-X, Song H, Liu S, Lu H, Jiang S, Ni J, Li H (2005) Chemoenzymatic synthesis of HIV-1 gp41 glycopeptides: effects of glycosylation on the anti-HIV activity and α-helix bundle-forming ability of peptide C34. ChemBioChem 6:1068–1074

158. Li B, Song H, Hauser S, Wang L-X (2006) A highly efficient chemoenzymatic approach toward glycoprotein synthesis. Org Lett 8:3081–3084

159. Zeng Y, Wang J, Li B, Hauser S, Li H, Wang L-X (2006) Glycopeptide synthesis through endo-glycosidase-catalyzed oligosaccharide yransfer of sugar pxazolines: probing substrate structural requirement. Chem Eur J 12:3355–3364

160. Ochiai H, Huang W, Wang L-X (2008) Expeditious chemoenzymatic synthesis of homogeneous N-glycoproteins carrying defined oligosaccharide ligands. J Am Chem Soc 130:13790–13803

161. Umekawa M, Huang W, Li B, Fujita K, Ashida H, Wang L-X, Yamamoto K (2008) Mutants of mucor hiemalis endo-β-N-acetylglucosaminidase show enhanced transglycosylation and glycosynthaselike activities. J Biol Chem 283:4469–4479

162. Li H, Singh S, Zeng Y, Song H, Wang L-X (2005) Chemoenzymatic synthesis of CD52 glycoproteins carrying native N-glycans. Bioorg Med Chem Lett 15:895–898

163. Wei Y, Li C, Huang W, Li B, Strome S, Wang L-X (2008) Glycoengineering of human IgG1-Fc through combined yeast expression and in vitro chemoenzymatic glycosylation. Biochemistry 47:10294–10304

164. Huang W, Li C, Li B, Umekawa M, Yamamoto K, Zhang X, Wang L-X (2009) Glycosynthases enable a highly efficient chemoenzymatic eynthesis of N-glycoproteins carrying intact natural N-glycans. J Am Chem Soc 131:2214–2223

165. Tomabechi Y, Krippner G, Rendle PM, Squire MA, Fairbanks AJ (2013) Glycosylation of pramlintide: synthetic glycopeptides that display in vitro and in vivo activities as amylin receptor agonists. Chem Eur J 19:15084–15088

166. Priyanka P, Parsons TB, Miller A, Platt FM, Fairbanks AJ (2016) Chemoenzymatic synthesis of a phosphorylated glycoprotein. Angew Chem Int Ed. doi:10.1002/anie.201600817

167. Fujita K, Takegawa K (2001) Tryptophan-216 is essential for the transglycosylation activity of endo-β-N-acetylglucosaminidase A. Biochem Biophys Res Commun 283:680–686

168. Umekawa M, Li CS, Higashiyama T, Huang W, Ashida H, Yamamoto K, Wang LX (2010) Efficient glycosynthase mutant derived from mucor hiemalis endo-b-N-acetylglucosaminidase capable of transferring oligosaccharide from both sugar oxazoline and natural N-glycan. J Biol Chem 285:511–521

169. Fan S-Q, Huang W, Wang L-X (2012) Remarkable transglycosylation activity of glycosynthase mutants of endo-D, an endo-β-N-acetylglucosaminidase from streptococcus pneumoniae. J Biol Chem 287:11272–11281

170. Huang W, Giddens J, Fan S-Q, Toonstra C, Wang L-X (2012) Chemoenzymatic glycoengineering of intact IgG antibodies for gain of functions. J Am Chem Soc 134:12308–12318

171. Parsons TB, Struwe WB, Gault J, Yamamoto K, Taylor TA, Raj R, Wals K, Mohammed S, Robinson CV, Benesch JLP, Davis BG (2016) Optimal synthetic glycosylation of a therapeutic antibody. Angew Chem Int Ed 55:2361–2367

172. Rosati S, van den Bremer ETJ, Schuurman J, Parren PWHI, Kamerling JP, Heck AJR (2013) In-depth qualitative and quantitative analysis of composite glycosylation profiles and other micro-heterogeneity on intactmonoclonal antibodies by high-resolution native mass spectrometry using a modified orbitrap. mAbs 5:917–924

173. Lin C-W, Tsai M-H, Li S-T, Tsai T-I, Chu K-C, Liu Y-C, Lai M-Y, Wu C-Y, Tseng Y-C, Shivatare SS, Wang C-H, Chao P, Wang S-Y, Shih H-W, Zeng Y-F, You T-H, Liao J-Y, Tu Y-C, Lin Y-S, Chuang H-Y, Chen C-L, Tsai C-S, Huang C-C, Lin N-H, Ma C, Wu C-Y, Wong C-H (2015) A common glycan structure on immunoglobulin G for enhancement of effector functions. Proc Nat Acad Sci USA 112:10611–10616

174. Saini KS, Azim HA Jr, Cocorocchio E, Vanazzi A, Saini ML, Raviele PR, Pruneri G, Peccatori FA (2011) Rituximab in Hodgkin lymphoma: is the target always a hit? Cancer Treat Rev 37:385–390

175. Ellis P (2008) Trastuzumab (Herceptin) a treatment for HER2-positive breast cancer. In: Handbook of therapeutic antibodies. Wiley-VCH Verlag GmbH, pp 1109–1130

176. Tomabechi Y, Odate Y, Izumi R, Haneda K, Inazu T (2010) Acceptor specificity in the transglycosylation reaction using endo-M. Carbohydr Res 345:2458–2463

177. Tomabechi Y, Squire MA, Fairbanks AJ (2014) Endo-b-N-acetylglucosaminidase catalysed glycosylation: tolerance of enzymes to structural variation of the glycosyl amino acid acceptor. Org Biomol Chem 12:942–955

178. Boltje TJ, Buskas T, Boons G-J (2009) Opportunities and challenges in synthetic oligosaccharide and glycoconjugate research. Nat Chem 1:611–622

179. Noguchi M, Tanaka T, Gyakushi H, Kobayashi A, Shoda S-I (2009) Efficient synthesis of sugar oxazolines from unprotected N-acetyl-2-amino sugars by using chloroformamidinium reagent in water. J Org Chem 74:2210–2212

180. Watt GM, Boons G-J (2004) A convergent strategy for the preparation of N-glycan core di-, tri-, and pentasaccharide thioaldoses for the site-specific glycosylation of peptides and proteins bearing free cysteines. Carbohydr Res 339:181–193

181. Seko A, Koketsu M, Nishizono M, Enoki Y, Ibrahim HR, Juneja LR, Kim M, Yamamoto T (1997) Occurrence of a sialylglycopeptide and free sialylglycans in hen's egg yolk. Biochim Biophys Acta Gen Subj 1335:23–32

182. Evers DL, Hung RL, Thomas VH, Rice KG (1998) Preparative purification of a High-mannose typeN-glycan from soy bean agglutinin by hydrazinolysis and tyrosinamide derivatization. Anal Biochem 265:313–316

183. Liener IE (1955) The photometric determination of the hemagglutinating activity of soyin and crude soybean extracts. Arch Biochem Biophys 54:223–231

184. Wang L-X, Ni J, Singh S, Li H (2004) Binding of high-mannose-type oligosaccharides and synthetic oligomannose clusters to human antibody 2G12. Cell Chem Biol 11:127–134

185. Specker D, Wittmann V (2007) Synthesis and application of glycopeptide and glycoprotein mimetics. In: Wittmann V (ed) Glycopeptides and glycoproteins: synthesis, structure, and application. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 65–107

186. Tornoe CW, Christensen C, Meldal M (2002)Peptidotriazoles on solid phase: [1,2,3]-triazoles by regiospecific copper(I)-catalyzed 1,3-dipolar cycloadditions of terminal alkynes to azides. J Org Chem 67:3057–3064

187. Rostovtsev VV, Green LG, Fokin VV, Sharpless KB (2002) A stepwise Huisgen cycloaddition process: Copper(I)-catalyzed regioselective "ligation" of azides and terminal alkynes. Angew Chem Int Ed 41:2596–2599

188. Trabocchi A, Guarna A (2014) Click chemistry: the triazole ring as a privileged peptidomimetic scaffold. In: Peptidomimetics in organic and medicinal chemistry. Wiley, New Jersey, pp 99–121

189. Miller N, Williams GM, Brimble MA (2010) Synthesis of fish antifreeze neoglycopeptides using microwave-assisted "Click Chemistry". Org Lett 12:1375–1376

190. Wojnar JM, Lee DJ, Evans CW, Mandal K, Kent SBH, Brimble MA (2013) Neoglycoprotein synthesis using the copper-catalyzed azide–alkyne click reaction and native chemical ligation. In: Click chemistry in glycoscience. Wiley, New Jersey, pp 251–270

191. Lim D, Brimble MA, Kowalczyk R, Watson AJA, Fairbanks AJ (2014) Protecting-group-free one-pot synthesis of glycoconjugates directly from reducing sugars. Angew Chem Int Ed 53:11907–11911

192. Grunberger GJ (2013) Novel therapies for the management of type 2 diabetes mellitus: part 1. Pramlintide and bromocriptine-QR. J Diabetes 5:110–117

193. Younk LM, Mikeladze M, Davis SN (2011) Pramlintide and the treatment of diabetes: a review of the data since its introduction. Expert Opin Pharmacother 12:1439–1451

194. Young AA, Vine W, Gedulin BR, Pittner R, Janes S, Gaeta LSL, Percy A, Moore CX, Koda JE, Rink TJ, Beaumont K (1996) Preclinical pharmacology of pramlintide in the rat: Comparisons with human and rat amylin. Drug Dev Res 37:231–248

195. Roberts AN, Leighton B, Todd JA, Cockburn D, Schofield PN, Sutton R, Holt S, Boyd Y, Day AJ, Foot EA, Willis AC, Reid KBM, Cooper GJS (1989) Molecular and functional-characterization of amylin, a peptide associated with type-2 diabetes-mellitus. Proc Nat Acad Sci USA 86:9662–9666

196. Maruyama K, Nagasawa H, Suzuki A (1999) 2,2′-bispyridyl disulfide rapidly induces intramolecular disulfide bonds in peptides. Peptides 20:881–884

197. Inazu T, Kobayashi K (1993) A new simple method for the synthesis of N(α)-Fmoc-N(b)-glycosylated-l-asparagine derivatives. Synlett, 869–870

198. Katayama H, Asahina Y, Hojo H (2011) Chemical synthesis of the S-linked glycopeptide, sublancin. J Pept Sci 17:818–821

199. Umekawa M, Higashiyama T, Koga Y, Tanaka T, Noguchi M, Kobayashi A, Shoda S, Huang W, Wang LX, Ashida H, Yamamoto K (2010) Efficient transfer of sialo-oligosaccharide onto proteins by combined use of a glycosynthase-like mutant of mucor hiemalis endoglycosidase and synthetic sialo-complex-type sugar oxazoline. Biophys Acta Gen Subj 1800:1203–1209

200. Hay DL, Christoulos G, Christopoulos A, Poyner DR, Sexton PM (2005) Pharmacological discrimination of calcitonin receptor: receptor activity-modifying protein complexes. Mol Pharmacol 67:1655–1665

201. Tilakaratne N, Christopoulos G, Zumpe ET, Foord SM, Sexton PM (2000) Amylin receptor phenotypes derived from human calcitonin receptor/RAMP coexpression exhibit pharma-cological differences dependent on receptor isoform and host cell environment. J Pharmacol Exp Ther 294:61–72

202. Gingell JJ, Burns ER, Hay DL (2014) Activity of pramlintide, rat and human amylin but not Ab1-42 at human amylin receptors. Endocrinology 155:21–26

203. Vilaseca M, Nicolas E, Capdevila F, Giralt E (1998) Reduction of methionine sulfoxide with NH4I/TFA: compatibility with peptides containing cysteine and aromatic amino acids. Tetrahedron 54:15273–15286

204. Wang Z, Chinoy ZS, Ambre SG, Peng W, McBride R, de Vries RP, Glushka J, Paulson JC, Boons G-J (2013) A general strategy for the chemoenzymatic synthesis of asymmetrically branched N-glycans. Science 341:379–383

Synthetic Antitumor Vaccines Through Coupling of Mucin Glycopeptide Antigens to Proteins

Markus Glaffig and Horst Kunz

Abstract The requirements for coupling reactions of carbohydrate molecules very much depend upon the biological recognition processes that should be investigated and upon the target structures of the desired carbohydrate ligand. If the carbohydrate conjugate itself is the recognized ligand, as for example, the binding site of a P-selectin ligand comprising sialyl-LewisX and a specific peptide sequence, the natural glycoside bond must be installed. A stereoselective and regioselective block glycosylation between a sialyl-LewisX trichloroacetimidate and a partially deprotected Thomsen–Friedenreich antigen derivative was developed to achieve this aim. In contrast, the coupling reactions by which glycopeptides from tumor-associated glycoproteins are conjugated to immune stimulating components in order to afford efficient vaccines can entail artificial linkages as long as they do not interfere with the immune reactions. For example, the coupling of glycophorin glycopeptides to bovine serum albumin was successfully achieved by carboxylic activation with a water-soluble carbodiimide in the presence of a supernucleophilic additive. This conjugation method is only recommendable if the glycopeptide does not contain several carboxylic and/or amino functions. The photochemically or radical initiator promoted thiol-ene coupling succeeded in couplings of MUC1 glycopeptide antigens to bovine serum albumin, however, is accompanied by oxidative disulfide formation. The conjugation of glycopeptide antigens from the tandem repeat region of the tumor-associated mucin MUC1 to bovine serum albumin or tetanus toxoid is efficiently accomplished using diethyl squarate as the coupling reagent. The intermediate squaric monoamide esters can be isolated and characterized, and then applied to a mild connecting process to the carrier proteins. The MUC1 glycopeptide-tetanus toxoid conjugates proved to be particularly useful vaccines. They induce extraordinarily strong immune responses in mice. The induced antibodies are prevailingly of the IgG1 isotype and show efficient binding to the glycoproteins exposed on epithelial tumor cells.

M. Glaffig · H. Kunz (✉)
Institut Für Organische Chemie, Johannes Gutenberg-Universität Mainz,
55099 Mainz, Germany
e-mail: hokunz@uni-mainz.de

1   Introduction

Since the discovery of the human blood group substances by Karl Landsteiner in 1901 [1], researchers have always been strongly interested in the question which functions do carbohydrates have in natural glycoconjugates, in particular in glycoproteins and glycolipids. For investigations of the role of the saccharide portions of natural glycoconjugates, model compounds are considered helpful in which a carbohydrate is linked to a suitable component in a well-known format. From interactions of these model conjugates with natural receptors or enzymes, conclusions can hopefully be drawn which shine light on the recognition processes amounting to the biological selectivity. Efficient coupling reactions often are a crucial prerequisite for the synthesis of these model glycoconjugates. The choice of the coupling method distinctly depends upon the nature of the functional molecule to which the carbohydrate is to be conjugated and upon the role, the carbohydrate does play in the considered biological process.

If the glycan merely influences the physicochemical properties of the glycoconjugate or protects it from degradation, an arsenal of coupling reactions can be exploited for forming a mimic of the natural prototype. These coupling processes can involve artificial linker structures. The same often applies to glycoconjugates in which the carbohydrate constitutes the solely recognized ligand, as for example, a blood group antigen [2], a bacterial cell wall saccharide antigen [3], or as the ligand of a lectin [4–6]. A number of efficient ligation reactions have been introduced during the past decade which can be applied to the decoration of peptides or proteins with carbohydrates through unnatural linkage structures, as for example, by thiol-ene couplings [7], cross-metathesis reactions [8], 1,3-dipolar cycloaddition (click) reactions [9], or Diels–Alder cycloaddition reactions with inverse electron demand [10]. The situation, however, is completely different if the carbohydrate constitutes only a part of the structure recognized by a receptor, enzyme or antibody, and the bound epitope comprises of both, carbohydrate and backbone structures to which the carbohydrate is conjugated. This also includes glycoprotein epitopes in which the carbohydrate exerts distinct influence onto the conformation of a recognized epitope even if this is a pure peptide structure [11–16]. In these cases, the artificial linker disarranges the recognized epitope, and observed effects may result in misleading conclusions.

Epitopes involving both peptide and carbohydrate structures are known from ligands of selectins [17] and tumor-associated mucin antigens [18]. In syntheses of model compounds representing these types of glycopeptide epitopes, the linkages between the carbohydrates and the peptide backbone should closely correspond to the nature prototype.

2 Glycan Coupling Through Glycosylation—Glycopeptide Ligands of Selectins

Selectins are important carbohydrate recognizing receptors involved in the early phase of recruitment of leukocytes into inflamed tissues [19]. P- and E-selectins are expressed on the apical surface of endothelial cells of the blood vessels within the inflamed area, while L-selectin is exposed on the leukocyte. The tetrasaccharide sialyl LewisX has been described as a ligand to these selectins [20]. Initial inter-action between P-selectin and its glycoprotein ligands occurs immediately after activation of P-selectin through inflammatory cytokines and results in rolling of the leukocytes. The expression of E-selectin requires several hours since the cells need to synthesize this glycoprotein after activation. Its interaction with the E-selectin ligands on the leucocytes results in a closer attraction of the rolling leukocytes to the endothelial cells [21]. The selective inhibition of these cell adhesion processes certainly is important for the treatment of inflammatory diseases and also for the prevention of tumor cell metastasis [22]. It was revealed that sialyl LewisX is a required, but only moderate ligand to P- and E-selectin. For an intensive binding, the cooperation of peptide segments of the glycoprotein ligands carrying the sialyl LewisX is mandatory. In the case of the P-selectin glycoprotein ligand 1 (PSGL-1) the *N*-terminal peptide, a sequence containing *O*-sulfated tyrosine residues (Fig. 1) was identified as an important ligand structure [23].

In fact, it was found that synthetic sialyl-LewisX glycopeptides with peptide sequences containing charged amino acids bind distinctly stronger to P-selectin than glycopeptides lacking charged groups [24, 25]. A sialyl-LewisX-RGD-peptide, for example, inhibited the binding of HL-60 tumor cell to an immobilized human P-selectin-IgG [24] with an IC_{50} value of 26 µmol. In this glycopeptide, the sialyl-LewisX tetrasaccharide was *N*-glycosidically coupled directly to *N*–4 of an asparagine. Thus, the distance between the saccharide ligand and the charged units did not correspond to the natural prototype (Fig. 1).

In order to obtain more active inhibitors of P-selectin, sialyl-LewisX trifluoroacetimidate **1** and the *N*-fluorenylmethoxycarbonyl-(Fmoc)-protected T-antigen threonine conjugate **2** [26], selectively deblocked in 4- and 6-position of the galactosamine unit were ligated by the activation of the trichloroacetimidate
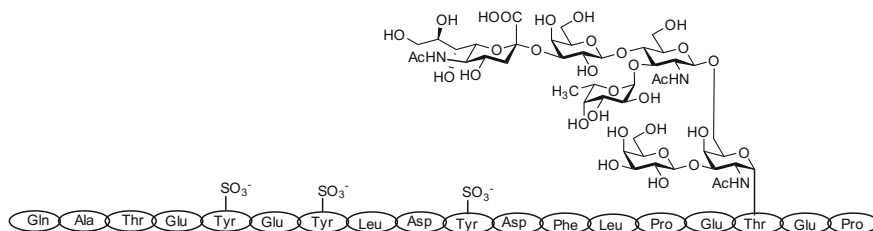

Fig. 1 P-selectin binding region of human P-selectin glycoprotein ligand 1 (Ref. [23] and lit. cited therein.)

with trimethylsilyl-trifluoromethanesulfonate (TMS-triflate) at low temperature. This coupling proceeded with high overall yield and high stereoselectivity but unexpected regioselectivity. The desired product of glycosylation at the primary 6-OH group **3** was the minor product. Fortunately, it was separable from the undesired regioisomer **3a** by flash-chromatography [27] and, thus, applicable to the solid-phase synthesis of a PSGL-1 binding domain glycopeptide (Scheme 1). In conjugate **3,** the natural linkage between sialyl-LewisX as a part of the ligand structure and the peptide backbone represented by the threonine is established, and thus the natural distance and orientation to the binding sites within a peptide sequence are guaranteed.

Alternatively, the mimic **4** of sialyl-LewisX containing cyclohexyl-lactic acid as the surrogate of sialic acid [28] was coupled to the partially deprotected T-antigen-threonine building block **2** [26].

After activation of the pseudo-tetrasaccharide trichloroacetimidate **4** with TMS-triflate, the β-glycoside bond to the acceptor was formed with high stereo-electivity and excellent overall yield. Again, the regioselective differentiation between the primary 6- and the secondary 4-hydroxy groups of the T-antigen threonine derivative was only moderate. However, in this case, the desired compound **5** was the major product. Flash-chromatographic separation gave the desired regioisomer **5** of the glycosyl amino acid in a yield of 55%, while the unexpected regioisomer **5a** was isolated in a yield of 33% (Scheme 2) [29].

The building block **5** after selective cleavage of the *tert*-butyl ester was introduced into the solid-phase synthesis of the PSGL-1 glycopeptide recognition domain. Acidolytic detachment of the assembled glycopeptide from the resin and
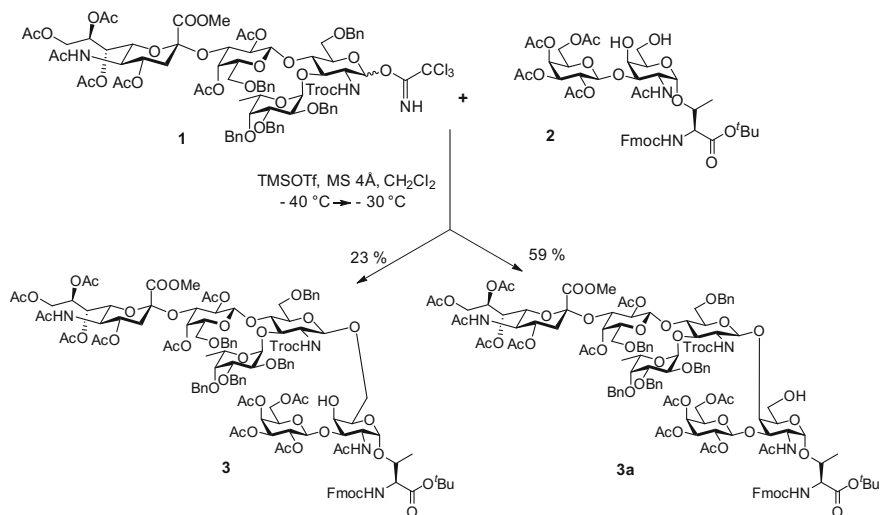
Scheme 1 Coupling of sialyl-LewisX to a T-antigen-threonine to install the natural β-glycoside linkage occurring in the PSGL-1 recognition domain [27]
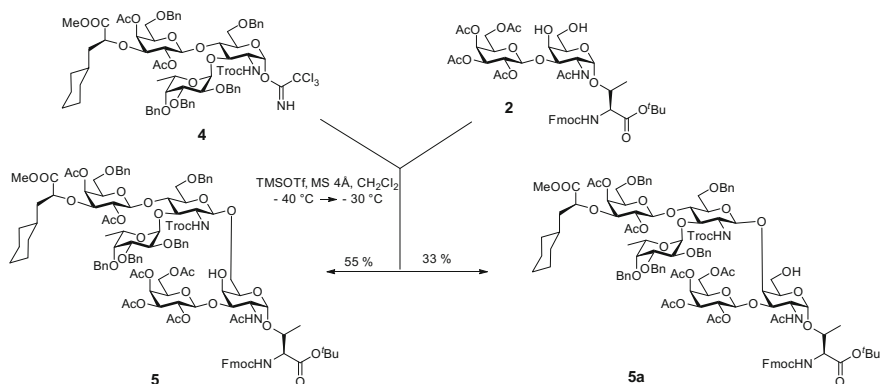
Scheme 2 Coupling of saccharide building blocks by stereoselective glycosylation to give the pseudo-sialyl-LewisX glycosyl threonine **5** [28]

concomitant removal of all acid-sensitive protecting groups from the amino acid side chains afforded the PSGL-1 glycopeptide binding sequence **6** still protected in the carbohydrate portion. Treatment of **6** with pyridine-SO$_3$ complex in pyridine/dimethylformamide, first at 0 °C, then at room temperature for 18 h, afforded the glycopeptide *O*-sulfated in all three tyrosine residues. After subsequent careful removal of the carbohydrate protecting groups under mild basic conditions preventing β-elimination of the carbohydrate, the sulfated PSGL-1 glycopeptide binding domain **7** was isolated by preparative RP-HPLC in pure form (Scheme 3). This compound inhibited the binding of murine granulocytes to a murine P-selectin-IgG fusion protein with an IC$_{50}$ of 20 µmol. The corresponding binding of human granulocytes was inhibited by **7** even more efficiently (IC$_{50}$ 5 µmol). This is not surprising since **7** represents the peptide sequence of human PSGL-1.

The potentiation of the binding effects of both, the pseudo-tetrasaccharide and the anionic centers, was only achieved with an inhibitor which contains the binding groups in optimal distance and steric arrangement. The coupling of the saccharide ligand to the peptide backbone in the natural structure is considered the prerequisite for this efficiency. This requirement would also be fulfilled for conjugates in which the whole glycopeptide recognition site is linked to a carrier molecule, for example, a protein, a dendrimer or a polymer.

3 Glycophorin-Derived Vaccines Containing Nature-Like Linkages

Small molecules and endogenous structures often are not sufficiently immunogenic in order to elicit an appropriate immune response. To obtain a vaccine, these compounds need to be coupled to immune stimulating components. A most general
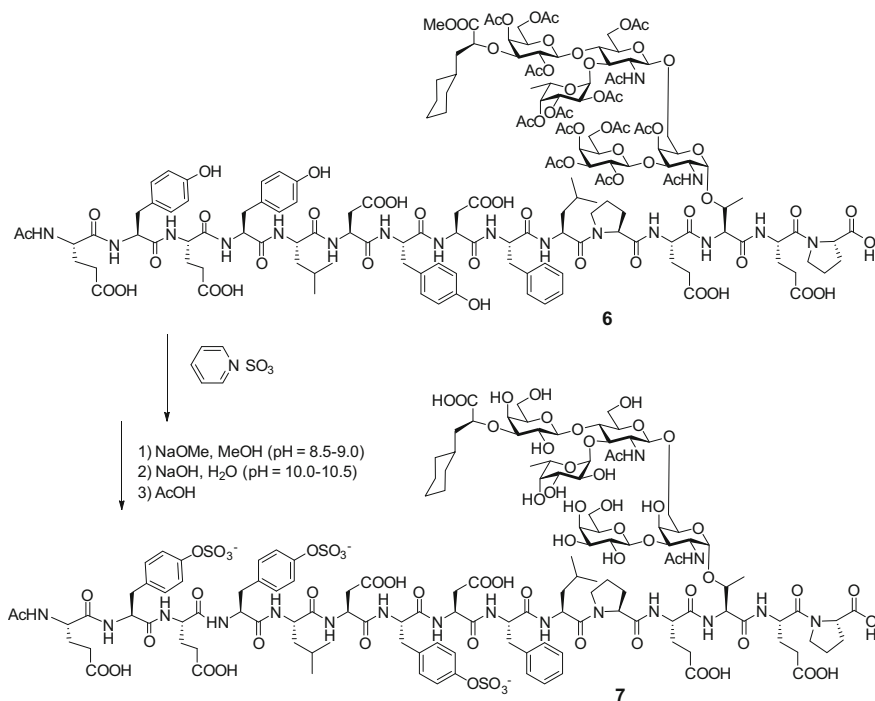
Scheme 3 Tyrosine *O*-sulfation and deprotection of the PSGL-1 glycopeptide recognition domain [29]

way for the construction of a vaccine consists of the conjugation to a carrier protein which provides T-cell epitopes that induce T-cell activation via major histocompatibility complex (MHC) pathways.

Inspired by reports of Springer [30] on the Thomsen-Friedenreich (T-)antigen as a tumor-associated carbohydrate antigen, we started the chemical synthesis of T- and Tn-antigen carrying glycopeptides in the early 80s. Springer and his coworkers had identified tumor-associated T-antigen containing glycoproteins on the membranes of epithelial tumor cells, while these molecules were absent on epithelial cells of normal tissues. The authors isolated these tumor-associated T-antigen and Tn-antigen glycoproteins from tumor cells and induced antibodies against these molecules. From cross reactivity of the obtained antibodies, they concluded that the tumor-associated T- and Tn-antigen glycoproteins should be structurally related to asialoglycophorin, the sialidase-treated form of glycophorin which is the major transmembrane glycoprotein on the red blood cells. Glycophorin occurs in two blood group forms M and N, which are decorated with the same *O*-glycan pattern, but are different in only two of the 132 amino acids (position 1 and 5). One of the differences concerns the N-terminal amino acid which serine in M-blood group and leucine in N-blood group glycophorin [30].

With the aim of constructing a vaccine against the tumor-associated epithelial glycoproteins, we synthesized the *N*-terminal glyco-tripeptide of M-blood group asialoglycophorin **8** [31]. The coupling of this glycopeptide to bovine serum albumin (BSA) as the carrier protein in water was achieved using the water-soluble *N*-ethyl-*N*′-(3-dimethylaminopropyl) carbodiimide (EDC) and 1-hydroxy-benzotriazole (HOBt, Scheme 4) [32].

This coupling process has become a reliable method for linking antigens to proteins. Instead of HOBt, *N*-hydroxy-succinimide (HO–Su or NHS) can be used as an additive for the conjugation of (glyco)peptides [33] as well as of small molecules, as for example, the designer opioid fentanyl [34]. The procedure takes profit from a remarkable chemoselectivity. In water at pH 6, the carboxy group of **8** reacts with EDC to furnish the corresponding *O*-acyl isourea which preferentially is attacked by the super-nucleophile HOBt to form the active ester **9** (Scheme 4). This active ester obviously is sufficiently resistant to hydrolysis, and therefore prevailingly undergoes aminolysis by the amino functions of the lysine side chains of BSA to afford the glycopeptide–protein conjugate **10**.

By immunizing mice with blood group M-asialoglycophorin-BSA vaccine **10** and cloning, a monoclonal antibody (82-A6, IgM subclass) was obtained [35]. Monoclonal antibody 82-A6 reacted with normal and tumor epithelial cells showing its binding to the T-antigen disaccharide. However, the monoclonal antibody 82-A6 exhibited distinctly higher affinity towards M-blood group asialoglycophorin having the identical *N*-terminal tripeptide sequence Ser-Ser(T-antigen)-Thr(T-antigen) as **10** than to N-blood group asialoglycophorin with the *N*-terminal sequence Leu-Ser(T-antigen)-Thr(T-antigen) (Fig. 2) [33b]. It appeared amazing that an antibody induced with a synthetic vaccine presenting only the small *N*-terminal part of the huge M-blood asialogroup glycophorin (132 amino acids) differentiates between the two M- and N-blood group asialoglycophorins which carry the
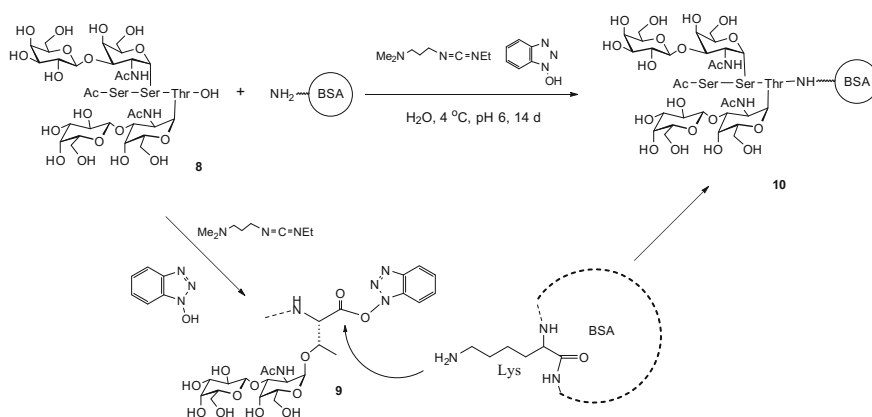
Scheme 4 Coupling of a T-antigen glycopeptide to bovine serum albumin in water through carboxy activation with a water-soluble carbodiimide/1-hydroxy-benzotriazole [31]
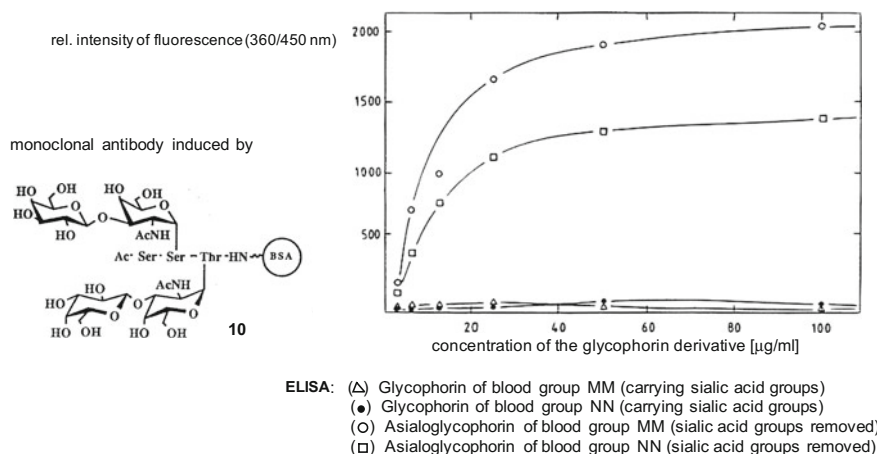
rel. intensity of fluorescence (360/450 nm)

monoclonal antibody induced by

10

concentration of the glycophorin derivative [µg/ml]

ELISA: (△) Glycophorin of blood group MM (carrying sialic acid groups)
(●) Glycophorin of blood group NN (carrying sialic acid groups)
(○) Asialoglycophorin of blood group MM (sialic acid groups removed)
(□) Asialoglycophorin of blood group NN (sialic acid groups removed)

Fig. 2 Recognition of M- and N-blood group asialoglycophorin by monoclonal IgM antibody induced in mice by vaccine **10**

identical glycan pattern (Fig. 2). This result suggested that not only the recognized saccharide (T-antigen) but also the amino acid sequence contributes to the recognized epitope. Nature obviously does not make differences between classes of natural products as they are organized in chapters of scientific text books, if selective recognition by the immune system is concerned.

It must also be concluded from the observed differentiation between M- and N-blood group asialoglycophorin by antibody 82-A6 induced through the glycopeptide vaccine that the tumor-associated carbohydrate antigen in a vaccine is not sufficient for the induction of tumor-selective antisera. It needs to be combined with a tumor-relevant peptide sequence in order to form a tumor-typical glycopeptide antigen. The peptide sequence of glycopeptide **8** obviously did not meet these criteria.

Biochemical and molecularbiological analyses of membrane glycoproteins from carcinoma cells, in particular, reported by the group of J Taylor-Papadimitriou [36], have revealed that the tumor-associated mucin MUC1 is a characteristic membrane glycoprotein occurring on many epithelial tumors.

4 The Tumor-Associated Mucin MUC1—A Promising Target for the Development of Antitumor Vaccines and Synthesis of Tumor-Associated MUC1 Glycopeptide Antigens

Mucin MUC1 is expressed on many epithelial tissues [37]. It is a large membrane-bound glycoprotein. In its extracellular portion, it contains an extended domain comprising a variable number (20–125) of tandem repeats of the amino acid

sequence HGVTSAPDTRPAPGSTAPPA [38]. The serines and threonines within these tandem repeats are the major glycosylation sites of this heavily glycosylated glycoprotein. MUC1 is over expressed in most of the epithelial tumors and characterized by a dramatically altered glycosylation (Fig. 3) [39]. In carcinoma cells, a glucosaminyl transferase C2GnT1 which transfers *N*-acetylglucosamine to O-6 of the galactosamine of the T-antigen (core 1) structure often is strongly downregulated. Thus, the formation of the branched core 2 trisaccharide [40] is suppressed. Therefore, the elongation of the oligosaccharides, which depends on the core 2 structure, cannot proceed. In addition, sialyltransferases are strongly upregulated in carcinomas [41].

As a result, tumor-associated MUC1 differs from MUC1 on normal epithelial cells by short, prematurely sialylated carbohydrate side chains. Because of these short glycan side chains, epitopes residing in the peptide backbone of tumor-associated MUC1, in particular within the tandem repeat region, are now accessible for the immune system, whereas these peptide sequences are shielded by the large glycan side chains of MUC1 on normal cells. As a consequence, tumor-selective glycopeptide antigens should be obtained by combination of tumor-associated saccharide antigens, as for example the T-antigen, with peptide sequences of the tandem repeat region of MUC1. Such molecules cannot be isolated
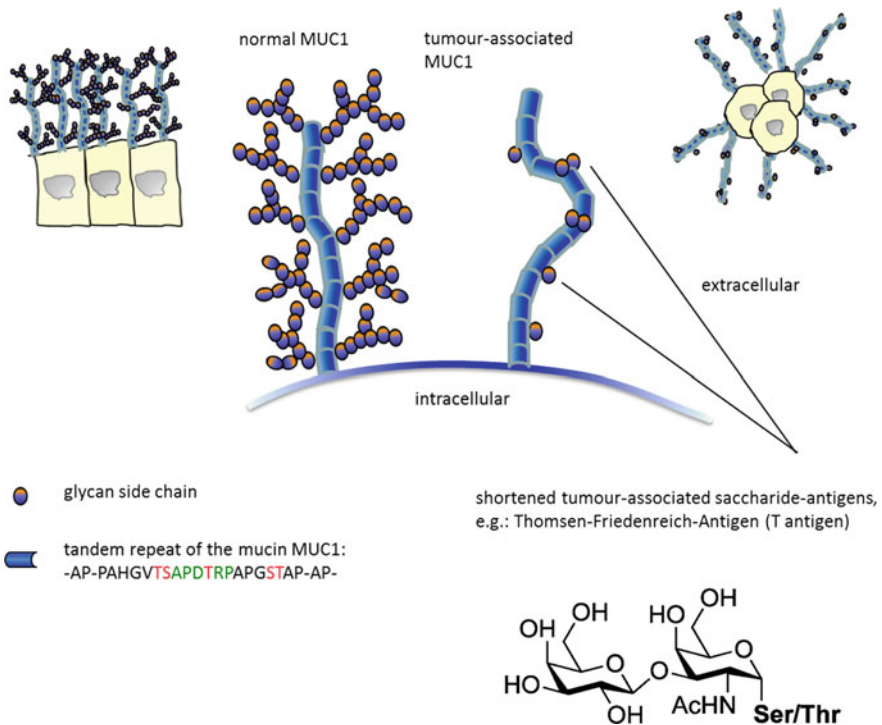
Fig. 3 Structural differences between MUC1 on normal and on epithelial tumor cells

from tumor cells. Because of the biological micro-heterogeneity, MUC1 molecules isolated from tumor cell membranes can carry long oligosaccharides typical for normal cells next to short, tumor-associated glycans even on a single protein chain. Therefore, sufficiently pure tumor-typical MUC1 glycopeptide antigens must be built up by synthesis.

Solid-phase synthesis of glycopeptides (SPPS) using glycosylated amino acid building blocks nowadays is a powerful, flexible method for the construction of exactly specified MUC1 glycopeptide antigens [18, 42]. The required Fmoc-protected glycosyl serine and threonine derivatives are obtained by stereo- and regioselective extension of the saccharide portion of the corresponding galactosamine (Tn-antigen) conjugates, as for example, **11** [43]. Due to the acetamido group and to the *O*-acyl protection, the glycoside bond of glycosylated amino acids and peptides are sufficiently stable to acids (carbonyl oxygens are prevailingly protonated thus creating a Faraday cage that protects the saccharide bonds [44]) the *tert*-butyl ester of **11**, as well as those of other Fmoc glycosyl amino acids, can selectively be cleaved using a cocktail of trifluoroacetic acid (TFA)/triisopropyl-silane(TIS)/water (10:1:1) or TFA/anisole. The extension of the saccharide portion, however, requires methods which do not affect the Fmoc- as well as the *tert*-butyl ester protection nor the glycoside bond (Scheme 5).

In this sense, the selective removal of the *O*-acetyl groups from **11** was achieved by transesterification in dry methanol at a pH of 8.5 to give the key substrate **12** for
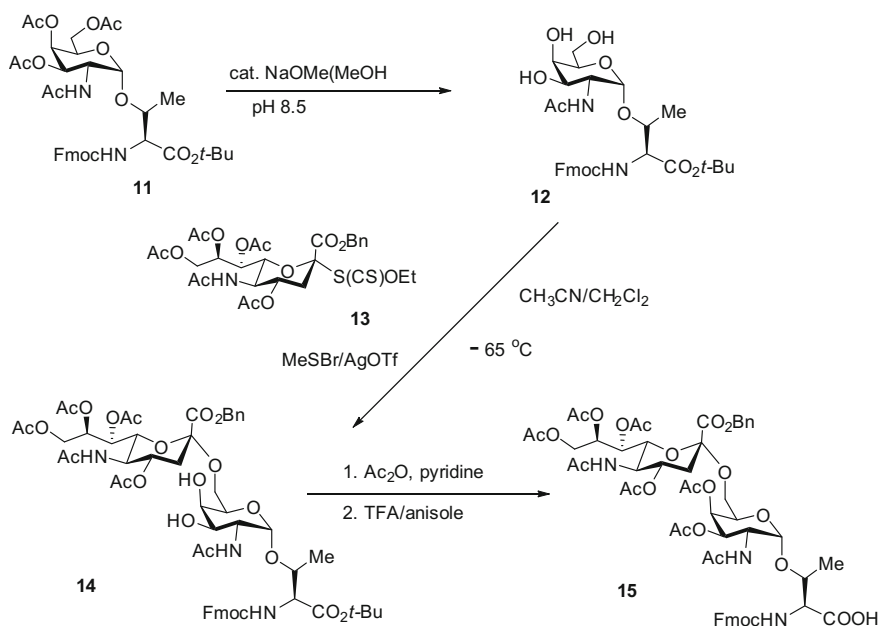

Scheme 5 Synthesis of Fmoc-protected glycosyl amino acid building blocks [18, 42]: Sialyl-Tn threonine

further glycosylation reactions [43, 45]. Regio- and stereoselective sialylation at O-6 of **12** using the sialyl ester xanthate **13** afforded the sialyl-Tn antigen-threonine derivative **14** which after acetylation and acidolytic removal of the *tert*-butyl ester furnished sialyl-Tn building block **15** applicable to solid-phase glycopeptide syntheses (Scheme 5). Further extensions of the carbohydrate to give T-antigen and sialyl-T-antigen building blocks are displayed in Refs. [18, 42]. Acid-catalyzed introduction of a 4,6-benzylidene protection to furnish **16** followed by galactosylation under Helferich conditions and careful methanolysis of the *O*-acetal groups resulted in the formation of the T-antigen structure **17** (Scheme 6). If Fmoc protection is lost under these conditions, it can selectively be reintroduced using *O*-Fmoc-hydroxy-succinimide.

Regioselective sialylation to give the 2,3-sialyl-T antigen threonine derivative **18** was achieved under the conditions described above [43]. *O*-Acetylation and final acidolysis of the *tert*-butyl ester yielded the Fmoc-sialyl-T-threonine buiding block **19** [46].

The solid-phase synthesis of the MUC1 glycopeptide antigens are usually conducted using Tentagel-resins equipped with trityl- or 2-chlorotrityl linkers according to the Fmoc strategy. Example **20**, in which the sequence is *N*-terminally extended with a triethylene glycol spacer and the tumor-associated carbohydrate antigen sialyl-Tn is incorporated at threonine-6 is shown in Scheme 7 [47]. The trityl anchor prevents the formation of a diketopiperazine on the level of the resin-linked dipeptide. The coupling of the glycosyl amino acid, because of its demanding preparations applied in only slight excess, was carried out manually using the more reactive *O*-(7-azabenzotrazolyl)-*N*,*N*,*N'*,*N'*-tetramethyluronium hexafluorophosphate (HATU) [48] instead of *O*-benzotrazolyl)-*N*,*N*,*N'*,*N'*-tetramethyluronium hexafluorophosphate (HBTU). After detachment from resin with
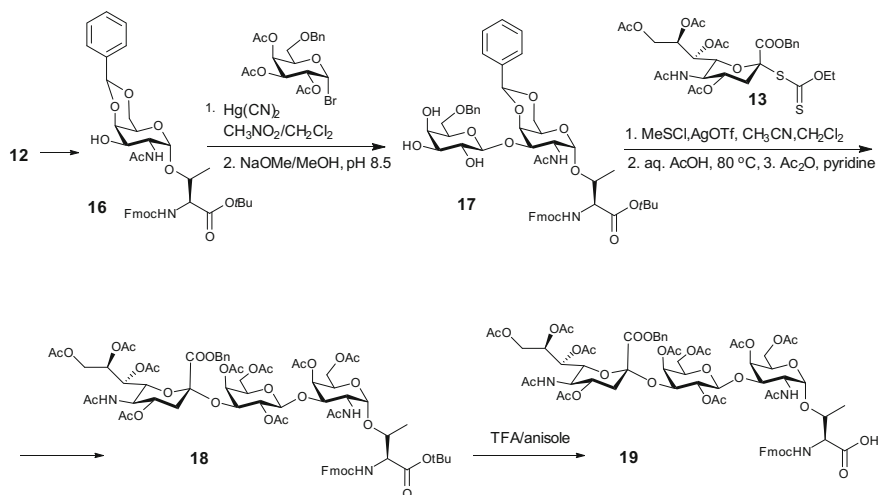
Scheme 6 Synthesis of core 1 di- and trisaccharide threonine building blocks
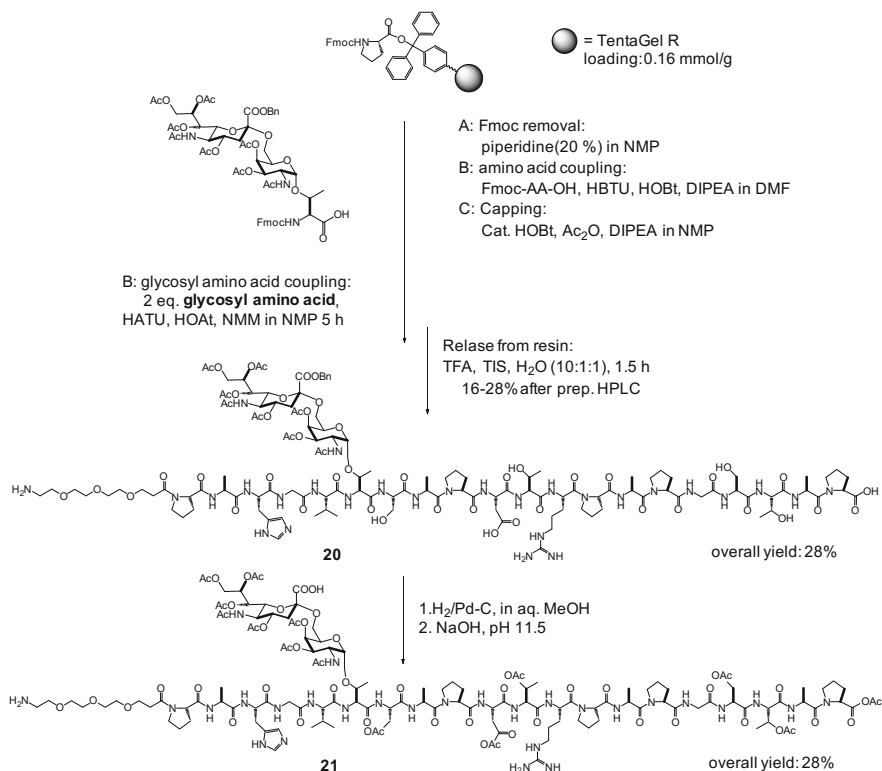
Scheme 7 Solid-phase synthesis of tumor-associated MUC1 glycopeptide antigens [47]

concomitant acidolytic removal of the side chain protection and purification by preparative HPLC glycopeptide **20** was obtained. Hydrogenation of the sialic benzyl ester and careful alkaline hydrolysis of the *O*-acetyl functions at pH 11.5 afforded after purification by preparative HPLC the MUC1 glycopeptide antigen **21** in a 30 mg scale.

The glycopeptide antigens, as for example **21**, represent endogenous structures and therefore, are of insufficient immunogenicity. In order to gain an efficient vaccine, these compounds must be conjugated to immune stimulating components. Most frequently, proteins serve as immune stimulants because they contain several T-cell activating peptide epitopes.

5 Thiol-Ene Coupling for the Formation of Glycopeptide Antigen–Protein Conjugates

The construction of vaccines demands coupling reactions which do not form immunogenic linker structures, as for example, electron-poor homo- or hetero-aromatic systems. Thioether linkages are considered immunologically silent

structures. Given the great number of functional groups in both, the deprotected glycopeptide antigen and the carrier protein to which the glycopeptide should be linked, a biorthogonal thioether formation is highly desirable. Thioether formation via heterobifunctional linkers, as for example N-succinimidyl-4-(maleimidomethyl) cyclohexane carboxylate [49], has frequently been applied for conjugation reactions. However, the immunologically less critical radical-type addition of thiol to nonactivated double bonds was not used for the coupling to proteins up to 2007, although the reaction was known since hundred years [50]. In order to prove whether sensitive amino acid derivatives, as for example, tyrosine **22a** or tryptophane **22b** are affected by photochemically or radical-type initiated thiyl radical addition, model reaction was performed [51] (Scheme 8).

The thiyl radicals generated from methyl N-acetyl cysteinate **23** either photochemically or by initiation with VA 044 as the initiator reacted with the amino acid N-allylamides **22** to furnish the thioether-linked conjugates **24**. The disulfide of **23** was the only side product. The electron-rich aromatic rings and the α-CH positions were not affected.

After these promising results of the model reactions bovine serum albumin (BSA) as the carrier protein was decorated with functional oligoethylene glycol spacer molecules **25** terminating either in S-acetylthio- or allyl ether functions. These reactions (Scheme 9) were carried out at room temperature through acylation at the lysine amino functions using the water-soluble carbodiimide in combination with HOBt or N-hydroxy-succinimide (HOSu) under weakly acidic conditions as described above (Scheme 4) for the glycophorin glycopeptide vaccine.

MALDI mass spectra gave evidence that in the conjugates **26** on average 25 of the 59 lysine residues of BSA were acylated with the functional spacers [51].

The protein molecules **26** decorated with functional side chains were subjected to radical-type thiol-ene coupling with MUC1 glycopeptide antigens. On the one hand, the MUC1 glycopeptide **27** carrying a T-antigen side chain was prepared on solid-phase as described above and N-terminally extended with a mono-allylamido-succinoyl group. On the other hand, the S-acetyl protections were removed from the
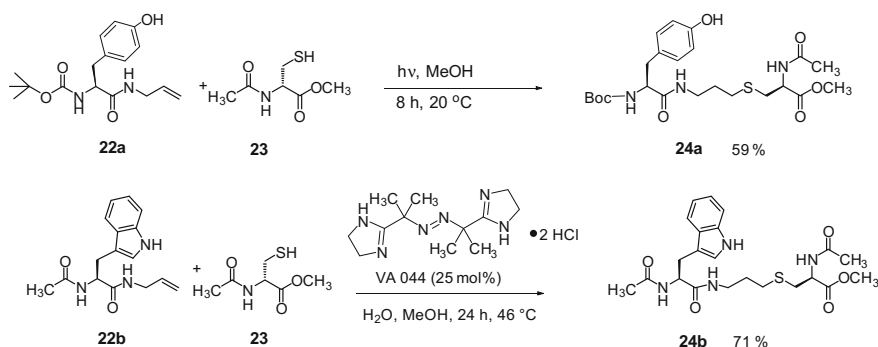
Scheme 8 Radical addition of N-actetyl cysteine methyl ester to N-protected tyrosine and tryptophane N-allylamide
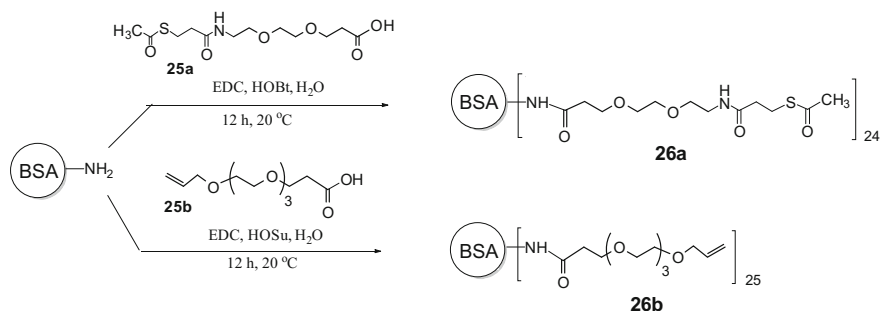
Scheme 9 Decoration of BSA with thiol- and allyl ether functionalized spacer molecules [51]

thiol-functionalized BSA **26a** using hydroxylamine solution under argon atmosphere. The set-free thiols were activated in water with the radical initiator ACVA (Scheme 10) under irradiation (λ = 254 nm) in the presence of **27**. Thus, the two components were linked to each other in the thiol-ene coupling to give the neoglycoprotein conjugate **28a**. According to its MALDI-TOF mass spectrum, conjugate **28a** carried on average 9 molecules glycopeptide per molecule protein.

The yield of the thiol-ene coupling amounts to about 35% [51]. Longer reaction time probably is recommendable.

The alternative coupling between the protein-bound olefin **26b** and the glycopeptide antigen equipped with the *N*-terminal thiol-functionalized spacer is displayed in Scheme 11. Glycopeptide **29** was assembled on solid-phase. The complex 2,3-sialyl-T antigen threonine building block **19** (Scheme 6) was introduced at position Thr-3. Finally, an *S*-acetylated triethylene glycol spacer acid was
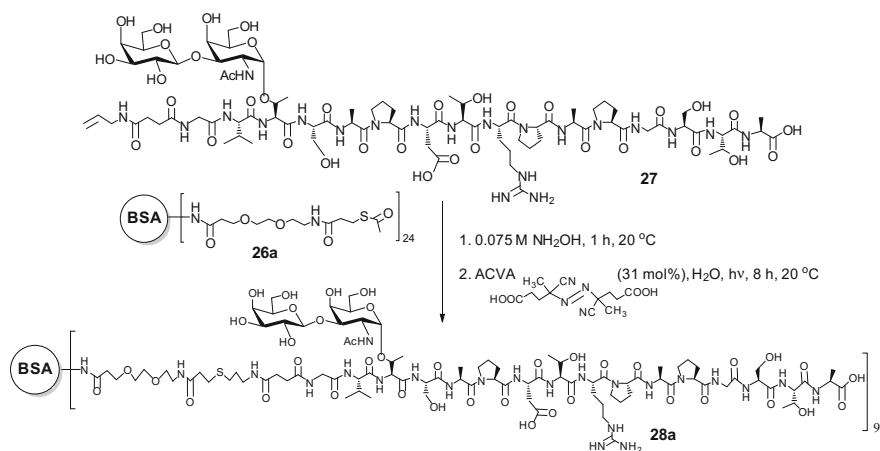
Scheme 10 Thiol-ene coupling of a glycopeptide antigen to thiol-functionalized BSA [51]
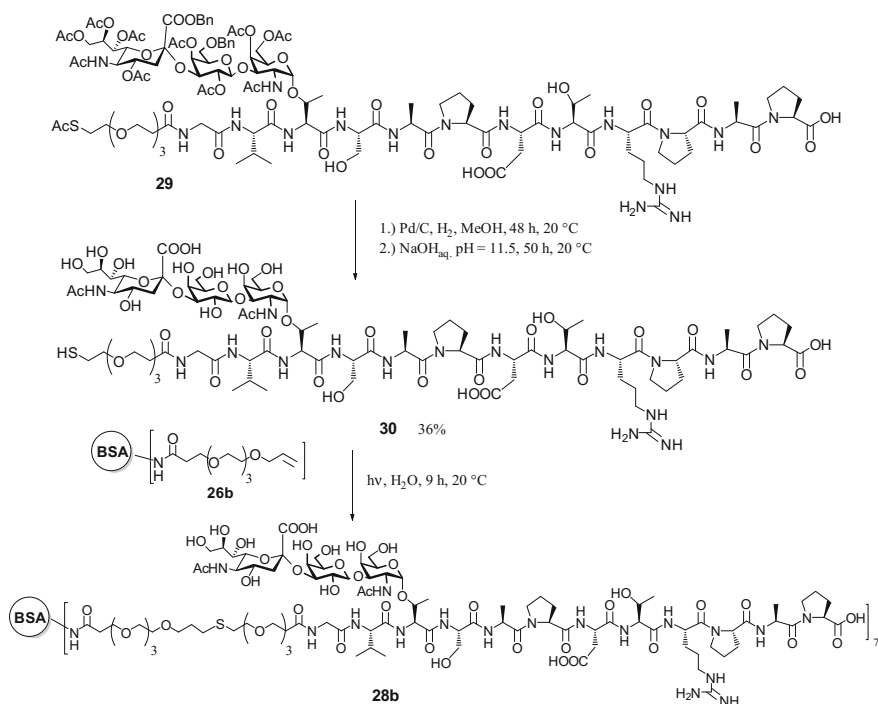
Scheme 11 Thiol-ene coupling of a glycopeptide antigen to olefin-decorated BSA [51]

condensed at the *N*-terminus before the glycopeptide **29** was detached from the resin. Removal of the sialic benzyl ester by hydrogenolysis also effected some desulfurization and, thus, reduced the yield of the desired glycopeptide **30** which was isolated after careful saponification at pH 11.5. It certainly would be more appropriate to remove the benzyl ester in this saponification step. Olefine-decorated BSA **26b** and glycopeptide antigen **30** were subjected to thiol-ene ligation by photochemical activation in water (Scheme 11).

This thiol-ene coupling with reverse arrangement of the functional groups appeared slightly less efficient. The number of glycopeptide antigens linked to a protein molecule was considered not optimal for the creation of an efficient vaccine. Oxidation to give the disulfide of glycopeptide **30** was observed as the major side reaction. Its amount increased with extended reaction time. The great number of different functional groups of the glycopeptides interferes with these coupling reactions proceeding via thiyl radicals. These factors are obviously less important in corresponding coupling reactions of carbohydrate ligands to proteins [52].

As the BSA conjugates of glycopeptide antigens exhibited only moderate immunogenicity, in further experiments they were not used as vaccines but served as coating material for probing antisera in ELISA analyses.

# 6 Coupling of Glycopeptides to Proteins Using Diethyl Squarate

Regarding the limited practicability of the photo-induced thio-ene coupling for the construction of glycopeptide-protein vaccines, the condensation reactions promoted by water-soluble carbodiimides, as for example EDC (see Schemes 4 and 9), may be considered useful alternatives. But orienting investigations with MUC1 tandem repeat peptides revealed that the activation of unprotected oligopeptides of this type with EDC and hydroxyl-benzotriazol (HOBt) or *N*-hydroxy-succinimide in dimethylformamide (DMF) or water gives the expected C-terminal active esters only in a mixture with a number of other products. Obviously, aspartic side chain carboxylic groups were also activated and, in particular, aspartimide rearrangement took place [53].

In contrast to these experiences, the differentiated reactivity of squaric diesters towards amines [54] appeared a promising tool for exploitation in glycopeptide-protein conjugate formation.

In fact, the sialyl-Tn glycopeptide MUC1 sequence **31** synthesized on solid-phase and completely deprotected reacted with diethyl sqarate in water at pH 8 to selectively furnish the glycopeptide squaric monoamide derivative **32** [55]. The compound can be purified by chromatography and characterized by NMR spectroscopy. In water at pH 9.0, it underwent further aminolysis by the amino functions of BSA lysine side chains to yield the BSA-glycopeptide conjugate **33** (Scheme 12) [55]. According to its MALDI-TOF spectrum, the conjugate contained on average 6 molecules glycopeptide per molecule BSA. Optimization of this coupling showed that the coupling rate is higher at pH 9.5, so that the second steps in further squarate couplings were conducted under these conditions.

The coupling via preformed squarate monoamides is also very valuable for the synthesis of complex, fully synthetic vaccines as was demonstrated for vaccines in which the tumor-associated MUC1 glycopeptide as the B-cell epitope was combined with three T-cell epitopes (Scheme 13) [56].

In a linear solid-phase synthesis, the combination **34** of the T-cell epitope peptide from *Yersinia pestis 1* with that from *Mycobacterium tuberculosis* is constructed. These components are connected by two triethylene glycol spacer molecules separated by a central lysine. The 4-amino group of this lysine was transformed into the squaric monoamide ester **34a** in water/ethanol at pH 8. This monoamide was then reacted with an analogously synthesized conjugate **35** of the MUC1 glycopeptide with another T-cell epitope from *Yersinia pestis 2* also connected to each other by two spacers linked through a lysine in the center. The coupling between the two large conjugates was achieved in high yield by squaric
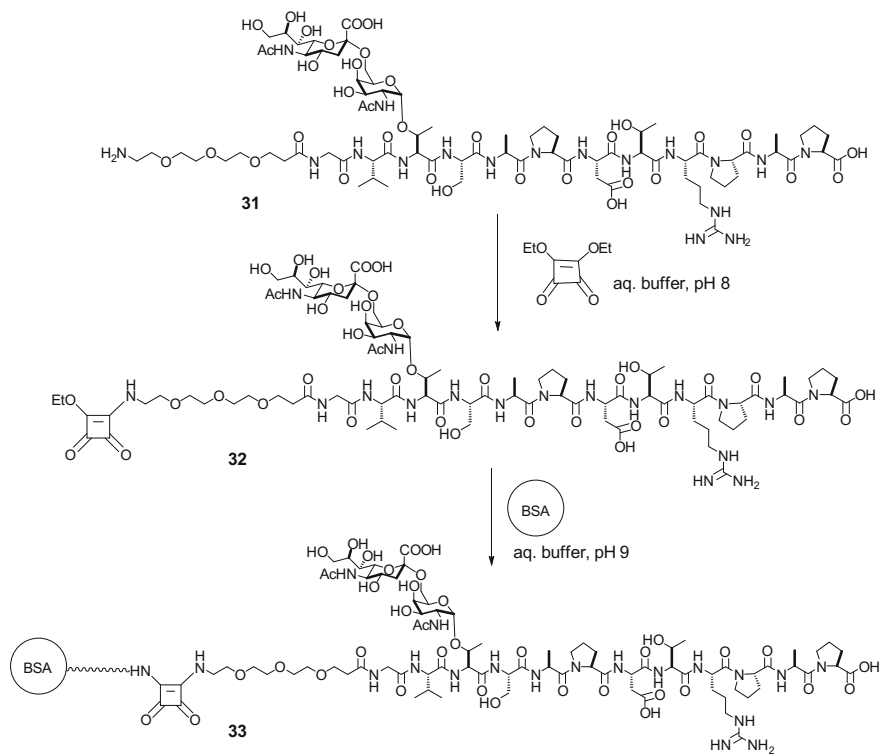
Scheme 12 Coupling of a glycopeptide antigen to BSA through squaric acid amide formation

diamide formation in water at pH 9.5 (Scheme 13) [56]. It is noteworthy that the vaccine **36** isolated after ultrafiltration was pure, and its structure was confirmed by high-resolution ESI mass spectra and 2D-COSY- HSQC- and HMBC NMR spectra.

Immunization of mice with vaccine **36** without application of any external immune stimulating adjuvant elicited significant immune responses of IgG antibodies. The antisera were much stronger than those induced by the Eastern part of **36** which is equal to conjugate **35** [56].
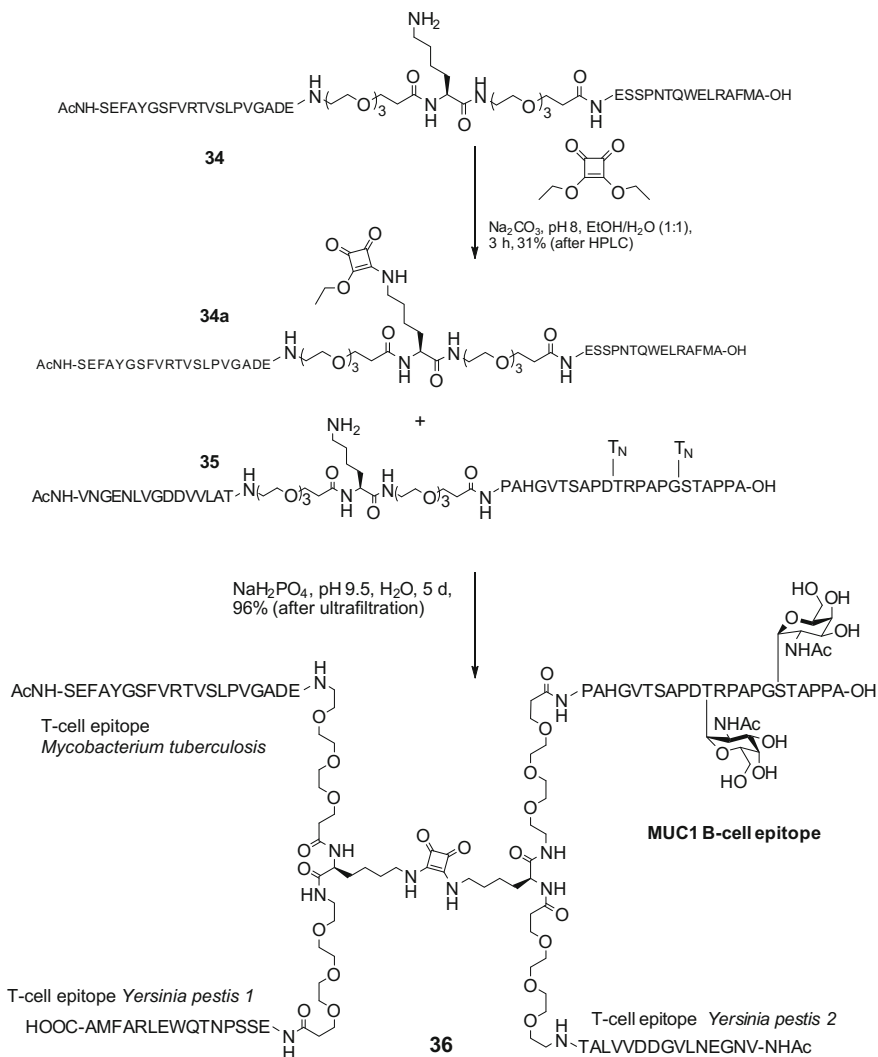
Scheme 13 A fully synthetic MUC1 glycopeptide vaccine in which the MUC1 B-cell epitope is covalently combined with three T-cell epitopes via squaric diamide coupling [56]

7 Glycopeptide-Tetanus Toxoid Vaccines

The fully synthetic vaccines as well as the glycopeptide conjugates with bovine serum albumin induced significant immune responses. However, overcoming the natural tolerance of the immune system towards glycopeptides as endogenous structures demands vaccines which elicit very strong immune reactions. Synthetic vaccines in which mucin glycopeptides are combined with T-cell epitopes from

tetanus toxoid had shown promising immunological properties [57–59]. The use of the tetanus toxoid itself was expected even more advantageous. Tetanus toxoid is frequently applied for vaccination in human medicine since many years. As a consequence, vaccine candidates consisting of MUC1 glycopeptide antigens and tetanus toxoid could also be used in human medicine as antitumor vaccines provided they are sufficiently effective and tumor selective. Tetanus toxoid is a much larger protein (molar weight ∼ 150 kDa) than BSA, and it is very expensive. As a rule, only small amounts of this carrier are available for academic studies.

Relying on the encouraging results of the coupling reaction with the aid of diethyl squarate, the MUC1 glycopeptide antigen **21** containing the tumor-associated sialyl-Tn antigen and the *N*-terminal spacer amino acid were treated with diethyl squarate in water/ethanol at pH 8 for 1.5 h (Scheme 14). After neutralization with 1 N acetic acid, the solution was lyophilized, and the product was purified by semi-preparative HPLC to give the squaric monoamide ester **37** in high yield. This compound was characterized by high-resolution ESI mass spectrometry and by NMR spectroscopy (COSY, HSQC). The thus isolated compound was dissolved together with tetanus toxoid in aqueous sodium phosphate buffer of pH 9.5 and stirred for 3 days. After ultrafiltration through a 30 kDa membrane using deionized water and lyophilization, the MUC1-glycopeptide-tetanus toxoid conjugate **38** was isolated as a colorless amorphous substance [47]. The loading of the protein with glycopeptide antigens could not be measured by MALDI-TOF mass spectrometry. Therefore, the loading was estimated by comparative ELISA tests using an antiserum of a mouse which had been immunized with fully synthetic vaccine containing the MUC1 glycopeptide antigen **31** (see Ref. [60]). Tetanus toxoid conjugate **38** and the corresponding BSA conjugate **39** (Fig. 4) served as coats in these comparative ELISAs. The bound antibodies were photometrically determined using a biotinylated secondary anti-mouse antibody and treatment with streptavidin-horseradish peroxidase via catalyzed oxidation to give a staining (see below). It was estimated from these experiments that the tetanus toxoid vaccine **38** carries on average at least 20 molecules of the glycopeptide antigen per molecule tetanus toxoid [47]. Gravimetric estimation indicated a larger loading of on average 44 molecules glycopeptide. Because of the hydrophilicity of tetanus toxoid and its conjugate **38**, this estimation certainly is less accurate.

The immunization of mice with MUC1 glycopeptide-tetanus toxoid vaccine **38** actually resulted in very strong immune reactions of the vaccinated mice. Ten
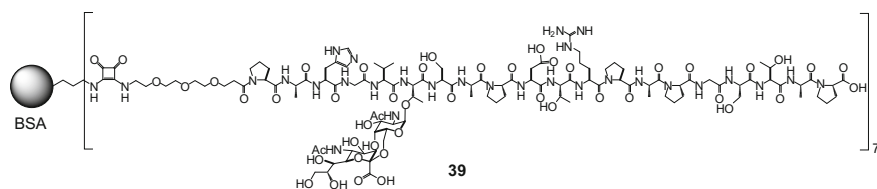

Fig. 4 BSA conjugate **39** of the MUC1 glycopeptide **21** used as coating material in ELISA analyses of antisera induced by vaccine **38**

BALB/c mice were vaccinated with 20 μg of vaccine **38** three times at intervals of 21 days. Five days after the third immunization, blood was drawn from the tail vein of each mouse. The obtained antisera were subjected to ELISA analyses in which the BSA conjugate **39** (Fig. 4) of the glycopeptide **21** served as the coating material placed in the wells of the micro-titer plates.

After washing with water, the antisera induced by vaccine **38** were added to the wells in a dilution series (Fig. 5). Incubation at 37 °C lasted for 1 h. After thoroughly washing with buffer solution the biotinylated sheep-anti-mouse antibody was added in order to determine the induced mouse antibodies which are bound to the BSA conjugate **39**. Coordination of biotin to streptavidin which is linked to the horseradish peroxidase (HPO) allowed for the quantitative determination of the bound antibodies through photometric analysis of the HPO-catalyzed oxidation of the colorless di-ammonium 2,2′-azino-bis(3-ethylbenzothiazolin-6-sulfonate) (ABTS) by hydrogen peroxide to give a green radical cation. The optical density of this color in relation to the dilution of the mouse antiserum affords the titer of antibodies induced by **38** which recognize the tumor-associated MUC1 antigen presented in the BSA conjugate **39** (Fig. 5) [47].

All ten mice showed dramatic immune responses certainly strong enough to override the natural tolerance of the immune system towards the endogenous MUC1 glycopeptide structure. End point titers (10% remaining absorption) of up to 1 million were observed. For a serum diluted by a factor of 10^6 still, 10% of the original serum antibody binding was recorded. The recognition of the glycopeptide structure in conjugate **39** is structure-selective. This was demonstrated by the complete neutralization of this binding after addition of a few μg of the glycopeptide antigen **21** to the antiserum [47].
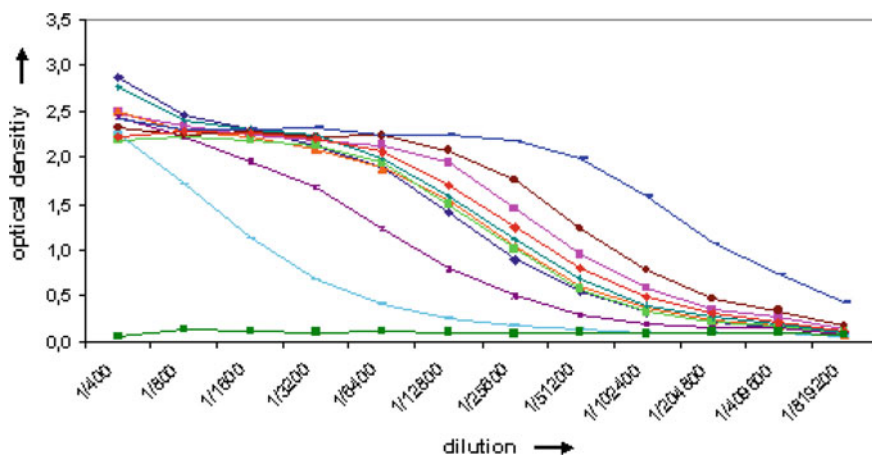

Fig. 5 ELISA analyses of antisera of ten mice immunized with MUC1 glycopeptide-tetanus toxoid vaccine **38**. The BSA conjugate **39** served as the coating material [47]; *green bottom line* negative control
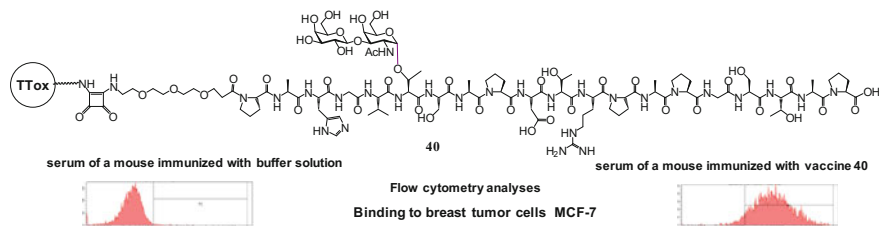
Fig. 6 Antiserum induced in mice by T-antigen MUC1 glycopeptide-tetanus toxoid vaccine **40** and its binding to MCF-7 breast tumor cells (picture on the *right*) [61]

In a second example of MUC1 glycopeptide-tetanus toxoid vaccine (Fig. 6) the sialyl-Tn antigen saccharide was substituted for the Thomsen-Friedenreich antigen (T-antigen). Vaccine **40** was prepared by solid-phase synthesis, deprotection, and coupling reactions in complete analogy to the construction of vaccine **38** [61]. Immunization of BALB/c mice with vaccine **40** also resulted in very strong immune responses, and the induced antiserum exhibited strong and selective recognition of the corresponding BSA conjugate which contains the T-antigen glycopeptide present in vaccine **40**. The induced antisera showed end point titers of about 800,000. The antiserum of one of the mice vaccinated with **40** was used for the investigation of its binding to breast tumor cells of cell line MCF-7. To this end, the MCF-7 tumor cells were incubated with the antiserum of this mouse, and after washing the antibodies bound to the tumor cells were detected with fluorescent-labeled goat-anti-mouse antibodies. The thus treated cells were passed through a flow cytometer in which all cells are counted by scattering a Laser beam, but analyzed according to their fluorescence. For comparison, the MCF-7 tumor cells were also incubated with the serum of a mouse which was just treated with buffer solution instead of a vaccine. The serum of this mouse showed no binding to the tumor cells (Fig. 6, left picture), whereas the 1000-fold diluted antiserum of the mouse vaccinated with vaccine **40** showed almost complete recognition of the membrane molecules present on the MCF-7 tumor cells [61].

Addition of the glycopeptide contained in the vaccine **40** to the elicited antiserum abolished this binding to the tumor cells to a large extent. This neutralization of the binding again gives evidence that the antibodies elicited by the synthetic vaccines actually bind to tumor-associated mucin MUC1 molecules exposed on the surface of the epithelial tumor cells.

In the synthesis of a third MUC1 glycopeptide antigen of this series, the tumor-associated sialyl-Tn antigen was introduced with the serine of the GSTA region of the tandem repeat, and the sequence was *C*-terminally extended by two amino acids in order to complete the STAPPA motif [62]. It had been found that glycosylation in this region favored a helix-type conformation of this segment which was considered characteristic for tumor-associated MUC1 [11]. The coupling of this glycopeptide antigen to tetanus toxoid was achieved via formation of the squaric monoamide ester **61** and its reaction with tetanus toxoid in aqueous sodium
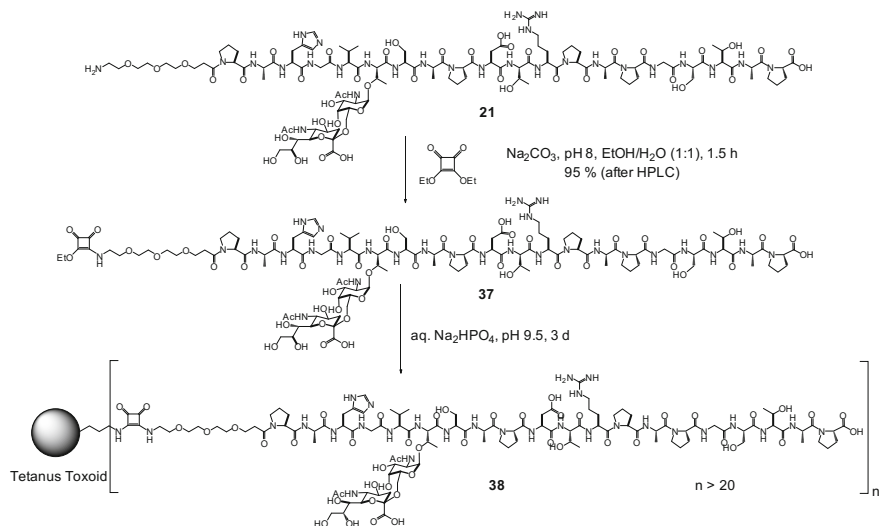
Scheme 14 Coupling of a tumor-associated MUC1 glycopeptide antigen to tetanus toxoid through squaric acid diamide formation [47]

phosphate buffer solution at pH 9.5 (Scheme 15A). Ultrafiltration and lyophilization afforded the MUC1 glycopeptide-tetanus toxoid vaccine **42** in which the glycosylation position is shifted and the STAPPA motif is completed in comparison to the vaccine **38** described above (Scheme 14).

Immunization of mice with this vaccine again elicited very strong immune responses in all animals. End point titers amounted to 500,000 and higher. The antiserum was investigated concerning the binding to the MCF-7 breast tumor cells (Scheme 15B). In comparison to the serum of a mouse which just was vaccinated with buffer solution (Scheme 15Ba), the antibodies in the antiserum of a mouse vaccinated with vaccine **42** showed more than 97% binding of the tumor cells (Scheme 15Bb). In contrast, antibodies induced against tetanus toxoid itself (Scheme 15Bc) exhibited no binding to the MCL-7 tumor cells.

ELISA with isotype-specific secondary antibodies revealed that prevailingly IgG1 antibodies had been induced (Fig. 7A) giving evidence that the switch to IgG-producing B-cells within an adaptive immune response had occurred.

In addition to probing the antibodies concerning their recognition of tumor cells in cell cultures (Scheme 15B), their capability of selective binding to tumor cells in tumor tissues was investigated [62]. Figure 7B shows two examples of mammary carcinoma tissue sections fixed with formalin and embedded in paraffin in a light microscope (magnification 1/100). On the one hand, the tissues were treated with an isotyping IgG antibody (Fig. 7B, a and c) and on the other hand with the antiserum of a mouse immunized with vaccine **42** (Fig. 7B, b and d). The antibodies bound to the tumor tissues were detected with a biotinylated secondary antibody. Its adhesion to the bound mouse antibodies was visualized with a streptavidin-horseradish
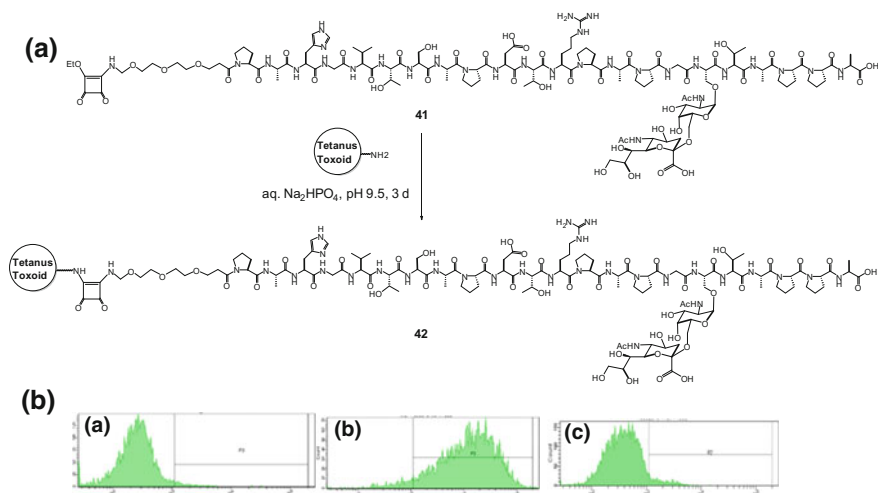
Scheme 15 **A** Formation of a MUC1 glycopeptide-tetanus toxoid vaccine (**42**) with glycosylation in the STAPPA sequence [62]; **B** Binding of antisera to MCF-7 breast tumor cells measured by flow cytometry: **a** serum of a mouse immunized with buffer solution (control); **b** antiserum of a mouse vaccinated with **42**; antiserum of a mouse vaccinated with tetanus toxoid

peroxidase conjugate that catalyzes the oxidation of 3-amino-9-ethyl-carbazole to a rose-colored dye.

While the tumor tissues in the early phase (G1 phase, Fig. 7B, a) after treatment with the antiserum of the mouse vaccinated with **42** indicated only weak binding (Fig. 7B, b) of the antibodies induced by the synthetic vaccine, the advanced tumor (Fig. 7B, c) after incubation with the antiserum elicited by vaccine **42** displayed very strong binding of the antibodies induced by vaccine **42** (Fig. 7B, d).
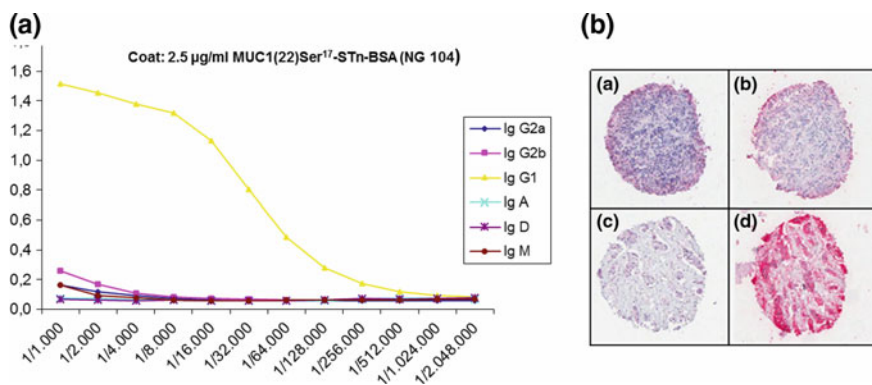
Fig. 7 a Antibody isotypes induced in mice through vaccination with MUC1 glycopeptide-tetanus toxoid vaccine **42**. **b** Recognition of tumor cells in mammary carcinoma tissue sections by the antiserum induced through vaccine **42** [62] in a light microscope: early stage tumor: *a*, prior to and *b*, after treatment with the antiserum; advanced tumor *c*, prior to and *d*, after treatment with the antiserum

In view of the very promising results achieved in vaccinations with the synthetic vaccine **42**, hybridomas were generated by fusing spleen cells from one of the immunized mice with murine myeloma cells. The experiments resulted in the production of a monoclonal antibody GGSK-1/30 which strongly binds to the MUC1 glycopeptide epitope presented in the vaccine **42** [63]. Flow cytometry experiments showed that monoclonal antibody GGSK-1/30 does bind with high rates to mammary carcinoma cells of cell lines T47-D and MCF-7 ($\sim$95%) and also recognize pancreas tumor cells PANC-1 with a rate of >73%, but does not recognize normal human mammary epithelial cells (HMEC) in analogous investigations. This selectivity in the binding to tumor cells was confirmed in fluorescence microscopy studies of MCF-7 and T47-D tumor cells and normal HMEC cells after treatment with monoclonal antibody GGSK-1/30 and subsequent staining with a fluorescent-labeled (Alexa Fluor 488) goat-anti-mouse secondary IgG1 antibody (Fig. 8).

The MCF-7 tumor cells (Fig. 8a) distinctly differ in their habitus from the normal epithelial cells HMEC (Fig. 8d). After incubation, these tumor cells with monoclonal antibody GGSK-1/30 and staining with the secondary antibody, their membranes brightly shine in green fluorescence (Fig. 8b) indicating the intense binding of mAb GGSK-1/30 to the molecules exposed on the membranes of the tumor cells. Since the same treatment of normal HMEC cells did not result in any effect (no fluorescence at all), the blue 4′,6-diamidino-2-phenylindole (DAPI) was added as a second fluorescene dye which colorizes the nuclei (Fig. 8e). Addition of DAPI to the MCF-7 cells already labeled with GGSK-1/30 and the Alexa Fluor 488-linked secondary antibody shows both, the green fluorescence of the membranes and the blue fluorescence of the nuclei (Fig. 8c).
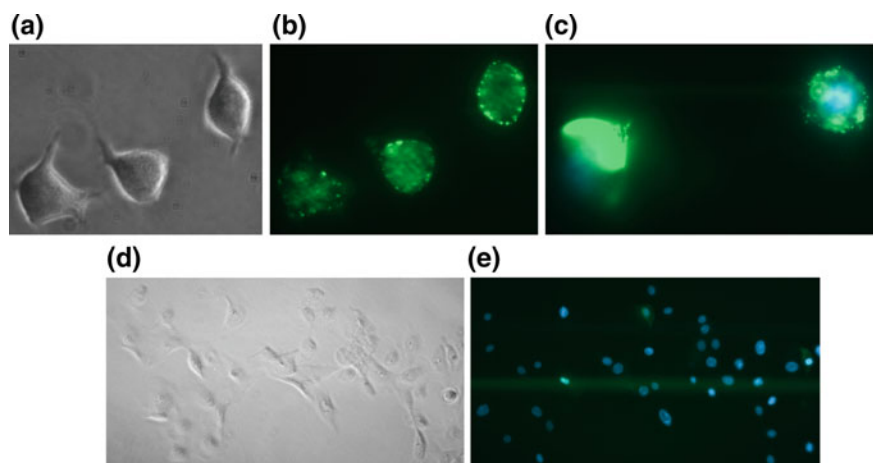

Fig. 8 Fluorescence microscopy of MCF-7 breast tumor cells (**a**) and normal human epithelial cells HMEC (**d**). **b** Staining of MCF-7 with monoclonal antibody GGSK-1/30 and an Alexa Fluor 488-labeled secondary antibody. **c** Same treatment as in **b** and addition of blue fluorescent DAPI visualizing the nuclei. **e** HMEC cells after treatment with monoclonal antibody GGSK-1/30 and an Alexa Fluor 488-labeled secondary antibody and addition of DAPI

These experiments give evidence that the monoclonal antibody induced in mice through the tumor-associated MUC1 glycopeptide-tetanus toxoid vaccine **42** completely differentiated between epithelial tumor cells and normal epithelial cells. This result is of particular importance because a strong immune response induced by the synthetic vaccine **42** which can overcome the natural tolerance against the endogenous structure should not cause severe autoimmune reactions.

8 Conclusion

Coupling reactions, in which the carbohydrate should be linked to a partner molecule to give a biological ligand structure comprising both components, are most demanding since the natural linkage needs to be installed. In a stereo- and regioselective glycosylation of a Thomsen-Friedenreich antigen with an sialyl-LewisX trichloroacetimidate as the glycosyl donor the natural β-glycoside connection was achieved. The obtained hexasaccharide-threonine building block was introduced into the synthesis of the natural glycopeptide binding site of the P-selectin glycoprotein ligand-1 (PSGL-1) which exhibited much stronger binding to P-selection than sialyl-LewisX itself.

Fortunately, for the construction of glycopeptide vaccines containing endogenous glycopeptide antigens, less strict requirements are applicable as far as the target antigen structure is appropriately presented to the immune system in the framework of the prepared vaccine. The carboxy-terminal activation of a glycopeptide by a water-soluble carbodiimide was realized in the preparation of vaccines comprising *N*-terminal glycophorin glycopeptides and bovine serum albumin. However, it can only be applied to molecules which have only one carboxy group of enhanced reactivity. More generally applicable ligation with proteins is achieved by radical-induced thiol-ene coupling reactions and conjugation reactions based on the differentiated electrophilic reactivity of squaric acid diesters. In particular, the latter have the advantage that amino functions of exposed reactivity on the glycopeptide selectively react with squaric diester in aqueous solutions at pH 8 to give the squaric monoamide ester which can be isolated and purified. It is then applicable to a mild coupling reaction in which it subject to aminolysis by amino groups of lysine side chains of the protein. According to this strategy, MUC1 glycopeptide-tetanus toxoid vaccines are accessible, which induce very strong immune responses in mice. These immune reactions have the potential to override the natural tolerance of the immune system towards endogenous structures. The induced antibodies selectively recognize epithelial tumor cells both in cell culture and in tumor tissues, as was demonstrated for mamma carcinomas. Since a monoclonal IgG1 antibody generated from one of the immunized mice completely differentiated between normal and tumor epithelial cells encouraging preconditions are fulfilled for the development of an active vaccination of patients against their tumor diseases.

References

1. Landsteiner K (1901) Ueber Agglutinationserscheinungen normalen menschlichen Blutes. Wiener Kl. Wochenschrift 14:1132–1134
2. Lemieux RU (1989) The origin of the specificity in the recognition of oligosaccharides by proteins. Chem Soc Rev 18:347–374
3. Bundle DA, Nitz M, Wu Y, Sadowska JM (2008) A uniquely small, protective carbohydrate epitope may yield a conjugate vaccine for candida albicans. ACS Symp Ser 989:163–183
4. Etzler ME, Kabat EA (1970) Purification and characterization of a lectin (plant hemagglutinin) with blood group A specificity from Dolichos biflorus. Biochemistry 9:869–877
5. Lis H, Sharon N (1998) Lectins: carbohydrate-specific proteins that mediate cellular recognition. Chem Rev 98:637–674
6. Simanek EE, McGarvey GJ, Jablonowski JA, Wong C-H (1998) Selectin-Carbohydrate interactions: from natural ligands to designed mimics. Chem Rev 98:833–882
7. Dondoni A, Massi A, Nanni P, Roda A (2009) A new ligation strategy for peptide and protein glycosylation: photoinduced thiol-ene coupling. Chem Eur J 15:11444–11449
8. Lin YA, Chalker JM, Davis BG (2009) Olefin metathesis for site-selective protein modification. *ChemBioChem* 10, 959–969
9. Lee DJ, Sung-Hyun Yang S-H, Williams GM, Brimble MA (2012) Synthesis of multivalent neoglyconjugates of MUC1 by the conjugation of carbohydrate-centered, triazole-linked glycoclusters to MUC1 peptides using click chemistry. J Org Chem 77:7564–7571
10. Doll F, Buntz A, Späte A-K, Schart VF, Timper A, Schrimpf W, Hauck CR, Zumbusch A, Wittmann V (2016) Visualization of protein-specific glycosylation inside living cells. Angew Chem Int Ed 55:2262–2266
11. Braun P, Davies GM, Price MR, Williams PM, Tendler SPJ, Kunz H (1998) Effects of glycosylation on fragments of tumor associated human epithelial mucin MUC1. Bioorg Med Chem 6:1531–1545
12. Coltart DM, Royyuru AK, Willaims LJ, Glunz PW, Sames D, Kuduk SD, Schwarz JB, Chen X-T, Danishefsky SJ, Live DH (2002) Principles of mucin architecture: structural studies on synthetic glycopeptides bearing clustered Mono-, Di-, Tri-, and hexasaccharide glycodomains. J Am Chem Soc 124:9833–9844
13. Corzana F, Busto JH, Garzia de Luis M, Jimenez-Barbero J, Avenoza A, Peregrina M (2009) The nature and sequence of the amino acid aglycone strongly modulates the conformation and dynamics effects of tn antigen's clusters. Chem Eur J 15:3863–3874
14. Kuhn A, Kunz H (2007) Saccharide-induced peptide conformation in glycopeptides of the recognition region of li-cadherin. Angew Chem Int Ed 46:454–458
15. Hashimoto R, Fujitani N, Takegawa Y, Kurogochi M, Matsushita T, Naruchi K, Ohyabu N, Hinou H, Gao XD, Manri N, Satake H, Kaneko A, Sakamoto T, Nishimura S-I (2011) An efficient approach for the characterization of mucin-typeglycopeptides: the effect of O-glycosylation on the conformation of synthetic mucin peptides. Chem Eur J 17:2393–2404
16. Bogert A, Heimburg-Molinaro J, Song X, Lasanjak Y, Ju T, Liu M, Thompson P, Raghupati G, Barany G, Smith DF, Cummings RD, Live D (2012) Deciphering structural elements of mucin glycoprotein recognition. ACS Chem Biol 7:1031–1039
17. Lijun X, Ramachandran V, McDaniel JM, Nguyen KN, Cummings RD, McEver RP (2003) N-terminal residues in murine P-selectin glycoprotein ligannd-1 required for binding to murine P-selectin. Blood 101:552–559
18. Gaidzik N, Westerlind U, Kunz H (2013) The development of synthetic antitumor vaccines from mucin glycopeptide antigens. Chem Soc Rev 42:4421–4442
19. Angiari S, Constantin G (2013) Selectins and their ligands as potential immunotherapeutic targets in neurological diseases. Immunotherapy 5:1207–1220
20. Phillips ML, Nudelman E, Gaeta FC, Perez M, Singhal AK, Hakomori S, Paulson JC (1990) ELAM-1 mediates cell adhesion by recognition of a carbohydrate ligand. Sialyl-Lex. Science 250:1130–1132

21. Springer A (1994) Traffic signals for lymphocyte recirculation and leukocyte emigration: the multistep paradigm. Cell 76:301–314
22. Coupland LA, Parish CR (2014) Platelets, selectins, and the control of tumor metastasis. Semin Oncol 41:422–434
23. Leppänen A, Mehtal P, Ouyang Y-B, Ju T, Helin J, Moore KL, van Die I, Canfield WM, McEver RP, Cummings RD (1999) A novel glycosulfopeptide binds to p-selectin and inhibits leukocyte adhesion to P-selectin. J BIol Soc 274:24838–24848
24. Sprengard U, Kretzschmar G, Bartnik E, Hüls C, Kunz H (1995) Synthesis of an RGD Sialyl-LewisX glycoconjugate: a new highly active ligand for P-selectin. Angew Chem Int Ed 34:990–993
25. Koeller KM, Smith MEB, Wong C-H (2000) Tyrosine sulfation on a PSGL-1 glycopeptide influences the reactivity of glycosyltransferases responsible responsible for synthesis of the attached O-Glycan. J Am Chem Soc 122:742–743
26. Brocke C, Kunz H (2004) Synthetic tumor-associated glycopeptide antigens from the tandem repeat of epithelial mucin MUC4. Synthesis, 525–542
27. Baumann K, Kowalczyk D, Kunz H (2008) Total synthesis of the glycopeptide recognition domain of the P-Selectin glycoprotein ligand 1. Angew Chem Int Ed 47:3445–3449
28. Kolb HC, Ernst B (1997) Development of tools for the design of selectin antagonists. Chem Eur J 3:1571–1578
29. Baumann K, Kowalczyk D, Gutjahr T, Pieczyk M, Jones C, Wild MK, Vestweber D, Kunz H (2009) Sulfated and non-sulfated glycopeptide recognition domain of P-selectin glycoprotein ligand 1 and their binding to P- and E-selectin. Angew Chem Int Ed 48:3174–3178
30. Review: Springer GF (1984) Tn and T, general carcinoma autoantigens. Science 224, 1198–1206
31. Kunz H, Birnbach S (1986) Synthesis of O-glycopeptides of the tumor-associated Tn- and T-antigen type and their binding to bovine serum albumin. Angew Chem Int Ed 25:360–362
32. König W, Geiger R (1970) Eine neue methode zur synthese von peptiden: aktivierung der carboxylgruppe mit dicyclohexylcarbodiimid unter zusatz von 1-hydroxy-benzotriazol. Chem Ber 103:788–798
33. Kunz H, von dem Bruch K (1994) Neoglycoproteins from synthetic glycopeptides. Methods Enzymol 247:3–30
34. Bremer PT, Kimishima A, Schlosburg JE, Zhou B, Collins KC, Janda KD (2016) Combatting synthetic designer opioids: a conjugate vaccine ablates lethal doses of fentanyl class drugs. Angew Chem Int Ed 55:3772–3775
35. (a) Dippold W, Steinborn A, Meyer zum Büschenfelde K-H (19990) The role of the thomsen-friedreich antigen as a tumor-associated molecule. Environ. Health Persp 88, 255–257. (b) Steinborn A (1990) Dissertation: definition von Proliferations- und Differenzierungsmolekülen auf menschlichen Tumorzellen. Universtät Mainz, p. 73
36. Review: Beatson RE, Taylor-Papadimitriou J, Burchell JM (2010) MUC1 Immunotherapy. Immunotherapy 2, 305–327
37. Zotter S, Hageman PC, Lossnitzer A, van den Tweel J, Hilkens J, Mooi WJ, Hilgers J (1988) Monoclonal antibodies to epithelial sialomucins recognize epitopes at different cellular sites in adenolymphomas of the parotid gland. Int J Cancer Suppl 3, 38–44
38. Gendler SJ, Lancaster CA, Taylor-Papadimitriou J, Duhig T, Peat N, Burchell J, Pemberton L, Lalani EN, Wilson D (1990) Molecular cloning and expression of human tumor-associated polymorphic epithelial mucin. J Biol Chem 265, 15286–15293
39. Burchell JM, Mungul A, Taylor-Papadimitiou J (2001) O-Linked glycosylation in the mammary gland: changes that occur during malignancy. J Mammary Gland Biol Neoplasia 6:355–364
40. Hanisch F-G, Peter-Katalinic J, Egge H, Dabrowski U, Uhlenbruck G (1990) Structures of acidic O-linked polylactosaminoglycans on human skim milk mucins. Glycoconjugate J 7:525–543

41. Brockhausen I, Yang JM, Burchell J, Whitehouse C, Taylor-Papadimitriou J (1995) Mechanisms underlying aberrant glycosylation of MUC1 mucin in breast cancer cells. Eur J Biochem 233:607–617

42. Review: Becker T, Dziadek S, Wittrock S, Kunz H (2006) Synthetic glycopeptides from the mucin family as potential tools in cancer immunotherapy. Curr Cancer Drug Targets 6, 491–517

43. Liebe B, Kunz H (1997) Solid-phase synthesis of a Sialyl-Tn glycoundecapeptide of the MUC1 repeating unit. Helv Chim Acta 80:1473–1482

44. Kunz H, Unverzagt C (1988) Protective group dependent stability of intersaccharide bonds. Synthesis of fucosyl chitobiose glycopeptides. Angew Chem Int Ed 27:1697–1699

45. Sjölin P, Elofsson M, Kihlberg J (1996) Removal of acyl protective groups from glycopeptides. Base does not epimerize peptide stereocenters, and β-elimination is slow. J Org Chem 61:560–565

46. Dziadek S, Brocke C, Kunz H (2004) Biomimetic synthesis of the tumor-associated (2,3)-sialyl-T antigen and ist incorporation into glycopeptide antigens from the mucins MUC1 and MUC4. Chem Eur J 10:4150–4162

47. Kaiser A, Gaidzik N, Westerlind U, Kowalczyk D, Hobel A, Schmitt E, Kunz H (2009) A synthetic vaccine consisting of a tumor-associated Sialyl-Tn-MUC1 tandem repeat glycopeptide and tetanus toxoid: induction of a strong and highly selective immune response. Angew Chem Int Ed 48:7551–7555

48. Carpino LA (1993) 1-Hydroxy-benzotriazole. An efficient peptide coupling additive. J Am Chem Soc 115:4397–4398

49. Yoshitake S, Imagawa M, Ishikawa E, Niitsu Y, Urushizaki I, Nishiura M, Kanazawa R, Kurosaki H, Tachibana S, Nakazawa N, Ogawa H (1982) Mild and efficient conjugation of rabbit Fab' and horseradish peroxidase using a maleimide compound and its use for enzyme immunoassay. J Biochem 92:1413–1424

50. Posner T (1905) Beiträge zur Kenntniss der ungesättigten Verbindungen. II. Ueber die Addition von Mercaptanen an ungesättigte Kohlenwasserstoffe. Ber Dtsch Chem Ges 38: 646–657

51. Wittrock S, Becker T, Kunz H (2007) Synthetic vaccines of tumor-associated glycopeptide antigens by immune-compatible thioether linkage to bovine serum albumin. Angew Chem Int Ed 46:5226–5230

52. Dondoni A, Massi A, Nanni P, Roda A (2009) A new ligation strategy for peptide and protein glycosylation: photoinduced thiol–ene coupling. Chem Eur J 15:11444–11449

53. Glaffig M, Kunz H (2016) unpublished experiments

54. Tietze LF, Arlt M, Beller M, Glüsenkamp K-H, Jäde E, Rajewski MF (1991) Squaric acid diethyl ester: a new coupling reagent for the formation of drug bio-polymer conjugates. Chem Ber 124:1215–1221

55. Dziadek S, Kowalzcyk D, Kunz H (2005) Synthetic vaccines consisting of tumor-associated MUC1 glycopeptide antigens and bovine serum albumin. Angew Chem Int Ed 44:7624–7630

56. Palitzsch B, Hartmann S, Stergiou N, Glaffig M, Schmitt E, Kunz H (2014) A fully synthetic four-component antitumor vaccine consisting of a mucin glycopeptide antigen combined with three different T-helper cell epitopes. Angew Chem Int Ed 53:14245–14249

57. Keil S, Claus C, Dippold W, Kunz H (2001) Towards the development of antitumor vaccines: a synthetic conjugate of a tumor-associated MUC1 glycopeptide antigen and a tetanus toxin epitope. Angew Chem Int Ed 40:366–369

58. Wilkinson BL, Day S, Malins LR, Apostolopoulos V, Payne RJ (2011) Self-adjuvanting multicomponent cancer vaccine candidates combining per-glycosylated MUC1 glycopeptides and the Toll-like receptor 2 agonist Pam₃CysSer. Angew Chem Int Ed 50, 1635–1639

59. Cai H, Chen M-S, Sun Z-Y, Zhao Y-F, Kunz H, Li Y-M (2013) Self-adjuvanting synthetic antitumor vaccine from MUC1 glycopeptides conjugated to T-cell epitopes from tetanus toxoid. Angew Chem Int Ed 52:6106–6110

60. Dziadek S, Hobel A, Schmitt E, Kunz H (2005) A fully synthetic vaccine consisting of a tumor-associated glycopeptide antigen and a T-cell epitope for the induction of a highly specific humoral immune response. Angew Chem Int Ed 44:7630–7635
61. Hoffmann-Röder A, Kaiser A, Wagner S, Gaidzik N, Kowalczyk D, Westerlind U, Gerlitzki B, Schmitt E, Kunz H (2010) Synthetic antitumor vaccines from tetanus toxoid conjugates of MUC1 glycopeptides with the Thomsen-Friedenreich antigen and a fluorine substituted analogue. Angew Chem Int Ed 49:8498–8503
62. Gaidzik N, Kaiser A, Kowalczyk D, Westerlind U, Gerlitzki B, Sinn HP, Schmitt E, Kunz H (2011) Synthetic antitumor vaccines containing MUC1 glycopeptides with two immunodominant domains—induction of a strong immune response against breast tumor tissues. Angew Chem Int Ed 50:99778–99981
63. Palitzsch B, Gaidzik N, Stergiou N, Stahn S, Hartmann S, Gerlitzki B, Teusch N, Flemming P, Schmitt E, Kunz H (2016) A synthetic glycopeptide vaccine for the induction of a monoclonal antibody that differentiates between normal and tumor mammary cells and enables the diagnosis of human pancreatic cancer. Angew Chem Int Ed 55:2894–2898

Recent Advances in the Stereochemical Outcome of Multicomponent Reactions Involving Convertible Isocyanides

Krishnakant Patel and Peter R. Andreana

Abstract Multi-component reactions (MCRs) have become an integral part of organic synthesis as short and very efficient routes to molecular diversity with included varying stereochemistry. MCRs have evolved rapidly in terms of the components used in the reactions and their role on the stereo chemical outcomes. This chapter focuses on covering recent contributions towards MCRs including targeted asymmetric control. Furthermore, advances in MCRs, as reported by many researchers, covering the utility of the convertible isocyanides, stereochemical advances of the Passerini and Ugi reactions and finally, the use of carbohydrates as chiral auxiliaries are examined and discussed.

1 Introduction

Carbohydrates are naturally occurring complex chiral molecules that have significant biological significance. They have been used in the laboratory setting as chirons for the transformation into multiple functionalities such as aldehydes, acids, amines, and isocyanides as well as employed in reaction settings such as multicomponent reactions. In recent years, there has been a marked increase in utilizing carbohydrates as key components in multicomponent coupling reactions such as the Ugi, and Passerini reactions to name a few [1, 2]. In the aforementioned reactions, sugars were used as the isocyanide and amine components for understanding diastereoselective roles in the key carbon–carbon bond forming reaction. However, over a period of time the use of carbohydrates in such multicomponent reactions has diminished. In the later

K. Patel · P.R. Andreana (✉)
Department of Chemistry and Biochemistry, School of Green Chemistry
and Engineering, University of Toledo, 2801 W. Bancroft St., Wolfe Hall 2232B, Toledo,
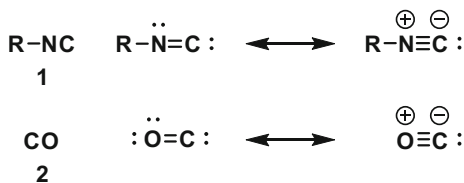OH 43606-3390, USA
e-mail: Peter.Andreana@utoledo.edu

K. Patel
e-mail: Krishnakant.Patel@rockets.utoledo.edu

Fig. 1 General representation of the isocyanide and carbon monoxide

$$R-NC \qquad R-\overset{..}{N}=C: \qquad \longleftrightarrow \qquad R-\overset{\oplus}{N}\equiv C:\overset{\ominus}{}$$

1

$$CO \qquad :\overset{..}{O}=C: \qquad \longleftrightarrow \qquad \overset{\oplus}{O}\equiv C:\overset{\ominus}{}$$

2

sections of this chapter, we attempt to discuss the role of carbohydrates in multi-component reactions as both isocyanide and amine functionalities.

An isocyanide, also known as an isonitrile or carbylamine is an organic compound characterized by the functional group $R–N\equiv C$. It is the isomer of closely related cyanide ($R–C\equiv N$), thus the prefix *iso*. Electronically, isocyanides are similar to carbon monoxide and therefore have been used to substitute carbon monoxide in organometallic transformations. Akin to carbon monoxide, the isocyanides are represented by two resonance structures, one with a triple bond between nitrogen and carbon and the other with a double bond (Fig. 1).

Isocyanides are adaptable functional groups having a divalent carbon which acts as a proton acceptor (Brønsted base). Very few compounds belong to this class of proton acceptors. Isocyanides are also stable to basic conditions but under highly acidic conditions, they tend to hydrolyze or polymerize. Some suspect that they may even undergo radical formations depending on reaction conditions [3–6].

The first reported synthesis of an isocyanide was serendipitously discovered in 1859 by Lieke [7] when the synthesis of allyl cyanide was attempted but instead an isocyanide was characterized. The first naturally occurring isocyanide, Xanthocillin, discovered by Rothe in 1950, was later used as an antibiotic (Fig. 2).

Ever since the development of synthetic routes for isocyanides established circa 1867 by Gautier and Hoffman [8, 9], there have been many reports describing further developments, not just in terms of different synthetic routes but also for obtaining varied isocyanides which have been used in the synthesis of biologically important scaffolds. The synthetic importance and utility of isocyanides rocketed following their use in multicomponent reactions (MCRs). MCRs involve the use of
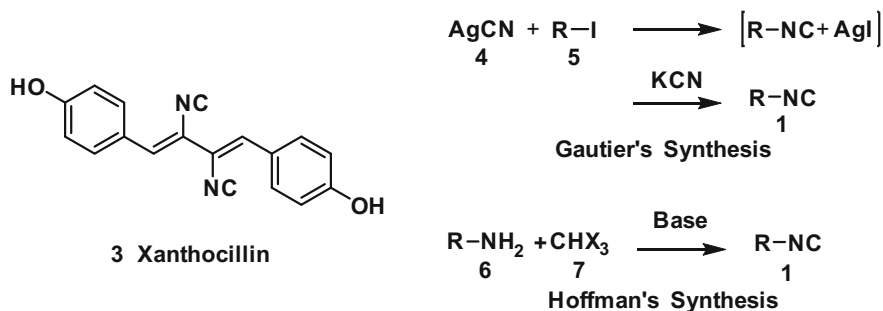
$$AgCN + R-I \longrightarrow \left[R-NC + AgI \right]$$

4   **5**

$$\xrightarrow{KCN} R-NC$$

1

Gautier's Synthesis

$$R-NH_2 + CHX_3 \xrightarrow{Base} R-NC$$

6   **7**   **1**

Hoffman's Synthesis

3 Xanthocillin

Fig. 2 First isolated naturally occurring isocyanide **3** and early attempts at isocyanide synthesis. (X = Halogen)

three or more different compounds as starting materials, which can react to give a product comprising the majority of atoms from the starting materials, making them atom economical reactions. Even though there has been an exponential increase in isocyanide published manuscripts over the past 15–20 years, there is a need for novel isocyanides that can be used in the synthesis of various natural products *via* well-known transformations. Isocyanides are used to introduce new functionality in the molecule, which is overshadowed by utilizing harsh reaction conditions to hydrolyze the amide bond formed. This creates complications in highly functionalized compounds and a major drawback for extended uses. To overcome this challenge a new class of isocyanides, now termed "Convertible Isocyanides" (CICs), have been developed. The name convertible is justified by the role it plays in the MCRs and the post-modifications of the products obtained. The labile amide can be easily removed in a single step and be "converted" into a variety of other moieties with augmented diastereoselectivity. These are, at times, referred to as "Universal Isocyanides".

2   Convertible Isocyanides (CICs)

CICs are isocyanides consisting of a moiety that allows for selective cleavage of the terminal amide; for example in an Ugi post-condensation product. They are widely used in MCRs like the Passerini reaction. The concept of CICs was first introduced by Armstrong in the year 1996. There has been a considerable increase in efforts for the development of CICs and some of the famous CICs are shown in Fig. 3.

Despite the fact that isocyanides **13** and **18** do not have the appearance of CICs, recent reports prove that they can be used as such [10, 11]. Recently, Orru and co-workers have reported on the use of 2-bromo-6-isocyanopyrudine **14** as a universal CIC for MCRs to synthesize some biologically important compounds [12].
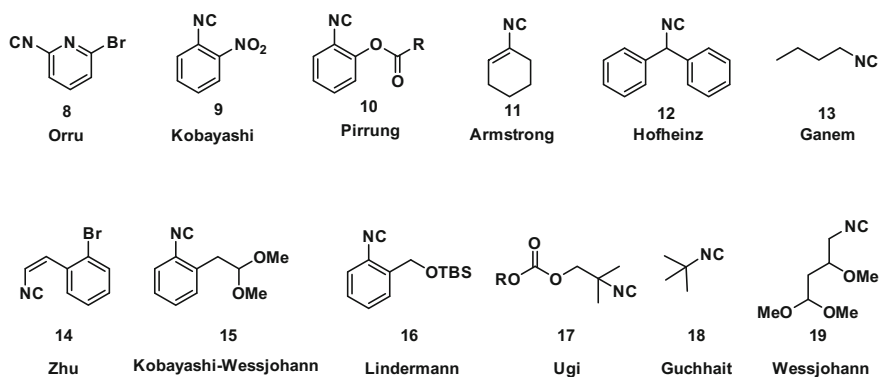

Fig. 3 Some of the more well-known convertible isocyanides

2.1 Methyl Isocyanides

Methyl isocyanides (MICs), can be considered a class of CICs where the substituent is present on the methyl group. Some of the better known MICs are *p*-tolylsulfonylmethyl isocyanide (TosMIC), diphenylmethyl isocyanide (DPMIC), benzotriazol-1-yl-methyl isocyanide (BetMIC), and methyl 2-isocyanoacetate (MICAc). Although the afore-mentioned MICs have been used as CICs in MCRs, the role of methyl isocyanide (MIC) alone as a CIC in an Ugi4-CR was only just first explored and reported by Andreana and co-workers [13] in 2016. The general outline for the reaction using MIC is given in Scheme 1. The required compound **24**, was obtained by mixing benzalde-hyde **22**, methyl isocyanide **20**, propionic acid **21**, and isopropyl amine **23** in methanol at room temperature for 24 h (Scheme 1). Then the dipeptide **24** was reduced to give compound **25** under Bechamp-type reduction conditions. Under microwave reactor conditions, the desired cyclized product **26** was readily obtained. Apart from the role as a CICs, MIC has also been used by the same group as a key component in the one-pot synthesis of diketopiperazine-based natural product Thaxtomin A (±)-*syn* (TA); well known for herbicidal activity [14]. First, the required starting materials 4-nitroindolylacetaldehyde **27**, methyl amine **28**, 3-hydroxyphenylpyruvic acid **29** and compound **20** were reacted in methanol followed by epimerization and cyclization under basic conditions using KOH at 70 °C to afford Thaxtomin A (±-*syn*-**30**) and the corresponding *anti*-diastereomer in a 4:1 ratio (Scheme 2).
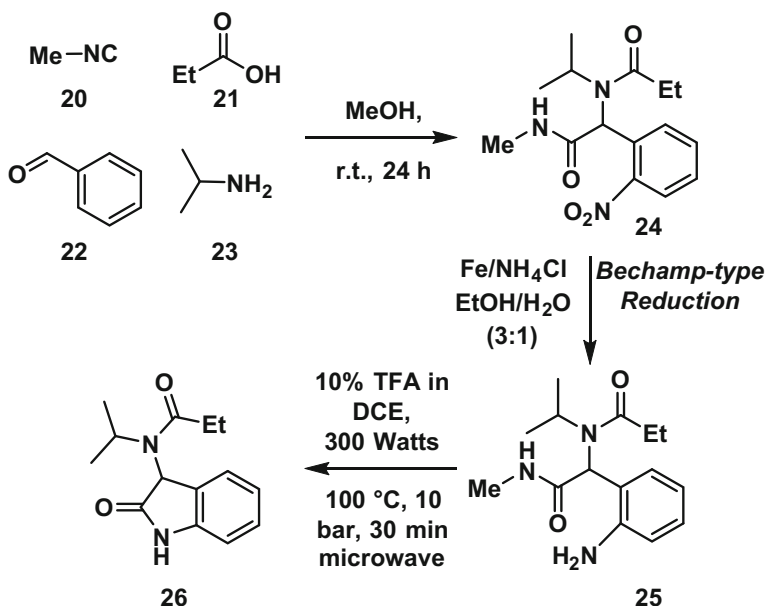
Scheme 1 Synthesis of 3-substituted indolinones
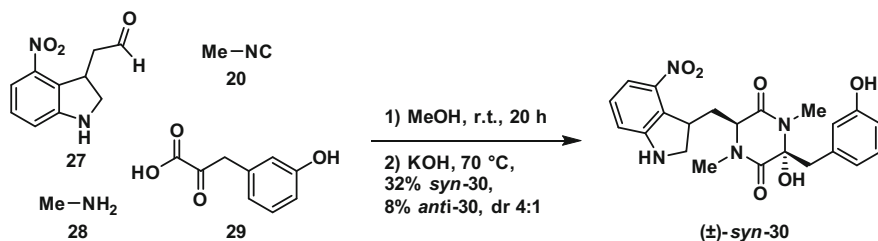
Scheme 2 Synthesis of (±)-Thaxtomin A

3 Asymmetric Control of Isocyanide-Based Multicomponent Reactions

Isocyanides have gained a lot of interest among the organic chemists in view of its importance in the various reactions used to generate a library of compounds and natural products *via* diversity-oriented synthesis (DOS). This is attributed to the generation of new chiral centers implicit in various multicomponent reactions like Passerini, Ugi to name a few. Admitting the fact, that the isocyanide is just one of the components in the MCR, the outcome of the reaction can be controlled by the types of isocyanides used along with the use of chiral catalysts as seen in some of the cases discussed hereafter.

3.1 Diastereoselective Additions

Although isocyanides have been widely used in the MCRs for the synthesis of various scaffolds including natural products and biologically active compounds, there has been a concern lurking over the stereoselectivity of the reactions involving isocyanides. Although there have been stereoselective reactions developed for a myriad of multicomponent coupling reactions, the isocyanide-based reactions seem to be a little more challenging and to date only the asymmetric Passerini reaction has been achieved.

Zhu and co-workers [15] reported the stereoselective synthesis of pyrrolo-pyridine *via* the formation of the oxanorbornene heterocyclic intermediate **35**, followed by the Diels–Alder reaction to give the desired product **36**. Isocyanide **31**, ester **32,** and aldehyde **33** were dissolved in methanol and stirred for 2 h at room temperature followed by the addition of triflouroacetic acid (TFA) at −78 °C and stirred for 20 min which afforded the pyrrolo-pyridine moiety **36** in 80% yield (Scheme 3). Intermediate **35** was obtained in high diastereoselectivity of >90% *de* with an excellent yield of 92%.

Furthermore, Kazmeier et al. [16, 17] reported the use of an enantiomerically pure isocyanide **38** to form oxazoline, ligands (**40a** and **40b**) which were used in

stereoselective additions of Et$_2$Zn to various aldehydes. Due to the instability of the oxazolines, they designed an alternate two-step synthesis of thiazoline ligands (**42a** and **42b**). In both the cases, an excellent stereoselectivity of >92% *ee* was obtained. The oxazoline ligand **40b** was synthesized in one step followed by the addition of Et$_2$Zn to benzaldehyde to yield **43** in 100% yield with a 94% *ee*. The thiazoline ligand **42b** was synthesized in a two-step process followed by the addition of Et$_2$Zn to benzaldehyde to yield **43** in 100% yield with a 92% *ee* (Scheme 4).
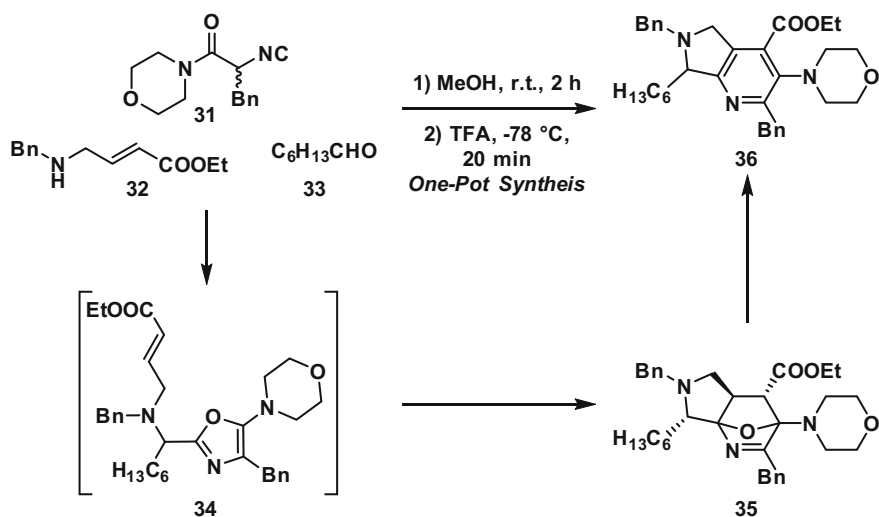
Scheme 3 Synthesis of pyrrolo-pyridine derivatives
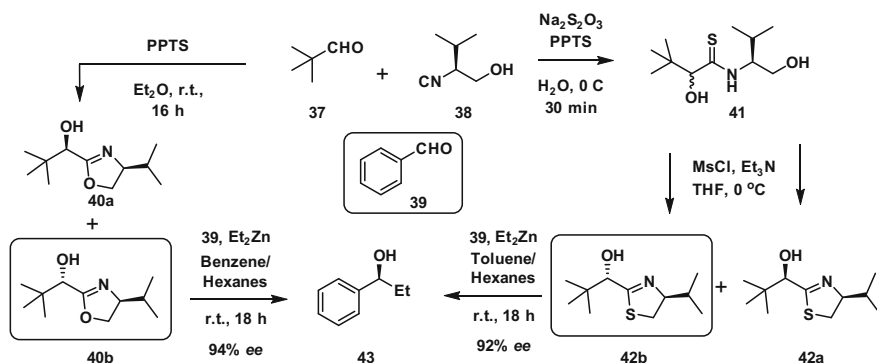
Scheme 4 Synthesis of oxazolines and thiazolines and their application in the addition of Et$_2$Zn to benzaldehydes

3.2 Enantioselective Additions

Denmark and Fan reported the synthesis of α-hydroxyl amides and esters *via* the use of the concept of Lewis base activation of Lewis acids which employs the use of $SiCl_4$ and a chiral bisphosphoramide (*R,R*)-catalyst [18, 19]. The reaction was carried out using various aldehydes **39, 44a–e** and isocyanide **18** in the presence of $SiCl_4$ and 5 mol% of the catalyst to give intermediates **46a–f**. It was determined that the intermediates could give two different products based on the workup route selected (Scheme 5). Under saturated aqueous $NaHCO_3$ workup conditions, α-hydroxyl amides **47a–f** were obtained in good to excellent yields of 76–92% with 82–99% *ee*. Under conditions where methanol and saturated aqueous $NaHCO_3$ workup conditions, it gives rise to α-hydroxyl esters **48a–f** in good to excellent yields of 70–92% with 80–98% *ee* (Table 1).
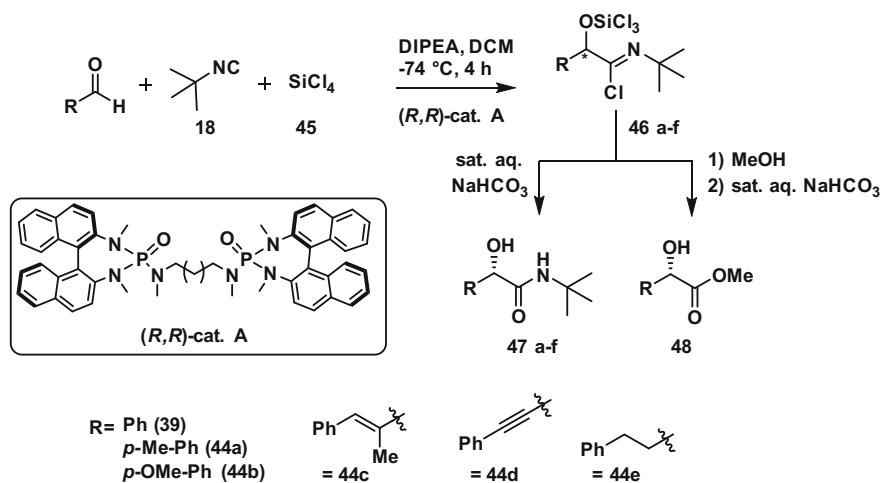
Scheme 5 Enantioselective addition of isocyanides to aldehydes using catalyst

Table 1 Reaction profile of Lewis base activation of Lewis acids

Product	Aldehyde	Yield (%)	er
46a	**39**	91	99.9:0.1
46b	**44a**	91	99.9:0.1
46c	**44b**	89	98.3:1.7
46d	**44c**	86	67.4:32.6
46e	**44d**	76	77.0:23.0
46f	**44e**	92	81.9:18.1

4 Stereoselectivity of Passerini Reactions

4.1 Diastereoselective Passerini Reactions

In recent times, there have been very few reports on the diastereoselective Passerini reaction. Berlozecki and co-workers have reported on Passerini reactions wherein they used N^α protected amino acids, cyclic ketones, and isocyanides [20]. They achieved exceptional diastereoselectivity which they attribute to the protecting groups used to protect the amine group of the amino acids. L-phenylalanine (Scheme 6) with *Boc-* or phthaloyl and trityl protected amine **49** gave a selectivity of 99% *de* with yields of 86, 84, and 66% respectively (Table 2). Whereas the D-valine (Scheme 7) with *Boc-* and phthaloyl protected amine **53** gave a selectivity of 85–96% *de* with the yields of 79 and 74% respectively (Table 3). Albeit high yields and diastereoselectivity was achieved, this approach had a drawback; ketones other than cyclohexanone were not tested for the transformation. They were later tested by Simila and Martin [21].

4.2 Enantioselective Passerini Reactions

Over the years many enantioselective Passerini reactions have been reported, but only a few of them have resulted in good enatioselectivity [22–25]. A recent report by Szymanski imposes the utilization of *wheat germ lipase* enzyme to resolve the Passerini product by hydrolysis of the racemic mixture (desymmetrization) [26]. Although two different products were obtained, the selectivity was excellent giving **58 (R)** in 70% *ee* and **59** in 99% *ee*.

The α-hydroxyl amide **59** was then converted into the corresponding amino acid by six additional steps. First, compound **59** was mesylated, followed by nucleophilic substitution of azide and hydrogenation using Pd/C in methanol afforded the amide **60** (Scheme 8). This was later subjected to hydrolysis under acidic and basic conditions followed by hydrogenation using Pd/C in methanol to yield the amino acid [27] in 71% with an excellent enatioselectivity (98% *ee*).

Similar to the Lewis base catalyzed Passerini reaction, Lewis acid catalyzed Passerini reactions were initially described by Domling [22] and Schreiber groups (2004) [28] using titanium- and copper-based catalysts respectively. However in each case, when these approaches were used, the drawback was lower enantioselectivity when non-chelating agents were employed. This was further explored by Zhu and co-workers (2008) using aluminum-based catalysts [29]. Several catalysts and reaction conditions were optimized which showed that the use of salen-aluminum catalyst could render good enantioselectivities. The reactions were carried out by mixing the aldehyde, carboxylic acid, and the isocyanide in toluene at −40 °C and stirring for 48 h (Scheme 9). The reaction yields were moderate to good (59–70%) affording the amide **65a–i** with moderate to excellent
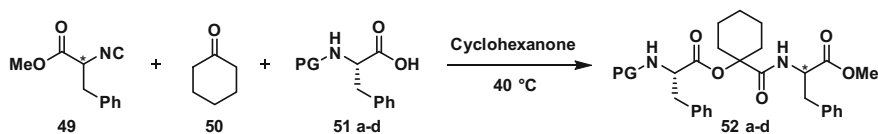
Scheme 6 Diastereoselective P-3CRs of L-phenylalanines

**Table 2** Reaction profile of P-3CRs of L-phenylalanines

Product	PG	Isocyanide configuration	Time (h)	Yield (%)	dr
52a	Bz	S	20	39	94:6
		R	20	48	7:93
52b	Trt	S	24	66	<1:99
		R	20	49	>99:1
52c	Phth	S	14	84	<1:99
		R	40	82	>99:1
52d	Boc	S	20	86	99:1
		R	16	>99	1:99
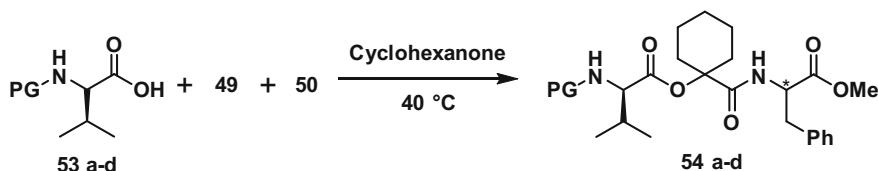
Scheme 7 Diastereoselective P-3CRs of D-valines

enatioselectivity of 63–99% *ee* (Table 4). Screening a set of diversified reagents revealed that the enatioselectivity is greatly dependent on the structures of the aldehydes and isocyanides. It was determined that the carboxylic acid also played a minor role in selectivity. The *S*-selectivity of the product is favored by the attack of the isocyanide on the *Re*-face of the aldehyde.

Table 3 Reaction profile of P-3CRs of D-valines

Product	PG	Isocyanide configuration	Time (h)	Yield (%)	dr
54a	Bz	S	20	30	16:84
		R	35	42	81:19
54b	Cbz	S	27	52	18:82
		R	25	89	87:13
54c	Phth	S	17	74	85:15
		R	20	78	18:82
54d	Boc	S	20	79	5:95
		R	23	87	96:4
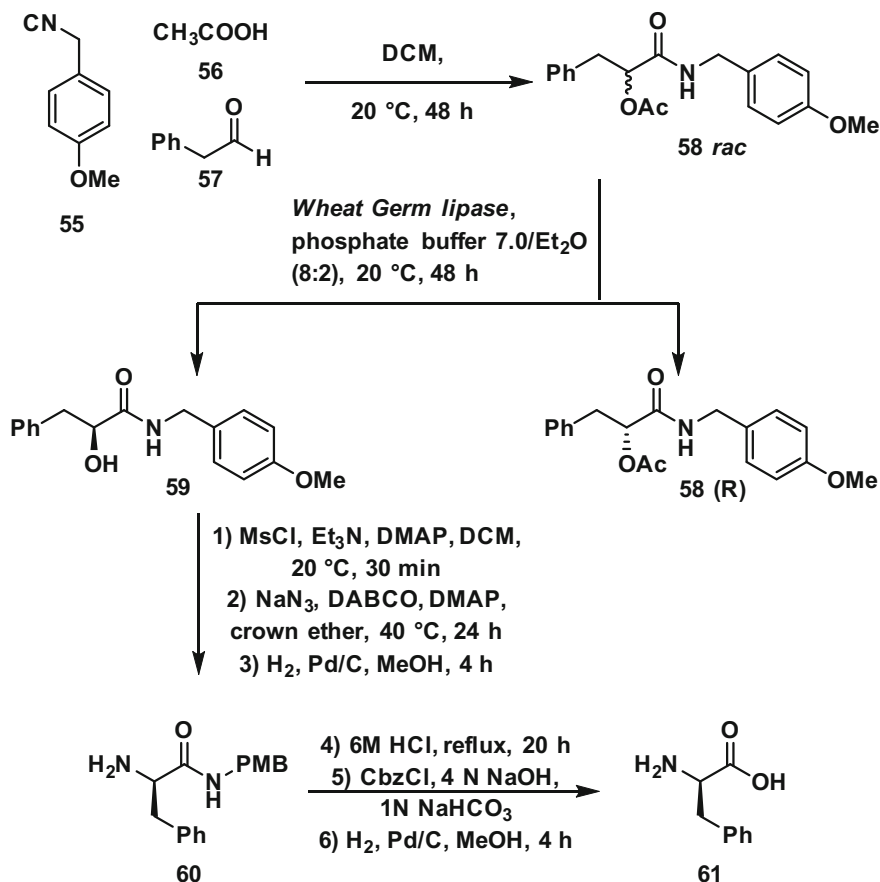
Scheme 8 Enzymatic resolution of the racemic Passerini product

Zhu and co-workers further modified the catalyst and replaced the carboxylic acid with the hydrazoic acid **67** which gave a cyclic product following a 1,5-dipolar cyclization (Scheme 10) instead of the acylation as observed in the previous case [24]. The products were achieved in high yields of 76–95% with excellent enatioselectivity up to 95% *ee* (Table 5). But this approach is also limited due to the fact that it is more suitable for aliphatic aldehydes and very few aromatic aldehydes.

A very recent report by Bin Tan's group has shown that the use of phosphoric acid based catalysts could actually enhance the enantioselectivity of Passerini reactions and could be applied to wide range of aldehydes and carboxylic acids. They have optimized the reaction conditions by screening a wide range of catalysts [30] and selecting the catalyst with the best results (**CP6**) followed by its use to further explore the reaction profile. Moderate to excellent yields of 45–99% with enantioselectivity of 84–99% *ee* were observed (Fig. 4).
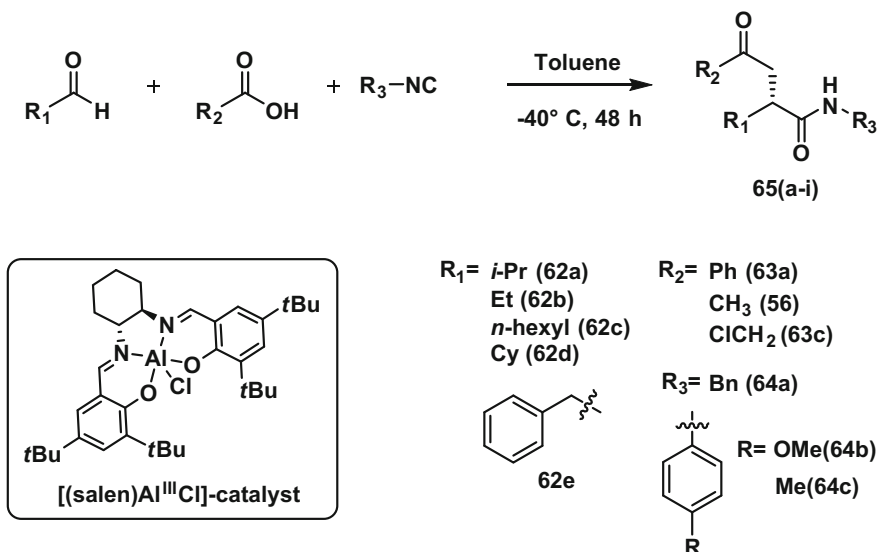
Scheme 9 Salen-aluminum catalyzed enantioselective P-3CR

Table 4 Reaction profile of Lewis acid catalyzed Passerini reaction

Product	Aldehyde	Acid	NC	Yield (%)	ee (%)
65a	62a	63a	64a	70	63
65b	62a	56	64a	59	63
65c	62a	56	64b	63	84
65d	62b	63c	64b	66	87
65e	62c	63c	64b	67	73
65f	62d	63c	64b	59	87
65 g	62e	63c	64b	68	71
65 h	62a	63c	64c	61	81
65i	62a	63c	64b	64	99

5 Stereoselective Ugi Reaction

Successful attempts at the diastereoselectivity and enatioselectivity of the Passerini reaction encouraged the organic chemists' community to look for ways to access similar success in Ugi reactions. List and co-workers [31] demonstrated the lack of enatioselectivity in Ugi reactions. Recent reports by Zhu and co-workers have given a ray of hope when they reported the enantioselective Ugi reaction [32, 33].

In the past few years, efforts from various research groups have shown that the diastereoselective outcome of the Ugi reaction has been fruitful. Kobayashi and co-workers [34, 35] demonstrated the synthesis of Omuralide, a proteasome
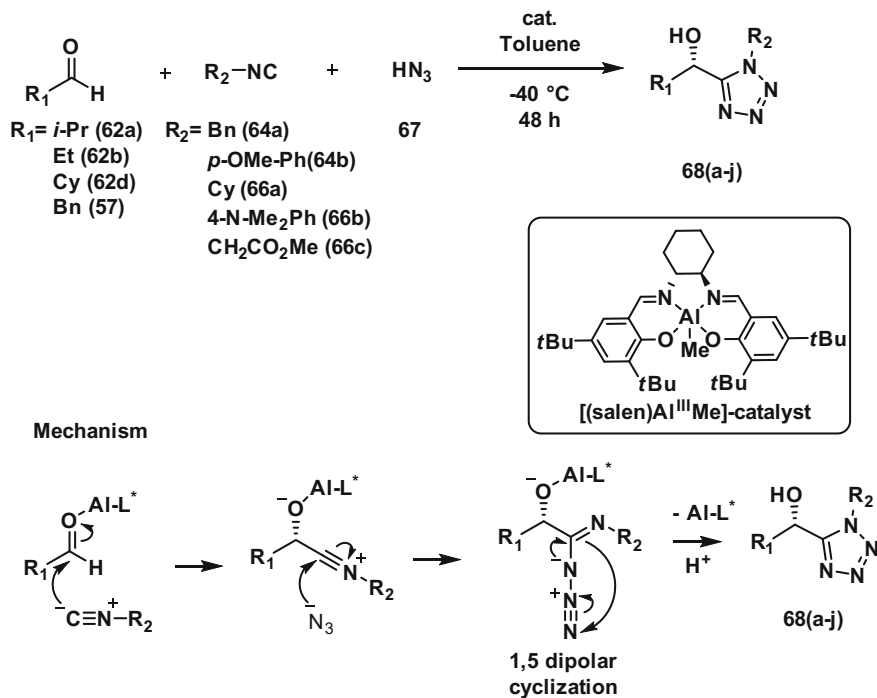
Scheme 10 Enantioselective synthesis of tetrazoles *via* P-3CR and the proposed mechanism

Table 5 Reaction profile of synthesis of tetrazoles *via* P-3CR

Product	Aldehyde	Isocyanide	Yield (%)	*ee* (%)
68a	**62a**	**64b**	90	85
68b	**62a**	**66a**	90	95
68c	**62b**	**64b**	88	87
68d	**62c**	**64b**	90	91
68e	**62a**	**64a**	76	92
68f	**57**	**64b**	97	64
68 g	**62c**	**66a**	93	84
68 h	**62c**	**66b**	95	94
68i	**62c**	**64a**	91	83
68j	**62c**	**66c**	93	75

inhibitor. The synthesis involved the use of beta-hydroxy gamma-keto acid. The starting materials {protected β-hydroxyγ-keto acid intermediate (**I**), isocyanide **15** and amine **71**} underwent the Ugi four center three-component reaction (Ugi-4C-3CR) to afford anti-**72** predominantly with a yield of 78% with the formation of single diastereomer (Scheme 11). This compound **72** further was used to complete the synthesis of Omuralide **74**.
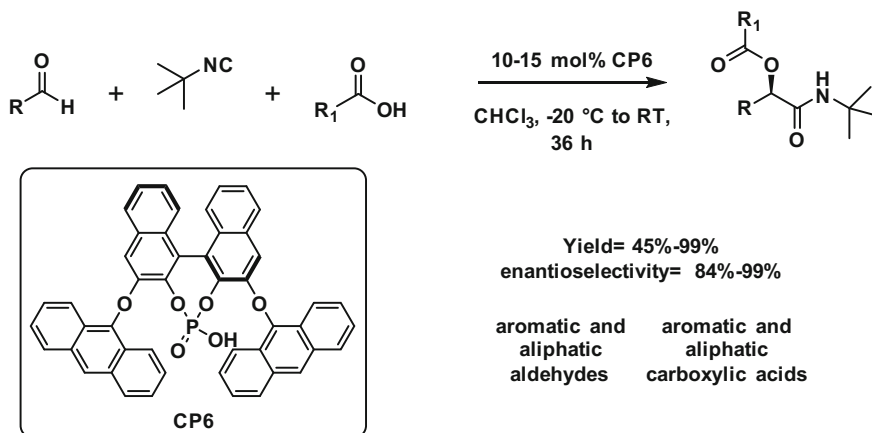
Fig. 4 Phosphoric acid catalyzed asymmetric Passerini reaction
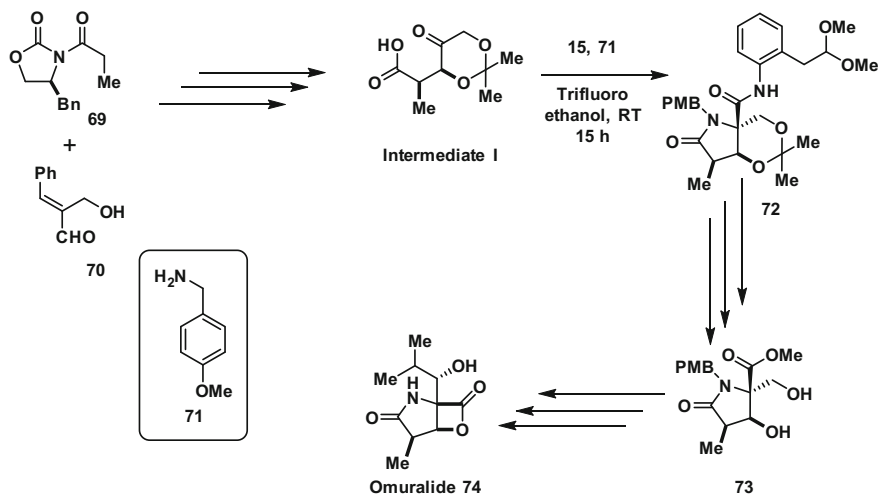
Scheme 11 Synthesis of Omuralide **74** *via* Ugi reaction

Use of chiral isocyanides was reported by Nenjadenko and co-workers to achieve diastereoselectivity for the Ugi MCRs. They have used the 4-methyl-2,6,7-trioxabicyclo[2.2.2]octyl derivatives or the chiral oxabicyclooctyl (OBO)esters **78a–e** (previously synthesized by them) to study the role of chiral isocyanide in the stereoselectivity [36]. The OBO esters played a key role in controlling the racemization. The general reaction module for this reaction is, stirring a mixture of amine, aldehyde or ketone, carboxylic acid, and the OBO esters of the isocyanide in methanol at room temperature. They observed the retention of configuration of the Ugi products **80a–**k (Scheme 12). To further
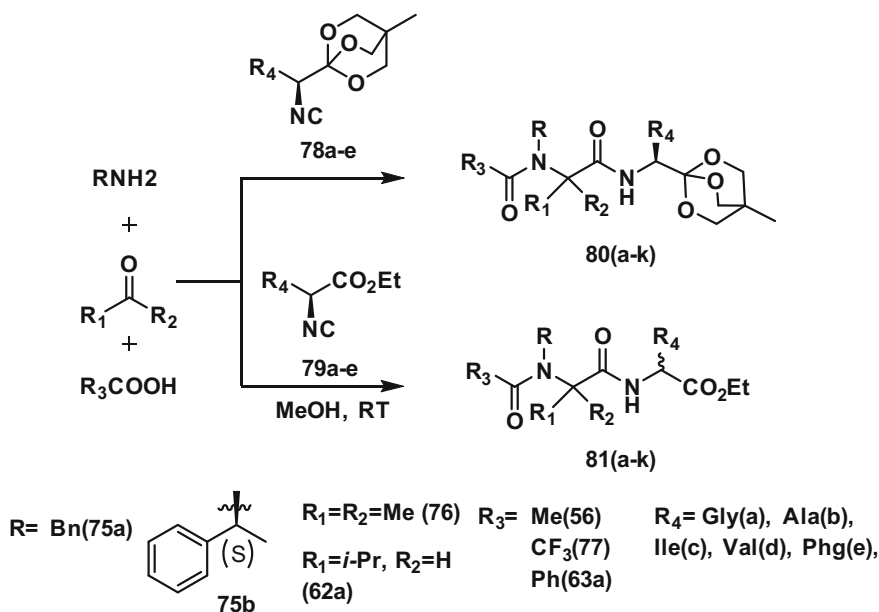
Scheme 12 Diastereoselective Ugi-4CR using chiral OBO esters
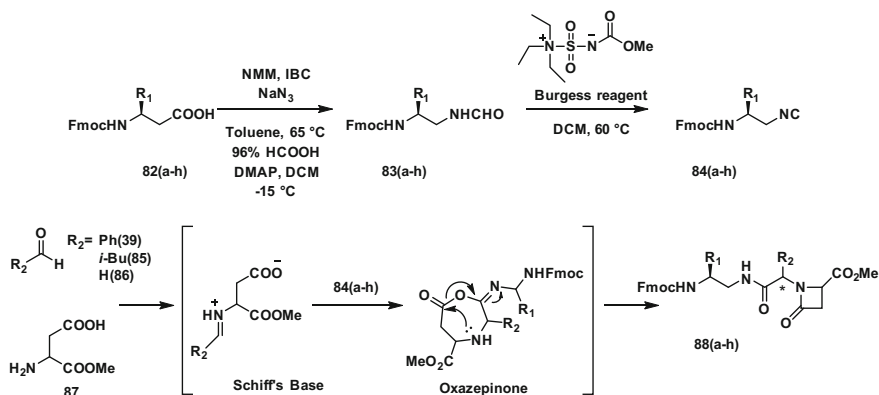
Scheme 13 Synthesis of β-lactams using chiral amino alkyl isocyanides

investigate the reaction profile they used the esters **79a–e**, surprisingly racemization at the chiral center on the isocyanide was observed. This observation suggests that the OBO plays an important role to prevent the racemization.

Along the same lines of using the chiral isocyanides, Sureshbabu and co-workers [37] attempted the utilization of the N^β-Fmoc amino alkyl isocyanides, which were previously synthesized by the same group. They employed those isocyanides for the

synthesis of a series of β-lactam peptidemimics *via* Ugi 4C-3CR (Scheme 13). The reactions were conducted using the isocyanides **84a–h**, aldehydes and methyl ester of the aspartic acid **87** to give the peptidomimics **88a–h** in moderate to good yields (53–78%) with excellent diastereoselectivity {*de* between 85 and 100% (Table 6)}.

The high selectivity can be attributed to the formation of the 7-membered oxazeoinone intermediate [38]. Furthermore, the Fmoc group on the final product can be deprotected and be used in other reactions to synthesize various complex natural products.

A collaborative work between Orru and van der Eycken [39] accomplished a new level of diastereoselectivity. They utilized an aldehyde-carboxylic acid derivative of biphenyl **89**, amine (**71**) and the isocyanide (**18**) for the microwave assisted Ugi 4C-3CR (Scheme 14). The reaction gave a dibenzo[c,e]azepinone derivative as a single diastereomer with an excellent yield of 82%. The reaction proceeds through the formation of iminium ion which is attacked by the isonitrile to form nitrilium ion followed by the intramolecular acylation to give the intermediate IV. This intermediate upon Mumm rearrangement forms the (R_a, S) isomer of the product **90**.

After all the years of hard work put in by various research groups to achieve the enantioselectivity for the Ugi reactions, the initial breakthrough was provided by Zhu's group in 2012, wherein they used chiral phosphoric acids (CPAs) as catalysts for the one-pot four-component Ugi type reaction [32]. The reactions were conducted by adding the first three components (aldehyde, amine and isocyanide) in dichloromethane and stirring for 24 h at −35 °C followed by addition of the fourth component (unsaturated acid chloride) and base and refluxing the mixture for 5 h in toluene. The process gave the desired single diastereomer in yields ranging from 41 to 94% with the enantioselectivity of 81–94% *ee*.

Following up on this work they further explored the role of the CPAs in achieving excellent enatioselectivity for Ugi reactions [33]. There were two different approaches that they followed in order to achieve the desired product. In the first approach, they followed a two-component version wherein they reacted the preformed furan (from carboxaldehyde and amine) and the isocyanide followed by the addition of the catalyst (**CP–NO₂**). Second approach was mixing all the components in a single pot followed by addition of the catalyst (**CP–NO₂**). In both the

Table 6 Reaction profile for the synthesis of β-lactams **88a–h**

Product	R_1	Aldehyde	Yield (%)	*dr*
88a	–H	**39**	65	99:1
88b	–Me	**39**	78	96:4
88c	–Bn	**85**	69	85:15
88d	–*i*-Pr	**86**	72	90:10
88e	–CH₂–OBn	**39**	76	89:11
88f	–CH₂–COOBn	**85**	63	100:0
88 g	–(CH₂)₄–NHZ	**86**	59	98:2
88 h	Pyrrolidine NC	**39**	53	95:5
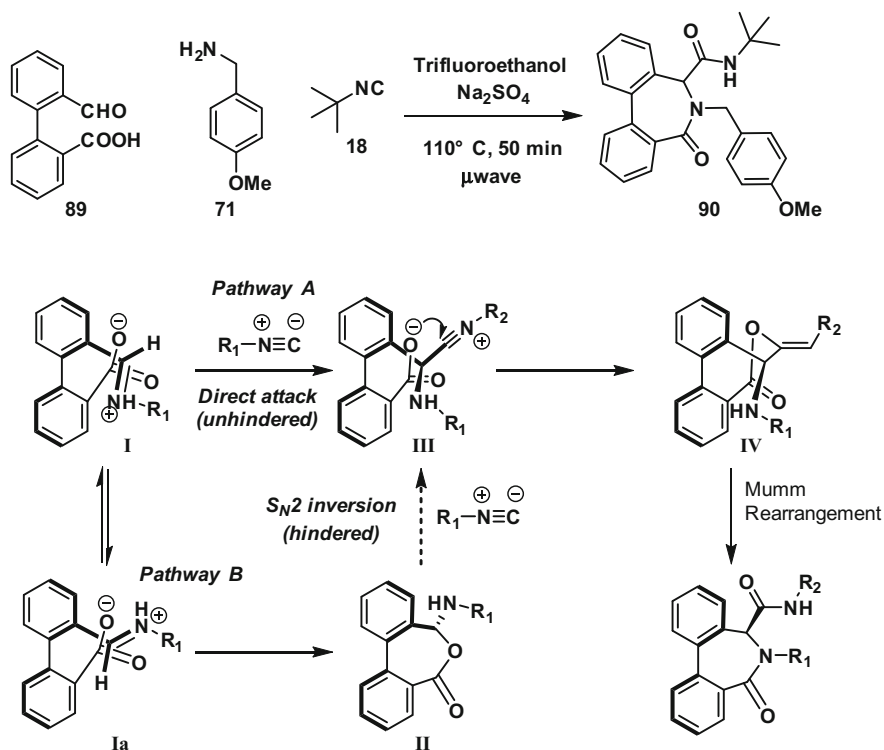
Scheme 14 Microwave assisted Ugi-4C-3CR and the proposed mechanism

cases excellent yields (74–98%) and enantioselectivities (82–97% *ee*) were obtained (Fig. 5).

6 Carbohydrates as Chiral Auxiliaries

There have been efforts in improving the stereoselectivity in the Ugi 4CRs to synthesize the substituted amino acids. Horst Kunz and co-workers have reported the use of carbohydrates as chiral auxiliaries keeping other constituents achiral [1, 2]. The α-D and L-amino acids can be achieved by simply switching the amines between the galactosamine and the arabinosylamine respectively.

When the galactosamine **91** was reacted with isocyanide **18** and formic acid **94**, variety of aldehydes in the presence of ZnCl$_2$ and tetrahydrofuran at lower temperatures, N-galactosyl amino acid amide derivative **95a–i** were obtained in excellent yields ranging from 80 to 93% and high enantioselectivity (only one chiral center) of 91–94% (Table 7). The derivatives were further processed for the cleavage of the N-glycosidic bond following a two-step acidic hydrolysis
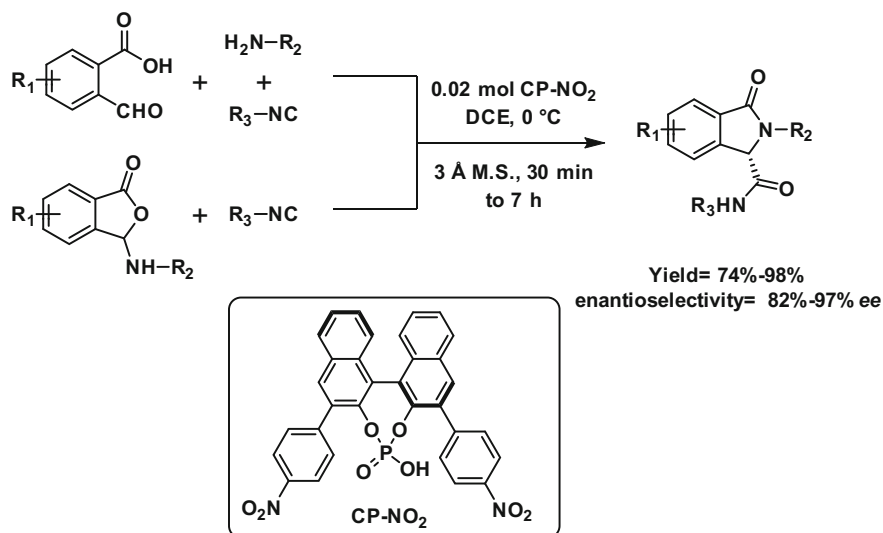
Fig. 5 Chiral phosphoric acid catalyzed enantioselective Ugi reaction

Table 7 Reaction profile for Ugi reaction using D-galactosamine auxiliary

Product	R_1	Aldehyde	Reaction temp (°C)/time	R:S	Yield (%) of pure R
95a	C_3H_7	**92a**	−78/2d	94:6	80
95b	-i-Pr	**62a**	−78/2d	95:5	86
95c	t-Bu	**92b**	−25/3d	96:4	80
95d	Bn	**57**	−78/8	95:5	80
95e	(furan-CH₃)	**92c**	−25/24 h	95:5	90
95f	(thiophene-CH₃)	**92d**	-25/24 h	96:4	93
95 g	Ph	**92e**	0/8d	91:9	81
95 h	p-Cl–Ph	**92f**	−25/24 h	97:3	92
95i	p–NO₂–Ph	**92 g**	0/4 h	94:6	91

(Scheme 15) giving the galactose template **96** and the desired free α-D-amino acids **98a–i** in quantitative yields [1].

When the Ugi reaction was conducted using arabinosylamine as the chiral amine, following similar reaction conditions, the derivatives [2] N-formyl-N-arabinosyl amino acid amides **100a–e** were obtained in excellent yields ranging from 85 to 95% and enantioselectivity of >96% respectively as observed in the case of galactosamine (Table 8). The desired L-amino acid derivatives **103a–e** and the arabinose template **101** were obtained in quantitative yields *via* a two-step acidic hydrolysis (Scheme 16).
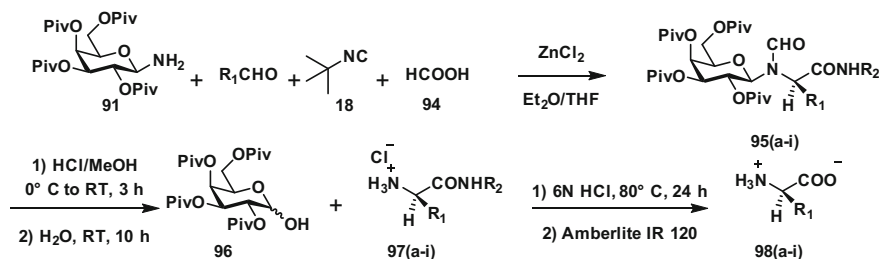
Scheme 15 Galactose derived chiral auxiliary based asymmetric Ugi reaction

Furthermore efforts were put in by Pellicciari's group [40], synthesizing the glycine derivatives with good to moderate yields and excellent diastereoselectivity (Fig. 6) and Ugi and co-workers [41] used a thiosugar to synthesize the D-leucine derivatives with excellent yields and enantioselectivity (Fig. 7). The selectivity could have been achieved due to the formation of the intermediate which directs the attack of the isocyanide at the imine center.

Use of sugars as chiral auxiliaries was further explored by Zeigler and co-workers [42] by employing carbohydrate-based isocyanides in the Ugi and Passerini MCRs to synthesize a library of glycopeptides. The reactions were performed using anomeric glucosyl isocyanides **104a, b** and protected

Table 8 Reaction profile for Ugi reaction using D-arabinosamine auxiliary

Product	R$_1$	Aldehyde	Reaction temp (°C)./time	R:S	Yield (%) of pure S
100a	t-Bu	**92b**	−25/3d	97:3	85
100b	Bn	**57**	−78/24 h	97:3	87
100c	p-Cl–Ph	**92f**	−25/24 h	98:2	91
100d	(furyl)	**92c**	−25/24 h	96:4	85
100e	(thienyl)	**92d**	−25/24 h	4:96	95 (2R)
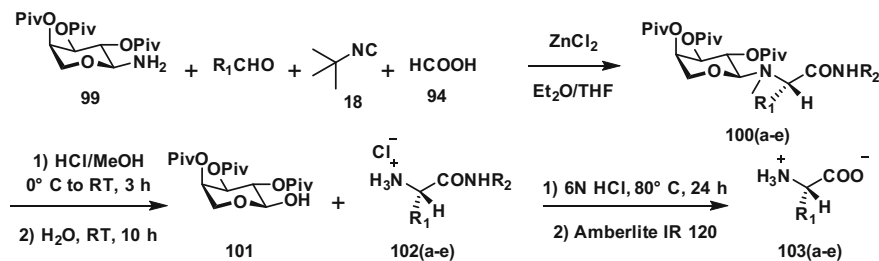
Scheme 16 Stereoselective synthesis of L-amino acid derivatives *via* D-arabinose auxiliary
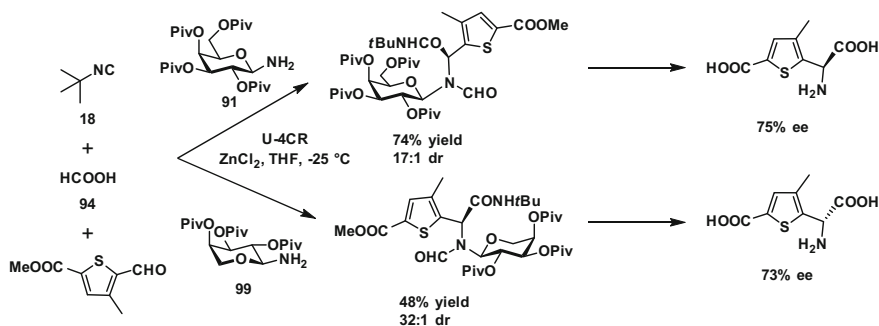
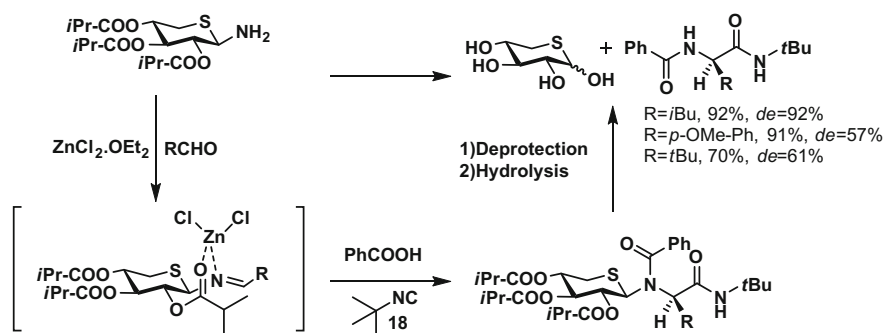Fig. 6 Pellicciari's stereoselective synthesis of glycone derivatives

Fig. 7 Stereoselective synthesis of D-Leucine derivatives using thiosugars as chiral auxiliary

2-deoxy-2-isocyano-β-D-glucopyranose **104c** with various combinations of aldehydes, carboxylic acids, and amines (Fig. 8). The reactions were characterized by longer reaction time, lower yields and more importantly poor diastereoselectivity with a couple of exceptions as evident from Tables 9 to 10.

The 2-isocyano compound faired well in the Passerini reactions with higher yields and better reaction times. Although the reactions did not result in desired
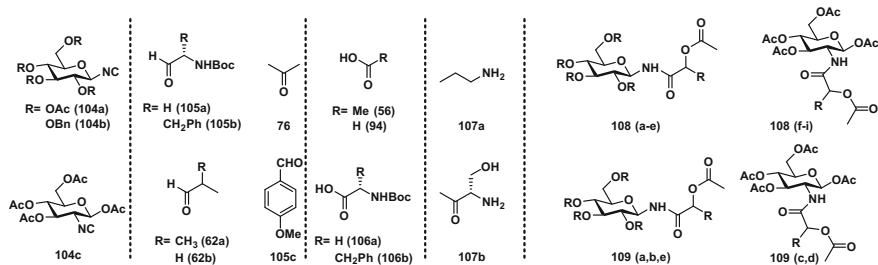
Fig. 8 Passerini and Ugi reactions of isocyanoglucoses

Table 9 Reaction profile for Passerini reaction leading to products **108a-i**

Product	Isocyanide	Aldehyde	Acid	Reaction temp/time	Yield %, *dr*
108a	**104a**	**105a**	**56**	RT/3d	23%, 55:45
108b	**104a**	**105b**	**56**	RT/3d	41%, 58:42
108c	**104a**	**62b**	**56**	RT/24 h	80%, 50:50
108d	**104b**	**105a**	**56**	RT/6d	31%, 57:43
108e	**104b**	**105b**	**56**	RT/8d	35%, 52:48
108f	**104c**	**105b**	**56**	RT/4d	57%, 60:40
108 g	**104c**	**62a**	**56**	RT/8d	23%, 53:47
108 h	**104c**	**62a**	**106a**	RT, 28 h	90%, 53:47
108i	**104c**	**76**	**94**	0 °C, 3 h	14%, –

Table 10 Reaction profile for Ugi reaction leading to products **109a–e**

Product	Isocyanide	Aldehyde	Acid	Amine	Reaction temp/time	Yield %, *dr*
109a	**104a**	**62a**	**106a**	**107a**	RT/37d	22%, 55:45
109b	**104b**	**62a**	**106a**	**107a**	RT/25d	35%, 60:40
109c	**104c**	**62a**	**106a**	**107a**	RT/30d	31%, –
109d	**104c**	**62a**	**106b**	**107a**	RT/18d	19%, –
109e	**104a**	**62a**	**107b**	**107b**	55 °C/11 h	15%, 63:37

diastereoselectivity, there was room for improvement but unfortunately not many researches have followed this work, and thus far not much work has been done in this area. This approach could be a useful tool for synthesizing complex glycopeptides.

7　Conclusions

This chapter describes the recent developments in the application of isocyanides in the multicomponent reactions (MCR) forming chiral products It also describes the first ever application of methyl isocyanide as a convertible isocyanide, which could be useful in the synthesis of various natural product intermediates via Ugi and Passerini reactions involving the multicomponent transformations. In order to achieve higher stereoselectivity, various catalysts were used giving a sign of importance of catalysis in chemical tranformations *en route* to the synthesis of natural products of biological significance. The post-modifications of Ugi and Passerini products can give rise to esters, acids, and thioesters.

Recent reports have shown that the high enatioselectivity can be achieved in the Ugi reactions, opening up the avenues to further explore the results that can be implemented in the synthesis of intermediates for complex natural products.

Use of carbohydrates as chiral auxiliaries has revolutionized the utility of the multicomponent reactions owing to their ease of access and further modifications in those reactions. The carbohydrates have been used for the synthesis of various amino acid derivatives and can further be used to synthesize biologically important amino acid based molecules.

We summarize the chapter with an anticipation that this will galvanize the synthetic chemists to focus on further development of applications of isocyanides in various fields of chemistry to synthesize compounds of biological significance and compounds that can make an impact on environment in a positive way.

References

1. Kunz H, Pfrengle W (1988) Asymmetric synthesis on carbohydrate templates: stereoselective Ugi-synthesis of.alpha.-amino acid derivatives. J Am Chem Soc 110(2):651–652. doi:10.1021/ja00210a084
2. Kunz H, Pfrengle W, Sager W (1989) Carbohydrates as chiral templates: diastereoselective Ugi synthesis of (S)-amino acids using O-acylated D-arabinopyranosylamine as the auxiliary. Tetrahedron Lett 30(31):4109–4110. doi:10.1016/S0040-4039(00)99334-1
3. Rüchardt C, Meier M, Haaf K, Pakusch J, Wolber EKA, Müller B (1991) The isocyanide-cyanide rearrangement; mechanism and preparative applications. Angew Chem Int Ed Engl 30(8):893–901. doi:10.1002/anie.199108933
4. Ryu I, Sonoda N, Curran DP (1996) Tandem radical reactions of carbon monoxide, isonitriles, and other reagent equivalents of the geminal radical acceptor/radical precursor synthon. Chem Rev 96(1):177–194. doi:10.1021/cr9400626
5. Josien H, Ko S-B, Bom D, Curran DP (1998) A general synthetic approach to the (20s)-camptothecin family of antitumor agents by a regiocontrolled cascade radical cyclization of aryl isonitriles. Chemistry A Eur J 4(1):67–83. doi:10.1002/(SICI)1521-3765(199801)4:1<67:AID-CHEM67>3.0.CO;2-F
6. Kobayashi Y, Fukuyama T (1998) Development of a novel indole synthesis and its application to natural products synthesis. J Heterocycl Chem 35(5):1043–1056. doi:10.1002/jhet.5570350504
7. Lieke W (1859) Ueber das cyanallyl. Justus Liebigs Annalen der Chemie 112(3):316–321. doi:10.1002/jlac.18591120307
8. Gautier A (1867) Ueber die Einwirkung des Chlorwasserstoffs u. a. auf das Aethyl- und Methylcyanür. Justus Liebigs Annalen der Chemie 142(3):289–294. doi:10.1002/jlac.18671420304
9. Hofmann AW (1867) Ueber eine neue Reihe von Homologen der Cyanwasserstoffsäure. Justus Liebigs Annalen der Chemie 144(1):114–120. doi:10.1002/jlac.18671440116
10. Le HV, Fan L, Ganem B (2011) A practical and inexpensive 'convertible' isonitrile for use in multicomponent reactions. Tetrahedron Lett 52(17):2209–2211. doi:10.1016/j.tetlet.2010.11.156
11. Guchhait SK, Madaan C (2010) Towards molecular diversity: dealkylation of tert-butyl amine in Ugi-type multicomponent reaction product establishes tert-butyl isocyanide as a useful convertible isonitrile. Org Biomol Chem 8(16):3631–3634. doi:10.1039/C0OB00022A
12. van der Heijden G, Jong JAW, Ruijter E, Orru RVA (2016) 2-Bromo-6-isocyanopyridine as a Universal Convertible Isocyanide for Multicomponent Chemistry. Org Lett 18(5):984–987. doi:10.1021/acs.orglett.6b00091
13. Maddirala AR, Andreana PR (2016) Synthesis of 3-Substituted 2-Indol-inones by a multicomponent coupling isocyanide-dependent microwave-assisted intramolecular transamidation process. Eur J Org Chem 1:196–209. doi:10.1002/ejoc.201501273

14. Bourgault JP, Maddirala AR, Andreana PR (2014) A one-pot multicomponent coupling/cyclization for natural product herbicide (+/−)-thaxtomin A. Org Biomol Chem 12(41):8125–8127. doi:10.1039/C4OB01148A

15. Gámez-Montaño R, González-Zamora E, Potier P, Zhu J (2002) Multicomponent domino process to oxa-bridged polyheterocycles and pyrrolopyridines, structural diversity derived from work-up procedure. Tetrahedron 58(32):6351–6358. doi:10.1016/S0040-4020(02)00634-8

16. Bauer M, Kazmaier U (2006) A new, modular approach towards 2-(1-hydroxyalkyl) oxazolines, effective bidentate chiral ligands. J Organomet Chem 691(10):2155–2158. doi:10.1016/j.jorganchem.2005.10.048

17. Bauer M, Maurer F, Hoffmann SM, Kazmaier U (2008) Hydroxyalkyl thiazolines, a new class of highly efficient ligands for carbonyl additions. Synlett 20:3203–3207. doi:10.1055/s-0028-1087366

18. Denmark SE, Fan Y (2003) The first catalytic, asymmetric α-additions of isocyanides. lewis-base-catalyzed, enantioselective passerini-type reactions. J Am Chem Soc 125 (26):7825–7827. doi:10.1021/ja035410c

19. Denmark SE, Fan Y (2005) Catalytic, enantioselective α-additions of isocyanides: lewis base catalyzed passerini-type reactions. J Org Chem 70(24):9667–9676. doi:10.1021/jo050549m

20. Berłożecki S, Szymanski W, Ostaszewski R (2008) α-Amino acids as acid components in the Passerini reaction: influence of N-protection on the yield and stereoselectivity. Tetrahedron 64 (41):9780–9783. doi:10.1016/j.tet.2008.07.064

21. Simila STM, Martin SF (2008) Applications of the Ugi reaction with ketones. Tetrahedron Lett 49(29–30):4501–4504. doi:10.1016/j.tetlet.2008.05.073

22. Kusebauch U, Beck B, Messer K, Herdtweck E, Dömling A (2003) Massive parallel catalyst screening: toward asymmetric MCRs. Org Lett 5(22):4021–4024. doi:10.1021/ol035010u

23. Enders D, Grondal C, Hüttl MRM (2007) Asymmetric organocatalytic domino reactions. Angew Chem Int Ed 46(10):1570–1581. doi:10.1002/anie.200603129

24. Yue T, Wang M-X, Wang D-X, Zhu J (2008) Asymmetric synthesis of 5-(1-Hydroxyalkyl) tetrazoles by catalytic enantioselective passerini-type reactions. Angew Chem Int Ed 47 (49):9454–9457. doi:10.1002/anie.200804213

25. Cao C-L, Sun X-L, Zhou J-L, Tang Y (2007) Enantioselectively organocatalytic michael addition of ketones to alkylidene malonates. J Org Chem 72(11):4073–4076. doi:10.1021/jo070070p

26. Szymanski W, Ostaszewski R (2006) Multicomponent diversity and enzymatic enantiose-lectivity as a route towards both enantiomers of α-amino acids—a model study. Tetrahedron Asymmetry 17(18):2667–2671. doi:10.1016/j.tetasy.2006.09.014

27. Szymanski W, Zwolinska M, Ostaszewski R (2007) Studies on the application of the Passerini reaction and enzymatic procedures to the synthesis of tripeptide mimetics. Tetrahedron 63 (32):7647–7653. doi:10.1016/j.tet.2007.05.044

28. Andreana PR, Liu CC, Schreiber SL (2004) Stereochemical control of the passerini reaction. Org Lett 6(23):4231–4233. doi:10.1021/ol0482893

29. Wang S-X, Wang M-X, Wang D-X, Zhu J (2008) Catalytic enantioselective passerini three-component reaction. Angew Chem 120(2):394–397. doi:10.1002/ange.200704315

30. Zhang J, Lin S-X, Cheng D-J, Liu X-Y, Tan B (2015) Phosphoric acid-catalyzed asymmetric classic passerini reaction. J Am Chem Soc 137(44):14039–14042. doi:10.1021/jacs.5b09117

31. Pan SC, List B (2008) Catalytic three-component Ugi reaction. Angew Chem Int Ed 47 (19):3622–3625. doi:10.1002/anie.200800494

32. Su Y, Bouma MJ, Alcaraz L, Stocks M, Furber M, Masson G, Zhu J (2012) Organocatalytic enantioselective one-pot four-component Ugi-type multicomponent reaction for the synthesis of epoxy-tetrahydropyrrolo[3,4-b]pyridin-5-ones. Chemistry A Eur J 18(40):12624–12627. doi:10.1002/chem.201202174

33. Zhang Y, Ao Y-F, Huang Z-T, Wang D-X, Wang M-X, Zhu J (2016) Chiral phosphoric acid catalyzed asymmetric Ugi reaction by dynamic kinetic resolution of the primary multicom-ponent adduct. Angew Chem Int Ed 55(17):5282–5285. doi:10.1002/anie.201600751

34. Gilley CB, Buller MJ, Kobayashi Y (2008) Synthesis of functionalized pyroglutamic acids, part 2. The stereoselective condensation of multifunctional groups with chiral levulinic acids. Synlett 15:2249–2252
35. Gilley CB, Kobayashi Y (2008) 2-Nitrophenyl isocyanide as a versatile convertible isocyanide: rapid access to a fused γ-lactam β-lactone bicycle. J Org Chem 73(11):4198–4204. doi:10.1021/jo800486k
36. Zhdanko AG, Nenajdenko VG (2009) Nonracemizable isocyanoacetates for multicomponent reactions. J Org Chem 74(2):884–887. doi:10.1021/jo802420c
37. Sureshbabu VV, Narendra N, Nagendra G (2009) Chiral N-Fmoc-β-Amino alkyl isonitriles derived from amino acids: first synthesis and application in 1-substituted tetrazole synthesis. J Org Chem 74(1):153–157. doi:10.1021/jo801527d
38. Vishwanatha TM, Narendra N, Sureshbabu VV (2011) Synthesis of β-lactam peptidomimetics through Ugi MCR: first application of chiral Nβ-Fmoc amino alkyl isonitriles in MCRs. Tetrahedron Lett 52(43):5620–5624. doi:10.1016/j.tetlet.2011.08.090
39. Mehta VP, Modha SG, Ruijter E, Van Hecke K, Van Meervelt L, Pannecouque C, Balzarini J, Orru RVA, Van der Eycken E (2011) A microwave-assisted diastereoselective multicomponent reaction to access dibenzo[c, e]azepinones: synthesis and biological evaluation. J Org Chem 76(8):2828–2839. doi:10.1021/jo200251q
40. Costantino G, Marinozzi M, Camaioni E, Natalini B, Sarichelou I, Micheli F, Cavanni P, Faedo S, Noe C, Moroni F, Pellicciari R (2004) Stereoselective synthesis and preliminary evaluation of (+)- and (−)-3-methyl-5-carboxy-thien-2-yl-glycine (3-MATIDA): identification of (+)-3-MATIDA as a novel mGluR1 competitive antagonist. Il Farmaco 59(2):93–99. doi:10.1016/j.farmac.2003.11.008
41. Ross G, Ugi I (2001) Stereoselective syntheses of & #945;-amino acid and peptide derivatives by the U-4CR of 5-desoxy-5-thio-D-xylopyranosylamine. Can J Chem 79(12):1934–1939
42. Ziegler T, Kaisers H-J, Schlömer R, Koch C (1999) Passerini and Ugi reactions of benzyl and acetyl protected isocyanoglucoses. Tetrahedron 55(28):8397–8408. doi:10.1016/S0040-4020(99)00461-5

Glycoconjugate-Based Inhibitors of *Mycobacterium Tuberculosis* GlgE

Sri Kumar Veleti and Steven J. Sucheck

Abstract Tuberculosis (TB) is the leading cause of death globally as a result of a single infectious disease. A staggering 6 million new cases were reported in the 2014. In order to eradicate the ongoing threat of TB and combat rising rates of TB drug resistance new therapeutics must be developed. In this chapter, we briefly review the history of TB, *Mycobacterium tuberculosis* (*Mtb*) cell wall structure, the enzymes involved in synthesizing cell wall, and the trehalose utilization pathways (TUP). We focus on the recent discovery of enzyme *Mtb* GlgE, a glycosyl hydrolase-like phosphorylase, which has been found to be essential for *Mtb* viability and the ongoing efforts to design inhibitors against this target.

1 Mycobacterium Tuberculosis

1.1 Introduction

Dr. Koch discovered the functioning agent of tuberculosis (TB), *Mycobacterium tuberculosis* (*Mtb*) in 1882 [1]. Over 100 years later, TB is the leading cause of death in the world due to an infectious disease. In 2014, 1.5 million people were killed by TB. Mortality, due to TB has decreased by 47% since 1990, and effective diagnosis and treatment of TB saved an estimation of 43 million lives between 2000 and 2014 [2]. Globally, the incident rates of TB have fallen by an average of 1.5% per year since 2000. Most of the deaths are caused by drug-resistant strains [2], which have evolved due to selective pressure, co-infection with HIV, and inconsistent adherence to treatment regimens [3]. In 2014, it was estimated that 480,000 new multidrug-resistant TB (MDR-TB) cases occurred worldwide and approximately 190,000 deaths resulted from MDR-TB. Thus, 3.3% of new TB cases were MDR-TB in 2014 and it is estimated that 9.7% of people with MDR-TB have

S.K. Veleti · S.J. Sucheck (✉)
Department of Chemistry and School of Green Chemistry and Engineering,
The University of Toledo, 2801 W. Bancroft Street, Toledo, OH 43606, USA
e-mail: steve.sucheck@utoledo.edu

extensively drug-resistant-TB (XDR-TB). Clearly, new drugs are desperately nee-
ded to eradicate TB and combat the rising numbers of drug-resistant infections
including MDR-TB and XDR-TB [4–7].

1.2 Present Drugs for Tuberculosis

TB drug therapy began after Koch's discovery became widely accepted [1]. The
major era of anti-TB drug discovery occurred in the mid 1960s [8]. Streptomycin

Table 1 First-line and second-line drugs used to treat TB

was the first antibiotic used for treating *Mtb,* beginning in 1944 [9]. Unfortunately, the pattern of *Mtb* developing resistance to antibiotics was discovered shortly after [10]. First- and second-line drugs that are commonly utilized to fight *Mtb.* infections are shown in Table 1 [11, 12].

The main reason for identifying new drugs for tuberculosis is the increased prevalence of resistance strains. Further, the current treatment is very complex which results in patient non-compliance. The treatment consists of two phases, the intensive phase and the continuous phase. During the intensive phase patients are administered four first-line drugs for approximately two months. The continuous phase spans four months on isoniazid and rifampin with a three times a week dosing regimen. Situations such as improper medical attention, lack of supplies, and a poor understanding of the strict dosing regimen all contribute to the development of resistant strains.

Drug-resistant strains are categorized as single-drug resistant, MDR and XDR strains. Single-drug resistant should be treatable by a combination of first-line drugs. However, MDR strains are resistant to both isoniazid and rifampin, whereas XDR strains are resistant to isoniazid and rifampin along with any fluoroquinolone and at least one of three injectable second-line drugs (i.e., amikacin, kanamycin, or capreomycin) [13]. Active *Mtb* infections are also more prominent with the patients affected with HIV, and TB treatments can tremendously interfere with the anti-retroviral drugs used to treat HIV [14]. Presently, researchers are focusing on inhibiting critical cellular process that occurs in either actively dividing or dormant *Mtb* [15]. Recently GlgE, a maltosyl transferase, in *Mtb* was identified as a potential new drug target [16].

1.3 Cell Wall of Mycobacterium Tuberculosis (Mtb)

The cell wall of *Mtb* has two layers. One is the inner layer which consists of membrane proteins in the plasma membrane (Fig. 1), whereas, the outer layer consists of four important main components: the mycolylarabinogalactan (mAG), the peptidoglycan, free glycolipids, and the capsule. The cell wall of *Mtb* is a sturdy structure which makes it difficult for drugs to pass through [17]. The arabinogalactan and the associated free glycolipid layer together form the mycobacterial outer membrane (MOM). The mAG is made up of D-arabinan, mycolic acids, and D-galactan which are covalently connected [18, 19]. The galactan fragment contains repeating units of galactofuranose in alternating β-(1 → 5) and β-(1 → 6) linkages. Arabinan is attached to the galactan at the fifth position of galactofuranose residue and itself with α-(1 → 3), α-(1 → 5), and β-(1 → 2) linkages. The arabinogalactan is esterified with mycolic acids and further covalently linked to peptigoglycan (PG) by a phosphoryl-*N*-acetylglucosaminosyl-rhamonse linker [20] which together form the mycolyl arabinogalactan peptidoglycan (mAGP) complex [21–23]. The peptidogycan is comprised of alternating units of *N*-acetylglucosamine and muramic acid residues, which are often acetylated or glycolated [20].
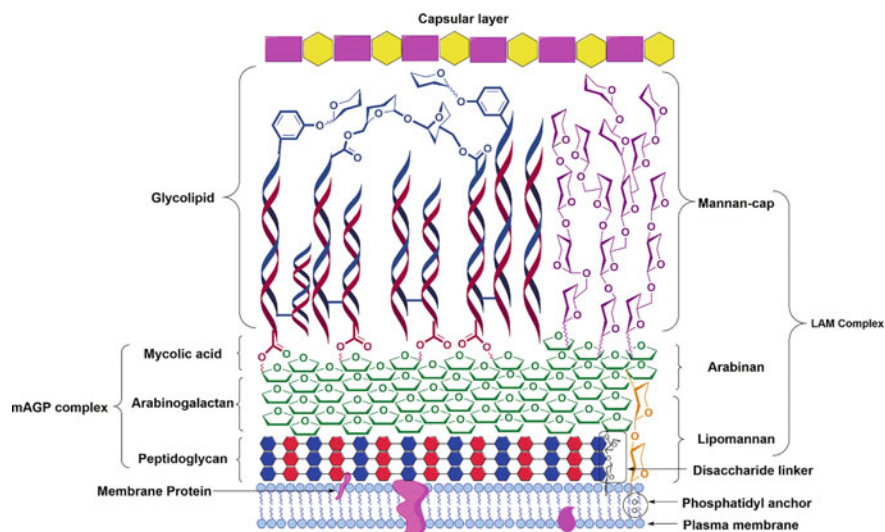
Fig. 1 Schematic representation of *Mtb* cell wall

Mycolic acids are made of long chain α-branched-β-hydroxyl fatty acids. The α-branches generally consist of more than 20 carbons, but 60 carbons can also be accomodated by longer meromycolate moities. These are usually functionalized by cis and trans double bonds, cyclopropane rings, methoxy, and keto esters [23–25]. These mycolic acids are attached to trehalose monomycolate (TMM), which can also be converted to trehalose dimycolate (TDM). Mycolic acids also form complexes with various sulfolipids, phosphotidylinositol mannoside (PIM), and lipoarabino-mannan (LAM) [26]. All these connections and functionalities contribute to the hydrophobicity and impermeability of the mycobacterial cell wall [27].

The LAM complex consists of a PIM anchor, D-mannan, and D-arabinan polysaccheride backbones [28]. These structures are formed by various mannosyl tranferases (ManTs) by the addition of mannose residues to phosphotidylinositol (PI) to form PIM [29, 30]. PIMs can be extended by ManTs to form linear and mature branched lipomannans (LM). LM is subsequently arabinosylated to form mature LAM. The terminal end of LAM can be further glycosylated with mannose to form an mannose cap refered to as manLAM. Together the structures are called the LAM complex. The PI anchor present in the membrane helps to connect the LAM complex with the membrane. This LAM complex also plays a crucial role in maintaining the integrity of the cell wall [31–33].

The outer part of the mycobacterial cell wall consists of the capsular layer which is made of polysaccherides and proteins, such as superoxidase dismutase, glutamine synthase, and thioredoxin. The characteristics of the capsular enzymes play crucial roles in bacterium's response to oxidative stress from the environment.

1.4 Trehalose Utilization Pathways (TUP)

Trehalose (1-*O*-α-D-glucopyranosyl-α-D-glucopyranoside) is a non-reducing disaccharide which is present in microbial cells, plants, bacterial, insects, and fungi. It protects the organism from stress, typically temperature, dehydration, and oxidative stress [34]. Its mechanism of protecting the cells is not well understood but depends on concentration [35]. Trehalose utilization pathways in *Mtb* contain several possible drug targets, for example TPP2 [36], Pks13 [37], GlgE [16], MmpL3 [38], and Ag85s [39, 40].

   Mtb uses trehalose for protection, energy, and as a precursor for building components of the cell wall [35]. Trehalose can be synthesized in three different pathways (Fig. 2) [41]. The first pathway is the OtsA/TPS pathway, it transforms glucose-6-phosphate to trehalose. OtsA encodes a trehalose phosphate synthase (TPS) which transfers glucose units from UDP-glucose to glucose-6-phosphate yielding trehalose-6-phosphate [42, 43]. In the next step, OtsB2 encodes TPP which dephosphorylates trehalose-6-phosphate forming trehalose. According to mutational studies, OtsA and OtsB genes are essential for the viability [44, 45].

   The second pathway is the TreYZ pathway which involves three enzymes TreX, TreY, and TreZ and transforms glycogen to trehalose [46]. Trehalose is also
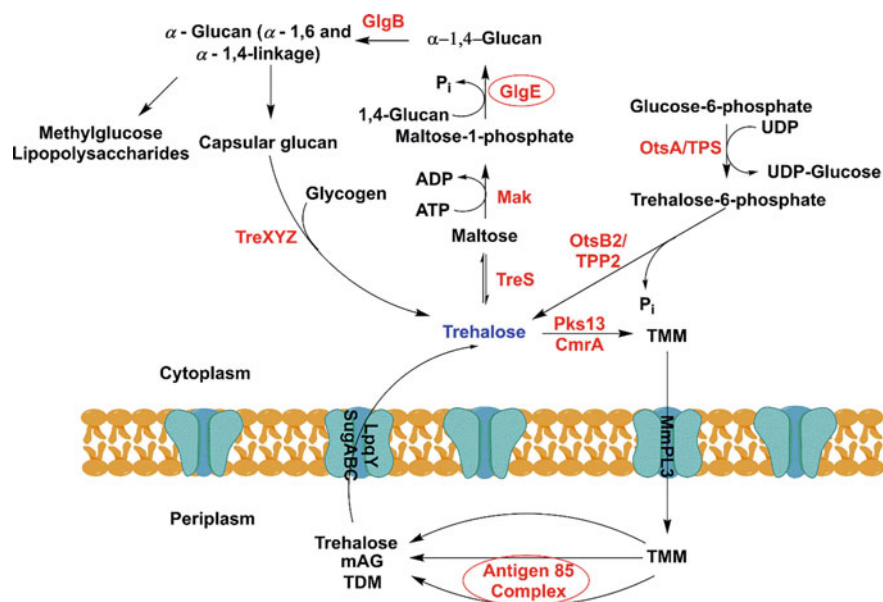

Fig. 2 Trehalose utilization pathways (TUP)

involved in production of TMM via Pks13 [47]. The newly synthesized TMM is passed from the cytoplasm to the periplasm which is taken up by the Ag85 complex and converted into trehalose, mycolylarabinogalactan (mAG), and TDM. TDM is a cell wall envelope glycolipid required for *Mtb* virulence [40].

The third pathway contains the genetically validated drug target GlgE. This pathway involves the conversion of trehalose into α-1,6-glucans. Briefly, trehalose is converted to maltose by TreS. The resulting maltose is phosphorylated by Mak yielding phosphorylated sugar, maltose-1-phosphate (M1P) [48]. M1P is the substrate for GlgE which forms a linear α-1,4-glucan [16, 49]. The linear α-1,4-glucans are modified with α-1,6 branches by GlgB which is transformed to capsular glucan.

1.5 Importance of GlgE as Drug Target

Kalscheuer et al. showed evidence that GlgE is an essential enzyme present in the α-glucan biosynthesis and showed by four ways that the enzyme meets the characteristic features of a promising new anti-TB target. First, by blocking GlgE activity with a chemical genetic strategy they inhibited the growth of *Mtb* and established that cell death was caused by a self-poisoning feedback loop. Second, GlgE inhibition lead to *Mtb* cell death by in vivo studies in lungs and spleens of infected mice. Third, GlgE is not present in humans or in normal gut flora. This implies the drugs made to inhibit GlgE will specifically inhibit *Mtb* with limited off target effects. Fourth, up to now no one has targeted the α-glucan pathway by chemotherapeutics in the treatment of *Mtb*. The self-poisoning response is a novel mode of action when compared to the mechanism of action of drugs used presently [16, 49]. Hopefully, a compound can be developed to potently inhibit GlgE which might solve the problem of MDR- and XDR-TB.

1.5.1 Mechanism of GlgE

GlgE is a member of the glycoside hydrolase subfamily GH13_3. The enzyme works by an α-retaining double-displacement mechanism and catalyzes the transfer of maltosyl units as shown in Fig. 3 [50]. The mechanism begins with the side chain of Asp418 attacking M1P generating a *β*-glycosyl enzyme intermediate. The incoming acceptor glucan is deprotonated by the general acid/base Glu447 side chain and attacks the *β*-glycosyl enzyme forming linear α-1,4 glucans [51]. Notably, release of phosphate was not observed when *β*-M1P was used as substrate.
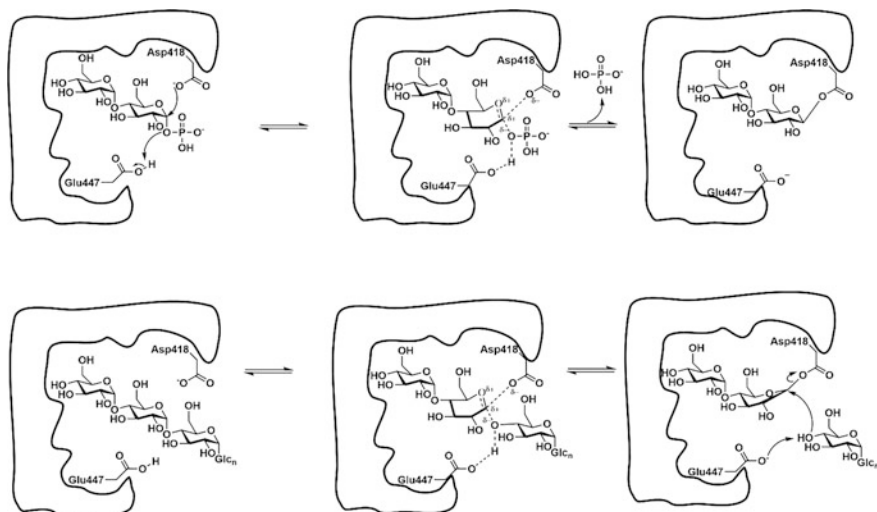
Fig. 3 Mechanism of GlgE. The numbering Asp418 and Glu447 are based on the numbering of *Mtb* GlgE

1.6 Synthesis of Ligands for Studying the Mtb GlgE Binding Site

1.6.1 Substrate Analog Maltose-C-Phosphonate

Veleti et al. have synthesized a non-hydrolysable and isosteric analog of M1P, maltose-C-phosphonate (**2**, MCP) which interacts sufficiently with the substrate binding pocket to produce inhibitory activity. Co-crystallization of the MCP with the homologous enzyme *Streptomyces coelicolor* (*Sco*) GlgEI provided information via an X-ray crystal structure that aids in the development of new anti-tuberculosis drugs. The molecule was prepared starting from maltose (**1**) and employed Wittig and Micheal-Arbuzov chemistry that led to the first inhibitor MCP. MCP moderately inhibited *Mtb* GlgE with an $IC_{50} = 230 \pm 24$ μM [52]. MCP was complexed with *Sco* GlgEI-V279S, an enzyme surrogate for *Mtb* GlgE, and the crystal structure was solved to 1.9 Å resolution [53]. The results helped to define important interactions between the enzyme and MCP (Fig. 4).

1.6.2 2-Deoxy-2-Fluoro Substrate Analogue

Syson and co-workers [51] reported on the interaction between *Sco* GlgEI and α-M1P. They performed mutational studies and substituted Asp394 with Ala in order to eliminate hydrolysis of α-M1P over the time scale of protein crystallization. Then they complexed *Sco* GlgEI-D394A with α-M1P, referred to as the Michaelis complex, and
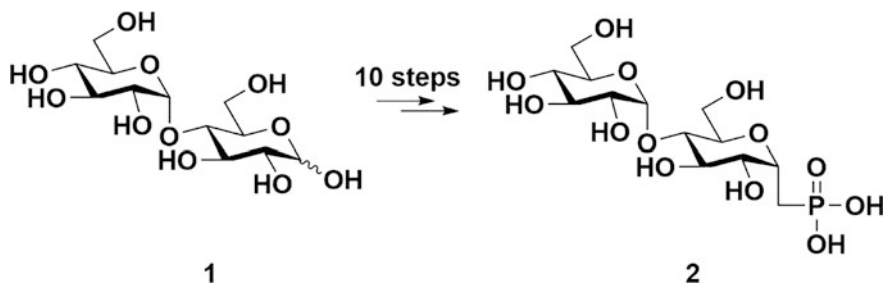
Fig. 4 Conversion of maltose (**1**) to maltose-C-phosphonate (**2**)

determined the X-ray structure for the complex to 2.55 Å resolution. The structural
studies help to define the assignment of the binding subsites. To trap the putative
enzyme intermediate they mutated Glu423 with Ala. *Sco* GlgEI-E423A was com-
plexed with 2-deoxy-2-fluoro-α-D-maltosyl fluoride (**1**). The X-ray crystal structure
was determined to 2.5 Å resolution and showed covalent bond formation between the
Asp394 and the C1 carbon. They also used mass spectrometry to show the successive
extension of the acceptor by incubating protein with substrate analogue. These crystal
structures were the first to show trapping of the glycosyl enzyme intermediate. Their
findings provide strong evidence to the support a double-displacement mechanism
and assignment of the catalytic nucleophile (Fig. 5).

1.6.3 Transition-State Like Inhibitor

In an attempt to mimic the transition state for the reaction catalyzed by GlgE,
Veleti, and co-workers designed an inhibitor with a pyrrolidine moiety shown to
inhibit other glycosyl hydrolases in the micromolar range [54, 55]. To synthesize
2,5-dideoxy-3-*O*-α-D-glucopyranosyl-2,5-imino-D-mannitol (**5**, DDGIM) a con-
vergent synthesis was followed by coupling thioglycoside (**2**) with 5-azido-3-*O*-
benzyl-5-deoxy-1,2-*O*-isopropylindene-β-D-fructopyranose (**4**) [56] followed by
global deprotection to afford the target molecule. Pyrrolidine **5** inhibited both *Mtb*

Fig. 5 2-Deoxy-2-fluoro-α-
D-maltosyl fluoride (**1**)
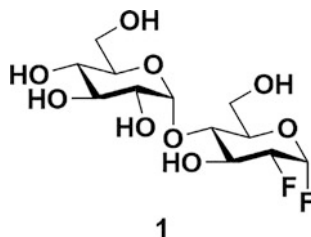
GlgE and *Sco* GlgEI-V279S with $Ki = 237 \pm 27$ μM and $Ki = 102 \pm 7.52$ μM, respectively [57]. Further studies complexing DDGIM with *Sco* GlgEI-V279S were conducted by Lindenberger et al. They solved the crystal structure to a resolution of 2.5 Å [53]. The structure showed numerous interactions in the binding pocket, specifically a strong ionic interaction between D394 and the putative secondary ammonium ion of DDGIM (Fig. 6).

1.6.4 2-Deoxy-2,2-Difluoro Substrate Analogue

Extending the use of fluoro sugars for studying structural insight of GlgE, Thanna et al. synthesized α-2-deoxy-2,2-difluoromaltosyl fluoride (**2,** α-MTF) and *β*-2-deoxy-2,2-difluoromaltosyl fluoride (**3,** *β*-MTF) [58]. These molecules did not inhibit GlgE; however, 2-deoxy-2,2-difluoro-α-maltosyl fluoride provided useful insight from its X-ray crystal structure complex with *Sco* GlgEI-V279S at 2.3 Å resolution. The complex provided evidence that Glu423 functions as proton donor by showing a hydrogen bond interaction between Glu423 and C1F. Further, Arg392 and axial C2 difluoromethylene moiety of α-MTF interacts by hydrogen bonding suggesting that C2 substitution can be tolerated with hydrogen bond acceptors (Fig. 7).

1.6.5 Proline and Pyrrolidine-Based Phosphonates as Transistion State Inhibitors

Ongoing work in our lab based on iminosugars as potential transition-state inhibitor against GlgE includes proline-based phosphonates (**1** and **3**) and pyrrolidine-based phosphonates (**2** and **4**). These compounds inhibited *Sco* GlgEI-V279S in a range of 45–95 μM. Addition of phosphonates improved enzyme inhibition 2-fold when compared to the previously synthesized compounds [59]. Crystallizing these target compounds with *Sco* GlgEI-V279S will be helpful for advancing the structure-based approach for identifying improved inhibitors against GlgE (Fig. 8).
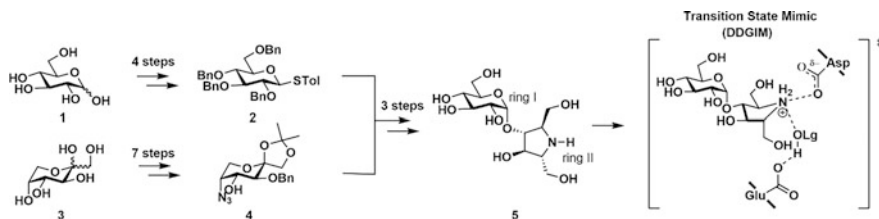

Fig. 6 Synthesis of poly-hydroxypyrrolidine-based inhibitor **5** and illustration of expected binding interactions in the enzyme active site
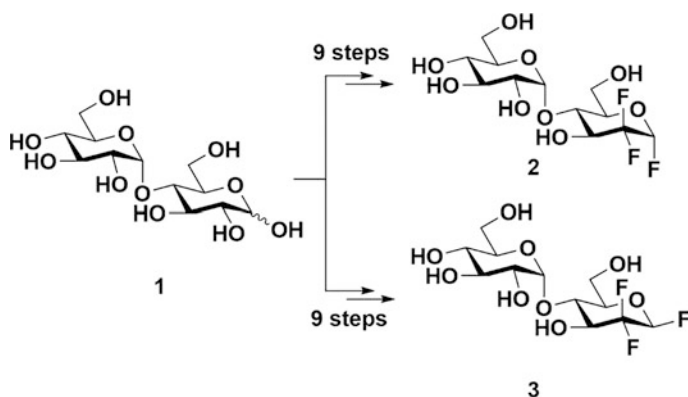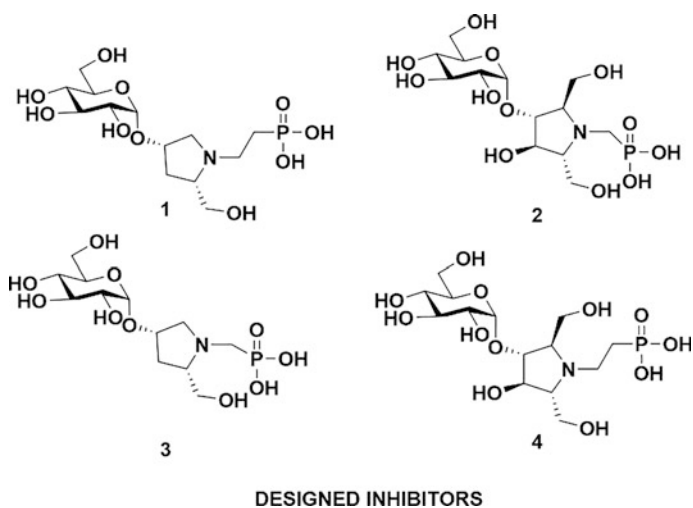
Fig. 7 Conversion of maltose to α-2-deoxy-2,2-difluoromaltosyl fluoride (**2**) and β-2-deoxy-2,2-difluoromaltosyl fluoride (**3**)

DESIGNED INHIBITORS

Fig. 8 Designed transition-state-like inhibitors

1.7 Inhibitors for Mtb GlgE Based on Docking Studies

Structure-based drug design frequently includes the use of molecular docking. It has become increasingly important and complementary to wet laboratory experiments in that it aids in evaluating the complex information of target enzyme with small ligands. The tools sample optimal geometrical arrangements and quantitate the strength of the bonding forces. By using molecular docking, the binding energy between the ligand and the enzyme binding site can be calculated as shown in Eq. 1 [60]. Recent advances in high-performance computational screening methods have

contributed to the design of several drugs that have advanced in clinical trials [61, 62]. For example, computational chemistry has informed lead compound development for compounds that prevent myocardial infraction, treat HIV infection, rheumatoid arthritis, and other diseases [63, 64].

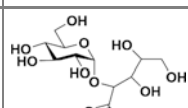$$E_{binding} = E_{target-ligand} - (E_{target} + E_{ligand})$$ (1)

1.7.1 Screening of α-Amylase Family of Protein Binders Against Homology Model of *Mtb* GlgE

Sengupta et al. studied the binding affinity of an α-amylase family of protein binders to a homology model of *Mtb* Glg. Since an X-ray structure of *Mtb* GlgE was unavailable at that time they used the homologue *Sco* GlgEI [50] to prepare a homologous model of *Mtb* GlgE. They retrieved the amino acid sequence of *Mtb* GlgE from the protein database of National Center of Biotechnology Information. They utilized *Sco* GlgEI (PDB ID: 3ZSS) as a template to superimpose the homology model of *Mtb* GlgE. Subsequently, the energy of homologous model was minimized using CHARMm. The model structure has three binding cavities with one primary binding site (PBS) and two secondary binding site (SBS1 and SBS2) [60]. Sengupta et al. used CDOCKER, which is a molecular dynamics simulated annealing based docking program, and LibDock, which is a fast feature-based algorithm for molecular docking. 3-*O*-α-D-gluco-pyranosyl-D-fructose (OTU) docked with the best binding affinity to the PBS among all the computationally screened substrates that were taken from the CHEMBL database as shown in Table 2. This study revealed insights into the active site, substrate binding affinities of ligands, and provide the first homologous 3D structure of *Mtb* GlgE.

1.7.2 Screening the ZINC Database Against *Sco* GlgEI
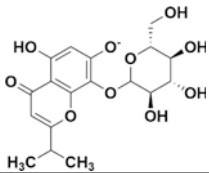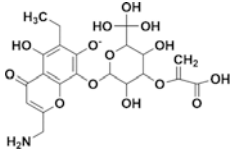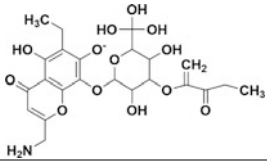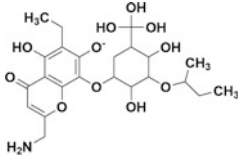
Billones et al. worked on structure-based inhibitors by screening virtual natural products against GlgE. They used the X-ray data from closely related, *Sco* GlgEI [65] complexed with maltose (PDB code: 3ZT5). The natural products catalogs

Table 2 Two-dimensional structure and interaction energies of a known α-amylase family protein binder to PBS and SBS1 of the homology model of *Mtb* GlgE

PDB ID	PDB ligand ID	Ligand structure	CDOCKER[a]		LibDock	Score
			PBS	SBS1	PBS	SBS1
3UEQ	OTU		−156.0	−114.5	148.0	120.6

[a]CDOCKER interaction energies are shown in kcal/mole

Table 3 Strucutres of ZINC39010596 and the top 3 modified ligands
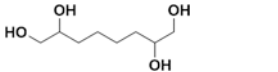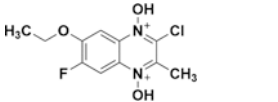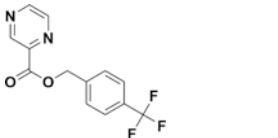
Compound ID	Structure	Binding energy (kcal/mol)
ZINC39010596		−309.58
Ligand 1		−426.00
Ligand 2		−424.65
Ligand 3		−414.68

were obtained from the ZINC database and CDOCKER software was used to dock the molecules. To modify the chosen lead compounds, the de novo evolution protocol was used. They also studied the carcinogenicity, mutagenicity, and biodegradability using a toxicity prediction (extensible) protocol. Among all the screened compounds, they observed three compounds with higher binding energy when compared to natural substrate M1P. The compounds were derived from ZINC39010596, these compounds found to be non-carcinogenic, non-mutagenic, and biodegradable (Table 3).

1.7.3 Screening the ZINC Database and Anti-Tuberculosis Compounds Database Against the *Mtb* GlgE
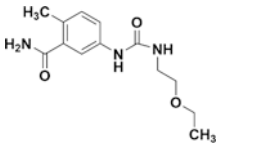
Sengupta et al. reported on a pharmacophore-based virtual screening, docking, and molecular dynamics simulations on *Mtb* GlgE [66]. They presented ligand and structure-based pharmacophore models showing a PBS and SBS2 which were constructed based on the 3D homology model of *Mtb* GlgE. The structure were screened against the ZINC database and an anti-tuberculosis compounds database (ATD). They identified 23 molecules from ZINC and ATD with better binding

Table 4 Structures and binding free energies of top scored ligands

Ligands	Structure	Binding free energy (kcal/mol)
ZINC40525623		−4467.5438
ZINC02539424		−4502.40983
AZI-OM2K2-49-6I		−5407.3600
AZI-OM2K2-29-5		−5454.94883

affinity than the natural substrate, M1P. The four top molecules that are the best at binding the PBS or SBS2 are shown in Table 4.

2 Summary

Tuberculosis is a communicable disease which is taking millions of lives every year. Current treatment options are complex, lengthy, and clearly difficult to complete. As result, drug-resistance rates are rising. Discovering new drug targets like GlgE, which may rapidly kill the organism, will help in eradicating TB. So far, efforts to identify inhibitors against GlgE have been structure-based. It is expected that gaining an understanding of the detailed binding interactions between the enzyme and its various ligands will inform improved inhibitor design and ultimately lead to new antibiotics.

References

1. Koch R (1882) Lecture on the discovery of *Mycobacterium tuberculosis*
2. World Health Organisation (2015) The Global tuberculosis report 2015. World Health Organization, Geneva
3. Brennan P, Young D, Robertson B (2008) Handbook of anti-tuberculosis agents. Tuberculosis 88(2):85–170

4. Chiang CY, Centis R, Migliori GB (2010) Drug-resistant tuberculosis: past, present, future. Respirology 15(3):413–432

5. Crofton J (1960) Tuberculosis undefeated. Brit Med J 2(5200):679

6. Mitchison D (2000) Role of individual drugs in the chemotherapy of tuberculosis. Int J Tuberc Lung Dis 4(9):796–806

7. Udwadia ZF, Amale RA, Ajbani KK, Rodrigues C (2012) Totally drug-resistant tuberculosis in India. Clin Infect Dis 54(4):579–581

8. Ginsberg A (2011) Research spotlight: The TB alliance: overcoming challenges to chart the future course of TB drug development. Future Med Chem 3(10):1247–1252

9. Jones D, Metzger H, Schatz A, Waksman SA (1944) Control of gram-negative bacteria in experimental animals by streptomycin. Science 100(2588):103–105

10. Waksman SA, Reilly HC, Schatz A (1945) Strain specificity and production of antibiotic substances: V. strain resistance of bacteria to antibiotic substances, especially to streptomycin. P Natl Acad Sci USA 31(6):157

11. Lehmann J (1946) Para-aminosalicylic acid in the treatment of tuberculosis. The Lancet 247 (6384):15–16

12. Janin YL (2007) Antituberculosis drugs: ten years of research. Bioorg Med Chem 15(7):2479–2513

13. Zhang Y, Yew W (2009) Mechanisms of drug resistance in Mycobacterium tuberculosis [State of the art series. Drug-resistant tuberculosis. Edited by CY. Chiang. Number 1 in the series]. Intl J Tuberc Lung Dis 13(11):1320–1330

14. Falzon D, Jaramillo E, Schünemann H, Arentz M, Bauer M, Bayona J, Blanc L, Caminero J, Daley C, Duncombe C (2011) WHO guidelines for the programmatic management of drug-resistant tuberculosis: 2011 update. Eur Respir J 38(3):516–528

15. Lamichhane G (2011) Novel targets in M. tuberculosis: search for new drugs. Trends Mol Med 17(1):25–33

16. Kalscheuer R, Syson K, Veeraraghavan U, Weinrick B, Biermann KE, Liu Z, Sacchettini JC, Besra G, Bornemann S, Jacobs WR Jr (2010) Self-poisoning of Mycobacterium tuberculosis by targeting GlgE in an [alpha]-glucan pathway. Nat Chem Biol 6(5):376–384

17. Daffé M, Draper P (1997) The envelope layers of mycobacteria with reference to their pathogenicity. Adv Microb Physiol 39:131–203

18. Barsom EK, Hatfull GF (1996) Characterization of a Mycobacterium smegmatis gene that confers resistance to phages L5 and D29 when overexpressed. Mol Microbiol 21(1):159–170

19. Tam P-H, Lowary TL (2009) Recent advances in mycobacterial cell wall glycan biosynthesis. Curr Opin Chem Biol 13(5):618–625

20. Crick DC, Mahapatra S, Brennan PJ (2001) Biosynthesis of the arabinogalactan-peptidoglycan complex of Mycobacterium tuberculosis. Glycobiology 11 (9):107R–118R

21. Kremer L, Dover LG, Morehouse C, Hitchin P, Everett M, Morris HR, Dell A, Brennan PJ, McNeil MR, Flaherty C (2001) Galactan biosynthesis in Mycobacterium tuberculosis Identification of a bifunctional UDP-galactofuranosyltransferase. J Biol Chem 276(28): 26430–26440

22. Alderwick L, Birch H, Mishra A, Eggeling L, Besra G (2007) Structure, function and biosynthesis of the Mycobacterium tuberculosis cell wall: arabinogalactan and lipoarabinomannan assembly with a view to discovering new drug targets. Biochem Soc T 35(5):1325–1328

23. Besra GS, Brennan PJ (1997) The mycobacterial cell envelope. J Pharm Pharmacol 49(S1):25–30

24. Dover LG, Alderwick LJ, Brown AK, Futterer K, Besra GS (2007) Regulation of cell wall synthesis and growth. Curr Mol Med 7(3):247–276

25. Besra GS, Khoo K-H, McNeil MR, Dell A, Morris HR, Brennan PJ (1995) A new interpretation of the structure of the mycolyl-arabinogalactan complex of Mycobacterium tuberculosis as revealed through characterization of oligoglycosylalditol fragments by

fast-atom bombardment mass spectrometry and 1H nuclear magnetic resonance spectroscopy. Biochemistry 34(13):4257–4266

26. Kaur D, Guerin ME, Škovierová H, Brennan PJ, Jackson M (2009) Biogenesis of the cell wall and other glycoconjugates of *Mycobacterium tuberculosis*. Adv Appl Microbiol 69:23–78

27. Favrot L, Ronning DR (2012) Targeting the mycobacterial envelope for tuberculosis drug development. Expert Rev Anti Infect Ther 10(9):1023–1036

28. Nigou J, Gilleron M, Puzo G (2003) Lipoarabinomannans: from structure to biosynthesis. Biochimie 85(1):153–166

29. Kaur D, Berg S, Dinadayala P, Gicquel B, Chatterjee D, McNeil MR, Vissa VD, Crick DC, Jackson M, Brennan PJ (2006) Biosynthesis of mycobacterial lipoarabinomannan: role of a branching mannosyltransferase. P Natl Acad Sci 103(37):13664–13669

30. Besra GS, Morehouse CB, Rittner CM, Waechter CJ, Brennan PJ (1997) Biosynthesis of mycobacterial lipoarabinomannan. J Biol Chem 272(29):18460–18466

31. Mishra AK, Driessen NN, Appelmelk BJ, Besra GS (2011) Lipoarabinomannan and related glycoconjugates: structure, biogenesis and role in *Mycobacterium tuberculosis* physiology and host–pathogen interaction. FEMS Microbiol Rev 35(6):1126–1157

32. Strohmeier GR, Fenton MJ (1999) Roles of lipoarabinomannan in the pathogenesis of tuberculosis. Microbes Infect 1(9):709–717

33. Gaitonde V, Sucheck SJ (2015) Antitubercular drugs based on carbohydrate derivatives. carbohydrate chemistry: state of the art and challenges for drug development: an overview on structure, biological roles, synthetic methods and application as therapeutics, 441

34. Singer MA, Lindquist S (1998) Multiple effects of trehalose on protein folding in vitro and in vivo. Mol Cell 1(5):639–648

35. Erdei É, Molnár M, Gyémánt G, Antal K, Emri T, Pócsi I, Nagy J (2011) Trehalose overproduction affects the stress tolerance of Kluyveromyces marxianus ambiguously. Bioresource Technol 102(14):7232–7235

36. Paul MJ, Primavesi LF, Jhurreea D, Zhang Y (2008) Trehalose metabolism and signaling. Annu Rev Plant Biol 59:417–441

37. Gavalda S, Bardou F, Laval F, Bon C, Malaga W, Chalut C, Guilhot C, Mourey L, Daffé M, Quémard A (2014) The polyketide synthase Pks13 catalyzes a novel mechanism of lipid transfer in mycobacteria. Chem Bio 21(12):1660–1669

38. Li W, Upadhyay A, Fontes FL, North EJ, Wang Y, Crans DC, Grzegorzewicz AE, Jones V, Franzblau SG, Lee RE (2014) Novel insights into the mechanism of inhibition of MmpL3, a target of multiple pharmacophores in *Mycobacterium tuberculosis*. Antimicrob Agents Chemother 58(11):6413–6423

39. Harth G, Lee B-Y, Wang J, Clemens DL, Horwitz MA (1996) Novel insights into the genetics, biochemistry, and immunocytochemistry of the 30-kilodalton major extracellular protein of *Mycobacterium tuberculosis*. Infect Immun 64(8):3038–3047

40. Jackson M, Raynaud C, Lanéelle M A, Guilhot C, Laurent-Winter C, Ensergueix D, Gicquel B, Daffé M Inactivation of the antigen 85C gene profoundly affects the mycolate content and alters the permeability of the *Mycobacterium tuberculosis* cell envelope. Mol Microbiol 31(5):1573–1587

41. De Smet KA, Weston A, Brown IN, Young DB, Robertson BD (2000) Three pathways for trehalose biosynthesis in mycobacteria. Microbiology 146(1):199–208

42. Gibson RP, Turkenburg JP, Charnock SJ, Lloyd R, Davies GJ (2002) Insights into trehalose synthesis provided by the structure of the retaining glucosyltransferase OtsA. Chem Biol 9(12):1337–1346

43. Cole S, Brosch R, Parkhill J, Garnier T, Churcher C, Harris D, Gordon S, Eiglmeier K, Gas S, Barry CR (1998) Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. Nature 393(6685):537–544

44. Sassetti CM, Boyd DH, Rubin EJ (2003) Genes required for mycobacterial growth defined by high density mutagenesis. Mol Microbiol 48(1):77–84

45. Murphy HN, Stewart GR, Mischenko VV, Apt AS, Harris R, McAlister MS, Driscoll PC, Young DB, Robertson BD (2005) The OtsAB pathway is essential for trehalose biosynthesis in *Mycobacterium tuberculosis*. J Biol Chem 280(15):14524–14529

46. Kalscheuer R, Weinrick B, Veeraraghavan U, Besra GS, Jacobs WR (2010) Trehalose-recycling ABC transporter LpqY-SugA-SugB-SugC is essential for virulence of *Mycobacterium tuberculosis*. P Natl Acad Sci 107(50):21761–21766

47. Portevin D, de Sousa-D'Auria C, Houssin C, Grimaldi C, Chami M, Daffé M, Guilhot C (2004) A polyketide synthase catalyzes the last condensation step of mycolic acid biosynthesis in mycobacteria and related organisms. P Natl Acad Sci 101(1):314–319

48. Fraga J, Maranha A, Mendes V, Pereira PJB, Empadinhas N, Macedo-Ribeiro S (2015) Structure of mycobacterial maltokinase, the missing link in the essential GlgE-pathway. Sci Rep 5

49. Kalscheuer R, Jacobs WR Jr (2010) The significance of GlgE as a new target for tuberculosis. Drug News Perspect 23(10):619–624

50. Syson K, Stevenson CE, Rejzek M, Fairhurst SA, Nair A, Bruton CJ, Field RA, Chater KF, Lawson DM, Bornemann S (2011) Structure of Streptomyces maltosyltransferase GlgE, a homologue of a genetically validated anti-tuberculosis target. J Biol Chem 286(44):38298–38310

51. Syson K, Stevenson CE, Rashid AM, Saalbach G, Tang M, Tuukkanen A, Svergun DI, Withers SG, Lawson DM, Bornemann S (2014) Structural insight into how streptomyces coelicolor maltosyl transferase GlgE binds α-Maltose-1-phosphate and forms a maltosyl-enzyme intermediate. Biochemistry 53(15):2494–2504

52. Veleti SK, Lindenberger JJ, Ronning DR, Sucheck SJ (2014) Synthesis of a C-phosphonate mimic of maltose-1-phosphate and inhibition studies on *Mycobacterium tuberculosis* GlgE. Bioorg Med Chem 22(4):1404–1411

53. Lindenberger JJ, Veleti SK, Wilson BN, Sucheck SJ, Ronning DR (2015) Crystal structures of *Mycobacterium tuberculosis* GlgE and complexes with non-covalent inhibitors. Sci Rep 5

54. Provencher L, Steensma DH, Wong C-H (1994) Five-membered ring azasugars as potent inhibitors of α-L-rhamnosidase (naringinase) from Penicillium decumbens. Bioorg Med Chem 2(11):1179–1188

55. de Melo EB, da Silveira Gomes A, Carvalho I (2006) α- and β-Glucosidase inhibitors: chemical structure and biological activity. Tetrahedron 62(44):10277–10302

56. Izquierdo I, Plaza MT, Yáñez V (2007) Polyhydroxylated pyrrolidines: synthesis from d-fructose of new tri-orthogonally protected 2, 5-dideoxy-2, 5-iminohexitols. Tetrahedron 63(6):1440–1447

57. Veleti SK, Lindenberger JJ, Thanna S, Ronning DR, Sucheck SJ (2014) Synthesis of a Poly-hydroxypyrolidine-Based inhibitor of *Mycobacterium tuberculosis* GlgE. J Org Chem 79(20):9444–9450

58. Thanna S, Lindenberger JJ, Gaitonde VV, Ronning DR, Sucheck SJ (2015) Synthesis of 2-deoxy-2, 2-difluoro-α-maltosyl fluoride and its X-ray structure in complex with Streptomyces coelicolor GlgEI-V279S. Org Biomol Chem 13(27):7542–7550

59. Veleti SK, Petit C, Ronning DR, Sucheck SJ (2016) Synthesis and inhibition studies of proline and pyrolidene-based phosphonates to inhibit *Streptomyces Coelicolor* (Sco) GlgE (manuscript in preparation)

60. Sengupta S, Roy D, Bandyopadhyay S (2014) Modeling of a new tubercular maltosyl transferase, GlgE, study of its binding sites and virtual screening. Mol Biol Rep 41(6):3549–3560

61. Arooj M, Sakkiah S, Kim S, Arulalapperumal V, Lee KW (2013) A combination of receptor-based pharmacophore modeling & QM techniques for identification of human chymase inhibitors. PLoS One 8(4):e63030

62. Silverman RB, Holladay MW (2014) The organic chemistry of drug design and drug action
63. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ (2012) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv Drug Deliv Rev 64:4–17
64. Cramer CJ (2013) Essentials of computational chemistry: theories and models. Wiley, New Jersey
65. Billones JB, Valle AMF (2014) Structure-based design of inhibitors against maltosyltransferase GlgE. Orient J Chem 30(3):1137–1145
66. Sengupta S, Roy D, Bandyopadhyay S (2015) Structural insight into *Mycobacterium tuberculosis* maltosyl transferase inhibitors: pharmacophore-based virtual screening, docking, and molecular dynamics simulations. J Biomol Struct Dyn 33(12):2655–2666

Selective Transformations of the Anomeric Centre in Water Using DMC and Derivatives

David Lim and Antony J. Fairbanks

Abstract 2-Chloro-1,3-dimethylimidazolinium chloride (DMC) and its derivatives are useful for numerous synthetic transformations, which involve selective activation of the anomeric centre of unprotected reducing sugars in water. This chapter summarises research reported to date using DMC and derivatives, such as 2-azido-1,3-dimethylimidazolinium hexafluorophosphate (ADMP). DMC has been successfully employed for the synthesis of glycosyl oxazolines, 1,6-anhydro-, 1-azido-, and a variety of thioglycosides. The use of ADMP allows the one-pot synthesis of glycosyl triazoles in water *via* the Cu-catalysed azide-alkyne Huisgen cycloaddition reaction. This latter methodology can be applied to a wide variety of carbohydrates and is also amenable to convergent glycopeptide synthesis in which oligosaccharides are directly conjugated to peptides that contain propargyl glycine residues. Such protecting group free methodologies, particularly when applied to complex oligosaccharides isolated from natural sources, may allow ready access to a wide variety of biologically interesting glycoconjugates.

1 Introduction

Protecting groups are typically unavoidable in synthetic carbohydrate chemistry. However, their use inevitably involves additional synthetic steps, resulting in loss of materials, and the generation of a significant amount of waste; not only chemical waste, but also in terms of cost and time. Thus, there is an increasing need for the development of methods for the synthetic manipulation of carbohydrates that do not involve protecting groups, but yet still furnish the desired products in high yield and with the requisite selectivity.

D. Lim · A.J. Fairbanks (✉)
Department of Chemistry, University of Canterbury, Private Bag 4800, Christchurch 8140, New Zealand
e-mail: antony.fairbanks@canterbury.ac.nz

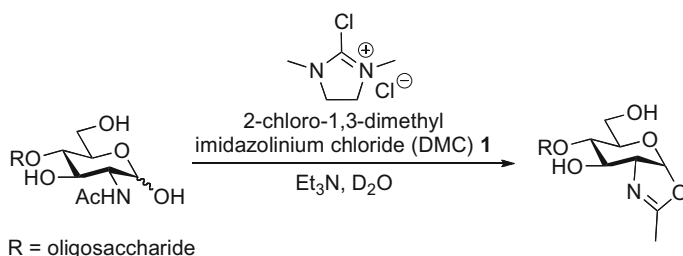# 2 Synthesis of Glycosyl Oxazolines in Water Using DMC

Glycosyl oxazolines have found wide application in carbohydrate chemistry, first in their protected forms as donors for oligosaccharide synthesis [1, 2], and subsequently, and perhaps more importantly as activated donor substrates for the enzymatic synthesis of glycopeptides and glycoproteins [3, 4].

In 2004, Kadokawa et al. [5] came close to developing a useful method for the direct synthesis of glycosyl oxazolines in water when they used a series of carbodiimides to activate the anomeric centre. However, these reactions only afforded the desired products in very low yield (<30%). Even so, the fact that this type of highly selective transformation was possible at all drew the attention of the Glycoscience community.

The dehydrative properties of the reagent, 2-chloro-1,3-dimethylimidazolium chloride **1** (DMC), were originally reported by Isobe and Ishikawa [6, 7]; however these reported transformations were all performed in organic solvents. In 2009, Shoda and co-workers [8] made a significant breakthrough in the field when they reported the use of DMC for the direct synthesis of glycosyl oxazolines from reducing sugars in D_2O in high yield (Scheme 1).

Prior to this report, the synthesis of glycosyl oxazolines typically required peracetylated 2-acetamido-2-deoxy sugars to be prepared. These could then be converted to the corresponding oxazolines using Lewis acids, such as ferric(III) chloride, tin(IV) chloride, boron trifluoride, or trimethylsilyl triflate, [9–12] before the final removal of acetate protecting groups under basic conditions. However, the use of strong Lewis acids can damage glycosidic linkages and, typically in the cases of larger oligosaccharides, may result in the formation complex reaction mixtures and low yields. This report, therefore, represented a significant advance in the production of *N*-glycan oxazolines, particularly as donor substrates for enzymatic synthesis.

The proposed mechanism [8] of oxazoline formation involves preferential attack of the hemiacetal hydroxyl on DMC; in the case of the β-anomer this yields reactive intermediate **2** (Scheme 2). Intramolecular attack of the 2-acetamido group at the anomeric centre, followed by abstraction of a proton by a suitable base, then affords

R = oligosaccharide

Scheme 1 Synthesis of glycosyl oxazolines directly from reducing sugars in water by Shoda et al. [8]
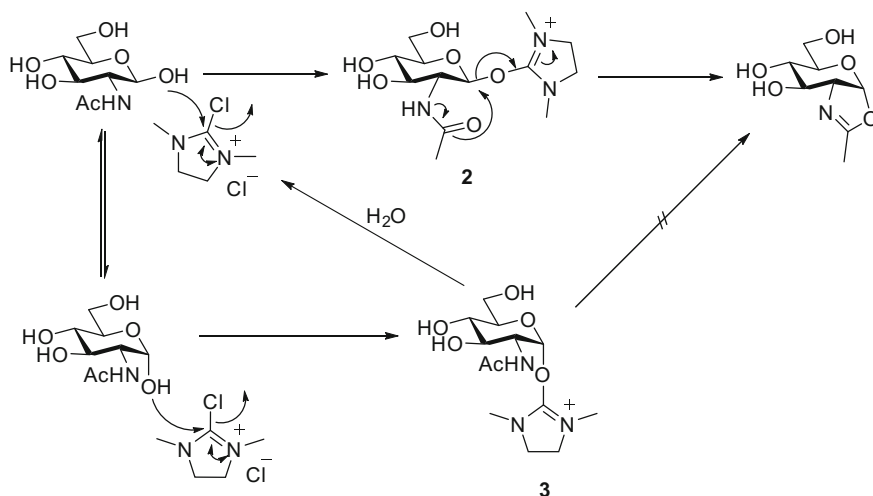
Scheme 2 Proposed mechanism for sugar oxazoline formation using DMC [8]

the oxazoline. The α-imidazolinium intermediate **3** can also be formed by attack of the α-hemiacetal hydroxyl on DMC. However, since **3** cannot directly form an oxazoline it is then probably hydrolysed to regenerate the β-anomer of the free sugar; this β-anomer then follows the reaction pathway *via* intermediate **2** to the oxazoline product. Although this mechanistic pathway is plausible, Shoda has stated [8] that the intermediacy of a β-glycosyl chloride during the conversion of **3** into the oxazoline product cannot be ruled out.

The use of DMC for the selective activation of reducing sugars in water has subsequently proved pivotal in facilitating the use of glycosyl oxazolines as donors for enzymatic glycosylation reactions catalysed by endo-β-*N*-acet-ylglucosaminidases (ENGases), such as those reported by the groups of Fairbanks [13–16], Wang [17–19], and Yamamoto [20–22].

Subsequently, Shoda and co-workers [23] reported another method for the direct synthesis of glycosyl oxazolines using the DMC analogue, 2-chloro-1,3-dimethyl-1*H*-benzimidazol-3-ium chloride **4**, (CDMBI) as the dehydrative agent (Fig. 1).

Fig. 1 DMC-analogue CDMBI **4**, and the urea DMBI **5** formed by its hydrolysis [23]
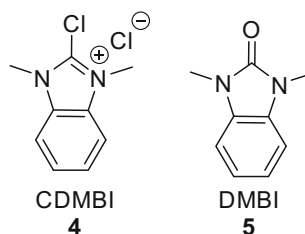
CDMBI was alleged to be superior to DMC, as it is less hygroscopic and easier to handle, although most applications still use the commercially available DMC for oxazoline formation. Shoda also stated that replacing the imidazolidine ring of DMC with the more electron-rich benzimidazole would reduce the reactivity of the chloroformamidinium group toward nucleophilic attack by water, and that additionally it made the reactive glycosyl-imidazolidinium intermediate more stable. It was also postulated that the introduction of the aromatic ring would also reduce the aqueous solubility of the hydrolysed product, namely the urea **5** (1,3-dimethylbenzimidazol-2-one, DMBI) (Fig. 1). Indeed, when the complex bi-antennary *N*-glycan decasaccharide **6** (derived from a sialyl glycopeptide isolated from egg yolks [24]) was converted to the corresponding oxazoline **7** using CDMBI **4**, DMBI **5** precipitated from the mixture as the reaction proceeded, and was easily removed by filtration (Scheme 3). Moreover, the filtrate containing the oxazoline could then be directly used for a subsequent enzymatic glycosylation of the acceptor *p*-nitrophenyl 2-acetamido-2-deoxy-β-D-glucopyranoside **8** using the glycosynthase, Endo-M N175Q [21], without the need for further purification or isolation of the oxazoline intermediate.
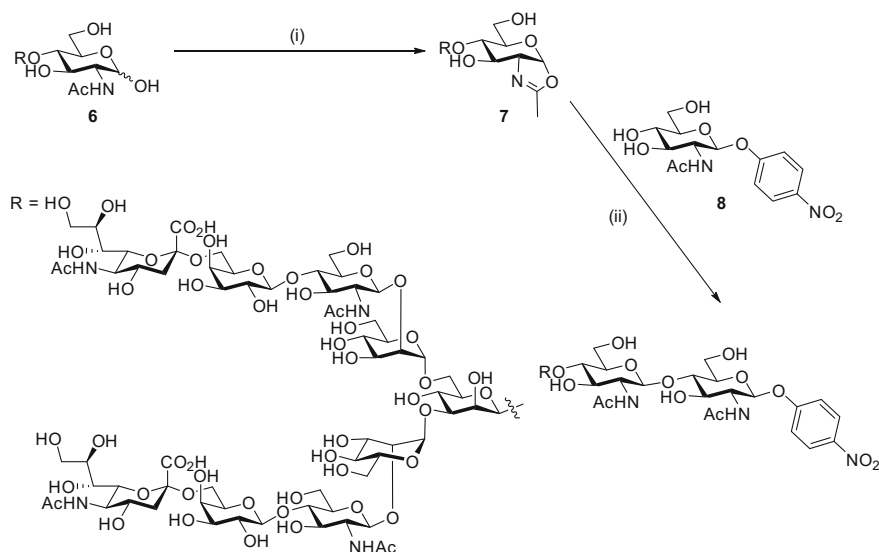
Scheme 3 Enzymatic glycosylation of GlcNAc-β-pNP with the sialoglycan-oxazoline without isolation of the intermediate oxazoline [23]. Reaction conditions: **i** CDMBI **4**, Na$_3$PO$_4$, H$_2$O; **ii 8**, Endo-M N175Q

3   Synthesis of 1,6-Anhydro Sugars in Water Using DMC

Subsequent to the first report on oxazoline formation, DMC and its derivatives have been employed for the synthesis of a variety of glycosides formed by intercepting the α-imidazolinium intermediate **3** (Scheme 2) with different nucleophiles. In the absence of any external nucleophile, the 6-hydroxyl group may attack the anomeric centre. Thus, Shoda et al. [25] reported the use of DMC for the formation of 1,6-anhydro sugars from unprotected glycopyranoses in water in almost quantitative yield. The DMC procedure mitigated the requirement for the harsh reaction conditions that had previously been used for the synthesis of 1,6-anhydro sugars, such as pyrolysis [26, 27] or thermal degradation [28, 29]. The proposed reaction mechanism is illustrated in Scheme 4, and is similar to that previously suggested [8] for the formation of glycosyl oxazolines. Herein the first step is a nucleophilic attack of the anomeric hydroxyl group of the predominant β-anomer of glucose **9** on DMC **1**, giving rise to intermediate **10**. Intramolecular attack of the 2-hydroxyl group of **10** at the anomeric carbon affords a 1,2-anhydro intermediate **11**, which is subsequently converted to the 1,6-anhydro sugar **12** via intramolecular nucleophilic attack by the 6-hydroxyl group. The α-anomer of glucose **13** also reacts with DMC to give the corresponding α-intermediate **14**. Although direct conversion of **14** into the 1,6-anhydro sugar **12** is possible by an attack of the 6-hydroxyl at the anomeric centre, Shoda suggests that in fact **14** is hydrolysed by an attack of water to regenerate β-glucose **9**. β-Glucose **9** then follows the above pathway via **11** to give the 1,6-anhydro sugar **12**. Evidence for this overall mechanistic pathway, and the requirement for a 1,2-anhydro intermediate such as **11** *en route* to the 1,6-anhydro sugar, was the fact that the corresponding 1,6-anhydro sugars were *not* formed when D-mannose, 2-deoxy-D-glucose, and 2-fluoro-2-deoxy-D-glucose were used as substrates.
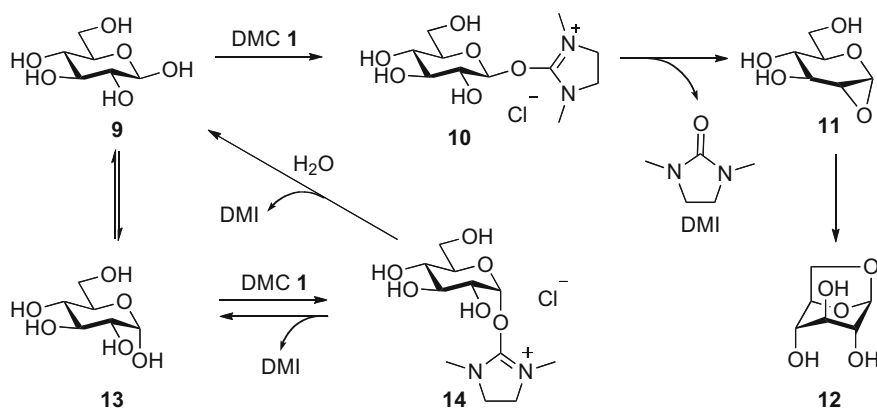
**Scheme 4** Proposed mechanism for 1,6-anhydro sugar formation [25]

The published optimised method uses a large excess of triethylamine to drive 1,6-anhydro sugar formation. In cases where other nucleophiles are to be introduced at the anomeric centre, e.g., for glycosyl azide synthesis using DMC (see Sect. 5) [30], fewer equivalents of base are commonly used. However, in these instances, 1,6-anhydro sugar formation can still be an undesired competing side reaction, which reduces the yields of desired products and can be particularly problematic when oligosaccharides, synthesised using multi-step methods and obtained in small quantities, are used. In these cases, 1,6-anhydro sugar formation can typically be minimised by the use of a large excess of the external nucleophile. Alternatively, Shoda has also reported that acetonitrile may be used as a co-solvent to suppress 1,6-anhydro sugar formation [31].

4 Synthesis of Thioglycosides Using DMC

Thioglycosides have found widespread application throughout the carbohydrate field as glycosyl donors for chemical oligosaccharide synthesis [32–35]. There has also been interest in the synthesis of de-protected thioglycosides as stable analogues of *O*-glycosides and as potential enzyme inhibitors [36, 37]. The introduction of sulfur at the anomeric centre often involves the reaction of a peracetylated sugar with a thiol in the presence of a Lewis acid [38], or alternatively substitution of a glycosyl halide with a thiolate [39]. In general, these methods require multi-step reaction sequences, involving protection/de-protection strategies, and are usually performed in organic solvents. Although direct methods for the preparation of thioglycosides from hemiacetals have been reported using trifluoroacetic acid as an activator, these reactions gave poor anomeric selectivities and also resulted in the formation of dithioacetal by-products [40, 41].

4.1 Aryl Thioglycosides

In 2009, Tanaka et al. [31] reported a simple method for the direct synthesis of aryl thioglycosides from reducing sugars in an aqueous solvent system using DMC and excess triethylamine (Scheme 5).
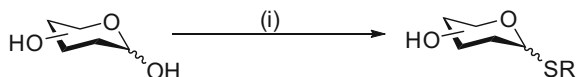
Scheme 5 Synthesis of aryl thioglycosides from unprotected sugars [31]. Reaction conditions: **i** DMC **1**, R-SH, Et$_3$N, H$_2$O/MeCN

The use of benzenethiol, *p*-toluenethiol, and 4-methoxybenzenethiol as the thiol component afforded the corresponding aryl thioglycosides as mixtures of anomers, in high yield. However, when 4-nitrobenzenethiol was reacted with D-glucose and DMC, the corresponding thioglycoside product was formed exclusively as the β-anomer, in 90% yield. It was found that the procedure could also be applied to disaccharides, and the corresponding products were formed in quantitative yield, and as exclusively the β-anomer.
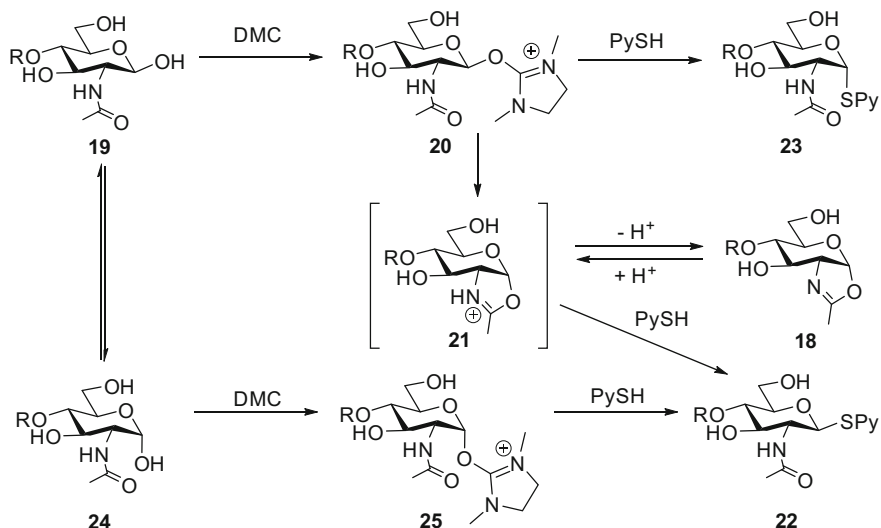
4.2   2-Pyridyl 1-Thio-Glycosides

Shoda et al. [31, 42] then demonstrated that it was also possible to glycosylate 2-mercaptopyridine **16** with oligosaccharides having an *N*-acetylglucosamine unit at the reducing end **15** using DMC to give the corresponding 2-pyridylthioglycosides **17** (Scheme 6).

Interestingly, the glycosyl oxazoline **18** was obtained as a by-product from these reactions. However, it was found that treating oxazoline **18** with 1 M HCl led to the opening of the ring and an increased yield of the desired 2-pyridylthioglycoside **17**. A plausible mechanism, which explains the formation of both anomers of the 2-pyridylthioglycoside products as well as the oxazoline, was suggested and is shown in Scheme 7. Nucleophilic attack of the β-hemiacetal of **19** on DMC affords **20**, which may then cyclise to give the oxazolinium intermediate **21**. Attack on intermediate **21** by mercaptopyridine (PySH) will yield the β-2-pyridylthioglycoside **22**. However, if the proton on the nitrogen atom of **21** is abstracted by a base, the intermediate will be converted to the corresponding oxazoline **18**, which is itself unreactive towards nucleophilic opening. The addition of acid to oxazoline **18** leads to the re-protonation of the nitrogen to regenerate **21**, which can subsequently react with mercaptopyridine to give **22**. Alternatively, imidazolidinium intermediate **20** can react directly with mercaptopyridine to give the α-2-pyridylthioglycoside **23**.

The α-anomer of the starting material **24** can also react with DMC to give intermediate **25**, which can be directly attacked by mercaptopyridine to give the β-2-pyridylthioglycoside **22**.
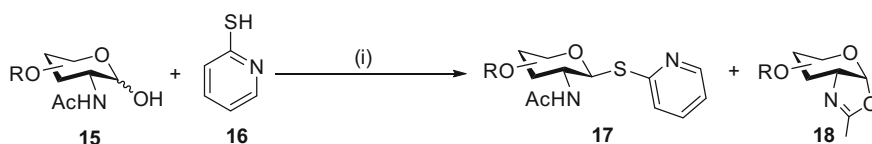
Scheme 6 Glycosylation of mercaptopyridine **16** with unprotected sugars using DMC reported Shoda et al. [42] Reaction conditions: **i** DMC **1**, Et$_3$N, H$_2$O/MeCN (4:1), 0 °C, 1 h

Scheme 7 Proposed mechanism for 2-pyridylthioglycoside formation using PySH and DMC [42]

4.3 4-Methyl-7-Thioumbelliferyl-Labelled Glycosides (MUS-Labelled Glycosides)

There is a growing need to develop quantitative ways to label sugars with chromophores. The most common current method of labelling is reductive animation, in which an oligosaccharide is reacted with an amine containing the chromophore, and the ensuing imine is then reduced by a source of hydride [43]. However, sugars labelled in this manner contain an open-chain 1-aminoalditol moiety, essentially rendering the reducing end inert for further useful transformations. This ring opening also reduces the structural relevance of that part of the oligosaccharide, for example with respect to binding or conformational studies. Furthermore, the conversion of 1-aminoalditols back into reducing sugars requires harsh conditions, such as the use of hydrogen peroxide [44, 45]. A method for introducing a detectable chromophore that could also be easily detached would, therefore, be highly advantageous.

In 2013, Shoda et al. [46] reported the direct introduction of 4-methyl-7-thioumbelliferone **26** (MUS) at the anomeric centre of monosaccharides and at the reducing terminus of oligosaccharides using DMC activation (Scheme 8). The fluorescence spectrum of the MUS-labelled glycosides produced indicated that the labelled derivatives showed a high sensitivity for fluorescence detection based on the maximum wavelengths for excitation and emission at 330 and 395 nm, respectively.

The yields of the MUS-sugars made by this DMC labelling method were sufficiently high to suggest its application in quantitative studies. Indeed, analysis of a mixture of laminari-oligosaccharides of known composition, which were then modified by both MUS-labelling and pyridylamination, showed very similar
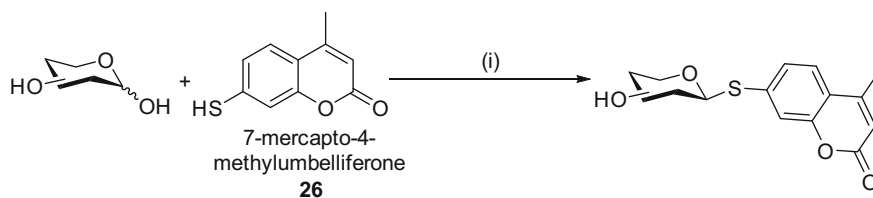
Scheme 8 Synthesis of MUS-labelled sugars using DMC by Shoda et al. [46]. Reaction conditions: **i** DMC **1**, Et$_3$N, H$_2$O/MeCN (1:1), 0 °C, 1.5 h

Table 1 Quantitative analysis of MUS-labelled lamnari-oligosaccharide structures by Shoda et al. [46]

DP	Actual amount (mol%)	MUS-labelling (area%)	Pyridylamination (area%)
1	36.9	34.5	37.6
2	19.4	21.6	19.0
3	13.2	14.4	12.7
4	10.0	10.6	10.0
5	8.0	8.2	8.2
6	6.7	6.2	7.0
7	5.8	4.5	5.5

quantitative values (Table 1). The MUS-glycosides were also easily de-protected by conventional halosuccinimide-mediated thioglycoside activation (e.g. *N*-bromo-, *N*-chloro-, and *N*-iodosuccinimide) to re-form the corresponding reducing sugars, which could then be used for further functionalisation; a significant advantage compared to the pyridylamination labelling method.

4.4 Glycopolymerisation

Although the binding of carbohydrates to proteins is a fundamental process, with widespread importance throughout Biology, it is now appreciated that the majority of carbohydrate–protein interactions are of low affinity. Nature has compensated for this apparent paradox by the use of multivalent receptor-ligand presentation, which amplifies the affinity of single interactions. It is fair to say that all aspects of this so-called 'glycocluster effect' [47, 48] are not yet completely understood, although detailed discussions and theories have been presented [49, 50]. Despite shortcomings in our understanding, there have been many reports on the design and production of synthetic glycoclusters, such as glycopolymers [51], glycodendrimers [52], and glyconanoparticles [53]. Many of these synthetic constructs have demonstrated an amplification in carbohydrate-mediated binding, in a similar manner to that achieved by multivalent carbohydrate presentations of natural glycopeptides [54] and glycoproteins [55].
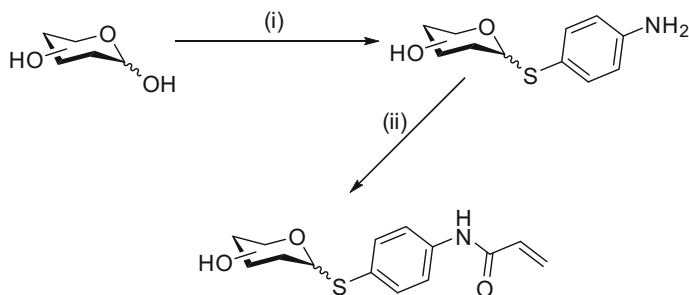
Scheme 9 One-pot synthesis of glycomonomers [58]. Reaction conditions: **i** DMC **1**, HS–C$_6$H$_4$–NH$_2$, Et$_3$N, H$_2$O/MeCN; **ii** acryloyl chloride, Et$_3$N, H$_2$O/THF

The preparation of glycopolymers often requires laborious, multi-step procedures involving protection and de-protection of a saccharide and which must also include the introduction of a polymerizable group, such as vinyl or norbornene, at the anomeric centre [56, 57]. Only after multiple synthetic steps can the resulting gly-comonomers be used in a polymerization reaction. Clearly, more efficient synthetic methodologies for the preparation of glycomonomers would be advantageous.

In 2014, Tanaka et al. [58] reported a one-pot method for the production of glycomonomers directly from unprotected sugars, by the DMC-mediated synthesis of 4-aminophenyl 1-thio-glycosides in water in the presence of triethylamine, and their subsequent acrylamidation (Scheme 9).

These acylamide containing glycomonomers were then subjected to reversible addition-fragmentation chain transfer (RAFT) living radical polymerisation to give glycopolymers, which were subsequently immobilised onto gold nanoparticles for investigations into glycocluster effects.

4.5 S-*Linked Glycopeptides*

S-Linked peptide and protein glycosylation, where the anomeric oxygen of *O*-linked glycosides has been replaced by sulfur, has attracted a significant amount of interest from the Glycoscience community [59–61]. *S*-Linked glycopeptides are well known for their functional surrogacy, in addition to their chemical and biological stability as compared to their *O*-linked counterparts, especially towards glycosidase-catalysed cleavage [62].

S-Linked glycopeptides can be accessed by several strategies. The most common involve either conjugate addition or nucleophilic substitution reactions, with a glycosyl thiol acting as the nucleophile [63–66]. However, more recent approaches, such as free radical thiol-ene 'Click' reactions [67], desulfurative rearrangements [68], and the opening of 1,6-anhydrosugars [69], are alternatives that may possess certain advantages. Nevertheless, all of these methods still require multi-step methodologies in order to prepare the required glycosyl thiol.

Winssinger et al. [70] recently reported the use of DMC for the solid phase synthesis of *S*-linked glycopeptides. By using a H₂O/dioxane solvent mixture and cooling to −10 °C, they demonstrated that unprotected sugars could be 'pre-activated' by DMC, and found the glycosyl-imidazolinium intermediate that was formed was stable for up to 1 h. The addition of a polymer-bound peptide containing a thiol to the reaction mixture during that time period then led to the formation of the desired polymer-bound *S*-linked glycopeptide **27** (Scheme 10).

An attempt was also made to glycosylate a polymer-bound tripeptide containing cysteine (Fmoc-Cys-Phe-Phe-Rink) with various mono- and disaccharides. Interestingly, direct reaction with D-glucose with DMC in the presence of the tripeptide **28** (i.e. without the pre-activation process) led to the exclusive formation of the thioimidazolidinium by-product **29** (Scheme 11a). However, when the conditions previously identified for pre-activation were used, the desired glycopeptide **30** was obtained in >90% yield (Scheme 11b).

The process was applied to other monosaccharides (D-galactose, D-mannose, L-fucose) and also to disaccharides (lactose, mellibiose, and Galα(1 → 4)Glc), though the yields were lower in the cases of the disaccharides. An elegant method, involving an iterative addition strategy, was found to increase efficiency of glycopeptide formation, and led to conversions to the desired product of >99%. The process was then exemplified by the solid phase synthesis of an analogue of the repeat unit of the cancer-associated MUC1 glycopeptide, in which the two natural *O*-linked carbohydrates sites were replaced with *S*-linked glycans.
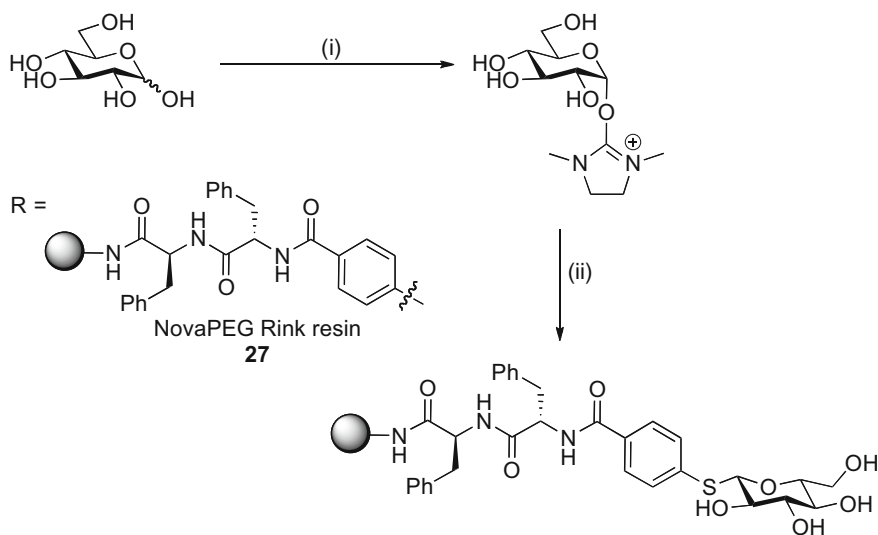
Scheme 10 Glycosylation of a thiol-containing resin with D-glucose using the DMC methodology [70]. Reaction conditions: **i** DMC **1**, Et₃N, H₂O/dioxane (1:1), −10 °C; **ii** RSH
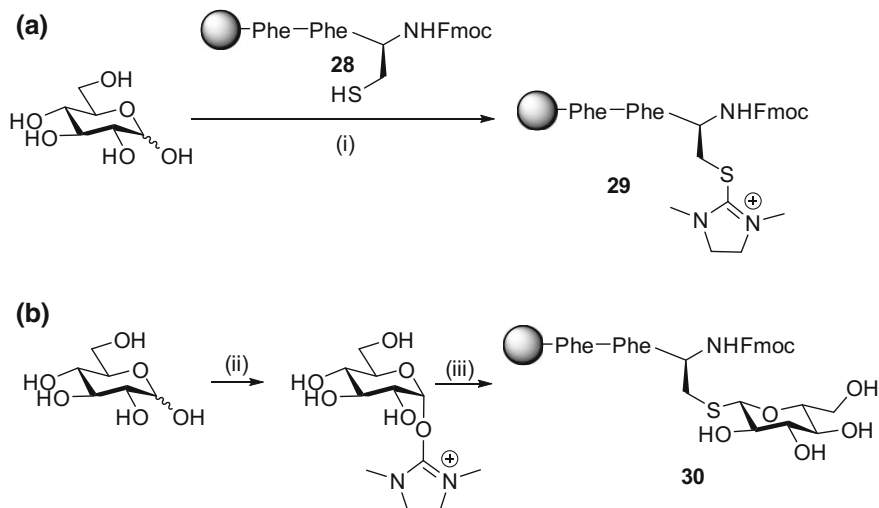
Scheme 11 Solid phase *S*-glycosylation of a resin-bound tripeptide **28** by Winssinger et al. [70]. Reaction conditions: **28**, DMC **1**, Et₃N, H₂O/dioxane (1:1), −10 °C; **ii** DMC **1**, Et₃N, H₂O/dioxane, −10 °C, 15 min; **iii 28**

5 Synthesis of Glycosyl Azides in Water Using DMC

Glycosyl azides are highly useful synthetic intermediates in sugar chemistry [71]. The traditional method for their formation involves at least three or four steps, invariably involving protecting group manipulations, to furnish the desired de-protected product. For example, a typical sequence involves the conversion of the sugar, e.g. *N*-acetyl-D-glucosamine **31**, into a protected glycosyl halide **32**, followed by nucleophilic displacement of the anomeric leaving group by azide to give the glycosyl azide **33** (Scheme 12). Finally, the ester protecting groups are removed, for example by Zemplén de-acetylation, to give the de-protected glycosyl azide **34**.

In addition to the inherent inefficiency of these multi-step reaction sequences, this type of synthetic route can also occasionally become problematic when applied to higher oligosaccharides due to cleavage of inter-glycosidic bonds during the synthetic pathway.

Following on from earlier studies in the Shoda group, Tanaka et al. [30] reported that glycosyl azides of monosaccharides could be formed directly when unprotected sugars were treated with excess DMC in water in the presence of a large excess (typically 10 equivalents of each) of azide and triethylamine (Scheme 13).

Interestingly, when the reaction was applied to disaccharides using triethylamine as the base, 1,6-anhydro sugars were formed as by-products. It was suggested that, for some reason, intramolecular nucleophilic attack of the 6-hydroxy group at the anomeric centre was increased in the presence of triethylamine [30]. A screen of various other bases was performed, and it was found that the use of either *N,N*-diisopropylethylamine (Hunig's base, DIPEA) or 2,6-lutidine led to reduced
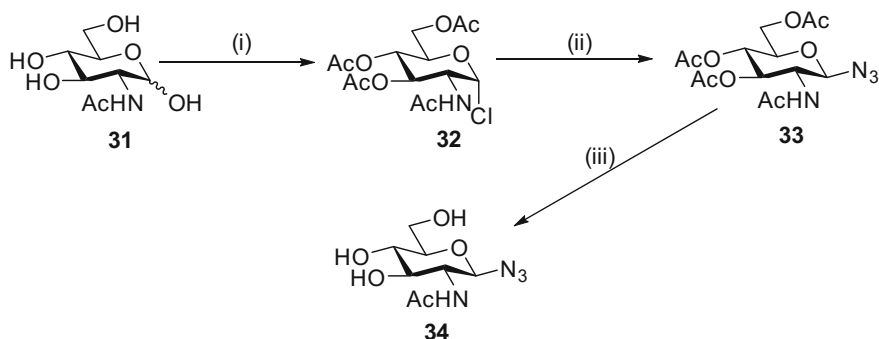
Scheme 12 Synthesis of de-protected glycosyl azides. Reaction conditions: **i** AcCl; **ii** NaN$_3$, sat. aq. NaHCO$_3$, tBuNH$_4$·HSO$_4$, DCM; **iii** Na, MeOH
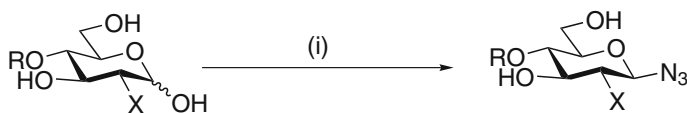
where X = NHAc or OH
base = Et$_3$N, DIPEA, or 2,6-lutidine

Scheme 13 Direct synthesis of glycosyl azides from unprotected sugars using DMC [30]. Reaction conditions: **i** DMC **1**, NaN$_3$, base, H$_2$O

formation of the 1,6-anhydro derivative, presumably an effect that was related to their greater steric bulk. The use of 2,6-lutidine as the base was also found to be essential in the case of 2-acetamido sugars to avoid significant oxazoline formation.

Tanaka et al. [30] then demonstrated that the reaction could be applied to larger oligosaccharides, including a sialic acid terminated complex biantennary decasaccharide, which was converted to the corresponding glycosyl azide in 87% yield, though this reaction did require the use of 20 equivalents of DMC and 40 equivalents of 2,6- lutidine. Glycosyl azides were generally produced stereoselectively as the 1,2-*trans* glycosides, except in the cases of 2-deoxy sugars, which were formed as anomeric mixtures. To date, a detailed mechanism has not been presented, though the results suggest that reaction of sugars containing a 2-hydroxy group probably proceeds via 1,2-anhydro sugar intermediates.

6 Synthesis of Glycosyl Triazoles in a One-Pot Reaction

The original Shoda method [30] for glycosyl azide synthesis, like all DMC activation procedures, potentially generates up to two equivalents of HCl per equivalent of DMC used; one equivalent is produced from de-protonation of the anomeric hydroxyl group, and a second one may additionally be produced from water in the

case of a hydrolysis reaction. This production of acid during the course of the reaction necessitates the use of excessive amounts of triethylamine. Furthermore, the use of a large excess of azide is also required.

As an alternative approach which would avoid the requirement for this large excess of reagents, we envisaged replacement of the chlorine on the dimethylimidazolinium ring of DMC with azide; such a reagent should still be able to activate a reducing sugar, would only result in the formation of one equivalent of acid, and would additionally itself act as the source of azide. 2-Azido-1,3-dimethylimidazolinium hexafluoro phosphate (ADMP), has been previously used as an efficient agent for diazo-transfers [72–75], migratory aminations [76, 77], and azide transfer [78] reactions. Although ADMP contains a significant amount of nitrogen, impact sensitivity and friction sensitivity tests have demonstrated that it is not explosive [73], making it in fact safer to handle than sodium azide. Additionally unlike DMC, which is hygroscopic [23, 73], ADMP is isolated as a stable crystalline solid and is easily handled. Although, unlike DMC, it is not currently commercially available, ADMP can be very easily synthesised, for example following procedures originally reported by Kitamura et al. [72, 79] (Scheme 14), simply by conversion of DMC **1** to the hexafluorophosphate salt **35**, and then treatment with sodium azide to give ADMP **36**.

In 2014, we [80] reported an alternative method for the synthesis of glycosyl azides in water. N-Acetylglucosamine was reacted with ADMP **36** and triethylamine in a $D_2O/MeCN$ (4:1) solvent mixture to give the glycosyl azide **33**, as well as oxazoline **37** (Scheme 15). The use of D_2O as solvent, as first reported by Shoda in the original oxazoline formation study [8], resulted in higher yields of products, presumably due to reduced rates of competitive solvent-mediated hydrolysis.
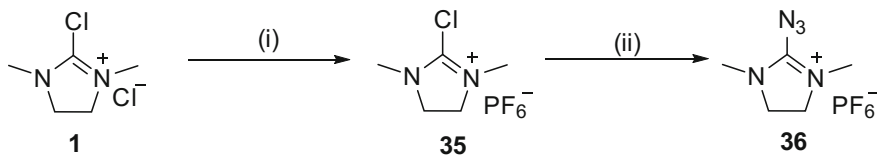
Scheme 14 Synthesis of ADMP **36** [72, 79]. Reaction conditions: **i** $NaPF_6$, MeCN, rt, 30 min, 99%; **ii** NaN_3, MeCN, 0 °C, 3 h, quant
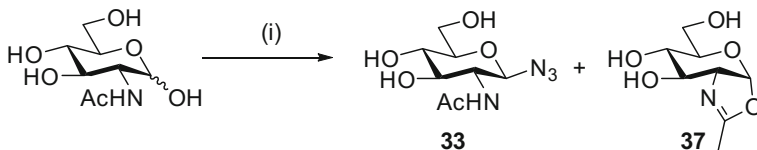

Scheme 15 Reaction of GlcNAc with ADMP to yield glycosyl azide **33** and oxazoline **37** [80]. Reaction conditions: **i** ADMP **36**, Et_3N, $D_2O/MeCN$ (4:1), 0 °C

However, applying the idea reported by Shoda et al. [42] that the subsequent addition of acid would result in oxazoline ring opening, we found that the desired azide **33** was obtained as the sole reaction product when the crude reaction mixture was simply acidified with 1 M aqueous HCl.

Glycosyl azides are obvious substrates for further elaboration by the modified Huisgen cycloaddition [81, 82], the most well known 'Click' reaction. This high yielding transformation may be performed in water, leading to numerous applications, as reaction conditions are compatible with biological systems. In the carbohydrate field, 'Click chemistry' has already been applied to the synthesis of a wide range of sugar derivatives [83] having interesting biological properties [84], including inhibitory activity against glycosidases [85–90] and glycosyltransferases [91]. The development of a one-pot process for glycosyl azide and then triazole formation by Click reaction with an alkyne was, therefore, an obvious avenue for investigation.

We found that after the formation of the glycosyl azide was complete, the addition of propargyl alcohol, $CuSO_4 \cdot 5H_2O$, and L-ascorbic acid, and then heating at 50 °C for 14 h, led to the formation of the glycosyl triazole **38** in excellent yield and with complete stereoselectivity; the 1,2-*trans* glycosyl triazole being formed in all cases (Scheme 16). This procedure was then applied to a variety of sugars ranging from in size from mono- to trisaccharides [80]. In all cases, the use of D_2O as co-solvent led to a variable, but minor, the amount of deuterium incorporation into the triazole ring during the course of the reaction.

The generality of the one-pot process with respect to the alkyne was explored. The reaction was found to be widely applicable, including with interesting and potentially biologically relevant substrates such as propargyl glycine. All alkynes investigated were successfully Clicked with monosaccharides, and products formed in high yield and with complete stereoselectivity. The only limitation appeared to be in terms of the solubility of the alkyne-coupling partner in the aqueous reaction medium. The one-pot Click reaction also allowed the direct conjugation of reducing sugars to a variety of other carbohydrates which themselves contained an alkyne

38
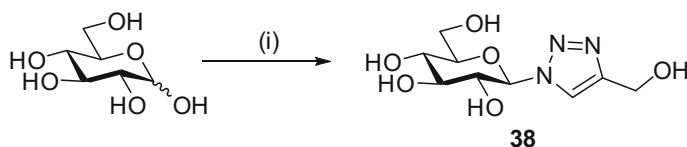
Scheme 16 Direct synthesis of glycosyl triazoles from reducing sugars in water by Fairbanks et al. [80]. Reaction conditions: **i** (1) ADMP **36**, Et_3N, $D_2O/MeCN$ (4:1), 0 °C, 3 h, then add propargyl alcohol, $CuSO_4 \cdot 5H_2O$, L-ascorbic acid, 50 °C for 14 h

functionality, giving rise to a variety of di-, tri- and pentasaccharide mimics. All reactions were high yielding and completely stereoselective [80].

Click chemistry has previously been applied to access different glycopeptides, [92, 93] and also protein scaffolds decorated with oligosaccharides [94]. However,

Scheme 17 Conjugation of GalNAc with 10-mer **39** and 20-mer **40** by Fairbanks et al. [80]; Reaction conditions: **i** GalNAc, ADMP, Et$_3$N, D$_2$O, MeCN, 0 °C, 3 h, then add alkyne, CuSO$_4$·5H$_2$O, L-ascorbic acid, and heat to 50 °C for 14 h [80]
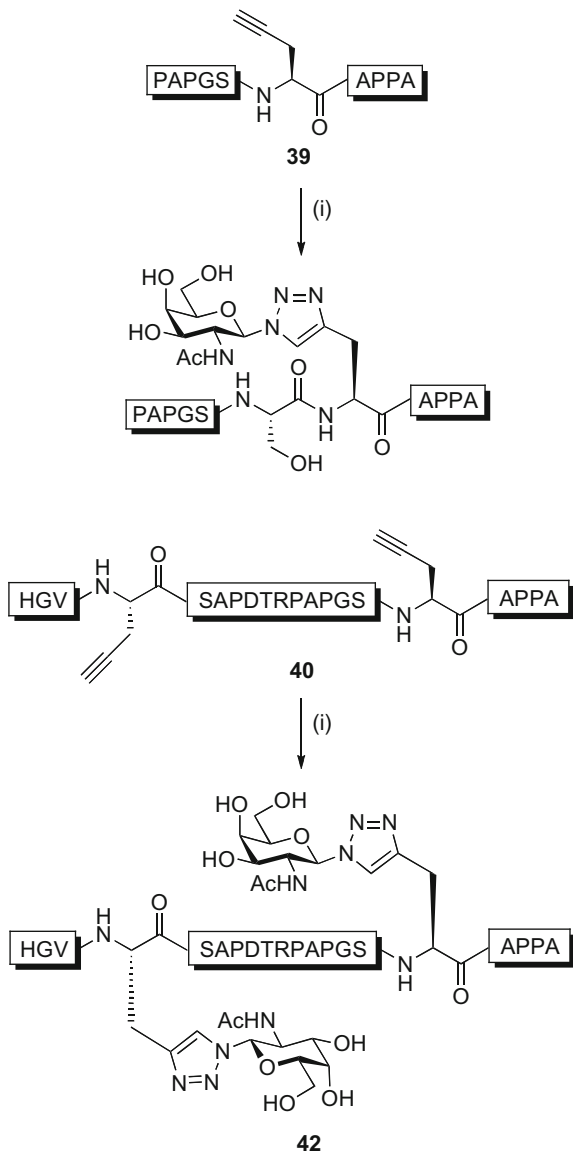
Scheme 18 Reaction of tetrasaccharide **43** with 10-mer **39** by Fairbanks et al. [80]. Reaction conditions: **i** ADMP, Et$_3$N, D$_2$O, MeCN, 0 °C, 3 h, then add 10-mer **39**, CuSO$_4$·5H$_2$O, L-ascorbic acid, and heat to 50 °C for 14 h
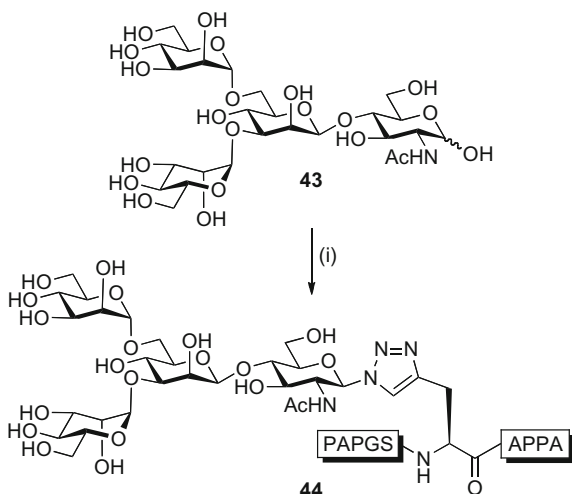
reported procedures require multiple protecting group manipulations. In terms of a biologically interesting application, glycosylated versions of the tandem repeat domain of the cancer-associated mucin MUC1 [95, 96] have shown potential as components of synthetic anti-cancer vaccines [97, 98]. A further exemplification of the one-pot azide/Click reaction, we, therefore, conjugated a variety of sugars to two synthetic MUC1 peptides that incorporated propargyl glycine (Pra) residues. Conversion of GalNAc to its azide using ADMP and then direct reaction with peptides **39** and **40**, comprising one and two propargyl glycines, respectively, in the presence of CuSO$_4$·5H$_2$O, gave the corresponding glycopeptides **41** and **42** in good yield (Scheme 17).

Similarly tetrasaccharide **43** [99], which corresponds to a core region of *N*-glycans, was also directly reacted with peptide **39** and gave glycopeptide **44** in 42% yield (Scheme 18).

As a final example of the utility of the method, the complex biantennary *N*-glycan **6** decasaccharide [21, 100] was converted to the glycosyl azide using ADMP, and then directly reacted with peptide **39** in the presence of CuSO$_4$·5H$_2$O to furnish glycopeptide **45** in 42% yield (Scheme 19).
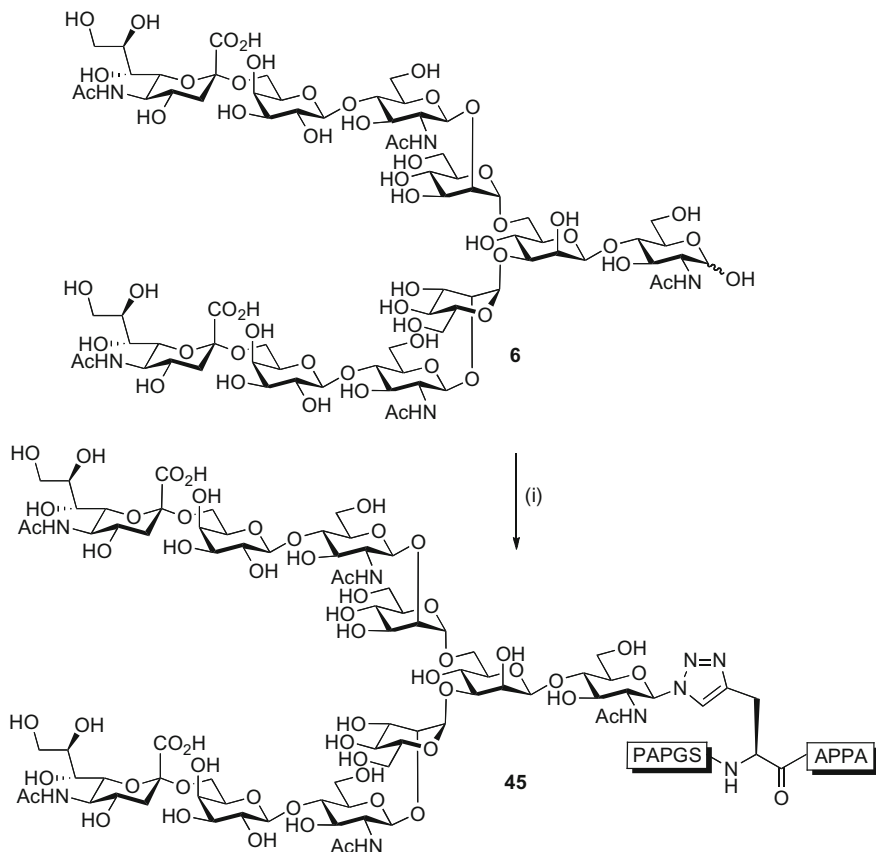
Scheme 19 Conjugation of decasaccharide **6** to 10-mer **45** by Fairbanks et al. [80]. Reaction conditions: **i 6**, ADMP, Et$_3$N, D$_2$O, MeCN, 0 °C, 3 h, then add 10-mer **39**, CuSO$_4$·5H$_2$O, L-ascorbic acid, and heat to 50 °C for 14 h

7 Conclusion

Traditional methods of accessing glycoconjugates typically involve multi-step syntheses and require complex and protracted protecting group strategies. These methods are generally technically demanding, inefficient, expensive, and the production of significant amounts of the target material is usually logistically difficult to achieve.

Since the initial report by Shoda and co-workers on the use of 1,3-dimethylimidazolinium chloride (DMC) for the selective conversion of unprotected 2-acetamido sugars to glycosyl oxazolines in water, various synthetic targets have been accessed using this, or related, highly selective activating agent. In particular, selective activation of the anomeric centre under aqueous conditions

and direct reaction with a diverse range of nucleophiles allows ready access to a range of glycoconjugates without recourse to any protecting group manipulations. Additionally, the products of nucleophilic substitution at the anomeric centre may be further derivatised in the same reaction vessel, for example by Click chemistry.

As a wide variety of structurally complex reducing oligosaccharides are available from natural sources, protecting group free methods based on DMC and similar activating agents show significant promise for rapid access to a broad range of biologically interesting glycoconjugates.

References

1. Donohoe TJ, Logan JG, Laffan DDP (2003) Trichloro-oxazolines as activated donors for aminosugar coupling. Org Lett 5:4995–4998
2. Blatter G, Beau J-M, Jacquinet J-C (1994) The use of 2-deoxy-2- trichloroacetamido-D-glucopyranose derivatives in syntheses of oligosaccharides. Carbohydr Res 260:189–202
3. Fairbanks AJ (2011) Endohexosaminidase catalysed glycosylation with oxazoline donors: the development of robust biocatalytic methods for synthesis of defined homogeneous glycoconjugates. C R Chim 14:44–58
4. Fairbanks AJ (2013) Endohexosaminidase-catalyzed synthesis of glycopeptides and proteins. Pure Appl Chem 85:1847–1863
5. Kadokawa J, Mito M, Takahashi S et al (2004) Direct conversion of 2-Acetamido-2-deoxysugars to 1,2-Oxazoline derivatives by dehydrative cyclization in water. Heterocycles 63:1531–1535
6. Isobe T, Ishikawa T (1999) 2-Chloro-1,3-dimethylimidazolinium chloride. 2. Its application to the construction of heterocycles through dehydration reactions. J Org Chem 64:6989–6992
7. Isobe T, Ishikawa T (1999) 2-Chloro-1,3-dimethylimidazolinium chloride. 1. A powerful dehydrating equivalent to DCC. J Org Chem 64:6984–6988
8. Noguchi M, Tanaka T, Gyakushi H et al (2009) Efficient synthesis of sugar oxazolines from unprotected N-acetyl-2-amino sugars by using chloroformamidinium reagent in water. J Org Chem 74:2210–2212
9. Matta KL, Johnson EA, Barlow JJ (1973) A simple method for the synthesis of 2-acetamido-2-deoxy-β-D-galactopyranosides. Carbohydr Res 26:215–218
10. Srivastava VK (1982) A facile synthesis of 2-methyl-(3,4,6-tri-O-acetyl-1,2-dideoxy-α-D-glucopyrano)-[2,1-e]-2-oxazoline. Carbohydr Res 103:286–292
11. Nakabayashi S, Warren CD, Jeanloz RW (1986) A new procedure for the preparation of oligosaccharide oxazolines. Carbohydr Res 150:c7–c10
12. Colon M, Staveski MM, Davis JT (1991) Mild conditions for the preparation of high-mannose oligosaccharide oxazolines: entry point for β-glycoside and neoglycoprotein syntheses. Tetrahedron Lett 32:4447–4450
13. Kowalczyk R, Brimble MA, Tomabechi Y et al (2014) Convergent chemoenzymatic synthesis of a library of glycosylated analogues of pramlintide: structure–activity relationships for amylin receptor agonism. Org Biomol Chem 12:8142–8151
14. McIntosh JD, Brimble MA, Brooks AES et al (2015) Convergent chemo-enzymatic synthesis of mannosylated glycopeptides; targeting of putative vaccine candidates to antigen presenting cells. Chem Sci 6:4636–4642

15. Tomabechi Y, Krippner G, Rendle PM et al (2013) Glycosylation of pramlintide: synthetic glycopeptides that display in vitro and in vivo activities as amylin receptor agonists. Chem Eur J 19:15084–15088

16. Tomabechi Y, Squire MA, Fairbanks AJ (2014) Endo-β-N-acetylglucosaminidase catalysed glycosylation: tolerance of enzymes to structural variation of the glycosyl amino acid acceptor. Org Biomol Chem 12:942–955

17. Lomino JV, Naegeli A, Orwenyo J et al (2013) A two-step enzymatic glycosylation of polypeptides with complex N-glycans. Bioorg Med Chem 21:2262–2270

18. Orwenyo J, Huang W, Wang L-X (2013) Chemoenzymatic synthesis and lectin recognition of a selectively fluorinated glycoprotein. Bioorg Med Chem 21:4768–4777

19. Smith EL, Giddens JP, Iavarone AT et al (2014) Chemoenzymatic Fc glycosylation via engineered aldehyde tags. Bioconjug Chem 25:788–795

20. Umekawa M, Huang W, Li B et al (2008) Mutants of mucor hiemalis endo-β-N-acetylglucosaminidase show enhanced transglycosylation and glycosynthase-like activities. J Biol Chem 283:4469–4479

21. Umekawa M, Higashiyama T, Koga Y et al (2010) Efficient transfer of sialo-oligosaccharide onto proteins by combined use of a glycosynthase-like mutant of Mucor hiemalis endoglycosidase and synthetic sialo-complex-type sugar oxazoline. Biochim Biophys Acta Gen Subj 1800:1203–1209

22. Umekawa M, Li C, Higashiyama T et al (2010) Efficient glycosynthase mutant derived from mucor hiemalis endo-β-N-acetylglucosaminidase capable of transferring oligosaccharide from both sugar oxazoline and natural N-glycan. J Biol Chem 285:511–521

23. Noguchi M, Fujieda T, Huang WC et al (2012) A practical one-step synthesis of 1,2-oxazoline derivatives from unprotected sugars and its application to chemoenzymatic β-N-acetylglucosaminidation of disialo-oligosaccharide. Helv Chim Acta 95:1928–1936

24. Sun B, Bao W, Tian X et al (2014) A simplified procedure for gram-scale production of sialylglycopeptide (SGP) from egg yolks and subsequent semi-synthesis of Man3GlcNAc oxazoline. Carbohydr Res 396:62–69

25. Tanaka T, Huang WC, Noguchi M et al (2009) Direct synthesis of 1,6-anhydro sugars from unprotected glycopyranoses by using 2-chloro-1,3-dimethylimidazolinium chloride. Tetrahedron Lett 50:2154–2157

26. Köll P, Metzger J (1978) Thermal degradation of cellulose and chitin in supercritical acetone. Angew Chem Int Ed 17:754–755

27. Miura M, Kaga H, Yoshida T, Ando K (2001) Microwave pyrolysis of cellulosic materials for the production of anhydrosugars. J Wood Sci 47:502–506

28. Sasaki M, Takahashi K, Haneda Y et al (2008) Thermochemical transformation of glucose to 1,6-anhydroglucose in high-temperature steam. Carbohydr Res 343:848–854

29. Köll P, Borchers G, Metzger JO (1991) Thermal degradation of chitin and cellulose. J Anal Appl Pyrolysis 19:119–129

30. Tanaka T, Nagai H, Noguchi M, et al. (2009) One-step conversion of unprotected sugars to β-glycosyl azides using 2-chloroimidazolinium salt in aqueous solution. Chem Commun 3378–3379

31. Tanaka T, Matsumoto T, Noguchi M et al (2009) Direct Transformation of unprotected sugars to Aryl 1-Thio-β-glycosides in aqueous media using 2-Chloro-1,3-dimethylimidazolinium chloride. Chem Lett 38:458–459

32. Sarkar S, Sucheck SJ (2011) Comparing the use of 2-methylenenapthyl, 4-methoxybenzyl, 3,4-dimethoxybenzyl and 2,4,6-trimethoxybenzyl as N-H protecting groups for p-tolyl 2-acetamido-3,4,6-tri-O-acetyl-2-deoxy-1-thio-β-D-glucosides. Carbohydr Res 346:393–400

33. Milhomme O, Dhénin SGY, Djedaïni-Pilard F et al (2012) Synthetic studies toward the anthrax tetrasaccharide: alternative synthesis of this antigen. Carbohydr Res 356:115–131

34. Ennis SC, Fairbanks AJ, Slinn CA et al (2001) N-Iodosuccinimide-mediated intramolecular aglycon delivery. Tetrahedron 57:4221–4230

35. Yasomanee JP, Demchenko AV (2014) Hydrogen bond mediated aglycone delivery: synthesis of linear and branched α-glucans. Angew Chem Int Ed 53:10453–10456

36. Rye CS, Withers SG (2004) The synthesis of a novel thio-linked disaccharide of chondroitin as a potential inhibitor of polysaccharide lyases. Carbohydr Res 339:699–703
37. Rempel BP, Withers SG (2008) Covalent inhibitors of glycosidases and their applications in biochemistry and biology. Glycobiology 18:570–586
38. Drouin L, Cowley AR, Fairbanks AJ, Thompson AL (2008) 4-Methoxyphenyl 2,3,4,6-tetra-O-acetyl-1-thio-α-D-mannopyranoside. Acta Crystallogr E 64:o1401–o1401
39. Pei Z, Dong H, Caraballo R, Ramström O (2007) Synthesis of positional thiol analogs of β-D-galactopyranose. Eur J Org Chem 4927–4934
40. Funabashi M, Arai S, Shinohara M (1999) Novel syntheses of diphenyl and/or trimethylene dithioacetals of mono- and oligosaccharides in 90% trifluoroacetic acid. J Carbohydr Chem 18:333–341
41. Yanase M, Funabashi M (2000) Stereoselective 1,2-cis-1-thioglycosidation of aldohexoses with tert-butyl mercaptan in 90% trifluoroacetic acid. J Carbohydr Chem 19:53–66
42. Yoshida N, Noguchi M, Tanaka T et al (2011) Direct dehydrative pyridylthio-glycosidation of unprotected sugars in aqueous media using 2-chloro-1,3-dimethylimidazolinium chloride as a condensing agent. Chem Asian J 6:1876–1885
43. Hase S (2010) Pyridylamination as a means of analyzing complex sugar chains. Proc Jpn Acad Ser B 86:378–390
44. Kallin E, Lonn H, Norberg T (1988) Derivatization procedures for reducing oligosaccharides, part 2: chemical transformation of 1-Deoxy-1-(4-trifluoroacetamidophenyl)aminoalditols. Glycoconj J 5:145–150
45. Suzuki S, Fujimori T, Yodoshi M (2006) Recovery of free oligosaccharides from derivatives labeled by reductive amination. Anal Biochem 354:94–103
46. Yoshida N, Fujieda T, Kobayashi A et al (2013) Direct introduction of detachable fluorescent tag into oligosaccharides. Chem Lett 42:1038–1039
47. Lee YC, Lee RT (1995) Carbohydrate-protein interactions: basis of glycobiology. Acc Chem Res 28:321–327
48. Lundquist JJ, Toone EJ (2002) The cluster glycoside effect. Chem Rev 102:555–578
49. Lee RT, Lee YC (2000) Affinity enhancement by multivalent lectin-carbohydrate interaction. Glycoconj J 17:543–551
50. Dam TK, Brewer CF (2010) Multivalent lectin—carbohydrate interactions, pp 139–164
51. Le Droumaguet B, Nicolas J (2010) Recent advances in the design of bioconjugates from controlled/living radical polymerization. Polym Chem 1:563
52. Tanaka K, Siwu ERO, Minami K et al (2010) Noninvasive imaging of dendrimer-type N-glycan clusters. In Vivo dynamics dependence on oligosaccharide structure. Angew Chem Int Ed 49:8195–8200
53. Poonthiyil V, Nagesh PT, Husain M et al (2015) Gold nanoparticles decorated with sialic acid terminated Bi-antennary N-glycans for the detection of influenza virus at nanomolar concentrations. ChemistryOpen 4:708–716
54. Glunz PW, Hintermann S, Williams LJ et al (2000) Design and synthesis of Le y -bearing glycopeptides that mimic cell surface le y mucin glycoprotein architecture. J Am Chem Soc 122:7273–7279
55. Yamamoto N, Tanabe Y, Okamoto R et al (2008) Chemical synthesis of a glycoprotein having an intact human complex-type sialyloligosaccharide under the boc and fmoc synthetic strategies. J Am Chem Soc 130:501–510
56. Roy R, Tropper FD, Romanowska A (1992) New strategy in glycopolymer synthesis. Preparation of antigenic water-soluble poly(acrylamide-co-p-acrylamidophenyl beta-lactoside). Bioconjug Chem 3:256–261
57. Fraser C, Grubbs RH (1995) Synthesis of glycopolymers of controlled molecular weight by ring-opening metathesis polymerization using well-defined functional group tolerant ruthenium carbene catalysts. Macromolecules 28:7248–7255
58. Tanaka T, Inoue G, Shoda S-I, Kimura Y (2014) Protecting-group-free synthesis of glycopolymers bearing thioglycosides via one-pot monomer synthesis from free saccharides. J Polym Sci A 1(52):3513–3520

59. Gamblin DP, Garnier P, van Kasteren S et al (2004) Glyco-SeS: selenenylsulfide-mediated protein glycoconjugation—a new strategy in post-translational modification. Angew Chem Int Ed 116:846–851

60. Bernardes GJL, Marston JP, Batsanov AS et al. (2007) A trisulfide-linked glycoprotein. Chem Commun 3145–3147

61. Brimble MA, Edwards PJ, Harris PWR et al (2015) Synthesis of the antimicrobial s-linked glycopeptide, glycocin F. Chem Eur J 21:3556–3561

62. Driguez H (2001) Thiooligosaccharides as tools for structural biology. ChemBioChem 2:311–318

63. Levengood MR, van der Donk WA (2007) Dehydroalanine-containing peptides: preparation from phenylselenocysteine and utility in convergent ligation strategies. Nat Protoc 1:3001–3010

64. Galonić DP, van der Donk WA, Gin DY (2003) Oligosaccharide-peptide ligation of glycosyl thiolates with dehydropeptides: synthesis of S-linked mucin-related glycopeptide conjugates. Chem Eur J 9:5997–6006

65. Thayer DA, Yu HN, Galan MC, Wong C-H (2005) A general strategy toward S-linked glycopeptides. Angew Chem Int Ed 44:4596–4599

66. Bernardes GJL, Grayson EJ, Thompson S et al (2008) From disulfide- to thioether-linked glycoproteins. Angew Chem Int Ed 47:2244–2247

67. Dondoni A, Massi A, Nanni P, Roda A (2009) A new ligation strategy for peptide and protein glycosylation: photoinduced thiol-ene coupling. Chem Eur J 15:11444–11449

68. Crich D, Yang F (2008) Synthesis of neoglycoconjugates by the desulfurative rearrangement of allylic disulfides. J Org Chem 73:7017–7027

69. Zhu X, Dere RT, Jiang J et al (2011) Synthesis of α-glycosyl thiols by stereospecific ring-opening of 1,6-anhydrosugars. J Org Chem 76:10187–10197

70. Novoa A, Barluenga S, Serba C, Winssinger N (2013) Solid phase synthesis of glycopeptides using Shoda's activation of unprotected carbohydrates. Chem Commun 49:7608–7610

71. Györgydeák Z, Thiem J (2006) Synthesis and transformation of glycosyl azides. Adv Carbohydr Chem Biochem 60:103–182

72. Kitamura M, Tashiro N, Miyagawa S, Okauchi T (2011) 2-Azido-1,3-dimethylimidazolinium salts: Efficient diazo-transfer reagents for 1,3-dicarbonyl compounds. Synthesis 1037–1044

73. Kitamura M, Kato S, Yano M et al (2014) A reagent for safe and efficient diazo-transfer to primary amines: 2-azido-1,3-dimethylimidazolinium hexafluorophosphate. Org Biomol Chem 12:4397–4406

74. Kitamura M, Yano M, Tashiro N et al (2011) Direct synthesis of organic azides from primary amines with 2-Azido-1,3-dimethylimidazolinium hexafluorophosphate. Eur J Org Chem 2011:458–462

75. Kitamura K, Shigeta M, Maezawa Y et al (2013) Preparation of L-vancosamine-related glycosyl donors. J Antibiot 66:131–139

76. Kitamura M, Murakami K, Shiratake Y, Okauchi T (2013) Synthesis of α-arylcarboxylic acid amides from silyl enol ether via migratory amidation with 2-Azido-1,3-dimethylimidazolinium hexafluorophosphate. Chem Lett 42:691–693

77. Kitamura M, Miyagawa S, Okauchi T (2011) Synthesis of α, α-diarylacetamides from benzyl aryl ketones using 2-azido-1,3-dimethylimidazolinium hexafluorophosphate. Tetrahedron Lett 52:3158–3161

78. Kitamura M, Koga T, Yano M, Okauchi T (2012) Direct synthesis of organic azides from alcohols using 2-Azido-1,3-dimethyl-imidazolinium hexafluorophosphate. Synlett 23:1335–1338

79. Kitamura M (2015) Synthesis Of 2-Azido-1,3-dimethylimidazolinium hexafluorophosphate (ADMP). Org Synth 92:171–181

80. Lim D, Brimble MA, Kowalczyk R et al (2014) Protecting-group-free one-pot synthesis of glycoconjugates directly from reducing sugars. Angew Chem Int Ed 53:11907–11911

81. Tornøe CW, Christensen C, Meldal M (2002) Peptidotriazoles on solid phase: [1,2,3]-triazoles by regiospecific copper(I)-catalyzed 1,3-dipolar cycloadditions of terminal alkynes to azides. J Org Chem 67:3057–3064

82. Rostovtsev VV, Green LG, Fokin VV, Sharpless KB (2002) A stepwise huisgen cycloaddition process: copper(i)-catalyzed regioselective "ligation" of azides and terminal alkynes. Angew Chem Int Ed 41:2596–2599

83. Dondoni A (2007) Triazole: the keystone in glycosylated molecular architectures constructed by a click reaction. Chem-Asian J 2:700–708

84. Wilkinson BL, Long H, Sim E, Fairbanks AJ (2008) Synthesis of arabino glycosyl triazoles as potential inhibitors of mycobacterial cell wall biosynthesis. Bioorg Med Chem Lett 18:6265–6267

85. El Akri K, Bougrin K, Balzarini J et al (2007) Efficient synthesis and in vitro cytostatic activity of 4-substituted triazolyl-nucleosides. Bioorg Med Chem Lett 17:6656–6659

86. Rossi LL, Basu A (2005) Glycosidase inhibition by 1-glycosyl-4-phenyl triazoles. Bioorg Med Chem Lett 15:3596–3599

87. Wilkinson BL, Innocenti A, Vullo D et al (2008) Inhibition of carbonic anhydrases with glycosyltriazole benzene sulfonamides. J Med Chem 51:1945–1953

88. Wilkinson BL, Bornaghi LF, Houston TA et al (2006) A novel class of carbonic anhydrase inhibitors: glycoconjugate benzene sulfonamides prepared by "click-tailing". J Med Chem 49:6539–6548

89. De las Heras FG, Alonso R, Alonso G (1979) Alkylating nucleosides. 1. Synthesis and cytostatic activity of N-glycosyl(halomethyl)-1,2,3-triazoles. A new type of alkylating agent. J Med Chem 22:496–501

90. De las Heras FG, Camarasa M-J (1982) Synthesis of Alkylating 1-Glycosyl-5-substituted 1,2,4-Triazoles 1. Nucleos Nucleot 1:45–56

91. Yeoh KK, Butters TD, Wilkinson BL, Fairbanks AJ (2009) Probing replacement of pyrophosphate via click chemistry; synthesis of UDP-sugar analogues as potential glycosyl transferase inhibitors. Carbohydr Res 344:586–591

92. Li H, Aneja R, Chaiken I (2013) Click chemistry in peptide-based drug design. Molecules 18:9797–9817

93. Tomabechi Y (2015) Synthesis of glycopeptides by click chemistry. Trends Glycosci Glycotechnol 27:63–65

94. Wang H, Huang W, Orwenyo J et al (2013) Design and synthesis of glycoprotein-based multivalent glyco-ligands for influenza hemagglutinin and human galectin-3. Bioorg Med Chem 21:2037–2044

95. Hanisch F-G, Muller S (2000) MUC1: the polymorphic appearance of a human mucin. Glycobiology 10:439–449

96. Sherblom AP, Moody CE (1986) Cell surface sialomucin and resistance to natural cell-mediated cytotoxicity of rat mammary tumor ascites cells. Cancer Res 46:4543–4546

97. Kaiser A, Gaidzik N, Westerlind U et al (2009) A synthetic vaccine consisting of a tumor-associated sialyl-T N-MUC1 tandem-repeat glycopeptide and tetanus toxoid: induction of a strong and highly selective immune response. Angew Chem Int Ed 48:7551–7555

98. Lakshminarayanan V, Thompson P, Wolfert MA et al (2012) Immune recognition of tumor-associated mucin MUC1 is achieved by a fully synthetic aberrantly glycosylated MUC1 tripartite vaccine. Proc Natl Acad Sci 109:261–266

99. Rising TWDF, Heidecke CD, Moir JWB et al (2008) Endohexosaminidase-catalysed glycosylation with oxazoline donors: fine tuning of catalytic efficiency and reversibility. Chem Eur J 14:6444–6464

100. Seko A, Koketsu M, Nishizono M et al (1997) Occurrence of a sialylglycopeptide and free sialylglycans in hen's egg yolk. Biochim Biophys Acta - Gen Subj 1335:23–32

[3, 3]-Sigmatropic Rearrangement as a Powerful Synthetic Tool on Skeletal Modification of Unsaturated Sugars

**Gour Chand Daskhan, Malyasree Giri
and Narayanaswamy Jayaraman**

Abstract The development of practical, atom-economical and stereocontrolled routes to modify monosaccharides occupy a central importance in synthetic carbohydrate chemistry. This review provides an account on carbon–carbon and carbon–heteroatom bond formation as a skeleton modification on monosaccharides, through [3, 3]-sigmatropic rearrangement reactions on unsaturated sugar synthons, the generality of the reaction conditions and synthetic utilization of the resulting functionalized sugar building blocks. Major emphasis is laid on thermal rearrangement reactions, namely, Claisen, Ireland–Claisen, aza–Claisen and Johnson–Claisen rearrangements on carbohydrate-derived allyl vinyl ethers, silyl ketene acetals, allylic trichloroacetamidates and allylic orthoesters, respectively. These reactions offer a very promising prospect and permit a straightforward approach to access a large variety of biologically important, densely functionalized and novel carbohydrate mimetics. Further, examples of bioactive complex natural products, secured in high yields and profound stereoselectivity through such thermal rearrangement reactions on monosaccharides are described herein.

1 Introduction

The stereocontrolled carbon–carbon bond formation represents one of the most significant and challenging goals in organic synthesis. Over the past two decades, extensive efforts were made to develop newer synthetic methods for carbon–carbon bond formation, involving the major class of carbohydrates as synthons. Carbohydrate-derived synthons, in turn, permitted successful synthesis of structurally complex, diverse, highly oxygenated natural products [53, 63, 22, 23, 40] and important carbohydrate mimetics [47, 54, 37, 38, 15]. The classic Claisen rearrangement, pioneered by Claisen in 1912 [9] representing the general category of the [3, 3]-sigmatropic reaction, emerged as one of the most convenient

G.C. Daskhan · M. Giri · N. Jayaraman (✉)
Department of Organic Chemistry, Indian Institute of Science, Bangalore 560 012, India
e-mail: jayaraman@orgchem.iisc.ernet.in

stereoselective carbon–carbon bond-forming synthetic methods. The thermal Claisen rearrangement of allyl vinyl ether **1** is understood to proceed through a chair-like transition state giving rise to γ,δ-unsaturated carbonyl system **2**, in a stereocontrolled fashion (Scheme 1). Analogous rearrangement reaction involving all-carbon 1,5-diene synthon **3** transforming to **4** is known as the Cope rearrangement. Both these reactions find greater utility to derive a diverse range of sugar mimetics and branched sugars.

A number of modifications were introduced to expand the scope of the above rearrangement reactions, that include the method developed by Ireland and co-workers in 1972 [25, 26, 72] on the rearrangement of silyl ketene acetal **6** to derive δ,γ-unsaturated carboxylic acid. Base-promoted *O*-sialylation of allyl ester **5** is a facile modification to afford silyl ketene acetal **6**, which underwent smooth rearrangement under ambient condition to provide the δ,γ-unsaturated silyl ester **7** (Scheme 2). The rearrangement proceeds via a highly ordered transition state and the geometry of the resulting unsaturated silyl ester can be finely tuned depending on the solvents used. For example, ester **8**, upon treatment with hexamethylene phosphoramide/THF mixture afforded *Z*-silyl ketene acetal **9** preferably, whereas *E*-silyl ketene acetal **10** formed in the presence of THF alone. Implementation of the Claisen and Claisen–Ireland rearrangements in carbohydrate-derived allyl vinyl ethers and silyl ketene acetals is discussed in an early report by Werschkun and Thiem [80].

Aza–Claisen rearrangement, known as the Overman rearrangement [50, 51]; is a thermal [3, 3]-sigmatropic rearrangement of the allylic trichloroacetimidates for the stereoselective synthesis of nitrogen containing natural products and iminosugars. Thus, rearrangement of allylic acetimidate **11** obtained from allylic alcohol **10** upon reaction with NaH/Cl$_3$CCN, in xylene under reflux conditions, offer a valuable route to allyl amide **12** (Scheme 3). This rearrangement has also been utilized for the synthesis of various bioactive azasugars, as described later in this article.
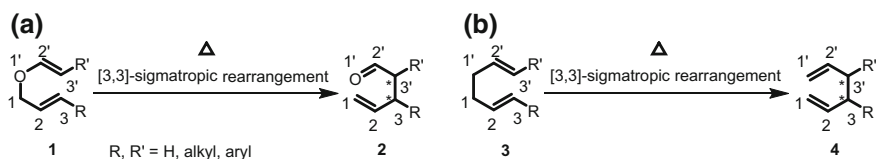
Scheme 1 **a** Claisen and **b** Cope rearrangement reactions
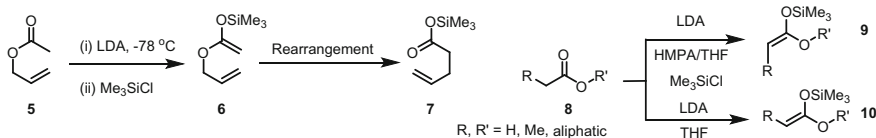
Scheme 2 The Ireland–Claisen rearrangement of silyl ketene acetal

Scheme 3 Aza–Claisen rearrangement of allylic trichloroacetimidate
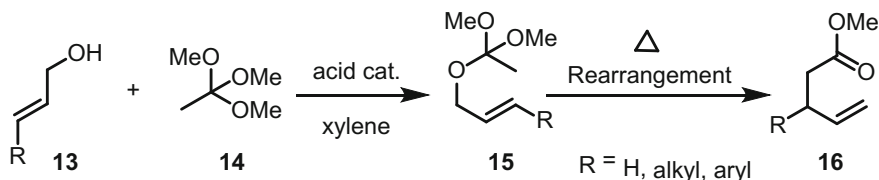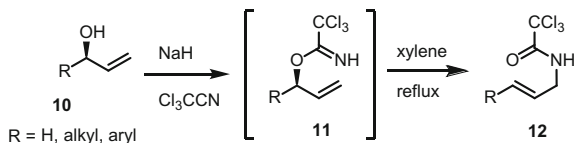
Scheme 4 Johnson–Claisen rearrangement of allylic orthoester

The Johnson–Claisen rearrangement [13, 25, 28, 83] on allylic orthoester **15**, derived from the reaction of allylic alcohol **13** with trimethyl orthoacetate **14** in the presence of catalytic propionic acid, afforded δ,γ-unsaturated carboxylic ester, with the olefin moiety in **16**, with exclusive *E*-geometry, in the case of substituted olefins (Scheme 4). This thermal rearrangement reaction was exploited for the construction of diverse bioactive azasugar analogues.

Application of the above general class of [3, 3]-sigmatropic rearrangement reactions to the synthesis of varied carbohydrate analogues and mimetics are discussed in the following sections.

2 Monosaccharide Modifications

The skeletal rearrangement reactions of various *endo-* and *exo-*cyclic 1,5-dienes led to the formation of diverse *C*-glycosides including 2-*C*-and 3-*C*-branched sugars in good yields and stereoselectivity.

2.1 Synthesis of C-Branched Carbohydrates

2.1.1 Synthesis of *C*-Glycosides

C-Glycoside mimetics [15, 37, 38, 61] are non-hydrolysable analogues of naturally occurring *O*-glycosides and play an important role in a wide range of biological events [75, 16, 69, 76]. *C*-Glycosides have emerged as an attractive synthetic target owing to their therapeutic applications on the one hand and their applicability to serve as chiral building blocks for general synthetic transformations on the other [1,

8, 52, 77]. Synthesis of such derivatives primarily relies on either Claisen [9] or Ireland-Claisen [25, 26, 72] rearrangement on sugar-derived allyl vinyl ethers or silyl ketene acetals, respectively [80].

Fairbank and co-workers [19] developed a method of sequential Tebbe methylation [71] in combination with Claisen rearrangement to synthesize diverse range of β- and α-C-glycosides. Tebbe reagent mediated methylation of glycal esters **17a–c** afforded derivatives **18a–c**, that subsequently underwent thermal rearrangement at 180 °C in benzonitrile or tributyl amine to afford β-C-glycosides **19a–c**, in 84–89% yields and high stereoselectivity. Under identical reaction condition or in the presence of various Lewis acid promoters including AlCl₃, BF₃.OEt₂, Yb (OTf)₃ and TiCl₄, compounds **21a–c**, prepared through Tebbe methylenation of **20**, afforded a mixture of α- and β-C-glycoside products. Synthesis of pure α-anomer over β-anomer was successfully accomplished by performing thermal reaction at 195 °C in either xylene or benzene solvent in a sealed tube. Indeed, α-C-glycosides **22a-c** were derived through rearrangement of benzylidene protected *allo*-derivatives **21a–c** at 195 °C in benzene or xylene in a sealed tube, in excellent yields (Scheme 5).

By employing a similar approach, stereoselective synthesis of α-C-glycosides from substituted glycal esters was accomplished [20]. However, the isomerization of enol ether derivatives prior to rearrangement led to a mixture of rearranged products. Thus, thermal rearrangement of palmitic enol ether derivative **23** at 195 °C in xylene afforded **24** and **25**, in a ratio of 1.3:1, in 85% yield. Product **25** was synthesized to form via a partial isomerization of enol ether **23** to more substituted tautomer, before the rearrangement reaction. However, treatment of derivative **26** in xylene as solvent under similar condition afforded only product **27a** (Scheme 6), in 85% yield. Whereas, using as benzene solvent led **27b** to form in almost equal amount as that of **27a**.
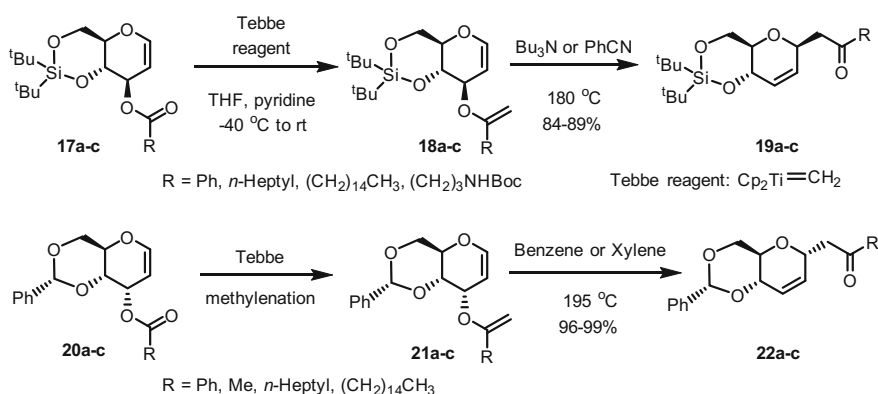
**Scheme 5** Synthesis of β-and α-C-glycosides
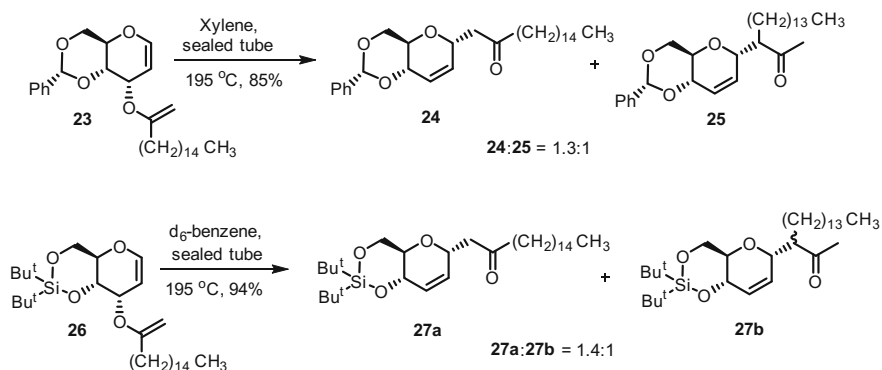
**Scheme 6** α-*C*-Glycosides prepared via tandem Tebbe/Claisen rearrangement

The Tebbe–Claisen rearrangement in amino acid-substituted vinyl ethers provided an easy access to β-*C*-glycosyl amino acid derivatives [7]. An illustration of this sequence is the methylenation of glycal esters **28a–d** to **29a–d**, followed by thermal rearrangement, leading to a variety of β-*C*-glycosyl amino acids **30a-d**, in 47–97% yields and in high stereoselectivity (Scheme 7). However, a similar methylenation of glycal esters of substituted α-amino acids was unsuccessful.

2-*C*-Branched carbohydrates are important subunit of many bioactive natural products, and antibiotics [34, 22, 23]. Among many methods [55, 59, 6, 39, 60, 62, 65, 73, 82], the stereoselective synthesis of 2-*C*-branched sugars primarily relies heavily on thermal rearrangement of glycal-derived allyl vinyl ethers.

A facile synthesis of 2-deoxy-2-*C*-alkyl glycal derivative was developed, which offers a novel route to prepare a series of 2-deoxy-2-*C*-branched sugars from 2-hydroxy glycal ester-derived 1,5-diene precursor [12]. The synthetic strategy consists of following key reactions: (i) *C*-allylation at *C*-1; (ii) Wittig methylenation at *C*-2 and (iii) implementation of the Cope rearrangement of 1,5-diene functionality. Treatment of vinyl ester **31** with allyl bromide in the presence of NaH/aq DMF afforded *O*-allylated derivative **32** and *C*-allylated derivative **33**, in a ratio of 9:1 as an inseparable mixture. Heating the crude reaction mixture at 165 °C in a sealed tube transformed *O*-allyl derivative **32** to *C*-allyl derivative **33**, in an overall yield of 63% starting from derivative **31**. The *C*-allyl derivative **33** remained to be an inseparable anomeric mixture, in a ratio of 55:45. Wittig methylenation of derivative **33** afforded *C*-2-methylene α-*C*-glycoside **34** and *C*-2-methylene β-*C*-glycoside **35**, the anomers could be separated through chromatography (Scheme 8).
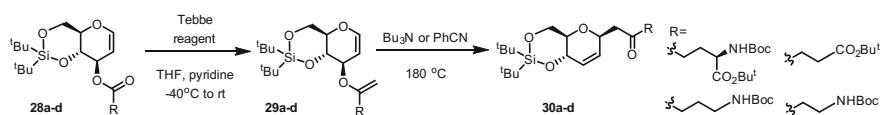
**Scheme 7** Synthesis of β-*C*-glycosyl amino acids
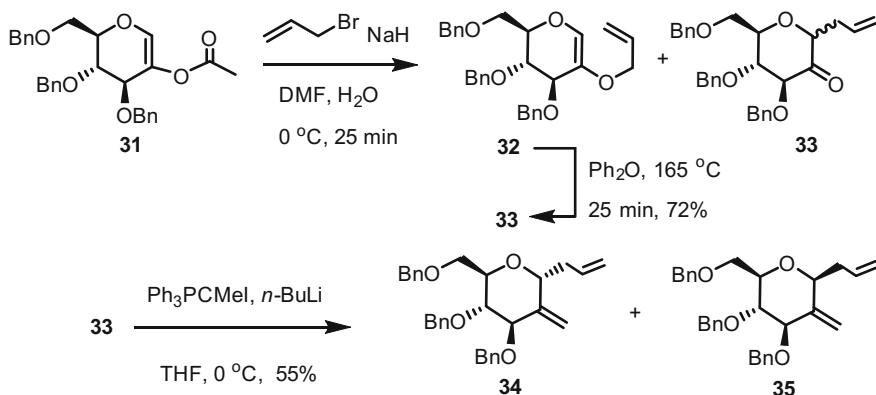
Scheme 8 Synthesis of 2-hydroxy glycal ester-derived 1,5-diene derivatives

When heated at 240 °C in diphenyl ether, the α-anomer **34** smoothly underwent rearrangement to afford 2-deoxy-2-*C*-alkyl glycal **36**, in 72% yield (Scheme 9). On the other hand, thermal rearrangement did not occur with the β-anomer **35**, illustrating the topological requirement for the rearrangement to be effective on a 1,5-diene synthon. Glycosyl donor properties of glycal derivative **36** were uncovered through further synthetic modifications, for which the terminal olefin was first subjected hydration selectively, with the aid of hydroboration-oxidation reactions using 9-BBN in THF and H_2O_2 (30%)/aq. NaOH (3 N) to afford **35** [12]. Terminal hydroxyl group in **37** was protected as an acetate **38** (Scheme 9).

The haloglycosylation strategy [11] was employed on derivative **38**, in the presence of appropriate aglycosyl and glycosyl acceptors, in the presence of NIS in CH_2Cl_2 at room temperature. In the event, haloglycosylation afforded 2-*C*-branched glycosides **39-47**, in good yields and with α-anomeric configuration (Scheme 10). Haloglycosylated disaccharide **43** and **44** were also derived using glycosyl acceptor alcohols under similar reaction conditions. In addition, an intramolecular iodoglycosylation with terminal olefin-derived glycal alcohol **37** afforded the oxepine **45**, with exclusive β-anomeric configuration. Subsequent deiodination reaction of glycosides **43** and **45**, using Bu₃SnH/AIBN

in benzene afforded 2-*C*-branched glycosides **46** and **47**, respectively, in good yields and stereoselectivity (Scheme 1).
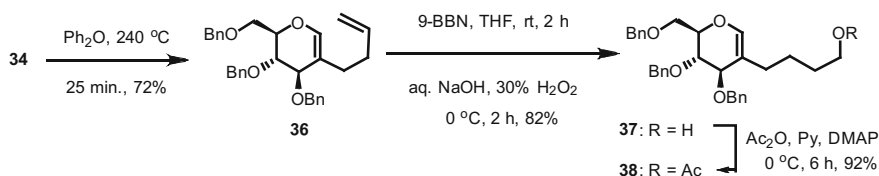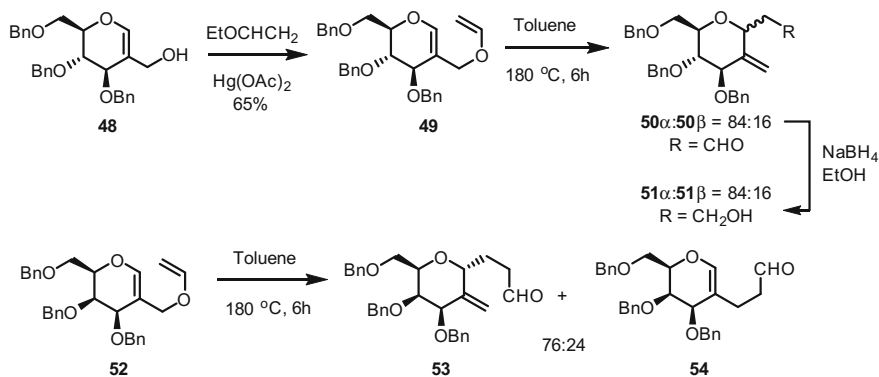

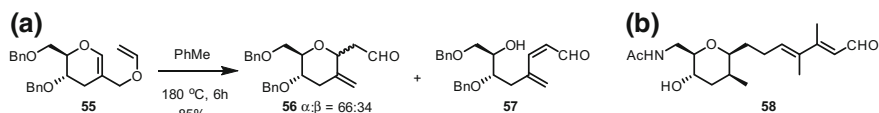**Scheme 9** Synthesis of 2-*C*-branched glycal through Cope rearrangement
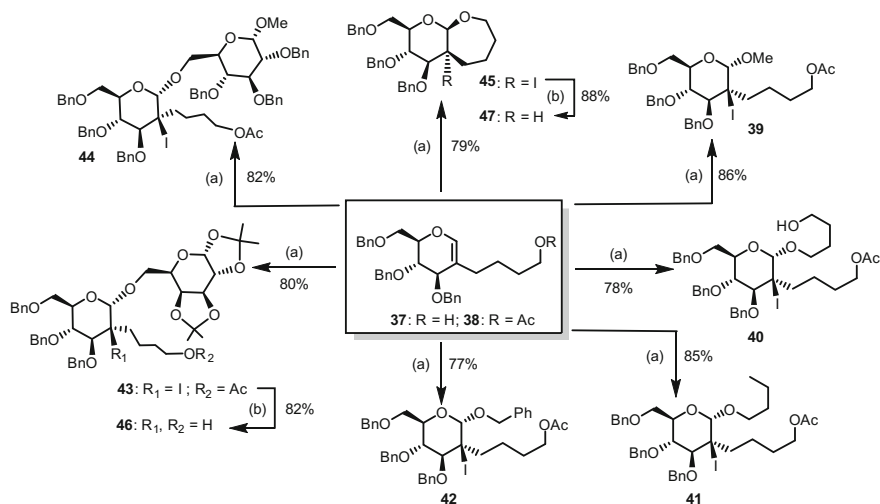
Scheme 10 Synthesis of a series of 2-deoxy-2-*C*-alkyl glycosides **39-47**. Reagents and conditions: **a** NIS, CH$_2$Cl$_2$, rt, appropriate aglycosyl and glycosyl acceptors; **b** Bu$_3$SnH, AIBN, C$_6$H$_6$, 75 °C, 2 h

2-Vinyloxymethyl glycal has also been employed to implement Claisen rearrangement so as to derive 2-*C*-branched carbohydrate derivatives [66]. For example, the thermal rearrangement of 2-vinyloxymethyl-3,4,6-tri-*O*-benzyl-D-glucal **49**, derived from *endo*-glycal alcohol **48**, afforded derivatives **50α** and **50β**, in 84:16 ratio, respectively. A NaBH$_4$ mediated reduction of aldehyde functionality afforded inseparable mixture of 2-*C*-methylene-β-*C*-glycoside **51α** and 2-*C*-methylene-β-*C*-glycoside **51β**. Interestingly, anomerization of α-*C*-glycoside to β-*C*-glycoside was also realized through ring opening, followed by intramolecular *oxa*-Michael-addition reaction, when the crude reaction mixture was purified through silica gel chromatography. In contrary, under similar reaction conditions, derivative **52** furnished **53**, as a single diastereomer along with unexpected product **54**, in 84:16 ratio, respectively (Scheme 11).

Encouraged by the above results, the opportunity to derive dideoxy-2-*C*-branched sugars from 2-vinyloxymethyl glycals was also investigated [68]. When heated at 180 °C in toluene in a sealed tube, derivative **55** underwent rearrangement to afford inseparable mixture of 3-deoxy-2-*C*-methylene-α/β-*C*-glycosides **56**, in 35% yield, along with ring-opened product **57**, in 50% yield (Scheme 12). Anomers were separated after reduction of the aldehyde group, using ethanolic NaBH$_4$ at −10 °C. Furthermore, Zn(OAc)$_2$-mediated anomerization of α-anomer to β-anomer was also utilized to prepare 2-*C*-branched-β-*C*-glycosides. Synthesis of a series of dideoxy-2-*C*-methylene-α- and β-*C*-glycosides was accomplished in good yields and diastereoselectivity. β-Anomer of **56** facilitated formal stereoselective synthesis of (-)-brevisamide **58** [36, 74] (Scheme 12).

Scheme 11 Synthesis of 2-*C*-branched glycosides via Claisen rearrangement

Scheme 12 **a** Synthesis of 3-deoxy-2-*C*-methylene-*C*-glycoside **56**; **b** Molecular structure of (-)-brevisamide **58**

Thermal Claisen rearrangement of 2,3-unsaturated vinyl glycosides was also exploited to derive 3-*C*-branched glycal derivatives [67]. A base-promoted thermal fragmentation of selenone glycoside **59** afforded allyl vinyl ether **60**, which underwent rearrangement at 160 °C in the presence of *N,N'*-dimethyl aniline/nitrobenzene to afford 3-*C*-branched aldehyde **61**, in 85% yield (Scheme 13). Similarly, allyl vinyl ethers derived from the corresponding acetate protected derivative of D-galactal, L-rhamnal, L- and D-arabinal were also transformed to 3-*C*-branched aldehyde derivatives, in good yields. The methodology was further exploited through the synthesis of a series of *cis*-fused perhydrofuro[2,3-b]furan and *cis*-fused perhydro-5-oxo-furo[2,3-b]furan derivatives from 3-*C*-branched sugar analogues. For example, perhydrofuro[2,3-b]furan **62** was derived from **60**, through ozonolysis of the olefin, followed by acid-catalysed acetalization. The synthetic
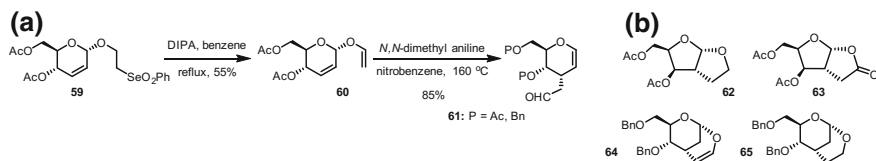
Scheme 13 **a** Synthesis of glucal-derived 3-*C*-branched sugars; **b** molecular structures of furan and nonane derivatives prepared from **61**

strategy was extended further to prepare 3-*C*-branched carboxylic acid derivatives, which were then transformed to *cis*-fused perhydro-5-oxo –furo[2,3-b]furan derivative **63**.

In contrast, 3-*C*-branched glycal building block **61** was successfully used for synthesis of 2,8-dioxabicyclo[3.3.1]nonene **64** and 2,8-dioxabicyclo[3.3.1]nonane **65**, in good yields and in a stereocontrolled fashion, upon (i) reduction of aldehyde moiety and (ii) TMSOTf-mediated cyclo-esterification reactions (Scheme 13b) [56, 57].

Stereocontrolled synthesis of natural products, such as, isoneosemburin **69** and neosemburin **70** were accomplished using [3, 3]-sigmatropic rearrangement as a key reaction. Dess–Martin periodinane oxidation of **66**, followed by Wittig homologation furnished allyl vinyl ether **67**, which upon heating at 160 °C in *N,N′*-dimethyl aniline and nitrobenzene, afforded desired 3-*C*-branched derivative **68**, in 62% yield (Scheme 14). Functional group reduction, followed by a series of transformations on **68** led to the synthesis of **69** and **70** (Scheme 14b) [56, 57].

2.2 Synthesis of Azasugars

Polyhydroxylated nitrogen containing cyclic and bicyclic frameworks are known as azasugars. Azasugars are important as glycosidase inhibitors [4, 10, 70] and serves as potential therapeutic agents [64, 79]. Thermal rearrangements were used in the facile synthesis of a variety of azasugars, for example, Johnson–Claisen orthoester rearrangement was employed for the synthesis of tetrahydroxy perhydroaza azulenes **73** and **74**, from D-glucose-derived epimeric allylic alcohols **71** [42, 43]. The epimeric allylic alcohols **71**, on treatment with trimethyl orthoacetate in the presence of catalytic propionic acid at 145 °C afforded derivative **72**, in 90% yield (Scheme 15), which was transformed to derivatives **73**–**76**, through a series of further synthetic manipulations [42, 43].

Aza–Claisen rearrangement on 2-(hydroxymethyl) glycal derivatives was implemented to prepare 2-*C*-methylene-*N*-glycosyl amides and their potential utilization for the construction of several azasugar analogues [21]. Thus, rearrangement of 3,4,6-tri-*O*-benzyl-2-*C*-(hydroxymethyl) galactal **77**, in the presence of NaH and trichloroacetonitrile at room temperature, afforded α- and β-*N*-glycosyl tricholoroacetimidates **79a** and **79b**, in 9:1 ratio, respectively (Scheme 16). Glucal derivative **78** also rearranged to glycosyl amides **80a** and **80b**, in a ratio of 3.7:1,
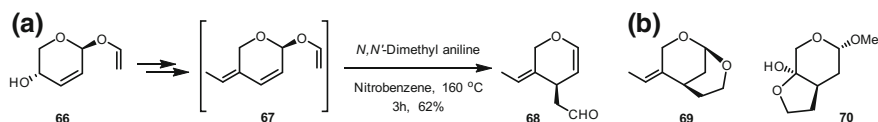
Scheme 14 a Stereoselective synthesis of 3-*C*-branched derivative **68**; **b** natural products **69** and **70** derived from **68**
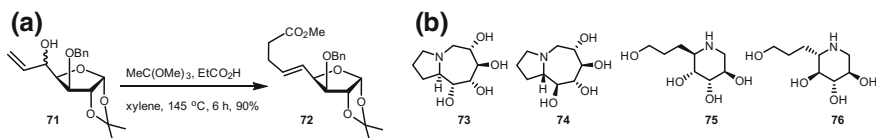
Scheme 15 Johnson–Claisen orthoester rearrangement of **71**–**72**; **b** Molecular structures of glycosidase inhibitors **73**–**76**, derived from **72**

respectively. Further synthetic manipulations on glycosyl amides **79a** and **79b** afforded L-*altro*-deoxynojirimycin **81**. Similarly, **80a** and **80b** were also transformed to L-*ido*-deoxynojirimycin **82** and L-*allo*-deoxynojirimycin **83** (Scheme 16). Hydroxymethyl modified compound **82** exhibited selective and moderate binding potency against α-galactosidase (IC$_{50}$ = 8.5 mM), in contrast to its parent analogue L-*ido*-deoxynojirimycin, which did not exhibit any inhibition.

More recently, the Aza–Claisen rearrangement was also employed to derive bioactive azasugar analogues from 2-*C*-hydroxymethyl glycals that were obtained from the corresponding D-arabinal, L-arabinal, and D-xylal precursors [58]. Thermal rearrangements of 3,4-di-*O*-benzyl-2-*C*-hydroxymethyl-D-arabinal, L-arabinal and 3,4-di-*O*-benzyl-2-*C*-hydroxymethyl-D-xylal derivatives provided inseparable mixture of α/β-*N*-glycosyl trichloroacetamidates **84a/b**, **85a/b** and **86a/b**, respectively, as described in Scheme 17. Glycosyl amides **84a/b** and **85a/b** were further transformed to isogalactofagomine **87** and *ent*-isogalactofagomine **88**, respectively. Additionally, synthesis of isofagomine **89** and 5-*epi*-isofagomine **90** were also prepared from glycosyl amides **86a/b**, through a series of synthetic transformations (Scheme 17).

2.3 Synthesis of Carbasugars

Carbasugars are hydrolytically stable unnatural carbocyclic sugar mimetics, form as important component of many native oligosaccharides, nucleosides and glycoside mimetics and posses promising antiviral, antitumor, and antibiotic activities [3]. Thermal or catalyst-promoted sugar-to-carbocycle rearrangements provided a
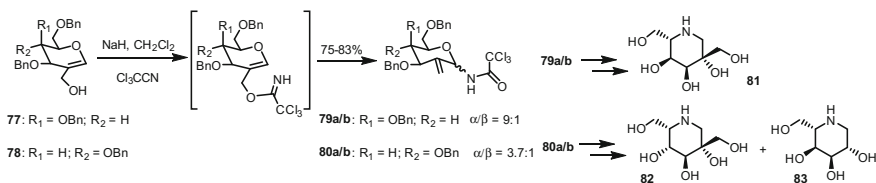
Scheme 16 Synthesis of 2-*C*-methylene-*N*-glycosyl amides and azasugars
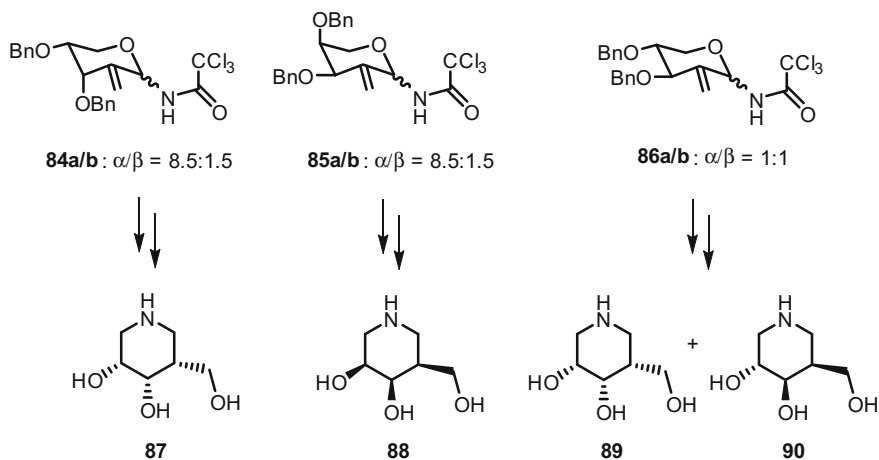
Scheme 17 Molecular structures of bioactive azasugars **87–90** derived from α/β-*N*-glycosyl trichloroacetamidates **84–86**

straightforward access to several six-, seven- and eight-membered carbasugar scaffolds.

Triisobutylaluminum (TIBAL)-promoted reductive [3, 3]-sigmatropic rearrangement was utilized by Sinaÿ and co-workers to prepare *gem*-difluoro carbasugar analogues from *gem*-difluoroalkyne precursors [14]. Treatment of alkyne **91** with dicobalt octacarbonyl, followed by a reductive TIBAL-mediated rearrangement of **92** in toluene afforded **93**, in a one-pot fashion, in 75% overall yield (Scheme 18). Ceric ammonium nitrate mediated decomplexation and oxidative cleavage of the alkyne moiety of **93** afforded *gem*-difluoro carba-α-D-glucopyranose **94**. Ti(O*i*Pr)Cl$_3$-promoted rearrangement of **92** at −78 °C in THF provided cyclohexanone derivative **95,** which upon a stereoselective reduction of ketone to alcohol **96** and a series of further synthetic manipulations afforded β-anomer **97** (Scheme 18).
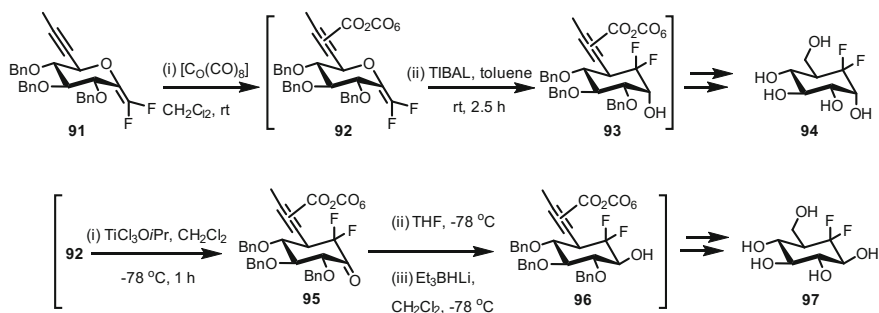
Scheme 18 Synthesis of *gem*-difluorocarba sugar

Stereoselective synthesis of carbasugar analogues was also accomplished through thermal Claisen rearrangement on glucal derived allyl vinyl ether, namely, 3,4-di-*O*-PMB-D-glucal derivative **98** [18]. Thermal treatment of derivative **98** at 250 °C afforded carbasugar **99**, in a good yield (Scheme 19). Subsequent ethanolic NaBH₄-mediated reduction of unstable aldehyde functionality to alcohol **100**, followed by a series of synthetic manipulations afforded enantiopure vinyl epoxides **101** and **102**. Pseudoglycosyl donor properties of epoxide **101** was realized through a glycosylation with 1,2;3,4-di-*O*-isopropylidine-α-D-galoctopyranose under acidic conditions, so as to form a stereocontrolled *O*-linked carba-disaccharide **103** (Scheme 19). These findings open-up the versatility of thermal rearrangement for the construction of a wide range of carbasugar skeletons and related *O*-linked carba-oligosaccharide analogues.

A reliable and straightforward route to synthesize Tamiflu, a potent inhibitor of viral neuraminidase, was devised from a carbohydrate precursor by using thermal Claisen rearrangement as a key reaction [41]. Derivative **104** under reflux condition underwent the key rearrangement reaction to afford cyclohexene derivative **105**, which was followed by a series of synthetic manipulations to afford compound **106** (Scheme 20). Binding studies against neuroendocrine PC12 cells showed a significant inhibition of vesicular excocytosis.
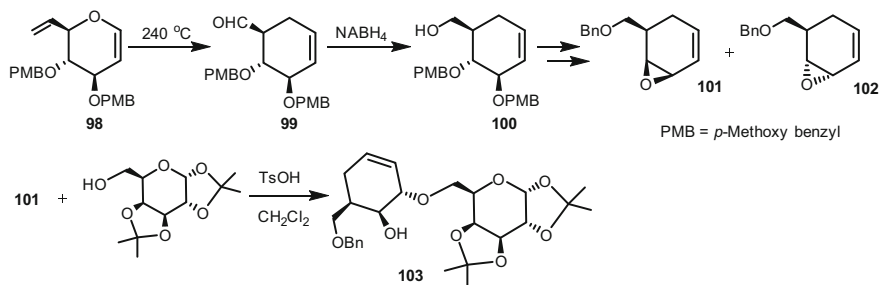
Scheme 19 Synthesis of carbasugar analogues
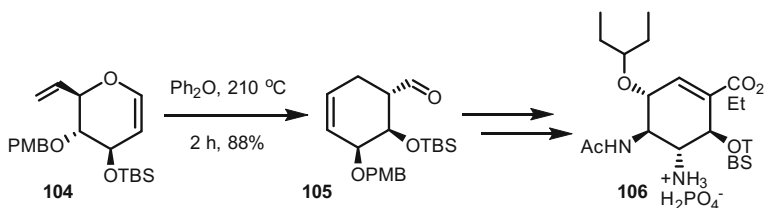
Scheme 20 Synthesis of key intermediate **105** via [3, 3]-sigmatropic rearrangement, towards synthesis of Tamiflu **106**

Scheme 21 Synthesis of cycloheptanic mimetics
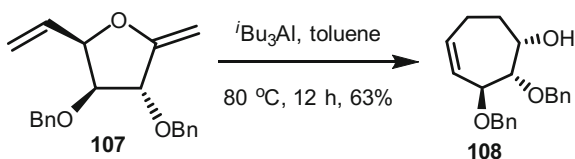
Triisobutylaluminium (TIBAL)-promoted thermal rearrangement protocol is also reliable to synthesize seven-membered carbasugars in a pronounced stereoselective fashion [27]. Thus, heating 3,4-*O*-di-benzyl-2-methylene-5-vinyl-tetrahydrofuran **107** at 80 °C in the presence of TIBAL in toluene afforded **108**, in 63% yield (Scheme 21).

Eight-membered carbocycles are novel class of unnatural cyclic sugar analogues. A facile synthesis of cyclooctanic derivatives **110** through thermal or triisobutylaluminium-promoted Claisen rearrangement of allyl vinyl ether, namely, 2-methylene-6-vinyltetrahydropyrans **109** was pioneered early on by Paquette and co-workers (Scheme 22) [33].

Thiem and co-workers developed the Claisen rearrangement either stand alone reaction or in combination with a glycosylation reaction to prepare various enantiopure cyclooctenyl glycosides [29]. Thermal rearrangement of D-glucose derived enol ether **111** at 185 °C in nitrobenzene afforded 5-cyclooctenone **112**, which upon reduction followed by a glycosylation with β-trichloroacetimidate **113** furnished α-glycoside **114**, in 51% yield and enantiopure form (Scheme 23). Further, Glcα1 → 4Man disaccharide-derived 1,5-diene precursor **115** underwent thermal

Scheme 22 Claisen rearrangement of allyl vinyl ether **109** to cyclooctanone derivative **110**
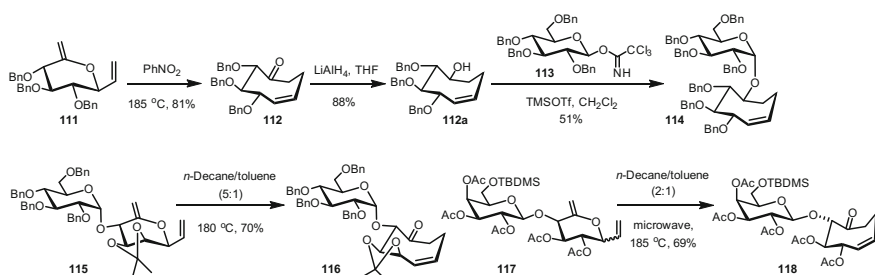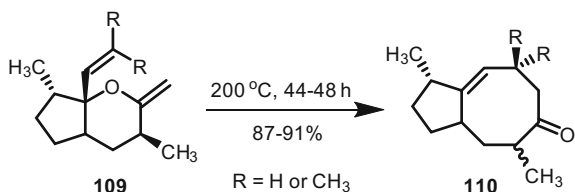
Scheme 23 Synthesis of cyclooctenyl-glycosylated scaffolds
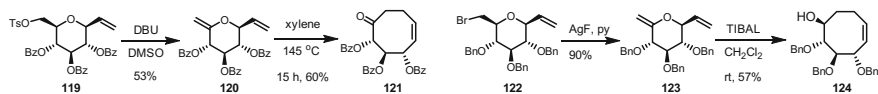
Scheme 24 Synthesis of eight-membered carbocycles **121** and **124**

rearrangement at 180 °C in *n*-decane/toluene (5:1) to provide glycosylated cyclooctenone derivative **116**, in 70% yield.

Synthesis of galtosylated cyclooctenone derivative using Claisen rearrangement protocol was also achieved following a similar protocol [32]. Thus, thermal rearrangement of 1,5-diene **117** at 180 °C in *n*-decane/toluene (2:1) under microwave irradiation led to the formation of glycosylated product **118**, in 69% yield (Scheme 23).

Unlike *endo*-glycals, the thermal rearrangement of *exo*-cyclic allyl vinyl ether primarily relies on altered reaction condition to afford desired rearranged products, implying that the thermal rearrangement depends entirely on the activation parameters employed. In this context, Thiem and co-workers studied both thermal and catalyst-induced thermal rearrangements of D-glucose and D-mannose-derived allyl vinyl ethers to prepare enantiopure and highly oxygenated eight-membered ring carbocycles [31]. When 2,3,4-tri-*O*-benzoyl derivative **120**, derived from tosylated derivative **119**, heated at 145 °C in xylene, compound **121** resulted in 60% yield (Scheme 24). Triisobutylaluminum-promoted rearrangement of benzylated analogue **123**, secured from tosylate **122** at rt in dichloromethane afforded rearrangement product **124**, in 57% yield (Scheme 24). Shorter reaction times and higher temperatures are found to be better for the thermal rearrangement reaction. X-ray crystallography and NMR spectroscopy studies revealed that the 5-cyclooctenone derivatives exhibited enhanced conformational flexibility, covering chair, boat and twist-chair conformations.

3   Synthesis of *C-C*-Linked Disaccharides

Hydrolytically stable *C*-linked disaccharides are particularly interesting carbohydrate analogues of the native *O*-glycosides [5, 17, 49]. Claisen rearrangement was regarded as a highly efficient and reliable synthetic route to access *C-C*-linked disaccharide analogues.

A Claisen rearrangement-based straightforward approach to prepare various hydrolytically stable *C-C*-linked disaccharide analogues [81] was developed by Thiem and co-workers. The methylene-bridged disaccharide analogues were prepared through ketene acetal [3, 3]-sigmatropic rearrangement as the key reaction. Simultaneous enolization and in situ silylation to the corresponding D-arabinose-derived ketene acetal **125** afforded rearranged epimeric disaccharides **126** and **127**, in a 1:1 diastereomeric ratio and in 15% overall yield (Scheme 25).
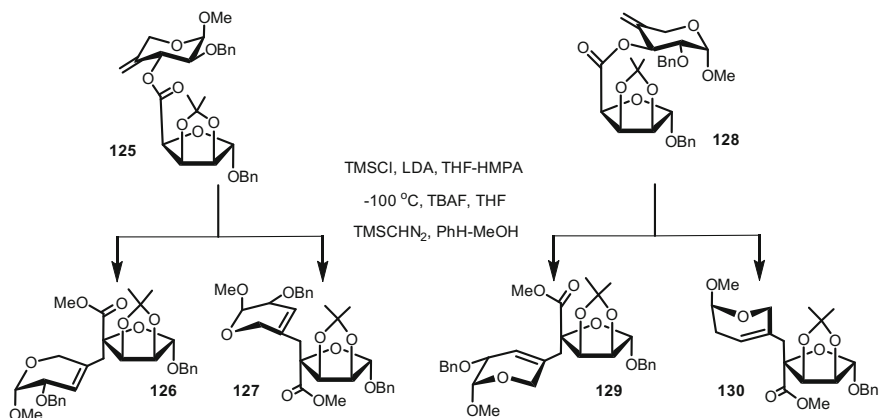
Scheme 25 Synthesis of *C-C*-linked disaccharide analogues

Similarly, rearrangement of L-arabinose-derived derivative **128** afforded epimeric products **129** and **130**, in an improved overall yield of 40%.

Allyl ketene acetal rearrangement in *endo*-cyclic olefin was exploited further as precursors to derive *C*-disaccharide derivatives. The strategy consists of thermal rearrangement of disaccharide allyl ketene acetal displaying both the double bonds of 1,5-diene system in *endo*-cyclic position [30]. The derivative **132** was obtained from glucal **131**, through few steps including thermal rearrangement at 130 °C. In the presence of *n*-butyl vinyl ether and diisopropylamine, derivative **132** afforded *C-C*-linked disaccharides **133** and **134**, with a diastereomeric ratio of 10:3, respectively (Scheme 26). A conformational analysis in solution by NMR spectroscopy revealed that compounds **133** and **134** possessed a boat ($B_{2,5}$) and half-chair ($^4H_3$) conformation, respectively.

In order to synthesize nonhydrolyzable ganglioside mimetics, Sodeoka and co-workers developed a stereoselective and highly efficient route involving Ireland-Claisen rearrangement to derive CF_2-linked ganglioside GM4 **137** [24]. Difluoromethylene-bridged sialosides are analogues to naturally occurring *O*-sialosides and possess interesting conformational flexibility and biological properties. Treatment of compound **135** in the presence of LHMDS and TMSCl at room temperature followed by reaction with TMSCHN$_2$ at −78 °C furnished desired derivative CF_2-linked α-(2 → 3) sialylgalactose **136**, in 86% yield (Scheme 27). Synthesis of difluoromethylene-linked ganglioside GM4 **137** was also achieved from sialoside **136**. Preliminary investigations showed that derivative **137** exhibited
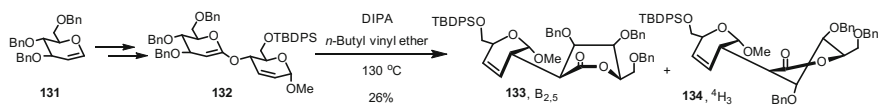
Scheme 26 Synthesis of C-C-linked disaccharides
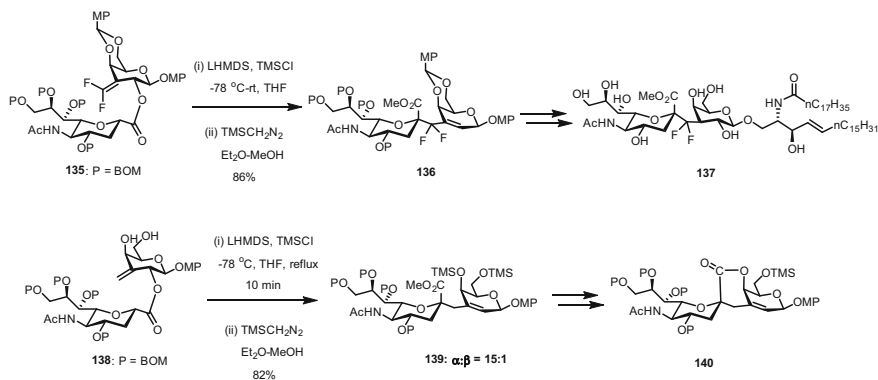
Scheme 27 Synthesis of ganglioside mimetics

similar biological properties as that of native GM4 and moderate inhibition against human neuraminidases, namely, NEU 2 (IC_{50} = 754 μM) and NEU 4 (IC_{50} = 930 μM).

In an effort to relate the chemical reactivity and conformational properties of CF_2-linked sialosides, a stereoselective synthesis of CH_2-linked α-(2 → 3)sialylgalactose lactone [78] was conducted. The investigation showed that benzylidene deprotected compound **138** underwent facile Ireland-Claisen rearrangement under reflux in THF, to afford the desired product **139**, in an α–β ratio of 15:1 (Scheme 27). This observation suggested that the rearrangement proceeded via chair-like transition state of the (Z)-silyl ketene acetal intermediate. Derivative **139** was subsequently transformed to CH_2-linked α-(2 → 3)sialylgalactose lactone **140**, through a series of synthetic manipulations.

4 Miscellaneous

The Claisen–Johnson rearrangement can also be utilized for the stereocontrolled synthesis of sugar amino acids [44]. Thus derivative **141** on treatment with triethyl orthoacetate in the presence of propanoic acid and hydroquinone under reflux afforded derivative **142**, as the only product, through suprafacial allyl rearrangement pathway (Scheme 28). Displacement of *tert*-butyldiphenylsilyl group with azide group provided sugar amino acid derivative **143**. Incorporation of building block **143** onto peptidomimetics provided the encephalin analogue **144**, after subsequent reactions.

Overman rearrangement has been utilized to derive 2-*C*-amino glycosides from hex-2,3-enopyranosides [35]. Allylic trichloroacetimidate **146**, derived from *C*-allyl glycoside **143**, underwent a facile rearrangement upon refluxing in 1,2-dichlorobenzene in the presence of K_2CO_3, to afford 2-*C*-amino allyl glycoside
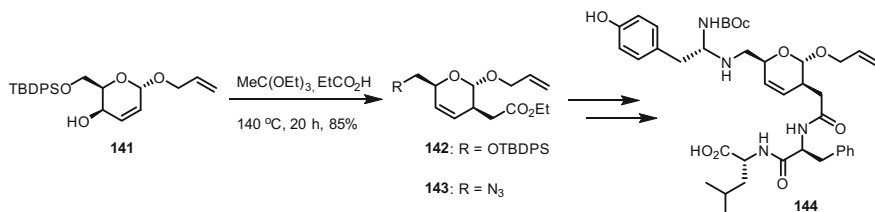
Scheme 28 Synthesis of key intermediate **143**, towards synthesis of encephalin analogue **144**

147, in 96% yield (Scheme 29). However, a similar rearrangement of epimeric *O*-allyl glycoside **149** led to the formation of desired 2-*C*-amino glycoside **150**, in 23% yield and allylic chloride **151** formed as a side-product, in 37% yield [45]. Interestingly, it has been observed that the product distribution can be tuned by using chromatographically pure allylic trichloroacetimidate. Thus, rearrangement of pure imidate **149**, derived from **148**, under similar reaction condition afforded expected *O*-allyl amide **150**, in 82% yield. Formation of only allylic chloride **151**, in 50% yield, was realized when the crude reaction mixture was subjected to rearrangement at 140–150 °C in xylene, in the presence of hydroquinone as a radical scavenger (Scheme 29) [46].

Thermal rearrangement on 2-*C*-*O*- and 3-*C*-*O*-allyl-L-ascorbic acid derivatives facilitated a convenient and stereocontrolled access to 2-*C*- and 3-*C*-branched aldono-1,4-lactone derivatives [48]. Thermal rearrangement of 5,6-*O*-isopropylidine-3-*O*-allyl-L-ascorbic acid **152** and 5,6-*O*-isopropylidine-2-*O*-allyl-L-ascorbic acid **154** in toluene under reflux condition afforded the corresponding 5,6-*O*-isopropylidine-3-allyl-2-keto-L-galactono-γ-lactone **153** and 5,6-*O*-isopropylidine-2-allyl-3-keto-L-galactono-γ-lactone **155**, respectively, in good yields (Scheme 30).

A facile [3, 3]-sigmatropic rearrangement was also utilized to construct subunit **160** of adriatoxin **156**, which is a marine ladder toxin natural product [2]. The derivative **160** was successfully derived through extended synthetic manipulations on 2-deoxy ribose **157**, which was used to prepare the advanced level intermediate **158**, in a multistep reaction sequence. Treatment of **158** with pyridinium *p*-toluenesulfonate/pyridine under reflux afforded intermediate **159**. Rearrangement
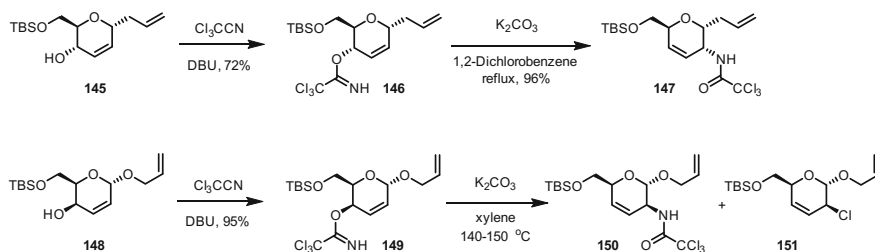
Scheme 29 Synthesis of 2-*C*-amino glycosides
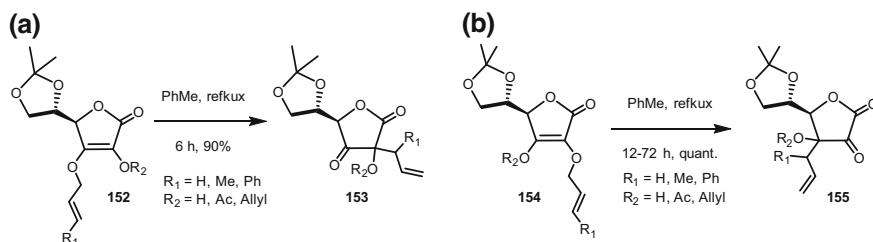
Scheme 30 **a** Rearrangement of 3-*O*-allyl derivative **152**; **b** 2-*O*-allyl derivative **154**, to 2-*C*- (**153**) and 3-C-branched (**155**) furanoses, respectively

reaction of **159**, followed by reduction of ketone moiety afforded EF ring **160**, in 94% yield (Scheme 31).

5 Conclusion and Perspective

Coupling–decoupling is an important strategy in the chemical synthesis of carbo-hydrate derivatives and mimetics, given the nature of inherent complexities of the molecular scaffolds. Rearrangement reactions, particularly [3, 3]-sigmatropic reactions, provide an important route towards stereoselective carbon–carbon bond formation. The foregoing discussion emphasizes the elegance of [3, 3]-sigmatropic rearrangements, in the form of Claisen, Ireland–Claisen, aza–Claisen and Johnson–Claisen routes, to modify the substituents around unsaturated carbohydrate frameworks, through either thermal or catalyst-promoted conditions. Especially, the implementation of these methods in *endo*- and *exo*-glycal derived unsaturated sugar synthons allow facile access to numerous valuable synthetic targets and interme-diates, which are used gainfully for further synthetic modifications. This article covers the advancement during the past over a decade on the rearrangements as expeditious tools for the stereoselective and straightforward access to large panel of densely functionalized *C*-glycosides, *C*-branched sugars, bioactive natural products, azasugars, carbasugar analogues, *C*-linked disaccharides and other important sugar
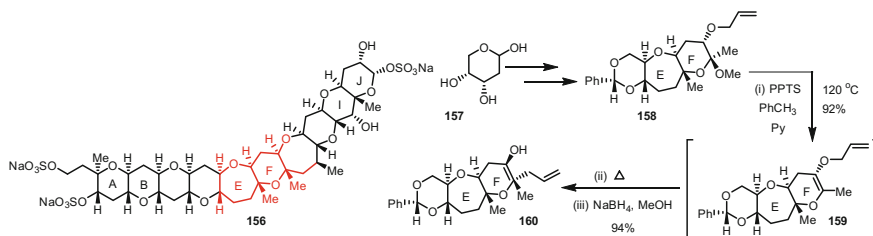
Scheme 31 Claisen rearrangement route to EF ring of adriatoxin

mimetics. Reliability and efficiency coupled with topological control of [3, 3]-sigmatropic rearrangements shall remain to ensure the sought-after nature of these reactions in C–C and C–heteroatom bond formation in monosaccharides and their derivatives.

References

1. Abe H, Shuto S, Matsuda A (2001) Highly α- and β-selective radical C-glycosylation reactions using a controlling anomeric effect based on the conformational restriction strategy. A study on the conformation-anomeric effect stereoselectivity relationship in anomeric radical reactions. J Am Chem Soc 123:11870–11882
2. Akoto CO, Rainier JD (2008) Harnessing glycal-epoxide rearrangements: the generation of the AB, EF, and IJ rings of Adriatoxin. Angew Chem Int Ed 47:8055–8058
3. Arjona O, Gómez AM, López C, Plumet J (2007) Synthesis and conformational and biological aspects of carbasugars. Chem Rev 107:1919–2036
4. Asano N, Nash RJ, Molyneux RJ, Fleet GWJ (2000) Sugar-mimic glycosidase inhibitors: natural occurrence, biological activity and prospects for therapeutic application. Tetrahedron Asymmetry 11:1645–1680
5. Babirad SA, Wang Y, Kishi Y (1987) Synthesis of C-disaccharides. J Org Chem 52:1370–1372
6. Beyer J, Skaanderup PR, Madsen RJ (2000) Platinum-catalyzed ring opening of 1,2-cyclopropanated sugars with O-nucleophiles. Convenient synthesis of 2-C-branched carbohydrates. J Am Chem Soc 122:9575–9583
7. Chambers DJ, Evans GR, Fairbanks AJ (2005) Synthesis of C-glycosyl amino acids: scope and limitations of the tandem Tebbe/Claisen approach. Tetrahedron Asymmetry 16:45–55
8. Chiara JL, Sesmilo E (2002) Samarium diiodide-mediated reductive coupling of epoxides and carbonyl compounds: a stereocontrolled synthesis of C-glycosides from 1,2-Anhydro sugars. Angew Chem Int Ed Engl 41:3242–3246
9. Claisen L (1912) Ber Dtsch Chem Ges 45:3157–3166
10. Compain PE (2007) Iminosugars from Synthesis to Therapeutic Applications. Wiley VCH Weinheim, Germany
11. Danishefsky SJ, Bilodeau MT (1996) Glycals in organic synthesis: the evolution of comprehensive strategies for the assembly of oligosaccharides and glycoconjugates of biological consequence. Angew Chem Int Ed Engl 35:1380–1419
12. Daskhan GC, Jayaraman N (2012) Synthesis of 2-Deoxy-2-C-alkyl glycal and glycopyranosides from 2-Hydroxy glycal ester. J Org Chem 77:2185–2191
13. Daub GW, Teramura DH, Bryant KE, Burch T (1981) Ortho ester Claisen rearrangements using trimethyl methoxyorthoacetate. J Org Chem 46:1485–1486
14. Deleuze A, Menozzi C, Sollogoub M, Sinaÿ P (2004) Synthesis of gem-Difluorocarba-D-glucose: a step further in sugar mimetics. Angew Chem Int Ed 43:6680–6683
15. Du Y, Linhardt RJ, Vlahov IR (1998) Recent advances in stereoselective C-glycoside synthesis. Tetrahedron 54:9913–9959
16. Dwek RA (1996) Glycobiology: toward understanding the function of sugars. Chem Rev 96:683–720

17. Ferritto R, Vogel P (1994) Synthesis of a-D-(1 + 3) and a-D-(1 + 4)-C-linked Galactosides of D-Mannose Derivatives. Conformation of a-C-Ga1actosides. Tetrahedron Asymmetry 5:2077–2092

18. Frau I, Bussolo VD, Favero L, Pineschi M, Crotti P (2011) Stereodivergent synthesis of diastereoisomeric carba analogs of glycal-derived vinyl epoxides: a new access to carbasugars. Chirality 23:820–826

19. Godage HY, Chambers DJ, Evans GR, Fairbanks AJ (2003) Stereoselective synthesis of C-glycosides from carboxylic acids: the tandem Tebbe-Claisen approach. Org Biomol Chem 1:3772–3786

20. Godage HY, Fairbanks AJ (2003) Synthesis of α-C-glycosides via tandem Tebbe methylenation and Claisen rearrangement. Tetrahedron Lett 44:3631–3635

21. Gupta P, Vankar YD (2009) Facile aza-Claisen rearrangement of glycals: application in the synthesis of 1-Deoxy-L-iminosugars. Eur J Org Chem 1925–1933

22. Hanessian S (1993) Reflection on the total synthesis of natural products: art, craft, logic, and the chiron approach. Pure & Appl Chem 65:1189–1204

23. Hanessian S (1983) Total synthesis of natural products: The "Chiron" approach. Pergamon Press, Oxford, pp 40–183

24. Hirai G, Watanable T, Yamaguchi K, Miyagi T, Sodeoka M (2007) Stereocontrolled and convergent entry to CF_2-Sialosides: synthesis of CF_2-linked ganglioside GM4. J Am Chem Soc 129:15420–15421

25. Ireland RE, Mueller RH (1972) The Claisen rearrangement of allyl esters. J Am Chem Soc 94:5897–5898

26. Ireland RE, Mueller RH, Willard AK (1976) The ester enolate Claisen rearrangement. stereochemical control through Stereoselective enolate formation. J Am Chem Soc 98:2868–2877

27. Jia C, Zhang Y, Zhang L (2003) Highly stereoselective synthesis of a seven-membered carbasugar via triisobutylaluminium promoted Claisen rearrangement. Tetrahedron Asymmetry 14:2195–2199

28. Johnson WS, Werthemann I, Bartlett WR, Brockson TJ, Li T, Fulkner DJ, Petersen (1970) A Simple stereoselective version of the Claisen rearrangement leading to trans-trisubstituted olefinic bonds. Synthesis of squalene. J Am Chem Soc 92:741–743

29. Jürs S, Thiem J (2005) Alternative approaches towards glycosylated eight-membered ring compounds employing Claisen rearrangement of mono and disaccharide allyl vinyl ether precursors. Tetrahedron Asymmetry 16:1631–1638

30. Jürs S, Thiem J (2006) Novel approach towards C–C-linked carbohydrate dimers via Claisen rearrangement of a disaccharide allyl ketene acetal. Synthesis 2117–2120

31. Jürs S, Werschkum B, Thiem (2006) Claisen rearrangement of carbohydrate-derived precursors towards highly functionalized cyclooctenones with L-xylo, D-arabino and L-lyxo configuration and their diastereoselective transformations. Eur J Org Chem 4451–4462

32. Jürs S, Thiem J (2007) From lactose towards a novel galactosylated cyclooctenone. Carbohydr Res 342:1238–1243

33. Kinney WA, Coghlan MJ, Paquette LA (1985) General approach to annulated 4-Cyclooctenones by aliphatic claisen rearrangement. Stereospecific total synthesis of (f)-Precapnelladiene. J Am Chem Soc 107:7352–7360

34. Kirschning A, Bechthold AFW, Rohr J (1997) Top Curr Chem 188:1–84

35. Koester DC, Holkenbrink A, Werz DB, (2010) Recent advances in the synthesis of carbohydrate mimetics Synthesis 19:3217–3242

36. Kriek NMAJ, van der Hout E, Kelly P, van Meijgaarden KE, Geluk A, Ottenhoff THM, van der Marel GA, Overhand M, van Boom JH, Valentijn ARPM (2003) Synthesis of novel tetrahydropyran-based dipeptide isosters by Overman rearrangement of 2,3-Didehydroglycosides. Eur J Org Chem 2418–2427

37. Levy D, Tang C (1995) The chemistry of C-glycosides. Pergamon, Oxford

38. Levy D, Tang EC (1995) The chemistry of C-glycosides tetrahedron organic chemistry series, vol 13. Pergamon Press, Oxford

39. Linker T, Sommermann Kahlenberg TF (1997) The addition of malonates to glycals: a general and convenient method for the synthesis of 2-C-branched carbohydrates. J Am Chem Soc 119:9377–9384

40. Lundt I (1997) Topics Curr Chem 187:117–153

41. Ma J, Zaho Y, Ng S, Zhang J, Zeng J, Than A, Chen P, Liu XW (2010) Sugar-based synthesis of tamiflu and its inhibitory effects on cell secretion. Chem Eur J 16:4533–4540

42. Markad SD, Karanjule NS, Sharma T, Sabharwal SG, Puranik VG, Dhavale DD (2006) Synthesis of tetrahydroxy perhydroaza-azulenes: tandem Johnson-Claisen rearrangement of D-glucose-derived allylic alcohols. Org Biomol Chem 4:2549–2555

43. Markad SD, Karanjule NS, Sharma T, Sabharwal SG, Puranik VG, Dhavale DD (2006) polyhydroxylated homoazepanes and 1-deoxy-homonojirimycin analogues: synthesis and glycosidase. Org Biomol Chem 4:3675–3680

44. Montero A, Mann E, Herradón B (2004) Preparation of sugar amino acids by Claisen-Johnson rearrangement: synthesis and incorporation into Enkephalin analogues. Eur J Org Chem 3063–3073

45. Montero A, Mann E, Herradón B (2005) The Overman rearrangement in carbohydrate chemistry: stereoselective synthesis of functionalized 3-amino-3,6-dihydro-2H-pyrans and incorporation in peptide derivatives. Tetrahedron Lett 46:401–405

46. Montero A, Benito E, Herradón B (2010) Synthesis and applications of a chiral-oxygenated 3-chloro-3,6-dihydro-2 Hpyran obtained under Overman rearrangement conditions. Tetrahedron Lett 51:277–280

47. Nicotra F (1997) Synthesis of C-glycosides of biological interest. Topics Curr Chem 187:55–83

48. Olabisi AO, Mahindaratne MPD, Wimalasena K (2005) A convenient entry to C2- and C3-substituted Gulono-γ-lactone derivatives from L-Ascorbic acid. J Org Chem 70:6782–6789

49. O'Leary DJ, Kishi Y (1994) C-Sucrose vs. O-Sucrose: different conformational behavior in methanol solutions containing Ca^{2+}. Tetrahedron Lett 35:5591–5594

50. Overman LE (1974) Thermal and mercuric Ion catalyzed [3,3]-sigmatropic rearrangement of allylic trichloroacetimidates. The 1, 3 transposition of alcohol and amine functions. J Am Chem Soc 96:597–598

51. Overman LE, Carpenter E (2005) Organic reactions 66:1–107 (Wiley-VCH Weinheim)

52. Paterson DE, Griffin FK, Alcaraz ML, Taylor RJK (2002) A Ramberg-Bäcklund approach to the synthesis of C-glycosides, C-linked disaccharides, and C-glycosyl amino acids Eur J Org Chem 1323–1336

53. Paterson L, Keown LE (1997) Studies in marine macrolide synthesis: stereocontrolled synthesis of the F-ring subunit of Spongistatin 1 (Altohyrtin A). Tetrahedron Lett 38:5727–5730

54. Postema MHD (1992) Recent developments in the synthesis of C-glycosidase. Tetrahedron 48:8545–8599

55. Ramesh NG, Balasubramanian KK (1991) Vilsmeier-Haack reaction of glycals-a short route to C-2-formyl glycals. Tetrahedron Lett 32:3875–3878

56. Reddy GM, Sridhar PR (2014) Stereoselective construction of 2,8-Dioxabicyclo[3.3.1] nonane/nonene systems from 3- C-branched glycals. Eur J Org Chem 1496–1504

57. Reddy GM, Sridhar PR (2014) The first stereoselective total synthesis of neosemburin and isoneosemburin. Org Biomol Chem 12:8408–8414

58. Reddy YS, Kancharla PK, Roy R, Vankar YD (2012) Aza-Claisen rearrangement of 2-C-hydroxymethyl glycals as a versatile strategy towards synthesis of isofagomine and related biologically important azasugars. Org Biomol Chem 10:2760–2773

59. Schmidt R, Kast RJ (1986) Direct lithiation of glycals. Synthesis of C-2 branched sugars. Tetrahedron Lett 27:4007–4010

60. Scott RW, Heathcock CH (1996) An efficient synthesis of 3,4,6-tri-O-benzyl-2-C-methyl-D-glucal. Carbohydr Res 291:205–208

61. Sears P, Wong CH (1999) Carbohydrate Mimetics: a new strategy for tackling the problem of carbohydrate-mediated biological recognition. Angew Chem Int Ed 38:2300–2324

62. Shao H, Ekthawatchai S, Wu SH, Zou W (2004) 2-C-branched glycosides from 2′-Carbonylalkyl 2-O-Ms(Ts)-C-glycosides. A tandem $S_N2 - S_N2$ reaction via 1,2-Cyclopropanated sugars. Org Lett 6:3497–3499

63. Smith AB III, Zhuang L, Brook CS, Boldi AM, McBriar MD (1997) Spongistatin synthetic studies. 1. Construction of A C(29-48) subtarget. Tetrahedron Lett 38:8667–8670

64. Somsak L, Nagya V, Hadady Z, Dosca T, Gergely P (2003) Curr Pharm Des 9:117–1189

65. Sridhar PR, Ashalu KC, Chandrasekaran S (2004) Efficient methodology for the synthesis of 2-C-branched glyco-amino acids by ring opening of 1,2-Cyclopropanecarboxylated sugars. Org Lett 6:1777–1779

66. Sridhar PR, Sudharani C (2012) Stereoselective synthesis of C-2-methylene and C-2-methyl-C-glycosides by Claisen rearrangement of 2-vinyloxymethyl glycals. RSC Adv 2:8596–8598

67. Sridhar PR, Reddy GM, Seshadri K (2012) Stereoselective synthesis of cis-fused perhydrofuro [2,3-b]furan derivatives from sugar-derived allyl vinyl ethers. Eur J Org Chem 6228–6235

68. Sudharani C, Venukumar P, Sridhar PR (2014) Stereoselective synthesis of C-2-methylene and C-2-methyl α-and β-C-glycosides from 2-C-branched glycals: formal total synthesis of (–)-brevisamide. Eur J Org Chem 8085–8093

69. Stanley P, Bertozzi CR, Hart GW, Etzler ME (2009) Cold Spring Harbor, NY, 2nd edn, ch 6:75–88

70. Stutz AE (1999) Iminosugars as glycosidase inhibitors nojirimycin and beyond. Wiley VCH Weinheim, Germany

71. Tebbe FN, Parshall GW, Reddy GS (1978) Olefin homologation with titanium methylene compounds. J Am Chem Soc 100:3611–3613

72. Thiem J (2005) Alternative approaches towards glycosylated eight-membered ring compounds employing Claisen rearrangement of mono and disaccharide allyl vinyl ether precursors. Tetrahedron Asymmetry 16:569–576

73. Tian Q, Xu L, Ma X, Zou W, Shao H (2010) Stereoselective synthesis of 2-C-Acetonyl-2-Deoxy-D-Galactosides using 1,2-Cyclopropaneacetylated sugar as novel glycosyl donor. Org Lett 12:540–543

74. Tsutsumi R, Kuranaga T, Wright JLC, Baden DG, Ito E, Satake M, Tachibana K (2010) An improved synthesis of (-)-brevisamide, a marine monocyclic ether amide of dinoflagellate origin. Tetrahedron 66:6775–6782

75. Varki A (1993) Biological roles of oligosaccharides: all of the theories are correct. Glycobiology 3:97–130

76. Varki A, Lowe JB (2009) Biological roles of glycans. In: Varki A et al. (eds) Essentials of glycobiology. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, pp 75–88

77. Wang Q, Linhardt R (2003) Synthesis of a serine-based neuraminic acid C-glycoside. J Org Chem 68:2668–2672

78. Watanable T, Hirai G, Kato M, Hashizume D, Miyagi T, Sodeoka M (2008) Synthesis of CH_2-Linked α(2,3)Sialylgalactose analogue: on the stereoselectivity of the Key Ireland-Claisen Rearrangement. Org Lett 10:4167–4170

79. Weiss M, Hettmer S, Smith P, Ladish S (2003) Inhibition of melanoma tumor growth by a novel inhibitor of glucosylceramide synthase. Cancer Res 63:3654–3658

80. Werschkun B, Thiem J (2001) Topics Curr Chem 215:293–325

81. Werschkum B, Thiem J (2005) Synthesis of novel types of divalent saccharide structures by a ketene acetal Claisen rearrangement. Tetrahedron Asymmetry 16:569–579

82. Wong G, Fraser-Reid B (1994) Some routes to C2-alkoxymethyl hex-2-enopyranoside. Can J Chem 72:69–74

83. Ziegler FE (1977) Stereo- and regiochemistry of the Claisen rearrangement: applications to natural products synthesis. Acc Chem Res 10:227–232

Recognition of Thiols in Coupling Reactions to Organic and Carbohydrate Acceptors

Zbigniew J. Witczak and Roman Bielski

Abstract Reactions of thiols in thio-click coupling processes with various reactive systems (including carbohydrates) are compiled. A selection of simple and complex thiols in stereoselective and non-stereoselective approaches recognizing their reactivity is also reviewed. Solvents, employed in the discussed processes including water, are briefly discussed as well.

1   Introduction

The coupling reaction forming C–S, C–C, or C–N bonds usually requires an activation of the existing functional group. To control the steric outcome of the coupling reaction, the reactive functionalities forming the new asymmetric center should be easily accessible and kinetically favor the formation of only one stereoisomer.

The addition of thiols to conjugated or nonconjugated multiple bonds belongs to a few processes applicable to our coupling and decoupling (CAD) methodology [1]. The practical utility of our CAD methodology to many strategic targets is outstanding. The strategy can be easily adopted for many complex reactive systems provided that the activation step is strictly followed and the intermediate adducts will be actively involved and be compatible with a specific protocol of activation.

The click reactions involving the addition of thiols have been reviewed. Thiol-ene click chemistry, particularly applicable to polymer chemistry, was reviewed by Hoyle and Bowman [2]. The Bowman team [3] also reviewed the thio-Michael addition click reaction as another powerful tool in material chemistry.

Our laboratory reviewed [4] the synthesis of carbohydrate thiols as universal coupling agents applicable in our CAD [1] methodology. The synthetic procedures explored the versatility, stereochemical outcome of thio-click coupling reactions,

Z.J. Witczak (✉) · R. Bielski
Department of Pharmaceutical Sciences, Nesbitt School of Pharmacy,
Wilkes University, 84 W. S. Street, Wilkes-Barre, PA 18766, USA
e-mail: zbigniew.witczak@wilkes.edu

and synthetic aspects of thiols as universal starting materials. Additionally, the thiol activation under alkaline pH (8–9.5), polarity of the reaction solvent, and thiol/acceptor ratio were discussed.

Two other review articles [5, 6] explore the sensitivity and specificity of organic thiols recognition and detection in biological systems.

2 Thiol Recognition in Synthetic Approaches in Thio-Click Addition Reactions

The thiol recognition in biological systems is a primary factor of efficacy of many biologically important molecules containing –SH functionality. A rapidly developing area of synthetic organic and carbohydrate chemistry is exploring many specific tools for biological ligation of natural thiols and their peptide and protein systems.

It also utilizes a variety of other tools such as metalo-organic catalysts, effects of polar solvents, and enzymatic systems to form stereoselectively non-hydrolyzable C–S bonds. These sulfur bonds often are resistant to multiple enzymatic systems widely present in living organisms. Consequently, the sulfur-linked derivatives may escape any enzymatic intervention by scavenging the –SH group.

As already mentioned, Yoon and co-workers reviewed [6] fluorescent and colorimetric probes to detect three important thiols present in living organisms— cysteine, (Cys) homocysteine (Hcy), and glutathione (GSH). Their similar chemical character and structural composition derive from the presence of three reactive functional groups, –SH, –COOH, and –NH$_2$ capable of forming specifically labeled molecules.

All these three biomolecules (Fig. 1) equipped with the mercapto group play a crucial role in maintaining functionality of biological systems. Their low cellular levels are linked or implicated in many diseases. Therefore, the development of fluorescent and colorimetric probes for their detection is of utmost importance. Shiu and co-workers reported [7] a highly selective FRET-based fluorescent probe to detect cysteine (Cys) and homocysteine (Hcy).

Huo and co-workers reported [8] the chemistry of the functionalized chromene moiety as a "lock", the thiol as a "key", and a mercury (II) ion as a "hand", a single
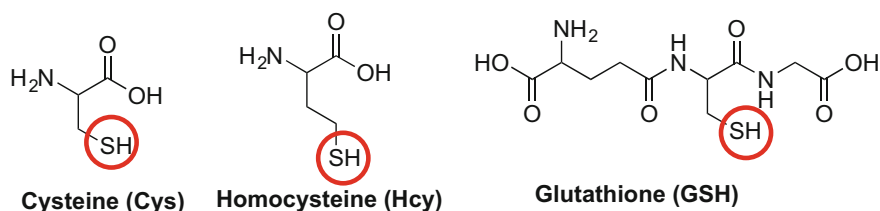
Fig. 1 Natural Thiols; Cysteine (Cys), Homocysteine (Hcy) and Glutathione (GSH)
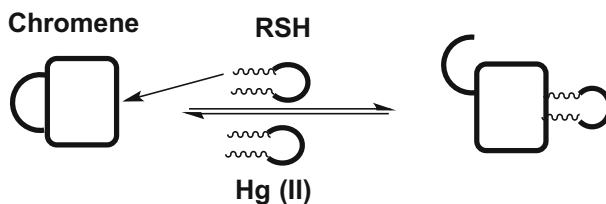
Scheme 1 A single molecular machine thiol recognition system

molecular automated recognition system. The simplified aspect of the thiols recognition mechanism is shown in Scheme 1.

Clayden and MacLellan [9] reported the asymmetric synthesis of specifically designated tertiary thiols and their functionalized thioethers. Selected strategic coupling approaches are shown in Scheme 2.

Whereas tertiary thiols are important synthetic templates, their dominant reactive character must be recognized during coupling reactions, including thiol-yne. In the presence of metal catalysts, their reactivity increases and the reaction time is significantly shortened as, compared to catalyst-free methodologies [10]. Among many metal catalysts, the following primary catalysts were employed for construction of C–S bonds during synthesis of *S*-thioglycosides: nano indium oxide [11], iridium complex (Ir(COD)$_2$BF$_4$), [12], and palladium diacetate [Pd(OAc)$_2$] phosphine ligand system [13, 14]. The highly efficient palladium diacetate catalyzed synthesis of thioglycosides is depicted in Scheme 3.

Misra and co-workers [15] reported a green chemistry approach toward synthesis of 3-thio-2-deoxy and 3-dithiocarbamate sugar derivatives. This efficient methodology uses no catalyst and is performed in water as a polar solvent. The products high yields and purities are impressive as compared to other methods. Examples are shown in Scheme 4.
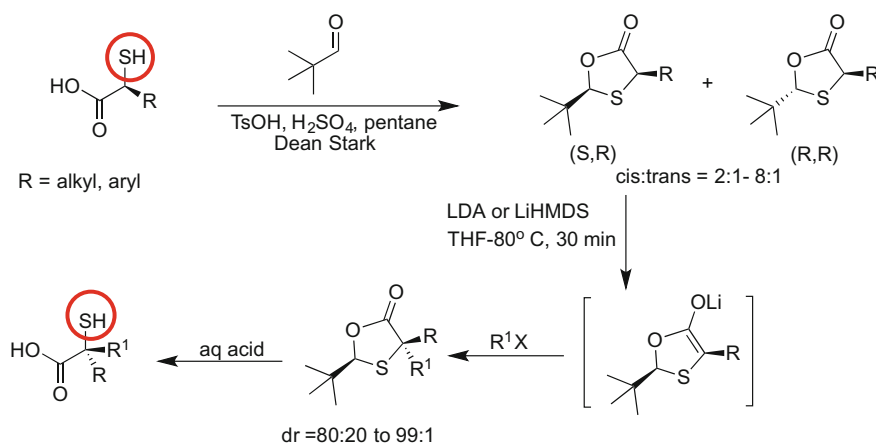
Scheme 2 Asymmetric synthesis of tertiary thiols
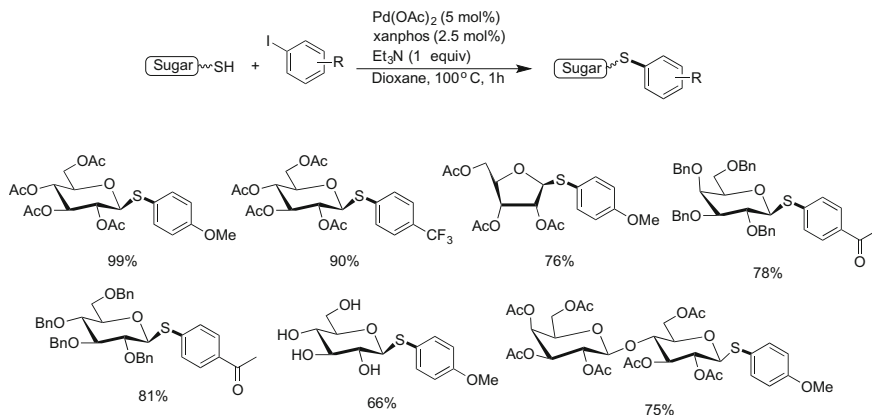
Scheme 3 Synthesis of thioglycosides catalyzed by Pd(OAc)$_2$/Xanphos system
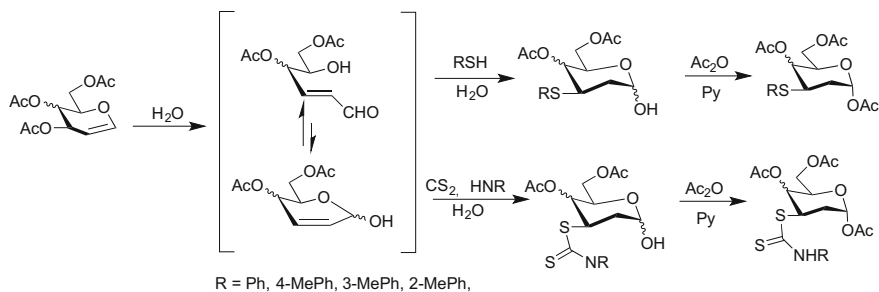
Scheme 4 Catalyst-free green chemistry synthetic approach to 3-thio-2-deoxy, and 3-dithiocarbamate sugar derivatives

Misra and co-workers [16] developed also an odorless methodology of preparing 1-thio-sugars and thio-Michael adducts of carbohydrate derivatives as intermediates for the advanced syntheses of thio-sugars. Selected synthetic routes to these intermediates are shown in Scheme 5.
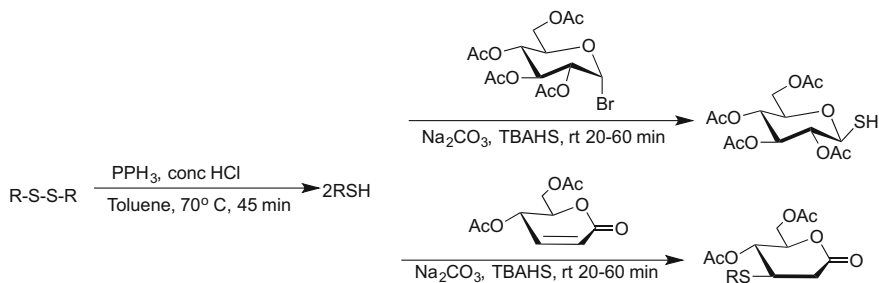
Scheme 5 Synthesis of carbohydrate thiols and 3-Thio-Michael adducts
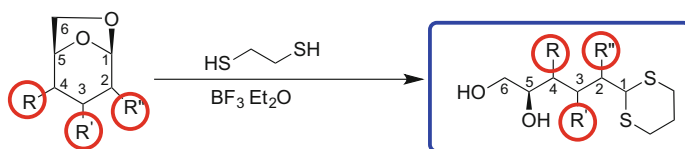
Scheme 6 Lewis acid-catalyzed Opening of 1,6-Anhydrosugars with 1,3-propanethiol

Krohn and co-workers [17] developed a successful methodology of Lewis acid-catalyzed opening of 1,6-anhydro sugars with 1,3-propanedithiol to produce open chain aldehydes protected as 1,3-dithianes. The methodology constitutes a simple route to the complex macrolide building blocks, which are difficult to synthesize. The primary example of this strategy is illustrated in Scheme 6.

Joshi and Anslyn [18] developed a novel approach to dynamic library of thiol exchange with β-sulfido-α,β-unsaturated carbonyl compounds.

The equilibrium between thiols and β-sulfido a,β-unsaturated carbonyls is observed within a few hours. These particular time scales make this system ideal for creation of dynamic combinatorial library.

The team cleverly utilizes the previously well-established thio addition to β-sulfido-conjugated system, as shown in Scheme 7.

Anslyn and co-workers [19] also recently developed a unique thio-click coupling and decoupling approach, which utilizes a reversible amino and thiol coupling via a conjugate acceptor. Scheme 8 illustrates this elegant methodology.

Interestingly, Shi and Greaney [20] reported earlier (in 2005) a similar reversible Michael addition approach. The authors developed specific reaction conditions for subsequent decoupling. The synthetic approach is shown in Scheme 9.

Among approaches used for the synthesis of macrolide thiols and disulfides, Otto and co-workers [21] constructed a dynamic combinatorial library. Dynamic libraries of macrocyclic disulfides form spontaneously upon stirring a mixture of three selected dithiols at pH 7–9 in an open flask. Oxygen from the air is sufficient to effectively oxidize thiols to disulfides. The simplified aspect of these thio functionalization reactions is shown in Scheme 10.

Rim and co-workers [22] discovered a 1,3,5-triacryloylhexahydro-1,3,5-triazine (TAT) system for an ionic thiol-ene click reaction with the formation of functional tripodal thioethers. The authors continue to explore previously unknown chemistry of the TAT moiety and its potential biological importance and applications. According to the authors, thiol-ene reactions tolerate a wide range of functionalities including amino, hydroxyl, carboxylate, and trimethoxysilyl groups. Commercially available aliphatic and aromatic thiols efficiently reacted with TAT to produce thio adducts in high yields (63–96%) and high purity. Some of the aspects of thiol-ene click reactions are shown in Scheme 11.

Gothelf and co-workers [23] developed a cleavable amino-thiol linker for reversible linking of amines to DNA. This discovery has a great practical potential for the exploration of various protection techniques of functionalized DNA derivatives. Some aspects of this new methodology are depicted in Scheme 12.
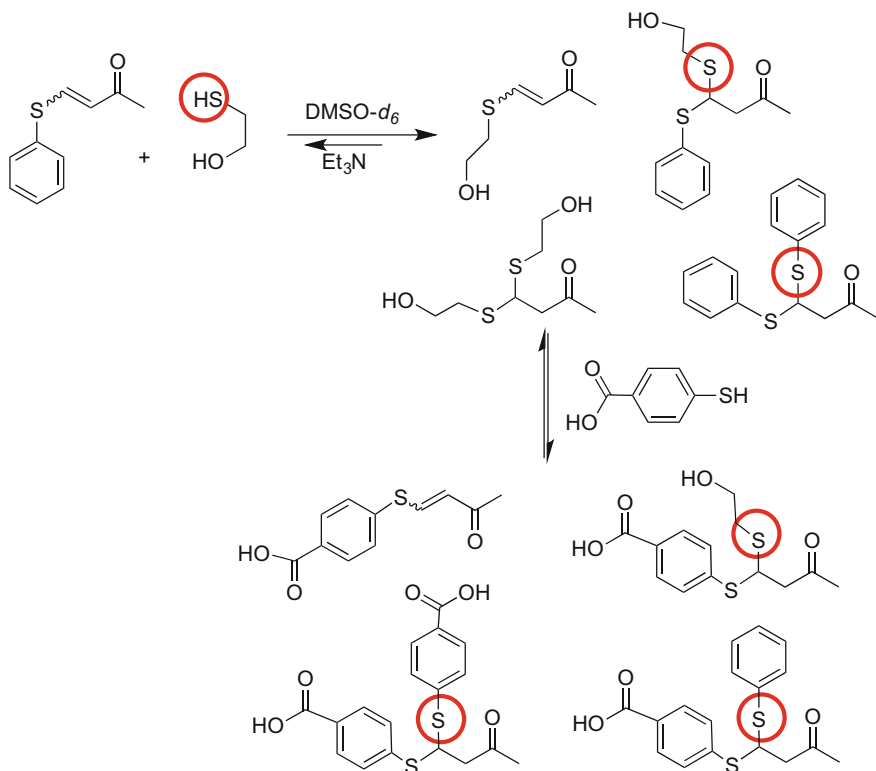
Scheme 7 Dynamic thiol exchange with β-sulfido-α,β-unsaturated carbonyl compounds
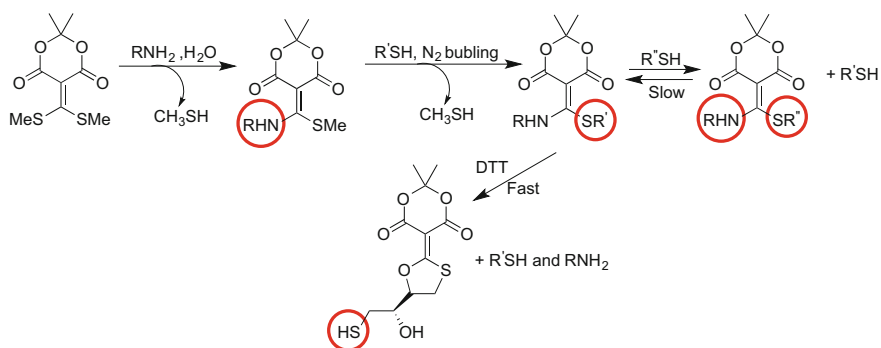
Scheme 8 Click and chemically triggered declick strategy by reversible amine and thiol coupling via a conjugate acceptor
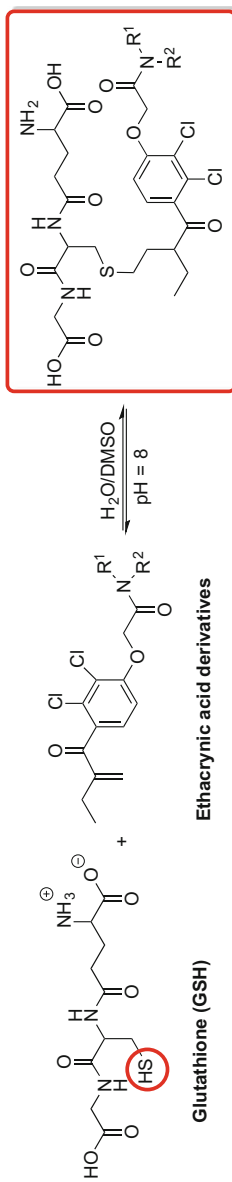
Scheme 9 Reversible Thio-Michael addition of thiols to ethacrynic acid enone
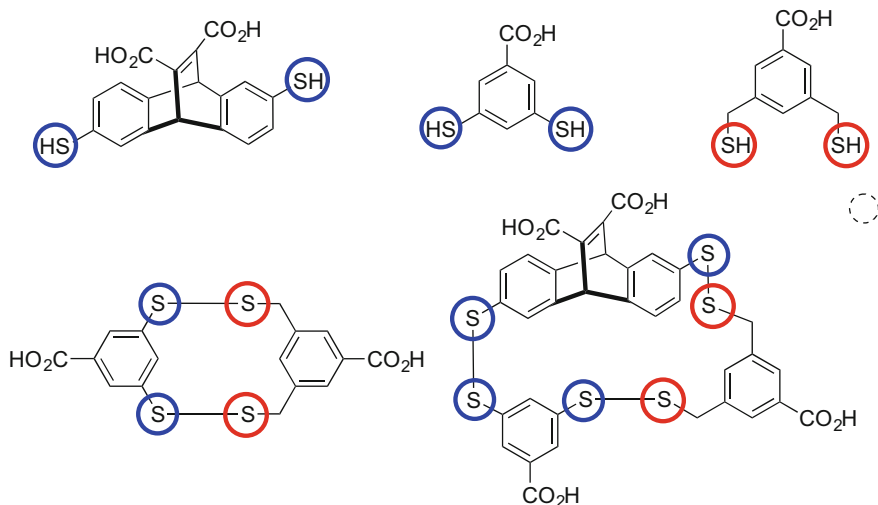
Scheme 10 Dynamic combinatorial library of macrocyclic disulfides
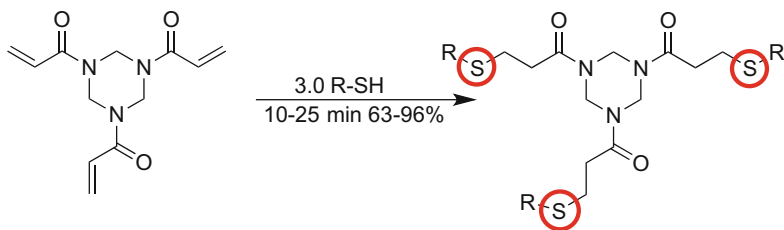
R = -CH$_2$CH$_2$OH, -CH$_2$CH(OH)CH$_2$OH, -CH$_2$CO$_2$CH$_3$, -CH$_2$CH$_2$CH$_2$Si (OCH$_3$)$_3$,
    -CH$_2$CH$_2$NH$_2$,-CH$_2$CH$_2$CO$_2$H, -CH$_2$CH(NH$_2$)CO$_2$H, -o-C$_6$H$_4$NH$_2$, -o-C$_6$H$_4$CO$_2$H

Scheme 11 Thiol-ene reactions of 1,3,5-triacryloylhexahydro-1,3,5-triazine (TAT)
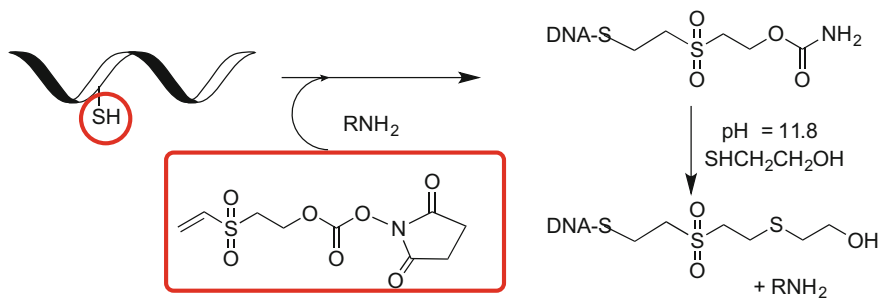
Scheme 12 A cleavable amino-thiol linker for reversible linking of amines to DNA
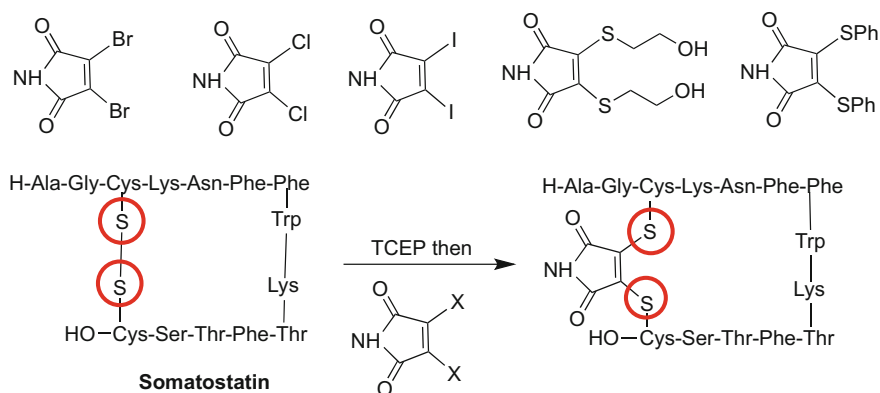
Scheme 13 Maleimides coupling with peptide disulfides

Schumacher and co-workers [24] created a new approach to the protein PEGylation utilizing maleimide bridging of disulfides. The highly reactive conjugate system of selected functionalized maleimides can be coupled with disulfides to form C–S functionalized derivatives as shown in Scheme 13.

Campopiano and co-workers [25] discovered novel glutathione S-transferase (GST) inhibitors using a dynamic, combinatorial chemistry approach. The synthetic approach to this new class of inhibitors begins exploration of these fascinating and medicinally important molecules. Some of the examples are depicted in Scheme 14.

Shiu and co-workers [26] developed cleavable reagents to modify cysteine-containing peptides in an aqueous medium. Highly reactive alkynes were used as starting materials, as depicted in Scheme 15.

Finally, Dondoni and co-workers [27] highlighted a critically important approach in thio-click chemistry; thio-ene cluster formation, and thiol-yne click reaction [28, 29]. Dondoni [30] further developed a thiol-yne strategy applied to diagnostic aspects of serum albumin. Massi and Nani [31] reviewed previously reported methods of thiol-yne click chemistry, creating an up-to-date chronology.
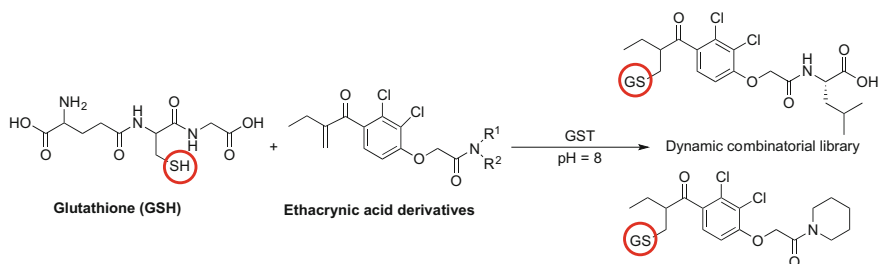
Scheme 14 Glutathione S-transferase inhibitors synthesized via dynamic combinatorial chemistry
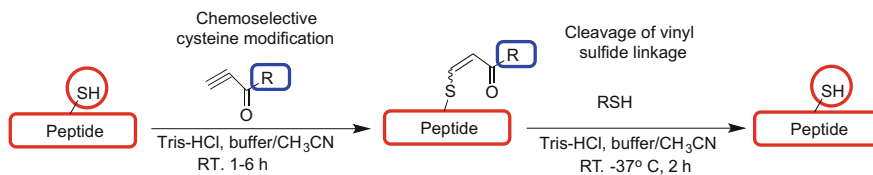
Scheme 15 Cleavable reagent for the modification of cysteine-containing peptides

3   Conclusion

The new strategical developments in the construction of protected diverse molecular targets of biological importance are growing steadily. Among the important areas, the glycoscience is one with particularly enormous growth. Other areas including biomolecular and macromolecular chemistry are developing as well. Among many strategic approaches, thiol directed functionalization and coupling reactions are of great importance and applicability.

When applied to thiols the CAD methodology will always utilize the convenience of four essential factors: the reacting system, catalysts, solvents, and thiol reactivity. We hope that the CAD strategy of creation of sacrificial unit will be further developed into conventionally applicable approach to many new targets of biological importance.

Additionally, other multiple approaches were developed for forming C–S bonds via methodologies utilizing thiol-ene and thiol-yne additions, providing the desired coupling C–S products in a highly stereoselective manner. Moreover, specific reaction conditions are compatible with the stability of the functionalized substrates and products, so yields of desired coupling products are not compromised. The demonstrated thiol-ene and thiol-yne sequences indicate the great potential of functionalized organic and carbohydrate thiols in the synthesis of highly functionalized biomimetic structure motifs by operationally simple protocols. It is worth adding that some of the discussed processes became competitive to the addition of thiols to conjugates multiple bonds such as the Michael addition. All of the new strategies currently available or under development constitute a significant milestone in the area of glycoscience. These new synthetic methodologies are of utmost importance and will be closely followed, as many new biological targets will constitute promising prospects for the future syntheses.

References

1. Bielski R, Witczak ZJ (2013) Strategies of coupling molecular units when later decoupling is needed. Chem Rev 113:2205–2243
2. Hoyle CE, Bowman CN (2010) Thiol-ene click chemistry. Angew Chem Int Ed 49:1540–1573

3. Nair DP, Podgórski M, Chatani S, Gong T, Xi W, Fenoli CR, Bowman CN (2014) The thiol-michael addition click reaction: a powerful and widely used tool in materials chemistry. Chem Mater 26:724–744

4. Witczak ZJ (2013) Thio-click chemistry in glycosciences: overview and perspectives. In: Witczak ZJ, Bielski R (eds) Click chemistry in glycoscience, new development and strategies. Willey, Hoboken, pp 33–43

5. Yin C, Huo F, Zhang J, Martınez-Manez R, Yang Y, Haigang Lv, Li S (2013) Thiol-addition reactions and their applications in thiol recognition. Chem Soc Rev 42:6032–6059

6. Chen X, Zhou Y, Peng X, Yoon J (2010) Fluorescent and colorimetric probes for detection of thiols. Chem Soc Rev 39:2120–2135

7. Shiu H-Y, Chong H-C, Leung Y-C, Wong MK, Che C-M (2010) A highly selective FRET-based fluorescent probe for detection of cysteine and homocysteine. Chem Eur J 16:3308–3313

8. Huo FJ, Sun Y-Q, Jing Su J, Yang Y-T, Yin C-X, Chao J-B (2010) Chromene "Lock", Thiol "Key", and Mercury (II) Ion "Hand": a single molecular machine recognition system. Org Lett 12:4756–4759

9. Clayden J, MacLellan P (2011) Asymmetric synthesis of tertiary thiols and thioethers. Beilstein J Org Chem 7:582–595

10. Khatik GL, Kumar R, Chakraborti AK (2006) Catalyst-free conjugated addition of thiols to α-β-unsaturated carbonyl compounds in water Org. Lett 8:2433–2436

11. Reddy VP, Kumar AV, Swapna K, Rao KR (2009) Nanoindium oxide as recyclable catalyst for C-S cross-coupling of thiols with aryl halides under ligand free conditions Org Lett 11:1697–1700

12. Li S, Lu Z, Meng L, Wang J (2016) Iridium-catalyzed asymmetric addition of thiophenols to oxabenzonorbornadienes Org Lett 18:5276–5279

13. Brachet E, Brion J-D, Messaoudi S, Alami M (2013) Palladium-catalyzed cross-coupling reaction of thioglycosides with (hetero)aryl halides Adv. Synth Catal 355:477–490

14. Brachet E, Brion J-D, Alami M, Messaoudi S (2013) Stereoselective palladium-catalyzed alkenylation and alkynylation of thioglycosides Adv. Synth Catal 355:2627–2636

15. Mandal PK, Maiti GH, Misra AK (2008) Catalyst-free efficient synthesis of 3-thio-2-Deoxysugar derivatives in water. J Carb Chem 27:238–257

16. Mukherjee C, Misra AK (2007) Odorless preparation of thioglycosides and thio Michael adducts of carbohydrate derivative. J Carb Chem 26:213–221

17. Krohn K, Ahmed I, Gehle D, Al Sahli M (2009) Open chain chiral macrolide building blocks by opening of deoxygenated 1,6-Anhydrosugars with 1,3-Propanedithiol. J Carb Chem 28:238–257

18. Joshi G, Anslyn EV (2012) Dynamic thiol exchange with β-sulfido-α, β-unsaturated carbonyl compounds and dithianes. Org Lett 14:4714–4717

19. Diehl KL, Kolesnichenko IV, Robotham SA, Bachman JL, Zhong Y, Brodbelt JS, Anslyn EV (2016) Click and chemically triggered declick reactions through reversible amine and thiol coupling via a conjugate acceptor. Nat Chem 8:968–973

20. Shi B, Greaney MF (2005) Reversible Michael addition of thiols as a new tool for dynamic combinatorial chemistry. Chem Commun 886–888

21. Otto S, Furlan RLE, Sanders JKM (2002) Selection and amplification of hosts from dynamic combinatorial libraries of macrocyclic disulfides. Science 297:590–593

22. Rim C, Lahey LJ, Patel VG, Zhang H, Son DY (2009) Thiol-ene reactions of 1,3,5-triacryloylhexahydro-1,3,5-triazine (TAT): facile access to functional tripodal thioethers. Tetrahedron Lett 50:745–747

23. Højfeldt JW, Blakskjaer P, Gothelf KV (2006) A cleavable amino-thiol linker for reversible linking of amines to DNA. J Org Chem 71:9556–9559

24. Schumacher FF, Nobles M, Ryan CP, Smith MEB, Tinker A, Caddick S, Baker JR (2011) In situ maleimide bridging of disulfides and a new approach to protein PEGylation. Bioconjug Chem 22:132–136

25. Shi B, Stevenson R, Campopiano DJ, Greaney MF (2006) Discovery of glutathione S-transferase inhibitors using dynamic combinatorial chemistry. J Am Chem Soc 128:8459–8467
26. Shiu H-Y, Chan T-Y, Ho C-M, Liu Y, Wong M-K, Che C-M (2009) Electron-deficient alkynes as cleavable reagents for the modification of cysteine-containing peptides in aqueous medium. Chem Eur J 15:3839–3850
27. Fiore M, Chambery A, Marra A, Dondoni A (2009) Single and dual glycoside clustering around calix [4] arene scaffolds *via* click thiol-ene coupling and azide-alkyne cycloaddition Org. Biomol Chem 7:3910–3913
28. Conte ML, Pacifico S, Chambery A, Marra A, Dondoni A (2010) Photoinduced addition of glycosyl thiols to alkynyl peptides: use of free-radical thiol-yne coupling for post-translational double glycosylation of peptides. J Org Chem 75:4644–4647
29. Dondoni A, Marra A (2012) Recent applications of thiol-ene coupling as a click process for glycoconjugation. Chem Soc Rev 41:573–586
30. Dondoni A (2011) Multi-molecule reaction of serum albumin can occur through thiol-yne coupling. Chem Commun 47:11086–11088
31. Massi A, Nanni D (2012) Thiol-yne coupling; revisiting old concepts as a breakthrough for up-to date applications Org. Biomol Chem 10:3791–3807

Chemical Approaches Towards Neurodegenerative Disease Prevention: The Role of Coupling Sugars to Phenolic Biomolecular Entities

Catarina Dias, Ana M. Matos and Amélia P. Rauter

Abstract Polyphenols are natural molecular entities exhibiting a wide variety of bioactivities including anticholinergic and/or antiamyloidogenic activities. Their low solubility is recognized as a key factor for bioavailability and their glycosylation is indeed relevant to improve the bioaccess to these molecules. In this chapter, chemical and enzymatic syntheses of glycosylated flavonoids, stilbenoids, phenylethanoids and phenylpropanoids are illustrated, covering examples that demonstrate the impact of coupling sugars to bioactive aglycones in their bioavailability and in their pharmacological activity. The chapter is focused particularly on glycosyl polyphenols with promising activities against neurodegenerative impairments, given their potential to intervene in biological processes that cause catastrophic diseases, namely the Alzheimer's disease.

1 Introduction

Polyphenols are plant secondary metabolites present in the common human diet and known to play important roles in human health. They are poorly absorbed, resulting in a very low concentration in the circulatory streams [69]. The modification of their physicochemical properties such as solubility and partition coefficient by glycosylation seems to exert a positive influence on the entry of polyphenols into enterocytes [69]. The low solubility of most of the polyphenol aglycones may also result from their tendency to form aggregates via hydrophobic interactions with aromatic rings, and hydrogen bonding by the hydroxy groups [3]. In nature, polyphenols occur often as glycosylated derivatives. The sugar moiety of polyphenol glycosides plays a major role in their absorption [69] but polyphenol

---

Catarina Dias and Ana M. Matos have both contributed equally to the chapter.

C. Dias (✉) · A. M. Matos (✉) · A. P. Rauter (✉)
Faculdade de Ciências, Centro de Química E Bioquímica,
Universidade de Lisboa, Ed C8, Piso 5, 1749-016 Lisbon, Portugal
e-mail: aprauter@fc.ul.pt

glycosylation may also exert other benefits by improving bioavailability or preventing oxidation by masking phenolic groups. In this chapter, synthetic strategies via chemical or enzymatic methodologies to access biologically active glycosyl polyphenols are illustrated, covering flavonoids, stilbenoids, phenylpropanoids and phenylethanoids. Natural occurrence and compound bioactivities are also reviewed for the promising polyphenol molecular entities described that exhibit neuroprotective activities.

2　Glycosylated Flavonoids

Flavonoids are polyphenolic secondary metabolites in the plant kingdom whose structural feature is based on derivatives of a phenyl-substituted 1-phenylpropane possessing a C_{15} skeleton. In this chapter, the given examples focus particularly on flavones, whose structure is that of a 1-benzopyran (chromene), in which the aromatic ring is designated as ring A and the pyran as ring C, along with the (substituted) phenyl group (ring B) on ring C at position 2 (flavone) or position 3 (isoflavone). Thousands of different scaffolds have been isolated and structurally identified over the past decades, and many have been reported due to their wide-range bioactive profiles often associated with very potent antioxidant and anti-inflammatory effects [70]. They may occur as aglycones or as the corresponding glycosylated forms, either as *O*-glycosides or *C*-glycosyl derivatives; yet, the advantages of glycosyl flavones over the corresponding aglycones have been highlighted in the context of Alzheimer's disease with respect to their ability to remodel and inactivate neurotoxic amyloid β (Aβ) aggregates [36], again reinforcing the importance of the sugar moiety for optimized anti-neurodegenerative activity.

The growing interest in the therapeutic potential of glycosyl flavonoids has motivated organic and medicinal chemists to develop efficient synthetic and bioenzymatic routes involving a diverse collection of sugar coupling reactions. By describing the synthesis of some of the most promising molecular entities with neuroprotective activities, we will provide an overview of the most useful methodologies for the generation of flavones bearing in their structure O-linked or C-linked sugars, covering both chemical and enzymatic synthesis reported in the literature.

2.1　Flavone Glycosides

The 7-*O*-β-glucuronide of baicalein, baicalin (**1**), is one of the most abundant compounds in *Scutellaria baicalensis Georgi*, a plant extensively used in traditional Chinese medicine for the treatment of inflammatory disorders, bacterial infections, among others [6]. Baicalin (**1**) itself was recently found to improve Aβ-induced learning and memory deficits in rats by attenuating hippocampal injury and neuron apoptosis [11]. Its anti-inflammatory activity has actually been proposed as a

paramount mechanism underlying these neuroprotective effects [6], namely by inhibiting microglial activation and inflammatory cytokine secretion [83]. Moreover, baicalin (**1**) could also upregulate antioxidant enzymes such as superoxide dismutase, catalase and glutathione peroxidase, thus contributing to decreased oxidative injury in the brain of diseased animals [11].

Based on its promising therapeutic potential, Li and co-workers described an efficient route for baicalin (Scheme 1) starting from the selectively acetylated aglycone **3**, which was accessed after a series of simple protection–deprotection reactions [45], and using the first type of sugar donor ever applied in the synthesis of flavonoid glycosides: a glycosyl bromide [62]. In this method, 6-OTBDPS protected bromide **2** was coupled to the aglycone in an Ag_2O-promoted reaction that afforded only the β-glucoside in 92% yield due to acyl neighbouring group participation, thus overpowering the otherwise dominant anomeric effect that would have given the α-anomer as the major product. After deprotection with TBAF, position 6″ was then submitted to Widlanski oxidation using TEMPO and BAIB to give the glucuronic acid derivative **4**, followed by a deacylation reaction that led to the desired product, baicalin (**1**).

The 4′-hydroxy analogue of baicalin (**1**), scutellarin (**5**), is the major component of *the Erigeron breviscapus* Hand-Mazz flavonoid extract, also used in traditional Chinese medicine for the treatment of cerebral infarction and other cardiovascular diseases [54]. Similarly, this compound was found to attenuate neuroinflammation through the suppression of microglial activation [15], and was indeed associated with major improvements in neuronal injury and behaviour of rats with cerebral ischemia [14, 68]. Moreover, it is able to inhibit Aβ aggregation in vitro, while preventing Aβ-mediated neuronal cell death [91].

Nonetheless, pharmacokinetic studies have revealed that scutellarin (**5**) displays a rather poor bioavailability due to the action of endogenous β-glucuronidase enzymes that readily hydrolyze the glycosidic bond [5, 17, 23]. To surpass this problem, Li and co-workers designed and synthesized the scutellarin β-*O*-glucosyl analogue **6** with improved physicochemical properties and an even more
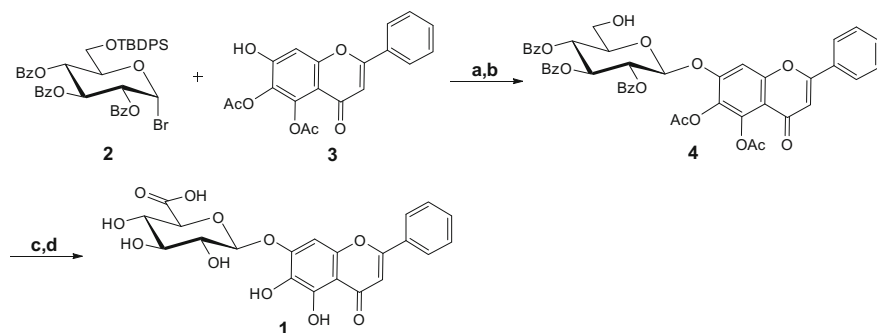
Scheme 1 Reagents and conditions: **a** Ag_2O, 4 ÅMS, quinoline, r.t. (92%); **b** TBAF, AcOH, THF, 4 h (86%); **c** TEMPO, BAIB, DCM/H_2O, r.t. (87%), **d** $Mg(OMe)_2$, MeOH, r.t. (85%) [45]
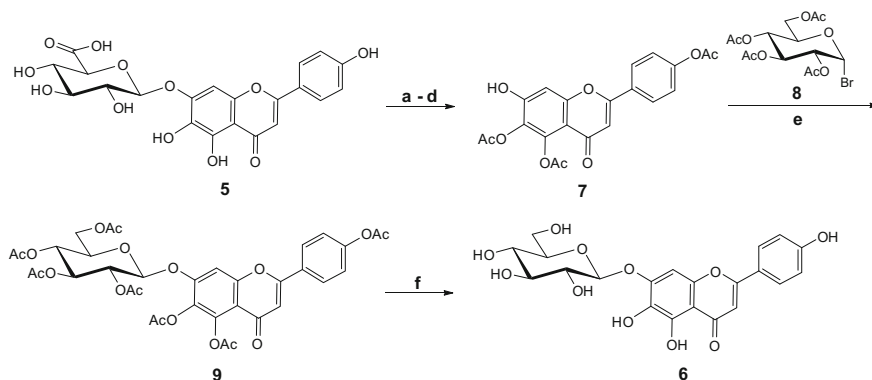
Scheme 2 Reagents and conditions: **a** 6 N HCl, EtOH, 120 °C (17%); **b** pyridine, Ac$_2$O, DMAP, 25 °C (79%); **c** BnBr, K$_2$CO$_3$, KI, acetone, reflux (70%); **d** Pd/C, H$_2$, DCM/EtOH, 25 °C (95%); **e** CuSO$_4$, AgO, quinoline, 25 °C (40%); **f** NaOH, CHCl$_3$, 0 °C (41%) [40]

pronounced attenuation of H$_2$O$_2$-induced neuronal damage when compared to scutellarin (**5**) [40]. Using scutellarin itself as the aglycone source, the authors coupled compound **7** to the glucosyl bromide **8** with Ag$_2$O and CuSO$_4$ as promotors; yet, the β-*O*-glucoside **6** was achieved in only moderate yield (40%) (Scheme 2).

Among the most promising *O*-glucosyl flavonoid leads against neurodegenerative diseases is quercetin 3-β-*O*-glucoside (**10**) (trivial name: isoquercetin), which has been isolated from a variety of sources, including mangos or medicinal plants such *Serjania erecta Radlk* (Sapindaceae) or *Psidium guajava* L. [18, 46, 48]. In addition to its antioxidant and anti-inflammatory activities, this compound is able to prevent hippocampal neuronal apoptosis after cerebral ischemia and reperfusion injury [76, 77], and displays protective effects against Aβ-induced cytotoxicity. Importantly, it was also found to inhibit both BACE-1 and AChE with IC$_{50}$ values of 41.2 and 66.9 μM, respectively [26, 27]. Furthermore, a comparative study between polyphenolic glycosides and their respective aglycones has shown that whilst quercetin (**11**) acts by remodelling Aβ toxic oligomers into large aggregates, isoquercetin (**10**) rapidly disaggregates the amyloid structures into soluble polypeptides as a result of a synergistic action between the sugar and the aglycone [36].

Isoquercetin (**10**) can be obtained from quercetin (**11**) by the action of UGT78D1, a flavonoid-specific uridine diphosphate glycosyltransferase (Scheme 3) that catalyzes the in vitro regioselective transfer of a glucose or a rhamnose unit from UDP-glucose or UDP-rhamnose, respectively, to flavonoid glycosyl acceptors containing a hydroxyl group in position 3, as reported by Ren and co-workers [59]. This study was able to clarify the substrate specificity of this enzyme in detail, showing that only flavones hydroxylated in both rings A and B are recognized by UGT78D1, highlighting 2′-OH flavones as exceptions to this rule.

Tiliroside (**12**) is a kaempferol 3-β-*O*-glycoside that can be found in *Agrimonia pilosa* or *Potentilla chinesis* for instance, and displayed stronger AChE inhibitory
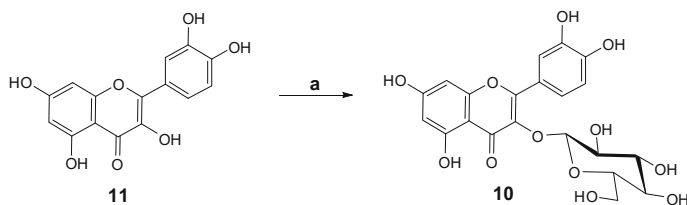
Scheme 3 Reagents and conditions: **a** Recombinant UGT78D1, 50 mM Tris-HCl pH 7.2, UDP-Glucose, 30 °C [59]

activity when compared to isoquercetin (**10**), with an IC_{50} value of 25.5 μM [27, 57]. It was also found to inhibit neuroinflammation in activated microglial cells by modulating pro-inflammatory intracellular pathways, which was at least in part attributed to its antioxidant properties [71].

The synthesis of tiliroside (**12**) was described in 1981 by Vermes and co-workers (Scheme 4) [72]. In the first step of this route, glucosyl bromide **13** and 4',7-*O*-dibenzyl kaempferol (**14**) were coupled in a reaction promoted by Ag_2CO_3 to afford the β-anomer in 54% yield. After debenzylation followed by acetylation and selective removal of the 6″-*O*-chloroacetyl protecting group, intermediate **15** was generated and subsequently esterified by *p*-coumaroyl chloride in pyridine. Further deprotection directly afforded tiliroside (**12**) in good yield. Many other phenyl-propanoid glycosides with neuroprotective activities such as this one will be presented and their synthetic routes described in detail in Sect. 3.
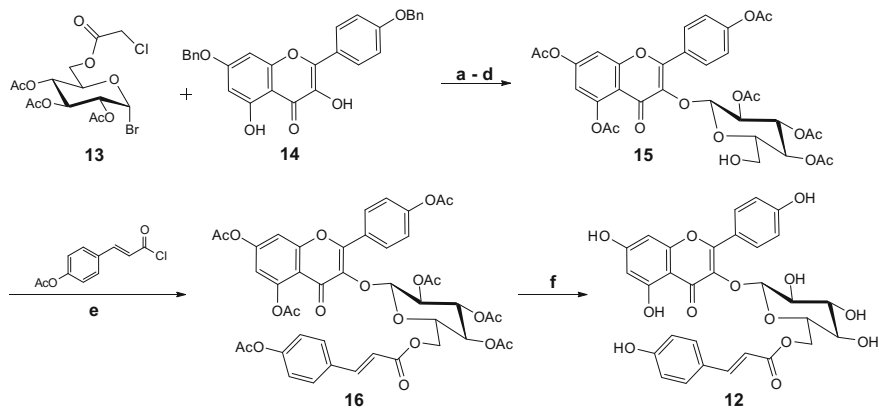
Scheme 4 Reagents and conditions: **a** Ag_2CO_3, pyridine, drierite, 0 °C (54%); **b** H_2, Pd/C, EtOH, r.t. (90%); **c** Ac_2O, pyridine, r.t. (88%); **d** MeOH, thiourea, r.t. (90%); **e** DCM, pyridine (62%); **f** $CHCl_3$, NaOMe, MeOH (67%) [72]

2.2   *C-Glycosyl Flavones and Isoflavones*

In spite of the remarkable neuroprotective effects associated with the flavone O-glycosides described above, *C*-glycosyl flavonoids have been receiving growing attention for their insusceptibility to in vivo hydrolysis by glucosidases, allowing them to remain intact in the blood circulation following oral administration [9]. Even though the synthesis of the C–C bond usually requires stronger conditions when compared to the formation of the C–O bond in O-glycosides, a variety of methods have been reported in the literature over the past few decades, offering a wide range of options for regio- and stereoselective reactions using different glycosyl donors and acceptors when the time comes to design a synthetic route for the target compound [64].

Vitexin (**17**) and isovitexin (**18**), the 8-β-*C*- and 6-β-*C*-glucosyl derivatives of apigenin, respectively, are good examples of the potential of natural *C*-glycosyl flavonoid derivatives against neurodegenerative disorders. These compounds can be found in *Serjania erecta Radlk* (Sapindaceae) [18] or the flour from the *Prosopis alba* seed [4], for instance, and both were able to inhibit AChE and BChE with IC_{50} values ranging from 6.2 to 12.2 μM, although vitexin was substantially more effective as a BACE-1 inhibitor than isovitexin (IC_{50} = 51.1 μM vs. >100 μM), thus indicating a preference for the sugar moiety to be in position 8 for improved affinity towards the enzyme [7, 8]. Vitexin (**17**) has also been described to exert neuroprotective effects in cerebral ischemia and reperfusion injury by positively and negatively modulating cell proliferation and apoptosis pathways, respectively [79]. In addition, vitexin (**17**) was found to have a more pronounced impact in reversing Aβ-induced cytotoxicity not only when compared to isovitexin (**18**), but also when put alongside with the earlier presented isoquercetin (**10**) [18].

Back in 1995, Mahling and co-workers developed a synthetic route for both vitexin (**17**) and isovitexin (**18**) by taking advantage of the Fries-type rearrangement, described to occur in *O*-aryl glycosides with high regio- and stereoselectivity to afford the corresponding *ortho*-hydroxy *C*-glycosyl phenolic derivative [32, 47, 64]. Hence, in the first step of this synthesis (Scheme 5), the reaction of the glycosyl trichloroacetimidate **19** with the silyl-protected acetophenone **20** was catalyzed by TMSOTf at-30 °C and afforded the α-*O*-glycoside **21** in 85% yield. After cleavage of the remaining TBS group followed by regioselective benzylation in position 4, a Fries-type rearrangement took place in another TMSOTf-catalyzed reaction, this time at room temperature, to afford the corresponding β-*C*-glycosyl derivative in 57% yield. Subsequent acylation converted this derivative into intermediate **17** and, at this point, the Baker–Venkataraman rearrangement was carried out and resulted in a mixture of compounds **23** and **24**, which were both cyclized and further deprotected after separation to give vitexin (**17**) and isovitexin (**18**).

Vitexin (**17**) and isovitexin (**18**) were also obtained as protected intermediates in a more recent and concise synthesis developed by Furuta and co-workers [16] with the ultimate goal of accessing compound **25**, an anti-inflammatory glycosyl flavone
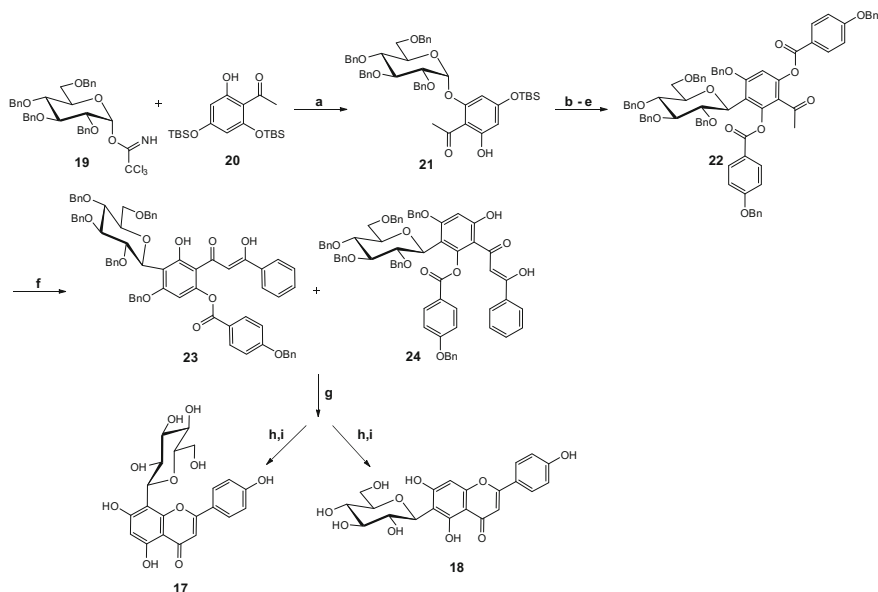
Scheme 5 Reagents and conditions: **a** TMSOTf, DCM, −30 °C (85%); **b** TBAF, THF (86%); **c** BnBr, NaH, DMF (89%); **d** TMSOTf, DCM, r.t. (57%); **e** WSC, DMAP, 4-(benzyloxy)benzoic acid, DCM (81%); **f** TBA·HSO₄, K₂CO₃, H₂O/Benzene, 60 °C (39%); **g** TMSOTf, DCM, r.t. (17%); **h** NaOMe, MeOH, 50 °C (62 and 100%, respectively); **i** H₂, Pd/C, EtOAc/MeOH (88 and 89%, respectively) [47]

isolated from oolong tea extract [24]. In this route (Scheme 6), trichloroacetimidate **19** was directly coupled with the monobenzyl-protected acetophenone **26** to afford the desired β-*C*-glucosyl derivative in 69% yield, which was further acylated to give intermediate **27**. In contrast with the previous work by Mahling et al., this procedure involves the initial formation of a glycoside at low temperature, which then undergoes, by warming up, the $O \rightarrow C$ Fries-type rearrangement in situ [64].

In another one-pot reaction using potassium carbonate in pyridine under reflux, intermediate **27** was converted into both protected isovitexin (**28**) and protected vitexin (**29**); yet, to accomplish the synthesis of the target compound, only **28** proceeded in this route. After debenzylation, it was submitted to an intramolecular Mitsunobu reaction using modified experimental conditions in which inversion of the configuration of carbon 2″ led to the transformation of the *gluco* derivative into the desired *manno* derivative in tandem with the formation of a fused tetracyclic system with the aglycone. Further deprotection afforded the target compound **25**, but it is still interesting to note that this was the major product of Mitsunobu reaction regardless of the presence of a primary alcohol and another phenolic group in its precursor, thus highlighting reaction regiospecificity.

Orientin (**31**) and isoorientin (**32**) are another pair of *C*-glucosyl flavonoid derivatives extensively studied for their potential against neurodegenerative
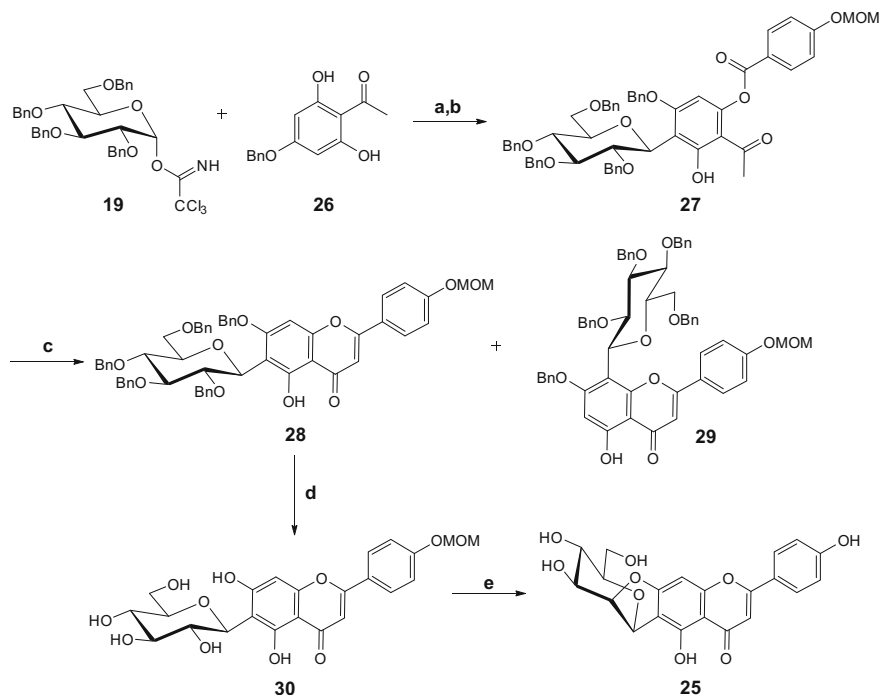
Scheme 6 Reagents and conditions: **a** TMSOTf, DCM, 0 °C → r.t. (69%); **b** 4-(methoxymethoxy)benzoic acid, DCC, DMAP, DCM, r.t. (82%); **c** K$_2$CO$_3$, pyridine, reflux (17 and 15%, respectively); **d** Pd(OH)$_2$, H$_2$, EtOH, 35 °C (95%); **e** 1,1′-azobis(N,N-dimethylformamide), Bu$_3$P, THF, 50 °C; then HCl-dioxane, MeOH, r.t. (32%) [16]

processes and can both be found, for instance, in buckwheat bran [92], *Glochidion hypoleucum* (Miq.) Boerl leaves, or *Stellaria holostea* [1]. Orientin (**31**) was able to alleviate cognitive deficits in mice with Alzheimer's disease, while attenuating mitochondrial dysfunction induced by Aβ [84]. Moreover, it exerted neuroprotective effects by inhibiting the activity of three members of the caspase family, including caspase 3, which is directly involved in synaptic loss and cognitive dysfunction in Alzheimer's disease [37, 10]. Both orientin (**31**) and isoorientin (**32**) are BACE-1 inhibitors with IC$_{50}$ values of 16.0 and 20.9 μM, respectively, showing that the presence of the additional hydroxy group in position 3′ when compared to vitexin (**17**) and isovitexin (**18**) positively affects the affinity towards the enzyme, especially in the case of 6-β-C-glucosyl derivatives. Furthermore, they are also AChE and BChE inhibitors and seem to be slightly selective towards the later, with similar IC$_{50}$ values of roughly 11 μM [7, 8].

In contrast to the described synthetic approaches for vitexin (**17**) and isovitexin (**18**), each of these luteolin *C*-glucosyl derivatives has been accessed individually in more effective, regioselective routes. Kumazawa and co-workers reported, on the one hand, the synthesis of orientin (Scheme 7) in which the glucosyl fluoride **33**
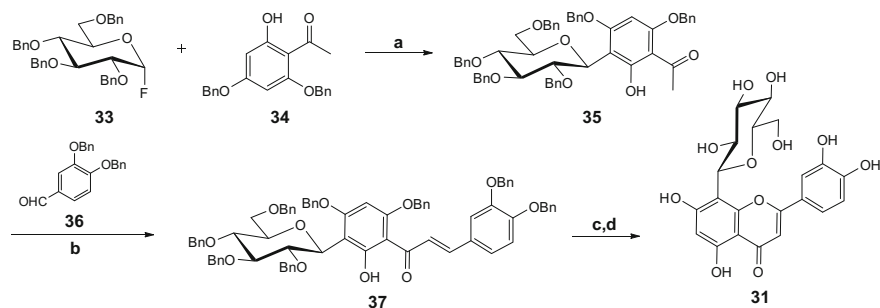
Scheme 7 Reagents and conditions: **a** BF$_3$·Et$_2$O, MS4Å, DCM, −78 °C → r.t. (96%); **b** 1,4-dioxane, aq. NaOH 50%, r.t. (84%); **c** I$_2$, DMSO, reflux (84%); **d** Pd/C, H$_2$, EtOH, r.t. (quantitative yield) [33]

was coupled with acetophenone **34** in a BF$_3$·Et$_2$O-promoted reaction to afford the β-*C*-glucosyl derivative **35** in 96% yield [33]. Subsequently, an aldol condensation led to chalcone **37** which, after cyclization and deprotection, gave orientin (**31**) in very good yield.

On the other hand, these authors were able to develop a synthetic path towards isoorientin (Scheme 8) using the same coupling methodology but taking advantage of differences in hydrogenolysis rates between benzyl and 2-methylbenzyl protecting groups [34]. Indeed, after a series of protection-deprotection reactions, the free hydroxy group was *para* to the sugar moiety in intermediate **40**, and after aldol condensation, cyclization and deprotection, isoorientin (**32**) were successfully generated. It is noteworthy that, in this route, the *C*-glycosylation step was significantly less effective (75% yield) than the one described in the synthesis of orientin (**31**, 95% yield), even though the coupling method applied was the same in both cases. Given that the only difference between the two glycosyl acceptors was the 2-methylbenzyl group in compound **39**, this result highlights the impact of protecting groups on the efficiency of this type of coupling reactions.

More recently, the biosynthesis of vitexin (**17**), isovitexin (**18**), orientin (**31**) and isoorientin (**32**) was accomplished by Hao and co-workers *using Desmodium incanum* root proteins, starting from the corresponding 2-hydroxyflavanones, the required substrates of *C*-glycosyltransferases existent in *Desmodium spp.* [21, 22]. As clarified in a previous report [29], Wessely–Moser isomerization is responsible for the interconversion between the corresponding 8-β-*C*- (**45** and **46**) and 6-β-*C*-glucosyl derivatives (**49** and **50**), as 2-hydroxyflavanones may exist in solution in either open chain or cyclized structures (Scheme 9). In spite of the consequent lack of regioselectivity, these intermediates afforded the respective flavones in overall excellent yields after acid-promoted chemical dehydration.

Puerarin (**51**), the major component of *Puerariae Lobatae Radix* [82], is another *C*-glucosyl flavonoid with potential against neurodegenerative disorders and has received particular attention in regard to its ability to act against diabetes-induced cognitive dysfunction, complementing its known antidiabetic activity [44, 82, 84].
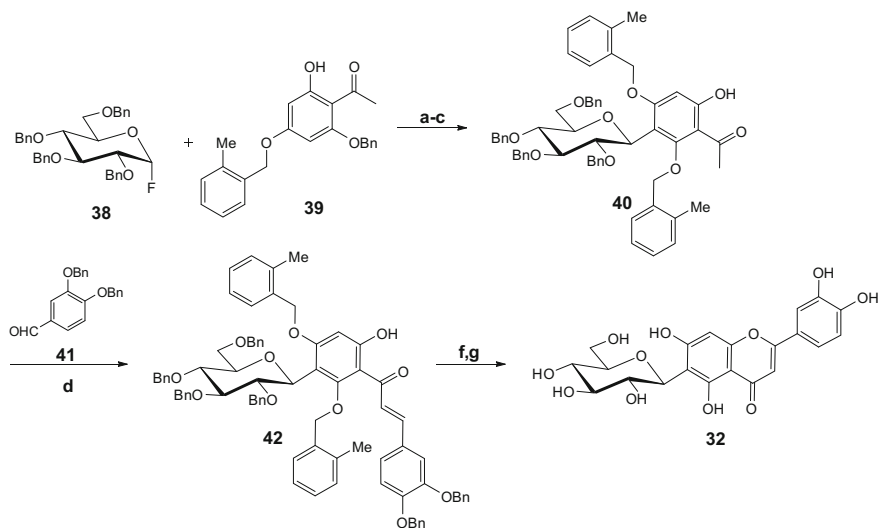
Scheme 8 Reagents and conditions: **a** BF$_3$·Et$_2$O, MS4Å, DCM, −78 °C → r.t. (75%); **b** anhydrous K$_2$CO$_3$, 2-methylbenzyl chloride, DMF, 80 °C (quantitative yield); **c** Pd/C, H$_2$, EtOAc, r.t. (40%); **d** 1,4-dioxane, aq. NaOH 50%, r.t. (91%); **f** I$_2$, DMSO, 200 °C (76%); **g** Pd/C, H$_2$, EtOAc/ EtOH, r.t. (quantitative yield) [34]
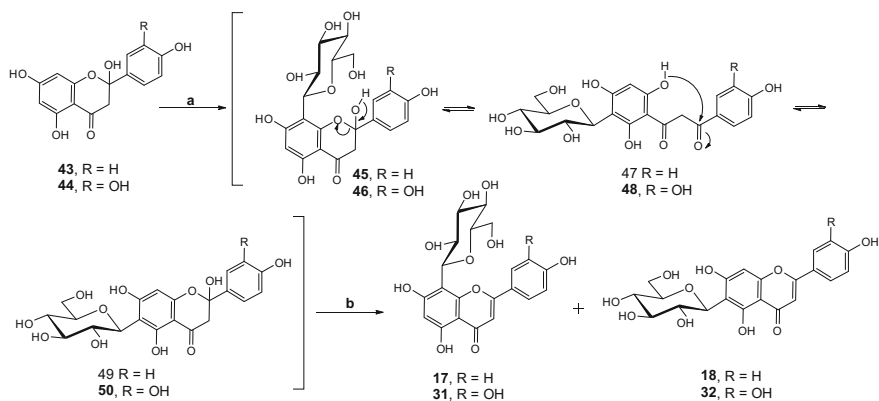
Scheme 9 Reagents and conditions: **a** *D. Incanum* root protein, 100 mM HEPES, 2 mM DTT, pH 7.5, UDP-Glucose, 30 °C; **b** 1 M HCl (vitexin, **17**, 44% and isovitexin, **18**, 53%; orientin, **31**, 43% and isoorientin, **32**, 56%) [22, 29]

This is of particular importance due to the well-established relationship between type 2 diabetes and Alzheimer's disease [31] and, in fact, puerarin (**51**) was found to have neuroprotective activity in STZ-induced diabetic rodents with learning and memory deficits by exerting antioxidant, anti-inflammatory and anti-apoptotic effects [42, 89]. In addition, this *C*-glucosyl isoflavone was able to attenuate Aβ-induced oxidative

stress, cell injury and resulting cognitive impairment [39, 41, 78, 86, 90], and could also improve learning and memory functions in rats with vascular dementia by activating cellular antioxidant defense mechanisms [87].

The total synthesis of puerarin (**51**) was firstly reported by Lee and co-workers in 2003 [38] (Scheme 10). In this approach, the benzyl protected glycopyranolactone **52** was coupled to the lithiated glycosyl acceptor **53** at low temperature, followed by reduction with triethylsilane and $BF_3 \cdot Et_2O$ to give the β-anomer in 56%. After a couple of protection-deprotection reactions, a Friedel-Crafts reaction catalyzed by $AlCl_3$ and subsequent deacetylation gave intermediate **55**, which then underwent aldol condensation with *p*-methoxybenzaldehyde (**56**) to afford chalcone **57**. After acetylation, TTN-promoted oxidative rearrangement of ring B followed by closure of ring C and demethylation gave puerarin (**51**) in moderate overall yield.

The trihydroxyisoflavone analogue of puerarin (**51**) is the 8-β-D-glucosylgenistein (**58**), the main component of the ethyl acetate extract of *Genista tenera*, a plant found in Madeira island and used in folk medicine to treat diabetes [25]. In addition to its potent antidiabetic activity, 8-β-D-glucosylgenistein (**58**) was found to interact with $A\beta_{1-42}$ polypeptides, suggesting potential neuroprotective effects as well. In this study, the binding epitope of 8-β-D-glucosylgenistein (**58**) with Aβ was disclosed, confirming the already expected key role of both aromatic rings in the resulting interaction, and reinforcing the importance of the sugar moiety in the antiamyloidogenic activity of this compound.

The synthesis of 8-β-D-glucosylgenistein (**58**) (Scheme 11) was accomplished by coupling the commercially available glucopyranoside **59** and acetophenone **60** catalyzed by TMSOTf, to give the desired C-glycosylation product in 56% yield, which was selectively benzylated to afford intermediate **61** [25]. Then aldol condensation with *p*-benzyloxybenzaldehyde followed by acetylation led to the formation of chalcone **63** and subsequent TTN-promoted oxidative rearrangement, ring closure and deprotection afforded the target compound, **58**.
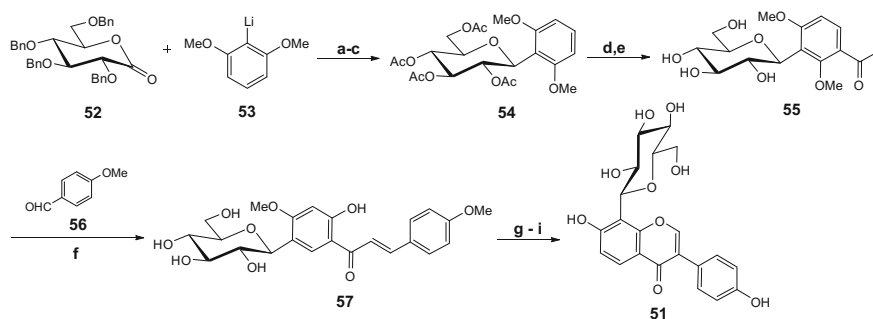
Scheme 10 Reagents and conditions: **a** THF, −78 °C → −10 °C; then TESH, $BF_3 \cdot Et_2O$, DCM, −78 °C → r.t. (56%); **b** Pd/C, H_2, MeOH, r.t. (quantitative yield); **c** pyridine, Ac_2O, r.t. (96%); **d** $AlCl_3$, AcCl, Et_2O, r.t. (69%); **e** 1: Na, MeOH; 2: Dowex 50WX8-200 (97%); **f** NaOH, EtOH, r.t. (92%); **g** pyridine, Ac_2O, r.t. (85%); **h** 1: $Tl(NO_3)_3$, MeOH/$CH(OCH_3)_3$; 2: 10% HCl, MeOH, reflux (84%); **i** TMSI, ACN (35%) [38]

In a nutshell, the coupling of sugars with polyphenols to generate bioactive glycosyl flavonoids may involve a variety of different strategies and experimental conditions which primarily depend upon the available starting materials, reaction promoters or catalysts, and the nature of the pursued C–C or C–O bond. Regio- and stereoselectivity can be achieved with the use of the appropriate sugar protecting groups and glycosyl acceptor, while temperature is a key factor in the formation of either *O*- or *C*-glycosyl derivatives, particularly when a Fries-type rearrangement is involved in the reaction mechanism. Also, by covering the synthesis of structurally complex compounds such as the presented bioactive glycosyl flavones and isoflavones, this section enclosed a number of useful protection–deprotection strategies, interesting rearrangement reactions and cyclization approaches, which may be convenient for the synthesis of new nature-inspired glycosylated molecules towards neurodegenerative disease prevention.

3   Stilbenoid Glycosides

Stilbenoids are natural compounds occurring in a number of plant families, particularly in grapevine [2]. Amongst them, the most well known is resveratrol (*E*)-3,4′,5-trihydroxystilbene, (**64**), possessing anti-inflammatory, antioxidant and chemopreventive activities. This powerful compound is present in wine and has been speculated to be responsible for so-called French paradox, where the saturated fat rich French diet correlates with a low mortality from coronary heart disease [60, 81]. Resveratrol also occurs ubiquitously in nature as resveratrol 3-β-glucoside (piceid, **67**) (Fig. 1). Other stilbenes include pterostilbene (*E*)-4′-hydroxy-3,5-
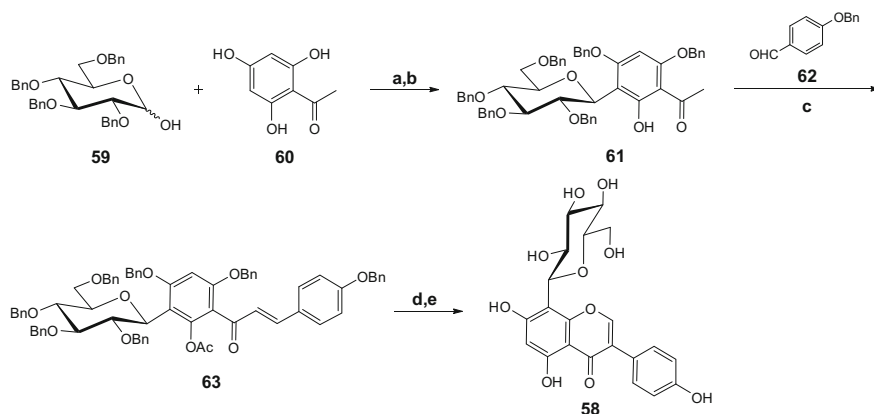
Scheme 11 Reagents and conditions: **a** 1. TMSOTf, DCM/ACN, drierite, −40 °C → r.t. (56%); **b** BnBr, K$_2$CO$_3$, r.t. (74%), **c** 1: 1,4-dioxane, aq. NaOH 50%, reflux; 2: pyridine, Ac$_2$O, DMAP, r.t. (60%); **d** 1: Tl(NO$_3$)$_3$, MeOH/CH(OCH$_3$)$_3$, 40 °C; 2: THF/MeOH, aq. NaOH 50%, r.t. (63%); **e** Pd/C, H$_2$, EtOAc/MeOH, r.t. (96%) [25]
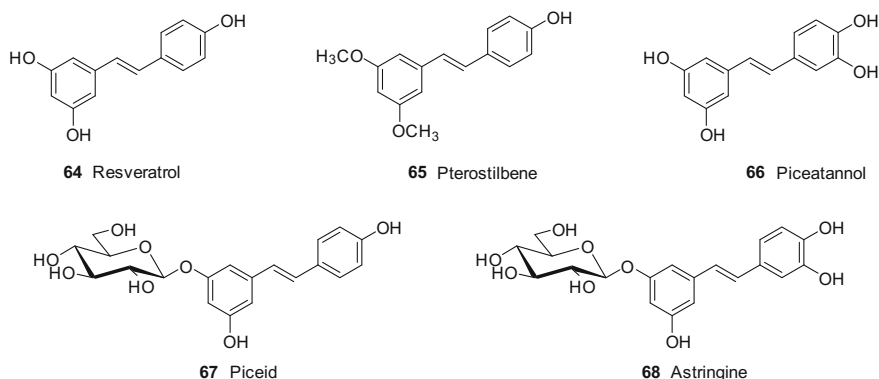
Fig. 1 Structure of natural stilbenes and stilbene glycosides

dimethoxystilbene, (**65**), piceantannol (*E*)-3,3′,4′,5-tetrahydroxystilbene, (**66**) (Fig. 1) and astringine (**68**) which biological activities have been reviewed [63].

Much attention has also been paid to stilbenes potential ability to protect from neurodegeneration. In fact, research points resveratrol as neuroprotective, not only due to the already mentioned antioxidant and anti-inflammatory activities, but also due to its ability to inhibit Aβ oligomeric cytotoxicity and to reduce neuronal cell death [58]. In a comparative study, the inhibitory activity of a series of stilbenes against Aβ (25–35) fibril formation was assessed. Both resveratrol **64** and piceid **67** effectively and dose dependently inhibited Aβ more extensively than curcumin [61].

Despite the promising activities of resveratrol and its glycoside piceid, their bioavailability in humans is quite poor [73, 81]. Indeed, the oral bioavailability of resveratrol is less than 1% as a consequence of quick and extensive metabolism, mainly through glucuronidation and sulfation, although it is not known whether resveratrol metabolites have a positive biological impact. The water-insolubility of stilbenes such as resveratrol, pterostilbene and piceatannol limits their further pharmacological exploitation. Literature shows a number of efforts to develop new stilbene analogues with higher solubility and bioavailability, and glycochemistry has definitely played a very relevant role. Glycosylation allows water-insoluble and unstable organic compounds to be converted into the corresponding water-soluble and stable compounds.

The synthesis of piceid itself was first described by Orsini and co-workers, in an attempt to obtain this natural product more efficiently (Scheme 12) [53]. The synthetic strategy aimed at building the stilbene skeleton first, by Wittig reaction of the aldehyde **69** and phosphonium ylide **70**, followed by desilylation. Methyl protected intermediary **71** was then glycosylated in the aqueous base under the phase transfer catalyst benzyltriethylammonium bromide (BTEAB) which afforded glucoside **73** in 32%. The diglucoside was also formed and isolated in 13% yield. Further deprotection (2 steps) afforded piceid [66] in 60% yield (13% overall yield)
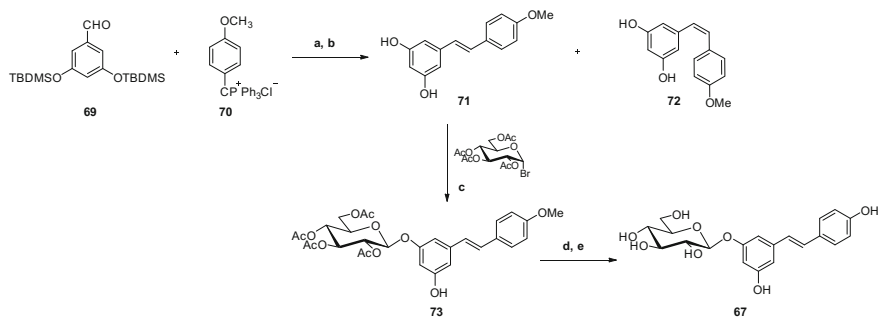
Scheme 12 Reagents and conditions: **a** BuLi, THF, −20 °C (98%, Z/E 2.3:1); **b** TBAF, THF, quant.; **c** BTEAB, NaOH, CHCl₃, 60 °C, 32%; **d** MeONa, MeOH, 25 °C, quant.; **e** EtSNa, DMF (60%) [53]

[53]. This methodology was also employed to the synthesis of other stilbene glycosides such as combretastatin analogues.

More recent efforts towards glycosylation of resveratrol take advantage of biotransformation for a simpler and more efficient synthesis. Glucosyltransferase *Pa*GT3 from *Phytolacca americana* expressed in *Bacillus subtilis* was used to convert resveratrol into its 3- and 4′-β-glucosides (**67** and **74**), as well as pterostilbene and piceatannol into their 4′-β-glucosides **75** and **76,** respectively, (Scheme 13). Glucosylation reactions were performed at 37 °C, in potassium phosphate buffer supplemented with UDP-glucose and enzyme. Although the procedure was not very effective towards piceid (12% yield), it afforded the 4′-β-glucosides in yields ranging from 50 to 76% [20].

Glycosylation of stilbenes was also performed using cultured cells from *P. Americana* and glucosyltransferase (*Pa*GT). This biocatalytic glycosylation
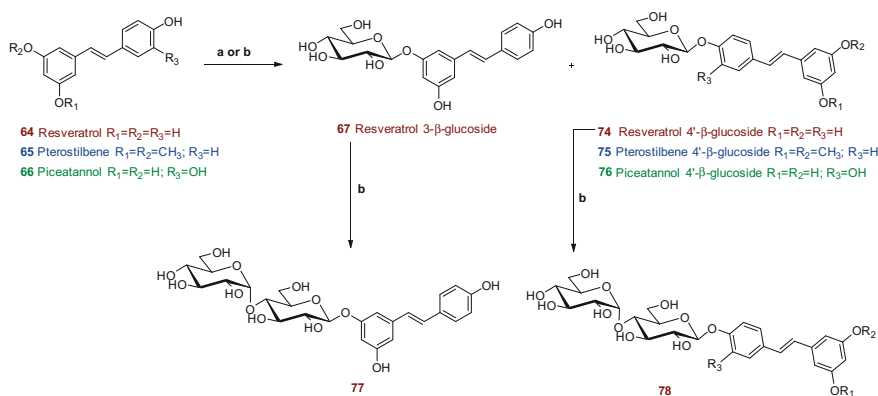
Scheme 13 Glycosylation of stilbenes resveratrol, pterostilbene and piceantannol. Reagents and conditions: **a** *Pa*GT3, potassium phosphate buffer (pH 7.2), UDP-glucose [20]. **b** cyclodextrin glucanotransferase (CGTase) [20, 65]

using cultured cells, in opposition to the direct use of the extracted enzyme, afforded resveratrol glucosides **67** and **74** in 35 and 22% yield, respectively, [65], favouring the formation of piceid, and 77% of piceatannol glucoside **76**, which proved to be the best substrate for this enzyme. Pterostilbene was only slightly converted into **75**.

In addition, resveratrol 3- and 4′-β-glucosides were further glycosylated using cyclodextrin glucanotransferase (CGTase) to afford resveratrol 3- and 4′-β-maltosides (**77** and **78**), respectively, with yields of 17 and 27%. The phospho-diesterase (PDE) inhibitory activity or resveratrol and pterostilbene was enhanced by glycosylation, since resveratrol 3- and 4′-β-glucosides, resveratrol 4′-β-maltoside and pterostilbene 4′-β-glucoside were better PDE inhibitors than their corresponding aglycone. This is particularly relevant as PDE inhibitors could be used in the treatment of neurodegenerative disorders such as Alzheimer's disease as they show potential to exert a neuroprotective role. Interestingly, piceatannol 4′-β-glucoside revealed also potent histamine release inhibitory activity (anti-allergic activity) [20, 65].

Enzymatic synthesis has also been employed in further glycosylation of the natural piceid, generating more soluble piceid glycosides such as **79**, which was obtained after incubation of piceid with maltosyltransferase from *Caldicellulosiruptor bescii* and maltotriose at 70 °C, in 18% yield. The water sol-ubility of maltosyl piceid **79** is 8540 and 1860 times greater than that of resveratrol and piceid, respectively [55]. Since the α-1,4-glycosidic linkages present in **79** can be easily hydrolyzed in vivo by α-glucosidase, this piceid glycoside could poten-tially be a resveratrol prodrug, with increased bioavailability and delayed meta-bolism [55]. Several piceid glucosides have also been obtained using cyclodextrin glucanotransferase from *Bacillus macerans* [49].

More recently, a sucrose phosphorylate from *Thermoanaerobacterium ther-mosaccharolyticum* (TtSPP) was engineered envisioning quantitative glycosylation of resveratrol in aqueous media (Scheme 14). Desmet and co-workers were able to identify a residue particularly important in the active site of TtSPP, which normally does not have a pocket deep enough for the binding resveratrol. Such residue, R134, was replaced by a smaller residue aiming at leaving an opening in the enzyme's closed conformation, enabling the accommodation of larger substrates. Indeed, the variant R134A, where arginine 134 was replaced by alanine, proved to have a reasonable affinity for resveratrol and to be very effective in the glycosylation of resveratrol at
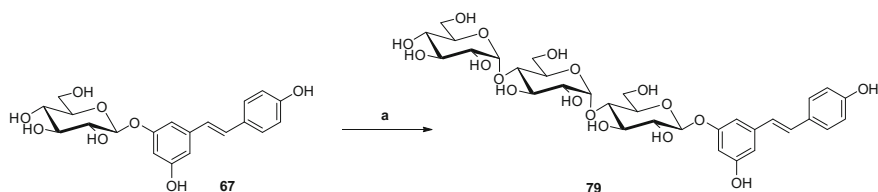
Scheme 14 Enzymatic synthesis of maltosyl piceid. **a** Maltosyltransferase (MTase), maltotriose [55]

gram scale, allowing the quantitative production of resveratrol 3-α-glucoside in an aqueous system, using sucrose as a cheap glycosyl donor [12].

Cyclodextrin glucanotransferase was also used to convert resveratrol and starch to α-glucosylated resveratrol products at 3-OH, at 4′-OH and at both 3-OH and 4′-OH, with increased water solubility when compared to that of resveratrol [69]. Interestingly, while the water solubility of piceid is 0.37 g/L, its alpha anomer presented solubility higher than 2 g/L [69]. Nevertheless, it would be interesting to compare the bioactivity of **82** with its anomer piceid, as configuration may play an important role in bioactivity and bioavailability, as demosntarted for the solubility. To the best of our knowledge, no bioactivity studies were conducted on resveratrol 3-α-glucoside so far.

Keeping in mind the challenge of resveratrol low water solubility, a new resveratrol analogue was developed, where glucosyl units were added to a resveratrol core with a succinate linker. It was speculated that the presence of glycosyl groups may also improve bioavailability by influencing phenomena taking place upstream of entry into erythrocytes, as occurs for quercetin 3-*O*-glucoside [3]. For the construction of the resveratrol analogue, a succinyl linker was firstly attached to the 3-hydroxy group of diacetoneglucose (**83**) (Scheme 15). The resulting succinyl ester (**85**) was used for the transesterification with resveratrol hydroxy groups using EDC. Hydrolysis of the isopropylidene protecting groups afforded the resveratrol derivative **87** in a 57% overall yield. This compound is relatively stable in acidic conditions but can be converted into resveratrol by blood esterases. Pharmacokinetics parameters were also improved, as its administration resulted in a blood concentration versus time curve shifted to longer times in comparison to resveratrol. This chemical transformation is particularly attractive as it may be employed in other bioactive polyphenols with poor water solubility. In addition, coating the hydroxy groups with sugar moieties can even make them more palatable [3]. Thus, it would be interesting to study this kind of modification in other polyphenols and, in addition to pharmacokinetics properties, study its influence in their organoleptic characteristics.

In summary, selected examples of glycosyl stilbenoids chemical and enzymatic synthesis were presented and described the reported benefits of stilbene
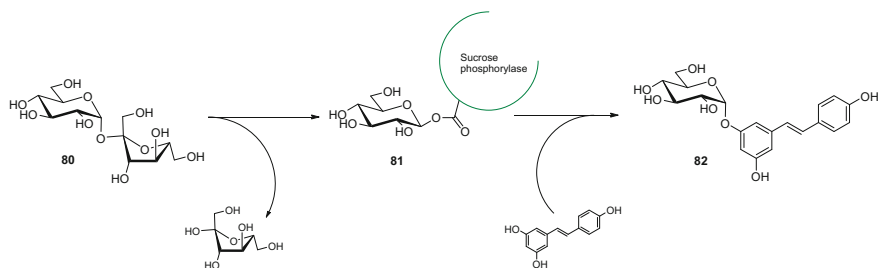
Scheme 15 Synthesis of resveratrol 3-α-glucoside with R134A variant of sucrose phosphorylase [12]

glycosylation for the usefulness of such biomolecular entities towards prevention of neurodegenerative impairments and related diseases.

4   Phenylethanoid and Phenylpropanoid Glycosides

A number of phenylethanoid and phenylpropanoid glycosides, either of natural or synthetic origin, have been described to possess neuroprotective activities, tackling both the amyloid cascade and the cholinergic system.

Acteoside (**95**) is a natural phenylpropanoid glycoside, also known as verbascoside, first isolated from the plant *Verbascum sinuatum* in the 1960s [28]. Meanwhile, a number of relevant bioactivities have been described, including its neuroprotective properties. Acteoside, isolated from *Orobanche minor,* strongly inhibits the aggregation of Aβ1-42, with an IC$_{50}$ of 8.9 μM [35] and protects against Aβ-induced cell injury by attenuation of reactive oxygen species production, by modulation of the apoptotic signal pathway through Bcl-2 family [75] and by upregulation of heme oxygenase-1 [74]. However, even before the mechanisms of neuroprotection were unravelled, the unsatisfactory extraction of this natural product from plant sources, prompted Sakuno and co-workers to develop the total synthesis of acteoside [28]. The synthetic strategy involves reaction of glucosyl chloride **88** (which was prepared from the peracetylated corresponding sugar) with the phenyletyl derivative **89** by the Koenigs-Knorr method in the presence of silver carbonate (Scheme 16). The presence of an acetyl group at position 2 directs to the formation of the 1,2-*trans* glycosidic bond through a neighbouring group participation mechanism [28]. A series of protection and deprotection steps to afford glycoside **91** is followed by the introduction of the caffeoyl moiety by esterification. Oxidative cleavage of the 3-*O*-allyl group and rhamnosylation, performed with 2,3,4-tri-*O*-acetyl-α-ʟ-rhamnopyranosyl trichloroacetimidate in the presence of boron trifluoride diethyl etherate gives the expected α-rhamnoside **94** in 73% yield. Finally, two considerable challenges lie on both the selective deacetylation over the cleavage of caffeoyl ester, and the selective removal of benzyl groups while keeping the double bond. Acetyl cleavage was consummate with methylamine in methanol (MeNH$_2$–MeOH), after which catalytic transfer hydrogenation of benzyl ethers using 1,4-cyclohexadiene as a hydrogen source afforded acteoside (**95**) successfully, in an overall yield of 3.5% [28].

More recently, an alternative and more efficient route towards phenylethanoid glycosides, such as acteoside, has been described [51], using a low substrate concentration and *N*-formylmorpholine modulated glycosylation for the construction of β- and α-glycosidic bonds. Interestingly, contrary to what was reported by Kawada and co-workers, the coupling of the β-glucoside **98** with the protected caffeic acid furnished not only the (*E*)-isomer of **99**, but also trace amounts of the
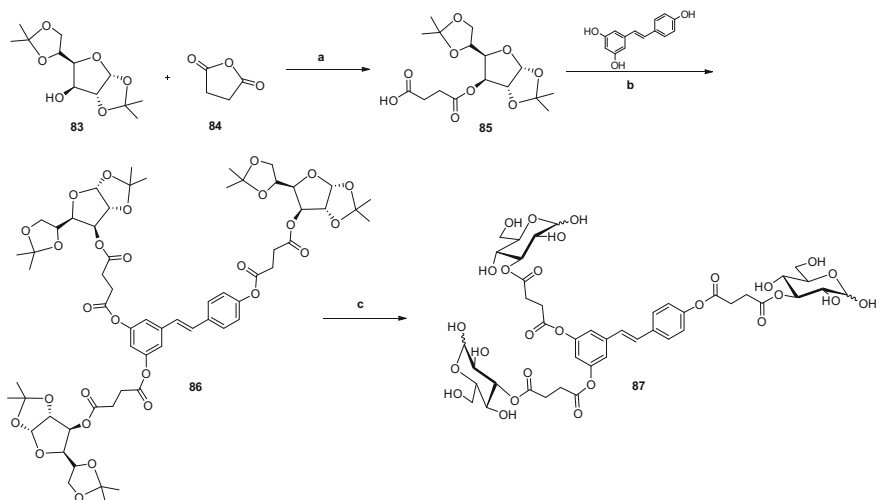
Scheme 16 Reagents and conditions: **a** Pyr. DMAP, rt (78%); **b** Pyr, DMAP, EDC, rt. (74%);
c TFA, rt (98%) [3]

(Z)-isomer. Nevertheless, acteoside was obtained in an overall yield of 10.8% (*E/Z*
12:1) (Scheme 17) [51].

It is well-established that enhancing cholinergic transmission by blocking the
activity of acetylcholinesterase (AChE) slows down the AD-associated decline in
behaviour and cognition. The natural phenylpropanoid diglycoside rosavin (**107**)
and its analogues (*E*)-3-phenylprop-2-en-1-yl β-D-xylopyranosyl-(1→6)-β-D-
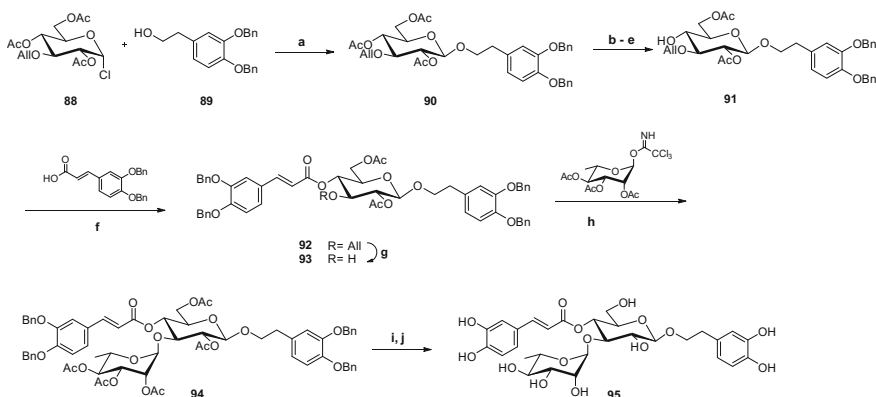
Scheme 17 Synthesis of acteoside by Kawada et al. Reagents and conditions: **a** Ag$_2$CO$_3$, DCE
(72%); **b** NaOMe, MeOH (94%); **c** TrCl, C$_5$H$_5$N (86%); **d** Ac$_2$O, C$_5$H$_5$N (92%); **e** AcOH, dioxane
(79%); **f** DCC, DMAP, DMAP·HCl, DCM (79%); **g** SeO$_2$, AcOH-dioxane (66%); **h** BF$_3$·Et$_2$O, DCM
(73%); **i** MeNH$_2$, MeOH-DCM (49%); **j** 1,4-ciclohexadiene, 5% Pd/C, DMF-EtOH (44%) [28]

glucopyranoside (**115**), (*E*)-3-(4-methoxyphenyl)prop-2-en-1-yl α-L-arabinopyranosyl-(1→6)-β-D-glucopyranoside (**116**) and (*E*)-3-phenylprop-2-en-1-yl α-L-rhamnopyranosyl-(1→6)-β-D-glucopyranoside (**117**) (Scheme 19) displayed a remarkable anti-AChE with IC_{50} values of 1.72, 3.71, 4.23, 2.05 μM, respectively [43]. Indeed, rosavin displayed the most potent AChE inhibition out of the natural compounds described so far. This natural product was firstly synthesized as shown in Scheme 18. The disaccharide **104** was first constructed by reaction of the glycosyl bromide **102** with the isopropylidene protected glucose **103**. After cleavage of the isopropylidene groups, acetylation and the introduction of the anomeric sulfanyl group, the glycosyl donor **105** was obtained. Activated by iodine, this donor reacted with the cinnamyl alcohol to afford the acetylated precursor **106**, which further deprotection gave rosavin (**107**).

Rosavin, along with its natural analogues **115–117**, was also synthesized by an alternative methodology, where the phenylpropanoid monoglycosides **113a** and **113b** were first synthesized, and then coupled with the appropriate glycosyl trichloroacetimidate (**118**, **119** or **120**), promoted by TMSOTf. Further deprotection afforded the natural rosavin analogues (Scheme 20). The same procedure was employed for the synthesis of a small library of phenylpropanoid glycosides, with derivatives incorporating substituted phenyl groups with F, Cl and Br, and varying the methoxy and hydroxy substitution patterns. However, none of the synthesized derivatives was as active as the natural diglycosides **107**, **115–117** [43]. Other methodologies for the synthesis of rosavin and its counterparts can be found in the literature, including the use of Mizoroki–Heck type reaction, involving the coupling of phenylboronic acid and allyl glycosides [30].

Other examples of powerful anti-AChE glycosides are the derivatives of the natural antidepressant helicid, synthesized starting from 4-hydroxybenzaldehyde,
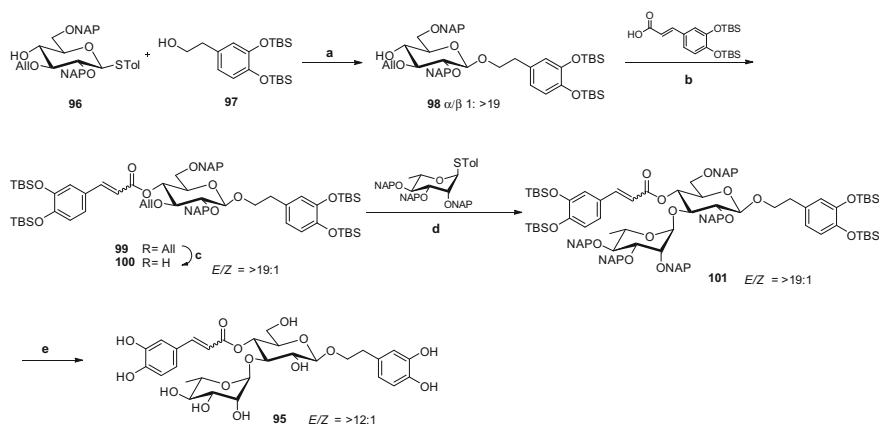
Scheme 18 Synthesis of acteoside by Mulani et al. **a** TMSOTf, −60 °C, 1:2:1 DCM-ACN-EtCN (63%); **b** NIS, TMSOTf, −40 °C, 1:2:1 DCM-ACN-EtCN (60%); **c** $PdCl_2$, NaOAc, AcOH, acetone (70%); **d** i. NFM; ii. NIS, TMSOTf, DCM (60%); **e** i. DDQ, 3:1 MeOH, CH2Cl2; ii. TREAT-HF, ET_3N, pyr. (two steps 68%) [51]
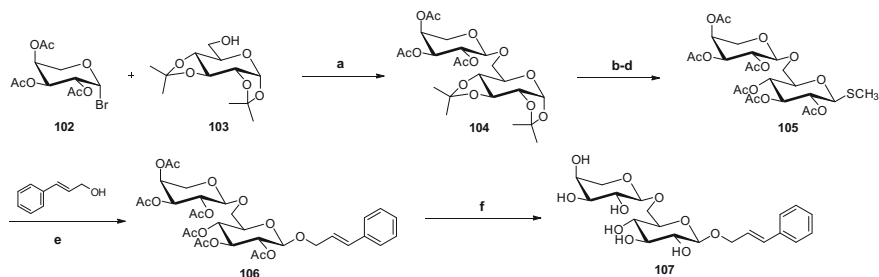
Scheme 19 **a** AgClO$_4$, CaCO$_3$, acetone (18%); **b** AcOH, 60 °C; **c** Ac$_2$O, HClO$_4$; **d** HSCH$_3$, BF$_3$-Et$_2$O, CHCl$_3$ (64%); **e** I$_2$, CHCl$_3$ (30%); **f** MeONa, MeOH (95%) [43]

followed by glycosylation, deprotection and condensation with amines, as depicted in Scheme 21 [80]. These transformations afforded noteworthy AChE inhibitors with IC$_{50}$ under 10 μM, three of them even under 0.55 μM. The synthetic approach was based on the reaction of glycosyl bromides with 4-hydroxybenzaldehyde in the presence of TBAB to afford the corresponding protected phenyl glycosides. Subsequent Zemplén deacetylation yielded the sugar-linked helicid analogues **122** and **124**. Schiff base derivative **127** was synthesized by reaction of **126** with methoxyamine (Scheme 21) [80]. Although an extensive library of helicid derivatives was obtained by this method, only the most active ones are depicted in Scheme 21. Interestingly, while helicid was not active up to 500 μM, its epimer at C-3 (**122**) presented an IC$_{50}$ of 0.45 μM. However, the most potent inhibitor is the 4-formylphenyl β-D-ribopyranoside (**124**). It exhibits the same configuration of carbons 2, 3 and 4 as helicid but its hydroxymethyl group is replaced by a hydrogen atom, presenting an IC$_{50}$ value of 0.20 μM on electric eel AChE, twice more active than galantamine. Also the Schiff base **127** has an IC$_{50}$ value of 0.49 μM. These results highlight the close correlation of the bioactivity with the sugar structure [80].

Structurally similar to the compounds discussed so far is also salidroside (**132**) (Scheme 22), a phenylpropanoid glycoside isolated from *Rhodiola* species that is one of the active principles responsible for plant antidepressant and anxiolytic activities. The low content of salidroside in *Rhodiola sachalinensis*, the unsustainable overexploitation of this species, and the need to fully exploit its potential clinical applications, have encouraged chemists to the develop a synthetic approach towards 2-(4-hydroxyphenyl)ethyl β-glucopyranoside. Various examples in the literature show the preference for the silver carbonate promoted glycosylation of tyrosol (**128**), which aromatic hydroxy group can be protected or unprotected, using peracetylated glucosyl bromide as a glycosyl donor [19, 66, 67]. In 2011, a multi-kilogram scale-up of salidroside was reported, featuring the selective acetylation of tyrosol aromatic hydroxy group in aqueous media, and affording the target natural glycoside in 72% yield (Scheme 22) [66].

Outstandingly, this natural glucoside protects neurons from glutamate-induced oxidative stress and apoptosis and was shown to be therapeutically effective against cognitive decline during ageing. Salidroside also intervenes in the amyloid cascade
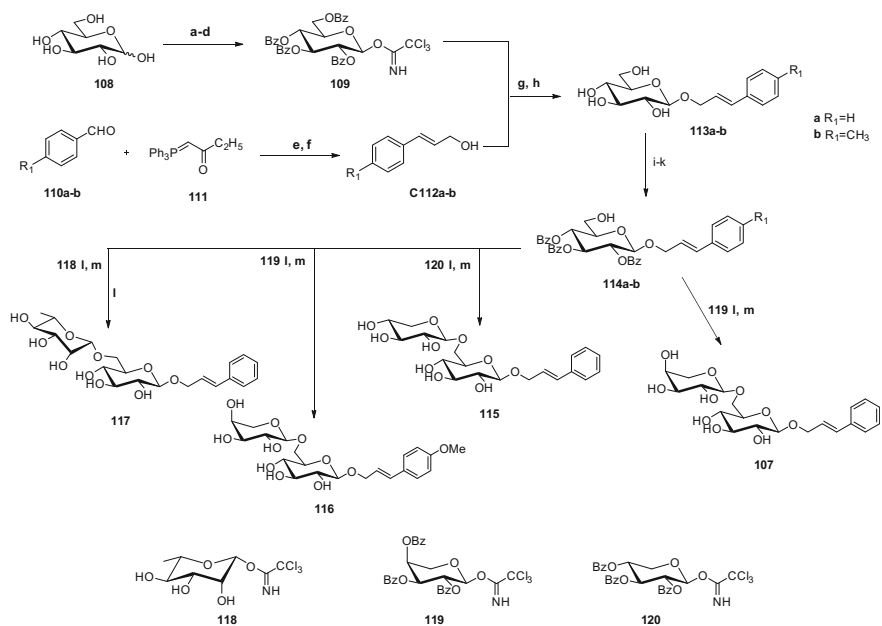
Scheme 20 Reagents and conditions: **a** BzCl/Py, rt; **b** HBr-AcOH/Ac$_2$O, rt; **c** NaI/H$_2$O/acetone, 30 °C (85% three steps); **d** DBU/CCl$_3$CN, rt (80%); **e** PhMe 80°C; **f** DIBAlH, PhMe, −20 °C; **g** TMSOTf, DCM, −20 °C; **h** NaOCH$_3$, MeOH, 0 °C (41, 61%, two steps); **i** TrCl, DMAP, TEA, DMF, MS 4 Å (56, 64%); **j** BzCl, Py (87, 90%); **k** 90% TFA, DCM (84, 86%); **l** TMSOTf, DCM, −20 °C; **m** NaOMe, MeOH (78–83%, two steps). (Yields in parentheses separated by comma are given for compounds type **a** and **b**, respectively) [43]
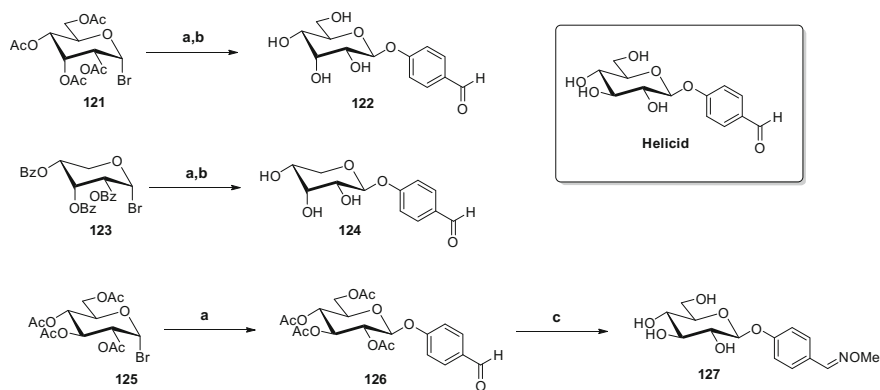
Scheme 21 Preparation of helicid derivatives. Reagents and conditions: **a** Preparation of helicid derivatives TBAB, NaOH, CHCl$_3$/H2O, 45 °C. **b** NaOMe, rt, 3 h; **c** amine, EtOH, reflux or 45 °C, 2–6 h [80]

events, as it protects against Aβ25–35-induced oxidative stress. In fact, pretreatment with salidroside noticeably attenuated Aβ25–35-induced loss of cell viability and apoptosis in a dose-dependent manner [88]. A fairly recent study also supports these findings and further attests the activity of this tyrosol glycoside by showing that it protects four different Drosophila models of AD against Aβ-induced neurotoxicity. The study also reveals that salidroside decreased Aβ levels and Aβ deposition in the fly's brain and ameliorated toxicity in Aβ-treated primary neuronal culture [85].

However, perhaps one of the most well-known phenylpropanoid derivatives with well-documented neuroprotective activities is curcumin (**133**, Scheme 22), an active ingredient in the spice turmeric consisting of two cinnamoyl units linked by a methylene group. Curcumin has been reported to act on several biochemical pathways associated with the onset and progression of AD. It disrupts amyloid-β and tau peptide aggregation, inhibits inflammation and protects against oxidative stress [50, 52]. However, its pharmaceutical use is restricted due to its poor water and plasma solubility and consequent low bioavailability [52, 56]. Considering that the addition of a sugar moiety would significantly increase the water/plasma solubility of the molecule while retaining all the characteristics of the curcumin pharmacophore, a clicked galactose–curcumin conjugate was developed using click-chemistry [13].

Such soluble "clicked" sugar conjugate of curcumin (SC) was synthesized as depicted in Scheme 22. The curcumin monoalkyne **134** was coupled with an acetyl-protected galactoside bearing an azide and, after removal of acetyl groups, a galactose-curcumin conjugate possessing a triazole-based linker was obtained. This non-toxic curcumin derivative is ca. 1000 times more soluble than curcumin in water, and exhibits enhanced ability to inhibit both amyloid-β and tau aggregation, at concentrations as low as 8 and 0.1 nM, respectively [13] (Scheme 23).
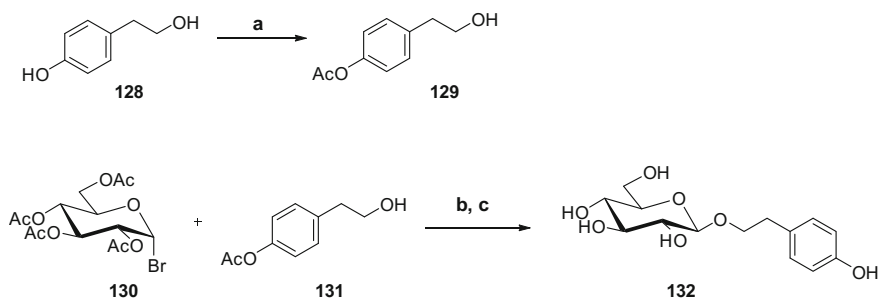
Scheme 22 Synthesis of salidroside. Reagents and conditions: **a** Ac$_2$O, NaOH, H2O (90%); **b** Ag$_2$CO$_3$, DCM, molecular sieves (4 Å); **c** NaOMe, MeOH (80%, two steps) [66]
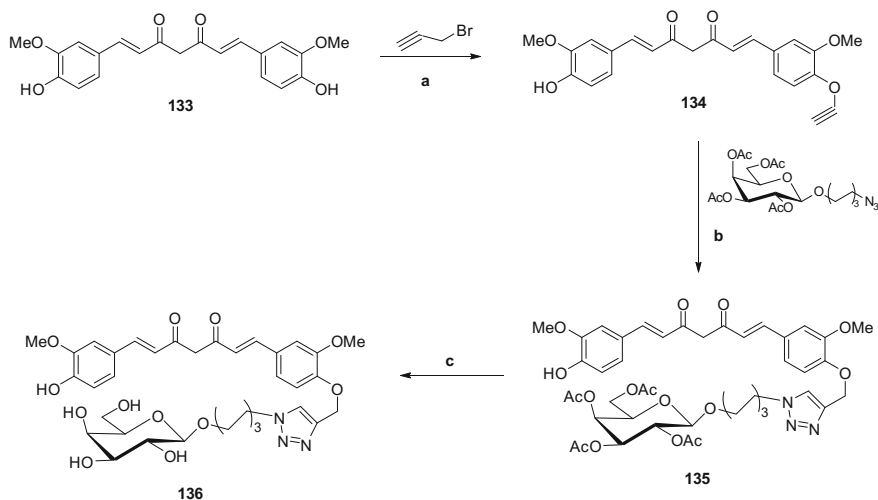
Scheme 23 **a** K_2CO_3, DMF, r.t; **b** i. $CuSO_4 5H_2O$, sodium ascorbate, t-BuOH/THF/H_2O, r.t (76%); **c** NaOMe, MeOH, r.t (70%) [13]

5 Conclusion

This chapter is devoted to highlight the biological importance of linking sugars to polyphenols and to the methodologies described for this purpose. Examples of polyphenol glycosides, that exhibit an increased neurodegenerative protective effect when compared to their aglycones, are given in this chapter. The role of sugar binding to improve polyphenol solubility and ameliorating its bioavailability is also clearly illustrated with examples. Chemical and enzymatic approaches to glycoside synthesis are described for various families of polyphenols, namely stilbenoids that include the well-known resveratrol, phenylethanoids and propanoids covering also the "dimeric analogue" curcumin and the flavonoids, covering only flavone glycosides. To flavone and isoflavone C-glycosylation is given a particular attention, given the relevance of the C–C bond, that is not hydrolytically cleaved, allowing *C*-glycosyl flavonoids to remain intact in the blood circulation following oral administration.

We really hope that this chapter will encourage chemists and biochemists to further investigate the role of sugar binding to polyphenols, not only to diversify and optimize coupling conditions but also to discover new biomolecular entities to effectively prevent neurodegenerative impairments with clinical applications.

References

1. Ancheeva E, Daletos G, Muharini R, Lin W, Teslov L, Proksch P (2015) Flavonoids from Stellaria nemorum and Stellaria holostea. Nat Prod Commun 10(3):437–440

2. Bavaresco L, Fregoni C, Cantù E, Trevisan M (1999) Stilbene compounds: from the grapevine to wine. Drugs Exp Clin Res 25(2–3):57–63

3. Biasutto L, Marotta E, Bradaschia A, Fallica M, Mattarei A, Garbisa S, Zoratti M, Paradisi C (2009) Soluble polyphenols: Synthesis and bioavailability of 3,4′,5-tri (α-D-glucose-3-O-succinyl) resveratrol. Bioorg Med Chem Lett 19(23):6721–6724

4. Cattaneo F, Costamagna M, Zampini I, Sayago J, Alberto M, Chamorro V, Pazos A, Thomas-Valdés S, Schmeda-Hirschmann G, Isla M (2016) Flour from Prosopis alba cotyledons: a natural source of nutrient and bioactive phytochemicals. Food Chem 208:89–96

5. Chen X, Cui L, Duan X, Ma B, Zhong D (2006) Pharmacokinetics and metabolism of the flavonoid scutellarin in humans after a single oral administration. Drug Metab Dispos 34 (8):1345–1352

6. Chen C, Li X, Gao P, Tu Y, Zhao M, Li J, Zhang S, Liang H (2015) Baicalin attenuates Alzheimer-like pathological changes and memory deficits induced by amyloid β1-42 protein. Metab Brain Dis 30(2):537–544

7. Choi J, Islam M, Ali M, Kim E, Kim Y, Jung H (2014) Effects of C-glycosylation on anti-diabetic, anti-Alzheimer's disease and anti-inflammatory potential of apigenin. Food Chem Toxicol 64:27–33

8. Choi J, Islam M, Ali M, Kim Y, Park H, Sohn H, Jung H (2014) The effects of C-glycosylation of luteolin on its antioxidant, anti-Alzheimer's disease, anti-diabetic, and anti-inflammatory activities. Arch Pharm Res 37(10):1354–1363

9. Courts F, Williamson G (2015) The occurrence, fate and biological activities of C-glycosyl flavonoids in the human diet. Crit Rev Food Sci Nutr 55(10):1352–1367

10. D'Amelio M, Cavallucci V, Middei S, Marchetti C, Pacioni S, Ferri A, Diamantini A, De Zio D, Carrara P, Battistini L, Moreno S, Bacci A, Ammassari-Teule M, Marie H, Cecconi F (2011) Caspase-3 triggers early synaptic dysfunction in a mouse model of Alzheimer's disease. Nat Neurosci 14(1):69–76

11. Ding H, Wang H, Zhao Y, Sun D, Zhai X (2015) Protective effects of baicalin on $A\beta_{1-42}$-induced learning and memory deficit, oxidative stress, and apoptosis in rat. Cell Mol Neurobiol 35(5):623–632

12. Dirks-Hofmeister ME, Verhaeghe T, De Winter K, Desmet T (2015) Creating space for large acceptors: rational biocatalyst design for resveratrol glycosylation in an aqueous system. Angew Chem Int Ed 54:9289–9292

13. Dolai S, Shi W, Corbo C, Sun C, Averick S, Obeysekera D, Farid M, Alonso A, Banerjee P, Raja K (2011) "Clicked" sugar − curcumin conjugate: modulator of amyloid-β and tau peptide aggregation at ultralow concentrations. Chem Neurosci 2:694–699

14. Fang M, Yuan Y, Lu J, Li H, Zhao M, Ling E, Wu C (2016) Scutellarin promotes microglia-mediated astrogliosis coupled with improved behavioral function in cerebral ischemia. Neurochem Int [epub ahead of print]

15. Fang M, Yuan Y, Rangarajan P, Lu J, Wu Y, Wang H, Wu C, Ling E (2015) Scutellarin regulates microglia-mediated TNC1 astrocytic reaction and astrogliosis in cerebral ischemia in the adult rats. BMC Neurosci 16:84

16. Furuta T, Kimura T, Kondo S, Mihara H, Wakimoto T, Nukaya H, Tsuji K, Tanaka K (2004) Concise total synthesis of flavone C-glycoside having potent anti-inflammatory activity. Tetrahedron 60:9375–9379

17. Gao C, Chen X, Zhong D (2011) Absorption and disposition of scutellarin in rats: a pharmacokinetic explanation for the high exposure of its isomeric metabolite. Drug Metab Dispos 39(11):2034–2044

18. Guimarães C, Oliveira D, Valdevite M, Saltoratto A, Pereira S, França S, Pereira A, Pereira P (2015) The glycosylated flavonoids vitexin, isovitexin, and quercetrin isolated from Serjania

erecta Radlk (Sapindaceae) leaves protect PC12 cells against amyloid-β25-35 peptide-induced toxicity. Food Chem Toxicol 86:88–94

19. Guo Y, Zhao Y, Zheng C, Meng Y, Yang Y (2010) Synthesis, biological activity of salidroside and its analogues. Chem Pharm Bull 58(12):1627–1629

20. Hamada H, Shimoda K, Shimizy N, Akagi M (2014) Synthesis of glycosides of resveratrol, pterostilbene, and piceatannol by glucosyltransferase from phytolacca americana expressed in bacillus subtilis and their chemopreventive activity against cancer, allergic, and alzheimer's diseases. Glycobiol Insights 4:1–6

21. Hamilton M, Caulfield J, Pickett J, Hooper A (2009) C-Glucosylflavonoid biosynthesis from 2-hydroxynaringenin by Desmodium uncinatum (Jacq.)(Fabaceae). Tetrahedron Lett 50:5656–5659

22. Hao B, Caulfield J, Hamilton M, Pickett J, Midega C, Khan Z, Wang J, Hooper A (2016) Biosynthesis of natural and novel C-glycosylflavones utilising recombinant Orzya sativa C-glycosyltransferase (OsCGT) and Desmodium incanum root proteins. Phytochemistry 125:73–87

23. Huang J, Weng W, Huang X, Ji Y, Chen E (2005) Pharmacokinetics of scutellarin and its aglycone conjugated metabolites in rats. Eur J Drug Metab Pharmacokinet 30(3):165–170

24. Ishikura Y, Tsuji K, Nukaya H (2002) Novel derivative of flavone C-glycoside and composition containing the same. WO 2004005296

25. Jesus A, Dias C, Matos A, Almeida R, Viana A, Marcelo F, Ribeiro R, Macedo M, Airoldi C, Nicotra F, Martins A, Cabrita E, Jiménez-Barbero J, Rauter A (2014) Exploiting the therapeutic potential of 8-β-d-glucopyranosylgenistein: synthesis, antidiabetic activity, and molecular interaction with islet amyloid polypeptide and amyloid β-peptide (1–42). J Med Chem 57(22):9463–9476

26. Jung H, Karki S, Kim J, Choi J (2015) BACE1 and cholinesterase inhibitory activities of Nelumbo nucifera embryos. Acta Pharm Res 38(6):1178–1187

27. Jung M, Park M (2007) Acetylcholinesterase inhibition by flavonoids from Agrimonia pilosa. Molecules 12(9):2130–2139

28. Kawada T, Asano R, Hayashida S, Sakuno T (1999) Total synthesis of the phenylpropanoid glycoside, acteoside. J Org Chem 64:9268–9271

29. Kerscher F, Franz G (1986) Biosynthesis of vitexin and isovitexin: enzymatic synthesis of the C-glucosylflavones vitexin and isovitexin with an enzyme preparation from Fagopyrum esculentum M. seedlings. Z. Naturforsch. 42c:519–524

30. Kishida M, Akita H (2005) Synthesis of rosavin and its analogues based on a mizoroki-heck type reaction. Tetrahedron Asymmetry 16:2625–2630

31. Koekkoek P, Kappelle L, Van den Berg E, Rutten G, Biessels G (2015) Cognitive function in patients with diabetes mellitus: guidance for daily care. Lancet Neurol 14(3):329–340

32. Komentani T, Kondo H, Fujimori Y (1988) Boron trifluoride-catalyzed rearrangement of 2-aryloxytetrahydropyrans: a new entry to C-arylglycosidation. Synthesis 1005–1007

33. Kumazawa T, Kimura T, Matsuba S, Sato S, Onodera J (2011) Synthesis of 8-C-glucosylflavones. Carbohydr Res 334(3):183–193

34. Kumazawa T, Minatogawa T, Matsuba S, Sato S, Onodera J (2000) An effective synthesis of isoorientin: the regioselective synthesis of a 6-C-glucosylflavone. Carbohydr Res 329(3):507–513

35. Kurisu M, Miyamae Y, Murakami K, Han J, Isoda H, Irie K, Shigemori H (2013) Inhibition of amyloid β aggregation by acteoside, a phenylethanoid glycoside. Biosci Biotechnol Biochem 77(6):1329–1332

36. Ladiwala A, Mora-Pale M, Lin J, Bale S, Fishman Z, Dordick J, Tessier P (2011) Polyphenolic glycosides and aglycones utilize opposing pathways to selectively remodel and inactivate toxic oligomers of amyloid β. ChemBioChem 12(11):1749–1758

37. Law B, Ling A, Koh R, Chye S, Wong Y (2014) Neuroprotective effects of orientin on hydrogen peroxide–induced apoptosis in SH-SY5Y cells. Mol Med Rep 9(3):947–954

38. Lee D, Zhang W, Karnati V (2003) Total synthesis of puerarin, an isoflavone C-glycoside. Tetrahedron Lett 44(36):6857–6859

39. Li J, Wang G, Liu J, Zhou L, Dong M, Wang R, Li X, Li X, Lin C, Niu Y (2010) Puerarin attenuates amyloid-beta-induced cognitive impairment through suppression of apoptosis in rat hippocampus in vivo. Eur J Pharmacol 649(1–3):195–201

40. Li N, Shen M, Wang Z, Tang Y, Shi Z, Fu Y, Shi Q, Tang H, Duan J (2013) Design, synthesis and biological evaluation of glucose-containing scutellarein derivatives as neuroprotective agents based on metabolic mechanism of scutellarin in vivo. Bioorg Med Chem Lett 23 (1):102–106

41. Lin F, Xie B, Cai F, Wu G (2012) Protective effect of Puerarin on β-amyloid-induced neurotoxicity in rat hippocampal neurons. Arzneimittelforschung 62(4):187–193

42. Liu X, Mo Y, Gong J, Li Z, Peng H, Chen J, Wang Q, Ke Z, Xie J (2016) Puerarin ameliorates cognitive deficits in streptozotocin-induced diabetic rats. Metab Brain Dis 31 (2):417–423

43. Li XD, Kang ST, Li X, Wang JH (2011) Synthesis of some Phenylpropanoid Glycosides (PPGs) and their acetylcholinesterase/xanthine oxidase inhibitory activities. Molecules 16:3580–3596

44. Li Z, Shangguan Z, Liu Y, Wang J, Li X, Yang S, Liu S (2014) Puerarin protects pancreatic β-cell survival via PI3 K/Akt signaling pathway. J Mol Endocrinol 53(1):71–79

45. Li Y, Biao Y, Sun J, Wang R (2015) Efficient synthesis of baicalin and its analogs. Tetrahedron Lett 56:3816–3819

46. Lozoya X, Meckes M, Abou-Zaid M, Tortoriello J, Nozzolillo C, Arnason J (1994) Quercetin glycosides in Psidium guajava L. leaves and determination of a spasmolytic principle. Arch Med Res 25(1):11–15

47. Mahling J, Jung K, Schmidt R (1995) Synthesis of flavone C-glycosides vitexin, isovitexin and isoembigenin. Leibigs Ann Chem 461–466

48. Masibo M, He Q (2008) Major mango polyphenols and their potential significance to human health. CRFSFS 7(4):309–319

49. Mathew S, Hedström M, Adlercreutz P (2012) Enzymatic synthesis of piceid glycosides by cyclodextrin glucanotransferase. Process Biochem 47:528–532

50. Mishra S, Palanivelu K (2006) The effect of curcumin (turmeric) on Alzheimer's disease: An overview. Life Sci 78(18):2081–2087

51. Mulani SK, Guh JH, Mong KKT (2014) general synthetic strategy and the anti-proliferation properties on prostate cancer cell lines for natural phenylethanoid glycosides. Org Biomol Chem 12:2926–2937

52. Noorafshan A, Ashkani-Esfahani S (2013) A review of therapeutic effects of curcumin. Curr Pharm Des 19(11):2032–2046

53. Orsini F, Pelozzoni F, Bellini B, Miglierini G (1997) Synthesis of biologically active polyphenolic glycosides (combretastatin and resveratrol series). Carb Res 301:95–109

54. Pan Z, Feng T, Shan L, Cai B, Chu W, Niu H, Lu Y, Yang B (2008) Scutellarin-induced endothelium-independent relaxation in rat aorta. Phytother Res 22(11):1428–1433

55. Park H, Kim J, Choi KH, Hwand S, Yang SJ, Baek NI, Cha J (2012) Enzymatic synthesis of piceid glucosides using maltosyltransferase from caldicellusiruptor bescii DSM 6725. J Agric Food Chem 60:8183–8189

56. Prasad S, Tyagi AK, Aggarwal BB (2014) Recent developments in delivery, bioavailability, absorption and metabolism of curcumin: the golden pigment from golden spice. Cancer Res Treat 46(1):2–18

57. Qiao W, Zhao C, Qin N, Zhai H, Duan H (2011) Identification of trans-tiliroside as active principle with anti-hyperglycemic, anti-hyperlipidemic and antioxidant effects from Potentilla chinesis. J Ethnopharmacol 135(2):515–521

58. Rege SD, Geetha T, Griffin GD, Broderick TL, Babu JR (2014) Neuroprotective effects of resveratrol in Alzheimer disease pathology. Front Aging Neurosci 6(1):1–12

59. Ren G, Hou J, Fang Q, Sun H, Liu X, Zhang L, Wang P (2012) Synthesis of flavonol 3-O-glycoside by UGT78D1. Glycoconj J 29(5–6):425–432

60. Renaud S, Lorgeril M (1992) Wine, alcohol, platelets, and the French paradox for coronary heart disease. Lancet 339(8808):1523–1526

61. Rivière C, Richard T, Quentin L, Krisa S, Mérillon JM, Monti JP (2007) Inhibitory activity of stilbenes on Alzheimer's β-amyloid fibrils in vitro. Bioorg Med Chem 15:1160–1167
62. Robertson A, Robinson R (1926) Experiments on the synthesis of anthocyanins. Part I. J Chem Soc 129:1713–1720
63. Roupe KA, Remsberg CM, Yáñez JA, Davies NM (2006) Pharmacometrics of stilbenes: guing towards the Clinic. Curr Clin Pharmacol 1:81–101
64. Santos R, Jesus A, Caio J, Rauter A (2011) Fries-type reactions for the C-glycosylation of phenols. Curr Org Chem 15:128–148
65. Sato D, Shimizu N, Shimizu Y, Akagi M, Eshita Y, Ozaki S, Nakajima N, Ishihara K, Sato D, Shimizu N, Shimizu Y, Akagi M, Eshita Y, Ozaki S-I, Nakajima N, Ishihara K, Masuoka N, Hamada H, Shimoda K, Kubota N (2014) Synthesis of glycosides of resveratrol, pterostilbene, and piceatannol, and their anti-oxidant, anti-allergic, and neuroprotective activities. Biosci Biotech Biochem 78:1123–1128
66. Shi T, Chen H, Jing L, Liu X, Sun X, Jiang R (2011) Development of a kilogram-scale synthesis of salidroside and its analogs. Synth Commun 41:2594–2600
67. Shi LF, Cai Z, Yao B (2004) Preparation of salidroside derivatives as antioxidants. *Faming Zhuanli Shenqing Gongkai Shuomingshu*. CN Patent 1475492 A 20040218
68. Tang H, Tang Y, Li N, Shi Q, Guo J, Shang E, Duan J (2014) Neuroprotective effects of scutellarin and scutellarein on repeatedly cerebral ischemia-reperfusion in rats. Pharmacol Biochem Behav 118:51–59
69. Torres P, Poveda A, Jimenez-Barbero J, Parra JL, Comelles F, Ballesteros AO, Plou FJ (2011) Enzymatic Synthesis of α-Glucosides of Resveratrol with surfactant activity. Adv Synth Catal 353:1077–1086
70. Uriarte-Pueyo I, Calvo M (2011) Flavonoids as acetylcholinesterase inhibitors. Curr Med Chem 18(34):5289–5302
71. Velagapudi R, Aderogba M, Olajide O (2014) Tiliroside, a dietary glycosidic flavonoid, inhibits TRAF-6/NF-κB/p38-mediated neuroinflammation in activated BV2 microglia. Biochim Biophys Acta 1840(12):3311–3319
72. Vermes B, Chari V, Wagner H (1981) Structure elucidation and synthesis of flavonol acylglycosides. The synthesis of tiliroside. Helv Chim Acta 64(6):1964–1967
73. Walle T (2011) Bioavailability of resveratrol. Ann NY Acad Sci 1215:9–15
74. Wang HQ, Xu YX, Zhu CK (2012) Upregulation of heme oxygenase-1 by acteoside through ERK and PI3 K/Akt pathway confer neuroprotection against beta-amyloid-induced neurotoxicity. Neurotoxic Res 21:368–378
75. Wang HQ, Xu YX, Yan J, Zhao XY, Sun XB, Zhang YP, Guo JC, Zhu CK (2009) Acteoside protects human neuroblastoma SH-SY5Y cells against beta-amyloid-induced cell injury. Brain Res 1283:139–147
76. Wang C, Li J, Zhang L, Zhang X, Yu S, Liang X, Ding F, Wang Z (2013) Isoquercetin protects cortical neurons from oxygen-glucose deprivation-reperfusion induced injury via suppression of TLR4-NF-κB signal pathway. Neurochem Int 63(8):741–749
77. Wang C, Shi Y, Tang M, Zhang X, Gu Y, Liang X, Wang Z, Ding F (2016) Isoquercetin ameliorates cerebral impairment in focal ischemia through anti-oxidative, anti-inflammatory, and anti-apoptotic effects in primary culture of rat hippocampal neurons and hippocampal CA1 region of rats. Mol Neurobiol [epub ahead of print]
78. Wang C, Xie N, Zhang H, Li Y, Wang Y (2014) Puerarin protects against β-amyloid-induced microglia apoptosis via a PI3 K-dependent signalling pathway. Neurochem Res 39(11):2189–2196
79. Wang Y, Zhen Y, Wu X, Jiang Q, Li X, Chen Z, Zhang G, Dong L (2015) Vitexin protects brain against ischemia/reperfusion injury via modulating mitogen-activated protein kinase and apoptosis signalling in mice. Phytomedicine 22(3):379–384
80. Wen H, Lin C, Que L, Ge H, Ma L, Cao R, Wan Y, Peng W, Wang Z, Song H (2008) Synthesis and biological evaluation of helicid analogues as novel acetylcholinesterase inhibitors. Eur J Med Chem 43:166–173

81. Wenzel E, Somoza V (2005) Metabolism and bioavailability of trans-resveratrol. Mol Nutr Food Res 49:472–481
82. Wu K, Liang T, Duan X, Xu L, Zhang K, Li R (2013) Anti-diabetic effects of puerarin, isolated from Pueraria lobata (Willd.), on streptozotocin-diabetogenic mice through promoting insulin expression and ameliorating metabolic function. Food Chem Toxicol 60:341–347
83. Xiong J, Wang C, Chen H, Hu Y, Tian L, Pan J, Geng M (2014) Aβ-induced microglial cell activation is inhibited by baicalin through the JAK2/STAT3 signaling pathway. Int J Neurosci 124(8):609–620
84. Yu L, Wang S, Chen X, Yang H, Li X, Xu Y, Zhu X (2015) Orientin alleviates cognitive deficits and oxidative stress in Aβ1-42-induced mouse model of Alzheimer's disease. Life Sci 121:104–109
85. Zhang B, Wang Y, Li H, Xiong R, Zhao Z, Chu X, Li Q, Sun S, Chen S (2016) Neuroprotective effects of salidroside through P13 K/Akt pathway activation in Alzheimer's disease models. Drug Des Dev Ther 10:1335–1343
86. Zhang H, Liu Y, Wang H, Xu J, Hu H (2008) Puerarin protects PC12 cells against beta-amyloid-induced cell injury. Cell Biol Int 32(10):1230–1237
87. Zhang J, Guo W, Tian B, Sun M, Li H, Zhou L, Liu X (2015) Puerarin attenuates cognitive dysfunction and oxidative stress in vascular dementia rats induced by chronic ischemia. Int J Clin Exp Pathol 8(5):4695–4704
88. Zhang L, Yu H, Zhao X, Lin X, Tan C, Cao G, Wang Z (2010) Neuroprotective effects of salidroside against beta-amyloid-induced oxidative stress in SH-SY5Y human neuroblastoma cells. Neurochem Int 57(5):547–555
89. Zhao S, Yang W, Jin H, Ma K, Feng G (2015) Puerarin attenuates learning and memory impairments and inhibits oxidative stress in STZ-induced SAD mice. Neurotoxicology 51:166–171
90. Zhou Y, Xie N, Li L2, Zou Y, Zhang X, Dong M (2014) Puerarin alleviates cognitive impairment and oxidative stress in APP/PS1 transgenic mice. Int J Neuropsychopharmacol 17 (4):635–644
91. Zhu J, Choi R, Li J, Xie H, Bi C, Cheung A, Dong T, Jiang Z, Chen J, Tsim K (2009) Estrogenic and neuroprotective properties of scutellarin from Erigeron breviscapus: a drug against postmenopausal symptoms and Alzheimer's disease. Planta Med 75(14):1489–1493
92. Żmijewski M, Sokół-Łętowska A, Pejcz E, Orzeł D (2015) Antioxidant activity of rye bread enriched with milled buckwheat groats fractions. Rocz Panstw Zakl Hig 66(2):115–121

Probing for *Trypanosoma cruzi* Cell Surface Glycobiomarkers for the Diagnosis and Follow-Up of Chemotherapy of Chagas Disease

Nathaniel S. Schocker, Susana Portillo, Roger A. Ashmus, Carlos R.N. Brito, Igor E. Silva, Yanira Cordero Mendoza, Alexandre F. Marques, Erika Y. Monroy, Andrew Pardo, Luis Izquierdo, Montserrat Gállego, Joaquim Gascon, Igor C. Almeida and Katja Michael

Abstract *Trypanosoma cruzi* is a protozoan parasite that causes Chagas disease in humans. Linear and branched *O*-glycans with non-reducing, terminal α-galactosyl (α-Gal) glycotopes located on cell surface glycosylphosphatidylinositol (GPI)-anchored mucins of the infective trypomastigote form of the parasite are foreign to humans and elicit high levels of anti-α-Gal antibodies in Chagas disease patients (Ch anti-α-Gal antibodies). These antibodies have the capability to lyse the parasite in a complement-dependent or -independent manner. Ch anti-α-Gal antibodies have a considerably higher reactivity to the parasitic surface α-Gal glycotopes than the normal human serum (NHS) anti-α-Gal antibodies, which are present in every healthy human being. A series of ten mercaptopropyl saccharides with α-Gal moieties at the non-reducing end, all connected to another galactose unit, and five non-α-Gal-containing glycan controls were synthesized, and conjugated to maleimide-derivatized bovine serum albumin. This produced neoglycoproteins (NGPs), which were assembled into glycoarrays for the interrogation with sera of chronic Chagas disease patients and healthy individuals using chemiluminescent

N.S. Schocker · R.A. Ashmus · E.Y. Monroy · A. Pardo · K. Michael (✉)
Department of Chemistry, University of Texas at El Paso, El Paso, TX 79968, USA
e-mail: kmichael@utep.edu

S. Portillo · C.R.N. Brito · I.E. Silva · Y.C. Mendoza · A.F. Marques · I.C. Almeida (✉)
Border Biomedical Research Center, Department of Biological Sciences, University of Texas at El Paso, El Paso, TX 79968, USA
e-mail: icalmeida@utep.edu

C.R.N. Brito · A.F. Marques
Department of Parasitology, Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, MG, Brazil

L. Izquierdo · M. Gállego · J. Gascon
Barcelona Institute for Global Health (ISGlobal), Hospital Clinic—Universitat de Barcelona, 08036 Barcelona, Spain

enzyme-linked immunosorbent assay (CL-ELISA). This study identified the terminal Galα(1,3)Galβ disaccharide as an immunodominant *T. cruzi* glycotope and bio-marker, which shows a considerable binding differential between Ch and NHS anti-α-Gal antibodies. Therefore, this glycotope is suitable for the diagnosis of Chagas disease, and could also be potentially used for follow-up studies for the effectiveness of chemotherapy in Chagas disease patients.

1 Introduction

Chagas disease is an infectious disease caused by the protozoan parasite *Trypanosoma cruzi*, which is transmitted by blood-sucking insect vectors of the Reduviidae family, popularly known as "kissing bugs". Transmission can also occur by blood transfusion, organ transplantation, ingestion of tainted foods and liquids, and by the congenital route. Currently, approximately 8–10 million people are infected, and approximately 12,000 die every year, mostly as a consequence of cardiomyopathy [1, 2]. Chagas disease is endemic to South and Central America; however, due to the migration of many thousands of chronically infected, asymptomatic individuals from endemic areas, Chagas disease now exists in non-endemic regions such as the U.S., Europe, Australia, New Zealand, and Japan [2–5].

The cell surface of *T. cruzi* has heavily *O*-glycosylated glycosylphosphatidyl-inositol (GPI)-anchored mucins, which are the major constituents of a particularly dense glycocalyx [6–8]. Unlike the *O*-glycans of GPI-mucins from the insect-derived developmental stages of the parasite, that have been characterized in several strains and genotypes, the structural details of the *O*-glycans of mucins from the mammal-dwelling trypomastigote form (tGPI-mucins) are mostly unknown [6, 9, 10]. However, through partial structural analysis and immunoassays it is known that the majority of these tGPI-mucin *O*-glycans contain highly immunogenic, non-reducing terminal α-galactopyranosyl residues [11], here abbreviated as α-Gal, αGal, or Galα, omitting the explicit designator "*p*" for pyranose. Interestingly, these terminal α-Gal residues are completely absent in GPI-mucins derived from the insect vector-dwelling epimastigote and metacyclic trypomastigote forms or stages [6–10]. The αGal residues are partial structures of most likely several immuno-dominant glycotopes and are recognized by the highly abundant, protective anti-α-Gal antibodies present in the sera of patients in the acute and chronic phases of Chagas disease [11, 12]. These antibodies are responsible for controlling the parasitemia in both stages of the infection [11–13]. The *O*-glycosylation of the *T. cruzi* mucins is a posttranslational modification in which α-*N*-acetylglucosamine (GlcNAcα) is added to a threonine side chain by the UDP-GlcNAc:polypeptide α-*N*-acetylglucosaminyltransferase in the Golgi apparatus [14]. This α-GlcNAc moiety is heavily 4,6-di-*O*-substituted, albeit 4-*O* monosubstitution also exists [11–16]. In addition, 2-*O*-substituted Gal, 3-*O*-substituted Gal, 4-*O*-substituted Gal, 6-*O*-substituted Gal, and 2,6-di-*O*-substituted Gal motives also exist, indicating that

tGPI-mucin *O*-glycans are galactose-rich and predominantly branched [11]. Nevertheless, the only tGPI-mucin *O*-glycan that has been fully characterized to date is the linear trisaccharide Galα(1,3)Galβ(1,4)GlcNAcα. It is strongly recognized by Ch anti-α-Gal Abs, but only weakly by anti-α-Gal Abs from healthy individuals [normal human serum (NHS) anti-α-Gal Abs] [11], which are produced mainly against Gram-negative enterobacteria of the human flora [17]. These enterobacteria (e.g., *E. coli*, *Serratia* spp., *Enterobacter* spp., *Klebsiella* spp., *Salmonella* spp., *Citrobacter* spp., and *Shigella* spp.) have a number of different non-reducing, terminal αGal-linked glycans, mostly Galα(1,2)-R, Galα(1,4)-R and Galα(1,6)-R (where R is the remaining side chain or core glycan) on the lipopolysaccharide (LPS) core oligosaccharides or *O*-antigens [18]. The glycotope Galα(1,3)Galβ(1,4)GlcNAcα has so far not been reported in enterobacteria.

Despite the existence of intraspecies polymorphism in the *O*-glycans of the GPI-mucins, the expression of highly immunogenic, non-reducing terminal αGal residues seems to be highly conserved in tGPI-mucins from at least four major *T. cruzi* genotypes or discrete typing units (DTUs) known to infect humans (i.e., TcI, TcII, TcV, and TcVI) [11, 19–21]. This is supported by numerous studies showing the abundant presence of high levels of protective Ch anti-α-Gal Abs in patients from different endemic and nonendemic regions [11, 12, 19, 22–30].

2 Results and Discussion

Here we present the identification of an immunodominant glycotope present on the *T. cruzi* cell surface that is strongly recognized by anti-α-Gal antibodies from chronic Chagas disease (CCD) patients. In order to identify potential *T. cruzi* αGal-containing glycotopes, we synthesized a biased library of ten glycans consisting of mono-, di-, and trisaccharides with terminal αGal moieties based on the partial structural information available for tGPI-mucin *O*-glycans [11]. The synthetic glycans were conjugated to bovine serum albumin (BSA) to produce neo-glycoproteins (NGPs), which were assembled into glycoarrays and interrogated with pooled sera of *T. cruzi*-infected and healthy individuals using chemiluminescent enzyme-linked immunosorbent-assay (CL-ELISA) [24]. We reasoned that with a glycan library in hand, polyclonal anti-α-Gal Abs from CCD patients could be used as probes for the identification of immunodominant *T. cruzi* glycotopes. Antibody recognition of certain saccharides would suggest that these saccharides are immunogenic glycotopes or glycotope partial structures that exist on the cell surface of the infective trypomastigote form that lives in the human host. Due to the structural diversity of the trypomastigote cell surface *O*-glycans, several glycans with αGal moieties differently connected to another underlying sugar, most likely another galactose unit, could potentially be identified as immunogenic glycotopes.

For the synthesis of a potential αGal-containing library, three factors had to be considered: (a) the size and connectivity of the saccharide targets; (b) a suitable linker allowing for the conjugation to BSA; and (c) a versatile synthetic strategy
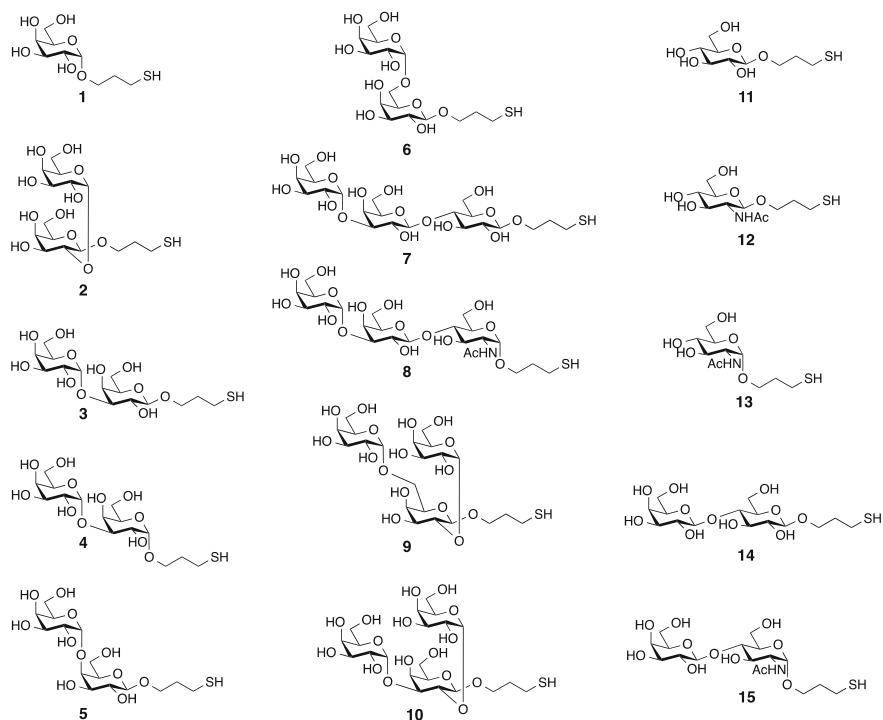
Fig. 1 A synthetic library of saccharides with mercaptopropyl linkers: ten saccharides with non-reducing αGal (**1–10**), and five saccharides that lack αGal (**11–15**) [32, 33]

allowing for the use of common precursors. We synthesized a total of ten αGal-containing mercaptopropyl saccharides (**1–10**, Fig. 1) suitable for conjugation to commercially available maleimide-derivatized BSA. In order to shine light on whether Ch anti-α-Gal Abs can recognize a single monosaccharide, mercaptopropyl Galα (**1**) [31, 32] was included in the study. The mercaptopropyl glycosides of the following five αGal-containing disaccharides were also included: Galα(1,2)Galβ (**2**) [32]; Galα(1,3)Galβ (**3**) [32]; Galα(1,3)Galα (**4**); Galα(1,4)Galβ (**5**) [32]; and Galα (1,6)Galβ (**6**) [32]. In addition, the mercaptopropyl glycosides of Galα(1,3)Galβ (1,4)Glcα (**7**) [32], and the linear tGPI-mucin trisaccharide Galα(1,3)Galβ(1,4) GlcNAcα (**8**) [33] were selected as targets. Trisaccharide **8** serves as a positive control, and trisaccharide **7** may shine light on the importance of the third sugar at the reducing end for antibody recognition. In addition, the trisaccharides Galα(1,2) [Galα(1,6)]Galβ (**9**) and Galα(1,2)[Galα(1,3)]Galβ (**10**) were included because they represent branched *O*-glycans with two terminal αGal units. Lastly, six different putative negative controls that lack terminal αGal units were also included in the library: the mercaptopropyl glycosides of monosaccharides Galβ (**11**) [32], GlcNAcβ (**12**) [33], GlcNAcα (**13**) [33], the disaccharides Galβ(1,4)Glcβ (**14**) [32], and Galβ(1,4)GlcNAcα (**15**) [33], as well as cysteine [32, 33]. Including compounds **13** and **15** in the study will provide information on whether Ch anti-α-Gal

antibodies have the ability to recognize the reducing end mono and disaccharide partial structures of the known *T. cruzi* tGPI mucin glycan Galα(1,3)Galβ(1,4) GlcNAcα. Lastly, including both anomers of GlcNAc (compounds **12** and **13**) may provide information on the importance of the configuration at the reducing end of Galα(1,3)Galβ(1,4)GlcNAcα.

The assembly of the glycans into a glycoarray requires their immobilization in microtiter plate wells. This can be accomplished by conjugation of the glycan to a protein, which adheres to Nunc MaxiSorp® microtiter plate wells made from polystyrene by non-covalent interactions. Initially, we chose maleimide-derivatized keyhole limpet hemocyanin (KLH) for the conjugation due to its high-loading capacity, but this protein showed poor water solubility and made accurate microplate-well loading impossible. Therefore, we decided to conjugate all gly-cosides (**1–15**) to BSA instead, which has superior solubility properties. The mercaptopropyl glycosides were conjugated to commercially available maleimide-activated BSA by 1,4-addition in aqueous solution at pH 7.2, which produced NGPs **1-BSA–15-BSA**. The conjugation was carried out in the presence of a water-soluble phosphine to reduce any sugar disulfides. The NGPs prepared in this manner have the tendency to aggregate over a time period of several months when kept in solution at 4 °C, which is in accordance with a recent study on the stability of BSA and lysozyme that had been exposed to reducing conditions [34]. Therefore we recommend storing the NGPs frozen at −20 to −80 °C, in small working aliquots to avoid repeated freeze-thaw cycles. Scheme 1 shows the con-jugation of trisaccharide **7** to BSA producing the NGP **7-BSA** as an example. The average glycan load per BSA molecule can be determined by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF-MS).

Figure 2 illustrates the glycan load determination of **2-BSA** by MALDI-TOF-MS as a representative example.

Schemes 2, 3, and 4 illustrate the syntheses of mercaptopropyl glycosides **2**, **7**, and **8** as examples of synthetic strategies applied for the generation of the αGal-containing glycan library. The most important synthetic feature that we applied for the synthesis of all αGal-containing disaccharides and trisaccharides [32, 33], is the use of Kiso's di-*tert*-butylsilylidene galactosyl donor **18** (Scheme 2) allowing for the stereoselective α-galactosylation despite the presence of a benzoyl protecting group at position 2 [35, 36]. For the synthesis of mercaptopropyl di-saccharide **2**, unprotected allyl β-galactoside **16** [37], synthesized from a per-acetylated allyl galactoside precursor [38], was protected as its isopropylidene ketal [39] followed by silylation with *tert*-butyl-diphenylsilylchloride to give acceptor **17**
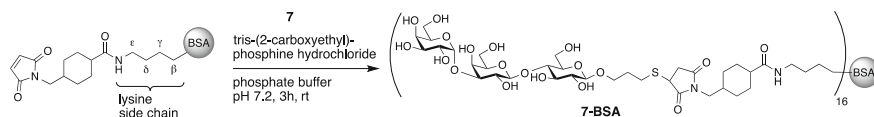
Scheme 1 Conjugate addition of mercaptopropyl saccharide **7** to Imject™ maleimide-activated BSA from Thermo Fisher Scientific produced NGP **7-BSA** with a glycan load of 16 glycans per BSA molecule
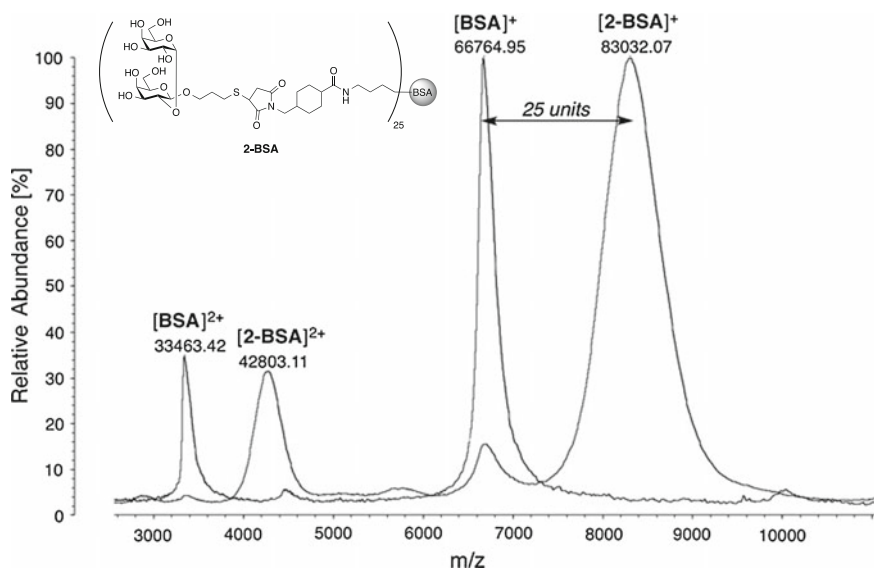
Fig. 2 MALDI-TOF mass spectra of BSA and **2-BSA**: in this case, the average loading is 25 disaccharide units per BSA molecule, based on the m/z difference of [BSA]$^+$ and [**2-BSA**]$^+$
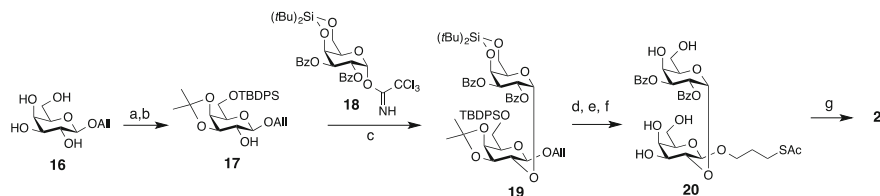
Scheme 2 Synthesis of mercaptopropyl disaccharide **2**: **a** 2,2-dimethoxypropane, pTsOH, DMF (85%); **b** TBDPSCl, DMAP, DCM (95%); **c** TMSOTf, DCM, 4 Å MS (82%); **d** HF-pyr, THF, 0 °C (91%); **e** DCM, H$_2$O, TFA (98%); **f** AcSH, AIBN, THF, 350 nm (83%); **g** NaOMe, MeOH (quant.)
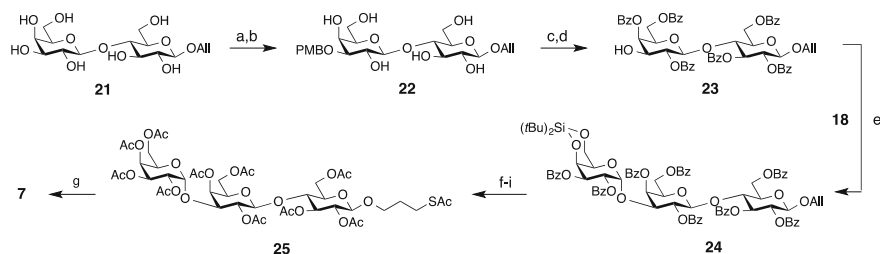
Scheme 3 Synthesis of the αGal-containing mercaptopropyl trisaccharide **7**: **a** Bu$_2$SnO, MeOH, 65°C; **b** PMBCl, Bu$_4$NBr, benzene, 65°C (54% two steps); **c** BzCl, pyr (quant.); **d** DDQ, CH$_2$Cl$_2$, H$_2$O (72%); **e** TMSOTf, DCM, 4 Å MS (85%); **f** HF-pyr, THF (90%); **g** NaOMe, MeOH (quant.); **h** Ac$_2$O, pyr; **i** AcSH, AIBN, THF, 350 nm (83% over two steps)
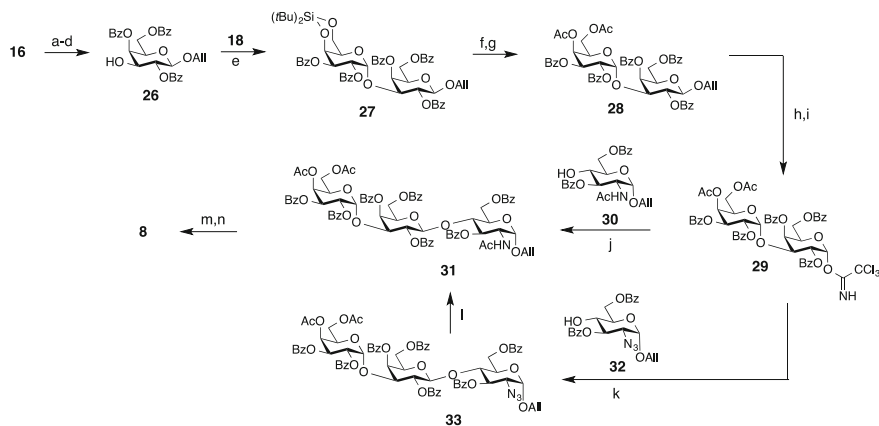
Scheme 4 Synthesis of Galα1,3Galβ1,4GlcNAcα-(CH$_2$)$_3$SH (**8**): **a** Bu$_2$SnO, MeOH; **b** PMBCl, Bu$_4$NCl, benzene (75% two steps); **c** BzCl, pyr (91%); **d** DDQ, DCM, H$_2$O (98%); **e** TMSOTf, DCM, 4 Å MS (92%); **f** HF-pyr, THF (90%); **g** Ac$_2$O, pyr (89%, two steps); **h** PdCl$_2$, MeOH (87%); **i** CCl$_3$CN, DCM, DBU (84%); **j** TMSOTf, DCM, 4 Å MS (30% α/β 1:4, separable by FPLC); **k** TMSOTf, DCM, 4 Å MS (46%); **l** AcSH (77%); **m** AcSH, AIBN, THF, 350 nm (89%); **n** NaOMe, MeOH (quant.)

[40]. α-Galactosylation with the Kiso donor **18** [35, 36] afforded the fully protected allyl disaccharide **19**. Removal of the di-*tert*-butylsilylidene group with tetrabutyl-ammonium fluoride or potassium fluoride in the presence of 18-crown-6 failed, but its removal was accomplished with hydrofluoric acid-pyridine complex, which simultaneously also removed the *tert*-butyldiphenylsilyl group to furnish the dibenzoylated allyl disaccharide. Radical addition of thioacetic acid to the double bond [31, 41, 42] of the allyl glycoside gave the thioester **20**, which was depro-tected under Zemplén conditions to afford mercaptopropyl disaccharide **2** (Scheme 2).

Scheme 3 illustrates the synthesis of mercaptopropyl trisaccharide **7** from allyl lactoside **21** [43], which was p-methoxybenzylated at position 3 in the galactose ring via its tin-acetal to give compound **22**. It was converted into glycosyl acceptor **23** by perbenzoylation, followed by oxidative removal of the p-methoxybenzyl group using 2,3-dichloro-5,6-dicyano-1,4-benzoquinone (DDQ). α-Galactosylation of acceptor **23** with the Kiso donor **18** furnished the fully protected allyl trisac-charide **24**, which was treated with hydrofluoric acid-pyridine complex to remove the di-*tert*-butylsilylidene group. Debenzoylation followed by acetylation, followed by radical addition of thioacetic acid by a thiol-ene reaction afforded trisaccharide **25**. The reason for the debenzoylation-acetylation strategy was a more straight-forward purification of the thiol-ene reaction product by silica gel column chro-matography. Upon deesterification of the fully esterified trisaccharide **25** using Zemplén conditions the target mercaptopropyl trisaccharide **7** was obtained.

The synthesis of mercaptopropyl trisaccharide **8** is much more challenging than the synthesis of its analog **7** (Scheme 3) due to the presence of *N*-acetylglucosamine

instead of a glucose moiety at the reducing end. The glycan moiety Galα(1,3)Galβ (1,4)GlcNAcα present in the tGPI mucins of *T. cruzi* [11] is an anomer of the well-known Galili epitope Galα(1,3)Galβ(1,4)GlcNAcβ [44] present on all mammalian cell surfaces with the exception of humans and Old World monkeys. For the Galili epitope and its glycosides, several chemical and chemo-enzymatic syntheses have been reported [45–52], but to the best of our knowledge there are no reports on the synthesis of mercaptopropyl trisaccharide **8** and analogs thereof with its unusual α-configuration at the reducing end. In order to develop a synthesis for **8** we envisioned to utilize the Kiso donor **18**, and a strategy that makes use of predominantly acyl protecting groups that can be easily installed and cleanly removed (Scheme **4**). Allyl β-galactoside **16** was converted into its tin acetal followed by p-methoxybenzylation at position 3. Benzoylation of the remaining hydroxyl groups and oxidative removal of the p-methoxybenzyl group with DDQ afforded acceptor **26**. This acceptor was glycosylated with the donor **18**, using trimethylsilyl trifluoromethanesulfonate (TMSOTf) catalysis to give disaccharide **27**. The di-*tert*-butylsilylidene group was cleaved with hydrofluoric acid-pyridine complex in tetrahydrofuran (THF), followed by acetylation of the two hydroxyls to afford the peracylated allyl disaccharide **28**. Treatment with palladium(II) chloride in methanol gave the hemiacetal, which was filtered immediately after consumption of the starting material to avoid the formation of a polar by-product that is observable after 2 h of reaction, and was converted into the α-trichloroacetimidate **29** with trichloroacetonitrile in the presence of 1,8-diazabicycloundec-7-ene (DBU). This donor was at first used to glycosylate the allyl GlcNAc acceptor **30** [33, 53], obtained from allyl 2-deoxy-2-acetamido-α-D-glucopyranoside [54], by selective benzoylation with *N*-benzoylimidazole [55, 56]. The glycosylation was accomplished with TMSOTf, but trisaccharide **31** was obtained as an anomeric mixture (1:4 α/β) of low yield, most likely due to the well-known poor nucleophilicity of the 4-OH of GlcNAc acceptors [57]. The separation of the two diastereomeric trisaccharides by column chromatography proved to be difficult, but can be accomplished by FPLC. Replacement of the acceptor **30** with the allyl 2-deoxy-2-azido-Glc acceptor **32** [33], produced from allyl 2-deoxy-2-azido-α-D-glucopyranoside [54] by selective benzoylation [53], furnished trisaccharide **33** in 46% yield, which was purified by flash chromatography. Reduction of the azide and installation of an *N*-acetyl group with neat thioacetic acid (AcSH) gave the trisaccharide **31**. Radical addition of AcSH in the presence of azobisisobutyronitrile (AIBN) in THF under UV light gave the thioester (not shown). All ester-protecting groups were removed under Zemplén conditions to afford the target trisaccharide **8**.

For glycoarray interrogation, pooled sera from ten CCD patients (ChSP) and pooled sera from ten healthy individuals (NHSP), obtained from the ISGlobal, Hospital Clinic, Universitat de Barcelona, were used. CCD patients had been diagnosed by two conventional ELISA tests, one with a lysate of *T. cruzi* parasites (Ortho-Clinical Diagnostics, Raritan, NJ, USA), and the other one with recombinant antigens (BioELISA Chagas, Biokit S.A., Barcelona, Spain). Healthy individuals tested negative in these two ELISAs. The synthetic αGal-containing NGPs **1-BSA**–**10-BSA**, cysteine conjugated to BSA (Cysteine-BSA), as well as the five

non-αGal-containing NGPs (**11-BSA–15-BSA**), were immobilized in microtiter plate wells, and the resulting glycoarrays were subjected to CL-ELISA [24] in two glycoarray sets. Unlike our previously published results, in which 125 ng of NGP was immobilized per well, and sera dilutions of 1:100 and 1:300 were used [32], here we decreased the quantity of NGP per well to 24 ng (Fig. 3) and 12 ng (Fig. 4), and increased the serum dilution to 1:800, because under these conditions most NGPs show higher reactivity differentials (ratios) between ChSP and NHSP.

Figure 3 represents the first of two glycoarray studies and shows the CL-ELISA responses for **1-BSA–11-BSA**, **14-BSA** and **Cysteine-BSA**, most of which had been previously studied under different conditions [32], using ChSP and NHSP at sera dilutions of 1:800. It is noticeable that the NGPs **3-BSA**, **7-BSA**, and **8-BSA** show strong CL-ELISA reactivities with ChSP, and the greatest differentials between ChSP and NHSP of ∼11 fold. All three NGPs share a common non-reducing terminal Galα(1,3)Galβ moiety, which is highly indicative for the disaccharide Galα(1,3)Galβ being an immunodominant *T. cruzi* trypomastigote cell surface glycotope.

Fig. 3 CL-ELISA reactivities of ChSP and NHSP to αGal- and βGal-containing NGPs. Each NGP was tested with ChSP and NHSP in duplicate. **Cysteine-BSA** was used as a negative control. The three NGPs with strong reactivities and large differentials between ChSP and NHSP are highlighted (dashed boxes). The amount of NGP and Cysteine-BSA loading in each well was 24 ng; sera were diluted at 1:800. RLU, relative luminescence units
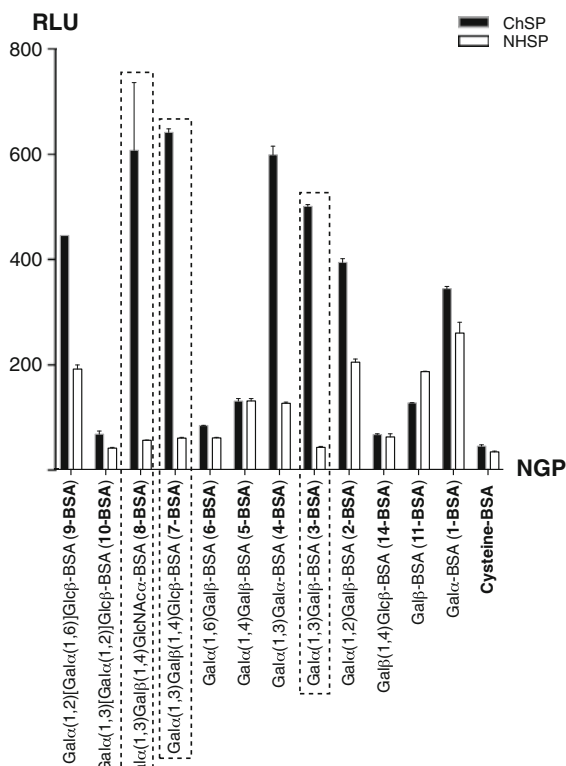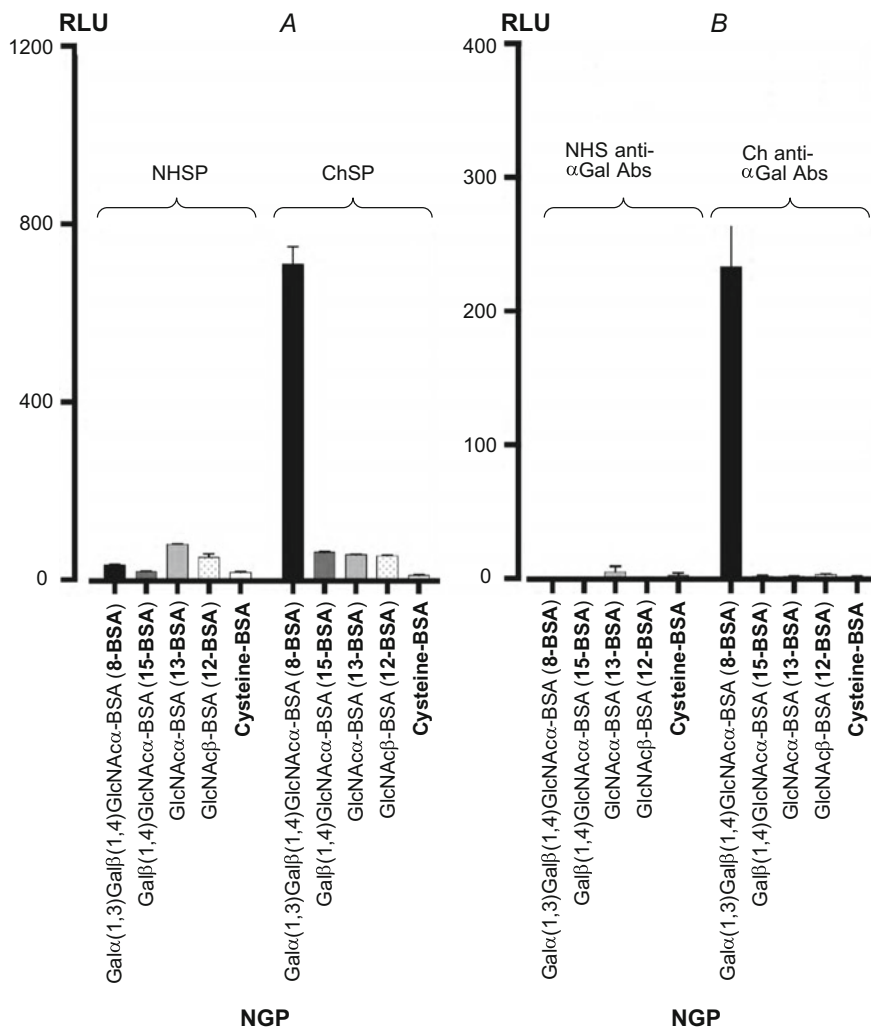
Fig. 4 **A** CL-ELISA reactivity of NHSP versus ChSP to NGPs and Cysteine-BSA at 12 ng/well; sera dilution 1:800. **B** CL-ELISA reactivity of purified NHS anti-α-Gal Abs versus Ch anti-α-Gal to NGPs and Cysteine-BSA at 12 ng/well; antibody concentration: 1 μg/mL. RLU, relative luminescence units. Figure modified from Ref. [33]

None of these NGPs are significantly recognized by NHSP. Interestingly, the branched trisaccharide Galα(1,3)[Galα(1,2)]Galβ in **10-BSA** is practically not recognized by these antibodies. It resembles the blood group B antigen Galα(1,3)[Fucα (1,2)]GalβR, which is also not recognized by NHS anti-α-Gal Abs [58] or Ch anti-α-Gal Abs (Almeida I.C., unpublished data). The Galα(1,3)Galα containing

NGP **4-BSA**, which differs from the Galα(1,3)Galβ containing NGP **3-BSA** only in its anomeric configuration at the reducing end, also shows a strong reactivity with ChSP, but a somewhat smaller differential between ChSP and NHSP of only ∼ 5:1 is observed. The strong reactivity with ChSP exhibited by **4-BSA** could be due to the fact that the Ch anti-α-Gal antibodies cannot distinguish between the two possible configurations of the galactose moiety at the reducing end. Another potential reason could be that *T. cruzi* trypomastigotes express both disaccharides, however, so far the Galα(1,3)Galα glycotope has not been identified by analysis of *T. cruzi*-derived lysates, and its existence remains unconfirmed. The Galα(1,2)Galβ-containing NGPs **2-BSA** and **9-BSA** also have strong reactivities with ChSP, which matches reports of high levels of anti-α-Gal Abs with specificity toward the Galα(1,2)Gal epitope in the sera of CCD patients [30]. However, the differentials between ChSP and NHSP reactivity to **2-BSA** and **9-BSA** are much lower (∼ 2 fold) than those observed for **3-BSA**, **7-BSA**, and **8-BSA**. The NGP **1-BSA**, which contains only the monosaccharide αGal, gave a medium strong CL-ELISA response with a poor differential of ∼ 1.3 fold between ChSP and NHSP. **5-BSA** [Galα(1,4)Galβ BSA] and **6-BSA** [Galα(1,6)Galβ BSA] show weak reactivities with ChSP, indicating that neither of them is a major immunogenic *T. cruzi* glycotope. Their reactivities are similar to the ones displayed by the βGal-containing NGP **11-BSA** and the lactose-containing NGP **14-BSA**, which only have slightly higher reactivities than the negative control **Cysteine-BSA** (Fig. 3).

The second glycoarray included the NGPs **8-BSA**, **12-BSA**, **13-BSA**, and **15-BSA**, as well as the negative control **Cysteine-BSA** [33]. Unlike the first glycoarray in which 24 ng of NGP was immobilized in each well, the second glycoarray was constructed immobilizing only 12 ng of NGP per well. Figure 4A shows CL-ELISA results of the glycoarray interrogation using ChSP and NHSP. Galα(1,3) Galβ(1,4)GlcNAcα-BSA (**8-BSA**) shows a 20-fold differential between ChSP and NHSP, whereas the NGPs GlcNAcβ-BSA (**12-BSA**), GlcNAcα-BSA (**13-BSA**) and Galβ(1,4)GlcNAcα-BSA (**15-BSA**) are only weakly recognized by antibodies of ChSP or NHSP. As expected, **Cysteine-BSA** showed practically no reactivity. In addition, the glycoarray was interrogated with Ch and NHS anti-α-Gal Abs (Fig. 4B), which had been purified by affinity chromatography using immobilized Galα(1,3)Galβ(1,4)GlcNAcβ [13]. Galα(1,3)Galβ(1,4)GlcNAcα-BSA (**8-BSA**) displays a 230-fold differential between Ch-αGal Abs and NHS-αGal Abs, while the other NGPs (**12-BSA**, **13-BSA**, and **15-BSA**) and **Cysteine-BSA** remain practically unrecognized by either Abs. Our results show that the terminal αGal moiety of Galα (1,3)Galβ(1,4)GlcNAcα is essential for Ch antibody recognition. Although GlcNAcα and Galβ(1,4)GlcNAcα, which are underlying partial structures of Galα (1,3)Galβ(1,4)GlcNAcα, are nonself glycotopes for humans, there are only very weak antibody responses against them in ChSP. The αGal-containing glycoarray/CL-ELISA method presented here, especially when carried out with only 12 ng or 24 ng of antigen/well under dilute conditions is highly suitable for the differentiation between *T. cruzi*-infected and non-infected sera.

Based on the CL-ELISA results illustrated in Fig. 3, the question of specific glycotope recognition by Ch anti-α-Gal Abs arises: Does each of the three Galα (1,3)Galβ-containing NGPs **3-BSA**, **7-BSA**, and commercial Galα(1,3)Galβ(1,4) GlcNAcβ conjugated to BSA (purchased from V-Labs), which is also strongly recognized by ChSP and shows a favorable differential between ChSP and NHSP [32], recruit its own set of antibodies from the ChSP, or are the three glycotopes recognized by the same anti-α-Gal Abs? To address this question, we compared the CL-ELISAs of the individual NGPs immobilized in microtiter plate wells at different quantities, with that of the combined NGPs immobilized in the same microtiter plate wells at different quantities between 5 and 125 ng, and CL-ELISA reactivities were measured at three different sera dilutions (1:100, 1:200, and 1:400) (Fig. 5) [32]. We hypothesized if each NGP was recognized by a different set of antibodies, one would expect significantly higher RLU readings for the combined NGPs. However, the curves of the four experimental sets and the RLU readings resemble each other quite closely (Fig. 5). As expected, NHSP showed almost no binding to **3-BSA**, **7-BSA**, or Galα(1,3)Galβ(1,4)GlcNAcβ-BSA up to 125 ng. With ChSP, the titration curves for **3-BSA**, **7-BSA**, and Galα(1,3)Galβ(1,4) GlcNAcβ-BSA were similar. This indicates that antibodies from CCD patients recognize all three saccharides (Galα(1,3)Galβ, Galα(1,3)Galβ(1,4)Glcβ, and Galα (1,3)Galβ(1,4)GlcNAcβ) to a similar extent. As can be seen in the titration curves of Fig. 5, right panel, when the NGPs **3-BSA**, **7-BSA**, and Galα(1,3)Galβ(1,4) GlcNAcβ-BSA are combined in the same wells, no significant increase in the fluorescence signal was observed at a serum dilution of 1/100, showing that only a very small or no synergistic effect exists. This experiment suggests that for the most part the same antibodies recognize all three NGPs (**3-BSA**, **7-BSA**, and Galα(1,3) Galβ(1,4)GlcNAcβ-BSA) confirming that Galα(1,3)Galβ is the immunodominant disaccharide glycotope that is specifically recognized, regardless of whether it is linked to a short alkyl chain, or whether it has a 1,4 linkage to βGlc, or βGlcNAc.
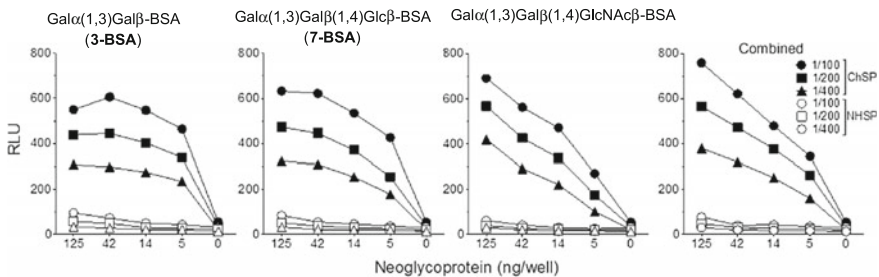
Fig. 5 CL-ELISA reactivity of ChSP and NHSP to the Galα(1,3)Galβ-containing NGPs **3-BSA**, **7-BSA**, and the commercial NGP Galα(1,3)Galβ(1,4)GlcNAcβ-BSA, alone or combined. Each NGP (alone or combined) was tested with ChSP and NHSP in duplicate. Figure modified from Ref. [32]

3   Conclusions

Synthetic strategies were developed for the synthesis of a biased library of ten *T. cruzi* glycans with terminal αGal moieties, which can be interrogated with sera of CCD patients to identify potentially immunodominant glycotopes. One of the glycans synthesized is the mercaptopropyl glycoside of Galα(1,3)Galβ(1,4) GlcNAcα, which is the only structurally defined *T. cruzi* tGPI mucin glycan. It was prepared in 12 steps from known monosaccharide building blocks. The two key features utilized throughout all glycan syntheses is the stereoselective installation of the terminal αGal moiety with the di-*tert*-butylsilylidene protected "Kiso donor" **18**, and the installation of allyl glycosides at the reducing ends. We have performed thiol-ene reactions by radical addition of thioacetic acid to the double bond of these allyl glycosides, and saponification of the resulting thioester afforded mercapto-propyl thioglycosides. However, allyl glycosides are very versatile, as they can be easily converted into hemiacetals, or into aldehydes by ozonolysis, which can then be further derivatized.

All glycans were conjugated to BSA, and the resulting NGPs were immobilized in wells of microtiter plates, thus generating a glycoarray that was subjected to CL-ELISA using pooled sera from CCD patients (ChSP), and healthy individuals (NHSP). ChSP strongly recognized the terminal disaccharide Galα(1,3)Galβ indicating that this glycotope is immunodominant. No matter if this disaccharide is connected to a short alkyl residue or to a glucose or *N*-acetylglucosamine moiety, the same set of Ch anti-α-Gal antibodies seem to recognize this disaccharide. The terminal disaccharide Galα(1,3)Galβ is specifically recognized by Ch anti-α-Gal antibodies with a large differential of 10–20 fold between CCD sera and the NHS when only 12 or 24 ng of the NGP is immobilized per well, and when the sera are diluted at 1:800. All other saccharides synthesized and conjugated to BSA gave poor reactivity differentials between ChSP and NHSP or had significantly weaker reactivity with ChSP. Interestingly, the nonself Galβ(1,4)GlcNAcα and GlcNAcα glycotopes are not recognized by ChSP, which stresses that the terminal αGal residue is essential for ChSP binding. Our data indicate that based on the large differential in reactivity, fully synthetic, structurally defined Galα(1,3) Galβ-containing NGPs could be used as biomarkers for the diagnosis of Chagas disease and could potentially be used for follow-up of chemotherapy, thus replacing purified and heterogeneous tGPI-mucins currently used for these purposes [19, 21, 23–25, 59]. In addition, synthetic Galα(1,3)Galβ-containing glycoconjugates could potentially be suitable for the development of glycan-based therapeutic and/or preventive vaccines for experimental vaccination against *T. cruzi* infection, as we recently proposed [33].

References

1. Rassi A Jr, Rassi A, Marcondes de Rezende J (2012) American trypanosomiasis (Chagas disease). Infect Dis Clin North Am 26:275–291
2. Carod-Artal FJ, Gascon J (2010) Chagas disease and stroke. Lancet Neurol 9:533–542
3. Gascon J, Bern C, Pinazo MJ (2010) Chagas disease in Spain, the United States and other non-endemic countries. Acta Trop 115:22–27
4. Jackson Y, Pinto A, Pett S (2014) Chagas disease in Australia and New Zealand: risks and needs for public health interventions. Trop Med Int Health 19:212–218
5. Imai K, Maeda T, Sayama Y, Mikita K, Fujikura Y, Misawa K, Nagumo M, Iwata O, Ono T, Kurane I, Miyahira Y, Kawana A, Miura S (2014) Mother-to-child transmission of congenital Chagas disease. Japan Emerg Infect Dis 20:146–148
6. Acosta-Serrano A, Almeida IC, Freitas-Junior LH, Yoshida N, Schenkman S (2001) The mucin-like glycoprotein super-family of Trypanosoma cruzi: structure and biological roles. Mol Biochem Parasitol 114:143–150
7. Buscaglia CA, Campo VA, Frasch AC, Di Noia JM (2006) *Trypanosoma cruzi* surface mucins: host-dependent coat diversity. Nat Rev Microbiol 4:229–236
8. Acosta-Serrano A, Hutchinson C, Nakayasu ES, Almeida IC, Carrington M (2007) Comparison and evolution of the surface architecture of trypanosomatid parasites. In: Barry JD, Mottram JC, McCulloch R, Acosta-Serrano A (eds) Trypanosomes: after the genome. Horizon Scientific Press, Norwich, UK, pp 319–337
9. de Lederkremer RM, Agusti R (2009) Glycobiology of *Trypanosoma cruzi*. Adv Carbohydr Chem Biochem 62:311–366
10. Mendonca-Previato L, Penha L, Garcez TC, Jones C, Previato JO (2013) Addition of alpha-O-GlcNAc to threonine residues define the post-translational modification of mucin-like molecules in *Trypanosoma cruzi*. Glycoconj J 30:659–666
11. Almeida IC, Ferguson MA, Schenkman S, Travassos LR (1994) Lytic anti-alpha-galactosyl antibodies from patients with chronic Chagas' disease recognize novel O-linked oligosaccharides on mucin-like glycosyl-phosphatidylinositol-anchored glycoproteins of *Trypanosoma cruzi*. Biochem J 304(Pt 3):793–802
12. Gazzinelli RT, Pereira ME, Romanha A, Gazzinelli G, Brener Z (1991) Direct lysis of *Trypanosoma cruzi*: a novel effector mechanism of protection mediated by human anti-gal antibodies. Parasite Immunol 13:345–356
13. Almeida IC, Milani SR, Gorin PA, Travassos LR (1991) Complement-mediated lysis of *Trypanosoma cruzi* trypomastigotes by human anti-alpha-galactosyl antibodies. J Immunol 146:2394–2400

14. Previato JO, Sola-Penna M, Agrellos OA, Jones C, Oeltmann T, Travassos LR, Mendonca-Previato L (1998) Biosynthesis of O-N-acetylglucosamine-linked glycans in *Trypanosoma cruzi*. Characterization of the novel uridine diphospho-N-acetylglucosamine: polypeptide N-acetylglucosaminyltransferase-catalyzing formation of N-acetylglucosamine alpha1– > O-threonine. J Biol Chem 273:14982–14988

15. Serrano AA, Schenkman S, Yoshida N, Mehlert A, Richardson JM, Ferguson MA (1995) The lipid structure of the glycosylphosphatidylinositol-anchored mucin-like sialic acid acceptors of *Trypanosoma cruzi* changes during parasite differentiation from epimastigotes to infective metacyclic trypomastigote forms. J Biol Chem 270:27244–27253

16. Previato JO, Jones C, Goncalves LP, Wait R, Travassos LR, Mendonca-Previato L (1994) O-glycosidically linked N-acetylglucosamine-bound oligosaccharides from glycoproteins of *Trypanosoma cruzi*. Biochem J 301(Pt 1):151–159

17. Galili U, Wang L, Temple DCL, Radic MZ (1999) The natural anti-Gal antibody. Sub-Cell. Biochem 32:79–106

18. Wilkinson SG (1996) Bacterial lipopolysaccharides—themes and variations. Prog Lipid Res 35:283–343

19. Almeida IC, Krautz GM, Krettli AU, Travassos LR (1993) Glycoconjugates of *Trypanosoma cruzi*: a 74 kD antigen of trypomastigotes specifically reacts with lytic anti-alpha-galactosyl antibodies from patients with chronic Chagas disease. J Clin Lab Anal 7:307–316

20. Soares RP, Torrecilhas AC, Assis RR, Rocha MN, Moura e Castro FA, Freitas GF, Murta SM, Santos SL, Marques AF, Almeida IC, Romanha AJ (2012) Intraspecies variation in Trypanosoma cruzi GPI-mucins: biological activities and differential expression of alpha-galactosyl residues. Am J Trop Med Hyg 87:87-96

21. Izquierdo L, Marques AF, Gallego M, Sanz S, Tebar S, Riera C, Quinto L, Aldasoro E, Almeida IC, Gascon J (2013) Evaluation of a chemiluminescent enzyme-linked immunosorbent assay for the diagnosis of *Trypanosoma cruzi* infection in a nonendemic setting. Mem Inst Oswaldo Cruz 108:928–931

22. De Marchi CR, Di Noia JM, Frasch AC, Amato Neto V, Almeida IC, Buscaglia CA (2011) Evaluation of a recombinant *Trypanosoma cruzi* mucin-like antigen for serodiagnosis of Chagas' disease. Clin Vaccine Immunol 18:1850–1855

23. Andrade AL, Martelli CM, Oliveira RM, Silva SA, Aires AI, Soussumi LM, Covas DT, Silva LS, Andrade JG, Travassos LR, Almeida IC (2004) Short report: benznidazole efficacy among *Trypanosoma cruzi*-infected adolescents after a six-year follow-up. Am J Trop Med Hyg 71:594–597

24. Almeida IC, Covas DT, Soussumi LM, Travassos LR (1997) A highly sensitive and specific chemiluminescent enzyme-linked immunosorbent assay for diagnosis of active *Trypanosoma cruzi* infection. Transfusion 37:850–857

25. de Andrade AL, Zicker F, de Oliveira RM, Almeida Silva S, Luquetti A, Travassos LR, Almeida IC, de Andrade SS, de Andrade JG, Martelli CM (1996) Randomised trial of efficacy of benznidazole in treatment of early *Trypanosoma cruzi* infection. Lancet 348:1407–1413

26. Avila JL, Rojas M, Galili U (1989) Immunogenic Gal alpha 1–3Gal carbohydrate epitopes are present on pathogenic American Trypanosoma and Leishmania. J Immunol. 142:2828–2834

27. Antas PR, Medrano-Mercado N, Torrico F, Ugarte-Fernandez R, Gomez F, Correa Oliveira R, Chaves AC, Romanha AJ, Araujo-Jorge TC (1999) Early, intermediate, and late acute stages in Chagas' disease: a study combining anti-galactose IgG, specific serodiagnosis, and polymerase chain reaction analysis. Am J Trop Med Hyg 61:308–314

28. Medrano-Mercado N, Luz MR, Torrico F, Tapia G, Van Leuven F, Araujo-Jorge TC (1996) Acute-phase proteins and serologic profiles of chagasic children from an endemic area in Bolivia. Am J Tro p Med Hyg 54:154–161

29. Gonzalez J, Neira I, Gutierrez B, Anacona D, Manque P, Silva X, Marin S, Sagua H, Vergara U (1996) Serum antibodies to *Trypanosoma cruzi* antigens in Atacamenos patients from highland of northern Chile.Acta.Tro p 60:225–236

30. Avila JL, Rojas M, Velazquez-Avila G (1992) Characterization of a natural human antibody with anti-galactosyl(alpha 1-2)galactose specificity that is present at high titers in chronic *Trypanosoma cruzi* infection. Am J Trop Med Hyg 47:413–421

31. Houseman BT, Gawalt ES, Mrksich M (2003) Maleimide-functionalized self-assembled monolayers for the preparation of peptide and carbohydrate biochips. Langmuir 19:1522–1531

32. Ashmus RA, Schocker NS, Cordero-Mendoza Y, Marques AF, Monroy EY, Pardo A, Izquierdo L, Gallego M, Gascon J, Almeida IC, Michael K (2013) Potential use of synthetic alpha-galactosyl-containing glycotopes of the parasite *Trypanosoma cruzi* as diagnostic antigens for Chagas disease. Org Biomol Chem 11:5579–5583

33. Schocker NS, Portillo S, Brito CR, Marques AF, Almeida IC, Michael K (2016) Synthesis of Galα(1,3)Galβ(1,4)GlcNAcα-, Galβ(1,4)GlcNAcα- and GlcNAc-containing neoglycoproteins and their immunological evaluation in the context of Chagas disease. Glycobiology 26:39–50

34. Yang M, Dutta C, Tiwari A (2015) Disulfide-bond scrambling promotes amorphous aggregates in lysozyme and bovine serum albumin. J Phys Chem B 119:3969–3981

35. Imamura A, Kimura A, Ando H, Ishida H, Kiso M (2006) Extended application of di-tert-butylsilylene-directed alpha-predominant galactosylation compatible with C2-participating groups toward the assembly of various glycosides. Chem Eur J 12:8862–8870

36. Kimura A, Imamura A, Ando H, Ishida H, Kiso M (2006) A novel synthetic route to alpha-galactosyl ceramides and iGb3 using DTBS-directed alpha-selective galactosylation. Synlett 2379–2382

37. Stevenson DE, Furneaux RH (1996) Synthesis of allyl β-D-galactopyranoside from lactose using Streptococcus thermophilus β-D-galactosidase. Carbohydr Res 284:279–283

38. Khamsi J, Ashmus RA, Schocker NS, Michael K (2012) A high-yielding synthesis of allyl glycosides from peracetylated glycosyl donors. Carbohydr Res 357:147–150

39. Balcerzak AK, Ferreira SS, Trant JF, Ben RN (2012) Structurally diverse disaccharide analogs of antifreeze glycoproteins and their ability to inhibit ice recrystallization. Bioorg Med Chem Lett 22:1719–1721

40. Murata S, Ichikawa S, Matsuda A (2005) Synthesis of galactose-linked uridine derivatives with simple linkers as potential galactosyltransferase inhibitors. Tetrahedron 61:5837–5842

41. Matsuoka K, Oka H, Koyama T, Esumi Y, Terunuma D (2001) An alternative route for the construction of carbosilane dendrimers uniformly functionalized with lactose or sialyllactose moieties. Tetrahedron Lett 42:3327–3330

42. Posner T (1905) Beiträge zur Kenntniss der ungesättigten Verbindungen. II. Ueber die Addition von Mercaptanen an ungesättigte Kohlenwasserstoffe. Ber 38:646–657

43. Kononov LO, Kornilov AV, Sherman AA, Zyrianov EV, Zatonskiĭ GV, Shashkov AS, Nifant'ev NE (1998) Synthesis of oligosaccharides related to HNK-1 antigen. 2. Synthesis of 3'''-O-(3-O-sulfo-beta-D-glucuronopyranosyl)-lacto-N-neotetraose beta-propylglycoside. Bioorg Khim 24:608–822

44. Galili U, Shohet SB, Kobrin E, Stults CL, Macher BA (1988) Man, apes, and Old World monkeys differ from other mammals in the expression of alpha-galactosyl epitopes on nucleated cells. J Biol Chem 263:17755–17762

45. Litjens REJN, Hoogerhout P, Filippov DV, Codee JDC, Bos LJvd, Berg RJBHNvd, Overkleeft HS, Marel GAvd (2005) Synthesis of an alpha-Gal epitope α-D-Galp-(1-3)-β-D-Galp-(1-4)-β-GlcpNAc-lipid conjugate. J Carbohydr Chem 24:755–769

46. Brinkmann N, Malissard M, Ramuz M, Römer U, Schumacher T, Berger EG, Elling L, Wandrey C, Liese A (2001) Chemo-enzymatic synthesis of the galili epitope Galα(1→3)Galβ(1→4)GlcNAc on a homogeneously soluble PEG polymer by a multi-enzyme system. Bioorg Med Chem Lett 11:2503–2506

47. Dahmén J, Magnusson G, Hansen HC (2002) Synthesis of the linear B type 2 trisaccharide Galα3Galβ4GlcNAcβOTMSEt, and coupling of the corresponding 2-carboxyethyl β-thioglycoside to sepharose. J Carbohydr Chem 21:1–12

48. Plaza-Alexander P, Lowary TL (2013) Synthesis of trisaccharides incorporating the α-Gal antigen functionalized for neoglycoconjugate preparation. Arkivoc ii:112–122

49. Vic G, Hao T, Scigelova M, Crout DHG (1997) Glycosidase-catalysed synthesis of oligosaccharides: a one step synthesis of lactosamine and of the linear B type 2 trisaccharide alpha-D-Gal-(1,3)-beta-D-Gal-(1,4)-beta-D-GlcNAcSEt involved in the hyperacute rejection response in xenotransplantation from pigs to man and as the specific receptor for toxin A from Clostridium difficile. Chem Commun 169–170

50. Wang Y, Yan Q, Wu J, Zhang L-H, Ye X (2005) A new one-pot synthesis of α-Gal epitope derivatives involved in the hyperacute rejection response in xenotransplantation. Tetrahedron 61:4313–4321

51. Hanessian S, Saavedra OM, Mascitti V, Marterer W, Oehrlein R, Mak C-P (2001) Practical syntheses of B disaccharide and linear B type 2 trisaccharide—non-primate epitope markers recognized by human anti-α-Gal antibodies causing hyperacute rejection of xenotransplants. Tetrahedron 57:3267–3280

52. Fang J, Li J, Chen X, Zhang Y, Wang J, Guo Z, Zhang W, Yu L, Brew K, Wang PG (1998) Highly efficient chemoenzymatic synthesis of α-galactosyl epitopes with a recombinant α(1 → 3)-galactosyltransferase. J Am Chem Soc 120:6635–6638

53. Danac R, Ball L, Gurr SJ, Muller T, Fairbanks AJ (2007) Carbohydrate Chain Terminators: Rational Design of Novel Carbohydrate-Based Antifungal Agents. ChemBioChem 8:1241–1245

54. Gavard O, Hersant Y, Alais J, Duverger V, Dilhas A, Bascou A, Bonnaffé D (2003) Efficient preparation of three building blocks for the synthesis of heparan sulfate fragments: towards the combinatorial synthesis of oligosaccharides from hypervariable regions. Eur J Org Chem 3603–3620

55. Varela O, Cicero D, de Lederkremer RM (1989) A convenient synthesis of 4-thio-D-galactofuranose. J Org Chem 54:1884–1890

56. Gallo-Rodriguez C, Varela O, de Lederkremer RM (1996) First synthesis of beta-D-Galf(1-4) GlcNAc, a structural unit attached O-glycosidically in glycoproteins of *Trypanosoma cruzi*. J Org Chem 61:1886–1889

57. Crich D, Dudkin V (2001) Why are the hydroxyl groups of partially protected N-acetylglucosamine derivatives such poor glycosyl acceptors, and what can be done about it? J Am Chem Soc 123:6819–6825

58. Galili U, Macher BA, Buehler J, Shohet SB (1985) Human natural anti-alpha-galactosyl IgG. II. The specific recognition of alpha (1–3)-linked galactose residues. J Exp Med 162:573–582

59. Pinazo MJ, Posada Ede J, Izquierdo L, Tassies D, Marques AF, de Lazzari E, Aldasoro E, Munoz J, Abras A, Tebar S, Gallego M, de Almeida IC, Reverter JC, Gascon J (2016) Altered Hypercoagulability Factors in Patients with Chronic Chagas Disease: Potential Biomarkers of Therapeutic Response. PLoS Negl. Trop. Dis. 10:e0004269

Syntheses and Functions
of Glycosaminoglycan Mimicking
Polymers

Yoshiko Miura, Tomohiro Fukuda, Hirokazu Seto and Yu Hoshino

Abstract Glycosaminoglycans (GAGs) are important polysaccharides in the living system. Though total syntheses of GAGs oligosaccharides have been reported, it is still difficult to obtain GAGs. In this investigation, GAGs mimetics were prepared by polymerization of vinyl sugars instead of total synthesis. Glycopolymers are polymers with pendant saccharides, and exhibit a strong molecular recognition due to the multivalency. GAGs mimicking glycopolymers were prepared by polymerizing acrylamide derivatives carrying sulfated *N*-acetyl glucosamine (GlcNAc). The polymer interactions with proteins were investigated. The glycopolymer libraries were prepared with varying molecular weight, sugar structure, and sugar ratios. The GAG glycopolymers with biodegradable backbone, dendrimer, and arrays were also prepared. The inhibitory activity of Alzheimer amyloid beta peptides by the glycopolymers was investigated in detail. The glycopolymer with sulfated GlcNAc inhibited the aggregation of amyloid beta (Aβ) and the multivalency of sulfated GlcNAc was the key of the interaction. The activity depends on the chemical structure of glycopolymers. Also, the sulfated saccharide function was correlated to the functions of native GAGs.

Y. Miura (✉) · T. Fukuda · H. Seto · Y. Hoshino
Department of Chemical Engineering, Graduate School of Engineering, Kyushu University, 744 Motooka, Nishi-ku, Fukuoka 819-0395, Japan
e-mail: miuray@chem-eng.kyushu-u.ac.jp

T. Fukuda
National Institute of Technology, Toyama College, 13 Hogo-machi, Toyama City, Toyama 939-8630, Japan

H. Seto
Department of Engineering, Fukuoka University, 8-19-1 Nanakuma, Jonan-ku, Fukuoka 814-0180, Japan

1 Introduction

Saccharides play important roles in living systems. Compounds containing saccharides include glycolipids, glycoproteins, and polysaccharides [1]. Polysaccharides have various functions in living organisms. They can act as the source of energy, cause compounds to be water soluble, are building units of cell walls, and can act as biological ligands for molecular recognition. Among them, proteoglycans and GAGs are extracellular and known to relate to the important molecular recognition and tuning of biological activities [2]. The proteoglycans also contain GAGs with the backbone protein. GAGs are involved in various biological phenomena such as cell differentiation, cell proliferation (activation of growth factor), immune response, and angiogenesis, antithrombus, and protein structures [2–4].

GAGs are polysaccharides having alternating structure of amino-sugars [*N*-acetyl glucosamine (GlcNAc) and *N*-acetyl galactosamine (GalNAc)] and uronic acids [iduronic acid (IdoA), and glucuronic acid (GlcA)] [2, 5] (Fig. 1). With the exception of hyaluronic acid, most GAGs are highly sulfated, and their sulfated patterns and saccharide structures are specific. GAGs are natural polymers of poly anions with relatively high molecular weights. The chemical structures of GAGs are classified into some categories: hyaluronic acid is non-sulfated polysaccharide with GlcNAc and GlcA. Heparin with GlcNAc and IdoA/GlcA is a highly sulfated GAG exhibiting antithrombogenicity. Chondroitin sulfate with GalNAc and GlcA is a component of cartilage responsible for the compression resistance. The chemical structures of GAGs are complicated due to the possibility of various sulfation patterns, various chiral patterns, and two stereoisomers at the glycosidic (anomeric) position.
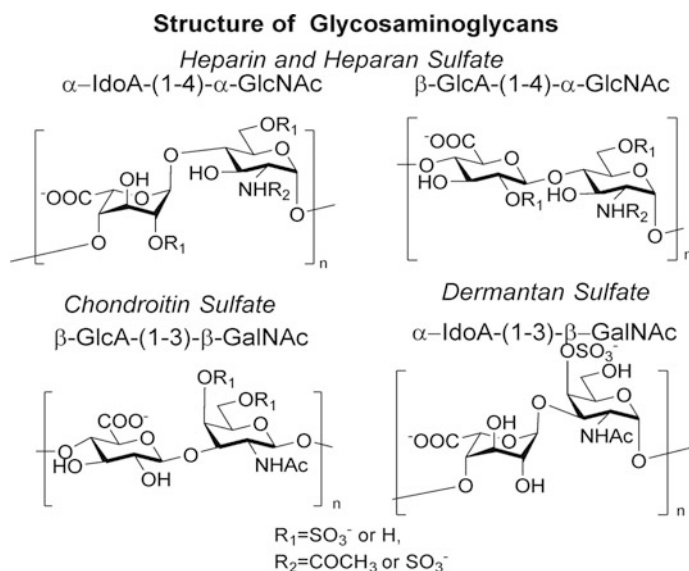
Fig. 1 Representative chemical structures of glycosaminoglycans

Since GAGs exhibit various important functions and possess prominent properties important in biochemistry and biomaterials, many groups have tried total syntheses of GAGs [5]. It would be of great value for research and development if there was a stable supply of GAGs. The effective total synthesis would offer a stable supply of GAGs and enable detailed biochemical studies. Additionally, the methods of retro synthesis have been studied. In the synthesis of heparin and heparan sulfate, the important GAGs oligosaccharide like IdoA(2S)(α1–4)GlcNS(6S) were specially investigated. Even though the biological activity of formed oligosaccharides was much weaker than that of original GAGs, oligosaccharides showed some of the GAGs function such as blood platelet binding and antithrombin III binding [6, 7]. The Kusumoto group reported the total synthesis of a pentasaccharide of GAGs [8], and the Huang group reported the synthesis of octasaccharides by the sophisticated glycosylation and protective group approach [9]. The Seeberger group reported the library of GAGs oligosaccharide and microarrays [10]. The synthesis of GAGs oligosaccharides was also accomplished by biological methods [11].

In spite of the enthusiastic researches by organic chemists, the attained syntheses are limited to oligosaccharides. Since GAGs are polysaccharides with large molecular weight, the biological activities of the oligosaccharides are much weaker. In general, the saccharide–protein interactions are weak and can be amplified by taking advantage of multivalency. The multivalent compound can amplify the sugar–protein interaction. The Hsieh-Wilson group reported the glycopolymers with GAGs disaccharide and tetrasaccharide [12]. The Suda group has reported the glycodendrimer with oligosaccharides of GAG that can amplify the interaction with growth factor and platelet [13]. The strong interaction of native GAGs is contributed by multivalent effect. Many groups have reported the synthesis of multivalent sugar compounds such as glycodendrimers, glycopeptides, and glycopolymers [14]. Among them, glycopolymers show the strongest multivalent effects to proteins and cells. In addition, the glycopolymers have the advantage of being applicable to biomaterials. Taking into account the availability and a potential for strong interactions, a glycopolymer with GAGs functionality is very attractive. In this chapter, we describe the method of GAG mimicking when employing a glycopolymer.

2 Glycopolymer and Multivalency

There are many kinds of multivalent compounds. We define glycopolymers as polymers with pendant saccharides. The Kobayashi group [15] and the Whitesides group [16] demonstrated their practical significance and application in the early studies from 1980s. The glycopolymers can be applied as biomaterials for cell cultivation, sugar-recognition proteins (lectins), and virus neutralization, indicating the importance of multivalency.

Multivalent effects of glycopolymers are based on the plural binding of glycopolymers to proteins, and the increase of the binding possibilities due to the

multivalent display. The plural binding of a glycopolymer has an advantage of the enthalpy gain, and the increase of the binding possibilities is preferable to the entropy gain. The glycopolymer is thermodynamically advantageous due to both enthalpy and entropy factors. This contributes to strong biological activities. The previous reports suggest that the polysaccharide of GAGs has a multivalent structure for protein binding [12, 13]. The structure of the saccharide and its multivalency are the key to GAGs biological activities.

The preparation of glycopolymers was mostly accomplished by polymerizing sugar derivatives. The sugar derivatives with acryl amide, methacrylate, acrylate, and norbornene were reported to produce glycopolymers [14]. The glycopolymers were prepared by the orthogonal reaction to sugars such as radical polymerization. We have reported the preparation of various glycopolymers by employing radical and living radical polymerizations [14]. The Kiessling group reported the synthesis of glycopolymers with a controlled molecular weight by ring-opening metathesis polymerization (ROMP) [17]. Recently, the technology of polymer synthesis was much improved. The controlled polymerizations were attained by living radical polymerization and ROMP. Polymers with special structures like glycodendrimers and glycostar polymers have been also reported. Since the native saccharides have precisely controlled structures, the incorporation of recent synthetic polymer techniques can enable the well-controlled glyco-mimetics.

We propose a new approach to mimicking the function of GAGs. The GAGs structure can be divided into the functional saccharide component and polymerizable sugars. The monomer (polymerizable) sugars were polymerized by radical polymerization. The glycopolymers can show the strong interaction to the target protein due to multivalency. The advantage of the GAGs mimicking polymer is the easy preparation as compared to total synthesis, which can lead to the preparation of GAGs mimicking libraries (Fig. 2).

3   Sulfated Saccharides and Polymers

Since sulfated sugars are specific components in GAGs, we mimicked the GAGs with glycopolymer with sulfated sugars [18]. The synthesis of polymerizable sugars requires usually a tedious procedure. In order to simplify the synthetic procedures, *p*-nitrophenyl glucosides were used as starting materials. Acrylamide phenyl 6-sulfo-GlcNAc was prepared as the monomer. It is worth noting that 6-sulfo-GlcNAc was reported to play a significant role in living systems [19]. The polymer was prepared by using an initiator. The advantage of employing the glycopolymer was an easy preparation of molecular libraries. First, we prepared the GAGs mimicking polymer library of 6-sulfo-GlcNAc with varying sugar density (Fig. 3), where densities were adjusted by the copolymerization with acrylamide. The function of sulfated glycopolymers was investigated by studying the interaction to Alzheimer amyloid beta proteins [Aβ (1-40) and Aβ (1-42)] that are pathogens of the Alzheimer disease [20, 21] (Fig. 3).
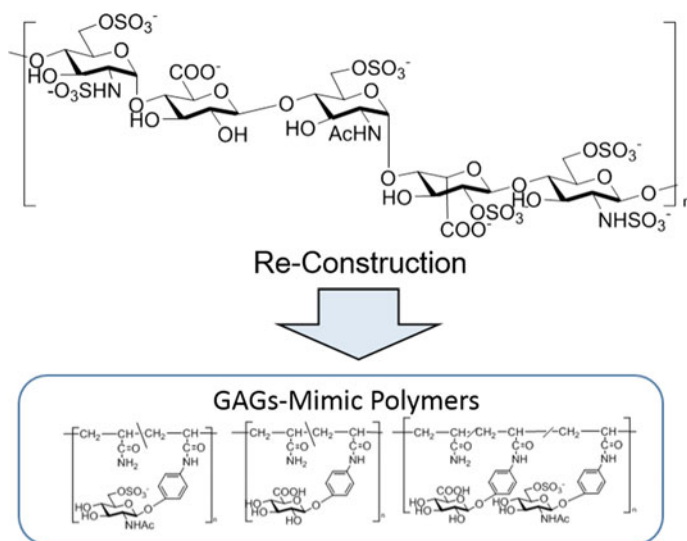
Fig. 2 The concept of GAGs mimicking glycopolymers
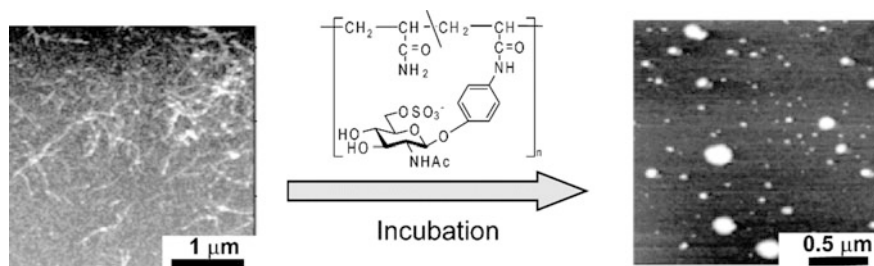
Fig. 3 The toxic nanofibrils of Aβ (1-42) were changed into nontoxic objects by addition of a GAGs mimicking polymer carrying 6-sulfo-GlcNAc

Aβ (1-42) spontaneously aggregates to form protein amyloid. The aggregation of Aβ (1-42) was examined in the presence and in the absence of sulfated glycopolymers. The aggregation of Aβ was monitored by fluorescent intensity of thioflavain T (ThT) that binds to protein amyloid and shows strong fluorescence [22]. The fluorescence of ThT decreased in the presence of the sulfated glycopolymer. The decrease of ThT fluorescence depended on the sugar content of the glycopolymer. The glycopolymer with the modest sugar (10–30%) content showed the inhibition effect on the Aβ (1-42) aggregation. The addition of sulfated glycopolymer also affected the conformation of Aβ. Aβ (1-42) has a β-sheet structure in the aggregated state, but the glycopolymer addition decreased the β-sheet content. In addition, the addition of sulfated glycopolymer changed the morphology of

Fig. 4 GAGs mimicking polymers synthesized by the RAFT polymerization with different molecular weights (*top*). Aβ (1-40) aggregation in the presence of glycopolymers (*bottom*)
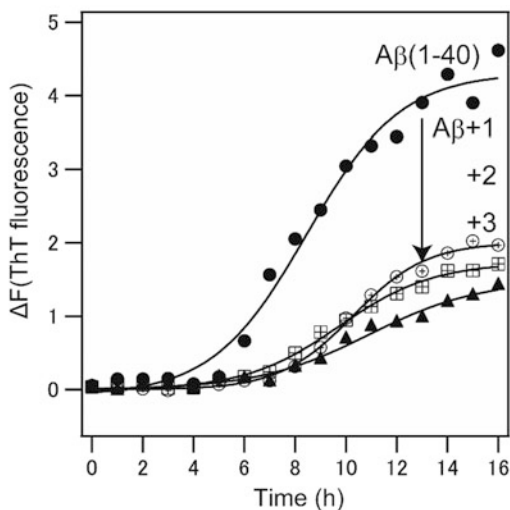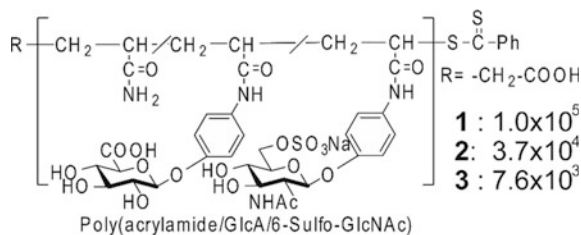
Aβ aggregates from toxic nanofibrils to spherical aggregates. The cytotoxicity of Aβ was examined by in vitro assay of HeLa cells.

It has been reported that GAGs are related the various protein amyloidosis including Alzheimer, prion disease, and Creutzfeldt-Jakob disease [23]. Protein amyloidosis occurs on the cell surfaces, and GAGs have been reported to play important roles in the amyloidosis as triggers and catalysts. Since GAGs are involved in the protein amyloidosis, the GAGs mimicking polymer of poly (6-sulfo-GlcNAc/acrylamide) controlled the amyloidosis of Aβ.

4   Molecular Weights of GAGs Mimicking Polymers and Functions

Functions of GAGs often depend on the molecular weight due to the different physical properties and cell permeability. For example, it has been reported that low-molecular weight heparin has a longer half-life in antithrombus than usual high molecular weight heparin. It is because heparin can exhibit different mechanisms of

antithrombus activity at different molecular weights [24]. It has been also reported that the GAGs molecular weight in protein affects amyloidosis [25].

Since the molecular weight of the glycopolymer can be controlled by living radical polymerization, the glycopolymers with different molecular weights were prepared [26]. There are many methods of living radical polymerization. We employed a reversible addition-fragmentation chain transfer (RAFT) polymerization because it offers an advantage of controlled polymerization in bulky monomers. Despite the known difficulties of the RAFT method to control the molecular weight of polyacrylamide, the glycopolymers with different molecular weights (10^5, 10^4 and 10^3) were prepared. In order to clarify the role of uronic acid, the glycopolymers were prepared using 6-sulfo-GlcNAc and GlcA, with the similar sugar contents.

The function of the glycopolymers was investigated using Aβ (1-40). The velocity and the degree of amyloidosis were measured with fluorescent ThT and the morphology was evaluated by AFM (Fig. 3). The addition of the glycopolymer with 6-sulfo-GlcNAc delayed the beginning of amyloidosis, and the glycopolymer with a low molecular weight induced the stronger delay effect on amyloidosis. The addition of glycopolymer with GlcA had no effect on the starting point and velocity of amyloidosis but reduced the degree of amyloidosis. The glyco-terpolymer with both 6-sulfo-GlcNAc and GlcA showed strong inhibitory effects on the decrease of velocity and degree of amyloidosis. In all cases, the glycopolymers with lower molecular weight showed stronger inhibitory effect, which suggested the validity of the discussed mimicking method (Fig. 4).

The role of each sugar in GAGs was analyzed employing glycopolymers. The influence of nucleation and amyloid fibril elongation were analyzed based on the kinetics of amyloidosis. The glycopolymer with sulfo-GlcNAc inhibited the
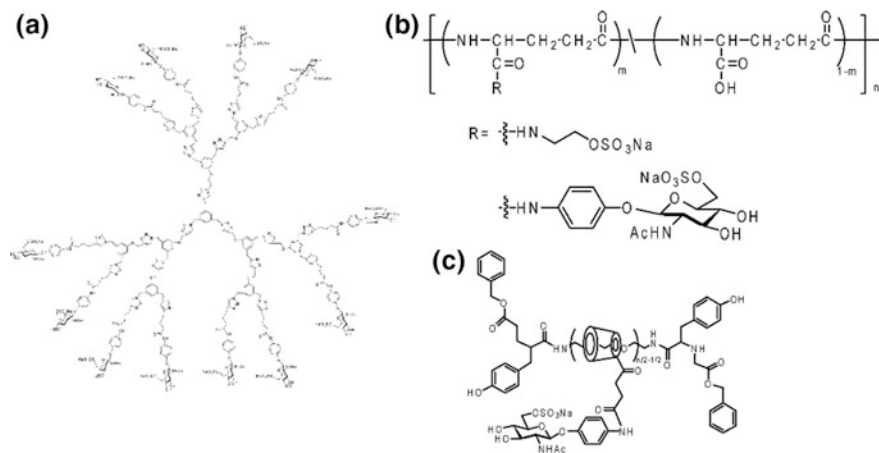
Fig. 5 GAGs mimicking polymers with various polymer backbone. **a** Dendrimer, **b** biodegradable poly(γ-glutamate) and **c** polyrotaxane
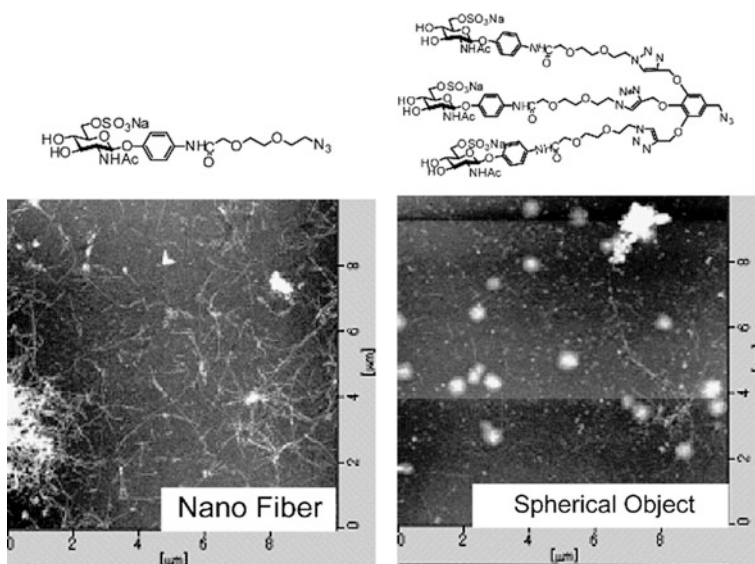
Fig. 6 The structure of GAGs mimicking saccharides and the morphology of Aβ (1-42). The monomeric saccharide induced the nanofiber and trimeric saccharide induced the spherical object

nucleation of amyloid protein by the interaction with the sulfate group and Aβ. The glycopolymer with GlcA modestly inhibited the elongation of the amyloid. The conjugation of the two effects in the terpolymer effectively inhibited the protein amyloidosis. Interestingly, the terpolymer with sulfated sugar and uronic acid offers the best mimic of GAGs in the inhibition of amyloid.

The inhibitory effect was also estimated by a computational chemistry. The interaction of the glycopolymer with Aβ(1-42) is an electrostatic interaction and His13 and His14 are considered to interact with the glycopolymer sulfate group.

5   Glycopolymer Library with Other Sulfated Sugars

The functionalities of GAGs are strongly related to the sugar structure. Since various sugars are involved in the natural GAGs, it is of interest to expand GAGs mimic libraries onto other sulfated sugar monomers. The sugar monomers with region selective sulfation were prepared using a protective group chemistry. 3-Sulfo-, 4-sulfo-, 6-sulfo-, and 3,4,6-sulfo-GlcNAc were prepared, and glycopolymers with those sugar were prepared. The monomers have little different chemical components from the natural GAGs, but the glycopolymer showed the biological activities based on the monomer sugar structure.

Glycopolymers were evaluated for their inhibitory activity on Alzheimer's β-secretase (BACE-1), which is the enzyme producing Aβ proteins from amyloid

precursor protein [26]. It has been reported that GAGs regulate the activity of BACE-1. The Turnbull group reported that the addition of heparin oligosaccharides reduced the activity of BACE-1 [27]. The glycopolymers with sulfated GlcNAc and GlcA were prepared by radical polymerization, and the activity on BACE-1 was measured using fluorescence spectra of a FRET labeled peptide.

The addition of sulfated sacccharides can change the enzyme activity. While the glycopolymer with 3-sulfo-GlcNAc or 4-sulfo-GlcNAc activated the enzyme activity up to 150%, the glycopolymer with 3,4,6-GlcNAc and 6-sulfo-GlcNAc reduced the activity up to 20%. It has been reported that GAGs regulate the orientation of BACE-1 on the cell surface. The results suggest that the regio-selective sulfation of polysaccharide is a key to control the enzyme activity. On the other hand, the sugar monomers were also investigated, but the results did not show significant differences, which suggests the importance of the multivalent effect. Glycopolymers only having 6-sulfo-GlcNAc structures showed the strong inhibitory effect on BACE-1. The strong activity of 6-sulfo-GlcNAc was based on the easy accessibility due to the flexible sulfate group at the 6 position. The glycopolymer with 3,4,6-sulfo-GlcNAc showed the strongest activity.

The glycopolymer binding to BACE-1 was analyzed by docking simulation. The glycopolymers with 3,4,6-sulfo-GlcNAc bound to positively charged protein sites, including the active site of enzyme with electrostatic interactions. The glycopolymers with 3-sulfo- and 4-sulfo- moieties bound to other sites of enzyme also by electrostatic interaction, which increases the enzyme activity.

6 GAGs Mimic with Other Backbones

Considering the potential biological application of a polymer as a drug, the biocompatible and biodegradable polymers are more appropriate. The glycopolymers with 6-sulfo-GlcNAc were prepared by conjugation to poly(γ-glutamate) by taking advantage of the amide formation [28]. The addition of the sulfated glycopolymer with poly(γ-glutamate) inhibited the aggregation of Aβ (1-40) (Fig. 5a). The inhibitory effect was comparable to that of the glycopolymer with a polyacrylamide backbone. Also, another GAGs mimicking polymer of poly(γ-glutamate) with a sulfonate group was prepared. It turns out that GAGs mimic without the sugar moiety does not work on the Aβ aggregation. This result proves that though the chemical structures of glycopolymers were totally different from the native GAGs, the saccharide moiety is indispensable to GAGs mimicking functions.

Various polymer backbones provide the diversity of GAGs mimics (Fig. 5). The advantage of using the glycopolymer is a facile preparation of the molecular library. However, it is difficult to obtain the glycocluster with a clear structure. In order to obtain the uniform GAGs mimics, the glycodendrimers with 6-sulfo-GlcNAc were prepared by utilizing the click chemistry approach [29]. The glycodendrimer provides the defined sugar cluster, and there is no polydispsesity in the molecules. The

glycodendrimer with higher generation (2nd and 3rd) displayed the densely packed sugars, and the glycodendrimer with higher generation showed the inhibitory effect on the Aβ aggregation.

7 GAGs Mimic Array on Gold Substrate

GAGs mimicking compounds in this study were polymers in which their chemical structure can be modified and diverse. The conjugation of the glycopolymer to the interface enables the detailed analyses with biosensing techniques such as surface plasmon resonance (SPR) and quartz crystal microbalance (QCM). The Seeberger group reported the use of the GAGs oligosaccharide arrays on the substrate [10]. Suda et al. [13] reported the use of GAGs mimicking arrays with dendrimer to analyze the protein–sugar interaction.

We prepared GAGs mimicking dendrimer arrays and analyzed the interaction with Alzheimer Aβ (1-42) (Fig. 6). The glycodendrimer arrays were prepared by orthogonal reactions belonging to the click chemistry [30]. Tri-, di- and monovalent sulfated sugars were synthesized and immobilized on the gold substrate. The interaction with Aβ (1-42) was investigated by AFM, SPR, and a cell cytotoxicity assay. The interaction with Aβ was weak in monovalent state but amplified by divalent and trivalent sugars. The morphology of Aβ was specific to sugar arrays. While Aβ formed a nano-fibril on monovalent sugar, it formed a toxic spherical object on divalent and trivalent sugars. The results showed that the morphology of Aβ was controlled by the interaction with GAGs, which suggests that the multivalency of sulfated sugars is important in Alzheimer disease.

Thus, it can be said that the GAGs mimicking arrays are effective when investigating the sugar structure, the multivalency, and the biochemical mechanism.

8 Conclusion

Glycosaminoglycans (GAGs) are important polysaccharides in the living systems. While successful total syntheses of GAGs oligosaccharides have been reported, it is still difficult to obtain GAGs. In this investigation instead of total synthesis to form GAGs, glycopolymers mimicking GAGs were prepared. Glycopolymers are polymers with pendant saccharides and exhibit a strong molecular recognition due to the multivalency. GAGs mimicking glycopolymers were prepared by polymerization of acrylamide derivatives carrying sulfated *N*-acetyl glucosamine (GlcNAc) and the resulting products' interaction with proteins was investigated. The glycopolymer libraries were prepared with varying molecular weight, sugar structure, and sugar ratios. The GAGs glycopolymer with biodegradable backbone, dendrimer, and arrays were also prepared. The inhibitory activity of the glycopolymers on the Alzheimer amyloid beta peptides was studied in detail. The glycopolymer with

sulfated GlcNAc inhibited the aggregation of Aβs [(1-40) and (1-42)] and the multivalency of sulfated GlcNAc was the key to the interaction. The activity depends on the chemical structure of glycopolymers. The sulfated saccharide function was correlated to the functions of native GAGs.

References

1. Taylor ME, Drickamaer K (2002) Introduction to glycobiology. Oxford Press, London
2. Rudd TR, Skidomore MA, Guerrini M, Hricovini M, Powell AK, Siligardi G, Yates EA (2010) The conformation and structure of GAGs: recent progress and perspectives. Curr Opin Struct Biol 20:567–574
3. Kaplan KL, Francis CW (2002) Direct thrombin inhibitors. Semin Hematol 39:187–196
4. Farrell E, O'Brien FJ, Doyle P, Fischer J, Yannas I, Harley BA, O'Connell B, Prendergast PJ, Campbell VA (2006) A collagen-glycosaminoglycan scaffold supports adult rat mesenchymal stem cell differentiation along osteogenic and chondrogenic routes. Tissue Eng 12:459–468
5. Tamura J (2001) Recent advances in the synthetic studies of glycosaminoglycans. Trends Glycosci Glycotech 13:65–89
6. Suda Y, Marques D, Kermode JC, Kusumoto S, Sobel M (1993) Structural characterization of heparin's binding domain for human platelets. Thromb Res 69:501–550
7. Maccarana M, Casu B, Lindahl U (1993) Minimal sequence in heparin/heparan sulfate required for binding of basic fibroblast growth factor. J Biol Chem 268:23898–23905
8. Koshida S, Suda Y, Sobel M, Ormsby J, Kusumoto S (1999) Synthesis of heparin partial structures and their binding activities to platelets. Bioorg Med Chem Lett 9:3127–3132
9. Hu YP, Lin SY, Huang CY, Zulueta MML, Liu JY, Chang Y, Hung SC (2011) Synthesis of 3-O-sulfonated heparan sulfate octasaccharides that inhibit the herpes simplex virus type 1 host–cell interaction. Nat Chem 3:557–563
10. de Paz JL, Noti C, Seeberger PH (2006) Microarrays of synthetic heparin oligosaccharides. J Am Chem Soc 128:2766–2767
11. Habuchi H, Habuchi O, Kimata K (1998) Biosynthesis of heparan sulfate and heparin: how are the multifunctional glycosaminoglycans built up? Trends Glycosci Glycotechnol 10:65–80
12. Rawat M, Gama CI, Matson JB, Hsieh-Wilson LC (2008) Neuroactive chondroitin sulfate glycomimetics. J Am Chem Soc 130:2959–2961
13. Suda Y, Arano A, Fukui Y, Koshida S, Wakao M, Nishimura T, Kusumoto S, Sobel M (2006) Immobilization and clustering of structurally defined oligosaccharides for sugar chips: an improved method for surface plasmon resonance analysis of protein-carbohydrate interactions. Bioconjug Chem 17:1125–1135
14. Miura Y, Hoshino Y, Seto H (2016) Glycopolymer nanobiotechnology. Chem Rev 116:1673–1692
15. Kobayashi K, Sumitomo H, Ina Y (1985) Synthesis and functions of polystyrene derivatives having pendant oligosaccharides. Polymer J 17:567–575
16. Cairo GW, Gestwicki JE, Kanai M, Kiessling LL (2002) Control of multivalent interactions by binding epitope density. J Am Chem Soc 124:1615–1619
17. Miura Y, Yasuda K, Yamamoto K, Koike M, Nishida Y, Kobayashi K (2007) Inhibition of Alzheimer amyloid aggregation with sulfated glycopolymers. Biomacromol 8:2129–2134
18. Sasaki K, Nishida Y, Uzawa H, Kobayashi K (2003) N-Acetyl-6-sulfo-D-glucosamine as a promising mimic of N-acetyl neuraminic acid. Bioorg Med Chem Lett 13:2821–2823
19. Selkoe DJ (1994) Normal and abnormal biology of the beta-amyloid precursor protein. Annu Rev Nuerosci 17:489–517

20. Glenner GG, Wong CW (1984) Alzheimer's disease and down's syndrome: Sharing of a unique cerebrovascular amyloid fibril protein. Biochem Biophys Res Commun 122:1131–1135
21. Levine HIII (1999) Quantification of b-sheet amyloid fibril structures with thioflavin T. Methods Enzymol 309:274–284
22. Snow AD, Willmer J, Kisilevsky R (1987) Sulfated glycosaminoglycans: a common constituent of all amyloids? Lab Invest J Tech Methods Pathol 56:120–123
23. Hirsh J, Levine M (1992) Low molecular weight heparin. Blood 79:1–17
24. Bergamaschini L, Rossi E, Storini C, Pizzimenti S, Distaso M, Perego C, Luigi AD, Vergani C, de Simoni MG (2004) Peripheral treatment with enoxaparin, a low molecular weight heparin, reduces Plaques and β-amyloid accumulation in a mouse model of Alzheimer's disease. J Neurosci 24:4181–4186
25. Miura Y, Mizuno H (2010) Interaction analyses of amyloid BETA peptide (1-40) with glycosaminoglycan model polymers. Bull Chem Soc Jpn 83:1004–1009
26. Nishimura Y, Shudo H, Seto H, Hoshino Y, Miura Y (2013) Syntheses of sulfated glycopolymers and analyses of their BACE-1 inhibitory activity. Bioorg Med Chem Lett 23:6390–6395
27. Patey SJ, Edwards EA, Yates EA, Turnbull JE (2006) Heparin derivatives as inhibitors of BACE-1, the Alzheimer's β-secretase, with reduced activity against factor Xa and other proteases. J Med Chem 49:6129–6132
28. Fukuda T, Kawamura M, Mizuno H, Miura Y (2013) Glycosaminoglycan model polymers with poly (γ-glutamate) backbone to inhibit aggregation of β-amyloid peptide. Polymer J 45:359–362
29. Miura Y, Onogi S, Fukuda T (2012) Syntheses of sulfo-glycodendrimers using click chemistry and their biological evaluation. Molecules 17:11877–11896
30. Fukuda T, Matsumoto E, Onogi S, Miura Y (2010) Aggregation of Alzheimer amyloid β peptide (1–42) on the multivalent sulfonated sugar interface. Bioconjugate Chem 21:1079–1086

Methods for the High Resolution Analysis of Glycoconjugates

Christopher Gray and Sabine L. Flitsch

Abstract Glycans and their conjugates form the largest and most diverse class of biological molecules found in nature. These glycosides are vital for numerous cellular functions including recognition events, protein stabilisation and energy storage. Additionally, abnormalities within these structures are associated with a wide range of disease states. As a result, robust analytical techniques capable of in depth characterisation of carbohydrates and their binding partners are required. This chapter provides an overview of currently used analytical techniques, focussing on chromatographic and mass spectrometry-based methods.

1 Introduction

Glycosylation involves the enzymatic transfer of a carbohydrate from a donor molecule to a substrate such as a protein, lipid or another carbohydrate, forming elongated and often branched glycoconjugate structures. Diverse varieties of these glycoconjugates coat the surface of all cells and act as receptors for glycan-binding species such as lectins, antibodies or pathogens [1, 2]. Glycosylation is the most prevalent post-translational modification (PTM) observed within nature; in fact, it is thought that greater than half of all known proteins are glycosylated [3]. It is widely reported that these glycans are vital for regulation of many biological interactions such as cell–cell recognition, [4] cell adhesion [5], immune response [6], infection [7, 8] and fertilisation [9–11]. It has also been shown that aberrant glycosylation is related to several diseases including cancer [12–17], muscular dystrophy [5] and pancreatitis [18, 19]. Even subtle changes in carbohydrate structure can result in vastly different interactions, make them susceptible to proteolysis or alter glyco-conjugate tertiary structures and thus affects the observed biological response [5, 15, 17, 20–25]. It remains unclear a priori how changes within a glycan structure

C. Gray · S.L. Flitsch (✉)
School of Chemistry and Manchester Institute of Biotechnology, The University of Manchester, 131 Princess Street, Manchester M1 7DN, UK
e-mail: sabine.flitsch@manchester.ac.uk

will affect the resultant biological function [26]. Knowledge of these structure–function relationships could enable the development of novel therapeutics or diagnostics [27–29]. Additionally, carbohydrates are foreseen as routes to novel fuels and materials [30–32], providing further requirement for rigorous analytical approaches to characterise them.

However, there are inherent difficulties in analysing the 'glycome', the entire complement of glycans and their glycoconjugates produced by an organism under specified conditions of time, space and environment [2], compared to the widely researched genome and proteome. Firstly, glycan structures are not directly encoded from genetic information. Moreover, glycosylated macromolecules tend to exist in multiple glycoforms and are often low abundant compared to their non-glycosylated counterparts, thus requiring enrichment prior to analysis. These glycoforms vary depending on conditions, such as disease state [18], age [33], and gender [9, 34] further increasing the complexity of glycome analysis. Finally, the elucidation of glycan chemical structures using most analytical techniques is extremely challenging. Glycans can be composed of a far greater number of natural monosaccharide building blocks (several hundred) compared to the 4 nucleotides and 20 essential amino acids found in DNA and proteins, respectively [2], although there is a core of 10 major monosaccharide units found within vertebrates (Fig. 1).

These monosaccharide building blocks are often stereo- or regio-isomers of one another (unlike nucleotides or amino acids) making their characterisation more
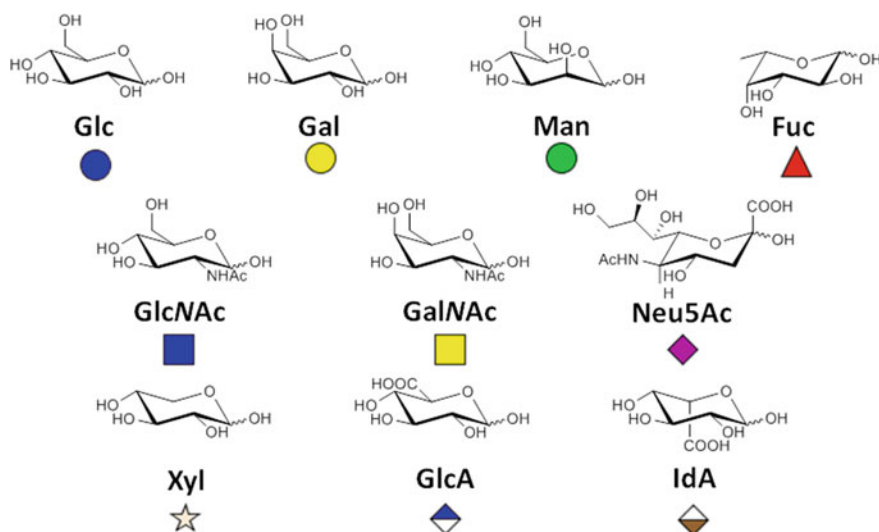
Fig. 1 The common monosaccharides found in vertebrates including the shorthand used by the Consortium for Functional Glycomics. Glc = Glucose, Gal = Galactose, Man = Mannose, Fuc = Fucose, GlcNAc = N-Acetylglucosamine, GalNAc = N-Acetylgalactosamine, Xyl = Xylose, GlcA = Glucuronic acid, IdA = Iduronic acid and Neu5Ac = N-acetylneuraminic acid. Glc, Gal and Man isomers form the hexoses (Hex), GlcNAc and GalNAc isomers the N-acetylhexosamines and GlcA and IdA isomers the hexuronic acids

analytically challenging [35]. Additionally, the glycosidic bond formed between monosaccharides can adopt two different configurations, namely α- and β-, where the binding monosaccharide lies either above the plane of ring, or planar to the ring, respectively (Fig. 2). This linkage can also form at several positions resulting in the formation of branched structures, which are primarily linear combinations of amino acids and nucleotides, respectively [2, 36, 37]. Finally, these carbohydrate rings can potentially exist as furanose or pyranose forms.

There are two common types of protein glycosylation, (not to be confused with the chemically similar process glycation which occurs non-enzymatically): *N*-glycosylation and *O*-glycosylation. Typical eukaryotic *N*-glycosylation involves enzymatic *en bloc* transfer of the glycan Glc$_3$Man$_9$GlcNAc$_2$ from dolichol phosphate to asparagine residues within in the consensus peptide sequence Asn-Xxx-Ser/Thr (where Xxx is any amino acid except proline) [2, 38, 39]
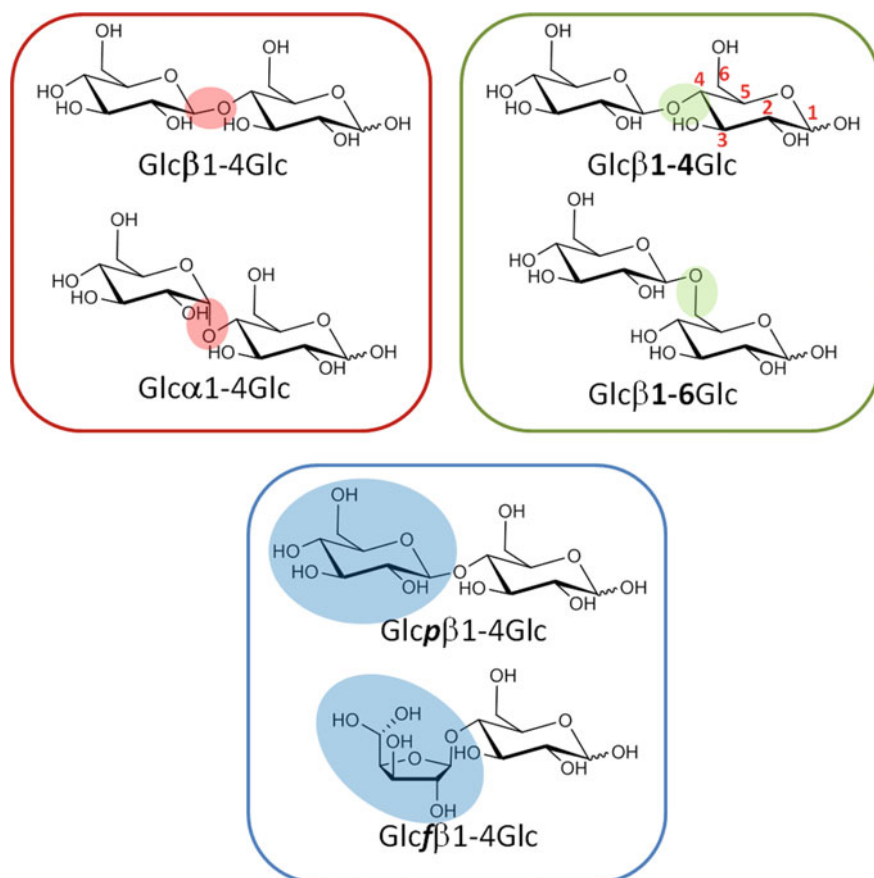
Fig. 2 Scheme depicting two α-/β-stereoisomers (*red*), two regioisomers (*green*) and two pyranose/furanose isomers (*blue*)

although this is not always the case [3]. The $Glc_3Man_9GlcNAc_2$ modification is synthesised first on the cytoplasmic face of the endoplasmic reticulum (ER) membrane in a stepwise manner from dolichol phosphate to $Man_5GlcNAc_2$, This is then 'flipped' so that it faces into the ER lumen, where additional glycosyl transferases modify this species producing $Glc_3Man_9GlcNAc_2$ that is then transferred in a co-translational event *en bloc*, mediated by an oligosaccharyltransferase, to the asparagine acceptor of the forming polypeptide chain [2, 40]. Various glycosidases and glycosyltransferases further modify this glycoconjugate within the Golgi apparatus, generating a structure that can be classed as one of three classes (Fig. 3): **high mannose**, where the *N*-glycan core ($Man_3GlcNAc_2$) is comprised of only branched Man and GlcNAc residues; **complex** where the cores antennae are functionalised with other saccharides including galactose, fucose and sialic acid; and **hybrid** where a single antennae from the core structure is functionalised with mannose residues (α1–6 arm) and the others with complex structures. *N*-glycosylation primarily occurs on secreted or membrane-bound proteins within eukaryotes or archaea [41] and was later shown to occur on proteins within the Gram-negative bacterium *Campylobacter jejuni* [42]. Unlike *N*-glycosylation in eukaryotes, there is no conserved carbohydrate sequence transferred *en bloc* to proteins (i.e. $Glc_3Man_9GlcNAc_2$ for eukaryotes), although certain motifs seem to be important such as the presence of a reducing acetamido group for bacterial *N*-oligosaccharide transferases [40]. Also archaea and bacterial *N*-glycans are structurally different compared to eukaryotic *N*-glycans possessing, for example, hexuronic acids or bacilosamine residues, respectively [40].

O-glycosylation occurs post-translationally onto the –OH group of typically serine or threonine residues within proteins. Unlike *N*-glycans, *O*-glycans are synthesised stepwise on glycoproteins and tend to be much shorter than their *N*-glycan counterparts consisting of just a single residue in some cases [38]. Within Eukaryotes, initial transfer of α-GalNAc, which is the most common addition, and α-Fuc to proteins occurs in the Golgi, whereas *O*-mannosylation is initiated in the ER and *O*-GlcNAcylation and glucosylation occurs in the cytosol (and nucleus for *O*-GlcNAc) [2, 38, 43]. These carbohydrates may then be extended by a series of glycosyltransferases in the Golgi. For the most common type of *O*-glycosylation, *O*-GalNAc, a series of eight common core structures exist (Fig. 4). Unlike *N*-glycosylation no consensus motif has yet been identified for *O*-glycosylation, making *O*-glycosylation site identification more challenging [44]. For specific classes of *O*-glycosylation though, protein sequence preferences have been identified allowing prediction of potential *O*-glycosylation sites [45]. Certain proteins, known as mucins, are heavily *O*-glycosylated with GalNAc: these will be the focus of Chap. "[3,3]-Sigmatropic Rearrangement as a Powerful Synthetic Tool on Skeletal Modification of Unsaturated Sugars".

Glycan structures are routinely assigned using minimal analytical information, based on the assumption that their biosynthetic pathways are highly conserved, which is not necessarily always the case [19, 46, 47]. Alternatively, structures can be directly elucidated to varying degrees by NMR, tandem mass spectrometry or sequential glycosidase treatment followed by chromatographic separation [15, 46, 48–50]. However, these approaches are limited by either sensitivity or specificity.
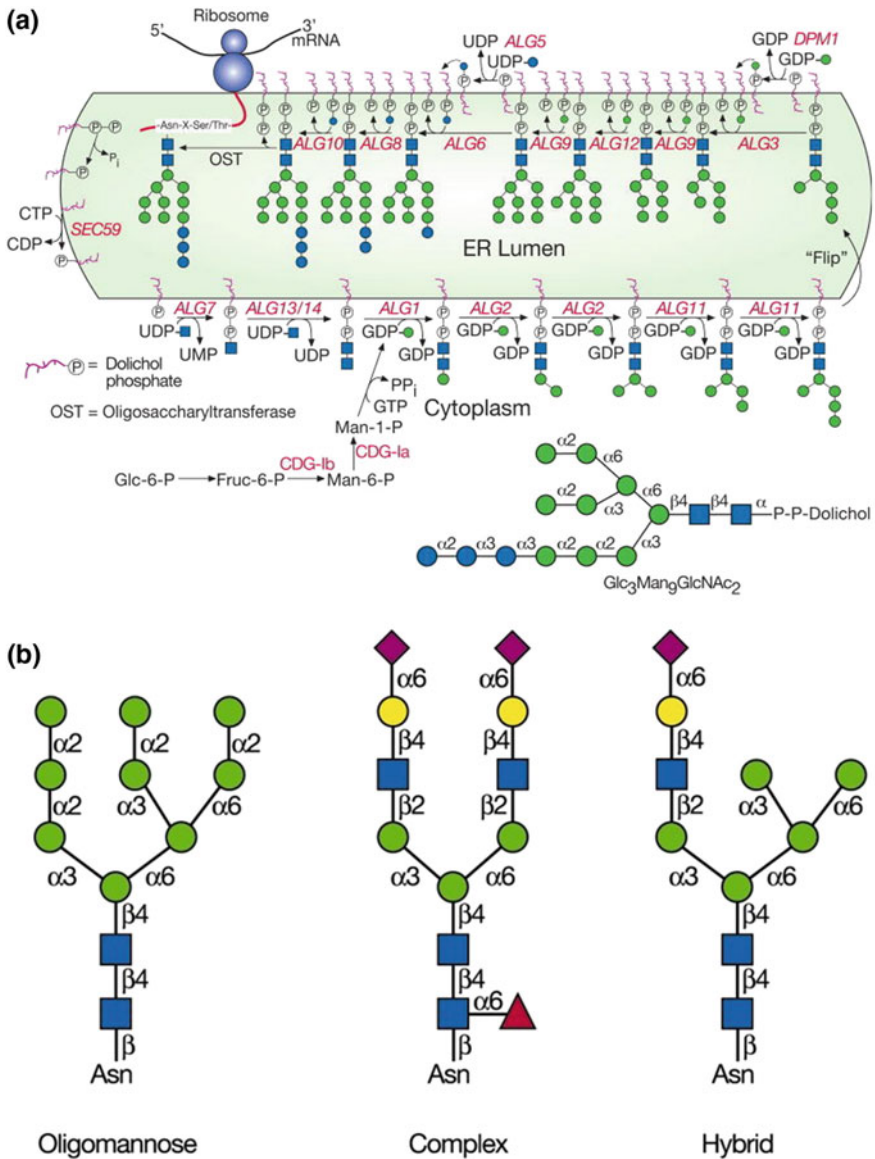
Fig. 3 The biosynthetic pathway by which the oligomannose structure transferred to proteins forming *N*-glycans is generated (**a**). Various glycosidases and glycosyltransferases then act on this glycan to produce the different types of *N*-glycans observed within eukaryotic organisms (**b**). Figures are taken from Varki et al. [2]

Fig. 4 Basic core *O*-Gal*N*Acylated structures [44]
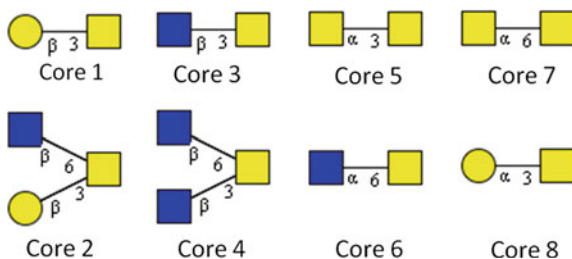
Consequently, high-throughput analytical methodologies capable of sequencing glycan structures and further identifying protein binding partners are highly sought after. Mass spectrometry (MS)-based techniques offer the capability of elucidating structural information on a wide range of biological analytes including peptides, proteins and glycans in a high-throughput manner. However, they are often unable to unambiguously assign carbohydrate structures given that most monomeric building blocks are simple epimers of one another. Integration of ion mobility spectrometry, a technique that separates ions by their rotationally averaged collision cross section, with (tandem) mass spectrometry offers the potential capability to separate these isomeric precursors and product ions, which would greatly aid glycan characterisation. Of particular interest is the ability to separate and determine the cross section of isomeric mono-/disaccharide product ions, which crucially could be indicative of the stereochemistry of the residue, the anomeric configuration, ring size and regiochemistry. Therefore, IM-MS has the potential to fill a gap in the Glycomics community, namely a high-throughput carbohydrate sequencing strategy.

High-throughput strategies to identify proteins that bind to glycans are also of great interest. Most high-throughput strategies involve arraying thousands of glycans to a solid support, incubating them with purified proteins, washing and then visualisation by fluorescence or radiation. However, these approaches rely on purified material and incorporation of a fluorescent or radioactive tag. (Tandem) Mass spectrometry of proteins adhered to the arrays, or peptides resulting from on-chip proteolysis, offers a rapid and direct means to unambiguously characterise unlabelled bound proteins, even from endogenous mixtures.

Multiple MS-based strategies underpin the ability to characterise glycans and their binding partners. Glycans and their conjugates are initially ionised, principally by matrix-assisted laser desorption ionisation (MALDI) or electrospray ionisation (ESI), as either cation or anion adducts and then can be made dissociate if required [51]. These precursor or product ions are then separated based on their mass-to-charge (*m/z*) ratio mass and subsequently detected.

2 Mass Spectrometry Techniques Applied to Glycoconjugate Analysis

2.1 Ionisation Techniques

2.1.1 Electrospray Ionisation (ESI)

Electrospray ionisation (ESI), or nanospray ESI (nESI) for nanoflow, is a soft ionisation technique, i.e. it causes little to no fragmentation of the molecules during ionisation, including large molecules such as proteins [52] and protein assemblies [53, 54]. It is also the most commonly used ionisation strategy in biomolecule analysis as it can be operated at atmospheric pressure and can be easy coupled to liquid chromatography meaning that a mixture of analytes can be separated on-line prior to ESI-MS analysis. However, samples can be directly infused by syringe pumps or static nESI tips. ESI and nESI operate at $\mu L\ min^{-1}$ and $nL\ min^{-1}$ flow rates, respectively, permitting analysis of low amounts of analyte suitable for biological samples. (n)ESI emitters are subjected to a high potential electric field (1–5 kV) [55]. At the capillary tip a large electric field is formed ($\sim 10^6\ V\ m^{-1}$) resulting in the charged analyte solution being polarised, i.e. in positive ion mode a negative potential is applied across the capillary therefore positively charged molecules are drawn to the end of the tip. This field also causes the meniscus at the capillary nozzle to be perturbed forming a cone (Taylor cone) as shown in stage 1 Fig. 5.

When the force the electric field exerts is high enough, the tip of the Taylor cone is destabilised resulting in the formation of a jet of charged droplets that repel one another, causing them to spread out orthogonally and accelerate towards the counter electrode by electrostatic attraction [52, 57, 58]. These droplets evaporate, which may be aided by an inert drying gas such as nitrogen and elevated temperatures, until the point where the Coulombic repulsive forces of the charged analyte destabilise the droplet. This occurs slightly below the Rayleigh limit where repulsion equals the surface tension. At this point, the droplets 'explode' forming even smaller droplets. The precise mechanisms of ion formation from these droplets are not clear and are thought to be dependent on the nature of the analyte ion. Three postulated mechanisms are the ion evaporation model (IEM) [59], the charged residue model (CRM) [60] and the chain ejection model (CEM) [56, 61]. The IEM, believed to be prevalent for small molecules (e.g. glycans), involves ejection of a small analyte ion from the surface of a small (nm radius) charged droplet when the Rayleigh limit is sufficiently high. This process is kinetically disfavored for large molecule ions, which are thought to form by the CRM [56]. In the CRM, solvent continually evaporates from the droplet containing the analyte. As the final solvent shell evaporates, the remaining charge within the droplet is transferred to the analyte. During this process, analytes are charge reduced so they maintain the Rayleigh limit by IEM ejection of solvated protons and small ions [56]. Finally the CEM, which suggests unfolded 'chains' migrate to the droplet surface when the
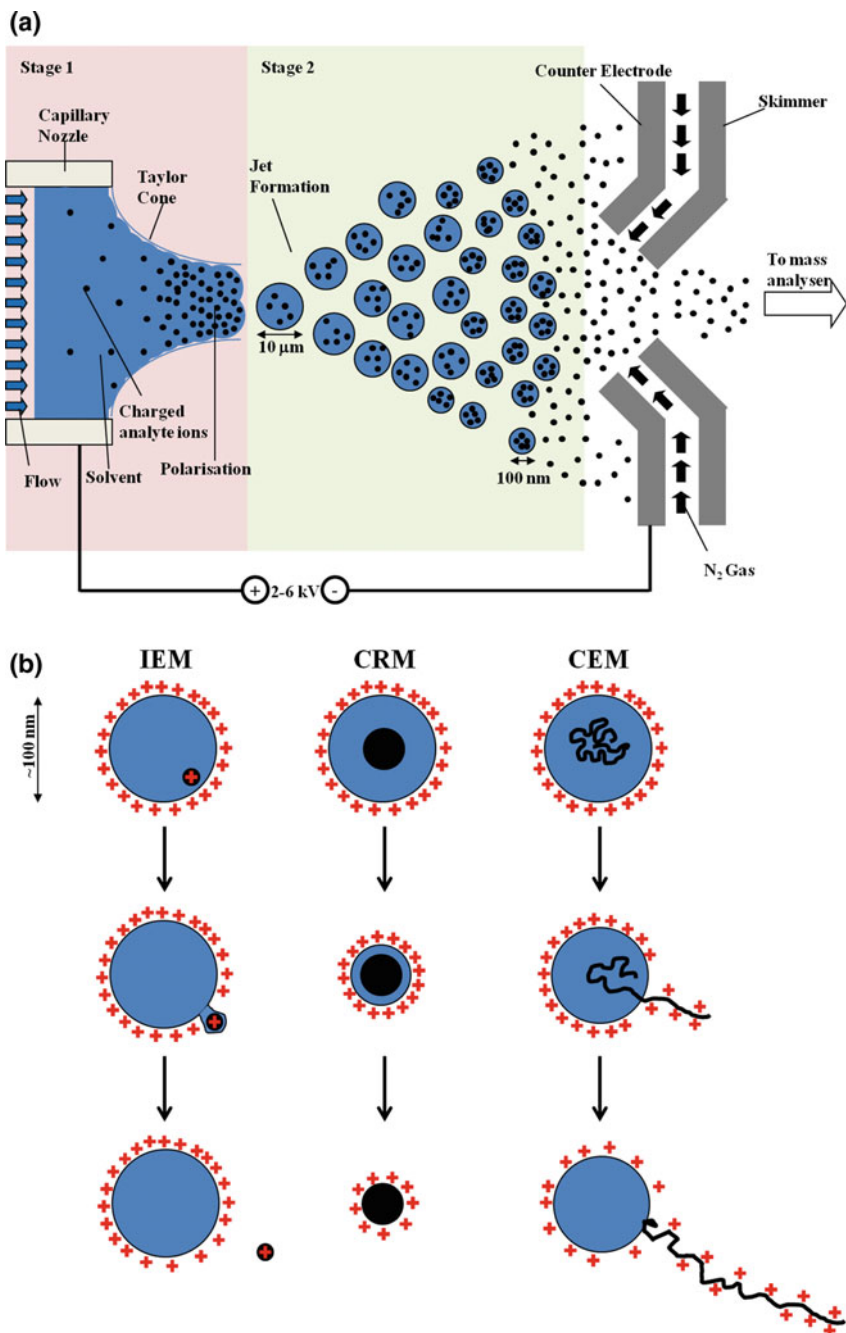
Fig. 5 Schematic depicting the ESI ionisation mechanism (**a**) and schemes depicting ion formation models for the ion evaporation model (IEM), charged residue model (CRM) and the chain ejection model (CEM) (**b**) (redrafted version of figure from Ref. [56])

Rayleigh limit is reached. This charged chain is then liable to be ejected from the surface of the droplet. Sequentially, more and more of the charged chain is ejected from the droplet until the remainder of the droplet evaporates or the entire analyte is ejected [62]. This mechanism is believed to occur for large disordered analytes, such as polymers or disordered proteins and accounts for the observed higher analyte charge states than the CRM would predict [56]. Regardless of the ion formation mechanism, the composition of the sprayed solution will impact the formed ions and as a result can perturb the measured mass spectrum. ESI solutions often contain organic solvents such as methanol or acetonitrile to lower the surface tension of electrospray droplets improving desolvation and thus ionisation. Furthermore, these solutions tend to contain volatile acids, such as formic acid, for positive ion mode or bases, such as ammonia, for negative ion mode providing a source of protons or a proton sink, respectively. Conversely, molecules can be promoted to form metal adducts by doping the solution with salts such as sodium formate or lithium chloride. Structured proteins can also be stabilised through buffering with volatile salts such as ammonium acetate [53].

ESI is advantageous as it can produce large (kDa to MDa) ions, even from non-volatile, thermally labile compounds and is typically compatible with conventional liquid chromatography techniques [52]. Also given that these ionised species are typically multiply charged, their m/z values tend to fall within the operating range of most mass spectrometers [58]. However, ESI, like most ionisation techniques, suffers from a low tolerance towards salts, therefore, samples must be desalted prior to analysis. Finally, as the technique is very sensitive, the spray chamber must be kept very clean to avoid contamination and signal suppression [63].

2.1.2 MALDI

Matrix-assisted laser desorption ionisation (MALDI) is, like ESI, a soft ionisation technique, although involves desorption of analyte ions from the solid phase induced by irradiation with a pulsed UV laser at 337 or 355 nm for 1–10 ns [58, 64, 65]. Interestingly, unlike ESI, analyte ions generated tend to only be singly charged, greatly simplifying mass spectra. Analytes are co-crystallised with an excess of a UV absorbing organic acid matrix (1:5000), on an inert metal (typically a stainless steel or gold) target, forming ideally homogenous crystals. The matrix has two main purposes, firstly it separates analyte molecules preventing analyte–analyte interactions that may hinder desorption and ionisation, and secondly and more importantly, it absorbs most of the UV radiation from the laser pulse protecting the analyte (Fig. 6). The matrix may also act as proton sources or sinks depending on the operation mode [66]. Different matrices work better for ionisation of different classes of analytes, so the choice of the matrix depends on the analyte being studied. For example sinapinic acid (SA) works well with proteins, α-cyano-4-hydroxycinnamic acid (CHCA) is good for small proteins and peptides, 2,4,6-trihydroxyacetophenone (THAP) is good for glycans and 2,5-dihydroxybenzoic acid (DHB) is commonly
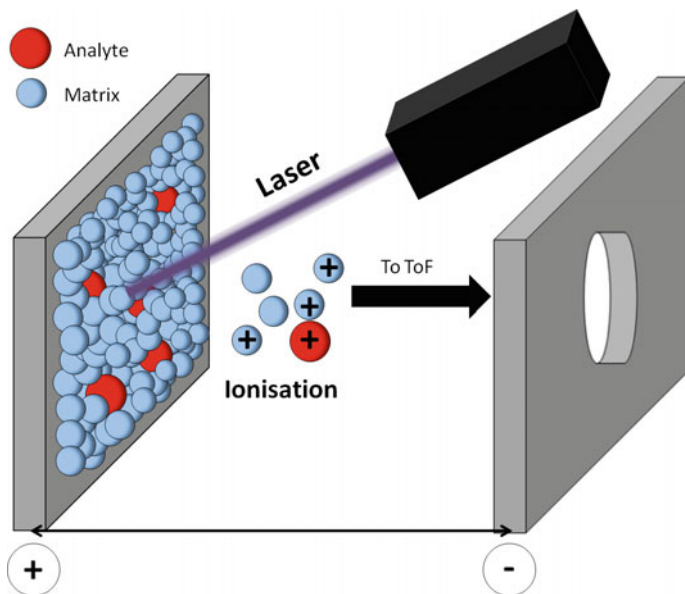
Fig. 6 Schematic of MALDI ionisation of positively charged analytes (*red*) co-crystallised with an excess of UV-absorbing matrix (*blue*)

used for all organic molecules and is particularly useful for those analytes which are labile for example, covalently modified peptides. Upon UV absorption the matrix sublimes *in vacuo* causing rapid expansion, which results in analyte and photoionised matrix molecules being ejected into the gas phase, although most remain uncharged (primary ionisation) [58, 67]. There are several models for ionisation with none being completely accepted. The two most accepted are the cluster '(refined) Lucky Survivors' and the gas-phase proton transfer [64, 68–72]. The 'Lucky Survivors' model states that the singly and multiply charged analyte ions are pre-formed in the matrix solution and retain their solution charge during co-crystallisation. Following a UV laser pulse, the matrix and analyte sublime and the multiply charged analyte ions undergo secondary neutralisation reactions with free electrons until they become singly charged [73]. The gas-phase protonation model suggests secondary collisions, within the MALDI plume, between the neutral analyte and charged UV matrix molecules in the gas plume results in ionisation of the analyte (secondary ionisation). The matrix charge is either preformed in solution or resulting from the laser energy being absorbed by the matrix (pooling—Coupled Chemical and Physical Dynamics model) [70, 74]. Recently, it has been reported that it is highly likely that both ionisation mechanisms are involved [73]. This ionisation technique is advantageous as it is quick and produces little fragmentation of analyte ions. However, ions produced by MALDI are metastable, and therefore are liable to dissociating during traversing the field-free region of the Time-of-Flight (ToF) mass analyser that is typically coupled with a MALDI source. Metastable

decay is more prominent for larger species whose time-of-flight is greater. This causes a loss in sensitivity due to a reduction in the number of intact analyte ions reaching the detector [55]. Nevertheless, this post-source decay can be exploited in MS^2 studies. Another disadvantage of MALDI is the limitation of analysing ions below *m/z* of 500 due to the excess number of matrix-derived ions that can saturate the signal. Salts and other buffers also hinder MALDI, although to a lesser extent compared to ESI [58, 63].

2.2 Mass Analysers

2.2.1 Quadrupole

Quadrupoles are composed of four parallel rods, ideally with a hyperbolic cross section, arranged in a diamond shape. Each opposing rod is electrically connected to one another and has the opposite polarity to the pair of rods perpendicular to them (Fig. 7). Both pairs possess a direct current (DC) potential and are overlaid with an alternating radio-frequency (RF) potential. In the case where the analyte ions are positive, when the RF and DC potential in the *x–z* plane is positive the analyte ions are focussed into the centre of the quadrupole.
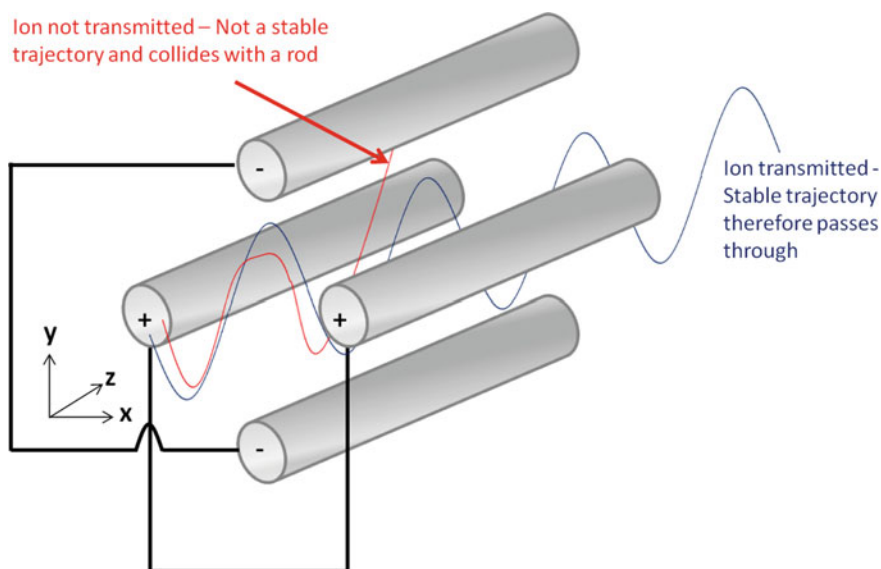
Fig. 7 Scheme depicting transmission though a quadrupole mass analyser and the respective potentials applied to the rods at a certain point in time. Depending on the magnitude and polarity of the applied potentials, certain ions of a given *m/z* posses stable trajectories and are transmitted through the quadrupole (*blue*) unlike others that collide with the rods and annihilate (*red*)

When the RF polarity switches to negative (more than the DC positivity), the ions are accelerated towards the *x–z* plane rods. Ions of low mass and higher charge will be accelerated more than heavier ions of lower charge and may collide with the electrodes at which point they will be discharged and pumped off as neutral species. As a result, only ions with high *m/z* are transferred (referred to as the high pass mass filter). The *y–z* plane rods operate at an opposite DC potential to those in the *x–z* plane (for this example negative). In this case, ions of large *m/z* are more likely to collide with the rods as they are less likely to respond to the focussing that occurs when the alternating RF potential becomes positive unlike ions with lower *m/z*. Therefore, only ions with a low *m/z* are transferred (the low pass mass filter). The net effect of these two filters is that only a narrow range of ions will have a stable trajectory and thus pass through the quadrupole [75]. If the DC and RF voltage is increased whilst keeping the DC to RF ratio constant, new stable trajectories will be created for ions with an increasing *m/z* value, allowing a range of ions of different *m/z* values to be scanned [58]. Controlling the DC and RF ratio plays a significant role in the resolution of this mass analyser with a lower DC to RF ratio producing a lower mass resolution. The stability of ions passing through a (hyperbolic) quadrupole field possessing both DC and RF potential can also be expressed in terms of the Mathieu equations (Eqs. 1 and 2) [75].

$$a_x = -a_y = \frac{4eU}{mr_0^2\omega^2} \tag{1}$$

$$q_x = -q_y = \frac{2eV}{mr_0^2\omega^2} \tag{2}$$

where U and V are the DC and alternating RF potential, respectively, ω is angular frequency of the applied RF, e is the electronic charge, m is the mass, r_0 is this distance from the z-axis (i.e. centre of the four rods) and a and q represent points in space. Plotting solutions of a and q against one another generates a graph defining the a and q values at which ions possess a stable trajectory through the quadrupole (Fig. 8).

Quadrupole mass analysers are advantageous as they are small, inexpensive, have a high scanning speed and when three are coupled together (triple quadrupole instrument) they are capable of selective reaction monitoring experiments (SRM). In a SRM experiment, Q1 and Q3 only allow transmission of a specified precursor and product ion, respectively, improving detection of a specific ion. On the other hand, they have a limited mass resolution and have a finite mass range unlike linear ToFs (Sect. 1.2.2.3) [76].

2.2.2 Quadrupole Ion Trap

Unlike quadrupole (and the later discussed ToF) mass analysers, quadrupole ion traps enable trapping and storage of gaseous ions that can be ejected after a defined
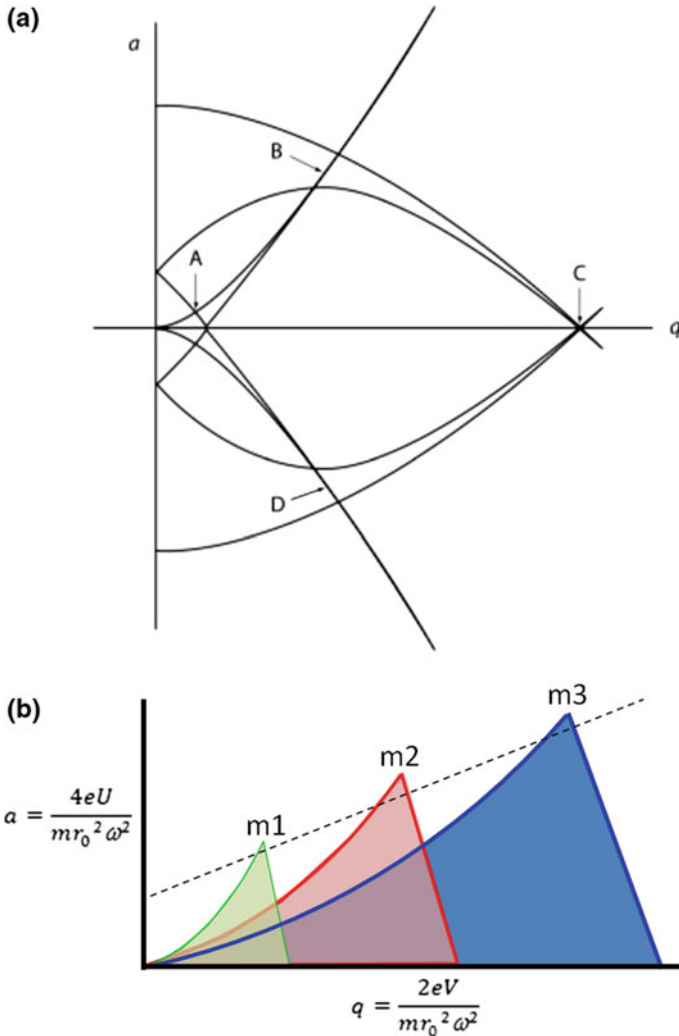
Fig. 8 Plot of solutions to the Mathieu equations for a given m/z (**a**). Regions highlighted as A, B, C and D are trajectories with both x and y stability. The majority of quadrupoles operate in stability region A (lower voltages) [77]. Also shown is a zoomed in diagram of region A for three ions with different m/z, (m1, m2 and m3) (**b**). The dashed line indicates the scan line for which trajectories for m1, m2 then m3 become stable as the magnitude of U:V is increased

period of time. For 3D quadrupole ion traps (Paul trap), ions are gated into the trap to prevent ion escape [65]. The trap itself consists of three hyperboloidal electrodes, two end-cap electrodes and a ring electrode between them (Fig. 9) [78]. DC and RF potentials are applied to the end-cap and ring electrodes, respectively, resulting in the formation of a parabolic potential well (shaped like a saddle in 3D), which ions

become trapped in [79]. Stable m/z ion trajectories are dictated by the Mathieu equations akin to quadrupoles (Fig. 8). An inert gas (~ 1 mTorr helium) is added to the trap to dampen the kinetic energy of the ions in the trap through collisions and as a result stabilises the ions trajectory and confines it to the centre of the trap. Ions can be ejected from the trap by linearly increasing the RF amplitude causing the ion trajectories to eventually become unstable, with low m/z becoming unstable first then higher m/z-ions, when the voltage reaches the resonant frequency of the ion, at which point the ions are ejected from the trap and are detected externally. Therefore, ions are not discarded prior to detection unlike beam type instruments [65]. RF voltages can also be applied to the end-cap electrodes in resonance with the periodicity for specific or multiple m/z ion(s), resulting in these trapped ions gaining kinetic energy and moving away from the centre of the trap. When this energy is high enough ions are ejected from the trap [78, 80]. This enables ejection of larger ions which would otherwise require impractically high voltages by raising the RF amplitude of the ring electrode alone [80]. Crucially, for tandem MS experiments, ions of a given m/z can also be isolated by this method. This resonance can also be exploited to raise the trapped ions kinetic energy enough so that they undergo CID without being ejected from the trap. This enables multiple ion dissociation stages to be performed (MS^n) [65, 80].

The main disadvantage of 3D ion traps is poor resolving power since small changes in the RF potential ramp can result in ejection of a 1 m/z window—under normal operating conditions 3D quadrupole ion traps can identify singly charged ions from m/z 500–2500 with a resolution of 0.2–0.5 Da [78]. They also have longer ion detection times given they trap ions for MS, as opposed to the μs required for ions to reach the detector from the source of beam type instruments.

Ions can also be trapped in linear quadrupole ion traps, which function in a similar manner to the 3D quadrupole ion traps except with elongation of all electrodes (additionally a buffer gas is not always essential for trapping). As a result, a greater number of ions can be trapped compared to the 3D ion traps. Furthermore, the ejection efficiency of linear quadrupole ion traps is much greater and therefore the sensitivity is improved [81]. However, the mass resolution is also much poorer and is typically limited to studying species $m/z < 2000$ [58]. It is also important for both types of ion traps to not overfill the trap, otherwise, space charge effects can become apparent, which will further reduce the resolving power and mass accuracy [81].

2.2.3 ToF

Linear and reflectron time-of-flight (ToF) mass analysers are capable of determining the m/z ratio of an ion from the time it takes that ion to travel through a field-free region in a vacuum after being accelerated by an electrical potential. Ions with lower m/z values reach the detector before those of high m/z values as they are accelerated to a greater velocity (v) [65]. This can be easily represented through equations for linear ToFs (linear drift region). The kinetic energy ($\frac{1}{2}mv^2$) gained by the ion is equal to the accelerating potential energy (eV where e is the charge of an
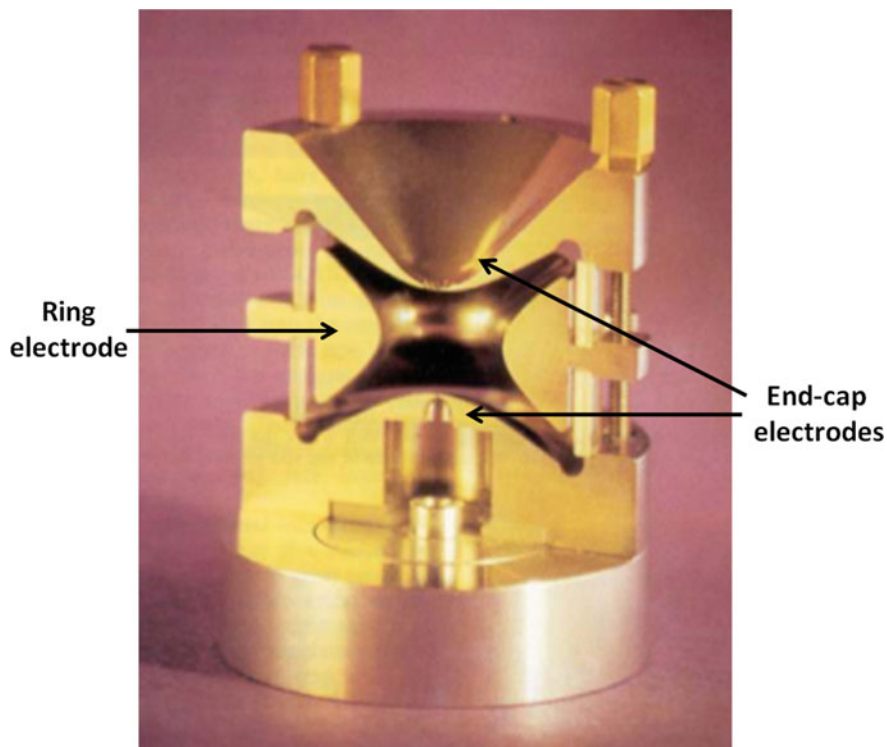
Fig. 9 Photograph of a transverse section of a 3D quadrupole ion trap with the electrodes labelled. Ions are trapped within the volume between the electrodes. Amended from March [78]

electron and V is the accelerating potential) (Eq. 3) and the time-of-flight (ToF) is related the length of the flight tube (L) (Eq. 4), the *m/z* value can thus be determined using Eq. 5 below:

$$zeV = \frac{1}{2}mv^2 \tag{3}$$

$$TOF = \frac{L}{v} \tag{4}$$

$$TOF = L\left(\frac{m}{2zeV}\right)^{\frac{1}{2}} \tag{5}$$

where z is the ion charge [58].

Reflectrons work around the same principle except that ions are reflected by an 'ion mirror' (*a* series of electrostatic ion guides with increasing potential) angled to reflect these ions towards a secondary detector preventing them from being reflected
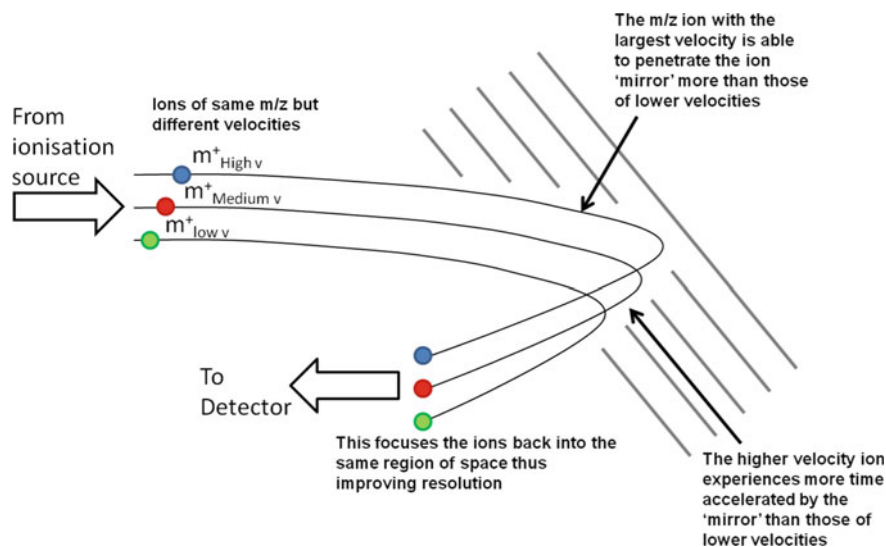
Fig. 10 Diagram depicting the reflectron mode ion trajectory for three ions, with identical *m/z* but different velocities, through a time-of-flight mass analyser

back to the source (Fig. 10) or colliding with other ions coming from the source. The purpose of the reflectron is to correct for differences in kinetic energy of ions of the same *m/z* value which would otherwise result in slight differences in computed mass and thus loss of resolution [82]. Ions of greater velocities penetrate into the reflectron 'mirror' more than those of a lower velocity. These ions are then accelerated towards the detector by the 'mirror'. The ions that had penetrated the mirror the most (higher kinetic energy) spends more time under an accelerating potential than the lower penetrating ions. This results in focussing the ions as the ones with a higher velocity have had to travel a longer distance. This v-optic reflectron is shown schematically in Fig. 10. A second reflection (w-optic) can increase the resolution even further, although this comes at the expense of sensitivity (factor of 3 compared to v-optic) as the longer drift-times increases the likelihood of metastable decomposition [58].

ToFs are advantageous as they can have a high resolving power and linear ToFs theoretically have no upper mass limit, as all that is required is the ion to be capable of travelling intact through the drift tube. Also spectra can be acquired rapidly, which allows for spectral averaging. However, reflectrons require a higher vacuum than quadrupole analyzers and the ambient temperature needs to be controlled very carefully so that the thermal energy of the ions is not increased [58]. MALDI ionisation is usually used in combination with a ToF mass analyser as they complement one another, since MALDI produces a defined pulse of ions allowing the packets time-of-flight to be measured. Also MALDI and the ToF analyser can be placed in-line reducing ion loss.

2.3   Tandem Mass Spectrometry for the Analysis of Glycopeptides

Fragmentation of (glyco)peptides can be used to determine the sequence of the peptide and also possibly the site that is glycosylated. There are three methods for fragmentation that shall be discussed, firstly collision-induced dissociation (CID), secondly electron transfer dissociation (ETD) and finally infrared multiphoton dissociation (IRMPD). There is a nomenclature for naming the ions produced by peptide and glycan fragmentation known as the Domon–Costello nomenclature (A-, B-, C-, X-, Y- and Z-ions) and the Roepstorff–Fohlmann–Biemann (a-, b-, c-, x-, y- and z-ions), respectively (Fig. 11) [51, 83].

2.3.1   CID

Collision-induced dissociation (CID) (previously referred to as collisionally activated dissociation or CAD) is the most common method used for fragmentation of glycans and peptides for sequence elucidation. For peptides, CID often generates b- and y-product ions derived from the *N*- and *C*-terminus, respectively, as a result of the amide bond dissociation (Fig. 11). CID of glycans (approximately 0.1–1 eV or 10–100 eV in ion trap or beam type analysers, respectively [84]) commonly results in fragmentation either side of the glycosidic bond generating B/C- and Z/Y-ions from the non-reducing and reducing termini, respectively. At higher energies, fragmentation across the ring is also observed. During CID, analyte ions are accelerated through
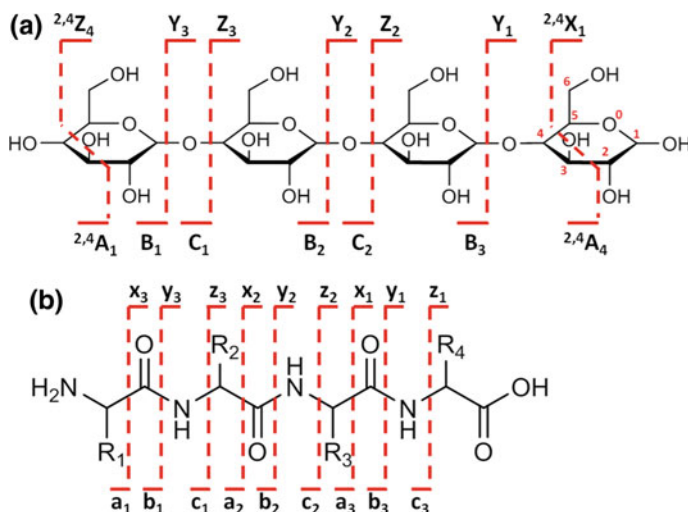
Fig. 11 Domon-Costello   nomenclature   for   glycan   fragmentation   (**a**)   [51]   and Roepstorff-Fohlmann-Biemann nomenclature for peptide backbone fragmentation (**b**). [83]

an inert neutral gas such as helium or argon. The energy the analyte ion gains from colliding with the gas is converted into internal energy that is distributed throughout the structure. It results in dissociation of typically labile bonds although charge directed mechanisms may result in dissociation of thermodynamically more stable moieties. For protonated peptides, this charge directed cleavage is best described by the mobile proton model [85, 86]. The mobile proton model states that at high internal energy mobile protons are able to protonate less thermodynamically favourable positions, such as the amide nitrogen in the peptide backbone compared to when the internal energy is lower and protonates only basic residues. Protonation of the amide nitrogen weakens the bond and thus allows the peptide backbone to fragment [87]. Protonated glycoconjugates often fragment to yield primarily glycosidic product ions [88], whereas deprotonated species also generate cross-ring fragments [88–90]. However, since free-carbohydrates typically lack any basic groups (proton sinks), they often preferentially form metal adducts with sodium or potassium without an additional metal dopant. Doping carbohydrate samples with other metals such as lithium, silver [91], or manganese [92] greatly affects the propensity of specific product ions, which has been proposed to be a result of differential metal binding sites facilitating dissociation by differing pathways [92]. Typically, large ions with low charge densities, such as potassium and rubidium, tend to dissociate to lose the metal adduct over fragmenting the carbohydrate structure, whereas for small highly charged ions like lithium extensive fragmentation is observed [93–95]. At higher energies, cross-ring fragmentation also becomes prevalent compared to glycosidic fragmentation. Also the nature of glycosidic fragmentation depends on the metal adduct. It has been reported for example that $[M + Mg]^{2+}$ adducted carbohydrates dissociate to yield more cross-ring fragments compared to the $[M + 2Li]^{2+}$ equivalent, whilst $[M + Ag]^+$ adducts produce almost no cross-ring fragments [95]. Therefore, studying glycans and glycoconjugates with different adducts can produce complementary structural information [88, 96]. Additionally, chemical derivatization such as permethylation/acetylation appears to favour formation of glycosidic fragment ions [97]. Finally, the presence of other product ions depends on the identity of the glycan and its regiochemistry [98]. The stereochemistry of the glycosidic bond often results in different ion yields rather than the formation of different product ions. Currently, there is little knowledge of these product ion structures [51, 94, 99] and, unlike for peptides, no definitive mechanisms for glycan dissociation have been elucidated [85–87, 100, 101].

CID of glycoconjugates commonly yields diagnostic product ions losses such as m/z 162, 164, 180, 203 and 221 corresponding to dehydrated hexose, deoxy hexose, hexose, dehydrated N-acetylhexosamine and N-acetylhexosamine respectively. As a result, neutral loss scanning of these diagnostic species facilitates the identification of precursor glycoconjugates. Conversely, product ions analysis of glycan fragment ions (e.g. m/z 204 for [Hex$N$Ac–H$_2$O + H]$^+$) can also facilitate glycan precursor identification.

There are issues using CID fragmentation when studying glycans and their conjugates. Firstly, CID typically fragments the most labile bonds which for gly-copeptides tend to be the glycosidic linkages between the peptide and the glycan

(although for high mannose containing *N*-glycans the glycosidic bond can remain intact). Loss of the glycan from *O*-glycopeptides goes via a mechanism that regenerates the alcohol group on the serine/threonine residues [102], preventing characterisation of the glycosylation site. This loss of the labile glycan moiety is elevated for beam instruments like quadrupole-ToFs [103]. Furthermore, this fragmentation pathway often comes at the expense of backbone fragmentation, therefore higher collisional energies are often required to fragment the peptide backbone [103]. In addition, the glycan chains themselves also fragment by CID typically producing B/Z- and C/Y-ions in positive ion mode, but A/X-ions can also be observed where the ring itself fragments greatly complicating tandem mass spectra [51]. CID-induced migrations of Fuc [104–106], hexose [107] and sulphate [108] groups have been observed for protonated and deprotonated carbohydrates, further complicating analysis [105, 107, 109, 110], although this has not yet been reported for metal adducted analytes [109, 111].

2.3.2 ETD

Electron transfer dissociation (ETD), not to be mistaken with electron capture dissociation (ECD), is a less widespread fragmentation technique compared to CID. Dissociation is achieved by an electron transfer reaction between a gaseous radical such as fluoranthene or anthracene and the analyte ions of interest. For peptides, the resultant radical ion is unstable and results in the fragmentation of the bond with the lowest barrier of dissociation, which tends to be the N–C$_\alpha$ backbone bond generating c- and z-ions (also fragmentation of the amine within proline has been reported) [112, 113]. ETD fragmentation is much more efficient for peptides with a high charge density as lower charge allows for the possibility of non-covalent intrapeptide interactions that could contain the extra electron, thus preventing dissociation [114]. Although overall fragmentation of smaller peptides by ETD is less efficient than CID, ETD has a major advantage as the glycosidic linkage does not preferentially undergo fragmentation thereby enabling the observation of fragment peptide ions with the glycan still attached [103]. Application of ETD, therefore, increases the likelihood of being able to pinpoint the site of glycosylation. However, for highly charged free glycans (such as glycosaminoglycans or high valency metal cations), ETD (and electron detachment dissociation) has been shown to be capable of causing primarily cross-ring fragmentation whereas CID produces primarily glycosidic dissociation [115–119].

2.3.3 IRMPD

Infrared multiphoton dissociation (IRMPD), first reported by Beauchamp and co-workers [120], involves irradiation of trapped mass-selected ions, typically achieved in quadrupole ion traps or Penning traps, with infrared radiation from a tunable infrared laser (Fig. 12). The wavelength of the laser is ramped over a period
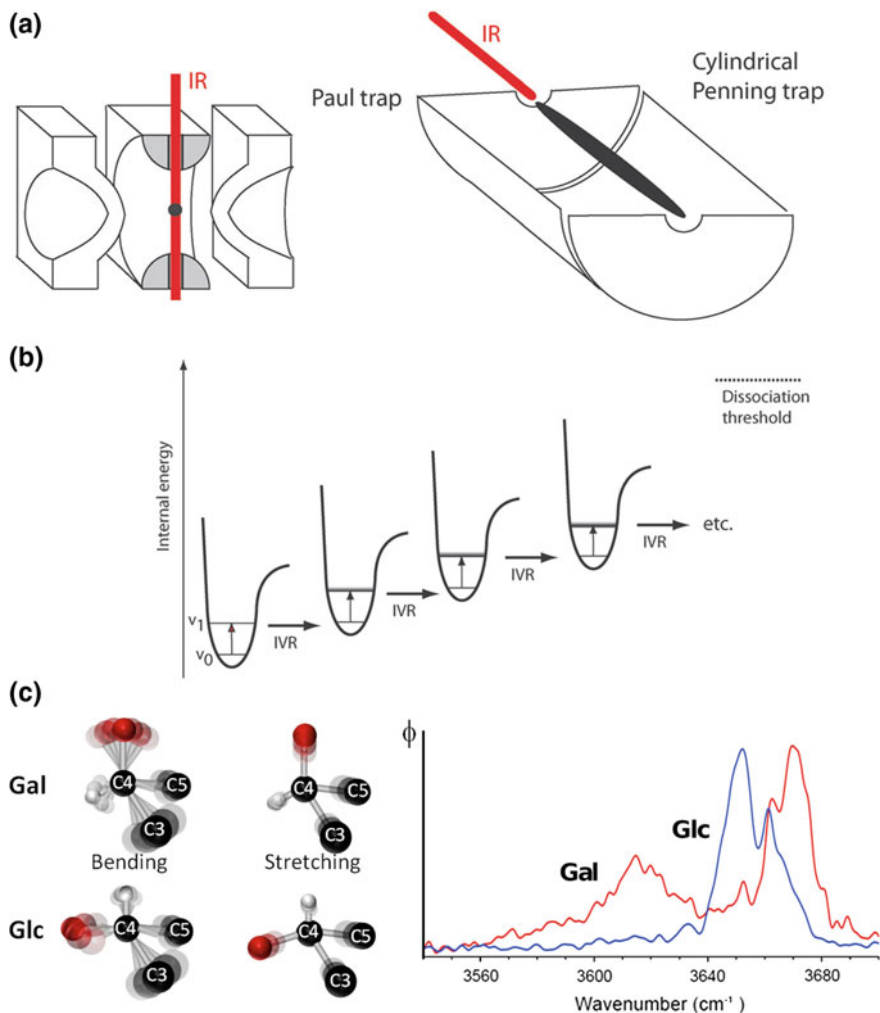
Fig. 12 Scheme depicting the typical orientation of the tuneable IR laser as it passes through the ion cloud in a Paul (3D-quadrupole ion) trap and Penning trap (**a**). Schematic of the IRMPD mechanism, the internal energy the ion gains resulting from photon absorption at the fundamental vibrational frequency is dissipated by sequential intramolecular vibrational re-distribution (IVR) events (**b**). Also depicted are cartoon bending and stretching modes for a region within hexose isomers Glc and Gal and the action IR spectra of these two molecules (**c**). Figure parts **a** and **c** are taken from Polfer [123]

of time causing unimolecular fragmentation (for carbohydrates this dissociation produces CID like product ions [121, 122]) of the trapped ions only when they possess resonant vibrational modes [123]. Therefore a plot of the wavelength against the fragmentation yield, ϕ, normalised to the respective laser power at each given frequency, generates an action IR spectrum of the gas-phase ion. Additional

molecular dynamic (MD) simulations can yield key atomistic or conformation information for small molecules or secondary structural features for peptides or proteins that would be unobtainable by convention MS-based methods alone [121–126]. Therefore IRMPD-MS could be an extremely powerful technique to aid characterisation of often isomeric carbohydrates [124, 126], though this approach is limited by the availability of bench top IR lasers that cover a wide and tunable range [123]. To achieve action IR spectroscopy spanning the majority of the IR range, free electron laser facilities are required [127]. Theoretical and experimental comparisons are also extremely challenging since theoretical calculations typically assume harmonic vibrational spectra, whereas experimental spectra are anharmonic. Also the absorption of IR radiation is not necessarily linearly scaled with the imparted energy, therefore experimental intensities may vary as compared to the theoretical spectra. Finally intramolecular vibrational re-distribution results in band broadening and red-shifting, which needs to be accounted for the experimental spectrum [123]. Generation of reasonable gas-phase candidate structures, from which theoretical IR spectra may be determined, is also not trivial given the typical inherent conformational diversity within molecules.

3   Glycan Sequencing Strategies

Glycan characterisation can be achieved to various degrees by a number of techniques. Most of these techniques are limited to studying solely glycan moieties that have been removed from their conjugates prior to characterisation, although methods exist that are capable of characterising the glycan constituents of intact glycoproteins [128], glycopeptides [46] or other glycosides [37, 129]. For *N*-glycans, glycan release is normally achieved enzymatically using one of three enzymes-peptide-*N*-glycosidase F (PNGase F), which selectively cleaves the bond between the reducing Glc*N*Ac residue and the Asn residue (preferably without a core α1-3 Fuc moiety); endoglycosidase F (Endo F) that cleaves between the chitobiose core; or endoglycosidase H (Endo H), which also cleaves between the chitobiose unit of only high mannose or hybrid *N*-glycans (Fig. 13) [2, 50]. As Endo F/H cleavage leaves a glycan unit attached to the protein/peptide, sites of glycosylation can also be characterised using typical proteomic strategies (e.g. proteolysis followed by LC-MS^2).

In comparison, the glycan moiety is completely removed when using PNGase F, also converting the Asn residue to Asp, making glycosite characterisation more challenging. This can be overcome by performing the transformation in ^{18}O-water which results in the formation of an ^{18}O-Asp residue [130]. No similar enzymes are known for *O*-glycan deglycosylation making them more challenging to process. As a result, *O*-glycans are typically cleaved chemically using either reductive β-elimination [48, 50, 131, 132] (Fig. 14) or hydrazinolysis (Fig. 15) [133–135]. However, hydrazinolysis may result in sequential loss of the reducing-terminal residues (peeling), inhibiting complete sequence elucidation [133, 134]. After
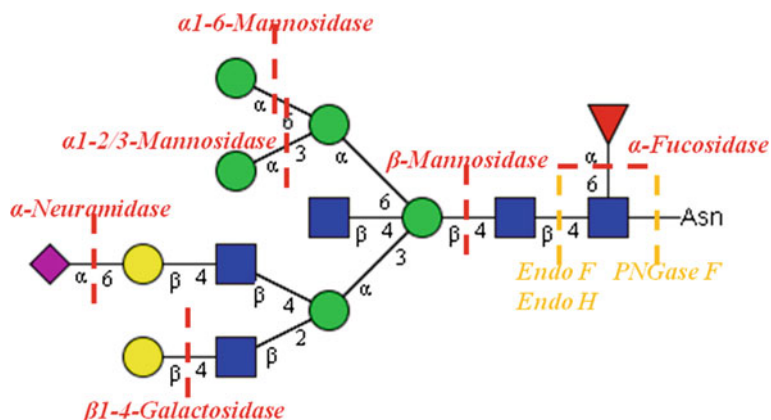
Fig. 13 Representative structure for a hybrid type *N*-glycan and the potential endoglycosidases (*orange*) and example exoglycosidases (*red*) that can act on it
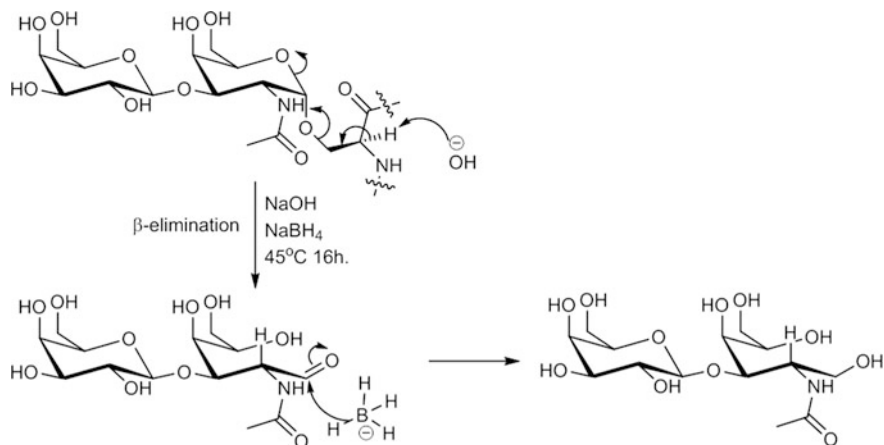
Fig. 14 Proposed mechanism for sodium borohydride induced reductive β-elimination of *O*-glycans [50]

release, the glycans can be enriched and purified using either lectin affinity [136–138], porous graphitised carbon (PGC) [139–141] or size exclusion columns [142] and/or tagged with a fluorescent label [49, 143, 144]. Since these glycans normally exist in multiple glycoforms, the released glycans are typically fractionated by chromatography prior to their characterisation [15, 145–149]. Conversely, glycans can be hydrolysed into their constituent monosaccharides by TFA or sulphuric acid treatment [150, 151] followed by re-*N*-acetylation [152]. The relative abundances and composition of these monosaccharides are then elucidated.
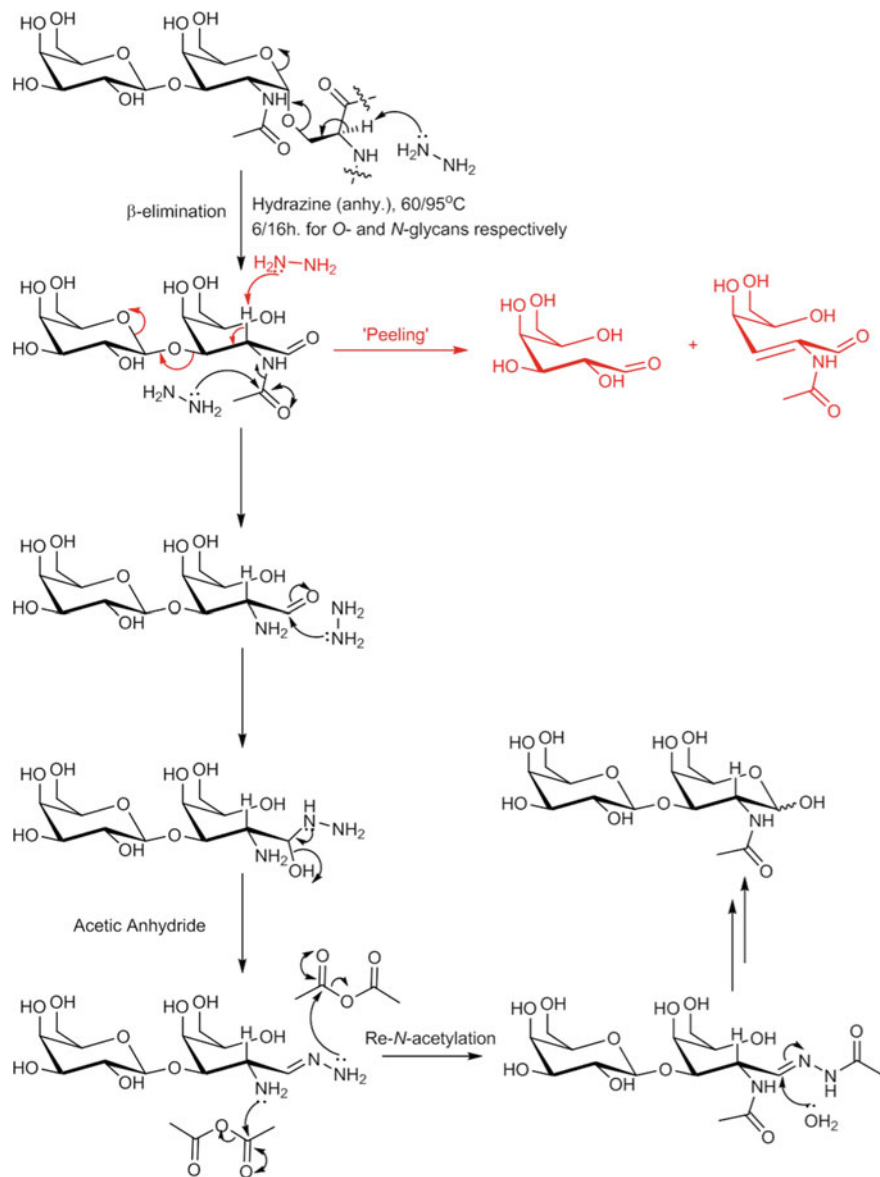
Fig. 15 Potential reaction mechanism for hydrazinolysis of an O-glycan including the 'Peeling' side reactions [135]

3.1 Glycosidase Treatment

Treatment of unknown glycans with cocktails of glycosidases, whose specificity and activity are well characterised, is one of the most routinely employed strategies to facilitate elucidating monosaccharide sequence and glycosidic linkage information [49, 50, 153]. Glycosidases fall into two groups, endoglycosidases that cleave non-terminal monosaccharides or glycosidic bonds to aglycons (such as proteins and lipids) and exoglycosidases that cleave monosaccharides from the non-reducing terminus in a stepwise manner. Both types are typically extremely specific for a given monosaccharide and linkage type. Therefore, sequential treatment of unknown glycans with specific exoglycosidases enables deduction of both the monosaccharide sequence and the linkage information. A wide variety of exoglycosidases have been characterised that cleave Galβ (Bovine testes) [154], Galβ1-4 (*Streptococcus pneumoniae*) [155], GlcNAcβ (*Streptococcus pneumoniae*) [156], Neu5Acα (*Streptococcus pneumoniae*) [157], Manα1-2/3 (*Xanthomonas manihotis*) [158], Manα1-6 (*Xanthomonas manihotis*) [158] and Fucα1-2 (*Xanthomonas manihotis*) [158] to name a few (Fig. 13). Detection of enzymatic hydrolysis products is facilitated by other analytical techniques (usually MS and/or LC) [15]. This approach is routinely employed in *N*-glycan sequencing. *N*-glycans consist of well-characterised core oligosaccharide structures and the enzymatic transformations that act on them all relatively well characterised (at least for mammalian systems) [15, 26, 153]. To enable characterisation, however, glycans must be purified prior to hydrolysis so that changes resulting from the enzymatic action can be monitored. Therefore, a priori fractionation techniques are required which can be challenging for oligosaccharides since they are often chemically similar. Additionally, glycosides capable of cleaving all potential glycosidic linkages have yet to be elucidated limiting this approach. Finally, the sequential enzymatic processes involved are low-throughput and often tedious [88].

3.2 Glycan-Binding Probes (GBPs)

One of the earliest approaches to identify carbohydrate structures was to employ well-characterised glycan-binding proteins (GBPs), such as lectins or antibodies, as affinity probes (Fig. 16) [138, 159–162]. A wide variety of GBP specificities has been characterised that recognise a diverse range of glycan moieties [20, 163–168]. This enables elucidation of a range of divergent carbohydrate functionalities. For example, the lectin Concanavalin A (ConA) preferably binds high mannose *N*-glycans (α-Man specific), whereas *Lotus tetragonolobus* lectin (LTL) binds to fucosylated glycans [163]. The difference in structural specificities of two GBPs can also be more subtle. For example, *Sambucus nigra* (SNA) only binds to α-2,6-sialylated carbohydrates [20], whereas *Maackia amurensis* II (MAL II) binds preferentially to terminal α-2,3-sialylated *N*-glycans [164]. Unlike most other
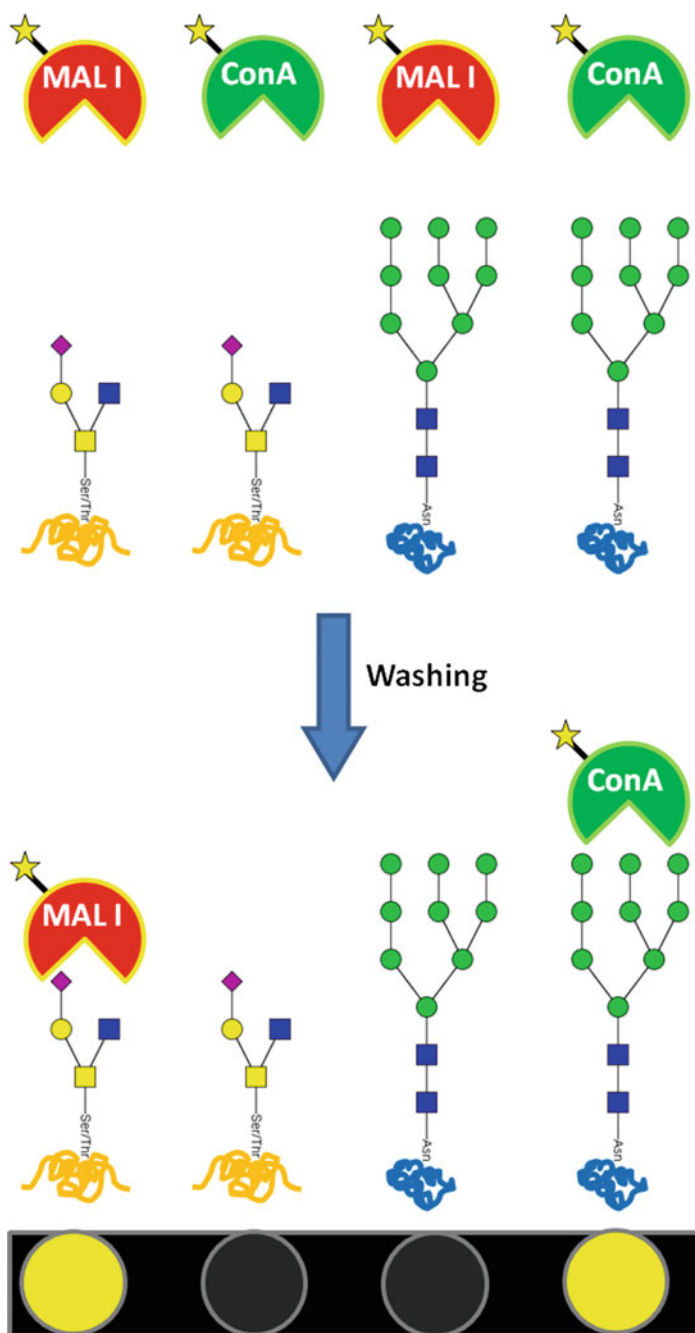
Fig. 16 Scheme depicting incubation of fluorescently tagged MAL I (sialic acid specific lectin) and ConA (high mannose *N*-glycan specific) with an array of *O*- and *N*-glycans and their resultant fluorescent response after washing

sequencing strategies, the use of GBPs allows for glycan profiling at the macroscopic level in the form of tissue staining [169, 170] and microscopic in the form of a cell [171, 172], as well as being capable of characterising glycoconjugates or free glycans [173–177]. These GBPs are typically labelled with a fluorescent or radioactive tags to facilitate detection of binding [138, 174, 178], which is disadvantageous compared to other label-free strategies, although label-free detections techniques are also viable in certain situations such as MALDI-ToF MS [161, 162] and surface plasmon resonance (SPR) spectroscopy [179, 180] can also be arrayed enabling high-throughput glycan profiling of glycoconjugates [36, 161, 162, 181–183] or even cells [172] and bacteria [184, 185]. Similarly, lectins can be attached to supports, enabling selective enrichment of glycosides with a given structural motif (lectin affinity chromatography) [186]. Unlike other profiling strategies, however, lectin arrays yield few molecular structural features and because of promiscuous lectin binding specificities, are liable to produce misleading information.

3.3 Liquid Chromatography

For carbohydrate analysis, liquid chromatography (LC) is exploited primarily as a separation and to a lesser extent as a characterisation tool [49, 50, 147, 187–190]. LC separates analytes based on their differing affinity towards a defined stationary phase. Analytes that display poor retention (or are highly soluble in the mobile phase) are sequentially eluted before those that have higher retention. Detection of eluted molecules is normally achieved by spectroscopic means if the analyte possess a chromophore or by MS when the LC is placed in-line with the mass spectrometer. Confirmation of an analytes identity can be facilitated through comparing its retention time to a known standard [188, 189]. This strategy has been greatly facilitated by the development of a database (GlycoBase) consisting of experimental chromatograms (and mass spectra) for various separation techniques (e.g. HPLC, capillary electrophoresis etc.) under a specified set of standard conditions [188, 189]. For glycans, monitoring alterations in glycan retention times (often described in 'glucose units' (GU) with reference to a ladder of dextran polymers [145]) after incubation with various glycosidase 'cocktails', whose specific activities are well characterised, can enable comprehensive elucidation of stereochemical sequences [15, 26, 37, 145]. This approach relies heavily on access to glycosidases capable of removing all natural glycosidic linkages, which currently is not achievable. Characterising unknown glycans with a series of glycosidases would also be time-consuming. LC techniques to separate carbohydrates have lagged behind other biomolecules, primarily because carbohydrates lack a distinct chromaphore to facilitate their detection [148]. Additionally, they display poor retention and separation on classical C_{18} reversed-phase (RP) columns given their highly hydrophilic nature [147, 191]. Recently, it has been reported though that separation of underivatized carbohydrates on RP columns can be improved through

incorporation of an orthogonal separation tools such as ion mobility-mass spectrometry (IM-MS) [147]. As a result of poor separation on RP columns, carbohydrates are typically derivatised prior to LC analysis to add a chromaphore and increase hydrophobicity. This is typically achieved through coupling of an aminated chromophore, such as 2-aminobenzamide, 2-aminobenzoic acid or 2-aminopyridine, to the reducing aldehyde functionality of carbohydrate moieties (reductive amination) [144, 192]. Permethylation of carbohydrates also improves retention on C_{18} RP-LC supports and facilitates MS-based detection [193, 194], although this relies on quantitative derivatization which can be challenging for larger sterically more hindered structures. Despite the application of these strategies, separation of glycan isomers remains challenging by C_{18} RP-LC methodology alone [148, 191, 195]. As a result several other LC approaches consisting of novel stationary phases have been developed, such as high pH (also referred to as high-performance) anion-exchange chromatography (HPAEC) [196], hydrophilic interaction chromatography (HILIC) [15, 197] and porous graphitic carbon (PGC) [140, 198, 199], which do not rely on derivatization strategies (Fig. 17).

HPAEC is well established for the high resolution separation of underivatized carbohydrates [196, 200, 201]. HPAEC columns typically consist of an agglomeration of non-porous polystyrene-divinylbenzene and smaller polystyrene-divinylbenzene beads possessing quaternary amine groups [200]. Carbohydrate samples are deprotonated with high pH mobile phases (pH > 12, linear gradient ∼10–100 mM sodium hydroxide to sodium acetate) resulting in separation by anion-exchange with the quaternary amine stationary phase [200]. Classically glycans eluted from HPAEC columns are quantitatively detected by pulsed amperometric detection (PAD), since carbohydrate possess no chromophore.
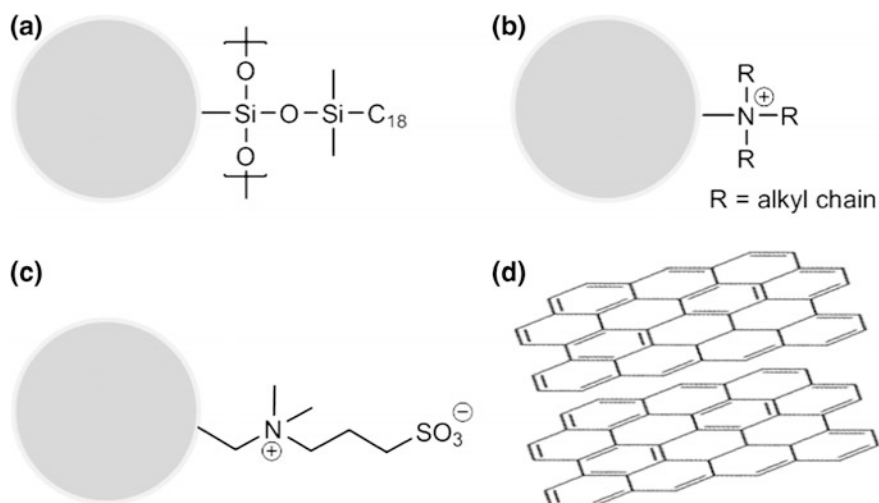

Fig. 17 Examples of common stationary phases for RP (C18) (**a**), HPAEC (CarboPac, dionex) (**b**), HILIC (ZIC-HILIC SeQuant®) (**c**) and PGC (**d**)

Conveniently, most of the eluted analyte can be recovered after detection. Unfortunately, HPAEC is often limited to small oligosaccharides (trisaccharides or less) and is thus typically used for monosaccharide composition analysis. Also due to the high-levels of salt, HPAEC is not amenable to MS-based analytical techniques without carbohydrate desalting [139]. HPAEC-PAD also requires a relatively large amount of analyte ($\sim$ mg).

The separation mechanisms of HILIC chromatography are much more complex compared to RP and HPAEC [202] and are still not fully understood. The specific details of these mechanisms are beyond the scope of this chapter. HILIC employs polar stationary phases to which polar substrates are retained preferentially compared to the less-polar mobile phase (typically composed of acetonitrile). Polar carbohydrates are eluted by increasing the concentration of water within the mobile phase. This enables separation of a wide range of potentially isomeric carbohydrates including small oligosaccharides [203], N-glycans [15, 204] and glycopeptides [197]. Also unlike HPAEC, the mobile phase employed is compatible with label-free analytical techniques such as MS [204] and UV/Vis absorbance for oligosaccharides labelled with a chromaphore [15]. As a result, the use of HILIC separation within research science is rapidly growing [202].

PGC stationary phases are increasingly being employed for the desalting and separating underivatized carbohydrate samples [139, 140, 199, 204–206]. As separation of aqueous analytes occurs by both hydrophobic and electronic interactions with the PGC stationary phase, PGC chromatography is amenable to studying both polar and non-polar substrates [141]. The precise mechanisms associated with this separation behaviour are also poorly understood similar to HILIC [205]. Unlike most stationary phases, PGC is extraordinarily resilient to extreme operating conditions including pH 0–14, allowing separation of acidic and basic analytes in their neutral forms, respectively; and high temperatures (200 °C) that result in improved peak symmetry and potentially improved separation [207] and additionally accommodate for larger flow rates with little increase in back-pressure given the loss in viscosity associated with the mobile phase [205]. Like HILIC, PGC chromatography has been reported to be capable of separating branched and linear glycan regio- and stereoisomers [140, 198, 199, 208]. Few studies have been undertaken comparing PGC and HILIC, although a large study comparing 141 unique metabolites found HILIC (aminopropyl column at pH 9.45) provided the greatest separation of these species [209]. Estimation of PGC retention times from the glycan structure is also more challenging compared to HILIC and RP [204].

An issue for LC approaches for carbohydrate identification is that an internal standard is required to validate detected carbohydrates, although this requires a priori characterisation of that given structure for full structural elucidation [199]. Conversely, the retention times of a purified known reference library whose structures have been characterised can be compared to unknown structures. However, the isolation or synthesis of such a vast library would be extremely challenging given the challenges associated with their chemical synthesis [210, 211] and the huge number of potential structures glycans that can be formed from only a small subset of monosaccharide residues [212]. Finally, even with all these

LC methodologies, separation of certain glycan isomers remains a significant challenge [213], although this may be alleviated by using orthogonal tandem LC strategies [208].

3.4 Mass Spectrometry and Hyphenated Mass Spectrometry Methods

MS and tandem MS techniques are the most commonly employed approach to structurally characterise glycans or glycopeptides given their speed and sensitivity [88, 214, 215]. Additionally, unlike most of the other characterisation tools, mass spectrometry enables direct glycosite characterisation at the glycopeptide level.

However, the primary limitation associated with studying glycans by conventional MS approaches is that most of the common natural monosaccharide building blocks are simple epimers of one another and therefore possess identical m/z (Figs. 1 and 18). Therefore, MS alone, is limited to characterising the monosaccharide class (i.e. hexose, N-acetylhexosamine, deoxy hexose, etc.) and is insufficient to directly identify these monosaccharide units without the use of an orthogonal sequencing approach, and are (???) often achieved through characterising the glycan processing pathways [46]. Glycan regiochemistry and branching can be elucidated either from diagnostic fragment ions [98, 215, 216] or though analysis of tandem mass spectra of permethylated or peracetylated glycans, since glycosidic fragments will have a number methyl/acetyl groups equal to the number of branches from that residue [194]. However, incomplete peralkylation or peracetylation complicates identification of branching and they (????) often require an additional purification step *c.f.* their underivatized equivalents; although, as previously mentioned, peralkylation or peracetylation strategies are also advantageous as they improve chromatographic retention on classical RP columns and increases MS response [190, 194]. Peralkylation and peracetylation also increase the volatility and thermal stability of monosaccharides allowing their analysis by gas-chromatography MS [217]. Stereochemical assignment of the monosaccharide units and of the glycosidic bond linking them is much more challenging. As discussed previously in Sect. 1.3.1, shifts in m/z after application of specific well-characterised exoglycosidases allows for certain moieties to be elucidated depending on the availability of the exoglycosidase [218]. Spectral matching glycan [88, 98, 219–223] or glycopeptide [220, 224] experimental tandem mass spectra to those of synthesised reference standards have also shown promise in being able to discern stereochemical (and regiochemical) information. A drawback of spectral matching is the requirement of a characterised synthetic standard, which is not feasible for large oligosaccharides given the number of potential isomers and challenges associated with synthesising these standards. Although it is not clear how well spectral similarities observed for small oligosaccharide standards will extend to larger structures. Konda et al. have reported that MS[3] (???) of a diagnostic

fragment ion (*m/z* 221) corresponding to deprotonated Hex-glycolaldehyde produces a tandem mass spectrum that is diagnostic of the stereochemistry of the monosaccharide and anomeric configuration [215, 225]. This approach would require of a relatively small reference library to allow identification of all species, although the approach has currently only been developed on a relatively small reference library of Glc, Gal and Man. It also requires the formation of specific fragments that produce these MS^n structurally rich spectra, which may be challenging for larger structures. Further recent strategies include energy-resolved mass spectrometry (ER-MS), where the product ion yields are recorded against the specific collision energy imparted since different diastereoisomers may require different activation energies [226–230]. For carbohydrates, these approaches often require purification of glycan mixtures prior to analysis otherwise you are liable to record chimeric ER-MS of iso-bars/mers that cannot be readily deconvoluted. Using a variation of ER-MS, Nagy et al. reported the first separation of all natural aldohexose, ketohexose and pentose monosaccharide stereoisomers including enantiomers. This was achieved by initially forming diastereomeric complexes with the monosaccharide of interest, a chiral reference molecule (e.g. *L*-serine) and Cu^{2+}, fragmenting them by collision-induced dissociation (CID) at a specific energy and comparing the relative intensities of the product ions associated with loss of the chiral reference and the glycan (R_{fixed}) [228, 230]. However, this strategy cannot generate regio-/stereochemical information in regards to the glycosidic linkages and requires glycan hydrolysis to the monosaccharide species prior to analysis; the sequence in which the monosaccharides appear in the glycan cannot thus be determined directly. Combining ion mobility spectrometry, an orthogonal gas-phase separation technique, to ER-MS further improves discrimination of isomeric carbohydrates [229].

Ion mobility (IM) spectrometry is a growing technique to characterise or separate carbohydrates and is especially powerful when coupled to conventional MS and LC strategies. This recent surge in ion mobility analysis of glycans is primarily a result of the recent commercialisation of hybrid IM-MS instrumentation (2006).
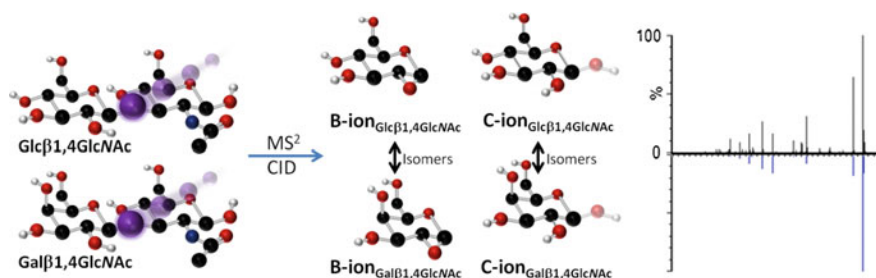
Fig. 18 Example of the spectral similarity within the tandem mass spectrum of two disaccharide isomers that have been fragmented by collision-induced dissociation (CID). Resulting product ions of these isomeric species are also isomers of one another. CID corresponds to collision-induced dissociation

IM separates species based on their rotationally averaged cross-sectional area-to-charge (Ω/z) ratio, a parameter that is intrinsic to a given molecule under a defined set of experimental conditions. Therefore, this technique possesses the capability of separating isomeric and isobaric glycan and glycoconjugates precursors and product ions, indistinguishable by MS alone and crucially when combined with molecular dynamics may facilitate elucidation of structural and conformational information [46, 129, 213, 229, 231–233].

Another drawback to using MS as an approach to characterise glycans is the reported rearrangement of glycan during CID of protonated carbohydrates, as previously mentioned in Sect. 1.2.3.1, making structure elucidation ambiguous [104, 109]. Although this migration has never been observed for metal cationised carbohydrates [109].

3.5 Alternative Approaches to Characterise Glycan Structures

NMR spectroscopy is currently the only analytical technique that can directly yield full three-dimensional structure and conformation elucidation of completely novel carbohydrates in solution [5, 234, 235]. Given the typical biological complexity of carbohydrates, 2D NMR experiments are normally acquired which improves spectral dispersion and can also yield important conformational information such as bond distances through-space between different or identical atoms for (Hetero) Nuclear Overhauser Effect spectroscopy (H/NOESY), respectively, and connectivity and stereochemical information for Heteronuclear Multiple-Quantum Correlation spectroscopy (HMQC) [5, 235], Correlation spectroscopy (COSY) [236] and Total Correlation spectroscopy (TOCSY) [46, 234]. The main disadvantages of NMR are that it is low-throughput, requires a relatively large amount of material which is typically not-amenable to biological samples and due to the spectra complexity of carbohydrates, samples need to be purified and isolated which is challenging for similar glycoforms. Additionally, spectra are normally extremely challenging to interpret and thus require computer assisted assignments [234].

X-ray crystallography also possesses the ability to give atomistic information on crystallised glycan and glycoconjugates structures [237, 238]. This technique like NMR is low-throughput and additionally requires regular crystals or carbohydrate structures, which underivatized carbohydrates rarely produce presumably due to their inherent flexibility. Additionally, due to their flexibility, carbohydrates tend to give poorly resolved crystal structures [237, 239].

4 Characterisation of Glycan-Binding Proteins

(Micro)array approaches are eminently suited for the elucidation of glycan-glycan-binding protein (GBP) partners as they can potentially offer the ability to screen >1000 reactions in a high-throughput manner [174] and require sub µL amounts of material. The application of array technology within both basic and applied (e.g. clinical) research science is highly diverse and has previously been used to study antigen binding [240], enzymatic transformations [175, 241, 242], bacterial binding [243] and glycan-GBP binding partners [138, 159, 174] to name a few. It is therefore unsurprising that there is a plethora of both array technologies and analytical techniques employed. A comprehensive list of array technologies and the analytical techniques used to screen them have been reviewed by us very recently [244]. Whilst this review focuses on the applications of array technologies to screen enzymatic transformations, the principles described in the review extend to studying substrate–ligand binding partners.

For applications immobilising glycans or GBPs to self-assembled monolayers (SAMs) on gold offer an ideal platform to study glycan–GBP partners. SAMs are relatively thermally and chemically stable and, to some extent, mimic the fluidity of the lipid bilayer motif present at the cell surface membrane [245]. These arrays can be screened directly by matrix-assisted laser desorption/ionisation (MALDI) time-of-flight (ToF) mass spectrometry (MS), a rapid label-free technique that reveals structural information on the bound analyte. When combined with routine bottom-up proteomics strategies, namely proteolytic digestion followed by MS and/or MS/MS of the generated peptides and subsequent database screening [246], the identity of unknown proteins can be determined [162, 179, 183]. Such arrays can, therefore, be utilised to characterise glycan-GBP partners directly from complex biofluids, without the need for a challenging labelling step as is often required for other common visualisation techniques such as fluorescence or radiation [138, 247].

5 Conclusions

Mass spectrometry-based methodologies have become key analytical tools both for carbohydrate sequencing and the characterisation of glycan-binding proteins. The key limitation of classical mass spectrometry as a 'two-dimensional technique' has been the lack of stereochemical information of glycan structure. Such shortcomings can be overcome by combining mass spectrometry with enzymatic digestion protocols and chromatographic techniques. Some very exciting recent developments are the adding of ion mobility spectrometry and IR spectroscopy in-line with mass spectrometry, used as hyphenated analytical techniques that give very high structural resolution.

References

1. Shin I, Park S, Lee MR (2005) Carbohydrate microarrays: An advanced technology for functional studies of glycans. Chem-Eur J 11:2894–2901
2. Varki A, C.R, Esko JD, et al. (2009) Essentials of glycobiology, 2 edn. Cold Spring Harbor, New York
3. Apweiler R, Hermjakob H, Sharon N (1999) On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. Biochim Biophys Acta-Gen Subj 1473:4–8
4. Ohtsubo K, Marth JD (2006) Glycosylation in cellular mechanisms of health and disease. Cell 126:855–867
5. Yoshida-Moriguchi T et al (2010) O-Mannosyl phosphorylation of alpha-dystroglycan is required for laminin binding. Science 327:88–92
6. Sperandio M, Gleissner CA, Ley K (2009) Glycosylation in immune cell trafficking. Immunol Rev 230:97–113
7. Bode L (2012) Human milk oligosaccharides: every baby needs a sugar mama. Glycobiology
8. Poulin MB et al (2014) Specificity of a UDP-GalNAc pyranose-furanose mutase: a potential therapeutic target for campylobacter jejuni infections. ChemBioChem 15:47–56
9. Morris HR et al (1996) Gender-specific glycosylation of human glycodelin affects its contraceptive activity. J Biol Chem 271:32159–32167
10. Benoff S (1997) Carbohydrates and fertilization: an overview. Mol Hum Reprod 3:599–637
11. Schröter S, Osterhoff C, McArdle W, Ivell R (1999) The glycocalyx of the sperm surface. Human Reprod Update 5:302–313
12. Paszek MJ et al (2014) The cancer glycocalyx mechanically primes integrin-mediated growth and survival. Nat Adv Online Pub (2014)
13. Jacob F et al (2012) Serum antiglycan antibody detection of nonmucinous ovarian cancers by using a printed glycan array. Int J Cancer 130:138–146
14. Springer GFT, Tn (1984) General carcinoma auto-antigens. Science 224:1198–1206
15. Bones J, Mittermayr S, O'Donoghue N, Guttman A, Rudd PM (2010) Ultra performance liquid chromatographic profiling of serum N-glycans for fast and efficient identification of cancer associated alterations in glycosylation. Anal Chem 82:10208–10215
16. Chen YT et al (2013) Identification of novel tumor markers for oral squamous cell carcinoma using glycoproteomic analysis. Clin Chim Acta 420:45–53
17. Lauc G et al (2013) Loci associated with N-glycosylation of human immunoglobulin G show pleiotropy with autoimmune diseases and haematological cancers. PLoS Genet 9
18. Gornik O, Gornik I, Gasparovic V, Lauc G (2008) Change in transferrin sialylation is a potential prognostic marker for severity of acute pancreatitis. Clin Biochem 41:504–510
19. Hennet T (2012) Diseases of glycosylation beyond classical congenital disorders of glycosylation. Biochimica et Biophysica Acta (BBA)—General Subjects 1820:1306–1317
20. Iskratsch T, Braun A, Paschinger K, Wilson IBH (2009) Specificity analysis of lectins and antibodies using remodeled glycoproteins. Anal Biochem 386:133–146
21. Kumar SR, Sauter ER, Quinn TP, Deutscher SL (2005) Thomsen-friedenreich and Tn antigens in nipple fluid: carbohydrate biomarkers for breast cancer detection. Clin Cancer Res 11:6868–6871
22. Anthony RM et al (2008) Recapitulation of IVIG Anti-inflammatory activity with a recombinant IgG Fc. Science 320:373–376
23. Kanda Y et al (2007) Comparison of biological activity among nonfucosylated therapeutic IgG1 antibodies with three different N-linked Fc oligosaccharides: the high-mannose, hybrid, and complex types. Glycobiology 17:104–118
24. Thanabalasingham G et al (2013) Mutations in HNF1A result in marked alterations of plasma glycan profile. Diabetes 62:1329–1337
25. Lodish HF et al (2000) Molecular cell biology, 4th edn. New York

26. Royle L et al (2003) Secretory IgA N- and O-glycans provide a link between the innate and adaptive immune systems. J Biol Chem 278:20140–20153
27. Clardy J, Walsh C (2004) Lessons from natural molecules. Nature 432:829–837
28. Dalziel M, Crispin M, Scanlan CN, Zitzmann N, Dwek RA (2014) Emerging principles for the therapeutic exploitation of glycosylation. Science 343
29. Thaker MN, Wright GD (2015) Opportunities for synthetic biology in antibiotics: expanding glycopeptide chemical diversity. ACS Synth Biol 4:195–206
30. Ragauskas AJ et al (2006) The path forward for biofuels and biomaterials. Science 311:484–489
31. Azadi P, Inderwildi OR, Farnood R, King DA (2013) Liquid fuels, hydrogen and chemicals from lignin: a critical review. Renew Sust Energ Rev 21:506–523
32. Brinchi L, Cotana F, Fortunati E, Kenny JM (2013) Production of nanocrystalline cellulose from lignocellulosic biomass: technology and applications. Carbohydr Polym 94:154–169
33. Krištić J et al (2014) Glycans are a novel biomarker of chronological and biological ages. J Gerontol Series A: Biol Sci Med Sci 69:779–789
34. Wuhrer M et al (2006) Gender-specific expression of complex-type N-glycans in schistosomes. Glycobiology 16:991–1006
35. Krishnamoorthy L, Mahal LK (2009) Glycomic analysis: an array of technologies. ACS Chem Biol 4:715–732
36. Katrlik J, Svitel J, Gemeiner P, Kozar T, Tkac J (2010) Glycan and lectin microarrays for glycomics and medicinal applications. Med Res Rev 30:394–418
37. Marino K, Bones J, Kattla JJ, Rudd PM (2010) A systematic approach to protein glycosylation analysis: a path through the maze. Nat Chem Biol 6:713–723
38. Spiro RG (2002) Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds. Glycobiology 12:43R–56R
39. Mellquist JL, Kasturi L, Spitalnik SL, Shakin-Eshleman SH (1998) The amino acid following an asn-X-Ser/Thr sequon is an important determinant of N-linked core glycosylation efficiency. Biochemistry 37:6833–6837
40. Dell A, Galadari A, Sastre F, Hitchen P (2010) Similarities and differences in the glycosylation mechanisms in prokaryotes and eukaryotes. Int J Microbiol 2010:14
41. Mescher MF, Strominger JL (1976) Purification and characterization of a prokaryotic glucoprotein from the cell envelope of Halobacterium salinarium. J Biol Chem 251:2005–2014
42. Young NM et al (2002) Structure of the N-Linked glycan present on multiple glycoproteins in the gram-negative bacterium, Campylobacter jejuni. J Biol Chem 277:42530–42539
43. Stanley P (2011) Golgi glycosylation. Cold Spring Harbor Perspect. Biol 3
44. Roth Z, Yehezkel G, Khalaila I (2012) Identification and quantification of protein glycosylation. Int J Carbohydrate Chem 2012:10
45. Steentoft C et al (2013) Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. The EMBO J 32:1478–1488
46. Both P et al (2014) Discrimination of epimeric glycans and glycopeptides using IM-MS and its potential for carbohydrate sequencing. Nat Chem 6:65–74
47. Hart GW, Copeland RJ (2010) Glycomics hits the big time. Cell 143:672–676
48. Jensen PH, Karlsson NG, Kolarich D, Packer NH (2012) Structural analysis of N- and O-glycans released from glycoproteins. Nat Protocols 7:1299–1310
49. Rudd PM et al (1997) Oligosaccharide sequencing technology. Nature 388:205–207
50. Morelle W, Michalski J-C (2007) Analysis of protein glycosylation by mass spectrometry. Nat Protocols 2:1585–1602
51. Domon B, Costello CE (1988) A systematic nomenclature for carbohydrate fragmentations in FAB-MS/MS spectra of glycoconjugates. Glycoconjugate J 5:397–409
52. Kebarle P, Verkerk UH (2009) Electrospray: from ions in solution to ions in the gas phase, what we know now. Mass Spectrom Rev 28:898–917
53. Ruotolo BT, Benesch JLP, Sandercock AM, Hyung SJ, Robinson CV (2008) Ion mobility-mass spectrometry analysis of large protein complexes. Nat Protoc 3:1139–1152

54. Hall Z, Politis A, Bush MF, Smith LJ, Robinson CV (2012) Charge-state dependent compaction and dissociation of protein complexes: insights from ion mobility and molecular dynamics. J Am Chem Soc 134:3429–3438

55. Mann M, Hendrickson RC, Pandey A (2001) Analysis of proteins and proteomes by mass spectrometry. Annu Rev Biochem 70:437–473

56. Konermann L, Ahadi E, Rodriguez AD, Vahidi S (2013) Unraveling the mechanism of electrospray ionization. Anal Chem 85:2–9

57. Liuni P, Wilson DJ (2011) Understanding and optimizing electrospray ionization techniques for proteomic analysis. Expert Rev Proteomics 8:197–209

58. Watson JT, Sparkman OD (2007) Introduction to mass spectrometry: instrumentation, applications and strategies for data interpretation. Wiley, New Jersey, 2007

59. Iribarne JV, Thomson BA (1976) Evaporation of small ions from charged droplets. J Chem Phys 64:2287–2294

60. Fernandez de la Mora J (2000) Electrospray ionization of large multiply charged species proceeds via Dole's charged residue mechanism. Anal Chim Acta 406:93–104

61. Yue XF, Vahidi S, Konermann L (2014) Insights into the mechanism of protein electrospray ionization from salt adduction measurements. J Am Soc Mass Spectrom 25:1322–1331

62. Chung JK, Consta S (2012) Release mechanisms of poly(ethylene glycol) macroions from aqueous charged nanodroplets. J Phys Chem B 116:5777–5785

63. Simpson RJ (2003) Proteins and proteomics: a laboratory manual. Cold Spring Harbor Laboratory Press, New York

64. Karas M, Gluckmann M, Schafer J (2000) Ionization in matrix-assisted laser desorption/ionization: singly charged molecular ions are the lucky survivors. J Mass Spectrom 35:1–12

65. Yates JR (1998) Mass spectrometry and the age of the proteome. J Mass Spectrom 33:1–19

66. El-Aneed A, Cohen A, Banoub J (2009) Mass Spectrometry, review of the basics: electrospray, MALDI, and commonly used mass analyzers. Appl Spectrosc Rev 44:210–230

67. Lewis JK, Wei J, Siuzdak G (2006) Encyclopedia of analytical chemistry. Wiley, New Jersey

68. Knochenmuss R (2013) MALDI ionization mechanisms: the coupled photophysical and chemical dynamics model correctly predicts 'temperature'-selected spectra. J Mass Spectrom 48:998–1004

69. Awad H, Khamis MM, El-Aneed A (2015) Mass spectrometry, review of the basics: ionization. Appl Spectrosc Rev 50:158–175

70. Knochenmuss R (2006) Ion formation mechanisms in UV-MALDI. Analyst 131:966–986

71. Karas M, Kruger R (2003) Ion formation in MALDI: the cluster ionization mechanism. Chem Rev 103:427–439

72. Ehring H, Karas M, Hillenkamp F (1992) Role of photoionization and photochemistry in ionization processes of organic-molecules and relevance for matrix-assisted lader desorption ionization mass-spectrometry. Org Mass Spectrom 27:472–480

73. Jaskolla TW, Karas M (2011) Compelling evidence for lucky survivor and gas phase protonation: the unified MALDI analyte protonation mechanism. J Am Soc Mass Spectrom 22:976–988

74. Batoy S, Akhmetova E, Miladinovic S, Smeal J, Wilkins CL (2008) Developments in MALDI mass spectrometry: the quest for the perfect matrix. Appl Spectrosc Rev 43:485–550

75. Miller PE, Denton MB (1986) The quadrupole mass filter—basic operating concepts. J Chem Educ 63:617–622

76. Shaw LM, Kwong TC (2001) The clinical toxicology laboratory: contemporary practice of poisoning evaluation. American Association for Clinical Chemistry, Incorporated, USA

77. Barner-Kowollik C, Gruendling T, Falkenhagen J, Weidner S (2012) Mass spectrometry in polymer chemistry. Wiley, New Jersey

78. March RE (1997) An introduction to quadrupole ion trap mass spectrometry. J Mass Spectrom 32:351–369

79. Jonscher KR, Yates Iii JR (1997) The quadrupole ion trap mass spectrometer—a small solution to a big challenge. *Anal. Biochem.* 244:1–15
80. Payne AH, Glish GL (2005) In: Methods in enzymology, vol 402, Academic Press, pp 109–148
81. Douglas DJ, Frank AJ, Mao DM (2005) Linear ion traps in mass spectrometry. Mass Spectrom Rev 24:1–29
82. Wiley WC, McLaren IH (1955) Time-of-flight mass spectrometer with improved resolution. Rev Sci Instrum 26:1150–1157
83. Roepstorff P, Fohlman J (1984) Proposal for a common nomenclature for sequence ions in mass-spectra of peptides. Biomedical Mass Spectrometry 11:601–601
84. Dodds ED (2012) Gas-phase dissociation of glycosylated peptide ions. Mass Spectrom Rev 31:666–682
85. Dongré AR, Jones JL, Somogyi Á, Wysocki VH (1996) Influence of peptide composition, gas-phase basicity, and chemical modification on fragmentation efficiency: evidence for the mobile proton model. J Am Chem Soc 118:8365–8374
86. Burlet O, Orkiszewski RS, Ballard KD, Gaskell SJ (1992) Charge promotion of low-energy fragmentations of peptide ions. Rapid Commun Mass Spectrom 6:658–662
87. Paizs B, Suhai S (2005) Fragmentation pathways of protonated peptides. Mass Spectrom Rev 24:508–548
88. Chen XY, Flynn GC (2009) Gas-phase oligosaccharide nonreducing end (GONE) sequencing and structural analysis by reversed phase hplc/mass spectrometry with polarity switching. J Am Soc Mass Spectrom 20:1821–1833
89. Domann P, Spencer DIR, Harvey DJ (2012) Production and fragmentation of negative ions from neutral N-linked carbohydrates ionized by matrix-assisted laser desorption/ionization. Rapid Commun Mass Spectrom 26:469–479
90. Zaia J, Miller MJC, Seymour JL, Costello CE (2007) The role of mobile protons in negative ion CID of oligosaccharides. J Am Soc Mass Spectrom 18:952–960
91. Harvey DJ (2005) Ionization and fragmentation of N-linked glycans as silver adducts by electrospray mass spectrometry. Rapid Commun Mass Spectrom 19:484–492
92. Zhu F, Glover M, Shi H, Trinidad J, Clemmer D (2015) Populations of metal-glycan structures influence MS fragmentation patterns. J Am Soc Mass Spectrom 26:25–35
93. Cancilla MT, Penn SG, Carroll JA, Lebrilla CB (1996) Coordination of alkali metals to oligosaccharides dictates fragmentation behavior in matrix assisted laser desorption ionization/Fourier transform mass spectrometry. J Am Chem Soc 118:6736–6745
94. Cancilla MT, Wang AW, Voss LR, Lebrilla CB (1999) Fragmentation reactions in the mass spectrometry analysis of neutral oligosaccharides. Anal Chem 71:3206–3218
95. Zaia J (2004) Mass spectrometry of oligosaccharides. Mass Spectrom Rev 23:161–227
96. Deguchi K et al (2006) Complementary structural information of positive- and negative-ion MSn spectra of glycopeptides with neutral and sialylated N-glycans. Rapid Commun Mass Spectrom 20:741–746
97. Domon B, Mueller DR, Richter WJ (1994) Tandem mass-spectrometric analysis of fixed-charge derivatized oligosaccharides. Org Mass Spectrom 29:713–719
98. Azenha CSR, Coimbra MA, Moreira ASP, Domingues P, Domingues MRM (2013) Differentiation of isomeric β-(1–4) hexose disaccharides by positive electrospray tandem mass spectrometry. J Mass Spectrom 48:548–552
99. Brown DJ et al (2011) Direct evidence for the ring opening of monosaccharide anions in the gas phase: photodissociation of aldohexoses and aldohexoses derived from disaccharides using variable-wavelength infrared irradiation in the carbonyl stretch region. Carbohydr Res 346:2469–2481
100. Polfer NC, Bohrer BC, Plasencia MD, Paizs B, Clemmer DE (2008) On the dynamics of fragment isomerization in collision-induced dissociation of peptides. J Phys Chem A 112:1286–1293

101. Riba-Garcia I, Giles K, Bateman RH, Gaskell SJ (2008) Evidence for structural variants of a- and b-type peptide fragment ions using combined ion mobility/mass spectrometry. J Am Soc Mass Spectrom 19:609–613

102. Darula Z, Chalkley RJ, Baker P, Burlingame AL, Medzihradszky KF (2010) Mass spectrometric analysis, automated identification and complete annotation of O-linked glycopeptides. Eur J Mass Spectrom 16:421–428

103. Zaia J (2010) Mass Spectrometry and Glycomics. Omics 14:401–418

104. Franz AH, Lebrilla CB (2002) Evidence for long-range glycosyl transfer reactions in the gas phase. J Am Soc Mass Spectrom 13:325–337

105. Harvey DJ et al (2002) "Internal residue loss": Rearrangements occurring during the fragmentation of carbohydrates derivatized at the reducing terminus. Anal Chem 74:734–740

106. Wuhrer M, Koeleman CAM, Hokke CH, Deelder AM (2006) Mass spectrometry of proton adducts of fucosylated N-glycans: fucose transfer between antennae gives rise to misleading fragments. Rapid Commun Mass Spectrom 20:1747–1754

107. Wuhrer M, Koeleman CAM, Deelder AM (2009) Hexose rearrangements upon fragmentation of N-glycopeptides and reductively aminated N-glycans. Anal Chem 81:4422–4432

108. Kenny DT, Issa SMA, Karlsson NG (2011) Sulfate migration in oligosaccharides induced by negative ion mode ion trap collision-induced dissociation. Rapid Commun Mass Spectrom 25:2611–2618

109. Wuhrer M, Deelder AM, van der Burgt YEM (2011) Mass spectrometric glycan rearrangements. Mass Spectrom Rev 30:664–680

110. Brull LP et al (1997) Loss of internal 1-> 6 substituted monosaccharide residues from underivatized and per-O-methylated trisaccharides. J Am Soc Mass Spectrom 8:43–49

111. Brüll LP, Kovácik V, Thomas-Oates JE, Heerma W, Haverkamp J (1998) Sodium-cationized oligosaccharides do not appear to undergo 'internal residue loss' rearrangement processes on tandem mass spectrometry. Rapid Commun Mass Spectrom 12:1520–1532

112. Chi A et al (2007) Analysis of phosphorylation sites on proteins from Saccharomyces cerevisiae by electron transfer dissociation (ETD) mass spectrometry. Proc. Natl. Acad. Sci. USA. 104:2193–2198

113. Syka JEP, Coon JJ, Schroeder MJ, Shabanowitz J, Hunt DF (2004) Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. Proc. Natl. Acad. Sci. USA. 101:9528–9533

114. Molina H, Horn DM, Tang N, Mathivanan S, Pandey A (2007) Global proteomic profiling of phosphopeptides using electron transfer dissociation tandem mass spectrometry. Proc. Natl. Acad. Sci. USA 104:2199–2204

115. Han L, Costello C (2011) Electron transfer dissociation of milk oligosaccharides. J Am Soc Mass Spectrom 22:997–1013

116. Leach FE III et al (2012) Hexuronic Acid stereochemistry determination in chondroitin sulfate glycosaminoglycan oligosaccharides by electron detachment dissociation. J Am Soc Mass Spectrom 23:1488–1497

117. Leach FE III et al (2011) Negative electron transfer dissociation Fourier transform mass spectrometry of glycosaminoglycan carbohydrates. Eur J Mass Spectrom 17:167–176

118. Wolff JJ et al (2010) Negative electron transfer dissociation of glycosaminoglycans. Anal Chem 82:3460–3466

119. Wolff JJ, Chi L, Linhardt RJ, Amster IJ (2007) Distinguishing glucuronic from iduronic acid in glycosaminoglycan tetrasaccharides by using electron detachment dissociation. Anal Chem 79:2015–2022

120. Woodin RL, Bomse DS, Beauchamp JL (1978) Multi-photon dissociation of molecules with low-power continuous wave infrared-laser radiation. J Am Chem Soc 100:3248–3250

121. Polfer NC et al (2006) Differentiation of isomers by wavelength-tunable infrared multiple-photon dissociation-mass spectrometry: application to glucose-containing disaccharides. Anal Chem 78:670–679

release in intact and unreduced form both N-linked

122. Stefan SE, Eyler JR (2010) Differentiation of glucose-containing disaccharides by infrared multiple photon dissociation with a tunable CO2 laser and Fourier transform ion cyclotron resonance mass spectrometry. Int J Mass Spectrom 297:96–101

123. Polfer NC (2011) Infrared multiple photon dissociation spectroscopy of trapped ions. Chem Soc Rev 40:2211–2221

124. Schindler B et al (2014) Distinguishing isobaric phosphated and sulfated carbohydrates by coupling of mass spectrometry with gas phase vibrational spectroscopy. Phys Chem Chem Phys 16:22131–22138

125. Barnes L et al (2015) Anharmonic simulations of the vibrational spectrum of sulfated compounds: application to the glycosaminoglycan fragment glucosamine 6-sulfate. Phys. Chem. Chem, Phys

126. Tan Y, Polfer N (2015) Linkage and anomeric differentiation in trisaccharides by sequential fragmentation and variable-wavelength infrared photodissociation. J Am Soc Mass Spectrom 26:359–368

127. Oepts D, van der Meer AFG, van Amersfoort PW (1995) The free-electron-laser user facility FELIX. Infrared Phys Technol 36:297–308

128. Wang H et al (2014) Multiplex profiling of glycoproteins using a novel bead-based lectin array. Proteomics 14:78–86

129. Pičmanová M et al (2015) A recycling pathway for cyanogenic glycosides evidenced by the comparative metabolic profiling in three cyanogenic plant species. Biochem J 469:375–389

130. Gonzalez J et al (1992) A method for determination of N-glycosylation sites in glycoproteins by collision-induced dissociation analysis in fast atom bombardment mass spectrometry: Identification of the positions of carbohydrate-linked asparagine in recombinant α-amylase by treatment with peptide-N-glycosidase F in 18O-labeled water. Anal Biochem 205:151–158

131. Duk M, Ugorski M, Lisowska E (1997) β-Elimination of O-glycans from glycoproteins transferred to immobilon P membranes: method and some applications. Anal Biochem 253:98–102

132. Stalnaker SH et al (2010) Site mapping and characterization of O-glycan structures on α-dystroglycan Isolated from rabbit skeletal muscle. J Biol Chem 285:24882–24891

133. Patel T et al (1993) Use of hydrazine to release in intact and unreduced form both N-linked and O-linked oligosaccharides from glycoproteins. Biochemistry 32:679–693

134. Kozak RP, Royle L, Gardner RA, Fernandes DL, Wuhrer M (2012) Suppression of peeling during the release of O-glycans by hydrazinolysis. Anal Biochem 423:119–128

135. Merry AH et al (2002) Recovery of Intact 2-aminobenzamide-labeled O-glycans released from glycoproteins by hydrazinolysis. Anal Biochem 304:91–99

136. Choi E, Loo D, Dennis JW, O'Leary CA, Hill MM (2011) High-throughput lectin magnetic bead array-coupled tandem mass spectrometry for glycoprotein biomarker discovery. Electrophoresis 32:3564–3575

137. Angeloni S et al (2005) Glycoprofiling with micro-arrays of glycoconjugates and lectins. Glycobiology 15:31–41

138. Kuno A et al (2005) Evanescent-field fluorescence-assisted lectin microarray: a new strategy for glycan profiling. Nat Methods 2:851–856

139. Packer N, Lawson M, Jardine D, Redmond J (1998) A general approach to desalting oligosaccharides released from glycoproteins. Glycoconjugate J. 15:737–747

140. Itoh S et al (2002) Simultaneous microanalysis of N-linked oligosaccharides in a glycoprotein using microbore graphitized carbon column liquid chromatography-mass spectrometry. J Chromatogr A 968:89–100

141. Hennion M-C (2000) Graphitized carbons for solid-phase extraction. J Chromatogr A 885:73–95

142. Powell AK, Ahmed YA, Yates EA, Turnbull JE (2010) Generating heparan sulfate saccharide libraries for glycomics applications. Nat Protoc 5:821–833

143. Harvey DJ (2000) Electrospray mass spectrometry and fragmentation of N-linked carbohydrates derivatized at the reducing terminus. J Am Soc Mass Spectrom 11:900–915

144. Bigge JC et al (1995) Nonselective and efficient fluorescent labeling of glycans using 2-amino benzamide and anthranilic acid. Anal Biochem 230:229–238

145. Guile GR, Rudd PM, Wing DR, Prime SB, Dwek RA (1996) A rapid high-resolution high-performance liquid chromatographic method for separating glycan mixtures and analyzing oligosaccharide profiles. Anal Biochem 240:210–226

146. Takegawa Y et al (2006) Simple separation of isomeric sialylated N-glycopeptides by a zwitterionic type of hydrophilic interaction chromatography. J Sep Sci 29:2533–2540

147. Lareau NM, May JC, McLean JA (2015) Non-derivatized glycan analysis by reverse phase liquid chromatography and ion mobility-mass spectrometry. The Analyst 140:3335–3338

148. Alley WR, Mann BF, Novotny MV (2013) High-sensitivity analytical approaches for the structural characterization of glycoproteins. Chem Rev 113:2668–2732

149. Stadlmann J, Pabst M, Kolarich D, Kunert R, Altmann F (2008) Analysis of immunoglobulin glycosylation by LC-ESI-MS of glycopeptides and oligosaccharides. Proteomics 8:2858–2871

150. Lohmander LS (1986) Analysis by high-performance liquid chromatography of radioactively labeled carbohydrate components of proteoglycans. Anal Biochem 154:75–84

151. Pettolino FA, Walsh C, Fincher GB, Bacic A (2012) Determining the polysaccharide composition of plant cell walls. Nat. Protocols 7:1590–1607

152. Byers HL, Tarelli E, Homer KA, Beighton D (1999) Sequential deglycosylation and utilization of the N-linked, complex-type glycans of human a1-acid glycoprotein mediates growth of Streptococcus oralis. Glycobiology 9:469–479

153. Kannicht C, Grunow D, Lucka L (2008) In: Kannicht C (ed) Post-translational modifications of proteins, Vol 446, pp 255–266. Humana Press

154. Distler JJ, Jourdian GW (1973) The purification and properties of β-galactosidase from bovine testes. J Biol Chem 248:6772–6780

155. Zahner D, Hakenbeck R (2000) The Streptococcus pneumoniae beta-galactosidase is a surface protein. J Bacteriol 182:5919–5921

156. Clarke VA, Platt N, Butters TD (1995) Cloning and expression of the β-N-acetylglucosaminidase gene from Streptococcus pneumoniae: generation of truncated enzymes with modified aglycon specificity. J Biol Chem 270:8805–8814

157. Camara M, Boulnois GJ, Andrew PW, Mitchell TJ (1994) A neuraminidase from Streptococcus-pneumoniae was the features of surface protein. Infect Immun 62:3688–3695

158. Wongmadden ST, Landry D (1995) Purification and characterization of novel glycosidases from the bacterial genus Xanthomonas. Glycobiology 5:19–28

159. Pilobello KT, Krishnamoorthy L, Slawek D, Mahal LK (2005) Development of a lectin microarray for the rapid analysis of protein glycopatterns. ChemBioChem 6:985–989

160. Zheng T, Peelen D, Smith LM (2005) Lectin arrays for profiling cell surface carbohydrate expression. J Am Chem Soc 127:9982–9983

161. Hu S, Wong DT (2009) Lectin microarray. Proteom. Clin. Appl. 3:148–154

162. Beloqui A et al (2013) Analysis of Microarrays by MALDI-TOF MS. Angew Chem Int Ed 52:7477–7481

163. Debray H, Decout D, Strecker G, Spik G, Montreuil J (1981) Specificity of 12 lectins towards oligosaccharides and glycopeptides relted to N-glycosylproteins. Eur J Biochem 117:41–55

164. Geisler C, Jarvis DL (2011) Letter to the Glyco-Forum: Effective glycoanalysis with Maackia amurensis lectins requires a clear understanding of their binding specificities. Glycobiology 21:988–993

165. Porter A et al (2010) A motif-based analysis of glycan array data to determine the specificities of glycan-binding proteins. Glycobiology 20:369–380

166. Yan LY et al (1997) Immobilized Lotus tetragonolobus agglutinin binds oligosaccharides containing the Le(x) determinant. Glycoconjugate J 14:45–55

167. Kletter D, Singh S, Bern M, Haab BB (2013) Global Comparisons of Lectin-Glycan Interactions Using a Database of Analyzed Glycan Array Data. Mol Cell Proteomics 12:1026–1035

168. Consortium for Functional Glycomics, http://www.functionalglycomics.org/. (2001)
169. Akimoto Y, Kawakami H (2014) In: Hirabayashi J (ed) Lectins, Vol. 1200, pp 153–163. Springer, New York
170. Mannoji H, Yeger H, Becker LE (1986) A specific histochemical marker (lectin Ricinus-communis agglutinin-1) for normal human microglia, and apllication to routine histopathology. Acta Neuropathol 71:341–343
171. Landemarre L, Cancellieri P, Duverger E (2013) Cell surface lectin array: parameters affecting cell glycan signature. Glycoconjugate J. 30:195–203
172. Nishijima Y et al (2012) Glycan profiling of endometrial cancers using lectin microarray. Genes Cells 17:826–836
173. Furukawa J-I et al (2008) Comprehensive Approach to Structural and Functional Glycomics Based on Chemoselective Glycoblotting and Sequential Tag Conversion. Anal Chem 80:1094–1101
174. Blixt O et al (2004) Printed covalent glycan array for ligand profiling of diverse glycan binding proteins. Proc. Natl. Acad. Sci. U S A 101:17033–17038
175. Blixt O et al (2008) Glycan microarrays for screening sialyltransferase specificities. Glycoconjugate J. 25:59–68
176. Blixt O et al (2010) A high-throughput O-glycopeptide discovery platform for seromic profiling. J Proteome Res 9:5250–5261
177. Ban L et al (2012) Discovery of glycosyltransferases using carbohydrate arrays and mass spectrometry. Nat Chem Biol 8:769–773
178. Paulson JC, Blixt O, Collins BE (2006) Sweet spots in functional glycomics. Nat Chem Biol 2:238–248
179. Remy-Martin F et al (2012) Surface plasmon resonance imaging in arrays coupled with mass spectrometry (SUPRA–MS): proof of concept of on-chip characterization of a potential breast cancer marker in human plasma. Anal Bioanal Chem 404:423–432
180. Munoz FJ et al (2009) Glycan tagging to produce bioactive ligands for a surface plasmon resonance (SPR) study via immobilization on different surfaces. Bioconjugate Chem. 20:673–682
181. Etxebarria J, Calvo J, Martin-Lomas M, Reichardt NC (2012) Lectin-array blotting: profiling protein glycosylation in complex mixtures. ACS Chem Biol 7:1729–1737
182. Koshi Y, Nakata E, Yamane H, Hamachi I (2006) A fluorescent lectin array using supramolecular hydrogel for simple detection and pattern profiling for various glycoconjugates. J Am Chem Soc 128:10413–10422
183. Li C et al (2009) Pancreatic cancer serum detection using a lectin/glyco-antibody array method. J Proteome Res 8:483–492
184. Hsu K-L, Mahal LK (2006) A lectin microarray approach for the rapid analysis of bacterial glycans. Nat. Protocols 1:543–549
185. Hsu KL, Pilobello KT, Mahal LK (2006) Analyzing the dynamic bacterial glycome with a lectin microarray approach. Nat Chem Biol 2:153–157
186. Zhao J, Simeone DM, Heidt D, Anderson MA, Lubman DM (2006) Comparative serum glycoproteomics using lectin selected sialic acid glycoproteins with mass spectrometric analysis: application to pancreatic cancer serum. J Proteome Res 5:1792–1802
187. Vanderschaeghe D, Festjens N, Delanghe J, Callewaert N (2010) Glycome profiling using modern glycomics technology: technical aspects and applications. Biol Chem 391:149–161
188. Campbell MP, Royle L, Radcliffe CM, Dwek RA, Rudd PM (2008) GlycoBase and autoGU: tools for HPLC-based glycan analysis. Bioinformatics 24:1214–1216
189. Royle L et al (2008) HPLC-based analysis of serum N-glycans on a 96-well plate platform with dedicated database software. Anal Biochem 376:1–12
190. Harvey DJ (2011) Derivatization of carbohydrates for analysis by chromatography; electrophoresis and mass spectrometry. J Chromatogr B 879:1196–1225
191. Chen X, Flynn GC (2007) Analysis of N-glycans from recombinant immunoglobulin G by on-line reversed-phase high-performance liquid chromatography/mass spectrometry. Anal Biochem 370:147–161

192. Hase S (1993) Analysis of sugar chains by pyridylamination. Methods in molecular biology (Clifton, N.J.) 14, 69–80

193. Harvey DJ (1999) Matrix-assisted laser desorption/ionization mass spectrometry of carbohydrates. Mass Spectrom Rev 18:349–450

194. Viseux N, de Hoffmann E, Domon B (1998) Structural assignment of permethylated oligosaccharide subunits using sequential tandem mass spectrometry. Anal Chem 70:4951–4959

195. Higel F, Demelbauer U, Seidl A, Friess W, Sörgel F (2013) Reversed-phase liquid-chromatographic mass spectrometric N-glycan analysis of biopharmaceuticals. Anal Bioanal Chem 405:2481–2493

196. Rocklin RD, Pohl CA (1983) Determination of carbohydrates by anion-exchange chromatography with pulsed amperometric detection. J Liq Chromatogr 6:1577–1590

197. Hagglund P, Bunkenborg J, Elortza F, Jensen ON, Roepstorff P (2004) A new strategy for identification of N-glycosylated proteins and unambiguous assignment of their glycosylation sites using HILIC enrichment and partial deglycosylation. J Proteome Res 3:556–566

198. Itoh S et al (2006) N-linked oligosaccharide analysis of rat brain Thy-1 by liquid chromatography with graphitized carbon column/ion trap-Fourier transform ion cyclotron resonance mass spectrometry in positive and negative ion modes. J Chromatogr A 1103:296–306

199. Pabst M, Bondili JS, Stadlmann J, Mach L, Altmann F (2007) Mass + retention time = structure: a strategy for the analysis of N-glycans by carbon LC-ESI-MS and its application to fibrin N-glycans. Anal Chem 79:5051–5057

200. Rohrer JS, Basumallick L, Hurum D (2013) High-performance anion-exchange chromatography with pulsed amperometric detection for carbohydrate analysis of glycoproteins. *Biochem.-Moscow* 78:697–709

201. Lee YC (1996) Carbohydrate analyses with high-performance anion-exchange chromatography. J Chromatogr A 720:137–149

202. Buszewski B, Noga S (2012) Hydrophilic interaction liquid chromatography (HILIC)—a powerful separation technique. Anal Bioanal Chem 402:231–247

203. Antonio C et al (2008) Hydrophilic interaction chromatography/electrospray mass spectrometry analysis of carbohydrate-related metabolites from Arabidopsis thaliana leaf tissue. Rapid Commun Mass Spectrom 22:1399–1407

204. Melmer M, Stangler T, Premstaller A, Lindner W (2011) Comparison of hydrophilic-interaction, reversed-phase and porous graphitic carbon chromatography for glycan analysis. J Chromatogr A 1218:118–123

205. West C, Elfakir C, Lafosse M (2010) Porous graphitic carbon: a versatile stationary phase for liquid chromatography. J Chromatogr A 1217:3201–3216

206. Karlsson NG et al (2004) Negative ion graphitised carbon nano-liquid chromatography/mass spectrometry increases sensitivity for glycoprotein oligosaccharide analysis. Rapid Commun Mass Spectrom 18:2282–2292

207. Heinisch S, Puy G, Barrioulet M-P, Rocca J-L (2006) Effect of temperature on the retention of ionizable compounds in reversed-phase liquid chromatography: application to method development. J Chromatogr A 1118:234–243

208. Ortmayr K, Hann S, Koellensperger G (2015) Complementing reversed-phase selectivity with porous graphitized carbon to increase the metabolome coverage in an on-line two-dimensional LC-MS setup for metabolomics. Analyst 140:3465–3473

209. Bajad SU et al (2006) Separation and quantitation of water soluble cellular metabolites by hydrophilic interaction chromatography-tandem mass spectrometry. J Chromatogr A 1125:76–88

210. Castagner B and Seeberger PH (2007) In: Brase S (ed) Combinatorial chemistry on solid supports, vol 278, pp 289–309. Springer-Verlag Berlin, Berlin

211. Bernardes GJL, Castagner B, Seeberger PH (2009) Combined approaches to the synthesis and study of glycoproteins. ACS Chem Biol 4:703–713

212. Hellerqvist CG (1990) In: Methods in Enzymology, vol 193, pp 554–573. Academic Press

213. Creese AJ, Cooper HJ (2012) Separation and identification of isomeric glycopeptides by high field asymmetric waveform ion mobility spectrometry. Anal Chem 84:2597–2601
214. Spina E et al (2004) New fragmentation mechanisms in matrix-assisted laser desorption/ionization time-of-flight/time-of-flight tandem mass spectrometry of carbohydrates. Rapid Commun Mass Spectrom 18:392–398
215. Konda C, Londry F, Bendiak B, Xia Y (2014) Assignment of the stereochemistry and anomeric configuration of sugars within oligosaccharides via overlapping disaccharide ladders using MSn. J Am Soc Mass Spectrom 25:1441–1450
216. Konda C, Bendiak B, Xia Y (2014) Linkage determination of linear oligosaccharides by MSn (n > 2) Collision-induced dissociation of Z(1) ions in the negative ion mode. J Am Soc Mass Spectrom 25:248–257
217. Merkle RK, Poppe I (1994) In: William GWH, Lennarz J (ed) Methods in enzymology, vol 230, pp 1–15, Academic Press
218. Xie Y, Tseng K, Lebrilla C, Hedrick J (2001) Targeted use of exoglycosidase digestion for the structural elucidation of neutral O-linked oligosaccharides. J Am Soc Mass Spectrom 12:877–884
219. Maslen S, Sadowski P, Adam A, Lilley K, Stephens E (2006) Differentiation of isomeric N-glycan structures by normal-phase liquid chromatography—MALDI-TOF/TOF tandem mass spectrometry. Anal Chem 78:8491–8498
220. Frolov A, Hoffmann P, Hoffmann R (2006) Fragmentation behavior of glycated peptides derived from D-glucose, D-fructose and D-ribose in tandem mass spectrometry. J Mass Spectrom 41:1459–1469
221. Xue J et al (2004) Determination of linkage position and anomeric configuration in Hex-Fuc disaccharides using electrospray ionization tandem mass spectrometry. Rapid Commun Mass Spectrom 18:1947–1955
222. El-Aneed A, Banoub J, Koen-Alonso M, Boullanger P, Lafont D (2007) Establishment of mass spectrometric fingerprints of novel synthetic cholesteryl neoglycolipids: the presence of a unique C-glycoside species during electrospray ionization and during collision-induced dissociation tandem mass spectrometry. J Am Soc Mass Spectrom 18:294–310
223. Xue J, Laine RA, Matta KL (2015) Enhancing MSn mass spectrometry strategy for carbohydrate analysis: A b2 ion spectral library. J. Proteomics 112:224–249
224. Ito H et al (2006) Direct structural assignment of neutral and sialylated N-glycans of glycopeptides using collision-induced dissociation MSn spectral matching. Rapid Commun Mass Spectrom 20:3557–3565
225. Konda C, Bendiak B, Xia Y (2012) Differentiation of the stereochemistry and anomeric configuration for 1-3 linked disaccharides via tandem mass spectrometry and O-18-labeling. J Am Soc Mass Spectrom 23:347–358
226. Toyama A et al (2012) Quantitative structural characterization of local N-glycan microheterogeneity in therapeutic antibodies by energy-resolved oxonium ion monitoring. Anal Chem 84:9655–9662
227. Daikoku S, Widmalm G, Kanie O (2009) Analysis of a series of isomeric oligosaccharides by energy-resolved mass spectrometry: a challenge on homobranched trisaccharides. Rapid Commun Mass Spectrom 23:3713–3719
228. Nagy G, Pohl NB (2015) Complete hexose isomer identification with mass spectrometry. J Am Soc Mass Spectrom 26:677–685
229. Hoffmann W, Hofmann J, Pagel K (2014) Energy-resolved ion mobility-mass spectrometry —a concept to improve the separation of isomeric carbohydrates. J Am Soc Mass Spectrom 1–9
230. Nagy G, Pohl NLB (2015) Monosaccharide identification as a first step toward de novo carbohydrate sequencing: mass spectrometry strategy for the identification and differentiation of diastereomeric and enantiomeric pentose isomers. Anal Chem 87:4566–4571
231. Lee S, Wyttenbach T, Bowers MT (1997) Gas phase structures of sodiated oligosaccharides by ion mobility/ion chromatography methods. Int J Mass Spectrom Ion Process 167–168:605–614

232. Fenn LS, McLean JA (2011) Structural resolution of carbohydrate positional and structural isomers based on gas-phase ion mobility-mass spectrometry. Phys Chem Chem Phys 13:2196–2205
233. Williams JP et al (2010) Characterization of simple isomeric oligosaccharides and the rapid separation of glycan mixtures by ion mobility mass spectrometry. Int J Mass Spectrom 298:119–127
234. Lundborg M, Widmalm G (2011) Structural analysis of glycans by nmr chemical shift prediction. Anal Chem 83:1514–1517
235. Barb AW, Prestegard JH (2011) NMR analysis demonstrates immunoglobulin G N-glycans are accessible and dynamic. Nat Chem Biol 7:147–153
236. Van Calsteren M-R, Gagnon F, Nishimura J, Makino S (2015) Structure determination of the neutral exopolysaccharide produced by Lactobacillus delbrueckii subsp. bulgaricus OLL1073R-1. Carbohydr Res 413:115–122
237. Wormald MR et al (2002) Conformational studies of oligosaccharides and glycopeptides: complementarity of NMR, X-ray crystallography, and molecular modelling. Chem Rev 102:371–386
238. Perez S et al (1996) Crystal and molecular structure of a histo-blood group antigen involved in cell adhesion: the Lewis x trisaccharide. Glycobiology 6:537–542
239. Rudd PM et al (1999) Roles for glycosylation of cell surface receptors involved in cellular immune recognition. J Mol Biol 293:351–366
240. Huang CY et al (2006) Carbohydrate microarray for profiling the antibodies interacting with Globo H tumor antigen. Proc. Natl. Acad. Sci. U. S. A. 103:15–20
241. Serna S, Yan S, Martin-Lomas M, Wilson IBH, Reichardt NC (2011) Fucosyltransferases as synthetic tools: glycan array based substrate selection and core fucosylation of synthetic N-glycans. J Am Chem Soc 133:16495–16502
242. Sanchez-Ruiz A, Serna S, Ruiz N, Martin-Lomas M, Reichardt NC (2011) MALDI-TOF mass spectrometric analysis of enzyme activity and lectin trapping on an array of N-glycans. Angew. Chem.-Int. Edit. 50:1801–1804
243. Bundy J, Fenselau C (1999) Lectin-based affinity capture for MALDI-MS analysis of bacteria. Anal Chem 71:1460–1463
244. Gray CJ, Weissenborn MJ, Eyers CE, Flitsch SL (2013) Enzymatic reactions on immobilised substrates. Chem. Soc, Rev
245. Love JC, Estroff LA, Kriebel JK, Nuzzo RG, Whitesides GM (2005) Self-assembled monolayers of thiolates on metals as a form of nanotechnology. Chem Rev 105:1103–1169
246. Perkins DN, Pappin DJC, Creasy DM, Cottrell JS (1999) Probability-based protein identification by searching sequence databases using mass spectrometry data. Electrophoresis 20:3551–3567
247. Hirabayashi J, Yamada M, Kuno A, Tateno H (2013) Lectin microarrays: concept, principle and applications. Chem Soc Rev 42:4443–4458

Masking Strategies for the Bioorthogonal Release of Anticancer Glycosides

**Belén Rubio-Ruiz, Thomas L. Bray, Ana M. López-Pérez
and Asier Unciti-Broceta**

Abstract Significant progress in the bioorthogonal field has resulted in the advent of a new type of prodrug: *bioorthogonal prodrugs*, i.e. metabolically stable precursors of therapeutic agents that are specifically activated by non-native, non-biological, non-perturbing physical or chemical stimuli. The application of such unique drug precursors in conjunction with their corresponding activating source is under preclinical experimentation as a novel way to elicit site-specific activation of cytotoxic drugs, with particular emphasis on anticancer glycosides. In this chapter, the strategies developed for the masking and bioorthogonal release of cytotoxic nucleosides using benign electromagnetic radiations, biocompatible click chemistry and bioorthogonal organometallic (BOOM) catalysis will be discussed in detail.

1 Basic Principles

1.1 Glycosides

Formally, a glycoside is any molecule in which a sugar group is bonded through its anomeric carbon to another group via a glycosidic bond [1]. The sugar residue of the glycoside is termed the glycone, which consist of a single sugar moiety (monosaccharide) or multiple sugar moieties (oligosaccharide), and the non-sugar part is the aglycone or genin part. Glycosides were originally defined as mixed acetals derived from cyclic forms of saccharides. Nowadays, the term 'glycoside' encompasses not only O-glycosides, i.e. compounds with the anomeric OH replaced by a –OR group, but also thioglycosides (–SR), selenoglycosides (–SeR), glycosylamines (–NR_1R_2), and C-glycosyl compounds (–$CR_1R_2R_3$) [2, 3].

B. Rubio-Ruiz · T.L. Bray · A.M. López-Pérez · A. Unciti-Broceta (✉)
Edinburgh Cancer Research UK Centre, Institute of Genetics and Molecular Medicine,
University of Edinburgh, Crewe Road South, Edinburgh EH4 2XR, UK
e-mail: Asier.Unciti-Broceta@igmm.ed.ac.uk

Many bioactive small molecules, including hormones, antibiotics, chemotherapeutics, etc., contain sugar residues. Although establishing a general rule on the functional role of such glycones in their respective glycosides is difficult, some general trends derived from selected examples show the potential applications of glycosylation. Normally, glycosides are more hydrophilic than the respective aglycones. Such glycosylation strongly influences the pharmacodynamic and/or pharmacokinetic properties of the parent compound, e.g., circulation, elimination, and distribution across body compartments, including transport through important barriers as the hematoencephalic or placental barrier. There are, however, glycosides with specific biological functions that do not derive from the biological activity of the corresponding aglycone, i.e. the final activity is then given by the overall molecular structure [4].

1.2 Anticancer Glycosides

In the vast anticancer therapeutic arsenal, glycosides have played a prominent role as chemotherapeutic agents, either in their natural, semi-synthetic, or synthetic forms. Among the natural glycosides, anthracyclines produced by bacteria of the genus *Streptomyces* and closely related genera stand out because of their broad spectrum of activity against blood and solid malignancies [5]. The first compounds discovered of this structural class of natural products were doxorubicin (**DOX**) and daunorubicin (**DNR**), isolated from the pigment-producing *Streptomyces peucetius* in the early 1960s [6]. Structurally, both are *O*-glycosides containing a tetracyclic ring system that consists of an anthraquinone fused to a cyclohexyl ring—the aglycone part—and an aminosugar (daunosamine) attached by a glycosidic bond at the C7 position of the ring (Fig. 1). The only structural difference between **DOX** and **DNR** is the presence of a hydroxyl group at the C14 position in **DOX**, which
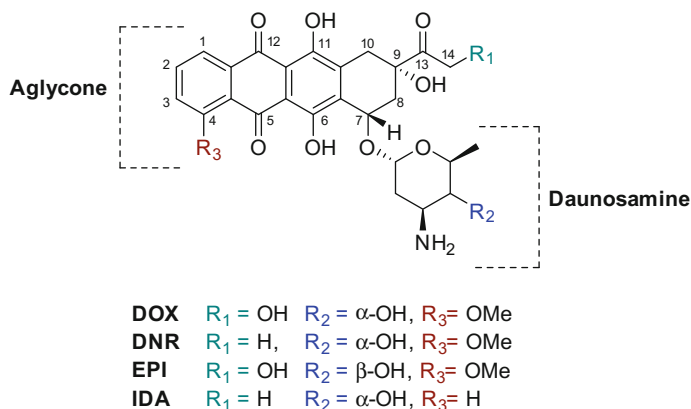

DOX	$R_1 = OH$	$R_2 = \alpha\text{-OH}$,	$R_3 = OMe$
DNR	$R_1 = H$,	$R_2 = \alpha\text{-OH}$,	$R_3 = OMe$
EPI	$R_1 = OH$	$R_2 = \beta\text{-OH}$,	$R_3 = OMe$
IDA	$R_1 = H$	$R_2 = \alpha\text{-OH}$,	$R_3 = H$

Fig. 1 Chemical structure of clinically approved anticancer anthracyclines

dramatically affects their spectrum of activity; whereas **DOX** is used against breast cancer, childhood solid tumors, soft tissue sarcomas, and aggressive lymphomas, **DNR** shows superior activity against acute lymphoblastic or myeloblastic leukemias [6].

Several modes of action have been postulated for anthracyclines. Inhibition of topoisomerase II is considered the principal cytotoxic mechanism which occurs by intercalation of the aglycone portion of the anthracycline between adjacent DNA base pairs. While the planar ring system is important for intercalation into DNA, the cyclohexane ring A and the aminosugar residue play a major role in the formation and stabilization of the anthracycline–DNA–topoisomerase II ternary complex. This stabilization results in the production of double-strand DNA breaks. Other reported modes of action include induction of apoptosis, inhibition of RNA synthesis, a one- and two-electron reduction to reactive compounds, and the formation of reactive oxygen species (ROS) that are able to damage DNA and cell membranes [7].

Despite their medical importance (**DOX** and **DNR** are listed on the WHO Model List of Essential Medicines), the clinical efficacy of both **DOX** and **DNR** is hampered by the development of chemoresistant tumor cells and by a cumulative dose-dependent cardiotoxicity, which can cause irreversible heart failure. The search for novel anthracyclines with enhanced activity and cardiac tolerability has resulted in many analogs, e.g. epirubicin (EPI) and idarubicin (**IDA**) that exhibit improved therapeutic index, but the risk of inducing cardiomyopathy is not mitigated [6].

Included in the WHO Model List of Essential Medicines is also etoposide, a glycoside derivative of the non-alkaloid natural product Podophyllotoxin [8]. As a topoisomerase II inhibitor unrelated to the anthracycline family, etoposide poisons the binary covalent complex DNA-Topoisomerase II, resulting in the formation of double-strand DNA breaks [9]. Approved by FDA in 1983, etoposide is currently used for the treatment of several malignancies such as testicular cancer, non-small cell lung cancer (NSCLC), non-Hodgkin's lymphoma or Ewing sarcoma, although its therapeutic use is limited by toxicity involving mainly myelosuppression.

Other important group of anticancer glycosides is constituted by nucleoside analogs, some of which have been in clinical use for almost 50 years. Since the initial approval of cytarabine in 1969 for the treatment of acute myeloid leukemia, numerous nucleoside analogs have been developed as antimetabolites and approved by the US Food and Drug Administration (FDA) for the treatment of different types of cancers [10]. These compounds are chemically modified derivatives of nucleosides that mimic their physiological counterparts, disrupting cell growth, and survival either by incorporation into the DNA and RNA macromolecules, by interference with various enzymes involved in the synthesis of nucleic acids, and/or by modification of the metabolism of physiological nucleosides [11].

Subtle modifications in the nitrogen base and, importantly, in the carbohydrate moiety, are the basis of the chemical and clinical diversity of these compounds (Fig. 2). Fluoropyrimidines, such as floxuridine (**FUDR**) and capecitabine (Fig. 3), mostly exert their cytotoxic effect by inhibition of thymidylate synthase which

Fig. 2 General structure of nucleoside analogs (*N*-glycosides) and their main chemical modifications
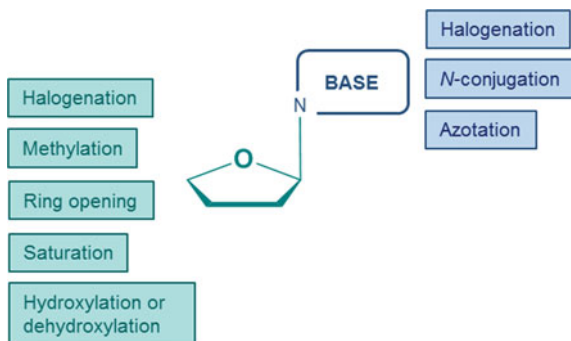
induces a status of deoxythimidine monophosphate deficiency and imbalance in the nucleotide pool that impairs DNA replication and repair [12].

Other nucleoside analogs such as cytarabine (cytidine epimer) or gemcitabine (Fig. 3) are however converted to their respective nucleotide analogs, which in turn inhibit DNA synthesis by inhibition of DNA polymerases and/or ribonucleotide reductase. Nonetheless, there are differences in the interaction of these agents and their cytotoxic metabolites with the various metabolic pathways and intracellular targets, thereby imparting unique properties to each of these agents and resulting in unique clinical activity [13].

Despite the availability of several nucleoside analogs in the clinic, the development of new and improved agents is justified by the need to overcome issues of resistance, poor oral bioavailability, long-term toxicity, and inter-individual variability requiring dose adaptation. New nucleosides that have important modifications in their base and sugar moieties are currently in clinical trials. Sapacitabine, currently in Phase 2 trials in elderly patients with myelodysplastic syndromes, NSCLC and chronic lymphocytic leukemia, is an example [10].

1.3 Bioresponsive Versus Bioorthogonal Prodrugs

In spite of their medical importance, the clinical application of the chemotherapeutic drugs described above is limited by lack of selectivity [14, 15]. The mode of action of these drugs renders them particularly harmful to healthy tissues with a high rate of cell regeneration which leads to dose-limiting adverse effects such as myelosuppression. To mitigate unwanted side effects while increasing the levels of drug in the disease area, many efforts have been made towards the design of cancer-specific strategies. One such strategy consists of transforming chemotherapeutic agents into latent prodrugs, which are derivatives devoid of pharmacological activity that are converted into their active forms by specific stimuli [16, 17]. Most popular prodrug strategies used for anticancer glycosides take advantage of the biochemical differences between healthy and malignant tissues. High expression
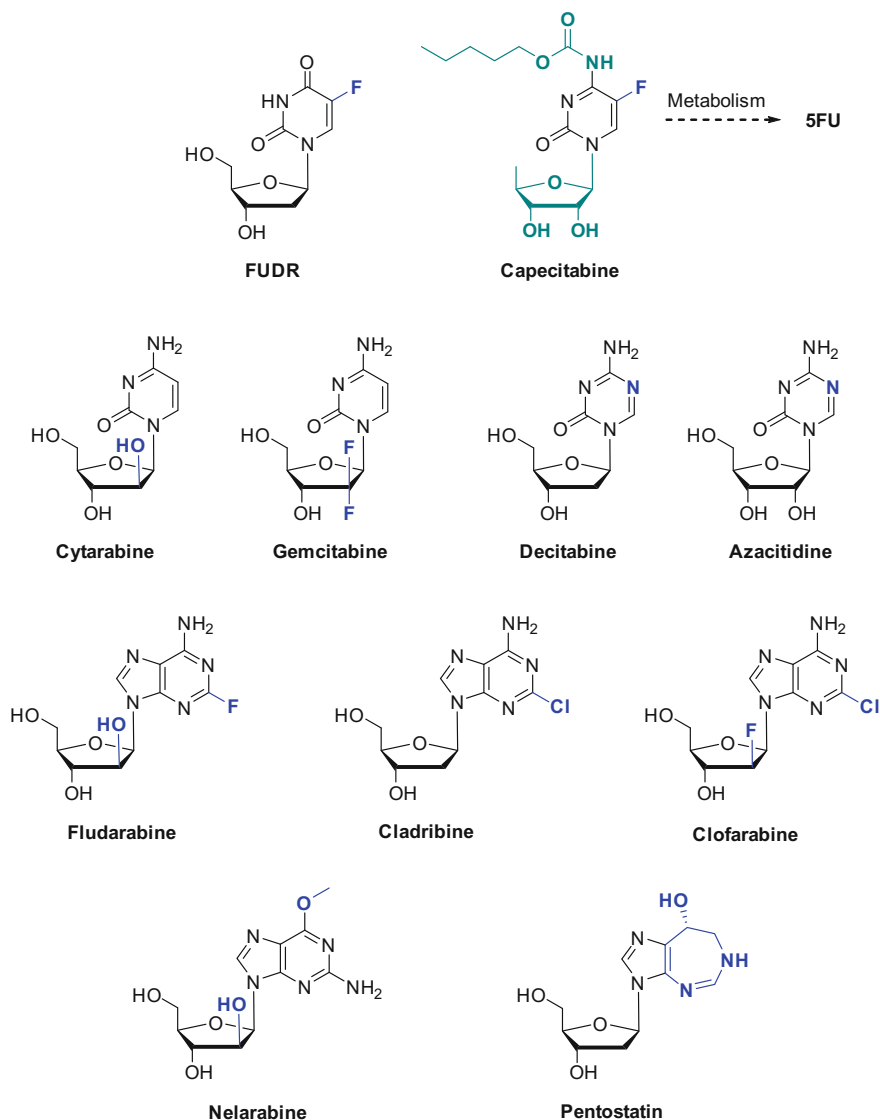
Fig. 3 FDA-approved nucleoside analogs. Chemical modifications are highlighted in *blue*

levels of specific intracellular enzymes, tumor-specific antigens expressed on tumor cell surfaces, the hypoxic environment inside the tumor or low extracellular pH have been widely employed to achieve cancer-specific prodrug activation [18]. Drug precursors that are designed to exploit such biological or physiological stimuli are herein classified as *bioresponsive prodrugs*.

Using enzyme-specific substrates, prodrugs can become preferentially activated at the tumor site. Many tumor-associated enzymes belonging to four IUPAC classes —oxidoreductases, transferases, hydrolases, and lyases—are considered optimal targets to catalyze drug release in a stringently controlled fashion [19, 20]. One of them, β-glucuronidase—an enzyme that hydrolyzes the glycosidic bond of glucuronides—have been intensively explored over the past three decades for the selective activation of several glucuronidated prodrugs of anthracyclines [21]. As β-Glucuronidase is overexpressed in a wide range of tumor types and is particularly localized in necrotic areas, this strategy was conceived to overcome the dose-related cardiotoxicity of anthracyclines.

With few exceptions (e.g. **EPI**-glucuronide [22]), the majority of the reported anthracycline prodrugs activated by human β-glucuronidase present a self-immolative linker between the carbohydrate masking group and the drug which enhances the hydrolysis efficiency by the enzyme. **HMR 1826** [23] and **DOX-GA3** [24] (glucuronidated prodrugs of **DOX**) and **DNR-GA3** [25, 26] (obtained from **DNR**) respond to this general structure (Fig. 4).

Although these glucuronide antitumor prodrugs can be applied in monotherapy, their effectiveness was increased using an antibody-directed enzyme prodrug therapy (ADEPT). Unlike anticancer prodrug monotherapy in which prodrugs are designed for direct activation by a tumor-associated factor (e.g. an enzyme), in the ADEPT concept, antigens expressed on tumors cells are used to target and deliver exogenous prodrug-activating enzymes to the tumor site [27]. Upon localization of such antibody–enzyme conjugates in the tumor, prodrugs are administrated. Using this strategy, prodrugs **HMR 1826**, **DOX-GA3**, and **DNR-GA3** showed favorable therapeutic effects compared with the parent drug, but the administration of very high doses was usually required to achieve a significant therapeutic effect [21].

Tumor hypoxia is another well-characterized feature that has been explored for targeted prodrug activation. Solid tumors often contain an inefficient microvascular system and, as result, the presence of regions with acute or chronic hypoxia in tumors is common in many cancers. Although hypoxia is a major problem in
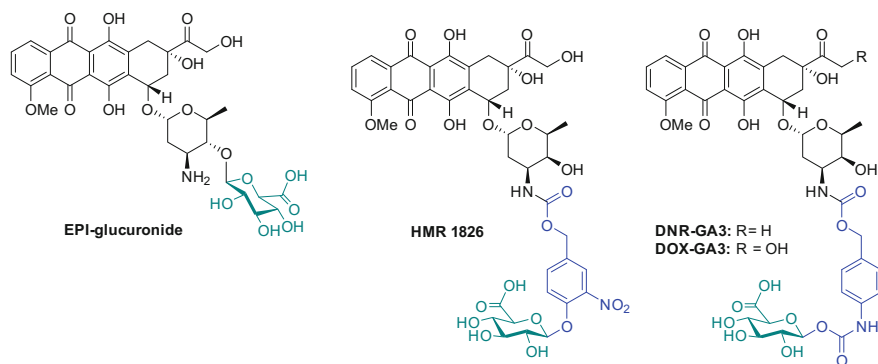

Fig. 4 Examples of glucuronidated (bioresponsive) prodrugs of anthracyclines

radiotherapy and chemotherapy, these low oxygen levels have been turned into a therapeutic advantage by designing hypoxia-responsive prodrugs that are able to release active cytotoxic agents upon reduction under hypoxic conditions [28]. Most hypoxia selective strategies are designed for reduction by endogenous reducing enzymes such as cytochrome P450 reductase or DT-diaphorase. Although these oxidoreductases are also present in normal aerobic cells, the selectivity of hypoxia-responsive prodrugs is favored in tumor cells because the reduced derivatives are rapidly oxidized back to the inactive prodrug in normoxia. Alternative strategies based on the expression of exogenous reductive enzymes into tumors for prodrug activation have also been developed, e.g. gene directed enzyme prodrug therapy (GDEPT) and virus directed enzyme prodrug therapy (VDEPT). Examples of reductive enzymes used in such approaches include cytochrome P450s, *E. coli* nitroreductase, and DT-diaphorase [29].

To date, the main classes of bioreductive compounds that undergo enzymatic reduction include quinones, nitro heterocyclic compounds, aromatic *N*-oxides, aliphatic *N*-oxides, and metal complexes [30]. In a related approach, some of these reducible functions are used as promoieties to effectively release anticancer cyto-toxic agents upon bioreduction. This is the case of the nitroaromatic group shown in Fig. 5, whose reduction induces an intramolecular cyclization reaction and subse-quent release of a conjugated drug. Following this strategy, different prodrugs of **FUDR** were designed [31, 32]. Although cyclization occurs rapidly upon reduction of the nitro group, these compounds are not good substrates of nitroreductase which impeded an efficient enzymatic reduction.

As previously mentioned, the great majority of anticancer prodrugs reported to date are *bioresponsive prodrugs*, i.e. drug precursors that become active through a biochemical process. In the last few years, however, a number of novel methods originated from the Chemical Biology field have emerged to explore novel site-specific activation of cytotoxic drugs. Since Carolyn Bertozzi showcased the potential of performing chemical modifications in living biological systems without interfering with the host biochemistry, many labs around the globe have felt inspired by the concept of bioorthogonality [33–36]. Significant progress on the use of benign non-biological means to activate drug precursors [37] have resulted in the advent of a new type of prodrug: *bioorthogonal prodrugs*, i.e. biologically stable drug precursors that are specifically activated by non-native, non-biological,
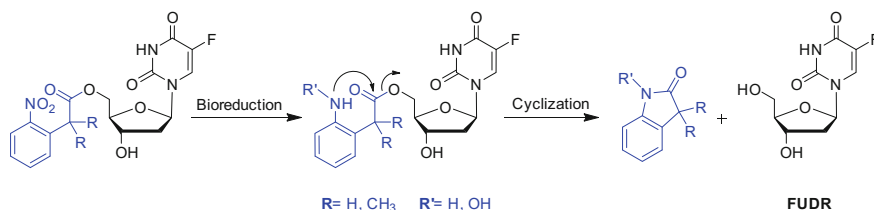

Fig. 5 Bioreductively activated prodrug of **FUDR**. The strategy is based on the use of a nitro aromatic group that upon reduction becomes an intramolecular nucleophilic trigger [31, 32]

non-perturbing means [38]. During the subsequent subchapters, the strategies used for the masking and bioorthogonal release of cytotoxic glycosides using benign electromagnetic radiations, biocompatible click chemistry and bioorthogonal organometallic (BOOM) catalysis will be discussed in detail.

2 Photoactivated Prodrugs

2.1 Introduction to Phototherapy

Generally speaking, phototherapy relates to the use of light in the treatment of disease. The systemic administration of a photosensitizer followed by the application of light on a specific area was undertaken for the first time in 1903 by von Tappeiner and Jesionek, using a combination of topical eosin and white light to treat skin tumors [39]. Subsequent studies in man have validated the clinical use of so-called photodynamic therapy (PDT) in a variety of cancers [40–46], a technique that results in the destruction of tissue after focal irradiation of benign light as a consequence of the in situ excitation of systemically administered photosensitizers and subsequent generation of reactive oxygen species [47]. One of the drawbacks of PDT, however, is that it causes indiscriminate cell damage to both tumor and healthy tissue within the area that is exposed to light [48]. Photoinduced activation of cytotoxic prodrugs endows a higher grade of tumor specificity by combining site-specific activity with discriminatory cytotoxicity towards cells that divide rapidly, thus providing an attractive means of targeting malignancies whilst reducing systemic side effects.

Photoactivated prodrugs are generated by coupling a photolabile protecting motif on functional groups that are directly involved in drug–target interactions, thus disrupting its bioactivity. In the absence of radiation, the caged drug will induce minimal cytotoxicity. Irradiation of benign light with an appropriate wavelength will uncage the active therapeutic agent thus restoring its pharmacological activity. A number of photoresponsive masking groups have been developed for wavelength-specific drug-uncaging applications, and the extent of caging and photoinduced uncaging is currently under investigation in a wide range of chemotherapeutics [49–51].

2.2 A Photoactivated Prodrug of Doxorubicin

Most anthracycline prodrugs are developed by modification of the primary amino group of its daunosamine sugar residue. This group forms key interactions between DNA strands within the drug–DNA adduct [52] and along the minor groove of DNA during intercalation [53], and consequently is essential for the formation of

the topoisomerase II–drug–DNA complex. The efficacy of this strategy has been demonstrated in a number of bioresponsive prodrugs of **DOX** and **EPI** [21, 22, 54, 55], and presents an attractive strategy for the reduction of systemic toxicity (particularly cardiotoxicity) and increased tumor-specific activity.

Ibsen et al. reported a **DOX** prodrug capable of undergoing photoinduced activation in vivo [56, 57]. The caging group disguises the amino group of **DOX** as a carbamate, and incorporates a photocleavable *o*-nitrobenzyloxycarbonyl linker coupled to a biotin molecule through a bis-(ethylene glycol) spacer arm (Fig. 6). The authors reported that cellular uptake of the prodrug was very fast. However, in contrast to **DOX**, it did not localize in the nucleus, proving that DNA intercalation was inhibited. Cell viability studies showed that the chosen masking strategy (photocleavable biotin (PCB) [58]) produced a derivative that displayed over 200-fold cytotoxicity reduction in lung cancer cells relative to **DOX**.

The authors claimed that the PEG-biotin moiety was introduced to enhance the clearance rate of the freely circulating prodrug to minimize prodrug activation in the light-exposed tumor bloodstream and, thereby, to reduce the systemic release of active **DOX** [57]. The masked amino group proved to be resistant to enzymatic cleavage under incubation with human liver microsomes, and efficient release of free **DOX** was verified in cancer cell culture upon photo-irradiation at 350 nm for
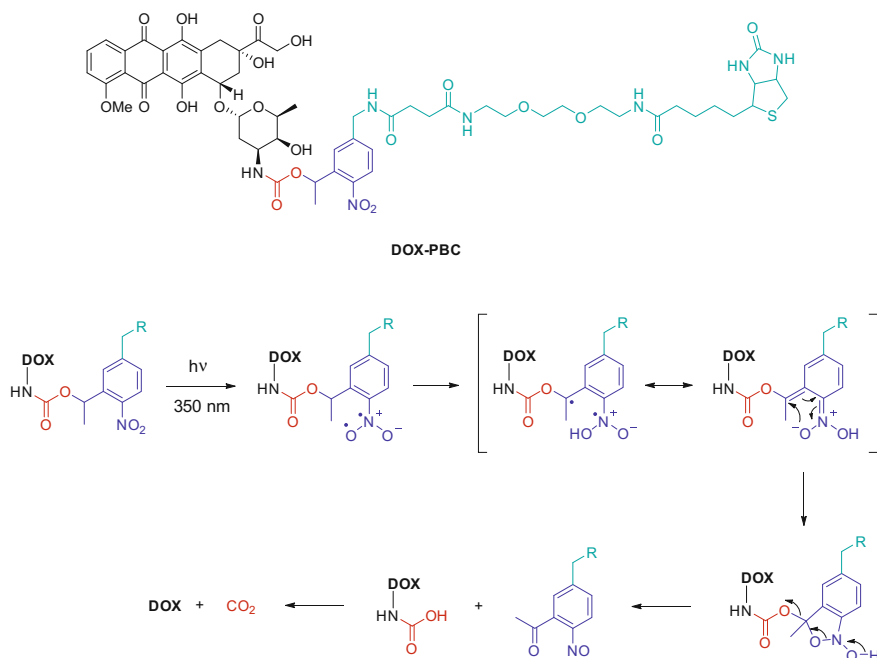

Fig. 6 Prodrug **DOX–PBC** and its proposed photolysis mechanism [56]

60 min [56]. Ex vivo analysis of mouse tumor tissue irradiated with 365 nm light by a custom designed LED fiber optic cable verified a good distribution of scattered light throughout the tumor sample. Subsequent LCMS analysis of an ex vivo tumor sample dosed with the prodrug confirmed **DOX** release within the tissue after 365 nm light irradiation for 30 min [57]. In addition, **DOX–PBC** was administered to mice and studied in the absence and presence of light exposure. While the drug precursor was detected in both tumor samples, light-exposed tumors presented roughly 6-fold higher concentration of **DOX**. The cytotoxic effect of **DOX** was localized within the tumor, with minimal systemic circulation of free **DOX** [57].

2.3 A Photoactivated Prodrug of Floxuridine

FUDR is an anticancer antimetabolite used to treat advanced colorectal, stomach, and kidney cancers, and also employed to treat unresectable gastrointestinal malignancies that have spread to the liver [59]. **FUDR** is also one of the cytotoxic metabolites intracellularly generated after administration of 5-fluorouracil (**5FU**). As most cytotoxic nucleoside analogs, **FUDR** is converted into highly cytotoxic species by phosphorylation of the 5′-OH of the deoxyribose ring. The **FUDR**-5′-monophosphate derivative inhibits thymidylate synthase, repressing DNA synthesis upstream, whereas triphosphorylated **FUDR** incorporates into DNA molecules and further contributes to cytotoxicity by DNA damage [60]. Chemical blockade of this essential position prevents **FUDR** phosphorylation into its active metabolites, thus precluding its cytotoxic activity [61].

Pei and Gong reported a **FUDR** prodrug that possesses the primary alcohol position 5′-OH caged as a carbonate group using a 4,5-dimethoxy-2-nitrobenzyloxycarbonyl [62], and demonstrated prodrug photolysis into **FUDR** and 4,5-dimethoxy-2-nitrosobenzaldehyde upon light irradiation at 350 nm by HPLC. In the absence of the light source, *E. coli* DH5A cell growth was only slightly inhibited by the prodrug, while complete inhibition of bacterial proliferation was observed upon exposure to 350 nm light. Unfortunately, the authors did not provide any data in human cancer cells (Fig. 7).
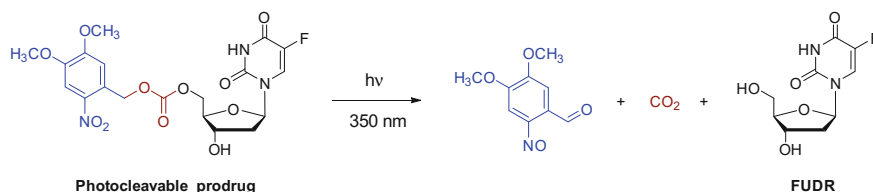

Fig. 7 Photoactivated prodrug of **FUDR** and the products resulting from its photolysis [62]. The cleavage mechanism is analogous to the one shown in Fig. 6

2.4 Photoactivated Prodrugs of 5-Fluorouracil

5FU is an antineoplastic antimetabolite that targets the uracil anabolic pathway. Although **5FU** is not a glycoside, it is metabolized inside cells into cytotoxic nucleotides by glycosidation of its N1 position with either ribose or deoxyribose and subsequent phosphorylation of the 5′-OH [63]. While deoxyribose incorporation is disfavored, **FUDR-5′-**monophosphate is the main contributor to the anticancer properties of **5FU** by inhibition of thymidylate synthase [64]. **5FU** possesses a narrow therapeutic window and >80% is catabolized in the liver into ineffective dihydrofluorouracil [65]. Consequently, much effort has focused on the development of **5FU** precursors with improved ADME-Tox. Because the position N1 of **5FU** needs to become glycosylated for its conversion into cytotoxic metabolites, most prodrug strategies have focused on blocking such a position with a broad range of chemical groups [63, 66–68].

Zhang et al. [68] developed a photocleavable **5FU** prodrug with and without tumor-targeting properties by conjugating an o-nitrobenzyl group at the N1 position of **5FU** (Fig. 8). Tumor-targeting capabilities were introduced by coupling a cyclic Cys–Asn–Gly–Arg–Cys (CNGRC) motif to the masking group. The CNGRC motif is a well-established tumor-homing peptide that specifically recognizes membrane-bound aminopeptidase N (also called CD13) on the surface of the tumor vasculature, a protein that is involved in angiogenesis [69]. Incorporation of such a motif through a photoresponsive linker was aimed to increase prodrug accumulation in the cancerous tissue prior to light irradiation. The **5FU-ONB** prodrug showed stability in aqueous solution after 24 h in the dark and efficient cleavage after exposure to UVA light for 60 min. However, the photoinduced cleavage of the peptide-labeled derivative **5FU-CNGRC** was significantly slower, requiring 5 h to achieve 50% release of **5FU**.
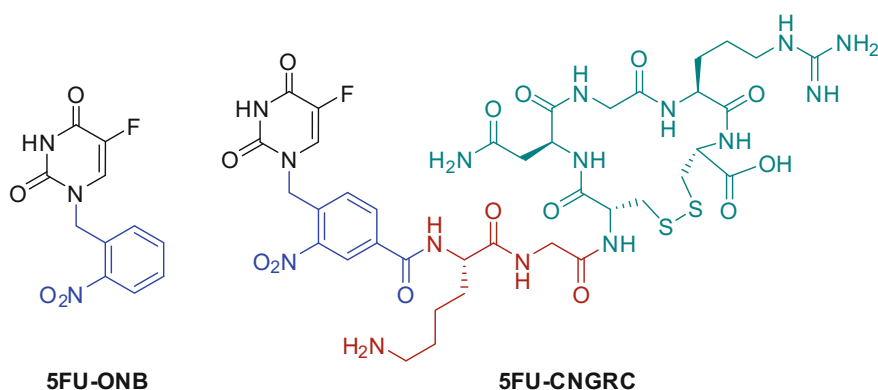

5FU-ONB   **5FU-CNGRC**

Fig. 8 Photo-cleavable prodrugs of **5FU**. The lysine-containing linker (in *red*) of **5FU-CNGRC** was used to increase water solubility [67]. The cleavage mechanism is analogous to the one shown in Fig. 6

Using a similar strategy, Lin et al. [70] developed a "dual bioorthogonal/bioresponsive" prodrug approach for the photoinduced release of the clinically approved **5FU** prodrug Tegafur followed by metabolic release of **5FU**. Tegafur contains a 2-tetrahydro-2-furanyl group at the *N*1 position of **5FU** and is metabolized into **5FU** by the enzyme Cytochrome P450 2A6 (CYP2A6). Tegafur is employed with moderate efficacy in many gastrointestinal cancers [71]. The introduction of the porphyrin group was proposed to accumulate the prodrug in cancer and, potentially, be used as a photosensitizer for photodynamic therapy to potentiate the efficacy of Tegafur [70].

Alkylation of the *N*3 position of Tegafur with a photosensitive *p*-triphenylporphyrin-*o*-nitrobenzyl group resulted in a stable, biologically inactive derivative (Fig. 9). Upon 350 nm photo-irradiation for 15 min, >60% of Tegafur was efficiently released. In vitro studies in MCF7 breast cancer cells showed that the double prodrug mediated very low cytotoxicity in the absence of light, whereas a significant increase in cytotoxic effect was observed after UVA light irradiation, indicating in situ photo-triggered Tegafur release and subsequent metabolic generation of **5FU**.

Agarwal and co-workers [72] developed a photoinduced **5FU**-releasing polymeric system (Fig. 10). **5FU** molecules were covalently caged to a methylcoumarin-functionalized polymer by light-promoted [2 + 2] cycloaddition. Reversion of this process was achieved by irradiation at 266 nm. Such a highly energetic radiation, however, represents the major disadvantage of the strategy, as wavelengths below 320 nm induce severe sub-lethal and lethal damage to
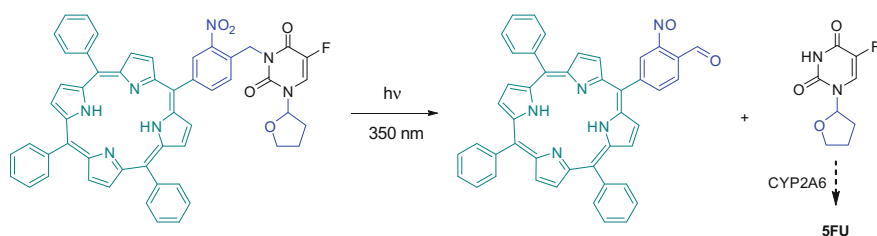

Fig. 9 Photoresponsive porphyrin-Tegafur prodrug and the resulting products from its photolysis [70]
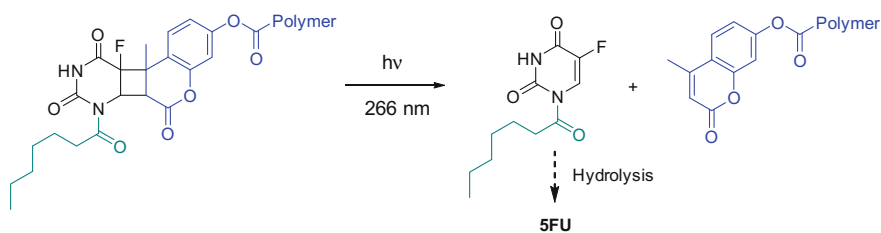

Fig. 10 Non-bioorthogonal UVC light-induced release of hydrolysable **5FU**

DNA/RNA both in normal and tumoral cells. Since the mechanism that triggers **5FU** release also mediates a direct biological effect, this photoactivated masking strategy cannot be considered as bioorthogonal.

3 Prodrugs Activated by Bioorthogonal Click Chemistry

3.1 Introduction to Click Chemistry

The concept of "click chemistry" was formally described by Nobel Laureate Barry Sharpless in 2001 [73]. This synthetic paradigm relies on reactions offering high values of efficiency, selectivity, functional group tolerance and atom economy to build or decorate complex molecules in biocompatible conditions. The quintessential example of this type of chemistry is the Cu(I)-catalyzed azide-alkyne cycloaddition (CuAAC) [74, 75]. Since its introduction, click chemistry has become one of the most widely employed and reliable methods for the covalent assembly of macromolecules, and it have found applications in many disciplines including nanotechnology, chemical biology, drug delivery and medicinal chemistry [76–81]. Shortly after, Carolyn Bertozzi coined the concept of "bioorthogonal chemistry" to define chemical reactions that occur in living environments without interacting or interfering with biological components [82]. The notion of bioorthogonality has organically grown into an intrinsic property of matter and energy and, for many scientists including us, nowadays encompasses any molecule (or functional group), material or electromagnetic radiation that is biologically inactive and metabolically stable. Since most click reactions including CuAAC are not compatible with life, the stringent chemical requirements derived from the bioorthogonal concept demanded the search for novel chemical processes. The first recognized example of a bioorthogonal reaction was the Staudinger ligation [35], where alkyl azides react with triphenylphosphine reagents with an electrophilic trap to form a stable amide bond, thus facilitating the covalent labeling of cell surface components both in vitro [33] and in vivo [35]. However, the slow kinetics of this reaction restricted the study of fast-occurring biological processes. Bertozzi's lab reported in 2004 a bioorthogonal (copper-free) version of the CuAAC click reaction using spring-loaded cyclooctyne reagents, which are capable of undergoing 1,3-dipolar cycloadditions with alkylazides in the absence of catalysts and in living organisms with increased reaction rates [34, 35]. In the last decade, a wide range of bio-compatible chemical ligations has been reported using different bioorthogonal partners such as azide-oxanorbornadiene [83], nitrone-cyclooctyne [84], and the fast inverse Diels–Alder reaction between tetrazines and strained alkenes [85, 86].

Although bioorthogonal click reactions have been mainly used for labeling biomolecules in their native environment, their scope has been recently expanded beyond molecular imaging and diagnostics. In recent years, a number of research labs have studied the potential use of bioorthogonal click chemistry for the selective

targeting and release of chemotherapeutics in tumors to improve its clinical efficacy and reduce side effects. For example, Kim and co-workers used metabolic glyco-engineering to label the surface of cancer cells with unnatural sialic acids labeled with azide groups. Through copper-free click chemistry with the azide-labeled cell surface components, dibenzylcyclootyne-functionalized nanoparticles were specifically taken up by tagged cancer cells both in vitro and in vivo [87]. Artemov and co-workers proposed an alternative strategy where breast cancer cells are pre-targeted with azide-functionalized trastuzumab (anti-HER2 humanized monoclonal antibody). Subsequent treatment with dibenzylcyclooctyne-functionalized albumin conjugated with paclitaxel resulted in the bioorthogonal formation of cross-linked clusters on the cell surface that facilitated complex internalization and intracellular drug release [88]. Brudno et al. demonstrated the use of bioorthogonal click chemistry for the selective in vivo targeting of azide and tetrazine-functionalized hydrogels injected in a disease site with drug surrogates [89]. All these studies have shown the potential of bioorthogonal click chemistry to target-specific tissues, but they did not explore its direct use as a triggering mechanism to release bioactive small molecules. In the following sections, we will focus on different strategies recently reported for the chemical cleavage of small molecule prodrugs of **DOX** by bioorthogonal click reactions.

3.2 Prodrugs Activated by Staudinger-Type Reactions

In 2006, Azoulay et al. reported the use of a modified Staudinger ligation to trigger the uncaging of an inactive triphenylphosphine-functionalized precursor of **DOX** [90] (Fig. 11). To facilitate **DOX** release, the authors introduced an electrophilic trap designed to capture the aza-ylide intermediary and, upon intramolecular migration, generate an unstable *para*-hydroxybenzyloxycarbonyl group that undergoes 1,6-elimination [91]. **DOX** release was monitored by HPLC at 37 °C in aqueous THF, demonstrating no drug liberation in the absence of the alkylazide compared to >90% release after 3 h incubation with alkylazide. Despite the apparent efficacy of the strategy, **DOX** release was not proved either in cell culture or in vivo, thus questioning the actual bioorthogonality of the strategy.

A similar concept was later proposed by Robillard and coworkers [92], who described the first in situ activation of an azido-functionalized **DOX** prodrug by means of a Staudinger reduction in cell culture. In Robillard's design, the **DOX** precursor features a *para*-azidobenzyloxycarbonyl group on the primary amine of daunosamine (Fig. 12). Reaction with a water-soluble triphenylphosphine reagent results in the reduction of the azido group to amino, thus triggering a 1,6-elimination reaction and subsequent formation of CO_2 and **DOX**. The prodrug showed a significant reduction of cytotoxicity compared to **DOX** (IC_{50} = 15 μM *vs.* 0.09 μM, respectively, in human vulvar skin squamous carcinoma cells), while the triphenylphosphine reagent did not induce any reduction of cell viability up to 60 μM concentration. Incubation of **DOX–N₃** with triphenylphosphine reagent in
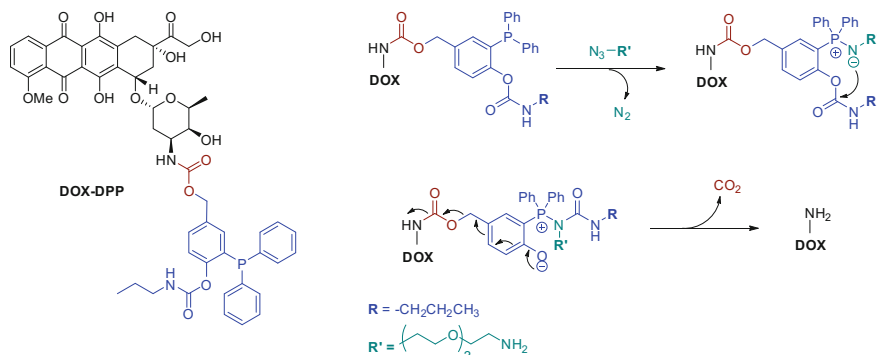
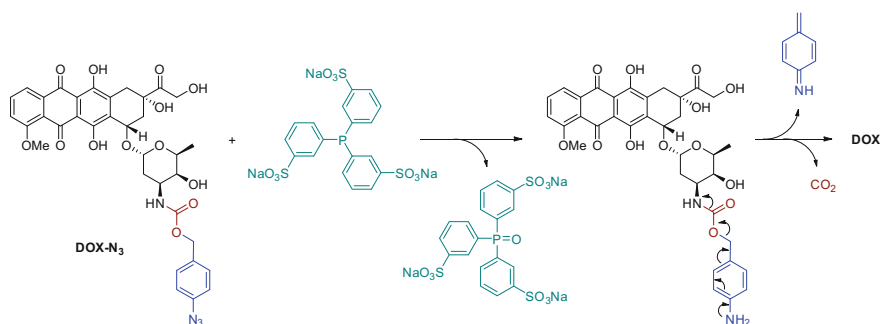**Fig. 11** Prodrug **DOX–DPP** and its activation mechanism upon reaction with an alkylazide

Fig. 12 Prodrug **DOX–N$_3$** and its activation mechanism upon reaction with a water-soluble triphenylphosphine reagent

excess (60 μM) led to complete restoration of **DOX** cytotoxic properties (IC$_{50}$ = 0.074 μM). In an analogous strategy, Winssinger demonstrated that the use of two peptide nucleic acid molecules, each of them functionalized with either a triphenylphosphine moiety or an azidobenzyloxycarbonyl–**DOX**, can mediate the release of **DOX** in the presence of a DNA template of complementary sequence by bringing both bioorthogonal reactive partners to close proximity [93].

3.3 A Prodrug Activated by Tetrazine Ligation

Robillard reported in 2013 a novel drug-uncaging strategy triggered by a bioorthogonal tetrazine ligation, thereby resulting in rapid drug release [94]. Conceptually, it was proposed that such an approach could be exploited as a coupling and decoupling strategy [95] for the functionalization and bioorthogonally controlled release of antibody–drug conjugates (Fig. 13). By covalent linking of
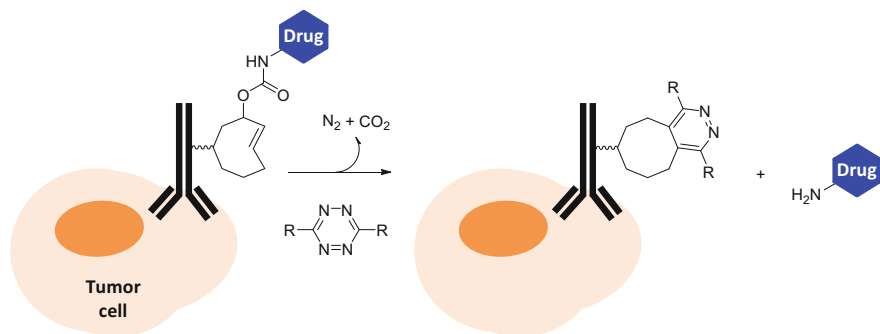
Fig. 13 Bioorthogonal decoupling of tumor-targeting antibody–drug conjugates by inverse Diels–Alder ligation [94]
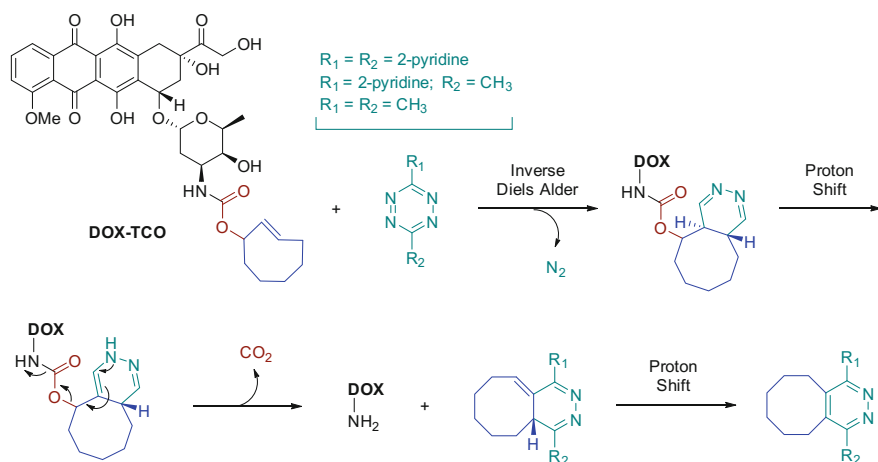

Fig. 14 Prodrug **DOX–TCO** and its mechanism of uncaging triggered by tetrazine ligation [94]

tumor-targeting antibodies with cytotoxic drugs through a carbamate-TCO (*trans*-cyclooctene) linker, the drug would be selectively transported to the cancer site in an inactive form. Upon inverse Diels–Alder reaction with a systemically administered tetrazine, the drug would be released at the periphery of cancer cells [94].

The proof of concept of this novel approach was demonstrated with **DOX** (Fig. 14). The authors showed that ligation of **DOX–TCO** with tetrazine reagents generates an unstable cyclooctane-dihydropyridazine adduct that undergoes spontaneous degradation and release of **DOX**. The yield of this reaction was between 55 and 79% after 4 h incubation depending on the tetrazine employed (Fig. 14). Importantly, cell viability studies with A431 human vulvar skin squamous carcinoma cells showed that the prodrug (**DOX–TCO**) displayed significantly lower cytotoxic effect than **DOX** (EC_{50} = 3.8 μM vs. 0.049 μM, respectively). Effective

restoration of **DOX** cytotoxicity ($EC_{50} = 0.037$ μM) was achieved by combining **DOX–TCO** with 10 μM of tetrazine, indicating that **DOX** was effectively released in cell culture [94].

3.4 A Prodrug Activated by Strain-Promoted Cycloaddition

In 2015, Gamble and co-workers [96] reported a bioorthogonal prodrug activation approach triggered by a 1,3-dipolar cycloaddition reaction between a TCO and a phenylazide (e.g. **DOX–N₃**). The uncaging strategy was based on the formation of an unstable 1,2,3-triazoline that undergoes ring opening and spontaneous rearrangement with the release of molecular nitrogen and generation of an aldimine derivative (Fig. 15). Subsequent imine hydrolysis in neutral-to-slightly acidic conditions generates a *p*-aminobenzyloxycarbonyl derivative that, upon 1,6-elimination and release of CO_2, will give rise to **DOX**. The authors proposed that this cleavage would be enhanced by the acidity of the tumor microenvironment (pH 6.0–7.4). NMR analysis confirmed the reaction favored the synthesis of the desired labile aldimine ($\geq 90\%$ conversion) and subsequent degradation, whereas HPLC studies showed the cycloaddition rate was 0.020 M^{-1} s^{-1}. As expected [92], **DOX–N₃** showed reduced cytotoxic effect against murine melanoma cells relative
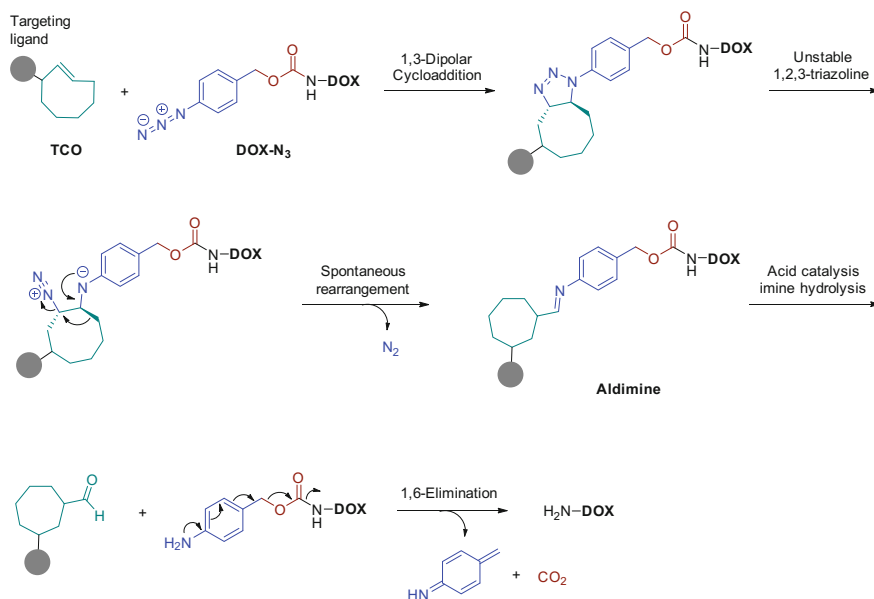

Fig. 15 Mechanism of **DOX–N₃** uncaging promoted by a bioorthogonal 1,3-dipolar cycloaddition [96]. The targeting ligand could be an antibody, peptide or small molecule

to **DOX** (70-fold difference). Incubation of **DOX–N$_3$** with 100 μM *trans*-cyclooctenol induced a 50-fold increment of its cytotoxic activity, thus demonstrating the in situ bioorthogonal release of **DOX** [96].

4 Prodrugs Activated by Bioorthogonal Transition Metals

4.1 Introduction to BOOM Chemistry

Transition metals are defined by the IUPAC as "elements whose atoms have an incomplete *d* sub-shell, or which can give rise to cations with an incomplete *d* sub-shell" [3]. Partial filling of the *d* orbitals confers their characteristic properties; e.g. ability to form five or more chemical bonds, multiple accessible oxidation states and a tendency to accept electron pairs to form coordination complexes [97]. The use of transition metals to catalyze organic reactions constitutes one of the most powerful methods to address selectivity (chemoselectivity, regioselectivity, diastereoselectivity, and enantioselectivity) and atom economy, key features to perform highly efficient synthetic processes. Hydrogenation, isomerization, asymmetric synthesis, oxidation, and C–C coupling reactions are just a few examples of the great variety of chemical reactions that can be carried out utilizing transition metals [98].

The synthetic versatility offered by transition metals also provides very attractive possibilities in chemical biology. Several transition metals are used by living systems to form complexes with proteins—so-called metalloproteins—and catalyze essential biochemical processes [99]. Notably, in recent years several research groups have tried to exploit the catalytic properties of non-biological transition metals to mediate BOOM reactions [100]. To achieve such kind of chemistry in living systems is a formidable challenge because of problems with the biocompatibility, stability, and reactivity of the metals: the catalyst has to be innocuous and stable in a concentrated solution containing a myriad of biological components and, at the same time, has to be highly reactive towards its designated substrate which is typically present at very low concentration. Despite these challenges, researchers have partially addressed these issues and successfully used abiotic metals for different applications in, on and outside cells, such as the synthesis of small molecules (e.g. fluorophores), the functionalization and uncaging of enzymes and the in situ activation of toxigenic prodrugs [101, 102].

A decade ago, Streu and Meggers described the first application of a ruthenium-based catalyst to mediate an uncaging reaction (allylcarbamate cleavage of protected amines) under physiologically relevant conditions and inside living cells [103]. Using an Alloc-protected fluorophore precursor, they demonstrated the catalytic properties of the organometallic half-sandwich complex [Cp*Ru(COD)Cl] (**Ru1**) within HeLa cells. Although **Ru1** showed no toxic effect during the life of the experiment (minutes), the short duration of the assay and the need for toxic

thiophenol as an additive severely limited the general applicability of this catalyst [103]. Nevertheless, this work was one of the earliest uses of non-biological metals to mediate artificial chemical reactions in living cells. A photoactivatable version of this catalytic system was reported in 2012, namely [Cp*Ru(η[6]-pyrene)]PF$_6$ (**Ru2**), which slightly improved the efficiency of its predecessor and enabled spatiotemporal control [104].

In 2011 Yusop et al. expanded the manifold applications of the chemistry of Palladium (Pd) into the living cell [105]. This was achieved through the development of a bio-friendly heterogeneous catalyst of subcellular size consisting of small polystyrene microspheres (0.5 μm in diameter) internally functionalized with 5 nm Pd nanoparticles, which were generated and physically captured within the microspheres. These devices were capable of entering HeLa cells (>75% uptake rate after 24 h) and staying harmlessly within the cytosol, as verified by the low levels of necrosis (<4%) and excellent cell viability (>91%) after 48 h incubation [105]. The catalytic properties of the Pd devices were verified by mediating intracellular Alloc deprotection of the pro-fluorophore developed by Meggers [102], demonstrating for the first time Pd-mediated catalysis inside a cell [105]. The versatility of this dual catalytic cell delivery system to perform Pd chemistries was shown with the intracellular generation of the fluorescent dye anthofluorescein [106] via Suzuki–Miyaura cross-coupling [105]. A year later, Unciti-Broceta et al. proposed the concept of BOOM chemistry to describe the use of non-biotic transition metals to mediate bioorthogonal chemistry in living systems and provided the basis for its general application in chemical biology and experimental pharmacology [99].

4.2 Prodrugs Activated by Palladium

In 2014 Weiss et al. reported a novel bioorthogonal strategy for the local generation of cytotoxic drugs from a biologically inert precursor by heterogeneous Pd catalysis [107]. This first-in-class approach, so-called Palladium-Activated Prodrug Therapy, consisted of using Pd-functionalized PEG-polystyrene macrospheres of 150 microns in diameter (so-called **Pd-resins**) as an extracellular heterogeneous catalytic system to drive the spatially controlled conversion of a biochemically stable prodrug into its active form. To achieve full control over the area where the chemotherapeutic drugs are generated, the authors stated that such prodrugs had to possess three essential characteristics: lack of pharmacological activity, minimal susceptibility to enzymatic cleavage and high sensitivity to Pd-mediated uncaging. **5FU**, a potent clinically approved antimetabolite drug whose main medical limitation is its narrow therapeutic index, was the first drug used to implement such a novel strategy. By N-propargylation of the position 1 of **5FU**, the authors created a prodrug (**Pro-5FU**) possessing all three properties above described [107]. Alkylation of the N1 position blocked the enzymatic glycosylation of the drug, thereby impeding the generation of its cytotoxic metabolites. Because of the high

metabolic stability of the propargyl group, the prodrug exhibited >500-fold reduction of cytotoxicity relative to the parent drug in colorectal and pancreatic cancer cells. Importantly, **Pro-5FU** showed high sensitivity to Palladium catalysis in biocompatible conditions (pH = 7.4, 37 °C, isotonicity), resulting in complete prodrug-into-drug conversion in less than 24 h (Figs. 16 and 17). The remarkable efficiency and biocompatibility of such a deprotection reaction, which represented the discovery of a novel bioorthogonal process, was due to the particular properties
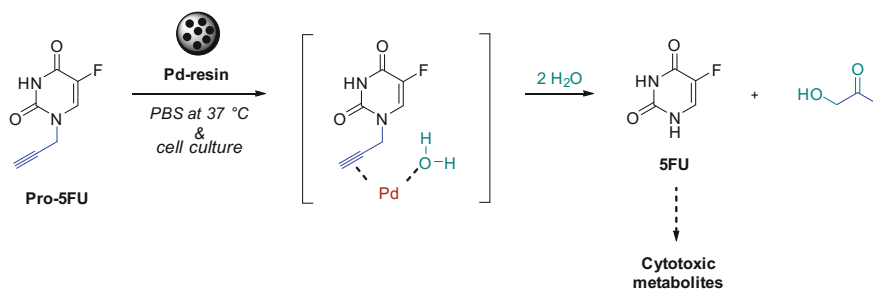

Fig. 16 Pd-activated prodrug of **5FU** and the products resulting from its bioorthogonal release in the presence of **Pd-resins**, including harmless 1-hydroxyacetone [107]
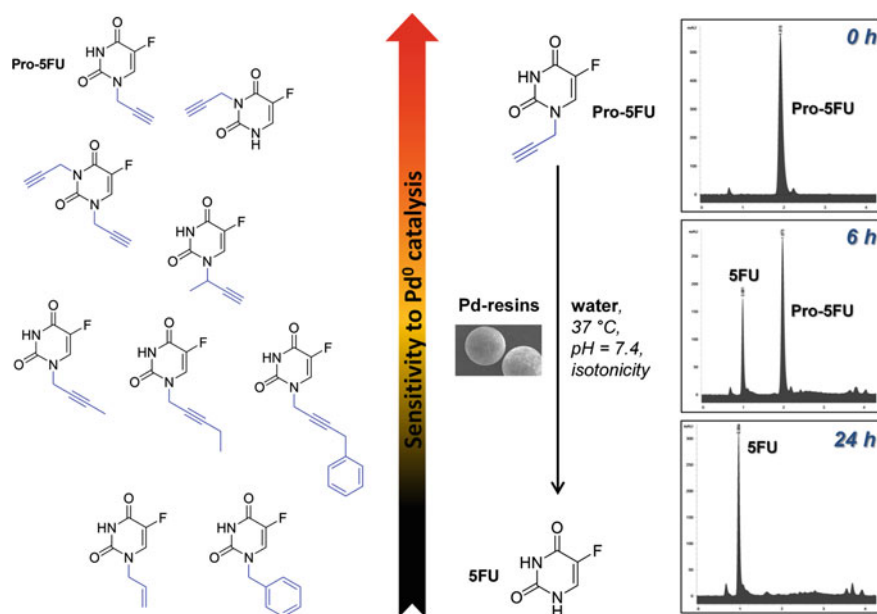

Fig. 17 *N*-Alkynyl precursors of **5FU** and their relative reactivity to Pd-catalyzed cleavage [38, 107]. HPLC traces of the reaction between **Pro-5FU** and **Pd-resins** in biocompatible conditions [107]

of the **5FU** heterocycle which undergoes tautomeric shifts in aqueous media. As a result, the *N*1 position possesses a relatively low pK_a that promotes its leaving group properties [107]. Functionalization of the *N*1 position with an allyl or benzyl group, however, did not generate Pd-cleavable derivatives. Importantly, whereas separately neither **Pd–resins** nor **Pro–5FU** exhibited biological activity, its combination in culture led to a toxigenic effect equivalent to **5FU** cytotoxicity. Through reverse phase protein microarray analysis and western blotting, the authors demonstrated that the antiproliferative effect mediated by the **Pd–resins/Pro–5FU** combination was indeed a consequence of the in situ generation of **5FU** in cell culture by BOOM chemistry [107]. In addition, the authors showed that the **Pd–resins** were both biocompatible and catalytically active in zebrafish [107].

To provide additional insights on the properties and scope of this novel bioorthogonal deprotection method, Weiss et al. reported a follow-up study on the development of a set of *N*-alkynyl derivatives of **5FU** and investigated their sensitivity to Pd catalysis under biocompatible conditions [38]. HPLC analysis demonstrated that the larger the size of the *N*-alkynyl group, the lower the sensitivity of the prodrug to Pd catalysis (Fig. 17). Additionally, a comparative study was carried out to determine the difference in bioorthogonality and Pd sensitivity of the *N*1- versus the *N*3-propargyl derivatives of **5FU**. While both prodrugs were found equally innocuous to cells, **Pro-5FU** cleavage occurred at higher rate, thus indicating that not only the pK_a but also steric and conformational effects are important factors modulating the reaction kinetics. The study of the influence of pH changes on the *N*-dealkylation demonstrated that the reaction is compatible with the range of pH expected to be found in both normal and cancerous tissues [38].

To expand the applicability of this novel focal therapeutic strategy to drugs with higher structural complexity, Weiss et al. subsequently investigated the compatibility of BOOM chemistry with the uridine analog **FUDR**, which is significantly more potent than **5FU** [37]. **FUDR** is a cytotoxic nucleoside which presents a *NH* group at the position 3 of the uracil ring that plays a fundamental role in the substrate recognition by both anabolic and catabolic enzymes. Weiss et al. hypothesized that the propargylation of the *N*3 position of **FUDR** would not only block the cytotoxic activity of the drug but also protect it from systemic metabolization before reaching the target (Fig. 18). In addition, since such a position undergoes lactam/lactim tautomery shifts, the authors expected the resulting prodrug would be sensitive to Pd catalysis [37]. **Pro-FUDR** displayed a vast reduction of cytotoxicity relative to **FUDR** in colorectal and pancreatic cancer cells: >6000-fold reduction, confirming the essential role of such a hydrogen donor in the drug's biological properties and the high bioorthogonality of the *N*-propargyl group. *N*-Dealkylation of **Pro-FUDR** in biocompatible conditions and cell culture demonstrated the bioorthogonal generation of **FUDR** in the presence and absence of oxygen and in normal and slightly acidic conditions, thus indicating that the Pd-mediated *N*-dealkylation of **Pro-FUDR** is compatible with the environment found in both early (normoxia, pH = 7–7.4) and late-stage (hypoxia, pH = 6–7) cancers [37].
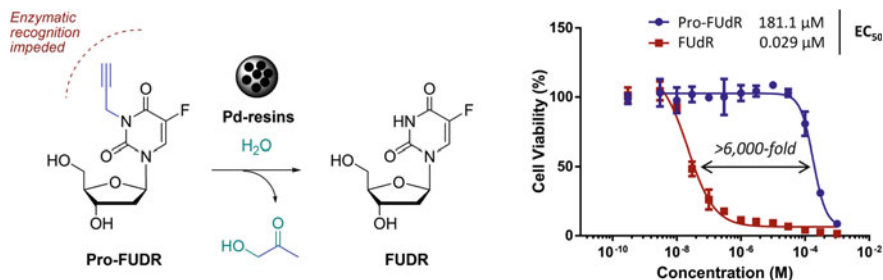
Fig. 18 *Left* **Pro-FUDR** and its Pd-mediated cleavage. *Right* Dose response study of HCT116 colorectal cancer cell viability for **Pro-FUDR** and **FUDR** [37]

Alongside the discovery and application of the *N*1-depropargylation of **Pro-5FU** as a bioorthogonal prodrug activation strategy, Weiss et al. reported in May 2014 an investigation on Pd-activated prodrugs of the cytotoxic nucleoside gemcitabine (**GEM**) [108]. This drug, which lacks of endocyclic NH groups with lactam/lactim tautomery, was masked using standard alkyloxycarbonyl chemistry. Two types of prodrugs were generated by carbamoylation of the NH_2 group at the position 4 of the cytosine ring or by carbonate formation at the 5′-OH group of the difluorodeoxyribose ring (Fig. 19). Since **GEM** is used as first-line therapy in the treatment of cancer of pancreas, the authors evaluated the bioorthogonality of the prodrugs with pancreatic cancer cells. This study confirmed the more resilient nature of the carbamate bond relative to the carbonate one. The carbonate-masked prodrugs (**GEM–CBO**) elicited similar cytotoxic effect than the unmodified drug, suggesting rapid bioactivation inside cells, while the carbamate-protected prodrugs (**GEM–CBA**) displayed a significant decrease of cytotoxic activity (>23-fold relative to that of the parent drug) (Fig. 19). Among the prodrugs investigated, the **GEM–CBA** prodrug masked with a propargyloxycarbonyl (Poc) group showed superior sensitivity to Pd-triggered drug release in biocompatible media and cell culture. This is of relevance because, whereas Pd-mediated Alloc deprotection is a well-established method used in organic chemistry, the cleavage of Poc groups with Pd had not been previously reported prior to this study and an independent study from Chen and coworkers published two months earlier [109]. Immunofluorescence studies using phosphorylated γ-H2AX as a marker of DNA damage confirmed that the cytotoxic activity elicited by the **GEM-CBA/Pd-resins** combination was indeed due to the release of **GEM**.

To expand the scope of intracellular BOOM chemistry, in 2015 Rotello and co-workers reported the development of artificial nanozymes through the encapsulation of hydrophobic metal catalysts (Pd and Ru complexes) into self-assembled lipophilic monolayers immobilized on gold nanoparticles [110]. Using a guest-host recognition strategy, the authors protected the surface of the devices with macrocyclic cucurbiturils, thus blocking the access to the catalyst. The "inactive" nanozymes were able to enter cells by endocytosis and, in the presence of guest molecule 1-adamantylamine, become catalytically active and
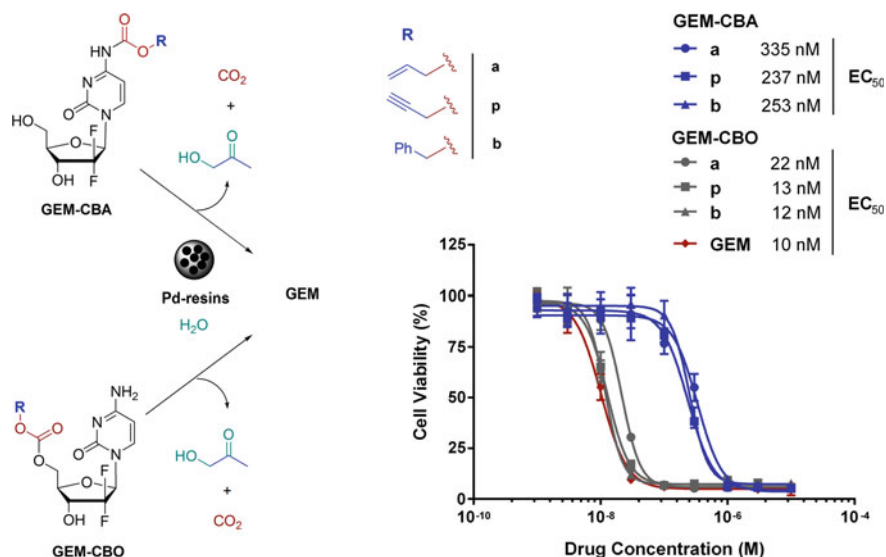
Fig. 19 Carbamate and carbonate-based prodrugs of **GEM** (*left*) and its Pd-mediated release. Dose response study of pancreatic cancer cell viability for carbamate and carbonate prodrugs relative to **GEM** (*right*) [108]

mediate the intracellular release of **5FU** from **Pro-5FU**, the bioorthogonal prodrug developed by Weiss et al. [107]. Such a multifunctional device, which has been considered an early prototype of a medical nanomachine [111], could be "up-graded" to incorporate targeting capabilities and operate as a tumor-specific drug nanofactory.

4.3 A Prodrug Activated by Ruthenium

In Aug 2014, Meggers and co-workers reported the study and optimization of a set of ruthenium half-sandwich complexes to mediate Tsuji–Trost reaction in bio-compatible conditions [112]. Among the initial set of complexes tested, only [CpRu(QA)(allyl)]PF$_6$ (**Ru3**) catalyzed the deallylation reaction of *N*-(allyloxycarbonyl)-aminocoumarin with high efficiency under biologically relevant conditions. Further modifications of the bidentate ligand by introducing a π-donating methoxy ([CpRu(QA-OMe)(allyl)]PF$_6$, **Ru4**) or dimethylamino ([CpRu(QA-NMe$_2$)(allyl)] PF$_6$, **Ru5**) group into the quinoline moiety led to a significant boost in catalytic activity, reaching high turnover numbers (**Ru3** = 90 < **Ru4** = 150 < **Ru5** = 270). Incubation of Alloc-protected rhodamine 110 with HeLa cells followed by addition of **Ru5** generated up to 130-fold increase of fluorescence intensity within the cytoplasm. Inspired by previous work from Weiss et al. [107], the authors studied
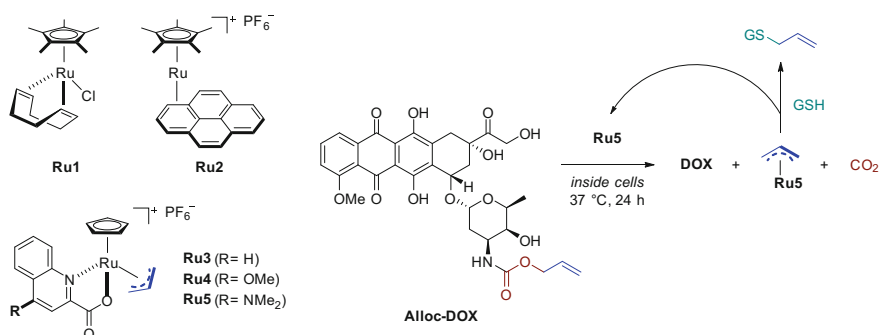
Fig. 20 Ru-activated prodrug of **DOX** and the products resulting from its bioorthogonal release in the presence of **Ru5** [112]

the use of **Ru5** as a bioorthogonal tool for the in situ generation of chemothera-peutic agents. The anticancer glycoside **DOX** was modified by masking the primary amino group of its daunosamine sugar residue with a bioorthogonal allyl carbamate protecting group (Fig. 20). N-(allyloxycarbonyl)doxorubicin (**Alloc-DOX**) showed no negative effect on the survival rate of HeLa cells at several concentrations (20, 50, and 100 µM) that were found highly cytotoxic for the unmodified drug [112]. To assess the bioorthogonal uncaging of the prodrug inside cells, HeLa cells were incubated with **Alloc-DOX** (100 µM) for 3 h, washed with PBS buffer, and then incubated with catalyst **Ru5** (20 µM) for 24 h. **Alloc-DOX/Ru5** combination led to a reduction in cell viability equivalent to free **DOX**. Since neither catalyst nor the prodrug displayed any toxicity at these concentrations, Meggers and co-workers concluded that the observed loss in cell viability was due to the bioorthogonal activation of the prodrug within the cell cytoplasm catalyzed by ruthenium-mediated chemistry [112]. Given the possibility of introducing tumor-targeting ligands into such ruthenium complexes, this strategy could be potentially used for increasing the selectivity of chemotherapeutics towards cancer cells.

5 Conclusions and Perspectives

The exquisite spatiotemporal control achieved through the use of bioorthogonal uncaging methods has drawn the attention of many research groups working in the interface of chemistry and biomedicine. In the search to optimize the selective targeting of cancers with therapeutic agents, various bioorthogonal strategies are under examination for the controlled release of cytotoxic drugs in specific anatomical areas. Contrary to the classical activation mechanisms undergone by bioresponsive prodrugs, the use of bioindependent means to activate drug precur-sors provides a way to bypass the need for inherent metabolic processes, which are

often inconsistent from animals to humans (thus deterring preclinical development) and even among patients due to genetic differences. Although the exploitation of bioorthogonal strategies in medicine is still in their infancy (only photodynamic therapy can be considered a clinically validated treatment based on a bioorthogonal activation method), our ever-growing understanding and control of chemical phenomena in complex systems and the permanent evolution of biotechnology and medical engineering are paving the way for the next therapeutic revolution. From advanced LED devices to next-generation photonic technologies, light sources will offer a highly accurate way to activate photoresponsive prodrugs with single cell precision. Progress in the performance, manufacture (e.g. 3D printing) and surgical insertion of biocompatible materials, implant technologies and heterogeneous catalysts will provide optimal means to mediate long-term transition metal-based activation of small molecule prodrugs in any tissue or organ [107, 113]. Advances in the development and functionalization of tumor-targeting antibodies and ligands will facilitate the specific release of therapeutics in nominated cell types via bioorthogonal click chemistry or BOOM uncaging reactions. Finally, even if nowadays it looks unachievable, the upcoming creation of purely artificial nanoscale devices with bioorthogonal functionalities (**nano**engineered **bi**oorthogonal **t**echnologies or **nanobots** [111]) will enable full control of drug release in selected microenvironments targeted through remotely controlled mechanisms.

As summarized in this chapter, anticancer glycosides have and will continue to have a prominent role in the development of drug precursors that are activated by bioorthogonal means. New opportunities with classical and non-classical bioactive glycosides (e.g. etoposide) will be likely exploited in the near future alongside alternative bioorthogonal means. Feeding from seemly unrelated disciplines such as organometallic chemistry, medicinal chemistry, implant technologies, photonics, engineering, nanotechnology, robotics, wireless navigation systems, etc., advanced bioorthogonal techniques will provide diagnostic and therapeutic tools with unprecedented selectivity. This chapter is just a prelude of what is about to come.

References

1. Stipanuk MH, Caudill MA (2012) Biochemical, physiological, and molecular aspects of human nutrition, 3rd edn. Saunders/Elsevier, Philadelphia
2. McNaught AD (1996) Nomenclature of carbohydrates (IUPAC Recommendations 1996). Pure Appl Chem 68:1919–2008
3. McNaught AD, Wilkinson A (eds) (1997) Compendium of chemical terminology the "gold book", 2nd edn. International Union of Pure and Applied Chemistry. Blackwell Scientific Publications, Oxford

4. Kren V, Martínková L (2001) Glycosides in medicine: "The role of glycosidic residue in biological activity". Curr Med Chem 8:1303–1328

5. Weiss RB (1992) The anthracyclines: will we ever find a better doxorubicin? Semin Oncol 19:670–686

6. Minotti G, Menna P, Salvatorelli E, Cairo G, Gianni L (2004) Anthracyclines: molecular advances and pharmacologic developments in antitumor activity and cardiotoxicity. Pharmacol Rev 56:185–229

7. Gewirtz DA (1999) A critical evaluation of the mechanisms of action proposed for the antitumor effects of the anthracycline antibiotics adriamycin and daunorubicin. Biochem Pharmacol 57:727–741

8. Hande KR (1998) Etoposide: four decades of development of a topoisomerase II inhibitor. Eur J Cancer 34:1514–1521

9. van Maanen JM, Retèl J, de Vries J, Pinedo HM (1988) Mechanism of action of antitumor drug etoposide: a review. J Natl Cancer Inst 80:1526–1533

10. Jordheim LP, Durantel D, Zoulim F, Dumontet C (2013) Advances in the development of nucleoside and nucleotide analogues for cancer and viral diseases. Nat Rev Drug Discov 12:447–464

11. Shelton J, Lu X, Hollenbaugh JA, Cho JH, Amblard F, Schinazi RF (2016) Metabolism, biochemical actions, and chemical synthesis of anticancer nucleosides, nucleotides, and base analogs. Chem Rev 116:14379–14455

12. Wilson PM, Danenberg PV, Johnston PG, Lenz HJ, Ladner RD (2014) Standing the test of time: targeting thymidylate biosynthesis in cancer therapy. Nat Rev Clin Oncol 11:282–298

13. Parker WB (2009) Enzymology of purine and pyrimidine antimetabolites used in the treatment of cancer. Chem Rev 109:2880–2893

14. Chabner BA, Roberts TG Jr (2005) Chemotherapy and the war on cancer. Nat Rev Cancer 5:65–72

15. DeVita VT Jr, Chu E (2008) A history of cancer chemotherapy. Cancer Res 68:8643–8653

16. Rautio J, Kumpulainen H, Heimbach T, Oliyai R, Oh D, Järvinen T, Savolainen J (2008) Prodrugs: design and clinical applications. Nat Rev Drug Discov 7:255–270

17. Huttunen KM, Raunio H, Rautio J (2008) Prodrugs-from serendipity to rational design. Pharmacol Rev 63:750–771

18. Kratz F, Müller IA, Ryppa C, Warnecke A (2008) Prodrug strategies in anticancer chemotherapy. ChemMedChem 3:20–53

19. Rooseboom M, Commandeur JN, Vermeulen NP (2004) Enzyme-catalyzed activation of anticancerprodrugs. Pharmacol Rev 56:53–102

20. Yang Y, Aloysius H, Inoyama D, Chen Y, Hu L (2011) Enzyme-mediated hydrolytic activation of prodrugs. Acta Pharmaceutica Sinica B 11:143–159

21. Tranoy-Opalinski I, Legigan T, Barat R, Clarhaut J, Thomas M, Renoux B, Papot S (2014) β-Glucuronidase-responsive prodrugs for selective cancer chemotherapy: an update. Eur J Med Chem 74:302–313

22. Haisma HJ, Boven E, van Muijen M, de Jong J, van der Vijgh WJ, Pinedo HM (1992) A monoclonal antibody-beta-glucuronidase conjugate as activator of the prodrug epirubicin-glucuronide for specific treatment of cancer. Br J Cancer 66:474–478

23. Mürdter TE, Sperker B, Kivistö KT, McClellan M, Fritz P, Friedel G, Linder A, Bosslet K, Toomes H, Dierkesmann R, Kroemer HK (1997) Enhanced uptake of Doxorubicin into bronchial carcinoma: β-glucuronidase mediates release of Doxorubicin from a glucuronide prodrug (HMR 1826) at the tumor site. Cancer Res 57:2440–2445

24. Houba PH, Boven E, van der Meulen-Muileman IH, Leenders RG, Scheeren JW, Pinedo HM, Haisma HJ (2001) A novel doxorubicin-glucuronide prodrug DOX-GA3 for tumour-selective chemotherapy: distribution and efficacy in experimental human ovarian cancer. Br J Cancer 84:550–557

25. Houba PH, Leenders RG, Boven E, Scheeren JW, Pinedo HM, Haisma HJ (1996) Characterization of novel anthracycline prodrugs activated by human beta-glucuronidase for use in antibody-directed enzyme prodrug therapy. Biochem Pharmacol 52:455–463

26. Houba PH, Boven E, Erkelens CA, Leenders RG, Scheeren JW, Pinedo HM, Haisma HJ (1998) The efficacy of the anthracycline prodrug daunorubicin-GA3 in human ovarian cancer xenografts. Br J Cancer 78:1600–1606

27. Bagshawe KD (1987) Antibody directed enzymes revive anti-cancer prodrugs concept. Br J Cancer 56:531–532

28. Brown JM, Wilson WR (2004) Exploiting tumour hypoxia in cancer treatment. Nat Rev Cancer 4:437–447

29. McKeown SR, Cowen RL, Williams KJ (2007) Bioreductive drugs: from concept to clinic. Clin Oncol 19:427–442

30. Chen Y, Hu L (2009) Design of anticancer prodrugs for reductive activation. Med Res Rev 29:29–64

31. Hu L, Liu B, Hacking DR (2000) 5′-[2-(2-Nitrophenyl)-2-methylpropionyl]-2′-deoxy-5-fluorouridine as a potential bioreductively activated prodrug of FUDR: synthesis, stability and reductive activation. Bioorg Med Chem Lett 10:797–800

32. Liu B, Hu L (2003) 5′-(2-Nitrophenylalkanoyl)-2′-deoxy-5-fluorouridines as potential prodrugs of FUDR for reductive activation. Bioorg Med Chem 11:3889–3899

33. Saxon E, Bertozzi CR (2000) Cell surface engineering by a modified Staudinger reaction. Science 287:2007–2010

34. Agard NJ, Prescher J, Bertozzi CR (2004) A strain-promoted [3 + 2] azide-alkyne cycloaddition for covalent modification of biomolecules in living systems. J Am Chem Soc 126:15046–15047

35. Sletten EM, Bertozzi CR (2011) From mechanism to mouse: a tale of two bioorthogonal reactions. Acc Chem Res 44:666–676

36. Bertozzi CR (2011) A decade of bioorthogonal chemistry. Acc Chem Res 44:651–653

37. Weiss JT, Carragher NO, Unciti-Broceta A (2015) Palladium-mediated dealkylation of N-propargyl-floxuridine as a bioorthogonal oxygen-independent prodrug strategy. Sci Rep 5:9329

38. Weiss JT, Fraser C, Rubio-Ruiz B, Myers SH, Crispin R, Dawson JC, Brunton VG, Patton EE, Carragher NO, Unciti-Broceta A (2015) N-alkynyl derivatives of 5-fluorouracil: susceptibility to palladium-mediated dealkylation and toxigenicity in cancer cell culture. Front Chem 2:56

39. Von Tappeiner H, Jesionek A (1903) Therapeutische Versuche mit fluoreszierenden Stoffen. Münchner Med Wochenschr 47:2042–2044

40. Diamond I, Granelli SG, McDonagh AF, Nielsen S, Wilson CB, Jaenicke R (1972) Photodynamic therapy of malignant tumours. Lancet 2:1175–1177

41. Doiron DR, Gomer CJ (1984) Porphyrin localization and treatment of tumors. AR Liss Inc, New York

42. Ward BG, Forbes IJ, Cowled PA, McEvoy MM, Cox LW (1982) The treatment of vaginal recurrences of gynecological malignancy with phototherapy following hematoporphyrin derivative pre-treatment. Am J Obstet Gynecol 142:356–357

43. Gomer CJ, Doiron DR, Jester JV, Szirth BC, Murphree AL (1983) Hematoporphyrin derivative photoradiation therapy for the treatment of intraocular tumors: examination of acute normal ocular toxicity. Cancer Res 43:721–727

44. Hill JS, Kaye AH, Sawyer WH, Morstyn G, Megison PD, Stylli SS (1990) Selective uptake of hematoporphyrin derivative into human cerebral glioma. Neurosurgery 26:248–254

45. Wenig BL, Kurtzman DM, Grossweiner LI, Mafee MF, Harris DM, Lobraico RV, Prycz RA, Appelbaum EL (1990) Photodynamic therapy in the treatment of squamous cell carcinoma of the head and neck. Arch Otolaryngol Head Neck Surg 116:1267–1270

46. Barr H, Krasner N, Boulos PB, Chatlani P, Bown SG (1990) Photodynamic therapy for colorectal cancer: a quantitative pilot study. Br J Surg 77:93–96

47. Dougherty TJ, Gomer CJ, Henderson BW, Jori G, Kessel D, Korbelik M, Moan J, Peng Q (1998) Photodynamic therapy. J Natl Cancer Inst 90:889–905

48. Hendersonand B, Dougherty T (1992) How does photodynamic therapy work? Photochem Photobiol 55:145–157

49. Tietze LF, Müller M, Duefert SC, Schmuck K, Schuberth I (2013) Photoactivatable prodrugs of highly potent duocarmycin analogues for a selective cancer therapy. Chem Eur J 19:1726–1731

50. Horbert R, Pinchuk B, Davies P, Alessi D, Peifer C (2015) Photoactivatable prodrugs of anti-melanoma agent vemurafenib. ACS Chem Biol 10:2099–2107

51. Hossion AML, Bio M, Nkepang G, Awuah SG, You Y (2013) Visible light controlled release of anticancer drug through double activation of prodrug. ACS Med Chem Lett 4:124–127

52. Forrest RA, Swift LP, Rephaeli A, Nudelman A, Kimura K, Phillips DR, Cutts SM (2012) Activation of DNA damage response pathways as a consequence of anthracycline-DNA adduct formation. Biochem Pharmacol 83:1602–1612

53. Agudelo D, Bourassa P, Bérubé G (2014) Intercalation of antitumor drug doxorubicin and its analogue by DNA duplex: structural features and biological implications. Int J Biol Macromol 66:144–150

54. Pawar SK, Badhwar AJ, Kharas F, Khandare JJ, Vavia PR (2012) Design, synthesis and evaluation of N-acetyl glucosamine (NAG)-PEG-doxorubicin targeted conjugates for anticancer delivery. Int J Pharm 436:183–193

55. Mita MM, Natale RB, Wolin EM, Laabs B, Dinh H, Wieland S, Levitt DJ, Mita AC (2015) Pharmacokinetic study of aldoxorubicin in patients with solid tumors. Invest New Drugs 33:341–348

56. Ibsen S, Zahavy E, Wrasdilo W, Berns M, Chan M, Esener S (2010) A novel doxorubicin prodrug with controllable photolysis activation for cancer chemotherapy. Pharm Res 27:1848–1860

57. Ibsen S, Zahavy E, Wrasidlo W, Hayashi T, Norton J, Su Y, Adams S, Esener S (2013) Localized *in vivo* activation of a photoactivatable doxorubicin prodrug in deep tumor tissue. Photochem Photobiol 89:698–708

58. Olejnik J, Sonar S, Krzymanska-Olejnik E, Rothschild KJ (1995) Photocleavable biotin derivatives: a versatile approach for the isolation of biomolecules. Proc Natl Acad Sci USA 92:7590–7594

59. Power DG, Kemeny NE (2009) The role of floxuridine in metastatic liver disease. Mol Cancer Therapeutics 8:1015–1025

60. Galmarini CM, Mackey JR, Dumontet C (2002) Nucleoside analogues and nucleobases in cancer treatment. Lancet Oncol 3:415–424

61. Tobias SC, Borch RF (2001) Synthesis and biological studies of novel nucleoside phosphoramidate prodrugs. J Med Chem 44:4475–4480

62. Wei Y, Yan Y, Pei D, Gong B (1998) A photoactivated prodrug. Bioorganic Med Chem Lett 8:2419–2422

63. Longley DB, Harkin DP, Johnston PG (2003) 5-fluorouracil: mechanisms of action and clinical strategies. Nat Rev Cancer 3:330–338

64. Schwartz EL, Baptiste N, Wadler S, Makower D (1995) Thymidine phosphorylase mediates the sensitivity of human colon-carcinoma cells to 5-fluorouracil. J Biol Chem 270:19073–19077

65. Dobritzsch D, Ricagno S, Schneider G, Schnackerz KD, Lindqvist Y (2002) Crystal structure of the productive ternary complex of dihydropyrimidine dehydrogenase with NADPH and 5-iodouracil. Implications for mechanism of inhibition and electron transfer. J Biol Chem 277:13155–13166

66. Nishimoto S, Hatta H, Ueshima H, Kagiya T (1992) 1-(5′-Fluoro-6′-hydroxy-5′,6′-dihydrouracil-5′-yl)-5-fluorouracil, a novel N(1)-C(5′)-linked dimer that releases 5-fluorouracil by radiation activation under hypoxic conditions. J Med Chem 35:2711–2712

67. Ito T, Tanabe K, Yamada H, Hatta H, Nishimoto S (2008) Radiation- and photo-induced activation of 5-fluorouracil prodrugs as a strategy for the selective treatment of solid tumors. Molecules 13:2370–2384

68. Zhang Z, Hatta H, Ito T, Nishimoto S (2005) Synthesis and photochemical properties of photoactivated antitumor prodrugs releasing 5-fluorouracil. Org Biomol Chem 3:592–596

69. Pasqualini R, Koivunen E, Kain R, Lahdenranta J, Sakamoto M, Stryhn A, Ashmun RA, Shapiro LH, Arap W, Ruoslahti E (2000) Aminopeptidase N is a receptor for tumor-homing peptides and a target for inhibiting angiogenesis. Cancer Res 60:722–727

70. Lin W, Peng D, Wang B, Long L, Guo C, Yuan J (2008) A model for light-triggered porphyrin anticancer prodrugs based on an o-nitrobenzyl photolabile group. Eur J Org Chem 793–796

71. Takiuchi H, Ajani JA (1998) Uracil-tegafur in gastric carcinoma: a comprehensive review. J Clin Oncol 16:2877–2885

72. Sinkel C, Greiner A, Agarwal S (2008) Synthesis, characterization, and properties evaluation of methylcoumarin end-functionalized poly(methyl methacrylate) for photoinduced drug release. Macromolecules 41:3460–3467

73. Kolb HC, Finn MG, Sharpless KB (2001) Click chemistry: diverse chemical function from a few good reactions. Angew Chem Int Ed Engl 40:2004–2021

74. Meldal M, Tornøe CW (2008) Cu-catalyzed azide-alkyne cycloaddition. Chem Rev 108:2952–3015

75. Agard N, Baskin J, Prescher J, Lo A, Bertozzi C (2006) A comparative study of bioorthogonal reactions with azides. ACS Chem Biol 1:644–648

76. Binder WH (2008) "Click"—chemistry in polymer and material science: the update. Macromol Rapid Commun 29:951

77. Hou J, Liu X, Shen J, Zhao G, Wang PG (2012) The impact of click chemistry in medicinal chemistry. Expert Opin Drug Discov 7:489–501

78. Kolb HC, Sharpless KB (2003) The growing impact of click chemistry on drug discovery. Drug Discov Today 8:1128–1137

79. Lahann J (2009) Click chemistry for biotechnology and materials science. In: Click chemistry for biotechnology and materials science. Wiley, Chichester

80. Neibert K, Gosein V, Sharma A, Khan M, Whitehead MA, Maysinger D, Kakkar A (2013) "Click" dendrimers as anti-inflammatory agents: with insights into their binding from molecular modeling studies. Mol Pharm 10:2502–2508

81. Sevenson S, Tomalia DA (2012) Dendrimers in biomedical applications-reflections on the field. Adv Drug Delivery Rev 64:102–115

82. Sletten EM, Bertozzi CR (2009) Bioorthogonal chemistry: fishing for selectivity in a sea of functionality. Angew Chem Int Ed 48:6974–6998

83. van Berkel SS, Dirks AT, Debets MF, van Delft FL, Cornelissen JJ, Nolte RJ, Rutjes FP (2007) Metal-free triazole formation as a tool for bioconjugation. ChemBioChem 8:1504–1508

84. McKay CS, Moran J, Pezacki JP (2010) Nitrones as dipoles for rapid strain-promoted 1,3-dipolar cycloadditions with cyclooctynes. Chem Commun 46:931–933

85. Blackman ML, Royzen M, Fox JM (2008) Tetrazine ligation: fast bioconjugation based on inverse-electron-demand Diels-Alder reactivity. J Am Chem Soc 130:13518–13519

86. Devaraj NK, Weissleder R (2011) Biomedical applications of tetrazine cycloadditions. Acc Chem Res 44:816–827

87. Koo H, Lee S, Na JH, Kim SH, Hahn SK, Choi K, Kwon IC, Jeong SY, Kim K (2012) Bioorthogonal copper-free Click Chemistry in vivo for tumor-targeted delivery of nanoparticles. Angew Chem Int Ed 51:11836–11840

88. Hapuarachchige S, Zhu W, Kato Y, Artemov D (2014) Bioorthogonal, two-component delivery systems based on antibody and drug-loaded nanocarriers for enhanced internalization of nanotherapeutics. Biomaterials 7:2346–2354

89. Brudno Y, Desai RM, Kwee BJ, Neel SJ, Aizenberg M, Mooney DJ (2015) In vivo targeting through Click Chemistry. ChemMedChem 10:617–620

90. Azoulay M, Tuffin G, Sallem W, Floret JC (2006) A new drug-release method using the Staudinger ligation. Bioorg Med Chem Lett 16:3147–3149

91. Carl PL, Chakravarty PK, Katzenellenbogen JA (1981) A novel connector linkage applicable in prodrug design. J Med Chem 24:479–480

92. van Brakel R, Vulders RC, Bokdam RJ, Grull H, Robillard MS (2008) A doxorubicin prodrug activated by the Staudinger reaction. Bioconjugate Chem 19:714–718

93. Gorska K, Manicardi A, Barluenga S, Winssinger N (2011) DNA-templated release of functional molecules with an azide-reduction-triggered immolative linker. Chem Commun 47:4364–4366

94. Versteegen RM, Rossin R, ten Hoeve W, Janssen HM, Robillard MS (2013) Click to release: instantaneous doxorubicin elimination upon tetrazine ligation. Angew Chem Int Ed 52:14112–14116

95. Bielski R, Witczak Z (2013) Strategies for coupling molecular units if subsequent decoupling is required. Chem Rev 113:2205–2243

96. Matikonda SS, Orsi DL, Staudacher V, Jenkins IA, Fiedler F, Chen J, Gamble AB (2015) Bioorthogonal prodrug activation driven by a strain-promoted 1,3-dipolar cycloaddition. Chem Sci 6:1212–1218

97. Crabtree RH (2014) The organometallic chemistry of the transition metals, 6th edn. Wiley, Hoboken

98. Beller M, Bolm C (2008) Transition metals for organic synthesis: building blocks and fine chemicals, 2nd edn. Wiley-VCH Verlag GmbH, Weinheim

99. Waldron KJ, Rutherford JC, Ford D, Robinson NJ (2009) Metalloproteins and metal sensing. Nature 460:823–830

100. Unciti-Broceta A, Johansson EM, Yusop RM, Sánchez-Martín RM, Bradley M (2012) Synthesis of polystyrene microspheres and functionalization with Pd(0) nanoparticles to perform bioorthogonal organometallic chemistry in living cells. Nat Protocols 7:1207–1218

101. Völker T, Meggers E (2015) Transition-metal-mediated uncaging in living human cells—an emerging alternative to photolabile protecting groups. Curr Opin Chem Biol 25:48–54

102. Li J, Chen PR (2016) Development and application of bond cleavage reactions in bioorthogonal chemistry. Nat Chem Biol 12:129–137

103. Streu C, Meggers E (2006) Ruthenium-induced allylcarbamate cleavage in living cells. Angew Chem Int Ed 45:5645–5648

104. Sasmal PK, Carregal-Romero S, Parak WJ, Meggers E (2012) Light-triggered ruthenium-catalyzed allylcarbamate cleavage in biological environments. Organometallics 31:5968–5970

105. Yusop RM, Unciti-Broceta A, Johansson EM, Sánchez-Martín RM, Bradley M (2011) Palladium-mediated intracellular chemistry. Nat Chem 3:239–243

106. Unciti-Broceta A, Yusop RM, Richardson PR, Walton JGA, Bradley M (2009) A fluorescein-derived anthocyanidin-inspired pH sensor. Tetrahedron Lett 50:3713–3715

107. Weiss JT, Dawson JC, Macleod KG, Rybski W, Fraser C, Torres-Sánchez C, Patton EE, Bradley M, Carragher NO, Unciti-Broceta A (2014) Extracellular palladium-catalysed dealkylation of 5-fluoro-1-propargyl-uracil as a bioorthogonally activated prodrug approach. Nat Commun 5:3277

108. Weiss JT, Dawson JC, Fraser C, Rybski W, Torres-Sánchez C, Bradley M, Patton EE, Carragher NO, Unciti-Broceta A (2014) Development and bioorthogonal activation of palladium-labile prodrugs of gemcitabine. J Med Chem 57:5395–5404

109. Li J, Yu J, Zhao J, Wang J, Zheng S, Lin S, Chen L, Yang M, Jia S, Zhang X, Chen PR (2014) Palladium-triggered deprotection chemistry for protein activation in living cells. Nat Chem 6:352–361

110. Tonga GY, Jeong Y, Duncan B, Mizuhara T, Mout R, Das R, Kim ST, Yeh YC, Yan B, Hou S, Rotello VM (2015) Supramolecular regulation of bioorthogonal catalysis in cells using nanoparticle-embedded transition metal catalysts. Nat Chem 7:597–603

111. Unciti-Broceta A (2015) Bioorthogonal catalysis: rise of the nanobots. Nat Chem 7:538–539

112. Völker T, Dempwolff F, Graumann PL, Meggers E (2014) Progress towards bioorthogonal catalysis with organometallic compounds. Angew Chem Int Ed Engl 53:10536–10540

113. Pérez-López AM, Rubio-Ruiz B, Sebastián V, Hamilton L, Adam C, Bray TL, Irusta S, Brennan PM, Lloyd-Jones GC, Sieger D, Santamaría J, Unciti-Broceta A (2017) Gold-triggered uncaging chemistry in living systems. Angew Chem Int Ed Engl 56. 10.1002/anie. 201705609

Example of Sacrificial Unit Using Two Different Click Reactions in Coupling and Decoupling (CAD) Chemistry

Roman Bielski, Zbigniew J. Witczak and Donald Mencer

Abstract An example of a specific coupling and decoupling (CAD) chemistry is described. It takes advantage of propargyl acrylate as a sacrificial unit (SU). The addition of a selected compound representing a molecular unit equipped with an azide functionality to the terminal triple bond of the SU and another compound acting as a molecular unit equipped with a thiol functionality to the conjugated double bond of the SU proceeded at very good yields. The construct containing two molecular units can be decoupled using a few different reactions and the decoupling can take place at two positions.

A few years ago, we committed two articles [1, 2] in which we emphasized the growing importance of decoupling of connected molecular units and a need for developing novel cleaving techniques. The point is that there are many circumstances when it is necessary to disconnect two (or more) components of a larger construct. They include the need for effective methods to:

- terminate chemo or radiotherapy,
- decouple moieties that were introduced to enable the employment of specific analytical procedures,
- decouple molecular units from the surface,
- decouple the final product of the synthesis from the solid support.

While there are magnificent methods of coupling various molecular units, the number of good decoupling methodologies is rather limited. There is a clear need

R. Bielski (✉) · Z.J. Witczak (✉)
Department of Pharmaceutical Sciences, Nesbitt School of Pharmacy, Wilkes University, 84 W. South Street, Wilkes-Barre, PA 18707, USA
e-mail: bielski1@verizon.net

Z.J. Witczak
e-mail: zbigniew.witczak@wilkes.edu

D. Mencer
Department of Chemistry and Biology, Wilkes University, 84 W. South Street, Wilkes-Barre, PA 18707, USA

for more and better chemistries applicable to disconnecting molecular units. Very often we know in advance that at a certain point we will have to disconnect the coupled units. In such a case, we should design the whole process wisely. We proposed a general approach to such situations which we call coupling and decoupling (CAD). It seems that the CAD chemistry should become an important tool of bioorthogonal chemistry, a wonderful concept introduced several years ago by Carolyn Bertozzi [3–5].

The coupling chemistry should belong to the category of click chemistry [6, 7]. The decoupling chemistry should be as simple and easy as possible and designed for a specific circumstance. For example, experiments performed in vivo and in the laboratory may require very different chemistries.

Of course, when the coupling process can be reversed at the desired moment, there is no reason not to use the "reversible" reactions. For example, if we can use the formation of ester as a coupling process and its hydrolysis as a decoupling process, we should. Unfortunately, such approach is usually impossible, particularly when the molecular units comprise of a variety of functional groups. Scheme 1 shows the use of "reversible" reactions.

Our methodology (CAD) asks for employing a multifunctional compound that we call a sacrificial unit (SU). The Scheme 2 shows the concept. The sacrificial unit **4** is equipped with two, Z and W, (different) functionalities both capable of forming click products. The two click reactions should not be the same. Additionally, somewhere within the SU (between the functionalities capable of clicking), a cleavable moiety (AB, sacrificial functionality) must be present. After connecting the clickable functionalities Z and W to molecular units (**1** and **2**), the formed construct **5** contains both units. After performing the necessary chemistry and

Scheme 1 Reversible reactions applicable to coupling and decoupling
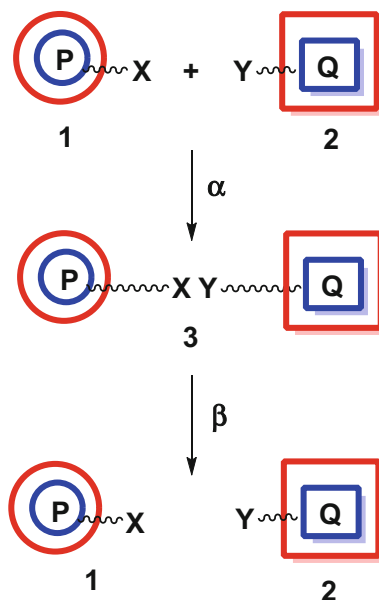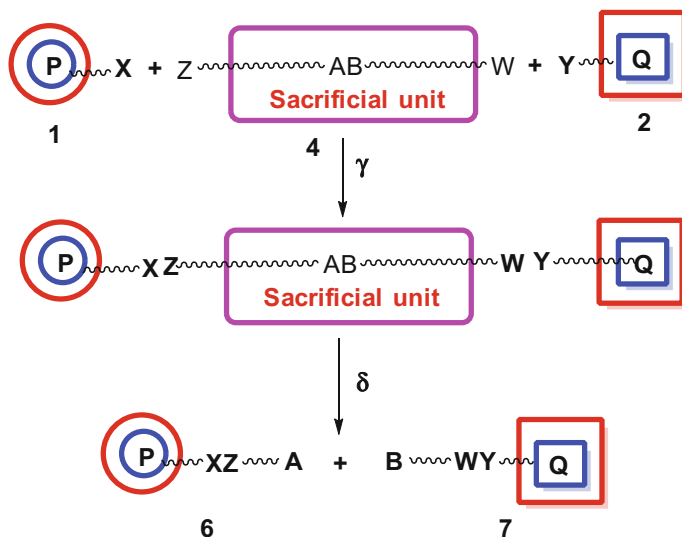
Scheme 2 The use of sacrificial units in the CAD chemistry

accomplishing the desired goals, the molecular units are decoupled by splitting the cleavable moiety (AB).

In most cases, the coupling chemistry is much less demanding because almost always it can be performed in the laboratory. The decoupling must often take place in vivo. However, many decoupling procedures can be performed in the lab as well. Then, many more processes are applicable. The only important requirement is that the molecular units or at least one of them are not degraded by the decoupling procedure.

Since the publication of our paper introducing the CAD concept, several new, effective methods and approaches to the CAD chemistry have been added to the existing repertoire. We find particularly interesting novel decoupling processes. Recently, scientists from the Granada University in Spain [8] offered wonderful examples of a well-designed CAD chemistry. They take advantage of a vinyl sulfonate-based coupling (Michael addition to vinyl sulfonate). To accomplish the decoupling, they use a nucleophilic substitution of sulfonate acting as a leaving group. Another approach to the CAD chemistry was offered by the Xavier group [9]. They use the Huisgen addition of azide to the terminal triple bond as a coupling process and very successfully hydrolyze an amide in the decoupling process. Professor Unciti-Bronceta's group from Edinburgh describes spectacular results of decoupling using metallic palladium [10]. Another very interesting methodology for decoupling has been described by scientists from the Peking University who employ kinases both in vitro and in vivo [11].

The present paper shows a possible execution of the concept of the CAD sacrificial unit. We wanted the unit to be as non-expensive, simple, easy to synthesize,

and robust as possible. We selected propargyl ester of *E*-cinnamic acid as our first target. We expected the propargyl group to be a useful functionality capable of entering a click reaction with an azide group of one of molecular units to form a cyclic triazole. Additionally, the cinnamic acyl functionality can act as an acceptor of a thiol to form a Michael adduct (Scheme 3). While it is not a perfect Michael addition acceptor, we expected ester to be sufficiently effective for our purposes. Thus, we planned to employ a (click) Huisgen reaction between the terminal triple bond of the SU and the azide group of the molecular unit and a (click) Michael addition of the mercaptan present in the other molecular unit to a double bond of the conjugated ester functionality of the SU. The thioether **11** resulting from the Michael addition should be an object of a facile hydrogenation to decouple the unit connected to the thiol from the unit connected to the azide. Alternatively, the hydrolysis or LAH reduction of the ester functionality could result in decoupling at the cinnamic ester location.

The addition of azides to propargyl groups is very well established as a classical click chemistry reaction [7]. Almost equally well established as a click chemistry process is a Michael addition of thiols to α,β-unsaturated carbonyl (or sulfonyl) compounds [12]. Utilizing cinnamic esters as Michael acceptors is less common. Primo, esters are poorer Michael acceptors than corresponding aldehydes or ketones and, secundo, cinnamates are poorer Michael acceptors than corresponding esters without the benzene ring—acrylates. Our reasoning for employing a thiol addition to cinnamic ester was based on the fact that cinnamic esters are very non-expensive, easily available and there are very many substituted derivatives of cinnamic esters. Moreover, today there is a plurality of catalysts available for the Michael addition. They include such catalysts as Lewis acids [13–15], nucleophiles and bases [12, 16], fluorapatite [17], KF-alumina [18], organocatalysts [19], and even bifunctional catalysts [20]. There are examples of successful catalytic Michael additions to cinnamic acid derivatives. A few years ago, it has been shown that in the presence
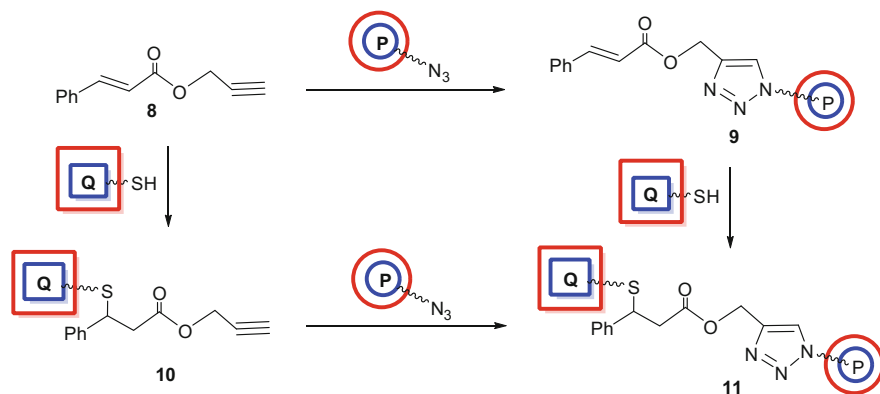
Scheme 3 The possible use of propargyl cinnamate as the sacrificial unit

of iodine even cinnamic acids give excellent yields of Michael addition with thiols and other nucleophiles [21].

We selected a monosaccharide derivative, 1-azido-1-deoxy-2,3,4,6- tetraacetyl-β-D-glucopyranose, to act as a molecular unit equipped with the azide group. It is easily available and should be perfectly visible in the NMR spectrum both before and after the decoupling. We selected thiophenol as a moiety representing a molecular unit capable of adding to conjugated, unsaturated ester. Again, it has NMR (aromatic) signals easily distinguishable from aromatic signals of protons (and carbon atoms) of the cinnamic unit.

Thus, commercially available cinnamoyl chloride was reacted with propargyl alcohol to give the quantitative yield of expected ester (^{1}H NMR; 400 MHz, CDCl$_3$, δ 2.43 ppm, t, 1H, J = Hz, acetylenic; 4.72 ppm, d, 2H; methylene; 6.36 ppm, d, 2H, J = 3.5 Hz, 7.3 ppm, m, 3H, aromatic; 7.4 ppm, m, 2H, aromatic; 7.64 ppm, d, 2H, cinnamic). The reaction of the triple bond of propargyl ester with peracetylated glucose equipped with the azido group at C-1 proceeded smoothly in the presence of copper sulfate and ascorbic acid to give the addition product at a yield of 83% (after column chromatography). The presence of a sharp singlet at 7.85 ppm (^{1}H NMR: 400 MHz, CDCl$_3$, δ 1.80 ppm, s, 3H, Ac; 1.96, s, 3H, Ac; 2.00, s, 3H, Ac; 2.02, s, 3H, Ac; 3.95, m, 1H; 4.08, pd, 1H; 4.23, pd, 1H; 5.18, m, 1H; 5.3, m, 2H; 5.38, m, 2H; 5.83, d, 1H, C–1; 6.38, d, 1H, cinnamic; 7.3, m, 3H; 7.44, m, 2H; 7.66, d, 1H; 7.85, s, 1H, triazole) clearly indicated the formation of the cyclic triazole.

The purified cycloaddition product was reacted with various thiols (thiophenol, acetylcysteine, thiosalicylic acid, 2-thiazoline-2-thiol) in the presence of a catalyst (triethylamine {TEA}, ethyldiisopropylamine) in dichloromethane or chloroform at room or elevated temperatures. While the reagent slowly disappeared, we were not able to isolate satisfactory quantities of the expected addition products. The same can be said about the reaction between propargyl cinnamate and phenylthiol in the presence of TEA. No addition product could be isolated. Surprisingly, even reaction of cyclic triazole containing cinnamic ester with phenylthiol in the presence of iodine gave no addition product. However, it must be noted that trying to prove the generality of the method, we performed our reactions in a solvent (dichloromethane or chloroform) while it is recommended [21] to run the iodine catalyzed Michael addition in a solvent-free system.

In the light of these results, we decided to abandon the cinnamic ester and replace it with a similar compound but a substantially better Michael acceptor, commercially available propargyl acrylate **12**. It still contains a terminal triple bond (capable of reacting with the azide functionality), but the double bond is conjugated to the ester moiety only, and not to the aromatic ring (Scheme 4).

Propargyl acrylate was reacted with peracetylated glucose azide **16** in methylene chloride in the presence of copper (II) acetate/ascorbic acid to give the expected, crystalline click product (cyclic triazole) **17** (Scheme 5). The NMR (^{1}H) spectrum contained a singlet at 7.85 ppm which is characteristic of the proton attached to the triazole ring, four singlets representing protons of acetyl groups and typical carbohydrate signals. Additionally, it contained olefinic signals of the unreacted acrylic
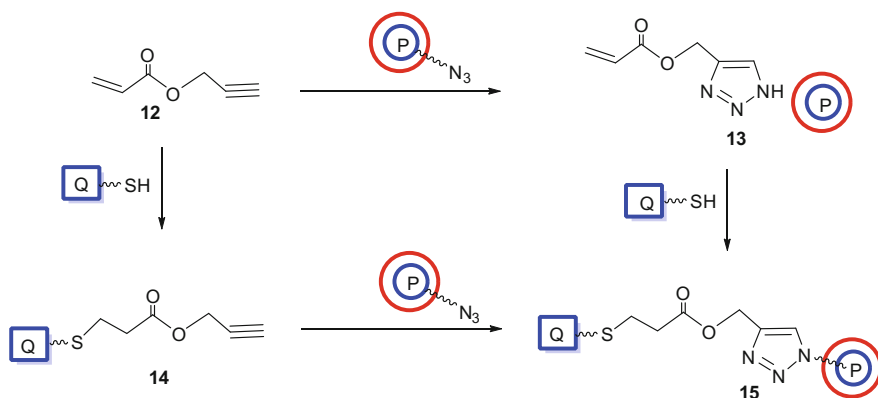
Scheme 4 The possible use of propargyl acrylate as the sacrificial unit
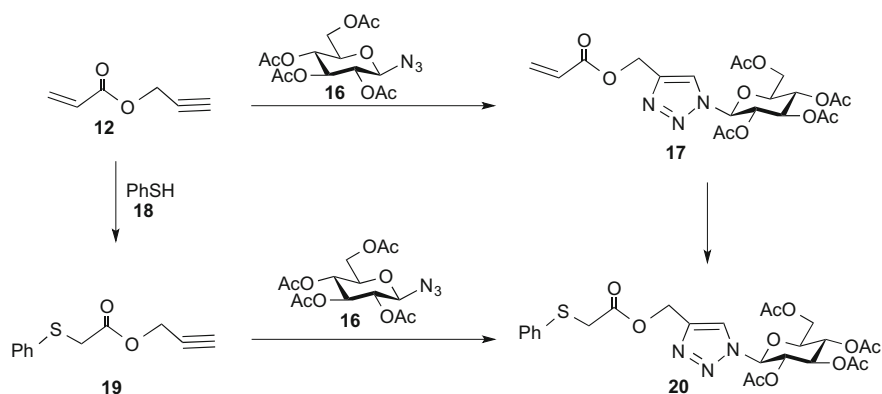
Scheme 5 The example of using propargyl acrylate as the sacrificial unit

double bond. The yield was 88%. The purified product was further reacted with phenylthiol **18** in the presence of tetramethylguanidine (TMG) in chloroform to produce the expected addition product **20**. The reaction mixture was worked up and purified using a column chromatography to give the expected sulfide in 78% yield. The ^{1}H NMR spectrum indicated the presence of both phenyl and glucose signals (^{1}H NMR: 400 MHz, CDCl$_3$, δ 1.79 ppm, s, 3H, Ac; 1.97, s, 3H, Ac; 2.00, s, 3H, Ac; 2.02, s, 3H, Ac; 2.58, t, 2H, "acrylic" methylene; 3.1, broad t, 2H, "acrylic" methylene; 3.92, m, 1H; 4.08, pd, 1H; 4.14, pd, 1H; 5.18, s + m, 2H + 1H, "propargyl" methylene; 5.35, t, 2H; 5.81, t, 1H; 7.12–7.32, m, 5H, aromatic; 7.82, s, 1H, triazole).

The same product was synthesized when a one-pot approach was used. Thus, propargyl acrylate was reacted with a small excess of phenylthiol in methylene chloride in the presence of catalytic amounts of tetramethylguanidine (TMG). After

overnight stirring of the reaction mixture at room temperature, the TLC indicated a disappearance of the starting material. The mixture was treated with 1-azido-1-deoxy-2,3,4,6-tetraacetyl-β-D-glucopyranose and equimolar amounts of copper (II) acetate and ascorbic acid. The color of the reaction mixture changed after about 1 h. The stirring was continued for 3 more days. The work up (aqueous hydrogen carbonate solution, extraction, drying) followed by chromatography gave the crystalline product in 73% yield. The reaction of propargyl acrylate with tetraacetyl glucose azide {in the presence of Cu (I)} followed by the addition of phenylthiol and TMG gave the same product.

There are a few possible ways of decoupling the molecular unit (P) connected to the alkyl (propargyl) part of the SU from the molecular unit (Q) coupled to the acyl (acrylic) part of the SU. The decision which one is the most convenient must be based on the structure of the molecular units P and Q. One option is to hydrolyze the ester functionality provided that the applied conditions will not degrade the molecular units (or at least one of them). In our case of acetylated glucose, most methods will hydrolyze not only (modified) acrylic ester but also the acetate groups present in the carbohydrate unit. Usually, such deprotection will not be of any concern since decoupling is usually performed after all necessary steps had been already performed. Alternatively, one can achieve decoupling at the same position by reducing the ester functionality with lithium aluminum hydride. Of course, no other ester functionalities including acetates will survive. The LAH treatment will reduce carbon sulfur bonds as well but in most cases, it does not matter. The real question is if the molecular units will survive the treatment.

Another point of possible decoupling is the sulfur atom. It seems that some of the methods applicable to decoupling at the sulfur atom require milder conditions. Useful hydrogenation methods include the hydrogenation with Raney nickel and utilizing tributyltin hydride. Most other applicable methods take advantage of transition metal-mediated reactions [22] or sodium and lithium [23, 24].

The experiment using the methods described above and other possible chemistries enabling the decoupling is still performed. The results of this research will be published elsewhere.

1　Conclusion

In the described procedures we used peracetylated glucose azide and phenyl mercaptan as examples of molecular units. Of course, they are to represent much larger units such as proteins or polysaccharides. We expect that practically any molecule can replace peracetylated glucose and a phenyl group provided that they are equipped with the thiol and azide functionalities. The described methodology seems robust and applicable to a variety of molecular units. Moreover, all the starting materials, intermediates, and final products should be safe and applicable to many circumstances. We believe that the methodology is very easy to simple, employs non-expensive, easy available reactants, and the procedures are very simple.

References

1. Bielski R, Witczak Z (2013) Strategies for coupling molecular units if subsequent decoupling is required. Chem Rev 113:2205–2222
2. Bielski R, Witczak ZJ (2013) Paradigm and advantage of carbohydrate click chemistry strategy for future decoupling. In: Witczak ZJ, Bielski R (eds) Click chemistry in glycoscience: new developments and strategies, Wiley, New Jersey, pp 3–30
3. Sletten EM, Bertozzi CR (2009) Bioorthogonal chemistry: fishing for selectivity in a sea of functionality. Angew Chem Int Ed Engl 48(38):6974–6998
4. Bertozzi CR (2011) A decade of bioorthogonal chemistry. Acc Chem Res 44(9):651–653
5. Knight JC, Cornelissen B (2014) Bioorthogonal chemistry: implications for pretargeted nuclear (PET/SPECT) and therapy. Am J Nucl Med Mol Imaging 4(2):96–113
6. Kolb HC, Finn MG, Sharpless KB (2001) Click chemistry: diverse chemical function from a few good reactions. Angew Chem Int Ed 40:2004–2021
7. Thirumurugan P, Matosiuk D, Jozwiak K (2013) Click chemistry for drug development and diverse chemical—biology applications. Chem Rev 113:4905–4979
8. Cruz CM, Ortega-Muñoz M, López-Jaramillo FJ, Hernández-Mateo F, Blanco V, Santoyo-González F (2016) Vinyl Sulfonates. A click function for coupling and decoupling chemistry and their applications. Adv Synth Catal 358:3394–3413
9. Xavier NM, Lucas SD, Jorda R, Schwarz S, Loesche A, Csuk R, Oliveira MC (2015) Synthesis and evaluation of the biological profile of novel analogues of nucleosides and of potential mimetics of sugar phosphates and nucleotides. Synlett 26:2663–2672
10. Weiss JT, Carragher NO, Unciti-Broceta A (2015) Palladium-mediated dealkylation of N-propargyl-floxuridine as a bioorthogonal oxygen-independent prodrug strategy. Sci Rep 5:9329–9335
11. Zhang G, Li J, Xie R, Fan X, Liu Y, Zheng S, Ge Y, Chen PR (2016) Bioorthogonal chemical activation of kinases in living systems. ACS Cent Sci 2:325–331
12. Nair DP, Podgórski M, Chatani S, Gong T, Xi W, Fenoli CR, Bowman CN (2013) The thiol-michael addition click reaction: a powerful and widely used tool in materials chemistry. Chem Mater 26:724–744
13. Kobayashi S, Ogawa C, Kawamura M, Sugiura M (2001) A Ligand-accelerated chiral lewis acid catalyst in asymmetric michael addition of thiols to α,β-unsaturated carbonyls. Synlett 2001:983–985
14. Ranu BC, Dey SS, Samanta S (2005) Indium(III) chloride—catalyzed Michael addition of thiols to chalcones: a remarkable solvent effect. ARKIVOC 2005:44–50
15. Abe AMM, Sauerland SJK, Koskinen AMP (2007) Highly enantioselective conjugate addition of thiols using mild scandium triflate catalysis. J Org Chem 72:5411–5413
16. Li GZ, Randev RK, Soeriyadi AH, Rees G, Boyer C, Tong Z, Davis TP, Becer CR, Haddleton DM (2010) Investigation into thiol-(meth)acrylate Michael addition reactions using amine and phosphine catalysts. Polymer Chem 1:1196–1204
17. Zahouily A, Abrouki Y, Rayadh A, Sebti S, Dhimane H, David M (2003) Fluorapatite: efficient catalyst for the Michael addition. Tet Lett 44:2463–2465
18. Lenardão EJ, Trecha DO, Ferreira PdC, Jacob RG, Perin G (2009) Green Michael addition of thiols to electron deficient alkenes using KF/Alumina and recyclable solvent or solvent-free conditions. J Braz Chem Soc 20:93–99
19. Fang X, Li J, Wang Ch-J (2013) Organocatalytic asymmetric sulfa-michael addition of thiols to α,β-unsaturated hexafluoroisopropyl esters: expeditious access to (R)-thiazesim. Org Lett 15:3448–3451
20. Liu Y, Sun B, Wang B, Wakem M, Deng L (2009) Catalytic asymmetric conjugate addition of simple alkyl thiols to α,β-unsaturated N-acylated oxazolidin-2-ones with bifunctional catalysts. J Am Chem Soc 131:418–419

21. Gao S, Tzeng T, Sastry MNV, Chu C-M, Liu JT, Lin Ch, Yao C-F (2006) Iodine catalyzed conjugate addition of mercaptans to α, β-unsaturated carboxylic acids under solvent-free conditions. Tet Lett 47:1889–1893
22. Luh T-Y, Ni Z-J (1990) Transition-metal-mediated C–S bond cleavage reactions. Synthesis 1990:89–103
23. Farmer SC, Berg SH (2008) Ring contracting sulfur extrusion from oxidized phenothiazine ring systems. Molecules 13:1345–1352
24. Morales DP, Taylor AS, Farmer SC (2010) Desulfurization of dibenzothiophene and oxidized dibenzothiophene ring systems. Molecules 15:1265–1269

Increased Efficacy of NKT Cell-Adjuvanted Peptide Vaccines Through Chemical Conjugation

Colin M. Hayman, Ian F. Hermans and Gavin F. Painter

Abstract Through vaccination infectious diseases such as smallpox, polio, measles, and tetanus have either been eradicated or significantly restricted. However, there remain many diseases for which no effective vaccine exists, and therefore new vaccine approaches are still needed. Current vaccine approaches that generate strong immune responses are often based on ill-defined immunogens such as heat-killed or live-attenuated biological products that suffer from concerns related to safety, stability, and lengthy or complex manufacturing processes. For these reasons, there is a strong push toward vaccines that elicit immune responses to defined structures within the targeted pathogen or tissue, which can be achieved by injecting defined antigenic proteins or peptides. On their own, proteins or peptides are generally poorly immunogenic and they must be combined with immune stimulants known as adjuvants to drive antigen-specific immune responses. Recent studies have shown that the direct conjugation of adjuvant compounds to protein or peptide antigens can enhance the magnitude and quality of induced immune responses. In this chapter, we will discuss the chemical approaches our group has used to synthesize a new class of vaccines based on conjugation of peptides with lipid structures that activate innate-like T cells. The stimulatory milieu created by these structures helps drive potent T cell-mediated immune responses that can prevent infectious disease, or can act therapeutically in noncommunicable conditions as diverse as cancer and allergy.

C.M. Hayman (✉) · G.F. Painter
The Ferrier Research Institute, Victoria University of Wellington,
Wellington, New Zealand
e-mail: colin.hayman@vuw.ac.nz

I.F. Hermans
Malaghan Institute of Medical Research, Wellington, New Zealand

1   Introduction

The control or eradication of infectious disease through vaccination is arguably one of mankind's major achievements. Effective prophylactic vaccines have been developed for many indications including smallpox, polio, and measles. As illustrated in Fig. 1, the impact of vaccination can be dramatic when used widely and in the case of smallpox, the disease was declared eradicated in 1980. However, there remain a number of situations where current vaccine technologies simply do not exist or are cumbersome. For example, the Bacillus Calmette–Guérin (BCG) vaccine for tuberculosis only works in certain groups and there are no vaccines for malaria or human immunodeficiency virus (HIV). In other situations, such as influenza, vaccines show limited cross-strain reactivity creating the need for regular updates which require seasonal, strain-specific vaccines to be manufactured and administered every year. The situation arises because influenza vaccines generate antibody immune responses directed at surface protein structures that change seasonally. In contrast to this, vaccines that generate T cell responses targeted at conserved epitopes across different strains are expected to show cross-strain reactivity. However, there are currently no precisely defined vaccine technologies that drive T cell responses.

Furthermore, vaccines that generate T cell responses that can also be used therapeutically against noncommunicable diseases such as cancer and allergy are now emerging. In these situations, vaccines that generate humoral (antibody) immune responses are likely to be less effective. For example, a promising new vaccine concept for allergy that targets the antigen-presenting cells that drive allergic responses for destruction exploits T cell-mediated cytotoxicity rather than antibody-mediated mechanisms [1]. Since allergy is maintained by antibody production (IgE), vaccines that generate antibody responses may even exacerbate disease. Whether it be to treat infectious disease or noncommunicable disease, for the field to make rapid progress, it desperately needs new precisely defined vaccine technologies that drive T cell responses.

Fig. 1 The number of measles cases in the US before and after the introduction of widespread vaccination
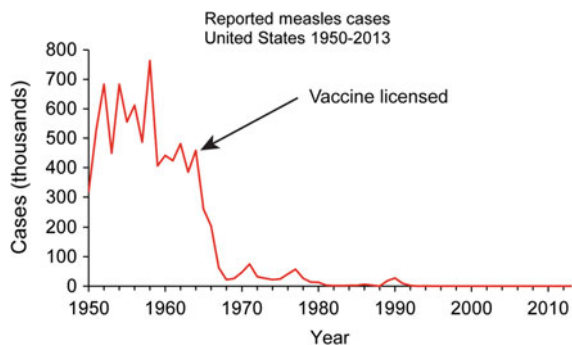
Source data: Centers for Disease Control and Prevention. Epidemiology and Prevention of Vaccine-Preventable Diseases. Hamborsky J, Kroger A, Wolfe S, eds. 13th ed. Washington D.C. Public Health Foundation, 2015.

Through the action of T cells, the immune system can recognize and respond to fragments of protein antigens, known as peptide epitopes, presented by major histocompatibility complex (MHC) molecules on the surface of infected or neoplastic (cancerous) cells through a process of antigen presentation. In the priming phase of an immune response, these antigens must first be acquired from the affected tissues by antigen-presenting cells (APCs), and then peptide epitopes presented by these cells to T cells in the lymphoid tissues. Peptides presented on MHC class I molecules are capable of stimulating CD8$^+$ T cells, whereas peptides presented on MHC class II molecules are capable of stimulating CD4$^+$ T cells. The peptide-recognition receptors expressed by T cells, called T cell receptors (TCRs), are generated through a random somatic recombination process. Within the vast repertoire of T cells generated, there is a high likelihood that some cells will recognize a given antigenic peptide structure in the context of MHC molecules. An adaptive immune response is initiated when these few antigen-specific T cells are stimulated to undergo clonal differentiation and proliferation into effector T cells that have the ability to control infection, or eliminate infected or neoplastic tissue (Fig. 2). A powerful feature of this whole process is that any abnormally expressed protein can potentially be recognized by T cells, regardless of location, be that protein derived from a pathogenic agent or resulted from genetic instability in neoplastic cells.

In contrast to T cells, B cells, the other major component of the adaptive immune system, are not cellular effectors in themselves, but secrete their antigen receptors in the form of antibodies. Antibodies can recognize extracellular antigens or components expressed on the outer surface of cells; they cannot generally recognize antigens in intracellular locations, or epitopes that are concealed within larger structures. To initiate responses to extracellular antigens, recognition is initiated through surface-expressed B cell receptors (BCRs). Once engaged, antigen-specific BCRs can trigger antigen internalization and processing resulting in peptide
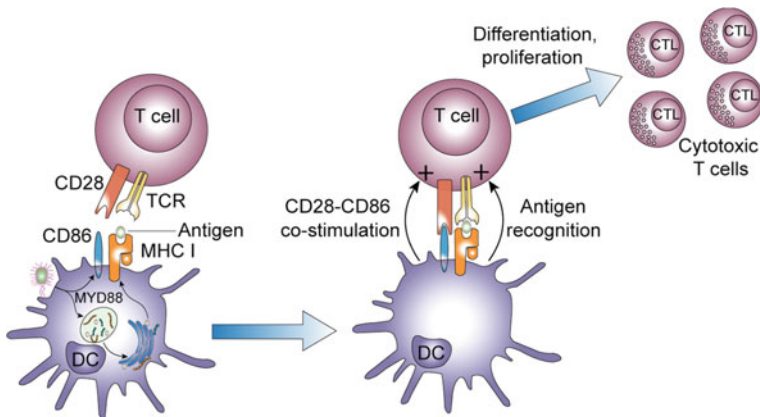

Fig. 2 MHC class I presentation and the generation of cytotoxic T lymphocytes
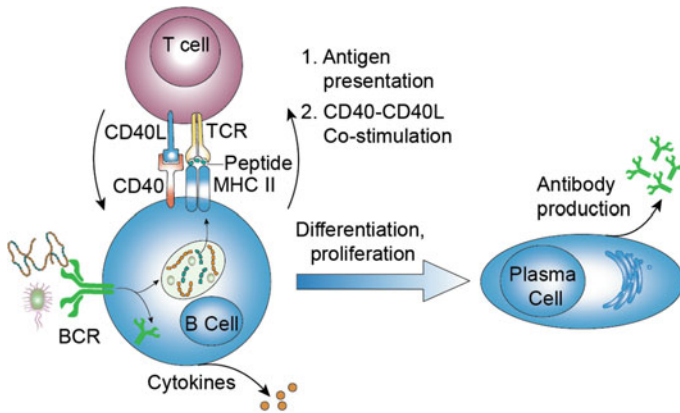
Fig. 3 B cell (humoral) immune response

fragments being presented by MHC class II molecules to CD4$^+$ T cells. This process, known as T cell help, contributes to the further activation and differentiation of B cells into plasma cells—B cells that are able to secrete large quantities of a soluble form of their BCR, namely antibodies (Fig. 3).

Our understanding of precisely what type of immune response will be most effective in certain disease situations is still rudimentary, but nonetheless it is clear that responses have to be tailored in each case; an inappropriate response can be ineffective, or even detrimental. There is therefore a corresponding need to generate new vaccines that induce specifically tailored immune responses. Modifying immune responses to whole cell, live-attenuated or vaccines that contain complex ill-defined components, is challenging and somewhat empirical. In addition to this, concerns around safety and lengthy manufacturing processes are driving a strong push toward vaccines that comprise well-defined components such as antigenic proteins, or peptides that encompass specific epitopes.

Due to their ease of manufacture and relatively simple characterization, peptide vaccines are particularly attractive vaccine candidates from a chemical manufacturing perspective [2]. However, and in contrast to whole microorganisms, a major disadvantage of peptide vaccines is their lack immunogenicity [3]. The intrinsic immunogenicity of vaccines based on microorganisms is due largely to the presence of immunostimulatory molecules that trigger pattern recognition receptors (PRRs). Expression of PRRs is largely confined to cells of the innate arm of the immune response, including most APCs. These receptors have evolved to recognize a variety conserved chemical structures that are features of microorganisms, or various tissue-derived structures associated with damage or inflammation. Examples include toll-like receptor 2 (TLR-2), which binds lipopeptide structures typically found in bacterial cell walls, and TLR-9, which binds unmethylated CpG structures typically seen in prokaryotic DNA. Once engaged, PRRs such as the TLRs activate intracellular signaling cascades that trigger immediate effector functions, such as

respiratory burst or cytokine release, or that can orchestrate downstream adaptive immune responses. Importantly, APCs positioned in infected or perturbed tissue, or in the lymphoid tissues associated with these sites, can efficiently acquire antigens in addition to receiving stimulation via their PRRs. This pattern recognition is significant in that it triggers dendritic cells (DCs), a specialized APC subpopulation, to release soluble factors such as pro-inflammatory cytokines and chemokines, and upregulates surface expression of costimulatory and antigen presentation molecules that enable effective stimulation of T cells (Fig. 4). In order to accomplish this, changes in the surface expression of adhesion molecules facilitate migration to T cell areas of the secondary lymphoid organs (e.g., draining lymph nodes or spleen) where they present antigen. Responding T cells can differentiate into effector cells that recirculate into the affected tissues, or can provide help to B cells to make antibodies.

Because the triggering of PRR pathways (i.e., the TLRs) leads to increased antigen-specific immune responses, compounds that engage these pathways can be used as vaccine adjuvants. However, despite the huge potential that vaccine adjuvants of this type have to modify and shape vaccine response, there are very few products currently licensed for use in humans. Examples include the TLR-4 agonist monophosphoryl lipid A (MPL), which is incorporated into GSK's Engerix-B vaccine, Dynavax's CpG-rich motifs that are included in various products, including their vaccine for Hepatitis B (i.e., HEPLISAV-B), and the Pam_3CSK_4 series of compounds recognized via TLR-2 that were originally derived from N-terminally lipidated proteins isolated from the cell wall of mycobacterial species (Fig. 5).

Importantly, the addition of TLR agonists (along with other PRR agonists) serves to increase the immunogenicity of poorly immunogenic peptide vaccines [4]. Furthermore, it has recently been shown that this effect can be further enhanced by chemically conjugating adjuvant compounds directly to peptide antigens. This observation can be explained, at least in part, by the requirement of APCs to not only present peptide antigens to T cells but also be appropriately activated at the same time. For example, appropriate activation results in the provision of costimulatory molecules (e.g., CD86) providing additional T cell signals that result in its differentiation of $CD8^+$ T cells into cytotoxic T lymphocytes (CTLs) that have the capacity to kill target loaded cells (Fig. 2).
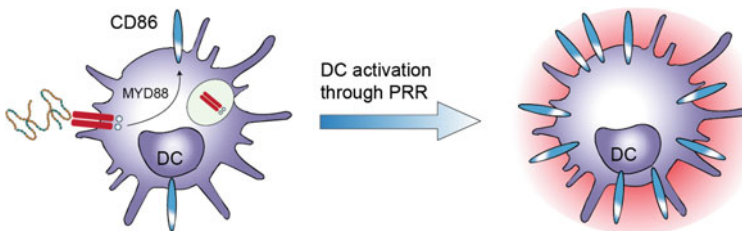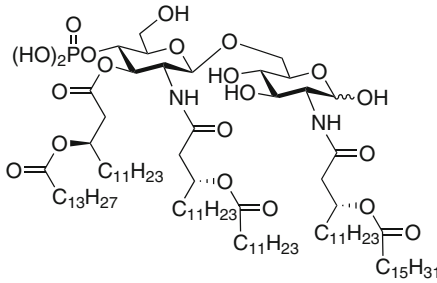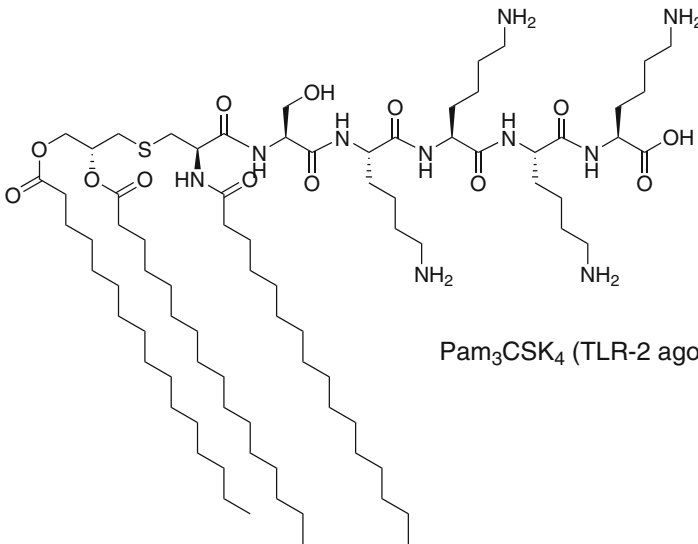

Fig. 4 Pattern recognition receptor signaling leading to DC activation (represented by cell surface CD86 upregulation)

Main active component of MPL® (TLR-4 agonist)

5'-TCGGCGC-HEG-AACGTTC-HEG-TCGGCGC-3'

CpG ODN DV230 (TLR-9 agonist)

Pam₃CSK₄ (TLR-2 agonist)

Fig. 5 Chemical structures of MPL®, DV230 and Pam$_3$CSK$_4$

In an exciting body of work, TLR ligands covalently bound to peptide antigens have been shown to increase the magnitude of either T (cellular) [5] or B (humoral or antibody) cell [6] responses. For example, conjugation has been reported to increase cellular responses through improved antigen targeting to DCs—APCs that are highly specialized in stimulating T cell responses. Conjugated compounds ultimately require cleavage within APCs, a processing step that effectively leads to the buildup of an intracellular antigen depot that supports prolonged antigen presentation and improves immune response [5, 7]. Many of these observations were

described for conjugates that utilize Pam_3CSK_4 as a TLR-2 agonist, and however conjugate vaccines have also been prepared based on a number of alternative TLR agonists such as a monoacyl TLR-2 ligand [8], the TLR-9 ligand CpG [9], and others [10–13].

The conjugation of larger sequences that include T-helper peptide epitopes can drive increased antibody production [6]. For example, a fully synthetic three-component vaccine containing the TLR-2 agonist Pam_3CSK_4, a T-helper epitope from polio virus and a B cell epitope derived from mucin was able to drive increased IgG antibody responses (Fig. 6). All three components were required to be conjugated together to induce the maximal response, and a liposomal delivery vehicle was needed.

Compared to pattern recognition receptor (PRR) agonists such as the TLR agonists, a less-explored approach to activating APCs is the use of ancillary cells that provide stimulatory signals to the APC. The most studied cells of this type are a sub-class of innate-like T cell known as type I natural killer T (NKT) cells, which are defined on the basis of a largely invariant TCR structure, expression of surface molecules in common with classic innate cells like Natural Killer (NK) cells, and a semi-activated phenotype that enables them to respond rapidly to antigen recognition [14]. Compared to classical T cells that are very rare, NKT cells are in relative abundance in many mammalian species, including humans and mice, and their TCRs structures are conserved between individuals (and across species) making the antigens they recognize attractive adjuvants for vaccine development. Whereas classical T cells recognize MHC-restricted peptide antigens (as discussed above), NKT cells respond to *lipid antigens* presented by the non-classical MHC class I molecule CD1d. The CD1d molecule contains a hydrophobic binding groove that sits between two α-helixes and above an antiparallel β-sheet that can accommodate lipids. Associated polar groups such as sugars or phosphates are situated at
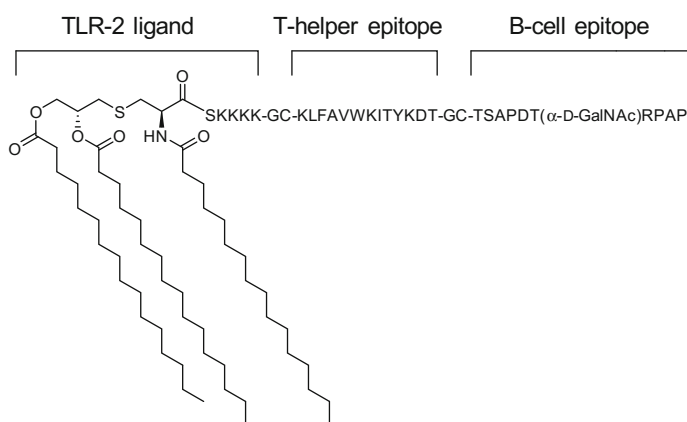

Fig. 6 The chemical structure of a TLR-2 self-adjuvanting lipopeptide vaccine (see text for details)

the surface for interaction with the TCR. A number of synthetic and natural antigens have been reported that consist of a lipid backbone and polar head group [15]. The prototypical antigen is α-galactosylceramide (α-GalCer) (Fig. 7), a synthetic glycolipid compound that was derived from structure–activity relationship studies on a class of marine sponge-derived glycolipids; it remains the most studied of all the reported compounds [16]. Its lipid chains sit deep in the CD1d hydrophobic binding pockets, whereas its polar head group forms a number of hydrogen bonds with CD1d and the TCR [17].

Lipid antigens, such as α-GalCer, presented in this fashion by APCs serve to activate NKT cells (Fig. 8). Activated NKT cells have been implicated in a variety of biological activities including antitumor and antimicrobial activities. The antitumor activity has been attributed to several activities, including NKT-mediated cytotoxicity toward CD1d[+] tumor cells, trans-activation of NK cells, and release of antiangiogenic cytokines like IFN-γ. In addition to this (and of direct relevance to this review), NKT cells are able to activate APCs through the provision of CD40L and soluble factors leading to the enhancement of T and B cell responses [18]. Therefore, in this context, lipid antigens such as α-GalCer can also be considered vaccine adjuvants since when added to vaccines they can drive increased antigen-specific immune responses.
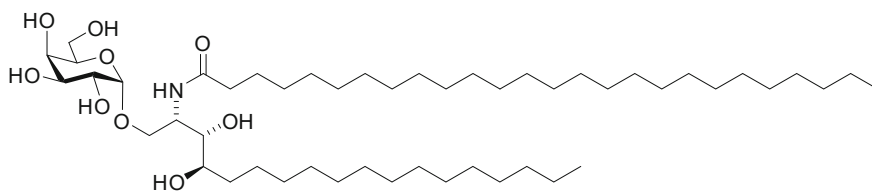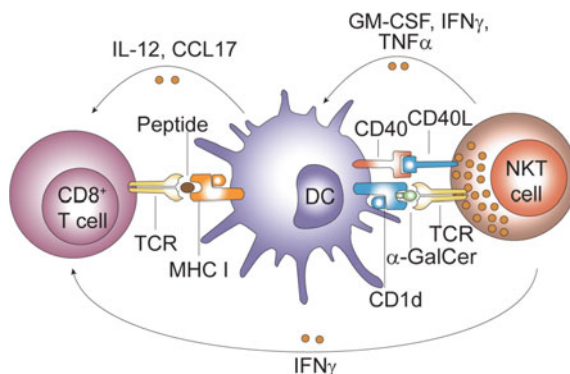

Fig. 7 α-GalCer chemical structure

Fig. 8 Alternative APC activation through NKT cell help

2 Transforming α-GalCer into a Proadjuvant

At the heart of the vaccine technology to be described, herein is the utilization of NKT cells as "cellular adjuvants" through provision of glycolipid antigens combined with the concept that conjugation of vaccine adjuvants through covalent linkages to antigens can lead to increased adaptive immune responses and efficacy in disease challenge models. We favored the idea of utilizing NKT cells over the direct stimulation of APCs through PRRs because highly potent responses can be achieved in animal models; there was some reported adjuvant activity in in vitro cultures of human leucocytes, and evidence that this mode of APC activation can be complementary to TLR ligation, suggesting opportunities to combine with other adjuvants to maximize impact [18].

Although there are a number of highly potent NKT cell agonists (examples include 7-DW8-5 [19] and ABX196 [20]) that can drive adjuvant activity, we choose to use the original compound α-GalCer (KRN7000) because this compound has already been shown to be safe when administered in humans [21]. However, using α-GalCer as the NKT cell ligand, we required a suitable attachment site that was readily accessible and would allow the formation of stable adjuvant–antigen conjugates. Due to their high reactivity, free amines have been widely exploited as a convenient functionality for conjugation to suitably activated coupling partners; however, the only nitrogen of α-GalCer is present as an amide and is not available for functionalization. As others have done successfully, we considered exchanging the 6-hydroxyl group to an amine [22, 23] or modifying with a reactive appended group [24]; however, these approaches would not release an adjuvant compound that was known to be safe. Other potential sites for antigen conjugation to α-GalCer are the four hydroxyl groups of the galactosyl moiety which we did not consider attractive conjugations sites as they are not readily accessed in a selective manner and are less reactive than amines.

Practical access to the 2′-N atom as a synthetically useful amine was realized by carefully characterizing a reaction by-product. During an unpublished synthesis of α-GalCer, we isolated **1** as a minor component during the final deprotection step of **2** which had been accessed by Gervay-Hague's [25] glycosylation methodology (Scheme 1).

In this instance, hydrogenation of the alkene and hydrogenolysis of the two benzyl protecting groups was undertaken using Pearlman's catalyst under an atmosphere of H_2 using a mixed solvent system of chloroform/methanol. Using these conditions along with silica gel chromatography for purification, a modest yield of α-GalCer was isolated along with a significant amount of **1** as a low-R_f by-product. We reasoned that the use of chloroform as a co-solvent led to the in situ formation of HCl and promoted the $N \rightarrow O$ migration of the C26 acyl group. Strongly acidic conditions have been reported to promote the $N \rightarrow O$ migration of acyl groups; however, as far as we are aware the use of Pd-catalyzed HCl generation from chloroform had not.

The discovery that an apparently acid-promoted $N \rightarrow O$ acyl migration gave an isolable analogue of α-GalCer with a free amine group immediately raised the intriguing possibility that **1** could form the basis of novel class of adjuvant–antigen conjugates applicable in the field of immunotherapy.

To progress this idea, we established conditions which provided ready access to **1** directly from α-GalCer using HCl in dioxane. We were also able to establish that the free amine of **1** was amenable to further modification by preparing the acetamide **2** (Scheme 2).

We had also envisaged that if **1** was released in a biological setting that the reverse $O \rightarrow N$ acyl migration would be facile under physiological conditions. A dilute solution of **1** in 99:1 PBS/DMSO was monitored over time at ambient
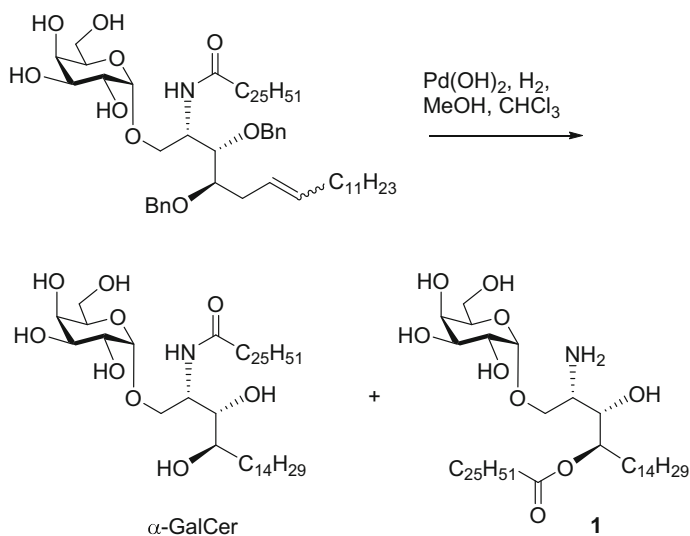
Scheme 1 The discovery of amine **1** during a synthesis of α-GalCer
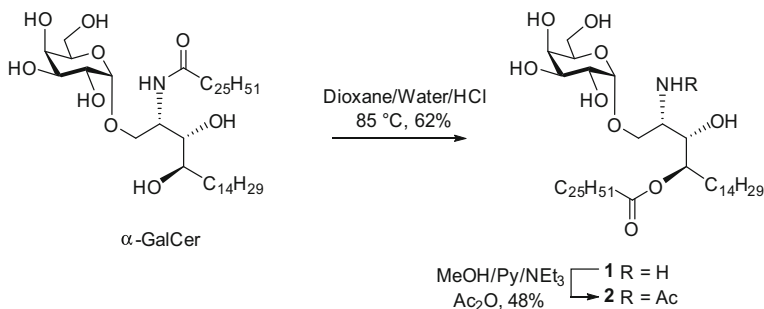
Scheme 2 The preparation of **1** and **2**
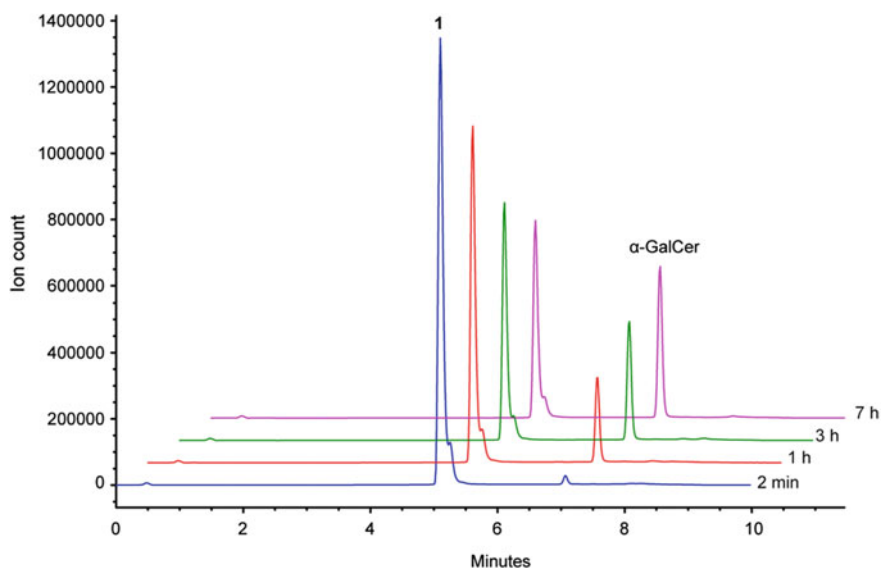
Fig. 9 HPLC-MS chromatograms of **1** in 99:1 PBS/DMSO monitored at the indicated time points

temperature by HPLC-MS. The MS was operated in single ion monitoring (SIM) mode which clearly demonstrated that **1** reverted to α-GalCer over a number of hours (Fig. 9).

With this insight into the fate of **1** at physiological pH, it was expected that its biological activity would be similar to α-GalCer and this was shown to be the case in the context of NKT cell activation in vivo. Compound **1** was able to induce the NKT cell-mediated maturation of DCs as measured by the upregulation of cell surface CD86 on DCs (Fig. 10) [26]. However, with compound **2** the $O \rightarrow N$ acyl migration is unable to take place and this correlated with an inability for **2** to activate NKT cells.

Taken together the in vivo data and the $O \rightarrow N$ acyl migration data suggested that we had a solution to a key part of the puzzle of how to make a conjugate vaccine based on α-GalCer. We proposed that capping the nitrogen atom of **1** with a stable traceless linker provided a proadjuvant which was inactive until enzymatically cleaved allowing for $O \rightarrow N$ acyl migration and generation of the active adjuvant α-GalCer. Combining the proadjuvant concept into a conjugate which comprised an antigen may give a construct which was inactive until enzymatic processing released **1** and a modified peptide antigen for further enzymatic processing to effectively deliver α-GalCer and an antigen to the same site.
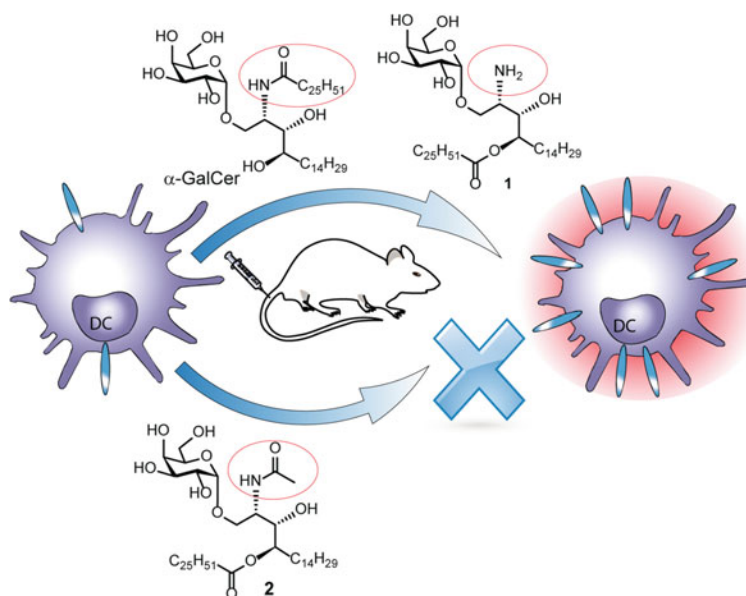
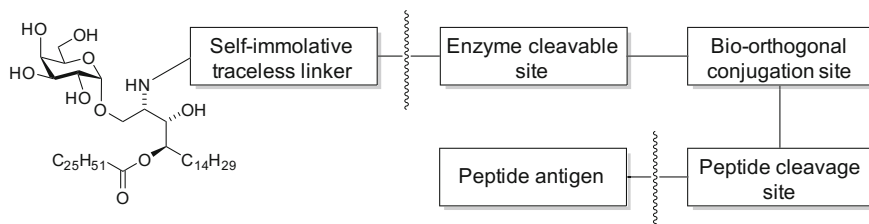**Fig. 10** α-GalCer and **1** activate DCs in vivo, whereas **2** does not

Fig. 11 The design concept for an α-GalCer proadjuvant–antigen conjugate

3 Vaccine Conjugate—Concept and Enablement

To prepare a practical proadjuvant–antigen conjugate, we need to consider a number of aspects of the overall linking strategy which would allow an antigenic peptide to be coupled to **1**. A summary of our design concept is described diagrammatically in Fig. 11.

The first aspect of the design is that the linker attached directly to the N atom must be traceless. So enzymatic processing must either release a free amine directly or reveal an unstable moiety which in turn self-immolates under physiological conditions and collapses to reveal the free amine. Second, the linker needs to have a suitable site for conjugating an antigenic peptide of choice. For this purpose, there is a large body of work in the field of bioorthogonal chemistry where the reactive

chemical moieties couple together under benign conditions and do not react with those functional groups commonly found in biological systems that typically include peptides and proteins [27]. The third aspect we considered was a peptide cleavage sequence appended to the peptide antigen of interest which would encourage peptide release from residual linker components enabling further processing so that peptide epitopes of interest may be released and presented by MHC molecules.

3.1 Esterase Cleavable Linker—Acyloxymethyl Carbamate

Our first family of conjugate vaccines was based on the acyloxymethyl carbamate moiety which has precedent as a self-immolative linker system for amines that can be triggered by enzymatic activity [28]. The general mode of collapse in the context of our vaccine design is described in Fig. 12.

The acyloxymethyl carbamate contains an ester which was expected to be susceptible to enzymatic cleavage where indicated. This releases the hydroxymethyl carbamate moiety which is known to collapse to release formaldehyde and carbon dioxide, and in this case reveal the free amine **1**. To be most effective, we envisage intracellular cleavage will be required for efficacy, however, when administered in vivo this approach may allow for some extracellular cleavage to also occur. Other drugs have demonstrated successful intracellular, esterase-dependent, and linker collapse [29].
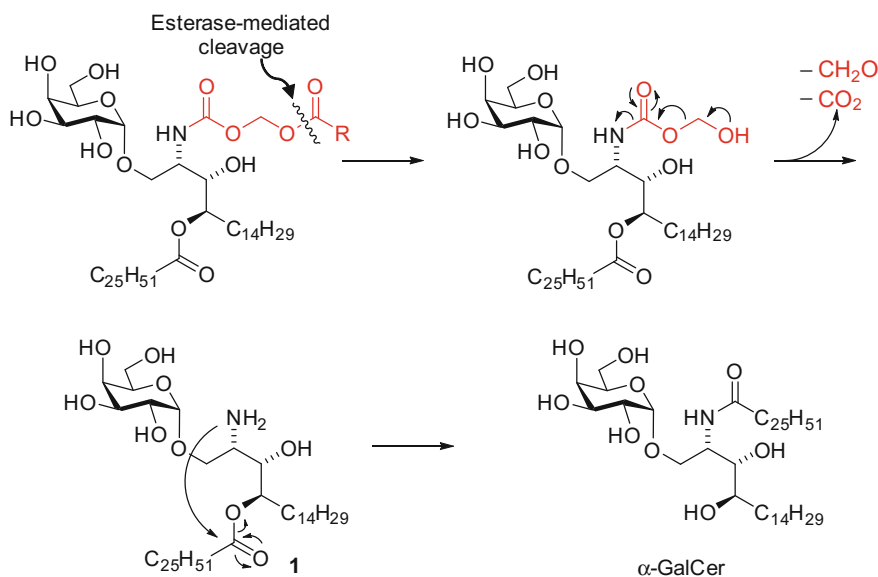

Fig. 12 The mode of collapse of the acyloxymethyl carbamate linker and the generation of α-GalCer proceeding via **1**

The specific acyloxymethyl carbamate derivative of **1** that we decided to make first was **3** where the linker is derived from levulinic acid (Scheme 3) [26].

As was the case for migrated α-GalCer compound **1** intravenous administration of **3** to C57BL/6 mice demonstrated the ability of **3** to induce DC maturation to a similar extent as α-GalCer (Fig. 13). The activity of **3** or α-GalCer was not significantly different to the PBS control when administered to CD1d knockout mice confirming the activity was dependent upon NKT cells since CD1d knockout mice do not have any NKT cells.

To test whether enzymatic processing was required to achieve this activity, an in vitro experiment was conducted whereby plate-bound CD1d was treated with **3** and tested for the ability of any potential CD1d-**3** complex to activate NKT cells (Fig. 14) [26]. Using α-GalCer as a positive control, this experiment demonstrated that **3** was not a good activator of NKT cells in the vitro experiment.
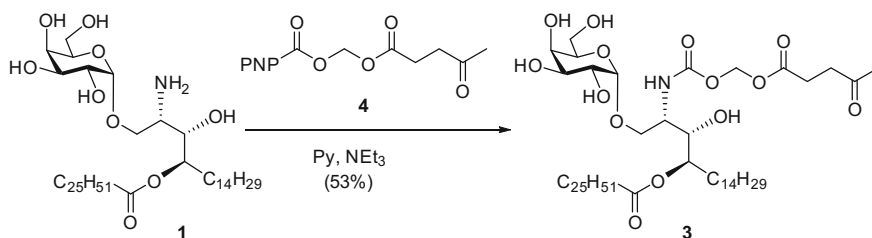
Scheme 3 The preparation of **3**
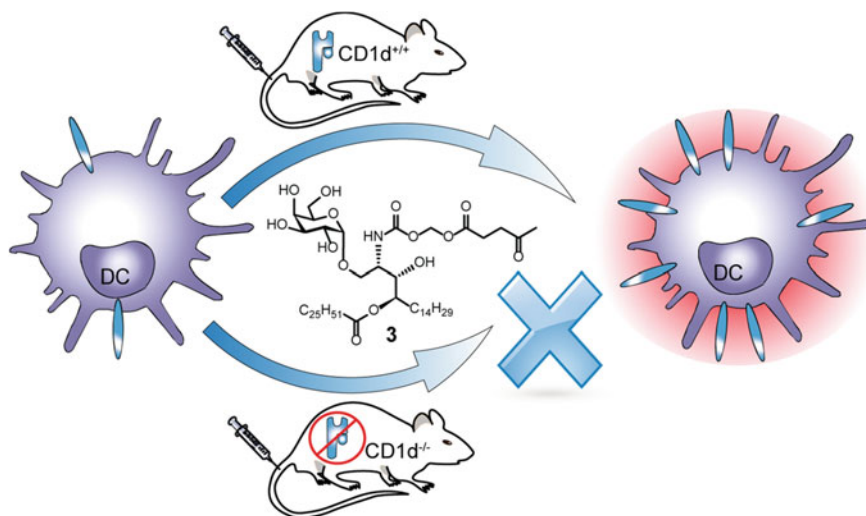

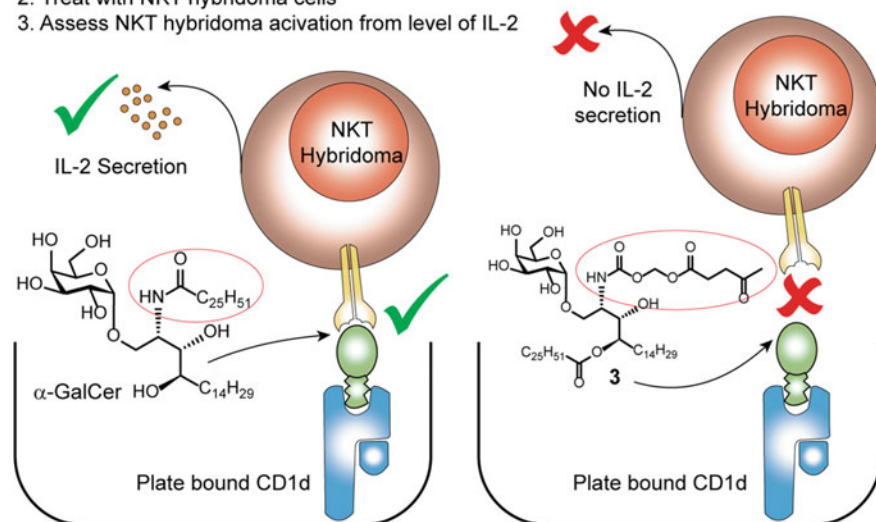**Fig. 13** Compound **3** activates DCs in a CD1d-dependent manner

Fig. 14 Compound **3** is an inactive CD1d ligand in vitro

The observation that **3** cannot activate NKT cells in vitro whereas it is observed in whole animal studies that the injection of **3** elicits a strong NKT cell-dependent response is a good demonstration of **3** acting as a proadjuvant.

3.2 Conjugation of the Peptide Antigen—Oxime Chemistry

To complete the construction of the conjugate vaccine, we prepared the peptide antigen flanked by a peptide cleavage site which was in turn capped at the *N*-terminus with an aminooxy group as the bioorthogonal functionality (**5**). The reaction of aminooxy groups and ketones or aldehydes to form oximes is widely reported due to the reliability of the chemistry and the stability of the oxime moiety. Although oximes do hydrolyze slowly at physiological pH (half-life is measured in the order of weeks), they are well suited for bioconjugation of peptides or proteins to synthetic components free of ketones and aldehydes other than at the site of conjugation. Oxime formation is actually so efficient that solvents of high purity must be used to avoid the scavenging of the hydroxylamine component by low levels of carbonyl impurities common in many solvents [30].

The immunodominant CD8 epitope from ovalbumin, Ser-Ile-Ile-Asn-Phe-Glu-Lys-Leu (SIINFEKL) (OVA$_{257}$), was selected as the peptide antigen. This was appended at the *N*-terminus with the sequence Phe-Phe-Arg-Lys (FFRK) which has been shown to promote antigen presentation of CD8 epitopes via MHC class I
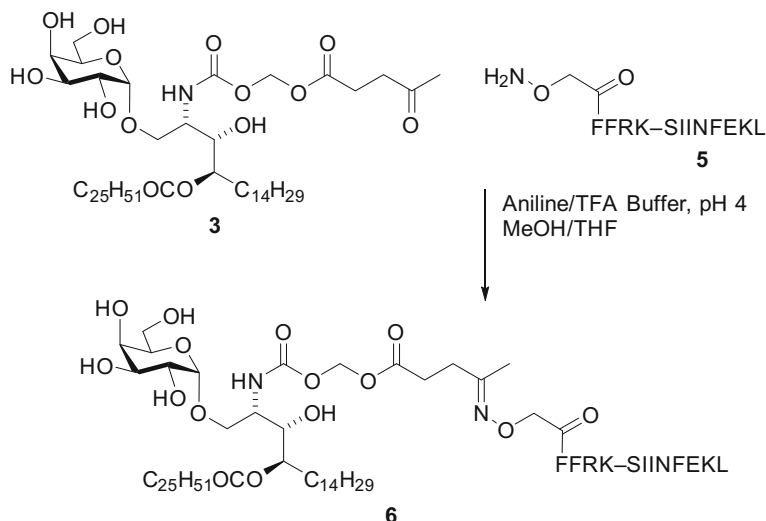
Scheme 4 The conjugation of ketone **3** and oxime **5** to give the vaccine **6**

molecules. The oxime-capped peptide **5** was prepared using Fmoc solid-phase peptide synthesis and reacted with ketone **3** to give the conjugate vaccine **6** (Scheme 4).

A number of features of this conjugation meant that significant optimization of the reaction conditions was required. The degree of difference in the solubilities of the glycolipid and the peptide meant that the reaction had to be carried out at a moderately high dilution (1.6 mM) which required the use of aniline as a catalyst to achieve complete reaction after 48 h. Oxime formations are also generally undertaken at mildly acidic pH to enhance the rate of reaction. In this case, we found that it was necessary to use aqueous aniline which had been buffered to pH 4. At the higher pH of 4.5–5, loss of peptide to its unreactive formaldehyde adduct was observed by LCMS which we attributed to collapse of the linker component of **3** and generation of formaldehyde in situ. With these conditions optimized, however, the reaction was monitored for loss of **3** by HPLC and purified by semi-preparative HPLC to give **6** in an isolated yield of 33%.

The biological activity of **6** was studied extensively in mice which established that it was effective as a vaccine. Proliferation of a transgenic population of peptide-specific $CD8^+$ T cells was observed by cell flow cytometry analysis of blood 7 days after injection with **6** (Fig. 15). The increase as a percentage of total $CD8^+$ T cells was a significant enhancement in the population treated with **6** compared to that treated with an admixture of α-GalCer and OVA_{257} peptide (i.e., approx. 3.5 vs. 0.5%). In addition to the proliferation of transgenic $CD8^+$ T cells, vaccination with **6** produces endogenous T cell population with a markedly enhanced cell-killing ability over the co-administration of α-GalCer and various peptide derivatives [26].
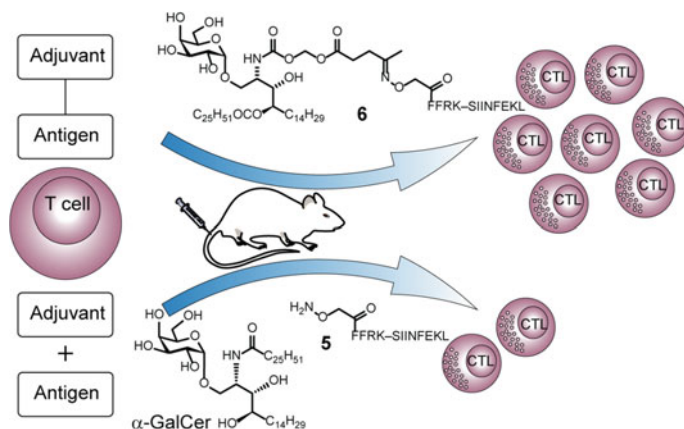
Fig. 15 Administration of conjugate vaccine **6** induces greater proliferation of antigen-specific CD8$^+$ T cells compared to admixed controls

The combined data showed a clear advantage that vaccination with a conjugated adjuvant–antigen constructs provided as compared to co-administration of the components. Intracellular processing of peptide can enhance the efficiency of MHC class I presentation through a cross-presentation mechanism; however, the advantage of the conjugate is also seen for the modified peptides which demonstrate that the effect is not simply due to modification of the peptide.

3.3 NKT Cell Conjugate Vaccine and Asthma

Allergic asthma is characterized by shortness of breath, wheezing, coughing, and airway hyperresponsiveness affecting large numbers of people—particularly in developed countries. Common treatments for the disease such as steroids target the symptoms of the disease, not the underlying cause. At a cellular level, allergen-specific CD4$^+$ T cells that produce type 2 cytokines (i.e., interleukins 4, 5, and 13), known as T$_H$2 cells, have an essential role in driving and maintaining allergic responses. This includes the recruitment of eosinophils and activation of B cells to secrete allergen-specific IgE antibody which in turn leads to mast cell activation [31]. It has been reported that by acquiring and presenting allergen to T$_H$2 cells, DCs are required in the effector phase of the inflammatory response [1]. It has also been reported that adoptively transferred allergen-specific CD8$^+$ T cells can reduce inflammation by mechanisms that potentially include CTL elimination of allergen-specific DCs [32]. For these reasons, we hypothesized that the increased T cell responses generated by our conjugated peptide vaccines might be able to suppress allergic inflammation. As reported [26], we showed our conjugate vaccine could completely suppress eosinophil infiltration, whereas admixed controls

(i.e., α-GalCer and peptide) were ineffective. Further studies are still required to determine if DC killing is the primary mechanism but the finding has opened up an exciting area of ongoing research and development.

3.4 Protease-Cleavable Linker—Valine-Citrulline-p-Aminobenzyl Carbamates

The use of esterase cleavable linkers to covalently attach α-GalCer as an adjuvant with a peptide antigen enabled us to demonstrate a clear advantage over co-administration. We decided an improved vaccine that may be able to be prepared by incorporating a traceless linker which was more directed toward intracellular enzymolysis. Valine-citrulline p-aminobenzyl (VC-PAB) carbamates are susceptible to enzymolysis by cathepsin B which is an abundant lysosomal enzyme in antigen-presenting cells. The VC-PAB linkage has been shown to have good plasma stability and has found application in the clinic in the field of antibody–drug conjugates [33].

The VC-PAB linkage undergoes proteolytic cleavage to reveal the amine of the PAB carbamate which in turn is self-immolative under physiological conditions to give carbon dioxide and azaquinone methide. The collapse cascade in the context of **1** is shown in Fig. 16.

To prepare a conjugate vaccine comprising the VC-PAB linker, we started with the known VC-PAB precursor **7** and capped this at the N-terminus with the active ester of levulinic acid **8** and subsequent formation of the activated carbonate gave **9**. This was coupled to **1** which was followed by oxime formation in a manner analogous to **6** to provide **10** as a second conjugate vaccine (Scheme 5) [34].
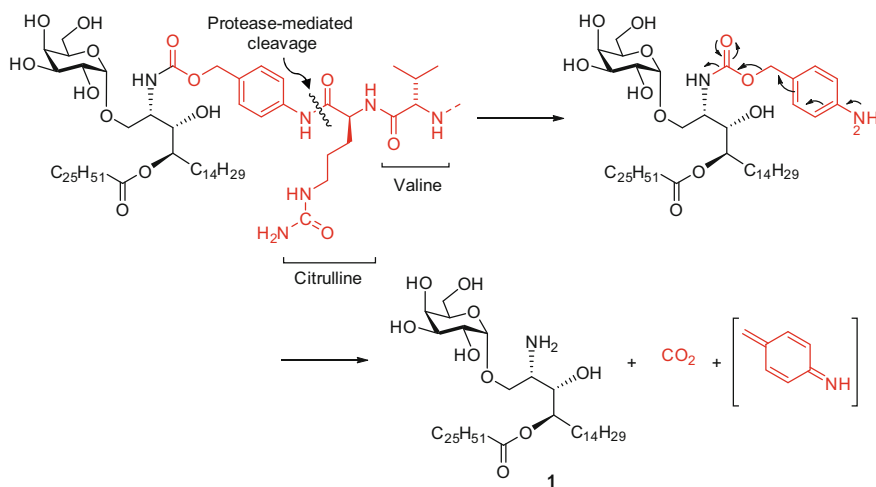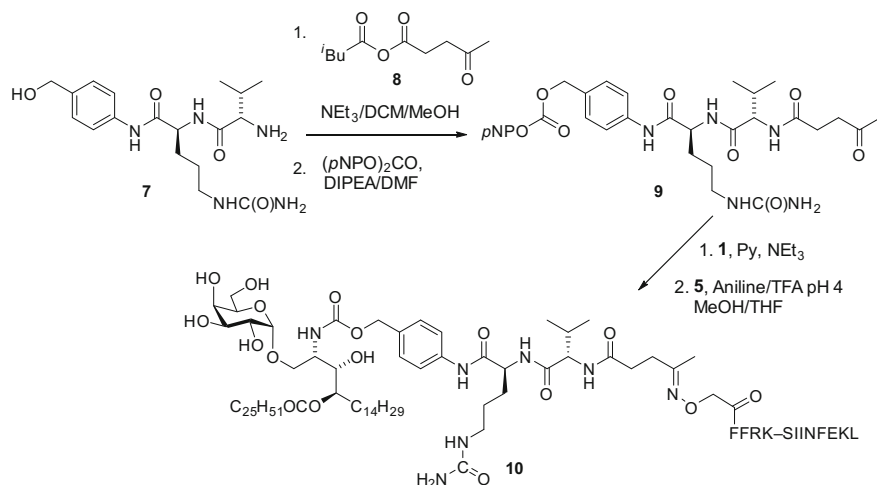

Fig. 16 The mode of collapse of the valine-citrulline-p-aminobenzyl carbamate linker
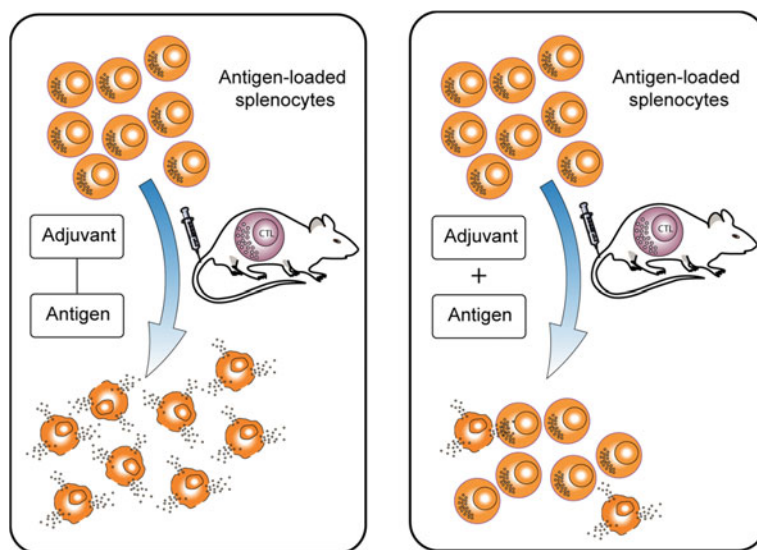
Scheme 5 The preparation of **10**

Fig. 17 Comparison of the ability of vaccination with conjugate compared to admixed controls to generate an endogenous antigen-specific cytotoxic response

The biological activity of **10** was found to be very similar to **6** in a side-by-side comparison in a cell lysis assay. The ability of **6** and **10** to elicit an endogenous antigen-specific cytolytic activity against splenocytes which had been pulsed with different concentrations of OVA_{257} peptide was measured which established a clear advantage of the conjugate vaccines over admix vaccination with α-GalCer and peptide (Fig. 17).

3.5 Alternative Linkage to the Peptide—Cu(I)-Mediated Azide–Alkyne Cycloaddition

Conjugates **6** and **10** had validated the concept of preparing a vaccine comprising covalently bound adjuvant and antigen. A limitation of oxime ligation, however, is the ease of access to a range of peptide antigens. The extreme reactivity of the aminooxy group toward adventitious aldehydes and ketones renders the preparation of aminooxy-capped peptides unreliable, and in our experience, we have found that this can limit the ability of commercial providers to readily supply them.

Therefore, we investigated the use of the Cu(I)-mediated azide–alkyne cycloaddition (CuAAC). The CuAAC reaction is one of the most widely used widely reported ligations in the area of bioconjugation. This is due to a number of reasons which include essentially complete bioorthogonality which azides and alkynes have within biological systems, the relative chemical stability of azides and alkynes, and the ease of preparation of the conjugation partners.

In the context of our vaccine conjugates, we opted to use alkyne-capped peptide **11** (Scheme 6). The glycolipid coupling partners to this were the azides **12** and **14** comprising the esterase- and protease-cleavable linkers, respectively. These were constructed in a manner analogous to the ketone-containing coupling partners above, except the linker portion was built up from 6-azidohexanoic acid instead of levulinic acid.

Similar to the development of the oxime ligations, the CuAAC chemistry required significant optimization in order for the efficient reaction of our coupling partners. Satisfactory conditions were found which used copper foil and copper sulfate as the redox couple (to generate the Cu(I)) along with the accelerating ligand
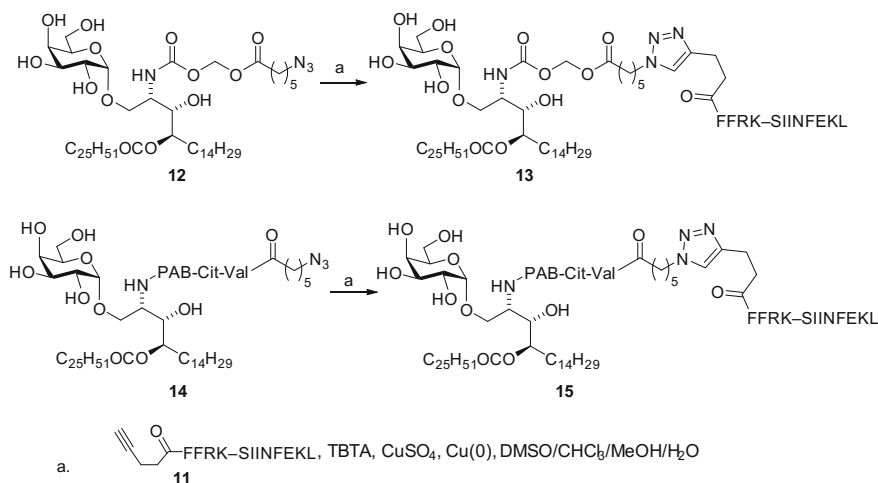
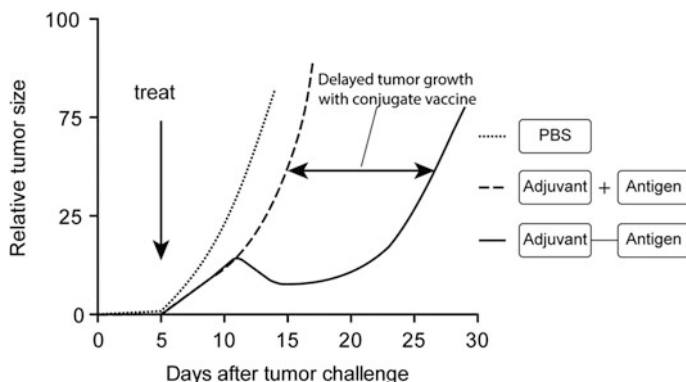**Scheme 6** Synthesis of **13** and **15**

Fig. 18 Comparison of tumor growth over time of mice vaccinated with conjugate vaccines versus admixed controls

Tris-[(1-benzyl-1*H*-1,2,3-triazol-4-yl)methyl]amine (TBTA) to give the conjugates **13** and **15**. The disparate solubility of the reaction components again dictated the use of a complex mixture of solvents to minimize heterogeneity. We also kept the concentration of DMSO at 30% due to the detrimental effect it has on reaction rate when using TBTA as an accelerating ligand. We were unable to get the reaction to proceed using the commonly reported $CuSO_4$ with sodium ascorbate as the reductant [35].

The therapeutic activity of all four vaccine conjugates (**6**, **10**, **13** and **15**) was shown to confer significant antitumor activity against the aggressive mouse melanoma B16.OVA (Fig. 18) [34]. In this test, a single dose of candidate vaccine was administered once the tumors had become engrafted and palpable. The effect of conjugation was demonstrated by the beneficial delay in tumor outgrowth observed when animals were treated with the conjugates as compared to the co-administration of α-GalCer and either capped or uncapped OVA_{257} peptide.

A reproducible feature of the comparison of these four vaccines (i.e., **6**, **10**, **13**, and **15**) was that conjugate **13** gave a less-pronounced antitumor activity than the others. This was an interesting demonstration that the combination of the choice of linker chemistries was altering the quality of the immunological response in vivo.

4 Decoupling

The biological activity presented above demonstrates the clear advantage that the administration of the conjugated adjuvant and antigen have over the co-administration of the components. This was shown for the ability of vaccinated mice to generate a cytotoxic response against antigen-specific cells as well as vaccination having strong antitumor activity in vivo. In vitro studies have also demonstrated that constructs **2** and **3** do not induce CD1d-dependent DC activation.

However, we wanted to check if our vaccine design was releasing the active components in the way we had designed it to. For this purpose, we used the CD1d monomer NKT cell hybridoma in vitro assay (see Fig. 14).

Treating CD1d monomer-loaded plate wells with solutions of conjugate and incubating with NKT hybridoma cells resulted in no appreciable IL-2 release (i.e., no NKT cell activation), whereas the positive control, α-GalCer, induced IL-2 release as expected. In contrast, pre-treatment of the conjugate with cathepsin B at pH 5.5 recovered much of the 'α-GalCer' activity. This demonstrates that the treatment with cathepsin B generates a ligand for CD1d capable of stimulating NKT cell activity. This is consistent with the expected mode of action, whereby the cathepsin B-mediated cleavage of the VC-PAB linker generates **1** which in turn undergoes $O \rightarrow N$ acyl migration to produce α-GalCer (see Fig. 16).

This change in peptide ligation chemistry also gave vaccines with the ability to generate CD8$^+$ T cells with antigen-specific cell lysis activity comparable to the oxime-ligated conjugates **6** and **10** (data not shown).

4.1 Early Release of α-GalCer

HPLC-MSMS was used to monitor for early release of α-GalCer which may arise from hydrolysis of the self-immolative linker. Collapse of the acyloxymethyl carbamate or the VC-PAB linker would initially lead to **1** which we had already shown will give α-GalCer over time in aqueous conditions. Therefore, monitoring solutions of conjugates for the level of α-GalCer over time serves as a proxy for the susceptibility of the two linkers to nonspecific hydrolysis.

Aqueous samples of vaccines **6**, **10**, **13,** and **15** each buffered to pH 3.0, 5.0, 7.4, and 9.0 were stored at 20 °C for 8 days and monitored periodically for α-GalCer level. This study demonstrated that at pH 7.4 and 9.0 that the acyloxymethyl carbamate conjugates **6** and **13** were both susceptible to nonspecific hydrolysis, whereas no α-GalCer was detected in the solutions of the VC-PAB conjugates **10** and **15** (Fig. 19) [34].
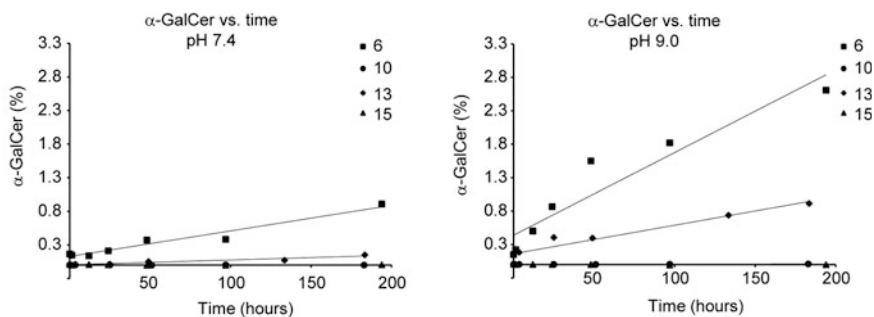

Fig. 19 α-GalCer levels monitored over time by HPLC-MSMS in dilute aqueous solutions of the conjugate vaccines **6**, **10**, **13,** and **15** at pH 7.4 and pH 9.0

4.2 Conjugates Induce Responses in Human Blood

Another important consideration in developing an NKT cell-based vaccine technology is the ability to transfer an effective vaccine response from murine models to humans. Human and mice NKT cells react to the same glycolipid antigens; however, humans have fewer NKT cells than mice. NKT cells represent 4% of the circulating T cell population in mice [36]; however, their abundance in human blood is typically <0.1% [37]. To test whether our conjugate vaccines were applicable in a human setting, we assessed the effect of a conjugate comprising a human-relevant peptide sequence from cytomegalovirus (CMV) pp65 protein [26]. The prevalence of CMV exposure in the local population enabled ready access to blood from donors which had been exposed to CMV. The peptide sequence chosen was NLVPMVATV (NLV) due to its ability to be detected in the peripheral blood mononuclear cells of HLA-A*02$^+$ CMV seropositive donors. With the epitope chosen, the conjugate **16** (Fig. 20a) was prepared which made use of the more stable VC-PAB linker in combination with employing CuAAC chemistry to allow the use of a more readily available peptide coupling partner. The peptide also contained the FFRK peptide cleavage sequence designed to aid MHC class I presentation of the NLV epitope.
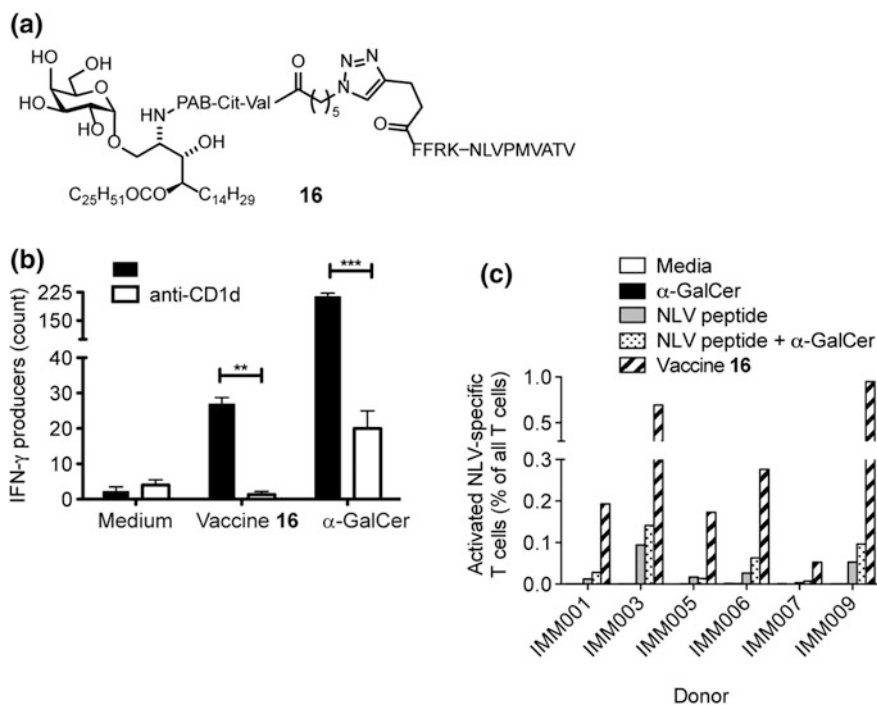

Fig. 20 **a** The chemical structure of the CMV conjugate **16**, **b** Conjugate **16** activates human NKT cells, **c** Conjugate **16** enhances peptide-specific CTL in human blood in vitro. (Reproduced from [26] with permission from the Royal Society of Chemistry)

It was initially very interesting to determine whether compound **16** was able to activate human NKT cells. To do this, **16** was incubated with PBMCs from an HLA-A*02 negative donor in the presence or absence of a CD1d blocking antibody, and the release of interferon-γ (IFN-γ) was monitored (Fig. 20b). HLA-A*02 negative donors were used for this experiment to ensure the IFN-γ was not derived from antigen-specific T cells responding to the NLV peptide. The results from this did indeed indicate that the IFN-γ release was CD1d-dependent which supports our hopes that the conjugate vaccines are active in a human setting. Further assessment of the ability of **16** to promote the proliferation of NLV antigen-specific CD8$^+$ T cells was determined by the upregulation of CD137 which is used as a marker of T cell activation. The enhancement in the percentage of antigen-specific T cells was markedly more pronounced after exposure to the conjugate vaccine **16** as opposed to α-GalCer, the NLV peptide, or co-administration of α-GalCer and NLV peptide (Fig. 20c). Importantly, this proliferation of NLV-specific CD8$^+$ T cells was observed in several donors, which is encouraging when considering the low abundance of NKT cells in humans.

5   Summary and Future Directions

In this work, we showed that treating the NKT cell agonist α-GalCer with acid caused the C26 fatty acid to 'miss-orientate' to form an inactive compound with a reactive amine handle. To stop the facile reformation of α-GalCer in physiological conditions, the amine was capped with a blocking group that was designed to be removed in vivo through enzymatic processes. The chemical design of the capping group also included provision of a chemical reporter group such as a ketone, azide, or alkyne that facilitated bioorthagonal conjugation to suitably N-terminally modified peptide antigens. Although both oxime ligation and copper catalyzed cycloaddition methodologies were equally efficient, the cycloaddition protocol was preferred due to the ease of synthesis of the alkyne-modified peptides versus the alkoxyamine-modified peptides. In terms of self-immolative linkers, the valine-citrulline p-aminobenzyl group was preferred over the acyloxy group due to increased chemical stability.

It was demonstrated that prodrugs (or proadjuvants) and conjugates were active in vivo but inactive when tested in plate-bound (CD1d monomer) assays. Importantly, NKT cell activity could be recovered in the in vitro experiments when the proteolytic enzyme cathepsin B was added. Upon conjugation with peptide antigen, increased T cell responses in terms of numbers and functional activity (i.e., antigen-specific cellular cytotoxicity) were measured in animal models. Taken together, these data provided good evidence that the conjugate vaccine activity was, at least in part, due to the designed mechanism (i.e., co-delivery of adjuvant and antigen to the same cell). Further studies are required to fully elucidate that the impact conjugation has with regard to altered biodistribution and how this contributes to biological activity. Increased T cell responses due to conjugation also

translated into better activity in animal models for allergy and cancer. Using a synthetic vaccine that induces antigen-specific cellular cytotoxicity to suppress acute allergic inflammation is a highly novel and interesting concept; however, further studies into the scope and limitations of this approach are needed.

Although it was pleasing to observe increased antitumor activity in an aggressive melanoma model, delayed tumor growth is unlikely to translate well into the clinic, so ongoing research is still needed to find therapies that induce long-term regression. There are many considerations such as repeated dosing, combination with other vaccine technologies (such as those that contain TLR agonists), the use of synthetic long peptides, analysis of T cell populations, and consideration of tumor microenvironment-induced immunosuppression. Given the stunning success in clinical trials of monoclonal antibodies (mAbs) such as nivolumab and ipilimumab [38] that block so-called checkpoint mechanisms, cancer vaccine technologies can no longer be considered in isolation. In particular, for metastatic disease cancer, vaccines must now be considered in the context of checkpoint blockade treatments and how they may enhance these. Thankfully, because checkpoint blockade treatments work by releasing T cell responses and vaccines by inducing T cell responses, the approaches are likely to be synergistic. However, much research and clinical evaluation are still required to validate this concept and other treatment combinations. Other situations where vaccination may be appropriate include pre-malignant disease [39]. This setting is attractive because vaccines tend to work better in healthier patients with lower disease burden. Furthermore, due to the potential of checkpoint blockade drugs to cause autoimmunity, these treatments are less suited to pre-malignant conditions. Finally, vaccine therapies may also work well as an adjunct therapy after tumor reduction surgery. An excellent recent review discusses these and related points in detail [40].

References

1. van Rijt LS, Jung S, Kleinjan A, Vos N, Willart M, Duez C, Hoogsteden HC, Lambrecht BN (2005) In vivo depletion of lung CD11c + dendritic cells during allergen challenge abrogates the characteristic features of asthma. J Exp Med 201(6):981–991. doi:10.1084/jem.20042311
2. Silverman E (2012) Can we afford the war on cancer? Biotechnol Healthc 9(4):13–16
3. Jones LH (2015) Recent advances in the molecular design of synthetic vaccines. Nat Chem 7 (12):952–960. doi:10.1038/nchem.2396
4. Reed SG, Orr MT, Fox CB (2013) Key roles of adjuvants in modern vaccines. Nat Med 19 (12):1597–1608. doi:10.1038/nm.3409
5. Zom GG, Khan S, Britten CM, Sommandas V, Camps MG, Loof NM, Budden CF, Meeuwenoord NJ, Filippov DV, van der Marel GA, Overkleeft HS, Melief CJ, Ossendorp F (2014) Efficient induction of antitumor immunity by synthetic toll-like receptor ligand-peptide conjugates. Cancer Immunol Res 2(8):756–764. doi:10.1158/2326-6066.CIR-13-0223
6. Ingale S, Wolfert MA, Gaekwad J, Buskas T, Boons GJ (2007) Robust immune responses elicited by a fully synthetic three-component vaccine. Nat Chem Biol 3(10):663–667. doi:10.1038/nchembio.2007.25
7. Zom GG, Filippov DV, van der Marel GA, Overkleeft HS, Melief CJ, Ossendorp F (2014) Two in one: improving synthetic long peptide vaccines by combining antigen and adjuvant in one molecule. Oncoimmunology 3(7):e947892. doi:10.4161/21624011.2014.947892

8. Wright TH, Brooks AES, Didsbury AJ, MacIntosh JD, Williams GM, Harris PWR, Dunbar PR, Brimble MA (2013) Direct peptide lipidation through thiol-ene coupling enables rapid synthesis and evaluation of self-adjuvanting vaccine candidates (vol 52, pg 10616. Angew Chem Int Edit 52(45):11686. doi:10.1002/anie.201305620

9. Khan S, Bijker MS, Weterings JJ, Tanke HJ, Adema GJ, van Hall T, Drijfhout JW, Melief CJM, Overkleeft HS, van der Marel GA, Filippov DV, van der Burg SH, Ossendorp F (2007) Distinct uptake mechanisms but similar intracellular processing of two different Toll-like receptor ligand-peptide conjugates in dendritic cells. J Biol Chem 282(29):21145–21159. doi:10.1074/jbc.M701705200

10. Vecchi S, Bufali S, Uno T, Wu T, Arcidiacono L, Filippini S, Rigat F, O'Hagan D (2014) Conjugation of a TLR7 agonist and antigen enhances protection in the S. pneumoniae murine infection model. Eur J Pharm Biopharm 87(2):310–317. doi:10.1016/j.ejpb.2014.01.002

11. Jackson DC, Lau YF, Le T, Suhrbier A, Deliyannis G, Cheers C, Smith C, Zeng WG, Brown LE (2004) A totally synthetic vaccine of generic structure that targets Toll-like receptor 2 on dendritic cells and promotes antibody or cytotoxic T cell responses. Proc Natl Acad Sci USA 101(43):15440–15445. doi:10.1073/pnas.0406740101

12. Oh JZ, Kedl RM (2010) The capacity to induce cross-presentation dictates the success of a TLR7 agonist-conjugate vaccine for eliciting cellular immunity. J Immunol 185(8):4602–4608. doi:10.4049/jimmunol.1001892

13. Tighe H, Takabayashi K, Schwartz D, Marsden R, Beck L, Corbeil J, Richman DD, Eiden JJ Jr, Spiegelberg HL, Raz E (2000) Conjugation of protein to immunostimulatory DNA results in a rapid, long-lasting and potent induction of cell-mediated and humoral immunity. Eur J Immunol 30(7):1939–1947. doi:10.1002/1521-4141(200007)30:7<1939:AID-IMMU1939>3. 0.CO;2-#

14. Hermans IF, Silk JD, Gileadi U, Salio M, Mathew B, Ritter G, Schmidt R, Harris AL, Old L, Cerundolo V (2003) NKT cells enhance CD4 + and CD8 + T cell responses to soluble antigen in vivo through direct interaction with dendritic cells. J Immunol 171(10):5140–5147

15. Banchet-Cadeddu A, Henon E, Dauchez M, Renault JH, Monneaux F, Haudrechy A (2011) The stimulating adventure of KRN 7000. Org Biomol Chem 9(9):3080–3104. doi:10.1039/c0ob00975j

16. Morita M, Motoki K, Akimoto K, Natori T, Sakai T, Sawa E, Yamaji K, Koezuka Y, Kobayashi E, Fukushima H (1995) Structure-activity relationship of alpha-galactosylceramides against B16-bearing mice. J Med Chem 38(12):2176–2187

17. Borg NA, Wun KS, Kjer-Nielsen L, Wilce MC, Pellicci DG, Koh R, Besra GS, Bharadwaj M, Godfrey DI, McCluskey J, Rossjohn J (2007) CD1d-lipid-antigen recognition by the semi-invariant NKT T-cell receptor. Nature 448(7149):44–49. doi:10.1038/nature05907

18. Hermans IF, Silk JD, Gileadi U, Masri SH, Shepherd D, Farrand KJ, Salio M, Cerundolo V (2007) Dendritic cell function can be modulated through cooperative actions of TLR ligands and invariant NKT cells. J Immunol 178(5):2721–2729

19. Li X, Fujio M, Imamura M, Wu D, Vasan S, Wong CH, Ho DD, Tsuji M (2010) Design of a potent CD1d-binding NKT cell ligand as a vaccine adjuvant. Proc Natl Acad Sci USA 107 (29):13010–13015. doi:10.1073/pnas.1006662107

20. Tefit JN, Crabe S, Orlandini B, Nell H, Bendelac A, Deng SL, Savage PB, Teyton L, Serra V (2014) Efficacy of ABX196, a new NKT agonist, in prophylactic human vaccination. Vaccine 32(46):6138–6145. doi:10.1016/j.vaccine.2014.08.070

21. Giaccone G, Punt CJ, Ando Y, Ruijter R, Nishi N, Peters M, von Blomberg BM, Scheper RJ, van der Vliet HJ, van den Eertwegh AJ, Roelvink M, Beijnen J, Zwierzina H, Pinedo HM (2002) A phase I study of the natural killer T-cell ligand alpha-galactosylceramide (KRN7000) in patients with solid tumors. Clin Cancer Res Official J Am Assoc Cancer Res 8(12):3702–3709

22. Liu Y, Goff RD, Zhou D, Mattner J, Sullivan BA, Khurana A, Cantu C 3rd, Ravkov EV, Ibegbu CC, Altman JD, Teyton L, Bendelac A, Savage PB (2006) A modified alpha-galactosyl ceramide for staining and stimulating natural killer T cells. J Immunol Methods 312(1–2):34–39. doi:10.1016/j.jim.2006.02.009

23. Zhou XT, Forestier C, Goff RD, Li C, Teyton L, Bendelac A, Savage PB (2002) Synthesis and NKT cell stimulating properties of fluorophore- and biotin-appended 6″-amino-6″-deoxy-galactosylceramides. Org Lett 4(8):1267–1270

24. Pauwels N, Aspeslagh S, Vanhoenacker G, Sandra K, Yu ED, Zajonc DM, Elewaut D, Linclau B, Van Calenbergh S (2011) Divergent synthetic approach to 6"-modified alpha-GalCer analogues. Org Biomol Chem 9(24):8413–8421. doi:10.1039/c1ob06235b

25. Du W, Kulkarni SS, Gervay-Hague J (2007) Efficient, one-pot syntheses of biologically active alpha-linked glycolipids. Chem Commun 23:2336–2338. doi:10.1039/b702551c

26. Anderson RJ, Tang CW, Daniels NJ, Compton BJ, Hayman CM, Johnston KA, Knight DA, Gasser O, Poyntz HC, Ferguson PM, Larsen DS, Ronchese F, Painter GF, Hermans IF (2014) A self-adjuvanting vaccine induces cytotoxic T lymphocytes that suppress allergy. Nat Chem Biol 10(11):943–949. doi:10.1038/Nchembio.1640

27. Bertozzi CR (2011) A decade of bioorthogonal chemistry. Acc Chem Res 44(9):651–653. doi:10.1021/ar200193f

28. Sun XC, Zeckner DJ, Current WL, Boyer R, McMillian C, Yumibe N, Chen SH (2001) N-Acyloxymethyl carbamate linked prodrugs of pseudomycins are novel antifungal agents. Bioorg Med Chem Lett 11(14):1875–1879. doi:10.1016/S0960-894x(01)00333-X

29. Dinkel C, Moody M, Traynor-Kaplan A, Schultz C (2001) Membrane-Permeant 3-OH-Phosphorylated Phosphoinositide Derivatives. Angew Chem Int Ed Engl 40(16):3004–3008. doi:10.1002/1521-3773(20010817)40:16<3004:AID-ANIE3004>3.0.CO;2-O

30. Ulrich S, Boturyn D, Marra A, Renaudet O, Dumy P (2014) Oxime ligation: a chemoselective click-type reaction for accessing multifunctional biomolecular constructs. Chem-Eur J 20 (1):34–41. doi:10.1002/chem.201302426

31. Wills-Karp M (1999) Immunologic basis of antigen-induced airway hyperresponsiveness. Annu Rev Immunol 17:255–281. doi:10.1146/annurev.immunol.17.1.255

32. Wells JW, Cowled CJ, Giorgini A, Kemeny DM, Noble A (2007) Regulation of allergic airway inflammation by class I-restricted allergen presentation and CD8 T-cell infiltration. J Allergy Clin Immunol 119(1):226–234. doi:10.1016/j.jaci.2006.09.004

33. Perez HL, Cardarelli PM, Deshpande S, Gangwar S, Schroeder GM, Vite GD, Borzilleri RM (2014) Antibody-drug conjugates: current status and future directions. Drug Discov Today 19 (7):869–881. doi:10.1016/j.drudis.2013.11.004

34. Anderson RJ, Compton BJ, Tang CW, Authier-Hall A, Hayman CM, Swinerd GW, Kowalczyk R, Harris P, Brimble MA, Larsen DS, Gasser O, Weinkove R, Hermans IF, Painter GF (2015) NKT cell-dependent glycolipid-peptide vaccines with potent anti-tumour activity. Chem Sci 6(9):5120–5127. doi:10.1039/c4sc03599b

35. Yang M, Yang Y, Chen PR (2016) Transition-metal-catalyzed bioorthogonal cycloaddition reactions. Top Curr Chem 374(1):1–29. doi:10.1007/s41061-015-0001-3

36. Godfrey DI, Hammond KJ, Poulton LD, Smyth MJ, Baxter AG (2000) NKT cells: facts, functions and fallacies. Immunol Today 21(11):573–583

37. Weinkove R, Brooks CR, Carter JM, Hermans IF, Ronchese F (2013) Functional invariant natural killer T-cell and CD1d axis in chronic lymphocytic leukemia: implications for immunotherapy. Haematologica 98(3):376–384. doi:10.3324/haematol.2012.072835

38. Wolchok JD, Kluger H, Callahan MK, Postow MA, Rizvi NA, Lesokhin AM, Segal NH, Ariyan CE, Gordon RA, Reed K, Burke MM, Caldwell A, Kronenberg SA, Agunwamba BU, Zhang X, Lowy I, Inzunza HD, Feely W, Horak CE, Hong Q, Korman AJ, Wigginton JM, Gupta A, Sznol M (2013) Nivolumab plus ipilimumab in advanced melanoma. N Eng J Med 369(2):122–133. doi:10.1056/NEJMoa1302369

39. Beatty PL, Finn OJ (2013) Preventing cancer by targeting abnormally expressed self-antigens: MUC1 vaccines for prevention of epithelial adenocarcinomas. Ann N Y Acad Sci 1284. doi:10.1111/nyas.12108

40. van der Burg SH, Arens R, Ossendorp F, van Hall T, Melief CJ (2016) Vaccines for established cancer: overcoming the challenges posed by immune evasion. Nat Rev Cancer 16 (4):219–233. doi:10.1038/nrc.2016.16