Advances in Experimental Medicine and Biology 806

Alisa G. Woods
Costel C. Darie *Editors*

Advancements of Mass Spectrometry in Biomedical Research

Springer

Advances in Experimental Medicine and Biology

Volume 806

Editorial Board:

IRUN R. COHEN, *The Weizmann Institute of Science, Rehovot, Israel*
ABEL LAJTHA, *N.S. Kline Institute for Psychiatric Research, Orangeburg, NY, USA*
RODOLFO PAOLETTI, *University of Milan, Milan, Italy*
JOHN D. LAMBRIS, *University of Pennsylvania, Philadelphia, PA, USA*

For further volumes:
http://www.springer.com/series/5584

Alisa G. Woods • Costel C. Darie
Editors

Advancements of Mass Spectrometry in Biomedical Research

Springer

Editors
Alisa G. Woods
Biochemistry and Proteomics Group
Department of Chemistry and
Biomolecular Science
Clarkson University
Potsdam, NY, USA

Costel C. Darie
Biochemistry and Proteomics Group
Department of Chemistry and
Biomolecular Science
Clarkson University
Potsdam, NY, USA

ISSN 0065-2598 ISSN 2214-8019 (electronic)
ISBN 978-3-319-06067-5 ISBN 978-3-319-06068-2 (eBook)
DOI 10.1007/978-3-319-06068-2
Springer Cham Heidelberg New York Dordrecht London

Library of Congress Control Number: 2014939934

© Springer International Publishing Switzerland 2014
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

To Professor Dr. Vlad Artenie
(University of Iasi, Romania)
and the late Professor Dr. Wolfgang Haehnel
(University of Freiburg, Germany),
as well as the many supporters
of the Darie Lab. Also to Kenneth R. Woods, PhD, an inspirational biochemist

Foreword

The use of mass spectrometry is becoming increasingly important for biomedical research and for clinical applications. Detection of exogenous substances, such as toxins, can be performed by mass spectrometry and was perhaps the initial use of this technology in biomedicine. However, identification of endogenous, disease-related molecules is also possible. With the advent of the genetic revolution, the proteomic revolution has followed in close succession. Mass spectrometers are essential for proteomic discovery, and other "omic" fields (such as glycomics, metabolomics, lipidomics, and many more) are exploding with new information. The sensitivity of mass spectrometers and increasingly more sophisticated bioinformatics tools are opening up this field to untold possibilities for biomedical researchers and clinicians.

In the spirit of this revolution in biomedicine, we have assembled this comprehensive work, which largely focuses on the application of mass spectrometry to "omics" analysis in biomedicine. We start with broad descriptions of the field, definitions of the machinery, and then delve into the various methods and approaches that can be utilized. We consider aspects of molecular analysis and then discuss how mass spectrometry can be applied to our understanding of specific diseases and disorders.

Numerous MS-based methodologies are now available to researchers, and this text reviews many cutting-edge and relevant technologies. Dudley focuses on the application of matrix-assisted laser desorption/ionization (MALDI) in biomedicine, while colleagues from Waters Corporation examine how peak capacity can be best maximized. Brown describes quantitative shotgun proteomics with data-independent acquisition and traveling wave ion mobility spectrometry as a versatile life-science tool, while Hoedt and coworkers explore the use of stable isotope labeling by amino acids in cell culture (SILAC) for protein quantification using mass spectrometry. A chapter by Roy and coworkers discusses the complementary use of computational structural biology in mass spectrometry.

Different aspects of proteins and other molecules can be studied using MS. Petre explores protein structures and interactions, while Zamfir has tackled the use of mass spectrometry to understand gangliosides. Ngounou Wetie et al. examine the

analysis of protein posttranslational modifications and protein–protein interactions, and Samways focuses on mass spectrometry-based analysis of ion channel structure and function. Budayeva and Cristea further expand upon how protein–protein interactions can be understood via mass spectrometry, while Ferguson et al. focus on its use for understanding small molecules. Ckless elucidates how redox proteomics can be studied in the lab, for eventual application to clinical uses, and Luck looks at fluorinated proteins. Baral et al. explore mass spectrometry analysis using a model animal system, the zebrafish, while Miller and Spellman describe the workflow for biomarker discovery in a pharmaceutical company.

Focusing on health promotion, Andrei and colleagues examine how mass spectrometry can be used to analyze biomedically relevant stilbenes from wine. With regard to increased understanding of diseases and disorders, chapters by Monien, Schneider, Patel, and Sandu describe the use of mass spectrometry for analysis in cancer. Topics covered include quantifying DNA adducts, analysis of breast milk, apoptosis and cancer secretome analysis as identification of heat shock response. In the field of infectious disease, Branza-Nichita and colleagues use mass spectrometry to study the HBV life cycle, while Marrakchi et al. look at oxidative stress and antibiotic resistance in bacterial pathogens.

Deinhardt examines how mass spectrometry can be used to understand neuronal signaling, which could apply to numerous neurological and psychiatric conditions, while Woods et al. study how mass spectrometry can facilitate the understanding of a novel central nervous system protein. We then explore more applied uses of mass spectrometry in the central nervous system: specifically, how biomarker discovery may be directly performed for neurodevelopmental disorders and be used to understand and potentially diagnose depression.

In the realm of therapeutics, Heckman and coworkers have used mass spectrometry to localize and analyze the efficacy of nanoceria, a potential delivery system for a variety of medical conditions. We finally end with a particularly important chapter on bottlenecks in proteomics, topics that are encountered by almost all researchers but that are almost never discussed in publications.

We thus present to the reader a comprehensive text, examining the many uses of mass spectrometry in biomedicine, with the hope that this will be useful to both researchers and clinicians. As this exciting field further expands, so will the potential applications for using mass spectrometry to understand medical issues and to address them through exploration, as well as eventual clinical prognosis, diagnosis, and monitoring. We look forward to an exciting era of MS-based discovery and application.

Potsdam, NY

Alisa G. Woods
Costel C. Darie

Contents

Contributors

Silvana Andreescu Department of Chemistry and Biomolecular Science, Clarkson University, Potsdam, NY, USA

Veronica Andrei International Centre of Biodynamics, Bucharest, Romania

Kathleen F. Arcaro Department of Veterinary and Animal Sciences, University of Massachusetts Amherst, Amherst, MA, USA

Roshanak Aslebagh Biochemistry and Proteomics Group, Department of Chemistry and Biomolecular Science, Clarkson University, Potsdam, NY, USA

Reshica Baral Department of Biology, Clarkson University, Potsdam, NY, USA

Heather L. Bowling Skirball Institute of Biomolecular Medicine, New York University School of Medicine, New York, NY, USA

Norica Branza-Nichita Department of Viral Glycoproteins, Institute of Biochemistry of the Romanian Academy, Bucharest, Romania

Department of Analytical Chemistry and Environmental Engineering, Faculty of Applied Chemistry and Materials Science, Politehnica University of Bucharest, Bucharest, Romania

Lewis M. Brown Department of Biological Sciences, Quantitative Proteomics Center, Columbia University, New York, NY, USA

Hanna G. Budayeva Department of Molecular Biology, Princeton University, Princeton, NJ, USA

Karina Ckless Department of Chemistry, State University of New York at Plattsburgh, Plattsburgh, NY, USA

Bayard D. Clarkson Molecular Pharmacology and Chemistry Program, Memorial Sloan Kettering Cancer Center, New York, NY, USA

Joanne B. Connolly Waters Corporation, Manchester, UK

Ileana M. Cristea Department of Molecular Biology, Princeton University, Princeton, NJ, USA

Costel C. Darie Biochemistry and Proteomics Group, Department of Chemistry and Biomolecular Science, Clarkson University, Potsdam, NY, USA

Katrin Deinhardt Institute for Life Sciences and Centre for Biological Sciences, University of Southampton, Southampton, UK

Robert M. Dekroon Lineberger Comprehensive Cancer Center, University of North Carolina, Chapel Hill, NC, USA

Ed Dudley Institute of Mass Spectrometry, College of Medicine, Swansea University, Swansea, UK

Joseph Erlichman Department of Biology, St. Lawrence University, Canton, NY, USA

Carly N. Ferguson Department of Chemistry and Biochemistry, University of California-Los Angeles, Los Angeles, CA, USA

Joseph W.M. Fowler Department of Chemistry and Biochemistry, University of California-Los Angeles, Los Angeles, CA, USA

Richard A. Gatti Department of Pathology and Laboratory Medicine, David Geffen School of Medicine, University of California-Los Angeles, Los Angeles, CA, USA

Department of Human Genetics, David Geffen School of Medicine, University of California-Los Angeles, Los Angeles, CA, USA

Lee A. Gethings Waters Corporation, Manchester, UK

Karin L. Heckman Department of Biology, St. Lawrence University, Canton, NY, USA

Esthelle Hoedt Kimmel Center for Biology and Medicine at the Skirball Institute and Department of Biochemistry and Molecular Pharmacology, New York University School of Medicine, New York, NY, USA

Dan V. Iosifescu Mood and Anxiety Disorders Program, Mount Sinai School of Medicine, New York, NY, USA

Catalin Lazar Department of Viral Glycoproteins, Institute of Biochemistry of the Romanian Academy, Bucharest, Romania

Xiaobo Liu Department of Chemistry and Biomolecular Science, Clarkson University, Potsdam, NY, USA

Joseph A. Loo Departments of Chemistry and Biochemistry, University of California-Los Angeles, Los Angeles, CA, USA

Department of Biological Chemistry, David Geffen School of Medicine, University of California-Los Angeles, Los Angeles, CA, USA

Linda A. Luck Department of Chemistry, State University of New York at Plattsburgh, Plattsburgh, NY, USA

Mouna Marrakchi Department of Chemistry and Biomolecular Science, Clarkson University, Potsdam, NY, USA

Laboratory of Ecology and Microbial Technology, National Institute of Applied Sciences, Carthage University, Charguia Cedex, Tunisia

Iuliana Mihai International Centre of Biodynamics, Bucharest, Romania

Ronald A. Miller Molecular Biomarkers Department, Merck Research Laboratories, West Point, PA, USA

Mihaela Mocanu Lineberger Comprehensive Cancer Center, University of North Carolina, Chapel Hill, NC, USA

Bernhard H. Monien Lead, Genotoxic Food Contaminants Research Group, German Institute of Human Nutrition (DIfE) Potsdam Rehbrücke, Nuthetal, Germany

Thomas A. Neubert Kimmel Center for Biology and Medicine at the Skirball Institute and Department of Biochemistry and Molecular Pharmacology, New York University School of Medicine, New York, NY, USA

Armand G. Ngounou Wetie Biochemistry and Proteomics Group, Department of Chemistry and Biomolecular Science, Clarkson University, Potsdam, NY, USA

Sapan Patel Molecular Pharmacology and Chemistry Program, Memorial Sloan Kettering Cancer Center, New York, NY, USA

Biochemistry and Proteomics Group, Department of Chemistry and Biomolecular Science, Clarkson University, Postdam, NY, USA

Catalina Petrareanu Department of Viral Glycoproteins, Institute of Biochemistry of the Romanian Academy, Bucharest, Romania

Department of Analytical Chemistry and Environmental Engineering, Faculty of Applied Chemistry and Materials Science, Politehnica University of Bucharest, Bucharest, Romania

Brindusa Alina Petre Laboratory of Biochemistry, Department of Chemistry, Al. I. Cuza University of Iasi, Iasi, Romania

Laboratory of Analytical Chemistry and Biopolymer Structure Analysis, Department of Chemistry, University of Konstanz, Konstanz, Germany

Ken Reed Cerion NRx LLC, Rochester, NY, USA

Urmi Roy Structural Biology and Molecular Modeling Unit, Biochemistry and Proteomics Group, Department of Chemistry and Biomolecular Science, Clarkson University, Potsdam, NY, USA

Jeanne P. Ryan SUNY Plattsburgh Neuropsychology Clinic and Psychoeducation Services, Plattsburgh, NY, USA

Damien S.K. Samways Department of Biology, Clarkson University, Potsdam, NY, USA

Cristinel Sandu Strang Laboratory of Apoptosis and Cancer Biology, Howard Hughes Medical Institute, The Rockefeller University, New York, NY, USA

Sallie S. Schneider Pioneer Valley Life Sciences Institute, Springfield, MA, USA

Devon A. Shipp Polymer Science and Materials Group, Department of Chemistry and Biomolecular Science, Clarkson University, Potsdam, NY, USA

Matthew Skeels Department of Chemistry, St. Lawrence University, Canton, NY, USA

Izabela Sokolowska Biochemistry and Proteomics Group, Department of Chemistry and Biomolecular Science, Clarkson University, Potsdam, NY, USA

Daniel S. Spellman Molecular Biomarkers Department, Merck Research Laboratories, West Point, PA, USA

Hermann Steller Strang Laboratory of Apoptosis and Cancer Biology, Howard Hughes Medical Institute, The Rockefeller University, New York, NY, USA

Susan R. Sturgeon Division of Biostatistics and Epidemiology, School of Public Health and Health Sciences, University of Massachusetts Amherst, Amherst, MA, USA

Alina Vasilescu International Centre of Biodynamics, Bucharest, Romania

Kenneth N. Wallace Department of Biology, Clarkson University, Potsdam, NY, USA

Jonathan F. Waxer Department of Chemistry and Biochemistry, University of California-Los Angeles, Los Angeles, CA, USA

Alisa G. Woods Biochemistry and Proteomics Group, Department of Chemistry and Biomolecular Science, Clarkson University, Potsdam, NY, USA

Kelly Wormwood Biochemistry and Proteomics Group, Department of Chemistry and Biomolecular Science, Clarkson University, Potsdam, NY, USA

Alina D. Zamfir Faculty of Physics, West University of Timisoara, Timisoara, Romania

National Institute for Research and Development in Electrochemistry and Condensed Matter, Timisoara, Romania

Guoan Zhang Kimmel Center for Biology and Medicine at the Skirball Institute and Department of Biochemistry and Molecular Pharmacology, New York University School of Medicine, New York, NY, USA

Chapter 1
Mass Spectrometry for Proteomics-Based Investigation

Alisa G. Woods, Izabela Sokolowska, Armand G. Ngounou Wetie, Kelly Wormwood, Roshanak Aslebagh, Sapan Patel, and Costel C. Darie

Abstract Within the past years, we have witnessed a great improvement in mass spectrometry (MS) and proteomics approaches in terms of instrumentation, protein fractionation, and bioinformatics. With the current technology, protein identification alone is no longer sufficient. Both scientists and clinicians want not only to identify proteins but also to identify the protein's posttranslational modifications (PTMs), protein isoforms, protein truncation, protein–protein interaction (PPI), and protein quantitation. Here, we describe the principle of MS and proteomics and strategies to identify proteins, protein's PTMs, protein isoforms, protein truncation, PPIs, and protein quantitation. We also discuss the strengths and weaknesses within this field. Finally, in our concluding remarks we assess the role of mass spectrometry and proteomics in scientific and clinical settings in the near future. This chapter provides an introduction and overview for subsequent chapters that will discuss specific MS proteomic methodologies and their application to specific medical conditions. Other chapters will also touch upon areas that expand beyond proteomics, such as lipidomics and metabolomics.

A.G. Woods • I. Sokolowska • A.G. Ngounou Wetie • K. Wormwood
R. Aslebagh • C.C. Darie (✉)
Biochemistry & Proteomics Group, Department of Chemistry & Biomolecular Science, Clarkson University, 8 Clarkson Avenue, Potsdam, NY 13699-5810, USA
e-mail: cdarie@clarkson.edu

S. Patel
Biochemistry & Proteomics Group, Department of Chemistry & Biomolecular Science, Clarkson University, 8 Clarkson Avenue, Potsdam, NY 13699-5810, USA

Memorial Sloan Kettering Cancer Center, Molecular Pharmacology and Chemistry Program, New York, NY USA

A.G. Woods and C.C. Darie (eds.), *Advancements of Mass Spectrometry in Biomedical Research*, Advances in Experimental Medicine and Biology 806,
DOI 10.1007/978-3-319-06068-2_1, © Springer International Publishing Switzerland 2014

Abbreviations

BN-PAGE	Blue native PAGE
CI	Chemical ionization
CN-PAGE	Colorless native PAGE
DIGE	Differential gel electrophoresis
EI	Electron ionization
ESI	Electrospray ionization
ESI-MS	Electrospray ionization mass spectrometry
FT	Fourier transform
IT	Ion trap
LC–MS/MS	Liquid chromatography–mass spectrometry
m/*z*	Mass/charge
MALDI	Matrix-assisted laser desorption ionization
MALDI-MS	MALDI mass spectrometry
MS	Mass spectrometry
Mw	Molecular weight
Q	Quadrupole
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
TIC	Total ion current/chromatogram
TOF	Time of flight

1.1 Introduction

Proteomics is the large scale study of the protein complement, also known as the proteome. Proteomics is studied through mass spectrometry (MS) [1–8]. MS can be used to investigate a large variety of chemical and biological molecules, including products of chemical synthesis or degradation, biological molecules such as proteins, nucleic acids, lipids, or glycans, or various natural compounds of either large or small molecular mass. Depending on what type of molecule is being analyzed, there are various types of MS focus, such as small-molecule MS, large-molecule MS, and biological MS (when the molecules investigated are biomolecules). Within biological MS, there are also different MS subfields, such as proteomics, lipidomics, glycomics, and metabolomics. The focus of proteomics is to analyze proteins and protein derivatives (such as glycoproteins), peptides, posttranslational modifications (PTMs) within proteins, or protein–protein interactions (PPIs).

The standard workflow in a proteomics experiment starts with sample fractionation, involving the separation of proteins prior to their analysis by MS [9–17]. This can be done by one or more biochemical fractionation methods. For example, a one-dimensional separation can be achieved by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE); a two-dimensional separation can be performed by two-dimensional electrophoresis or by affinity purification followed by

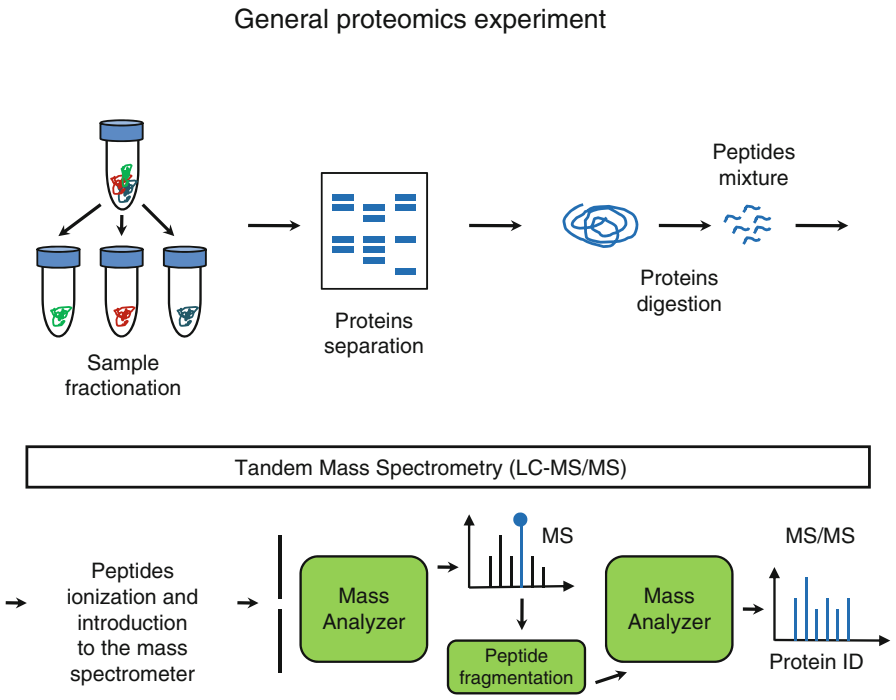


Fig. 1.1 General proteomic experiment workflow schematic. Reprinted and adapted with permission from the *Australian Journal of Chemistry CSIRO Publishing* http://www.publish.csiro.au/?paper=CH13137 [15]

SDS-PAGE. Biochemical fractionation is then followed by enzymatic digestion (usually trypsin), peptide extraction, and peptide fractionation by HPLC and MS analysis [1]. Data analysis leads to identification of one or more proteins and further simultaneous investigation or re-investigation of the results can extract additional information from the same MS experiment, such as PTMs and interaction partners of some proteins (PPIs) [18–26]. A schematic of a proteomics workflow is shown in Fig. 1.1 and a schematic of a proteomics experiment is shown in Fig. 1.2a.

Proteomic analysis can be performed using samples from various sources such as supracellular, subcellular, intracellular, or extracellular, as well as at the peptide level (peptidomics), protein (regular proteomics), PTMs ("PTM-omics"), or protein complex level (interactomics). Proteomics can also be classified as classical or functional, when one analyzes protein samples from two different conditions (for example, normal and cancer), and targeted proteomics, when one focuses on a particular sub-proteome, such as phosphoproteomics or glycoproteomics. Proteomics can also be classified based on the protein complement from a set of samples that is being analyzed such as proteomes (i.e., all proteins) or sub-proteomes (i.e., just the nuclei or mitochondria). A schematic of such classification is shown in Fig. 1.2b.

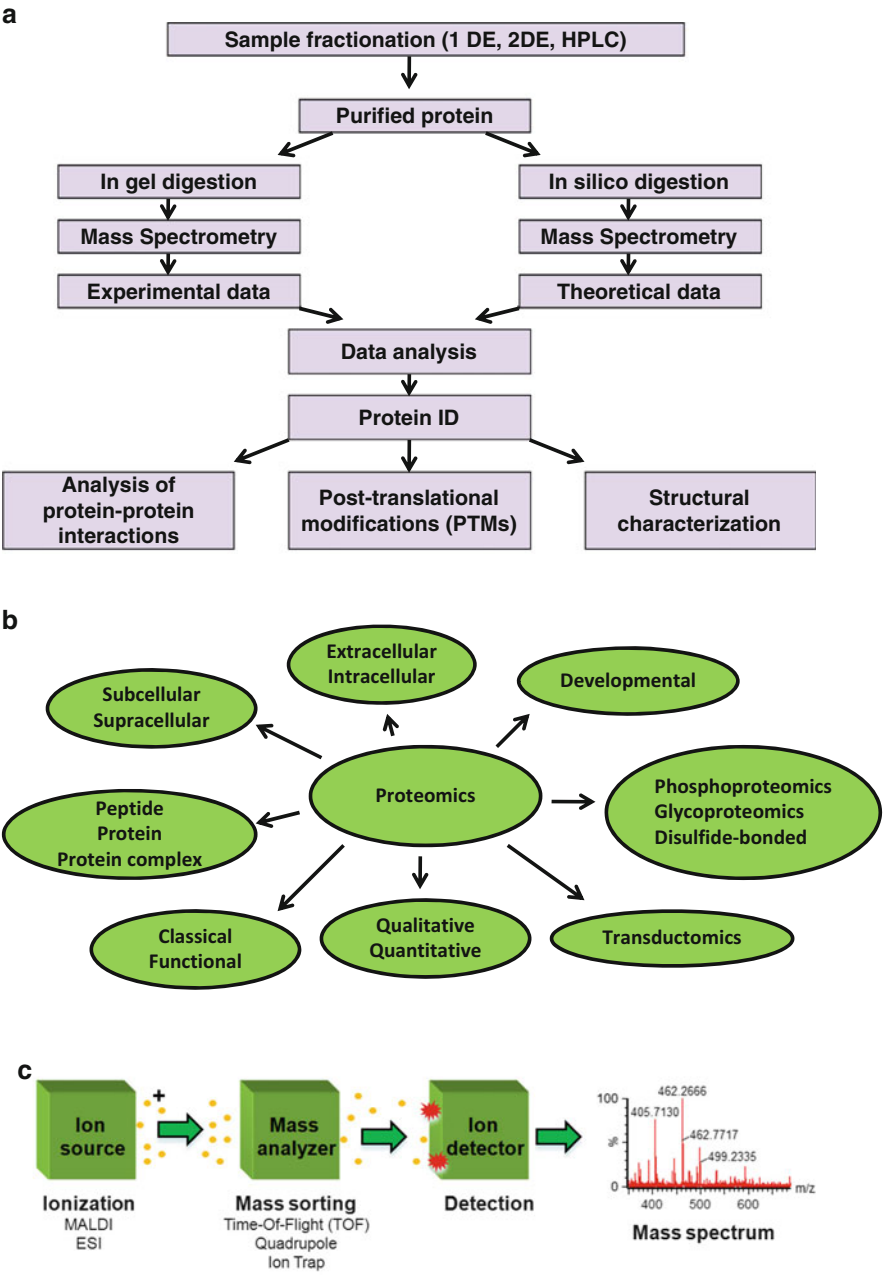


Fig. 1.2 General proteomics experiment. (**a**) Proteomics experiment workflow schematic. (**b**) Proteomics and applications schematic. (**c**) Mass spectrometer schematic. Reprinted and adapted with permission from the *Oxidative Stress*: *Diagnostics*, *Prevention*, *and Therapy*, S. Andreescu and M. Hepel, Editors. 2011, American Chemical Society: Washington, D.C [16]

Proteomic analysis can also focus on quality such as for protein identification, or the determination of protein amounts by quantitative proteomics. These analyses are usually performed using a mass spectrometer, the "workhorse" in a proteomics experiment. A mass spectrometer has three main components: the ionization source, a mass analyzer, and a detector (Fig. 1.2c). There are primarily two types of ionization sources on mass spectrometers: matrix-assisted laser desorption ionization (MALDI) and electrospray ionization (ESI). The mass spectrometers are consequently named MALDI mass spectrometry (MALDI-MS) and electrospray ionization mass spectrometry (ESI-MS). Here, we describe a proteomics experiment, specifically how proteins and peptides are analyzed by MS. We also describe the type of information that can be obtained from such an experiment.

1.2 Biochemical Fractionation

The first step in a proteomics experiment is biochemical fractionation, in which various proteins are separated from each other using their physicochemical properties. Biochemical fractionation usually depends on the goal of the experiment and it is perhaps the most important step in a proteomics experiment. A good sample fractionation usually leads to a good experimental outcome. A proteomics experiment can still be performed without biochemical fractionation, for example, when one analyzes the full proteome of a cell at once. However, without biochemical fractionation, the results in a proteomics experiment may not necessarily be optimal.

The physicochemical properties of proteins (or compounds of interest) that are used to achieve biochemical fractionation are, among others, molecular mass, isoelectric point, charge at various pH, and the protein's affinity to other compounds. These properties of the proteins are well exploited by biochemical fractionations such as electrophoresis, centrifugation, and chromatography. Types of chromatography can include affinity chromatography, ion exchange chromatography, and size-exclusion chromatography.

To give one example, proteins can be separated by electrophoresis, usually SDS-PAGE, reduced and denatured, and then separated according to their molecular mass. If the reduction step is not used, the disulfide bridges in a protein or between proteins remain intact, thus providing an additional fractionation principle: two proteins with low molecular mass (such as haptoglobin subunits) are kept together through disulfide bridges and are separated under SDS-PAGE under nonreducing conditions as a heterotetramer with a high molecular mass. In a different variant of SDS-PAGE, but not using the detergent (SDS), one may separate proteins under native conditions. Therefore, simply by adding one reagent (for example, SDS) or two (SDS and a reducing agent like dithiothreitol or DTT), separation of these proteins may have a totally different outcome. A variant of SDS-PAGE is tricine-PAGE [27, 28], which has a principle of separation similar to the SDS-PAGE, but it has the highest separating resolution in the low molecular weight (Mw) proteins and peptides (2–20 kDa), where SDS-PAGE has poor or very poor resolution. Therefore, SDS-PAGE and tricine-PAGE complement each other.

Other types of electrophoresis are blue native PAGE (BN-PAGE), colorless native PAGE (CN-PAGE), and detergent-less SDS-PAGE (native PAGE) [1, 4, 6, 18–22, 29–34], all native electrophoresis. BN-PAGE separates protein complexes by using the external charge induced by Coomassie dye; thus, the complexes will have the same charge and will separate according to their molecular weight. If the Coomassie dye is not used, the external charge is not induced and the separation does not take place according to the molecular weight of the complexes, but rather according to the internal charge of the protein complexes. This method, a variant on BN-PAGE, is named CN-PAGE. CN-PAGE is particularly useful when two protein complexes with identical mass must be separated from each other.

In addition to the techniques mentioned for biochemical fractionation, hyphenated techniques may also be used. The classical example is two-dimensional electrophoresis (2D-PAGE), which includes separation of proteins by isoelectric focusing and by SDS-PAGE [3, 7, 35–45], still used in some proteomics labs. In fact, a variant of 2D-PAGE is differential gel electrophoresis (DIGE), a powerful method for gel-based proteomics. Other fractionation methods such as pre-coated chips, centrifugal filters, and magnetic beads are also possible [46, 47].

1.3 Mass Spectrometry

A mass spectrometer has three main parts: an ion source, a mass analyzer, and a detector. Initially, the sample is ionized and the ions produced by MALDI or ESI source are separated in the mass analyzer based on their mass-to-charge (m/z) ratio. The ions are then detected by the detector. The end product is a mass spectrum, which is a plot of ion abundance versus m/z.

Ionization sources. Ionization of peptides is dependent on the electrical potential at the ion source and on the pH at which they are analyzed. At low pH, the peptides are protonated through the amino-containing amino acids such as Arg or Lys, while at high pH, the peptides are de-protonated through the carboxyl-containing amino acids such as Asp or Glu. When the electrical potential at the ion source is positive, ionization is in positive ion mode. Conversely, when the electrical potential is negative, ionization is in negative ion mode. Therefore, there are two types of ionization: positive, when peptides are analyzed at low pH and the Arg, Lys, and His are protonated, and negative ionization, when peptides are analyzed at high pH and the Asp and Glu are de-protonated. In the current chapter, we will focus only on positive ionization, because it is one of the most used ionization modes for analyzing peptides and proteins. In addition, the enzyme that is the most widely used in proteomics is trypsin which cleaves conveniently at the C-terminus of Arg and Lys and produces peptides that are, upon ionization, at least doubly charged (the peptide and the C-terminal amino acid) and produces a y product ion series upon collision-induced fragmentation (described later).

In addition to ESI and MALDI, there are several additional ionization methods, such as chemical ionization (CI), electron ionization (EI), or atmospheric pressure

chemical ionization (APCI) [48, 49]. EI is used for analysis of organic compounds and can be used for all volatile compounds with a mass smaller than 1,000 Da. EI provides good structural information derived from fragmentation. However, molecular mass determination is rather poor (poor signal or the absence of M^+ ions) [50]. Chemical ionization is the opposite: it is very good for the determination of the molecular mass of molecules, but it is not very good in providing structural information due to reduced fragmentation in comparison to EI. Therefore CI and EI could complement each other. In CI experiments, ionized species are formed when the gaseous molecules to be analyzed collide with primary ions present in the source under a high vacuum [51]. A variant of CI is negative CI used only for volatile analytes with a mass of less than 1,000 Da [52, 53]. Another ionization technique, APCI, is an alternative for analysis of compounds that do not ionize in ESI. During APCI, generally only singly charged ions are formed and it is usually applied to compounds with a molecular weight of less than 1,500 Da [54].

Mass analyzers. There are three main types of mass analyzers used for proteomics experiments: trapping type instruments (quadrupole ion trap—QIT, linear ion trap—LIT, Fourier transform ion cyclotron resonance—FT-ICR, and Orbitrap), quadrupole (Q), and time of flight (TOF) instruments.

Trapping type instruments first accumulate ions and then allow for mass measurement. The ion trap analyzers first capture ions in three-dimensional space (trap), and then electrostatic gate pulses to inject ions into the ion trap. The ion trap-based analyzers are relatively inexpensive, sensitive, and robust. They have been extensively used in proteomic analysis. However, a problem with these instruments is their accuracy for both precursor and product ions, partially overcome by an FT-ICR. Unfortunately, this instrument is not very often used in proteomics research because peptides do not fragment well and the instrument is expensive [55, 56].

In quadrupole mass analyzers, ions constantly enter the analyzers, which are separated based on their trajectory in the electric field applied to two pairs of charged cylindrical rods. There is an electric potential between each pair of rods drawing the ions towards one rod. These instruments provide good reproducibility and low cost, but their resolution and accuracy are limited [49, 57].

Instruments with TOF mass analyzers are popular for sample analysis in proteomics due to their high resolution and relatively low cost, speed of measurements, and high mass accuracy [49, 57]. In TOF mass analyzers, ions are accelerated by a known electric field and then travel from the ion source to the detector. The instrument measures the time it takes for ions with different masses to travel from the ion source to detector,

Mass spectrometers can have stand-alone analyzers or in combination, usually two or three analyzers within one instrument, thus taking advantage of the strength of all combined analyzers simultaneously. Examples of such instruments are Q-Trap, QQQ, Q-TOF, TOF–TOF, QQ-LIT; these instruments are also called hybrid mass spectrometers, and are highly sensitive and also have a high resolution [1, 57–59].

MS detectors. The MS detectors are usually electron multipliers, photodiode arrays, microchannel plates, or image current detectors.

1.4 MALDI-TOF MS

MALDI-TOF MS or MALDI-MS (Fig. 1.3a) is mostly used for determination of the mass of a peptide or protein and for identification of a protein using peptide mass fingerprinting. In MALDI-MS, the peptide mixture is co-crystallized under acidic conditions with a UV-absorbing matrix (for example, dihydrobenzoic acid, sinapinic acid, alpha-hydroxycinnamic acid) and spotted on a plate. A laser beam (usually nitrogen; 337 nm) then ionizes the matrix and peptides, which desorb and start to fly under an electrical field. The matrix molecules transfer a proton to peptides, which then become ionized, fly through the TOF tube, and are detected in the detector as a mass spectrum. Charged peptides fly through the mass analyzer as ions according to their mass-to-charge ratio (m/z) and to the formula: $[M+zH]/z$, where M is the mass of the peptide and z is the charge of the peptide; H is the mass of hydrogen (1.007825035 atomic mass units). In MALDI-MS analysis, the charge of peptides is almost always +1 and the peptides are mostly observed as singly charged; the formula is then $[M+1\times1]/1$ or $[M+1]/1$ or $[M+1]$. Therefore, the peptides are mostly detected as singly charged peaks or [MH]+ peaks (Fig. 1.3b).

In the MALDI-MS mass spectrum, one peak corresponds to one peptide and many peaks correspond to many peptides, either from one protein or from more proteins. Database search of the MALDI-MS spectra usually identifies that single protein or those proteins through a process named peptide mass fingerprinting (Fig. 1.3c).

1.5 ESI-MS

In contrast to MALDI-MS, in which peptides are ionized with the help of a matrix (and are in the solid phase), in ESI-MS (Fig. 1.4a) peptides are ionized in the liquid phase, under high electrical current. Also, while in MALDI-MS peptides are mostly singly charged, in ESI-MS peptides are mostly double or multiple charged. Regarding the ionization method, peptides fly as ions according to m/z and calculation of the molecular mass of the peptide is performed according to the same $[M+z]/z$ formula, where z is again the charge (z is 2 for doubly charged peptides, 3 for triply charged peptides, etc.).

When a peptide mixture is injected into the mass spectrometer, all or most peptides that ionize under the experimental conditions are detected as ions in an MS

Fig. 1.3 (continued) peptide mixture is analyzed by MALDI-MS and a spectrum is collected. A similar experiment is performed in silico (a theoretical experiment in computer), but the cleavage is performed in all proteins from a database. During the database search, the best match between the theoretical and the experimental spectra then lead to identification of a protein. Reprinted and adapted with permission from the *Oxidative Stress*: *Diagnostics*, *Prevention*, *and Therapy*, S. Andreescu and M. Hepel, Editors. 2011, American Chemical Society: Washington, D.C [16]

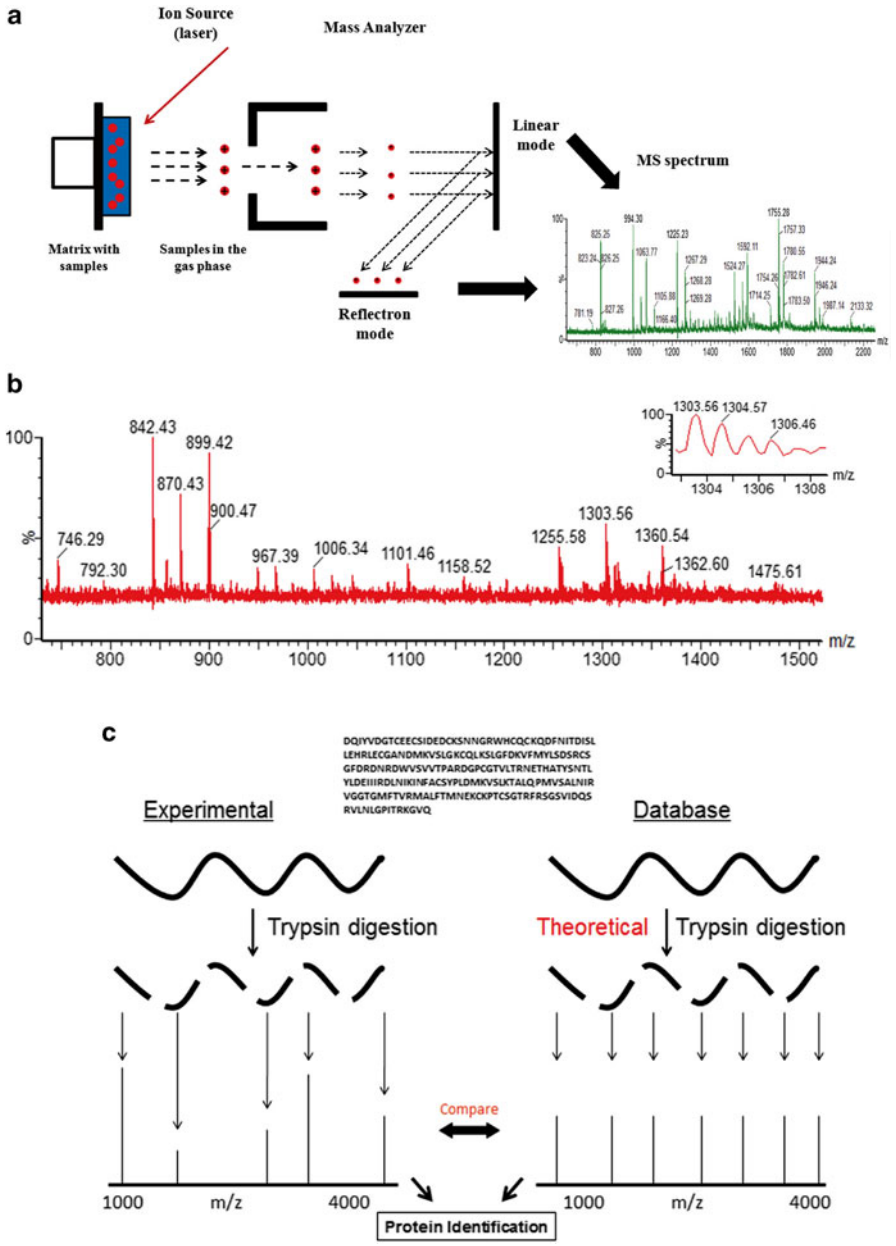


Fig. 1.3 MALDI-TOF MS. (**a**) MALDI-TOF mass spectrometer principle. An ion source, a mass analyzer, and detector are present on the instrument. At the detector the mass spectrum is detected/recorded. The mass analyzer is a TOF and can be used in linear mode or reflective mode. (**b**) A MALDI-MS spectrum primarily contains singly charged peaks; one example is shown (enlarged) to reveal the peak's charged state (single charged or +1). (**c**) Protein identification via MALDI-MS and peptide mass fingerprinting (PMF). A protein is digested into peptides using trypsin and the

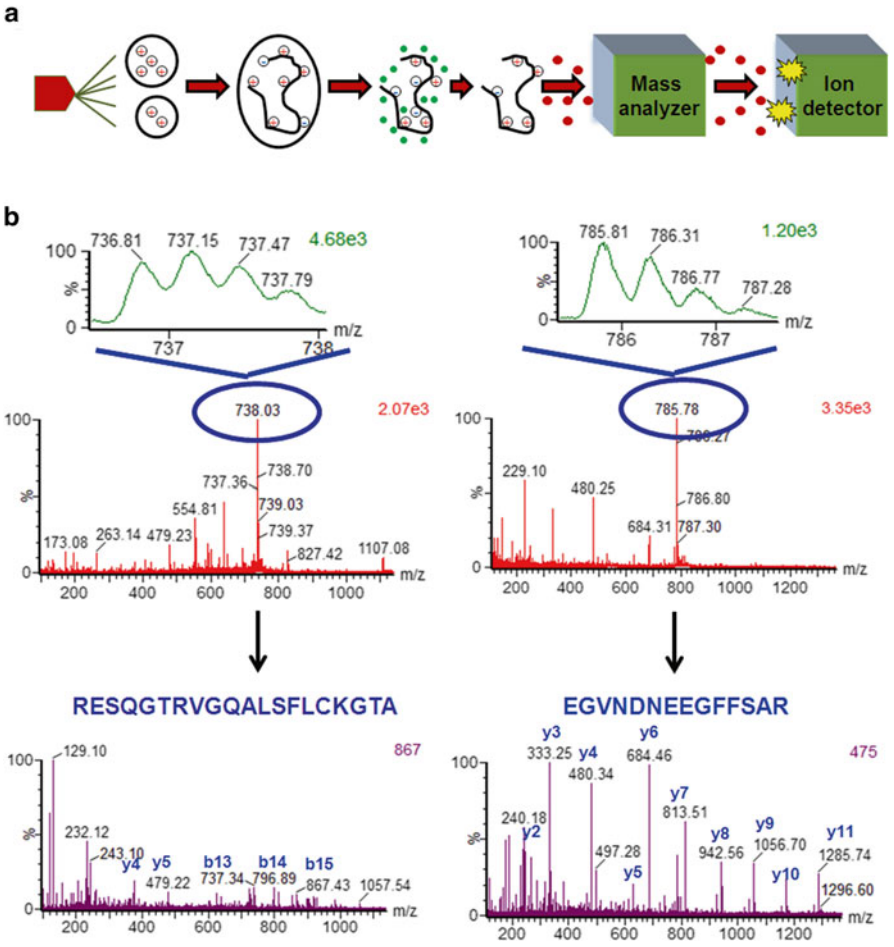


Fig. 1.4 ESI-MS of peptides. (**a**) An ESI-MS mass spectrometer. The ESI-MS has an ion source, in which the ions are ionized, a mass analyzer that ions travel through, as well as an ion detector, which records the mass spectrum. In ESI-MS, the sample is liquid, under high temperature and high electric current. The sample dehydrates and becomes protonated for positive ionization. (**b**) TOF MS spectra example, in which two different peaks, one triply charged peak with *m/z* of 736.81 (*left*) and one double charged with *m/z* of 785.81 (*right*, both circled and zoomed in), are selected for fragmentation and produce the MS/MS spectra whose data analysis led to identification of peptides with the amino acid sequence RESQGTRVGQALSFCKGTA (*left*) and EGVNDNEEGFFSAR (*right*). Note that when the protonation site (R) is on the N-terminus of the peptide, the quality of the MS/MS spectrum is not great and analysis of the b and y ions produced by the MS/MS fragmentation is difficult to interpret. However, when the protonation site is on the C-terminus of the peptide, the fragmentation produces a nice y ion series and the analysis of these ions can easily identify the amino acid sequence of the peptide. Reprinted and adapted with permission from the *Oxidative Stress*: *Diagnostics*, *Prevention*, *and Therapy*, S. Andreescu and M. Hepel, Editors. 2011, American Chemical Society: Washington, D.C [16]

spectrum in a process called direct infusion (ESI-MS mode). For example, if one has 10 peptides in an Eppendorf tube, one can identify all 10 peptides in one spectrum. However, in the MS one identifies only the masses of the peptides. In order to identify the sequence information about one particular peptide, one must isolate one peak that corresponds to one of the 10 peptides (precursor ion), fragment it in the collision cell using a neutral gas (for example, Argon gas), and record a spectrum (a sum of spectra) of the product ions that resulted from fragmentation of the precursor ion called MS/MS (ESI-MS/MS mode). Data analysis of the MS and MS/MS spectra usually leads to identification of the mass and sequence information about the peptide of interest. Examples of ESI-MS and ESI-MS/MS spectra are shown in Fig. 1.4b. As observed, the quality of the MS/MS spectra is directly dependent on the amino acid sequence, but more important, by the position of the proton-trapping amino acid (R, H, or K, in this case, R). For example, if the proton-trapping amino acid is on the N-terminus, low intensity b and y ions are observed (Fig. 1.4b, left). However, when the proton-trapping amino acid is located on the C-terminus, the fragments produced are almost always y ions of high quality. This is also the main reason for which most proteomics experiments use trypsin as an enzyme, since it cleaves the C-termini of R and K and produces peptides with an R or a K at the C-terminus.

Sometimes, when a peptide has more than one proton accepting amino acid such as Arg or Lys, the peptide may be protonated by more than two or three protons. Therefore, the same peptide may be identified with more than two or three charges. The advantage for these peptides is that if the precursor ion in a charge state of, e.g., 2+ does not fragment well in MS/MS, then the peak that corresponds to the same peptide but in a different charge state (e.g., 3+ or 4+) may fragment very well. One drawback for the multiply charged peptides is that they are usually longer (2,500–3,000 Da) than the regular peptides analyzed by MS (800–2,500 Da) and data analysis for these peptides may be more difficult than for regular peptides. However, overall, fragmentation of more than one peak corresponding to the same peptide but with different charge states may help in obtaining additional information about that peptide.

ESI-MS can be used not only for peptides but also for investigation of proteins and the information is particularly useful for determining the molecular mass of those proteins, of their potential PTMs, and of their conformation. In addition, the high molecular mass proteins can also be analyzed by ESI-MS in either positive mode (protonated) or negative mode (de-protonated), thus providing distinct, yet complementary, information regarding the distribution of charges on the surface of the protein investigated. Examples of MS spectra of a 16.9 kDa protein investigated by ESI-MS in both positive and negative mode are shown in Fig. 1.5.

1.6 LC–MS/MS

Analysis of peptide mixtures by ESI-MS for determination of the molecular mass of the peptides is usually a quick procedure. However if one wants to investigate the sequence information of more than one peptide, it is not the method of choice,

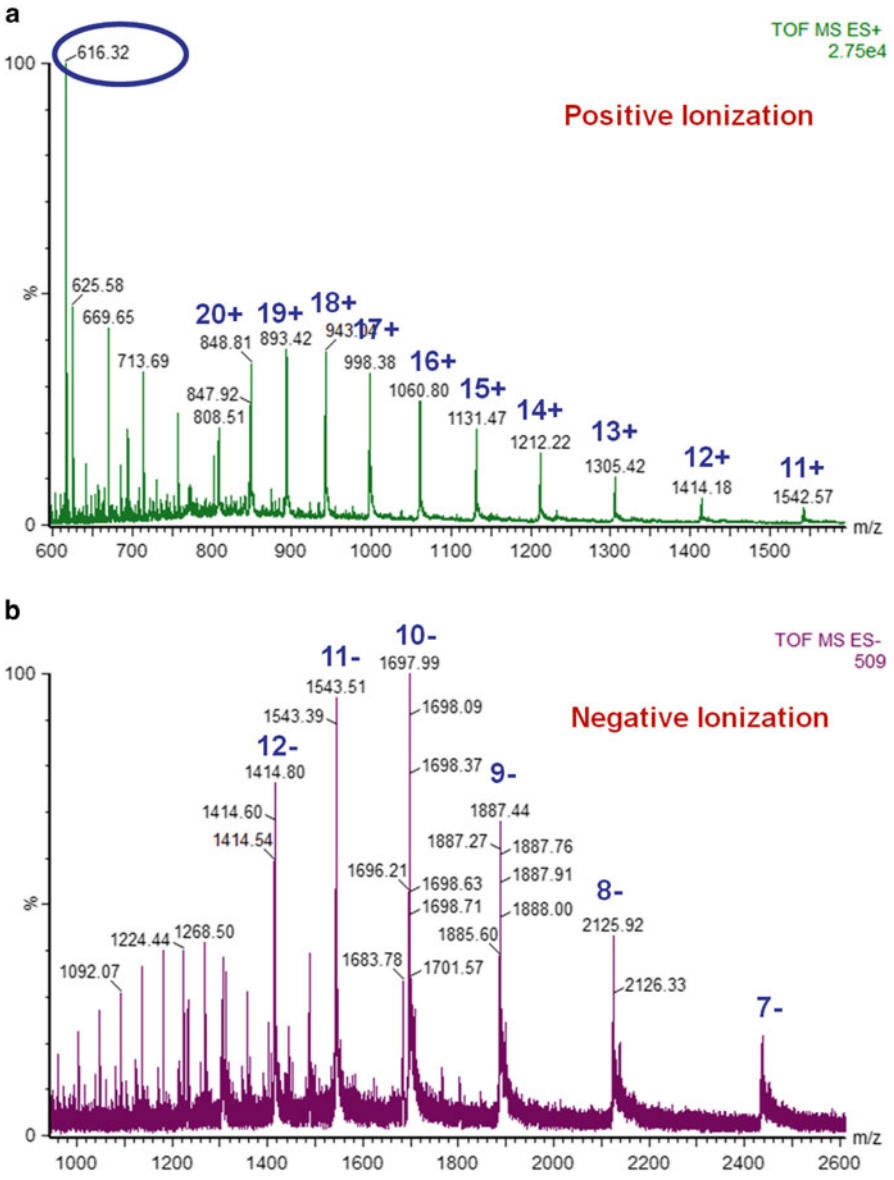


Fig. 1.5 ESI-MS proteins: ESI-MS spectra of intact 17 kDa protein, myoglobin, analyzed under acidic conditions (pH ~ 2). (**a**) MS spectrum in positive ionization; (**b**) MS spectrum analyzed in negative ionization. The positive (A) and negative (−) charges are indicated. The peak with *m*/*z* of 616.32 (1+) corresponds to the heme group, which is the prosthetic group of myoglobin. Reprinted and adapted with permission from the *Australian Journal of Chemistry CSIRO Publishing* http://www.publish.csiro.au/?paper=CH13137 [15]

since fragmentation of the ions that correspond to peptides happens manually; one peptide at the time. For example, if one has 4 peptides in a mixture, we can determine the molecular mass of all peptides in minutes, but to determine their amino acid sequence, the peptides must be selected for fragmentation one at the time. Therefore, to automate this process, an alternative approach is necessary. One option is to fractionate the peptides by column chromatography coupled to an HPLC, i.e., reversed phase-based HPLC (reversed phase columns are particularly compatible with MS). The combination of HPLC and ESI-MS is named HPLC–ESI-MS or LC–MS. In this setting, the peptides are fractionated by HPLC prior to MS analysis. They can also be selected for fragmentation and then fragmented by MS/MS. In a process called data-dependent analysis (DDA), usually 3–4 precursor peaks (which correspond to peptides) are selected for fragmentation from one MS scan and fragmented by MS/MS in a process called LC–MS/MS. In LC–MS/MS, the mass spectrometer analyzes fewer peptides per unit of time as compared with ESI-MS, simply because the HPLC fractionates the peptide mixture over a longer period of time (such as a 60 min gradient) and gives the mass spectrometer more time to analyze more peptides. A schematic of the LC–MS/MS is shown in Fig. 1.6a.

Various types of improvements can be done to increase the number of MS/MS spectra with high quality data which can lead to identification of additional proteins. One is at the flow rate of the HPLC. On a high flow rate, the mass spectrometer will have less time to analyze the peptide mixtures, as compared with lower flow rate. On a longer HPLC gradient (such as 120 min), the mass spectrometer will have more time to analyze more peptides, as compared with a shorted gradient. The number of MS/MS may also influence the number of peptides fragmented per minute. For example, a mass spectrometer has usually one MS survey followed by several MS/MS, for example, between 3 and 10 channels for MS/MS (newer instruments can be up to 30 MS/MS). If the method is set to have one MS survey scan and then to do MS/MS of the two most intense peaks, then the instrument will work as follows: one second MS survey, one second MS/MS (Peak 1), one second MS/MS (Peak 2), and then again one second MS survey (Fig. 1.6a).

Assuming that a mass spectrometer has a cycle of one MS and two MS/MS (such as 0.1 s for an MS survey followed by selection of two precursor peaks for fragmentation by MS/MS; 3 s per MS/MS), this means that in 1 min, the MS instrument can perform ~30 MS/MS that can lead to identification of ~15 proteins. In a 120 min gradient, the possible number of proteins that can be identified is $\sim 15 \times 120 = 1{,}800$ proteins, but keeping in mind that the real length of a 120 min gradient is about 90 min (the rest of 30 min in washing with organic), this means that an MS run can identify $\sim 15 \times 90 = 1{,}350$ proteins. If the length of an MS/MS decreases from 3 to 1 s and the number of precursors selected within MS survey for MS/MS increases to 6, then the number of proteins identified increases by sixfold ($\sim 1{,}350 \times 6 = 8{,}100$ proteins). Assuming that these results are at a flow rate of 0.5 μL/min, if we reduce the flow rate by ½, the number of proteins that can be identified increases by a factor of 2 (i.e., $8{,}100 \times 2 = 16{,}200$).

However, when we calculate the number of these proteins that can be identified, our assumption is that all the steps mentioned work perfectly. In practice, this is often

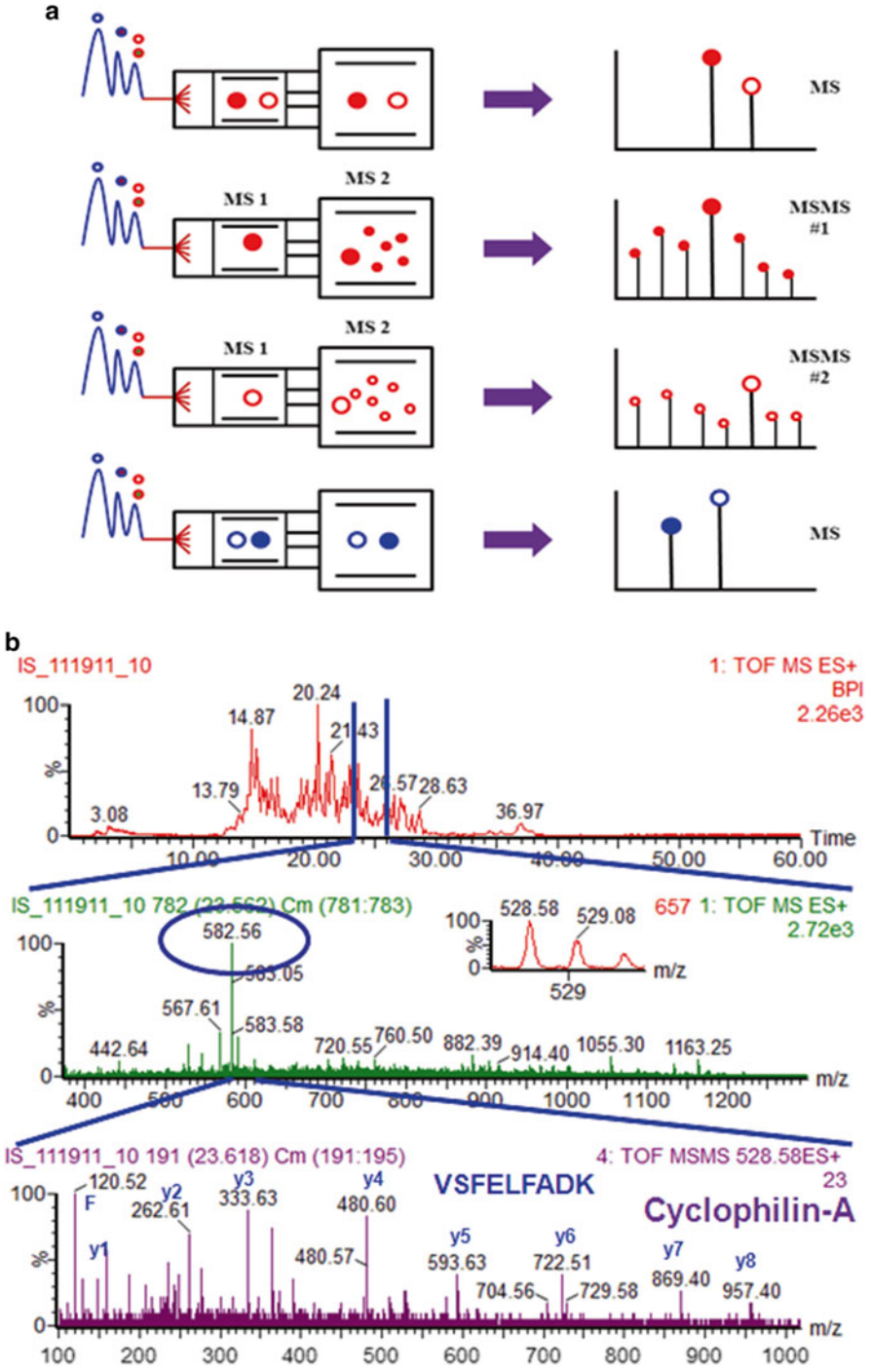


Fig. 1.6 LC–MS/MS experiment. (**a**) In each LC–MS/MS experiment, with elution of peptides from the HPLC gradually, the mass spectrometer analyzes corresponding ions via MS survey (recorded in an MS spectrum). Ions with highest intensity (typically 1–8 ions; two ions in this example) are selected for MS/MS fragmentation, fragmented, and then recorded as MS/MS #1 and MS/MS #2.

not the case. For example, the type and length of the gradient in HPLC (for example, sharp or shallow) do play an important role in peptide fractionation. An optimized versus a non-optimized nanospray will always play a role in the outcome of the proteomics experiment and the number of proteins identified. Obtaining a nanospray is just not good enough; "getting a good nanospray" is crucial to the success of a proteomics experiment. These and other known and/or unknown factors (not described here) that may influence the protein identification do indeed decrease the number of proteins identified in a proteomics experiment and in practice, a good LC–MS/MS run usually leads to identification of about 500–1,000 proteins. An example of a total ion current/chromatogram (TIC), MS, and MS/MS is shown in Fig. 1.6b.

1.7 Data Analysis

The raw data collected by a mass spectrometer are usually processed with software (for example, Protein Lynx Global Server, PLGS from Waters Corporation) and the output data (i.e., a peak list) is used for database search. There are many database search engines such as Sequest, X!Tandem, Mascot, or Phenyx. The results from the database search (such as from PLGS processing or Mascot search) can also be imported into a third-party software such as Scaffold (proteomesoftware.com) and further analyzed for protein modification, quantitation, and other factors.

MS may be not only qualitative but also quantitative and methods such as DIGE [60], isotope-coded affinity tag (ICAT) [5], stable isotope labeling by amino acids in cell culture (SILAC) [61], absolute quantitation (AQUA) [62], multiple reaction monitoring (MRM) [63], or spectral counting [64] have been successfully used in detection, identification, and quantification of proteins or peptides.

1.8 Protein Identification and Characterization

Determination of the molecular mass and amino acid sequence is the first step in protein identification. Once the protein is identified, then it is characterized. There are two methods for protein characterization using MS: a top-down approach when

Fig. 1.6 (continued) The mass spectrometer returns to the MS function at that point, recording an MS spectrum (MS survey). Once again ions with highest intensity are selected for fragmentation, fragmented, and recorded as MS/MS spectra. (**b**) An example of an LC–MS/MS experiment in which total ion current is recorded and at a specified time, an MS survey is recorded and one peak corresponding to a peptide (*m*/*z* of 582.56, doubly charged) is selected, and then fragmented in MS/MS. The fragmentation pattern (primarily b and y ions) from MS/MS provides sequence information regarding the peptide, leading to identification via database search. In this example, the peptide identified had the sequence VSFELFADK, identified as a component of human cyclophilin A. Reprinted and adapted with permission from the *Oxidative Stress*: *Diagnostics*, *Prevention*, *and Therapy*, S. Andreescu and M. Hepel, Editors. 2011, American Chemical Society: Washington, D.C [16]

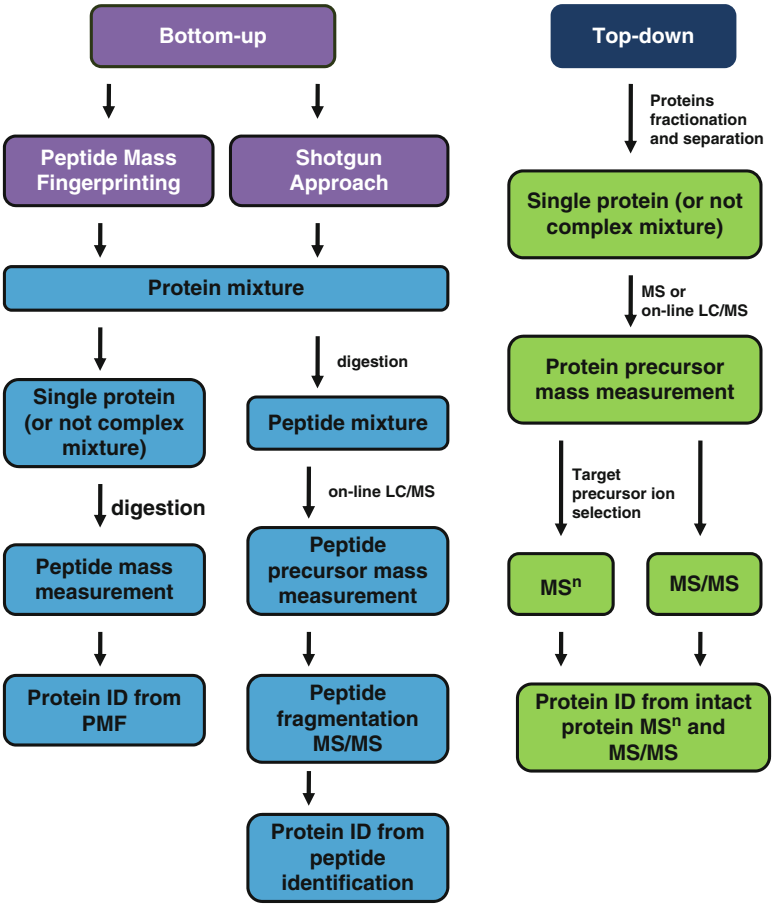


Fig. 1.7 Schematic workflow for bottom-up and top-down MS-based protein characterization and identification. Reprinted and adapted with permission from the *Australian Journal of Chemistry CSIRO Publishing* http://www.publish.csiro.au/?paper=CH13137 [15]

intact proteins are investigated and bottom-up approach when proteins are digested and the peptide mixture is analyzed (Fig. 1.7).

A top-down approach allows for the identification of protein isoforms or any potential PTMs within proteins [65]. In bottom-up approach, digested proteins are subjected to MS analysis using on-line tandem mass spectrometry (MS/MS). In the same bottom-up approach, peptide mass fingerprinting for protein identification is also used, particularly in MALDI-MS analyses.

In a variation of bottom-up proteomics, known as shotgun proteomics, a large protein mixture is digested, and the resulting peptides are fractionated by one-dimensional or multidimensional chromatography and further analyzed by MS/MS [66]. For maximum protein identification and characterization, a combination of bottom-up and top-down proteomics is/can be used [67, 68].

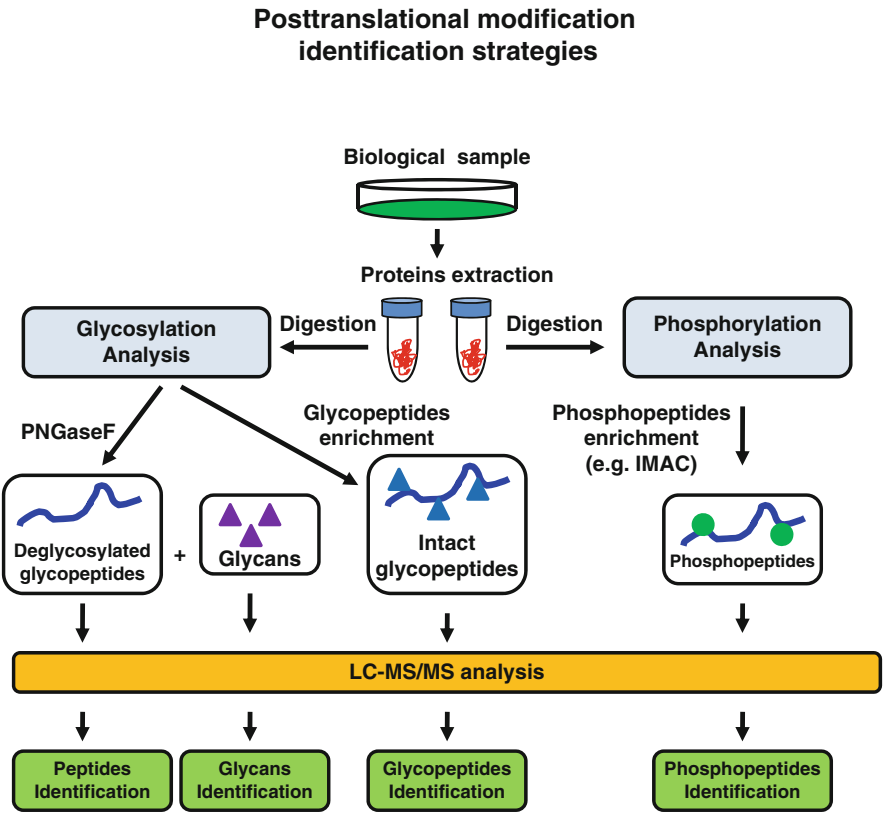


Fig. 1.8 MS-based characterization of protein PTMs (glycosylation and phosphorylation), general strategies. Reprinted and adapted with permission from the *Australian Journal of Chemistry CSIRO Publishing* http://www.publish.csiro.au/?paper=CH13137 [15]

Characterization of proteins is not easy, but it becomes even more complicated due to the intensive PTMs of proteins. It is very difficult to fully identify PTMs at a particular time point in cells, tissues, or organisms and to derive a meaningful interpretation and biological significance from these identified PTMs. So far, the only method that is appropriate for large scale identification of PTMs is MS-based proteomics [69]. PTMs are time- and site-specific events and are important to all biological processes. However, for a meaningful characterization, special enrichment strategies must be used. These strategies are able to characterize most stable modifications in proteins which include glycosylation, phosphorylation, disulfide bridges, acetylation, ubiquitination, and methylation. MS approaches for identification and characterization of proteins and PTMs are shown in Fig. 1.8.

Two common PTMs in proteins are glycosylation and phosphorylation. Glycosylation is commonly found in extracellular proteins or in the proteins that form the extracellular side and are responsible for biological processes such as cell–cell communication or ligand–lectin interaction [70, 71]. In the pharmaceutical and

biotechnology industry that focus on biotherapeutics, glycosylation is a critical modification of recombinant proteins, which influences their stability and solubility [72, 73]. Therefore, characterization of glycoproteins is difficult because the glycosylation is not uniform and usually more glycoforms are simultaneously produced by the cells and the accuracy in the MS-based identification and characterization of the glycoprotein isoforms is crucial [74].

Analysis of glycoproteins may be accomplished by LC–MS/MS analysis of tryptic digests. This method allows for identification of saccharide diagnostic fragments (i.e., hexoses), but its detection efficiency for glycoproteins is rather poor [75–77]. A better strategy involves glycoprotein enrichment by affinity chromatography (lectins), which facilitates its identification in subsequent LC–MS/MS analysis [78]. Another strategy involves the release of the glycans from glycopeptides, followed by targeted analysis of the glycans. N-linked glycans can be digested using peptide-*N*-glycosidase F (PNGase F), while O-linked glycans can be released by β-elimination. Change in mass units for peptides upon glycan removal allows for identification of the types of glycosylation (N- or O-linked), as well as the sites of glycosylation. For N-linked glycans, PNGase F treatment leads to an asparagine-to-aspartate conversion, with a net increase of 1 mass unit [14].

For O-glycans, conversion of serine to alanine and of threonine to aminobutyric acid results in a net loss of 16 mass units [79]. While the currently developed methods allow for fast and reliable identification of glycosylation sites, their characterization is still a great challenge, mostly due to the presence of glycan positional isomers [74].

Phosphorylation is a common and reversible PTM that plays a role in modulating many cellular processes [80]. Abnormal phosphorylation in various proteins and the phosphorylation patterns in the proteomes have been connected to various diseases [81, 82]. Therefore identification of protein phosphorylation will allow us to understand many physiological processes such as the phosphorylation-based signal transduction pathways and hopefully may lead to the discovery of new therapeutic targets [83–85].

Identification and characterization of phosphorylation on peptides are usually accomplished by MS and scanning for neutral loss of HPO_3 (80 mass units) from phosphotyrosine and H_3PO_4 (98 mass units) from phosphoserine and phosphothreonine residues [86, 87] usually allows the identification of phosphopeptides and the amino acid that is phosphorylated. Complete methylation of peptide sample followed by MALDI-MS analysis in both positive and negative ionization modes was also successfully employed [88]. However, since phosphorylation is a transient event and phosphorylation–dephosphorylation events may have opposite biological effects, data verification, data validation, and data interpretation may be difficult. Therefore, enrichment of phosphopeptides using TiO_2, metal-oxide-based resins (MOAC), a combination of TiO_2 and IMAC (TiMAC), and antibody affinity purifications [89, 90] is advised.

Another important protein PTM is disulfide bridges [11, 12, 15, 91], formed through the oxidation of cysteine residues, with an important role in maintaining the

three-dimensional conformation of proteins and inherently their physiological function. Disulfide bridges are usually found in extracellular and membrane-bound proteins, both as homodimers and homopolymers, but also as heterodimers and heteropolymers. Correct disulfide bridge formation is essential for proteins in adopting their optimal three-dimensional structure and assignment of the disulfide connectivities allows researchers to understand the structure and function of these proteins under physiological conditions and to predict problems in normal functioning of proteins when the disulfide bridges are scrambled or misconnected [92, 93].

Assignment of disulfide bridges in proteins may be accomplished by many approaches. For example, separation of disulfide-linked proteins or peptides by SDS-PAGE or tricine-PAGE under nonreducing and reducing conditions, followed by Coomassie staining and MS analysis, is one option. This MS analysis involves digestion of reduced and non-reduced aliquots of the same peptide mixture, followed by comparison of the masses of peptides that contain one cysteine, versus the disulfide-linked peptides in their oxidized form using MALDI-MS or ESI-MS [94, 95]. This task (assignment of disulfide bridges) is difficult when only one protein is analyzed and it is even more difficult when there are more cysteine residues per protein. In addition, there is no particular approach that allows one to identify disulfide bridges on a large scale. While mass spectrometry is capable of simultaneously analyzing many disulfide bridges, there are no bioinformatics means to interpret the MS data. Here we discussed only three types of PTMs, but there are many additional PTMs with biological significance and a similar number of challenges that are yet to be solved for each PTM and for automation of high-throughput identification and characterization of PTMs. However, it is clear to us that MS-based proteomics is perhaps the best and only option to accomplish both identification and characterization of PTMs.

1.9 Mass Spectrometry-Based Peptide and Protein Profiling and Quantitation

In addition to qualitative proteomics, another dimension in a proteomics experiment is quantitative analysis of proteins from the samples analyzed. Changes in protein expression between different physiological states or between physiological and pathological ones are common and qualitative proteomics without quantitative interrogation is at its best, partial proteomics. Therefore adding an extra dimension (quantitative) to MS-based proteomics expands its capabilities and advantages [96].

There are many workflows that have been developed and optimized to interrogate two or more proteomes or particular proteins from these proteomes using quantitative analysis, some of which are depicted in Fig. 1.9. Traditional quantitative analysis compares two sets of proteomes or a proteome from two different physiological states or physiological and pathological states and is a gel-based protein profiling technology which employs 2D-PAGE [97]. The protein spots that have different

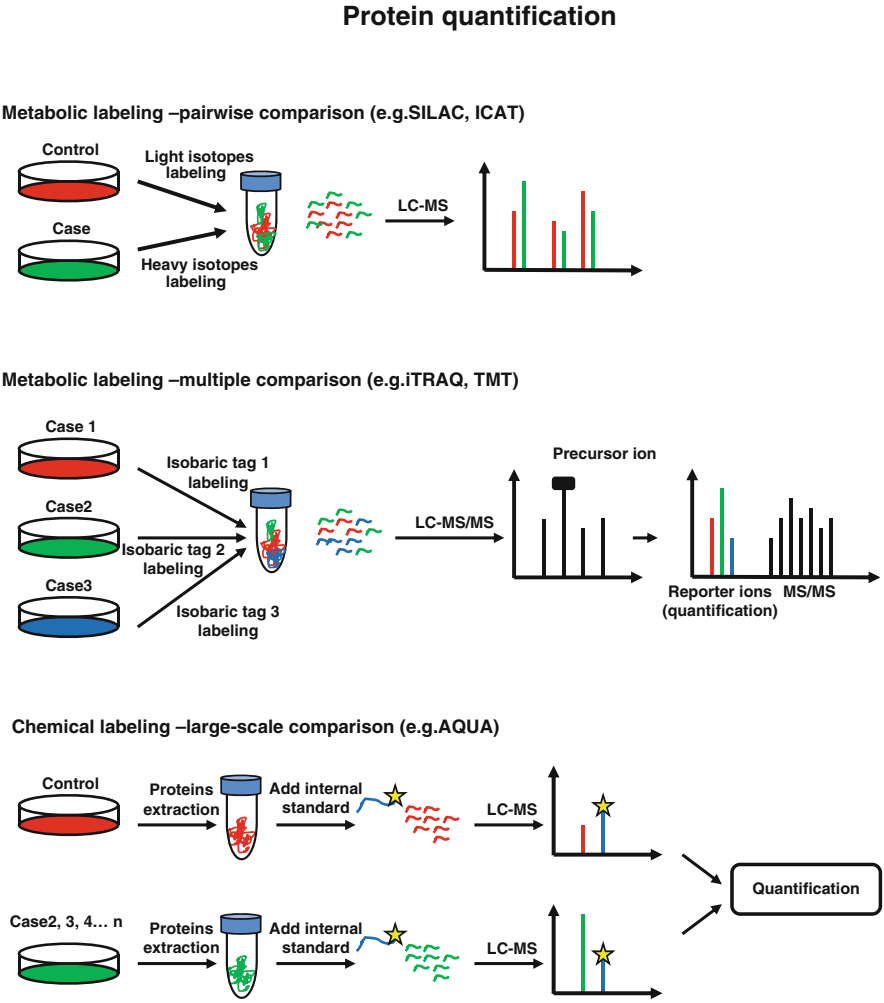


Fig. 1.9 MS-based protein quantification workflow strategies via stable isotope labeling. Reprinted and adapted with permission from the *Australian Journal of Chemistry CSIRO Publishing* http://www.publish.csiro.au/?paper=CH13137 [15]

intensities are usually excised, digested, and analyzed by MS, usually accomplished by instruments capable of MS/MS fragmentation (i.e., triple quadruple or ion trap mass spectrometers), simply because many compounds have similar masses and it may be difficult to monitor them in complex matrices. In addition, combination of MS and MS/MS allows one to use a combination of precursor ion for MS and fragment ions for MS/MS, thus providing a more selective monitoring of peptide/protein quantity [98–100].

Protein quantitation can be made using label-based or label-free techniques. Label-free methods are often used in many proteomic measurements because they

are simple and the cost is low [99]. Proteins also do not require special handling, such as tag or isotope labeling. With current advancements in software technology, there is no limit to the number of samples that can be analyzed.

Among the approaches used for label-free protein quantification are spectral counting and measurement of MS precursor ion intensity (or chromatographic peak area) [101, 102]. In spectral counting, one measures the number of spectra that correspond with peptides that are part of one protein [64, 103], while MS precursor ion intensity approach interrogates the chromatographic peaks corresponding to particular peptides at a normalized elution time. The protein or peptide quantity is then calculated using a standard curve or area under the curve as compared with another sample. All these label-free methods, although fast and cheap, also have disadvantages: they depend on analytical and biological reproducibility and any variation in sample preparation or sample analysis can lead to technical and instrumental errors, well reflected in the quantitation outcome [104, 105].

In addition to label-free quantitation, label-based quantitation strategies have emerged that use stable isotopes (^{13}C, ^{15}N, ^{18}O, or ^{2}H) [106], in which native and labeled samples are combined and analyzed simultaneously. This isotope-based quantitation is also called absolute quantitation, as opposed to label-free, relative quantitation.

Using the absolute quantification method, synthetic peptides or proteins are used, which are labeled with stable isotopes on one or more amino acids [6, 107]. The peptides are used as internal labeled standards, which are added directly to the samples to be analyzed. Due to the difference in the isotope pattern (and the mass difference) during MS analysis, quantitation can be performed [108–110]. One label-based quantitation method is stable isotope labeling of select amino acids (usually arginine or lysine) in cell culture (SILAC), used for metabolic labeling [61]. This method can be used in many applications such as investigation of signaling pathways [8, 22, 111–119], but it is mostly restricted to cell culture and it cannot be used to investigate biological fluids (i.e., blood, urine, saliva) [120]. However, recently, Matthias Mann's group created labeled mice. Animals were fed with a ^{13}C arginine- and ^{13}C lysine-infused diet and then can be used to investigate biological fluids [121, 122]. Another method for stable isotope labeling is the incorporation of labeled/modified tags on specific amino acids. In one such method, cysteine residues may be labeled using an ICAT, in which two conditions (i.e., two proteomes) are investigated [5]. For analysis of more than two conditions, such as for time-course experiments, or for three or more different biological samples, labeling strategies using isobaric tags for relative and absolute quantification (iTRAQ) and tandem mass tags (TMT) have been successfully developed [123–125]. Although these methods have improved quantification capabilities, there are still limitations such as lack of reproducibility between individual analyses. Some of these issues are addressed through targeted quantification using approaches such as selected reaction monitoring (SRM) [126] and MRM [127], which have shown excellent reproducibility when used with stable isotope-labeled internal standards [128, 129].

With fast advancement in current proteomic methodologies and technologies, protein quantification and profiling will become a standard for use in clinical

diagnostic laboratories. However, before large scale proteomics experiments can be performed for clinical use, the reproducibility of large scale quantitation must be addressed and new, improved quantitation platforms developed and tested. Nevertheless, these new technologies indeed have the potential to be easily integrated in the set of tools that can perform disease- or disorder-specific protein profiling and that future is close to becoming a reality.

1.10 Identification of Protein–Protein Interactions (PPIs) Using MS

The molecules within a cell are not static, but rather dynamic. They form various types of interactions. These interactions can be static, as in protein complexes, or dynamic, as in transient protein interactions such as hormone–receptor interactions or substrate–enzyme interactions. All of these interactions within a cell form the interactomics network or interactome, and the proteins are a major component [130], modifying and controlling their own or other proteins' functions [131]. The dysregulation of these interactions, particularly PPIs, usually leads to a pathological state such as diseases or disorders and their investigation is essential to the current efforts to understand these diseases or disorders.

There are many methods for identifying PPIs such as size-exclusion chromatography (SEC) [132–134], sucrose gradient ultracentrifugation [135, 136], the yeast two-hybrid system (Y2H), or affinity purification MS (AP-MS) [3, 130, 137–140]. These methods allow identification of stable PPIs (i.e., by sucrose gradient, SEC, or AP-MS), of binary interactions (Y2H), or of transient PPIs (AP-MS). However, these current methods have limitations. Sucrose gradient ultracentrifugation and SEC are time-consuming and not suitable for automation, while Y2H and AP-MS can be automated, but have high rate of false positive identifications of PPIs [141–143]. Therefore, efforts are being made to reduce these limitations.

Native gel electrophoresis (clear native PAGE or CN-PAGE and BN-PAGE) are an alternative option and separate protein complexes according to their molecular mass (BN-PAGE) or according to their internal charge and independent of their mass (CN-PAGE) [21, 144]. The advantage of these methods is that they can separate all protein complexes from the whole proteome in one single experiment and can be combined with MS to identify protein complexes [145, 146]. However, the problem with these methods is that the gels are usually "home-made," are not always reproducible, and require extensive work and bioinformatics expertise.

As an alternative option, one may also use ESI-MS for direct measurement of stable and transient PPIs in a solution [147, 148]. However, we still do not have capability to fully comprehend and the means to fully investigate the PPIs, in particular transient PPIs and even more complicated transient PPIs that have transient or reversible PTMs. This field will perhaps be called something like PTM-ed-PPI-omics.

1.11 Recent Advances

Currently we have the means to create disease animal models or control animal models (such as the SILAC mouse for absolute quantitation) [149, 150]. Capabilities of mass spectrometers have recently increased and currently MS-based technology allows us to identify thousands of proteins. Current machines are rich in additional technology that allows one to identify not only proteins and their PTMs and PPIs but also their shape and configuration. Such instruments are commercially available [151], thus filling the need for analysis of proteins that are not easy to investigate using classical approaches such as X-ray and NMR [152]. Therefore, these and other methods not listed here not only provide a solution for analysis of challenging proteins but have also opened the doors to new fields such as structural proteomics [153, 154].

1.12 Challenges and Perspectives

Many genomes have been sequenced. Many proteins from various sources have been identified. However, it will be an enormous mistake to state that we have identified a full proteome of a whole cell. This milestone has not yet been achieved. We have identified many or most of proteins in specific cells, such as bacteria, expressed at a particular time point, under particular growing conditions. However, while we are close to identification of most proteins in a cell, we are far from identification of the full proteome, including all proteins, isoproteins, modified proteins (PTMs), and truncated proteins. After the sequencing of the human genome, humans realized that our genome contains only ~30,000 genes, but encodes for about 100,000 unique protein sequences [155, 156]. Adding the truncated proteins, splice isoforms, and mutated proteins, PTMs will give us a number of, between 1 and 2,000,000, proteins, many of them expressed either transiently or in a very low concentration, thus making the cell's proteome complexity more difficult to analyze and interpret [157]. To partially overcome some of these challenges, more and more advanced MS-based technologies can be combined with many fractionation, separation, and identification methods within one experiment, such as combining immunoaffinity with gel electrophoresis, liquid chromatography (LC), and MS, to increase sensitivity and dynamic range. Furthermore, there is plenty of room for optimization of MS-based methods to be successfully used in high-throughput analysis.

Therapeutic proteins are frequently membrane proteins [158–161]. However, the membrane proteomics (membranomics) is the most difficult task that can be achieved in proteomics. In addition, transmembrane proteins are the most modified proteins by PTMs and are simply very difficult to investigate by MS, although some progress has been made in this direction [162, 163].

1.13 Conclusions

Despite the many challenges that the MS and proteomics fields face, the impact of MS is stronger and stronger, year after year. The number of unknown proteins decreases and protein databases become more comprehensive over time. With new MS technology and with combinatorial approaches towards the simultaneous identification of proteins, isoproteins, and truncated PTMs, PPIs will hopefully allow us to completely characterize proteomes at both a qualitative and quantitative level.

Acknowledgements We would like to thank Ms. Laura Mulderig, Scott Nichols, and their colleagues (Waters Corporation) for their generous support in setting up the Proteomics Center at Clarkson University. CCD thanks Drs. Thomas A. Neubert (New York University), Belinda Willard (Cleveland Clinic), and Gregory Wolber and David Mclaughin (Eastman Kodak Company) for donation of a TofSpec2E MALDI-MS (each). CCD thanks his advisors, Vlad Artenie, Wolfgang Haehnel, Paul M. Wassarman, and Thomas A. Neubert, for their advice and support. This work was supported in part by the Keep a Breast Foundation (KEABF-375-35054), the Redcay Foundation (SUNY Plattsburgh), the Alexander von Humboldt Foundation, SciFund Challenge, private donations (Ms. Mary Stewart Joyce and Mr. Kenneth Sandler), the David A. Walsh fellowship, and by the U.S. Army research office (DURIP grant #W911NF-11-1-0304).

References

1. Aebersold R, Mann M (2003) Mass spectrometry-based proteomics. Nature 422:198–207
2. Aivaliotis M, Karas M, Tsiotis G (2006) High throughput two-dimensional blue-native electrophoresis: a tool for functional proteomics of cytoplasmatic protein complexes from Chlorobium tepidum. Photosynth Res 88:143–157
3. Blagoev B, Kratchmarova I, Ong SE, Nielsen M, Foster LJ, Mann M (2003) A proteomics strategy to elucidate functional protein-protein interactions applied to EGF signaling. Nat Biotechnol 21:315–318
4. Camacho-Carvajal MM, Wollscheid B, Aebersold R, Steimle V, Schamel WW (2004) Two-dimensional Blue native/SDS gel electrophoresis of multi-protein complexes from whole cellular lysates: a proteomics approach. Mol Cell Proteomics 3:176–182
5. Gygi SP, Rist B, Gerber SA, Turecek F, Gelb MH, Aebersold R (1999) Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. Nat Biotechnol 17:994–999
6. Ong SE, Foster LJ, Mann M (2003) Mass spectrometric-based approaches in quantitative proteomics. Methods 29:124–130
7. Shevchenko A, Wilm M, Vorm O, Mann M (1996) Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. Anal Chem 68:850–858
8. Zhang G, Spellman DS, Skolnik EY, Neubert TA (2006) Quantitative phosphotyrosine proteomics of EphB2 signaling by stable isotope labeling with amino acids in cell culture (SILAC). J Proteome Res 5:581–588
9. Darie C (2013) Mass spectrometry and proteomics: principle, workflow, challenges and perspectives. Mod Chem Appl 1:e105
10. Darie CC (2013) Mass spectrometry and its application in life sciences. Aust J Chem 66:1–2
11. Ngounou Wetie AG, Sokolowska I, Woods AG, Darie CC (2013) Identification of post-translational modifications by mass spectrometry. Aust J Chem 66:734–748
12. Ngounou Wetie AG, Sokolowska I, Woods AG, Roy U, Deinhardt K, Darie CC (2014) Protein-protein interactions: switch from classical methods to proteomics and bioinformatics-based approaches. Cell Mol Life Sci 71(2):205–228

13. Ngounou Wetie AG, Sokolowska I, Wormwood K, Michel TM, Thome J, Darie CC, Woods AG (2013) Mass spectrometry for the detection of potential psychiatric biomarkers. J Mol Psychiatry 1:8
14. Sokolowska I, Ngounou Wetie AG, Roy U, Woods AG, Darie CC (2013) Mass spectrometry investigation of glycosylation on the NXS/T sites in recombinant glycoproteins. Biochim Biophys Acta 1834:1474–1483
15. Sokolowska I, Ngounou Wetie AG, Woods AG, Darie CC (2013) Applications of mass spectrometry in proteomics. Aust J Chem 66:721–733
16. Sokolowska I, Woods AG, Wagner J, Dorler J, Wormwood K, Thome J, Darie CC (2011) Mass spectrometry for proteomics-based investigation of oxidative stress and heat shock proteins. In: Andreescu S, Hepel M (eds) Oxidative stress: diagnostics, prevention, and therapy. American Chemical Society, Washington, DC
17. Woods AG, Ngounou Wetie AG, Sokolowska I, Russell S, Ryan JP, Michel TM, Thome J, Darie CC (2013) Mass spectrometry as a tool for studying autism spectrum disorder. J Mol Psychiatry 1:6
18. Darie CC, Biniossek ML, Winter V, Mutschler B, Haehnel W (2005) Isolation and structural characterization of the Ndh complex from mesophyll and bundle sheath chloroplasts of Zea mays. Febs J 272:2705–2716
19. Darie CC, Janssen WG, Litscher ES, Wassarman PM (2008) Purified trout egg vitelline envelope proteins VEbeta and VEgamma polymerize into homomeric fibrils from dimers in vitro. Biochim Biophys Acta 1784:385–392
20. Schagger H, Cramer WA, von Jagow G (1994) Analysis of molecular masses and oligomeric states of protein complexes by blue native electrophoresis and isolation of membrane protein complexes by two-dimensional native electrophoresis. Anal Biochem 217:220–230
21. Schagger H, von Jagow G (1991) Blue native electrophoresis for isolation of membrane protein complexes in enzymatically active form. Anal Biochem 199:223–231
22. Spellman DS, Deinhardt K, Darie CC, Chao MV, Neubert TA (2008) Stable isotopic labeling by amino acids in cultured primary neurons: application to brain-derived neurotrophic factor-dependent phosphotyrosine-associated signaling. Mol Cell Proteomics 7:1067–1076
23. Darie CC, Shetty V, Spellman DS, Zhang G, Xu C, Cardasis HL, Blais S, Fenyo D, Neubert TA (2008) Blue Native PAGE and mass spectrometry analysis of the ephrin stimulation-dependent protein-protein interactions in NG108-EphB2 cells. Springer, Düsseldorf, Germany
24. Darie CC, Litscher ES, Wassarman PM (2008) Structure, processing, and polymerization of rainbow trout egg vitelline envelope proteins. Springer, Düsseldorf, Germany
25. Darie CC, Biniossek ML, Gawinowicz MA, Milgrom Y, Thumfart JO, Jovine L, Litscher ES, Wassarman PM (2005) Mass spectrometric evidence that proteolytic processing of rainbow trout egg vitelline envelope proteins takes place on the egg. J Biol Chem 280:37585–37598
26. Darie CC, Biniossek ML, Jovine L, Litscher ES, Wassarman PM (2004) Structural characterization of fish egg vitelline envelope proteins by mass spectrometry. Biochemistry 43: 7459–7478
27. Schagger H (2006) Tricine-SDS-PAGE. Nat Protoc 1:16–22
28. Schagger H, von Jagow G (1987) Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. Anal Biochem 166:368–379
29. Jovine L, Darie CC, Litscher ES, Wassarman PM (2005) Zona pellucida domain proteins. Annu Rev Biochem 74:83–114
30. Litscher ES, Janssen WG, Darie CC, Wassarman PM (2008) Purified mouse egg zona pellucida glycoproteins polymerize into homomeric fibrils under non-denaturing conditions. J Cell Physiol 214:153–157
31. Wassarman PM, Jovine L, Qi H, Williams Z, Darie C, Litscher ES (2005) Recent aspects of mammalian fertilization research. Mol Cell Endocrinol 234:95–103
32. Darie C (2013) Investigation of protein-protein interactions by blue native-PAGE & mass spectrometry. Mod Chem Appl 1:e111

33. Darie C (2013) Post-translational modification (PTM) proteomics: challenges and perspectives. Mod Chem appl 1:e114
34. Woods AG, Sokolowska I, Yakubu R, Butkiewicz M, LaFleur M, Talbot C, Darie CC (2011) Blue native page and mass spectrometry as an approach for the investigation of stable and transient protein-protein interactions. In: Andreescu S, Hepel M (eds) Oxidative stress: diagnostics, prevention, and therapy. American Chemical Society, Washington, DC
35. Ghezzi P, Bonetto V (2003) Redox proteomics: identification of oxidatively modified proteins. Proteomics 3:1145–1153
36. Li X, Pan W, Yang GZ, Di YN, Zhao F, Zhu LY, Jiang ZH (2011) Proteome analysis of differential protein expression in brain of rats with type 1 diabetes mellitus. Exp Clin Endocrinol Diabetes 119:265–270
37. Muroi M, Kazami S, Noda K, Kondo H, Takayama H, Kawatani M, Usui T, Osada H (2010) Application of proteomic profiling based on 2D-DIGE for classification of compounds according to the mechanism of action. Chem Biol 17:460–470
38. Polden J, McManus CA, Remedios CD, Dunn MJ (2011) A 2-D gel reference map of the basic human heart proteome. Proteomics 11(17):3582–3586
39. Stefanescu R, Iacob RE, Damoc EN, Marquardt A, Amstalden E, Manea M, Perdivara I, Maftei M, Paraschiv G, Przybylski M (2007) Mass spectrometric approaches for elucidation of antigenantibody recognition structures in molecular immunology. Eur J Mass Spectrom (Chichester, Eng) 13:69–75
40. Sun X, Jia HL, Xiao CL, Yin XF, Yang XY, Lu J, He X, Li N, Li H, He QY (2011) Bacterial proteome of Streptococcus pneumoniae through multidimensional separations coupled with LC-MS/MS. OMICS 15(7–8):477–482
41. Wang Y, Li R, Du D, Zhang C, Yuan H, Zeng R, Chen Z (2006) Proteomic analysis reveals novel molecules involved in insulin signaling pathway. J Proteome Res 5:846–855
42. Bauw G, Rasmussen HH, van den Bulcke M, van Damme J, Puype M, Gesser B, Celis JE, Vandekerckhove J (1990) Two-dimensional gel electrophoresis, protein electroblotting and microsequencing: a direct link between proteins and genes. Electrophoresis 11:528–536
43. Celis JE, Gromov P (1999) 2D protein electrophoresis: can it be perfected? Curr Opin Biotechnol 10:16–21
44. Celis JE, Gromov P, Ostergaard M, Madsen P, Honore B, Dejgaard K, Olsen E, Vorum H, Kristensen DB, Gromova I, Haunso A, Van Damme J, Puype M, Vandekerckhove J, Rasmussen HH (1996) Human 2-D PAGE databases for proteome analysis in health and disease: http://biobase.dk/cgi-bin/celis. FEBS Lett 398:129–134
45. Celis JE, Gromova I, Moreira JM, Cabezon T, Gromov P (2004) Impact of proteomics on bladder cancer research. Pharmacogenomics 5:381–394
46. Taurines R, Dudley E, Conner AC, Grassl J, Jans T, Guderian F, Mehler-Wex C, Warnke A, Gerlach M, Thome J (2010) Serum protein profiling and proteomics in autistic spectrum disorder using magnetic bead-assisted mass spectrometry. Eur Arch Psychiatry Clin Neurosci 260:249–255
47. Taurines R, Dudley E, Grassl J, Warnke A, Gerlach M, Coogan AN, Thome J (2011) Proteomic research in psychiatry. J Psychopharmacol 25:151–196
48. Dass C (2007) Fundamentals of contemporary mass spectrometry. Wiley-Interscience, Hoboken, NJ
49. Hoffmann ED, Stroobant V (2007) Mass spectrometry: principles and applications, 3rd edn. J. Wiley, Chichester, West Sussex, England
50. Abate S, Ahn YG, Kind T, Cataldi TR, Fiehn O (2010) Determination of elemental compositions by gas chromatography/time-of-flight mass spectrometry using chemical and electron ionization. Rapid Commun Mass Spectrom 24:1172–1180
51. Harrison AG (1992) Chemical ionization mass spectrometry, 2nd edn. CRC Press, Boca Raton, FL
52. Rivera-Rodriguez LB, Rodriguez-Estrella R, Ellington JJ, Evans JJ (2007) Quantification of low levels of organochlorine pesticides using small volumes (<or = 100 microl) of plasma of

wild birds through gas chromatography negative chemical ionization mass spectrometry. Environ Pollut 148:654–662
53. Dougherty RC (1981) Negative chemical ionization mass spectrometry: applications in environmental analytical chemistry. Biomed Mass Spectrom 8:283–292
54. Zaikin VG, Halket JM (2006) Derivatization in mass spectrometry—8. Soft ionization mass spectrometry of small molecules. Eur J Mass Spectrom (Chichester, Eng) 12:79–115
55. Marshall AG, Hendrickson CL, Jackson GS (1998) Fourier transform ion cyclotron resonance mass spectrometry: a primer. Mass Spectrom Rev 17:1–35
56. Martin SE, Shabanowitz J, Hunt DF, Marto JA (2000) Subfemtomole MS and MS/MS peptide sequence analysis using nano-HPLC micro-ESI fourier transform ion cyclotron resonance mass spectrometry. Anal Chem 72:4266–4274
57. Parker CE, Warren MR, Mocanu V (2010) Mass spectrometry for proteomics
58. Yates JR, Ruse CI, Nakorchevsky A (2009) Proteomics by mass spectrometry: approaches, advances, and applications. Annu Rev Biomed Eng 11:49–79
59. Domon B, Aebersold R (2006) Mass spectrometry and protein analysis. Science 312: 212–217
60. Viswanathan S, Unlu M, Minden JS (2006) Two-dimensional difference gel electrophoresis. Nat Protoc 1:1351–1358
61. Ong SE, Blagoev B, Kratchmarova I, Kristensen DB, Steen H, Pandey A, Mann M (2002) Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. Mol Cell Proteomics 1:376–386
62. Stemmann O, Zou H, Gerber SA, Gygi SP, Kirschner MW (2001) Dual inhibition of sister chromatid separation at metaphase. Cell 107:715–726
63. Anderson L, Hunter CL (2006) Quantitative mass spectrometric multiple reaction monitoring assays for major plasma proteins. Mol Cell Proteomics 5:573–588
64. Liu H, Sadygov RG, Yates JR III (2004) A model for random sampling and estimation of relative protein abundance in shotgun proteomics. Anal Chem 76:4193–4201
65. McLafferty FW, Breuker K, Jin M, Han X, Infusini G, Jiang H, Kong X, Begley TP (2007) Top-down MS, a powerful complement to the high capabilities of proteolysis proteomics. FEBS J 274:6256–6268
66. McDonald WH, Yates JR III (2003) Shotgun proteomics: integrating technologies to answer biological questions. Curr Opin Mol Ther 5:302–309
67. Wu S, Lourette NM, Tolic N, Zhao R, Robinson EW, Tolmachev AV, Smith RD, Pasa-Tolic L (2009) An integrated top-down and bottom-up strategy for broadly characterizing protein isoforms and modifications. J Proteome Res 8:1347–1357
68. Han X, Aslanian A, Yates JR III (2008) Mass spectrometry for proteomics. Curr Opin Chem Biol 12:483–490
69. Savitski MF, Savitski MM (2010) Unbiased detection of posttranslational modifications using mass spectrometry. Methods Mol Biol 673:203–210
70. Spiro RG (2002) Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds. Glycobiology 12:43R–56R
71. 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
72. Read EK, Park JT, Brorson KA (2011) Industry and regulatory experience of the glycosylation of monoclonal antibodies. Biotechnol Appl Biochem 58:213–219
73. Kamoda S, Kakehi K (2008) Evaluation of glycosylation for quality assurance of antibody pharmaceuticals by capillary electrophoresis. Electrophoresis 29:3595–3604
74. Leymarie N, Zaia J (2012) Effective use of mass spectrometry for glycan and glycopeptide structural analysis. Anal Chem 84:3040–3048
75. Pan S, Chen R, Aebersold R, Brentnall TA (2011) Mass spectrometry based glycoproteomics—from a proteomics perspective. Mol Cell Proteomics 10(R110):003251
76. Morelle W, Michalski JC (2007) Analysis of protein glycosylation by mass spectrometry. Nat Protoc 2:1585–1602

77. Wuhrer M, Catalina MI, Deelder AM, Hokke CH (2007) Glycoproteomics based on tandem mass spectrometry of glycopeptides. J Chromatogr B Analyt Technol Biomed Life Sci 849:115–128
78. Mechref Y, Madera M, Novotny MV (2008) Glycoprotein enrichment through lectin affinity techniques. Methods Mol Biol 424:373–396
79. Bond MR, Kohler JJ (2007) Chemical methods for glycoprotein discovery. Curr Opin Chem Biol 11:52–58
80. Tarrant MK, Cole PA (2009) The chemical biology of protein phosphorylation. Annu Rev Biochem 78:797–825
81. Blume-Jensen P, Hunter T (2001) Oncogenic kinase signalling. Nature 411:355–365
82. Cohen P (2001) The role of protein phosphorylation in human health and disease. The Sir Hans Krebs Medal Lecture. Eur J Biochem 268:5001–5010
83. Badiola N, Suarez-Calvet M, Lleo A (2010) Tau phosphorylation and aggregation as a therapeutic target in tauopathies. CNS Neurol Disord Drug Targets 9:727–740
84. Cohen P (2002) Protein kinases—the major drug targets of the twenty-first century? Nat Rev Drug Discov 1:309–315
85. Strebhardt K (2010) Multifaceted polo-like kinases: drug targets and antitargets for cancer therapy. Nat Rev Drug Discov 9:643–660
86. Le Blanc JC, Hager JW, Ilisiu AM, Hunter C, Zhong F, Chu I (2003) Unique scanning capabilities of a new hybrid linear ion trap mass spectrometer (Q TRAP) used for high sensitivity proteomics applications. Proteomics 3:859–869
87. Unwin RD, Griffiths JR, Leverentz MK, Grallert A, Hagan IM, Whetton AD (2005) Multiple reaction monitoring to identify sites of protein phosphorylation with high sensitivity. Mol Cell Proteomics 4:1134–1144
88. Xu CF, Lu Y, Ma J, Mohammadi M, Neubert TA (2005) Identification of phosphopeptides by MALDI Q-TOF MS in positive and negative ion modes after methyl esterification. Mol Cell Proteomics 4:809–818
89. Beltran L, Cutillas PR (2012) Advances in phosphopeptide enrichment techniques for phosphoproteomics. Amino Acids 43:1009–1024
90. Corthals GL, Aebersold R, Goodlett DR (2005) Identification of phosphorylation sites using microimmobilized metal affinity chromatography. Methods Enzymol 405:66–81
91. Ngounou Wetie AG, Sokolowska I, Woods AG, Wormwood KL, Dao S, Patel S, Clarkson BD, Darie CC (2013) Automated mass spectrometry-based functional assay for the routine analysis of the secretome. J Lab Autom 18:19–29
92. Gorman JJ, Wallis TP, Pitt JJ (2002) Protein disulfide bond determination by mass spectrometry. Mass Spectrom Rev 21:183–216
93. McAuley A, Jacob J, Kolvenbach CG, Westland K, Lee HJ, Brych SR, Rehder D, Kleemann GR, Brems DN, Matsumura M (2008) Contributions of a disulfide bond to the structure, stability, and dimerization of human IgG1 antibody CH3 domain. Protein Sci 17:95–106
94. Sokolowska I, Gawinowicz MA, Ngounou Wetie AG, Darie CC (2012) Disulfide proteomics for identification of extracellular or secreted proteins. Electrophoresis 33:2527–2536
95. Sokolowska I, Ngounou Wetie AG, Woods AG, Darie CC (2012) Automatic determination of disulfide bridges in proteins. J Lab Autom 17:408–416
96. Panchaud A, Affolter M, Moreillon P, Kussmann M (2008) Experimental and computational approaches to quantitative proteomics: status quo and outlook. J Proteomics 71:19–33
97. Berkelman T (2008) Quantitation of protein in samples prepared for 2-D electrophoresis. Methods Mol Biol 424:43–49
98. Bantscheff M, Schirle M, Sweetman G, Rick J, Kuster B (2007) Quantitative mass spectrometry in proteomics: a critical review. Anal Bioanal Chem 389:1017–1031
99. Xie F, Liu T, Qian WJ, Petyuk VA, Smith RD (2011) Liquid chromatography-mass spectrometry-based quantitative proteomics. J Biol Chem 286:25443–25449
100. Pan S, Aebersold R, Chen R, Rush J, Goodlett DR, McIntosh MW, Zhang J, Brentnall TA (2009) Mass spectrometry based targeted protein quantification: methods and applications. J Proteome Res 8:787–797

101. Negishi A, Ono M, Handa Y, Kato H, Yamashita K, Honda K, Shitashige M, Satow R, Sakuma T, Kuwabara H, Omura K, Hirohashi S, Yamada T (2009) Large-scale quantitative clinical proteomics by label-free liquid chromatography and mass spectrometry. Cancer Sci 100:514–519
102. Ono M, Shitashige M, Honda K, Isobe T, Kuwabara H, Matsuzuki H, Hirohashi S, Yamada T (2006) Label-free quantitative proteomics using large peptide data sets generated by nano-flow liquid chromatography and mass spectrometry. Mol Cell Proteomics 5:1338–1347
103. Qian WJ, Jacobs JM, Camp DG II, Monroe ME, Moore RJ, Gritsenko MA, Calvano SE, Lowry SF, Xiao W, Moldawer LL, Davis RW, Tompkins RG, Smith RD (2005) Comparative proteome analyses of human plasma following in vivo lipopolysaccharide administration using multidimensional separations coupled with tandem mass spectrometry. Proteomics 5:572–584
104. Petyuk VA, Jaitly N, Moore RJ, Ding J, Metz TO, Tang K, Monroe ME, Tolmachev AV, Adkins JN, Belov ME, Dabney AR, Qian WJ, Camp DG II, Smith RD (2008) Elimination of systematic mass measurement errors in liquid chromatography-mass spectrometry based proteomics using regression models and a priori partial knowledge of the sample content. Anal Chem 80:693–706
105. Strittmatter EF, Ferguson PL, Tang K, Smith RD (2003) Proteome analyses using accurate mass and elution time peptide tags with capillary LC time-of-flight mass spectrometry. J Am Soc Mass Spectrom 14:980–991
106. Zhang R, Regnier FE (2002) Minimizing resolution of isotopically coded peptides in comparative proteomics. J Proteome Res 1:139–147
107. Julka S, Regnier F (2004) Quantification in proteomics through stable isotope coding: a review. J Proteome Res 3:350–363
108. Bronstrup M (2004) Absolute quantification strategies in proteomics based on mass spectrometry. Expert Rev Proteomics 1:503–512
109. Gerber SA, Rush J, Stemman O, Kirschner MW, Gygi SP (2003) Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. Proc Natl Acad Sci U S A 100:6940–6945
110. Kirkpatrick DS, Gerber SA, Gygi SP (2005) The absolute quantification strategy: a general procedure for the quantification of proteins and post-translational modifications. Methods 35:265–273
111. Zhang G, Neubert TA (2011) Comparison of three quantitative phosphoproteomic strategies to study receptor tyrosine kinase signaling. J Proteome Res 10:5454–5462
112. Zhang G, Deinhardt K, Chao MV, Neubert TA (2011) Study of neurotrophin-3 signaling in primary cultured neurons using multiplex stable isotope labeling with amino acids in cell culture. J Proteome Res 10:2546–2554
113. Deinhardt K, Kim T, Spellman DS, Mains RE, Eipper BA, Neubert TA, Chao MV, Hempstead BL (2011) Neuronal growth cone retraction relies on proneurotrophin receptor signaling through Rac. Sci Signal 4:ra82
114. Zhang G, Ueberheide BM, Waldemarson S, Myung S, Molloy K, Eriksson J, Chait BT, Neubert TA, Fenyo D (2010) Protein quantitation using mass spectrometry. Methods Mol Biol 673:211–222
115. Neubert TA, Tempst P (2010) Super-SILAC for tumors and tissues. Nat Methods 7:361–362
116. Zhang G, Neubert TA (2009) Use of stable isotope labeling by amino acids in cell culture (SILAC) for phosphotyrosine protein identification and quantitation. Methods Mol Biol 527:79–92, xi
117. Zhang G, Fenyo D, Neubert TA (2009) Evaluation of the variation in sample preparation for comparative proteomics using stable isotope labeling by amino acids in cell culture. J Proteome Res 8:1285–1292
118. Zhang G, Fenyo D, Neubert TA (2008) Screening for EphB signaling effectors using SILAC with a linear ion trap-orbitrap mass spectrometer. J Proteome Res 7:4715–4726
119. Zhang G, Neubert TA (2006) Automated comparative proteomics based on multiplex tandem mass spectrometry and stable isotope labeling. Mol Cell Proteomics 5:401–411

120. Guo A, Villen J, Kornhauser J, Lee KA, Stokes MP, Rikova K, Possemato A, Nardone J, Innocenti G, Wetzel R, Wang Y, MacNeill J, Mitchell J, Gygi SP, Rush J, Polakiewicz RD, Comb MJ (2008) Signaling networks assembled by oncogenic EGFR and c-Met. Proc Natl Acad Sci U S A 105:692–697
121. Kruger M, Moser M, Ussar S, Thievessen I, Luber CA, Forner F, Schmidt S, Zanivan S, Fassler R, Mann M (2008) SILAC mouse for quantitative proteomics uncovers kindlin-3 as an essential factor for red blood cell function. Cell 134:353–364
122. Zanivan S, Krueger M, Mann M (2012) In vivo quantitative proteomics: the SILAC mouse. Methods Mol Biol 757:435–450
123. Muth T, Keller D, Puetz SM, Martens L, Sickmann A, Boehm AM (2010) jTraqX: a free, platform independent tool for isobaric tag quantitation at the protein level. Proteomics 10:1223–1225
124. Ross PL, Huang YN, Marchese JN, Williamson B, Parker K, Hattan S, Khainovski N, Pillai S, Dey S, Daniels S, Purkayastha S, Juhasz P, Martin S, Bartlet-Jones M, He F, Jacobson A, Pappin DJ (2004) Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. Mol Cell Proteomics 3:1154–1169
125. Dayon L, Hainard A, Licker V, Turck N, Kuhn K, Hochstrasser DF, Burkhard PR, Sanchez JC (2008) Relative quantification of proteins in human cerebrospinal fluids by MS/MS using 6-plex isobaric tags. Anal Chem 80:2921–2931
126. Holman SW, Sims PF, Eyers CE (2012) The use of selected reaction monitoring in quantitative proteomics. Bioanalysis 4:1763–1786
127. Wolf-Yadlin A, Hautaniemi S, Lauffenburger DA, White FM (2007) Multiple reaction monitoring for robust quantitative proteomic analysis of cellular signaling networks. Proc Natl Acad Sci U S A 104:5860–5865
128. Duncan MW, Aebersold R, Caprioli RM (2010) The pros and cons of peptide-centric proteomics. Nat Biotechnol 28:659–664
129. Addona TA, Abbatiello SE, Schilling B, Skates SJ, Mani DR, Bunk DM, Spiegelman CH, Zimmerman LJ, Ham AJ, Keshishian H, Hall SC, Allen S, Blackman RK, Borchers CH, Buck C, Cardasis HL, Cusack MP, Dodder NG, Gibson BW, Held JM, Hiltke T, Jackson A, Johansen EB, Kinsinger CR, Li J, Mesri M, Neubert TA, Niles RK, Pulsipher TC, Ransohoff D, Rodriguez H, Rudnick PA, Smith D, Tabb DL, Tegeler TJ, Variyath AM, Vega-Montoto LJ, Wahlander A, Waldemarson S, Wang M, Whiteaker JR, Zhao L, Anderson NL, Fisher SJ, Liebler DC, Paulovich AG, Regnier FE, Tempst P, Carr SA (2009) Multi-site assessment of the precision and reproducibility of multiple reaction monitoring-based measurements of proteins in plasma. Nat Biotechnol 27:633–641
130. Koh GC, Porras P, Aranda B, Hermjakob H, Orchard SE (2012) Analyzing protein-protein interaction networks. J Proteome Res 11:2014–2031
131. De Las Rivas J, Fontanillo C (2010) Protein-protein interactions essentials: key concepts to building and analyzing interactome networks. PLoS Comput Biol 6:e1000807
132. Cavanagh J, Thompson R, Bobay B, Benson LM, Naylor S (2002) Stoichiometries of protein-protein/DNA binding and conformational changes for the transition-state regulator AbrB measured by pseudo cell-size exclusion chromatography-mass spectrometry. Biochemistry 41:7859–7865
133. Wen J, Arakawa T, Philo JS (1996) Size-exclusion chromatography with on-line light-scattering, absorbance, and refractive index detectors for studying proteins and their interactions. Anal Biochem 240:155–166
134. Mayer CL, Snyder WK, Swietlicka MA, Vanschoiack AD, Austin CR, McFarland BJ (2009) Size-exclusion chromatography can identify faster-associating protein complexes and evaluate design strategies. BMC Res Notes 2:135
135. Berkowitz SA (2006) Role of analytical ultracentrifugation in assessing the aggregation of protein biopharmaceuticals. AAPS J 8:E590–E605
136. Phizicky EM, Fields S (1995) Protein-protein interactions: methods for detection and analysis. Microbiol Rev 59:94–123

137. Sokolowska I, Woods AG, Gawinowicz MA, Roy U, Darie CC (2012) Identification of potential tumor differentiation factor (TDF) receptor from steroid-responsive and steroid-resistant breast cancer cells. J Biol Chem 287:1719–1733
138. Ngounou Wetie AG, Sokolowska I, Woods AG, Roy U, Loo JA, Darie CC (2013) Investigation of stable and transient protein-protein interactions: past, present, and future. Proteomics 13(3–4):538–557
139. Sokolowska I, Woods AG, Gawinowicz MA, Roy U, Darie CC (2012) Identification of a potential tumor differentiation factor receptor candidate in prostate cancer cells. FEBS J 279:2579–2594
140. Chautard E, Fatoux-Ardore M, Ballut L, Thierry-Mieg N, Ricard-Blum S (2011) MatrixDB, the extracellular matrix interaction database. Nucleic Acids Res 39:D235–D240
141. Pflieger D, Gonnet F, de la Fuente van Bentem S, Hirt H, de la Fuente A (2011) Linking the proteins—elucidation of proteome-scale networks using mass spectrometry. Mass Spectrom Rev 30:268–297
142. Suter B, Kittanakom S, Stagljar I (2008) Two-hybrid technologies in proteomics research. Curr Opin Biotechnol 19:316–323
143. Suter B, Kittanakom S, Stagljar I (2008) Interactive proteomics: what lies ahead? Biotechniques 44:681–691
144. Krause F (2006) Detection and analysis of protein-protein interactions in organellar and prokaryotic proteomes by native gel electrophoresis: (Membrane) protein complexes and supercomplexes. Electrophoresis 27:2759–2781
145. Sokolova L, Wittig I, Barth HD, Schagger H, Brutschy B, Brandt U (2010) Laser-induced liquid bead ion desorption-MS of protein complexes from blue-native gels, a sensitive top-down proteomic approach. Proteomics 10:1401–1407
146. Darie CC, Deinhardt K, Zhang G, Cardasis HS, Chao MV, Neubert TA (2011) Identifying transient protein-protein interactions in EphB2 signaling by blue native PAGE and mass spectrometry. Proteomics 11:4514–4528
147. Heck AJ, Van Den Heuvel RH (2004) Investigation of intact protein complexes by mass spectrometry. Mass Spectrom Rev 23:368–389
148. Kaddis CS, Lomeli SH, Yin S, Berhane B, Apostol MI, Kickhoefer VA, Rome LH, Loo JA (2007) Sizing large proteins and protein complexes by electrospray ionization mass spectrometry and ion mobility. J Am Soc Mass Spectrom 18:1206–1216
149. Petyuk VA, Qian WJ, Chin MH, Wang H, Livesay EA, Monroe ME, Adkins JN, Jaitly N, Anderson DJ, Camp DG II, Smith DJ, Smith RD (2007) Spatial mapping of protein abundances in the mouse brain by voxelation integrated with high-throughput liquid chromatography-mass spectrometry. Genome Res 17:328–336
150. Qian WJ, Petritis BO, Kaushal A, Finnerty CC, Jeschke MG, Monroe ME, Moore RJ, Schepmoes AA, Xiao W, Moldawer LL, Davis RW, Tompkins RG, Herndon DN, Camp DG II, Smith RD (2010) Plasma proteome response to severe burn injury revealed by 18O-labeled "universal" reference-based quantitative proteomics. J Proteome Res 9:4779–4789
151. Zhong Y, Hyung SJ, Ruotolo BT (2012) Ion mobility-mass spectrometry for structural proteomics. Expert Rev Proteomics 9:47–58
152. Sali A, Glaeser R, Earnest T, Baumeister W (2003) From words to literature in structural proteomics. Nature 422:216–225
153. Jurneczko E, Barran PE (2011) How useful is ion mobility mass spectrometry for structural biology? The relationship between protein crystal structures and their collision cross sections in the gas phase. Analyst 136:20–28
154. Scarff CA, Thalassinos K, Hilton GR, Scrivens JH (2008) Travelling wave ion mobility mass spectrometry studies of protein structure: biological significance and comparison with X-ray crystallography and nuclear magnetic resonance spectroscopy measurements. Rapid Commun Mass Spectrom 22:3297–3304
155. Roberts GC, Smith CW (2002) Alternative splicing: combinatorial output from the genome. Curr Opin Chem Biol 6:375–383

156. Paulo JA, Kadiyala V, Banks PA, Steen H, Conwell DL (2012) Mass spectrometry-based proteomics for translational research: a technical overview. Yale J Biol Med 85:59–73
157. Angel TE, Aryal UK, Hengel SM, Baker ES, Kelly RT, Robinson EW, Smith RD (2012) Mass spectrometry-based proteomics: existing capabilities and future directions. Chem Soc Rev 41:3912–3928
158. Pierce KL, Premont RT, Lefkowitz RJ (2002) Seven-transmembrane receptors. Nat Rev Mol Cell Biol 3:639–650
159. Das N, Biswas B, Khera R (2013) Membrane-bound complement regulatory proteins as biomarkers and potential therapeutic targets for SLE. Adv Exp Med Biol 734:55–81
160. McMahon G (2000) VEGF receptor signaling in tumor angiogenesis. Oncologist 5(Suppl 1): 3–10
161. Zwick E, Bange J, Ullrich A (2001) Receptor tyrosine kinase signalling as a target for cancer intervention strategies. Endocr Relat Cancer 8:161–173
162. Whitelegge J, Halgand F, Souda P, Zabrouskov V (2006) Top-down mass spectrometry of integral membrane proteins. Expert Rev Proteomics 3:585–596
163. Souda P, Ryan CM, Cramer WA, Whitelegge J (2011) Profiling of integral membrane proteins and their post translational modifications using high-resolution mass spectrometry. Methods 55:330–336

Chapter 2
MALDI Profiling and Applications in Medicine

Ed Dudley

Abstract Matrix-assisted laser desorption ionisation (MALDI) mass spectrometry allows for the rapid profiling of different biomolecular species from both biofluids and tissues. Whilst originally focussed upon the analysis of intact proteins and peptides, MALDI mass spectrometry has found further applications in lipidomic analysis, genotyping, micro-organism identification, biomarker discovery and metabolomics. The combining of multiple profiles data from differing locations across a sample, furthermore, allows for spatial distribution of biomolecules to be established utilising imaging MALDI analysis. This chapter discusses the MALDI process, its usual applications in the field of protein identification via peptide mass fingerprinting before focusing upon advances in the application of the profiling potential of MALDI mass spectrometry and its various applications in biomedicine.

2.1 MALDI Mass Spectrometry

Matrix-assisted laser desorption ionisation (MALDI) mass spectrometry (MS) represents an ionisation process first described in the mid-1980s and was developed by two independent research groups to the point at which MALDI mass spectrometers became commercially available in the early 1990s. The MALDI process involves the mixture of the analyte in solution with a "matrix" solution and co-crystallisation of both matrix and analyte together. During ionisation a laser (usually with a UV-nanometre wavelength) is fired at the mixture and causes desorption of the

E. Dudley (✉)
Institute of Mass Spectrometry, College of Medicine, Swansea University, Singleton Park, Swansea, UK
e-mail: e.dudley@swansea.ac.uk

A.G. Woods and C.C. Darie (eds.), *Advancements of Mass Spectrometry in Biomedical Research*, Advances in Experimental Medicine and Biology 806,
DOI 10.1007/978-3-319-06068-2_2, © Springer International Publishing Switzerland 2014

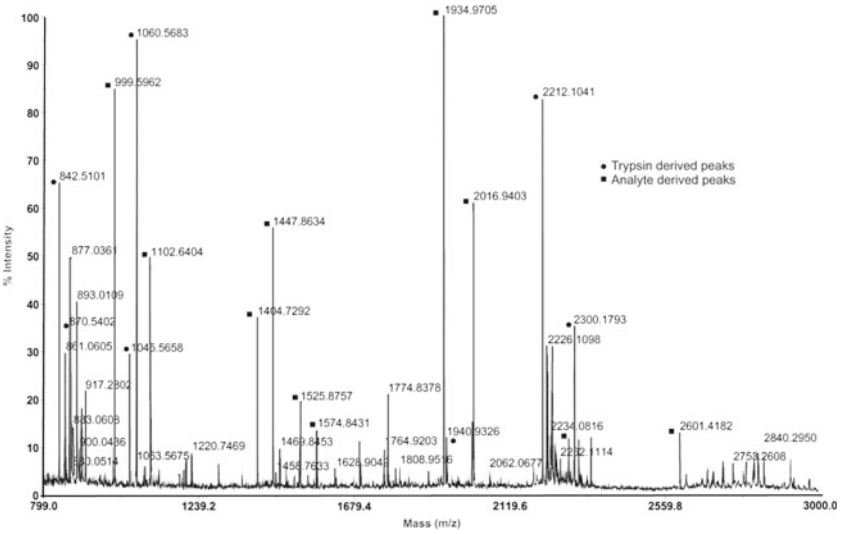
matrix and analyte into the gas phase. It is thought that the matrix is protonated first and thus protonates the analyte during the MALDI process. One major advantage of MALDI over other forms of ionisation, such as electrospray ionisation (ESI), is that usually singly charged ions are detected for the analyte even at increased mass/charge ratios. Comparatively ESI generates many multiply charged ion species for higher molecular weight protonated molecules which complicate the mass spectra generated and require deconvolution of the mass spectra in order to determine the molecular weight of the protonated molecule. MALDI ion sources are usually combined with time of flight (ToF) tubes for ion separation. The ToF tubes are utilised either in linear mode (in which ions pass once the full length of the ToF tube) or in reflectron mode (in which ions are reflected at the far end of the ToF tube and so pass through two times the length of ToF tube before reaching the detector). Reflectron mode analysis will provide a greater mass resolution but is limited to 5–10 kDa *m*/*z* ions whilst linear mode analysis can detect masses over 100 kDa and has a greater sensitivity. Considerations when undertaking a MALDI MS experiment are the choice of matrix to be mixed with the sample, how the sample and matrix solutions are mixed and the solvents used to solubilise both matrix and analyte. With reference to protein/peptide profiling and analysis, α-cyanohydroxycinnaminic acid (CHCA) is the most commonly utilised matrix for peptide analysis by MALDI MS whilst proteins are thought to ionise more effectively using sinapinic acid (SA) as the matrix. Both are used in concentrations of around 10 mg/mL and therefore represent a considerable excess of matrix molecules to analyte molecules in the resulting co-crystallised mixture. As a result of this the lower end of the mass spectrum (less than 800 Da) is dominated by ions arising from the matrix itself and is therefore usually not studied during analysis. The solvent used for solubilising the matrix and mixing with analyte also differs between peptide and protein analysis with sinapinic acid dissolved in a solvent with a lesser acetonitrile content for protein analysis compared to CHCA for peptide analysis. The main purpose of this is to prevent proteins potentially precipitating due to the acetonitrile present which would prevent effective co-crystallisation of matrix and analyte. More recently, novel matrix materials have been investigated in order to determine whether improvements could be gained in ionisation efficiency. Much of the research has focussed on using alternative matrices in order to avoid low *m*/*z* ions which arise from the matrix itself and therefore extend the mass window of analysis for MALDI MS further into lower *m*/*z* regions. For this purpose graphene (a few layers of graphite) was utilised and allowed for the analysis of nonpolar polymethylmethacrylate—approximately 650 Da [12], whilst hydroxyflavones allowed for low molecular weight lipid profiling and also required less laser energy for efficient ionisation of the analytes [90]. Further advances in the analysis of peptides by MALDI MS have considered whether improvements in sensitivity can be accomplished by chemical derivatisation of the peptides prior to analysis. 1-(3-Aminopropyl)-3-butylimidazolium bromide (BAPI) was utilised to derivatise the C terminal carboxyl groups of peptides prior to analysis and was reported to result in a 42-fold improvement in sensitivity and the derivatisation product was demonstrated to be stable for at least 1 week [78]. As an alternative, the phosphorylation of the N terminus of the peptides was also investigated as this

was selective for the N terminal lysine of tryptic peptides and due to the improved proton affinity provided by the additional phosphate group attached to the peptides an increase in sensitivity was again reported [25].

In relation to proteomic analysis, the most common application of MALDI MS is in undertaking peptide mass fingerprinting in order to identify an unknown protein and whilst this is not the topic of this chapter it is worth quickly reviewing. Peptide mass fingerprinting analyses the tryptic digest (usually) of a single protein and measures the *m*/*z* of the resulting tryptic peptides accurately (within at least 10 ppm). The resulting mass list of peptide *m*/*z* values can then be compared to a database of theoretical protein digest endproduct peptides and their singly charged *m*/*z* values and a protein identification suggested alongside a probability score for the identification—an example is given in Fig. 2.1. More recently, MALDI sources have been attached to ToF–ToF mass analysers which allow tandem mass spectrometry to be undertaken on peptides. This generates fragment ions (B and Y ions) alongside the mass of each peptide and therefore allows for more accurate identification of more complex protein mixtures; however, this area is not the focus of this chapter. In this chapter, we discuss the application of MALDI MS analysis in providing profiles of peptides and proteins from biological samples and how this type of profiling analysis has been more recently expanded to nucleic acid and metabolite analysis.

2.2 Sample Preparation for Protein/Peptides for Protein Profiling

Protein profiling has been applied to various biological fluids; however, the majority of studies have utilised plasma or serum as their starting material. Whilst serum and plasma are reasonably easy to collect in a clinical setting they do present the researcher with a large dynamic range of protein concentrations with up to 70 % of the protein present being represented by serum albumin and up to 25 % comprising globulin proteins [56]. These high abundance proteins are of little value for profiling purposes and can obscure the ability of analytical techniques to study low abundance proteins within the sample and therefore various sample preparation protocols have been developed and tested in order to be able to study lower abundance proteins by MALDI MS protein profiling. One obvious solution to the issue of high abundance proteins is to selectively remove these from the sample prior to analysis. A recent study compared immunodepletion columns that allowed for the removal of the 7 and 14 most abundant serum proteins prior to MALDI MS analysis and found that the removal of the 14 most abundant proteins and application of 2′,4′-dihydroxyacetophenone as the matrix utilised for ionisation allowed for the most reliable protein profile to be produced [21]. One issue regarding such immunodepletion columns is their cost and the time consuming nature of the protocol. One column can usually be used to deplete up to 200 serum samples one at a time and each depletion requires a number of steps to prepare the column, deplete the sample,



Protein identification: Phosphoglycerate kinase

m/z submitted	m/z matched	Delta ppm	Sequence of peptide
999.5959	999.5991	3.1855	(K) FSLAPLVPR (L)
1102.6402	1102.637	2.6429	(K) RPFAAIVGGSK (V)
1220.7472	1220.747	0.6248	(K) VILSTHLGRPK (G)
1404.7299	1404.737	5.3725	(K) ELDYLVGAVSNPK (R)
1447.8636	1447.867	2.8167	(K) FLKPSVAGFLLQK (E)
1525.8751	1525.884	5.9406	(K) GVSLLLPTDVVVADK (F)
1573.8428	1573.843	0.6144	(K) GVTTIIGGGDSVAAVEK (V)
1933.9712	1933.977	3.0853	(K) LASLADLYVNDAFGTAHR (A)
2015.9488	2015.952	1.6474	(R) ADLNVPLDDNQTITDDTR (I)
2285.1848	2285.181	1.5930	(K)VGVAGVMSHISTGGGASLELLEGK (V)
2600.4121	2600.403	3.3117	(K) AQGLSVGSSLVEEDKLELATELLAK (A)

Fig. 2.1 An example of peptide mass fingerprinting spectra and the interpretation returned upon database searching

elute bound proteins and regenerate the column prior to the next sample being processed. A further complication may arise due to potential co-depletion of other proteins associated with albumin in serum samples [5]. The most common alternative approaches to serum preparation prior to MALDI MS analysis are to utilise

microscale tips with peptide/protein binding stationary phases incorporated or magnetic beads with a range of immobilised chemistries. The most commonly applied technologies are C18 reverse phase stationary phases for microscale tip purification which allows usually for the study of low molecular weight peptides and C8 reverse phase magnetic bead purification of low-to-medium molecular weight proteins. Whilst reverse phase stationary phases are the most commonly utilised for these types of sample preparation, other chemistries are available which exhibit differing affinities for different subsets of proteins. These include cation and anion exchange stationary phases, mixed ion exchange resins, immobilised lectin stationary phases (for purification of glycopeptides/proteins) and immobilised metal affinity technologies (for purification of phosphopeptides/proteins and anionic moieties). Example MALDI spectra after microscale tip and bead separation for peptide and protein profiling, respectively, are given in Fig. 2.2. In theory, the combination of all of these different preparation approaches would allow for the fractionation of serum samples and a more comprehensive survey of proteins present by MALDI MS analysis would result. These technologies have not only been applied to serum but also been optimised for the analysis of peptide/protein profiles from other biofluids. A recent study optimised the application of reverse phase microscale tip purification of peptides from serum and found that reproducible protein profiles detailing over 400 individual peaks could be determined [9]. Whilst the application of magnetic bead sample preparation to urinary peptide/protein profiling analysis noted that for ion exchange and immobilised metal purification the pH of the urine to 7 was required in order to allow reproducible recovery of analytes and whilst some inter-day variation was noted between samples, this variability was reduced by normalising the protein content of the urinary sample purified to set amount of protein (3.5 μg per sample) [22]. Similarly, magnetic bead purification protocols were applied to profiling of cerebrospinal fluid and it was determined that samples were stable for 6 h at room temperature and 3 days at 4 °C prior to purification and analysis; however, contamination with high albumin or globulin levels had a detrimental effect on the protocol and resulting mass spectra [8]. Of the two techniques (microscale tip and magnetic bead purification), the microscale tips have been reported as offering an improved and more reproducible protein exhibiting an average CV of 10 % of the recorded relative abundance for over 100 peptide/protein peaks in any given profile [84]. As an alternative approach, the fact that the majority of the high abundance proteins in serum are of comparatively high molecular weight has been used as a characteristic upon which to base a depletion protocol. Ultrafiltration of serum with filters with molecular weight cut-offs of 50 and 30 kDa has been applied as a methodology in order to allow for the study of low abundance proteins/peptides [3]. Such protocols commonly dilute the serum with acidic solutions (2 % trifluoroacetic acid for example) in order to dissociate peptides and proteins bound to albumin. Further processing of the sample after ultracentrifugation, such as desalting, has been reported to further improve protein coverage especially within the 3–20 kDa mass range [4, 26]. As an alternative approach, differential precipitation utilising ammonium sulphate and different organic solvents was tested for the preparation of calf serum for MALDI analysis [85]. The different fractions obtained presented different

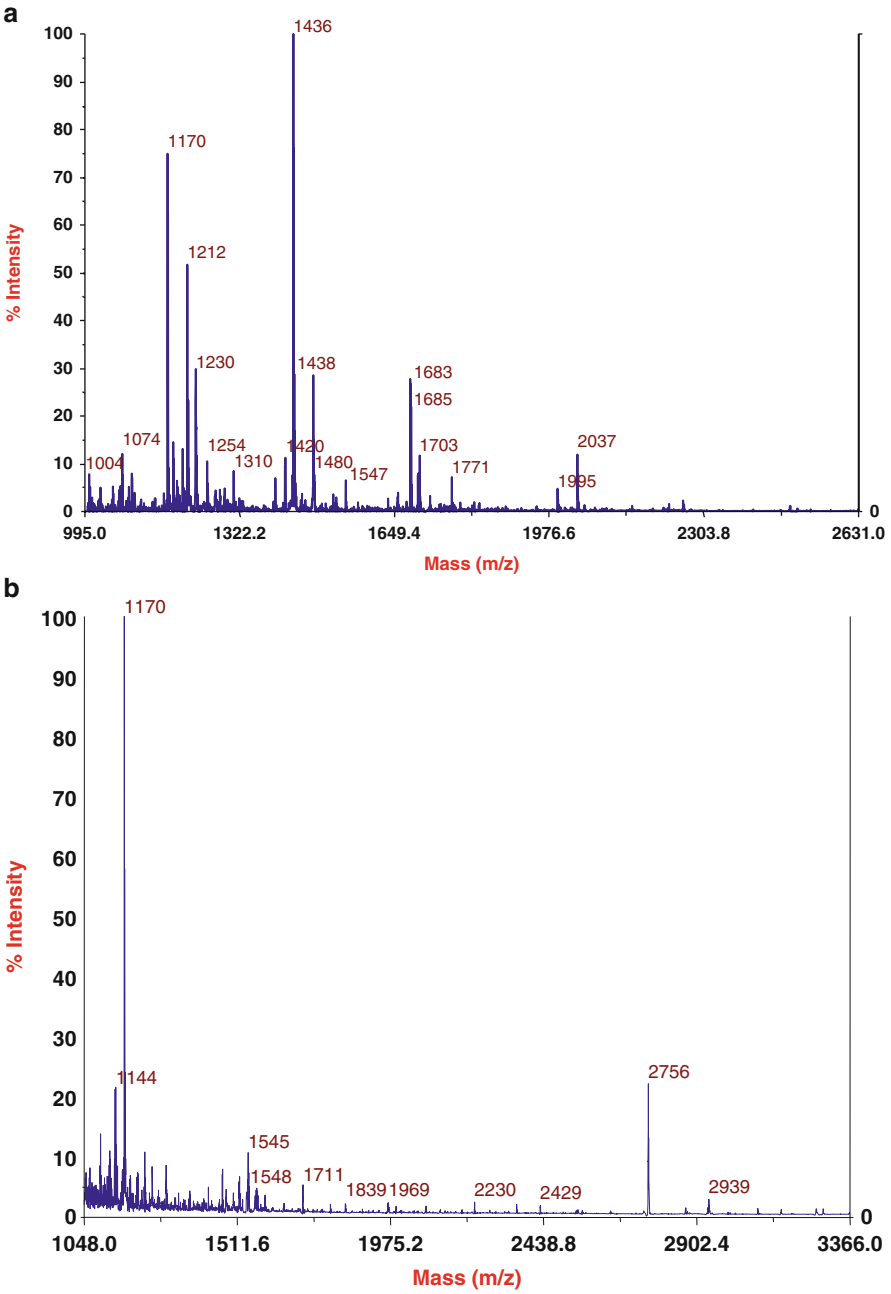


Fig. 2.2 MALDI protein and peptide profiling. (**a**) Global peptide profiling. (**b**) MALDI phosphopeptide profiling. (**c**) Global protein profiling. (**d**) MALDI phosphoprotein profiling

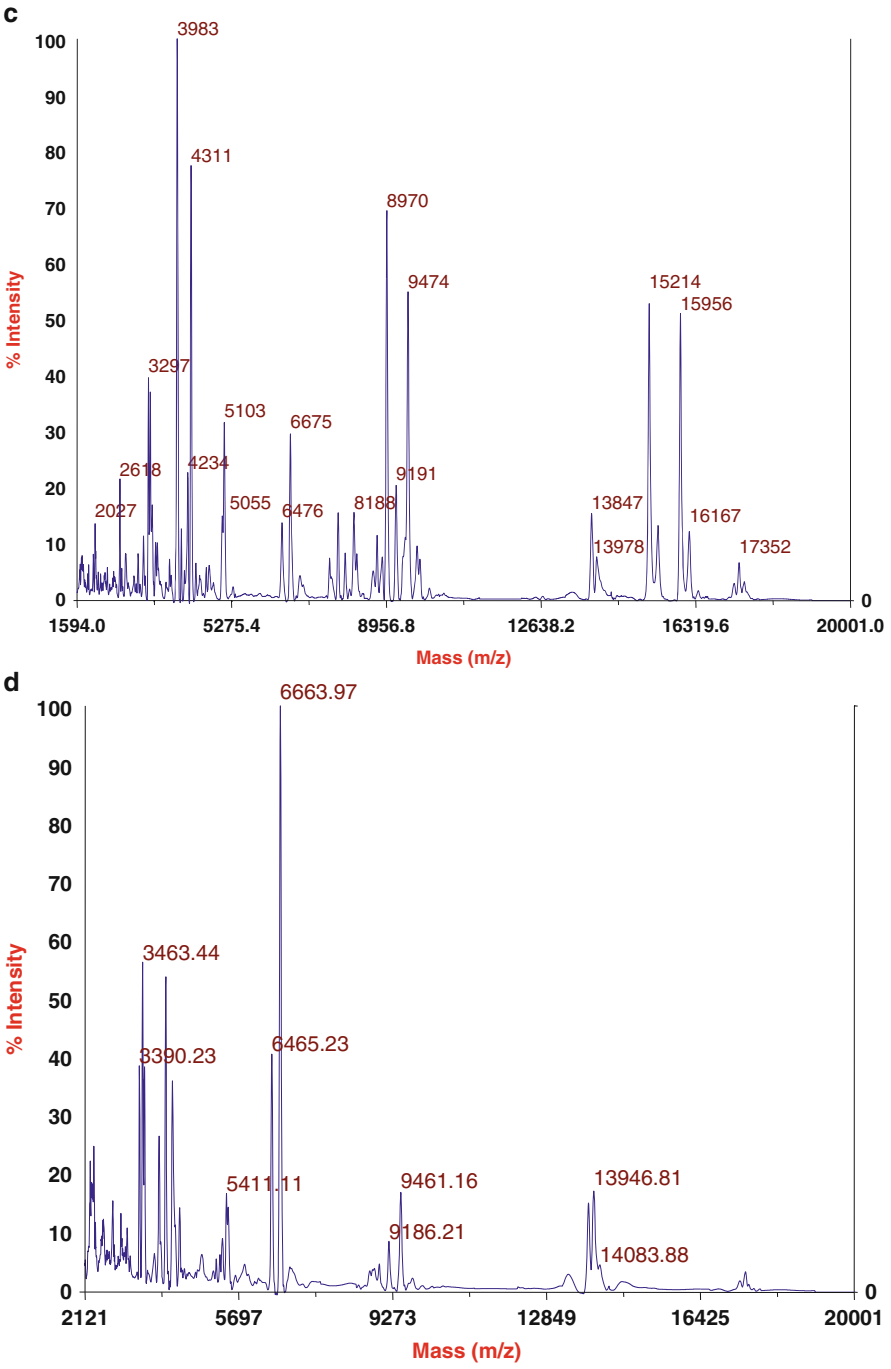


Fig. 2.2 (continued)

peptide profiles, numbering in the hundreds with optimal signal-to-noise ratios being derived from samples resulting from water solubilisation after organic solvent precipitation. Similarly, acetonitrile precipitation was shown to remove more than 99 % of the protein from serum and reduced the presence of albumin in serum from being the most abundant protein to being the 20th most abundant [43]. Further treatment of the precipitated serum sample by acid hydrolysis has also been shown to improve detection of major serum peptides (transferrin and fibrinogen) as well as other peptide species between 4 and 10 kDa [51]. Other approaches involve preparation of the sample when applied to the sample plate and alternative ionisation means other than MALDI being applied. One such method describes nanostructure initiator mass spectrometry (NIMS) in which cysteine containing peptides are captured by a maleimide chemistry and ionised without the need for a matrix to be applied, thereby reducing background noise signals and improving sensitivity [50]. An alternative, but related, approach to MALDI protein profiling is surface-enhanced laser desorption ionisation (SELDI) mass spectrometry. In SELDI preparation protocols, the biofluid in question is applied directly to the SELDI plate (equivalent to the MALDI sample plate) which has a surface chemistry similar to that found on magnetic beads and microscale tips as described previously. The sample is then washed on plate removing salts and unbound proteins and the remaining proteins analysed and presented either as raw spectra data or as interpreted gel-like images prior to statistical analysis. SELDI is most commonly utilised to identify differential protein peaks as prospective biomarkers; however, it has also been applied to the validation of previously determined HPLC–mass spectrometry results as a rapid validation method for studying levels of specific protein biomarkers in breathe condensate as indicators of chronic obstructive pulmonary disease [24]. Furthermore, SELDI has been applied in tandem with magnetic bead sample preparation utilising serum as the starting biofluid to improve peptide/protein detection [76] and advances in statistical data analysis have attempted to address some determined issues of reproducibility and accurate biomarker identification when using SELDI analysis [18].

2.3 Protein Profiling and MALDI Imaging

After sample preparation, a common application of protein profiling from biofluids is the characterisation of peptide and protein profiles followed by statistical examination of the profiles in order to determine any potentially diagnostic differences between different cohorts that might represent biomarkers of any condition under study. The analysis itself does not usually allow for the identification of the protein from which any diagnostic protein or peptide peaks arise and so only the m/z ratio of any prospective biomarker is determined by the analysis. Despite this, protein profiling of biofluids has been applied successfully in a wide range of biomedical areas. The approach has been applied widely in cancer biomarker research and protein/peptide peaks have been identified which allow for the discrimination of patients from control cohorts but also differentiate pre- and post-operative patients

and those at risk of metastasis and reoccurrence [88]. Furthermore, MALDI protein profiling of urine of patients allowed for the identification of proteins whose urinary expression could be used to identify patients at risk of liver injury due to chemotherapeutic treatments such as methotrexate [86]. Another area of well-studied interest is the application of such protocols in diabetes and kidney-related disorders. Magnetic bead preparation and MALDI analysis of urine samples were successfully shown to be able to identify three proteins/peptides with reduced expression in type two diabetes patients and further mass spectrometric analysis allowed these to be identified as three proteins without a known previous association with the disorder [13]. Peptide profiling was also applied to patients with primary nephrotic syndrome and relevant control groups and a group of 14 peptide ions were shown to have good diagnostic value offering the potential of replacing a time consuming and invasive biopsy procedure for diagnosis [34]. The analysis has also been utilised in order to further understand the potential beneficial effects of drug regimes utilised after kidney transplant. The effect of the drug paricalcitol on the serum protein profile was determined and the protein expression changes suggested to represent alterations in the expression levels of parathyroid hormone, alkaline phosphatase, bradykinin and complement factor C4, thereby suggesting secondary effects of the drugs administration [64]. More recently, protein and peptide profiling by MALDI mass spectrometry has also been applied to conditions for which the biological cause is less certain and/or variable in different cases. One key example of this is the application of the technique in psychiatric disorder diagnosis. The principle of applying different approaches to the preparation of the biofluid has been utilised predominantly to study serum protein and peptide changes. One study utilised acid hydrolysis in order to allow for peptide profiling to be undertaken in depression patients and control subjects whilst others have utilised both magnetic bead and microscale tip preparations for the same purpose [1, 35, 51]. Both allowed for the identification of different prospective biomarkers for depression with significant area under the curve values determined after the creation of receptor operator characteristic (ROC) curves. The latter analysis identified three peptide ions which had diagnostic value and whilst the accumulated relative abundance of the three biomarkers did not further improve diagnostic ability, the combined study of all three with alterations in two of the three being used as a diagnostic tool significantly improved diagnostic accuracy. This finding demonstrates the potential benefit of the MALDI peptide/protein profiling approach over techniques such as ELISAs which can only determine the expression of a single protein. The collation of the entire protein profile in a single analysis offers the utilisation of the expression of multiple species within the profile to be utilised without further analysis being undertaken and therefore various expression levels can be studied and ratios of individual proteins/peptides also utilised. Protein profiling after magnetic bead sample preparation has also been utilised in other psychiatric disorders; the technique was utilised for the analysis of serum samples from adolescent patients with autism spectrum disorder (ASD) with and without co-morbid attention deficit hyperactivity disorder (ADHD) compared to control subjects. Protein signatures were identified which differentiated all patients from control subjects and also could differentiate the ASD patients into two

groups (with and without co-morbid ADHD) [82]. A recent study considered the differentiation of obese subjects compared to control subjects utilising ultrafiltration to remove high abundance proteins and identified peptide signatures which differentiated not only patients from control subjects but also patients with and without diabetic symptoms. Such analyses suggest that the process of an unbiased protein profiling experiment can allow for the discrimination of patients with disorders whose biological cause is less certain and also allow for the differentiation of similar cohorts with different disorder diagnoses.

Whilst the majority of analyses in biomedical research utilising MALDI protein and peptide profiling have studied biofluids after sample preparation of one form or other, some have studied biological tissues and utilised a process of MALDI imaging mass spectrometry for spatial analysis of protein and peptide signals. The process of MALDI imaging requires spectra to be taken from tissues applied to a MALDI sample plate after the incorporation of a suitable matrix. Multiple spectra are obtained with high precision across the tissue and the data collated in order to demonstrate the distribution of proteins and peptides within different cells within the tissue examined. Such analysis has been utilised to study protein/peptide differences in cancer tumour cells compared to adjacent "normal" cells from formalin-fixed tissue samples and identified specific signatures associated with early stage disease and tumours with metastatic potential [80]. The analysis has also been used to determine the altered regulation of candidate proteins involved in cancer progression identified by other studies with a resolution of 30 μm [52]. Whilst cellular proteins can be easily determined by the imaging MALDI approach, it has been reported that membrane proteins are less well represented by such analyses and one study utilised a pretreatment approach to overcome this issue, allowing for the study of two known membrane proteins and their acylation status within tissues [60]. Whilst most imaging MALDI analysis allows for a single matrix to be applied and therefore one set of data to be obtained a recent approach utilised a frozen mouse brain section to be analysed with two matrices—one utilised to study proteins/peptides and a second used to emphasise lipidomic data—thereby allowing multiple analyses to be undertaken and more information to be extracted from a single tissue subsection [74]. Protein profiling from both biofluids and tissues therefore allows for diagnostic differences to be determined and whilst the biological identity of the protein or peptide in question is not obtained from the data, the unbiased analysis of profiles allows for novel biomarkers to be identified within biomedicine.

2.4 Post-translational Modification Analysis

A further application of MALDI protein and peptide profiling is in the analysis of the post-translational modification status of either the entire proteome or selected proteins or peptides. Of the available post-translational modifications, the two which have shown the greatest application to MALDI profiling are glycosylation and phosphorylation; however, the latter of these is more commonly studied utilising

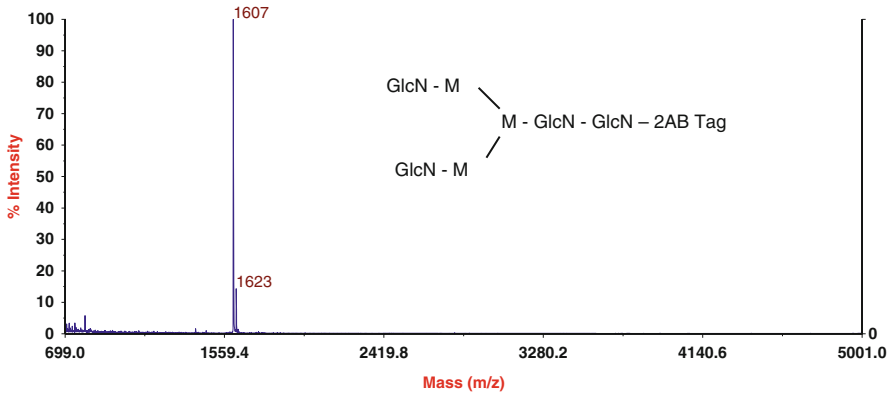


Fig. 2.3 MALDI analysis of an AB-modified glycan

HPLC–mass spectrometric methods and selective tandem mass spectrometry analysis such as constant neural loss and precursor ion discovery. Glycosylation therefore has shown the greatest application with regard to MALDI protein and peptide profiling, with the availability of magnetic bead technologies with immobilised lectin chemistries in order to allow for the selective capture of glycoproteins and glycopeptides prior to MALDI mass spectrometric analysis. The glycans themselves can be analysed after derivatisation and release from the peptide with dihydroxybenzoic acid being utilised as a matrix as shown in Fig. 2.3. Further advances in the pre-analysis preparation of glycoproteins/peptides has also encompassed the utilisation of a microscale tip purification approach which uses a 50:50 mixture of graphite and activated graphite (termed GA) in order to enrich for glycopeptides from protein digests and the utilisation of oxidised ordered mesoporous carbon as a pre-analysis purification approach [79, 95]. Larger scale pre-purification of glycoproteins commonly utilises offline separation techniques such as high-performance anion exchange chromatography with pulsed amperometric detection (HPAEC-PAD) followed by fraction collection and further MALDI mass spectrometric analysis [40]. The matrix choice when analysing glycoproteins and peptides has also previously been studied and 4-chloro-cinnaminic acid suggested as a matrix for the analysis of labile groups in negative ionisation mode and a combination of CHCA and 3-aminoquinoline as a matrix applicable to complex mixture analysis when studying glycoproteome samples [68, 91]. For comparative analysis of glycoproteins between different samples a simple derivatisation method has also been developed in which glycoproteins are reacted with differently labelled anthranilic acid and then mixed, digested and analysed by MALDI profiling (in a similar manner to ITRAQ technology) with sub-picomolar limits of detection [83]. Furthermore, comparative analysis of the glycosylation modification structure itself is also possible either by separation from its proteins/peptides and chemical reaction or via in-source decay processes being utilised during the MALDI process [31, 57]. The differential glycosylation has been studied in a number of diseases, including congenital orders specifically of glycosylation and ovarian cancer [7, 94]. In the ovarian cancer study, the

glycans themselves were cleaved from the glycoproteins and analysed by MALDI profiling and these profiles exhibited improved diagnostic potential when compared to the commonly utilised CA-125 biomarker. Further applications have also been reported for the MALDI analysis of glycoproteins within a biomedical context. The glycosylation status of proteins from stem cells during differentiation to adipose tissue (adipogenic differentiation) was studied and specific changes in the glycoproteome identified which accurately allowed the differentiation event to be monitored within the cells [30]. Other studies have focussed upon different glycosylation forms of specific proteins of interest. The application of high accurate mass MALDI profiling after purification allowed for the identification of six new apolipoprotein CIII isoforms and their glycosylation status which may have a role in lipid metabolism and transport within the body and ApoCIII isoforms have been identified as easy to purify by microscale tip preparation and able to allow for diagnoses in chronic hepatitis C and alcoholic liver cirrhosis [61]. A different approach sought to study the glycosylation status of membrane-type 1 matrix metalloproteinase (MT1-MMP) via MALDI mass spectrometric analysis after immunoprecipitation from cell lines as the glycosylation is required for cancer cell invasive processes [81]. The glycosylation status of haemoglobin has also been undertaken by MALDI profiling protocols encompassing the study of the non-enzymatic glycosylation as a result of diabetes and also differential glycosylation of different isoforms of haemoglobin suggesting that the isoform presented by individuals may act as a risk factor for glycation occurring in prospective patients of diabetes [15, 47].

The second commonly analysed post-translational modification of proteins utilising MALDI protein/peptide profiling is phosphorylation status. Numerous approaches to MALDI matrix preparation and application have been suggested as enhancing phosphopeptide profiling. Combining the matrix of 2,6-dihydroxyacetophenone (DHAP) or 2′,4′,6′-trihydroxyacetophenone (THAP) with the additive diammonium hydrogen citrate (DAHC) has been investigated as to their potential to improve phosphopeptide analysis and suggested less suppression by non-phosphorylated peptides during profiling without selective purification of phosphopeptides [33, 96]. However different groups have reported improved signals utilising a further matrix (dihydroxybenzoic acid) in combination with phosphoric acid as a matrix additive or combinations of two matrices (CHCA and 3-hydroxypinolinic acid (3-HPA)) [45, 100]. Further approaches to enhance phosphoprotein and peptide analysis by MALDI profiling have focussed on chemically altering the MALDI plate in order to selectively trap phosphate containing species and these have mainly utilised titanium dioxide as the linker to trap the phosphate group [89]. Phosphopeptide capture has been achieved on a sintered material before elution and subsequent MALDI profiling analysis and also by on-plate capture using photocatalytically patterned titanium dioxide distributions on plate using nanoparticles [20, 77]. Each of these utilises metal affinity to enhance the amount of phosphate containing peptide or protein in the sample prior to analysis and has been demonstrated to enhance coverage of expected phosphopeptides within tryptic digests of known phosphoproteins. Commonly, casein has been utilised as a test protein in the purification of phosphorylated peptides as it provides a known phosphopeptide in order to gauge the success of the purification technique (Fig. 2.4).

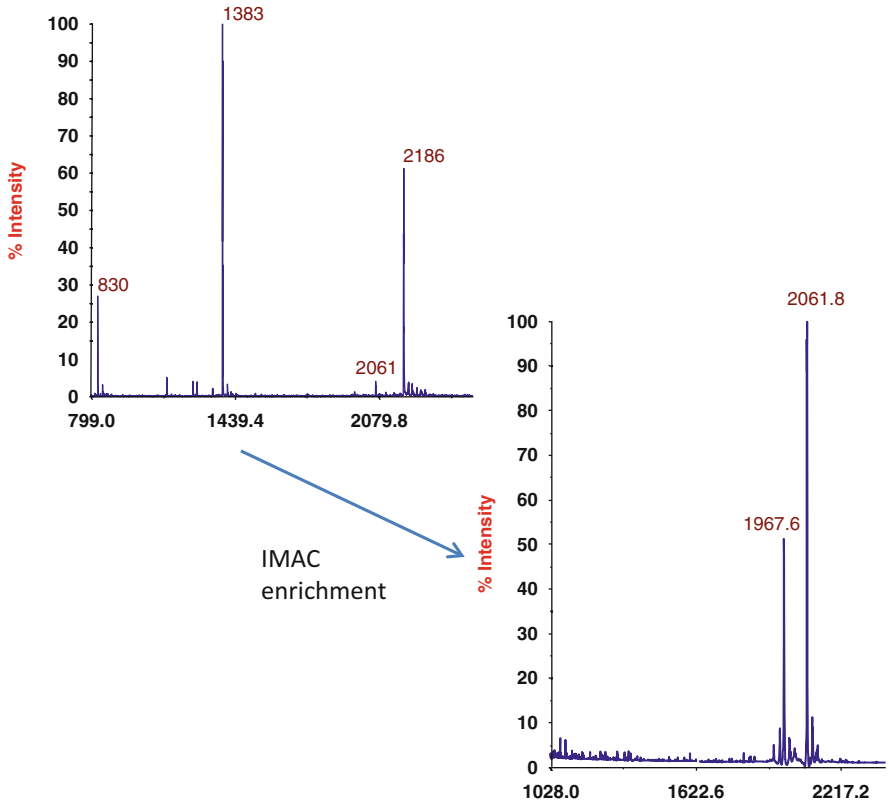


Fig. 2.4 MALDI spectra of tryptic peptides from casein before and after IMAC enrichment of phosphopeptides

2.5 Specific Protein Analysis

As well as protein profiling of entire proteomes, MALDI profiling has also been utilised in order to assay for specific proteins and variations in specific proteins in relation to different diseases. The approach has been applied to the study of diagnostic protein signatures for type 1 dengue virus, involving immunoaffinity capture on magnetic beads and analysis of the proteome by MALDI profiling allowing for differentiation of the virus from other similar viruses [10, 11]. Amyloid beta peptides have also been immunopurified (from cerebrospinal fluid) and the different isoforms studied by SELDI profiling allowing for the differentiation of 15 isoforms of the peptide of which the expression levels of many were indicative of cognitive decline in many age-related mental deterioration conditions [2]. MALDI profiling in combination with immunopurification has also been utilised to study the status of the circulating hormone Ghrelin [29]. The hormone can be acetylated and exists therefore in multiple forms and these could be studied and relative ratios determined by MALDI profiling following immunoprecipitation. In another study the

phosphorylation status of a specific protein (a target for protein kinase C) was studied from serum samples from mice with and without tumours utilising MALDI analysis and shown to provide diagnostic potential in relation to the ratio of phosphorylated and unphosphorylated peptide [41]. As well as studying natural protein modifications by MALDI profiling, the formation of organophosphorus adducts of human butyrylcholinesterase was also undertaken by MALDI analysis [37]. After exposure to the pollutant, a protein extraction was utilised and proteins digested, purified by titanium dioxide and analysed. The enzyme has the pollutant attached to it at a specific serine residue within its structure and therefore the bound and unbound mass of the tryptic peptide can be analysed by MALDI analysis as a ratio and an indicator of exposure to the pollutant.

2.6 Organism Identification

A further area in which MALDI protein profiling has been successfully applied is in the identification of organisms, primarily micro-organisms. In relation to micro-organism analysis, the protocol usually requires a sample of the micro-organism to be added to either a solvent or matrix solution. This is then added to the MALDI sample plate with additional matrix, allowed to dry and the MALDI process allows for the production of a protein profile from the organism itself. A number of groups and mass spectrometer manufacturers have developed databases of the protein profiles exhibited by known organisms and the protein profile of an unknown organism can therefore be compared to the database and an identification suggested as a result [16]. The approach has been reported as successfully identifying characteristics such as resistance in species and strain level identification within species as well as differentiating different species based on the protein profile obtained. Figure 2.5 shows spectra generated from the comparative analysis of four species of bacteria after simple plating on the MALDI plate with the addition of sinapic acid as a matrix. The process has been applied to various bacterial species including those responsible for acne, in which the differentiation of different phylotypes was also shown to be possible [59]. The method of sample preparation—in relation to both the culturing of the bacterial species and the subsequent preparation of the sample for MALDI analysis—has been investigated with broth media growth followed by a protein extraction method before analysis being reported as being optimal [27]. One group suggested that ribosomal proteins accounted for 10 of 13 species-specific peaks obtained by MALDI protein profiling and even strain-specific differences, confirming the expected mass shift by genome sequencing the genes responsible for the proteins themselves and identifying the amino acid differences resulting in the mass shifts exhibited [75]. A further study utilised a similar approach in order to confirm the identity of the peaks responsible for the differentiation of clonal strains of Staphylococcus aureus utilising sequencing and antisense RNA knockdown experiments in order to confirm the peptides responsible [38]. MALDI has also been shown to be able to differentiate enteropathogenic and

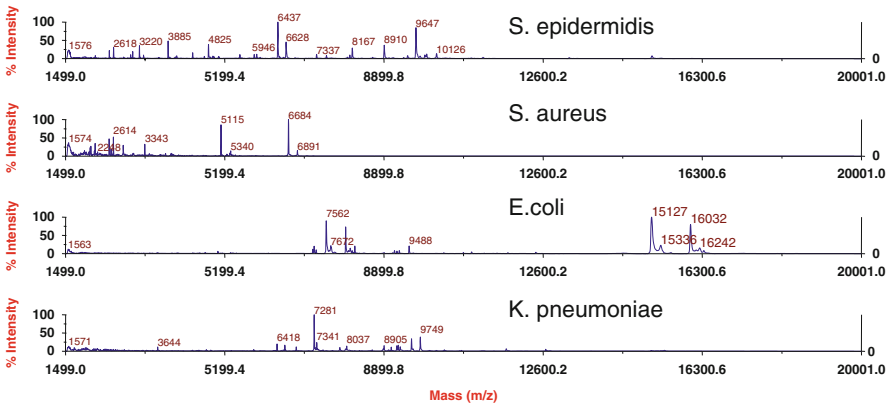


Fig. 2.5 MALDI protein profiles of four different bacteria

non-enteropathogenic strains of bacteria suggesting a role in fine-tuning the treatment of patients with bacterial infections [69]. Other studies have utilised blood culture bottles as the source of the micro-organisms under study and one study developed an optimised spin lysis/formic acid extraction method for the subsequent analysis [28, 49, 55, 71]. Such analyses would have a significant benefit to patients suffering from sepsis. However, a separate report identified a direct transfer of the sample to the plate and formic acid sample preparation as being optimal for subsequent analysis. The analysis was shown to be possible and both analysis from solid colonies and from blood cultures have been suggested to provide improvements in patient management in a hospital setting, being superior to existing blood culture methods and reducing analysis time [72].

The MALDI protein profiling approach has also been utilised to characterise rarer pathogenic bacteria and it was noted that of the various organisms studied fewer bacteria needed a second confirmation analysis when using MALDI compared to conventional approaches (50 compared to 620) [73]. Besides its application in identifying clinical bacterial infections, the MALDI protein profiling approach has also been reported in being able to confirm the identification of Bacillus anthracis spores in suspicious powder in less than half an hour [17]. Due to the promise suggested by such laboratory investigations inter-laboratory comparisons have been performed in order to show the reproducibility of the technique between centres, to compare competing MALDI mass spectrometers and commercial databases available and to compare the technique to existing methods of identification [23, 36, 42, 48, 53, 66, 93]. The method has also been applied to the identification of medically important fungal species such as candida, allowing for phenotypical differences to be determined in 1,383 isolates examined and also for antifungal susceptibility to be accurately determined, thereby allowing for a more informed treatment of patients [46, 87]. Beyond micro-organisms, the MALDI protein profiling approach has more recently been applied to other organisms. The identification of Tsetse (*Glossina* spp.) which are vectors of the so-called sleeping sickness was undertaken after formic

acid/acetonitrile extraction from the insects allowing for species identification but not sex determination [32]. Furthermore, a similar approach was used in the study of Anopheles mosquito species in order to accurately differentiate between differing subspecies and was shown to be an improved approach compared to the traditional polymerase chain reaction (PCR)-based techniques commonly applied [58]. MALDI analysis has also been utilised to study and characterise venoms from different species, mainly focused on scorpions. The so-called venom mass fingerprinting (VMF) was developed to study low levels of venom (nanograms of starting material) and specific toxin components, such as short chain peptides which act on potassium channels could be easily identified within venoms [54, 65].

2.7 Single-Nucleotide Polymorphism Analysis

Whilst a major focus of MALDI profiling has been concerned with the analysis of proteins and peptides as biomolecular species, the ability of MALDI to provide good mass accuracy signals for other biomolecules at high molecular weight has also been studied. One of the major focuses of this has been in the analysis of short sequence nucleotides, especially when applied to the analysis of single-nucleotide polymorphisms (SNPs). SNPs are variations in genetic sequence in specific genes which give rise to phenotypical variations in humans and can include susceptibility to diseases and response to specific treatments. They are therefore useful as indicators of disease risk and in the development of the most appropriate treatment regimes for patients. For SNP analysis, DNA is taken from an individual and the PCR used to amplify the DNA to a sufficient quantity. A primer is then designed in order to reproduce the gene but lacking one of the four required bases for nucleotide polymer extension. The length of the extended replication product of the gene is therefore dependent upon how soon the replication process requires the lacking nucleobase in order to further extend the replicating DNA. An example is shown in Fig. 2.6. With careful primer design and knowing the expected SNP site this can allow for differing length (and as a result differing mass) of the extended gene product dependent upon (a) whether the SNP is absent from both gene copies, (b) whether the SNP is present in only one gene copy and (c) whether it is present in both gene copies. Given that the expected resulting mass shift is usually a few hundred mass units, this can be clearly detected by the MALDI approach in sequences up to 20–30 nucleotides in length. Furthermore, different genes and different SNPs can be analysed simultaneously as long as the mass of the primer sequences and their extension products is sufficiently different to be determined separately by the MALDI profiling process. The approach has been applied successfully in a number of different disorders including cancer, lupus, arthritis, cirrhosis, insulin sensitivity and cholesterol metabolism disorders [10, 11, 14, 39, 63, 70, 97-99]. Usually after extension, the oligonucleotides are desalted using cation exchange beads (due to the fact that the PCR requires the addition of magnesium chloride which then interferes with the MALDI process) and mixing with a hydroxypicolinic acid matrix with ammonium citrate as an additive.

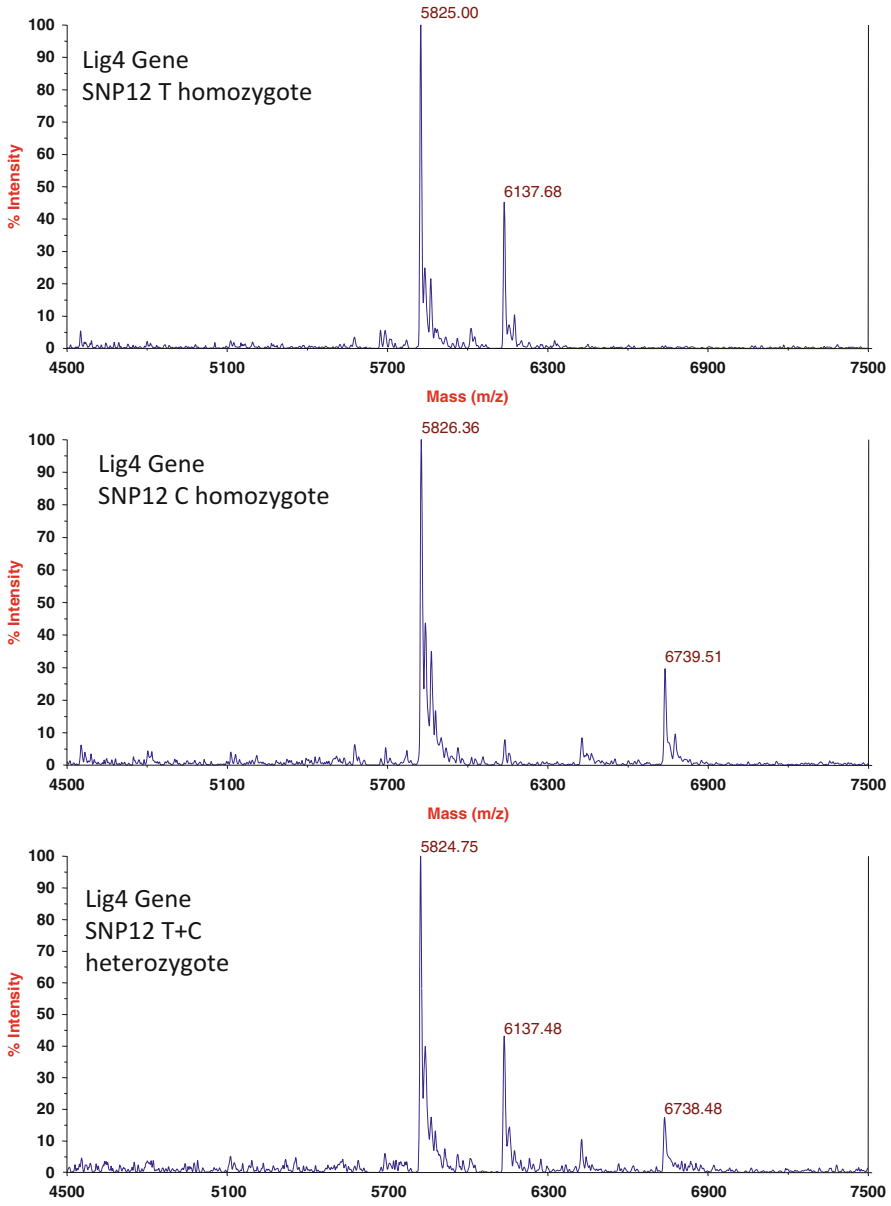


Fig. 2.6 Single-nucleotide polymorphism analysis utilising MALDI mass spectrometry

2.8 Metabolomics

Metabolomic analysis does not commonly utilise the MALDI approach due to matrix ions in the low mass range (less than 500 Da) which complicate the resulting

spectra. However, more recently, the approach has been utilised to study cell metabolism in cancer cells, metabolites from blood spots and amino acids present in the natural secretions of Lucilia sericata (the larvae of the greenbottle fly, utilised in maggot debridement therapy) [6, 44, 62]. Of the study of different metabolite profiles, the most common application of MALDI profiling is in the field of lipidomics in which phospholipids and associated species can be studied in a high-throughput manner using the mass accuracy provided by the ToF tube to allow for the grouping of phospholipids by head group and the revealing of the total length of both fatty acid chains combined and the total number of double bonds presented by both fatty acids combined. Usually, phosphatidylcholine lipids are analysed in positive ionisation mode due to the easily protonated head group associated with this class whilst the other lipid groups provide an improved signal in negative ionisation mode. The matrices utilised for the two ionisation modes differ, with dihydroxybenzoic acid in trifluoroacetic acid and methanol allowing for positive ionisation and *p*-nitroalanine in 2:1 chloroform:methanol being utilised for negative phospholipid analysis. Example spectra are shown in Fig. 2.7, alongside the effect of dosing the matrix with caesium chloride prior to analysis (causing a mass shift of 132 amu) which allows for the differentiation of different lipids and their sodium adducts in positive ionisation analyses. MALDI profiling in phospholipid analysis has found applications in the study of the species in eye lens, the lipid profiling of cancer cells, studies as to embryo optimisation and studying changes in the biochemistry of spermatozoa in obese patients [19, 44, 67]. In order to improve such analysis polystyrene spheres have been suggested as a matrix additive in order to improve matrix heterogeneity during crystallisation and precoating the MALDI plate with 1,5-diaminonaphthalene has been used in order to allow MALDI imaging of phospholipid distributions [92, 97]. Therefore, it can be seen that metabolite profiling, especially subclasses of metabolites such as phospholipid profiling, is rapidly improving based on advances already applied to protein and peptide profiling.

2.9 Conclusion

MALDI as a technique first found a niche role in allowing for the mass spectrometric analysis of large biomolecules (proteins predominantly) as a singly charged species. However it has been recognised as a potentially high-throughput method for the analysis of a variety of other biochemical entities. With improvements in matrix choices, heterogeneity of matrix crystallisation with analytes and post-acquisition data analysis tools, MALDI has found applications in studying both the presence and, in MALDI imaging, the distribution of many more biological moieties and applications in low molecular weight metabolites are also increasing in their scope and applications.

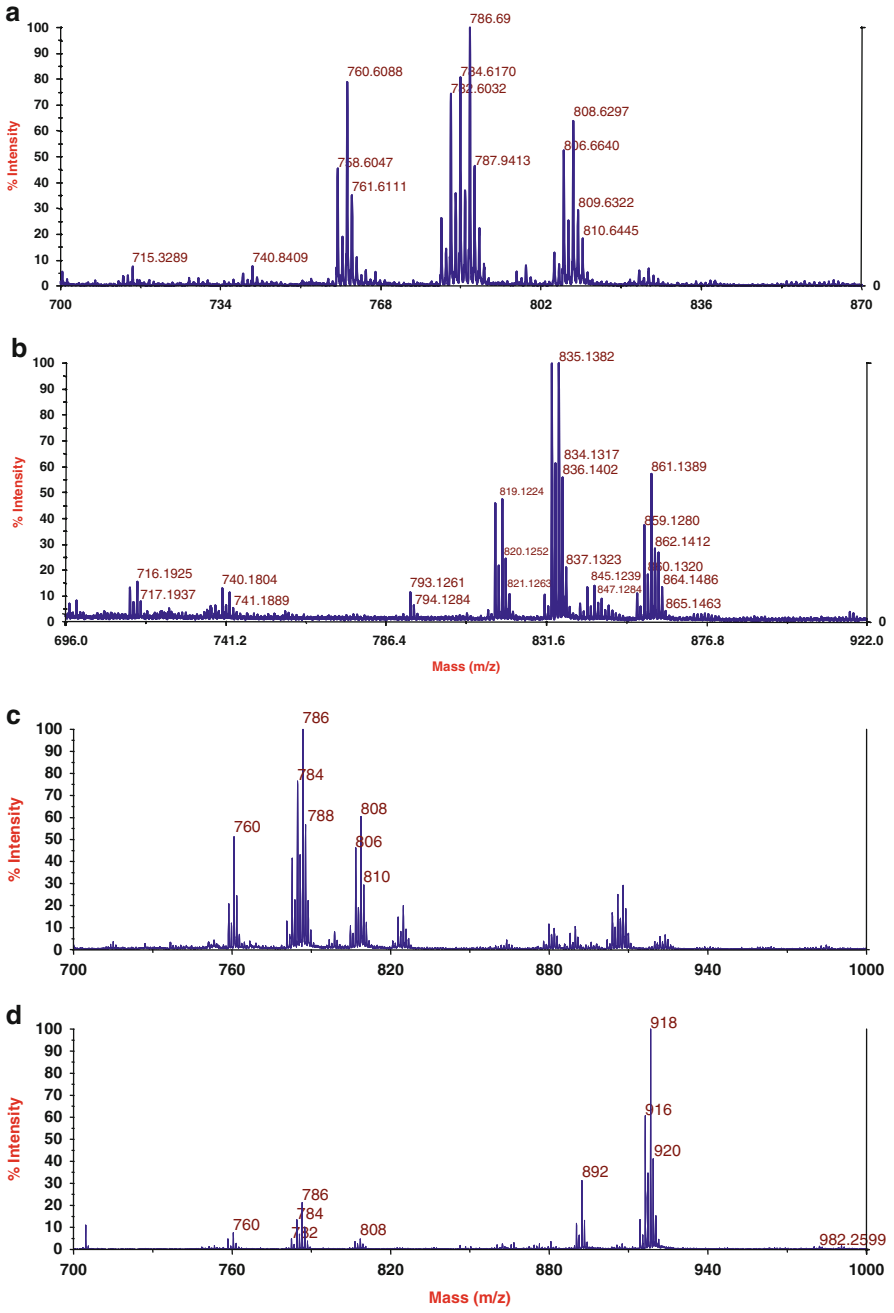


Fig. 2.7 Analysis of phospholipids by MALDI mass spectrometry. (**a**) Positive ionisation analysis. (**b**) Negative ionisation analysis. (**c**) Positive ionisation analysis without caesium chloride. (**d**) Positive ionisation analysis with caesium chloride

References

1. Alawam K, Dudley E, Donev R, Thome J (2012) Protein and peptide profiling as a tool for biomarker discovery in depression. Electrophoresis 33(24):3830–3834
2. Albertini V, Benussi L, Paterlini A, Glionna M, Prestia A, Bocchio-Chiavetto L, Amicucci G, Galluzzi S, Adorni A, Geroldi C, Binetti G, Frisoni GB, Ghidoni R (2012) Distinct cerebrospinal fluid amyloid-beta peptide signatures in cognitive decline associated with Alzheimer's disease and schizophrenia. Electrophoresis 33(24):3738–3744
3. An Y, Goldman R (2013) Analysis of peptides by denaturing ultrafiltration and LC-MALDI-TOF-MS. Methods Mol Biol 1023:13–19
4. Aresta A, Calvano CD, Palmisano F, Zambonin CG, Monaco A, Tommasi S, Pilato B, Paradiso A (2008) Impact of sample preparation in peptide/protein profiling in human serum by MALDI-TOF mass spectrometry. J Pharm Biomed Anal 46(1):157–164
5. Bellei E, Monari E, Bergamini S, Ozben T, Tomasi A (2011) Optimizing protein recovery yield from serum samples treated with beads technology. Electrophoresis 32(12): 1414–1421
6. Bexfield A, Bond AE, Morgan C, Wagstaff J, Newton RP, Ratcliffe NA, Dudley E, Nigam Y (2010) Amino acid derivatives from Lucilia sericata excretions/secretions may contribute to the beneficial effects of maggot therapy via increased angiogenesis. Br J Dermatol 162(3):554–562
7. Biskup K, Braicu EI, Sehouli J, Fotopoulou C, Tauber R, Berger M, Blanchard V (2013) Serum glycome profiling: a biomarker for diagnosis of ovarian cancer. J Proteome Res 12(9):4056–4063
8. Bruegel M, Planert M, Baumann S, Focke A, Bergh FT, Leichtle A, Ceglarek U, Thiery J, Fiedler GM (2009) Standardized peptidome profiling of human cerebrospinal fluid by magnetic bead separation and matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. J Proteomics 72(4):608–615
9. Callesen AK, Christensen RD, Madsen JS, Vach W, Zapico E, Cold S, Jørgensen PE, Mogensen O, Kruse TA, Jensen ON (2008) Reproducibility of serum protein profiling by systematic assessment using solid-phase extraction and matrix-assisted laser desorption/ionization mass spectrometry. Rapid Commun Mass Spectrom 22(3):291–300
10. Chen CD, Yu ZQ, Chen XL, Zhou JQ, Zhou XT, Sun XH, Chu RY (2013) Evaluating the association between pathological myopia and SNPs in RASGRF1. ACTC1 and GJD2 genes at chromosome 15q14 and 15q25 in a Chinese population. Ophthalmic Genet. [Epub ahead of print]
11. Chen WH, Hsu IH, Sun YC, Wang YK, Wu TK (2013) Immunocapture couples with matrix-assisted laser desorption/ionization time-of-flight mass spectrometry for rapid detection of type 1 dengue virus. J Chromatogr A 1288:21–27
12. Cho D, Hong S, Shim S (2013) Few layer graphene matrix for matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. J Nanosci Nanotechnol 13(8):5811–5813
13. Chu L, Fu G, Meng Q, Zhou H, Zhang M (2013) Identification of urinary biomarkers for type 2 diabetes using bead-based proteomic approach. Diabetes Res Clin Pract 101(2):187–193
14. Clifford AJ, Rincon G, Owens JE, Medrano JF, Moshfegh AJ, Baer DJ, Novotny JA (2013) Single nucleotide polymorphisms in CETP, SLC46A1, SLC19A1, CD36, BCMO1, APOA5, and ABCA1 are significant predictors of plasma HDL in healthy adults. Lipids Health Dis 12:66
15. D'Alessandro A, Mirasole C, Zolla L (2013) Haemoglobin glycation (Hb1Ac) increases during red blood cell storage: a MALDI-TOF mass-spectrometry-based investigation. Vox Sang 105(2):177–180
16. Demarco ML, Ford BA (2013) Beyond identification: emerging and future uses for MALDI-TOF mass spectrometry in the clinical microbiology laboratory. Clin Lab Med 33(3):611–628
17. Dybwad M, van der Laaken AL, Blatny JM, Paauw A (2013) Rapid identification of Bacillus anthracis spores in suspicious powder samples by using matrix-assisted laser desorption

ionization-time of flight mass spectrometry (MALDI-TOF MS). Appl Environ Microbiol 79(17):5372–5383
18. Emanuele VA II, Panicker G, Gurbaxani BM, Lin JM, Unger ER (2012) Sensitive and specific peak detection for SELDI-TOF mass spectrometry using a wavelet/neural-network based approach. PLoS One 7(11):e48103
19. Estrada R, Puppato A, Borchman D, Yappert MC (2010) Reevaluation of the phospholipid composition in membranes of adult human lenses by (31)P NMR and MALDI MS. Biochim Biophys Acta 1798(3):303–311
20. Eriksson A, Bergquist J, Edwards K, Hagfeldt A, Malmström D, Hernández VA (2011) Mesoporous TiO_2-based experimental layout for on-target enrichment and separation of multi- and monophosphorylated peptides prior to analysis with matrix-assisted laser desorption-ionization mass spectrometry. Anal Chem 83(3):761–766
21. Fania C, Vasso M, Torretta E, Robach P, Cairo G, Lundby C, Gelfi C (2011) Setup for human sera MALDI profiling: the case of rhEPO treatment. Electrophoresis 32(13):1715–1727
22. Fiedler GM, Baumann S, Leichtle A, Oltmann A, Kase J, Thiery J, Ceglarek U (2007) Standardized peptidome profiling of human urine by magnetic bead separation and matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. Clin Chem 53(3): 421–428
23. Ford BA, Burnham CA (2013) Optimization of routine identification of clinically relevant Gram-negative bacteria by use of matrix-assisted laser desorption ionization-time of flight mass spectrometry and the Bruker Biotyper. J Clin Microbiol 51(5):1412–1420
24. Fumagalli M, Ferrari F, Luisetti M, Stolk J, Hiemstra PS, Capuano D, Viglio S, Fregonese L, Cerveri I, Corana F, Tinelli C, Iadarola P (2012) Profiling the proteome of exhaled breath condensate in healthy smokers and COPD patients by LC-MS/MS. Int J Mol Sci 13(11): 13894–13910
25. Gao X, Bi X, Wei J, Peng Z, Liu H, Jiang Y, Wei W, Cai Z (2013) N-phosphorylation labeling for analysis of twenty natural amino acids and small peptides by using matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. Analyst 138(9):2632–2639
26. Gatlin CL, White KY, Tracy MB, Wilkins CE, Semmes OJ, Nyalwidhe JO, Drake RR, Malyarenko DI (2011) Enhancement in MALDI-TOF MS analysis of the low molecular weight human serum proteome. J Mass Spectrom 46(1):85–89
27. Goldstein JE, Zhang L, Borror CM, Rago JV, Sandrin TR (2013) Culture conditions and sample preparation methods affect spectrum quality and reproducibility during profiling of Staphylococcus aureus with matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. Lett Appl Microbiol 57(2):144–150
28. Gray TJ, Thomas L, Olma T, Iredell JR, Chen SC (2013) Rapid identification of Gram-negative organisms from blood culture bottles using a modified extraction method and MALDI-TOF mass spectrometry. Diagn Microbiol Infect Dis pii:S0732-8893(13)00348-9
29. Kirchner H, Gutierrez JA, Solenberg PJ, Pfluger PT, Czyzyk TA, Willency JA, Schürmann A, Joost HG, Jandacek RJ, Hale JE, Heiman ML, Tschöp MH (2009) GOAT links dietary lipids with the endocrine control of energy balance. Nat Med 15(7):741–745
30. Hamouda H, Ullah M, Berger M, Sittinger M, Tauber R, Ringe J, Blanchard V (2013) N-Glycosylation profile of undifferentiated and adipogenically differentiated human bone marrow mesenchymal stem cells—towards a next generation of stem cell markers. Stem Cells Dev 22(23):3100–3113
31. Hanisch FG (2011) Top-down sequencing of O-glycoproteins by in-source decay matrix-assisted laser desorption ionization mass spectrometry for glycosylation site analysis. Anal Chem 83(12):4829–4837
32. Hoppenheit A, Murugaiyan J, Bauer B, Steuber S, Clausen PH, Roesler U (2013) Identification of Tsetse (Glossina spp.) using matrix-assisted laser desorption/ionisation time of flight mass spectrometry. PLoS Negl Trop Dis 7(7):e2305
33. Hou J, Xie Z, Xue P, Cui Z, Chen X, Li J, Cai T, Wu P, Yang F (2010) Enhanced MALDI-TOF MS analysis of phosphopeptides using an optimized DHAP/DAHC matrix. J Biomed Biotechnol 2010:759690

34. Huang LT, Wen Q, Zhao MZ, Li ZB, Luo N, Wang YT, Dong XQ, Yu XQ (2012) Serum peptidome profiling for identifying pathological patterns in patients with primary nephrotic syndrome. Chin Med J (Engl) 125(24):4418–4423
35. Huang TL, Cho YT, Su H, Shiea J (2013) Principle component analysis combined with matrix-assisted laser desorption ionization mass spectrometry for rapid diagnosing the sera of patients with major depression. Clin Chim Acta 424C:175–181
36. Jamal W, Saleem R, Rotimi VO (2013) Rapid identification of pathogens directly from blood culture bottles by Bruker matrix-assisted laser desorption laser ionization-time of flight mass spectrometry versus routine methods. Diagn Microbiol Infect Dis 76(4):404–408
37. Jiang W, Murashko EA, Dubrovskii YA, Podolskaya EP, Babakov VN, Mikler J, Nachon F, Masson P, Schopfer LM, Lockridge O (2013) Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry of titanium oxide-enriched peptides for detection of aged organophosphorus adducts on human butyrylcholinesterase. Anal Biochem 439(2):132–141
38. Josten M, Reif M, Szekat C, Al-Sabti N, Roemer T, Sparbier K, Kostrzewa M, Rohde H, Sahl HG, Bierbaum G (2013) Analysis of the matrix-assisted laser desorption ionization-time of flight mass spectrum of Staphylococcus aureus identifies mutations that allow differentiation of the main clonal lineages. J Clin Microbiol 51(6):1809–1817
39. Justenhoven C, Pentimalli D, Rabstein S, Harth V, Lotz A, Pesch B, Brüning T, Dörk T, Schürmann P, Bogdanova N, Park-Simon TW, Couch FJ, Olson JE, Fasching PA, Beckmann MW, Häberle L, Ekici A, Hall P, Czene K, Liu J, Li J, Baisch C, Hamann U, Ko YD, Brauch H (2013) CYP2B6*6 is associated with increased breast cancer risk. Int J Cancer 134(2): 426–430
40. Kandzia S, Costa J (2013) N-Glycosylation analysis by HPAEC-PAD and mass spectrometry. Methods Mol Biol 1049:301–312
41. Kang JH, Mori T, Kitazaki H, Niidome T, Takayama K, Nakanishi Y, Katayama Y (2013) Serum protein kinase Cα as a diagnostic biomarker of cancers. Cancer Biomark 13(2):99–103
42. Karger A, Melzer F, Timke M, Bettin B, Kostrzewa M, Nöckler K, Hohmann A, Tomaso H, Neubauer H, Al Dahouk S (2013) Inter-laboratory comparison of intact cell matrix-assisted laser desorption/ionization time of flight mass spectrometry (MALDI-TOF MS) in the identification and differentiation of Brucella spp. J Clin Microbiol 51(9):3123–3126
43. Kay R, Barton C, Ratcliffe L, Matharoo-Ball B, Brown P, Roberts J, Teale P, Creaser C (2008) Enrichment of low molecular weight serum proteins using acetonitrile precipitation for mass spectrometry based proteomic analysis. Rapid Commun Mass Spectrom 22(20): 3255–3260
44. Kim IC, Lee JH, Bang G, Choi SH, Kim YH, Kim KP, Kim HK, Ro J (2013) Lipid profiles for HER2-positive breast cancer. Anticancer Res 33(6):2467–2472
45. Kjellström S, Jensen ON (2004) Phosphoric acid as a matrix additive for MALDI MS analysis of phosphopeptides and phosphoproteins. Anal Chem 76(17):5109–5117
46. Lacroix C, Gicquel A, Sendid B, Meyer J, Accoceberry I, François N, Morio F, Desoubeaux G, Chandenier J, Kauffmann-Lacroix C, Hennequin C, Guitard J, Nassif X, Bougnoux ME (2013) Evaluation of two matrix-assisted laser desorption ionization-time of flight mass spectrometry (MALDI-TOF MS) systems for the identification of Candida species. Clin Microbiol Infect 20(2):153–158
47. Lee BS, Jayathilaka GD, Huang JS, Vida LN, Honig GR, Gupta S (2011) Analyses of in vitro nonenzymatic glycation of normal and variant hemoglobins by MALDI-TOF mass spectrometry. J Biomol Tech 22(3):90–94
48. Lee TF, Lee H, Chen CM, Du SH, Cheng YC, Hsu CC, Chung MY, Teng SH, Teng LJ, Hsueh PR (2013) Comparison of the accuracy of matrix-assisted laser desorption ionization-time of flight mass spectrometry with that of other commercial identification systems for identifying Staphylococcus saprophyticus in urine. J Clin Microbiol 51(5):1563–1566
49. Leli C, Cenci E, Cardaccia A, Moretti A, D'Alò F, Pagliochini R, Barcaccia M, Farinelli S, Vento S, Bistoni F, Mencacci A (2013) Rapid identification of bacterial and fungal pathogens from positive blood cultures by MALDI-TOF MS. Int J Med Microbiol 303(4):205–209

50. Li J, Hu XK, Lipson RH (2013) On-chip enrichment and analysis of peptide subsets using a maleimide-functionalized fluorous affinity biochip and nanostructure initiator mass spectrometry. Anal Chem 85(11):5499–5505
51. Lo LH, Huang TL, Shiea J (2009) Acid hydrolysis followed by matrix-assisted laser desorption/ionization mass spectrometry for the rapid diagnosis of serum protein biomarkers in patients with major depression. Rapid Commun Mass Spectrom 23(5):589–598
52. Mahmoud K, Cole LM, Newton J, Mohamed S, Al-Enazi M, Quirke P, Clench MR (2013) Detection of the epidermal growth factor receptor, amphiregulin and epiregulin in formalin-fixed paraffin-embedded human placenta tissue by matrix-assisted laser desorption/ionization mass spectrometry imaging. Eur J Mass Spectrom (Chichester, Eng) 19(1):17–28
53. Mancini N, De Carolis E, Infurnari L, Vella A, Clementi N, Vaccaro L, Ruggeri A, Posteraro B, Burioni R, Clementi M, Sanguinetti M (2013) Comparative evaluation of the Bruker Biotyper and Vitek MS matrix-assisted laser desorption ionization-time of flight (MALDI-TOF) mass spectrometry systems for identification of yeasts of medical importance. J Clin Microbiol 51(7):2453–2457
54. Martin-Eauclaire MF, Granjeaud S, Belghazi M, Bougis PE (2013) Achieving automated scorpion venom mass fingerprinting (VMF) in the nanogram range. Toxicon 69:211–218
55. Martiny D, Debaugnies F, Gateff D, Gérard M, Aoun M, Martin C, Konopnicki D, Loizidou A, Georgala A, Hainaut M, Chantrenne M, Dediste A, Vandenberg O, Van Praet S (2013) Impact of rapid microbial identification directly from positive blood cultures using matrix-assisted laser desorption/ionization time-of-flight mass spectrometry on patient management. Clin Microbiol Infect 19(12):E568–E581
56. Merrell K, Southwick K, Graves SW, Esplin MS, Lewis NE, Thulin CD (2004) Analysis of low-abundance, low-molecular-weight serum proteins using mass spectrometry. J Biomol Tech 15(4):238–248
57. Morelle W, Faid V, Chirat F, Michalski JC (2009) Analysis of N- and O-linked glycans from glycoproteins using MALDI-TOF mass spectrometry. Methods Mol Biol 534:5–21
58. Müller P, Pflüger V, Wittwer M, Ziegler D, Chandre F, Simard F, Lengeler C (2013) Identification of cryptic Anopheles mosquito species by molecular protein profiling. PLoS One 8(2):e57486
59. Nagy E, Urbán E, Becker S, Kostrzewa M, Vörös A, Hunyadkürti J, Nagy I (2013) MALDI-TOF MS fingerprinting facilitates rapid discrimination of phylotypes I, II and III of Propionibacterium acnes. Anaerobe 20:20–26
60. Nicklay JJ, Harris GA, Schey KL, Caprioli RM (2013) MALDI imaging and in situ identification of integral membrane proteins from rat brain tissue sections. Anal Chem 85(15): 7191–7196
61. Nicolardi S, van der Burgt YE, Dragan I, Hensbergen PJ, Deelder AM (2013) Identification of new apolipoprotein-CIII glycoforms with ultrahigh resolution MALDI-FTICR mass spectrometry of human sera. J Proteome Res 12(5):2260–2268
62. Ostermann KM, Dieplinger R, Lutsch NM, Strupat K, Metz TF, Mechtler TP, Kasper DC (2013) Matrix-assisted laser desorption/ionization for simultaneous quantitation of (acyl-) carnitines and organic acids in dried blood spots. Rapid Commun Mass Spectrom 27(13): 1497–1504
63. Peng L, Guo J, Zhang Z, Liu L, Cao Y, Shi H, Wang J, Wang J, Friedman SL, Sninsky JJ (2013) A candidate gene study for the association of host single nucleotide polymorphisms with liver cirrhosis risk in Chinese hepatitis B patients. Genet Test Mol Biomarkers 17(9): 681–686
64. Pérez V, Sánchez-Escuredo A, Lauzurica R, Bayés B, Navarro-Muñoz M, Pastor MC, Cañas L, Bonet J, Romero R (2013) Magnetic bead-based proteomic technology to study paricalcitol effect in kidney transplant recipients. Eur J Pharmacol 709(1–3):72–79
65. Pimenta AM, Stöcklin R, Favreau P, Bougis PE, Martin-Eauclaire MF (2001) Moving pieces in a proteomic puzzle: mass fingerprinting of toxic fractions from the venom of Tityus serrulatus (Scorpiones, Buthidae). Rapid Commun Mass Spectrom 15(17):1562–1572

66. Rychert J, Burnham CA, Bythrow M, Garner OB, Ginocchio CC, Jennemann R, Lewinski MA, Manji R, Mochon AB, Procop GW, Richter SS, Sercia L, Westblade LF, Ferraro MJ, Branda JA (2013) Multicenter evaluation of the Vitek MS matrix-assisted laser desorption ionization-time of flight mass spectrometry system for identification of Gram-positive aerobic bacteria. J Clin Microbiol 51(7):2225–2231
67. Pyttel S, Zschörnig K, Nimptsch A, Paasch U, Schiller J (2012) Enhanced lysophosphatidylcholine and sphingomyelin contents are characteristic of spermatozoa from obese men-A MALDI mass spectrometric study. Chem Phys Lipids 165(8):861–865
68. Selman MH, Hoffmann M, Zauner G, McDonnell LA, Balog CI, Rapp E, Deelder AM, Wuhrer M (2012) MALDI-TOF-MS analysis of sialylated glycans and glycopeptides using 4-chloro-α-cyanocinnamic acid matrix. Proteomics 12(9):1337–1348
69. Schaumann R, Knoop N, Genzel GH, Losensky K, Rosenkranz C, Stîngu CS, Schellenberger W, Rodloff AC, Eschrich K (2013) Discrimination of Enterobacteriaceae and non-fermenting gram negative bacilli by MALDI-TOF mass spectrometry. Open Microbiol J 7:118–122
70. Shen L, Liu R, Zhang H, Huang Y, Sun R, Tang P (2013) Replication study of STAT4 rs7574865 G/T polymorphism and risk of rheumatoid arthritis in a Chinese population. Gene 526(2):259–264
71. Schneiderhan W, Grundt A, Wörner S, Findeisen P, Neumaier M (2013) Workflow analysis of around-the-clock processing of blood culture samples and integrated MALDI-TOF mass spectrometry analysis for the diagnosis of bloodstream infections. Clin Chem 59(11): 1649–1656
72. Schulthess B, Brodner K, Bloemberg GV, Zbinden R, Böttger EC, Hombach M (2013) Identification of Gram-positive cocci by use of matrix-assisted laser desorption ionization-time of flight mass spectrometry: comparison of different preparation methods and implementation of a practical algorithm for routine diagnostics. J Clin Microbiol 51(6): 1834–1840
73. Seng P, Abat C, Rolain JM, Colson P, Lagier JC, Gouriet F, Fournier PE, Drancourt M, La Scola B, Raoult D (2013) Identification of rare pathogenic bacteria in a clinical microbiology laboratory: impact of matrix-assisted laser desorption ionization-time of flight mass spectrometry. J Clin Microbiol 51(7):2182–2194
74. Steven RT, Bunch J (2013) Repeat MALDI MS imaging of a single tissue section using multiple matrices and tissue washes. Anal Bioanal Chem 405(14):4719–4728
75. Suarez S, Ferroni A, Lotz A, Jolley KA, Guérin P, Leto J, Dauphin B, Jamet A, Maiden MC, Nassif X, Armengaud J (2013) Ribosomal proteins as biomarkers for bacterial identification by mass spectrometry in the clinical microbiology laboratory. J Microbiol Methods pii:S0167-7012(13)00238-8
76. Qian JY, Mou SH, Liu CB (2012) SELDI-TOF MS combined with magnetic beads for detecting serum protein biomarkers and establishment of a boosting decision tree model for diagnosis of pancreatic cancer. Asian Pac J Cancer Prev 13(5):1911–1915
77. Qiao L, Roussel C, Wan J, Yang P, Girault HH, Liu B (2007) Specific on-plate enrichment of phosphorylated peptides for direct MALDI-TOF MS analysis. J Proteome Res 6(12): 4763–4769
78. Qiao X, Zhou Y, Hou C, Zhang X, Yang K, Zhang L, Zhang Y (2013) 1-(3-aminopropyl)-3-butylimidazolium bromide for carboxyl group derivatization: potential applications in high sensitivity peptide identification by mass spectrometry. Sci China Life Sci 56(3):240–245
79. Qin H, Zhao L, Li R, Wu R, Zou H (2011) Size-selective enrichment of N-linked glycans using highly ordered mesoporous carbon material and detection by MALDI-TOF MS. Anal Chem 83(20):7721–7728
80. Quaas A, Bahar AS, von Loga K, Seddiqi AS, Singer JM, Omidi M, Kraus O, Kwiatkowski M, Trusch M, Minner S, Burandt E, Stahl P, Wilczak W, Wurlitzer M, Simon R, Sauter G, Marx A, Schlüter H (2013) MALDI imaging on large-scale tissue microarrays identifies molecular features associated with tumour phenotype in oesophageal cancer. Histopathology 63(4):455–462. doi:10.1111/his.12193

81. Shuo T, Koshikawa N, Hoshino D, Minegishi T, Ao-Kondo H, Oyama M, Sekiya S, Iwamoto S, Tanaka K, Seiki M (2012) Detection of the heterogeneous O-glycosylation profile of MT1-MMP expressed in cancer cells by a simple MALDI-MS method. PLoS One 7(8):e43751
82. Taurines R, Dudley E, Conner AC, Grassl J, Jans T, Guderian F, Mehler-Wex C, Warnke A, Gerlach M, Thome J (2010) Serum protein profiling and proteomics in autistic spectrum disorder using magnetic bead-assisted mass spectrometry. Eur Arch Psychiatry Clin Neurosci 260(3):249–255
83. Tep S, Hincapie M, Hancock WS (2012) A MALDI-TOF MS method for the simultaneous and quantitative analysis of neutral and sialylated glycans of CHO-expressed glycoproteins. Carbohydr Res 347(1):121–129
84. Tiss A, Smith C, Camuzeaux S, Kabir M, Gayther S, Menon U, Waterfield M, Timms J, Jacobs I, Cramer R (2007) Serum peptide profiling using MALDI mass spectrometry: avoiding the pitfalls of coated magnetic beads using well-established ZipTip technology. Proteomics 7(suppl 1):77–89
85. Tucholska M, Scozzaro S, Williams D, Ackloo S, Lock C, Siu KW, Evans KR, Marshall JG (2007) Endogenous peptides from biophysical and biochemical fractionation of serum analyzed by matrix-assisted laser desorption/ionization and electrospray ionization hybrid quadrupole time-of-flight. Anal Biochem 370(2):228–245
86. van Swelm RP, Laarakkers CM, Kooijmans-Otero M, de Jong EM, Masereeuw R, Russel FG (2013) Biomarkers for methotrexate-induced liver injury: urinary protein profiling of psoriasis patients. Toxicol Lett 221(3):219–224. doi:10.1016/j.toxlet.2013.06.234
87. Vella A, De Carolis E, Vaccaro L, Posteraro P, Perlin DS, Kostrzewa M, Posteraro B, Sanguinetti M (2013) Rapid antifungal susceptibility testing by matrix-assisted laser desorption ionization time-of-flight mass spectrometry analysis. J Clin Microbiol 51(9): 2964–2969
88. Wan QL, Hou XS, Zhao G (2013) Utility of serum peptidome patterns of esophageal squamous cell carcinoma patients for comprehensive treatment. Asian Pac J Cancer Prev 14(5): 2919–2923
89. Wang H, Duan J, Cheng Q (2011) Photocatalytically patterned TiO_2 arrays for on-plate selective enrichment of phosphopeptides and direct MALDI MS analysis. Anal Chem 83(5): 1624–1631
90. Wang X, Han J, Chou A, Yang J, Pan J, Borchers CH (2013) Hydroxyflavones as a new family of matrices for MALDI tissue imaging. Anal Chem 85(15):7566–7573
91. Watanabe M, Terasawa K, Kaneshiro K, Uchimura H, Yamamoto R, Fukuyama Y, Shimizu K, Sato TA, Tanaka K (2013) Improvement of mass spectrometry analysis of glycoproteins by MALDI-MS using 3-aminoquinoline/α-cyano-4-hydroxycinnamic acid. Anal Bioanal Chem 405(12):4289–4293
92. Wei Y, Li S, Wang J, Shu C, Liu J, Xiong S, Song J, Zhang J, Zhao Z (2013) Polystyrene spheres-assisted matrix-assisted laser desorption ionization mass spectrometry for quantitative analysis of plasma lysophosphatidylcholines. Anal Chem 85(9):4729–4734
93. Wojewoda C, Education Committee of the Academy of Clinical Laboratory Physicians and Scientists (2013) Pathology consultation on matrix-assisted laser desorption ionization-time of flight mass spectrometry for microbiology. Am J Clin Pathol 140(2):143–148
94. Xia B, Zhang W, Li X, Jiang R, Harper T, Liu R, Cummings RD, He M (2013) Serum N-glycan and O-glycan analysis by mass spectrometry for diagnosis of congenital disorders of glycosylation. Anal Biochem pii:S0003-2697(13)00355-2
95. Xin L, Zhang H, Liu H, Li Z (2012) Equal ratio of graphite carbon to activated charcoal for enrichment of N-glycopeptides prior to matrix-assisted laser desorption/ionization time-of-flight mass spectrometric identification. Rapid Commun Mass Spectrom 26(3):269–274
96. Yang X, Wu H, Kobayashi T, Solaro RJ, van Breemen RB (2004) Enhanced ionization of phosphorylated peptides during MALDI TOF mass spectrometry. Anal Chem 76(5): 1532–1536
97. Yang J, Caprioli RM (2013) Matrix precoated targets for direct lipid analysis and imaging of tissue. Anal Chem 85(5):2907–2912

98. Yuan Y, Tong N (2013) Associations of the PTEN -9C>G polymorphism with insulin sensitivity and central obesity in Chinese. Gene pii:S0378-1119(13)00783-X
99. Zhang DM, Cheng LQ, Zhai ZF, Feng L, Zhong BY, You Y, Zhang N, Song ZQ, Yang XC, Chen FR, Hao F (2013) Single-nucleotide polymorphism and haplotypes of TNIP1 associated with systemic lupus erythematosus in a Chinese Han population. J Rheumatol 40(9): 1535–1544
100. Zhou LH, Kang GY, Kim KP (2009) A binary matrix for improved detection of phosphopeptides in matrix-assisted laser desorption/ionization mass spectrometry. Rapid Commun Mass Spectrom 23(15):2264–2272

Chapter 3
Simplifying the Proteome: Analytical Strategies for Improving Peak Capacity

Lee A. Gethings and Joanne B. Connolly

Abstract The diversity of biological samples and dynamic range of analytes being analyzed can prove to be an analytical challenge and is particularly prevalent to proteomic studies. Maximizing the peak capacity of the workflow employed can extend the dynamic range and increase identification rates. The focus of this chapter is to present means of achieving this for various analytical techniques such as liquid chromatography, mass spectrometry, and ion mobility. A combination of these methods can be used as part of a data-independent acquisition strategy, thereby limiting issues such as chimericy when analyzing regions of extreme analyte density.

3.1 Introduction

The term "proteomics" encompasses the large-scale qualitative and quantitative study of proteins in a complex biological sample. The potential to understand the nature of a biological system using this information, in conjunction with other multi-omic data, is extensive. Current mass spectrometric techniques allow large-scale, high-throughput analyses for the identification, characterization, and quantification of the proteome, but can encounter major limitations in effectively mining complex biological matrices. A typical bottom-up proteomic experiment, for example, can result in areas of extreme analyte density during liquid chromatography (LC) separation (Fig. 3.1). These highly concentrated analyte regions are not only problematic from a chromatography aspect but can also be an issue with respect to the mass domain separation and detection.

Overcoming such challenges and increasing separation capabilities of analytical systems involve finding means of increasing peak capacity and dynamic range of

L.A. Gethings (✉) • J.B. Connolly
Waters Corporation, Stamford Avenue, Altrincham Road,
Wilmslow, SK9 4AX, Manchester, UK
e-mail: lee_gethings@waters.com

A.G. Woods and C.C. Darie (eds.), *Advancements of Mass Spectrometry in Biomedical Research*, Advances in Experimental Medicine and Biology 806,
DOI 10.1007/978-3-319-06068-2_3, © Springer International Publishing Switzerland 2014

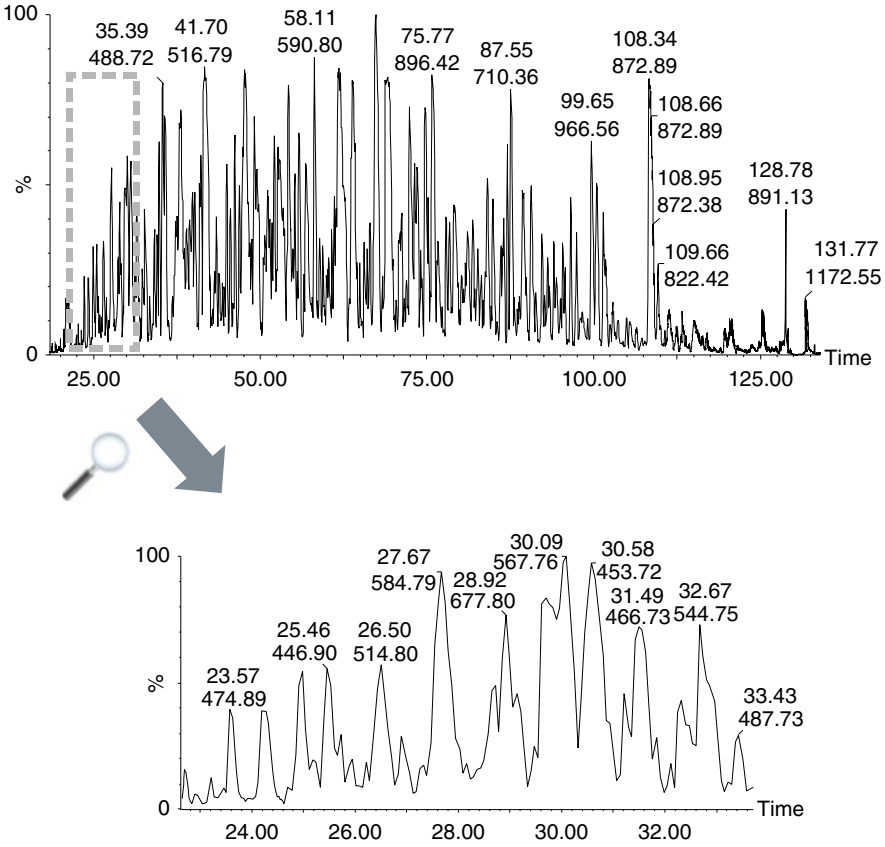


Fig. 3.1 Sample complexity and regions of analyte density represented by a 1,000 ng on-column loading of HeLa tryptic digest, separated over 150 min gradient. The lower figure shows the chromatogram expanded (23–33 min) with peak width half-height between 0.1 and 0.22 min

the workflow. Peak capacity is defined as the maximum number of components that can be separated to a specific resolution within a given separation space [1]. Throughout the course of this chapter, methodologies to increase system peak capacity will be introduced, including multidimensional chromatography and ion mobility (IM). Combined with a data-independent acquisition (DIA) strategy, these methods can provide enhanced sensitivity, specificity, and speed to biomarker discovery applications.

3.2 Expanding Peak Capacity from the LC Perspective

High performance liquid chromatography (HPLC) has and continues to be routinely used in combination with mass spectrometry (MS) for the separation of analytes. However, the performance of this technology has been augmented in recent years

with the introduction of ultra performance liquid chromatography (UPLC), making it possible to separate components from highly complex matrices with enhanced chromatographic resolution, sensitivity, and increased throughput. Innovative instrument design and advancements in column chemistries have contributed to these enhancements, particularly with the adoption of smaller column particle size for improved efficiency and pumping systems that can operate at elevated pressure, whilst maintaining the same selectivity and retention characteristics as HPLC equivalents [2]. Derivation of a mathematical expression which utilizes the van Deemter equation [3] illustrates how peak capacity is influenced by particle size, column length, diffusivity of the sample, linear velocity, and the length of gradient (Eq. 3.1). The terms used are defined as column length (L), equivalent theoretical plate height (H), analyte diffusion coefficient (D_m), packing material particle size (d_p), and empirical coefficients originating from the van Deemter equation (a, b, and c). The relationship between the logarithm of the capacity factor (k') and solvent composition is expressed as B, with Δc representing the difference in solvent composition over the course of the gradient. Finally, t_g/t_0 is defined as the gradient span.

$$n_{\mathrm{LC,gradient}} = 1 + \frac{1}{4}.\sqrt{\frac{L}{a.d_p + \frac{b.D_M}{u} + c.\frac{d_p^2}{D_M}.u}}.\frac{B.\Delta c}{B.\Delta c.\frac{t_0}{t_g} + 1} \tag{3.1}$$

If we consider the effect of reducing column particle size (d_p), transitioning from a 5 μm HPLC column packing to 1.7 μm stationary phase particles, efficiency and throughput can be improved up to a factor of three, but can generate backpressure increases by a factor of 27 [4]. Utilization of these high pressures results in enhanced separation efficiencies as represented by Fig. 3.2. Comparative chromatograms of glycogen phosphorylase B separated over the same length of gradient using HPLC and UPLC highlight increased column efficiency, resolution, and high throughput with UPLC. Sensitivity gains with UPLC show greater separation capability therefore limiting ion suppression and allow low intensity species to be readily identified. Factors other than d_p can influence the optimization of peak capacity from the LC platform. Three additional parameters, length of gradient, flow rate, and column length, also have roles to serve. An increased gradient length (t_g/t_0) provides higher peak capacities before eventually reaching a plateau where maximum peak capacity is attained. Within the scope of proteomics, gradients tend to routinely operate between 90 and 120 min depending on the complexity of sample being analyzed. Since longer gradients are employed with flow rates of less than 1 μL/min, the use of longer column lengths (L) to further increase peak capacity is viable. This is particularly important for the diverse range of biological samples discussed in the context of this review.

Multidimensional chromatography is a technique introduced as far back as 1953 [5], which described the use of two-dimensional (2D) paper chromatography and electrophoresis for peptide separation. It was soon adapted for 2D gel electrophoresis [6], but a number of drawbacks, such as the potential to modify proteins

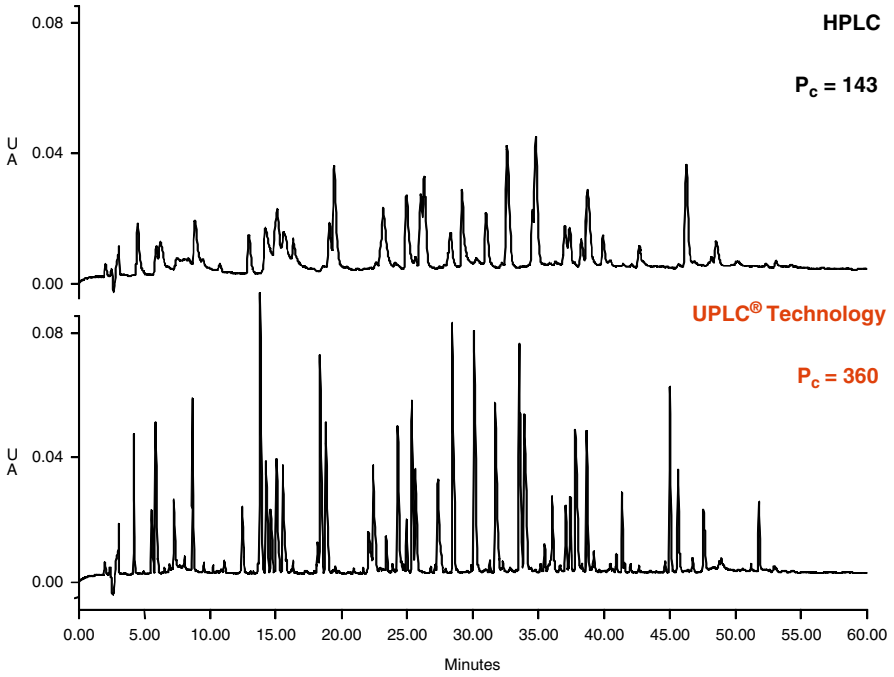


Fig. 3.2 Comparison of (**a**) HPLC (5 μm particles) with (**b**) UPLC (sub-2 μm particles) for the separation of a phosphorylase B tryptic digest

(i.e., free, unpolymerized acrylamide binding to the amino-terminal thereby preventing Edman degradation), technical reproducibility, and time constraints associated with performing gel analysis [7], make LC an attractive alternative. At the peptide level, as typically applied in bottom-up proteomic studies, 2D-LC provides high resolving power and ultimately increased peak capacity. Since the first and second dimensions have their individual peak capacities, $n_{c,1}$ and $n_{c,2}$, respectively, it is the product and not the sum of both [8, 9] that contributes to the overall peak capacity. This so-called product rule ($n_{c,2D}$) provides resolving powers that are not a true representation and often overestimate due to under-sampling effects of the first dimension. Accounting for the second-dimension separation cycle ($t_{c,2}$), the first-dimension peak capacity ($n_{c,1}$), and the first-dimension gradient time ($t_{g,1}$) [10], the influence of under-sampling on effective 2D-LC peak capacity shows no dependency on the first-dimension peak capacity (Eq. 3.2) with n_{2DLC} becoming less dependent on n_1.

$$n_{2DLC} = \frac{n_1 . n_2}{\sqrt{1 + 3.35 . \frac{t_{c,2} . n_{c,1}}{t_{g,1}}}} \tag{3.2}$$

In order to address the issue of under-sampling, an approximate model can be generated (Eq. 3.3), which assumes that the retention characteristics of the first and second dimensions are not correlated with little or no qualitative differences for non-orthogonal dimensions.

$$n_{2\mathrm{DLC}} \cong \frac{t_{g,1}.n_{c,2}}{1.83.t_{c,2}} \tag{3.3}$$

2D-LC experiments can be performed either off-line or on-line, depending on the configuration available. Off-line methods are often simpler, requiring fractions resulting from the first dimension to be collected prior to the second-dimension phase of separation, providing greater flexibility since the user is less restricted to the choice of eluent used in the first dimension. However, the risk of encountering peptide losses is increased. Implementing an on-line approach can work effectively, provided the selectivity between the two chemistries and compatibility of eluents are feasible. In most cases, the choice of the first-dimension conditions affords flexibility which can encompass a wide range of chemistries [11], such as strong cation exchange (SCX) [12–15], size exclusion chromatography (SEC) [16–19], immobilized metal ion affinity chromatography (IMAC) [20–24], or hydrophilic interaction chromatography (HILIC) [25–27]. The second dimension, however, tends to remain the same as that implemented for a 1D scenario (i.e., reversed phase). Of the examples provided, on-line SCX-RP is a common approach for peptide separation. The technique is partially orthogonal; however, distribution over multiple fractions is typically observed. An equally orthogonal technique, providing highly conserved peptide fractions and reducing the need for high salt concentration as is the case with SCX, is the use of a reversed phase–reversed phase (RP–RP) configuration [28]. Achieving RP–RP orthogonal separation is by differentiating on the basis of peptide interactions with the RP surface at differing charge states [29–32]. Considering acidic peptides, which are negatively charged at basic pH, this will result in early eluting fractions consisting of acidic peptides. Conversely for the second dimension at acidic pH, the acidic peptides eluting from the first dimension will now become "neutral" and hence switch their order of elution to becoming later eluting peptides. In the case of basic peptides the elution order is simply the reverse analogy described for acidic peptides. This scenario is clearly demonstrated in Fig. 3.3 which is a chromatogram of bovine haemoglobin tryptic digest. The chromatogram clearly shows orthogonal behavior of acidic and basic peptides across two dimensions operating at different pH. An RP–RP setup can be designed to comprise as many fractions as required. However, greater fractionation will compromise the analysis time and hence overall throughput, which also holds for off-line approaches. The higher peak capacity provided by RP–RP results in an increase of peptide and protein identifications [28, 33, 34]. To illustrate this, an example is shown in Fig. 3.4a, which represents comparative data from 1D, 2D 5-fraction, and 2D 10-fraction experiments for the bottom-up (2D) LC–DIA-MS analysis of cytosolic *Escherichia coli*. Sample loadings consisted of 750 ng (1D), 2.4 µg (2D 5-fraction), and 4.8 µg (2D 10-fraction). Confident peptides and proteins are shown, with significant gains

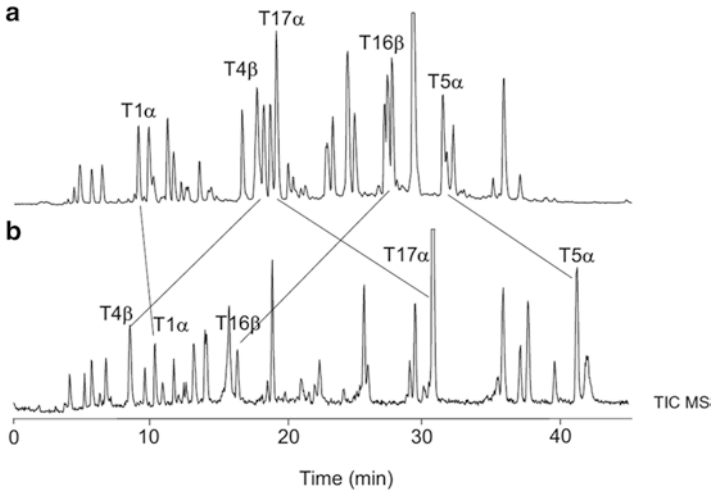


Fig. 3.3 The orthogonal nature of RP–RP demonstrated using a tryptic digest of bovine haemoglobin. The acidic (T16β, T4β), "neutral" (T1α), and basic (T17α, T5α) peptides clearly show a shift in retention order between pH 2.6 (**a**) and pH 10 (**b**) (Reprinted with permission from Gilar et. al., J. Sep. Sci., (2005), **28**, 1694–1703)

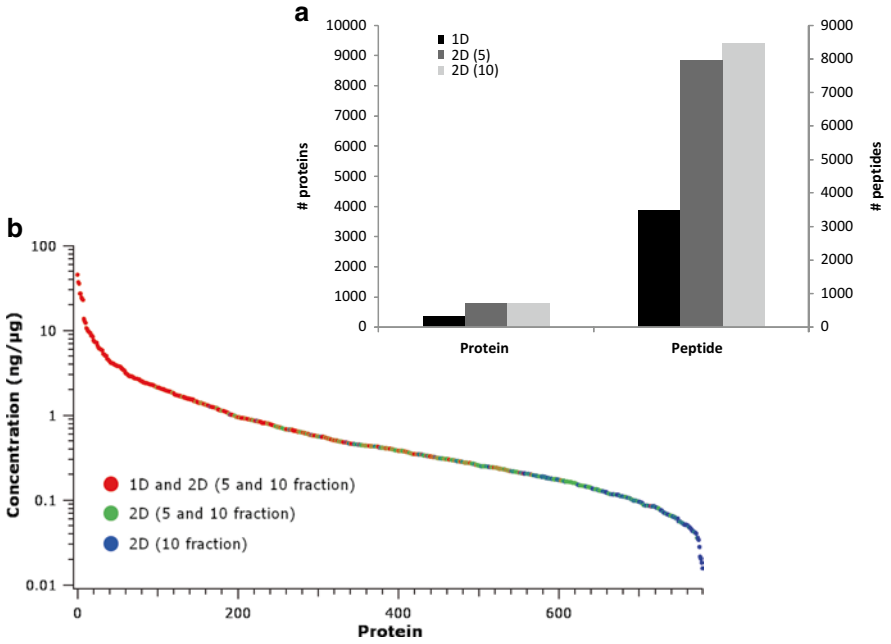


Fig. 3.4 Comparing 1D RP and 2D RP–RP (5- and 10-fraction) analysis of *Escherichia coli* shows (**a**) increased peptide and protein identifications with greater fractionation; (**b**) extended dynamic range gained with 2D RP–RP increasing the capability of identifying the less abundant peptides

achieved when 2D RP–RP is implemented. Transitioning from 1D to 2D (5-fraction) resulted in a 47 % increase in terms of protein identifications and an additional 11 % gain when fractionation is increased from 5 to 10 steps. This percentage gain can be accounted for with increased sample loadings and the ability to sample over a larger dynamic range (Fig. 3.4b).

3.3 Is Mass Resolution Alone Sufficient for Dealing with Sample Complexity?

Over recent years there have been significant gains in the mass resolution that can be achieved from mass spectrometers of various geometries. Mass resolution is an important parameter, providing a means of differentiating ions resulting from a complex sample. Resolution (R) is defined as a measure or capacity to distinguish ions of adjacent mass number, m and Δm, respectively (i.e., $R = m/\Delta m$). It is important however to differentiate between working resolution of an instrument and the resolution which can be obtained in practice, since acquisition speed and m/z are determining factors [35].

A range of high resolution mass analyzers are routinely used for biomarker discovery experiments including traps and quadrupole time-of-flight (Q-ToF) geometries. However, it is the geometry of a Q-ToF mass spectrometer which will be the discussed platform for the remainder of this chapter [36]. The geometry of a Q-ToF is a tandem version of the orthogonal accelerated ToF (oa-ToF), consisting of an MS1 quadrupole and collision region with reflecting oa-ToF MS as MS2. Data can be acquired in MS and MS/MS modes, providing high mass accuracy and sensitivity. These high levels of sensitivity are achieved compared with scanning instruments, since orthogonal geometry is applied for the detection of ions as opposed to sequential detection. A significant innovation to the original design saw the hybridization of ion mobility (IM) (Fig. 3.5) [37]. The principle of IM involves the separation of a packet of ions based on mobility differences as they drift through an inert gas under the influence of a weak electric field [38] and is a well-established technique for the structural analysis of proteins [39], but can also be applied to bottom-up proteomic applications [38]. Traditional IM experiments are typically performed using a drift tube platform, which achieves separation by applying a uniform, static electric field. In the case of IM-MS implemented within the geometry of a Q-ToF instrument, mobility separation is achieved by using an RF ion guide, termed travelling wave ion guide, to generate a travelling voltage wave, which has been thoroughly reviewed in the literature [40, 41]. Combining the travelling wave principle of the device with elevated inert gas pressures results in a proportion of traversing ions rolling back on the wave, whilst the average effect propels ions in the direction of the travelling wave. Ion species with low mobility traverse slower, whereas ions of higher mobility traverse faster, resulting in a shorter drift time. Comparative advantages associated with IM-MS compared to drift tube variants include higher sensitivity and sufficiently faster data acquisition. However, drift tubes still have a

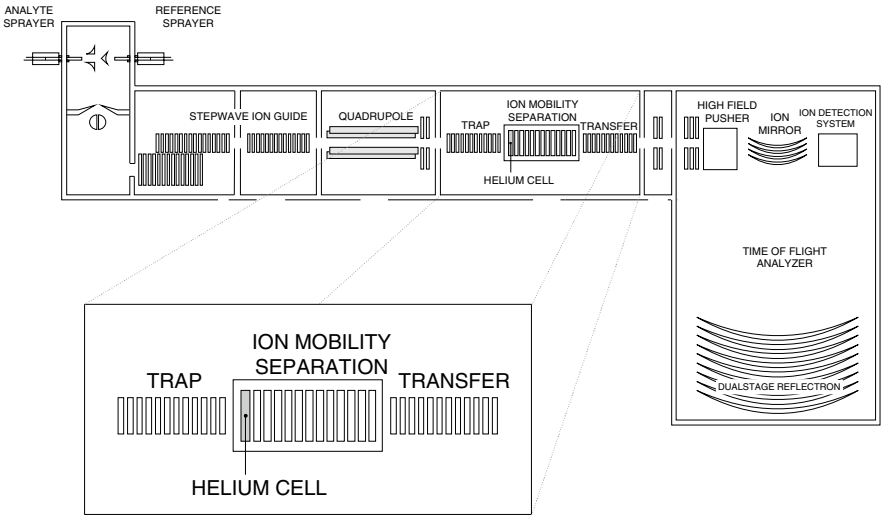


Fig. 3.5 A hybrid quadrupole-ion mobility-orthogonal acceleration time-of-flight mass spectrometer

distinct advantage in terms of the mobility resolution capabilities, which can be typically three orders of magnitude over that of current IM-MS [42].

Separation by IM operates within a millisecond timescale which fits perfectly between LC and ToF, which operate at the second and microsecond scale, respectively. The three domains are not completely orthogonal; however, the resulting system peak capacity that can be typically achieved is between 10- and 25-fold. It has been illustrated how peak capacity can be determined from a chromatographic standpoint. A similar approach can also be adopted for an oa-ToF mass analyzer (Eq. 3.4) and ion mobility separation (Eq. 3.5).

$$n_{\text{oaToF}} = \frac{\frac{R_{\text{oa-ToF}}}{m}.W_m}{n_1} \tag{3.4}$$

$$n_{\text{TWIM}} = \int_{k_{\min}}^{k_{\max}} \frac{R_{\text{IM,max}}}{\sqrt{k.k_{\max}}}\,dk \tag{3.5}$$

The peak capacity contribution made by the MS dimension, as outlined by Eq. 3.4, is defined by the mass spectrometer resolution at full width half maximum ($R_{\text{oa-ToF}}$), monoisotopic mass (m), mass distribution (W_m), and the number of identified isotopic peaks within a spectrum (n_l). Conversely, IM also contributes additional peak capacity provided by the travelling wave device [43], describing the relationship between resolution and ion mobility, where k represents mobility and $R_{\text{IM,max}}$ is an empirically derived maximum resolution travelling wave mobility separation. Based on the assumption that all three domains are completely orthogonal a product

describing the entire system peak capacity would be as defined by Eq. 3.6. In practice, however, the degree of orthogonality is compromised as discussed earlier, since the analytes rely on having the same physiochemical properties.

$$n_{\text{2DLC,IMS,oaToF}} \sim n_{\text{2DLC}} . n_{\text{TWIM}} . n_{\text{oaToF}} . f \tag{3.6}$$

3.4 Integrating Ion Mobility into a Data-Independent Strategy as Means of Increasing System Peak Capacity

A previous chapter introduced the concept of data-independent acquisition (DIA) for acquiring bottom-up proteomic data sets. Combining DIA with the strategies discussed here to increase peak capacity results in a sophisticated workflow that utilizes ion mobility to provide IM-DIA-MS. The principle of this method works in a similar manner as DIA, but with the additional degree of separation and specificity that is afforded by IM. Based on the instrument geometry shown in Fig. 3.5, IM-DIA-MS operates whereby the quadrupole mass analyzer is non-resolving. The collision cell is located within the travelling wave ion guide region and when operating in DIA mode, it is the primary stacked ring ion guide (SRIG) which is used to induce fragmentation. However, for IM-DIA-MS it is the secondary collision SRIG located after the ion mobility SRIG which is utilized for fragmentation. The primary SRIG maintains a static CE ensuring precursors are IM separated prior to fragmentation within the secondary SRIG, resulting in fragment ions sharing the same drift time as their associated precursors. This builds additional specificity into the analysis, since drift time and chromatographic retention time can now be used to correlate fragment ions with their respective precursors. For situations where the analyte density is large (i.e., the midpoint of a chromatographic gradient) the opportunity for multiple precursors to be present in the collision cell at a single time event is highly probable. The co-isolation of precursors sharing similar *m/z* and retention time is a phenomenon termed as chimericy. Since multiple MS/MS fragments are generated from multiple precursors the identification rate is significantly hampered resulting in unidentified fragments and hence under-sampling of the proteome. It has been reported that for highly complex samples, chimeras may be as high as 50 % of total spectra [44]. Relying solely on chromatography and mass resolution is not adequate to counter the effect. The additional separation capabilities provided by IM though do provide an opportunity to increase selectivity in those areas of extreme analyte density. Implementing IM-DIA overcomes some of the challenges associated with data-dependent analysis (DDA), such as the most abundant peptides typically being sampled and chimeric effects. Thoughts of increased sampling speed and sensitivity of instrumentation alone are deemed as insufficient and alternatively merging high resolution with a form of multiplexing (i.e., ion mobility) is thought to be necessary for comprehensive proteomic analysis [45]. Product ions are tentatively associated to precursors by means of a pre-database searching step. Isotope and charge state

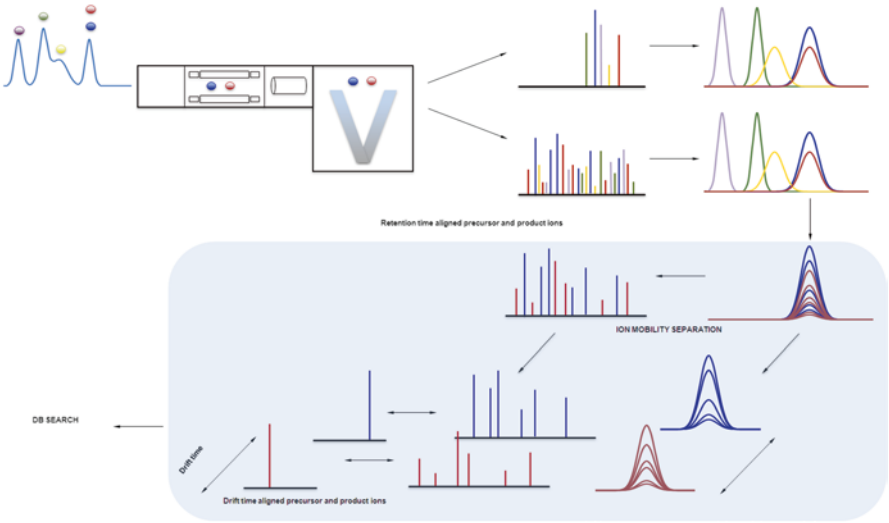


Fig. 3.6 Principle of IM-DIA showing retention time alignment for precursor and product ions is shown with additional drift time alignment for ion mobility workflows

information is collected for all precursors in addition to performing time and drift-alignment correlation (Fig. 3.6). Retention time-alignment correlation can be visualized by plotting the elution profiles of product ions from the elevated energy trace, resulting in identical elution profiles as their respective precursors. This chromatographic apex retention time principle forms the basis for associating precursor and product ions. The reasoning described for the alignment of chromatographic profiles can also be replicated with respect to the drift time domain, since fragment ions will share the same drift time as their associated precursor, thereby building additional specificity into the workflow.

Acquiring data sets using the methodologies outlined thus far is illustrated by a study involving the analysis of a *Rattus Norvegicus* exosome sub-proteome treated with either galactosamine (GalN) or lipopolysaccharide (LPS) culture media [46]. Comparative strategies were conducted involving 1D- and 2D (RP–RP)-LC in combination with DIA or IM-DIA. A three-way comparison of the results gained from 1D-LC–DIA-MS, 1D-LC–IM-DIA-MS, and 2D-LC–DIA-MS (Fig. 3.7) shows significant overlap with a large number of additional unique proteins identified with the implementation of IM. On average a 20 % increase is observed for the average number of peptides assigned to a protein suggesting the IM workflow to be more effective at sampling the proteome. Additional benefits include lower material consumption whilst increasing throughput and enhancing specificity of the identifications returned. The interaction between sensitivity and specificity has previously been addressed and has shown that increasing sensitivity alone does not result with increased identifications unless accompanied with additional specificity [46]. This is exemplified for the identified proteins of this study showing greater than twofold increase over the entire dynamic range (Fig. 3.8).

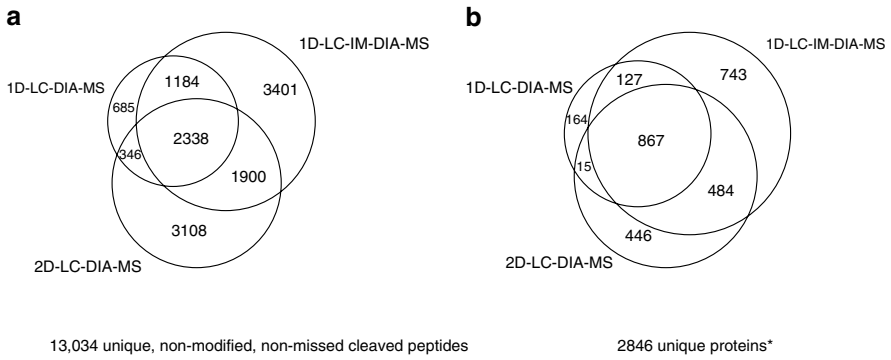


Fig. 3.7 A comparative evaluation of peptides (**a**) and proteins (**b**) identified from treated *Rattus Norvegicus* exosome samples. Both the peptide and protein results are subdivided into 1D (DIA), 1D (IM-DIA), and 2D (DIA) analyses. Identifications resulting from 1D were acquired at 10,000 FHWM resolution, whilst 2D was performed at 20,000 FHWM resolution (Rodriguez-Suárez et. al., Current Analytical Chemistry, 2013, 9, 199–211)

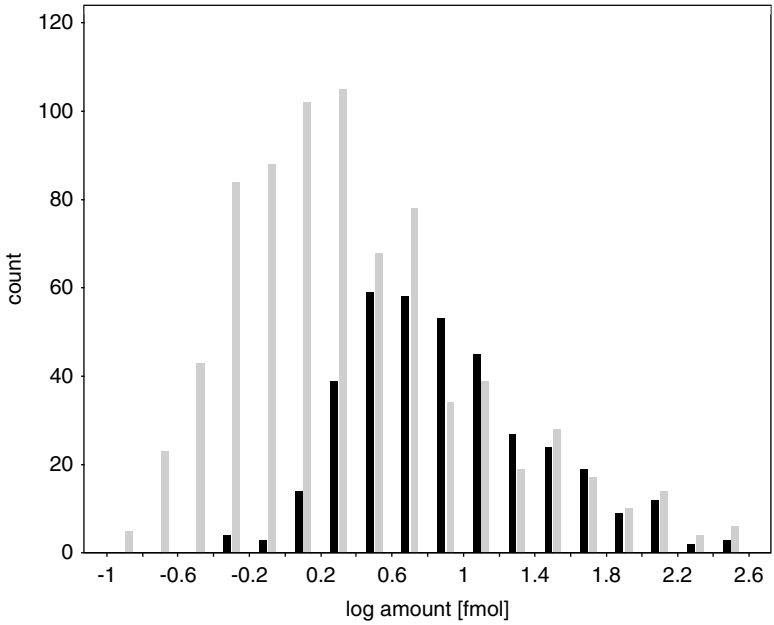


Fig. 3.8 Based on a 1 µg loading of control *Rattus Norvegicus* exosome, the normalized estimated molar amounts can be plotted to show a distribution expanding over two orders of magnitude of dynamic range. Data resulting from 1D-LC–DIA-MS (*black*) and 1D-LC–IM-DIA-MS (*gray*) is provided (Rodriguez- Suárez et. al., Current Analytical Chemistry, 2013, 9, 199–211)

The bulk of these discussions have focused on implementing analytical workflows, yet thought should also be given to the informatic strategy applied to such data sets. For example, the protein identification rate could be further improved by implementing the use of spectral libraries and/or DIA fragment ion repositories [47].

3.5 Informatics: Processing IM-DIA-MS Data Sets

The nature of DIA and IM-DIA strategies does not rely on instrument selection of the precursors as with DDA; the capacity to relate precursors with their fragment ions is therefore reliant on post-acquisition informatics using sophisticated algorithms for peak detection [48, 49], precursor–product ion correlation, and database searching [50].

3.5.1 Precursor–Product Ion Correlation

The tentative association of precursor and product ions as part of the pre-database search step is performed prior to the data being compiled into a list containing four dimensional attributes (ion mobility, retention time, *m*/*z*, and intensity) providing a multidimensional distribution. The presence of background noise can result in over-counting and therefore needs to be accounted for by applying a convolution filter to suppress the effect. Each of these deconvoluted measurements is referred to as accurate-mass, retention time (AMRT) components. For cases where multidimensional chromatography has been used, raw data files are processed individually before merging the processed outputs prior to database searching, ensuring that all peptides representing a single protein can be identified and quantified in a single event. Should peptides be shared between fractions, the summing process employed allows quantification.

3.5.2 Database Searching

A comprehensive overview discussing a search algorithm for qualitative identifications based on DIA data sets has been described in the literature [51]. Briefly, following peak detection and alignment of the low and elevated energy AMRTs a constructed list of precursor/product ion associations is interrogated for putative peptide identifications. A precursor/product subset list and a user-defined database are subjected to a pre-assessment search based on a set of physiochemical properties specific for tryptic peptides and proteins, which are applicable to both the liquid and gas phase dimensions. The database element consists of a decoy database (either randomized or reversed) being merged with the original user database, allowing a

false discovery rate and minimum protein score to be automatically determined. The search and score models are optimized using the previously mentioned physiochemical properties. The next stage of the algorithm search is divided into three steps. The first step queries the precursor/product ion entries against the protein database for sequences containing no tryptic-missed cleavages or variable modifications. A combination of precursor and fragment ion mass tolerances and accuracy of fit to the physiochemical properties provides a mechanism by which peptide identifications are scored and ranked. This process is continuously repeated until either the false-positive rate is exceeded or protein identifications no longer exceed the minimum score. The second step of the query specifically focuses on identifying modifications and nonspecific cleavages for protein entries resulting from the first pass as well as in-source fragments and precursor neutral losses. The final step of the query aligns any unidentified low and elevated energy AMRTs from the previous two steps before searching against the full database with no limitation on product ion intensity and restriction on the number of modifications per protein.

3.6 Importance of Peak Capacity for Systems Biology: A Multi-omic Biomarker Case Study

Complex data sets as described can be readily generated from an analytical perspective. The inherent difficulty arises with the search results and interpreting these into a biological context that can provide insight of both physiology and disease state of the model being studied. Data sets can often be large and originate from not only proteomics but also other omic areas such as metabolomics, lipidomics, genomics, and transcriptomics to provide a comprehensive understanding of biochemical processes and ultimately provide potential biomarkers. Since there is no single analytical system or workflow that can generate information for all the omic areas, gaining access to data of a common format for searching and mapping pathways can prove to be troublesome. Figure 3.9 represents an example workflow detailing how multi-omic data sets are integrated and interrogated for system networks analysis [52].

A case study representing multi-omic data using techniques to improve peak capacity for potentially identifying disease biomarkers is described here. The study focuses on a rare genetic kidney disorder termed idiopathic nephritic syndrome (INS), affecting only a small percentage of paediatric patients. The condition arises from a faulty glomerulus, resulting in proteinuria additionally characterized by edema, hypoalbuminemia, and hyperlipidemia, as well as increased levels of cholesterol and triglyceride [53]. The study cohort consisted of urine samples from paediatric subjects, control, and INS diagnosed. Samples were purified and prepared appropriately for proteomic or metabolomic analysis [54, 55] prior to an experimental strategy which combined LC and DIA, with the proteomic experiments specifically utilizing IM-DIA-MS as a means of generating label-free data. In order to establish differences between the two groups, statistical analysis using multivariate methods of both data sets showed significant deviations between both cohorts.

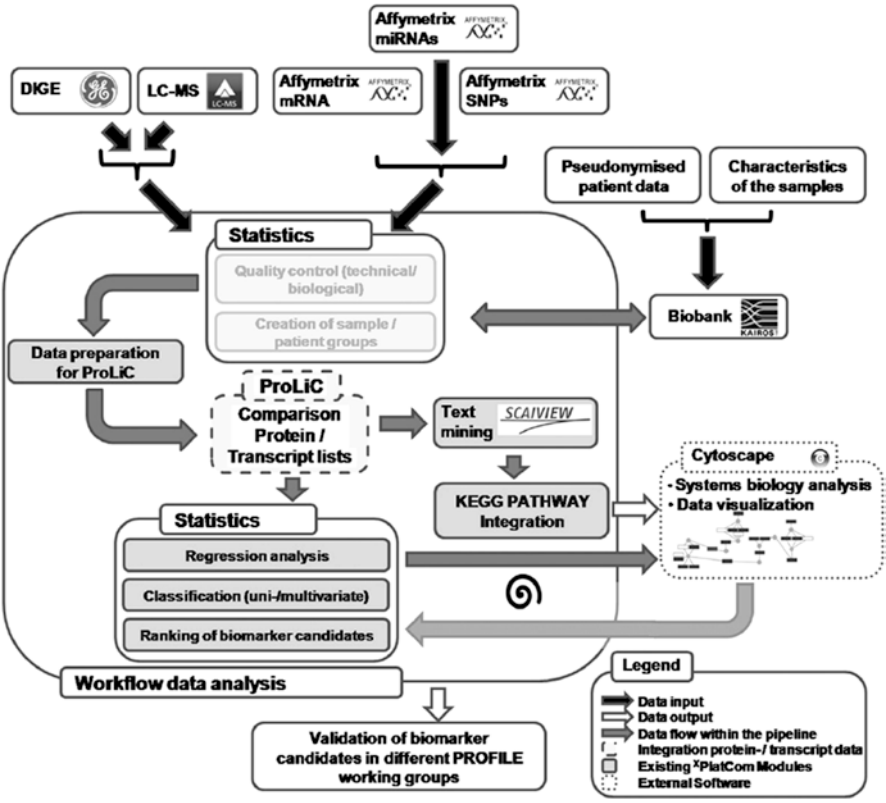


Fig. 3.9 Example pipeline for compiling and interrogating multi-omic data sets to pathway and network mapping studies (Kohl et. al., Biochim. Biophys. Acta, 2014, 1844, 52–62)

Identifying analytes responsible for such perturbation started by reviewing the protein data. In total more than 300 proteins were identified, with 80 % showing significant fold change (greater than two) and *p*-values of less than 5 % across all subjects. Hierarchical analysis (Fig. 3.10a) allows the visual identification of protein expression trends across conditions in addition to potential intra-subject variances. Regulated proteins showing distinct expression trends can be interrogated further as demonstrated with the example peptide associated to the prostaglandin receptor (Fig. 3.10b). The metabolite data was interpreted differently by means of constructing a loadings plot (Fig. 3.11) from previously derived OPLS-DA analysis. The loadings plot is constructed such that compounds contributing the greatest variance with the highest probability are target analytes for identification. In an effort to combine the data streams and understand their contributions in a biological context, pathway analysis was conducted to explore and visualize interactions and networks resulting from both omic data sets. The resultant network indicates relevant pathways such as chronic fatigue syndrome and neurological signs as major contributors.

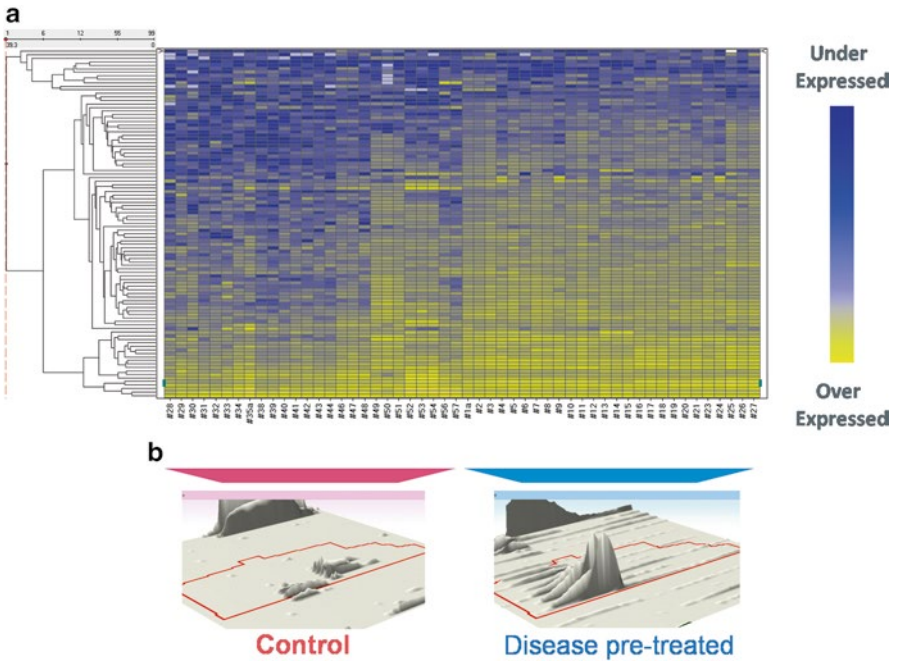


Fig. 3.10 (**a**) Hierarchical cluster analysis of statistically (ANOVA) significantly relevant regulated urinary proteins identified with 3 or more peptides and fold changes greater than 2. Subjects grouped within the pink banner are control, whilst disease pre-treated subjects correspond to the blue banner. The 3D montage images (**b**) are representative of prostaglandin receptor peptide TMLLQPAGSLGSYSYR

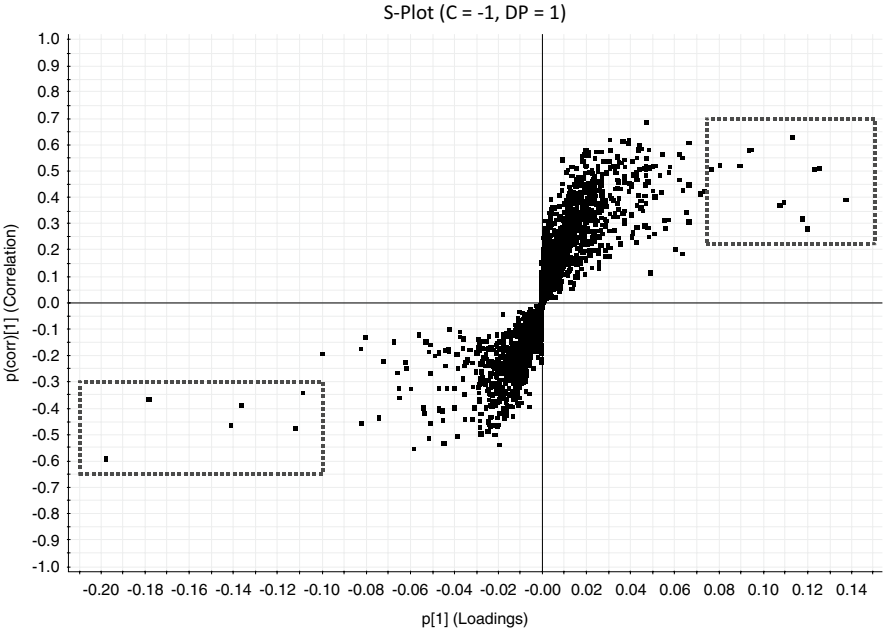


Fig. 3.11 Metabolite loadings plot resulting from OPLS-DA of disease pre-treated versus control subjects based in positive ion mode. Metabolites contributing the greatest variance such as hydroxyphenyl acetate and uridine are represented within the highlighted areas

3.7 Summary

The quest for identifying potential biomarkers is complicated and challenging with the majority of studies involving highly complex samples over wide dynamic ranges that require sophisticated analytical workflows to acquire, process, and biologically interpret data sets. The most interesting and often most significant proteins are those which are to be found at low abundance. From an LC–MS perspective, a number of potential strategies have been introduced to address how the proteome can be explored at greater depth, whilst limiting sample consumption and maximizing throughput. Multidimensional chromatography combined with ion mobility enabled DIA acquisition schema demonstrate how system peak capacity can be optimized. As technologies advance the volume of data increases in size and complexity; therefore, informatic requirements and capabilities should also be addressed. These combined efforts show not only increased peak capacity but also enhanced specificity, afforded by the chromatographic and drift time alignment of precursor and product ions, providing additional confidence to identifications.

Advancements towards finding relevant biomarkers are increasingly moving towards a system biology approach, whereby multi-omic data sets can provide a great insight into pathway information and the interaction of networks. Amalgamating techniques such as IM-enabled workflows with multi-omic biological pathway analysis has the potential to assist in our understanding of disease and its etiology for drug discovery.

Acknowledgements The authors would like to thank our collaborators who have provided figures and granted permission to use their collateral. In particular we would like to acknowledge Dr. Stefan Tenzer, Dr. Eva Rodriguez-Suárez, Dr. Sandra Kraljević Pavelić, Dr. Martin Gilar, Dr. John Shockcor, Kenneth Fountain, and Eric Grumbach. Finally Dr. Johannes P.C. Vissers is thanked for constructive comments during review.

References

1. Giddings JC (1984) Two-dimensional separations: concept and promise. Anal Chem 56: 1258A–1260A, 1262A, 1264A
2. Swartz ME (2007) Separation science & technology, vol 8. Academic, pp 145–147
3. Van Deemter JJ, Zuiderweg FJ, Klinkenberg A (1956) Longitudinal diffusion and resistance to mass transfer as causes of nonideality in chromatography. Chem Eng Sci 5:271
4. Plumb R, Castro-Perez J, Granger J, Beattie I, Joncour K, Wright A (2004) Ultra-performance liquid chromatography coupled to quadrupole-orthogonal time-of-flight mass spectrometry. Rapid Commun Mass Spectrom 18:2331–2337
5. Levy AL, Chung D (1953) Two-dimensional chromatography of amino acids on buffered papers. Anal Chem 25:396–399
6. Honneger CG (1961) Dünnschicht-Ionophorese und Dünnschicht-Ionophorese-Chromatographie. Helv Chim Acta 44:173
7. Rabilloud T (2002) Two-dimensional gel electrophopresis in proteomics: old, old fashioned, but it still climbs up the mountains. Proteomics 2:3–10

8. Neue UD, Mazzeo JR (2001) A theoretical study of the optimization of gradients at elevated temperature. J Sep Sci 24:921–929
9. Karger BL, Snyder LR, Horvarth C (1973) An introduction to separation science. Wiley, New York
10. Li X, Stoll DR, Carr PW (2009) Equation for peak capacity estimation in two-dimensional liquid chromatography. Anal Chem 81:845–850
11. Dowell JA, Frost DC, Zhang J, Li L (2008) Comparison of two-dimensional fractionation techniques for shotgun proteomics. Anal Chem 80:6715–6723
12. Washburn MP, Wolters D, Yates JR (2001) Large-scale analysis of the yeast proteome by mułtidimensional protein identification technology. Nat Biotechnol 19:242–247
13. Wolters DA, Washburn MP, Yates JR (2001) An automated multidimensional protein identification technology for shotgun proteomics. Anal Chem 73:5683–5690
14. Peng J, Elias JE, Thoreen CC, Licklider LJ, Gygi SP (2003) Evaluation of multidimensional chromatography coupled with tandem mass spectrometry (LC/LC-MS/MS) for large-scale protein analysis: the yeast proteome. J Proteome Res 2:43–50
15. Vollmer M, Horth P, Nagele E (2004) Optimization of two-dimensional off-line LC/MS separations to improve resolution of complex proteomic samples. Anal Chem 76:5180–5185
16. Opiteck GJ, Jorgenson JW, Anderegg RJ (1997) Two-dimensional SEC/RPLC coupled to mass spectrometry for the analysis of peptides. Anal Chem 69:2283–2291
17. Alvarez-Manilla G, James I, Guo Y, Warren NL, Orlando R, Pierce M (2006) Tools for glycoproteomic analysis: size exclusion chromatography facilitates identification of tryptic glycopeptides with N-linked glycosylation sites. J Proteome Res 5:701–708
18. Preud'homme H, Far J, Gil-Casal S, Lobinski R (2012) Large-scale identification of selenium metabolites by online size-exclusion-reversed phase liquid chromatography with combined inductively coupled plasma (ICP-MS) and electrospray ionization linear trap-orbitrap mass spectrometry (ESI-MSn). Metallomics 4:422–432
19. Barqawi H, Ostas E, Liu B, Carpenter JF, Binder WH (2012) Multidimensional characterization of α, ω-telechelic poly(Є-caprolactone)s via online coupling of 2D chromatographic methods (LC/SEC) and ESI-TOF/MALDI-TOF-MS. Macromolecules 45:9779–9790
20. Albuquerque CP, Smolka MB, Payne SH, Bafna V, Eng J, Zhou H (2008) A multidimensional chromatography technology for in-depth phosphoproteome analysis. Mol Cell Proteomics 7:1389–1396
21. Palma SD, Zoumaro-Djayoon A, Peng M, Post H, Preisinger C, Munoz J, Heck AJR (2013) Finding the same needles in the haystack? A comparison of phosphotyrosine peptides enriched by immuno-affinity precipitation and metal-based affinity chromatography. J Proteomics 91:331–337
22. Frantzi M, Zoidakis J, Papadopulos T, Zürbig P, Katafigiotis J, Stravodimos K, Lazaris A, Giannopoulou I, Ploumidis A, Mischak H, Mullen W, Vlahou A (2013) IMAC fractionation in combination with LC-MS reveals H2B and NIF-1 peptides as potential bladder cancer biomarkers. J Proteome Res 12:3969–3979
23. Liu S, Hughes C, Lajoie G (2012) Recent advances and special considerations for the analysis of phosphorylated peptides by LC-ESI-MS/MS. Curr Anal Chem 8:35–42
24. Stasyk T, Huber LA (2012) Mapping in vivo signal transduction defects by phosphoproteomics. Trends Mol Med 18:43–51
25. Boersema PJ, Divecha N, Heck AJR, Mohammed S (2007) Evaluation and optimization of ZIC-HILIC-RP as an alternative MudPIT strategy. J Proteome Res 6:937–946
26. Xie F, Smith RD, Shen Y (2012) Advanced proteomic liquid chromatography. J Chromatogr A 1261:78–90
27. Wang C, Yuan J, Wang Z, Huang L (2013) Separation of one-pot procedure released O-glycans as 1-phenyl-3-methyl-5-pyrazolone derivatives by hydrophilic interaction and reversed-phase liquid chromatography followed by identification using electrospray mass spectrometry and tandem mass spectrometry. J Chromatogr A 1274:107–117
28. Lau E, Lam MPY, Siu SO, Kong RPW, Chan WL, Zhou Z, Huang J, Lo C, Chu IK (2011) Combinatorial use of offline SCX and online RP-RP liquid chromatography for iTRAQ-based quantitative proteomics applications. Mol Biosyst 7:1399–1408

29. Gilar M, Olivova P, Daly AE, Gebler JC (2005) Orthogonality of separation in two-dimensional liquid chromatography. Anal Chem 77:6426–6434
30. Gilar M, Olivova P, Daly AE, Gebler JC (2005) Two-dimensional separation of peptides using RP-RP-HPLC system with different pH in first and second separation dimensions. J Sep Sci 28:1694–1703
31. Francois I, Cabooter D, Sandra K, Lynen F, Desmet G, Sandra P (2009) Tryptic digest analysis by comprehensive reversed phasextwo reversed phase liquid chromatography (RP-LCx2RP-LC) at different pH's. J Sep Sci 32:1137–1144
32. Francois I, de Villiers A, Tienpont B, David F, Sandra P (2008) Comprehensive two-dimensional liquid chromatography applying two parallel columns in the second dimension. J Chromatogr A 1178:33–42
33. Siu SO, Lam MPY, Lau E, Kong RPW, Lee SMY, Chu IK (2011) Fully automated two-dimensional reversed-phase capillary liquid chromatography with online tandem mass spectrometry for shotgun proteomics. Proteomics 11:2308–2319
34. Zhou F, Cardoza JD, Ficarro SB, Adelment GO, Lazaro JB, Marto JA (2010) Online nanoflow RP-RP-MS reveals dynamics of multicomponent Ku complex in response to DNA damage. J Proteome Res 9:6242–6255
35. Scigelova M, Hornshaw M, Giannokulas A, Makarov A (2011) Fourier transform mass spectrometry. MCP 1–19. doi:10.1074/mcp.M111.009431
36. Morris HR, Paxton T, Dell A, Langhorn B, Berg M, Bordoli RS, Hoyes J, Bateman RH (1996) High sensitivity collisionally-activated decomposition tandem mass spectrometry on a novel quadrupole/orthogonal-acceleration time-of-flight mass spectrometer. Rapid Commun Mass Spectrom 10:889–896
37. Pringle SD, Giles K, Wildgoose JL, Williams JP, Slade SE, Thalassinos K, Bateman RH, Bowers MT, Scrivens JH (2007) An investigation of the mobility separation of some peptide and protein ions using a new hybrid quadrupole/travelling wave IMS/oa-ToF instrument. Int J Mass Spectrom 261:1–12
38. Mason EA, McDaniel EW (1973) The mobility and diffusion of ions in gases. Wiley, New York
39. Ruotolo BT, Giles K, Campuzano I, Sandercock AM, Bateman RH, Robinson CV (2005) Evidence for macromolecular protein rings in the absence of bulk water. Science 310:1658
40. Harvey SR, Macphee CE, Barran PE (2011) Ion mobility mass spectrometry for peptide analysis. Methods 54:454–461
41. Giles K, Pringle SD, Worhtington KR, Little D, Wildgoose JL, Bateman RH (2004) Applications of a travelling wave-based radio-frequency-only stacked ring ion guide. Rapid Commun Mass Spectrom 18:2401–2414
42. Dugourd P, Hudgins RR, Clemmer DE, Jarrold MF (1997) High-resolution ion mobility measurements. Rev Sci Instrum 68:1122–1129
43. Shvartsburg AA, Smith RD (2008) Fundamentals of travelling wave ion mobility spectrometry. Anal Chem 80:9689–9699
44. Houel S, Abernathy R, Renganathan K, Meyer-Arendt K, Ahn NG, Old WM (2010) Quantifying the impact of chimera MS/MS spectra on peptide identification in large-scale proteomics studies. J Proteome Res 9:4152–4160
45. Michalski A, Cox J, Mann M (2011) More than 100,000 detectable peptide species elute in single shotgun proteomics runs but the majority is inaccessible to data-dependent LC-MS/MS. J Proteome Res 10:1785–1793
46. Rodriguez-Suarez E, Hughes C, Gethings L, Giles K, Wildgoose J, Stapels M, Fadgen KE, Geromanos SJ, Vissers JPC, Elortza F, Langridge JI (2012) An ion mobility assisted data independent LC-MS strategy for the analysis of complex biological samples. Curr Anal Chem 9:199–211
47. Geromanos SJ, Hughes C, Golick D, Ciavarini S, Gorenstein MV, Richardson K, Hoyes JB, Vissers JP, Langridge JI (2011) Simulating and validating proteomics data and search results. Proteomics 11:1189–1211

48. Thalassinos K, Vissers JP, Tenzer S, Levin Y, Thompson JW, Daniel D, Mann D, Delong MR, Moseley MA, America AH, Ottens AK, Cavey GS, Efstathiou G, Scrivens JH, Langridge JI, Germoanos SJ (2012) Design and application of a data-independent precursor and product ion repository. J Am Soc Mass Spectrom 23:1808–1820
49. Geromanos SJ, Vissers JPC, Silva JC, Dorschel CA, Li GZ, Gorenstein MV, Bateman RH, Langridge JI (2009) The detection, correlation, and comparison of peptide precursor and product ions from data independent LC-MS with data dependent LC-MS/MS. Proteomics 9:1683–1695
50. Silva JC, Denny R, Dorschel CA, Gorenstein M, Kass IJ, Li GZ, McKenna T, Nold MJ, Richardson K, Young P, Geromanos S (2005) Quantitative proteomic analysis by accurate mass retention time pairs. Anal Chem 77:2187–2200
51. Li GZ, Vissers JPC, Silva JC, Golick D, Gorenstein MV, Geromanos SJ (2009) Database searching and accounting of multiplexed precursor and product ion spectra from the data independent analysis of simple and complex peptide mixtures. Proteomics 9:1696–1719
52. Kohl M, Megger DA, Trippler M, Meckel H, Ahrens M, Bracht T, Weber F, Hoffmann AC, Baba HA, Sitek B, Schlaak JF, Meyer HE, Stephan C, Eisenacher M (2014) A practical data processing workflow for multi-OMICS projects. Biochem Biophys Acta 1844:52–62
53. Chesney RW (1999) The idiopathic nephrotic syndrome. Curr Opin Pediatr 11:158–161
54. Vaezzadeh AR, Briscoe AC, Steen H, Lee RS (2010) One-step sample concentration, purification, and albumin depletion method for urinary proteomics. J Proteome Res 9:6082–6089
55. Want EJ, Wilson ID, Gika H, Theodoridis G, Plumb RS, Shockcor J, Holmes E, Nicholson JK (2010) Global metabolic profiling procedures for urine using UPLC-MS. Nat Protoc 5:1005–1018

Chapter 4
Quantitative Shotgun Proteomics with Data-Independent Acquisition and Traveling Wave Ion Mobility Spectrometry: A Versatile Tool in the Life Sciences

Lewis M. Brown

Abstract Data-independent acquisition (DIA) implemented in a method called MS^E can be performed in a massively parallel, time-based schedule rather than by sampling masses sequentially in shotgun proteomics. In MS^E alternating low and high energy spectra are collected across the full mass range. This approach has been very successful and stimulated the development of variants modeled after the MS^E protocol, but over narrower mass ranges. The massively parallel MS^E and other DIA methodologies have enabled effective label-free quantitation methods that have been applied to a wide variety of samples including affinity pulldowns and studies of cells, tissues, and clinical samples. Another complementary technology matches accurate mass and retention times of precursor ions across multiple chromatographic runs. This further enhances the impact of MS^E in counteracting the stochastic nature of mass spectrometry as applied in proteomics. Otherwise significant amounts of data in typical large-scale protein profiling experiments are missing. A variety of software packages perform this function similar in concept to matching of accurate mass tags. Another enhancement of this method involves a variation of MS^E coupled with traveling wave ion mobility spectrometry to provide separations of peptides based on cross-sectional area and shape in addition to mass/charge (*m/z*) ratio. Such a two-dimensional separation in the gas phase considerably increases protein coverage as well as typically a doubling of the number of proteins detected.

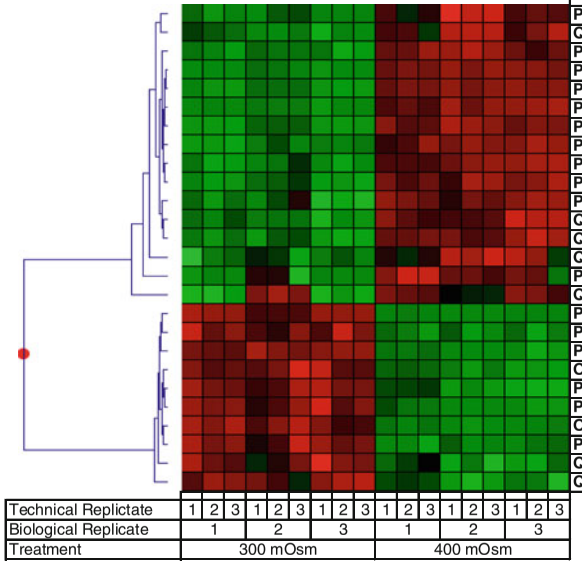
L.M. Brown (✉)
Quantitative Proteomics Center, Department of Biological Sciences,
Columbia University, New York, NY 10027, USA
e-mail: lb2425@columbia.edu

A.G. Woods and C.C. Darie (eds.), *Advancements of Mass Spectrometry in Biomedical Research*, Advances in Experimental Medicine and Biology 806,
DOI 10.1007/978-3-319-06068-2_4, © Springer International Publishing Switzerland 2014

These developments along with advances in ultrahigh pressure liquid chromatography have resulted in the evolution of a robust and versatile platform for label-free protein profiling.

4.1 MS^E and Other Data-Independent Acquisition Strategies

Shotgun proteomics is a strategy with broad applicability, and is based on digestion of proteins with proteolytic enzymes and analysis of the resulting peptides by liquid chromatography–mass spectrometry (LC/MS). The method of choice for data collection has long been a data-dependent acquisition (DDA) method in which acquisition parameters are modified in real time by selecting a narrow mass window in a quadrupole analyzer to allow precursor ions to pass through for fragmentation after an initial survey scan [1]. This was an important enhancement to the initial approach of using stored mass and retention time information from previous liquid chromatography runs [2]. More recently, a strategy has been proposed [3, 4] and finally implemented [5, 6] using data-independent acquisition (DIA) in a method called MS^E to perform data collection in a massively parallel, time-based schedule rather than attempting to sample masses sequentially. MS^E has been extensively developed as a method in which alternating low and high energy scans are recorded across the full mass range and was implemented on quadrupole time-of-flight mass spectrometers from Waters Corporation [5, 6]. The original approach has been very successful and has stimulated the development of variants modeled after the MS^E protocol. These include sequential proteomic method precursor acquisition independent from ion count (PAcIFIC) [7], windowed data-independent acquisition of total high-resolution (SWATH) [8], all-ion fragmentation (AIF) [9], and multiplexed data-independent acquisition (MSX) [10]. These recent adaptations can be utilized on other instrument platforms, although only MS^E is unique in that it covers the entire mass range in each scan. Our group has used MS^E effectively on a routine basis for cells, tissues, and affinity preparations in a label-free approach to large-scale protein profiling [11–15]. The validity, power, and effectiveness of MS^E are now firmly established and extensively validated independently by many groups [16–18]. One example is our work on human somatic stem cells, where we were able to use this technique to demonstrate clear differential expression of proteins such as aldose reductase and many other proteins (Fig. 4.1) which responded to the combination of growth factor priming and increased osmotic pressure [13]. The roles of some of these proteins are characteristic of the physiological states studied [13]. In that experiment, 5′-nucleotidase and transgelin were detected as differentially expressed, and have been previously linked to cell differentiation state. Data-independent label-free profiling was demonstrated in that work to be a useful tool in characterizing cellular responses to treatment regimes, and as an aid to optimization of cell priming protocols for cartilage tissue engineering.



Uniprot Accession Number	Protein Name	300 mOsm (Mean x 10^{-5})	400 mOsm (Mean x 10^{-5})	Ratio	*P-value*
P25705	ATP synthase subunit alpha, mitochondrial	0.32	0.59	1.79	2.52E-09
Q99879	Histone H2B type F-S	0.47	0.81	1.68	1.21E-10
P14618	Pyruvate kinase isozymes M1/M2	1.69	2.96	1.73	1.37E-30
P15121	Aldose reductase	8.99	32.14	3.52	<1.0E-45
P14618-2	Isoform M1 of Pyruvate kinase isozymes M1/M2	23.78	40.41	1.68	<1.0E-45
P52895	Aldo-keto reductase family 1 member C2	0.38	1.89	4.93	<1.0E-45
P04179	Superoxide dismutase [Mn], mitochondrial	0.16	0.72	4.30	4.20E-45
P40939	Trifunctional enzyme subunit alpha, mitochondrial	0.05	0.12	2.61	2.34E-31
P61026	Ras-related protein Rab-10	0.37	0.58	1.56	4.34E-20
P29401	Transketolase	0.27	0.45	1.62	2.09E-21
P30613	Pyruvate kinase isozymes R/L	1.48	2.49	1.65	2.03E-19
Q9NQC3	Reticulon-4	0.41	0.69	1.66	4.13E-18
Q99536	Synaptic vesicle membrane protein VAT-1 homolog	0.19	0.40	2.07	2.23E-31
Q16695	Histone H3.1	0.35	0.63	1.79	9.11E-07
P06703	Protein S100-A6	0.15	0.50	3.26	1.06E-08
Q15942	Zyxin	0.23	0.36	1.51	0.0088
P02751	Fibronectin	16.31	2.99	0.18	<1.0E-45
P19105	Myosin regulatory light chain 12A	0.73	0.46	0.63	9.90E-22
P42224	Signal transducer & activator of transcription 1-α/b	0.19	0.05	0.28	<1.0E-45
O00469-2	Procollagen-lysine,2-oxoglutarate 5-dioxygenase 2	0.17	0.06	0.35	7.05E-37
P21589	5'-nucleotidase	3.04	1.87	0.62	7.70E-40
P40261	Nicotinamide N-methyltransferase	0.26	0.12	0.46	4.46E-34
O15460	Prolyl 4-hydroxylase subunit alpha-2	0.55	0.21	0.38	<1.0E-45
P13674	Prolyl 4-hydroxylase subunit alpha-1	0.14	0.06	0.42	6.85E-27
Q9Y696	Chloride intracellular channel protein 4	0.61	0.39	0.65	6.10E-15
Q01995	Transgelin	7.32	4.34	0.60	9.63E-21

Fig. 4.1 Example of expression data for selected proteins in agglomerative hierarchical cluster of Z-score transformed intensity data as processed by the Rosetta Elucidator program. These proteins have at least 1.5-fold response to the treatment, and are represented by at least two peptides for both identification and quantification at ratio *P*-values as calculated by Elucidator. Mean relative abundance of each protein in control and treated as well as abundance ratio is listed. Cluster coloration indicates protein abundance in the sample (*red* indicates higher abundance; *green* indicates lower abundance and *black* equal abundance). Reprinted with permission from Oswald et al. [13]. Copyright (2011) American Chemical Society

4.2 DIA Strategies Enhanced by Accurate Mass and Retention Time Matching Across Multiple Chromatographic Runs

Our work on adipose-derived stem cells [13] uses the Elucidator Protein Expression Data Analysis System from Rosetta (and Ceiba Solutions), a commercial program to match accurate mass and retention time intensities of precursor ions across multiple chromatographic runs. This enabled measurement of the intensities of precursor peptides even if fragmentation is not particularly successful for a particular peptide in a particular chromatographic run. This approach to a large extent counters the limitation imposed by the stochastic nature of mass spectrometry that otherwise results in a large amounts of missing data in large-scale experiments. A similar strategy is used by the TransOmics Informatics for Progenesis QI (TOIP) from Waters Corporation [19], a program also specifically enhanced to take advantage of ion mobility separations (see below). Coupling mass and retention time matching programs such as these with DIA technologies including MS^E provides a particularly powerful label-free platform that enables large and complex experiments otherwise difficult to conduct by other methods.

These software approaches are enabled by new faster instruments capable of data-independent scanning, but owe their conceptual inspiration to the concept of an accurate mass tags database originally conceived by Smith's group [20].

4.3 Enhancement of MS^E: Coupling to Traveling Wave ion Mobility Spectrometry (TWIMS)

A variation of MS^E coupled with traveling wave ion mobility spectrometry (TWIMS) provides the capability to perform separations of peptides based on cross-sectional area and shape in addition to their mass/charge (m/z) ratio. The use of TWIMS for shotgun proteomics is now well established and has been validated independently by a number of groups [21–23]. A two-dimensional separation is achieved in the gas phase, providing considerably increased protein coverage, typically a doubling of the number of proteins detected, a significant advantage [12, 24]. The complex TWIMS data requires a powerful computing platform, typically with a graphics processing unit (GPU)-equipped computer with 448 cores or more [12, 24]. The TWIMS principle is illustrated in Fig. 4.2 where integrating ion mobility drift time enhances identification of peptides [25]. In a study of a bacteriophage virion proteome [26], MS^E with TWIMS was consistently more effective than more conventionally employed DDA method. Current versions of the ProteinLynx Global Server commercial program [5, 6] (Waters Corp.) have been enabled to process TWIMS data.

We have applied these techniques in many other systems [15, 26] including patient samples [27]. Such analyses include fold-change and p-value determinations, providing an unbiased view of phenotype or biological responses to experimental treatments. An example is shown in Fig. 4.3 (mouse hippocampus) [12].

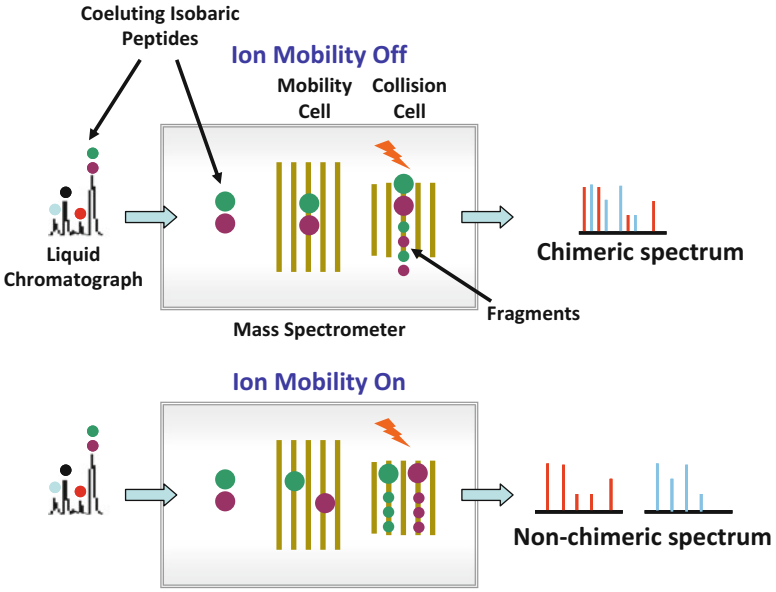


Fig. 4.2 Diagram illustrating the role of ion mobility spectrometry in addressing the challenge of peptides with both the same elution time in liquid chromatography (co-eluting) and having the same mass (isobaric). Without ion mobility, fragmentation of isobaric peptides in the collision cell results in a chimeric spectrum which is difficult to interpret. When ion mobility is activated, peptides are separated in the mobility cell on the basis of their drift time. The end result is cleaner spectra with reduced chimeracy [26]. Reprinted from Journal of Virological Methods, Vol 195, Moran, Deborah; Cross, Trevor; Brown, Lewis M.; Colligan, Ryan M; Dunbar, David, Data-independent acquisition (MS^E) with ion mobility provides a systematic method for analysis of a bacteriophage structural proteome, pp. 9–17, 2014, with permission from Elsevier

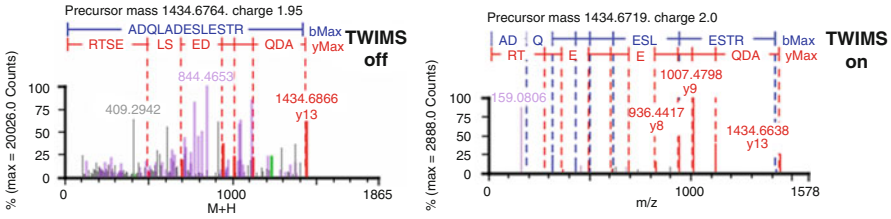


Fig. 4.3 Fragment ion spectra of peptide ADQLADESLESTR with TWIMS off and TWIMS on. *Purple* and *grey* peaks represent interferences and noise, respectively [12]

This figure represents fragmentation spectra collected with and without TWIMS activated. It can be clearly seen from this example that the interfering peaks from other peptides (derived from other proteins, magenta) and unassigned peaks (grey) are largely separated out by TWIMS. These sorts of comparisons can be made in data where spectra with or without TWIMS can be readily compared for the same peptide.

With this technique it is possible to quantify large numbers of proteins and generate abundances (examples in Table 4.1) for each LC/MS/MS run and for each biological replicate, and calculate means and standard deviations with coefficient of variation as low as 10–20 % for replicate analyses [12]. Large-scale experiments can be done on a routine basis with recording of many thousands of protein abundance values for multiple biological and technical replicates.

MS^E combined with TWIMS is being applied to a remarkable variety of practical biological problems including the interactome of the RNA-binding protein RALY [28], analysis of chaperonins and biosynthetic enzymes in the unculturable bacterial endosymbiont *Blochmannia*, [29] and quantitative analysis of human embryonic kidney cells proteome following sialic acid overproduction [30].

In another practical example of this technique, *Drosophila melanogaster* was evaluated for the role of dMyc in the larval fat body (Table 4.2). The fat body is a tissue that functions as a sensor of circulating nutrients to control the release of Drosophila insulin-like peptides (Dilps) influencing growth and development. Using MS^E and TWIMS, it was demonstrated that dMyc affected expression of hexokinase C and pyruvate kinase, key regulators of glycolysis, as well as of stearoyl-CoA desaturase (Desat1). Desat1 is an enzyme that is necessary for monosaturation and production of fatty acids, and its reduction affects dMyc and the ability to induce fat storage and resistance to animal survival.

These techniques can also be applied very effectively to chemoproteomic affinity experiments. For example, we have identified a protein target of a small molecule (RSL3) that is an inducer of a novel form of cell death (ferroptosis) and a potential novel anticancer agent [24]. In that work we also used DIA (MS^E) with TWIMS. A total of 1,353 proteins were detected in 27 $HDMS^E$ chromatograms. Analysis focused on 979 proteins with identification and quantitation supported by two or more unique peptides. Most detected proteins were not differentially bound to fluorescein-RSL3 beads compared to controls. Only one protein, glutathione peroxidase 4 (GPX4_HUMAN), was significantly enriched ($P<0.01$) and had the highest fold-change (26-fold in the affinity preparation compared to the inactive analogue and 13-fold compared to the control of preincubation with free RSL3). These studies identified GPX4 as an essential regulator of ferroptotic cancer cell death using data-independent profiling with TWIMS and these results were confirmed by an extensive series of corroborative experiments [24]. Isotopic clusters of an example peptide from GPX4 are given in Fig. 4.4

4.4 Conclusions

In mass spectrometry-based proteomics, there has been an evolution from initial pioneering data-dependent approaches where slower instruments could not collect enough precursor intensity data to allow robust label-free quantitation. The massively parallel MS^E and other DIA methodologies have enabled label-free quantitation and

Table 4.1 Example protein quantitation data from WT mouse hippocampus recorded using TWIMS and elucidator post-processing [12]

Protein	Mouse 1 LC/MS run 1	Mouse 1 LC/MS run 2	Mouse 1 LC/MS run 3	Mouse 2 LC/MS run 1	Mouse 2 LC/MS run 2	Mouse 2 LC/MS run 3	Mouse 3 LC/MS run 1	Mouse 3 LC/MS run 2	Mouse 3 LC/MS run 3	Protein abundance (mean)	Protein abundance (SD)	Protein abundance (coefficient of variation)	Number of peptides detected
Amyloid beta A4 protein	4.40	4.27	4.27	4.45	3.87	4.36	3.82	2.28	4.28	4.00	0.68	0.17	29
Isoform short of Abl interactor 1	0.89	0.89	0.92	0.80	0.73	0.77	0.79	0.90	0.78	0.83	0.07	0.08	6
Alpha-internexin	3.45	3.39	3.38	3.42	3.38	3.34	2.70	2.47	2.68	3.14	0.39	0.13	28
A-kinase anchor protein 12	0.85	0.77	0.80	0.85	0.82	0.83	0.92	0.85	0.90	0.84	0.05	0.06	23
Clathrin coat assembly protein AP180	8.25	8.06	7.93	8.36	5.99	7.89	7.41	6.61	7.22	7.53	0.80	0.11	25
Isoform 2 of brain-specific angiogenesis inhibitor 1-associated protein	2.14	2.05	2.00	2.15	2.17	2.12	1.65	1.58	1.69	1.95	0.24	0.12	22
Apoptosis regulator BAX	0.13	0.13	0.12	0.11	0.11	0.11	0.08	0.11	0.09	0.11	0.02	0.14	3
Isoform 2 of Bcl-2-associated transcription factor 1	0.67	0.62	0.62	0.73	0.66	0.71	0.72	0.50	0.77	0.67	0.08	0.12	10
Cdc42 effector protein 4	0.88	0.94	0.97	1.39	1.37	1.31	1.12	0.97	1.11	1.12	0.20	0.18	17
Transcription factor BTF3 homolog 4	0.50	0.49	0.50	0.52	0.52	0.51	0.54	0.57	0.56	0.52	0.03	0.05	3
Cadherin-2	1.82	1.78	1.80	2.03	1.98	1.95	1.60	1.56	1.65	1.80	0.17	0.09	19
Caprin-1	3.05	3.00	2.90	2.91	2.67	2.85	3.02	3.00	3.05	2.94	0.12	0.04	27

(continued)

Table 4.1 (continued)

Protein	Mouse 1 LC/MS run 1	Mouse 1 LC/MS run 2	Mouse 1 LC/MS run 3	Mouse 2 LC/MS run 1	Mouse 2 LC/MS run 2	Mouse 2 LC/MS run 3	Mouse 3 LC/MS run 1	Mouse 3 LC/MS run 2	Mouse 3 LC/MS run 3	Protein abundance (mean)	Protein abundance (SD)	Protein abundance (coefficient of variation)	Number of peptides detected
Chromobox protein homolog 1	0.85	0.82	0.83	0.69	0.71	0.70	0.76	0.86	0.77	0.78	0.07	0.08	7
CD166 antigen	0.85	0.82	0.83	0.70	0.58	0.66	0.76	0.75	0.74	0.74	0.09	0.12	14
Cell cycle exit and neuronal differentiation protein 1	2.62	2.30	2.26	2.79	2.78	2.68	1.76	0.69	1.66	2.17	0.70	0.32	9
Charged multivesicular body protein 4b	2.14	2.20	2.19	2.39	2.37	2.30	2.47	1.65	2.39	2.24	0.25	0.11	15
2′,3′-Cyclic-nucleotide 3′-phosphodiesterase	0.84	0.81	0.82	0.88	0.88	0.86	0.59	0.68	0.60	0.77	0.12	0.15	5
Isoform 2 of cellular nucleic acid-binding protein	1.04	1.10	1.12	0.56	0.56	0.56	1.41	1.25	1.42	1.00	0.36	0.36	5
CREB-regulated transcription coactivator 1	0.65	0.63	0.66	0.75	0.75	0.72	0.58	0.57	0.56	0.65	0.07	0.11	8
Isoform NGC-I of chondroitin sulfate proteoglycan 5	2.53	2.54	2.52	2.89	2.78	2.86	2.84	2.80	2.77	2.73	0.15	0.06	16
Density-regulated protein	0.60	0.61	0.60	0.58	0.57	0.55	0.68	0.70	0.68	0.62	0.05	0.09	4
Isoform 2 of disks large-associated protein 4	0.29	0.30	0.31	0.31	0.33	0.33	0.31	0.28	0.31	0.31	0.02	0.05	9
Band 4.1-like protein 3	0.54	0.52	0.55	0.59	0.63	0.61	0.38	0.30	0.31	0.49	0.13	0.26	4

Excitatory amino acid transporter 2	1.12	1.05	1.02	0.87	0.88	0.84	1.25	0.94	1.31	1.03	0.17	0.16	9
F-box only protein 22	1.12	1.17	1.19	1.28	1.28	1.28	1.36	1.22	1.29	1.24	0.07	0.06	6
Far upstream element-binding protein 2	2.80	2.76	2.75	2.90	2.85	2.76	3.13	2.95	3.26	2.91	0.18	0.06	34
Isoform 2 of FUS-interacting serine–arginine-rich protein 1	0.54	0.52	0.51	0.63	0.59	0.57	0.62	0.58	0.55	0.57	0.04	0.07	6
Guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit beta-1	1.05	1.05	1.03	0.60	0.60	0.63	0.77	0.74	0.73	0.80	0.19	0.24	8
Rho GDP-dissociation inhibitor 1	1.48	1.82	1.82	1.65	1.58	1.58	1.76	1.27	1.71	1.63	0.18	0.11	10
G protein-regulated inducer of neurite outgrowth 1	5.98	6.19	6.15	6.77	6.18	6.70	6.32	4.47	6.43	6.13	0.67	0.11	56
High mobility group protein B1	3.22	3.33	3.28	3.06	3.05	3.08	3.53	2.55	3.58	3.19	0.31	0.10	14
Homer protein homolog 1	3.38	3.28	3.32	3.39	3.43	3.39	3.25	2.52	3.27	3.25	0.28	0.09	31
Protein IMPACT	0.77	0.73	0.73	0.68	0.69	0.66	0.58	0.41	0.65	0.66	0.11	0.16	8
Neural cell adhesion molecule L1	3.98	3.81	3.78	3.32	3.20	3.24	3.82	3.20	3.79	3.57	0.32	0.09	44
Lin-7 homolog A	1.47	1.44	1.41	1.30	1.30	1.32	1.20	0.78	1.18	1.27	0.21	0.16	10
Myelin-associated glycoprotein	3.05	2.85	2.96	2.13	2.11	2.13	2.67	2.63	2.62	2.57	0.37	0.14	21

(continued)

Table 4.1 (continued)

Protein	Mouse 1 LC/MS run 1	Mouse 1 LC/MS run 2	Mouse 1 LC/MS run 3	Mouse 2 LC/MS run 1	Mouse 2 LC/MS run 2	Mouse 2 LC/MS run 3	Mouse 3 LC/MS run 1	Mouse 3 LC/MS run 2	Mouse 3 LC/MS run 3	Protein abundance (mean)	Protein abundance (SD)	Protein abundance (coefficient of variation)	Number of peptides detected
Isoform alpha of methyl-CpG-binding protein 2	2.39	3.64	3.59	4.18	4.04	4.07	3.95	2.51	3.92	3.59	0.67	0.19	15
Neural cell adhesion molecule 1	2.08	2.08	2.02	1.50	1.45	1.46	1.83	1.86	1.85	1.79	0.26	0.14	22
Neuronal calcium sensor 1	1.71	1.65	1.59	1.55	1.54	1.51	1.36	1.52	1.37	1.53	0.11	0.07	14
Neuronal growth regulator 1	4.18	4.26	4.32	3.79	3.58	3.79	4.40	3.94	4.32	4.06	0.29	0.07	20
Na(+)/H(+) exchange regulatory cofactor NHE-RF1	9.70	9.47	9.56	10.27	9.99	10.09	11.03	7.51	11.12	9.86	1.06	0.11	53
Isoform 3 of neuroplastin	4.91	4.80	4.65	3.97	3.84	3.89	4.44	4.16	4.45	4.35	0.40	0.09	13

Note the reproducible results for both biological (mouse-to-mouse) and technical (LC/MS run) replicates. These are just a few example quantitative determinations. Typical experiments yield lists well into 10^3 proteins

Table 4.2 Quantitation of some proteins in fat bodies from cg-dMyc and cg-control *D. melanogaster* larvae

Protein name	UniProt ID	Ratio fed	Ratio starved	*P*-value fed	*P*-value starved
Desat1	Q7K4Y0	1.2	2.6	0.0012	8.9E-43
Pyruvate kinase	KPYK	1.8	2.3	6.4E-07	2.7E-32
Hexokinase C	Q7JYW9	1.4	2.0	0.003	3.8E-28
Glutamine synthetase	GLNA1	1.4	1.4	1.5E-24	2.0E-21

Reprinted from Developmental Biology, Vol. 379, Parisi F, Riccardo S, Zola S, et al.: dMyc expression in the fat body affects DILP2 release and increases the expression of the fat desaturase Desat1 resulting in organismal growth, pp 64–75., 2013, with permission from Elsevier

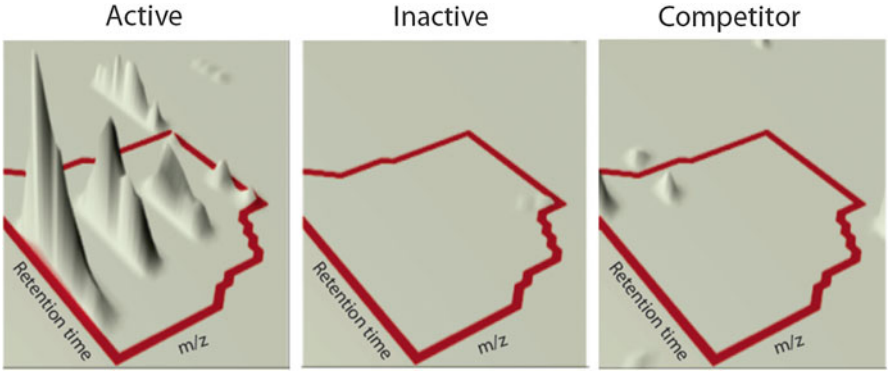


Fig. 4.4 3D visualization of an isotopic cluster of an example peptide ILAFPCNQFGK from GPX4 from an affinity preparation and two controls. Cell lysates were prepared from cells treated with active probe, inactive probe, or active probe in the presence of competitor. Reprinted from Cell Vol. 156, Yang, Wan S.; SriRamaratnam, R.; Welsch, Matthew E.; Shimada, K.; Skouta, R.; Viswanathan, Vasanthi S.; Cheah, Jaime H.; Clemons, Paul A.; Shamji, Alykhan F.; Clish, Clary B.; Brown, Lewis M.; Girotti, Albert W.; Cornish, Virginia W.; Schreiber, Stuart L.; Stockwell, Brent R., Regulation of Ferroptotic Cancer Cell Death by GPX4, pp., 317–331, 2014, with permission from Elsevier

have been applied to a wide variety of sample types from cells to tissues, and have been applied to viral, microbial, plant, animal, and patient samples. This evolution has continued with the introduction of software to match accurate mass and retention time data in large experiments, increasing sophistication of ultrapressure liquid chromatography separations and finally through orthogonal separation with technologies such as TWIMS that significantly increase the number of proteins detected in an experiment. This pipeline provides exceptional flexibility for large and complex experiments and complements isotopic labeling approaches described in other chapters of this volume.

References

1. Stahl DC, Swiderek KM, Davis MT, Lee TD (1996) Data-controlled automation of liquid chromatography/tandem mass spectrometry analysis of peptide mixtures. J Am Soc Mass Spectrom 7(6):532–540
2. Eng JK, McCormack AL, Yates JR III (1994) An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. J Am Soc Mass Spectrom 5(11):976–989
3. Masselon C, Anderson GA, Harkewicz R, Bruce JE, Pasa-Tolic L, Smith RD (2000) Accurate mass multiplexed tandem mass spectrometry for high-throughput polypeptide identification from mixtures. Anal Chem 72(8):1918–1924
4. Purvine S, Eppel J-T, Yi EC, Goodlett DR (2003) Shotgun collision-induced dissociation of peptides using a time of flight mass analyzer. Proteomics 3(6):847–850
5. Silva JC, Denny R, Dorschel C, Gorenstein MV, Li GZ, Richardson K, Wall D, Geromanos SJ (2006) Simultaneous qualitative and quantitative analysis of the Escherichia coli proteome—a sweet tale. Mol Cell Proteomics 5(4):589–607
6. Silva JC, Gorenstein MV, Li GZ, Vissers JPC, Geromanos SJ (2006) Absolute quantification of proteins by LCMSE—a virtue of parallel MS acquisition. Mol Cell Proteomics 5(1): 144–156
7. Panchaud A, Scherl A, Shaffer SA, von Haller PD, Kulasekara HD, Miller SI, Goodlett DR (2009) Precursor acquisition independent from ion count: how to dive deeper into the proteomics ocean. Anal Chem 81(15):6481–6488
8. Gillet LC, Navarro P, Tate S, Rost H, Selevsek N, Reiter L, Bonner R, Aebersold R (2012) Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis. Mol Cell Proteomics 11(6): O111.016717. doi:10.1074/mcp.O111.016717
9. Geiger T, Cox J, Mann M (2010) Proteomics on an Orbitrap benchtop mass spectrometer using all-ion fragmentation. Mol Cell Proteomics 9(10):2252–2261
10. Egertson JD, Kuehn A, Merrihew GE, Bateman NW, MacLean BX, Ting YS, Canterbury JD, Marsh DM, Kellmann M, Zabrouskov V, Wu CC, MacCoss MJ (2013) Multiplexed MS/MS for improved data-independent acquisition. Nat Methods 10(8):744–746
11. Brown LM, Boël G, Gangadhar N, Stockwell BR, Firestein S, Hunt JF (2009) Label-free proteomics with MSE: applications to protein functional biology and the biology of adult stem cells. In: Proceedings of the 57th annual conference on mass spectrometry and allied topics, May 31—June 4, 2009, Philadelphia, PA
12. Brown LM, Rayman JB, Gornstein A, Kandel ER (2011) Quantitative label-free protein profiling of mouse hippocampus. In: Proceedings of the 59th annual conference on mass spectrometry and allied topics, June 5–9, 2011, American Society for Mass Spectrometry, Denver, CO
13. Oswald ES, Brown LM, Bulinski JC, Hung CT (2011) Label-free protein profiling of adipose-derived human stem cells under hyperosmotic treatment. J Proteome Res 10(7):3050–3059
14. Oswald ES, Brown LM, Burke M, Schwartz S, Bulinski JC, Hung CT (2010) Proteomics of adipose-derived human stem cells under hyperosmotic treatment. In: Proceedings of the 58th annual conference on mass spectrometry and allied topics, May 23–27, 2010, Salt Lake City, UT
15. Parisi F, Riccardo S, Zola S, Lora C, Grifoni D, Brown LM, Bellosta P (2013) dMyc expression in the fat body affects DILP2 release and increases the expression of the fat desaturase Desat1 resulting in organismal growth. Dev Biol 379(1):64–75
16. Cyr DD, Lucas JE, Thompson JW, Patel K, Clark PJ, Thompson A, Tillmann HL, McHutchison JG, Moseley MA, McCarthy JJ (2011) Characterization of serum proteins associated with IL28B genotype among patients with chronic hepatitis C. PLoS One 6(7):e21854s
17. Levin Y, Hradetzky E, Bahn S (2011) Quantification of proteins using data-independent analysis (MS(E)) in simple and complex samples: a systematic evaluation. Proteomics 11(16): 3273–3287

18. Reidel B, Thompson JW, Farsiu S, Moseley MA, Skiba NP, Arshavsky VY (2011) Proteomic profiling of a layered tissue reveals unique glycolytic specializations of photoreceptor cells. Mol Cell Proteomics 10(3):M110.002469
19. Brown LM, Tang G, Colligan RM, Sulzer D (2013) Data-independent acquisition with ion mobility (HDMSE) for analysis of the mouse synaptosome proteome. In: Proceedings of the 61st annual conference on mass spectrometry and allied topics, American Society for Mass Spectrometry, Minneapolis, MN
20. Lipton MS, Pasa-Tolic' L, Anderson GA, Anderson DJ, Auberry DL, Battista JR, Daly MJ, Fredrickson J, Hixson KK, Kostandarithes H, Masselon C, Markillie LM, Moore RJ, Romine MF, Shen Y, Stritmatter E, Tolic' N, Udseth HR, Venkateswaran A, Wong K-K, Zhao R, Smith RD (2002) Global analysis of the Deinococcus radiodurans proteome by using accurate mass tags. Proc Natl Acad Sci 99(17):11049–11054
21. Ge H, Wu GS, Moore R, Rao N, Lee T, Gugiu G (2011) Label free quantitative proteomics of formalin fixed embedded (FFPE) tissue sections from temporal giant cell arteritis patients. In: Proceedings of the 59th annual conference on mass spectrometry and allied topics, June 5–9, 2011, American Society for Mass Spectrometry, Denver, CO
22. Ibrahim Y, Danielson WF, Prior D, Baker E, Kurulugama R, Anderson G, Seim T, Meka D, Smith RD, Belov M (2011) Performance of a new sensitive LC-IMS-QTOF platform for proteomics measurements. In: Proceedings of the 59th annual conference on mass spectrometry and allied topics, June 5–9, 2011, American Society for Mass Spectrometry, Denver, CO
23. Thompson JW, Geromanos S, Stapels MD, Dubois LD, Fadgen K, Chepanoske C, Soderblom EJ, Moseley MA (2011) Ion mobility-assisted data independent analysis with inter-analysis alignment provides improved depth of proteome coverage. In: Proceedings of the 59th annual conference on mass spectrometry and allied topics, June 5–9, 2011, American Society for Mass Spectrometry, Denver, CO
24. Yang WS, SriRamaratnam R, Welsch ME, Shimada K, Skouta R, Viswanathan VS, Cheah JH, Clemons PA, Shamji AF, Clish CB, Brown LM, Girotti AW, Cornish VW, Schreiber SL, Stockwell BR (2014) Regulation of ferroptotic cancer cell death by GPX4. Cell 156(1–2): 317–331
25. Valentine SJ, Ewing MA, Dilger JM, Glover MS, Geromanos S, Hughes C, Clemmer DE (2011) Using ion mobility data to improve peptide identification: intrinsic amino acid size parameters. J Proteome Res 10(5):2318–2329
26. Moran D, Cross T, Brown LM, Colligan RM, Dunbar D (2014) Data-independent acquisition (MSE) with ion mobility provides a systematic method for analysis of a bacteriophage structural proteome. J Virol Methods 195:9–17
27. Yang J, Li Y, Chan L, Tsai Y-T, Wu W-H, Nguyen HV, Hsu C-W, Li X, Brown LM, Egli D, Sparrow JR, Tsang SH (2014) Validation of genome-wide association study (GWAS)-identified disease risk alleles with patient-specific stem cell lines. Human Molecular Genetics. first published online February 4, 2014 doi:10.1093/hmg/ddu053
28. Tenzer S, Moro A, Kuharev J, Francis AC, Vidalino L, Provenzani A, Macchi P (2013) Proteome-wide characterization of the RNA-binding protein RALY-interactome using the in vivo-biotinylation-pulldown-quant (iBioPQ) approach. J Proteome Res 12(6):2869–2884
29. Fan Y, Thompson JW, Dubois LG, Moseley MA, Wernegreen JJ (2013) Proteomic analysis of an unculturable bacterial endosymbiont (Blochmannia) reveals high abundance of chaperonins and biosynthetic enzymes. J Proteome Res 12(2):704–718
30. Parviainen VI, Sakari J, Tohmola N, Renkonen R (2013) Label-free mass spectrometry proteome quantification of human embryonic kidney cells following 24 hours of sialic acid overproduction. Proteome Sci

Chapter 5
Stable Isotope Labeling by Amino Acids in Cell Culture (SILAC) for Quantitative Proteomics

Esthelle Hoedt, Guoan Zhang, and Thomas A. Neubert

Abstract Stable isotope labeling by amino acids in cell culture (SILAC) is a powerful approach for high-throughput quantitative proteomics. SILAC allows highly accurate protein quantitation through metabolic encoding of whole cell proteomes using stable isotope labeled amino acids. Since its introduction in 2002, SILAC has become increasingly popular. In this chapter we review the methodology and application of SILAC, with an emphasis on three research areas: dynamics of posttranslational modifications, protein–protein interactions, and protein turnover.

5.1 Introduction

Mass spectrometry (MS)-based quantitative proteomics has become a powerful tool for global functional protein analysis. Various quantitation techniques are now widely used to compare the relative abundances of proteins and posttranslational modifications (PTMs) on a large scale, allowing the addressing of biological questions from a systems biology perspective.

In principle, relative quantitation of a protein can be achieved by comparing its MS signal intensities across different samples. However, the measurement of MS signals can fluctuate, causing run-to-run errors in measurements of protein abundance. More importantly, parallel sample handling prior to MS analysis can

E. Hoedt • G. Zhang • T.A. Neubert (✉)
Kimmel Center for Biology and Medicine at the Skirball Institute and Department of Biochemistry and Molecular Pharmacology, New York University School of Medicine, 540 First Avenue, New York, NY 10016, USA
e-mail: Thomas.Neubert@med.nyu.edu

A.G. Woods and C.C. Darie (eds.), *Advancements of Mass Spectrometry in Biomedical Research*, Advances in Experimental Medicine and Biology 806,
DOI 10.1007/978-3-319-06068-2_5, © Springer International Publishing Switzerland 2014

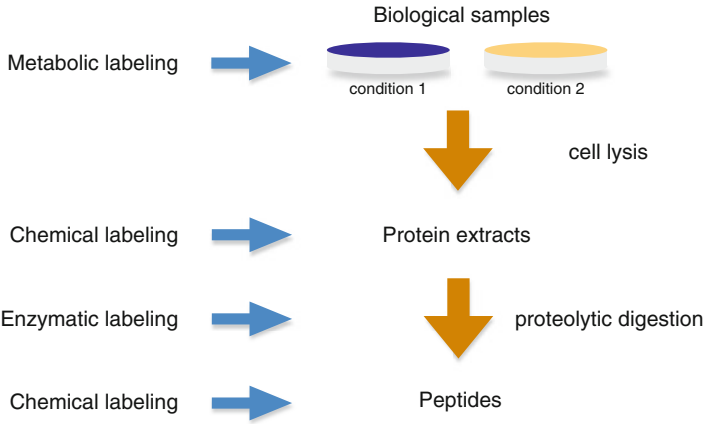


Fig. 5.1 Incorporation of stable isotopes for quantitative proteomics. The incorporation of stable isotopes can be performed directly in cell culture for metabolic labeling. The incorporation of stable isotopes into peptides can also be performed in vitro during proteolytic digestion using enzymatic labeling. Another approach for incorporating stable isotopes consists of modifying proteins or peptides with labeling reagents targeting specific amino acid functionalities by chemical labeling

introduce significant variations. To address these issues, stable isotope labeling approaches were developed.

In these methods, heavy stable isotopes are introduced into proteins/peptides in one or more of the experimental conditions to be compared. Then the heavy labeled sample is mixed with the unlabeled (light) sample and the mixture is analyzed by MS. The heavy label introduces a predictable mass shift compared to the light sample, which can be readily detected by MS. Because the labeled and unlabeled proteins/peptides have the same chemical and physical properties, they have the same signal response factor in the mass spectrometer. Thus the heavy/light signal pairs can be used to measure relative protein abundances. In addition, the sample pairs stay together during sample handling and therefore quantitation accuracy is not compromised by differences in sample preparation.

Stable isotope labeling-based approaches fall into three major categories depending on how heavy isotopes are introduced: (1) chemical labeling (e.g., ICAT, Isotope Coded Affinity Tag [1], iTRAQ, isobaric Tags for Relative and Absolute Quantitation [2] and TMT, Tandem Mass Tagging [3]), in which labels are attached to proteins/peptides through chemical derivatization, (2) enzymatic labeling (e.g., $H_2{}^{16}O/H_2{}^{18}O$) [4] in which labeling is introduced through enzymatic reaction and (3) metabolic labeling (e.g., SILAC, Stable Isotope Labeling by Amino Acids in Cell culture) [5, 6], in which labeling is incorporated into proteins during in vivo protein synthesis (Fig. 5.1). In this chapter, we focus on the methodology and applications of SILAC.

The metabolic labeling approach was first adopted for proteomics by Oda et al., who labeled bacterial proteins using nitrogen (^{15}N) [7]. The mass shift caused by ^{15}N labeling is dependent on the number of nitrogens present and thus is not constant for all peptides. This complicates peptide identification and quantification [8]. Later, Matthias Mann

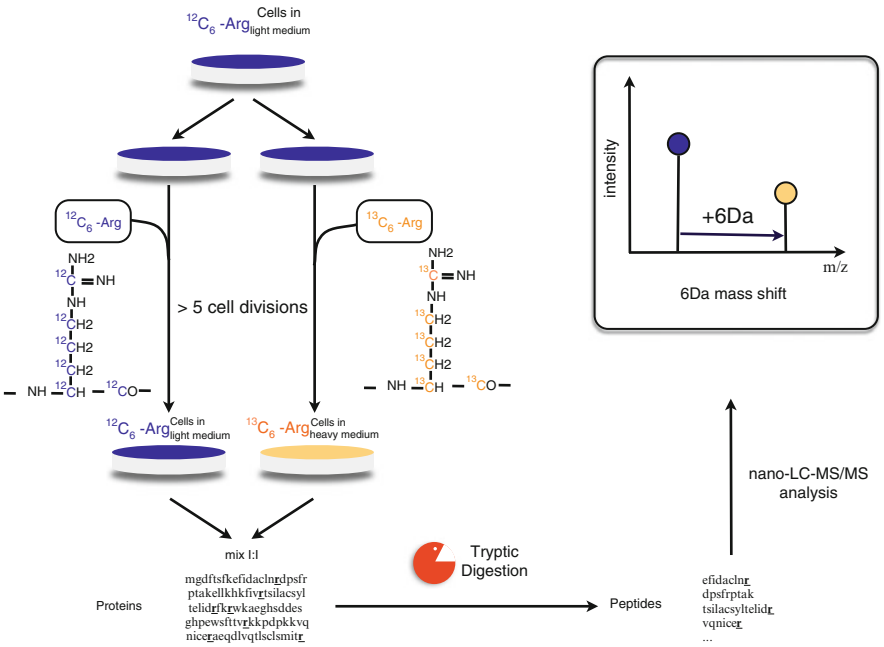


Fig. 5.2 Protein quantification using SILAC. Cells are differentially labeled in medium with normal arginine or medium with heavy arginine. After at least five cell divisions to ensure virtually complete labeling, the two cell populations are combined, digested with trypsin, and analyzed by nano LC-MS/MS. The tryptic cleavage creates pairs of peptides differing by 6 Da due to the molecular weight difference between $^{12}C_6$-arginine and $^{13}C_6$-arginine

and colleagues developed the SILAC approach for the study of eukaryote cells in which they labeled cells by culturing them in a medium supplemented with heavy isotope labeled amino acids [5].

Compared to chemical and enzymatic labeling, SILAC has an important advantage: it generally allows combining the light and heavy labeled samples early in the experimental workflow so that variability caused by parallel sample handling is minimized.

In a typical SILAC experiment, cells are labeled by growing in either medium with natural amino acids (light medium) or in medium containing one or more stable isotope labeled amino acids (heavy medium). In general, cells need to be maintained in SILAC media for at least five divisions to ensure virtually complete labeling of proteins. Then if needed, differential experimental treatments are performed. Equal amounts of the two differentially labeled cell populations or protein extracts are combined prior to further sample handling and processing steps as required by the experiment (Fig. 5.2). Because trypsin is the most commonly used protease in proteomics and it cleaves carboxy-terminal to lysine and arginine residues, double labeling with lysine and arginine ensures that every tryptic peptide (except C-terminal peptides) contains a heavy amino acid and can be used for protein quantitation (Fig. 5.2).

5.2 Applications and Examples of SILAC

5.2.1 Dynamics of Posttranslational Modifications

PTMs of proteins are known to play important roles in cell signaling. PTMs directly affect protein structure, protein localization, activity, and interaction. Many PTMs function as essential regulatory switches for various crucial signaling pathways. Therefore, characterization of PTM dynamics can generate rich information that is important for deciphering signal transduction mechanisms. The coupling of nano-flow HPLC to tandem MS has made it possible to profile PTM changes for whole proteomes.

Protein phosphorylation, catalyzed by kinases and reversed by phosphatases, is a major regulatory mechanism of cellular pathways, especially those involved in eukaryotic signal transduction. It is estimated to be the most abundant PTM, with about 3 % of human genes encoding proteins with kinase or phosphatase activities [9]. Relative changes in phosphorylation of proteins have traditionally been studied by methods such as phosphoamino acid analysis and use of ^{32}P-labeled ATP coupled to two-dimensional gel electrophoresis. However, these methods do not provide quantitative information about individual phosphorylation sites, which is especially important because different phosphorylation sites on the same protein can be differentially regulated. SILAC combined with MS allows for accurate, site-specific quantitation of phosphorylation. The quantitation can be performed for a single protein purified from cell lysates using immunoprecipitation [10–14]. But the real strength of SILAC in phosphorylation analysis is its ability to profile the whole phosphoproteome to obtain accurate site-specific changes. Recent advances in LC-MS and phosphopeptide enrichment techniques have enabled quantitation of >20,000 phosphosites from one SILAC experiment, allowing the description of serine/threonine/tyrosine-phosphorylation dependent global landscape of cellular signaling at the network level [15]. For large scale phosphoproteomic analyses, enrichment of phosphorylated peptides is indispensable and several such techniques have been developed and optimized in recent years [16–18]. The most popular techniques are affinity-based phosphopeptide enrichment technologies such as IMAC (Immobilized Metal Affinity Chromatography) [19, 20] and metal oxide chromatography including titanium dioxide (TiO_2) [15, 21]. Combined with peptide fractionation techniques such as Hydrophilic Interaction Liquid Chromatography (HILIC) and Strong Cation eXchange (SCX), phosphopeptide enrichment by these methods can enable large numbers of phosphosite identifications. However, coverage of tyrosine phosphorylation sites is often limited because of low occurrence of tyrosine phosphorylation compared to serine/threonine phosphorylation. To improve the coverage of the tyrosine phosphoproteome, tyrosine phosphorylated peptides can be immunoprecipitated using anti-phosphotyrosine antibodies [22, 23]. Gruhler et al. used SILAC with IMAC for phosphopeptide enrichment to analyze and quantify changes in protein phosphorylation of the G-protein-coupled receptor

in the yeast pheromone response [24]. This technique has been used efficiently in several studies [25–28]. The TiO_2 phosphopeptide enrichment method has been used to study the phosphoproteome of Drosophila Schneider cells following a knock down of a specific phosphatase by RNA interference [29], to analyze the bloodstream and procyclic forms of *Trypanosoma brucei* [30], to compare cellular phosphorylation levels upon epidermal growth factor (EGF) stimulation, and growth factor combined with kinase inhibitors [31] and in technical comparison with IMAC [23, 32].

An alternative strategy for analyzing phosphotyrosine signaling pathways is to combine SILAC with anti-phosphotyrosine immunoprecipitation of intact tyrosine phosphorylated proteins. Proteins that are differentially tyrosine phosphorylated during signal transduction, or that differentially associate with these proteins, will appear with changed SILAC ratios. This strategy has been successfully used for the analysis of Receptor Tyrosine Kinases (RTK) signaling networks, including the EGF receptor in HeLa cells [33] and the specific Her2/neu Receptor belonging to the EGFR family in NIH3T3 cells [34]. Other RTK signaling pathways have been studied using this method, such as the FGF receptor [35, 36], the PDGF receptor [37], the EphB2-receptor in mouse neuroblastoma x rat glioma hybrid [38], the insulin receptor in differentiated brown adipocytes [39], the TrkB in rat hippocampal primary cells [40], the Met receptor through HGF and EGF activation in A549 lung carcinoma cells [41], the Syk receptor in MCF-7 and MDA-MB-231 human breast cancer cell lines, IL-2 receptor in human T cell line kit 225 [42] and the T cell receptor in Jurkat T cells [43].

Because intracellular signal transduction is regulated temporally and spatially, characterization of temporal and spatial dynamics of phosphorylation can provide particularly useful data to help understand how signals are propagated in these dimensions. In this regard, SILAC has been used to globally analyze phosphosites after receptor stimulation, where the magnitude of change in phosphorylation, the timing of phosphorylation or the subcellular localization were analyzed within the same experiment [44–50].

In this review, our focus is on phosphorylation proteomics. However, other PTMs including acetylation [51–53], ubiquitination [54–59], methylation [60–64] and glycosylation [65–69] can be efficiently enriched, therefore amenable to large-scale analyses. Some studies even investigate several PTMs in the same analysis, yielding direct data on their crosstalk at a global level [13, 70–72].

5.2.2 Protein/Protein Interaction

In cells, proteins rarely function in isolation but instead interact with specific proteins to form complexes in order to perform particular cellular activities. Thus, deciphering the protein interaction network that dictates protein cellular function is essential for better understanding of these functions. Traditionally people have

used protein chips and two-hybrid systems for comprehensive protein–protein interaction studies. However, these techniques suffer from high false positive and false negative rates, because the assay is usually performed under non-physiological conditions and the subcellular localization and the posttranslational dynamics are not taken into account.

Advances in mass spectrometry for protein identification now allow identification of large numbers of protein partners from immunoprecipitated protein complexes. However, a major drawback of this type of assay is that sample-to-sample variations from immunopurification and downstream sample preparation can lead to high false positive rates for protein associations. Combining SILAC with affinity purification of protein complexes using various experimental setups can help to unambiguously distinguish specific partners from non-specific binding partners. Chait and colleagues have developed a method to distinguish contaminants from specific protein partners. In this method, cells containing an affinity tagged protein are grown in light isotopic medium while wild-type cells are grown in heavy isotopic medium. After mixing equal quantities of these two populations of cells, an immunoprecipitation is performed against the affinity tag. In MS, specific partners appear as isotopically light, while non-specific interaction partners appear as a mixture of isotopically light and heavy at a 1:1 ratio [73]. This strategy has been employed successfully for the analysis of growth factor signaling [74], Glut4 in the insulin signaling pathway [75, 76], integrin-linked kinase [77], and toll-like receptors [78]. SILAC has also been used to define tagged protein complexes in yeast [73, 79–83].

In addition, SILAC can be combined with RNA interference, allowing detection of protein–protein interactions at their endogenous levels. After knocking down the protein of interest by RNA interference in one of the two differentially labeled conditions, the target protein and its partners should be more abundant in the untreated cells than those in the knockdown cells, while contaminating proteins should be present in similar amounts in both untreated and knockdown cells. It has been first used for the identification of the partners of β-catenin and cbl [84]. This method has been successfully used to identify the 14-3-3ζ interacting proteins [85], the leucine-rich repeat kinase 2 interaction partners [86], to study the molecular functions of the ATP-dependent chromatin remodeling complex SWI/SNF in cell cycle control [87] and to investigate the role of Stat3 in multiple myeloma pathology [88].

SILAC can be used to identify the components of inducible protein complexes that are formed upon activation of signaling pathways. Indeed, Blagoev and coworkers have used SILAC for investigating the epidermal growth factor receptor (EGFR) pathway at an endogenous level, by using affinity purification of proteins that associates with the Src homology 2 (SH2) domain of the signaling adaptor protein Grb2 to identify components of the signaling complex induced by EGF stimulation of HeLa cells [89].

5.2.3 *Protein Synthesis and Turnover*

Protein degradation has a central role in a large number of cellular processes, including signal transduction, cell cycle regulation, and apoptosis. Misregulated protein degradation has been implicated in human diseases. Thus, global measurement of protein degradation rates is necessary to understand the principles that govern these processes. Cellular protein levels are determined by the balance of translation of new proteins and degradation of preexisting ones. In the past, two main approaches were used for the measurement of protein turnover. One approach relies on the incorporation of radiolabeled amino acids. This method, called pulse-chase, uses radioactive ^{35}S-methionine to label the protein of interest for a brief period of time ("the pulse"), and then the decay in radioactivity is measured over time ("the chase"). In the second approach, protein degradation rates are measured by quantitative Western blotting following translation inhibition by using protein synthesis inhibitors like cycloheximide. However, this technique is less suitable for measuring the half-life of long-lived proteins since the inhibitor could lead to severe cell stress [90].

SILAC is an efficient way to measure protein synthesis and degradation rates in response to stimuli [91–93]. When cells are switched from light culture medium to heavy medium, heavy amino acids are incorporated into newly synthesized proteins resulting in a gradual increase of heavy labeled peptide signal in MS, whereas the degradation of preexisting proteins is associated with the gradual loss of light peptide signal [94]. This gradual change of relative abundance of light and heavy proteins can be easily monitored by MS. Assessing the ratio between corresponding labeled and unlabeled peptides over time can then indicate the rate of protein turnover. Conversely, cells can first be labeled completely with heavy amino acids and then chased in light medium. Compared to traditional methods, SILAC cell culture is performed without radioactivity or protein synthesis inhibitors. It also allows measurement of large numbers of proteins within one experiment.

The first application of metabolic labeling for the analysis of protein turnover on a protein-by-protein basis was performed in *Saccharomyces cerevisiae* [95]. A similar approach has been applied to the bacterium *Escherichia coli* [96], human HeLa cells [97–99], human A549 lung carcinoma cells [92, 100], human stromal stem cells [101], mouse NIH3T3 and C2C12 cells [98, 102] and chicken [91]. Selbach and coworkers have compared protein translation rates between two samples by pulsed SILAC labeling with two different sets of stable isotopes, medium and heavy [102, 103]. Jayapal et al. used a multitagging proteomic approach by mixing iTRAQ and SILAC labeling to estimate protein turnover rate constants in *Streptomyces coelicolor*, a multicellular differentiated bacterium [104]. Recently, Ziv and colleagues have measured and analyzed the metabolic half-lives of synaptic proteins from rat cortical neurons [94].

As first described by Lamond and coworkers, cellular localization was taken into account to provide a high-throughput experiment for the unbiased analysis of changes in subcellular protein localization and responses to *stimuli*. Whereas the average protein turnover rate for HeLa cells is approximately 20 h, they found that synaptic proteins exhibit half-life times of 2–5 days [94, 99].

5.2.4 *Prospective*

Since its introduction in 2002, SILAC has become increasingly popular as an accurate relative quantitation technique for proteomics. During the last decade, this technique has been further refined and adapted to different types of applications. The efforts to improve SILAC have been mainly focused on: (1) quantitation of samples that are difficult to label and (2) to increase multiplexing.

Conventional SILAC requires that the cells must be completely labeled. This is difficult for postmitotic cells such as primary cells because they do not divide in culture. To circumvent this issue, it has been shown that by using two different sets of heavy amino acids for labeling, straightforward SILAC quantitation can be performed using partially labeled cells because the two cell populations are always equally labeled [102, 105]. Approaches have been developed to metabolically label model organisms like worms [106, 107], flies [108, 109], mice [110], and rats [111, 112] by long-term administration of heavy amino acid-enriched diets. For analysis of samples that cannot be SILAC labeled (such as clinical samples), Oda and colleagues have developed a strategy in which a SILAC labeled lysate is spiked into samples to be compared as an internal standard [113]. Mann and colleagues further improved this method by using a mixture of multiple SILAC labeled cell lines as the internal standard to more accurately represent the protein expression profiles of human tissues [114]. This approach has been used to classify cell lines derived from patients with various lymphoma subtypes [115] and adapted for global phosphosite quantification in tissues [116].

There has been a growing demand for higher multiplex SILAC to allow the comparison of more samples within one analysis to increase the throughput and quantification accuracy. But in conventional SILAC, to avoid MS signal overlap, the choices of labeling are limited. A single SILAC experiment can compare up to five samples at a time [117]. Recently, Coon and coworkers have described a new approach, called neutron encoding (NeuCode), which exploits the subtle (~6 mDa) mass defects in common stable isotopes to expand SILAC multiplexing without the increase in spectral complexity that accompanies traditional SILAC approaches [118]. This fresh SILAC variant offers the potential to achieve a quantitative capacity on par with isobaric labeling, while maintaining the measurement accuracy of full MS scan-based methods.

References

1. Gygi SP, Rist B, Gerber SA, Turecek F, Gelb MH, Aebersold R (1999) Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. Nat Biotechnol 17:994–999
2. Ross PL (2004) Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. Mol Cell Proteomics 3:1154–1169
3. Dayon L, Hainard A, Licker V, Turck N, Kuhn K, Hochstrasser DF, Burkhard PR, Sanchez J (2008) Relative quantification of proteins in human cerebrospinal fluids by MS/MS using 6-plex isobaric tags. Anal Chem 80:2921–2931

4. Stewart II, Thomson T, Figeys D (2001) 18O labeling: a tool for proteomics. Rapid Commun Mass Spectrom 15:2456–2465
5. Ong S, Blagoev B, Kratchmarova I, Kristensen DB, Steen H, Pandey A, Mann M (2002) Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. Mol Cell Proteomics 1:376–386
6. Ong S, Foster LJ, Mann M (2003) Mass spectrometric-based approaches in quantitative proteomics. Methods 29:124–130
7. Oda Y, Huang K, Cross FR, Cowburn D, Chait BT (1999) Accurate quantitation of protein expression and site-specific phosphorylation. Proc Natl Acad Sci U S A 96:6591–6596
8. Bindschedler LV, Cramer R (2011) Fully automated software solution for protein quantitation by global metabolic labeling with stable isotopes. Rapid Commun Mass Spectrom 25: 1461–1471
9. Manning G, Plowman GD, Hunter T, Sudarsanam S (2002) Evolution of protein kinase signaling from yeast to man. Trends Biochem Sci 27:514–520
10. Ibarrola N, Kalume DE, Gronborg M, Iwahori A, Pandey A (2003) A proteomic approach for quantitation of phosphorylation using stable isotope labeling in cell culture. Anal Chem 75:6043–6049
11. Liang X, Hajivandi M, Veach D, Wisniewski D, Clarkson B, Resh MD, Pope RM (2006) Quantification of change in phosphorylation of BCR-ABL kinase and its substrates in response to Imatinib treatment in human chronic myelogenous leukemia cells. Proteomics 6:4554–4564
12. Park K, Mohapatra DP, Misonou H, Trimmer JS (2006) Graded regulation of the Kv2.1 potassium channel by variable phosphorylation. Science 313:976–979
13. Wisniewski JR, Zougman A, Krüger S, Ziółkowski P, Pudełko M, Bebenek M, Mann M (2008) Constitutive and dynamic phosphorylation and acetylation sites on NUCKS, a hypermodified nuclear protein, studied by quantitative proteomics. Proteins 73:710–718
14. Lu X, Hamrahi VF, Tompkins RG, Fischman AJ (2009) Effect of insulin levels on the phosphorylation of specific amino acid residues in IRS-1: implications for burn-induced insulin resistance. Int J Mol Med 24:531–538
15. Olsen JV, Blagoev B, Gnad F, Macek B, Kumar C, Mortensen P, Mann M (2006) Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. Cell 127:635–648
16. Rogers LD, Foster LJ (2009) Phosphoproteomics—finally fulfilling the promise? Mol Biosyst 5:1122–1129
17. Nilsson CL (2012) Advances in quantitative phosphoproteomics. Anal Chem 84:735–746
18. Bodenmiller B, Mueller LN, Mueller M, Domon B, Aebersold R (2007) Reproducible isolation of distinct, overlapping segments of the phosphoproteome. Nat Methods 4:231–237
19. Stensballe A, Andersen S, Jensen ON (2001) Characterization of phosphoproteins from electrophoretic gels by nanoscale Fe(III) affinity chromatography with off-line mass spectrometry analysis. Proteomics 1:207–222
20. Ficarro S, Chertihin O, Westbrook VA, White F, Jayes F, Kalab P, Marto JA, Shabanowitz J, Herr JC, Hunt DF, Visconti PE (2003) Phosphoproteome analysis of capacitated human sperm. Evidence of tyrosine phosphorylation of a kinase-anchoring protein 3 and valosin-containing protein/p97 during capacitation. J Biol Chem 278:11579–11589
21. Larsen MR, Thingholm TE, Jensen ON, Roepstorff P, Jørgensen TJD (2005) Highly selective enrichment of phosphorylated peptides from peptide mixtures using titanium dioxide microcolumns. Mol Cell Proteomics 4:873–886
22. Rush J, Moritz A, Lee KA, Guo A, Goss VL, Spek EJ, Zhang H, Zha X, Polakiewicz RD, Comb MJ (2005) Immunoaffinity profiling of tyrosine phosphorylation in cancer cells. Nat Biotechnol 23:94–101
23. Zhang G, Neubert TA (2011) Comparison of three quantitative phosphoproteomic strategies to study receptor tyrosine kinase signaling. J Proteome Res 10:5454–5462
24. Gruhler A, Olsen JV, Mohammed S, Mortensen P, Faergeman NJ, Mann M, Jensen ON (2005) Quantitative phosphoproteomics applied to the yeast pheromone signaling pathway. Mol Cell Proteomics 4:310–327

25. Platt MD, Salicioni AM, Hunt DF, Visconti PE (2009) Use of differential isotopic labeling and mass spectrometry to analyze capacitation-associated changes in the phosphorylation status of mouse sperm proteins. J Proteome Res 8:1431–1440
26. Schreiber TB, Mäusbacher N, Soroka J, Wandinger SK, Buchner J, Daub H (2012) Global analysis of phosphoproteome regulation by the Ser/Thr phosphatase Ppt1 in Saccharomyces cerevisiae. J Proteome Res 11:2397–2408
27. Chen C, Wu D, Zhang L, Zhao Y, Guo L (2012) Comparative phosphoproteomics studies of macrophage response to bacterial virulence effectors. J Proteomics 77:251–261
28. Xiao K, Sun J, Kim J, Rajagopal S, Zhai B, Villén J, Haas W, Kovacs JJ, Shukla AK, Hara MR, Hernandez M, Lachmann A, Zhao S, Lin Y, Cheng Y, Mizuno K, Ma'ayan A, Gygi SP, Lefkowitz RJ (2010) Global phosphorylation analysis of beta-arrestin-mediated signaling downstream of a seven transmembrane receptor (7TMR). Proc Natl Acad Sci 107: 15299–15304
29. Hilger M, Bonaldi T, Gnad F, Mann M (2009) Systems-wide analysis of a phosphatase knockdown by quantitative proteomics and phosphoproteomics. Mol Cell Proteomics 8:1908–1920
30. Urbaniak MD, Martin DMA, Ferguson MAJ (2013) Global quantitative SILAC phosphoproteomics reveals differential phosphorylation is widespread between the procyclic and bloodstream form lifecycle stages of Trypanosoma brucei. J Proteome Res 12:2233–2244
31. Pan C, Olsen JV, Daub H, Mann M (2009) Global effects of kinase inhibitors on signaling networks revealed by quantitative phosphoproteomics. Mol Cell Proteomics 8:2796–2808
32. Liang X, Fonnum G, Hajivandi M, Stene T, Kjus NH, Ragnhildstveit E, Amshey JW, Predki P, Pope RM (2007) Quantitative comparison of IMAC and TiO2 surfaces used in the study of regulated, dynamic protein phosphorylation. J Am Soc Mass Spectrom 18:1932–1944
33. Blagoev B, Ong S, Kratchmarova I, Mann M (2004) Temporal analysis of phosphotyrosine-dependent signaling networks by quantitative proteomics. Nat Biotechnol 22:1139–1145
34. Bose R, Molina H, Patterson AS, Bitok JK, Periaswamy B, Bader JS, Pandey A, Cole PA (2006) Phosphoproteomic analysis of Her2/neu signaling and inhibition. Proc Natl Acad Sci U S A 103:9773–9778
35. Hinsby AM, Olsen JV, Mann M (2004) Tyrosine phosphoproteomics of fibroblast growth factor signaling: a role for insulin receptor substrate-4. J Biol Chem 279:46438–46447
36. Cunningham DL, Sweet SMM, Cooper HJ, Heath JK (2010) Differential phosphoproteomics of fibroblast growth factor signaling: identification of Src family kinase-mediated phosphorylation events. J Proteome Res 9:2317–2328
37. Kratchmarova I, Blagoev B, Haack-Sorensen M, Kassem M, Mann M (2005) Mechanism of divergent growth factor effects in mesenchymal stem cell differentiation. Science 308: 1472–1477
38. Zhang G, Spellman DS, Skolnik EY, Neubert TA (2006) Quantitative phosphotyrosine proteomics of EphB2 signaling by stable isotope labeling with amino acids in cell culture (SILAC). J Proteome Res 5:581–588
39. Krüger M, Kratchmarova I, Blagoev B, Tseng Y, Kahn CR, Mann M (2008) Dissection of the insulin signaling pathway via quantitative phosphoproteomics. Proc Natl Acad Sci 105:2451–2456
40. Spellman DS, Deinhardt K, Darie CC, Chao MV, Neubert TA (2008) Stable isotopic labeling by amino acids in cultured primary neurons: application to brain-derived neurotrophic factor-dependent phosphotyrosine-associated signaling. Mol Cell Proteomics 7:1067–1076
41. Hammond DE, Hyde R, Kratchmarova I, Beynon RJ, Blagoev B, Clague MJ (2010) Quantitative analysis of HGF and EGF-dependent phosphotyrosine signaling networks. J Proteome Res 9:2734–2742
42. Osinalde N, Moss H, Arrizabalaga O, Omaetxebarria MJ, Blagoev B, Zubiaga AM, Fullaondo A, Arizmendi JM, Kratchmarova I (2011) Interleukin-2 signaling pathway analysis by quantitative phosphoproteomics. J Proteomics 75:177–191
43. Størvold GL, Landskron J, Strozynski M, Arntzen MØ, Koehler CJ, Kalland ME, Taskén K, Thiede B (2013) Quantitative profiling of tyrosine phosphorylation revealed changes in

the activity of the T cell receptor signaling pathway upon cisplatin-induced apoptosis. J Proteomics 91:344–357
44. Zhang Y, Wolf-Yadlin A, Ross PL, Pappin DJ, Rush J, Lauffenburger DA, White FM (2005) Time-resolved mass spectrometry of tyrosine phosphorylation sites in the epidermal growth factor receptor signaling network reveals dynamic modules. Mol Cell Proteomics 4:1240–1250
45. Zhang L, Yu C, Vasquez FE, Galeva N, Onyango I, Swerdlow RH, Dobrowsky RT (2010) Hyperglycemia alters the schwann cell mitochondrial proteome and decreases coupled respiration in the absence of superoxide production. J Proteome Res 9:458–471
46. Matsumura T, Oyama M, Kozuka-Hata H, Ishikawa K, Inoue T, Muta T, Semba K, Inoue J (2010) Identification of BCAP-(L) as a negative regulator of the TLR signaling-induced production of IL-6 and IL-10 in macrophages by tyrosine phosphoproteomics. Biochem Biophys Res Commun 400:265–270
47. Brockmeyer C, Paster W, Pepper D, Tan CP, Trudgian DC, McGowan S, Fu G, Gascoigne NRJ, Acuto O, Salek M (2011) T cell receptor (TCR)-induced tyrosine phosphorylation dynamics identifies THEMIS as a new TCR signalosome component. J Biol Chem 286: 7535–7547
48. Azimifar SB, Böttcher RT, Zanivan S, Grashoff C, Krüger M, Legate KR, Mann M, Fässler R (2012) Induction of membrane circular dorsal ruffles requires co-signalling of integrin-ILK-complex and EGF receptor. J Cell Sci 125:435–448
49. Mäusbacher N, Schreiber TB, Machatti M, Schaab C, Daub H (2012) Proteome-wide analysis of temporal phosphorylation dynamics in lysophosphatidic acid-induced signaling. Proteomics 12:3485–3498
50. Pan X, Whitten DA, Wu M, Chan C, Wilkerson CG, Pestka JJ (2013) Global protein phosphorylation dynamics during deoxynivalenol-induced ribotoxic stress response in the macrophage. Toxicol Appl Pharmacol 268:201–211
51. Zhou Q, Chaerkady R, Shaw PG, Kensler TW, Pandey A, Davidson NE (2010) Screening for therapeutic targets of vorinostat by SILAC-based proteomic analysis in human breast cancer cells. Proteomics 10:1029–1039
52. Bennetzen MV, Larsen DH, Dinant C, Watanabe S, Bartek J, Lukas J, Andersen JS (2013) Acetylation dynamics of human nuclear proteins during the ionizing radiation-induced DNA damage response. Cell Cycle 12:1688–1695
53. Wu Q, Xu W, Cao L, Li X, He T, Wu Z, Li W (2013) SAHA treatment reveals the link between histone lysine acetylation and proteome in nonsmall cell lung cancer A549 Cells. J Proteome Res 12:4064–4073
54. Meierhofer D, Wang X, Huang L, Kaiser P (2008) Quantitative analysis of global ubiquitination in HeLa cells by mass spectrometry. J Proteome Res 7:4566–4576
55. Akimov V, Rigbolt KTG, Nielsen MM, Blagoev B (2011) Characterization of ubiquitination dependent dynamics in growth factor receptor signaling by quantitative proteomics. Mol Biosyst 7:3223–3233
56. Na CH, Peng J (2012) Analysis of ubiquitinated proteome by quantitative mass spectrometry. Methods Mol Biol 893:417–429
57. Udeshi ND, Mertins P, Svinkina T, Carr SA (2013) Large-scale identification of ubiquitination sites by mass spectrometry. Nat Protoc 8:1950–1960
58. Anania VG, Pham VC, Huang X, Masselot A, Lill JR, Kirkpatrick DS (2014) Peptide level immunoaffinity enrichment enhances ubiquitination site identification on individual proteins. Mol Cell Proteomics 13(1):145–156
59. Udeshi ND, Mani DR, Eisenhaure T, Mertins P, Jaffe JD, Clauser KR, Hacohen N, Carr SA (2012) Methods for quantification of in vivo changes in protein ubiquitination following proteasome and deubiquitinase inhibition. Mol Cell Proteomics 11:148–159
60. Ong S, Mittler G, Mann M (2004) Identifying and quantifying in vivo methylation sites by heavy methyl SILAC. Nat Methods 1:119–126
61. Ong S, Mann M (2006) A practical recipe for stable isotope labeling by amino acids in cell culture (SILAC). Nat Protoc 1:2650–2660

62. Zee BM, Levin RS, Xu B, LeRoy G, Wingreen NS, Garcia BA (2010) In vivo residue-specific histone methylation dynamics. J Biol Chem 285:3341–3350
63. Bartke T, Vermeulen M, Xhemalce B, Robson SC, Mann M, Kouzarides T (2010) Nucleosome-interacting proteins regulated by DNA and histone methylation. Cell 143: 470–484
64. Cao X, Zee BM, Garcia BA (2013) Heavy methyl-SILAC labeling coupled with liquid chromatography and high-resolution mass spectrometry to study the dynamics of site-specific histone methylation. Methods Mol Biol 977:299–313
65. Wang Z, Pandey A, Hart GW (2007) Dynamic interplay between O-linked N-acetylglucosaminylation and glycogen synthase kinase-3-dependent phosphorylation. Mol Cell Proteomics 6:1365–1379
66. Ostasiewicz P, Zielinska DF, Mann M, Wisniewski JR (2010) Proteome, phosphoproteome, and N-glycoproteome are quantitatively preserved in formalin-fixed paraffin-embedded tissue and analyzable by high-resolution mass spectrometry. J Proteome Res 9:3688–3700
67. Palmisano G, Lendal SE, Larsen MR (2011) Titanium dioxide enrichment of sialic acid-containing glycopeptides. Methods Mol Biol 753:309–322
68. Boersema PJ, Geiger T, Wisniewski JR, Mann M (2013) Quantification of the N-glycosylated secretome by super-SILAC during breast cancer progression and in human blood samples. Mol Cell Proteomics 12:158–171
69. Taga Y, Kusubata M, Ogawa-Goto K, Hattori S (2013) Site-specific quantitative analysis of overglycosylation of collagen in osteogenesis imperfecta using hydrazide chemistry and SILAC. J Proteome Res 12:2225–2232
70. Bonenfant D, Towbin H, Coulot M, Schindler P, Mueller DR, van Oostrum J (2007) Analysis of dynamic changes in post-translational modifications of human histones during cell cycle by mass spectrometry. Mol Cell Proteomics 6:1917–1932
71. Cuomo A, Moretti S, Minucci S, Bonaldi T (2011) SILAC-based proteomic analysis to dissect the "histone modification signature" of human breast cancer cells. Amino Acids 41:387–399
72. Guan X, Rastogi N, Parthun MR, Freitas MA (2013) Discovery of histone modification crosstalk networks by stable isotope labeling of amino acids in cell culture mass spectrometry (SILAC MS). Mol Cell Proteomics 12:2048–2059
73. Tackett AJ, DeGrasse JA, Sekedat MD, Oeffinger M, Rout MP, Chait BT (2005) I-DIRT, a general method for distinguishing between specific and nonspecific protein interactions. J Proteome Res 4:1752–1756
74. Zhong J, Chaerkady R, Kandasamy K, Gucek M, Cole RN, Pandey A (2011) The interactome of a PTB domain-containing adapter protein, Odin, revealed by SILAC. J Proteomics 74: 294–303
75. Foster LJ, Rudich A, Talior I, Patel N, Huang X, Furtado LM, Bilan PJ, Mann M, Klip A (2006) Insulin-dependent interactions of proteins with GLUT4 revealed through stable isotope labeling by amino acids in cell culture (SILAC). J Proteome Res 5:64–75
76. Hanke S, Mann M (2009) The phosphotyrosine interactome of the insulin receptor family and its substrates IRS-1 and IRS-2. Mol Cell Proteomics 8:519–534
77. Dobreva I, Fielding A, Foster LJ, Dedhar S (2008) Mapping the integrin-linked kinase interactome using SILAC. J Proteome Res 7:1740–1749
78. Sharma K, Kumar C, Kéri G, Breitkopf SB, Oppermann FS, Daub H (2010) Quantitative analysis of kinase-proximal signaling in lipopolysaccharide-induced innate immune response. J Proteome Res 9:2539–2549
79. Ranish JA, Yi EC, Leslie DM, Purvine SO, Goodlett DR, Eng J, Aebersold R (2003) The study of macromolecular complexes by quantitative proteomics. Nat Genet 33:349–355
80. Kito K, Kawaguchi N, Okada S, Ito T (2008) Discrimination between stable and dynamic components of protein complexes by means of quantitative proteomics. Proteomics 8: 2366–2370
81. Synowsky SA, van Wijk M, Raijmakers R, Heck AJR (2009) Comparative multiplexed mass spectrometric analyses of endogenously expressed yeast nuclear and cytoplasmic exosomes. J Mol Biol 385:1300–1313

82. Chao JT, Foster LJ, Loewen CJR (2009) Identification of protein complexes with quantitative proteomics in S. cerevisiae. J Vis Exp (25). pii: 1225
83. Bard-Chapeau EA, Gunaratne J, Kumar P, Chua BQ, Muller J, Bard FA, Blackstock W, Copeland NG, Jenkins NA (2013) EVI1 oncoprotein interacts with a large and complex network of proteins and integrates signals through protein phosphorylation. Proc Natl Acad Sci 110:E2885–E2894
84. Selbach M, Mann M (2006) Protein interaction screening by quantitative immunoprecipitation combined with knockdown (QUICK). Nat Methods 3:981–983
85. Ge F, Li W, Bi L, Tao S, Zhang Z, Zhang X (2010) Identification of novel 14-3-3ζ interacting proteins by quantitative immunoprecipitation combined with knockdown (QUICK). J Proteome Res 9:5848–5858
86. Meixner A, Boldt K, Van Troys M, Askenazi M, Gloeckner CJ, Bauer M, Marto JA, Ampe C, Kinkl N, Ueffing M (2011) A QUICK screen for Lrrk2 interaction partners—leucine-rich repeat kinase 2 is involved in actin cytoskeleton dynamics. Mol Cell Proteomics 10: M110.001172
87. Hah N, Kolkman A, Ruhl DD, Pijnappel WWMP, Heck AJR, Timmers HTM, Kraus WL (2010) A role for BAF57 in cell cycle-dependent transcriptional regulation by the SWI/SNF chromatin remodeling complex. Cancer Res 70:4402–4411
88. Zheng P, Zhong Q, Xiong Q, Yang M, Zhang J, Li C, Bi L, Ge F (2012) QUICK identification and SPR validation of signal transducers and activators of transcription 3 (Stat3) interacting proteins. J Proteomics 75:1055–1066
89. Blagoev B, Kratchmarova I, Ong S, Nielsen M, Foster LJ, Mann M (2003) A proteomics strategy to elucidate functional protein-protein interactions applied to EGF signaling. Nat Biotechnol 21:315–318
90. Belle A, Tanay A, Bitincka L, Shamir R, O'Shea EK (2006) Quantification of protein half-lives in the budding yeast proteome. Proc Natl Acad Sci U S A 103:13004–13009
91. Doherty MK, Whitehead C, McCormack H, Gaskell SJ, Beynon RJ (2005) Proteome dynamics in complex organisms: using stable isotopes to monitor individual protein turnover rates. Proteomics 5:522–533
92. Doherty MK, Hammond DE, Clague MJ, Gaskell SJ, Beynon RJ (2009) Turnover of the human proteome: determination of protein intracellular stability by dynamic SILAC. J Proteome Res 8:104–112
93. Milner E, Barnea E, Beer I, Admon A (2006) The turnover kinetics of major histocompatibility complex peptides of human cancer cells. Mol Cell Proteomics 5:357–365
94. Cohen LD, Zuchman R, Sorokina O, Müller A, Dieterich DC, Armstrong JD, Ziv T, Ziv NE (2013) Metabolic turnover of synaptic proteins: kinetics, interdependencies and implications for synaptic maintenance. PLoS ONE 8:e63191
95. Pratt JM, Robertson DHL, Gaskell SJ, Riba-Garcia I, Hubbard SJ, Sidhu K, Oliver SG, Butler P, Hayes A, Petty J, Beynon RJ (2002) Stable isotope labelling in vivo as an aid to protein identification in peptide mass fingerprinting. Proteomics 2:157–163
96. Cargile BJ, Bundy JL, Grunden AM, Stephenson JL (2004) Synthesis/degradation ratio mass spectrometry for measuring relative dynamic protein turnover. Anal Chem 76:86–97
97. Andersen JS, Lam YW, Leung AKL, Ong S, Lyon CE, Lamond AI, Mann M (2005) Nucleolar proteome dynamics. Nature 433:77–83
98. Cambridge SB, Gnad F, Nguyen C, Bermejo JL, Krüger M, Mann M (2011) Systems-wide proteomic analysis in mammalian cells reveals conserved, functional protein turnover. J Proteome Res 10:5275–5284
99. Boisvert F, Ahmad Y, Gierliński M, Charrière F, Lamont D, Scott M, Barton G, Lamond AI (2012) A quantitative spatial proteomics analysis of proteome turnover in human cells. Mol Cell Proteomics 11:M111.011429
100. Tilghman RW, Blais EM, Cowan CR, Sherman NE, Grigera PR, Jeffery ED, Fox JW, Blackman BR, Tschumperlin DJ, Papin JA, Parsons JT (2012) Matrix rigidity regulates cancer cell growth by modulating cellular metabolism and protein synthesis. PLoS ONE 7: e37231

101. Kristensen LP, Chen L, Nielsen MO, Qanie DW, Kratchmarova I, Kassem M, Andersen JS (2012) Temporal profiling and pulsed SILAC labeling identify novel secreted proteins during ex vivo osteoblast differentiation of human stromal stem cells. Mol Cell Proteomics 11: 989–1007
102. Schwanhäusser B, Busse D, Li N, Dittmar G, Schuchhardt J, Wolf J, Chen W, Selbach M (2011) Global quantification of mammalian gene expression control. Nature 473:337–342
103. Schwanhäusser B, Gossen M, Dittmar G, Selbach M (2009) Global analysis of cellular protein translation by pulsed SILAC. Proteomics 9:205–209
104. Jayapal KP, Sui S, Philp RJ, Kok Y, Yap MGS, Griffin TJ, Hu W (2010) Multitagging proteomic strategy to estimate protein turnover rates in dynamic systems. J Proteome Res 9: 2087–2097
105. Zhang G, Deinhardt K, Chao MV, Neubert TA (2011) Study of neurotrophin-3 signaling in primary cultured neurons using multiplex stable isotope labeling with amino acids in cell culture. J Proteome Res 10:2546–2554
106. Krijgsveld J, Ketting RF, Mahmoudi T, Johansen J, Artal-Sanz M, Verrijzer CP, Plasterk RHA, Heck AJR (2003) Metabolic labeling of C. elegans and D. melanogaster for quantitative proteomics. Nat Biotechnol 21:927–931
107. Larance M, Bailly AP, Pourkarimi E, Hay RT, Buchanan G, Coulthurst S, Xirodimas DP, Gartner A, Lamond AI (2011) Stable-isotope labeling with amino acids in nematodes. Nat Methods 8:849–851
108. Sury MD, Chen J, Selbach M (2010) The SILAC fly allows for accurate protein quantification in vivo. Mol Cell Proteomics 9:2173–2183
109. Xu P, Tan H, Duong DM, Yang Y, Kupsco J, Moberg KH, Li H, Jin P, Peng J (2012) Stable isotope labeling with amino acids in Drosophila for quantifying proteins and modifications. J Proteome Res 11:4403–4412
110. Krüger M, Moser M, Ussar S, Thievessen I, Luber CA, Forner F, Schmidt S, Zanivan S, Fässler R, Mann M (2008) SILAC mouse for quantitative proteomics uncovers kindlin-3 as an essential factor for red blood cell function. Cell 134:353–364
111. Wu CC, MacCoss MJ, Howell KE, Matthews DE, Yates JR (2004) Metabolic labeling of mammalian organisms with stable isotopes for quantitative proteomic analysis. Anal Chem 76:4951–4959
112. McClatchy DB, Dong M, Wu CC, Venable JD, Yates JR (2007) 15N metabolic labeling of mammalian tissue with slow protein turnover. J Proteome Res 6:2005–2010
113. Ishihama Y, Sato T, Tabata T, Miyamoto N, Sagane K, Nagasu T, Oda Y (2005) Quantitative mouse brain proteomics using culture-derived isotope tags as internal standards. Nat Biotechnol 23:617–621
114. Geiger T, Cox J, Ostasiewicz P, Wisniewski JR, Mann M (2010) Super-SILAC mix for quantitative proteomics of human tumor tissue. Nat Methods 7:383–385
115. Deeb SJ, D'Souza RCJ, Cox J, Schmidt-Supprian M, Mann M (2012) Super-SILAC allows classification of diffuse large B-cell lymphoma subtypes by their protein expression profiles. Mol Cell Proteomics 11:77–89
116. Monetti M, Nagaraj N, Sharma K, Mann M (2011) Large-scale phosphosite quantification in tissues by a spike-in SILAC method. Nat Methods 8:655–658
117. Tzouros M, Golling S, Avila D, Lamerz J, Berrera M, Ebeling M, Langen H, Augustin A (2013) Development of a 5-plex SILAC method tuned for the quantitation of tyrosine phosphorylation dynamics. Mol Cell Proteomics 12(11):3339–3349
118. Hebert AS, Merrill AE, Stefely JA, Bailey DJ, Wenger CD, Westphall MS, Pagliarini DJ, Coon JJ (2013) Amine-reactive neutron-encoded labels for highly plexed proteomic quantitation. Mol Cell Proteomics 12:3360–3369

Chapter 6
Utility of Computational Structural Biology in Mass Spectrometry

Urmi Roy, Alisa G. Woods, Izabela Sokolowska, and Costel C. Darie

Abstract Recent developments of mass spectrometry (MS) allow us to identify, estimate, and characterize proteins and protein complexes. At the same time, structural biology helps to determine the protein structure and its structure–function relationship. Together, they aid to understand the protein structure, property, function, protein–complex assembly, protein–protein interaction and dynamics. The present chapter is organized with illustrative results to demonstrate how experimental mass spectrometry can be combined with computational structural biology for detailed studies of protein's structures. We have used tumor differentiation factor protein/peptide as ligand and Hsp70/Hsp90 as receptor protein as examples to study ligand–protein interaction. To investigate possible protein conformation we will describe two proteins, lysozyme and myoglobin.

6.1 Introduction

6.1.1 What Is Structural Biology?

Structural biology is a rapidly developing field in science that determines the structure of macromolecules and how structure relates to its function. It is the structural and functional analysis of biomolecules and biological systems. X-ray crystallography and nuclear magnetic resonance (NMR) spectroscopy are the two most common experimental methods widely used to determine bio macromolecular structures [1].

U. Roy • A.G. Woods • I. Sokolowska • C.C. Darie (✉)
Department of Chemistry & Biomolecular Science, Biochemistry & Proteomics Group, Structural Biology & Molecular Modeling Unit, Clarkson University, 8 Clarkson Avenue, Potsdam, NY 13699-5810, USA
e-mail: cdarie@clarkson.edu

A.G. Woods and C.C. Darie (eds.), *Advancements of Mass Spectrometry in Biomedical Research*, Advances in Experimental Medicine and Biology 806,
DOI 10.1007/978-3-319-06068-2_6, © Springer International Publishing Switzerland 2014

X-ray crystallography is the most powerful method for protein structure determination and analysis [2]. During the last decade, the technique of X-ray crystallography has advanced significantly, and at this time, can provide very high structural resolution through rather routine measurements. Using X-ray crystallography we can assess the structure–function relationship as well as study the ligand–receptor interactions. We can also obtain atomic-level structural information of proteins by using solution phase NMR spectroscopy, where the experimental samples can be conveniently examined under normal physiological conditions. Using NMR spectroscopy proteins conformational changes, stability and dynamics could also be measured. Moreover, solid state NMR is used to study the membrane protein structures [3]. Cryo-electron microscopy (Cryo-EM) is another recently developed powerful experimental tool in structural biology [4]. The structural assembly in a macromolecular complex is studied using this technique. In recent years, there is a trend to combine Cryo-EM and X-ray crystallographic technique for detail understanding of the biological supramolecular complexes [4].

In silico or computational approach is another method to determine the 3D structures of biological macromolecules. Computational methods are supplemental to the experimental data and 3D models are particularly important when there are no readily available crystal or NMR structures. Homology modeling is a typical method for predicting protein structures. Using homology modeling we can predict the 3D structure of a protein based on the available template protein structure [5, 6]. Homology modeling typically follows four steps: identification of a template protein, alignment, building a model structure, and lastly assessing that model [7]. Structural or sequence homologies may suggest mutually similar biological functions. Other methods of protein structure predictions include *ab initio* and molecular threading approaches [8–10]. The *ab-initio* approach is based largely on a set of estimates derived from first-principle calculations, and does not require a template structure for execution. *Ab-initio* method works best for smaller protein structure prediction. Prediction by threading uses fold recognition technique and assesses a known protein template with the target protein sequence. Additionally, Molecular dynamics simulation (MD) may give some insight about proteins structural/conformational changes, folding/unfolding, and stability as functions of time [11]. Longer MD simulations may improve a model structure [12, 13]. The analytical capability of computational structural biology has significantly progressed in recent years by benefiting from the major advances occurred during the same time in the areas of modeling algorithms [14, 15]. Critical assessment of structure prediction (CASP) evaluates and assesses these structural protein model prediction methods [16]. Today's structural biology is immensely used for fundamental research—to predict and view biomolecular structure, to analyze their atomic level molecular interactions, and to study the dynamics of a bio-system.

In earlier years the biochemical studies were centered on small molecules, focusing largely on metabolites and metabolic pathways [17]. As the last two decade's efforts in this field have gradually moved toward the investigations of larger and complex molecules like proteins, a rather large number of protein

structures have become available. For instance, as of November 12, 2013, the RCSB protein data bank (PDB) contains 95,475 protein structures. As the new protein structures are discovered, the structural biology become more innovative and challenging [14]. Very recently, new project-like human proteome projects have been undertaken [18]. Broadly, development of these protein structures and design as well as screening of small molecules, their structural insights, and interaction analyses are the basis of today's drug discovery and pharmaceutical industry [19].

6.1.2 Structural Biology and Mass Spectrometry

Mass spectrometry (MS) is a high-throughput experimental technique that aids in the sorting and separation of ions based on their mass/charge (m/z) ratio. It is a useful method widely used in many areas including Biology, Chemistry, Geology, Medicine, and Environmental Science. Study of peptides and proteins using MS requires their existence as charged gas-phase peptide or protein ions, and thus their m/z ratios are measured [20–28]. The MS experiments are categorized based on the ionization source and mass analyzer. Two MS techniques are typically used in proteomics experiments: (1) Matrix Assisted Laser Desorption Ionization (MALDI) [29] and (2) Electrospray Ionization (ESI) [30]. MALDI-MS uses dried matrix and a laser. ESI-MS uses ionic peptides and proteins in liquid phase and is generally coupled with liquid chromatography. MS allowed researchers to characterize biomolecules and using the MS-based proteomics approaches we can identify possible cellular connections and interactions that involve signaling pathways, physiology, and disease development [25, 31–39].

During the last several years, MS progressed enormously to aid research in structural biology [40–42]. MS coupled experiments not only can predict biomolecular structures, sequences, and properties, but also help to monitor macromolecular assemblies as well as interactions among subunits and intricate dynamical systems [43]. In the following section we will briefly demonstrate how MS experimental data are used in proteins structural analysis and dynamics.

Affinity purification-mass spectrometry (AP-MS) is used to identify the structure and stoichiometry of protein subunits [42]. Combining AP-MS and proteomics techniques help to identify and estimate protein–protein association, interactome analysis, dynamics, and relevant functional components [44].

In recent years, mass spectrometry (MS) coupled hydrogen/deuterium exchange (HDX) method is emerged as a powerful tool in structural biology. HDX MS techniques are used to determine transitional conformational changes, comprehensive protein dynamics, inter-domain interaction as well as ligand-induced structural reordering [45–47]. Even residue-level structural insight, proteins property like solvent accessibility, and protein–protein interaction analysis are possible using HDX-MS [48].

Ion mobility-mass spectrometry (IM-MS) are used to examine biomolecular conformation and this procedure is used to determine the dissociation between liquid and gas phase and recently used to compare the shape and conformations between IM-MS gas phase experiment and the crystal structure [40, 41]. The structural architecture/assembly within proteins and protein complexes could be determined using ESI in combination with IM-MS [46].

Chemical cross-linking MS (CX-MS) has been used through several decades and has progressively advanced. CX-MS is mostly utilized to measure the distances between protein's subunits and to extract information about their structural details, protein–protein interactions, and neighboring residues [42]. As an example, CX-MS and modeling experiment together helps to expose the topology of TRiC/CCT chaperonin [42, 49, 50].

Unlike soluble proteins, structural studies of membrane proteins are non-trivial and difficult to examine using traditional methods like NMR or X-ray crystallography. With MS and imaging MS, membrane protein structures, their interactions and subunit dynamics are now being investigated and analyzed [51–53].

6.2 Examples of Structural Biology and Mass Spectrometry

6.2.1 Protein–Peptide and Protein–Protein Interactions

Perhaps the combination of structural biology and MS is at its best when complemented by each other and by other methods. For example, we recently used a combination of affinity purification, co-immunopurification, mass spectrometry, and structural biology to identify the potential receptor candidates for tumor differentiation factor (TDF), a protein produced by the pituitary and secreted through the blood stream [24, 25, 31, 54–57]. The target organs for this protein are breast and prostate, but its receptor was unknown. Therefore, the purification and identification of the TDF receptor candidates was the obvious step that had to be performed. To do so, we used TDF-P1, a peptide from the open reading frame of the TDF as bait for our affinity purification experiments. MS analysis of the affinity-purified proteins led to identification of GRP78, HSP70, and HSP90 proteins as potential TDF receptor (TDF-R) candidates. Example of MS/MS spectra that led to identification of these three proteins are shown in Fig. 6.1. Two types of interactions were identified in these experiments: peptide–protein interactions (TDF-P1-GRP78), as determined by AP-MS and further confirmed by structural biology and Western blotting (Figs. 6.2 and 6.3). The next logical question was whether TDF-P1 interacts directly with HSP90. Using two different types of structural biology software, we were then able to conclude that TDF-P1 also interacts with HSP90 (Figs. 6.4 and 6.5). Therefore, structural biology, when combined with MS, immuno-affinity purification and Western blotting, as well as with other complementary techniques, becomes particularly powerful.

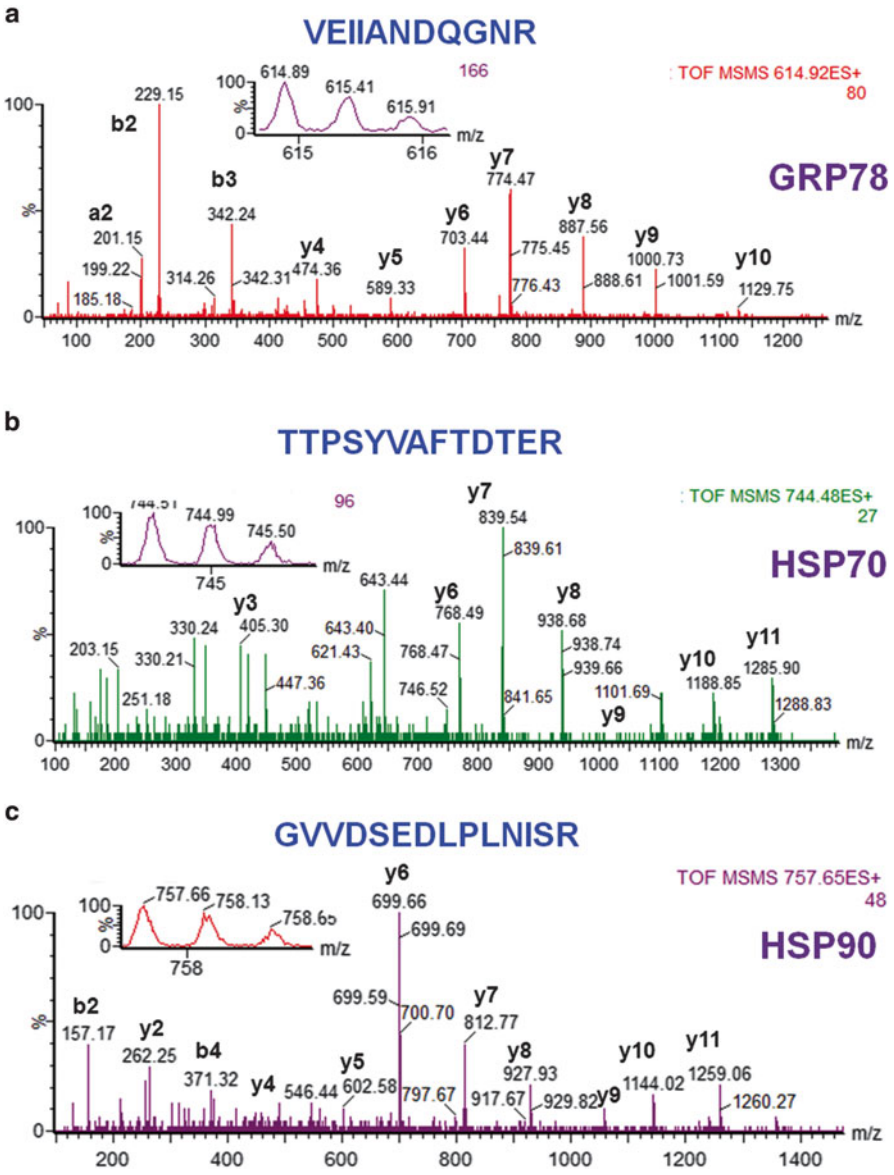


Fig. 6.1 Identification of TDF-R candidates in DU145 cells using AP and LC-MS/MS (AP-MS). The TDF-R candidates were purified from cell lysates by AP using TDF-P1 peptide crosslinked to agarose beads as a bait. The purified proteins were separated by SDS-PAGE and the gel bands were excised and digested by trypsin. The peptide mixtures were analyzed by LC-MS/MS to identify the purified proteins. (**a**) MS/MS spectrum of peptide VEIIANDQGNR that led to identification of GRP78 as TDF-R candidate. (**b**) MS/MS spectrum of peptide TTPSYVAFTDTER that led to identification of HSP70 as TDF-R candidate. (**c**) MS/MS spectrum of peptide GVVDSEDLPLNISR that led to identification of HSP90 as TDF-R candidate. Reprinted with permission from Roy et al., 2013 [55]

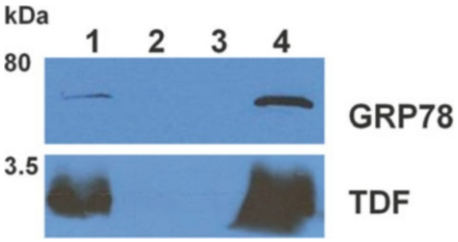


Fig. 6.2 Immuno-affinity precipitation (IAP) of GRP78 and TDF-P1 mixture. GRP78 protein was incubated with TDF-P1 peptide and then the mixture was precipitated by IAP using anti-GRP78 antibodies (mouse monoclonal). The input, flow through, control and eluate were then investigated by WB using anti-GRP78 (rat monoclonal) and anti-TDF (rabbit polyclonal) antibodies. Each WB contains input mixture of GRP78 protein and TDF-P1 (lane 1), flow through after IP (lane 2) negative control (lane 3) and eluate from IAP (lane 4). The molecular mass markers are shown (in kDa) on the *left*. Reprinted with permission from Sokolowska et al., 2012 [31]

6.2.2 Investigation of Protein Conformation

6.2.2.1 Case Study I: Lysozyme

We have taken hen egg-white lysozyme (PDB ID: 1HEW) as an example to study protein's structural properties [58]. 1HEW is the hen egg-white lysozyme with tri-N-acetylchitotriose inhibitor complex. We have used H++ to predict the lysozyme structure at pH 2.0 [59–61]. H++ is an automated server [http://biophysics.cs.vt.edu/H++] that protonates the residues of a side chain to a given pH and computes the pK values of ionizable groups in that protein structure. The relative accessibility of the amino acids in hen egg-white lysozyme is determined using Swiss- PDB viewer [6, 62].

1HEW is a single chain protein with 129 residues. Eight cysteine residues are present in 1HEW. These are: Cys6, Cys30, Cys64, Cys76, Cys80, Cys94, Cys115, and Cys127. Four disulfide bridges are found between Cys6-Cys127, Cys30-Cys115, Cys64-Cys80, and Cys76-Cys94. These cysteine residues and the cystine bridges are displayed in Fig. 6.6a, b. There are two glutamic acid residues in 1HEW. Pdb. These are Glu7 and Glu35. Glu7 is an exposed residue. Seven aspartic acid residues are present in 1HEW. Among them Asp18, Asp48, Asp87, Asp101, and Asp119 are surface-exposed residues [6, 62]. The glutamic acid and aspartic acid residues present in 1HEW are displayed in Fig. 6.6c. There are six lysine residues in 1HEW.Pdb. These are Lys1, Lys13, Lys33, Lys96, Lys97, and Lys116. Among them Lys1, Lys13, Lys33, Lys97, and Lys116 are exposed residues. Eleven arginine residues are present in 1HEW. These are Arg5, Arg14, Arg21, Arg45, Arg61, Arg68, Arg73, Arg112, Arg114, Arg125, and Arg128. All of them are surface-exposed residues [6, 62]. In a previous paper, Mehler and Guarnieri described the buried residues in hen egg-white lysozyme [63]. One histidine residue is present in 1HEW structure. This is His15. This is a buried residue. The lysine, arginine, and histidine residues present in 1HEW structure are displayed in Fig. 6.6d. Figure 6.6a–d are generated using Accelrys Discovery Studio 3.5.

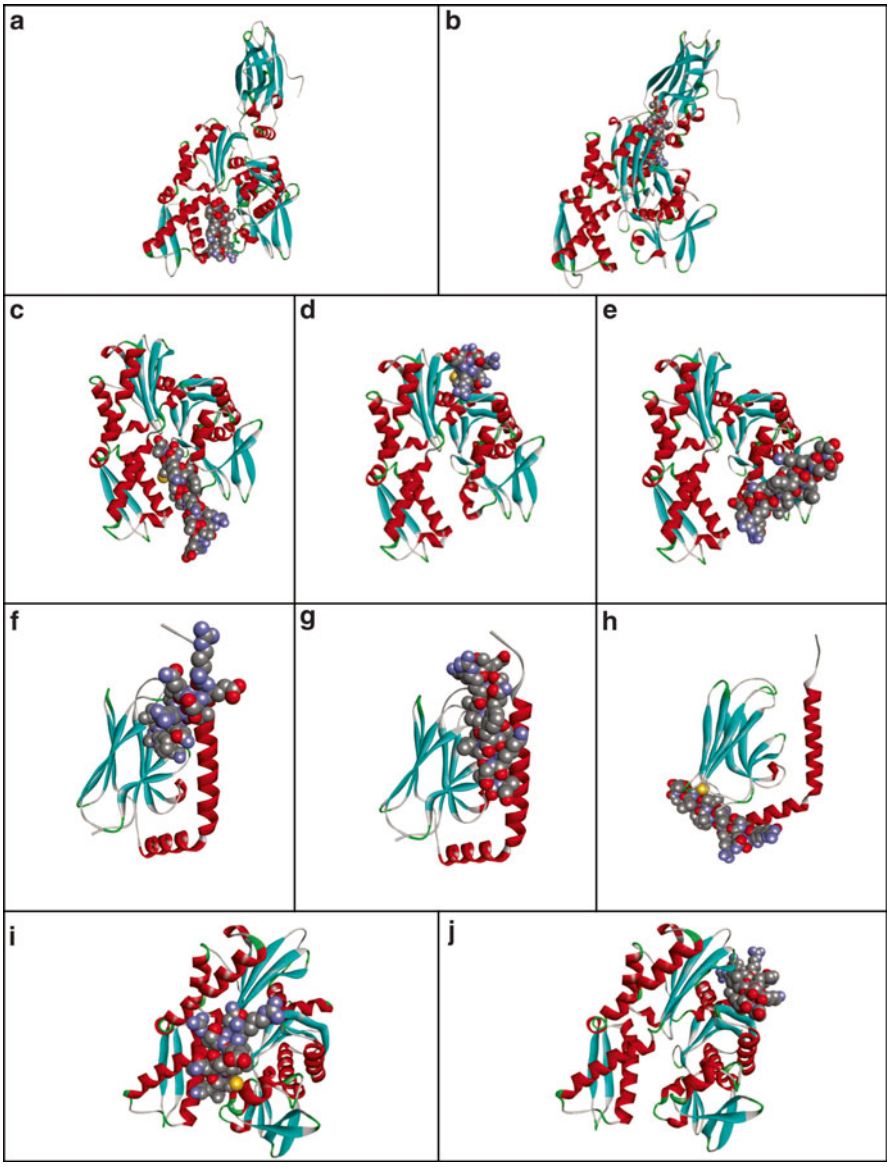


Fig. 6.3 Possible peptide binding pockets tentatively identified by GRAMM-X protein–protein docking web server v.1.2.0. The TDF-P1 peptide (P1) binding pockets were predicted using pdb structures 1YUW (**a**, **b**), 3LDN (**c**–**e**), 3N8E (**f**–**h**) and 2E88 (**i**, **j**). Receptor proteins are shown in the form of a ribbon diagram colored by secondary structure. P1 peptide is shown in space-filling mode colored by atom type. Reprinted with permission from Sokolowska et al., 2012 [31]

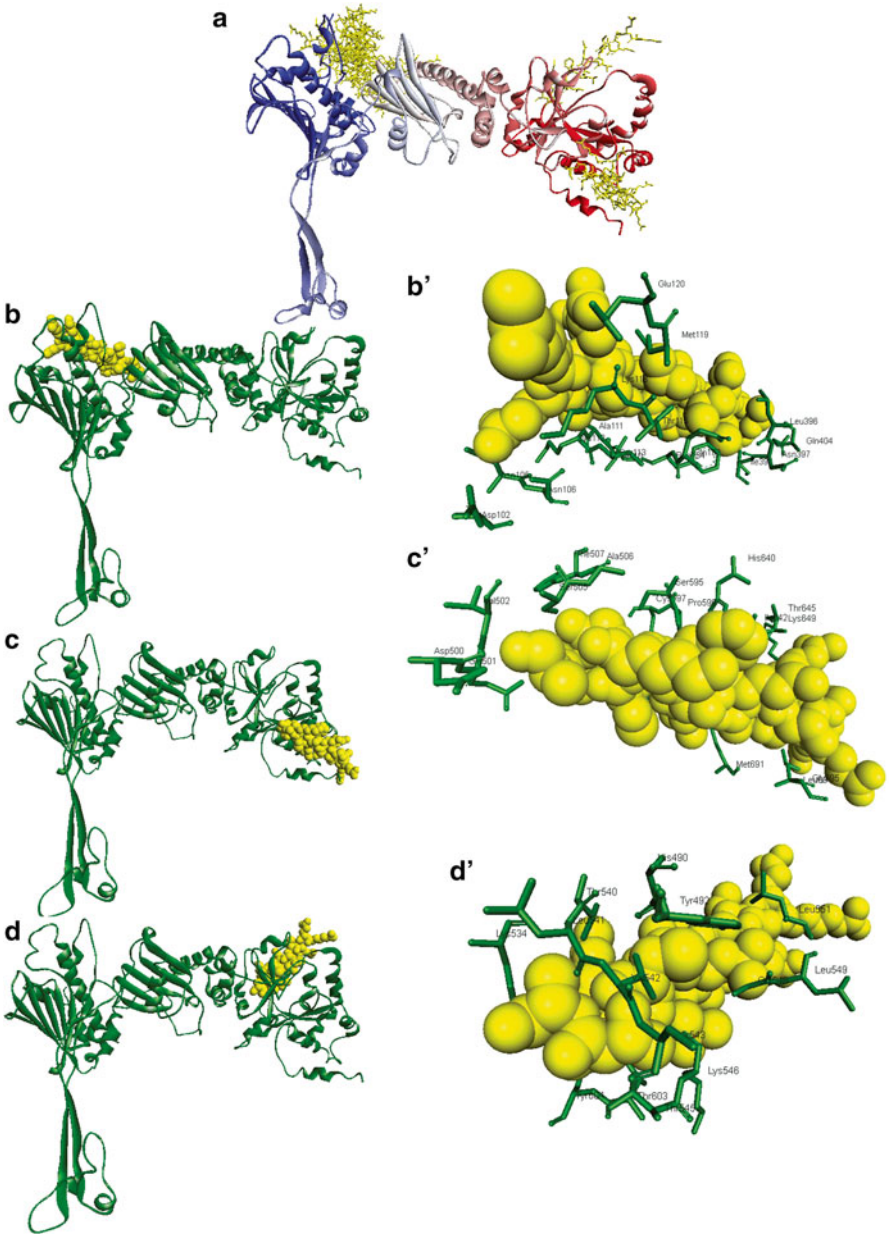


Fig. 6.4 Tentative docking sites of TDF-P1 on model HSP90 receptor protein as identified by "GRAMM-X." (**a**) Predicted docking sites and poses of docked peptide (*yellow*). The secondary structure of the model receptor protein is colored from N (*blue*) to C (*red*) terminus. (**b**–**d**) Predicted P1 binding sites, with the receptor protein displayed in *green* ribbon. The P1 peptide is shown in the *yellow* space-filled mode. (**b′**–**d′**) Neighboring amino acid residues of docked P1 on the model receptor protein. Reprinted with permission from Roy et al., 2013 [55]

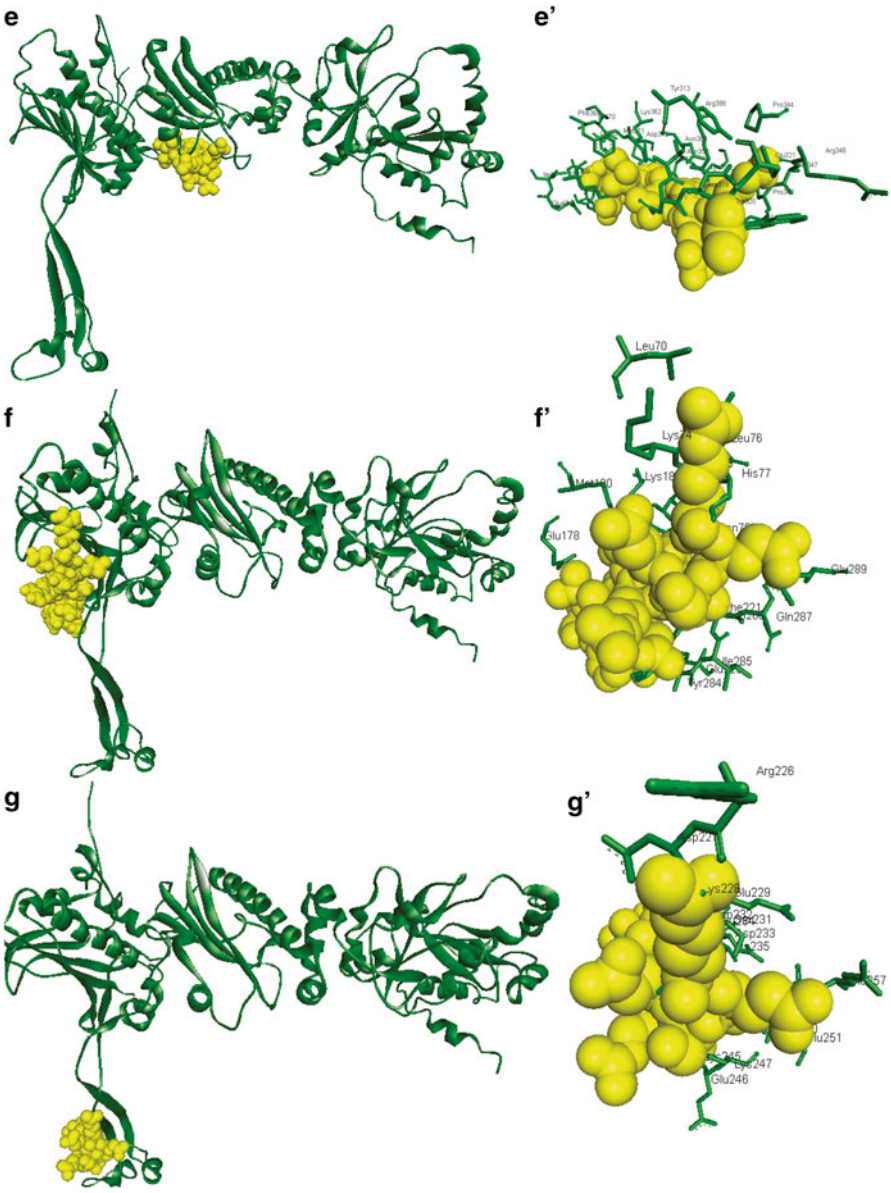


Fig. 6.5 Three additional potential docking sites on HSP90 for TDF-P1, as predicted by using "Patch dock" and "Fire dock." (**e–g**) Predicted P1 binding pockets. The receptor protein is displayed in *green* ribbon. The P1 peptide is shown in the *yellow* space filling mode. (**e'–g'**) Neighboring amino acid residues of docked P1 on the model receptor protein. Reprinted with permission from Roy et al., 2013 [55]

The relative accessibility of the amino acids of this protein at pH2 is displayed in Fig. 6.7. In Fig. 6.7a, the exposed Glu and Asp residues are displayed in space-filled mode. Figure 6.7c demonstrated the surface-exposed Lys and Arg residues in

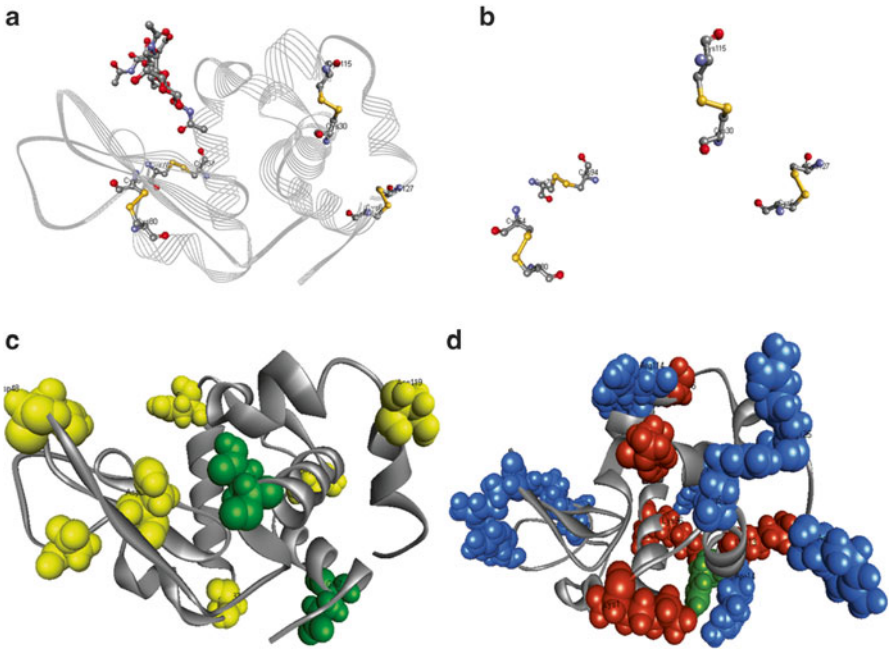


Fig. 6.6 (**a**, **b**) Cysteine residues and disulfide bridges in 1HEW. (**c**) The Glu (*green*) and Asp (*yellow*) residues of 1HEW. (**d**) The Arg (*blue*), Lys (*brown*), and His (*green*) residues of 1HEW. These residues are displayed in space-filled mode. Secondary structure of 1HEW is displayed in *gray* color

space-filled mode. The blue color represents buried residues; where the red represents the residues with greater surface accessibility [http://www.expasy.org/spdbv/]. Figure 6.7a–d are generated using Swiss Pdb Viewer/Deepview [6, 62].

A simple MS-based structural biology test for investigation of protein conformation is to analyze a protein in native and denatured states and to compare the two conformations and their MS spectra. This process can be easily performed using electrospray ionization MS (ESI-MS), simply by injecting the native or denatured protein using a syringe, in a process called direct infusion. Depending on the ionization mode, one may monitor the number and abundance of the protonated (positive ionization mode) or deprotonated species for this protein and therefore we may infer the conformation of a particular protein. Such an example is given in Fig. 6.8, where Lysozyme was analyzed by ESI-MS (positive mode) under acidic conditions (2 % Acetonitrile/ACN, 0.1 % Formic acid/FA in HPLC water). Here Lysozyme is present in 11 species, differently protonated. The number of charges varies from 7(+) (*m*/*z* of 2047.47) and 17(+) (*m*/*z* of 841.83). The most compact conformation of Lysozyme has the smallest number of charges because most chargeable amino acids (Arg, Lys, or His) are buried inside the three-dimensional conformation of the protein, while the most denatured (i.e., linear) Lysozyme molecule will have the highest number of charges, as upon denaturation, this protein's protonation sites (i.e., Arg, Lys, and His residues) become

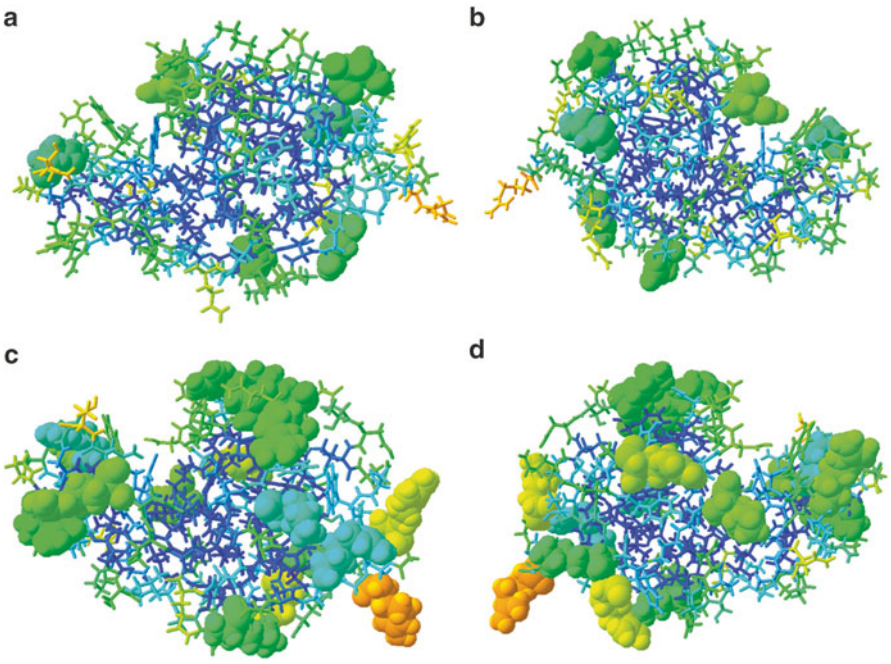


Fig. 6.7 This picture illustrates the exposed vs. buried residues of 1HEW.pdb. (**a**, **b**) The surface exposed Glu and Asp residues at pH 2 are displayed in space-filled mode. (**c**, **d**) The surface exposed Lys and Arg residues at pH 2 are displayed in space-filled mode. The relative accessibility of the amino acids of this protein is based on rainbow color- the *blue* color represents buried residues; and *red* residues are with higher surface accessibility. (**a**, **c**) Front view. (**b**, **d**) View from the back

available. This is well reflected in Fig. 6.8, where in the native conformation of this protein, the species with the highest abundance is the 10(+) (*m*/*z* of 1432.05), while in the denatured conformation of this protein (the protein was vortexed extensively prior ESI-MS analysis), the species with the highest abundance is the 13(+) (*m*/*z* of 1432.05), suggesting that upon denaturation, the protein was in average protonated to three additional protonation sites (Arg, Lys, and His residues), that were previously less accessible for protonation in the native state (Fig. 6.8). We would like to note that the "native" and "denatured" terms are absolute extremes, and should rather be viewed as two different conformations (i.e., conformation 1 and conformation 2).

When the ESI-MS and structural biology data are compared with each other, particularly important information about the conformation of a protein can be extracted. In our case, Lysozyme has a total of 5 Lys residues and 11 Arg residues, all exposed, and a His residue, buried inside the protein. This makes a total of 16 (out of 17 total) potential protonation sites. When we compare the conformation 1 (native, the most abundant ion species is 10(+)) and conformation 2 (denatured, most abundant ion species 13(+)), we can easily conclude that three (the difference between 10(+) and 13(+) ion species) out of the 16 exposed protonation sites are

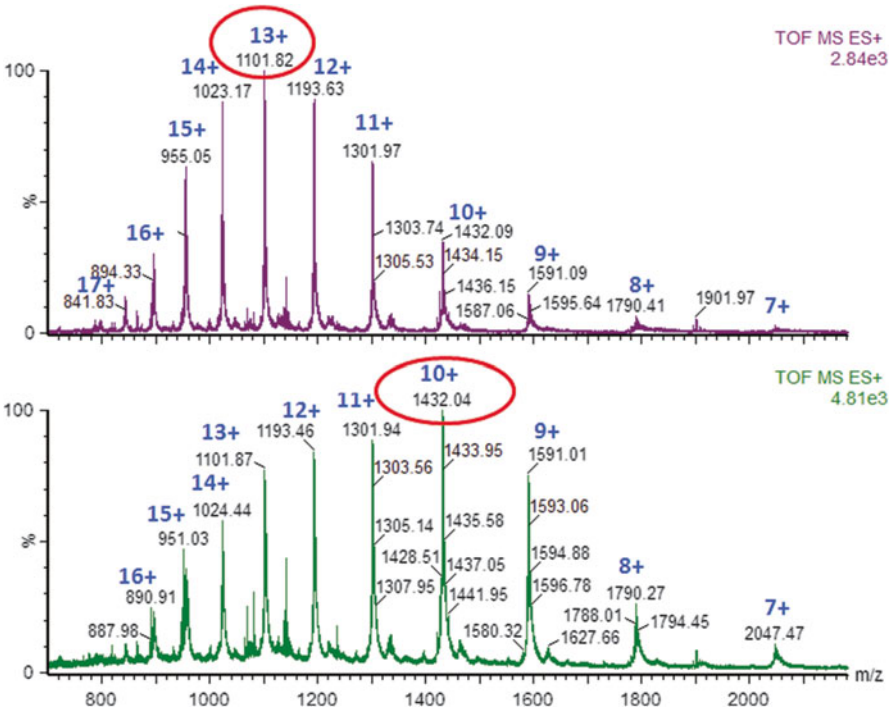


Fig. 6.8 ESI-MS spectra (positive mode) of lysozyme which has two main, different conformations (1 & 2 or denatured and native conformations). The spectrum with the 13(+) ion as the most abundant is the denatured conformation. The protein in the denatured conformation was obtained by extensive vortexing. The spectrum with 10(+) ion as the most abundant is the native conformation. During transition from the native to denatured conformation, more protonation sites are exposed to the environment and become protonated

either buried inside the protein or involved in H-bonds, thus keeping the protein in native state (in addition to the disulfide bridges present). Additional inspection of the spectra also allowed us to conclude that the apparition of the 17(+) ion in the denatured protein suggests that the H residue is protonated, and the protein is indeed denatured.

6.2.2.2 Case StudyII: Myoglobin

The crystal structure of the myoglobin (Mb) was described by John Kendrew and coworkers in 1958 [64]. Max Perutz and John Kendrew received Nobel Prize in Chemistry for discovering hemoglobin and myoglobin structures. Myoglobin is a single chain globular protein with eight alpha helices (helices A–H). Heme [protoporphyrin IX with Fe] is located in a hydrophobic cleft in Mb.

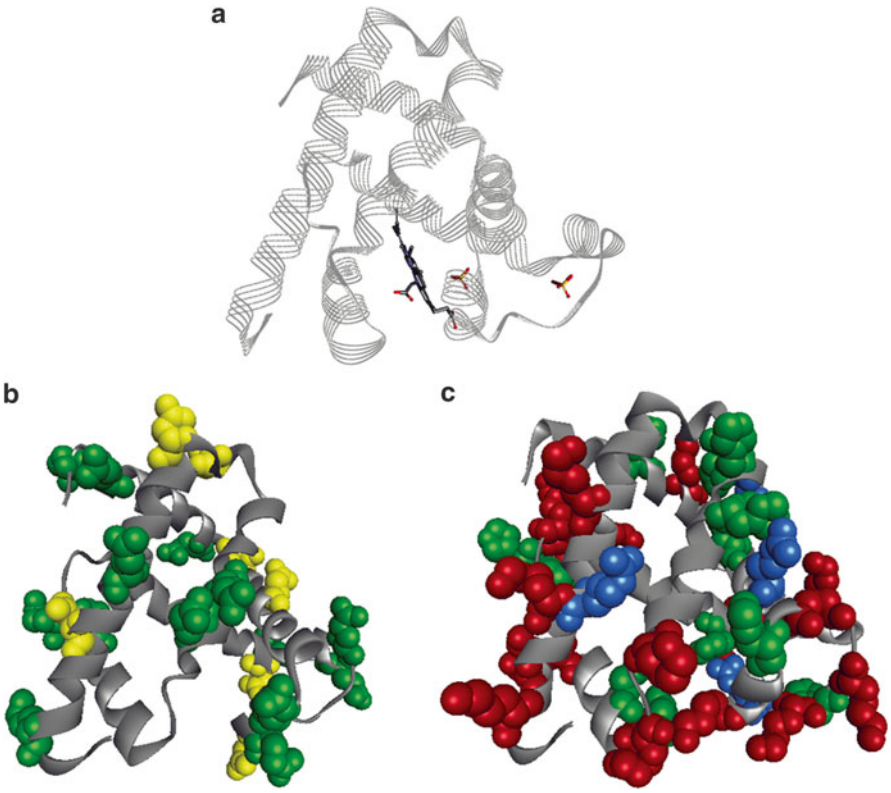


Fig. 6.9 (**a**) Secondary structure of myoglobin (PDB ID: 1VXB). (**b**) The Glu (*green*) and Asp (*yellow*) residues of 1VXB. (**c**) The Arg (*blue*), Lys (*brown*), and His (*green*) residues of 1VXB. These residues are displayed in space-filled mode

We have taken myoglobin (PDB ID: 1VXB) as an example to study its exposed and buried residues. 1VXB is a single chain protein that consists of 153 amino acid residues [65]. H++ server [http://biophysics.cs.vt.edu/H++] was used to predict this structure at pH 2.0 [59–61]. The buried residues and relative accessibility of the amino acids in 1VXB at pH 2.0 was determined using Swiss- PDB viewer [6, 62]. Mb has been extensively studied in the low-pH (2.0–4.0) regime [66–70], and as suggested by some of these studies [67, 69], the iron–His bond in myoglobin tends to break at such low pH values. Figure 6.9a displays the secondary structure of Mb. There are 14 glutamic acid and 7 aspartic acid residues in 1VXB. 19 lysine, 4 arginine, and 12 histidine residues are present in this structure. The glutamic acid and aspartic acid residues present in 1VXB are displayed in Fig. 6.9b. The lysine, arginine, and histidine residues present in Mb structure are displayed in Fig. 6.9c. Figure. 6.9a–c are generated using Accelrys Discovery Studio 3.5. Figure 6.10 displays the exposed and buried residues of Mb at pH2. Asp20, Asp 44, Asp 60, Asp 122, Asp 126 are exposed residues.Glu4, Glu6, Glu18. Glu38, Glu41, Glu52, Glu54, Glu59, Glu83, Glu105, Glu109, Glu136,

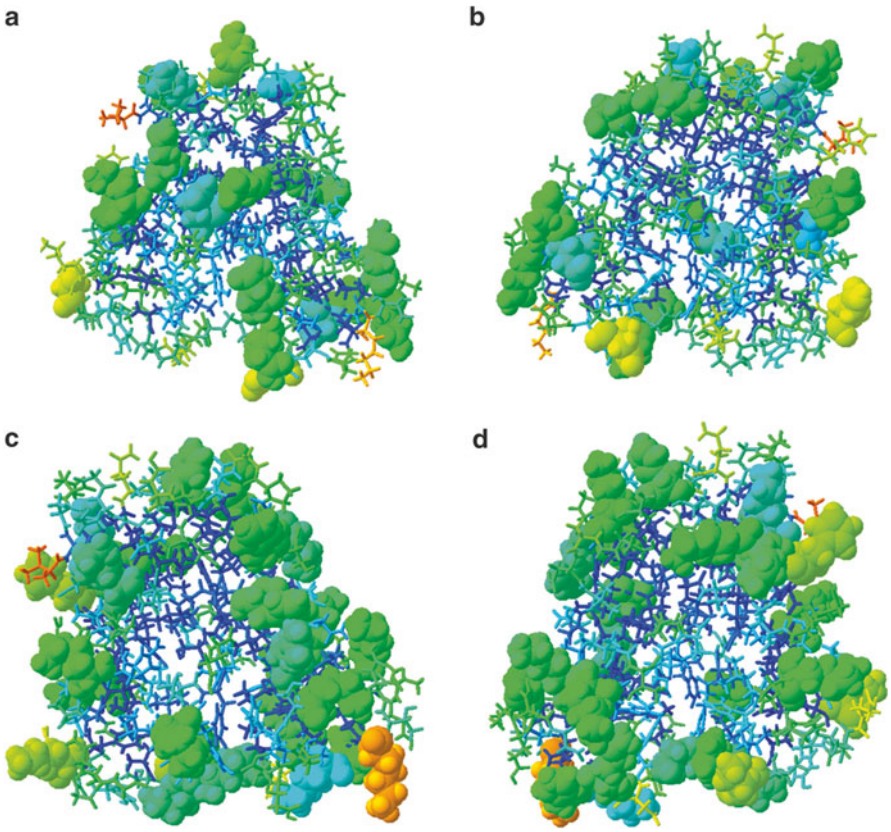


Fig. 6.10 This depiction illustrates the exposed vs. buried residues in 1VXB. (**a**, **b**) The surface exposed Glu and Asp residues at pH 2 are displayed in space-filled mode. (**c**, **d**) The surface exposed Arg, Lys, and His residues at pH 2 are displayed in space-filled mode. The relative accessibility of the amino acids of this protein is based on rainbow color- *blue* denotes buried residues; and *red* demonstrate residues with greater surface accessibility. (**a**, **c**) Front view. (**b**, **d**) View from the back

and Glu148 are exposed residues. Both Arg139 and Lys145 are buried residues. His12, His48, His81, His97, His113, and His116 are exposed residues.

Comparative surface accessibility of the amino acids of this protein at pH2 is displayed in Fig. 6.10. Figure 6.10a, c depict the exposed Glu/Asp and Arg/Lys/His residues in space filling mode [http://www.expasy.org/spdbv/]. Figure 6.10a–d are generated using Swiss Pdb Viewer/Deepview [6, 62].

A similar case as in the ESI-MS-based investigation of the two different conformations of Lysozyme can also be observed for myoglobin. Figure 6.11 shows such an example, where two different conformations are observed. For example, in conformation 1, the most abundant species is the 18(+) ion (Fig. 6.11a and zoomed in b), in conformation 2, the most abundant species is the 23(+) ion (Fig. 6.11a and zoomed in b), In "a," the heme prosthetic group (*m*/*z* of 616.33) is also shown.

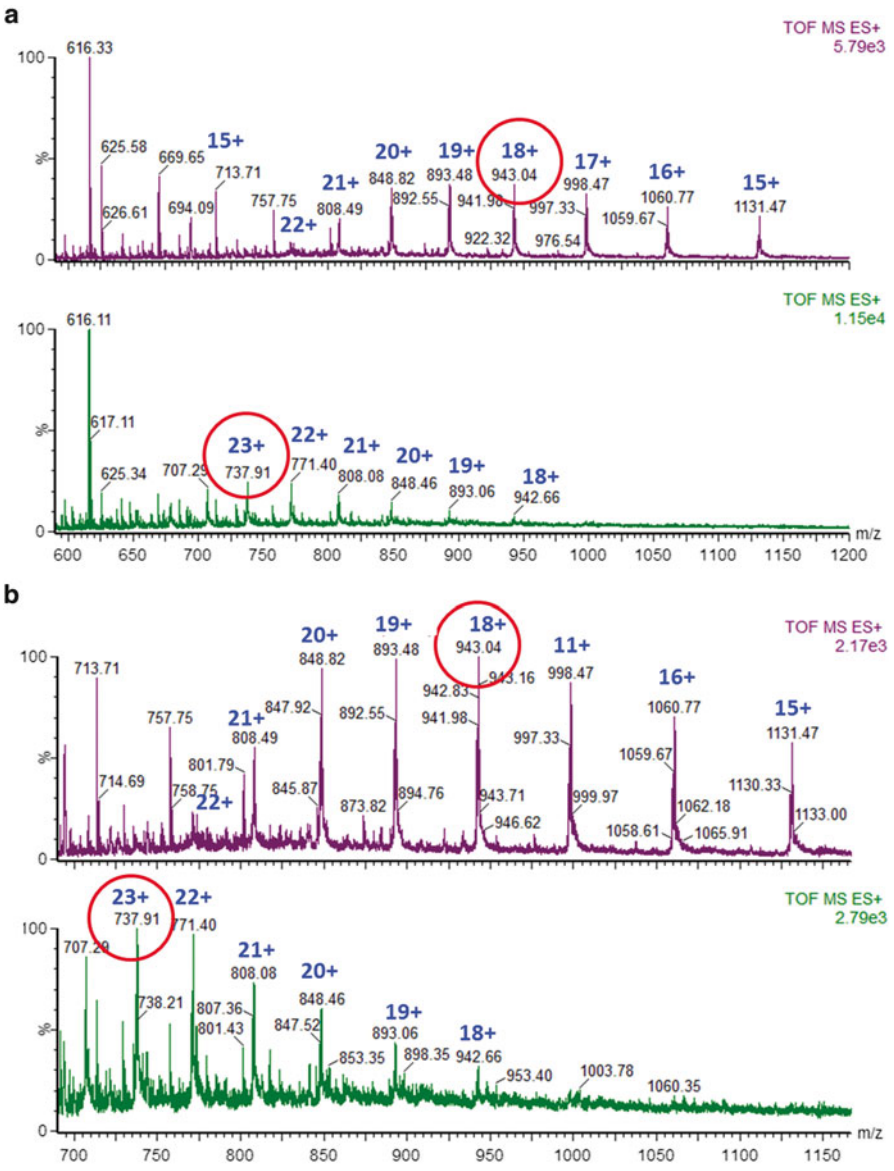


Fig. 6.11 ESI-MS spectra (positive mode) of myoglobin which has two main, different conformations (1 & 2 or denatured and native conformations). (**a**) The spectrum with the 23(+) ion as the most abundant is a more denatured conformation. The spectrum with 18(+) ion as the most abundant is a more native conformation. During transition from the native to denatured conformation, more protonation sites are exposed to the environment and become protonated. (**b**) The spectra shown in (**a**), but zoomed in

When the structural biology data and ESI-MS data for myoglobin are compared for the presence of the protonation site (R, K, or H) and the number of exposed residues within Myoglobin, we observe that we have a maximum of 19 Lys, 4 Arg, and 12 His residues, of which a maximum of 18 Lys, 3 Arg, and 6 His (total 27 residues) are exposed. Judging from Fig. 6.11, we can easily conclude that in conformation 1, where the most abundant species is the 18(+) ion, there are a total of 9 residues that are either involved in H bonds or are prevented to be protonated, thus buried inside the protein. In contrast, the protein in conformation 2 has the most abundant species as the 23(+) ion, which leaves only 4 residues that are either involved in H bonds or are prevented to be protonated. Therefore, protein in conformation 2 is the most elongated and/or denatured conformation of myoglobin.

When one investigates the structure of a protein, additional ESI-MS approaches can be used that, most of the time are confirmatory and/or complementary to the ESI-MS. For example, the spectra from Figs. 6.8 and 6.11 were recorded under acidic conditions and in positive ionization. However, one can use a neutral pH in positive mode, which allows deprotonation of His residues, as well as ESI-MS in negative mode at various pH. In this way, by analyzing the same protein under various conditions of pH and ionization, additional and sometimes complementary information can be extracted. Figure 6.12 shows such an example, where myoglobin was analyzed under both positive and negative ionization. Therefore, in positive ionization Arg, Lys, and His are protonated, while in negative ionization, amino acids Glutamate/Glu and Aspartate/Asp are deprotonated. As mentioned earlier and shown in Figs. 6.9 and 6.10, myoglobin contains 14 Glu and 7 Asp residues (21 total) of which only two Asp and one Glu are buried inside the protein. Yes, data from Fig. 6.12 show that myoglobin, when analyzed in negative mode has the ion species with 6(−) charges, meaning that only 6 Asp and Glu are deprotonated. By combining the information about the number and type of residues that are exposed (and buried inside the protein) under positive ionization (Arg, Lys, His) and the information about the number and type of residues that are exposed (and buried inside the protein) under negative ionization (Glu, Asp, and cysteine when it is not involved in disulfide bridges), additional information about the residues involved in H bridges are revealed.

Another type of information that can be extracted using ESI-MS of a protein is the identification of the loose end of a protein. When a protein has only two ends (i.e., N- and C- termini, unlike insulin, where there are two disulfide polypeptides with four ends), one can investigate the fragmentation of the intact protein by ESI-MS/MS. As such, the N-terminal part of the protein will produce b ions, while the C-terminal part of the protein will produce y ions. Therefore, identification of mostly y ions would suggest that the protein has the C-terminal part more flexible. Conversely, identification of the b ions would suggest that the N-terminal part of the protein is flexible. Such an example is shown in Fig. 6.13, where the 18(+) charged ion of myoglobin was selected for fragmentation. Analysis of the product ions within the MS/MS spectrum suggests that myoglobin has a flexible arm on the N-terminal end. Evidence for this statement comes from identification of the product ion with m/.z of 988.28 (Fig. 6.13b), which was produced from fragmentation of the

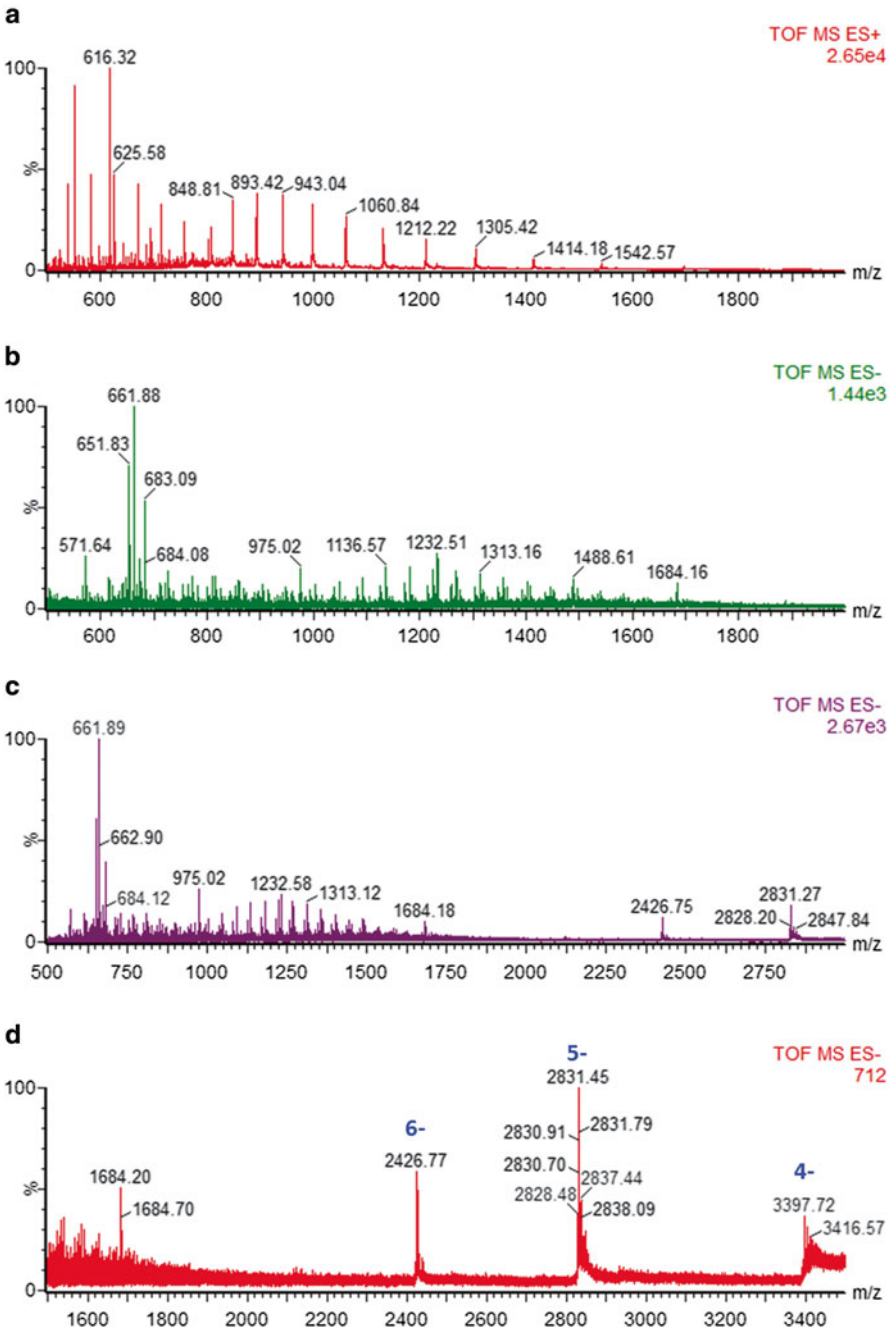


Fig. 6.12 ESI-MS spectra of myoglobin. (**a**, **b**) The spectra were recorded within the *m*/*z* range of 500–2,000 in positive (**a**) and negative (**b**) ionization mode. (**c**, **d**) The spectra were recorded within the *m*/*z* range of 500–3,500 (**c**) and 1,500–3,500 (**d**), both in negative ionization mode

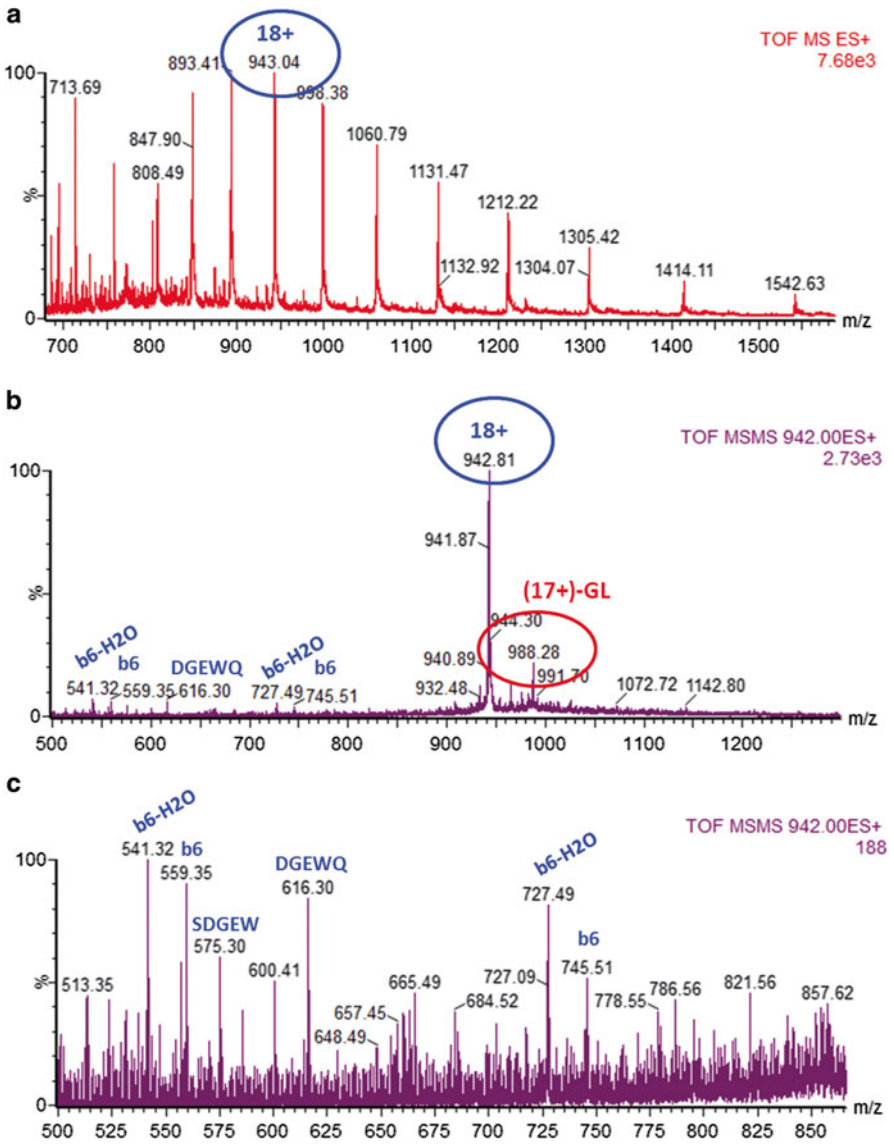


Fig. 6.13 ESI-MS and ESI-MS/MS of myoglobin. (**a**) ESI-MS spectrum of myoglobin. (**b**, **c**) The 18(+) ion with *m*/*z* 943.04 was selected for fragmentation and fragmented by ESI-MS/MS. Its fragmentation produced the 17(+) ion minus the first two N-terminal amino acids (GL) and a series of b and y ions that corresponded to the amino acid sequence from the N-terminal part of myoglobin

precursor ion with *m*/*z* of 943.04, which is an 18(+) ion. The ion with *m*/*z* of 988.28 corresponds to the 17(+) precursor ion, but without the first two amino acids from the N-terminus (GL). The theoretical mass of myoglobin is 16,951.48 Da and the loss of a GL dipeptide would leave a mass of 16,781.27, an almost a perfect match

for the 17(+) ion with *m*/*z* of 988.28. Further evidence that support the statement that the N-terminus of the myoglobin is labile/flexible come from identification of the b6 (*m*/*z* 559.35), b6-H_2O (*m*/*z* 541.2), b7 (*m*/*z* 745.51), b7-H_2O (*m*/*z* 727.3) and of several internal N-terminal fragments (*m*/*z*/616.23 which corresponds to DGEWQ or *m*/*z* 575.20 which corresponds to SDGEW), but of almost no C-terminal fragments (just y8 with *m*/*z* of 778.55).

6.3 Concluding Remarks

Structural mass spectrometry [32, 71] has now become an established experimental technique for the analysis of complex macromolecular structures and assemblies. The structures and subunit-interactions obtained through such experiments can be further detailed and substantiated with the aid of various techniques of structural biology [72]. In this regard, drug discovery, metabolism, and pharmacokinetics likely represent the most frequently explored fields where MS techniques are systematically coupled with computational structural biology.

Acknowledgements This work was supported in part by the Keep a Breast Foundation (KEABF-375-35054), the Redcay Foundation (SUNY Plattsburgh), the Alexander von Humboldt Foundation, SciFund Challenge, private donations (Ms. Mary Stewart Joyce & Mr. Kenneth Sandler), the David A. Walsh fellowship, and by the U.S. Army research office (DURIP grant #W911NF-11-1-0304).

References

1. Forster MJ (2002) Molecular modelling in structural biology. Micron 33(4):365–384
2. Alber F, Eswar N, Sali A (2004) Structure determination of macromolecular complexes by experiment and computation. In: Bujnicki J (ed) Practical bioinformatics. Springer, Berlin, Heidelberg, pp 73–96
3. Judge PJ, Watts A (2011) Recent contributions from solid-state NMR to the understanding of membrane protein structure and function. Curr Opin Chem Biol 15(5):690–695
4. Tang L, Johnson JE (2002) Structural biology of viruses by the combination of electron cryomicroscopy and X-ray crystallography†. Biochemistry 41(39):11517–11524
5. Sanchez R, Sali A (1997) Evaluation of comparative protein structure modeling by MODELLER-3. Proteins suppl 1:50–58
6. Guex N, Peitsch MC (1997) SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. Electrophoresis 18:2714–2723
7. Liu H-L, Hsu J-P (2005) Recent developments in structural proteomics for protein structure determination. Proteomics 5(8):2056–2068
8. Simons KT, Strauss C, Baker D (2001) Prospects for ab initio protein structural genomics. J Mol Biol 306(5):1191–1199
9. Bonneau R, Baker D (2001) Ab initio protein structure prediction: progress and prospects. Annu Rev Biophys Biomol Struct 30:173–189
10. Baker D, Sali A (2001) Protein structure prediction and structural genomics. Science 294(5540):93–96

11. Karplus M, Kuriyan J (2005) Molecular dynamics and protein function. Proc Natl Acad Sci U S A 102(19):6679–6685
12. Raval A et al (2012) Refinement of protein structure homology models via long, all-atom molecular dynamics simulations. Proteins 80(8):2071–2079
13. Chen J, Brooks CL III (2007) Can molecular dynamics simulations provide high-resolution refinement of protein structure? Proteins 67(4):922–930
14. Šali A, Kuriyan J (1999) Challenges at the frontiers of structural biology. Trends Cell Biol 9(12):M20–M24
15. Morin A, Sliz P (2013) Structural biology computing: lessons for the biomedical research sciences. Biopolymers 99(11):809–816
16. Moult J et al (2011) Critical assessment of methods of protein structure prediction (CASP)—round IX. Proteins Struct Funct Bioinform 79(S10):1–5
17. Wishart DS (2005) Bioinformatics in drug development and assessment. Drug Metab Rev 37(2):279–310
18. Aebersold R et al (2012) The biology/disease-driven human proteome project (B/D-HPP): enabling protein research for the life sciences community. J Proteome Res 12(1):23–27
19. Congreve M, Murray CW, Blundell TL (2005) Keynote review: structural biology and drug discovery. Drug Discov Today 10(13):895–907
20. Ngounou Wetie AG et al (2013) Identification of post-translational modifications by mass spectrometry. Aust J Chem 66:734–748
21. Ngounou Wetie AG et al (2014) Protein-protein interactions: switch from classical methods to proteomics and bioinformatics-based approaches. Cell Mol Life Sci 71(2):205–228
22. Ngounou Wetie AG et al (2013) Automated mass spectrometry-based functional assay for the routine analysis of the secretome. J Lab Autom 18(1):19–29
23. Sokolowska I et al (2012) Automatic determination of disulfide bridges in proteins. J Lab Autom 17(6):408–416
24. Sokolowska I et al (2013) Characterization of tumor differentiation factor (TDF) and its receptor (TDF-R). Cell Mol Life Sci 70(16):2835–2848
25. Sokolowska I et al (2012) Identification of a potential tumor differentiation factor receptor candidate in prostate cancer cells. FEBS J 279(14):2579–2594
26. Sokolowska I et al (2011) Mass spectrometry for proteomics-based investigation of oxidative stress and heat shock proteins. In: Andreescu S, Hepel M (eds) Oxidative stress: diagnostics, prevention, and therapy. American Chemical Society, Washington, DC
27. Woods AG et al (2012) Potential biomarkers in psychiatry: focus on the cholesterol system. J Cell Mol Med 16(6):1184–1195
28. Woods AG et al (2011) Blue native page and mass spectrometry as an approach for the investigation of stable and transient protein-protein interactions. In: Andreescu S, Hepel M (eds) Oxidative stress: diagnostics, prevention, and therapy. American Chemical Society, Washington, DC
29. Karas M, Hillenkamp F (1988) Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. Anal Chem 60(20):2299–2301
30. Fenn JB et al (1989) Electrospray ionization for mass spectrometry of large biomolecules. Science 246(4926):64–71
31. Sokolowska I et al (2012) Identification of potential tumor differentiation factor (TDF) receptor from steroid-responsive and steroid-resistant breast cancer cells. J Biol Chem 287(3): 1719–1733
32. Dyachenko A et al (2013) Allosteric mechanisms can be distinguished using structural mass spectrometry. Proc Natl Acad Sci U S A 110(18):7235–7239
33. Sokolowska I et al (2013) Characterization of tumor differentiation factor (TDF) and its receptor (TDF-R). Cell Mol Life Sci 70(16):2835–2848
34. Ngounou Wetie AG et al (2013) Investigation of stable and transient protein–protein interactions: past, present, and future. Proteomics 13(3–4):538–557
35. Sokolowska I et al (2013) Applications of mass spectrometry in proteomics. Aust J Chem 66(7):721–733

36. Darie CC (2013) Mass spectrometry and its applications in life sciences. Aust J Chem 66(7): 719–720
37. Florian PE et al (2013) Characterization of the anti-HBV activity of HLP1-23, a human lactoferrin-derived peptide. J Med Virol 85(5):780–788
38. Petrareanu C et al (2013) Comparative proteomics reveals novel components at the plasma membrane of differentiated HepaRG cells and different distribution in hepatocyte-and biliary-like cells. PLoS One 8(8):e71859
39. Sokolowska I et al (2012) Proteomic analysis of plasma membranes isolated from undifferentiated and differentiated HepaRG cells. Proteome Sci 10(1):47
40. Jurneczko E, Barran PE (2011) How useful is ion mobility mass spectrometry for structural biology? The relationship between protein crystal structures and their collision cross sections in the gas phase. Analyst 136(1):20–28
41. Benesch JLP, Ruotolo BT (2011) Mass spectrometry: come of age for structural and dynamical biology. Curr Opin Struct Biol 21(5):641–649
42. Walzthoeni T et al (2013) Mass spectrometry supported determination of protein complex structure. Curr Opin Struct Biol 23(2):252–260
43. Wintrode P (2013) Mass spectrometry in structural biology. Biochim Biophys-Acta Proteins Proteomics 1834(6):1187
44. Dunham WH, Mullin M, Gingras AC (2012) Affinity-purification coupled to mass spectrometry: basic principles and strategies. Proteomics 12(10):1576–1590
45. Kaltashov IA, Bobst CE, Abzalimov RR (2009) H/D exchange and mass spectrometry in the studies of protein conformation and dynamics: is there a need for a top-down approach? Anal Chem 81(19):7892–7899
46. Konijnenberg A, Butterer A, Sobott F (2013) Native ion mobility-mass spectrometry and related methods in structural biology. Biochim Biophys Acta 1834(6):1239–1256
47. Chalmers MJ et al (2011) Differential hydrogen/deuterium exchange mass spectrometry analysis of protein-ligand interactions. Expert Rev Proteomics 8(1):43–59
48. Kan Z-Y et al (2013) Protein hydrogen exchange at residue resolution by proteolytic fragmentation mass spectrometry analysis. Proc Natl Acad Sci 110(41):16438–16443
49. Leitner A et al (2012) The molecular architecture of the eukaryotic chaperonin TRiC/CCT. Structure 20(5):814–825
50. Kalisman N, Adams CM, Levitt M (2012) Subunit order of eukaryotic TRiC/CCT chaperonin by cross-linking, mass spectrometry, and combinatorial homology modeling. Proc Natl Acad Sci U S A 109(8):2884–2889
51. Barrera NP et al (2009) Mass spectrometry of membrane transporters reveals subunit stoichiometry and interactions. Nat Methods 6(8):585–587
52. Schey KL, Grey AC, Nicklay JJ (2013) Mass spectrometry of membrane proteins: a focus on aquaporins. Biochemistry 52(22):3807–3817
53. Laganowsky A et al (2013) Mass spectrometry of intact membrane protein complexes. Nat Protocols 8(4):639–651
54. Roy U et al (2012) Structural investigation of tumor differentiation factor. Biotechnol Appl Biochem 59(6):445–450
55. Roy U et al (2013) Tumor differentiation factor (TDF) and its receptor (TDF-R): is TDF-R an inducible complex with multiple docking sites? Mod Chem Appl 1(3):108
56. Roy U et al (2013) Structural evaluation and analyses of tumor differentiation factor. Protein J 32(7):512–518
57. Woods AG et al (2013) Identification of tumor differentiation factor (TDF) in select CNS neurons. Brain Struct Funct DOI: 10.1007/s00429-013-0571-1
58. Cheetham JC, Artymiuk PJ, Phillips DC (1992) Refinement of an enzyme complex with inhibitor bound at partial occupancy. Hen egg-white lysozyme and tri-N-acetylchitotriose at 1.75 A resolution. J Mol Biol 224(3):613–628
59. Gordon JC et al (2005) H++: a server for estimating pKas and adding missing hydrogens to macromolecules. Nucleic Acids Res 33:W368–W371

60. Myers J et al (2006) A simple clustering algorithm can be accurate enough for use in calculations of pKs in macromolecules. Proteins 63:928–938
61. Anandakrishnan R, Aguilar B, Onufriev AV (2012) H++ 3.0: automating pK prediction and the preparation of biomolecular structures for atomistic molecular modeling and simulation. Nucleic Acids Res 40(W1):W537–W541
62. Guex N, Peitsch MC (1996) Swiss-PdbViewer: a fast and easy-to-use PDB viewer for Macintosh and PC. Protein Data Bank Q Newslett 77:7
63. Mehler EL, Guarnieri F (1999) A self-consistent, microenvironment modulated screened coulomb potential approximation to calculate pH-dependent electrostatic effects in proteins. Biophys J 75:3–22
64. Kendrew JC et al (1958) A three-dimensional model of the myoglobin molecule obtained by x-ray analysis. Nature 181(4610):662–666
65. Yang F, Phillips GN Jr (1996) Crystal structures of CO-, deoxy- and met-myoglobins at various pH values. J Mol Biol 256(4):762–774
66. Urayama P, Phillips GN Jr, Gruner SM (2002) Probing substates in sperm whale myoglobin using high-pressure crystallography. Structure 10(1):51–60
67. Sage JT et al (1992) Low pH myoglobin photoproducts. Biophys J 61(4):1041–1044
68. Duprat AF et al (1995) Myoglobin-NO at low pH: free four-coordinated heme in the protein pocket. Biochemistry 34(8):2634–2644
69. Iben IE et al (1991) Carboxy Mb at pH 3. Time-resolved resonance Raman study at cryogenic temperatures. Biophys J 59(4):908–919
70. Han S et al (1990) Metastable intermediates in myoglobin at low pH. Proc Natl Acad Sci U S A 87(1):205–209
71. Kirshenbaum N, Michaelevski I, Sharon M (2010) Analyzing large protein complexes by structural mass spectrometry. J Vis Exp 40:1954
72. Sharon M (2013) Structural MS pulls its weight. Science 340(6136):1059–1060

Chapter 7
Affinity-Mass Spectrometry Approaches for Elucidating Structures and Interactions of Protein–Ligand Complexes

Brînduşa Alina Petre

Abstract Affinity-based approaches in combination with mass spectrometry for molecular structure identification in biological complexes such as protein–protein, and protein–carbohydrate complexes have become popular in recent years. Affinity-mass spectrometry involves immobilization of a biomolecule on a chemically activated support, affinity binding of ligand(s), dissociation of the complex, and mass spectrometric analysis of the bound fraction. In this chapter the affinity-mass spectrometric methodologies will be presented for (1) identification of the epitope structures in the Abeta amyloid peptide, (2) identification of oxidative modifications in proteins such as nitration of tyrosine, (3) determination of carbohydrate recognition domains, and as (4) development of a biosensor chip-based mass spectrometric system for concomitant quantification and identification of protein–ligand complexes.

Abbreviations

ACN	Acetonitrile
AD	Alzheimer's disease
APP	Amyloid precursor protein
CE	Equivalent carbon
CRDs	Carbohydrate recognition domains

B.A. Petre (✉)
Laboratory of Biochemistry, Department of Chemistry, Al. I. Cuza University of Iasi, Carol I Boulevard, No. 11, 700506 Iasi, Romania

Laboratory of Analytical Chemistry and Biopolymer Structure Analysis, Department of Chemistry, University of Konstanz, 78457 Constance, Germany
e-mail: brindusa.petre@uaic.ro

A.G. Woods and C.C. Darie (eds.), *Advancements of Mass Spectrometry in Biomedical Research*, Advances in Experimental Medicine and Biology 806,
DOI 10.1007/978-3-319-06068-2_7, © Springer International Publishing Switzerland 2014

CREDEX	*C*arbohydrate *re*cognition *d*omain *ex*cision
2-DE	Two-dimensional electrophoresis
DHB	2,4-Dihydroxybenzoic acid
ECP	Eosinophil cationic protein
ELISA	Enzyme-linked immunosorbent assay
ESI	Electrospray ionization
FTICR	Fourier transform ion cyclotron resonance
HCCA	α-Cyano-4-hydroxy-cinnamic acid
HPLC	High-performance liquid chromatography
LC	Liquid chromatography
MALDI	Matrix-assisted laser desorption/ionization
MS	Mass spectrometry
3NT	3 Nitrotyrosine
PROFINEX	*Pro*teolytic af*fin*ity *ex*traction
SAS	Solvent accessible surface
SAW	Surface acoustic wave
TFA	Trifluoroacetic acid
ToF	Time of flight

7.1 Introduction

Proteins have the ability to specifically interact with one or more ligands. Protein–ligand complexes such as protein–protein, protein–carbohydrates, and protein–DNA play an essential role in a multitude of physiological and pathophysiological cellular processes. Determination of protein–ligand-binding affinity is crucial for the mechanistic understanding of protein's function and for the development of new biochemical, analytical, or biomedical applications.

Traditionally, the structure of proteins in complex with their ligand(s) have been determined by X-ray crystallography, nuclear magnetic resonance (NMR), fluorescence and IR spectroscopy, and surface plasmon resonance [11, 26, 31, 52]. However, these methods have a number of limitations such as: (1) they require relatively large amounts of pure protein, (2) many of these studies were performed under nonequilibrium conditions, and (3) they are time consuming, through elaborate and lengthy experimental protocols [52]. Also, the specificities of protein/peptide antigen–antibody interactions were used either to obtain purified proteins from complex mixtures (such as serum or cell lysates) or for antibody purification. The former is typically performed using monoclonal or polyclonal antibodies attached covalently to resins [49], while the latter is based on immunoglobulin-specific protein A or protein G affinity columns [10]. Immunoaffinity techniques such as Western blot [15, 21], immunohistochemistry [45], and ELISA [17, 23] rely on the use of antibodies of various specificities to detect antigen–antibody interactions, providing information about the overall binding and strength of the interaction. However, no details about the chemical structures of antibody–antigen binding can be determined.

Mass spectrometric methods can provide structural details on many levels for different classes of biomolecules; proteins can now be analyzed by mass spectrometry to reveal complete or partial amino acid sequences, posttranslational modifications, protein–protein interaction structure sites, and even provide insight into higher order structure [44, 55]. One key advantage of mass spectrometric measurements is its capability to analyze extremely small quantities of sample with high sensitivity. The "soft" ionization/desorption techniques such as electrospray (ESI) and matrix-assisted laser desorption/ionization (MALDI) were major tools used to developed affinity-mass spectrometric methods that provide key data to identify the specific interacting structures within protein–ligand interactions.

The combination of affinity-mass spectrometry (affinity-MS) was shown in the last years, as an established methodology for (1) direct protein identification from complex biological material [22], (2) identifications of epitope [9, 43, 46, 56], paratope [42], and peptibody [41, 54], and (3) the identification of carbohydrate recognition domains (CRDs) [29]. These results were possible by developing and applying new approaches of affinity-MS in combination with selective proteolytic digestion (epitope excision) and affinity selection of proteolytic generated fragments (epitope extraction) as shown in Fig. 7.1. The development of these two affinity approaches were based on the hypothesis that (1) in epitope excision, an antibody will protect the binding site(s) of a bound protein or peptide antigen from proteolytic cleavage, and (2) in epitope extraction, an antibody will bind a minimum epitope sequence generated by the proteolytic cleavage of the antigen molecule in solution. The epitope excision (Fig. 7.1a) consists in immobilizing the antibody on a Sepharose material in a micro-column and binding the antigen peptide or protein to this column. Then, a specific protease is added to the column, and the proteolytic digestion is performed. The non-bound fragments are washed away, and the epitope–antibody complex is dissociated by using acidic condition (elution). In the epitope extraction (Fig. 7.1b), the antigen molecule is first proteolytically digested in solution, and the resulted fragments are added on the antibody column and allowed to bind. Again, the non-bound fragments are washed away, and the epitope–antibody complex is eluted by using acidic condition. The last washing fraction and elution fraction are analyzed by mass spectrometry for the identification of epitope structure.

The direct, chemical epitope identification approach using epitope excision/extraction-mass spectrometry provided differential information about the epitope structures by comparison of shielding of specific proteolytic cleavage sites comprised in the epitope peptide; information that cannot be directly determined by affinity techniques such as Western blot, quantitative ELISA, and SPR. However, the epitope identity can be readily ascertained and verified by affinity studies using synthetic mutated peptides in affinity-mass spectrometric approaches.

This chapter describes recent progress made in the area of affinity-mass spectrometric-based approaches for (1) identification of epitope structures in Abeta amyloid peptide, (2) specific identification of oxidative modifications in proteins, (3) for determination of CRDs, and as (4) development of a biosensor chip-based system.

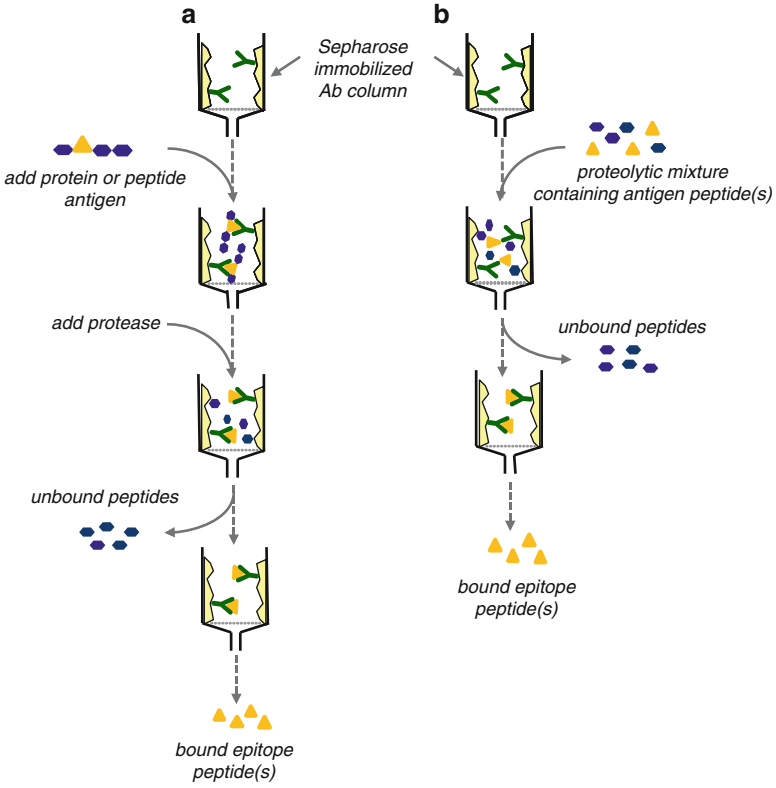


Fig. 7.1 Analytical scheme of affinity-mass spectrometric identification of antigen epitope(s). (**a**) In *epitope excision* the antigen is bound to the immobilized antibody column, and then proteolytical cleavage is performed using protease and (**b**) in *epitope extraction* the antigen is first degraded proteolytically in solution and then applied on to the immobilized antibody column. The resulted epitope containing fractions are analyzed by mass spectrometry

7.2 Materials and Methods

7.2.1 Immobilization of Antibodies to NHS-Activated Sepharose

For preparing the affinity column, an aliquot of 100 µg anti-Aβ(1-42) polyclonal antibody from TgCRND8 mouse or anti-3nitrotyrosine (3NT) monoclonal antibody (1 µg/µL) was mixed with coupling buffer (0.2 M $NaHCO_3$, 0.5 M NaCl, pH 8.3) and added to the appropriate amount of dry NHS-activated 6-aminohexanoic acid-coupled Sepharose (1 g Sepharose swelled in approximately 3 mL of coupling buffer). The coupling reaction was performed for 2 h at 25 °C under vigorous shaking. The reaction mixture was then loaded into a micro-column (MoBiTec, Göttingen, Germany) and extensively washed by alternating twice 10 mL blocking buffer

(0.1 M ethanolamine, 0.5 NaCl, pH 8.3), with 10 mL washing buffer (0.2 M NaOAc, 0.5 M NaCl, pH = 4.0). To block unreacted active groups, the affinity matrix was kept in blocking buffer for 1 h at room temperature; afterwards, the washing step was performed using 30 mL PBS buffer, pH = 7.4. The affinity column was stored in PBS buffer at 4 °C.

7.2.2 *Identification of the Abeta (1-40) Epitope by Proteolytic Epitope Excision/Extraction-Mass Spectrometry*

For *epitope excision* experiments 100 μg of antigen Aß(1-42) synthetic peptide were applied onto the immobilized antibody column. The micro-column was then gently shaken for 1 h to allow complete binding of antigen. The column was then washed with 20 mL PBS buffer for removal of unbound antigen peptide, and the remaining affinity-bound Aß(1-40) peptide was digested for 2 h at 37 °C by addition of 0.2 μg of protease (trypsin, or α-chymotrypsin and aminopeptidase M) in 200 μL PBS buffer. Supernatant non-epitope fragments were removed by washing with 20 mL PBS buffer. After removal of the unbound peptide fragments, the immune complex was dissociated from the immobilized antibody by addition of 500 μL 0.1 % TFA; the column was shaken gently for 15 min, and the released epitope peptides were collected into a micro-centrifuge tube. The samples were then lyophilized and stored until mass spectrometric analysis. The column was regenerated by washing with 10 mL 0.1 % TFA followed by 20 mL PBS buffer. When handled in this way, affinity micro-columns may be used at least 10–20 times without significant loss of antigen-binding capacity. With the immobilization and proteolytic digestion conditions employed, IgG antibodies generally are highly stable towards degradation, as established in previous studies [33, 43].

For the *epitope extraction* experiments, the antigen, Aß(1-40) synthetic peptide, was first digested with the chosen protease. After 2 h digestion time at 37 °C the resulted peptide mixture was added to the immobilized antibody column. The binding of the antigen peptide fragment to the antibody was performed for 1 h at RT. The antigen peptide retained on the column was eluted with 500 μL 0.1 % TFA. In all experiments the supernatant, the final milliliter of the washing fraction, and the elution fraction were lyophilized, desalted by Zip-Tip®, and analyzed by mass spectrometry.

7.2.3 *Proteolytic Affinity Extraction-Mass Spectrometry for Identification of Tyrosine Nitration Sites in Human Eosinophil Proteins*

For the proteolytic affinity peptide extraction, highly pure (>90 %) human eosinophil cationic protein (ECP) was denatured and digested in solution. The protein was dissolved first at a concentration of 1 μg/μL in 10 mM NH_4HCO_3 (pH 8) and then DTT in 10 mM NH_4HCO_3 (50-fold excess, relative to the number of

–S–S-bound in protein) was added. The reduction was carried out for 1 h at 56 °C under gentle shaking. After cooling to room temperature, a solution of IAA in MilliQ was added in 2.2-fold excess relating to DTT amount, and the reaction was performed for 45 min at ambient temperature in the dark with occasional vortexing. After lyophilization ECP was proteolytically digested by trypsin in 10 mM NH_4HCO_3, pH = 8. Trypsin was added to the sample at an enzyme to substrate ratio of 1:30 (w/w), and the digestion was carried out at 37 °C for 4 h and then quenched by freezing the sample with liquid nitrogen. The resulting peptide mixtures were analyzed directly by mass spectrometry or were used for immunoaffinity experiments. The resulted ECP peptides fragments were added onto the anti-3-NT-antibody column and incubated under gentle shaking for 2 h at room temperature. The non-bound peptides fragments (supernatant fraction) were removed by pushing air through the column using a 10 mL syringe. The matrix was subsequently washed with 200 mL PBS buffer for removal of unbound peptides. The immune complex between nitrated peptide and anti-3-NT antibody was dissociated by addition of 0.1 % TFA. The collected fractions were analyzed by MALDI-TOF and nano-ESI FT-ICR-MS. In a second experiment the elution fraction was analyzed by Edman sequencing.

7.2.4 Online SAW Biosensor–ESI-MS Coupling

A Bruker Esquire 3,000+ ion trap mass spectrometer (Bruker Daltonik, Bremen, Germany) was coupled online with an S-Sens K5 biosensor (Biosensor GmbH, Bonn, Germany) via an manual interface. The interface developed for online coupling of surface acoustic wave (SAW) biosensor with an ESI-ion trap mass spectrometer utilizes a six-port valve micro-column and micro-injector for desalting and in situ concentration of protein samples dissociated from the protein–ligand complex as illustrated in Fig. 7.6. The C_{18}-reversed phase guard column coupled to the microinjector unit was from Rheodyne (California, USA). The antibody was immobilized on carboxyl self-assembled monolayer (SAM) surface via available amide bond (N-terminal or lysine residues). SAM is formed first by activation of the carboxyl group with 200 mM *N*-(3-dimethylaminopropyl)-*N*-ethylcarbodiimide (EDC) and second by coupling reaction using 50 mM *N*-hydroxysuccinimide (NHS). A solution of 300 nM of antibody was immobilized on the preactivated-SAM surface, followed with capping of non-reacted NHS groups with 1 M ethanolamine pH 8.5. A 10 μM solution of peptide ligands mixture was added, and affinity binding was performed and recorded at a flow rate of 20 μL/min. All affinity binding experiments were performed at 20 °C in PBS binding buffer, pH 7.5. Following association of ligand(s), elution was carried out with glycine buffer, pH 2, cleaning of buffer salts and transfer of the eluted compound into the ESI source were performed using 0.3 % aqueous HCOOH (desalting step; approx. 300 μL at 70 μL/min) and with 0.3 % HCOOH/80 % acetonitrile (elution step; 70 μL/min), respectively, by switching of the microvalve injector. Determinations of dissociation constants were typically performed after

immobilization of peptide using a 10 μM solution on the activated SAM surface at a flow rate of 20 μL/min. The flow was then changed to PBS buffer, pH 7.5, and a solution of 350 nM anti-lysozyme antibody was added to block all unspecific binding sites, followed by regeneration with 150 μL glycine buffer. Association kinetics were determined with serial dilutions of 8–128 nM of antibody, followed with regeneration by glycine buffer, pH 2, and concentrations and volumes of ligands were optimized to obtain highest binding curves. Determination of antibody/protein association kinetics was performed by extracting the data from the sensor signals for all concentrations employed, using the 1:1 monomolecular growth model. Determination of the dissociation constant is provided after plotting the pseudo-first kinetic constant versus concentration, and applying a linear regression by $Kd = koff \times kon^{-1}$.

7.2.5 *Proteolytic Excision/Extraction-Mass Spectrometry of Ligand-Binding Peptides in Human Galectins-1 and Galectins-3*

The covalent coupling of lactose to divinylsulfone-activated Sepharose 4B was performed as previously described [6]. For proteolytic CRDs' *peptide excision*, a solution of 50 μg human galectin-1 or galectin-3 in 100 μL 50 mM PBS buffer, pH 7.5, was added to 200 μL affinity matrix and allowed to bind at 37 °C for 24 h. Any remaining unbound galectin was washed out with 30 mL binding buffer, and the washing fractions were collected and analyzed by MALDI-TOF-MS in order to determine the extent of binding. The digestion of bound galectin was performed using trypsin, in the binding buffer, with an enzyme to substrate ratio of 1:100, at 37 °C, for 3 h. After digestion the supernatant was analyzed by MALDI-FTICR-MS. The free tryptic peptides (not bound to the carbohydrate) were washed out with 30 mL PBS buffer, pH 7.5. The washing fractions were analyzed by MALDI-FTICR-MS. The affinity-bound peptides were eluted under strong shaking with 400 μL 0.3 M lactose in PBS buffer, pH 7.5, at 37 °C for 15 min, and this procedure was repeated twice. For galectin-1 protection of Cys residues from oxidation was performed using 0.6 mg DTT in the binding buffer and 1 g/L DTT in the washing buffer.

In a second experimental approach, proteolytic CRDs *peptide extraction*, 50 μg human galectin-1 or galectin-3 was digested in solution using trypsin (enzyme: substrate ratio, 1:20) for 5 h at 37 °C in 200 μL PBS buffer, 50 mM at pH 7.5. A sample aliquot from the digestion mixture was analyzed by MALDI-MS. The rest of the digest was added over an affinity column containing 200 μL matrix. After incubation for 12 h at 37 °C, the supernatant was removed and analyzed by MALDI-MS. Any traces of free peptides were washed out from the column with 10 mL PBS buffer, pH 7.5, and the washing fraction was analyzed by MALDI-FTICR-MS. The peptides bound to lactose were eluted under shaking with 400 μL ACN: 0.1 % TFA 2:1 at 37 °C for 15 min, repeating the procedure twice.

7.2.6 *Mass Spectrometry*

MALDI-ToF-MS was performed with a Bruker Biflex™ linear TOF mass spectrometer (Bruker Daltonics, Bremen, Germany) equipped with a nitrogen UV laser (337 nm) and a dual channel plate detector. A saturated solution of α-cyano-4-hydroxy-cinnamic acid (HCCA) in ACN: 0.1 % TFA (2:1) was used as matrix. Acquisition of spectra was carried out at an acceleration voltage of 20 kV and a detector voltage of 1.5 kV.

High-resolution FTICR-MS was performed with a Bruker Daltonics Apex II instrument equipped with a 7 T superconducting magnet, a cylindrical infinity ICR analyzer cell, and a Scout-100 MALDI or a nano-ESI ion source. A 50 mg/mL solution of 2,4-dihydroxybenzoic acid (DHB) in ACN: 0.1 % TFA (2:1) was used as matrix for MALDI-FTICR MS. Nano-ESI-FTICR mass spectra were obtained by accumulation of 15 single scans, with the capillary exit voltage set to 20 V and the skimmer 1 set to 10 V, while the capillary voltage was adjusted between −1,100 and −1,200 V until a stable spray was obtained. The ions were accumulated in the RF-only hexapole for 0.15 s before being transferred into the ICR cell. Calibration, acquisition, and processing of spectra were carried out with the Bruker XMASS Software.

7.3 Results and Discussions

7.3.1 *Affinity-Mass Spectrometric Approaches for Elucidation of a β-Amyloid-Plaque-Specific Epitope*

Amyloid-ß-peptide (Aß) fragment(s) of the amyloid precursor protein (APP) is the major peptide(s) constituent of senile amyloid plaques in brains, which represent hallmark of Alzheimer's disease (AD). The formation of neurotoxic oligomeric soluble intermediates precedes aggregation and is critical to overall neurotoxicity and progressive neurodegeneration [30]. Although Aß has been studied thoroughly, molecular details of pathophysiological degradation, chemical structure modification, Aß aggregation mechanism, and its cellular interactions remain unclear [2, 20]. Understanding pathophysiological amyloid-ß-peptide aggregation and its interactions with ligands such as antibodies or protective proteins is a key prerequisite for the development of drugs capable of (1) neutralizing or disaggregating amyloid-ß aggregates or (2) inhibiting the initial step(s) of aggregate formation.

Highly efficient affinity-MS methods such as proteolytic epitope excision and proteolytic epitope extraction have been developed and applied to elucidate the molecular mechanism of immunotherapeutic approaches involving Aß "plaque-specific" antibodies. These are produced in transgenic AD mouse models upon active immunization with Aß(1-42) and were found to promote disaggregation of Aß plaques and fibrils and to improve the memory impairment in AD mice [25, 27].

Epitope excision/extraction experiments were performed using the immobilized polyclonal antibody from TgCRND8 mouse (see Sect. 2.1) [14]. The antigen peptide Aß(1-40) was bound to the antibody column; the immuno-Aß(1-40) complex was allowed to form, and then trypsin was added onto the column as described in Sect. 2.2. After removal of the unbound fragments generated by proteolysis, the immune complex was dissociated by addition of 0.1 % TFA. In the MALDI-FTICR mass spectrum (Fig. 7.2) of the supernatant fraction (A), all expected tryptic fragments of ß-amyloid Aß(1-40) peptide were present. After washing off the unbound fragments, in the spectrum of the elution fraction (B), only a single peak corresponding to the Aß peptide [1-16] fragment was identified. This fragment was the only one which remained bound to the polyclonal antibody and proved that the epitope recognized by the polyclonal antibody from TgCRND8 transgenic mouse was in this N-terminal region of the peptide.

For epitope extraction, the Aß(1-40) was first digested using trypsin, and the tryptic peptide mixture was added on the same antibody column upon regeneration. The extraction experiment led to similar results, as the Aß(1-16) tryptic peptide was identified in the elution fraction. Therefore, further experiments were performed in order to determine in detail the minimum structure of the Aß(1-40) epitope.

A second epitope excision experiment with Aß(1-40) was carried out as described before but using endoprotease GluC instead of trypsin. This protease was chosen because it cleaves the peptide bonds after glutamic (E) and aspartic (D) acid residues, enabling determination of shorter peptide sequences if these still remain bound to the antibody. Analyzing the elution fractions by high-resolution MALDI-FTICR MS, only a single peptide fragment corresponding to the amino acid sequence Aβ (4-FRHDSGYE-11) was found bound to the polyclonal antibody column. A more detailed epitope elucidation experiment was carried out using a combination of α-chymotrypsin and aminopeptidase M. Amino peptidase M is an exopeptidase that cleaves all amino acids starting from the N-terminal part of the antigen sequence. For the epitope excision experiments, the antigen peptide, Aß(1-40), was bound to the polyclonal anti-Aß(1-42) antibody column, digested with α-chymotrypsin and afterwards with aminopeptidase M with washing steps in between. The MALDI-FTICR mass spectra of the supernatant and elution fractions are shown in Fig. 7.3. α-Chymotrypsin digestion of Aß(1-40) has generated the fragment (1-10), which was further digested by aminopeptidase M.

From the MALDI-FTICR mass spectrum of the elution fraction, the smallest fragment bound to the polyclonal anti-Aß(1-42) antibody column was Aß(4-10) peptide. This represents the minimal structure of the ß-amyloid peptide recognized by the polyclonal antibody. The specific N-terminal amino acid sequence Aß (4-FRHDSGY-10) was identified with high mass accuracy as the plaque-specific epitope, providing a basis for the development of new lead structures for AD vaccine development [39, 42]. The identification and quantification of Aß peptides bound to specific anti-Aß antibodies were provided by the newly introduced online bioaffinity-electrospray mass spectrometry [5]. In several further studies, proteolytic excision/extraction-mass spectrometry were successfully used for the identification of interacting Aß epitope sequences with aggregation-inhibiting polypeptides

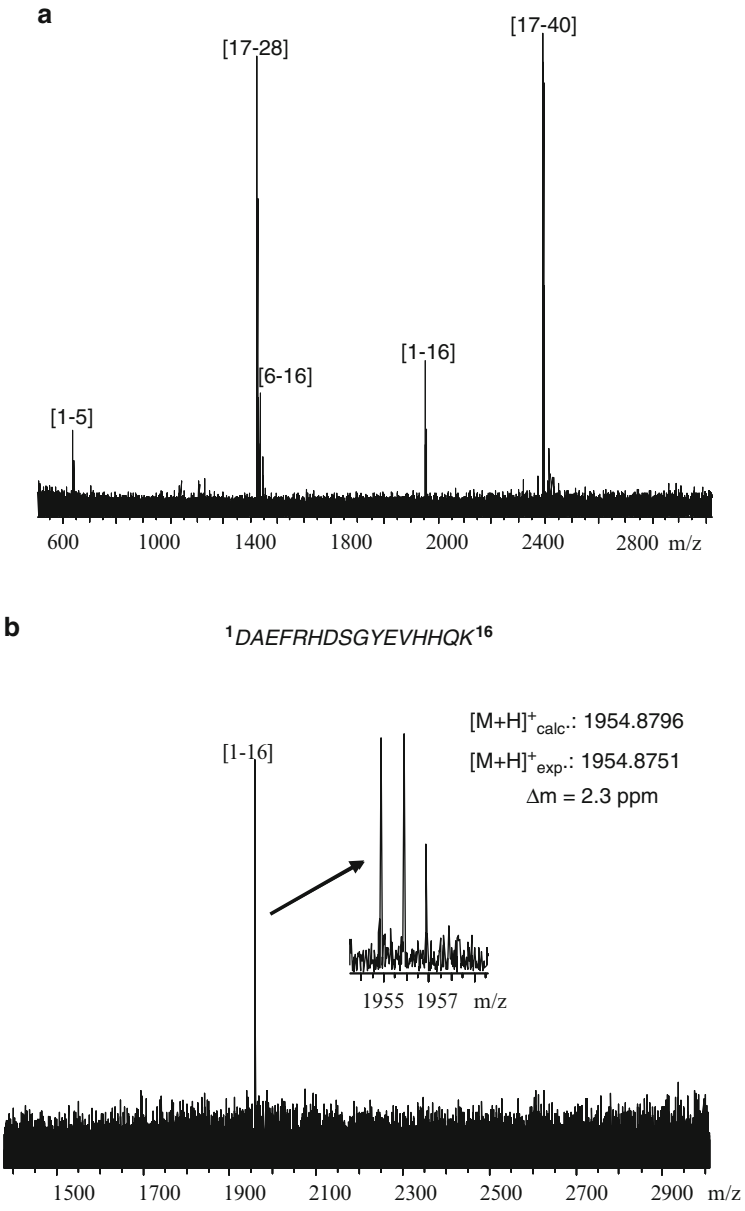


Fig. 7.2 Mass spectrometric epitope excision of Aß(1-40): (**a**) MALDI-FTICR MS of the supernatant fraction of Aß(1-40) after trypsin digestion covering the whole peptide sequence. (**b**) MALDI-FTICR MS of the elution fraction reveals only one Aß peptide fragment (1-16). Reprinted with author permission from [14]

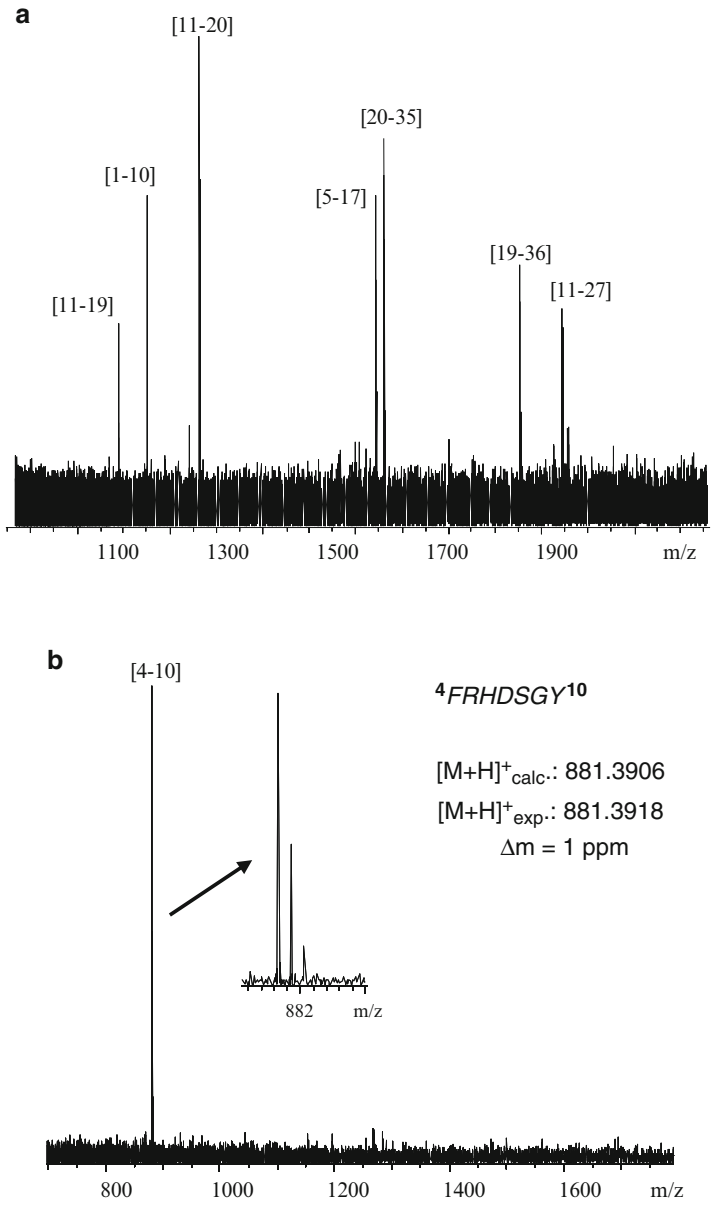


Fig. 7.3 Mass spectrometric epitope excision of Aß(1-40) using α-chymotrypsin and aminopeptidase M. (**a**) MALDI-FTICR MS of the supernatant fraction and (**b**) MALDI-FTICR MS of the elution fraction reveal minimum peptide fragment Aß (4-10) which remains bound to the antibody column. Reprinted with author permission from [14]

such as cystatin C [16], humanin [24], and Aß-nanobodies (single chain llama anti-ß-amyloid antibodies) providing also a potential for developing new diagnostic tools for AD as well as for designing new potential immunotherapeutic agents [34].

7.3.2 *Proteolytic Affinity Extraction-Mass Spectrometric Approach for Identification of Tyrosine Nitration Sites*

Tyrosine nitration represents an oxidative modification in proteins. It has been shown to occur under physiologic conditions, but found to be substantially enhanced under various pathophysiological conditions such as atherosclerosis, asthma and lung diseases, neurodegenerative diseases, and diabetes [2, 7]. For the molecular correlation of protein nitration with pathogenic mechanisms of human diseases and with animal or cellular models of diseases, it is essential to identify the protein targets of nitration and the individual modification sites. Most identifications of protein nitrations have been obtained using immunoanalytical methods such as Western blot, ELISA, and immuno-electron microscopy, employing different 3-NT-antibodies. These antibodies were poorly characterized regarding their specificities, providing only the overall detection of nitrations, and no identification of specific nitration sites and structures [12, 13]. A combination of electrophoresis, Western blotting, and mass spectrometry has been applied in several proteomics studies of nitrated proteins from biological material [1, 2, 8]. However, specific 3-nitrotyrosine modification sites were not identified, rendering the identification of nitration sites critically dependent on the specificity of NT-antibodies. The failure of many mass spectrometric-based approaches to characterize nitrated proteins may be due to multiple causes such as (1) low abundance of NT-containing proteins/peptides, (2) solubility problems, hydrophobicity, and/or extreme pI values of proteins, which may compromise the isoelectric-focusing in 2-DE separations, (3) insufficient recovery of NT-containing peptides from gels and/or HPLC columns during LC-MS analysis, and (4) instability of nitrated peptides due to photochemical decomposition of nitro-tyrosine residues under UV-MALDI-MS analysis condition [1, 2, 38]. The problems of identification of tyrosine nitration at biologically relevant levels in nitrated proteins have been overcome by the recent development of a proteolytic affinity extraction-MS "PROFINEX" [37], analogue to epitope extraction procedure described above (Sects. 2.2 and 3.1). The efficiency of the PROFINEX approach is shown here in the example of human ECP isolated from patients with abnormal elevate number of eosinophils in blood [47]. First ECP was denatured and digested in solution, and the resulted tryptic peptide mixture was added to the immobilized 3NT-antibody column. The 3NT-antibody–ECP peptide(s) complex was allowed to form, and unbound ECP tryptic peptides were removed by washing. Due to the high specificity of the 3NT-antibody for the antigen, only peptides containing the antigenic determinant (the 3-nitro-tyrosine) should interact with the antibody. Antibody-bound ECP peptide was eluted by dissociating the immune complex under slightly acidic condition (pH = 2.5). The supernatant, wash, (last mL of wash) and elution fractions were collected, lyophilized, and analyzed by mass

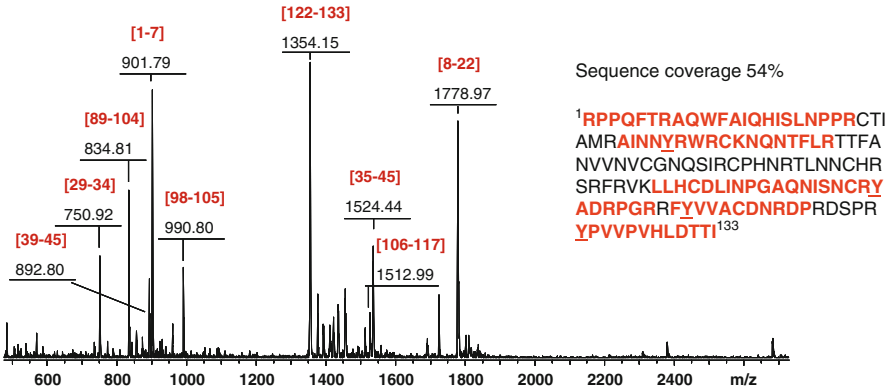


Fig. 7.4 Proteolytic affinity extraction-mass spectrometric approach. MALDI-TOF mass spectrum of unbound ECP tryptic peptides contained in the supernatant fraction. All four tyrosine residues in eosinophil cationic protein were found to be unmodified

spectrometry after desalting by C_{18} ZipTip. The unbound tryptic peptide contained in the supernatant fraction was analyzed by MALDI-TOF (Fig. 7.4). All four tyrosine residues contained in ECP were found unmodified. The identified peptides accounted for 54 % sequence coverage in the supernatant fraction.

The unequivocal identification of the peptide containing nitrated Tyr was obtained by nano-ESI-FT-ICR analysis of the eluate (Fig. 7.5). The spectrum contained a doubly protonated molecular ion at *m/z* 764.3618, corresponding to the ECP peptide ^{23}CTIAMRAINNY(NO_2)R^{34}, and nitrated at Tyr^{33} residue. The monoisotopic mass of the singly charged molecular ion was determined after deconvolution of the ESI mass spectrum using the XMASS software and was assigned to ECP peptide fragment (23-34) with a mass difference of 103 Da (45 Da for the nitro group and 58 Da for Cys^{23}—carbamidomethyl formation).

The missed cleavage at Arg^{28} indicates that this residue is part of an "epitope" structure within ECP only upon nitration of Tyr^{33}, because the same arginine residue was cleaved by trypsin in the non-nitrated Tyr^{33} containing peptide (Fig. 7.4). A spatial orientation of tyrosine residues in ECP was illustrated using the crystal structure available in the UniProt database (PBD accession number 1H1H) [28]. Using the molecular modeling software BallView 1.1.1, the secondary ribbon structure was rendered and is shown in Fig. 7.5. The location of Tyr^{33} within a loop structure is indicated by yellow, and all other three tyrosine residues were colored. The solvent accessible surface (SAS) defined by using this molecular modeling program showed that Tyr^{33} residue of ECP has the highest surface accessibility (94.4 Å^2), while the other three Tyr residues are embedded in the tertiary structure of the molecule. Moreover, the hydroxyl group of Tyr^{33} and the two equivalent *ortho* carbons CE_1 and CE_2 extremely protrude in the exterior space. These are critical for allowing tyrosine to be modified by nitration. The molecular SAS model shows that Tyr^{98} is completely buried in the protein core, and its hydroxyl group is not exposed to the surface, while for Try^{107} and Try^{122} the hydroxyl groups and the nearby carbons appear less surface exposed [35].

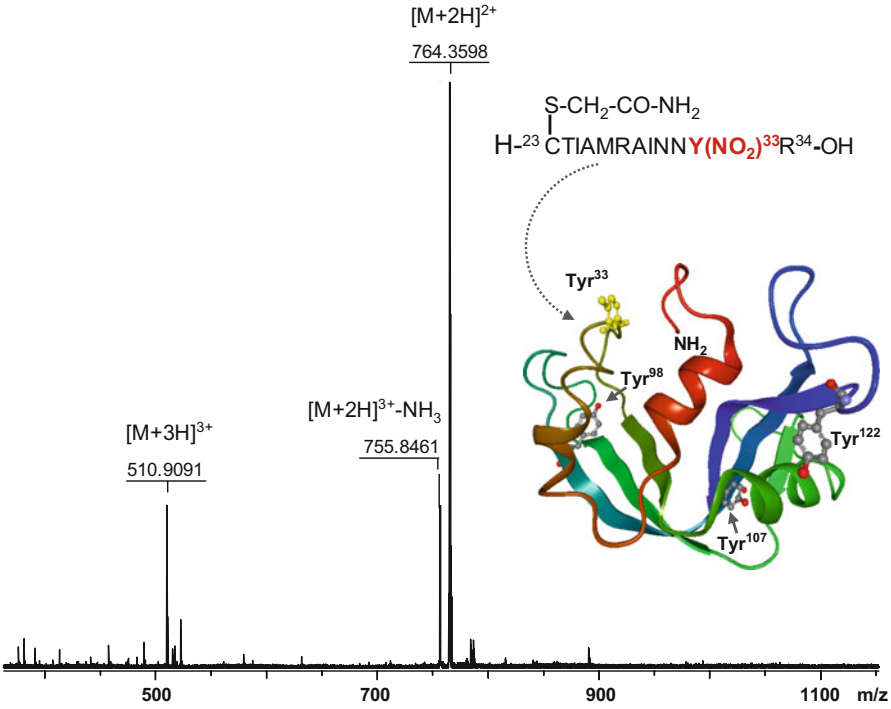


Fig. 7.5 Proteolytic affinity extraction-mass spectrometric approach. Nano-ESI FTICR mass spectrum of the elution fraction containing the nitrated ECP (23-34) peptide fragment, with nitro-Tyr^{33} residue. The inset shows the modeling structure of ECP indicating the accessible surface area of the protein rendered using BallView 1.1.1 program based on the available X-ray crystal structure of the ECP (PDB entry 1H1H)

Furthermore, the combination of proteolytic affinity peptide extraction and Edman sequencing provided quantitative information about the endogenous nitration of ECP in human eosinophil granules. Proteolytic affinity peptide extraction was performed for ECP protein as described above followed by analyzing the elution fraction by Edman sequencing. The quantification was obtained by determination of initial and repetitive yields in each sequencing cycle. From the ratios of sequenced nitrated peptides to intact ECP protein used for PROFINEX, approximately 15 % nitration level was determined for native nitrated ECP [35].

7.3.3 Online Bioaffinity-Mass Spectrometry for Molecular Recognition Specificity of Anti-3-nitrotyrosine Antibodies

Another recently developed analytical tool for measuring molecular interaction kinetics and identifying binding-specific structures is the online combination of SAW biosensor with electrospray ionization MS (SAW-ESI-MS). SAW technology uses the piezoelectric effect of amount differences loaded on a chip to detect bioaffinities with high sensitivity in dilute solutions, without labeling approaches or recalibration for

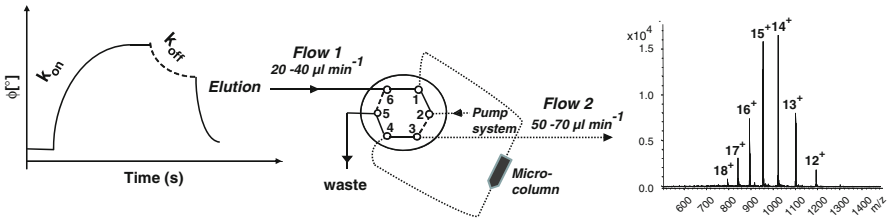


Fig. 7.6 Schematic representation of online combination of SAW biosensor instrument with ESI-MS using a six-port valve micro-column interface. The waste capillary from the SAW biosensor is connected to the interface in position 6 and the ESI-MS inlet capillary in position 3. Total dissociation of protein(s) from the antibody surface was performed by elution with acidic buffer to the micro-column interface (flow 1). This was followed by cleaning with washing solution, elution with aqueous acetonitrile/HCOOH, and transfer of the eluate into the ESI source by switching the injector position (flow 2). Adapted with kind permission of authors and printed with kind permission of Springer Science and Business Media from J. Am. Soc. Mass Spectrom. 21(10): 1643-8, Drăguşanu M., et al. (2010)

buffer changes [18, 40]. One of the first experiment we performed by SAW-ESI-MS online coupling enabled the direct detection, identification, and quantification of nitrated model peptides' affinities to a 3 NT-specific antibody [5]. A six-port valve inject unit system was used between the SAW biosensor and the ESI mass spectrometer equipped with a micro-guard column; the combination provides simultaneous sample concentration, desalting, and flow rate adjustment for mass spectrometric analysis of the dissociated ligand. The interface used for online coupling of SAW biosensor with an electrospray ion trap mass spectrometer is illustrated in Fig. 7.6 and described in detail by Dragusanu et al. [5]. After SAW biosensor measurement of the peptide–antibody association, the peptide is eluted into the guard column (Fig. 7.6, flow 1) and washed to remove PBS buffer salts. An HPLC-pumping system was used for performing the elution of the peptide from the guard column into a manual-assisted microvalve injector which further transferred the sample into the ESI source (Fig. 7.6, flow 2). This interface system provided routine mass spectrometric analysis of real time peptide/protein association from the biosensor surface and showed long-term operation stability with minimal contamination of the ESI-MS system [5].

As shown before (Sect. 3.2) the use of anti-3-nitro-tyrosine-specific antibodies has been essential for mass spectrometric identification of tyrosine nitrations using proteolytic affinity-MS. Poor information regarding the specificities of anti-3-nitrotyrosine antibodies has been reported before [4]. In a previous publication we described a molecular study of the recognition specificities and affinities of two commercially available, monoclonal anti-nitrotyrosine antibodies by dot-blot, ELISA, and affinity-MS, using different 3-nitrotyrosine-containing peptides [36]. The comparison of the three methods essentially provided consistent results to reveal differences in binding affinities and specificities by the anti-nitrotyrosine antibodies, suggesting that antibody binding may be influenced by the peptide structure adjacent to the nitrotyrosine modification.

Several nitrated peptides from ECP were synthesized by solid-phase peptide synthesis, purified by reversed phase-HPLC and characterized by mass spectrometry.

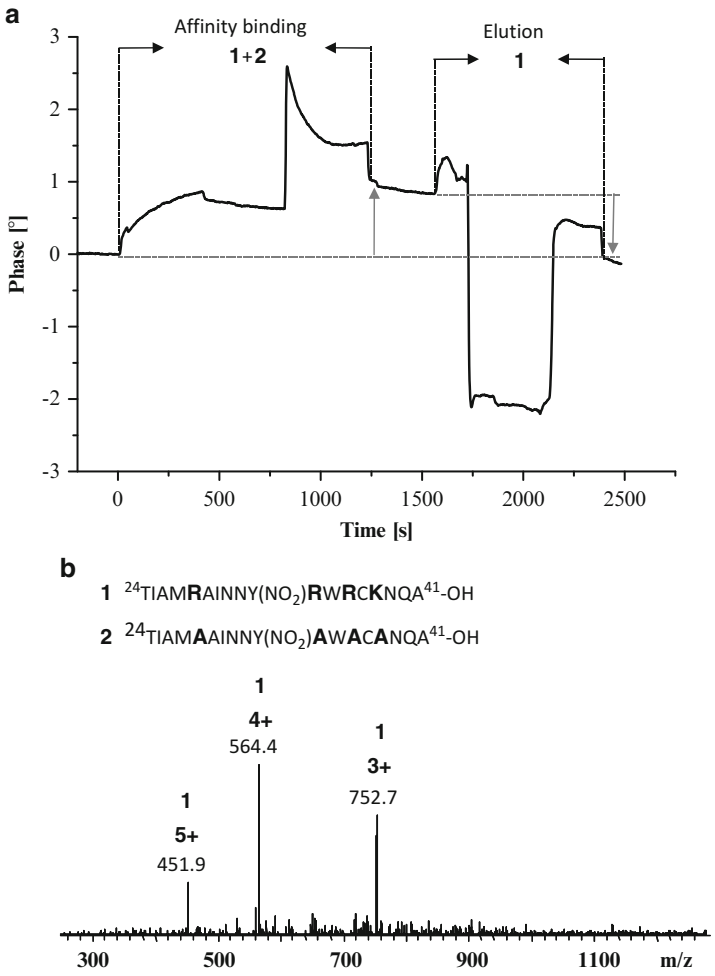


Fig. 7.7 (**a**) SAW binding and elution curves of the peptides mixture using online SAW-ESI-MS and (**b**) ESI-MS spectra of the eluted fraction illustrating the mass spectrometric identification of Tyr-nitrated ECP peptide **1** bound to the immobilized 3-NT antibody. Reprinted with kind permission of authors and Springer Science and Business Media from J. Am. Soc. Mass Spectrom. 23(11): 1831-1840, Petre B.A., et al. (2012)

These model-nitrated peptides have been used to study their affinity to monoclonal anti-3-NT antibodies by online-SAW-ESI-MS. For example, the immobilization of an anti-3NT antibody on the biosensor chip surface via a SAM is performed as a first step. Following association of nitrated peptides from an equimolar mixture (ECP-1 and ECP-2 peptides), dissociation was performed using an acidic glycine buffer which transferred also the sample into the interface for ESI-MS analysis (Fig. 7.6). Biosensor association curves and subsequent dissociation of affinity-bound ECP peptides from a mixture of nitrated peptides are shown in Fig. 7.7a.

The ESI mass spectra of the affinity-eluate showed exclusively ions of the nitrated ECP-1 peptide (Fig. 7.7b). Affinity of the nitrated ECP peptide was found to be completely abolished upon replacing the positively charged residues Arg^{28}, Arg^{34}, Arg^{36}, and Lys^{38} by alanine, thus confirming the importance of cationic amino acids adjacent to the nitration site. Affinity determinations of the nitrated peptide in comparison to the intact ECP protein using the SAW biosensor provided also dissociation constants of approximately 6 nM for the nitrated peptide ECP (24-41) and 28 nM for intact ECP [37]. In addition to the characterization of sequence variations of ECP-nitrated peptides, a preliminary epitope motif derived from the antibody-binding affinities of the tyrosine-nitrated prostacyclin synthase peptides was shown. All these results reveal remarkable similarities, suggesting that (1) stabilization by adjacent positively charged residues and (2) surface exposition of tyrosine-nitrated sites are important factors [4].

The capability of anti-nitrotyrosine antibodies to discriminate between nitrotyrosine in different environments in proteins may be useful for producing antibodies to specific motifs containing tyrosine residues and for the development of highly specific biomarkers.

7.3.4 Proteolytic Affinity Excision-Mass Spectrometric Approach for Identification of Carbohydrate Recognition Structure in Human Galectins

Interactions of proteins with carbohydrates play an important role in several physiological processes such as cellular recognition processes, intracellular regulation pathways, immunological reactions, and the transcriptional or posttranscriptional regulation of gene expression. Galectins are a class of lectin proteins that typically bind β-galactose-containing glycoconjugates and share primary structural homology in their CRDs. Each galectin CRD recognizes different types of glycan ligands and shows highest affinity binding to different structures such as natural glycoconjugate ligands expressed on cell surfaces or in the extracellular matrix providing a large diversity of functions [3, 48]. Galectin-1 and galectin-3 have been shown to be involved in pre-mRNA splicing [51]; galectin-1 can induce apoptosis of activated T-cells by binding to cell surface oligosaccharides [32, 50], and galectin-3 can activate neutrophils [53]. The precise structure of galectins–glycoconjugate ligands interactions is not well understood. The combination of the proteolytic affinity excision approach and mass spectrometry was used to obtain the structural peptide sequence from CRDs of human galectin-1 and galectin-3 [19]. The newly developed approach was named CREDEX-MS (carbohydrate recognition domain excision mass spectrometry) and was performed using lactosylated Sepharose 4B as an immobilization matrix support. Human galectin-1 was bound first to the affinity matrix in phosphate-buffered saline, and any unbound protein was removed by washing steps. In situ proteolytic digestion of remaining bound galectin-1 was performed using trypsin, followed by complete removal of the unbound tryptic generated peptides. Elution was performed using buffer solution containing 0.3 M lactose.

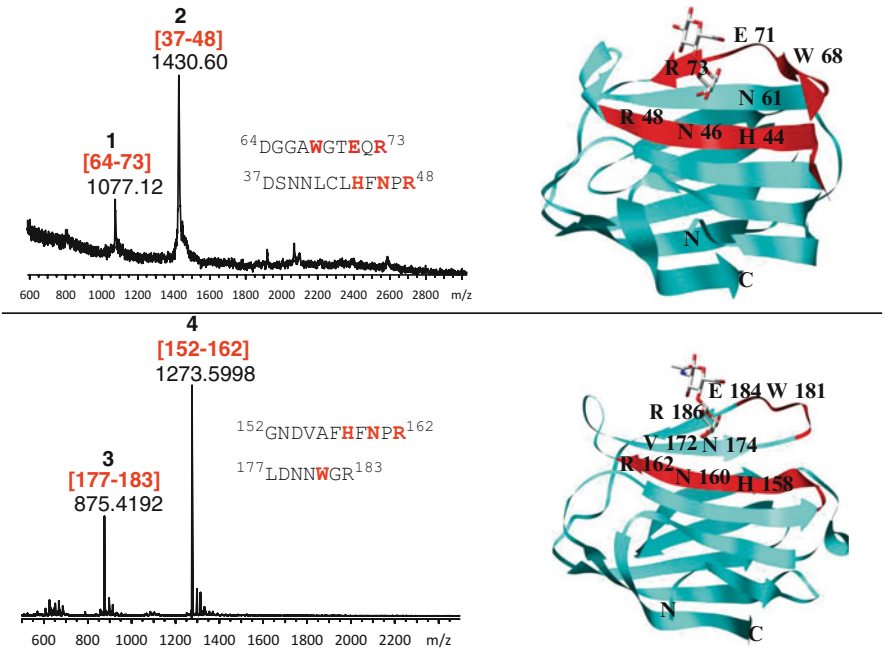


Fig. 7.8 Proteolytic CRDs' peptide excision for complexes of lactose with galectin-1 (*top*) and galectin-3 (*bottom*). MALDI-MS of elution fractions with signals of identified peptides. (*Right*) X-ray crystal structures (PDB entries 1W6O and 1A3K) showing the identified peptides in *red* and the amino acids that are in direct contact with the carbohydrate in *bold*. Reprinted with permission from Moise et al. J. Am. Chem. Soc. 2011; 133(38):14844-7. Copyright 2013 American Chemical Society

The MALDI-MS mass spectrum of the elution fraction shows two specific tryptic peptides (Fig. 7.8), which were displaced from bound galactin-1 by the presence of cognate sugar (lactose).

The two peptides (1 and 2) covered the galectin-1 (64-73) fragment, a sequence stretch with central tryptophan residue involved in C–H/π-interaction with galactose [19] and, respectively, galectin-1 (37-48) fragment [29]. CREDEX approach was applied for human galectin-3. Two specific peptides (3 and 4), galectin-3 (152-162) fragment, and galectin-3 (177-183) from the tryptic generated peptides were identified by MALDI-MS analysis of the lactose eluted fraction. The same two peptides were obtained by using proteolytic affinity extraction procedure, when galectin-3 was first digested in solution using trypsin; the resulted tryptic peptides were allowed to bind to the affinity matrix, and after washing step the bound peptides were eluted (Fig. 7.9).

Together these results provide evidence that the identified peptide fragments within CRDs of galectin-1 and galectin-3, respectively, comprise key amino acids in

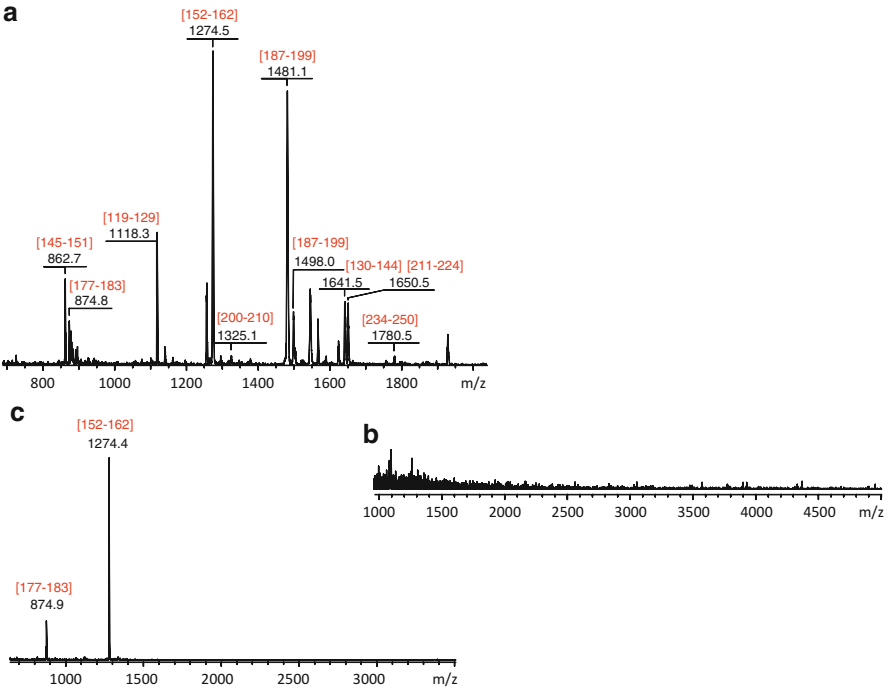


Fig. 7.9 Proteolytic CRDs' peptide extraction from galectin-3 digested peptide mixture. MALDI-MS of (**a**) supernatant fraction, (**b**) wash fraction, and (**c**) elution fraction identify the active galectin-3 peptide fragment (177-183) (peptide 3) and (152-162) galectin-3 peptide fragment (peptide 4). Reprinted with permission from Moise et al. J. Am. Chem. Soc. 2011; 133(38):14844-7. Copyright 2013 American Chemical Society

contact with the carbohydrate ligand. These peptide structures may be considered regions of biological importance for galectin-1 and galectin-3 and may be relevant for further studies of other types of sugar-binding proteins.

7.4 Summary

The results shown here provide evidence that the new affinity-mass spectrometric-based approach in combination with proteolytic digestion presents promising perspectives for elucidating (1) a high diversity of antigen–antibody interactions, (2) selective enrichment and identification of posttranslational modification in biological mixture, (3) extended binding sites, e.g., the interacting area for larger carbohydrates ligands. The new affinity-mass spectrometric approaches presented in this chapter may become increasingly utilized within interdisciplinary fields for

the development of (1) biomolecules that inhibit or propagate molecular cell processes, (2) structure-based drug design, and (3) rare diseases diagnostic approaches using defected enzyme–substrate affinities.

Acknowledgments The author would like to acknowledge Professor Michael Przybylski, Dr. Mihaela Stumbaum, Dr. Roxana Iacob, Dr. Irina Perdivara, Dr. Marilena Manea, and Adrian Moise for interesting topic, useful discussions, and expertise. Our research was supported in part by the Deutsche Forschungsgemeinschaft, Bonn, Germany (PR-175-14-1 and FG-753), EC for the GlycoHIT consortium, German Academic Exchange Service (DAAD), and the Romanian National Authority for Scientific Research, CNCS-UEFISCDI, project number PN-II-RU-TE-2011-3-0038.

References

1. Abello N, Kerstjens HA, Postma DS, Bischoff R (2009) Protein tyrosine nitration: selectivity, physicochemical and biological consequences, denitration, and proteomics methods for the identification of tyrosine-nitrated proteins. J Proteome Res 8:3222–3238
2. Castegna A, Thongboonkerd V, Klein JB, Lynn B, Markesbery WR, Butterfield DA (2003) Proteomic identification of nitrated proteins in Alzheimer's disease brain. J Neurochem 85:1394–1401
3. Cummings RD, Liu FT (2009) Galectins (Chapter 33). In: Varki A, Cummings RD, Esko JD et al (eds) Essentials of glycobiology, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY. http://www.ncbi.nlm.nih.gov/books/NBK1944/
4. Dragusanu M, Petre BA, Przybylski M (2011) Epitope motif of an anti-nitrotyrosine antibody specific for tyrosine-nitrated peptides revealed by a combination of affinity approaches and mass spectrometry. J Pept Sci 17(3):184–191
5. Dragusanu M, Petre BA, Slamnoiu S, Vlad C, Tu T, Przybylski M (2010) On-line bioaffinity-electrospray mass spectrometry for simultaneous detection, identification, and quantification of protein-ligand interactions. J Am Soc Mass Spectrom 21:1643–1648
6. Gabius HJ (1990) Influence of type of linkage and spacer on the interaction of beta-galactoside-binding proteins with immobilized affinity ligands. Anal Biochem 189:91–94
7. Garcia-Monzon C, Majano PL, Zubia I, Sanz P, Apolinario A, Moreno-Otero R (2000) Intrahepatic accumulation of nitrotyrosine in chronic viral hepatitis is associated with histological severity of liver disease. J Hepatol 32:331–338
8. Gokulrangan G, Zaidi A, Michaelis ML, Schoneich C (2007) Proteomic analysis of protein nitration in rat cerebellum: effect of biological aging. J Neurochem 100:1494–1504
9. Hager-Braun C, Hochleitner EO, Gorny MK, Zolla-Pazner S, Bienstock RJ, Tomer KB (2010) Characterization of a discontinuous epitope of the HIV envelope protein gp120 recognized by a human monoclonal antibody using chemical modification and mass spectrometric analysis. J Am Soc Mass Spectrom 21:1687–1698
10. Harlow E, Lane D (1988) Antibodies: a laboratory manual. Cold Spring Harbor Laboratory, USA, pp 53–224
11. Hearty S, Leonard P, O'Kennedy R (2012) Measuring antibody-antigen binding kinetics using surface plasmon resonance. Methods Mol Biol 907:411–442
12. Heijnen HF, van Donselaar E, Slot JW, Fries DM, Blachard-Fillion B, Hodara R, Lightfoot R, Polydoro M, Spielberg D, Thomson L, Regan EA, Crapo J, Ischiropoulos H (2006) Subcellular localization of tyrosine-nitrated proteins is dictated by reactive oxygen species generating enzymes and by proximity to nitric oxide synthase. Free Radic Biol Med 40:1903–1913
13. Hinson JA, Michael SL, Ault SG, Pumford NR (2000) Western blot analysis for nitrotyrosine protein adducts in livers of saline-treated and acetaminophen-treated mice. Toxicol Sci 53:467–473

14. Iacob RE (2004) Molecular recognition structures and antibody affinities of pathophysiological degradation products of Alzheimer's amyloid precursor protein elucidated by high resolution mass spectrometry. Konstanz University, Germany. ISBN 973-8076-85-4
15. Ino Y, Hirano H (2011) Mass spectrometric characterization of proteins transferred from polyacrylamide gels to membrane filters. FEBS J 278:3807–3814
16. Juszczyk P, Paraschiv G, Szymanska A, Kolodziejczyk AS, Rodziewicz-Motowidlo S, Grzonka Z, Przybylski M (2009) Binding epitopes and interaction structure of the neuroprotective protease inhibitor cystatin C with beta-amyloid revealed by proteolytic excision mass spectrometry and molecular docking simulation. J Med Chem 52:2420–2428
17. Khan J, Brennand DM, Bradley N, Gao B, Bruckdorfer R, Jacobs M (1998) 3-Nitrotyrosine in the proteins of human plasma determined by an ELISA method. Biochem J 330(Pt 2): 795–801
18. Lange K, Bender F, Voigt A, Gao H, Rappt M (2003) A surface acoustic wave biosensor concept with low flow cell volumes for label-free detection. Anal Chem 75:5561–5566
19. Lopez-Lucendo MF, Solis D, Andre S, Hirabayashi J, Kasai K, Kaltner H, Gabius HJ, Romero A (2004) Growth-regulatory human galectin-1: crystallographic characterisation of the structural changes induced by single-site mutations and their impact on the thermodynamics of ligand binding. J Mol Biol 343:957–970
20. Luhrs T, Ritter C, Adrian M, Riek-Loher D, Bohrmann B, Dobeli H, Schubert D, Riek R (2005) 3D structure of Alzheimer's amyloid-beta(1-42) fibrils. Proc Natl Acad Sci U S A 102:17342–17347
21. Luque-Garcia JL, Zhou G, Spellman DS, Sun TT, Neubert TA (2008) Analysis of electroblotted proteins by mass spectrometry: protein identification after Western blotting. Mol Cell Proteomics 7:308–314
22. Macht M, Marquardt A, Deininger SO, Damoc E, Kohlmann M, Przybylski M (2004) "Affinity-proteomics": direct protein identification from biological material using mass spectrometric epitope mapping. Anal Bioanal Chem 378:1102–1111
23. Maftei M, Thurm F, Leirer VM, von Arnim CA, Elbert T, Przybylski M, Kolassa IT, Manea M (2012) Antigen-bound and free beta-amyloid autoantibodies in serum of healthy adults. PLoS One 7:e44516
24. Maftei M, Tian X, Manea M, Exner TE, Schwanzar D, von Arnim CA, Przybylski M (2012) Interaction structure of the complex between neuroprotective factor humanin and Alzheimer's beta-amyloid peptide revealed by affinity mass spectrometry and molecular modeling. J Pept Sci 18:373–382
25. Manea M, Mezo G, Hudecz F, Przybylski M (2004) Polypeptide conjugates comprising a beta-amyloid plaque-specific epitope as new vaccine structures against Alzheimer's disease. Biopolymers 76:503–511
26. Maurer T (2005) NMR studies of protein-ligand interactions. Methods Mol Biol 305:197–214
27. McLaurin J, Cecal R, Kierstead ME, Tian X, Phinney AL, Manea M, French JE, Lambermon MH, Darabie AA, Brown ME, Janus C, Chishti MA, Horne P, Westaway D, Fraser PE, Mount HT, Przybylski M, St George-Hyslop P (2002) Therapeutically effective antibodies against amyloid-beta peptide target amyloid-beta residues 4-10 and inhibit cytotoxicity and fibrillogenesis. Nat Med 8:1263–1269
28. Mohan CG, Boix E, Evans HR, Nikolovski Z, Nogues MV, Cuchillo CM, Acharya KR (2002) The crystal structure of eosinophil cationic protein in complex with 2′,5′-ADP at 2.0 A resolution reveals the details of the ribonucleolytic active site. Biochemistry 41:12100–12106
29. Moise A, Andre S, Eggers F, Krzeminski M, Przybylski M, Gabius HJ (2011) Toward bioinspired galectin mimetics: identification of ligand-contacting peptides by proteolytic-excision mass spectrometry. J Am Chem Soc 133:14844–14847
30. Morgan D (2006) Immunotherapy for Alzheimer's disease. J Alzheimers Dis 9:425–432
31. Morrill PR, Millington RB, Lowe CR (2003) Imaging surface plasmon resonance system for screening affinity ligands. J Chromatogr B Analyt Technol Biomed Life Sci 793:229–251
32. Novelli F, Allione A, Wells V, Forni G, Mallucci L (1999) Negative cell cycle control of human T cells by beta-galactoside binding protein (beta GBP): induction of programmed cell death in leukaemic cells. J Cell Physiol 178:102–108

33. Papac DI, Hoyes J, Tomer KB (1994) Epitope mapping of the gastrin-releasing peptide/antibombesin monoclonal antibody complex by proteolysis followed by matrix-assisted laser desorption ionization mass spectrometry. Protein Sci 3:1485–1492
34. Paraschiv G, Vincke C, Czaplewska P, Manea M, Muyldermans S, Przybylski M (2013) Epitope structure and binding affinity of single chain llama anti-beta-amyloid antibodies revealed by proteolytic excision affinity-mass spectrometry. J Mol Recognit 26:1–9
35. Petre BA (2008) Analytical development and biochemical application of mass spectrometry in combination with immunoaffinity methods for identification and structural characterisation of protein nitration. Dissertation at http://kops.ub.uni-konstanz.de/handle/urn:nbn:de:bsz:352-opus-85026.
36. Petre BA, Dragusanu M, Przybylski M (2008) Molecular recognition specificity of anti-3-nitrotyrosine antibodies revealed by affinity-mass spectrometry and immunoanalytical methods. In: Popescu C, Zamfir AD, Dinca N (eds) Applications of mass spectrometry in life sciences. Springer, Netherlands. ISBN 978-1-4020-8811-7
37. Petre BA, Ulrich M, Stumbaum M, Bernevic B, Moise A, Doring G, Przybylski M (2012) When is mass spectrometry combined with affinity approaches essential? A case study of tyrosine nitration in proteins. J Am Soc Mass Spectrom 23:1831–1840
38. Petre BA, Youhnovski N, Lukkari J, Weber R, Przybylski M (2005) Structural characterisation of tyrosine-nitrated peptides by ultraviolet and infrared matrix-assisted laser desorption/ionisation Fourier transform ion cyclotron resonance mass spectrometry. Eur J Mass Spectrom (Chichester, Eng) 11:513–518
39. Przybylski M et al (2007) Molecular approaches for immunotherapy and diagnostics of Alzheimer's disease based on epitope-specific beta-amyloid-antibodies. Eur., US & PCT Patent application
40. Rocha-Gaso MI, March-Iborra C, Montoya-Baides A, Arnau-Vives A (2009) Surface generated acoustic wave biosensors for the detection of pathogens: a review. Sensors (Basel) 9:5740–5769
41. Shimamoto G, Gegg C, Boone T, Queva C (2012) Peptibodies: a flexible alternative format to antibodies. MAbs 4:586–591
42. Stefanescu R, Iacob RE, Damoc EN, Marquardt A, Amstalden E, Manea M, Perdivara I, Maftei M, Paraschiv G, Przybylski M (2007) Mass spectrometric approaches for elucidation of antigenantibody recognition structures in molecular immunology. Eur J Mass Spectrom (Chichester, Eng) 13:69–75
43. Suckau D, Kohl J, Karwath G, Schneider K, Casaretto M, Bitter-Suermann D, Przybylski M (1990) Molecular epitope identification by limited proteolysis of an immobilized antigen-antibody complex and mass spectrometric peptide mapping. Proc Natl Acad Sci U S A 87: 9848–9852
44. Tabb DL (2012) Evaluating protein interactions through cross-linking mass spectrometry. Nat Methods 9:879–881
45. Taylor CR, Levenson RM (2006) Quantification of immunohistochemistry—issues concerning methods, utility and semiquantitative assessment II. Histopathology 49:411–424
46. Tian X, Cecal R, McLaurin J, Manea M, Stefanescu R, Grau S, Harnasch M, Amir S, Ehrmann M, St George-Hyslop P, Kohlmann M, Przybylski M (2005) Identification and structural characterisation of carboxy-terminal polypeptides and antibody epitopes of Alzheimer's amyloid precursor protein using high-resolution mass spectrometry. Eur J Mass Spectrom (Chichester, Eng) 11:547–556
47. Ulrich M, Petre A, Youhnovski N, Promm F, Schirle M, Schumm M, Pero RS, Doyle A, Checkel J, Kita H, Thiyagarajan N, Acharya KR, Schmid-Grendelmeier P, Simon HU, Schwarz H, Tsutsui M, Shimokawa H, Bellon G, Lee JJ, Przybylski M, Doring G (2008) Post-translational tyrosine nitration of eosinophil granule toxins mediated by eosinophil peroxidase. J Biol Chem 283:28629–28640
48. van den Brule FA, Buicu C, Baldet M, Sobel ME, Cooper DN, Marschal P, Castronovo V (1995) Galectin-1 modulates human melanoma cell adhesion to laminin. Biochem Biophys Res Commun 209:760–767

49. van Sommeren APG, Machielsen PAGM, Gribnau TCJ (1993) Comparison of three activated agaroses for use in affinity chromatography: effects on coupling performance and ligand leakage. J Chromatogr A 639:23–31
50. Vespa GN, Lewis LA, Kozak KR, Moran M, Nguyen JT, Baum LG, Miceli MC (1999) Galectin-1 specifically modulates TCR signals to enhance TCR apoptosis but inhibit IL-2 production and proliferation. J Immunol 162:799–806
51. Vyakarnam A, Dagher SF, Wang JL, Patterson RJ (1997) Evidence for a role for galectin-1 in pre-mRNA splicing. Mol Cell Biol 17:4730–4737
52. Williams MA, Daviter T (2013) Protein-ligand interactions, 2nd edn. Humana, New York. ISBN 978-1-62703-397-8
53. Yamaoka A, Kuwabara I, Frigeri LG, Liu FT (1995) A human lectin, galectin-3 (epsilon bp/Mac-2), stimulates superoxide production by neutrophils. J Immunol 154:3479–3487
54. Yu L, Xiao G, Zhang J, Remmele RL Jr, Eu M, Liu D (2012) Identification and quantification of Fc fusion peptibody degradations by limited proteolysis method. Anal Biochem 428: 137–142
55. Zhao Y, Jensen ON (2009) Modification-specific proteomics: strategies for characterization of post-translational modifications using enrichment techniques. Proteomics 9:4632–4641
56. Zhao Y, Muir TW, Kent SB, Tischer E, Scardina JM, Chait BT (1996) Mapping protein-protein interactions by affinity-directed mass spectrometry. Proc Natl Acad Sci U S A 93:4020–4024

Chapter 8
Neurological Analyses: Focus on Gangliosides and Mass Spectrometry

Alina D. Zamfir

Abstract Gangliosides, sialylated glycosphingolipids, are particularly enriched in mammalian central nervous system where their expression is cell type-specific and changes particularly during brain development, maturation, aging, and diseases. For this reason, gangliosides are important diagnostic markers for various brain ailments, including primary and secondary brain tumors and neurodegenerative diseases. Among all biochemical and biophysical methods employed so far for ganglioside analysis, mass spectrometry (MS) emerged as one of the most reliable due to the sensitivity, accuracy, and speed of analysis as well as the possibility to characterize in details the molecular structure of the identified biomarkers.

This chapter presents significant achievements of MS with either electrospray (ESI), chip-based ESI, or matrix-assisted laser desorption/ionization (MALDI) in the analysis of gangliosides in normal and diseased human brain. Specifically, the chapter assesses the MS contribution in determination of topospecificity, filogenetic, and brain development stage dependence of ganglioside composition and structure as well as in discovery of ganglioside markers in neurodegenerative/neurodevelopmental conditions, primary and secondary brain tumors. The highlighted accomplishments in characterization of novel structures associated to severe brain pathologies show that MS has real perspectives to become a routine method for early diagnosis and therapy based on this biomolecule class.

A.D. Zamfir (✉)
Faculty of Physics, West University of Timisoara, 4 Boulevard V. Parvan, Timisoara, Romania

National Institute for Research and Development in Electrochemistry and Condensed Matter, 1 P. Andronescu Street, Timisoara, Romania
e-mail: alina.zamfir@uav.ro

A.G. Woods and C.C. Darie (eds.), *Advancements of Mass Spectrometry in Biomedical Research*, Advances in Experimental Medicine and Biology 806,
DOI 10.1007/978-3-319-06068-2_8, © Springer International Publishing Switzerland 2014

8.1 Introduction

Lipid bilayer membranes of the vertebrate cells contain a special class of glycosphingolipids (GSL) called gangliosides. Gangliosides consist of a hydrophobic ceramide (Cer) moiety through which the molecule is attached to the plasma membrane and a hydrophilic oligosaccharide chain. Cer includes a sphingoid base and a fatty acid chain, both with variable composition from a species to another. The carbohydrate core also exhibits variability of the length and structure. Besides, one or more characteristic sialic acid groups (i.e., *N*-acetylneuraminic acid, *N*-glycolylneuraminic acid) are attached to the oligosacchride core by an aketosidic linkage. This way, gangliosides are inserted in the outer layer of the plasma membrane [56] via the heterogeneous ceramide moiety, while their oligosaccharide chain faces the external medium; thus, the carbohydrate chain is free to interact with soluble extracellular molecules and with the hydrophilic portion of other membrane components.

Although this particular class of GSLs is expressed on basically all vertebrate cells, systematic investigations showed that central nervous system (CNS) contains a several times higher ganglioside concentration than the extraneural tissue. This feature indicates that ganglioside molecules are of particular importance for the development and function of CNS [58].

As plasma-membrane components, enriched in specialized microdomains, i.e., "glycosynapse," gangliosides interact with different signal transducers, thus mediating carbohydrate-dependent cell adhesion and inducing cell activation, motility, and growth. Several reports emphasize the participation of gangliosides in significant biological processes in the healthy CNS such as brain development, maturation, aging [58], and defined functions of each specialized brain region, as well as in pathological CNS, in particular neurodegenerative diseases [59] and malignant transformation [26]. The regulatory roles of gangliosides at the CNS level were clearly demonstrated by the observed dramatic change in their expression and structure during development of the nervous system and their region-specific distribution. On the other hand, certain structures were found to be differentially and specifically expressed in the brain affected by neurodegenerative diseases [77]. Changes in ganglioside composition and content have been observed also during neoplastic cell transformation. A decrease in the regular ganglioside profile and an increase in the structures detected only in small amounts in the normal brain tissue were found in primary and secondary (metastases) brain tumors, demonstrating a direct correlation between ganglioside composition and histological type and grade of the tumors [9, 17, 66]. This aspect provides the option to use gangliosides as biochemical markers in early histopathological diagnosis, grading, and prognosis of tumors. Therefore, nowadays gangliosides are considered valuable diagnostic markers of CNS ailments being in the current focus of research also as potential therapeutic agents.

In the last decade several biophysical methods have been developed for the investigation of ganglioside expression in severe brain diseases. Hence, for the assessment of ganglioside quantity and composition in human brain, cell- and tissue-staining analyses such as immunocytochemistry and immunohistochemistry and assays using chromatographic methods based on thin layer chromatography (TLC) and high-performance liquid chromatography (HPLC) were successfully employed [40, 69]. Ganglioside profiling, their quantification, and correlation to histomorphology and grading of human tumors, i.e., gliomas, were studied using a newly developed microbore HPLC method [84]. Also, the application of infrared spectroscopy as an adjunct to histopathology in detecting and diagnosing human brain tumors was demonstrated [36, 70]. Also, ganglioside expression in human glioblastoma was determined by confocal microscopy of immunostained brain sections using antiganglioside monoclonal antibodies [27]. However, to define the structure-to-function interrelationship of each particular ganglioside implicated in a physiological/pathological process and to improve their diagnostic significance, more sensitive and accurate approaches to enhance lipid-linked carbohydrate analysis under high-structural complexity were necessary. In this context, mass spectrometry (MS) has emerged lately as one of the most potent tools for the structural analysis of complex brain gangliosides. Implemented initially with fast atom bombardment (FAB) ionization [15], the capability of MS for sensitive structural analysis of gangliosides increased significantly after the introduction of matrix-assisted laser desorption / ionization (MALDI) and electrospray (ESI) ionization from one side and the possibility to perform tandem MS or multistage MS for detailed structural analysis.

Nevertheless, as compared to other classes of biomolecules, mass spectrometry of gangliosides has developed much slower. Although the value of ganglioside markers is universally acknowledged, the current inventory of species identified by MS is rather poor as compared to proteomics field, and the number of research laboratories constantly involved in MS of gangliosides is reduced to only a few, worldwide. This state-of-the-art is a result of the high structural diversity and heterogeneity of gangliosides and their inferior ionization efficiency as compared to peptides and proteins, which require reconsideration of all MS conditions for screening and sequencing. The experimental design must also be adapted to the type of the possibly present labile moieties (NeuAc, Fuc, *O*-Ac), the ionizability of the functional groups, and the special molecule constitution consisting of a hydrophilic sugar core and a hydrophobic ceramide part. All these characteristics made the class of gangliosides less amenable to MS. However, in the subsequent parts of this chapter it will be shown that adequate strategies may lead to a successful implementation of MS in ganglioside analysis, with important results not only in brain mapping but also in discovery of species associated to severe brain diseases.

8.2 Ganglioside Nomenclature

Gangliosides and the precursor glycosphingolipids are currently abbreviated according to the system introduced by Svennerholm and Fredman [71] and the recommendations of IUPAC-IUB Commission on Biochemical Nomenclature [30] as follows: LacCer, Galβ4Glcβ1Cer; GA2, Gg_3Cer, GalNAcβ4Galβ4Glcβ1Cer; GA1, Gg_4Cer, Galβ3GalNAcβ4Galβ4Glcβ1Cer; nLc_4Cer, Galβ4GlcNAcβ3Galβ4Glcβ1Cer; Lc_4Cer, Galβ3GlcNAcβ3Galβ4Glcβ1Cer; GM3, II^3-α-Neu5Ac-LacCer; GD3, II^3-α-$(Neu5Ac)_2$- LacCer; GT3, II^3-α-$(Neu5Ac)_{3\text{-}LacCer}$; GM2, II^3-α-Neu5Ac-Gg_3Cer; GD2, II^3-α-$(Neu5Ac)_2$-Gg_3Cer; GM1a or GM1, II^3-α-Neu5Ac-Gg_4Cer; GM1b, IV^3-α-Neu5Ac-Gg_4Cer; GalNAc-GM1b, IV^3-α-Neu5Ac-Gg5Cer; GD1a, IV^3-α-Neu5Ac,II^3-α-Neu5Ac-Gg_4Cer; GD1b, II^3-α-$(Neu5Ac)_2$-Gg_4Cer; GT1b, IV^3-α-Neu5Ac,II^3-α-$(Neu5Ac)_2$-Gg_4Cer; GQ1b, IV^3-α-$(Neu5Ac)_2$,II^3-α-$(Neu5Ac)_2$-Gg_4Cer; nLM1 or 3′-nLM1, IV^3-α-Neu5Ac-nLc4Cer; LM1 or 3′-isoLM1, IV^3-α-Neu5Ac-Lc_4Cer; nLD1, disialo-nLc_4Cer;

8.3 Gangliosides in Normal Brain

Prior to MS introduction in brain research, mapping of the gangliosides expressed in different regions of normal human brain was carried out using high affinity anti-ganglioside antibodies based on immunohistochemistry, TLC coupled with densitometric scanning and two-dimensional TLC. Because of the principle of detection, restricted to major components, and the low throughput, these methods offered a limited amount of information. Therefore, in last years many efforts to implement mass spectrometry with either ESI, nanoESI chip, or MALDI methodologies into CNS ganglioside analysis have been invested.

8.3.1 Cerebrum

In view of the technical performances and the accuracy of data with biological impact, high-resolution mass spectrometry on either Fourier transform ion cyclotron resonance (FTICR) MS or quadruple time of-flight (QTOF) MS proved a remarkable potential in brain ganglioside research. In comparison to all other MS methods, FTICR MS exhibits an ultra-high resolution exceeding 10^6 and a mass accuracy often below 1 ppm [53]. FTICR MS provides at the same time the advantage of several ion fragmentation techniques based on precursor dissociation such as CID, sustained off-resonance irradiation (SORI-CID), infrared laser multiphoton (IRMPD), or electron capture dissociation (ECD) as well as the possibility to perform multiple stage MS (MS^n).

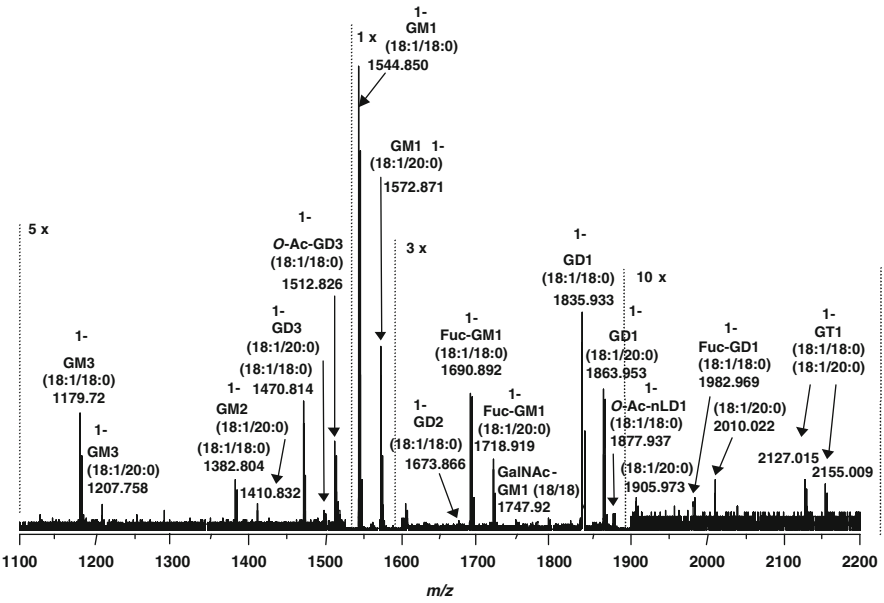


Fig. 8.1 NanoESI FTICR MS in negative ion mode of the ganglioside mixture from normal adult human cerebrum. Solvent Methanol. Sample concentration 5 pmol/μL. Capillary exit voltage 150 V. Number of scans 150. (Zamfir, Vukelić et al. unpublished results)

In brain ganglioside research nanoESI FTICR MS in the negative ion mode was shown [81] to be highly effective for screening and sequencing of the native ganglioside mixture extracted from human cerebrum (Fig. 8.1), for detailed structural analysis of an isolated GT1 (d18:1/20:0) fraction screened by nanoESI (Fig. 8.2) and sequenced using SORI-CID (Fig. 8.3). It was found that instrumental set-up in negative ion mode detection and the solvent systems play a critical role for ganglioside ion formation by nanoESI FTICR MS and for generation in SORI-CID MS^2 of product ions diagnostic for the polysialylated species.

Under adequate nanoESI FTICR MS screening conditions, cerebrum-associated ganglioside species decorated with labile attachments that readily cleave-off such as *O*-fucosylation and *O*-acetylation were detected in intact form (Fig. 8.1). Moreover, under thoroughly optimized SORI settings a set of fragment ions specific for the GT1 molecular form carrying the disialo ($Neu5Ac_2$) element at the inner galactose (Gal) residue were generated. Such a configuration is consistent with the GT1b (d18:1/20:0) isomer in human cerebrum, identified under an average mass accuracy below 5 ppm, at a sample concentration of only 5 pmol/μL. This enhancement of the sensitivity is of particular importance for characterization of material from biological sources wherefrom only minute quantities are available.

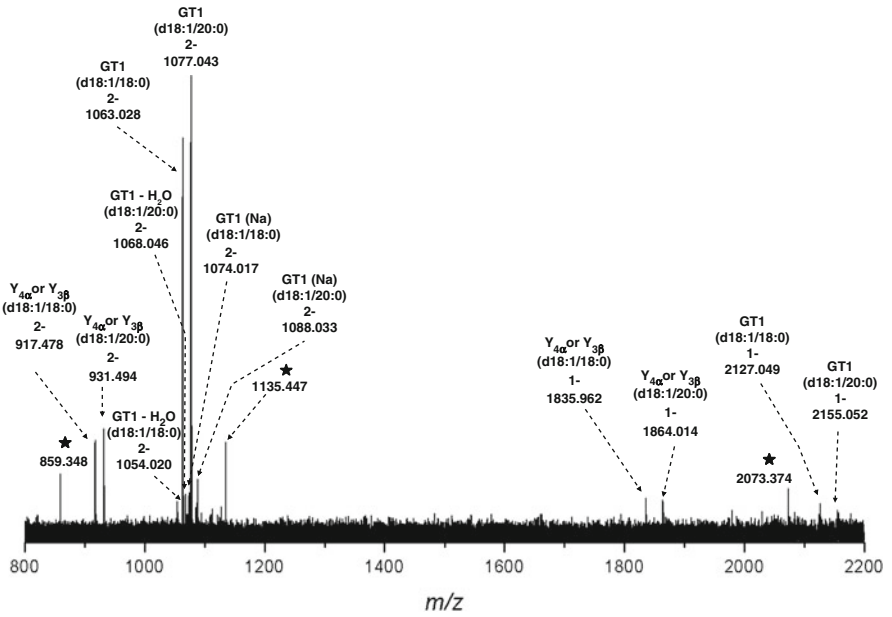


Fig. 8.2 NanoESI FTICR MS in negative ion mode of the purified GT1 fraction from normal adult human brain. Conditions as in Fig. 8.1. The assignment of the *in-source* fragmented ions is according to the nomenclature [14]. Reprinted with permission from Vukelić et al. [81]

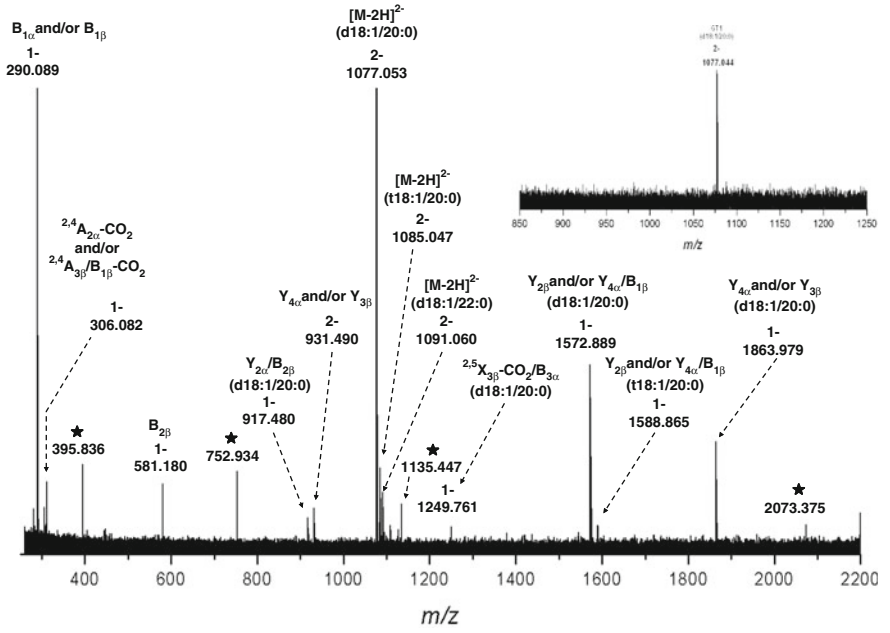


Fig. 8.3 NanoESI-FTICR SORI-CID MS^2 in negative ion mode of the doubly deprotonated molecule at *m/z* 1,077.043 corresponding to GT1 (d18:1/20:0) ganglioside species in human brain. Inset: Isolation of the precursor ion. Excitation pulse length 0.4 s. Collision gas Ar. Other conditions as in Fig. 8.1. The assignment of the fragment ions is according to the nomenclature [14]. Reprinted with permission from Vukelić et al. [81]

8.3.2 Cerebellum

Cerebellum is a morpho-anatomically and functionally highly specialized region of the brain with characteristic differences in the composition of its major ganglioside species, in comparison with cerebrum. The function, behavior, and even survival of cerebellar neurons are highly dependant on the cellular expression of the ganglioside species.

The first exhaustive determination of ganglioside expression in human cerebellum was accomplished by mass spectrometry using a high-resolution instrument (QTOF MS) in combination with silicon chip-based nanoESI [88]. Due to highly efficient ionization properties, silicon-based chips for nanoESI incorporated in a NanoMate robot (Advion Biosciences) preferentially form multiply charged ions. Furthermore, the *in-source* fragmentation of labile groups attached to the biomolecular core is minimized. With an internal diameter of the nozzle of 2.5 μm the system allows for flow rates down to 30–50 nL/min which represents a gain in sensitivity by at least five times as compared to classical nanospray capillaries. Additionally, the nozzles on the ESI chip provide a long-lasting steady spray, and elimination of sample-to-sample carryover.

In view of these advantages NanoMate incorporating chip ESI technology was coupled to QTOF MS and MS/MS and optimized in the negative ion mode, for the characterization of a complex cerebellar ganglioside mixture [88]. Negative ion mode screening of the cerebellum gray matter extracts enabled the identification of no less than 46 different ganglioside species exhibiting high heterogeneity in the ceramide motifs, sialylation status, and biologically relevant modifications (Fig. 8.4a, b, Table 8.1). This way, the first realistic representation of the ganglioside complexity, when compared against TLC and capillary-based ESI, was achieved. The mixture was found dominated by GD1 glycoforms: 19 different doubly deprotonated molecular variants of GD1 have been detected. Also identified in the mixture were GM1, GM2, GM3, GD2, GD3, GT1, and GQ1, expressing different ceramide portions. Biologically interesting *O*-acetylated and/or fucosylated GM1, GD1, GT1, and GQ1 exhibiting a high degree of heterogeneity in their ceramide motifs were also discovered (Table 8.1). For the first time, fully automated nanoESI chip MS infusion was combined also with automated precursor ion selection and fragmentation in data-dependant acquisition, allowing to obtain a sufficient set of "fingerprint" ions for the structural elucidation of a GT1b isomer having the configuration (d18:1/18:0), within 1 min of acquiring MS/MS data and only approximately 0.5 pmol of sample consumption.

In combination with ion trap mass spectrometry, chip-based nanoESI system was also used for determination of gangliosides expressed in fetal cerebellum [49]. A number of 56 ganglioside and asialo ganglioside species differing in either the composition of the glycan core and/or that of the ceramide were identified. Interestingly, *O*-Ac- and/or Fuc-modified gangliosides, found in the adult human cerebellum, were not identified in fetal human cerebellum. Ganglioside chains modified by such attachments are associated with the tissue during its

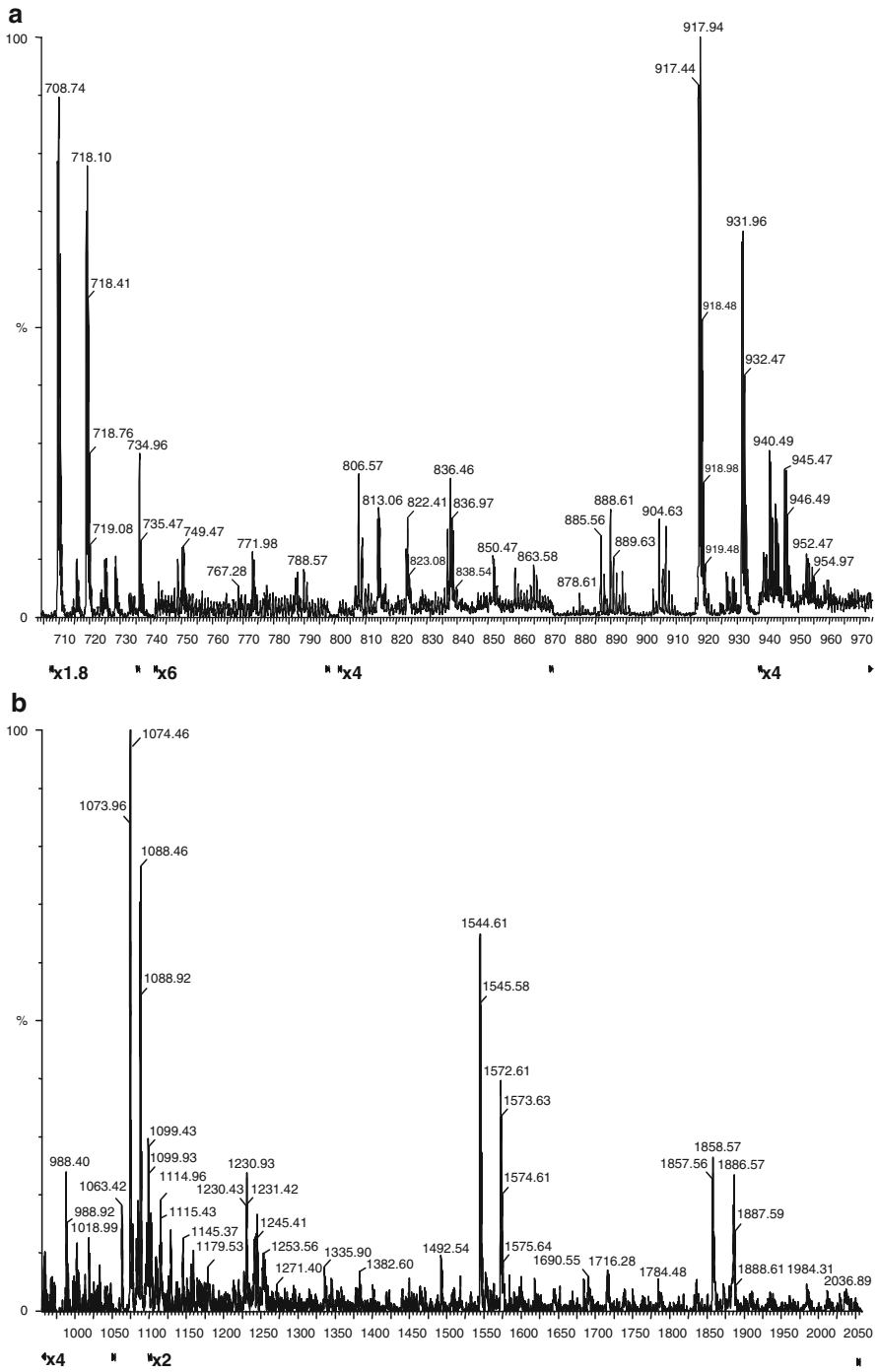


Fig. 8.4 Fully automated nanoESI chip QTOF MS in the negative ion mode of the ganglioside mixture extracted from the gray matter of adult cerebellum. Sample concentration: 5 pmol/μL in methanol. Acquisition time 3 min. Sampling cone potential 45–135 V. (**a**) *m/z* (700–980). (**b**) *m/z* (980–2,050) [88]

Table 8.1 Assignment of ganglioside components in the native mixture extracted from the gray matter of normal human adult cerebellum (male, 20 years) detected by nanoESI chip QTOF MS

Type of molecularion	*m/z* (monoisotopic) Detected	Calculated	Assigned structure
$[M+2Na-4H]^{2-}$	611.40	611.35	GM3 (d18:1/18:0)
$[M-H]^{-}$	1,179.57	1,179.74	
$[M-H]^{-}$	1,382.60	1,382.82	GM2 (d18:1/18:0)
$[M-2H]^{2-}$	734.96	734.91	GD3 (d18:1/18:0)
$[M+Na-2H]^{-}$	1,492.78	1,492.81	
$[M-2H]^{2-}$	748.99	748.93	GD3 (d18:1/20:0)
$[M-H]^{-}$	1,518.51	1,518.85	GM1, nLM1 and/or LM1 (d18:0/16:0)
$[M-2H]^{2-}$	771.98	771.93	GM1, nLM1 and/or LM1 (d18:1/18:0)
$[M-H]^{-}$	1,544.61	1,544.85	
$[M-2H]^{2-}$	786.00	785.92	GM1, nLM1 and/or LM1 (d18:1/20:0)
$[M-H]^{-}$	1,572.61	1,572.85	
$[M-H]^{-}$	1,690.55	1,690.93	Fuc-GM1 (d18:1/18:0)
$[M-H]^{-}$	1,716.56	1,716.94	Fuc-GM1 (d18:1/20:1)
$[M-2H]^{2-}$	836.46	836.45	GD2 (d18:1/18:0)
$[M-2H]^{2-}$	850.47	850.47	GD2 (d18:1/20:0)
$[M-2H]^{2-}$	917.44	917.48	GD1, nLD1 and/or LD1 (d18:1/18:0)
$[M+Na-3H]^{2-}$	928.45	928.47	
$[M-H]^{-}$	1,835.62	1,835.96	
$[M+Na-2H]^{-}$	1,857.56	1,857.95	
$[M-2H]^{2-}$	926.44	926.48	GD1, nLD1 and/or LD1 (t18:0/18:0)
$[M-2H]^{2-}$	924.44	924.49	GD1, nLD1 and/or LD1 (d18:1/19:0)
$[M-2H]^{2-}$	931.46	931.49	GD1, nLD1 and/or LD1 (d18:1/20:0)
$[M+Na-3H]^{2-}$	942.44	942.48	
$[M-H]^{-}$	1,885.60	1,885.98	
$[M-2H]^{2-}$	940.49	940.50	GD1, nLD1 and/or LD1 (t18:0/20:0)
$[M-2H]^{2-}$	938.44	938.50	GD1, nLD1 and/or LD1 (d18:1/21:0)
$[M-2H]^{2-}$	945.47	945.51	GD1, nLD1 and/or LD1 (d18:1/22:0)
$[M-2H]^{2-}$	954.46	954.51	GD1, nLD1 and/or LD1 (t18:0/22:0)
$[M-2H]^{2-}$	952.47	952.52	GD1, nLD1 and/or LD1 (d18:1/23:0)
$[M-2H]^{2-}$	958.46[a]	958.52	GD1, nLD1 and/or LD1 (d18:1/24:1)
$[M-2H]^{2-}$	966.44	966.53	GD1, nLD1 and/or LD1 (d18:1/25:0) or (d20:1/23:0)
$[M-2H]^{2-}$	988.40	988.49	Fuc-GD1 (d18:1/18:2)
$[M-2H]^{2-}$	990.40	990.51	Fuc-GD1 (d18:1/18:0)
$[M-2H]^{2-}$	999.41[a]	999.51	Fuc-GD1 (t18:0/18:0)
$[M-2H]^{2-}$	1,002.41	1,002.51	Fuc-GD1 (d18:1/20:2)
$[M-2H]^{2-}$	1,004.42	1,004.52	Fuc-GD1 (d18:1/20:0)
$[M-2H]^{2-}$	1,013.44[a]	1,013.53	Fuc-GD1 (t18:0/20:0)
$[M-2H]^{2-}$	1,018.99	1,019.02	GalNAc-GD1 (d18:1/18:0)
$[M-2H]^{2-}$	1,032.93[a]	1,033.03	GalNAc-GD1 (d18:1/20:0)
$[M-3H]^{3-}$	708.39	708.35	GT1 (d18:1/18:0)
$[M-2H]^{2-}$	1,062.96	1,063.03	
$[M+Na-3H]^{2-}$	1,073.92	1,074.02	
$[M+2Na-4H]^{2-}$	1,084.93	1,085.01	

(continued)

Table 8.1 (continued)

Type of molecularion	*m/z* (monoisotopic) Detected	Calculated	Assigned structure
$[M-3H]^{3-}$	714.41	714.35	GT1 (t18:0/18:0)
$[M+Na-3H]^{2-}$	1,082.92	1,083.02	
$[M-3H]^{3-}$	717.75	717.69	GT1 (d18:1/20:0)
$[M-2H]^{2-}$	1,076.97	1,077.04	
$[M+Na-3H]^{2-}$	1,087.95	1,088.03	
$[M+2Na-4H]^{2-}$	1,098.92	1,099.02	
$[M-3H]^{3-}$	723.75	723.70	GT1 (t18:0/20:0)
$[M+Na-3H]^{2-}$	1,096.93	1,097.04	
$[M+Na-3H]^{2-}$	1,094.95[a]	1,095.04	GT1 (d18:1/21:0)
$[M-3H]^{3-}$	727.11	727.04	GT1 (d18:1/22:0)
$[M+Na-3H]^{2-}$	1,101.92	1,102.05	
$[M+Na-3H]^{2-}$	1,108.92[a]	1,109.06	GT1 (d18:1/23:0)
$[M+Na-3H]^{2-}$	1,114.96	1,115.06	GT1 (d18:1/24:1)
$[M-3H]^{3-}$	722.39	722.35	*O*-Ac-GT1 (d18:1/18:0)
$[M-3H]^{3-}$	731.74	731.70	*O*-Ac-GT1 (d18:1/20:0)
$[M-2H]^{2-}$	1,128.95	1,129.05	Fuc-GT1 (d18:1/17:0)
$[M-2H]^{2-}$	1,144.89	1,145.06	Fuc-GT1 (t18:0/18:0)
$[M-2H]^{2-}$	1,159.89	1,159.08	Fuc-GT1 (t18:0/20:0)
$[M-3H]^{3-}$	805.40	805.38	GQ1 (d18:1/18:0)
$[M+Na-4H]^{3-}$	812.73	812.71	
$[M+2Na-4H]^{2-}$	1,230.43	1,230.56	
$[M+3Na-5H]^{2-}$	1,241.43	1,241.55	
$[M-3H]^{3-}$	814.74	814.72	GQ1 (d18:1/20:0)
$[M+Na-4H]^{3-}$	822.07	822.05	
$[M+2Na-4H]^{2-}$	1,244.42	1,244.57	
$[M-3H]^{3-}$	819.38[a]	819.38	*O*-Ac-GQ1 (d18:1/18:0)
$[M+Na-4H]^{3-}$	826.73[a]	826.71	

Reprinted with permission from [88]
d dihydroxy sphingoid base, *t* trihydroxy sphingoid base
[a]Low intensity ions

later developmental phase. Therefore, based on the MS data it was hypothesized that these modifications of cerebellum gangliosides start to reach a relevant level only later, during extrauterine brain development and maturation. By combining in high-throughput mode fully automated nanoESI chip MS infusion with CID MS^2 at variable RF signal amplitudes, a complete structural characterization of two cerebellum-associated ganglioside species, GD1 (d18:1/20:0) and GM2 (d18:1/19:0), was accomplished in only 1 minute of signal acquisition for each MS^2 experiment and with a sample consumption, situated in the femtomole range.

8.3.3 Defined Cerebrum Regions

8.3.3.1 Hippocampus

Hippocampus is a brain region situated inside the temporal lobe, under the cerebral cortex. This particular sector of the brain belongs to the limbic system and plays a vital role in learning, memory formation, and orientation [45]. Hippocampus is one of the first brain areas displaying the neurodegenerative pathology in Alzheimer's disease (AD). Consequently, the primary clinical symptoms of AD are due to the impairment of hippocampus functions and are characterized by disorientation and progressive loss of memory. From this perspective, the evaluation of ganglioside expression in this brain region has a particular clinical relevance for early detection of alteration that is attributable to Alzheimer's disease [55].

Mass spectrometry was introduced in human hippocampus ganglioside analysis in 2006 via an analytical platform encompassing fully automated chip-base nano-ESI on a NanoMate robot coupled to a high-resolution hybrid QTOF instrument [82]. Comparative analysis of gangliosides in fetal of different gestational weeks vs. adult hippocampus revealed a specific expression and structure of individual species, which depend on the brain development stage. Hence, polysialylated structures were found markers of the earlier fetal developmental stage. Trisialylated species were found better expressed in fetal hippocampus of 15 (FH15) and 17 (FH17) gestational weeks, while tetrasialylated components, in particular GQ1 forms, were found associated to FH15 (Fig. 8.5, Table 8.2) and not at all to FH17 and adult hippocampus of 20 years of age (AH20). Remarkably, tandem MS carried out using CID at low ion acceleration energies provided data consistent with GQ1b isomer in F15 hippocampus.

In contrast with sialylation, Fuc or GalNAc modifications of the ganglioside chain were found associated to the tissue in its either later fetal development stage or adulthood; fucosylated components were identified in FH17 (Fuc-GD1) and AH20 (Fuc-GD1 and Fuc-GT1) but not observed in FH15, while a GalNAc-GD1 species was found only in AH20. Actually, in this study using nanoESI chip MS characterized by high ionization efficient, sensitivity, and reproducibility, the topo-specific pattern and development-induced ganglioside modifications were for the first time reliably demonstrated and correlated with the biological expectations.

8.3.3.2 Caudate Nucleus

Caudate nucleus (*nucleus caudatus*, NC) is an essential part of the brain learning and memory system located within the basal ganglia of the brain of many animal species. Other functions which may be associated with this part of the brain include voluntary motor control and a regulating role in controlling the threshold potential

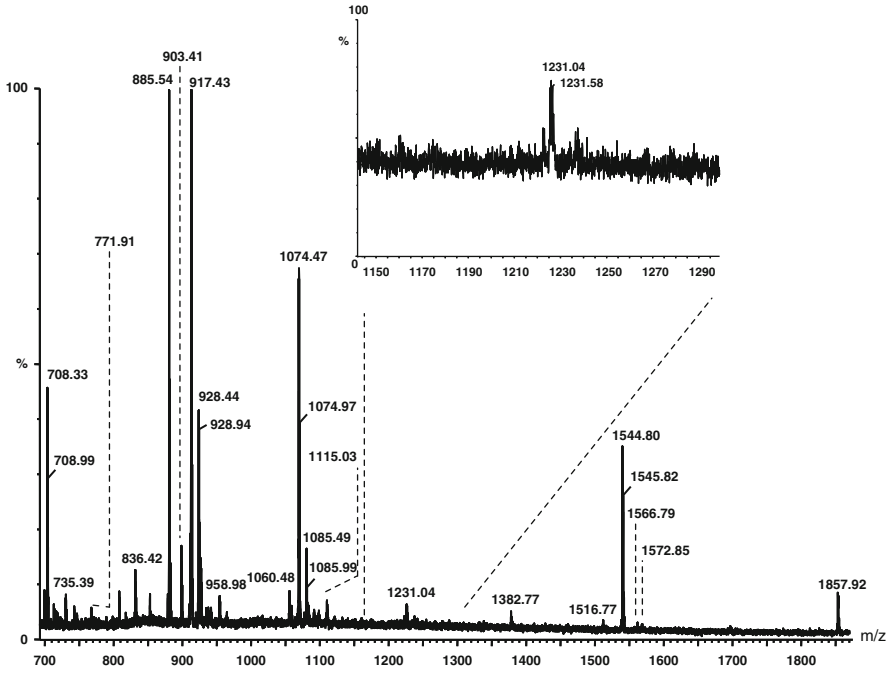


Fig. 8.5 NanoESI chip QTOF MS in the negative ion mode of the FH15 ganglioside mixture. ESI voltage 1.80 kV; cone voltage 40–100 V; Solvent MeOH. Sample concentration 3 pmol/μL; Inset: expanded area *m/z* (1,150–1,300). Reprinted with permission from Vukelić et al. [82]

Table 8.2 Comparative assignment of the ganglioside species identified by negative ion mode nanoESI chip QTOF MS in fetal and adult hippocampus

	m/z (monoisotopic)					
Proposed structure	Experimental	Theoretical	Molecular ion	FH15	FH17	AH20
GM3(d18:1/16:0)	1,151.94	1,151.71	$[M-H]^-$	+	–	–
GM2(d18:1/18:0)	1,382.77	1,382.82	$[M-H]^-$	+	+	+
GM1(d18:1/18:0)	771.91	771.93	$[M-2H]^{2-}$	+	+	+
	1,544.80	1,544.85	$[M-H]^-$			
	1,566.79	1,566.83	$[M+Na-2H]^-$			
GM1(d18:1/20:0)	786.05	785.92	$[M-2H]^{2-}$	+	+	+
	1,572.85	1,572.85	$[M-H]^-$			
GD3(d18:1/18:0)	734.88	734.91	$[M-2H]^{2-}$	+	+	+
GD2(d18:1/18:0)	836.42	836.45	$[M-2H]^{2-}$	+	+	+
GD1(d18:1/16:0)	903.41	903.46	$[M-2H]^{2-}$	+	+	–
GD1(d18:1/18:0)	917.43	917.48	$[M-2H]^{2-}$	+	+	+
	928.44	928.47	$[M+Na-3H]^{2-}$			
	1,857.92	1,857.95	$[M+Na-2H]^-$			
	1,879.94	1,879.93	$[M+2Na-3H]^-$			

(continued)

Table 8.2 (continued)

Proposed structure	*m/z* (monoisotopic)		Molecular ion	FH15	FH17	AH20
	Experimental	Theoretical				
GD1(d18:1/20:0)	931.44	931.49	$[M-2H]^{2-}$	+	+	+
	942.39	942.48	$[M+Na-3H]^{2-}$			
GD1(d18:1/24:1)	958.98	958.52	$[M-2H]^{2-}$	+	+	+
Fuc-GD1(t18:0/18:0)	999.47	999.51	$[M-2H]^{2-}$	–	+	+
Fuc-GD1(t18:0/20:0)	1,013.89	1,013.53	$[M-2H]^{2-}$	–	–	+
GalNAc-GD1(d18:1/18:0)	1,018.91	1,019.02	$[M-2H]^{2-}$	–	–	+
GT1(d18:1/18:1)	1,060.48	1,061.03	$[M-2H]^{2-}$	+	–	–
GT1(d18:1/18:0)	708.33	708.35	$[M-3H]^{3-}$	+	+	+
	1,074.47	1,074.02	$[M+Na-3H]^{2-}$			
	1,085.49	1,085.01	$[M+2Na-4H]^{2-}$			
GT1(d18:1/20:0)	717.91	717.69	$[M-3H]^{3-}$	+	+	+
	1,088.39	1,088.03	$[M+Na-3H]^{2-}$			
	1,099.40	1,099.02	$[M+2Na-4H]^{2-}$			
GT1(d18:1/24:1)	1,115.03	1,115.06	$[M+Na-3H]^{2-}$	+	+	+
Fuc- GT1(d18:1/17:0)	1,129.40	1,129.05	$[M-2H]^{2-}$	–	–	+
GQ1(d18:1/18:0)	1,231.04	1,230.56	$[M+2Na-4H]^{2-}$	+	–	–
	1,241.81	1,241.55	$[M+3Na-5H]^{2-}$			

Reprinted with permission from Vukelic et al. [81]
+ the structure was detected; – the structure was not detected

for general neuron activation and thereby preventing overload through positive feedback loops [21, 54]. Various disorders such as depression, Huntington's disease, Parkinson's disease, attention deficit disorder, schizophrenia, and obsessive compulsive disorder are strongly related with this part of the brain [6, 79, 86]. From this perspective, certainly, the evaluation of ganglioside expression in this brain region is essential for neurological studies. The literature data related to NC glycolipid investigation using other methods than MS is limited to a report on a few identified species. The first consistent compositional and structural analysis of gangliosides extracted from adult human NC was carried out using tandem mass spectrometry on a high-capacity ion trap (HCT) instrument equipped with nanoESI chip in the negative ion mode [65]. Over 80 components differing in the carbohydrate chain and ceramide composition, with a large range of fatty acids, were for the first time discovered in NC. NC was found dominated by mono-, di-, and trisialylated ganglioside structures. Twenty-two species were found to contain one sialic acid (Neu5Ac) moiety, 28 species disialylated, and 20 trisialylated. MS analysis revealed that the carbohydrate chains of gangliosides in human NC vary from one to four monosaccharide units, while the Cer moiety exhibits a high heterogeneity in the fatty acid and/or sphingoid base composition. Also, four tetrasialylated and two pentasialylated structures, i.e., a GP3 (d18:1/12:0) and a GP2 (d18:1/20:0), were detected and identified in the NC extract. The incidence of species with such a high sialylation grade in other adult brain areas was not reported before. Therefore, the identification in the NC of these structures has a particular biological and biochemical value. Besides, MS screening provided information on the incidence in NC of

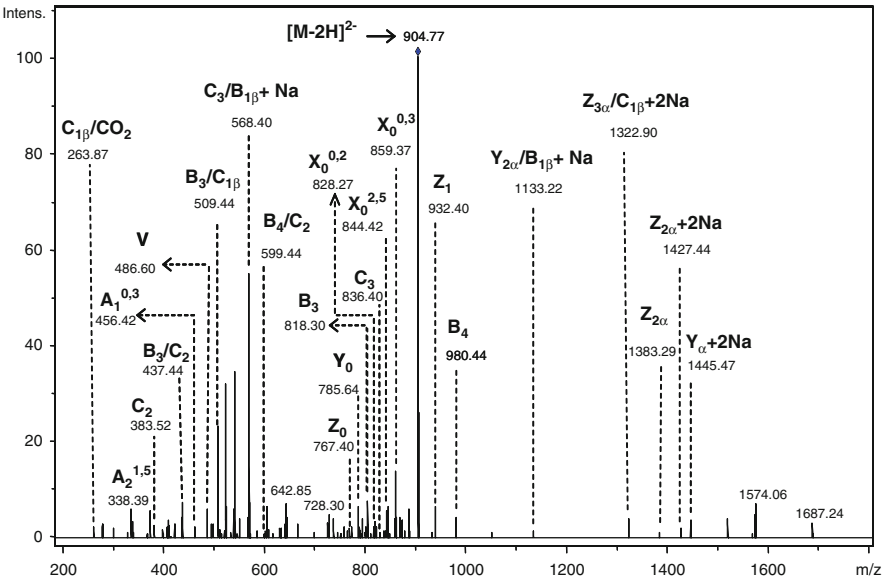


Fig. 8.6 NanoESI chip HCT CID MS^2 of the $[M+2Na-4H]^{2-}$ at *m/z* 904.77 corresponding to GM1 (d18:1/34:2) ganglioside species detected in adult human NC. Solvent: MeOH. Sample concentration: 3 pmol/μL. Acquisition time: 1.5 min. Chip ESI −0.85 kV; capillary exit −30 V. The assignment of the fragment ions is according to the nomenclature [14]. Reprinted with permission from Serb et al. [65]

several glycoforms modified by *O*-Fuc and *O*-Ac attachments and/or lactonization. In addition to these findings, as no less than 14 species with fatty acids chains exceeding 30 carbon atoms were discovered in NC, the presence of unusually long fatty acid chains in Cer composition is another feature, which distinguishes fundamentally NC from the other brain regions. In the second stage of NC ganglioside research by MS, a NC-associated structure exhibiting a rare ceramide composition with 34 carbon atoms in the fatty acid chain was fragmented by CID MS^2 (Fig. 8.6). Adequate fragmentation conditions induced the formation of parent ions diagnostic not only for the carbohydrate sequence but also for the highly specific and at the same time unusual fatty acid chain within the lipid part. Hence, the spectrum in Fig. 8.6 and the derived fragmentation scheme in Fig. 8.7 show that the fatty acid chain consisting of 34 carbon atoms bearing two double bounds is very well documented by the V-type fragment ion.

8.3.3.3 Sensory and Motor Cortex

Sensory cortex (SC) is a brain region located on the lateral postcentral gyrus of the parietal lobe and represents the main sensory receptive area. Damages of this region might affect the ability to recognize objects and, in an extreme condition, induce impairment in the contra-lateral side perception of the subject outer environment [11].

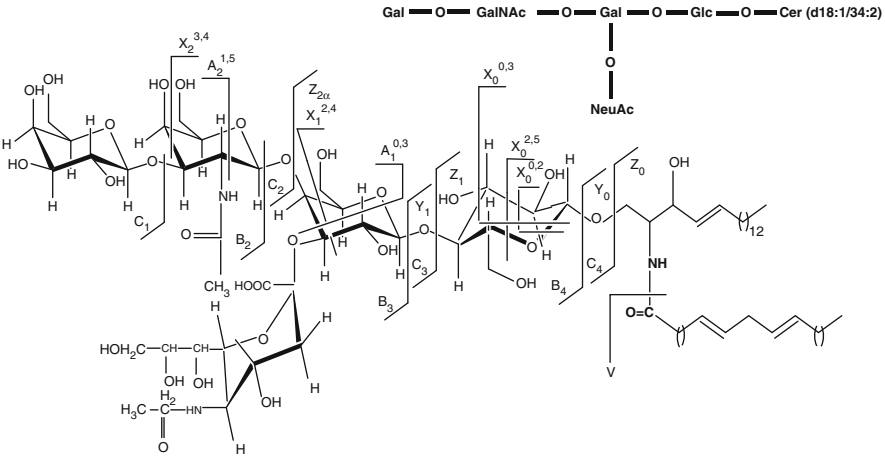


Fig. 8.7 Fragmentation scheme of the $[M+2Na-4H]^{2-}$ ion at *m/z* 904.77 corresponding to GM1 (d18:1/34:2) ganglioside species detected in adult human NC. Reprinted with permission from Serb et al. [65]

Motor cortex (MC) represents an area of the cerebral cortex responsible for the planning, control, and completing voluntary movements [12]. MC is the main contributor to generating neural impulses, which pass down to the spinal cord and control the execution of movements. Injury of this brain area creates spastic contralateral weakness, which is most prominent in the distal extremities [12].

Being situated in close proximity, MC and SC are often both affected by various disorders and accidents such as cerebral vascular accident, amyotrophic lateral sclerosis (ALS), multiple sclerosis, Ramsay Hunt syndrome, Parkinson disease (PD) and brain cancer [4, 34, 48]. Lately, positron emission tomography-PET [50], and functional magnetic resonance imaging-fMRI [10] emerged as highly efficient imaging techniques for the study of functional reorganization of MC and SC.

Several studies demonstrated that use of GM1 ganglioside resulted in significant symptom reduction in Parkinson's disease (PD) patients [63]. Also, marked aberrations in brain ganglioside profiles detected in MC were found related to ALS [76].

The actual trends in the management of MC and SC afflictions are oriented towards a combination of imaging methods for diagnostic and development of novel and effective drugs for treatment. In this regard, the last findings indicate that gangliosides are in the focus of research as possible bases for vaccines and drug therapies-oriented medication.

In a recent study high-resolution tandem MS on a QTOF instrument was developed and for the first time applied for the assessment of ganglioside composition and structure in specimens of human MC and SC [18]. Screening of the MC and SC extracts in the negative ion mode nanoESI, under identical conditions, disclosed the presence of 90 distinct gangliosides in both native mixtures, of which 48 in MC and 42 in SC. Remarkably, in MC, 2 tetra-, 11 tri-, and 20 disialylated ganglioside components were discovered in contrast with SC where only 8 tri- and 13 disialylated

molecules were detected. Of high importance is the identification in MC of tetrasialylated GQ1(d18:1/18:0) and GQ1(d18:1/20:0) species which were not observed in SC. MC mixture was also found to contain a higher number of fucosylated and acetylated species as well as gangliosides with trihydroxylated sphingoid bases. These findings demonstrated that the sialylation degree, fucosylation and acetylation pattern, and sphingoid base hydroxylation of the gangliosides expressed in the two cortex regions are to be correlated with the specific functions of MC and SC.

CID at low ion acceleration energies applied to elucidate the structure of fucosylated species Fuc-GM1 (d18:1/20:0) revealed the presence in MC of an isomer exhibiting both Neu5Ac and Fuc residues linked at the inner galactose. Since this structure was not reported previously, seemingly Fuc-GM1 (d18:1/20:0) species with sialylated and fucosylated inner galactose is associated to the motor cortex.

8.3.3.4 Neocortex

Neocortex is the newest area in brain development of mammals [29, 38]. In humans it accounts for about 76 % of the brain volume, as the outer layer of the cerebral hemispheres, made up of six layers. Among mammals human neocortex is the largest [31], allowing a new level of advanced behavior such as particular social behavior, language, and high-level consciousness. Neocortex is involved also in other elevated functions such as sensory perception, generation of motor commands, spatial reasoning, and all of the conscious thoughts. Basically, the neocortex enables the most complex mental activity, which is associated with humans.

Gangliosides extracted from fetal neocortex of 36 gestational weeks were analyzed by mass spectrometry using a HCT with nanaelectrospray chip-based infusion on a NanoMate robot [17]. Eighty-nine ganglioside species differing in their carbohydrate chain structure and ceramide composition were identified in the neocortex within only 2 min of signal acquisition for single MS screening experiments. A notable characteristic is the identification in neocortex of six GQ species and also asialo- components having the oligosaccharide composition $HexNAcHex_2$ and $HexHexNAcHex_2$.

Ganglioside chains decorated with peripheral attachments such as Fuc and/or *O*-Ac were previously reported as associated to the tissue in its advanced developmental phase. Indeed, MS investigation of neocortex provided relevant data documenting 12 fucosylated species, of which three GT1, two GT2, two GD2, two GM4, one GT3, one GD1 and one GA2, seven *O*-acetylated structures corresponding to GM1, GD3, GD1, GT1, and GQ1, exhibiting high heterogeneity in their ceramide motifs. Contrastingly, in other fetal brain regions a much lower number of *O*-acetylated species and no species modified by fucosylation were found. This particular aspect suggests that Fuc- and *O*-Ac- gangliosides are rather associated with the neocortex as the most recently developed human brain region, where the high level of thinking takes place and specialized integrative tasks are processed. The large number of ganglioside species identified in neocortex is to be correlated with the complexity of the functions coordinated by this region and with its central role in the higher mental functions as language, memory, thinking,

attention, abstraction, and perception in humans. The results obtained in direct comparative MS analyses of neocortex vs. brain lobes [17] have also shown that differences in ganglioside expression in fetal human brain are dependent rather on the phylogenetic development than topographic factors. This feature, discovered exclusively by MS, provides a novel insight into the major role of gangliosides in human brain evolution and advancement of its functions in comparison with the other mammals.

8.4 Gangliosides in Pathological Brain

8.4.1 Neurodegenerative Diseases

Neurodegenerative diseases represent a group of ailments characterized by the deterioration of the structure or function of neurons up to the death of neurons. Neurodegenerative diseases including Parkinson's, Alzheimer's, Huntington's diseases, ALS, anencephaly, spinocerebellar ataxia, and dementia arise as a consequence of neurodegenerative processes taking place from molecular to systemic level.

Gangliosides trigger a variety of neurodegenerative diseases [63, 64] from autoimmune-induced neuropathies caused by anti-ganglioside auto-antibodies to lipidoses, a group of inherited metabolic disorders caused by the accumulation of gangliosides in cell bodies due to a obstruction of their catabolic pathways [59].

Gangliosides were also shown to have neuritogenic and neuronotrophic activity [8] and to facilitate repair of neuronal tissue after mechanical, biochemical, or toxic injuries. In AD administration of GM1, having a high affinity for Aβ resulted in the reduction of Aβ level in the brain, suggesting that GM1 might serve as the therapeutic agent reducing/preventing brain amyloidosis by sequestering the plasma Aβ [3]. On the other hand, interactions of Aβ (1–40) with ganglioside-containing membranes, particularly with membrane rafts enriched in GM1 and GM3, were hypothesized to be involved in the pathogenesis of AD [52]. Using conventional high-performance thin-layer chromatographic separation/detection, immunochemical, and immunohistochemical detection methods, specific changes in ganglioside expression and quantity in investigated human brain regions in AD disease were discovered [32, 35].

MALDI TOF MS in combination with a method for transferring lipids separated on a TLC-plate to a poly-vinylidene difluoride membrane and direct mass spectrometric analysis of the individual lipids on the membrane was applied to the analysis of individual lipids and ganglioside molecular species in neural diseases [78]. In the case of Alzheimer's disease the expression of GD1b and GT1b gangliosides was found diminished in the hippocampal gray matter than in the hippocampal gray matter of patients with Parkinson's disease or the control patients [72]. The analysis of single ganglioside molecular species in patients with Alzheimer's disease demonstrated a major decrease of species exhibiting (20:1/18:0) ceramide composition.

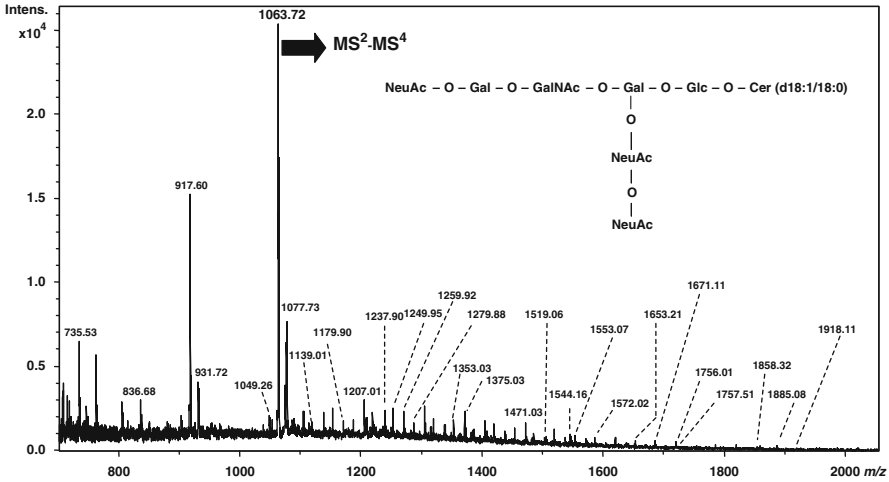


Fig. 8.8 NanoESI chip HCT MS in negative ion mode of native ganglioside mixture from glial islands of anencephalic fetus. Solvent MeOH. Sample concentration 5 pmol/μL. Acquisition time 7 min. Chip ESI −0.8 kV; capillary exit −50 V. Reprinted with permission from Almeida et al. [1]

These results imply that Alzheimer's disease is a ganglioside metabolic disease affecting the hippocampal area. Such findings explain in part the major symptoms of the disease, which are related to abilities promoted by hippocampus.

In anencephaly, a congenital malformation of the fetal brain occurring when the cephalic end of the neural tube fails to close [7], the first assessment of ganglioside composition by MS was reported in 2001 [80]. In this study nanoESI QTOF MS and tandem MS were applied for the first time for compositional and structural identification of native gangliosides from anencephalic cerebral residue. By this approach it was found that the total ganglioside concentrations in the anencephalic cerebral remnant and in cerebellum were significantly lower than in the corresponding regions of the age-matched brain used as control. In the cerebral remnant, GD3, GM2 and GT1b, GM1b nLM1, and nLD1 were found highly expressed. Oppositely, GD1a was found better expressed in the anencephalic cerebellum, while GQ1b was reduced in both anencephalic regions. In agreement with previously acquired information using immunochemical methods, by nanoESI MS, members of the neolactoseries gangliosides were also discovered in anencephalic brain tissues.

Supplementary information supporting a major modification of ganglioside expression in anencephalic vs. age-matched normal brain tissue were collected by a superior methodology based on fully automated nanoESI chip MS and multistage MS using the coupling of NanoMate robot to HCT MS tuned in the negative ion mode [1]. The ganglioside mixture extracted from glial islands of fetal anencephalic brain tissue was investigated in comparison with the gangliosides from a normal fetal frontal lobe. Under identical instrumental and solution conditions, 25 distinct species in the mixture from anencephalic tissue (Fig. 8.8, Table 8.3) vs. 44 of

Table 8.3 Assignment of the major ions detected by negative ion mode nanoESI chip HCT MS in the ganglioside mixture extracted from the glial islands of anencephalic fetus

m/z monoisotopic	Molecular ion	Proposed structure
735.12	$[M-2H]^{2-}$	GD3(d18:0/18:0)
836.71	$[M-2H]^{2-}$	GD2(d18:1/18:0)
851.60	$[M-H]^{-}$	GD2(d18:1/20:0)
918.08	$[M-2H]^{2-}$	GD1(d18:1/18:0)
931.72	$[M-2H]^{2-}$	GD1(18:1/20:0)
952.80	$[M-H]^{-}$	GD1(d18:1/23:0)
1,037.60	$[M-H]^{-}$	$HexNAcHex_2Cer$ (d18:0/14:0) or (d16:0/16:0)
1,041.60	$[M-H]^{-}$	GM4 (d18:1/20:2)
1,049.18	$[M-2H]^{2-}$	GT1(18:1/16:0)
1,063.33	$[M-2H]^{2-}$	GT1(d18:1/18:0)
1,065.63	$[M-H]^{-}$	$HexNAcHex_2Cer$ (d18:0/16:0)
1,077.71	$[M-2H]^{2-}$	GT1(d18:1/20:0)
1,104.78	$[M-2H]^{2-}$	GT1(d18:1/24:0)
1,139.01	$[M-H]^{-}$	GM3(d18:1/14:0) or (d18:1/h14:0) or $HexNAcHex_2Cer$ (d18:1/22:4)
1,151.71	$[M-H]^{-}$	GM3 (d18:1/16:0)
1,165.80	$[M-H]^{-}$	$HexNAcHex_2Cer$ (t18:0/22:0) or (d18:0/h22:0) or (d18:2/24:4)
1,167.82	$[M-H]^{-}$	GM3 (t18:1/16:0) or (d18:1/h16:0) or $HexNAcHex_2Cer$ (d18:1/24:4)
1,179.74	$[M-H]^{-}$	GM3 (d18:1/18:0)
1,181.75	$[M-H]^{-}$	GM3 (d18:0/18:0)
1,206.77	$[M-H]^{-}$	GM3(d18:1/20:0)
1,221.33	$[M-H]^{-}$	GM3(d18:1/18:0)
1,235.81	$[M-H]^{-}$	GM3 (d18:1/22:0)
1,237.81	$[M-H]^{-}$	GM3 (d18:0/22:0)
1,249.78	$[M-H]^{-}$	*O*-Ac-GM3 (d18:1/20:0) (or GM3 (18:1/23:0)
1,253.02	$[M-H]^{-}$	$HexHexNAcHex_2Cer$(d18:1/18:0)
1,259.79	$[M-H]^{-}$	GM3 (d18:1/24:2)
1,261.81	$[M-H]^{-}$	GM3 (d18:1/24:1)
1,263.83	$[M-H]^{-}$	GM3 (d18:1/24:0)
1,265.84	$[M-H]^{-}$	GM3 (d18:0/24:0)
1,275.80	$[M-H]^{-}$	*O*-Ac-GM3 (d18:1/22:1) (or GM3 (20:1/23:1)
1,277.80	$[M-H]^{-}$	*O*-Ac-GM3 (d18:1/22:0) (or GM3 (20:1/23:0)
1,279.81	$[M-H]^{-}$	*O*-Ac-GM3 (d18:0/22:0) (or GM3 (20:0/23:0)
1,301.82	$[M-H]^{-}$	*O*-Ac-GM3 (d18:1/24:2)
1,353.03	$[M-H]^{-}$	GM2 (d18:1/16:0)
1,354.79	$[M-H]^{-}$	GM2 (d18:1/16:0)
1,383.21	$[M-H]^{-}$	GM2 (d18:1/18:0)
1,384.81	$[M-H]^{-}$	GM2 (d18:0/18:0)
1,437.01	$[M-H]^{-}$	GM2 (d18:1/22:0)
1,442.78	$[M-H]^{-}$	GD3 (d18:1/16:0)
1,444.80	$[M-H]^{-}$	GD3 (d18:0/16:0)
1,468.79	$[M-H]^{-}$	GD3 (d18:1/18:1)

(continued)

Table 8.3 (continued)

m/z monoisotopic	Molecular ion	Proposed structure
1,471.03	$[M-H]^-$	GD3(d18:1/18:0)
1,472.83	$[M-H]^-$	GD3 (d18:0/18:0)
1,519.10	$[M-H]^-$	GM1, nLM1 and/or LM1 (d18:0/16:0)
1,544.16	$[M-H]^-$	GM1, nLM1 and/or LM1 (d18:1/18:0)
1,545.20	$[M-H]^-$	GM1, nLM1 and/or LM1 (d18:1/18:0)
1,805.23	$[M-H]^-$	GT3(d18:1/18:0)
1,836.40	$[M-H]^-$	GD1(d18:1/18:0)

Reprinted with permission from [1]

which 4 asialylated in the normal frontal lobe were for the first time identified. These results indicate that a high number of ganglioside species associated to anencephaly could be ionized and discriminated only by employing chip-based electrospray. Interestingly, GD3 (d18:1/18:0), GD2 (d18:1/18:0), GM1 (d18:1/18:0), and their neolacto or lacto-series isomers were detected as ions of similar low abundances in both mixtures, while GT1 (d18:1/18:0) and GD1 (d18:1/18:0) were found highly expressed in ancencephalic brain tissue (Table 8.4). Moreover, several structures such as GT1, GQ1, and GQ2 emerged clearly associated to anencephaly. This prominent occurrence of polysialylated structures in anencephaly is basically an effect to be used for the diagnosis of the brain development stagnation, characteristic to this disease [1]. In view of the results obtained by MS/MS, the earlier report [80] has postulated that GT1b is one of the disease markers; however, because of the limited information obtained by fragmentation analysis in a single CID stage, validation of sialylation sites could not be accomplished. To close this gap, a nanoESI chip CID MS^n protocol [1] for fine investigation of the anencephaly-specific GT1 (d18:1/18:0) species was elaborated (Fig. 8.9a–d). The beneficial combination of chip infusion, high capacity of ion storage, and multistage sequencing rendered ions diagnostic for the disialo ($Neu5Ac_2$) element localization at internal galactose moiety. Hence, chip MS results disclosed for the first time GT1b species in the cerebral remnant of anencephalic brain.

8.4.2 Primary Brain Tumors

Primary brain tumors account for approximately 2 % of all adult malignancies and are responsible for about 7 % of years of life lost prior to age 70. In childhood, 20 % of all malignancies identified prior to 15 years of age are primary brain tumors. This situates the primary brain tumors on the second place among the most frequent type of cancer identified in children. For primary brain tumors therapy, drastic surgical resection radiation and chemotherapy are especially applied [43]. These methods are frequently problematic firstly because of collateral brain tissue sensitivity to disruption and secondly because of the toxicity accompanied by side effects of the therapeutic agents.

Table 8.4 Comparative overview upon gangliosides and asialo-gangliosides detected by negative ion mode nanoESI chip HCT MS in the glial islands of anencephalic fetus and in the frontal lobe of healthy fetal brain

Ganglioside species	Proposed structure	Anecephaly	Normal frontal lobe
GM1	nLM1 and/or LM1 (d18:0/16:0)	+	+
	nLM1 and/or LM1 (d18:1/18:0)	+	+
	nLM1 and/or LM1 (d18:0/20:0)	+	−
	(d18:1/16:0)	+	+
	(d18:1/18:0)	−	+
	(d18:1/22:0)	−	+
GM3	(d18:1/14:0) or (d18:1/h14:0) or $HexNAcHex_2Cer$ (d18:1/22:4)	+	+
	(d18:1/16:0)	−	+
	(t18:1/16:0) or (d18:1/h16:0) or $HexNAcHex_2Cer$(d18:1/24:4)	−	+
	(d18:1/18:0)	+	+
	(d18:0/18:0)	−	+
	(d18:1/20:0)	+	+
	(d18:1/22:0)	−	+
	(d18:0/22:0)	+	+
	(d18:1/24:2)	+	+
	(d18:0/24:0)	+	+
	(d18:1/24:1)	−	+
	(d18:1/24:0)	−	+
	O-Ac-GM3 (d18:1/20:0) or GM3 (18:1/23:0)	−	+
	O-Ac-GM3 (d18:1/22:1) or GM3 (20:1/23:1)	−	+
	O-Ac-GM3 (d18:1/22:0) (or GM3 (20:1/23:0)	−	+
	O-Ac-GM3 (d18:0/22:0) (or GM3 (20:0/23:0)	+	+
	O-Ac-GM3 (d18:1/24:2)	−	+
GM4	(d18:1/20:2)	−	+
GD1	(d18:1/18:0)	+	+
	(d18:1/20:0)	+	+
	(d18:1/23:0)	−	+
	(d18:1/24:1)	+	−
	(d18:0/18:0)	+	−
GD2	(d18:1/18:0)	+	+
	(d18:1/18:1)	+	−
	(d18:1/24:1)	+	−
	(d18:1/24:0)	+	−
	(d18:1/20:0)	−	+
GD3	(d18:0/18:0)	+	+
	(d18:1/16:0)	−	+
	(d18:0/16:0)	−	+
	(d18:1/18:1)	−	+
	(d18:1/18:0)	+	+
	(d18:1/24:1)	+	−

(continued)

Table 8.4 (continued)

Ganglioside species	Proposed structure	Anecephaly	Normal frontal lobe
GT1	(18:1/16:0)	+	+
	(d18:1/18:0)	+	+
	(d18:0/20:0)	+	−
	(d18:1/20:0)	+	+
	(d18:1/24:0)	−	+
GT3	(d18:1/18:0)	−	+
	(d18:1/24:0)	−	+
	(d18:1/24:1)	+	−
GQ1	(d18:1/18:0)	+	−
Asialo-species	$HexNAcHex_2Cer$ (d18:0/14:0) or (d16:0/16:0)	−	+
	$HexNAcHex_2Cer$ (d18:0/16:0)	−	+
	$HexNAcHex_2Cer$ (t18:0/22:0) or (d18:0/h22:0) or (d18:2/24:4)	−	+
	$HexHexNAcHex_2Cer$ (d18:1/18:0)	−	+

Reprinted with permission from [1]

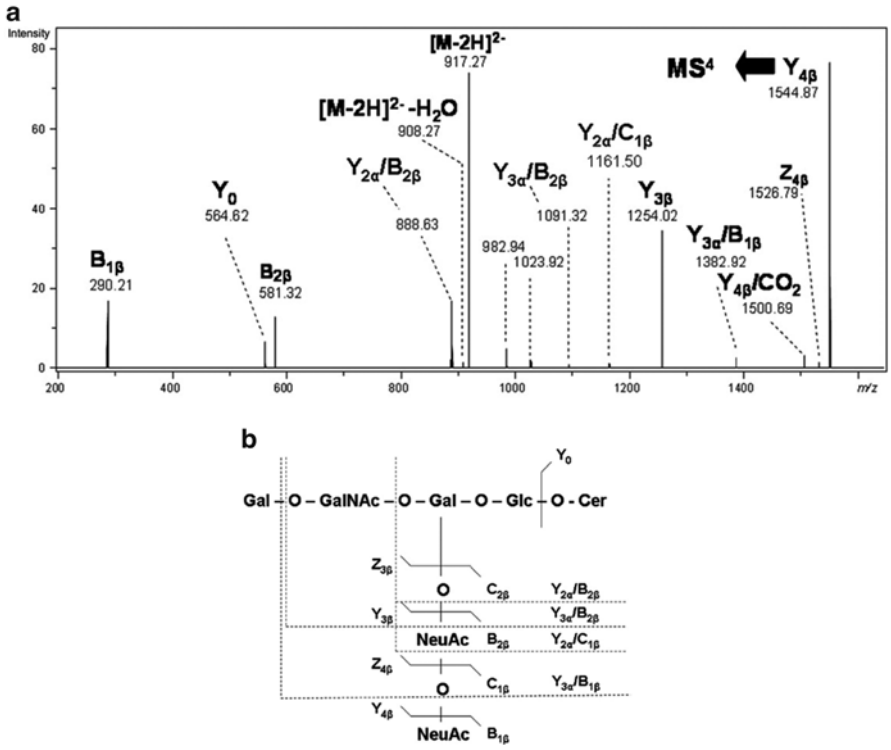


Fig. 8.9 NanoESI chip HCT multistage MS of the doubly charged ion at *m*/*z* 1,063.34 corresponding to GT1 (d18:1/18:0) species from glial islands of anencephalic fetus mixture: (**a**) MS^3. (**b**) fragmentation scheme of ion at *m*/*z* 917.32. (**c**) MS^4. (**d**) fragmentation scheme of the ion at *m*/*z* 1,544.87. The assignment of the fragment ions is according to the nomenclature [14]. Reprinted with permission from Almeida et al. [1]

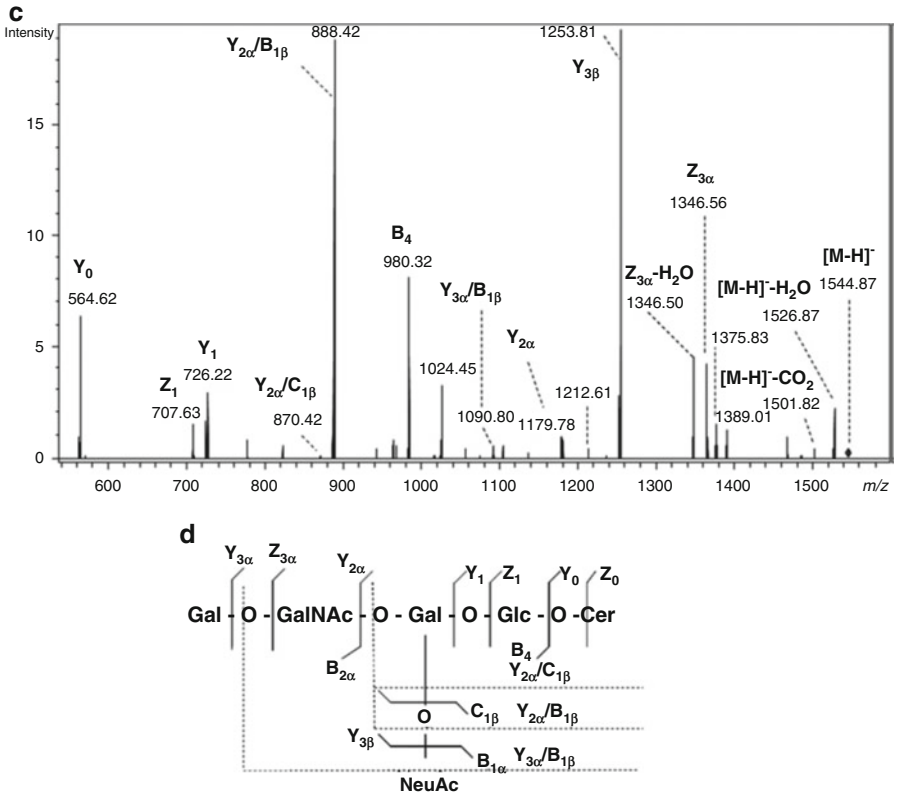


Fig. 8.9 (continued)

Tumorigenesis/malignant transformation is accompanied by aberrant cell surface composition, particularly due to irregularities in glycoconjugate glycosylation pathways. Various glycosyl epitopes constitute tumor-associated antigens [47, 73]. Some of them promote invasion/metastases, while some other suppress tumor progression. Among molecules bearing characteristic glycosyl epitopes causing such effects are also gangliosides. Gangliosides were recently shown to play an important role in brain tumor development, progression, and treatment [5, 37, 46, 66].

Glycosphingolipid-dependent cross-talk between glycosynapses interfacing tumor cells with their host cells has been even recognized as a basis to define tumor malignancy [23]. Specific changes of gangliosides pattern in brain tumors correlating with tumor histopathological origin, malignancy grade, invasiveness, and progression have been observed [67, 74]. Tumor cells of neuroectodermal origin may shed their gangliosides into circulation, resulting in higher ganglioside concentrations in serum [16]. This shedding of gangliosides into interstitial spaces and blood of oncological patients has been suggested to be involved in increased tumor cell growth and lack of immune cell recognition. Although higher concentrations of serum gangliosides in oncological patients comparing to healthy individuals have been reported [60], there have been no data on comparison of ganglioside

concentrations in serum or cerebrospinal fluid before and after surgical removal of tumor in individual patients. Glycoantigens and lipoantigens have been recognized as relevant and potentially valuable diagnostic and prognostic markers and tumor molecular targets for development/production of specific antitumor drugs [41] such as glycolipids-based vaccines, but their investigation in this regard has been neglected comparing to proteins.

Except for mass spectrometry, several biophysical and biochemical methods such as Raman spectroscopy, immunocytochemistry, flow cytometry, western blotting, transwell assays, and mid-infrared spectromicroscopy were developed and optimized for the determination of lipid/glycolipid molecular markers in brain tumors [17, 36, 69, 70]. Nevertheless, a large number of low abundant tumor-associated species could not be detected by these approaches. Using these methods, the data acquired on ganglioside composition in human brain tumors, sera and cerebrospinal fluid were restricted to several major species; many less abundant species have not been structurally characterized. This emphasized a need for detailed and systematic screening and structural characterization of brain tumor glycoconjugate composition, which could adequately be achieved only combining up-to-date, ultrasensitive, high-resolution methodological approaches of detection and sequencing of biomolecules, such as advanced mass spectrometry combined with bioinformatics for data interpretation.

8.4.2.1 Benign Tumors

Meningioma

Meningiomas are a group of tumors of the meninges, the membranes surrounding the brain and spinal cord. Meningiomas occur from the arachnoid cap cells, which enclose and adhere to the dura mater.

According to WHO [42] meningiomas are classified as benign (WHO grade I), atypical (WHO grade II), and malignant/anaplastic (WHO grade III). Ninety percent of all diagnosed meningiomas are benign. Due to the reduced incidence of the malignant type, in general meningiomas are regarded as tumors treatable by surgery [24]. Altogether, meningiomas represent about 13–30 % of the primary intracranial neoplasms, being among brain tumors only less frequent than gliomas.

If low-grade meningiomas are treated by surgical resection, which yields permanent healing, higher grade forms require radiation therapy following tumor removal [85]. Meningiomas can usually be surgically resected only if the tumor is superficial on the dural surface and accessible; if invasion of the adjacent bone/tissue already occurred, total removal is not feasible and the treatment is impeded. This particular aspect, which applies to tumors in their latter form, together with the discrepancy between the biological behavior and tumor grade opened lately new research directions in bioanalytics and molecular medicine towards discovery of molecular markers, to allow early tumor detection and development of novel and more efficient therapeutic schemes.

In this context, a complex strategy combining high-performance TLC, laser densitometry, and fully automated negative ion mode nanoESI chip coupled to QTOF MS was designed and applied for mapping of gangliosides in a specimen of human angioblastic meningioma [62]. Qualitative analysis of ganglioside pattern using TLC identified GM3, GM2, GM1, GD3, GD1a (nLD1, LD1), GD1b, and GD2 species in human meningioma. Moreover, comparative TLC analyses indicated considerable differences between the proportions of ganglioside species in meningioma vs. healthy brain tissue and showed for the first time that meningioma tissue is rich in GD1a (nLD1, LD1). However, by high resolution MS with chip-based ESI, 34 distinct species, of which two asialo, were identified. This inventory determined by MS contained many more species than previously reported in meningiomas using other biochemical or biophysical methods. Besides the unexpected number of species, an interesting characteristic revealed by MS is the exclusive expression of species with shorter glycan chains and reduced sialic acid content, i.e., maximum sialylation degree found in meningioma is 2.

Another characteristic of meningioma is the absence of *O*-fucosylation, *O*-acetylation, or *O*-GalNAc modification of the main oligosaccharides chain of ganglioside species found in other brain tumors or brains affected by neurodegenerative disorders. Furthermore, nanoESI chip MS evidenced that despite the low percentage of GM1 fraction in meningioma, a number of eight abundant ions were attributable to nine GM1 forms. CID MS/MS analysis documented that both GM1a and GM1b isomers are expressed in meningioma tissue. The findings related to GM1 advocate for the first time that, beside GM3 species, already known as an indisputable marker of meningioma, GM1 class is also associated to this type of tumor.

Hemangioma

Hemangioma represents a congenital benign tumor or vascular malformation of endothelial cells. The ailment features enlarged blood vessels with a single layer of endothelium accompanied by the absence of neuronal tissue within the lesions. Cavernous hemangioma is the most widespread form of brain hemangioma. This type can originate from any part of the brain and can also occur at any location along the vascular bed [28]. Frontal and temporal lobes are the most common sites of occurrence, with approximately 70 % of these lesions located in the supratentorial region of the brain; the remaining 30 % arise in the infratentorial region.

Hemangioma is most usually diagnosed by imaging techniques such as magnetic resonance imaging. A more specific detection can be accomplished by gradient-echo sequence MRI, able to expose even the tiny lacerations [39]. A practical choice to these methods is the early detection of hemangioma at an incipient stage, based on routine screening and cancer biomarker discovery before clinical symptoms arise. As bioindicators of brain cancer, gangliosides in hemangioma were analyzed by mass spectrometry in an attempt to discover the species associated to this type of tumors [61]. Native ganglioside mixture extracted from brain

hemangioma in the frontal cortex of an adult patient was screened by nanoESI chip high capacity ion trap MS in comparison with age-matched healthy frontal cortex. In contrast to the normal tissue, ganglioside mixture extracted from hemangioma was found dominated by species of short oligosaccharide chains with a reduced overall sialic acid content. From a total of 29 structures identified in hemangioma tissue, 13 were monosialylated species of GM1, GM2, GM3, and GM4-type and 13 disialylated species of GD1 and GD2-type bearing ceramides of variable structure. Only two polysialylated species namely GT1 having (d18:1/20:0) ceramide composition and GT3 with (d18:1/25:1) ceramide were detected. However, none of these trisialylated ganglioside forms were detected in the normal tissue; this aspect shows that such components are either absent or much lower expressed in normal frontal lobe, being associated to hemangioma tumor. Interestingly, two modified ganglioside structure were observed in the ganglioside mixture extracted from hemangioma: *O*-Ac-GD2 (d18:1/23:0) and *O*-Ac-GM4 (d18:0/29:0). The same *O*-Ac-GD2 (d18:1/23:0) species was identified also in the normal tissue of the frontal cortex, and was never reported in malignant tumors [61]. This suggested that *O*-Ac-GD2 structures might be markers of either benign cerebral tumors or tumors with reduced malignancy grade.

According to the presented data on meningioma, apparently the expression of polysialylated gangliosides is regulated in a growth- and development-dependent mode and associated with the type of normal/aberrant brain tissue status.

8.4.2.2 Malignant Tumors

Astrocytoma

Astrocytomas (AcTs) are collection of CNS neoplasms characterized by predominant cell type derived from an immortalized astrocyte. Situated in most parts of the brain, sporadically even in the spinal cord, astrocytomas can induce compression, invasion, and destruction of the neural tissue.

AcTs are categorized in four subtypes according to the growth rate and prospective for proliferation in the adjacent brain tissue [42]: (1) pilocytic astrocytoma (grade I), a slow-growing astrocytoma that usually does not spread to other parts of the CNS; (2) low grade astrocytoma (grade II), a relatively slow-growing type, which can invade the surrounding brain tissue and tends to reappear after treatment; (3) anaplastic astrocytoma (grade III) characterized by a rapidly growing rate, with incursion in the normal brain tissue and rapid recurrence after the treatment; (4) glioblastoma multiforme (grade IV), the most aggressive and highly invasive type displaying necrosis areas and a variety of cells including astrocytes and oligodendrocytes.

Astrocytomas have poor survival rates, considered from diagnosis and the beginning of the treatment as follows: 10 years for pilocytic form, 5 years for patients with low-grade diffuse astrocytomas, 2–5 years for anaplastic astrocytomas, and less than 1 year for patients with glioblastoma [87]. Regrettably, astrocytomas affect young ages, the most cases of pilocytic astrocytoma being discovered in the first two decades

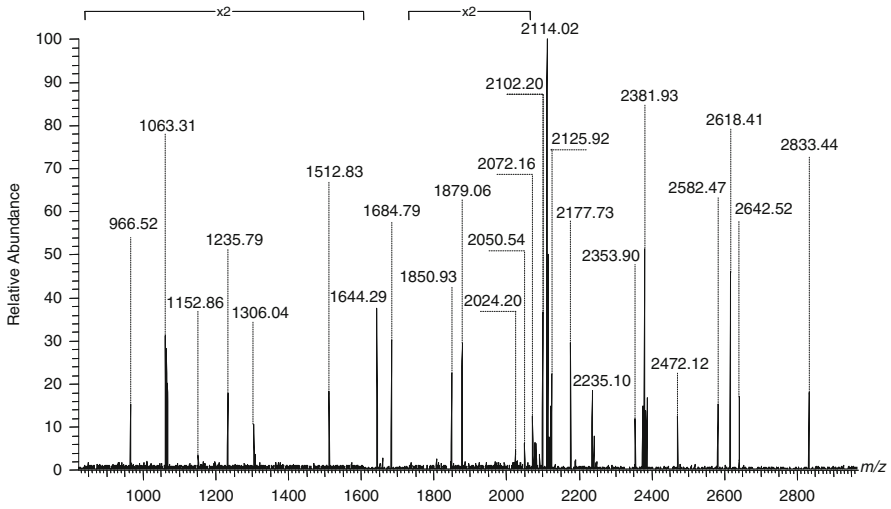


Fig. 8.10 NanoESI Orbitrap MS screening in the negative ion mode of the gangliosides from human astrocytoma. Solvent MeOH. Sample concentration: approximately 2 pmol/μL. Acquisition time 1 min. Flow rate 150 nL/min. Sample volume consumption 150 nL. Spray voltage 1.1 kV. Counterelectrode voltage −35 V. Reprinted with permission from Zamfir et al. [90]

of life [57]. In contrast, low-grade astrocytomas affect especially people aged 30–40 years and, only in low percentage, patients younger than 20 years of age.

Astrocytomas are usually detected by expensive diagnostic imaging methods often in a late phase, when only palliative treatment is possible. The only practical alternative is the early detection of tumor, at a stage when the resection is possible, which should be based on routine screening and biomarker discovery before clinical symptoms arise. Therefore, nowadays, the research is focused on the development of efficient analytical methods able to discover molecular fingerprints, among which the class of gangliosides is the most promising. Based on these premises, in 2013 the first investigation of gangliosides expressed in a low-grade astrocytoma by high-resolution MS on an Orbitrap instrument was initiated [90]. The research was conducted towards the first mapping of gangliosides in specimens from AcT, its surrounding tissue (ST), and a normal control brain tissue from the frontal lobe (NT) under identical conditions. Comparative (−) nanoESI MS screening at a sample concentration of only 2 pmol/μL of gangliosides from AcT (Fig. 8.10), ST (Fig. 8.11), and NT (Fig. 8.12) has led to the following findings summarized in Table 8.5a–c: (1) ganglioside compositions in AcT and ST are altered in comparison to the ganglioside expression in NT; (2) a number of 30 species are associated to the tumor; (3) 14 ganglioside species in AcT, 14 in ST, and only 5 in NT exhibit ceramide moieties with long chain fatty acids exceeding 25 carbon atoms. This finding represents another characteristic distinguishing AcT and ST from NT; (4) ST extract presents high levels of sialylation, fucosylation, and acetylation, typical for malignant transformation, a feature indicating the phenomenon of AcT cells protrusion in the ST.

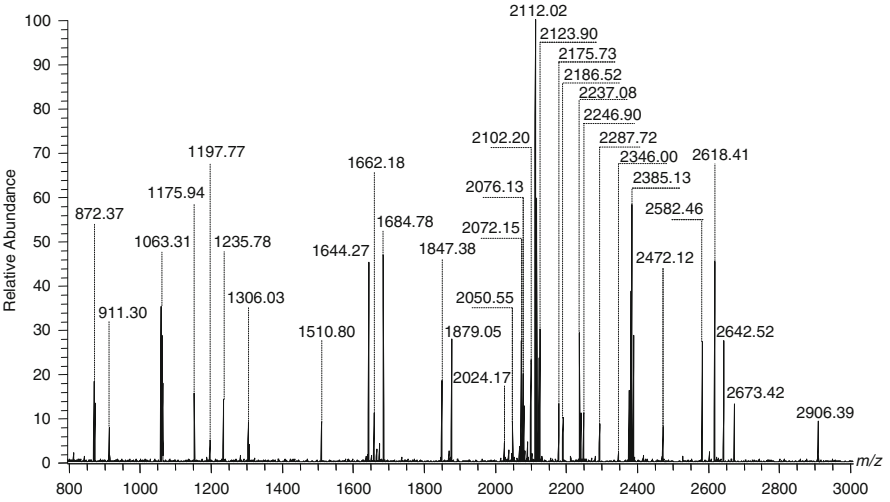


Fig. 8.11 NanoESI Orbitrap MS screening in the negative ion mode of the gangliosides from human astrocytoma surrounding tissue. Conditions as in Fig. 8.10. Reprinted with permission from Zamfir et al. [90]

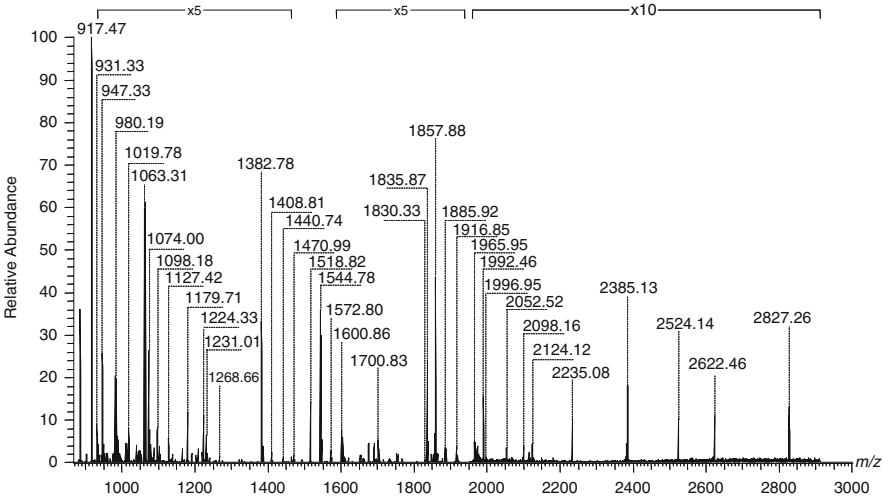


Fig. 8.12 NanoESI Orbitrap MS screening in the negative ion mode of the gangliosides from control tissue originating from the frontal lobe. Conditions as in Fig. 8.10. Reprinted with permission from Zamfir et al. [90]

The MS screening data indicated that AcT, ST, and NT present a common ganglioside structure: GT1(d18:1/18:0) or GT1(d18:0/18:1). CID MS^2–MS^4 experiments (Fig. 8.13) yielded a number of product ions corroborating for the presence of a GT1c isomer (Fig. 8.14) in AcT and ST, but not in NT. A GT1c isomer having

Table 8.5 Proposed composition of single components in the ganglioside mixture from (a) AcT. (b) ST. (c) NT as detected by negative ion mode nanoESI Orbitrap MS

m/z (monoisotopic) experimental	*m/z* (monoisotopic) theoretical	Mass accuracy (ppm)	Molecular ion	Proposed composition
966.52	966.53	10.35	$[M-2H]^{2-}$	GD1, nLD1 and/or LD1 (d18:1/25:0) or (d20:1/23:0)
1,063.31	1,063.33	18.81	$[M-2H]^{2-}$	GT1(d18:1/18:0) or GT1(d18:0/18:1)
1,152.86	1,152.88	17.34	$[M-H]^{-}$	GM3(d18:1/16:0) or GM3(d18:0/16:1)
1,235.79	1,235.81	16.18	$[M-H]^{-}$	GM3 (d18:1/22:0)
1,306.04	1,306.06	15.31	$[M-H]^{-}$	GM3(d18:1/27:0)
1,512.83	1,512.86	19.82	$[M-H]^{-}$	*O*-Ac-GD3 (d18:1/18:0)
1,644.29	1,644.31	12.16	$[M-H]^{-}$	GM1 (d18:0/25:0)
1,684.79	1,684.82	17.80	$[M-H]^{-}$	GD2(d18:1/19:2) or GD2(d18:0/19:3)
1,850.93	1,850.97	21.60	$[M-H]^{-}$	GD1(d18:1/19:0) or GD1(d18:0/19:1)
1,879.06	1,879.10	21.28	$[M-H]^{-}(-H_2O)$	Fuc-GT3(d18:0/17:0) or *O*-Ac-GT3 (d18:2/22:1)
	1,879.09	15.96	$[M+Na-2H]^{-}$	
2,024.20	2,024.24	19.76	$[M+2Na-3H]^{-}$	GD1 (d18:0/28:0)
2,050.54	2,050.60	29.26	$[M-H]^{-}$	GT2(d18:1/24:0) or GT2(d18:0/24:1)
2,072.16	2,072.21	24.13	$[M-H]^{-}$	*O*-Ac-GD1(d18:1/32:1)
2,102.20	2,102.26	28.54	$[M-H]^{-}$	*O*-Ac-GD1(d18:1/34:0)
2,114.02	2,114.09	33.11	$[M-H]^{-}$	GT1(d18:1/17:1) or GT1(d18:0/17:2)
2,125.92	2,125.97	23.51	$[M-H]^{-}$	GT1(d18:1/18:1)
2,177.73	2,177.80	32.13	$[M-H]^{-}$	GT2 (d18:0/33:0)
2,235.10	2,235.16	26.84	$[M-H]^{-}$	GT1(d18:1/26:2)
2,353.90	2,353.97	29.73	$[M+Na-2H]^{-}$	GQ1 (d18:1/12:1)
2,381.93	2,382.01	33.58	$[M+Na-2H]^{-}$	GQ1 (d18:1/14:1)
2,472.12	2,472.21	36.40	$[M-H]^{-}$	GQ1(d18:1/22:1)
2,582.47	2,582.56	34.84	$[M-H]^{-}(-H_2O)$	GQ1(d18:1/31:0)
2,618.41	2,618.51	38.18	$[M+2Na-3H]^{-}$	GQ1(d18:0/29:0)

(continued)

Table 8.5 (continued)

m/z (monoisotopic) experimental	*m/z* (monoisotopic) theoretical	Mass accuracy (ppm)	Molecular ion	Proposed composition
2,642.52	2,642.62	37.83	$[M-H]^-$	*O*-Ac-GQ1(d18:1/31:0)
2,833.44	2,833.55	38.82	$[M-H]^-$	GP1(d18:1/27:1)
872.37	872.38	11.46	$[M-2H]^{2-}$	*O*-Ac-GD2 (d18:1/20:0)
911.30	911.31	10.97	$[M-2H]^{2-}$	GD1, nLD1 or LD1 (d18:1/18:3)
1,063.31	1,063.33	18.81	$[M-2H]^{2-}$	GT1(d18:1/18:0) or GT1(d18:0/18:1)
1,175.94	1,175.96	17.00	$[M-H]^-$	GM3 (d18:1/18:2)
1,197.77	1,197.80	25.04	$[M-H]^-$	GM3(d18:0/19:0)
1,235.78	1,235.81	24.29	$[M-H]^-$	GM3 (d18:1/22:0)
1,306.03	1,306.06	22.97	$[M-H]^-$	GM3(d18:1/27:0)
1,510.80	1,510.84	26.47	$[M-H]^-$	*O*-Ac-GD3 (d18:1/18:1)
1,644.27	1,644.31	24.33	$[M-H]^-$	GM1 (d18:0/25:0)
1,662.18	1,662.22	24.06	$[M-H]^-$	GM1 (d18:1/29:2)
1,684.78	1,684.82	23.73	$[M-H]^-$	GD2(d18:1/19:2) or GD2(d18:0/19:3)
1,847.38	1,847.42	21.65	$[M-H]^-$	GT3(d18:0/24:0)
1,879.05	1,879.09	21.28	$[M-H]^-(-H_2O)$	Fuc-GT3(d18:0/17:0) or *O*-Ac-GT3 (d18:1/22:2)
	1,879.10	26.60	$[M+Na-2H]^-$	
2,024.17	2,024.23	29.64	$[M+2Na-3H]^-$	GD1 (d18:0/28:0)
2,050.55	2,050.60	24.37	$[M-H]^-$	GT2(d18:1/24:0) or GT2(d18:0/24:1)
2,072.15	2,072.21	28.95	$[M-H]^-$	*O*-Ac-GD1(d18:1/32:1)
2,076.13	2,076.19	28.90	$[M-H]^-$	*O*-Ac-GD1(d18:0/32:0)
2,102.20	2,102.26	28.54	$[M-H]^-$	*O*-Ac-GD1(d18:1/34:0)
2,112.02	2,112.09	33.14	$[M-H]^-$	GT1(d18:1/17:2) or GT1(d18:0/17:3)
2,123.90	2,123.97	32.95	$[M-H]^-$	GT1(d18:1/18:2)
2,175.73	2,175.80	32.16	$[M-H]^-$	GT2 (d18:1/33:0)
2,186.52	2,186.60	36.57	$[M-H]^-$	GT1(d18:0/22:0)
2,237.08	2,237.16	35.76	$[M-H]^-$	GT1(d18:1/26:1)

2,246.90	2,246.98	35.60	$[M-H]^-$	Fuc-GT1(d18:0/16:0)
2,287.72	2,287.80	34.96	$[M-H]^-$	Fuc-GT1 (d18:0/20:0)
2,346.00	2,346.09	38.36	$[M-H]^-$ $(-H_2O)$	GQ1(d18:1/13:1)
2,385.13	2,385.22	37.73	$[M+3Na-4H]^-$	GT1(18:1/32:2)
2,472.12	2,472.21	36.40	$[M-H]^-$	GQ1(d18:1/22:1)
2,582.46	2,582.56	38.71	$[M-H]^-$ $(-H_2O)$	GQ1(d18:1/31:0)
2,618.41	2,618.51	38.18	$[M+2Na-3H]^-$	GQ1(d18:0/29:0)
2,642.52	2,642.62	37.83	$[M-H]^-$	*O*-Ac-GQ1(d18:1/31:0)
2,673.42	2,673.52	37.41	$[M-H]^-$	GP2(d18:1/27:0)
2,906.39	2,906.50	37.85	$[M-H]^-$	Fuc-GP1(d18:1/23:1)
917.47	917.48	10.90	$[M-2H]^{2-}$	GD1, nLD1 or LD1 (d18:1/18:0)
931.33	931.34	10.74	$[M-2H]^{2-}$	GD1, nLD1 or LD1 (d18:1/20:0)
947.33	947.34	10.55	$[M-H]^-$	LacCer (d18:0/22:0)
980.19	980.21	20.40	$[M+Na-2H]^-$	GM4 (d18:1/14:2)
1,019.78	1,019.80	19.60	$[M-2H]^{2-}$	GT2 (d18:1/22:2)
1,063.31	1,063.33	18.81	$[M-2H]^{2-}$	GT1(d18:1/18:0) or GT1(d18:0/18:1)
1,074.00	1,074.02	18.62	$[M+Na-3H]^{2-}$	GT1 (d18:1/18:0)
1,098.18	1,098.20	18.21	$[M-2H]^{2-}$	GT1(d18:1/23:0) or GT1(d18:0/23:1)
1,127.42	1,127.45	26.61	$[M+Na-2H]^-$	GM3 (d18:1/13:2)
1,179.71	1,179.74	25.42	$[M-H]^-$	GM3(d18:1/18:0)
1,224.33	1,224.36	24.50	$[M-H]^-$	Fuc-GM3 (d18:1/12:0)
1,231.01	1,231.04	24.37	$[M+2Na-4H]^{2-}$	GQ1 (d18:1/18:0)
1,268.66	1,268.69	23.64	$[M-H]^-$	GM2 (d18:1/10:0)
1,382.78	1,382.82	28.92	$[M-H]^-$	GM2 (d18:1/18:0)
1,408.81	1,408.85	28.38	$[M-H]^-$	GM2 (d18:1/20:1)
1,440.74	1,440.78	27.75	$[M-H]^-$	GD3 (d18:1/16:1)
1,470.99	1,471.03	27.19	$[M-H]^-$	GD3 (d18:1/18:0)
1,518.82	1,518.86	26.33	$[M-H]^-$	GM1(d18:0/16:0)

(continued)

Table 8.5 (continued)

m/z (monoisotopic) experimental	*m/z* (monoisotopic) theoretical	Mass accuracy (ppm)	Molecular ion	Proposed composition
1,544.78	1,544.83	32.36	$[M-H]^-$	GM1, nLM1 or LM1 (d18:1/18:0)
1,572.80	1,572.85	31.78	$[M-H]^-$	GM1, nLM1 or LM1 (d18:1/20:0)
1,600.86	1,600.92	37.47	$[M-H]^-$	GM1,nLM1or LM1 (d18:1/22:0)
1,700.83	1,700.89	35.27	$[M+2Na-3H]^-$	GM1(d18:1/26:0)
1,830.33	1,830.40	38.25	$[M-H]^-$	GT3 (d18:1/23:1)
1,835.87	1,835.94	38.12	$[M-H]^-$	GD1, nLD1or LD1(d18:1/18:0)
1,857.88	1,857.95	37.67	$[M+Na-2H]^-$	GD1, nLD1 or LD1 (d18:1/18:0)
1,885.92	1,885.99	37.11	$[M+Na-2H]^-$	GD1(d18:1/20:0)
1,916.85	1,916.92	36.51	$[M-H]^-$	GD1 (d18:1/24:2)
1,965.95	1,966.02	35.60	$[M-H]^-$	GT2(d18:1/18:1) or GT2(d18:0/18:2)
1,992.46	1,992.53	35.12	$[M-H]^-$	Hex-HexNAc-nLM1(d18:1/24:1)
1,996.95	1,996.96	35.05	$[M-H]^-$	GD2–lactone (d18:1/22:2)
2,052.52	2,052.60	38.96	$[M-H]^-$	GT2 (d18:0/24:0)
2,098.16	2,098.24	38.13	$[M-H]^-$	*O*-Ac-GD1(d18:1/34:2)
2,124.12	2,124.20	37.66	$[M-H]^-$	GT1(d18:1/18:2) or GT1(d18:0/18:3)
2,235.08	2,235.16	35.79	$[M-H]^-$	GT1(d18:1/26:2)
2,385.13	2,385.22	37.73	$[M+3Na-4H]^-$	GT1(18:1/32:2)
2,524.14	2,524.24	39.61	$[M+Na-2H]^-$	GQ1(d18:1/24:0)
2,622.46	2,622.56	38.12	$[M+Na-2H]^-$	GQ1(d18:1/31:0)
2,827.26	2,827.37	38.91	$[M-H]^-(-H_2O)$	GP1(d18:1/28:2)

Reprinted with permission from [90]

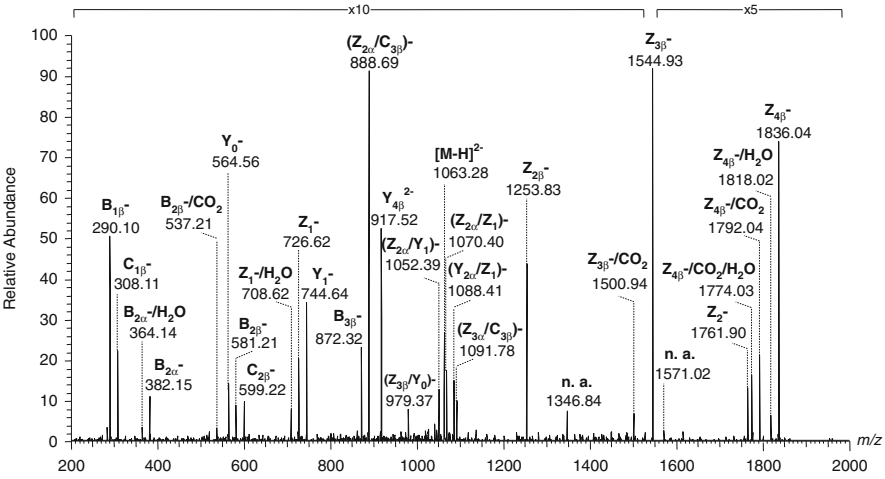


Fig. 8.13 Structural analysis by CID MS^2 of the $[M-2H]^{2-}$ corresponding to GT1(d18:1/18:0) or GT1(d18:0/18:1) detected at *m/z* 1.063.28 by MS screening of astrocytoma gangliosides. Collision energy within 35–50 eV range (*Elab*). The assignment of the fragment ions is according to the nomenclature [14]. Reprinted with permission from Zamfir et al. [90]

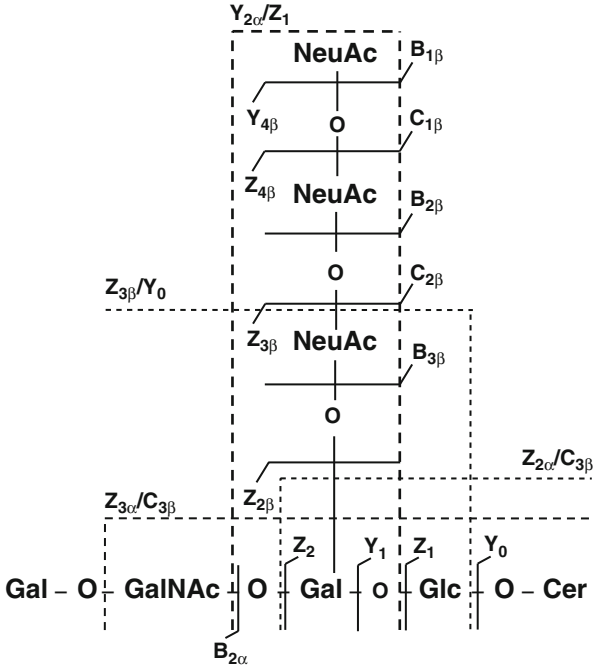


Fig. 8.14 Fragmentation scheme of GT1c (d18:1/18:0) or GT1c (d18:0/18:1) by CID MS^2. The assignment of the fragment ions is according to the nomenclature [14]. Reprinted with permission from Zamfir et al. [90]

(d18:1/18:0) ceramide composition was never identified in another brain tumor, which shows that under the high-resolution conditions and multiple stage fragmentation on Orbitrap MS, a novel species, biomarker of human astrocytoma was discovered and structurally characterized.

Gliosarcoma

Glioblastoma multiforme is a rare primary neoplasm of the CNS, ranked by the World Health Organization as a grade IV tumor [42]. Gliosarcoma is a variant of glioblastoma multiforme [25] defined by a biphasic tissue pattern consists of alternating areas displaying glial and mesenchymal (sarcomatous) differentiation. Gliosarcoma accounts for about 2 % of all the glioblastomas, usually affecting the adult population in the fourth to the sixth decade of life with a prevalence in male population (male:female, 1.8:1).

Gliosarcomas are usually located in the cerebral cortex involving the temporal, frontal, parietal, and occipital lobe in decreasing frequency. The typical clinical history of the patient is short and the presenting symptoms depend upon the location of the tumor [25]. Even after radical resection by surgery followed by chemo- and radiotherapy, the median survival time is usually 11.5 months with less than 10 % survival after 2 years following diagnosis [44]. The failure of the present therapeutic scheme in glioblastoma is the tumor aggressiveness which leads to rapid infiltration of the tumoral cells into the adjacent healthy brain tissue, which makes them rather inaccessible to treatment methods. Therefore, the present strategies investigated for treatment is to target the invading tumor cells by using specific binding ligands [68]. The critical point is, however, to identify the tumor-specific target molecules and characterize their structures in detail.

In the case of human gliomas, mono- and di-sialylated ganglioside species have been suggested as associated species and/or antigens [19]. Although by various methods a few gangliosides were found potential candidates as glioma-antigens, a more comprehensive mapping of this tumor biomarkers could be achieved by high-resolution mass spectrometry using chip-based nanoESI QTOF and ESI FTICR MS for screening and CID for structural elucidation [83]. By these accurate methods it was found that ganglioside expression in gliosarcoma is highly altered as compared to the normal brain. For instance, GD3 species having d18:1/18:0, d18:1/24:1, and d18:1/24:0 Cer compositions were found dominant in the ganglioside mixture extracted from human gliosarcoma (Fig. 8.15). Unexpectedly a high abundance of *O*-acetylated GD3 derivatives, particularly with d18:1/18:0 and d18:1/20:0 lipid moiety structure, were also observed. GM2, GM1, and/or their isomers nLM1 and LM1, as well as GD1 species characterized by heterogeneity in composition of their ceramide moieties, were found abundantly in the mixture, while the species with a higher sialylation degree were found poorly or not at all expressed. The reduction in the total ganglioside content and the altered pattern in gliosarcoma vs. control tissue is considered the result of both a lower overall biosynthetic rate, due to change in expression of certain glycosyltransferases, and higher turnover rate [83].

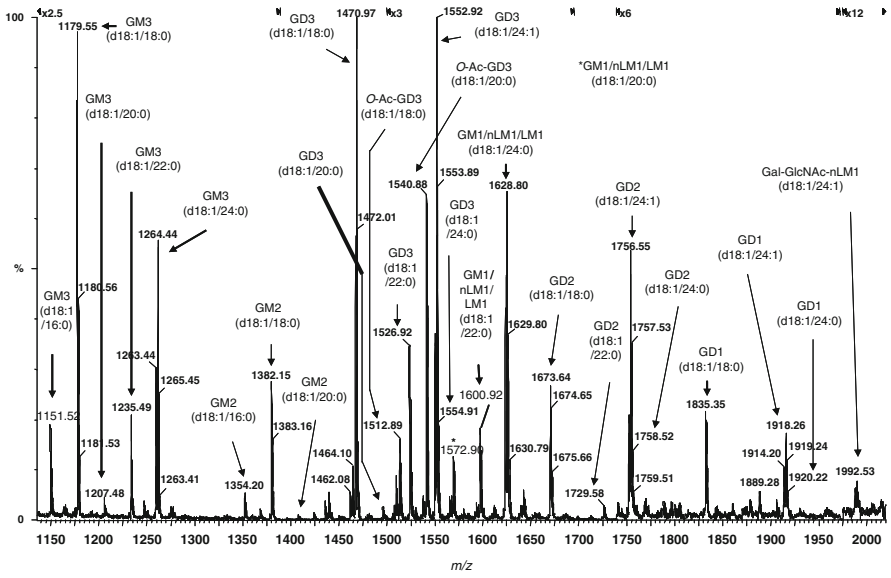


Fig. 8.15 NanoESI chip QTOF MS of the native gliosarcoma ganglioside mixture. ESI voltage 1.60 kV. Sampling cone 80 V. Acquisition 2 min. Average sample consumption 0.5 pmol. Reprinted with permission from Vukelić et al. [83]

A higher expression of sialyltransferase II (GD3 synthase) and a lower expression of galactosyltransferase II were found to be the most probable causes of the very high GD3 and very low GM1a, GD1a, and GD1b abundances. While the latter species were previously known as glioma biomarkers, GD3 and the corresponding acetylated forms were discovered only by mass spectrometry as tumor-associated. Tandem MS by CID using as the precursor *O*-Ac-GD3 (Fig. 8.16) provided sequence data consistent with the presence of gliosarcoma-associated isomer bearing *O*-acetylation at the inner Neu5Ac residue, a form previously not structurally confirmed. The information derived from the MS data according to which gliosarcoma, as the highest malignancy grade brain tumor, contains a higher amount of potentially proapoptotic GD3 than of the *O*-acetyl GD3 species suports the previous assumption [33] that the role of *O*-Ac-GD3 as tumor-specific component is the protection of tumor cells from apoptosis.

8.4.3 Secondary Brain Tumors

When highly expressed, some of the glycosyl epitopes promote invasion and metastasis and thus can lead to shorter survival rate of patients, while some others suppress tumor progression leading to higher postoperative survival period. Targeting carbohydrate antigens such as gangliosides expressed on metastatic tumor cells

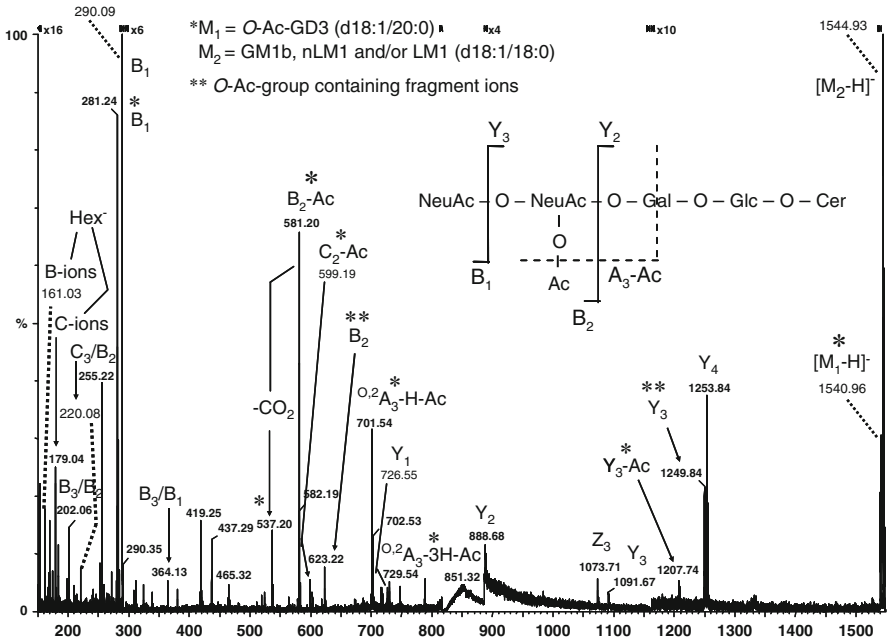


Fig. 8.16 NanoESI QTOF CID MS/MS of the $[M-H]^-$ at *m/z* 1,540.96 corresponding to the *O*-Ac-GD3 (d18:1/20:0). ESI voltage 1,000–1,250 V. Collision energy 25–40 eV. Collision gas pressure 5–10 psi. Acquisition time 11 min. Average sample consumption 3.5 pmol. Inset: the fragmentation scheme of *O*-Ac-GD3. Reprinted with permission from Vukelić et al. [83]

represents a priority for the immunotherapy of cancer since aberrant glycosylation exhibited by tumor cells is considered a factor of their uncontrolled growth, invasiveness, and increased metastatic potential.

8.4.3.1 Brain Metastasis of Lung Adenocarcinoma

Lung adenocarcinoma is a histological form of lung cancer that contains certain distinct malignant tissue architectural, cytological, or molecular features. Non-small cell lung cancer, the most frequent cause of cancer deaths in many countries, has a high risk of brain metastases that reportedly reaches 44 % in brain autopsy. As compared to other primary cancers, where brain spread is usually a later complication, lung cancer develops intracranial metastases relatively early and is often accompanied by neurologic symptoms on initial diagnosis [22].

Untreated brain metastases have a median survival of about 4 weeks with almost all patients dying from neurological rather than systemic causes [2]. Whole-brain radiation therapy and chemotherapy are currently the standard, unfortunately only palliative, treatments for patients with brain metastases. Even under treatment, the prognosis for patients with brain metastases is generally poor, with median survival time between 3 and 6 months [20].

Table 8.6 Ganglioside and asialo-ganglioside species from brain metastasis of lung adenocarcinoma (male, 73 years) detected by negative ion mode nanoESI chip QTOF MS analysis of complex native ganglioside mixture

m/z (monoisotopic) theoretical	*m/z* (monoisotopic) experimental	Mass accuracy (ppm)	Molecular ion	Proposed structure
875.19	874.91	33	$[M-H]^-$	LacCer(d18:1/17:0)
933.31	932.99	35	$[M-H]^-$	LacCer(d18:0/21:0)
947.34	947.19	16	$[M-H]^-$	LacCer(d18:0/22:0)
949.22	949.24	21	$[M+2Na-3H]^-$	LacCer(d18:0/19:0)
964.24	963.90	35	$[M-H]^-$	GM4(d18:0/14:0)
982.19	981.94	25	$[M+Na-2H]^-$	GM4(d18:1/14:1)
984.21	983.87	34	$[M+Na-2H]^-$	GM4(d18:1/14:0) or GM4(d18:0/14:1)
1,122.48	1,122.23	22	$[M-H]^-$	GA2(d18:0/20:0)
1,138.44	1,138.15	25	$[M-H]^-$	Fuc-GM4(d18:0/16:0) GA2(t18:0/20:0)
1,138.48		29		
1,150.49	1,150.17	28	$[M-H]^-$	GA2(d18:1/21:0)
1,150.40		20	$[M-H]^-$	GM3(d18:1/16:1)
1,168.42	1,168.01	35	$[M-H]^-$	GM3(t18:0/16:0)
1,178.46	1,178.14	27	$[M-H]^-$	GM3(d18:1/18:1)
1,179.74	1,180.10	30	$[M-H]^-$	GM3(d18:1/18:0)
1,182.49	1,182.21	24	$[M-H]^-$	GM3(d18:0/18:0)
1,184.37	1,184.08	24	$[M-H]^-$	*O*-Ac-GA1(d18:1/10:0)
1,194.50	1,194.15	29	$[M-H]^-$	GM3(d18:1/19:0) or GM3(d18:0/19:1)
1,206.51	1,206.33	15	$[M-H]^-$	GM3(d18:1/20:1)
1,206.64		26	$[M-H]^-$	GA2(d18:0/26:0)
1,222.51	1,222.19	26	$[M-H]^-$	*O*-Ac-GM3(d18:1/18:0)
1,222.55		29	$[M-H]^-$	GM3(d18:0/21:1) or GM3(d18:1/21:0)
1,234.56	1,234.22	27	$[M-H]^-$	GM3(d18:1/22:1)
1,234.52		24		*O*-Ac-GM3(d18:1/19:1)

(continued)

Table 8.6 (continued)

m/z (monoisotopic) theoretical	*m/z* (monoisotopic) experimental	Mass accuracy (ppm)	Molecular ion	Proposed structure
1,248.55	1,248.18	33	$[M-H]^-$	*O*-Ac-GM3(d18:1/20:1)
1,248.59		30		GM3(d18:1/23:1)
1,248.59	1,249.02	34	$[M-H]^-$	GM3(d18:1/23:0)
1,260.60	1,260.33	21	$[M-H]^-$	GM3(d18:1/24:2)
1,262.62	1,262.35	21	$[M-H]^-$	GM3(d18:1/24:1)
1,264.63	1,264.19	35	$[M-H]^-$	GM3(d18:1/24:0)
1,276.61	1,277.01	31	$[M-H]^-$	*O*-Ac-GM3(d18:1/22:0)
1,276.64		29	$[M-H]^-(-H_2O)$	GM3(d20:1/23:1)
1,276.69		25		GM3(d18:0/26:0)
1,278.66	1,278.21	35	$[M-H]^-$	GM3(d20:1/23:0)
1,278.61		31	$[M-H]^-$	*O*-Ac-GM3(d18:1/22:0) or *O-Ac-GM3*(d18:0/22:1)
1,288.67	1,289.04	29	$[M-H]^-$	GM3(d18:1/26:2) or GM3(d18:2/26:1)
1,288.67		29	$[M-H]^-$	GM3(d20:1/24:2)
1,292.68	1,292.23	35	$[M-H]^-$	GM3(d18:1/26:0) or GM3(d18:0/26:1)
1,296.54	1,296.24	23	$[M-H]^-$	Fuc-GM3(d18:1/16:1)
1,296.59		27	$[M-H]^-$	*O*-Ac-GA1(d18:1/18:0)
1,296.61		28	$[M-H]^-$	GA1(d18:0/21:0) or GA1(d18:0/21:0)
1,405.65	1,405.21	31	$[M+Na-2H]^-$	GM2(d18:1/18:0)
1,420.68	1,420.80	8	$[M-H]^-$	*O*-Ac-GM2(d18:2/18:2)
1,435.59	1,435.21	26	$[M+Na-2H]^-$	GD3(d18:1/14:1) or GD3(d18:0/14:2) or GD3(d18:2/14:0)
1,441.66	1,441.19	33	$[M-H]^-$	GD3(d18:1/16:1) or GD3(d18:0/16:2) or GD3(d18:2/16:0)
1,471.73	1,471.28	31	$[M-H]^-$	GD3(d18:1/18:0)
1,493.71	1,493.23	32	$[M+Na-2H]^-$	GD3(d18:1/18:0)
1,515.69	1,515.29	26	$[M+2Na-3H]^-$	GD3(d18:1/18:0) or GD3(d18:0/18:1)
1,515.74		30	$[M-H]^-$	GM1(d18:1/16:1) or GM1(d18:0/16:2) or GM1(d18:2/16:0)
1,515.78		32	$[M-H]^-$	*O*-Ac-GD3(d18:0/18:0)

1,515.71	1,516.01	20	$[M+Na-2H]^-$	GD3(d18:1/20:2) or GD3(d18:0/20:3) or GD3(d18:2/20:1)
1,515.75		17	$[M-H]^-$	GM1(d18:2/16:0) or GM1(d18:1/16:1) GM1(d18:0/16:2
1,527.83	1,528.16	22	$[M-H]^-$	GD3(d18:0/22:0)
1,541.79	1,542.19	26	$[M-H]^-$	GM1(d18:1/18:2) or GM1(d18:2/18:1) or GM1(d18:0/18:3)
1,569.78	1,570.29	32	$[M+2Na-3H]^-$	GD3(d18:1/22:0) or GD3(d18:0/22:1)
1,569.83		29	$[M-H]^-$	GM1(d18:1/20:1) or GM1(d18:0/20:2) or GM1(d18:2/20:0)
1,569.77		33	$[M+Na-2H]^-$	GD3(d18:0/24:2) or GD3(d18:1/24:1) or GD3(d18:2/24:0)
1,597.88	1,598.09	13	$[M-H]^-$	GM1(d18:0/22:2) or GM1(d18:1/22:1) or GM1(d18:2/22:0)
1,611.77	1,612.17	25	$[M+2Na-3H]^-$	GM1(d18:1/20:2)
1,625.89	1,625.40	30	$[M+2Na-3H]^-$	GD3(d18:1/26:1) or GD3(d18:0/26:2) or GD3(d18:2/26:0)
1,624.92		30	$[M-H]^-$	GM1(d18:1/24:2)
1,627.90	1,627.41	30	$[M+2Na-3H]^-$	GD3(d18:0/26:1) or GD3(d18:1/26:0)
1,626.93		29	$[M-H]^-$	GM1(d18:0/24:2) or GM1(d18:1/24:1) or GM1(d18:2/24:0)
1,629.92	1,629.42	31	$[M-H]^-$	GM1(d18:0/24:1) or GM1(d18:1/24:0)
1,628.94		29	$[M-H]^-$	di-*O*-Ac-GM1(d18:1/18:0)
1,659.79	1,660.18	23	$[M+3Na-4H]^-$	GM1(d18:1/22:3) or GM1(d18:0/22:4) or GM1(d18:2/22:2)
1,674.87	1,675.23	21	$[M+Na-2H]^-$ $(-H_2O)$	GD2 (d18:1/18:2)
1,748.97	1,749.39	24	$[M+Na-2H]^-$	GD2 (d18:1/22:1)
1,766.97	1,767.28	18	$[M-H]^-$ $(-H_2O)$	GT3 (d18:1/20:1)
1,785.07	1,785.37	17	$[M-H]^-$	*O*-Ac-GD2(d18:1/23:0) or *O-Ac-GD2*(d18:0/23:1)
1,833.81	1,833.28	29	$[M-H]^-$	GT3(d18:0/23:0)
1,833.07		11	$[M-H]^-$	*O*-Ac-GT3 (d18:0/20:0)
1,861.12	1,861.24	6	$[M-H]^-$	*O*-Ac-GT3-lactone(d18:0/22:0)
1,861.12		6	$[M-H]$ $(-H_2O)$	*O*-Ac-GT3(d18:0/22:0)
1,879.09	1,879.39	16	$[M+Na-2H]^-$	*O*-Ac-GT3 (d18:2/22:1)
1,879.10		15	$[M-H]^-(-H_2O)$	Fuc-GT3(d18:0/17:0)
1,879.99		32	$[M-H]^-$	GT2(d18:1/12:1) or GT2(d18:2/12:0)

(continued)

Table 8.6 (continued)

m/z (monoisotopic) theoretical	*m/z* (monoisotopic) experimental	Mass accuracy (ppm)	Molecular ion	Proposed structure
1,909.16	1,909.03	7	$[M-H]^-$	GD1 (d18:1/22:0)
1,960.21	1,959.84	19	$[M-H]^-$ ($-2H_2O$)	GT2(d18:0/20:0)
1,960.12		14	$[M-H]^-$	GT2(d18:1/18:3) or GT2(d18:2/18:2)
1,990.17	1,989.78	20	$[M+Na-2H]^-$	GT2(d18:0/18:0)
1,990.19		21	$[M-H]^-$	GT2(d18:0/20:3) or GT2(d18:1/20:2) or GT2(d18:2/20:1)
1,990.19	1,990.83	32	$[M-H]^-$	GT2(d18:1/20:1) or GT2(d18:0/20:2) or GT2(d18:2/20:0)
2,005.20	2,005.63	21	$[M-H]^-$	Fuc-GD1(d18:1/20:2)
2,006.19		28		*O*-Ac-GT2(d18:1/18:1)
2,048.23	2,048.80	28	$[M-H]^-$	di-*O*-Ac-GT2(d18:0/18:0)
2,048.10		34	$[M-H]^-(-H_2O)$	GT1(d18:2/14:2) or GT1(d18:3/14:1)

Reprinted with permission from [89]

Non-small cell lung cancer has been shown to exhibit elevated expression of GM3 and GM3 synthase (sialyltransferase-I or SAT-I) mRNA with a positive correlation between expression levels of SAT-I mRNA and GM3 in tumor tissues [51]. Overexpression of GM3 synthase was used to determine the effects of endogenous gangliosides on the metastatic process of 3LL Lewis lung carcinoma cells. It is also known that metastatic potential of lung cancer cells is regulated via GM1 ganglioside by modulating the matrix metalloproteinase-9. Low GM1 expressing cell lines showed increased proliferation, invasion, and metastatic potential [91]. Another ganglioside used in development of novel therapies for small cell lung cancer is fucosyl-GM1, which is specifically expressed in lung cancer cells. In the last year, a bidomainal fucosyl-GM1 ganglioside-based vaccine for the treatment of small cell lung cancer was developed [75]. It was also shown that an anti-ganglioside-based cancer vaccine containing 1E10 anti-idiotypic monoclonal antibody induces apoptosis and antiangiogenic effects in a metastatic lung carcinoma [13].

In 2011 high-performance MS was employed for the investigation of ganglioside expression and structure in secondary brain tumors, i.e., brain metastasis of lung adenocarcinoma [89]. Comparative nanoESI chip QTOF MS screening of gangliosides from metastatic (Table 8.6) vs. healthy tissue (Table 8.7) showed considerable discrepancy in expression, structure, and relative abundances of individual species. In contrast to healthy cerebellar tissue, the ganglioside mixture extracted from brain metastasis of lung adenocarcinoma was found to contain mostly species of short oligosaccharide chains and reduced overall sialic acid content. More than a half, from the total of 63 different ions detected and corresponding to 141 possible structures in brain metastatic tissue, represents monosialylated species of GM1, GM2, GM3, and GM4-type. Besides the large number of monosialylated components, six asialo species of GA1 and GA2-type bearing ceramides of variable constitution were discovered. The observed differences in ceramide structures and alteration of sialylation patterns were attributed to tumor-related changes in human carcinomas.

GD1, GD2, and GD3 as well as GT1, GT2, and GT3 with short carbohydrate chains, expressing different ceramide portions, were also identified in the mixture. Ganglioside components modified by Fuc or *O*-Ac could also be detected, but in a different pattern than in healthy brain; most *O*-acetylated gangliosides were found as monosialo species of GM3, GM2, and GM1 type, while fucosylated components were represented by monosialo species of GM3 and GM4 structure, di- and trisialylated GD1 and GT3 exhibiting high heterogeneity in their ceramide motifs [89].

By tandem MS using CID, brain metastasis-associated GD3 (d18:1/18:0) species was structurally elucidated (Fig. 8.17a, b). This structure was reported to enhance tumor cell proliferation, invasion, and metastasis in a variety of brain tumor cells, especially in glioma and neuroblastoma.

From the methodological point of view it is to be mentioned that nanoESI chip QTOF MS and CID MS/MS was able to provide compositional and structural characterization of native ganglioside mixtures from secondary brain tumors with remarkable analysis speed and sensitivity. Under the applied conditions a sample concentration of only 2.5 pmol/μL, which corresponds to 250 fmol biological extract consumption, was necessary for an experiment.

Table 8.7 Ganglioside and asialo-ganglioside species from the control healthy cerebellum (male, 79 years) detected by negative ion mode nanoESI chip QTOF MS analysis of complex native ganglioside mixture (from [89])

m/z (monoisotopic) theoretical	*m/z* (monoisotopic) experimental	Mass accuracy (ppm)	Molecular ion	Proposed structure
708.35	708.38	4	$[M-3H]^{3-}$	GT1(d18:1/18:0)
714.42	714.40	3	$[M-3H]^{3-}$	GT1(t18:0/18:0)
			$[M+Na-4H]^{3-}$	GT1(d18:1/18:2)
717.58	717.54	5	$[M-3H]^{3-}$	GT1(d18:1/20:0) or GT1(d18:0/20:1)
734.91	735.12	29	$[M-2H]^{2-}$	GD3(d18:1/18:0) or GD3(d18:0/18:1)
756.38	756.30	11	$[M-2H]^{2-}$	*O*-Ac-GD3(d18:1/18:0)
771.95	771.93	3	$[M-2H]^{2-}$	GM1(d18:0/18:1) or GM1(d18:1/18:0)
822.05	822.06	1	$[M+Na-4H]^{3-}$	GQ1(d18:1/20:0) or GQ1(d18:0/20:1)
835.95	835.69	31	$[M-2H]^{2-}$	GD2(d18:1/18:1) or GD2(d18:0/18:2) or GD2(d18:2/18:0)
844.96	844.69	32	$[M-2H]^{2-}$	*O*-Ac-GD2(d18:0/16:0)
850.47	850.22	29	$[M-2H]^{2-}$	GD2 (d18:1/20:0)
863.51	863.21	35	$[M-2H]^{2-}$	Fuc-GM1(d18:1/22:2) or Fuc-GM1(d18:0/22:3) or Fuc-GM1(d18:2/22:1)
863.00		24	$[M-2H]^{2-}$	GD2(d18:0/22:3) or GD2(d18:1/22:2) or GD2(d18:2/22:1)
862.99		26	$[M+Na-3H]^{2-}$	GD2(d18:0/20:0)
878.03	877.88	17	$[M-2H]^{2-}$	GD2(d18:1/24:1)
886.03	885.78	28	$[M-2H]^{2-}$	*O*-Ac-GD2(d18:1/22:0) or *O*-Ac-GD2(d18:0/22:1)
890.97	890.76	24	$[M+Na-3H]^{2-}$	GT3(d18:1/18:1)
905.01	905.11	11	$[M+Na-3H]^{2-}$	GT3(d18:1/20:0)
917.48	917.44	4	$[M-2H]^{2-}$	GD1(d18:1/18:0) or GD1(d18:0/18:1)
924.49	924.76	29	$[M-2H]^{2-}$	GD1(d18:1/19:0) or *O*-Ac-GT3(t18:1/20:0) or
924.53		25	$[M+2Na-4H]^{2-}$	*O*-Ac-GT3(d18:0/20:1)
931.49	931.45	4	$[M-2H]^{2-}$	GD1(d18:1/20:0) or GD1(d18:0/20:1)
940.50	940.46	4	$[M-2H]^{2-}$	GD1(t18:0/20:0)
940.19		29	$[M+2Na-4H]^{2-}$	GD1 (d18:1/18:0) or GD1(d18:0/18:1)
945.51	945.50	1	$[M-2H]^{2-}$	GD1(d18:1/22:0)

952.52	952.50	2	$[M-2H]^{2-}$	*O*-Ac-GD1 (d18:1/20:0) or *O*-Ac-GD1 (d18:0/20:1)
968.03	968.34	32	$[M-3H]^{3-}$	GH2(d18:1/24:0) or GH2(d18:0/24:1)
991.56	991.27	29	$[M+Na-3H]^{2-}$	GT2(d18:1/18:2) or GT2(d18:0/18:3) or GT2(d18:2/18:3)
1,005.58	1,005.28	30	$[M+2Na-4H]^{2-}$	GT2 (d18:0/18:0)
1,019.02	1,019.36	33	$[M-2H]^{2-}$	GalNAc-GD1(d18:0/18:0)
1,024.62	1,024.68	6	$[M-2H]^{2-}$	di-*O*-Ac-GT2 (d18:1/18:0)
1,024.57		11	$[M+2Na-4H]^{2-}$	*O*-Ac-GT2 (d18:1/18:1)
1,034.24	1,033.94	29	$[M+2Na-3H]^{-}$	GM4 (d18:1/16:0) or GM4 (d18:0/16:1)
1,042.60	1,042.51	9	$[M-2H]^{2}$	GT1 (t18:1/14:1) or GalNAc-GD1(t18:0/20:0)
1,042.66		14		
1,046.59	1,046.46	12	$[M+Na-3H]^{2-}$	GT1(d18:1/14:0) or GT1(d18:0/14:1)
1,049.62	1,049.51	10	$[M-2H]^{2-}$	GT1(d18:1/16:0) or GT1(d18:0/16:1)
1,059.61	1,059.28	31	$[M+Na-3H]^{2-}$	GT1(d18:1/16:1)
1,063.03	1,063.35	30	$[M-2H]^{2-}$	GT1(d18:1/18:0) or GT1(d18:0/18:1)
1,074.02	1,074.05	3	$[M+Na-3H]^{2-}$	GT1(d18:1/18:0)
1,077.04	1,077.37	31	$[M-2H]^{2-}$	GT1\(d18:1/20:0)
1,097.18	1,096.81	34	$M-2H]^{2-}$	*O*-Ac-GT1(d18:1/20:1) or *O*-Ac-GT1(d18:0/20:2) or *O*-Ac-GT1(d18:2/20:0)
1,097.04		21	$[M+Na-3H]^{2-}$	GT1(t18:0/20:0)
1,110.70	1,110.36	31	$[M-2H]^{2-}$	*O*-Ac-GT1(d18:1/22:2)
1,118.49	1,118.56	6	$[M-H]^{-}$	GM4(t18:1/24:0)
1,180.47	1,180.09	32	$[M-H]^{-}$	GM3(d18:1/18:0) or GM3(d18:0/18:1)
1,228.51	1,228.61	8	$[M-H]^{-}$	GA1(d18:0/16:0)
1,228.49		10	$[M+Na-2H]^{-}$	GM3(d18:1/20:1) or GM3(d18:0/20:2)
1,232.55	1,232.12	35	$[M-H]^{-}$	GM3(d18:1/22:2) or GM3(d18:0/22:3) or GM3(d18:2/22:1)
1,232.52		32	$[M+Na-2H]^{-}$	GM3(d18:0/20:0)
1,241.29	1,240.86	35	$[M+Na-3H]^{2-}$	*O*-Ac-GQ1(d18:1/18:0)

(continued)

Table 8.7 (continued)

m/z (monoisotopic) theoretical	*m/z* (monoisotopic) experimental	Mass accuracy (ppm)	Molecular ion	Proposed structure
1,252.58	1,252.19	31	$[M-H]^-$	*O*-Ac-GM3(d18:0/20:0)
1,252.60		33		GM3 (d18:0/23:0)
1,264.54	1,264.12	33	$[M-H]^-$	di-*O*-Ac-GM3(d18:1/18:0)
1,284.60	1,284.36	19	$[M+Na-2H]^-$	GM3(d18:1/24:1)
1,382.82	1,382.87	4	$[M-H]^-$	GM2(d18:1/18:0) or GM2(d18:0/18:1)
1,409.70	1,410.19	35	$[M-H]^-$	GM2(d18:1/20:0) or GM2(d18:0/20:1)
1,467.69	1,467.86	12	$[M-H]^-$	GD3(d18:1/18:2)
1,467.67		13	$[M+Na-2H]^-$	GD3(d18:0/16:0)
1,492.81	1,492.89	5	$[M+Na-2H]^-$	GD3(d18:1/18:0)
1,509.73	1,509.71	1	$[M-H]^-$	*O*-Ac-GD3(d18:1/18:1)
1,509.79		5		Fuc-GM2-lactone (d18:1/18:1)
1,513.76	1,513.29	31	$[M-H]^-$	*O*-Ac-GD3(d18:1/18:0)
1,516.84	1,517.30	30	$[M-H]^-$	GM1(d18:1/16:0) GD3(d18:2/20:2)
1,517.71		27	$[M+Na-2H]^-$	Fuc-GM2(d18:2/16:2)
1,517.70		26	$[M+Na-2H]^-$	
1,537.72	1,537.26	30	$[M+Na-2H]^-$	GM1(d18:1/16:1)
1,541.73	1,541.21	33	$[M+2Na-3H]^-$	GD3(d18:1/20:0)
1,544.87	1,544.92	3	$[M-H]^-$	GM1(d18:1/18:0) or GM1(d18:0/18:1)
1,561.80	1,561.36	28	$[M-H]^-$	*O*-Ac-GM1(d18:0/16:0)
1,566.85	1,566.68	11	$[M+Na-2H]^-$	GM1(d18:1/18:0) or GM1(d18:0/18:1)
1,572.90	1,572.92	1	$[M-H]^-$	GM1(d18:1/20:0) or GM1(d18:0/20:1)
1,589.83	1,589.28	35	$[M-H]^-$	Fuc-GD3(d18:1/16:0)
1,593.82	1,594.16	21	$[M+Na-2H]^-$	GM1(d18:1/20:0)
1,599.89	1,600.23	21	$[M-H]^-$	GM1(d18:0/22:0)
1,629.88	1,629.32	34	$[M-H]^-$	di-*O*-Ac-GM1(d18:1/18:0)
1,648.88	1,648.32	34	$[M-H]^-$	GD2(d18:0/16:0)
1,656.90	1,656.88	1	$[M-H]^-$	GD2-lactone (d18:1/18:0)

1,662.82	1,663.16	20	$[M+Na-2H]^-$	GD2(d18:2/16:2)
1,674.92	1,674.33	35	$[M-H]^-$	GD2(d18:1/18:0) or GD2(d18:0/18:1)
1,690.93	1,690.95	1	$[M-H]^-$	Fuc-GM1(d18:1/18:0)
1,700.96	1,701.42	27	$[M-H]^-$	GD2(d18:1/20:0) or GD2(d18:0/20:1)
1,708.88	1,709.36	27	$[M+Na-2H]^-$	*O*-Ac-GD2 (d18:1/16:1)
1,716.94	1,716.91	2	$[M-H]^-$	Fuc-GM1 (d18:1/20:1)
1,717.98	1,718.48	29	$[M-H]^-$	Fuc-GM1 (d18:1/20:0)
1,729.01	1,729.62	35	$[M-H]^-$	GD2(d18:1/22:0) or GD2(d18:0/22:1)
1,741.87	1,741.27	34	$[M+Na-2H]^-$	Fuc-GM1 (d18:0/20:0)
1,746.92	1,746.38	31	$[M+2Na-3H]^-$	GD2(d18:0/20:0)
1,775.04	1,775.57	30	$[M+Na-2H]^-$	GD2(d18:2/24:3)
1,787.98	1,787.76	12	$[M+2Na-3H]^-$	Fuc-GM1(d18:1/22:1)
1,796.07	1,796.50	24	$[M+Na-2H]^-$	Fuc-GM1(d18:1/24:0)
1,803.02	1,803.64	34	$[M-H]^-$	*O*-Ac-GT3(d18:1/18:1)
1,835.96	1,835.91	3	$[M-H]^-$	GD1(d18:1/18:0) or GD1(d18:0/18:1)
1,857.02	1,857.59	31	$[M+Na-2H]^-$	GD1(d18:0/18:0)
1,863.10	1,863.61	27	$[M-H]^-$	GD1(d18:1/20:0)
1,873.07	1,872.68	21	$[M+Na-2H]^-$	GD1(d18:1/19:0)
1,879.93	1,879.91	1	$[M+2Na-3H]^-$	GD1(d18:1/18:0)
1,887.79	1,886.98	5	$[M-H]^-$	Fuc-GT3-lactone(d18:1/18:2) Fuc-GT3-lactone(d18:0/18:3) Fuc-GT3-lactone(d18:2/18:1)
1,895.96	1,896.06	5	$[M+2Na-3H]^-$	GD1(d18:0/19:0)
1,901.05	1,901.71	35	$[M+2Na-3H]^-$	*O*-Ac-GT3(d18:1/22:1)
1,910.07	1,909.43	34	$[M-H]^-$	GT2(d18:0/14:0) or GT2(d18:0/14:1)
1,915.17	1,915.78	31	$[M-H]^-$	GD1(d18:1/24:2)
1,925.14	1,925.77	33	$[M+Na-2H]^-$	GD1(d18:1/23:1)
1,937.15	1,937.80	34	$[M+Na-2H]^-$	GD1(d18:1/24:2) or GD1(d18:0/24:3) or GD1(d18:2/24:1)
1,964.16	1,964.84	35	$[M-H]^-$	GT2(d18:0/18:0)

(continued)

Table 8.7 (continued)

m/z (monoisotopic) theoretical	*m/z* (monoisotopic) experimental	Mass accuracy (ppm)	Molecular ion	Proposed structure
1,983.20	1,982.55	33	$[M-H]^-$	Fuc-GD1(d18:1/18:0)
2,005.20	2,004.61	30	$[M-H]^-$	Fuc-GD1(d18:1/20:2)
2,010.16	2,010.78	31	$[M+Na-2H]^-$	GT2(d18:1/20:2) or GT2(d18:0/20:3) or GT2(d18:2/20:1)
2,032.23	2,032.63	20	$[M-H]^-$	*O*-Ac-GT2(d18:1/20:1) or *O*-Ac-GT2(d18:0/20:2) or *O*-Ac-GT2(d18:2/20:0)
2,032.14		24	$[M+2Na-3H]^-$	GT2(d18:1/20:2) or GT2(d18:0/20:3) or GT2(d18:2/20:1)
2,050.24	2,049.75	24	$[M-H]^-$	di-*O*-Ac-GT2 (d18:1/18:0)
2,059.27	2,059.78	29	$[M+Na-2H]^-$	Fuc-GD1(d18:1/22:0)
2,076.24	2,076.85		$[M-H]^-$	GQ3(d18:1/20:2) or GQ3(d18:0/20:3) or GQ3(d18:2/20:1)
2,106.27	2,105.78	23	$[M-H]^-$ $(-H_2O)$	GT1(d18:1/18:2)
2,166.32	2,165.56	35	$[M-H]^-$	*O*-Ac-GT1(d18:1/18:2) or *O*-Ac-GT1(d18:0/18:3) or *O*-Ac-GT1(d18:2/18:1)
2,165.03		24	$[M+Na-2H]^-$	GT1(t18:1/18:0)
2,166.29		34		*O*-Ac-GT1(d18:1/16:0)
2,172.19	2,172.04	7	$[M+2Na-3H]^-$	*O*-Ac-GT1(t18:1/14:1)
2,188.38	2,187.62	35	$[M+Na-2H]^-(H_2O)$	GT1(d18:1/22:0)
2,188.39		35	$[M+Na-2H]^-$	GT1(d18:1/21:1)
2,198.31	2,198.08	10	$[M+2Na-3H]^-$	GT1(d18:1/20:0)
2,214.25	2,215.78	24	$[M+Na-2H]^-$	*O*-Ac-GT1(d18:1/20:2)

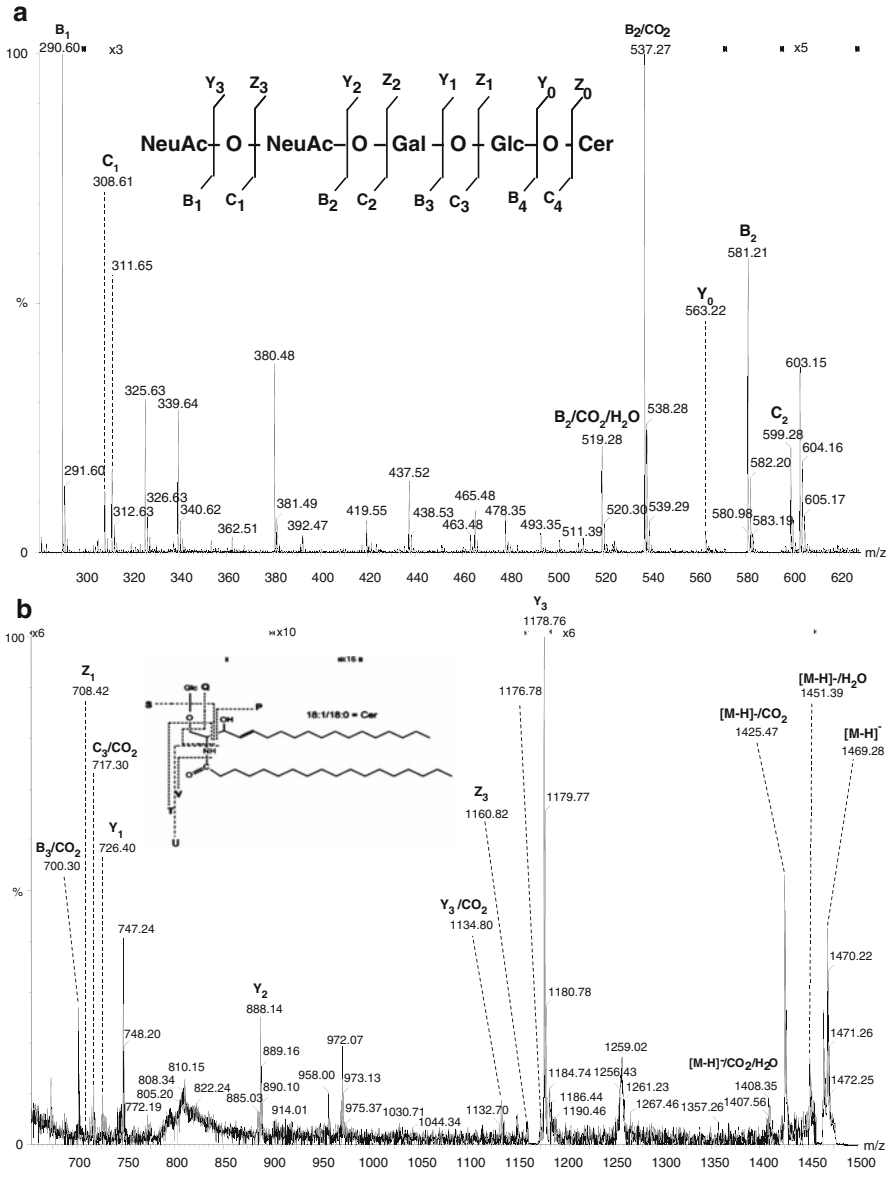


Fig. 8.17 NanoESI chip QTOF CID MS/MS of the singly charged ion at *m/z* 1,471.29 corresponding to GD3 (d18:1/18:0) from brain metastasis of lung adenocarcinoma. *m/z* range (**a**) 200–630. (**b**) 650–1,500. Acquisition time 1 min. *Insets*: fragmentation schemes of the oligosaccharide core and ceramide moiety. Reprinted with permission from Zamfir et al. [89]

The bioanalytical platforms based on mass spectrometry demonstrated here for the discovery of glycolipid/ganglioside molecular markers of brain in health and disease have real perspectives of development into routine, ultrafast, and sensitive methods applicable to other molecular markers of the pathologies at the neurological and other levels.

Acknowledgments This work was supported by the Romanian National Authority for Scientific Research, CNCS-UEFISCDI, projects PN-II-ID-PCE-2011-3-0047 PN-II-RU-2011-TE-0008 and PN-II-PCCA-2011-142 and by EU Commission, project FP7 Marie Curie-PIRSES-"MS-Life"-2010-269-256.

References

1. Almeida R, Mosoarca C, Chirita M et al (2008) Coupling of fully automated chip-based electrospray ionization to high-capacity ion trap mass spectrometer for ganglioside analysis. Anal Biochem 378:43–52
2. Ammirati M, Cobbs CS, Linskey ME et al (2010) The role of retreatment in the management of recurrent/progressive brain metastases: a systematic review and evidence-based clinical practice guideline. J Neurooncol 96:85–96
3. Augustinsson LE, Blennow K, Blomstrand C et al (1997) Intracerebroventricular administration of GM1 ganglioside to presenile Alzheimer patients. Dement Geriatr Cogn Disord 8:26–33
4. Bagce HF, Saleh S, Adamovich SV, Tunik E (2012) Visuomotor gain distortion alters online motor performance and enhances primary motor cortex excitability in patients with stroke. Neuromodulation 15:361–369
5. Birks SM, Danquah JO, King L et al (2011) Targeting the GD3 acetylation pathway selectively induces apoptosis in glioblastoma. Neuro Oncol 13:950–960
6. Bleich-Cohen M, Hendler T, Weizman R et al (2013) Working memory dysfunction in schizophrenia patients with obsessive-compulsive symptoms: an fMRI study. Eur Psychiatry 29(3): 160–166
7. Cameron M, Moran P (2009) Prenatal screening and diagnosis of neural tube defects. Prenat Diagn 29:402–411
8. Cannella MS, Oderfeld-Nowak B, Gradkowska M et al (1990) Derivatives of ganglioside GM1 as neuronotrophic agents: comparison of in vivo and in vitro effects. Brain Res 513:286–294
9. Chahlavi A, Rayman P, Richmond AL et al (2005) Glioblastomas induce T-lymphocyte death by two distinct pathways involving gangliosides and CD70. Cancer Res 65:5428–5538
10. Colorado RA, Shukla K, Zhou Y et al (2012) Multi-task functional MRI in multiple sclerosis patients without clinical disability. Neuroimage 59:573–581
11. Di Pietro F, McAuley JH, Parkitny L et al (2013) Primary somatosensory cortex function in complex regional pain syndrome: a systematic review and meta-analysis. J Pain 14:1001–1018
12. Di Pietro F, McAuley JH, Parkitny L et al (2013) Primary motor cortex function in complex regional pain syndrome: a systematic review and meta-analysis. J Pain 14:1270–1288
13. Diaz Y, Gonzalez A, Lopez A et al (2009) Anti-ganglioside anti-idiotypic monoclonal antibody-based cancer vaccine induces apoptosis and antiangiogenic effect in a metastatic lung carcinoma. Cancer Immunol Immunother 58:1117–1128
14. Domon B, Costello CE (1988) A systematic nomenclature for carbohydrate fragmentation in FAB-MS/MS spectra of glycoconjugates. Glycoconj J 5:397–409
15. Egge H, Peter-Katalinić J, Reuter G et al (1985) Analysis of gangliosides using fast atom bombardment mass spectrometry. Chem Phys Lipids 37:127–141

16. Fish RG (1996) Role of gangliosides in tumour progression: a molecular target for cancer therapy? Med Hypotheses 46:140–144
17. Flangea C, Serb A, Sisu E et al (2011) Chip-based nanoelectrospray mass spectrometry of brain gangliosides. Biochim Biophys Acta 1811:513–535
18. Flangea C, Fabris D, Vukelić Ž et al (2013) Mass spectrometry of gangliosides from human sensory and motor cortex. Aust J Chem 66:781–790
19. Fredman P, Hedberg K, Brezicka T (2003) Gangliosides as therapeutic targets for cancer. BioDrugs 17:155–167
20. Gaspar LE, Mehta MP, Patchell RA et al (2010) The role of whole brain radiation therapy in the management of newly diagnosed brain metastases: a systematic review and evidence based clinical practice guideline. J Neurooncol 96:17–32
21. Grahn JA, Parkinson JA, Owen AM (2008) The cognitive functions of the caudate nucleus. Prog Neurobiol 86:141–155
22. Gow CH, Chien CR, Chang YL (2008) Radiotherapy in lung adenocarcinoma with brain metastases: effects of activating epidermal growth factor receptor mutations on clinical response. Clin Cancer Res 14:162–168
23. Hakomori S, Handa K (2002) Glycosphingolipid-dependent cross-talk between glycosynapses interfacing tumor cells with their host cells: essential basis to define tumor malignancy. FEBS Lett 531:88–92
24. Hallinan JT, Hegde AN, Lim WE (2013) Dilemmas and diagnostic difficulties in meningioma. Clin Radiol 68:837–844
25. Han SJ, Yang I, Tihan T et al (2010) Primary gliosarcoma: key clinical and pathologic distinctions from glioblastoma with implications as a unique oncologic entity. J Neurooncol 96:313–320
26. Handa K, Hakomori SI (2012) Carbohydrate to carbohydrate interaction in development process and cancer progression. Glycoconj J 29:627–637
27. Hedberg KM, Mahesparan R, Read TA et al (2001) The glioma-associated gangliosides 3′-isoLM1, GD3 and GM2 show selective area expression in human glioblastoma xenografts in nude rat brains. Neuropathol Appl Neurobiol 27:451–464
28. Hegde AN, Mohan S, Lim CC (2012) CNS cavernous haemangioma: “popcorn” in the brain and spinal cord. Clin Radiol 67:380–388
29. Huffman K (2012) The developing, aging neocortex: how genetics and epigenetics influence early developmental patterning and age-related change. Front Genet 3:212–220
30. IUPAC-IUB Joint Commission on Biochemical Nomenclature (1998) Nomenclature of glycolipids. Eur J Biochem 257:293–298
31. Kaas JH (2012) The evolution of neocortex in primates. Prog Brain Res 195:91–102
32. Kalanj S, Kracun I, Rosner H, Cosović C (1991) Regional distribution of brain gangliosides in Alzheimer’s disease. Neurol Croat 40:269–281
33. Kniep B, Kniep E, Ozkucur N et al (2006) 9-O-acetyl GD3 protects tumor cells from apoptosis. Int J Cancer 119:67–73
34. Kojovic M, Bologna M, Kassavetis P et al (2012) Functional reorganization of sensorimotor cortex in early Parkinson disease. Neurology 78:1441–1448
35. Kracun I, Kalanj S, Talan-Hranilovic J et al (1992) Cortical distribution of gangliosides in Alzheimer’s disease. Neurochem Int 20:433–438
36. Krafft C, Neudert L, Simat T et al (2005) Near infrared Raman spectra of human brain lipids. Spectrochim Acta A Mol Biomol Spectrosc 61:1529–1535
37. Kroes RA, He H, Emmett MR et al (2010) Overexpression of ST6GalNAcV, a ganglioside-specific alpha2,6-sialyltransferase, inhibits glioma growth in vivo. Proc Natl Acad Sci U S A 107:12646–12651
38. Lefebvre L (2012) Primate encephalization. Prog Brain Res 195:393–412
39. Lehnhardt FG, von Smekal U, Rűckriem B et al (2005) Value of gradient-echo magnetic resonance imaging in the diagnosis of familial cerebral cavernous malformation. Arch Neurol 62:653–658

40. Levery SB (2005) Glycosphingolipid structural analysis and glycosphingolipidomics. Methods Enzymol 405:300–369
41. Lode HN, Schmidt M, Seidel D et al (2013) Vaccination with anti-idiotype antibody ganglidiomab mediates a GD(2)-specific anti-neuroblastoma immune response. Cancer Immunol Immunother 62:999–1010
42. Louis David N, Hiroko O, Wiestler D (2007) The 2007 WHO classification of tumours of the central nervous system. Acta Neuropathol 114:97–109
43. Lukas RV, Nicholas MK (2013) Update in the treatment of high-grade gliomas. Neurol Clin 31:847–867
44. Lutterbach J, Guttenberger R, Pagenstecher A (2001) Gliosarcoma: a clinical study. Radiother Oncol 61:57–64
45. Meck WH, Church RM, Matell MS (2013) Hippocampus, time, and memory—a retrospective analysis. Behav Neurosci 127:642–654
46. Messner MC, Cabot MC (2010) Glucosylceramide in humans. Adv Exp Med Biol 688:156–164
47. Mitsuzuka K, Handa K, Satoh M et al (2005) A specific microdomain ("glycosynapse 3") controls phenotypic conversion and reversion of bladder cancer cells through GM3-mediated interaction of alpha3beta1 integrin with CD9. J Biol Chem 280:35545–35553
48. Mochizuki Y, Mizutani T, Shimizu T et al (2011) Proportional neuronal loss between the primary motor and sensory cortex in amyotrophic lateral sclerosis. Neurosci Lett 503:73–75
49. Mosoarca C, Ghiulai RM, Novaconi CR et al (2011) Application of chip-based nanoelectrospray ion trap mass spectrometry to compositional and structural analysis of gangliosides in human fetal cerebellum. Anal Lett 44:1036–1049
50. Nelles G, Jentzen W, Bockisch A et al (2011) Neural substrates of good and poor recovery after hemiplegic stroke: a serial PET study. J Neurol 258:2168–2175
51. Noguchi M, Suzuki T, Kabayama K et al (2007) GM3 synthase gene is a novel biomarker for histologic classification and drug sensitivity against epidermal growth factor receptor tyrosine kinase inhibitors in non-small-cell lung cancer. Cancer Sci 98:1625–1632
52. Oikawa N, Yamaguchi H, Ogino K et al (2009) Gangliosides determine the amyloid pathology of Alzheimer's disease. Neuroreport 20:1043–1046
53. Park Y, Lebrilla CB (2005) Application of Fourier transform ion cyclotron resonance mass spectrometry to oligosaccharides. Mass Spectrom Rev 24:232–264
54. Pellizzaro Venti M, Paciaroni M et al (2012) Caudate infarcts and hemorrhages. Front Neurol Neurosci 30:137–140
55. Pernber Z, Blennow K, Bogdanovic N et al (2012) Altered distribution of the gangliosides GM1 and GM2 in Alzheimer's disease. Dement Geriatr Cogn Disord 33:174–188
56. Relini A, Marano N, Gliozzi A (2013) Probing the interplay between amyloidogenic proteins and membranes using lipid monolayers and bilayers. Adv Colloid Interface Sci. doi:10.1016/j.cis.2013.10.015
57. Rodriguez FJ, Lim KS, Bowers D et al (2013) Pathological and molecular advances in pediatric low-grade astrocytoma. Annu Rev Pathol 8:361–379
58. Ryan JM, Rice GE, Mitchell MD (2013) The role of gangliosides in brain development and the potential benefits of perinatal supplementation. Nutr Res 33:877–887
59. Sandhoff K, Harzer K (2013) Gangliosides and gangliosidoses: principles of molecular and metabolic pathogenesis. J Neurosci 33:10195–10208
60. Santin AD, Ravindranath MH, Bellone S et al (2004) Increased levels of gangliosides in the plasma and ascitic fluid of patients with advanced ovarian cancer. BJOG 111:613–618
61. Schiopu C, Flangea C, Capitan F et al (2009) Determination of ganglioside composition and structure in human brain hemangioma by chip-based nanoelectrospray ionization tandem mass spectrometry. Anal Bioanal Chem 395:2465–2477
62. Schiopu C, Vukelić Z, Capitan F et al (2012) Chip-nanoelectrospray quadrupole time-of-flight tandem mass spectrometry of meningioma gangliosides: a preliminary study. Electrophoresis 33:1778–1786

63. Schneider JS, Sendek S, Daskalakis C et al (2010) GM1 ganglioside in Parkinson's disease: results of a five year open study. J Neurol Sci 292:45–51
64. Schneider JS, Gollomp SM, Sendek S et al (2013) A randomized, controlled, delayed start trial of GM1 ganglioside in treated Parkinson's disease patients. J Neurol Sci 324:140–148
65. Serb A, Sisu E, Vukelić Z et al (2012) Profiling and sequencing of gangliosides from human caudate nucleus by chip-nanoelectrospray mass spectrometry. J Mass Spectrom 47:1561–1570
66. Seyfried TN, Mukherjee P (2010) Ganglioside GM3 is antiangiogenic in malignant brain cancer. J Oncol 2010:961243
67. Shibuya H, Hamamura K, Hotta H et al (2012) Enhancement of malignant properties of human osteosarcoma cells with disialyl gangliosides GD2/GD3. Cancer Sci 103:1656–1664
68. Shukla GS, Krag DN (2006) Selective delivery of therapeutic agents for the diagnosis and treatment of cancer. Expert Opin Biol Ther 6:39–54
69. Sisu E, Flangea C, Serb A et al (2011) High-performance separation techniques hyphenated to mass spectrometry for ganglioside analysis. Electrophoresis 32:1591–1609
70. Steiner G, Shaw A, Choo-Smith LP et al (2003) Distinguishing and grading human gliomas by IR spectroscopy. Biopolymers 72:464–471
71. Svennerholm L, Fredman P (1980) A procedure for the quantitative isolation of brain gangliosides. Biochim Biophys Acta 617:97–109
72. Taki T (2012) An approach to glycobiology from glycolipidomics: ganglioside molecular scanning in the brains of patients with Alzheimer's disease by TLC-blot/matrix assisted laser desorption/ionization-time of flight MS. Biol Pharm Bull 35:1642–1647
73. Tang T, Kmet M, Corral L et al (2005) Testisin, a glycosyl-phosphatidylinositol-linked serine protease, promotes malignant transformation in vitro and in vivo. Cancer Res 65:868–878
74. Tivnan A, Orr WS, Gubala V et al (2012) Inhibition of neuroblastoma tumor growth by targeted delivery of microRNA-34a using anti-disialoganglioside GD2 coated nanoparticles. PLoS One 7:e38129
75. Tokuda N, Zhang Q, Yoshida S (2006) Genetic mechanisms for the synthesis of fucosyl GM1 in small cell lung cancer cell lines. Glycobiology 16:916–925
76. Yamazaki T, Suzuki M, Irie T et al (2008) Amyotrophic lateral sclerosis associated with IgG anti-GalNAc-GD1a antibodies. Clin Neurol Neurosurg 110:722–724
77. Yu RK, Tsai YT, Ariga T (2012) Functional roles of gangliosides in neurodevelopment: an overview of recent advances. Neurochem Res 37:1230–1244
78. Valdes-Gonzalez T, Goto-Inoue N et al (2011) New approach for glyco- and lipidomics-molecular scanning of human brain gangliosides by TLC-Blot and MALDI-QIT-TOF MS. J Neurochem 116:678–683
79. Vriend C, Raijmakers P, Veltman DJ et al (2014) Depressive symptoms in Parkinson's disease are related to reduced [123I]FP-CIT binding in the caudate nucleus. J Neurol Neurosurg Psychiatry 85(2):159–164
80. Vukelić Z, Metelmann W, Müthing J et al (2001) Anencephaly: structural characterization of gangliosides in defined brain regions. Biol Chem 382:259–274
81. Vukelić Z, Zamfir AD, Bindila L et al (2005) Screening and sequencing of complex sialylated and sulfated glycosphingolipid mixtures by negative ion electrospray Fourier transform ion cyclotron resonance mass spectrometry. J Am Soc Mass Spectrom 16:571–580
82. Vukelić Z, Zarei M, Peter-Katalinić J et al (2006) Analysis of human hippocampus gangliosides by fully-automated chip-based nanoelectrospray tandem mass spectrometry. J Chromatogr A 1130:238–245
83. Vukelić Z, Kalanj-Bognar S, Froesch M et al (2007) Human gliosarcoma-associated ganglioside composition is complex and distinctive as evidenced by high-performance mass spectrometric determination and structural characterization. Glycobiology 17:504–515
84. Wagener R, Röhn G, Schillinger G et al (1999) Ganglioside profiles in human gliomas: quantification by microbore high performance liquid chromatography and correlation to histomorphology and grading. Acta Neurochir (Wien) 141:1339–13345

85. Walcott BP, Nahed BV, Brastianos PK et al (2013) Radiation treatment for WHO grade II and III meningiomas. Front Oncol 3:227
86. Walker FO (2007) Huntington's disease. Lancet 369:218
87. Wessels PH, Weber WE, Raven G et al (2003) Supratentorial grade II astrocytoma: biological features and clinical course. Lancet Neurol 2:395–403
88. Zamfir AD, Vukelić Z, Bindila L et al (2004) Fully-automated chip-based nanoelectrospray tandem mass spectrometry of gangliosides from human cerebellum. J Am Soc Mass Spectrom 15:1649–1657
89. Zamfir AD, Serb A, Vukelić Ž et al (2011) Assessment of the molecular expression and structure of gangliosides in brain metastasis of lung adenocarcinoma by an advanced approach based on fully automated chip-nanoelectrospray mass spectrometry. J Am Soc Mass Spectrom 22:2145–2159
90. Zamfir AD, Fabris D, Capitan F et al (2013) Profiling and sequence analysis of gangliosides in human astrocytoma by high-resolution mass spectrometry. Anal Bioanal Chem 405:7321–7335
91. Zhang Q, Furukawa K, Chen HH et al (2006) Metastatic potential of mouse Lewis lung cancer cells is regulated via ganglioside GM1 by modulating the matrix metalloprotease-9 localization in lipid rafts. J Biol Chem 281:18145–18155

Chapter 9
Mass Spectrometric Analysis of Post-translational Modifications (PTMs) and Protein–Protein Interactions (PPIs)

Armand G. Ngounou Wetie, Alisa G. Woods, and Costel C. Darie

Abstract Of the 25,000–30,000 human genes, about 2 % code for proteins. However, there are about one to two million protein entities. This is primarily due to alternative splicing and post-translational modifications (PTMs). Identifying all these modifications in one proteome at a particular time point during development or during the transition from normal to cancerous cells is a great challenge to scientists. In addition, identifying the biological significance of all these modifications, as well as their nature, such as stable versus transient modifications, is an even more challenging. Furthermore, interaction of proteins and protein isoforms that have one or more stable or transient PTMs with other proteins and protein isoforms makes the study of proteins daunting and complex. Here we review some of the strategies to study proteins, protein isoforms, protein PTMs, and protein–protein interactions (PPIs). Our goal is to provide a thorough understanding of these proteins and their isoforms, PTMs and PPIs and to shed light on the biological significance of these factors.

9.1 Introduction

9.1.1 Protein–Protein Interactions (PPIs)

Of the 25,000–30,000 human genes, about 2 % code for proteins. However, there are about one to two million protein entities. This is primarily due to alternative splicing, post-translational modifications (PTMs). After the completion of the human genome project, a wealth of information and data was gained which now

A.G. Ngounou Wetie • A.G. Woods • C.C. Darie (✉)
Biochemistry & Proteomics Group, Department of Chemistry & Biomolecular Science, Clarkson University, 8 Clarkson Avenue, Potsdam, NY 13699-5810, USA
e-mail: cdarie@clarkson.edu

A.G. Woods and C.C. Darie (eds.), *Advancements of Mass Spectrometry in Biomedical Research*, Advances in Experimental Medicine and Biology 806,
DOI 10.1007/978-3-319-06068-2_9, © Springer International Publishing Switzerland 2014

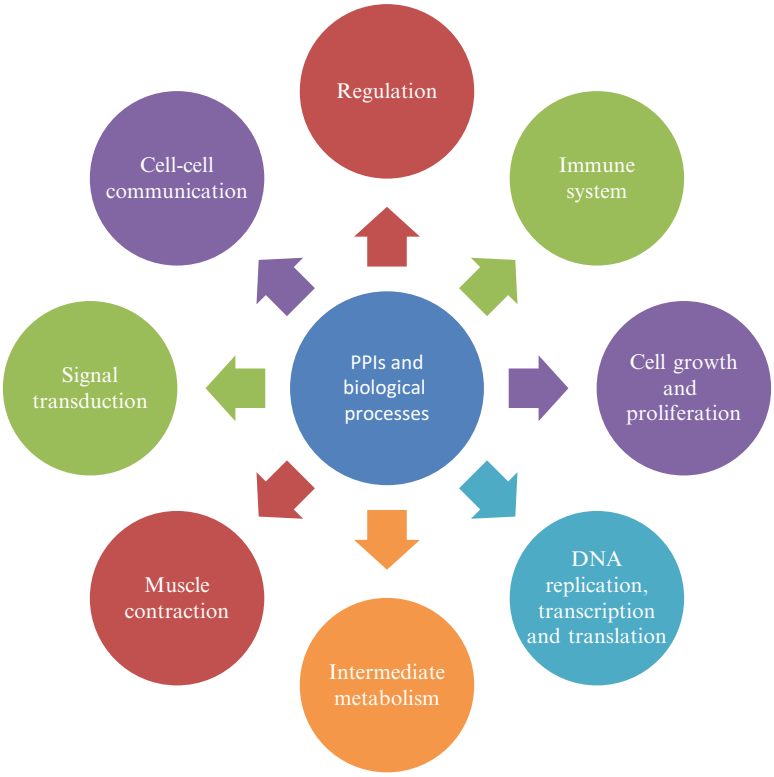


Fig. 9.1 PPIs play a role in almost every biological process and are thus important for life. Their investigation will result in the understanding of pathological and physiological processes. Reprinted and adapted with permission from *Proteomics* [4]

need to be translated into biological function. This genomic era led to the birth of a new discipline called proteomics [1]. Proteomics sets out to identify proteins, their sequence and known modifications as well as their quantitation in a biological sample for the purpose of understanding biological processes, protein cellular functions, and their physiological and pathological involvement in diseases [2, 3]. Proteins interact at the molecular level with other proteins, nucleic acids, lipids, carbohydrates, and metabolites to perform numerous cellular activities (Fig. 9.1). Interactions depend on many factors, such as cell-type, developmental stage of the cell, cell-cycle phase, external conditions, and obviously the presence of other proteins (interactors). Physical contact between proteins may trigger conformational changes or PTMs that modulate the activity of those proteins. Physical contact between proteins is not always static and permanent; there is continuous reassembly and turnover. Protein complexes can consist of sets of more stably (stable PPIs) and less stably (transient PPIs) interacting proteins or combination of both [4–11]. Examples of stable complexes include hemoglobin, RNA polymerase, ribosomes, tubulin α and β [12–16], whereas most interactions involving nuclear proteins and transcription factors are dynamic and transient [17]. One can also distinguish between

extracellular PPIs and intracellular PPIs. Particularly, extracellular PPIs involve secreted proteins and membrane-associated proteins and constitute promising therapeutic targets [18]. PPIs can occur by means of hydrophobic interactions, van der Waals forces, and ionic interactions [19, 20].

For example, the well-known p53 protein is activated by a diverse set of intra- and extracellular events such as DNA damage, oncogenes, oxidative stress, and then interact with a plethora of other molecules to mediate cell cycle regulation. However, mutations in p53 can lead to uncontrolled proliferation of cells and consequently to various types of cancer [21]. Based on this example, it can be seen that genotype, environment, and phenotype influence each other resulting in complex molecular interaction patterns and that disruption of these interactions can lead to pathologic conditions. Diseases such as cancers, cardiovascular disorders, diabetes, and neurological dysfunctions are multifactorial, i.e., they arise from the involvement of many genes and also possess multiple phenotypes. Therefore, the convergent mechanisms leading to disease generation must be understood for the development of efficient diagnostic and therapeutic approaches. PPIs span a wide range of cellular processes such as metabolism, cell cycle control, and signal transduction [22]. Perturbations of these PPIs, for instance, through mutations, can result in a pathology such as the impairment of the BRCA2 interaction with replication protein A that leads to the generation of carcinogenic DNA changes [23, 24]. The rationale behind monitoring proteins in general and PPIs in particular with regard to disease onset and progression is manifold. First, it has been shown that interacting proteins may have similar functions and thus may be involved in the same diseases [22]. Second, disease-relevant interactions have a higher affinity compared to disease-unrelated ones [25, 26]. In the past, mostly genetic and biochemical techniques were used for the investigation of PPIs. However, recently there has been a shift towards novel techniques such as mass spectrometry (MS) and increased use of bioinformatics tools [4, 5]. Two of the most-used techniques for PPI studies are the yeast two-hybrid (Y2H) system and tandem affinity purification coupled to mass spectrometry (TAP-MS). MS-based investigation of disease-related PPIs opens the door for the determination of novel protein functions and thereby the unraveling of yet undiscovered novel pathogenic mechanisms. Nevertheless, all the methods and techniques employed in the investigation of PPIs have advantages as well as disadvantages and therefore, it is important to analyze PPIs with more than one approach for both confirmation and validation in order to avoid false-positive results.

9.1.2 Post-translational Modifications (PTMs)

The presence of several functional groups on protein backbones allow them to be readily reversible, chemically modified following their biosynthesis. PTMs are more predominant in eukaryotes than prokaryotes, with about 5 % of the genome of eukaryotes coding for enzymes involved in protein modifications. PTMs are named according to the functional groups transferred onto the amino acid residues. Correspondingly, the transfer of phosphate, carbohydrate, methyl, and ubiquitin has

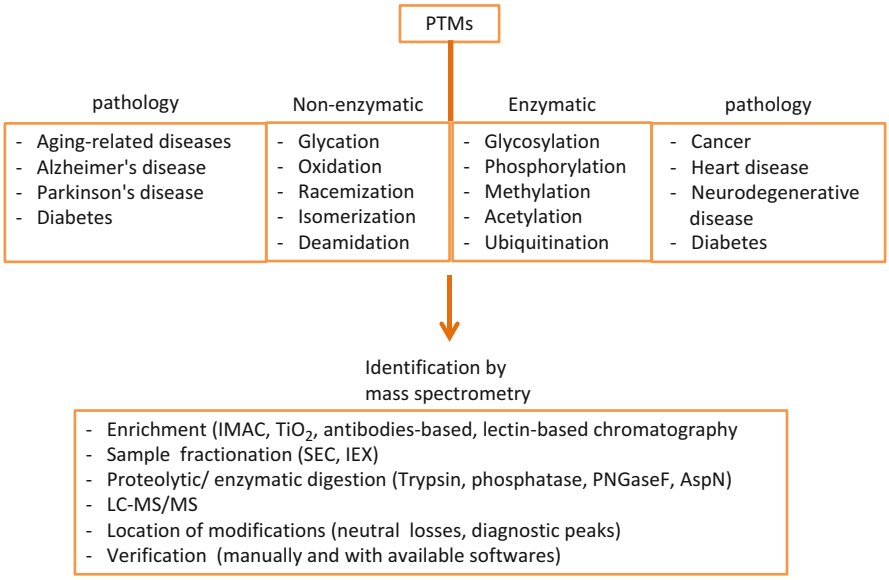


Fig. 9.2 Nonenzymatic and enzymatic PTMs and the respective diseases and disorders they have been related to as well as the strategy of identification of these PTMs by MS. Reprinted and adapted with permission from the *Australian Journal of Chemistry CSIRO Publishing* http://www.publish.csiro.au/?paper=CH13144 [35]

denoted phosphorylation, glycosylation, methylation, and ubiquitination, respectively. Two types of PTMs exist: enzyme-catalyzed modification (e.g., phosphorylation, glycosylation) and covalent cleavage (glycation, oxidation). In the case of enzyme-mediated PTMs, electrophilic groups are covalently added to the nucleophile side chain residues, whereas nonenzyme-catalyzed modifications occur simply through a breakage of the peptide backbones. Common groups of enzymatic units responsible for protein PTMs include kinases (phosphorylation), ubiquitin ligases (ubiquitination), and acetyl-transferases (acetylation) and these enzymes are tightly regulated and conserved across species. It has been well demonstrated that diseases such as cancer are characterized by a high rate of mutations of protein PTM effectors (kinases, ubiquitin E3 ligases) [27]. Further, many PTM-mediating enzymes are themselves regulated by other PTMs resulting in complex PTM networks [28, 29]. Of the 22 proteinogenic amino acids known, Leu, Ile, Val, Ala, and Phe do not undergo modifications in vivo, whereas some amino acids can harbor multiple distinct modifications for example, lysine can be methylated, acetylated, or ubiquitinated) [30–32]. Through these modifications, proteins gain in function and structure [30]. There are approximately 200 or more different PTMs that have been reported for the human proteome. Furthermore, protein PTMs control the activation of proteins and therefore the regulation of many processes [9, 33–45]. Their relevance in physiological and pathological processes no longer needs to be proven (Fig. 9.2). It is likely that there are still a huge amount of undetected protein PTMs

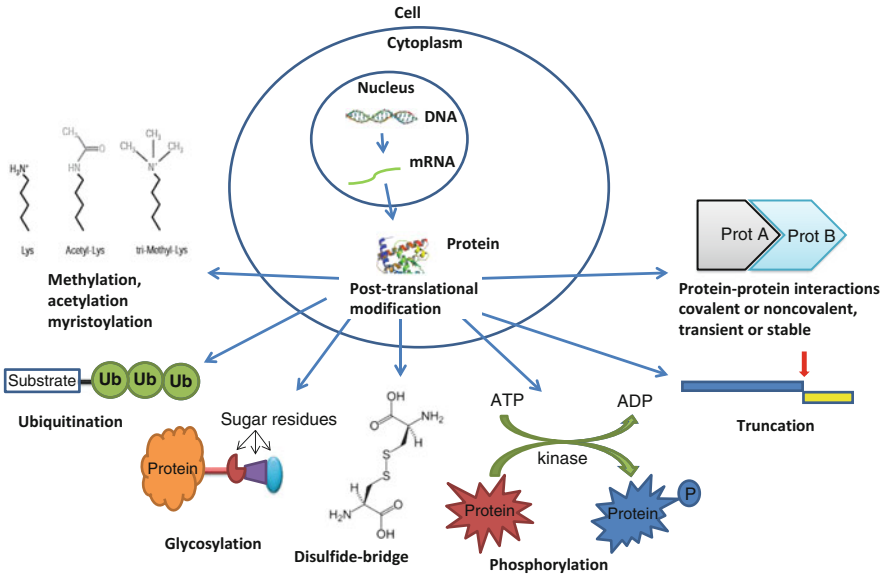


Fig. 9.3 Illustration of the modifications and interactions that protein undergo post-synthesis by the cell creating the huge diversity known among proteins. Reprinted and adapted with permission from the *Journal of Molecular Psychiatry* [37]

that may play a role in some diseases, but they are as of yet undiscovered due to limitations of sample preparations, analytical techniques, data analysis, or the short lifespan of these modifications. A review of the literature shows that PTMs have been analyzed mostly in immortalized cancer cell lines and in large-scale experiments in yeast and bacteria. In contrast, primary cells or tissues have been understudied so far likely due to the difficulty of handling them and because of their very low protein content and low protein yield [31, 46–49]. While immortalized cancer cell lines are characterized by mutational changes in regulatory proteins of the cell cycle, cell–cell communication and apoptosis; primary cells can build a more accurate picture of the physiological state in vivo [50–52]. There are a variety of techniques available for the analysis of protein PTMs. One of the most employed, most advanced methods used in the exploration of PTMs is MS-based proteomics. Pairing MS to in vitro or in vivo biological assays (cell culture-based or knock-out models) enables the extraction of biological function out of proteomic experiments [53].

9.1.3 Common PTMs

Common PTMs include oxidation, glycosylation, phosphorylation, methylation, acetylation, and ubiquitination (Fig. 9.3).

Oxidation of proteins is a very common process mediated by the so-called reactive oxygen species (ROS: superoxide, hydrogen peroxide, hydroxyl radical), reactive nitrogen species (RNS: nitric oxide, dinitrogen trioxide, peroxynitrite, nitroxyl anion), and reactive lipid species (RLS: lipid aldehydes) leading to DNA base modification, phospholipid damage, irreversible protein oxidation, cell death, or mutagenesis [54–56]. Amino acids that react mostly with reactive species include tyrosine and cysteine. For instance, protein nitrotyrosine have been associated with the development of several pathologies including heart failure, atherosclerosis, aging, and hypertension [57].

Glycosylation is the most prevalent protein modification (about 50 % of protein modifications) in the human proteome [58]. One can distinguish between N-linked glycosylation at a consensus sequence of NXS/T via the nitrogen of side chains of extracellular proteins and O-linked glycosylation via the hydroxyl oxygen of serine, threonine, tyrosine, hydroxylysine, or hydroxyproline side chains of extracellular, nuclear, and cytoplasmic proteins [59–63]. Enzymes called glycosyltransferases recognize specific protein motifs and then transfer the first monosaccharide (or preformed oligosaccharide for N-glycosylation) onto the recognition site. Other glycosyltransferases (and glycosidases for N-glycosylation) then sequentially elongate the glycan sequence. Of the two forms of glycosylation (N- and O-glycosylation), N-glycosylation is the most common. Glycosylation of proteins play a major role in the proper folding and stabilization of proteins as well as in cell–cell adhesion and communication [64]. In some cases of cancer and infectious diseases, deficient or absent glycosylation has been proposed as a possible leading cause of these diseases [65]. Moreover, numerous drugs on the markets are glycoproteins that have well-characterized glycan entities necessary for their function, efficacy, and safety [66]. Glycation, on the other hand, refers to the nonenzymatic addition of sugar aldehyde or ketone to the Ɛ-amino group of lysines or the N-terminal amino group of proteins. The glycan reaction, also known as the Maillard reaction, leads to the irreversible formation of advanced glycation end products (AGE), which can cross-link proteins thus rendering them detergent insoluble and protease resistant [67, 68].

Phosphorylation is the transfer of a phosphoryl group from adenosine triphosphate (ATP) or guanosine triphosphate (GTP) to primarily serine (86.4 %), threonine (11.8 %), and tyrosine (1.8 %) residues through formation of a phosphoester bond [31]. However, phosphorylation of histidine, aspartate, and arginine has also been observed [69, 70]. Hydrolysis of the phosphoester bond releases orthophosphate and is called dephosphorylation [71, 72]. It participates in a multitude of cellular processes including protein synthesis, activity and stability, inter-/intracellular signaling, gene expression, cell survival, and apoptosis [73, 74]. Consequently, aberrant phosphorylation status of proteins has been related to many diseases such as cancer. Enzymes that mediate phosphorylation or dephosphorylation are kinases and phosphatases, respectively. There are more than 500 kinase-coding human

genes and about 150 phosphatases, which points to their large involvement in diverse cellular and biological events [75–77].

Methylation has been shown to occur mostly on lysine and arginine residues and to a lesser extent also on histidine, cysteine, aspartic acid, glutamic acid, serine, and threonine [78, 79]. Biological processes such as transcriptional regulation, RNA processing, metabolism, and signal transduction are all shaped by methylation [78–80]. Methyltransferases are enzymes responsible for the transfer of one or more methyl groups from *S*-adenosyl-methionine (SAM) to methylation recognition sequence or motifs within proteins [81]. Methylation, acetylation, phosphorylation, and ubiquitination are able to cross-talk as shown by the prevention of protein ubiquitination through the methylation of lysine, thus prolonging the half-life of the protein [82, 83].

Ubiquitination is one of the main mechanisms by which protein homeostasis, cell cycle progression, gene transcription, receptor transport, and immune responses are controlled [84]. Ubiquitination is a complex process that requires the action of three different enzymes: ubiquitin-activating enzyme E1, ubiquitin-conjugating enzyme E2, and ubiquitin-ligating enzyme E3. E1 activates ubiquitin in an ATP-dependent manner leading to the transfer of ubiquitin to the active cysteine site of E1 through thioester bond formation and release of AMP. Subsequently, ubiquitin is conjugated to the active cysteine site of E2 and thus allowing E3 to attach it to a lysine residue on a particular protein through a thio-esterification reaction [85]. On the other hand, the so-called de-ubiquitinating enzymes (DUBs) can remove the covalently attached ubiquitin [86, 87]. Due to its comparable size (about 700 E1, E2, E3, and DUBs) to the phosphorylation system (650 kinases and phosphatases [88–90]), ubiquitination is considered a proteome-wide modification. A distinctive feature between phosphorylation and ubiquitination is the fact that contrary to phosphorylation, sequence recognition in ubiquitination is based on the accessibility of lysine residues or the specificity of the E3 enzyme or E2/E3 enzymes [29, 91–94].

Originally thought to take place only on histones, *acetylation* has been now reported on more than 80 transcription factors and other nuclear regulatory molecules. As such, it mainly affects transcription and metabolic regulation. Interestingly, enzymes involved in glycolysis, gluconeogenesis, the tricarboxylic acid (TCA) cycle, the urea cycle, fatty acid metabolism, and glycogen metabolism are acetylated [95–99]. The process of acetylation is catalyzed by histone acetyltransferases (HATs), which transfer the acetyl group from acetyl-coenzyme A to the amino group of lysine at the N-terminus of histones to form 3-*N*-acetyl lysine. This reaction can be reversed by the action of histone deacetylases (HDACs) [96]. Both, HATs and HDACs are potential drug targets for many disorders such as obesity, cancer, and neurodegenerative disorders [100–103].

9.2 Mass Spectrometry (MS)

9.2.1 Introduction

The technique of MS has revolutionized the biochemistry of biomolecules and is constantly increasing and expanding. It is practically inconceivable today to imagine the characterization of biomolecules (in particular proteins) and molecular complexes without this tool. Briefly, a mass spectrometer is built of the following main components: an ion source, where biomolecules are positively or negatively ionized; an analyzer for the separation of gas phase ions according to their mass-to-charge (m/z) ratio and a detector that registers the signal in the form of a spectrum. Classically, we distinguish primarily between bottom-up proteomics which consist in the digestion of proteins into short peptides prior to MS [37, 39, 104–107] and top-down proteomics, which is the direct analysis (without digestion) by MS of intact whole proteins [108]. Moreover, other approaches have been developed recently such as the middle-down approach, in which fragmentation of large polypeptides (3–20 kDa) is undertaken [109]. Depending on the instrumentation available and the questions to be answered, one can go with one approach or the other. The top-down approach is only possible with high mass accuracy instruments such as a Fourier transform ion cyclotron resonance mass spectrometer (FT-ICR-MS). As for bottom-up proteomics, there is a wide range of instruments that can be used. Furthermore, there are other terminologies employed in the field of proteomics that need to be defined here. The term shotgun proteomics or DDA (data-dependent acquisition) refers to a selection of peptide ions for fragmentation based on certain predefined criteria whereas targeted methods (e.g., selected reaction monitoring (SRM)) focus on the transitions from precursor to product ions for quantitative purposes [110]. Therefore, the latter approach exhibits lower detection limits and wider dynamic ranges compared to the former one. In addition to different distinctive approaches, there are also diverse fragmentation options available that can be used with particular instrumentation types: collision-induced dissociation (CID) [111], electron transfer dissociation (ETD) [112], and electron capture dissociation (ECD) [113]. Fragmentation in CID occurs by collision with an inert gas (e.g., argon or helium) and is mostly used in ion trap instruments (low-energy CID), triple quadrupoles (QqQ), and quadrupole time-of-flight (Q-TOF) instruments (beam-type CID). ETD and ECD fragmentations utilize anions with low-electron affinities to transfer electrons to the cationic peptide resulting in the release of a hydrogen radical to the backbone carbonyl group [112]. This fragmentation mode gives rise to c- and z-ions contrary to the b- and y-series obtained with CID fragmentation. Another difference between ECD/ETD and CID is the charge of the peptides preferred for fragmentation as the former has a preference for triply charged (3+) ions rather than doubly charged (2+) ions for the latter. Based on the observation that the majority of peptide ions are doubly charged and the more rapid CID reaction, it is hypothesized that CID is more suitable for large-scale proteomics than ETD [114].

After the samples are run, data analysis and processing are performed with a plethora of available software. For instance, through the comparison of generated fragment ion spectra against a database containing deposited MS/MS spectra of in silico protein digests, high-confident protein identifications are made [115].

9.2.2 MS-Based Strategies to Analyze PPIs and PTMs

Prior to the development and widespread usage of MS, PPIs and PTMs were analyzed by a variety of biochemical and genetic methods. PPIs can be investigated by a variety of non-proteomic methods including genetic (e.g., Y2H screen), biochemical (e.g., pull-down, Co-IP or TAP combined with MS), biophysical (e.g., analytical ultracentrifugation, microscopy, FRET) techniques, and as of recently more and more by bioinformatics [116–121]. In the case of PTMs, methods such as Edman degradation, amino acid analysis, and radioisotope labeling were applied. In both instances, these methods were characterized by a low-throughput, low specificity and a narrow dynamic range. All these shortcomings are addressed by proteomics-based MS, which is the reason it is the go-to method for PPIs and PTM analyses [115, 122].

Covalently linked PPIs. PPIs can be investigated by a variety of different proteomics-based MS methods depending on the information output desired. Disulfide-linked proteins constitute a special case of covalently linked PPIs, which can be investigated using reducing (R) and nonreducing conditions (NR). Disulfide linkages can occur intermolecularly or intramolecularly or a mixture of both (Fig. 9.4a–d). An example of an intermolecular disulfide linkage, as determined by direct infusion of a cysteine-containing peptide in both monomeric and homodimeric forms, is shown (Fig. 9.4e–h). In fact, this example is particularly useful as it investigates formation and identification of the homodimeric disulfide bridges. Here, the peptide of interest can be identified as a triply charged (3+) peptide with mass to charge ratio (*m*/*z*) of 761.05 and as a doubly charged (2+) peptide with *m*/*z* of 1,141.09 (Fig. 9.4e). If this peptide can easily be identified as a monomer or as a disulfide-linked homodimer, with the homodimer formed when the monomer was in either (3+) with *m*/*z* of 761.05 (Fig. 9.4f) or (2+) with *m*/*z* of 1,141.09 (Fig. 9.4g), it is not that easy to conclude that this peptide is in homodimeric form, as the only difference between the monomeric and homodimeric forms is the charge state of (3+) versus (6+) (Fig. 9.4f) or the (2+) versus (4+) (Fig. 9.4g). However, another alternative for easy identification of only the disulfide-linked homodimer is to look for a peptide with an uneven number of charges. For example, if a peptide has two charges, the disulfide-linked dimer has four charges, but if we look for the (3+) or (5+) charged disulfide-linked peptide, then it is easier to identify the homodimer and to exclude any monomer. For example, the peak with *m*/*z* of 913.05 (Fig. 9.4e) has a charge of (5+), as demonstrated in Fig. 9.4h.

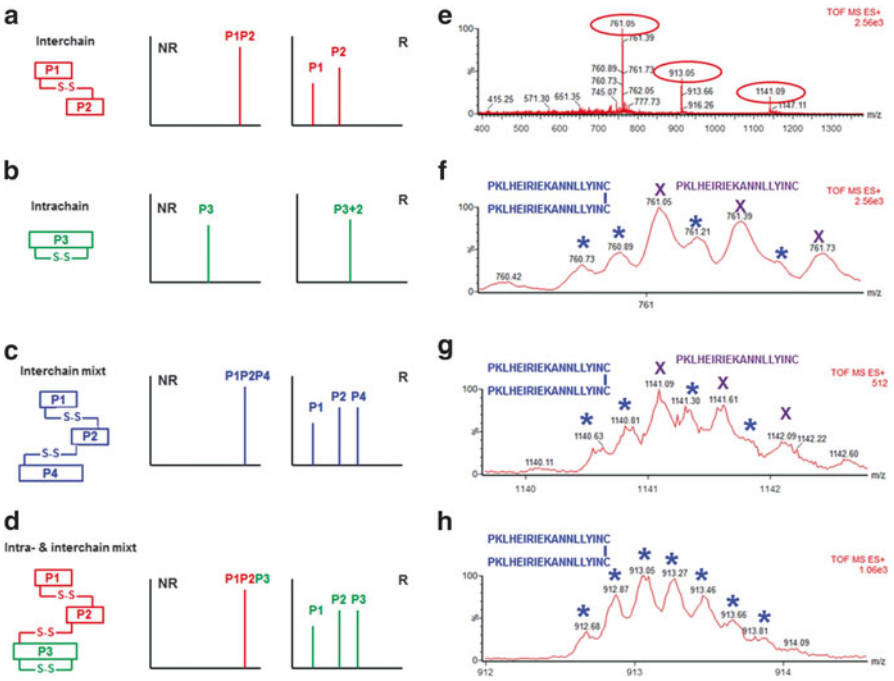


Fig. 9.4 Disulfide-linked proteomics as performed theoretically by matrix-assisted laser desorption-mass spectrometry (MALDI-MS) showing disulfide-linked peptides (P1, P2, P3, P4) connected by intrachain or interchain disulfide-bonds or a mixture of both and their analysis in reducing (R) and nonreducing (NR) conditions (**a**–**d**). ESI-MS of the cysteine-containing peptide PKLHEIRIEKANNLLYINC in monomeric (**e**–**g**) and dimeric disulfide-bonded (**e**, **h**) forms. The symbols x and "*"denote the monoisotopic peaks of the monomer (**f**, **g**) and of the dimer (**f**–**h**). Reprinted and adapted with permission from the *Journal of Laboratory Automation* [36]

For disulfide-linked proteins, usually a nonreducing SDS-PAGE (SDS-PAGE NR) in combination with MS can be undertaken (disulfide proteomics). The nonreducing gel can then be compared with a reducing gel (SDS-PAGE R) for the identification of disulfide-linked proteins (Fig. 9.5a) [38]. Based on this principle, three different sera samples were processed. While the protein bands in the SDS-PAGE R are identical for all the samples, differences between the three samples are visible in the SDS-PAGE NR (Fig. 9.5b). Those differences were determined by LC-MS/MS to be haptoglobin (45 kDa, Fig. 9.5c, d). These findings were also validated by Western blotting (WB) using antibodies against haptoglobin (Fig. 9.5e).

Transient and stable PPIs. Another gel-based technique for the investigation of protein complexes is the Blue Native-PAGE (BN-PAGE), which was developed by Schagger and von Jagow [123]. It was originally used for the isolation of membrane protein complexes from purified bovine mitochondria and has a dynamic range of 10–10,000 kDa [123, 124]. Later this method was used by others [7, 8, 11, 125, 126]. Figure 9.6 shows protein complexes that were separated by BN-PAGE (1D) and by Tricine (Fig. 9.6a–c) or SDS-PAGE (Fig. 9.6d) in a second dimension (2D)

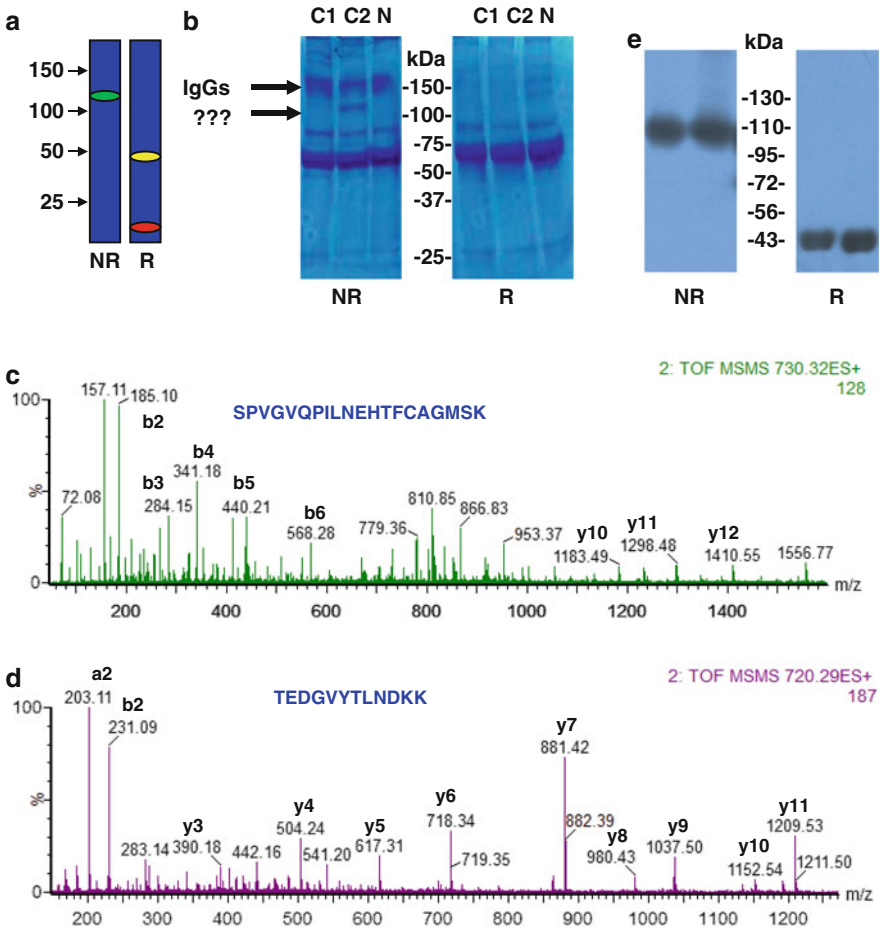


Fig. 9.5 Disulfide proteomics of sera samples in NR and R SDS-PAGE (**a**, **b**). The 110 kDa bands present in C1 and C2, but absent in N were analyzed by LC-MS/MS and the two peptides identified in these runs (**c**, **d**) were found to be part of the haptoglobin protein (Hp). Confirmation of these findings was done by WB (**e**). Reprinted and adapted with permission from *Electrophoresis* [38]

and stained to reveal the protein complexes (Fig. 9.6a–c, the horizontal band) or electroblotted and analyzed with several antibodies (with anti-Ndh antibodies (Fig. 9.6b, c) or anti-HRS antibodies (Fig. 9.6d)) by WB. The molecular mass of the intact protein complexes (1D) was determined using customized acrylamide-bisacrylamide gradients to accommodate different protein complexes of various sizes (Fig. 9.6a–c). Separation of the protein complexes in 1D followed by 2D (Fig. 9.6a) reveals the subunit composition of the protein complexes from thylakoid membranes: Photosystem I, NADH-dehydrogenase (Ndh complex), Ribulose bisphosphate carboxylase (RuBisCO), ATP synthase, or Cytochrome b6f complex (Cyt b6f complex). The Ndh complex (Mw of 550 kDa) divides into the 290-kDa

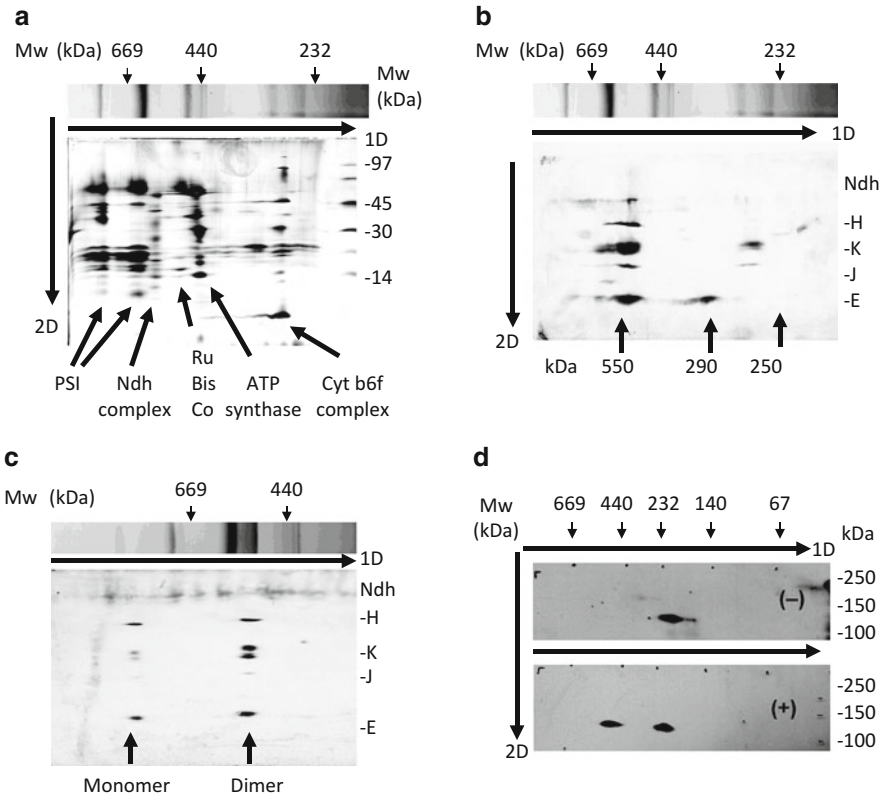


Fig. 9.6 Investigation of stable and transient PPIs of the thylakoid membranes and primary mouse neuronal cell cultures by BN-PAGE and WB. First, the proteins were separated on BN-PAGE 1D and Tricine PAGE 2D (**a**–**c**) or SDS-PAGE 2D (**d**). WB was undertaken using antibodies against Ndh (**b**, **c**) and HRS (**d**). (**a**) Subunit composition and molecular mass of protein complexes. (**b**) Disassembly of the Ndh complex (550 kDa) into the 290 kDa membrane Ndh subcomplex (monitored by anti-NdhE antibodies) and the 250 kDa soluble Ndh subcomplex (monitored by anti-NdhH, -K, and -J antibodies). (**c**) Monomeric and dimeric forms of the Ndh complex. (**d**) HRS protein complex formation upon stimulation of the cells with BDNF. Reprinted and adapted with permission from *Cellular Molecular Life Science* [5]

membrane Ndh subcomplex (monitored by anti-NdhE antibodies) and the 250-kDa soluble Ndh subcomplex (monitored by anti-NdhH, -K, and -J antibodies). Further, the Ndh complex is detected in monomeric and dimeric forms (Fig. 9.6c). Finally, BN-PAGE was used to investigate transient PPIs in primary neurons unstimulated (−) and BDNF stimulated (+). The appearance of an immunoreactive band in the BDNF stimulated (+) suggests that HRS protein interacts with other proteins and forms complexes of higher Mw (Fig. 9.6d). In another example, BN-PAGE followed by LC-MS/MS (Fig. 9.7A) as well as electron microscopy (Fig. 9.7B) was employed for the investigation of the multisubunit protein complexes (Fig. 9.7A) and

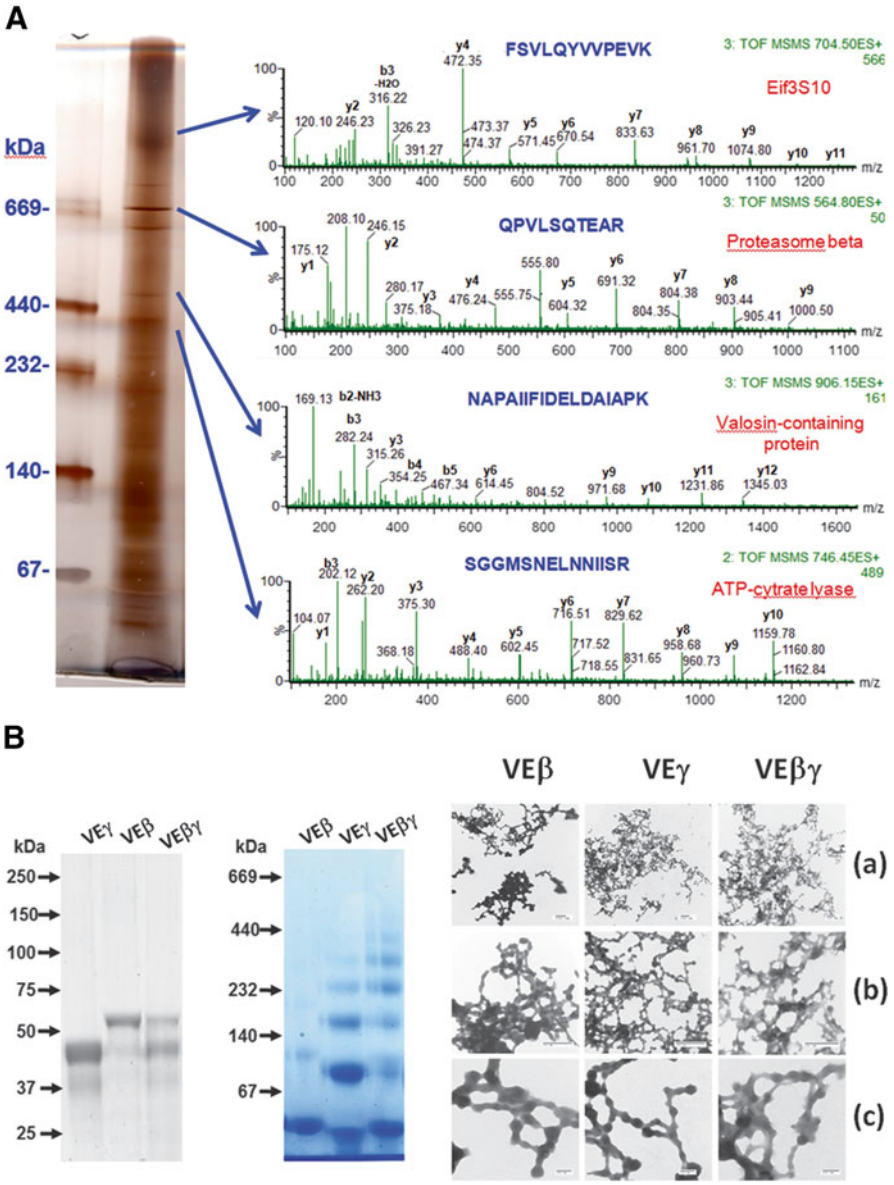


Fig. 9.7 Investigation of PPIs of cell-lysate from NG108 neuroblastoma×glioma cell. (**A**) Separation of cell-lysate by BN-PAGE and LC-MS/MS analysis of the gel bands corresponding to protein complexes. The sequences displayed are part of Eif3S10, proteasome beta, valosin-containing protein, and ATP citrate lyase. (**B**) Polymerization of vitelline envelope proteins (VE) beta and gamma was explored by SDS-PAGE (*left*), BN-PAGE (*middle*), and electron microscopy (EM) (*right*). Magnification is ×15,000 (**a**), ×60,000 (**b**), and ×150,000 (**c**). Reprinted and adapted with permission from *Proteomics* [4]

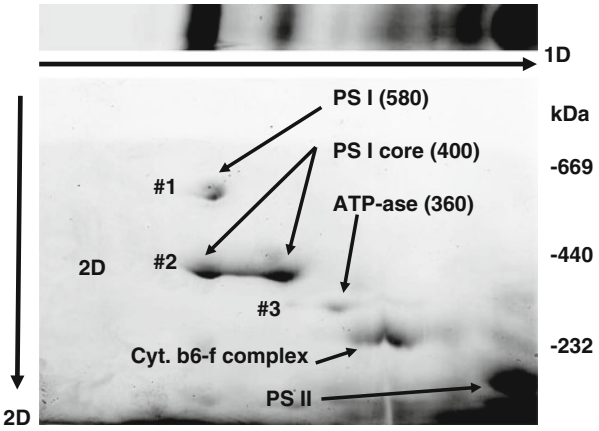


Fig. 9.8 Combination of CN-PAGE 1D and BN-PAGE 2D for the resolution of protein complexes from the thylakoid membranes with the same molecular weight or pI. Reprinted and adapted with permission from *Cellular Molecular Life Science* [5]

assembly of proteins into protein complexes (Fig. 9.7B). Heteroprotein complexes (Eif3S10 is part of an 800 kDa complex and proteasome beta is part of a 700 kDa complex) or homoprotein complexes (valosin-containing protein is a 540 kDa homohexamer and ATP citrate lyase is a 480 kDa homotetramer) were identified using BN-PAGE and MS (Fig. 9.7A). In a different experiment, BN-PAGE and electron microscopy were used to investigate assembly of the vitelline envelope (VE) proteins (Fig. 9.7B). Here, VE gamma and VE beta monomers were purified to homogeneity, as demonstrated by SDS-PAGE R and then their ability to polymerize was investigated by BN-PAGE to visualize the different monomeric and polymeric forms of these subunits. Polymerization of these proteins into VE alpha and VE beta homopolymers and of VE alpha–VE beta heteropolymers, with distinct polymerization patterns was demonstrated (Fig. 9.7B). Finally, electron microscopy confirmed the difference in polymerization patterns between VE beta, VE gamma, and a mixture of both.

A similar technique to BN-PAGE which has been used for the investigation and purification of cytochrome bc/bf complexes is the Colorless Native PAGE (CN-PAGE). This method separates protein complexes according to their internal charge, but independent of their molecular mass. Due to the nature of CN-PAGE, mostly acidic proteins (pI < 7) are run on a CN-PAGE [124, 127]. Figure 9.8 shows thylakoid membranes from mesophyll chloroplasts which were separated by CN-PAGE in 1D followed by BN-PAGE in 2D and led to the identification of protein complexes such as PSI, ATP-ase, Cyt b6f complex, or PSII. Since complexes #2 and #3 have the same mass (about 400 kDa), they can only be distinguished by CN-PAGE and not by BN-PAGE (Fig. 9.8). In the same sense, complexes #1 and #3 which co-migrate in CN-PAGE can only be separated in the second dimension due to their different Mw. The position of PSI, ATP-ase, Cyt b6f complex, and PSII is indicated. Note that there are no Mw markers for CN-PAGE because this method does not separate protein complexes according to their Mw. The direction of migration in CN-PAGE and BN-PAGE is indicated.

Native ESI-MS. MS has become the method of choice for the analysis of proteins, protein interactions, and protein modifications. To this end, it can be paired with other methods (AP-MS, TAP-MS, Co-IP-MS, cross-linking) to investigate PPIs. Nevertheless, it is also possible to perform MS on intact protein complexes without prior application of any other biochemical methods as it was observed that noncovalent receptor ligand complexes remained intact in the gas phase by using ESI-MS [128]. This type of MS was termed native MS or MS of intact assemblies. Native MS can reveal subunit stoichiometry, heterogeneity, and dynamical changes. Using this approach, large molecular assemblies such as the proteasome [129], the eukaryotic translation factor Elf3 [130], RNA polymerase III [131], intact viruses [132], and membrane-bound complexes have been analyzed [133, 134]. Further, quantitative MS-based proteomic strategies that are reliable and specific have also shed light on the study of protein complexes. These quantitative approaches are mainly based on stable isotope-labeling methods such as stable isotope labeling by amino acids in cell culture (SILAC), isobaric tags for relative and absolute quantification (iTRAQ) and isotope-coded affinity tags (ICAT) [135–140]. With these techniques, protein complex composition and subunit abundance can be determined by comparison of signal intensities or peak areas of isotope-labeled peptide pairs. They also make the spotting of false positive interactions because the latter have abundance ratio around 1 whereas true interaction partners have abundance ratios higher than 1 [141]. Using SILAC, interaction partners of GLUT4 [142], the integrin-linked kinase interactome [143], and the protein phosphatase 1 [144] were detected. In another study by Blagoev et al., SILAC was applied in conjunction with LC-MS/MS to identify components of the epidermal growth factor (EGF)-dependent signaling complex and these results were confirmed by confocal microscopy and FRET [145]. Another study successfully explored changes in PPIs and phosphorylation patterns in yeast and *D. melanogaster* cells by means of iTRAQ [146].

PTMs. For PTM studies by MS, there are several important considerations to take into account. Due to the characteristically low abundance of PTMs (lower fmol range) and their transient and labile nature, enrichment is an inevitable step. This can be done by a variety of techniques. One technique that has been employed for enrichment of modified phosphopeptides and O-glycosylated residues is chemical derivatization, which is based on β-elimination and subsequent Michael addition of the phosphate group and sugar residues [147–150]. Besides chemical derivatization, antibody-based methods (e.g., anti-phosphotyrosine antibodies) can also be utilized. Two other techniques mostly employed for the enrichment of phosphopeptides are TiO_2 affinity chromatography and immobilized metal affinity chromatography (IMAC). Sometimes prior to enrichment, fractionation of the samples by other chromatographic means (size exclusion chromatography or ion exchange chromatography) is undertaken in order to reduce sample complexity [46, 151, 152] (Fig. 9.2). For the digestion of enriched peptides, enzymes as varied as trypsin, LysC, GluC, AspN, or ArgC are used. The modified peptides can be selected in MS mode and fragmented to identify and localize the site of modification. In some other cases, neutral losses of the attached group (e.g., phosphate group) also serve as a

tool for the identification of the type of modification and its location. Modifications which present very close mass shifts can only be distinguished by using highly sensitive and accurate MS instrumentation such as FT-ICR-MS or LTQ Ion Trap Orbitrap. An example is the mass shifts generated by a tri-methylation (42.047 Da) or an acetylation (42.010 Da) [153]. An overview of basic information on all the PTMs discussed in this review is shown in Table 9.1.

Figures 9.9, 9.10, 9.11, 9.12, and 9.13 depict MS-based identification of PTMs discussed in this book chapter. Figure 9.9 treats one of the most frequently occurring protein alteration events in proteomics, namely methionine oxidation. This modification results in a 16 Da mass shift in an MS spectrum [154]. In the tryptic peptide mixture analyzed in this experiment, a peak with *m*/*z* of 769.42 (2+) corresponding to a peptide with the sequence VKEGMNIVEAMER was identified (Fig. 9.9a). The MS/MS fragmentation of this peak produced some b and y ions that allowed us to confirm the identity of the peptide in question, with both methionines oxidized to methionine-sulphoxide (Fig. 9.9b–d). A common process that follows oxidation of methionine to methionine-sulphoxide (a 16 Da gain) is a neutral loss of methane-sulfenic acid (a 64 Da loss). Therefore, we looked for both b and y ions and precursor ions, but also for the neutral loss of y and precursor ions (y-64). Identification in the MS/MS spectrum of the precursor ion with *m*/*z* of 769.44 (2+) and its neutral loss (64 Da) ion with *m*/*z* of 737.49 (2+) as well as identification of y12 and y12-64, y11 and y11-64, y10 and y10-64, y9 and y9-64, y8 and y8-64, y7 and y7-64, y4 and y4-64, and y12 (2+) and y12 (2+)-64 pairs, suggest that both the N-terminal and C-terminal methionines are oxidized.

In Fig. 9.10, the investigation of the occupancy of an NXS/NXT N-glycosylation site is shown. In this case, the peptide mixture resulted from a trypsin–AspN double digestion after initial treatment of the original protein (IgG heavy chain) with PNGase F. PNGase F specifically removes the N-linked glycosyl groups from the NXS/NXT glycosylation sites. If the NXS/NXT site is glycosylated, the PNGase F treatment would convert the asparagine from the NST site into aspartate; with a net gain of 1 Da. From this peptide mixture, a doubly charged peak with *m*/*z* of 595.20 corresponding to a peptide with the sequence EEQYNSTYR was detected. Its MS/MS fragmentation confirmed the amino acid sequence of this peptide. Therefore, the peptide EEQYNSTYR has an asparagine at the NST site and the NST site is not occupied by any oligosaccharides. The total ion chromatogram (TIC), the extracted ion chromatogram (XIC) for the precursor ion with the *m*/*z* of 595.20, the mass spectrum (TOF MS) showing the doubly charged precursor ion, as well as the MS/MS of the precursor ion that led to identification of the peptide with the sequence EEQYNSTYR is shown in Fig. 9.10. A different example of investigating the occupancy of the glycosylation sites is shown in Fig. 9.11. In this case, MS-based determination of the occupancy of the NSS site is not possible, unless the (1) NSS is glycosylated; (2) oligosaccharide group is removed by a PNGase F digestion and Asn from the sequence NSS is converted to Asp (NSS to DSS); (3) NSS to DSS conversion creates a new AspN cleavage site; and (4) trypsin–AspN double digestion produces the peptide containing the DSS, which is identified by MS. As observed in Fig. 9.11, the data indicate that the NSS site is glycosylated; the PNGase

Table 9.1 Summary of the principal post-translational modifications of proteins

PTMs	Electrophilic donor	Amino acids	Enzymes (+)	Enzymes (–)	Mass shift Δm (Da)	Neutral loss (Da)
Acetylation	Acetyl-CoA	Amino group (Lys)	Acetyltransferases	Deacetylases	42.010	–
Glycosylation	Glycans	Asn, Arg, Gly, Ser, Thr, Tyr	Glycosyltransferase	Glycosidase	–	–
Oxidation	Electrophilic reactive species	S (Cys, Met)	–	–	16	64
Methylation	*S*-adenosylmethionine	N (Lys, Gln, His, Arg) O (Asp, Glu) S (Cys) C (Arg, δC; Gln, αC)	Methyltransferase	–	14.0156	31.0422 73.064
Phosphorylation	ATP	O (Ser, Thr, Tyr, Asp) N (His)	Kinase	Phosphatase	79.99	98
Ubiquitination	ATP, ubiquitin	Ɛ-amino group (Lys) Cys	E1, E2, E3 ubiquitin ligases	DUBs	114.043	–

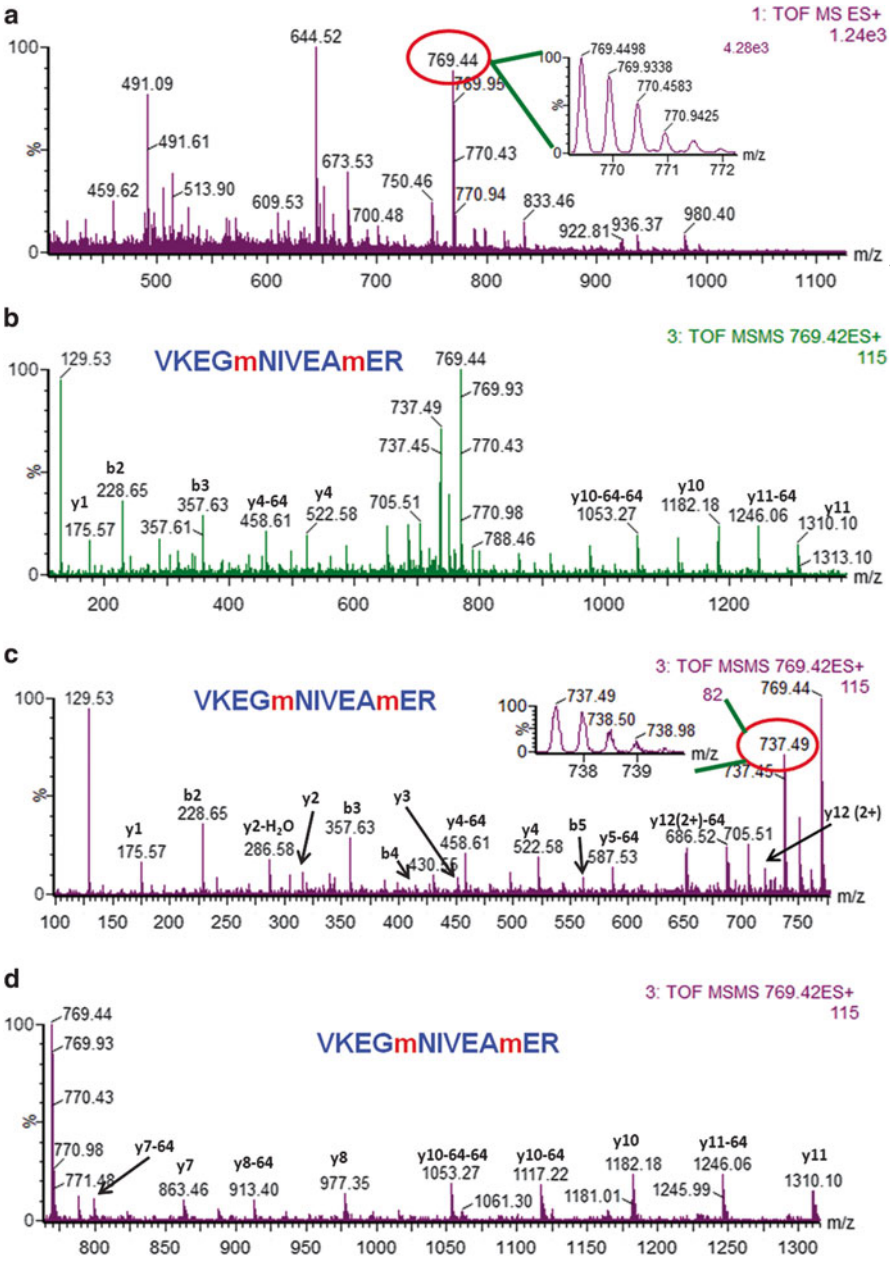


Fig. 9.9 Methionine oxidation of the peptide VKEGMNIVEAMER with *m/z* of 769.42 (2+) (**a**) as confirmed by the b and y ions series produced by the fragmentation of this peptide (**b**–**d**). A closer look at the b and y ions series from (**b**) and zoomed MS/MS spectra (**c**, **d**) showed that the peptide is oxidized at both methionines. Reprinted and adapted with permission from the *Australian Journal of Chemistry CSIRO Publishing* http://www.publish.csiro.au/?paper=CH13144 [35]

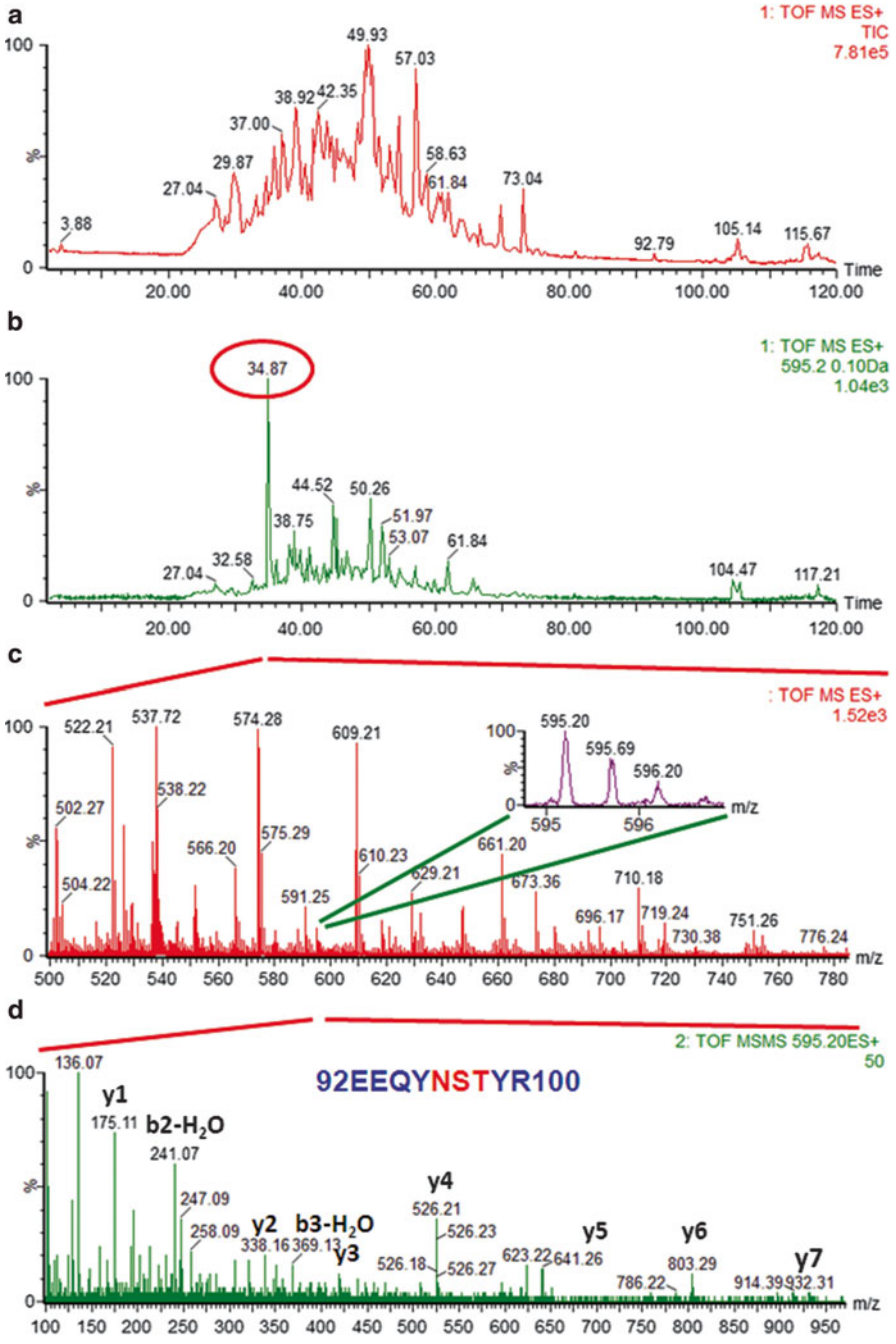


Fig. 9.10 Analysis of the occupancy of the NXS/T site of the IgG heavy chain by PNGase F treatment followed by trypsin–AspN double digestion. The peptide mixture was then run on LC-MS/MS. Total ion chromatogram (TIC) (**a**), extracted ion chromatogram (XIC) (**b**) for the doubly charged peptide EEQYNSTYR with *m/z* of 595.20 (2+) as well as its MS (**c**) and MS/MS spectrum (**d**) are shown. Identification of the peptide EEQYNSTYR, but not EEQYDSTYR suggests that the NST site is not occupied by glycans. Reprinted and adapted with permission from the *Australian Journal of Chemistry CSIRO Publishing* http://www.publish.csiro.au/?paper=CH13144 [35]

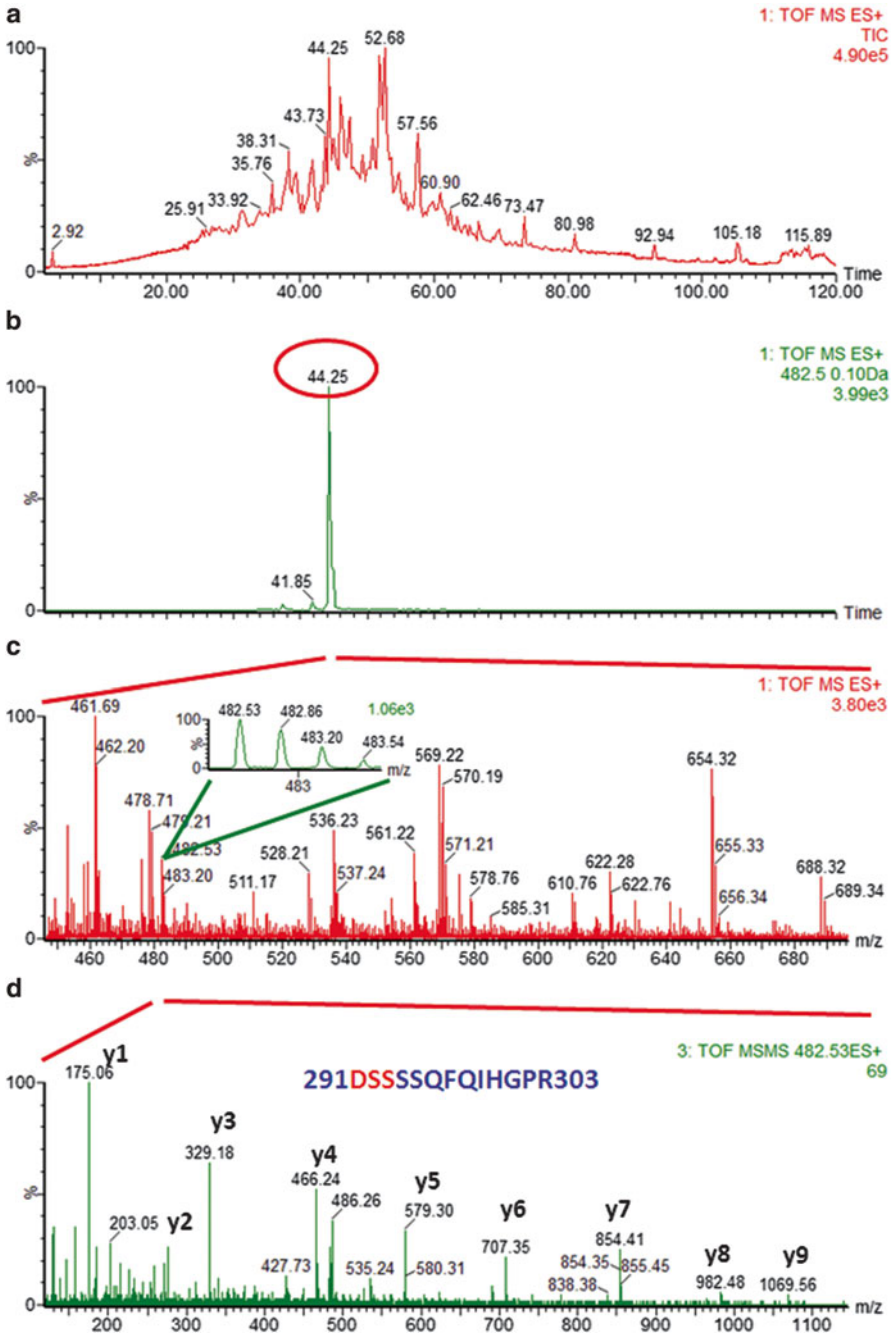
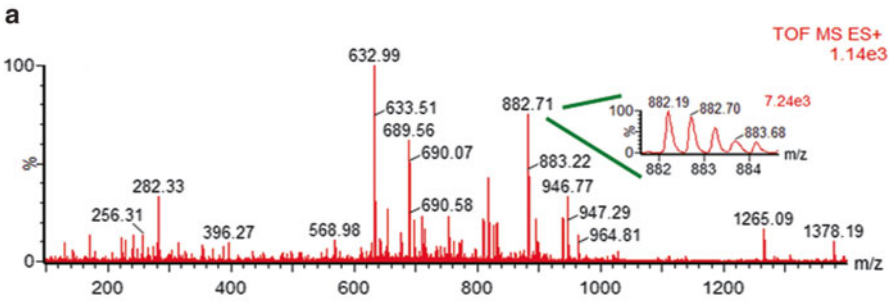
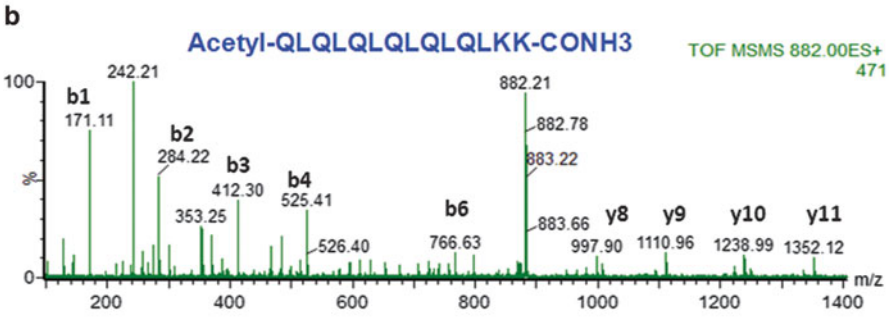


Fig. 9.11 Analysis of the occupancy of the NXS/T site of a peptide mixture treated by PNGase F followed by trypsin–AspN double digestion. The peptide mixture was then run on LC-MS/MS. TIC (**a**), XIC (**b**) for the triply charged peptide DSSSSQFQIHGPR with *m/z* of 482.53 (3+) as well as its MS (**c**) and MS/MS spectrum (**d**) are shown. Identification of the peptide DSSSSQFQIHGPR demonstrates that the asparagine of the original peptide NSSSSQFQIHGPR was occupied by an oligosaccharide. Reprinted and adapted with permission from the *Australian Journal of Chemistry CSIRO Publishing* http://www.publish.csiro.au/?paper=CH13144 [35]





c

MH^{+1} = 1721.08

b			y	
---	1	Q	14	---
242.1499	2	L	13	1593.0210
370.2085	3	Q	12	1479.9370
483.2926	4	L	11	1351.8784
611.3511	5	Q	10	1238.7943
724.4352	6	L	9	1110.7357
852.4938	7	Q	8	997.6517
965.5778	8	L	7	869.5931
1093.6364	9	Q	6	756.5090
1206.7205	10	L	5	628.4505
1334.7791	11	Q	4	515.3664
1447.8631	12	L	3	387.3078
1575.9581	13	K	2	274.2238
---	14	K	1	146.1288

Amidated

MH^{+1} = 1763.09

b			y	
Acetyl				
171.0764	1	Q	14	1721.0796
284.1605	2	L	13	1593.0210
412.2191	3	Q	12	1479.9370
525.3031	4	L	11	1351.8784
653.3617	5	Q	10	1238.7943
766.4458	6	L	9	1110.7357
894.5043	7	Q	8	997.6517
1007.5884	8	L	7	869.5931
1135.6470	9	Q	6	756.5090
1248.7310	10	L	5	628.4505
1376.7896	11	Q	4	515.3664
1489.8737	12	L	3	387.3078
1617.9687	13	K	2	274.2238
---	14	K	1	146.1288

Amidated

Fig. 9.12 ESI-MS and ESI-MS/MS of a synthetic peptide showing acetylation of the peptide at the N-terminus and amidation at the C-terminus. Mass spectrum of the peptide QLQLQLQLQLQLKK with *m/z* of 882.19 (+2) (**a**) as well as its MS/MS spectrum enabled the location of the amidation at the C-terminal amino acid and acetylation at the N-terminal amino acid. Data analysis was performed using de novo sequencing, and then confirmed by using the online software ProteinProspector v 5.10.9. Reprinted and adapted with permission from the *Australian Journal of Chemistry CSIRO Publishing* http://www.publish.csiro.au/?paper=CH13144 [35]

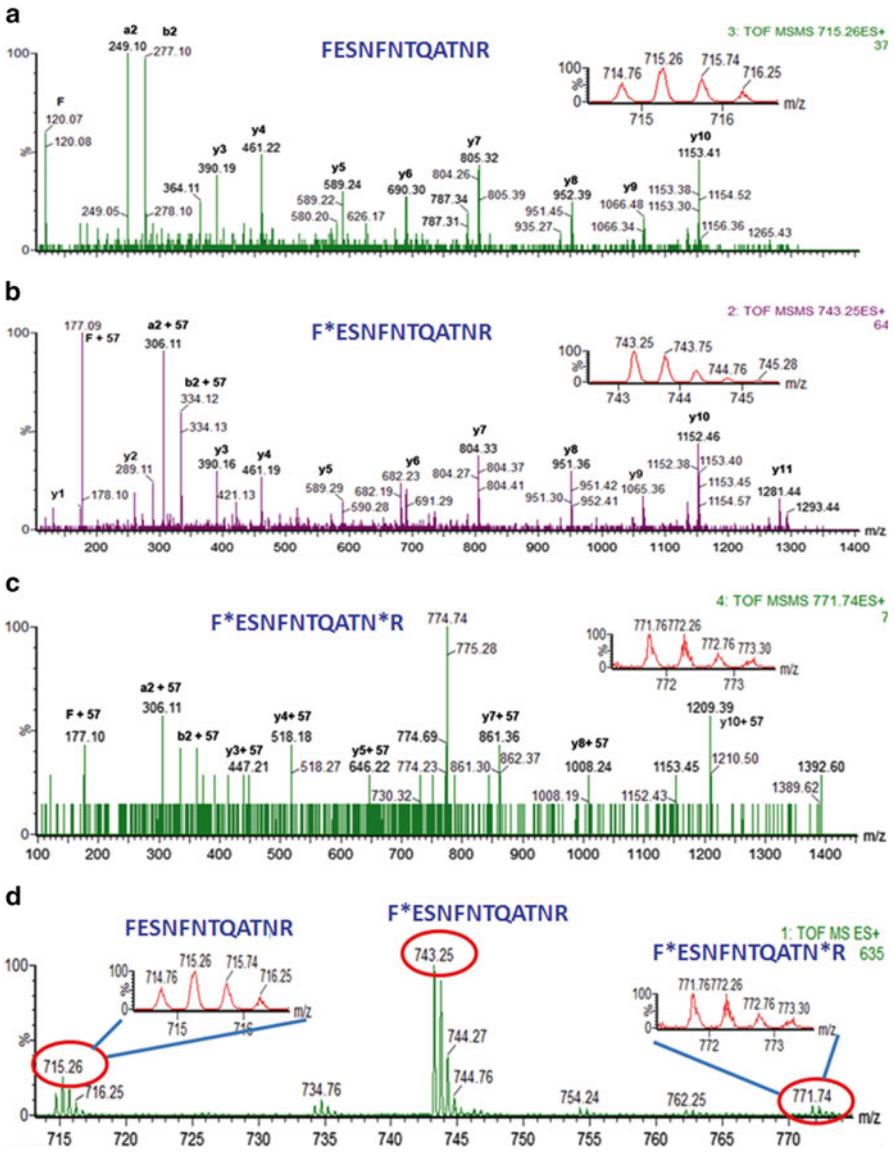


Fig. 9.13 Analysis of the alkylation of a non-cysteinyl peptide from a gel band treated with DTT followed by IAA prior to trypsin digestion. (**a**) MS/MS spectrum of the doubly charged peptide FESNFNTQATNR with *m/z* of 714.76 showing its b and y ion series. (**b**) The same peptide sequence was identified as a doubly charged precursor ion with *m/z* of 743.25 (mono-alkylated at the Phe) and (**c**) as a doubly charged precursor ion with *m/z* of 771.76 (di-alkylated at the Phe and Asn). (**d**) MS/MS spectrum of the three forms of the above peptide (un-, mono-, di-alkylated) and their ratios. Reprinted and adapted with permission from the *Biophysical & Biochemical Research Communications* [43]

with *m*/*z* of 771.76, which corresponds to the original peptide alkylated at two sites (Fig. 9.13d). MS/MS analysis of this precursor (Fig. 9.13c) and comparison with the previous MS/MS spectra (Fig. 9.13a, b) led to the identification of the modified Phe and of the a2 and b2 ions, as shown in Fig. 9.13b, but also to the identification of the alkylated y3–y5, y7–y8, and y10 ions, suggesting that the original peptide was alkylated at the N-terminal Phe residue, but also at one of the three C-terminal residues. A closer inspection of the intensities of the precursor ions for the unmodified (*m*/*z* of 714.76), monoalkylated (*m*/*z* of 743.25) and dialkylated (*m*/*z* of 771.76) peptides also indicates two concerning issues: (1) the unmodified peptide is in very low amounts and many times is not identified by database searches, and (2) alkylation of amino acids other than cysteine may lead to issues in quantitative analysis.

9.3 Conclusions

The relevance of studying and understanding protein PTMs and PPIs in order to elucidate biological processes and develop strategies for identifying and treating pathological disorders has been demonstrated throughout this chapter. Historically used methods for the study of protein interactions such as Y2H, Co-IP, AP-based methods, FRET are yielding to novel proteomics-based techniques such as MS.

MS is certainly the best available method to identify and quantify PPIs and PTMs in complex samples. It enables identification of proteins in the femto- to picomole range and is a powerful tool for the determination of proteins involved in complex formation. However, there are some limitations associated with the technique itself such as co-eluting peptides or overlapping *m*/*z*, a very broad dynamic range. Some of these issues are already being addressed through the development of mass spectrometers with greater resolution and sensitivity [159–161].

Though many protein modifications have been discovered and are well described, it is still believed that many are missed in proteomic experiments. A reason for this could be that experimental data are always matched against a database of known proteins. In this case, unassigned or unmatched experimental data are automatically considered "junk." Perhaps manual verification of these unmatched data could result in the discovery of novel protein modifications. Specific to the analysis of PTMs by MS is the single focus on one modification per experiment, without consideration of the cross-talk among different modifications. This is probably due to the lack of available techniques for the enrichment of two or more distinct modified proteins or peptides concomitantly. Multiply modified peptides do not occur as often in the proteome and are more labile than singly modified moieties. In addition, detection of manifold modified peptides is hindered by their multiple neutral losses and their little backbone fragmentation upon CID.

The study of noncovalent complexes by MS requires extensive optimization in order to get stable protein–protein or protein–ligand complexes in solution and in

the gas phase [162]. For example, high sample concentrations are required for sufficient MS signals since the total MS signal is split among isotopic patterns and differently charged protein complexes [163]. However, these high protein sample concentrations do not necessarily reflect physiological conditions and so false-positive interactions could be determined. Further, 2D-gels are time- and labor intensive, limited in their dynamic range, not highly reproducible between gels and biased due to the under-representation of membrane and "extreme" proteins (very high/low mass or pI) [164]. Another concern with gel electrophoresis is the low-throughput for protein identification, the difficulty to automate and the large amounts of reagents and proteins needed. It has been shown that with BN-PAGE, weak binding proteins might not be detected and PPIs are possibly incomplete [146]. Similarly, CN-PAGE is mostly applicable to protein complexes with high binding affinities and has even a lower resolution than BN-PAGE [165, 166]. Capillary electrophoresis (CE) has been suggested as an alternative due to its rapid, automatable, highly sensitive, quantitative, real-time high-resolution separation suitable for the investigation of biospecific interactions [167]. Unfortunately, the adsorption of basic proteins onto the capillary wall during CE causes loss of material and leads to a poor efficiency in separation [168].

A point that should be briefly discussed here for both PPIs and PTMs is the validation of their studies. The convention is that non-MS studies should be preferentially validated by MS experiments and vice versa to remove any doubts on the genuineness of the results of the studies. So far for the validation of MS experiments, WB experiments monitoring PTMs such as phosphorylation or glycosylation and interactions between two or more proteins are sufficient as confirmatory steps for the publication of the data.

Acknowledgments We would like to thank Ms. Laura Mulderig, Scott Nichols, and their colleagues (Waters Corporation) for their generous support in setting up the Proteomics Center at Clarkson University. C.C.D. thanks Drs. Thomas A. Neubert (New York University), Belinda Willard (Cleveland Clinic), and Gregory Wolber and David Mclaughin (Eastman Kodak Company) for donation of a TofSpec2E MALDI-MS (each). This work was supported in part by the Keep a Breast Foundation (KEABF-375-35054), the Redcay Foundation (SUNY Plattsburgh), the Alexander von Humboldt Foundation, SciFund Challenge, private donations (Ms. Mary Stewart Joyce, Mr. Kenneth Sandler, Bob Mattloff), and by the U.S. Army research office (DURIP grant #W911NF-11-1-0304).

References

1. Schmutz J et al (2004) Quality assessment of the human genome sequence. Nature 429(6990):365–368
2. Stein L (2001) Genome annotation: from sequence to biology. Nat Rev Genet 2(7):493–503
3. Eisenberg D et al (2000) Protein function in the post-genomic era. Nature 405(6788):823–826
4. Ngounou Wetie AG et al (2013) Investigation of stable and transient protein-protein interactions: past, present and future. Proteomics 13(3–4):538–557
5. Ngounou Wetie AG et al (2014) Protein-protein interactions: switch from classical methods to proteomics and bioinformatics-based approaches. Cell Mol Life Sci 71(2):205–228

6. Darie C (2013) Investigation of protein-protein interactions by Blue Native-PAGE & mass spectrometry. Mod Chem Appl 1(3):e111
7. Darie CC, Shetty V, Spellman DS, Zhang G, Xu C, Cardasis HL, Blais S, Fenyo D, Neubert TA (2008) Blue Native PAGE and mass spectrometry analysis of the ephrin stimulation-dependent protein-protein interactions in NG108-EphB2 cells. Applications of mass spectrometry in life safety, NATO science for peace and security series. Springer, Düsseldorf, Germany
8. Darie CC, Litscher ES, Wassarman PM (2008) Structure, processing, and polymerization of rainbow trout egg vitelline envelope proteins. Applications of mass spectrometry in life safety, NATO science for peace and security series. Springer, Düsseldorf, Germany
9. Darie CC (2013) Mass spectrometry and its application in life sciences. Aust J Chem 66:1–2
10. Darie CC et al (2006) Studies of the Ndh complex and photosystem II from mesophyll and bundle sheath chloroplasts of the C4-type plant Zea mays. J Plant Physiol 163(8):800–808
11. Darie CC et al (2011) Identifying transient protein-protein interactions in EphB2 signaling by blue native PAGE and mass spectrometry. Proteomics 11(23):4514–4528
12. Byrum S et al (2012) Analysis of stable and transient protein-protein interactions. Methods Mol Biol 833:143–152
13. Sadrzadeh SM, Bozorgmehr J (2004) Haptoglobin phenotypes in health and disorders. Am J Clin Pathol 121(Suppl):S97–S104
14. Bashor CJ et al (2010) Rewiring cells: synthetic biology as a tool to interrogate the organizational principles of living systems. Annu Rev Biophys 39:515–537
15. McNally FJ, Vale RD (1993) Identification of katanin, an ATPase that severs and disassembles stable microtubules. Cell 75(3):419–429
16. Dutcher SK (2001) The tubulin fraternity: alpha to eta. Curr Opin Cell Biol 13(1):49–54
17. Hemmerich P, Schmiedeberg L, Diekmann S (2011) Dynamic as well as stable protein interactions contribute to genome function and maintenance. Chromosome Res 19(1):131–151
18. Sanderson CM (2008) A new way to explore the world of extracellular protein interactions. Genome Res 18(4):517–520
19. DeLano WL (2002) Unraveling hot spots in binding interfaces: progress and challenges. Curr Opin Struct Biol 12(1):14–20
20. Krylov D, Mikhailenko I, Vinson C (1994) A thermodynamic scale for leucine zipper stability and dimerization specificity: e and g interhelical interactions. EMBO J 13(12):2849–2861
21. Vogelstein B, Lane D, Levine AJ (2000) Surfing the p53 network. Nature 408(6810):307–310
22. Ideker T, Sharan R (2008) Protein networks in disease. Genome Res 18(4):644–652
23. Schuster-Bockler B, Bateman A (2008) Protein interactions in human genetic diseases. Genome Biol 9(1):R9
24. Wong JM, Ionescu D, Ingles CJ (2003) Interaction between BRCA2 and replication protein A is compromised by a cancer-predisposing mutation in BRCA2. Oncogene 22(1):28–33
25. Jonsson PF, Bates PA (2006) Global topological features of cancer proteins in the human interactome. Bioinformatics 22(18):2291–2297
26. Soler-Lopez M et al (2011) Interactome mapping suggests new mechanistic details underlying Alzheimer's disease. Genome Res 21(3):364–376
27. Rikova K et al (2007) Global survey of phosphotyrosine signaling identifies oncogenic kinases in lung cancer. Cell 131(6):1190–1203
28. Manning G et al (2002) Evolution of protein kinase signaling from yeast to man. Trends Biochem Sci 27(10):514–520
29. Deshaies RJ, Joazeiro CA (2009) RING domain E3 ubiquitin ligases. Annu Rev Biochem 78:399–434
30. Ohtsubo K, Marth JD (2006) Glycosylation in cellular mechanisms of health and disease. Cell 126(5):855–867
31. Olsen JV et al (2006) Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. Cell 127(3):635–648
32. Bischoff R, Schluter H (2012) Amino acids: chemistry, functionality and selected non-enzymatic post-translational modifications. J Proteomics 75(8):2275–2296

33. Darie C (2013) Mass spectrometry and proteomics: principle, workflow, challenges and perspectives. Mod Chem Appl 1(2):e105
34. Darie C (2013) Post-translational modification (PTM) proteomics: challenges and perspectives. Mod Chem Appl 1:e114
35. Ngounou Wetie AG et al (2013) Identification of post-translational modifications by mass spectrometry. Aust J Chem 66:734–748
36. Ngounou Wetie AG et al (2013) Automated mass spectrometry-based functional assay for the routine analysis of the secretome. J Lab Autom 18(1):19–29
37. Ngounou Wetie AG et al (2013) Mass spectrometry for the detection of potential psychiatric biomarkers. J Mol Psychiatry 1:8
38. Sokolowska I et al (2012) Disulfide proteomics for identification of extracellular or secreted proteins. Electrophoresis 33(16):2527–2536
39. Sokolowska I et al (2013) Mass spectrometry investigation of glycosylation on the NXS/T sites in recombinant glycoproteins. Biochim Biophys Acta 1834(8):1474–1483
40. Sokolowska I et al (2013) Applications of mass spectrometry in proteomics. Aust J Chem 66:721–733
41. Sokolowska I et al (2013) Characterization of tumor differentiation factor (TDF) and its receptor (TDF-R). Cell Mol Life Sci 70(16):2835–2848
42. Sokolowska I et al (2011) Mass spectrometry for proteomics-based investigation of oxidative stress and heat shock proteins. In: Andreescu S, Hepel M (eds) Oxidative stress: diagnostics, prevention, and therapy. American Chemical Society, Washington, DC
43. Woods AG, Sokolowska I, Darie CC (2012) Identification of consistent alkylation of cysteine-less peptides in a proteomics experiment. Biochem Biophys Res Commun 419(2):305–308
44. Woods AG et al (2012) Potential biomarkers in psychiatry: focus on the cholesterol system. J Cell Mol Med 16(6):1184–1195
45. Woods AG et al (2011) Blue native page and mass spectrometry as an approach for the investigation of stable and transient protein-protein interactions. In: Andreescu S, Hepel M (eds) Oxidative stress: diagnostics, prevention, and therapy. American Chemical Society, Washington, DC
46. Olsen JV et al (2010) Quantitative phosphoproteomics reveals widespread full phosphorylation site occupancy during mitosis. Sci Signal 3(104):ra3
47. Zhang G et al (2006) Quantitative phosphotyrosine proteomics of EphB2 signaling by stable isotope labeling with amino acids in cell culture (SILAC). J Proteome Res 5(3):581–588
48. Rinschen MM et al (2010) Quantitative phosphoproteomic analysis reveals vasopressin V2-receptor-dependent signaling pathways in renal collecting duct cells. Proc Natl Acad Sci U S A 107(8):3882–3887
49. Cantin GT et al (2006) Quantitative phosphoproteomic analysis of the tumor necrosis factor pathway. J Proteome Res 5(1):127–134
50. Hanahan D, Weinberg RA (2000) The hallmarks of cancer. Cell 100(1):57–70
51. Pan C et al (2009) Comparative proteomic phenotyping of cell lines and primary cells to assess preservation of cell type-specific functions. Mol Cell Proteomics 8(3):443–450
52. Lee J et al (2006) Tumor stem cells derived from glioblastomas cultured in bFGF and EGF more closely mirror the phenotype and genotype of primary tumors than do serum-cultured cell lines. Cancer Cell 9(5):391–403
53. Malik R et al (2010) From proteome lists to biological impact—tools and strategies for the analysis of large MS data sets. Proteomics 10(6):1270–1283
54. Finkel T (2011) Signal transduction by reactive oxygen species. J Cell Biol 194(1):7–15
55. Hill BG et al (2010) What part of NO don't you understand? Some answers to the cardinal questions in nitric oxide biology. J Biol Chem 285(26):19699–19704
56. Higdon A et al (2012) Cell signalling by reactive lipid species: new concepts and molecular mechanisms. Biochem J 442(3):453–464
57. Pacher P, Beckman JS, Liaudet L (2007) Nitric oxide and peroxynitrite in health and disease. Physiol Rev 87(1):315–424
58. 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 1473(1):4–8

59. Kornfeld R, Kornfeld S (1985) Assembly of asparagine-linked oligosaccharides. Annu Rev Biochem 54:631–664
60. Stanley P (2011) Golgi glycosylation. Cold Spring Harb Perspect Biol 3(4):1–13
61. Halim A et al (2011) Site-specific characterization of threonine, serine, and tyrosine glycosylations of amyloid precursor protein/amyloid beta-peptides in human cerebrospinal fluid. Proc Natl Acad Sci U S A 108(29):11848–11853
62. Steentoft C et al (2011) Mining the O-glycoproteome using zinc-finger nuclease-glycoengineered SimpleCell lines. Nat Methods 8(11):977–982
63. Spiro RG (1969) Characterization and quantitative determination of the hydroxylysine-linked carbohydrate units of several collagens. J Biol Chem 244(4):602–612
64. Spiro RG (2002) Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds. Glycobiology 12(4):43R–56R
65. Reis CA et al (2010) Alterations in glycosylation as biomarkers for cancer detection. J Clin Pathol 63(4):322–329
66. Aggarwal S (2010) What's fueling the biotech engine—2009-2010. Nat Biotechnol 28(11):1165–1171
67. Hunt JV, Dean RT, Wolff SP (1988) Hydroxyl radical production and autoxidative glycosylation. Glucose autoxidation as the cause of protein damage in the experimental glycation model of diabetes mellitus and ageing. Biochem J 256(1):205–212
68. Smith MA et al (1994) Advanced Maillard reaction end products, free radicals, and protein oxidation in Alzheimer's disease. Ann N Y Acad Sci 738:447–454
69. Elsholz AK et al (2012) Global impact of protein arginine phosphorylation on the physiology of Bacillus subtilis. Proc Natl Acad Sci U S A 109(19):7451–7456
70. Laub MT, Goulian M (2007) Specificity in two-component signal transduction pathways. Annu Rev Genet 41:121–145
71. Barford D (1996) Molecular mechanisms of the protein serine/threonine phosphatases. Trends Biochem Sci 21(11):407–412
72. Zhang ZY (2002) Protein tyrosine phosphatases: structure and function, substrate specificity, and inhibitor development. Annu Rev Pharmacol Toxicol 42:209–234
73. Johnson LN, Barford D (1993) The effects of phosphorylation on the structure and function of proteins. Annu Rev Biophys Biomol Struct 22:199–232
74. Hunter T (2007) The age of crosstalk: phosphorylation, ubiquitination, and beyond. Mol Cell 28(5):730–738
75. Braconi Quintaje S, Orchard S (2008) The annotation of both human and mouse kinomes in UniProtKB/Swiss-Prot: one small step in manual annotation, one giant leap for full comprehension of genomes. Mol Cell Proteomics 7(8):1409–1419
76. Jackson MD, Denu JM (2001) Molecular reactions of protein phosphatases—insights from structure and chemistry. Chem Rev 101(8):2313–2340
77. Guan KL, Dixon JE (1991) Evidence for protein-tyrosine-phosphatase catalysis proceeding via a cysteine-phosphate intermediate. J Biol Chem 266(26):17026–17030
78. Paik WK, Paik DC, Kim S (2007) Historical review: the field of protein methylation. Trends Biochem Sci 32(3):146–152
79. Ishikawa Y, Melville DB (1970) The enzymatic alpha-N-methylation of histidine. J Biol Chem 245(22):5967–5973
80. Bedford MT, Clarke SG (2009) Protein arginine methylation in mammals: who, what, and why. Mol Cell 33(1):1–13
81. Wang C et al (2005) A general fluorescence-based coupled assay for S-adenosylmethionine-dependent methyltransferases. Biochem Biophys Res Commun 331(1):351–356
82. Erce MA et al (2012) The methylproteome and the intracellular methylation network. Proteomics 12(4–5):564–586
83. Darwanto A et al (2010) A modified "cross-talk" between histone H2B Lys-120 ubiquitination and H3 Lys-79 methylation. J Biol Chem 285(28):21868–21876
84. Haglund K, Dikic I (2005) Ubiquitylation and cell signaling. EMBO J 24(19):3353–3359
85. Pickart CM, Eddins MJ (2004) Ubiquitin: structures, functions, mechanisms. Biochim Biophys Acta 1695(1–3):55–72

86. Nijman SM et al (2005) A genomic and functional inventory of deubiquitinating enzymes. Cell 123(5):773–786
87. Bhoj VG, Chen ZJ (2009) Ubiquitylation in innate and adaptive immunity. Nature 458(7237):430–437
88. Manning G et al (2002) The protein kinase complement of the human genome. Science 298(5600):1912–1934
89. Alonso A et al (2004) Protein tyrosine phosphatases in the human genome. Cell 117(6):699–711
90. Shi Y (2009) Serine/threonine phosphatases: mechanism through structure. Cell 139(3):468–484
91. Danielsen JM et al (2011) Mass spectrometric analysis of lysine ubiquitylation reveals promiscuity at site level. Mol Cell Proteomics 10(3):M110.003590
92. Jin L et al (2012) Ubiquitin-dependent regulation of COPII coat size and function. Nature 482(7386):495–500
93. Pickart CM (2001) Mechanisms underlying ubiquitination. Annu Rev Biochem 70:503–533
94. Motegi A et al (2008) Polyubiquitination of proliferating cell nuclear antigen by HLTF and SHPRH prevents genomic instability from stalled replication forks. Proc Natl Acad Sci U S A 105(34):12411–12416
95. Zhao S et al (2010) Regulation of cellular metabolism by protein lysine acetylation. Science 327(5968):1000–1004
96. Wellen KE et al (2009) ATP-citrate lyase links cellular metabolism to histone acetylation. Science 324(5930):1076–1080
97. Ganesan A et al (2009) Epigenetic therapy: histone acetylation, DNA methylation and anticancer drug discovery. Curr Cancer Drug Targets 9(8):963–981
98. Li G, Reinberg D (2011) Chromatin higher-order structures and gene regulation. Curr Opin Genet Dev 21(2):175–186
99. Khan SN, Khan AU (2010) Role of histone acetylation in cell physiology and diseases: an update. Clin Chim Acta 411(19–20):1401–1411
100. Sato N et al (2003) Frequent hypomethylation of multiple genes overexpressed in pancreatic ductal adenocarcinoma. Cancer Res 63(14):4158–4166
101. Balasubramanyam K et al (2004) Curcumin, a novel p300/CREB-binding protein-specific inhibitor of acetyltransferase, represses the acetylation of histone/nonhistone proteins and histone acetyltransferase-dependent chromatin transcription. J Biol Chem 279(49):51163–51171
102. Aggarwal S et al (2006) Curcumin (diferuloylmethane) down-regulates expression of cell proliferation and antiapoptotic and metastatic gene products through suppression of IkappaBalpha kinase and Akt activation. Mol Pharmacol 69(1):195–206
103. Choudhary C et al (2009) Lysine acetylation targets protein complexes and co-regulates major cellular functions. Science 325(5942):834–840
104. Plazas-Mayorca MD et al (2010) Quantitative proteomics reveals direct and indirect alterations in the histone code following methyltransferase knockdown. Mol Biosyst 6(9):1719–1729
105. Sokolowska I et al (2012) Proteomic analysis of plasma membranes isolated from undifferentiated and differentiated HepaRG cells. Proteome Sci 10(1):47
106. Sokolowska I et al (2013) The potential of biomarkers in psychiatry: focus on proteomics. J Neural Transm 1–10)
107. Woods AG et al (2013) Mass spectrometry as a tool for studying autism spectrum disorder. J Mol Psychiatry 1:6
108. Garcia BA (2010) What does the future hold for top down mass spectrometry? J Am Soc Mass Spectrom 21(2):193–202
109. Cannon J et al (2010) High-throughput middle-down analysis using an orbitrap. J Proteome Res 9(8):3886–3890
110. Picotti P, Aebersold R (2012) Selected reaction monitoring-based proteomics: workflows, potential, pitfalls and future directions. Nat Methods 9(6):555–566
111. Lehmann WD et al (2007) Neutral loss-based phosphopeptide recognition: a collection of caveats. J Proteome Res 6(7):2866–2873

112. Syka JE et al (2004) Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. Proc Natl Acad Sci U S A 101(26):9528–9533
113. Kelleher NL et al (1999) Localization of labile posttranslational modifications by electron capture dissociation: the case of gamma-carboxyglutamic acid. Anal Chem 71(19):4250–4253
114. Good DM et al (2007) Performance characteristics of electron transfer dissociation mass spectrometry. Mol Cell Proteomics 6(11):1942–1951
115. Choudhary C, Mann M (2010) Decoding signalling networks by mass spectrometry-based proteomics. Nat Rev Mol Cell Biol 11(6):427–439
116. Fields S, Song O (1989) A novel genetic system to detect protein-protein interactions. Nature 340(6230):245–246
117. Monti M et al (2005) Interaction proteomics. Biosci Rep 25(1–2):45–56
118. Miernyk JA, Thelen JJ (2008) Biochemical approaches for discovering protein-protein interactions. Plant J 53(4):597–609
119. Berkowitz SA (2006) Role of analytical ultracentrifugation in assessing the aggregation of protein biopharmaceuticals. AAPS J 8(3):E590–E605
120. Miyashita T (2004) Confocal microscopy for intracellular co-localization of proteins. Methods Mol Biol 261:399–410
121. Berggard T, Linse S, James P (2007) Methods for the detection and analysis of protein-protein interactions. Proteomics 7(16):2833–2842
122. Jensen ON (2006) Interpreting the protein language using proteomics. Nat Rev Mol Cell Biol 7(6):391–403
123. Schagger H, von Jagow G (1991) Blue native electrophoresis for isolation of membrane protein complexes in enzymatically active form. Anal Biochem 199(2):223–231
124. Schagger H, Cramer WA, von Jagow G (1994) Analysis of molecular masses and oligomeric states of protein complexes by blue native electrophoresis and isolation of membrane protein complexes by two-dimensional native electrophoresis. Anal Biochem 217(2):220–230
125. Darie CC et al (2005) Isolation and structural characterization of the Ndh complex from mesophyll and bundle sheath chloroplasts of Zea mays. FEBS J 272(11):2705–2716
126. Spellman DS et al (2008) Stable isotopic labeling by amino acids in cultured primary neurons: application to brain-derived neurotrophic factor-dependent phosphotyrosine-associated signaling. Mol Cell Proteomics 7(6):1067–1076
127. Schagger H (1995) Native electrophoresis for isolation of mitochondrial oxidative phosphorylation protein complexes. Methods Enzymol 260:190–202
128. Ganem J, Li YT, Henion J (1991) Detection of noncovalent receptor-ligand complexes by mass spectrometry. J Am Chem Soc 113:6294–6296
129. Sakata E et al (2011) The catalytic activity of Ubp6 enhances maturation of the proteasomal regulatory particle. Mol Cell 42(5):637–649
130. Zhou M et al (2008) Mass spectrometry reveals modularity and a complete subunit interaction map of the eukaryotic translation factor eIF3. Proc Natl Acad Sci U S A 105(47): 18139–18144
131. Lorenzen K et al (2007) Structural biology of RNA polymerase III: mass spectrometry elucidates subcomplex architecture. Structure 15(10):1237–1245
132. Uetrecht C et al (2008) High-resolution mass spectrometry of viral assemblies: molecular composition and stability of dimorphic hepatitis B virus capsids. Proc Natl Acad Sci U S A 105(27):9216–9220
133. Barrera NP et al (2008) Micelles protect membrane complexes from solution to vacuum. Science 321(5886):243–246
134. Barrera NP et al (2009) Mass spectrometry of membrane transporters reveals subunit stoichiometry and interactions. Nat Methods 6(8):585–587
135. Xie F et al (2011) Liquid chromatography-mass spectrometry-based quantitative proteomics. J Biol Chem 286(29):25443–25449
136. Filiou MD et al (2012) To label or not to label: applications of quantitative proteomics in neuroscience research. Proteomics 12(4–5):736–747

137. Mann M (2006) Functional and quantitative proteomics using SILAC. Nat Rev Mol Cell Biol 7(12):952–958
138. Gygi SP et al (1999) Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. Nat Biotechnol 17(10):994–999
139. Ross PL et al (2004) Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. Mol Cell Proteomics 3(12):1154–1169
140. Ong SE et al (2002) Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. Mol Cell Proteomics 1(5):376–386
141. Oeljeklaus S, Meyer HE, Warscheid B (2009) New dimensions in the study of protein complexes using quantitative mass spectrometry. FEBS Lett 583(11):1674–1683
142. Foster LJ et al (2006) Insulin-dependent interactions of proteins with GLUT4 revealed through stable isotope labeling by amino acids in cell culture (SILAC). J Proteome Res 5(1):64–75
143. Dobreva I et al (2008) Mapping the integrin-linked kinase interactome using SILAC. J Proteome Res 7(4):1740–1749
144. Trinkle-Mulcahy L et al (2006) Repo-Man recruits PP1 gamma to chromatin and is essential for cell viability. J Cell Biol 172(5):679–692
145. Blagoev B et al (2003) A proteomics strategy to elucidate functional protein-protein interactions applied to EGF signaling. Nat Biotechnol 21(3):315–318
146. Pflieger D et al (2008) Quantitative proteomic analysis of protein complexes: concurrent identification of interactors and their state of phosphorylation. Mol Cell Proteomics 7(2):326–346
147. Knight ZA et al (2003) Phosphospecific proteolysis for mapping sites of protein phosphorylation. Nat Biotechnol 21(9):1047–1054
148. Oda Y, Nagasu T, Chait BT (2001) Enrichment analysis of phosphorylated proteins as a tool for probing the phosphoproteome. Nat Biotechnol 19(4):379–382
149. Wells L et al (2002) Mapping sites of O-GlcNAc modification using affinity tags for serine and threonine post-translational modifications. Mol Cell Proteomics 1(10):791–804
150. Li W et al (2003) Susceptibility of the hydroxyl groups in serine and threonine to beta-elimination/Michael addition under commonly used moderately high-temperature conditions. Anal Biochem 323(1):94–102
151. Dephoure N et al (2008) A quantitative atlas of mitotic phosphorylation. Proc Natl Acad Sci U S A 105(31):10762–10767
152. Tan CS et al (2009) Positive selection of tyrosine loss in metazoan evolution. Science 325(5948):1686–1688
153. Zhang K et al (2004) Differentiation between peptides containing acetylated or tri-methylated lysines by mass spectrometry: an application for determining lysine 9 acetylation and methylation of histone H3. Proteomics 4(1):1–10
154. Toda T et al (2010) Proteomic approaches to oxidative protein modifications implicated in the mechanism of aging. Geriatr Gerontol Int 10(Suppl 1):S25–S31
155. Lapko VN, Smith DL, Smith JB (2000) Identification of an artifact in the mass spectrometry of proteins derivatized with iodoacetamide. J Mass Spectrom 35(4):572–575
156. Lundell N, Schreitmuller T (1999) Sample preparation for peptide mapping—a pharmaceutical quality-control perspective. Anal Biochem 266(1):31–47
157. Windsor WT et al (1993) Disulfide bond assignments and secondary structure analysis of human and murine interleukin 10. Biochemistry 32(34):8807–8815
158. Yang Z, Attygalle AB (2007) LC/MS characterization of undesired products formed during iodoacetamide derivatization of sulfhydryl groups of peptides. J Mass Spectrom 42(2): 233–243
159. Michalski A, Cox J, Mann M (2011) More than 100,000 detectable peptide species elute in single shotgun proteomics runs but the majority is inaccessible to data-dependent LC-MS/MS. J Proteome Res 10(4):1785–1793
160. Cox J, Mann M (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat Biotechnol 26(12):1367–1372

161. Wu R et al (2011) A large-scale method to measure absolute protein phosphorylation stoichiometries. Nat Methods 8(8):677–683
162. Kool J et al (2011) Studying protein-protein affinity and immobilized ligand-protein affinity interactions using MS-based methods. Anal Bioanal Chem 401(4):1109–1125
163. Bunt J et al (2012) OTX2 directly activates cell cycle genes and inhibits differentiation in medulloblastoma cells. Int J Cancer 131(2):E21–E32
164. Brewis IA, Brennan P (2010) Proteomics technologies for the global identification and quantification of proteins. Adv Protein Chem Struct Biol 80:1–44
165. Wang F, Pan YC (1991) Structural analyses of proteins electroblotted from native polyacrylamide gels onto polyvinyldiene difluoride membranes. A method for determining the stoichiometry of protein-protein interaction in solution. Anal Biochem 198(2):285–291
166. Wang F et al (1993) Electroblotting proteolytic products from native gel for direct N-terminal sequence analysis: an approach for studying protein-protein interaction. Electrophoresis 14(9):847–851
167. Karger BL, Chu YH, Foret F (1995) Capillary electrophoresis of proteins and nucleic acids. Annu Rev Biophys Biomol Struct 24:579–610
168. Vergnon AL, Chu YH (1999) Electrophoretic methods for studying protein-protein interactions. Methods 19(2):270–277

Chapter 10
Applications for Mass Spectrometry in the Study of Ion Channel Structure and Function

Damien S.K. Samways

Abstract Ion channels are intrinsic membrane proteins that form gated ion-permeable pores across biological membranes. Depending on the type, ion channels exhibit sensitivities to a diverse range of stimuli including changes in membrane potential, binding by diffusible ligands, changes in temperature and direct mechanical force. The purpose of these proteins is to facilitate the passive diffusion of ions down their respective electrochemical gradients into and out of the cell, and between intracellular compartments. In doing so, ion channels can affect transmembrane potentials and regulate the intracellular homeostasis of the important second messenger, Ca^{2+}. The ion channels of the plasma membrane are of particular clinical interest due to their regulation of cell excitability and cytosolic Ca^{2+} levels, and the fact that they are most amenable to manipulation by exogenously applied drugs and toxins. A critical step in improving the pharmacopeia of chemicals available that influence the activity of ion channels is understanding how their three-dimensional structure imparts function. Here, progress has been slow relative to that for soluble protein structures in large part due to the limitations of applying conventional structure determination methods, such as X-ray crystallography, nuclear magnetic resonance imaging, and mass spectrometry, to membrane proteins. Although still an underutilized technique in the assessment of membrane protein structure, recent advances have pushed mass spectrometry to the fore as an important complementary approach to studying the structure and function of ion channels. In addition to revealing the subtle conformational changes in ion channel structure that accompany gating and permeation, mass spectrometry is already being used effectively for identifying tissue-specific posttranslational modifications and mRNA splice variants. Furthermore, the use of mass spectrometry for high-throughput proteomics analysis, which has proven so successful for soluble proteins, is already providing valuable insight into the functional interactions of ion channels within the context of the

D.S.K. Samways, Ph.D. (✉)
Department of Biology, Clarkson University, 8 Clarkson Avenue, Potsdam, NY 13699, USA
e-mail: samwaysds@slu.edu

A.G. Woods and C.C. Darie (eds.), *Advancements of Mass Spectrometry in Biomedical Research*, Advances in Experimental Medicine and Biology 806,
DOI 10.1007/978-3-319-06068-2_10, © Springer International Publishing Switzerland 2014

macromolecular-signaling complexes that they inhabit in vivo. In this chapter, the potential for mass spectrometry as a complementary approach to the study of ion channel structure and function will be reviewed with examples of its application.

10.1 Introduction

In stark contrast to the considerable volume of structural data obtained for water-soluble proteins over the last few decades, relatively few three-dimensional structures have been resolved for membrane proteins (<250, compared with >50,000 for soluble proteins) [1–3]. This is in large part due to the problems of enriching and solubilizing membrane proteins to the necessary degree required for high resolution structural determination by X-ray crystallography. This dearth of structural information poses a considerable deficit in our understanding when one considers that membrane proteins, principally ion channels and G protein-coupled receptors, make up more than 60 % of the current targets for therapeutic drugs [2]. Indeed, the promising expectations of de novo structure-based drug design [4, 5] have in large part been curbed by the continued lack of accurate structural data for these therapeutic targets.

Nevertheless, with advances in expression, purification, enrichment, and solubilization, recent years have seen some major successes, with crystal structures now resolved for a number of ion channels [6–13] and G protein-coupled receptors ([14–17] and for review, see [18]). These static structures have already proven valuable as sources for new testable hypotheses with regard to ion channel structure and function. Additionally, for many of these membrane receptors a substantial body of data exists from functional studies conducted on manufactured mutant proteins, and investigators now have the opportunity to review this work in the context of a three-dimensional model [19]. However, use of X-ray crystallography alone to determine membrane protein structures is not without its limitations. A general limitation applicable to soluble as well as membrane proteins is that static crystal structures reveal limited information about the conformational flexibility of a protein, which in reality are highly dynamic and able to reside in a range of conformations. In the case of ion channels, which are the subject of this chapter, one hopes to obtain at least two crystal structures representing the closed and open channel state, as then the conformational change linking the two can at least be crudely intuited [7]. However, this has by no means proven an easy undertaking. As it is, and like many proteins, ion channels can inhabit multiple inactive and active states [3].

Another difficulty is that many of the techniques designed to enhance the expression, solubilization, and crystallization of membrane proteins carry the risk of adversely influencing the structure of the membrane protein of interest. For example, expressing vertebrate proteins in invertebrate cells, yeast or bacteria, can lead to incorrect folding, incorrect mRNA splicing and/or inappropriate posttranslational modification of the ion channel [1]. Solubilizing a membrane protein necessitates exchanging the native lipid for a detergent or nonnative mimetic

lipid environment, either of which may coerce the transmembrane (TM) helices into nonnative conformational arrangements, as has been recently suggested to be the case for the open channel crystal structure for the ATP-gated zebrafish P2X4 receptor [7, 20]. This limitation is particularly problematic for ion channels, where knowing the orientation and position of the transmembrane-spanning helices in the closed and open channel state is critical to understanding stimulus-dependent gating, ion selection and conductance. In many cases, difficulties in resolving structures for mammalian membrane proteins have caused investigators to instead work with nonmammalian variants, in which there might not be full conservation of function. From the medical perspective, and particularly with regard to drug design, the ideal is to obtain accurate structures for the mammalian, if not the human, variant of the protein.

10.1.1 Mass Spectrometry as a Complementary Approach to Protein Structure

Mass spectroscopy has been used extensively to study secondary, tertiary, and quaternary structure in soluble proteins, complementing structural data obtained using other methods including X-ray crystallography, nuclear magnetic resonance (NMR) spectroscopy, electron microscopy (EM), and atomic force microscopy (AFM). X-ray crystallography achieves the highest structural resolution, albeit with the caveats described above. NMR can also provide high resolution structural determination, but is restricted to smaller proteins (<100 kDa) as increasing molecular weight slows the tumbling of the soluble protein, and this technique shares some of the problems with X-ray crystallography with respect to appropriate protein solubility. EM and AFM can provide information on stoichiometry and protein–protein interactions, but the resolution is low.

One area in which mass spectrometry has been particularly informative is in the assessment of solvent accessibility of protein domains, i.e., determining which parts of a protein form the water-contacting surfaces. For example, the use of hydrogen/deuterium exchange in combination with mass spectroscopy has helped provide a more accurate portrayal of proteins as dynamic molecules capable of occupying a multitude of conformational states [21] (See Section 10.3.1). Thus, where a static crystal structures might imply that a certain stretch of peptide backbone is embedded within the folded protein, removed from the aqueous environment, hydrogen/deuterium exchange has often revealed that structural changes occur that bring that backbone, however briefly, to the global surface of the protein [22]. As such, the depiction of a protein as a rigid and static structure implied by X-ray crystallography is superseded by the more accurate depiction of a protein as a highly dynamic and flexible molecule that exists in an equilibrium of numerous discreet conformational states. The likelihood of the protein inhabiting one state over another is a function of probability, the balance of which is amenable to tuning by numerous

factors. Another advantage of using mass spectroscopy is that relatively small, as low as nanomolar, concentrations of protein are required. This neatly overcomes a considerable limitation of both X-ray crystallography and NMR, allowing for analysis of mammalian protein isolated directly from native tissue, where correct folding and appropriate posttranslational modification is most assured. Indeed, mass spectrometry has been at the forefront of the high-throughput proteomics boom [23–28].

As has recently been reviewed elsewhere [29, 30], advances in techniques utilizing mass spectroscopy have allowed this tool to be increasingly applied to the problem of resolving the structure, conformational flexibility, and protein–protein interactions of membrane proteins. This chapter will focus on the potential for using mass spectrometry to unravel the relationship between structure and function for a particularly important class of membrane proteins, the ion channels. It should be noted that there is an extensive literature on the use of mass spectrometry to probe interactions between ligands and ion channels, but this will not be covered in this chapter (for review, see [31]).

10.2 Protein Channels as Passive Transporters of Water-Soluble Molecules

Protein channels are intrinsic membrane proteins that form pores across biological membranes permeable to water and/or ions. These proteins have a vital role in the regulated passive transport of aqueous solutes across the otherwise hydrophobic barrier of biological membranes. Although found on most biological membranes, including those of intracellular organelles such as the endoplasmic reticulum and the mitochondria, by far the most important families from a pharmacological perspective are those located on the plasma membrane surface of cells. With the exception of the aquaporins, whose main function is to passively transport water molecules across the plasma membrane [32], the primary function of most of these protein channels is to selectively transport ions across the plasma membrane, and these are thus termed collectively as “ion channels” (although it should be noted that many of these ion channels are also permeable to water, whether conducted alone or as part of the hydrated ion complex).

To fully appreciate the importance of ion channels, we must first understand two facts about a typical animal cell. First, by virtue of active transport proteins that use energy to pump ions against their electrochemical gradients, all animal cells possess a separation of charge across their plasma membrane, and possess a potential difference that is usually in the range of −40 to −90 mV [33]. Second, and also by virtue of active transport proteins, most animal cells maintain a very low resting cytosolic concentration of the vital second messenger Ca^{2+} relative to the extracellular environment and relative to the Ca^{2+}-containing stores of the endoplasmic reticulum within the cell. Due to the existence of these electrochemical gradients, the stimulus-dependent opening of ion channels positioned on the plasma membrane can affect

two main properties of a cell, depending on the identity of the permeable ions: (1) its membrane potential and (2) its cytosolic Ca^{2+} concentration. On the endoplasmic reticulum, second messenger activation of ion channels such as the ryanodine receptors (RyRs) and inositol 1,4,5-triphosphate ($InsP_3$) receptors, releases Ca^{2+} and provides an alternative means to elevate cytosolic-free Ca^{2+} concentrations [34]. Changes in plasma membrane potential have a role in many important cell functions, including action potential propagation in all excitable cells, excitation–contraction coupling in myocytes, and also cell proliferation, differentiation, and survival [33, 35]. Equally, elevations in Ca^{2+}, depending on the duration, intensity, and whether the effect is sustained or oscillatory, can have a myriad of effects on cell function ranging from secretion, muscle contraction, apoptosis, proliferation, and differentiation [34].

The ion channels are subdivided into numerous families that respond to different types of stimuli. A very important class from a pharmacological perspective is the ligand-gated ion channels, which commonly mediate changes in membrane potential and/or cytosolic Ca^{2+} concentration when opened upon binding by neurotransmitters released from neurons. Relevant examples include the nicotinic acetylcholine receptor (nAChR) and γ-aminobutyric acid-sensitive $GABA_A$ receptor of the Cys-loop family of ligand-gated ion channels [36], the glutamate-gated ionotropic receptors (iGluRs) [37], and the ATP-gated P2X receptors (P2XRs) [38]. By containing binding sites for diffusible extracellular ligands, these ion channels are, of course, prime targets for exogenously administered drugs. In addition to these are ion channels that sense and respond to changes in the plasma membrane potential, the voltage-gated ion channels. Two subfamilies, the voltage-gated Na^+ and K^+ channels, are vital for the initiation and propagation of action potentials in excitable cells of nerve and muscle [33]. Another group, the voltage-gated Ca^{2+} channels, is vital for regulating muscle contraction, the cardiac action potential, and stimulation of transmitter release upon action potential arrival at a terminal synapse [39]. As their names imply, the voltage-gated ion channels exhibit strong selectivity for conducting specific ions, whether Na^+, K^+, Ca^{2+} of Cl^-. The majority of ligand-gated ion channels exhibit less fidelity in this regard, but nevertheless tend to show preference between conducting either cations or anions [33].

Ever since the pioneering work of Hodgkin and Huxley on the squid giant axon [33, 40], there has been considerable interest in how ion channels function. Specifically, there are three questions we seek to answer with respect to relating the structure of these proteins to their functional properties. (1) By what means does the applied stimulus, whether a ligand, a change in membrane potential, a change in temperature, mechanical force etc., alter the conformation of the ion channel from a closed, ion-impermeant state to an open ion-permeant state? In short, how does the ion channel "gate"? (2) Which parts of the ion channel protein are responsible for forming the transmembrane ion-conducting pathway? (3) How do amino acids within the ion-conducting pathway confer selectivity, such that some ions are allowed to pass with preference over others?

Due to the previously highlighted difficulty in obtaining crystal structures for membrane proteins, inferences about the structure and function of ion channels

has relied heavily on indirect functional approaches employing mutagenesis of recombinant protein in combination with patch clamp electrophysiology and immunochemistry. These studies are adequate in identifying key amino acid residues within the primary sequence critical to ligand-binding, gating, and ion permeation, but provide limited insight as to the actual location of these residues in space beyond simply inferring that they are extracellular, intracellular, or somewhere within the lipid or water soluble parts of the transmembrane domains. However, the availability of crystal structures derived from a number of important ion channel families has allowed for a much greater degree of resolution in terms of identifying the key functional units within these membrane proteins [33]. Nevertheless, there is considerable room for further refinement of the available models, particularly considering the precise movements and conformational shifts that occur during signal transduction and gating of the transmembrane conduction pathway; movements and conformational changes that are often not captured with adequate temporal resolution by one or two static crystal structures. Focusing on recent experiments utilizing mass spectroscopy, a brief overview of the potential for applying this method in the study of ion channel structure and function will now be given.

10.3 Ion Channel Structure: Topology, Stoichiometry, and Solvent Accessibility

A key early goal in the characterization of an ion channel is determining its membrane topology. For a plasma membrane protein, a given amino acid side chain might be embedded in the protein as part of a protein–protein interface, protruding into either of the extracellular or cytosolic aqueous compartments, or facing the hydrophobic lipid environment of the bilayer itself. Knowing the whereabouts of water versus lipid-contacting residues not only assists in discriminating between the extracellular, membrane-spanning, and cytosolic parts of the protein, but also assists in the determination of the ion-permeable pore. Identifying the pore region is the first step towards unraveling the processes by which the channel opens and subsequently selectively conducts ions across the membrane (gating and permeability).

One straightforward application for mass spectrometry in ascertaining ion channel structure is the determination of likely membrane-spanning protein segments. Initially, predictions of membrane topology were largely made based on model-dependent primary sequence hydropathy analysis [41, 42], but these models could be inaccurate as demonstrated by their incorrect prediction that the C-terminal domain of the glutamate-gated NMDA receptor was extracellular rather than intracellular [43]. Such models must therefore be verified by additional experiments, such as immunolabeling [44] and glycosylation-linked scanning [45]. It has been shown that membrane-spanning domains tend to be protected from proteolytic cleavage by water-soluble enzymes. Thus, it is possible to expose membrane-embedded proteins to limited proteolysis and then separate the soluble and lipid-associated fragments [46]. The lipid-associated fragments can then be solubilized

by detergent and analyzed using mass spectrometry. Leite et al. [46] employed liquid chromatography in combination with electrospray mass spectrometry to test the validity of hydropathy plot-based topology predictions for the glycine-sensitive Cl– channel, GlyR, another member of the Cys-loop family and cousin of the nAChR and $GABA_A$ channels [36]. A limitation of this approach is the presence of the detergent, which can produce aggregates that reduce the signal-to-noise ratio of the mass spectrometry results [47] (Fig. 10.1a). Recently, Robinson et al. developed a significant enhancement of this approach in which whole heteromeric membrane protein complexes could be liberated from their detergent micelle by gas-phase activation [29, 47–49] (Fig. 10.1a). After first ionizing the detergent-embedded protein complexes via electrospray, the resulting charged particles are ejected from the detergent micelles into the gas phase through thermal activation, which was produced by delivery of the charged complexes into the collision cell of the mass spectrometer at very high acceleration voltages. Importantly, the initial presence of the detergent micelle was found to have a protective influence on the protein complex, preventing this thermal activation from destabilizing the non-covalent interactions holding the complex together. Thus, maintenance of both subunit stoichiometry and protein–lipid interactions was achieved [47, 49, 50]. The same group recently utilized this technique in combination with ion mobility mass spectrometry to validate the subunit stoichiometry of the tetrameric K^+ channel, KirBac3.1 [51] (Fig. 10.1b). The technique has also been adapted for high-throughput screening of detergents, with the goal of optimizing membrane protein solubilization in a manner that retains structural stability of the protein complex, which would provide useful information for X-ray crystallography as well as mass spectrometry [50].

10.3.1 Locating the Channel Pore

Previously, identifying water-accessible parts of an ion channel, including the ion-permeable transmembrane pore, has relied on the use of the Cysteine Scanning Accessibility Method (SCAM). This involves generating mutant ion channels in which single amino acids have been substituted for cysteine and then assessing the accessibility of these cysteines to covalent modification by water-soluble thiol-reactive agents. Assessment of accessibility is usually achieved by using patch clamp electrophysiology to investigate whether thiol-reactive agents modify the current profile of the ion channel [52], although recent studies have employed immunochemistry to monitor thiol-reactive biotin labeling [53, 54]. The advantage of this method is that it involves assessment of the functional ion channel in its native, membrane-associated environment. However, this method does have several limitations. First and foremost, inferences can only be drawn from positive functional data in which thiol modification of an introduced cysteine is shown to irreversibly alter the ion channel's transmembrane current profile (i.e., changing the current amplitude, activation or inactivation kinetics). If exposure of the mutant ion channel to thiol-reactive agents fails to have an effect on current profile, one cannot simply

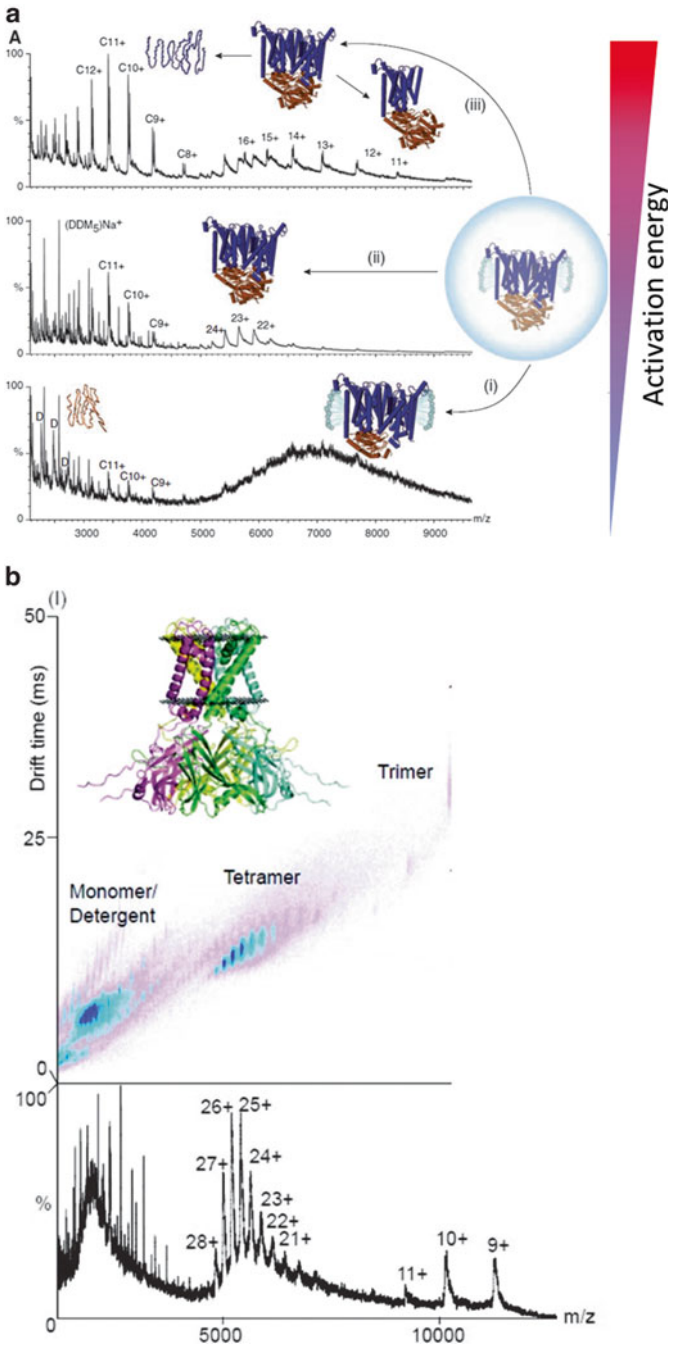


Fig. 10.1 A technique for liberating oligomeric protein complexes from their solubilizing detergent while minimally disturbing non-covalent protein–protein and protein–lipid interactions. (**a**) Schematic representation showing the progressive release of an intact membrane complex into the

conclude that the introduced cysteine is inaccessible; it may well be accessible to, and modified by, the water soluble thiol-reactive agent, but it might simply be that the modification has no effect on channel function. Second, the possibility of applied thiol-reactive agents modifying endogenous-free cysteine side chains has to be adequately controlled for, and if necessary such endogenous cysteines removed [55], although this carries the risk of fundamentally altering the structure of the ion channel. Third, the substitution of side chains for cysteine in the protein structure also carries the risk of changing the structure in some way, but possibly not in a way that immediately noticeable as an effect on channel function [55]. With respect to the first problem concerning the interpretation of negative SCAM data, mass spectrometry can of course be used to overcome this caveat because it allows for the simple determination of whether exposure of the ion channel to a thiol-reactive agent increases its mass or not, regardless of whether there is a functional consequence. Mass spectrometry has not been conducted in combination with SCAM, but has been successfully used in combination with a similar solvent-accessibility technique in which tyrosine residues are assessed for accessibility to tetranitromethane in glycine receptors [56, 57].

A similar conceptual paradigm to SCAM and related solvent-accessibility methods, mass spectrometry has been successfully employed in combination with hydrogen/deuterium exchange [21] and hydroxyl radical footprinting [58] to identify the parts of soluble proteins exposed to the aqueous environment. These approaches are now being successfully applied to membrane-embedded proteins. Hydrogen/deuterium exchange exploits the constant hydroxyl-catalyzed substitution that occurs between hydrogens in the backbone amides and side chains of amino acids with those of surrounding water molecules. Briefly, D_20 is included in the aqueous environment surrounding the protein such that the hydrogens of peptide stretches exposed to water are progressively substituted for deuterium. This can be executed with the protein embedded in a normal cell bilayer. The protein is then isolated and subjected to proteolysis, with the fragments subsequently analyzed by mass spectrometry. The rate of exchange for a given fragment can then be ascertained and thus the degree to which it has been exposed to water in the native protein determined. In addition to giving a straightforward assessment of which stretches of polypeptide form the water-soluble surface of the protein, this technique can also be used to monitor very subtle changes in the protein's conformational state, where burial of a polypeptide segment into protein or lipid correlates with a reduced rate

Fig. 10.1 (continued) gas phase from a detergent micelle within an electrospray droplet. At low thermal activation energies, detergent-protein aggregates are retained and can be observed above *m/z* 5,000 (1). Increasing the thermal activation energy releases the oligomeric protein complex from the detergent micelle, preserving stoichiometry and lipid interactions (2). Further increasing activation energy will begin to cause dissociation of the protein complex into lower order oligomers (3). Reproduced from [47] with permission. (**b**) Ion mobility contour plot (*top*) and mass spectra (*bottom*) of KirBac3.1 after release from detergent micelle by collisional activation (structure shown *inset*, (PDB 1XL6), with membrane depicted by the *black lines*). Adapted from [51] with permission

of deuterium exchange [59–61]. In addition, hydrogen/deuterium exchange has been used to show that individual hydrogen-bonds within membrane proteins tend to be fairly weak, and that structural stability is mostly conferred by numerous weak hydrogen bonding interactions rather than a small number of strong interactions [62]. Hydroxyl radical-mediated protein footprinting takes advantage of the ability for hydroxyl radicals to covalently modify backbone or amino acid side chains contributing to the water-accessible surface of a protein [30, 58]. The generation of hydroxyl residues can be achieved using either X-ray bombardment [63] or the use of chemical agents [64]. Similar to hydrogen/deuterium exchange, this technique is useful for mapping the aqueous interfaces of a protein generally, in addition to allowing the monitoring of changes in protein conformation.

Even in cases where X-ray crystal structures are available, these techniques have the potential to enhance our understanding of a number of important aspects of ion channel function. For all ion channels, protein movements accompanying gating could potentially be tracked according to the changes in aqueous accessibility of peptide segments. Also for all ion channels, the unpacking of protein in the gating region and removal of the barrier to ion permeation could be tracked, again according to changes in the aqueous accessibility of pore-forming side chains. The accessibility of peptide segments to hydroxyl radical modification or deuterium exchange could also be assessed in the absence and presence of ligands, in order to identify the putative ligand-binding sites for agonists and allosteric modulators of ion channels.

The feasibility of using hydroxyl radical-mediated protein footprinting for investigating channel gating has recently been demonstrated by Gupta et al. [63]. Using the K^+ channel, KirBac3.1., as their model, the investigators used focused synchrotron X-ray pulses to produce the hydroxyl radicals for modifying water accessible amino acids in both the closed and open channel states. They then probed the extent and rates of amino acid modification using high resolution mass spectrometry and from this determine key parts of the protein structure likely to move during gating (Fig. 10.2). Interpreting the data in the context of the available closed channel crystal structure, it was possible to predict the likely conformation of the open channel state, in the absence of an open state crystal structure. Thus, mass spectrometry used in combination with X-ray crystallographic data will likely prove invaluable in teasing out the nuanced and graded changes in protein conformation that underlie signal transduction and gating in ion channels, providing a considerable improvement to the lower temporal resolution of protein movement obtained from generating small numbers of static crystal structures in different gating states (states that may not even be accurate representations of the native channel, as has been discussed). Such mass spectrometry-assisted experiments also offer several advantages over SCAM for assessing the water-soluble surfaces of an ion channel in its various conformations. First, mutation of the ion channel is not required, thus circumnavigating the problem of introducing mutations that might alter the very structure and conformational stability of the open and closed channel states that the investigator is trying to resolve. Second, hydroxyl radical-accessible side chains are identified directly using mass spectroscopy, thus not relying on indirect assessments using functional assays in which negative data can be uninformative.

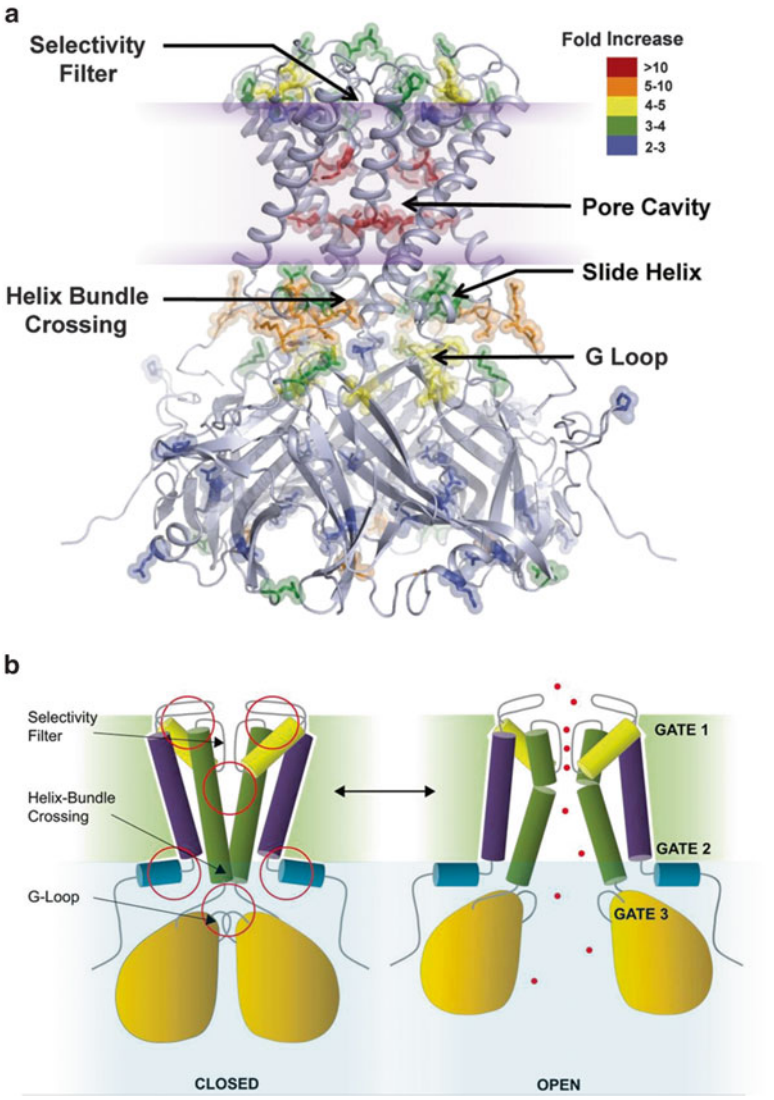


Fig. 10.2 Changes in solvent accessibility as determined using hydroxyl radical footprinting. (**a**) Closed state ribbon model of KirBac3.1 structure (PDB 1XL4) with hydroxyl-modified residues shown in *stick format* and the plasma membrane bilayer outlined in *purple*. Comparative rates of modification of these residues as the channel switches from the closed to open confirmation are *color coded* according to the color spectrum bar. Residues displaying less than twofold increase in modification rate are not highlighted. (**b**) Schematic model showing the predicted conformational change accompanies KirBac3.1 gating from the closed to the open state based on hydroxyl-radical footprinting. Reproduced from [63] with permission

10.3.2 Ion Channel–Lipid Interactions

Far from being a mere passive medium in which membrane proteins embed themselves, the lipids forming the plasma membrane can markedly affect structure and function by virtue of the manner in which they influence the structure of the membrane-spanning domains. The often necessary substitution of the native lipid bilayer with either a detergent micelle or a nonnative mimetic lipid environment, in order to perform structural studies with X-ray crystallography or NMR, can lead to inaccuracies in subsequently resolved structures for transmembrane protein segments [3, 65]. This issue is particularly acute for ion channels, where the transmembrane domains are not simply there to anchor the protein in the bilayer but have distinct functional responsibilities, whether it be the formation of the ion conduction pathway, gating or normal subunit assembly, and stoichiometry. For example, structural analysis of the viral proton channel, M2, using both X-ray crystallography and NMR, have shown differences in the arrangements of the transmembrane helices depending on whether the channel was reconstituted in detergent micelle or liposomes [3]. In early studies of the K^+ channel, KcsA, it was observed that the normal assembly of the full tetrameric protein was highly dependent on the presence of a lipid bilayer environment, with detergent solubilization causing considerable instability [66]. Analysis of crystal structures for the bacterial voltage-gated K^+ channel, KvAP, revealed the likelihood that reconstitution of the protein in detergent micelles, which do not form bilayers, was responsible for the nonnative positioning of the voltage-sensing domain relative to the channel pore [67]. Thus, the need to solubilize ion channel proteins for structural analysis poses a very real problem with regard to how closely the structure resolved from the nonnative environment corresponds to the structure in the native environment.

Using the KcsA channel as their model, Demmers et al. [68] investigated the how different lipids interacted with the channel. They reconstituted KcsA channels in liposomal bilayers of different lipid composition, exposed them to a mild solvent, and then employed electrospray mass spectrometry to attempt to identify preferences in non-covalent lipid binding with the protein. Using this method, they were able to ascertain a preference for KcsA interaction with the phospholipid, phosphatidylglycerol relative to another abundant membrane lipid, phosphatidylcholine. Conceptually, then, it was shown that mass spectrometry could be used to study the effects of lipid interactions on protein assembly and structure. However, limitations included the need for a solvent to be present with the protein–lipid complex, and the fact that only interactions between lipids and single KcsA subunits could be resolved, likely due to the harsh conditions required to ionize the complexes for subsequent mass spectrometry. As briefly mentioned earlier, Robinson et al. developed an approach by which protein complexes could first be liberated from their detergent micelle prior to analysis by mass spectrometry [29, 47–49]. Crucially, the removal of detergent in this manner has been shown to leave some protein–lipid interactions intact, as demonstrated for the ATP-binding cassette protein, MacB [69] (and see Fig. 10.1a), making this a very promising approach for studying ion channel–lipid interactions.

10.4 Identifying Posttranslational Modification of Ion Channel Subunits

As with all proteins, the functional properties of an assembled ion channel can be altered by posttranslational modification, in which a chemical group is covalently attached to an accessible amino acid side chain in the channel structure. The major examples of posttranslational modification are phosphorylation, glycosylation, ubiquitination, methylation, nitrosylation, and oxidation. Depending on the protein kinase, phosphorylation usually occurs on cytosolic serine, threonine, or tyrosine side chains of a membrane protein. Methylation can occur at lysine or arginine residues. Nitrosylation and oxidation occur on the thiol group of cysteine side chains. Ubiquitination occurs on lysine side chains. In contrast, glycosylation usually occurs on asparagine side chains that are ultimately located on the extracellular portions of a membrane protein. Clearly, the addition of such chemical groups to a protein will affect its mass accordingly, making mass spectrometry an appropriate technique for the identification and localization of posttranslational modifications, and this application has been used extensively in the study of soluble proteins [70].

In the case of ion channels, posttranslational modification is of considerable importance, not only with regard to the appropriate trafficking of the protein to the correct biological membrane compartments but with regard to the functioning of the channel as a passive transporter of ions. Indeed, posttranslational modification is a key mechanism by which the function of ion channels can be regulated by second messenger-signaling mechanisms, or rendered sensitive to changes in cell redox state. Gating sensitivity to stimulation, gating kinetics, inactivation, and even ion selection and conductance can all be modulated by the covalent modification of side chains in functionally important parts of the ion channel. For example, PKA-mediated phosphorylation of the voltage-gated Na^+ channel, NaV1.2, and PKC-mediated phosphorylation of the voltage-gated Ca^{2+} channel, Cav2.2 alter current profile [71, 72]. During sustained activation of TRPV1 channels, which mediate the responses of pain-processing sensory neurons to noxious heat and the chili pepper ingredient capsaicin, second messenger-signaling mechanisms engaged due to ion flux through the ion channel lead to kinase-mediated phosphorylation of amino acid side chains located on the intracellular surface of the channel resulting in channel desensitization (also termed inactivation) [73]. On the other hand, in pain-processing neurons the activity of G protein-coupled receptors activated by chemical mediators released by damaged cells and leukocytes during inflammation can lead to phosphorylation and subsequent sensitization of neighboring ion channels to their thermal and chemical stimuli, rendering the neuron much more sensitive to transmitting pain producing action potentials (peripheral hyperalgesia) [74]. The preference of the ion channel for selecting and conducting ions can also be subtly altered by phosphorylation, as has been reported for the ionotropic glutamate-gated NMDA receptors, where PKA-dependent phosphorylation alters relative Ca^{2+} permeability in a manner potentially relevant to the regulation of synaptic plasticity by these receptors [75].

When attempting to relate structure to function, it is of course vital to understand the role that posttranslational modification plays in ion channel function. Lack of knowledge here can severely limit the ability to draw accurate conclusions from crystal structures that might be obtained from ion channel proteins lacking modifications that are necessary for their normal functioning. Again, our objective as structural biologists is to determine the structure of the channel in a native setting, where correlations between structure and the channels actual functioning are most meaningful.

Identifying sites of posttranslational modification in ion channels has previously involved simply looking for known recognition motifs for modification enzymes, such as kinases, within the primary sequence of a channel subunit and then assessing whether site-directed mutagenesis at these positions affects the posttranslational modification. For example, potential recognition motifs for protein kinases are identified and amino acids within them substituted in order to either disrupt phosphorylation or mimic it (by substituting a side chain in the motif with a negatively charged Asp, for example). The mutant channels would then be assayed for incorporation of ^{32}P and compared to the wild type [76–79]. The functional significance of phosphorylation would then be determined pharmacologically by assessing the effect of protein kinase inhibitors and activators on channel function. A limitation of this strategy is the labor required to systematically assay mutant receptors one-by-one in order to identify which kinase recognition sites actually serve as a substrate for a kinase versus those that are inaccessible to modification. Another limitation to this approach is that some phosphorylation sites can arise only through alternative splicing of the ion channel subunit mRNA, meaning they are not identified by reading the original gene transcript [80].

The potential for using mass spectrometry to evaluate posttranslational modification of ion channels was actually realized relatively early, with the successful identification of glycosylated forms of the nAChR [81]. Much more recently is has been utilized to reveal the importance of glycosylation in regulating the function of voltage-gated ion channels in cardiac myocytes, and how aberrant glycosylation can impact the behavior of these cells [98].

10.4.1 Phosphorylation

Mass spectrometry is already established as a powerful technique for assessing protein phosphorylation [26, 82, 83], but it is only in recent years that it has been systematically applied to the identification of phosphorylation sites in ion channels. Usually, this has involved expressing the ion channel of interest in an expression system such as HEK293 or CHO cells, immunoaffinity purification of the protein, subjecting it to enzymatic cleavage and then assessing the fragments for phosphorylation using mass spectrometry. In this manner, a number of functionally significant phosphorylation sites have been discovered in voltage-gate ion channels, particularly K^+ channels [84]. For example, a functionally important

PKC-dependent phosphorylation site was discovered on the voltage-gated K^+ channel, Kv11.1, expressed and purified from CHO cells [85]. Subsequent functional studies revealed that phosphorylation of this site was important for maintaining a high density of the ion channel in the membrane and for regulating recovery from inactivation. Furthermore, dysfunction of this phosphorylation site due to single amino acid substitution was shown to be responsible for the Long QT syndrome associated with this particular naturally occurring polymorphism. In another study, mass spectrometry used in combination with a conventional ^{32}P phosphorylation assay identified a number of PKA phosphorylation sites on the hyperpolarization-activated, cAMP-sensitive nonselective cation channel, HCN4, expressed and purified from CHO cells [86], one of which was responsible for sifting the voltage sensitivity of the ion channel to more positive potentials (effectively sensitizing the ion channel). HCN4 is expressed in the specialized myocytes of the primary cardiac pace maker, the sinoatrial node, where it regulates action potential frequency and profile, and potentially mediates the effects of sympathetic adrenergic input to the heart.

In these cases, mass spectrometry was used simply to determine where phosphorylation can take place in the ion channel protein, with subsequent functional studies used to determine the specific kinase (or phosphatase) responsible for phosphorylating (or dephosphorylating) these sites. However, use of "mass-tagging" [26] in combination with mass spectrometry can allow relative phosphorylation levels of a protein to be compared between a control cell sample and cells in which a particular kinase or phosphatase has been activated/inhibited. The technique involves incorporating "light" molecular weight amino acids containing ^{12}C and ^{14}N into proteins expressed in one sample of cells, while incorporating "heavy" molecular weight amino acids containing ^{13}C and ^{15}N into proteins expressed in a second sample of cells, and then subjecting the different samples to different treatment regiments. This technique was used to compare the phosphorylation levels of the voltage-gated K^+ channel, KV2.1, in the absence and presence of the activated phosphatase, calcineurin [87, 88].

The channel could be isolated from an expression system, but crucially, due to the relatively low concentrations of protein required for mass spectrometry, also purified directly from tissue preparations, assuming selective antibodies are available. In a study focusing on the Ca^{2+}-activated large conductance potassium channel, BK_{Ca}, Yan et al. [80] used mass to identify phosphorylation sites on the pore-forming α subunit of the channel. In their study, Yan et al. immunopurified the BK_{Ca} a subunit from rat brain, allowing them to gain a better idea of the phosphorylation of the protein in vivo. They discovered that the protein is extensively phosphorylated, with phosphate groups located on no less than 30 serine and threonine residues. A small number of sites were found on the extracellular terminus, but the vast majority were located on the cytosolic side of the protein, explaining why BK_{Ca} function is so sensitive to regulation by protein kinases and phosphatases ([89–91] and see [92]).

More recently, high-throughput proteomics approaches conducted in mouse brain have yielded vast "global" data sets of phosphorylated proteins, of which many

ion channels are included [24, 27, 28]. This work has been recently reviewed [23]. The pore-forming α-subunits of the voltage-gated Na^+ and Ca^{2+} channels were both found to possess between 20 and 25 phosphorylation sites each. Between 13 and 15 phosphorylation sites were identified on voltage-gated K^+ channels, with ten located on Ca^{2+}-activated BK_{Ca} channels. Mass spectrometry continues to be utilized in the identification of ion channel phosphorylation sites, and has generated important findings in TRP channels [93, 94], ligand-gated P2XRs [95] and Cl^- channels [96].

10.4.2 Other Posttranslational Modifications

Glycosylation: In addition to its role in protein trafficking, glycosylation has also been found to be important for the normal functioning of the channel upon reaching its target compartment, with incorrect glycosylation of ion channels underpinning an array of congenital diseases [97]. Unlike the other reversible posttranslational modifications discussed, this modification is largely permanent. When one considers that up to 30 % of the mass of a voltage-gated ion channel can consist of glycans, it is not difficult to see why the extent of this modification could potentially influence how the channel senses voltage and conducts ions, and why aberrant glycosylation can cause damaging alterations of the action potential profile in excitable cells, particularly when those excitable cells are cardiac myocytes [98]. In a study by Montpetit et al. [98], MALDI mass spectrometry was used to screen the "glycome" of atrial and ventricular myocytes in order to compare differences between neonatal and adult cells, with the prediction that glycolysis mapping might change during cardiac development. These changes correlated with the expression of glycogenes, and the investigators found that down regulation of one of these glycogenes substantially reduced glycosylation of voltage-gated Na^+ channels in a manner that altered their gating characteristics and, subsequently, the profile of the cardiac action potential.

Nitrosylation: Although nitrosylation was originally shown to regulate NMDA receptor function 20 year ago [99], the extent to which this form of posttranslational modification is important for channel function is not nearly as well understood as for phosphorylation and glycosylation. More recently, mass spectrometry has been used to identify cysteines residues susceptible to covalent modulation by nitrosylation in the endoplasmic reticulum Ca^{2+} channel, the RyR, which is likely to impart the strong sensitivity of this ion channel to changes in redox state [100, 101].

Ubiquitination: Ubiquitination of proteins is an important step in their degradation pathway, and involves a multi-enzymatic process by which ubiquitin is covalently attached to lysine side chains in the target protein. Due to ubiquitin containing lysine residues, it too can be targeted for ubiquitination, giving rise to proteins modified by polyubiquitin chains of varying length. This label flags the protein for proteosomal degradation. Sliter et al. [101, 102] employed mass spectrometry to screen fragments of the Ca^{2+} permeable endoplastic reticulum ion channel, the $InsP_3$ receptor, discovering several lysine side chains at which ubiquitin had been attached.

Methylation: The methylation of proteins involves the enzymatic attachment of up to three methyl groups to the ε-amino group in the side chain of lysine or arginine residues [103]. This, of course, has a profound effect on the side chains solubility, increasing its hydrophobicity. Using a combination of both MALDI and liquid chromatography-tandem mass spectrometry, Beltran-Alvarez et al. [104] demonstrated posttranslational methylation of the voltage-gated Na^+ channel, NaV1.5, which is expressed in cardiac myocytes. Two of the three Arg residues shown to be methylated are known to contribute to the pathology of two congenital disorders of the heart, Brugada and Long QT Type 3 syndromes.

Polyester modification: In what may well emerge to be a significant finding, mass spectrometry has recently been used to study a novel and yet potentially important type of posttranslational modification involving the covalent attachment of naturally occurring polymers to ion channels. Polyphosphates and polyhydroxybutyrates were originally characterized in bacteria, but have since been found to be relatively widespread in eukaryotes, although their physiological significance is only just beginning to be investigated. Two members of the TRP channel superfamily of ion channels, TRPA1 and TRPM8, have been observed to be susceptible to covalent attachment by these polymers with clear functional consequences. In the case of TRPA1, an channel sensitive to cold temperatures and a range of pungent chemicals such as the wasabi ingredient allyl isothiocyanate, the channel actually requires modification by inorganic polyphosphates in order to function at all [105]. Recently, the use of mass spectrometry in combination with site-directed mutagenesis and patch clamp electrophysiology revealed that TRPM8 is modified by both polyphosphates and polyhydroxybutyrates, and that like TRPA1, the normal functioning of the channel is highly dependent on these modifications occurring [106, 107]. Given the diversity in size of the polyphosphate and polyhydroxybutyrate polymers, and the hydrophobicity of the latter, the potential for these molecules to fundamentally alter the properties of ion channels is considerable. In the event that crystal structures are obtained for TRP family members in the future, or indeed for any ion channels, caution is required in assuming that these represent the actual functioning units of the channel in their native setting.

Covalent modification of ion channels by agonists TRPA1 receptors are ion channels expressed in pain-processing neurons, particularly those of the trigeminal ganglion innervating the mouth and nasal cavities. They are polymodal cation channels that, when activated, depolarize the sensory neurons thereby facilitating transmission to the central nervous system. They are particularly important for sensing noxious environmental stimuli and very low temperatures, and it is these channels that are responsible for mediating the nasal pain associated with wasabi (the active ingredient of which is allyl thiocyanate) and inhaling very cold air. Whereas most ligand-gated ion channels are activated by a reversibly binding ligand, mass spectrometry was used to reveal that an unusual feature of TRPA1 ion channels is that some chemical compounds activate them by becoming covalently attached to the ion channel [108].

10.4.3 Protein–Protein Interactions

Of the currently characterized mammalian ion channels, the minimum functional unit usually consists of an oligomeric peptide complex. In the cases of the ligand-gated ion channels and some voltage-gated K^+ channels, the ion-permeable pore is formed by between three and five polypeptide subunits arranged symmetrically around the conduction pathway. In the case of some voltage-gated ion channels, a single polypeptide α subunit forms the pore, although this single peptide actually represents an evolutionary development in which four initially separate peptide subunits fused to form one. For the most part, expression of the pore-forming subunit(s) is sufficient to give a functioning gated ion channel. With respect to elucidating the stoichiometry of the minimum functional ion channel unit, mass spectrometry provides one useful approach [29, 109]. However, it has become apparent that in vivo the pore-forming subunits of ion channels are also in contact with multiple regulatory proteins that serve to influence ion channel expression, localization, gating, and internalization [110]. The best known example being the routine co-assembly of voltage-gated ion channel α subunits with regulatory β subunits [111, 112].

Traditionally, the identification of interacting membrane proteins has necessitated the use of co-immunoprecipitation pull-down and Western blot studies. In addition to advanced two-hybrid screening techniques for investigating protein–protein interactions involving membrane proteins [113], mass spectrometry is proving to be useful as part of a high-throughput approach to the mapping of protein–protein interactions involving ion channels within native cell macromolecular-signaling complexes (or "signalsomes") [25, 114, 115].

One strategy is to use purified ion channel protein fragments immobilized on agarose beads to screen for potential interacting proteins. This was employed by Bildl et al. [116] who identified a number of proteins from rat brain homogenates that could potentially co-assemble with the small conductance Ca^{2+}-activated K^+ (SK) channel. Alternatively, assuming selective antibodies are available, the ion channel of interest can be directly immunoprecipitated from tissue, and any attached proteins pulled down a long with identified using liquid chromatography-tandem mass spectrometry (Fig. 10.3), as has been employed for voltage-gated K^+ channels [114, 117] and voltage-gated Ca^{2+} channels [118], glutamate-gated AMPA receptors [119], and intracellular chloride channels [120]. In addition to the limitations of mass spectrometry, the use of immunoprecipitation has the additional well known caveats associated with antibody specificity, the likelihood that protein assemblies are preserved during the immunoprecipitation process, and that no abhorrent aggregates are formed involving proteins that do not ordinarily interact in the native setting [114].

10.4.4 Splice Variants

Mass spectrometry has recently been employed to localize the presence of a specific voltage-gated Ca^{2+} channel, CaV2.2, splice variant to the presynaptic terminal, where it presumably contributes to synaptic transmission [121]. Yan et al. [80]

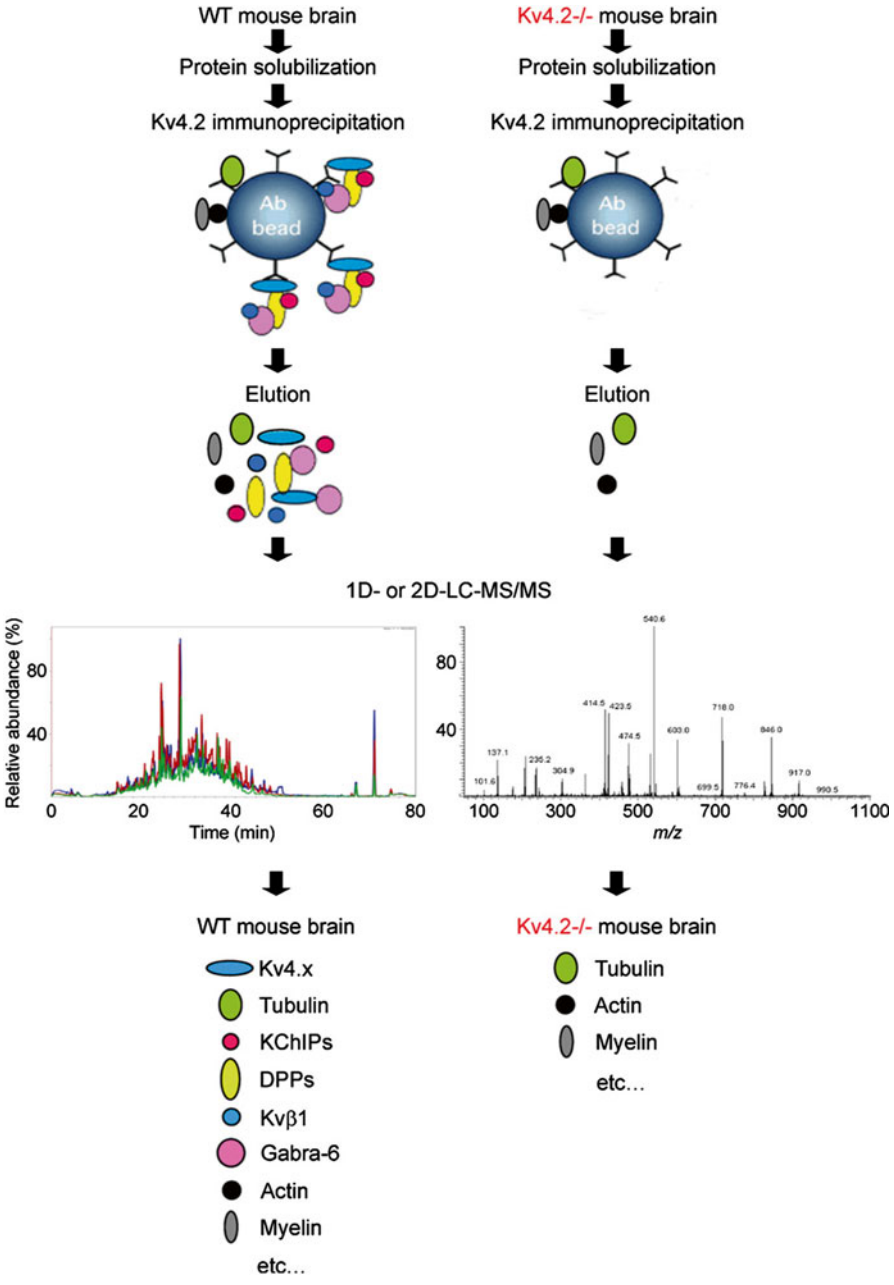


Fig. 10.3 In solution proteomics approach to resolving the macromolecular signaling domains occupied by a voltage-gated K^+ channel. Solubilized protein from brain homogenates extracted from wild type and Kv4.2$^{-/-}$ knockout mice were immunoprecipitated using an antibody for Kv4.2 under conditions designed to preserve protein–protein interactions. The resulting pulled down proteins were eluted, proteolysed, and the fragments sequenced by liquid chromatography and mass spectrometry and used to identify the parent proteins from a database. Reproduced from [114] with permission

employed nanoflow liquid chromatography and mass spectrometry to identify mRNA splice variants of the BK_{Ca} channel, with the objective of determining splice variant-dependent differences in protein phosphorylation sites. Subsequent functional studies on BK_{Ca} channels with mutations at these phosphorylation sites designed to disrupt or mimic phosphorylation revealed that they were important for regulating both their voltage-dependence and Ca^{2+} sensitivity. This perhaps explaining a recent finding that alternative splicing of BK_{Ca} in the blood pressure and gas-sensing cells of the carotid body confers sensitivity of the channel to inhibition by AMP-activated kinase [122].

10.5 Summary

Due to the inherent lack of aqueous solubility and low abundance of membrane proteins, resolving the structure of ion channels has been a considerable challenge. However, over the last 10 years, there have been numerous advances in each one of the principal techniques employed for determining the three-dimensional structure of membrane proteins, including X-ray crystallography, AFM and EM microscopy, NMR and, relevant to the topic of this chapter, mass spectrometry. Although not capable of the spatial resolutions achieved by X-ray crystallography and NMR, mass spectrometry is nevertheless proving to be valuable tool for the following reasons. First, by virtue of its compatibility with high-throughput analysis, mass spectrometry is revealing much about ion channel splice variants, their posttranslational modifications, and the interactions these proteins have with other proteins and the surrounding membrane lipid environment. Second, what mass spectrometry lacks in spatial resolution it more than makes up for in temporal resolution, with the use of solvent accessibility techniques allowing for a more refined view of the numerous conformational state transitions that accompany ion channel gating. Finally, in contrast to the common need to structurally manipulate ion channel proteins for the necessary expression and solubilization steps required for X-ray crystallography, mass spectrometry can potentially be used to probe the structure and function of native ion channels isolated straight from their native tissues. Taken together, mass spectrometry provides a vital and complementary approach to ion channel structure and function by effectively compensating for the limitations of alternative techniques. It is almost certain that subsequent advances in our understanding of ion channel function will be derived from a multipronged strategy that utilizes a combination of techniques, of which mass spectrometry is surely to be an invaluable component.

Acknowledgments I would like to thank Drs. Carol Robinson (Oxford University, UK), Mark Chance (Case Western, OH, USA), and Jeanne Nerbonne (Washington University, MO, USA) for permission to use published figures.

References

1. Carpenter EP et al (2008) Overcoming the challenges of membrane protein crystallography. Curr Opin Struct Biol 18(5):581–586
2. Lundstrom K (2004) Structural genomics on membrane proteins: the MePNet approach. Curr Opin Drug Discov Devel 7(3):342–346
3. Cross TA et al (2011) Influence of solubilizing environments on membrane protein structures. Trends Biochem Sci 36(2):117–125
4. Schneider G, Fechner U (2005) Computer-based de novo design of drug-like molecules. Nat Rev Drug Discov 4(8):649–663
5. Jorgensen WL (2004) The many roles of computation in drug discovery. Science 303(5665): 1813–1818
6. Baconguis I, Gouaux E (2012) Structural plasticity and dynamic selectivity of acid-sensing ion channel-spider toxin complexes. Nature 489(7416):400–405
7. Hattori M, Gouaux E (2012) Molecular mechanism of ATP binding and ion channel activation in P2X receptors. Nature 485(7397):207–212
8. Jasti J et al (2007) Structure of acid-sensing ion channel 1 at 1.9 A resolution and low pH. Nature 449(7160):316–323
9. Kawate T et al (2009) Crystal structure of the ATP-gated P2X(4) ion channel in the closed state. Nature 460(7255):592–598
10. Sobolevsky AI, Rosconi MP, Gouaux E (2009) X-ray structure, symmetry and mechanism of an AMPA-subtype glutamate receptor. Nature 462(7274):745–756
11. Unwin N (2005) Refined structure of the nicotinic acetylcholine receptor at 4A resolution. J Mol Biol 346(4):967–989
12. Long SB, Campbell EB, Mackinnon R (2005) Crystal structure of a mammalian voltage-dependent Shaker family K+ channel. Science 309(5736):897–903
13. Payandeh J et al (2011) The crystal structure of a voltage-gated sodium channel. Nature 475(7356):353–358
14. Cherezov V et al (2007) High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor. Science 318(5854):1258–1265
15. Rasmussen SG et al (2007) Crystal structure of the human beta2 adrenergic G-protein-coupled receptor. Nature 450(7168):383–387
16. Jaakola VP et al (2008) The 2.6 angstrom crystal structure of a human A2A adenosine receptor bound to an antagonist. Science 322(5905):1211–1217
17. Wu B et al (2010) Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonists. Science 330(6007):1066–1071
18. Katritch V, Cherezov V, Stevens RC (2013) Structure-function of the G protein-coupled receptor superfamily. Annu Rev Pharmacol Toxicol 53:531–556
19. Samways D, Li Z, Egan TM (2014) Principles and properties of ion flow in P2X receptors. Front Cell Neurosci 8:6
20. Heymann G et al (2013) Inter- and intrasubunit interactions between transmembrane helices in the open state of P2X receptor channels. Proc Natl Acad Sci U S A 110(42):E4045–E4054
21. Wales TE, Engen JR (2006) Hydrogen exchange mass spectrometry for the analysis of protein dynamics. Mass Spectrom Rev 25(1):158–170
22. Woodward CK, Hilton BD (1979) Hydrogen exchange kinetics and internal motions in proteins and nucleic acids. Annu Rev Biophys Bioeng 8:99–127
23. Cerda O, Baek JH, Trimmer JS (2011) Mining recent brain proteomic databases for ion channel phosphosite nuggets. J Gen Physiol 137(1):3–16
24. Lemeer S, Heck AJ (2009) The phosphoproteomics data explosion. Curr Opin Chem Biol 13(4):414–420
25. Liu XY et al (2008) Comparative proteomics and correlated signaling network of rat hippocampus in the pilocarpine model of temporal lobe epilepsy. Proteomics 8(3):582–603

26. Ong SE, Mann M (2005) Mass spectrometry-based proteomics turns quantitative. Nat Chem Biol 1(5):252–262
27. Wang H et al (2006) Characterization of the mouse brain proteome using global proteomic analysis complemented with cysteinyl-peptide enrichment. J Proteome Res 5(2):361–369
28. Yang X et al (2011) Comprehensive two-dimensional liquid chromatography mass spectrometric profiling of the rat hippocampal proteome. Proteomics 11(3):501–505
29. Barrera NP, Robinson CV (2011) Advances in the mass spectrometry of membrane proteins: from individual proteins to intact complexes. Annu Rev Biochem 80:247–271
30. Takamoto K, Chance MR (2006) Radiolytic protein footprinting with mass spectrometry to probe the structure of macromolecular complexes. Annu Rev Biophys Biomol Struct 35:251–276
31. Ojanpera I, Kolmonen M, Pelander A (2012) Current use of high-resolution mass spectrometry in drug screening relevant to clinical and forensic toxicology and doping control. Anal Bioanal Chem 403(5):1203–1220
32. Verkman AS (2012) Aquaporins in clinical medicine. Annu Rev Med 63:303–316
33. Hille B (2001) Ion channels of excitable membranes, 3rd edn. Sinauer Associates, Sunderland, MA
34. Berridge MJ, Lipp P, Bootman MD (2000) The versatility and universality of calcium signalling. Nat Rev Mol Cell Biol 1(1):11–21
35. Sundelacruz S, Levin M, Kaplan DL (2009) Role of membrane potential in the regulation of cell proliferation and differentiation. Stem Cell Rev 5(3):231–246
36. Sine SM et al (2010) On the origin of ion selectivity in the Cys-loop receptor family. J Mol Neurosci 40(1–2):70–76
37. Traynelis SF et al (2010) Glutamate receptor ion channels: structure, regulation, and function. Pharmacol Rev 62(3):405–496
38. North RA (2002) Molecular physiology of P2X receptors. Physiol Rev 82(4):1013–1067
39. Catterall WA (2011) Voltage-gated calcium channels. Cold Spring Harb Perspect Biol 3(8):a003947
40. Hodgkin AL, Huxley AF (1952) The components of membrane conductance in the giant axon of Loligo. J Physiol 116(4):473–496
41. Eisenberg D et al (1984) Analysis of membrane and surface protein sequences with the hydrophobic moment plot. J Mol Biol 179(1):125–142
42. Kyte J, Doolittle RF (1982) A simple method for displaying the hydropathic character of a protein. J Mol Biol 157(1):105–132
43. Hollmann M, Maron C, Heinemann S (1994) N-glycosylation site tagging suggests a three transmembrane domain topology for the glutamate receptor GluR1. Neuron 13(6):1331–1343
44. Bahouth SW, Wang HY, Malbon CC (1991) Immunological approaches for probing receptor structure and function. Trends Pharmacol Sci 12(9):338–343
45. Torres GE, Egan TM, Voigt MM (1998) Topological analysis of the ATP-gated ionotropic [correction of ionotrophic] P2X2 receptor subunit. FEBS Lett 425(1):19–23
46. Leite JF, Amoscato AA, Cascio M (2000) Coupled proteolytic and mass spectrometry studies indicate a novel topology for the glycine receptor. J Biol Chem 275(18):13683–13689
47. Barrera NP et al (2008) Micelles protect membrane complexes from solution to vacuum. Science 321(5886):243–246
48. Hopper JT et al (2013) Detergent-free mass spectrometry of membrane protein complexes. Nat Methods 10(12):1206–1208
49. Barrera NP et al (2009) Mass spectrometry of membrane transporters reveals subunit stoichiometry and interactions. Nat Methods 6(8):585–587
50. Laganowsky A et al (2013) Mass spectrometry of intact membrane protein complexes. Nat Protoc 8(4):639–651
51. Wang SC et al (2010) Ion mobility mass spectrometry of two tetrameric membrane protein complexes reveals compact structures and differences in stability and packing. J Am Chem Soc 132(44):15468–15470

52. Liapakis G, Simpson MM, Javitch JA (2001) The substituted-cysteine accessibility method (SCAM) to elucidate membrane protein structure. Curr Protoc Neurosci. Chapter 4:Unit 4. 15
53. Roberts JA, Evans RJ (2007) Cysteine substitution mutants give structural insight and identify ATP binding and activation sites at P2X receptors. J Neurosci 27(15):4072–4082
54. Roberts JA et al (2009) Contribution of the region Glu181 to Val200 of the extracellular loop of the human P2X1 receptor to agonist binding and gating revealed using cysteine scanning mutagenesis. J Neurochem 109(4):1042–1052
55. Li M et al (2008) Gating the pore of P2X receptor channels. Nat Neurosci 11(8):883–887
56. Leite JF, Cascio M (2001) Structure of ligand-gated ion channels: critical assessment of biochemical data supports novel topology. Mol Cell Neurosci 17(5):777–792
57. Leite JF, Cascio M (2002) Probing the topology of the glycine receptor by chemical modification coupled to mass spectrometry. Biochemistry 41(19):6140–6148
58. Xu G, Chance MR (2007) Hydroxyl radical-mediated modification of proteins as probes for structural proteomics. Chem Rev 107(8):3514–3543
59. Busenlehner LS, Armstrong RN (2005) Insights into enzyme structure and dynamics elucidated by amide H/D exchange mass spectrometry. Arch Biochem Biophys 433(1):34–46
60. Busenlehner LS et al (2004) Stress sensor triggers conformational response of the integral membrane protein microsomal glutathione transferase 1. Biochemistry 43(35): 11145–11152
61. Man P et al (2007) Defining the interacting regions between apomyoglobin and lipid membrane by hydrogen/deuterium exchange coupled to mass spectrometry. J Mol Biol 368(2): 464–472
62. Joh NH et al (2008) Modest stabilization by most hydrogen-bonded side-chain interactions in membrane proteins. Nature 453(7199):1266–1270
63. Gupta S et al (2010) Conformational changes during the gating of a potassium channel revealed by structural mass spectrometry. Structure 18(7):839–846
64. Zhu Y et al (2009) Elucidating in vivo structural dynamics in integral membrane protein by hydroxyl radical footprinting. Mol Cell Proteomics 8(8):1999–2010
65. Zhou HX, Cross TA (2013) Influences of membrane mimetic environments on membrane protein structures. Annu Rev Biophys 42:361–392
66. van Dalen A et al (2002) Influence of lipids on membrane assembly and stability of the potassium channel KcsA. FEBS Lett 525(1–3):33–38
67. Lee SY et al (2005) Structure of the KvAP voltage-dependent K+ channel and its dependence on the lipid membrane. Proc Natl Acad Sci U S A 102(43):15441–15446
68. Demmers JA et al (2003) Interaction of the K+ channel KcsA with membrane phospholipids as studied by ESI mass spectrometry. FEBS Lett 541(1–3):28–32
69. Lin HT et al (2009) MacB ABC transporter is a dimer whose ATPase activity and macrolide-binding capacity are regulated by the membrane fusion protein MacA. J Biol Chem 284(2): 1145–1154
70. Witze ES et al (2007) Mapping protein post-translational modifications with mass spectrometry. Nat Methods 4(10):798–806
71. Cantrell AR et al (1997) Dopaminergic modulation of sodium current in hippocampal neurons via cAMP-dependent phosphorylation of specific sites in the sodium channel alpha subunit. J Neurosci 17(19):7330–7338
72. Fang H et al (2006) Inhibitory role of Ser-425 of the alpha1 2.2 subunit in the enhancement of Cav 2.2 currents by phorbol-12-myristate, 13-acetate. J Biol Chem 281(29):20011–20017
73. Koplas PA, Rosenberg RL, Oxford GS (1997) The role of calcium in the desensitization of capsaicin responses in rat dorsal root ganglion neurons. J Neurosci 17(10):3525–3537
74. Julius D, Basbaum AI (2001) Molecular mechanisms of nociception. Nature 413(6852): 203–210
75. Skeberdis VA et al (2006) Protein kinase A regulates calcium permeability of NMDA receptors. Nat Neurosci 9(4):501–510
76. Vial C, Tobin AB, Evans RJ (2004) G-protein-coupled receptor regulation of P2X1 receptors does not involve direct channel phosphorylation. Biochem J 382(Pt 1):101–110

77. Siegel JN (2001) Preparation and analysis of phosphorylated proteins. Curr Protoc Immunol Chapter 11:Unit 11.2
78. Leonard AS et al (2003) cAMP-dependent protein kinase phosphorylation of the acid-sensing ion channel-1 regulates its binding to the protein interacting with C-kinase-1. Proc Natl Acad Sci U S A 100(4):2029–2034
79. Tan SE, Wenthold RJ, Soderling TR (1994) Phosphorylation of AMPA-type glutamate receptors by calcium/calmodulin-dependent protein kinase II and protein kinase C in cultured hippocampal neurons. J Neurosci 14(3 Pt 1):1123–1129
80. Yan J et al (2008) Profiling the phospho-status of the BKCa channel alpha subunit in rat brain reveals unexpected patterns and complexity. Mol Cell Proteomics 7(11):2188–2198
81. Poulter L et al (1989) Structure, oligosaccharide structures, and posttranslationally modified sites of the nicotinic acetylcholine receptor. Proc Natl Acad Sci U S A 86(17):6645–6649
82. Beausoleil SA et al (2006) A probability-based approach for high-throughput protein phosphorylation analysis and site localization. Nat Biotechnol 24(10):1285–1292
83. Jensen ON (2006) Interpreting the protein language using proteomics. Nat Rev Mol Cell Biol 7(6):391–403
84. Park KS et al (2008) Potassium channel phosphorylation in excitable cells: providing dynamic functional variability to a diverse family of ion channels. Physiology (Bethesda) 23:49–57
85. Donovan AJ et al (2012) Long QT2 mutation on the Kv11.1 ion channel inhibits current activity by ablating a protein kinase C alpha consensus site. Mol Pharmacol 82(3):428–437
86. Liao Z et al (2010) Phosphorylation and modulation of hyperpolarization-activated HCN4 channels by protein kinase A in the mouse sinoatrial node. J Gen Physiol 136(3):247–258
87. Mohapatra DP, Park KS, Trimmer JS (2007) Dynamic regulation of the voltage-gated Kv2.1 potassium channel by multisite phosphorylation. Biochem Soc Trans 35(Pt 5):1064–1068
88. Park KS, Mohapatra DP, Trimmer JS (2007) Proteomic analyses of K(v)2.1 channel phosphorylation sites determining cell background specific differences in function. Channels (Austin) 1(2):59–61
89. Kyle BD et al (2013) Specific phosphorylation sites underlie the stimulation of a large conductance, Ca(2+)-activated K(+) channel by cGMP-dependent protein kinase. FASEB J 27(5):2027–2038
90. Tian L et al (2004) Distinct stoichiometry of BKCa channel tetramer phosphorylation specifies channel activation and inhibition by cAMP-dependent protein kinase. Proc Natl Acad Sci U S A 101(32):11897–11902
91. Vetri F et al (2012) Impairment of neurovascular coupling in type 1 diabetes mellitus in rats is linked to PKC modulation of BK(Ca) and Kir channels. Am J Physiol Heart Circ Physiol 302(6):H1274–H1284
92. Schubert R, Nelson MT (2001) Protein kinases: tuners of the BKCa channel in smooth muscle. Trends Pharmacol Sci 22(10):505–512
93. Lee JE et al (2012) Mass spectrometric analysis of novel phosphorylation sites in the TRPC4beta channel. Rapid Commun Mass Spectrom 26(17):1965–1970
94. Voolstra O et al (2010) Light-dependent phosphorylation of the drosophila transient receptor potential ion channel. J Biol Chem 285(19):14275–14284
95. Roberts JA et al (2012) Mass spectrometry analysis of human P2X1 receptors; insight into phosphorylation, modelling and conformational changes. J Neurochem 123(5):725–735
96. Falin RA et al (2009) Identification of regulatory phosphorylation sites in a cell volume- and Ste20 kinase-dependent ClC anion channel. J Gen Physiol 133(1):29–42
97. Schachter H (2001) Congenital disorders involving defective N-glycosylation of proteins. Cell Mol Life Sci 58(8):1085–1104
98. Montpetit ML et al (2009) Regulated and aberrant glycosylation modulate cardiac electrical signaling. Proc Natl Acad Sci U S A 106(38):16517–16522
99. Lipton SA et al (1993) A redox-based mechanism for the neuroprotective and neurodestructive effects of nitric oxide and related nitroso-compounds. Nature 364(6438): p. 626–32

100. 100. Martinez-Moreno M et al (2005) Direct interaction between the reductase domain of endothelial nitric oxide synthase and the ryanodine receptor. FEBS Lett 579(14):3159–3163
101. Voss AA et al (2004) Identification of hyperreactive cysteines within ryanodine receptor type 1 by mass spectrometry. J Biol Chem 279(33):34514–34520
102. Sliter DA et al (2011) Activated inositol 1,4,5-trisphosphate receptors are modified by homogeneous Lys-48- and Lys-63-linked ubiquitin chains, but only Lys-48-linked chains are required for degradation. J Biol Chem 286(2):1074–1082
103. Sliter DA et al (2008) Mass spectrometric analysis of type 1 inositol 1,4,5-trisphosphate receptor ubiquitination. J Biol Chem 283(51):35319–35328
104. Paik WK, Paik DC, Kim S (2007) Historical review: the field of protein methylation. Trends Biochem Sci 32(3):146–152
105. Beltran-Alvarez P, Pagans S, Brugada R (2011) The cardiac sodium channel is post-translationally modified by arginine methylation. J Proteome Res 10(8):3712–3719
106. Kim D, Cavanaugh EJ (2007) Requirement of a soluble intracellular factor for activation of transient receptor potential A1 by pungent chemicals: role of inorganic polyphosphates. J Neurosci 27(24):6500–6509
107. Cao C et al (2013) Polyester modification of the mammalian TRPM8 channel protein: implications for structure and function. Cell Rep 4(2):302–315
108. Zakharian E et al (2009) Inorganic polyphosphate modulates TRPM8 channels. PLoS One 4(4):e5404
109. Macpherson LJ et al (2007) Noxious compounds activate TRPA1 ion channels through covalent modification of cysteines. Nature 445(7127):541–545
110. Hoffmann J et al (2010) Studying the stoichiometries of membrane proteins by mass spectrometry: microbial rhodopsins and a potassium ion channel. Phys Chem Chem Phys 12(14):3480–3485
111. Catterall WA et al (2006) Regulation of sodium and calcium channels by signaling complexes. J Recept Signal Transduct Res 26(5–6):577–598
112. Buraei Z, Yang J (2010) The β subunit of voltage-gated Ca2+ channels. Physiol Rev 90(4):1461–1506
113. Chahine M, O'Leary ME (2011) Regulatory role of voltage-gated Na channel beta subunits in sensory neurons. Front Pharmacol 2:70
114. Snider J et al (2010) Detecting interactions with membrane proteins using a membrane two-hybrid assay in yeast. Nat Protoc 5(7):1281–1293
115. Marionneau C, Townsend RR, Nerbonne JM (2011) Proteomic analysis highlights the molecular complexities of native Kv4 channel macromolecular complexes. Semin Cell Dev Biol 22(2):145–152
116. Wang HY, Malbon CC (2011) Probing the physical nature and composition of signalsomes. J Mol Signal 6(1):1
117. Bildl W et al (2004) Protein kinase CK2 is coassembled with small conductance Ca(2+)-activated K+ channels and regulates channel gating. Neuron 43(6):847–858
118. Zolles G et al (2009) Association with the auxiliary subunit PEX5R/Trip8b controls responsiveness of HCN channels to cAMP and adrenergic stimulation. Neuron 62(6):814–825
119. Khanna R, Zougman A, Stanley EF (2007) A proteomic screen for presynaptic terminal N-type calcium channel (CaV2.2) binding partners. J Biochem Mol Biol 40(3):302–314
120. Schwenk J et al (2009) Functional proteomics identify cornichon proteins as auxiliary subunits of AMPA receptors. Science 323(5919):1313–1319
121. Suginta W et al (2001) Chloride intracellular channel protein CLIC4 (p64H1) binds directly to brain dynamin I in a complex containing actin, tubulin and 14-3-3 isoforms. Biochem J 359(Pt 1):55–64
122. Gardezi SR, Taylor P, Stanley EF (2010) Long C terminal splice variant CaV2.2 identified in presynaptic membrane by mass spectrometric analysis. Channels (Austin) 4(1):58–62
123. Ross FA et al (2011) Selective expression in carotid body type I cells of a single splice variant of the large conductance calcium- and voltage-activated potassium channel confers regulation by AMP-activated protein kinase. J Biol Chem 286(14):11929–11936

Chapter 11
A Mass Spectrometry View of Stable and Transient Protein Interactions

Hanna G. Budayeva and Ileana M. Cristea

Abstract Through an impressive range of dynamic interactions, proteins succeed to carry out the majority of functions in a cell. These temporally and spatially regulated interactions provide the means through which one single protein can perform diverse functions and modulate different cellular pathways. Understanding the identity and nature of these interactions is therefore critical for defining protein functions and their contribution to health and disease processes. Here, we provide an overview of workflows that incorporate immunoaffinity purifications and quantitative mass spectrometry (frequently abbreviated as IP-MS or AP-MS) for characterizing protein–protein interactions. We discuss experimental aspects that should be considered when optimizing the isolation of a protein complex. As the presence of nonspecific associations is a concern in these experiments, we discuss the common sources of nonspecific interactions and present label-free and metabolic labeling mass spectrometry-based methods that can help determine the specificity of interactions. The effective regulation of cellular pathways and the rapid reaction to various environmental stresses rely on the formation of stable, transient, and fast-exchanging protein–protein interactions. While determining the exact nature of an interaction remains challenging, we review cross-linking and metabolic labeling approaches that can help address this important aspect of characterizing protein interactions and macromolecular assemblies.

H.G. Budayeva • I.M. Cristea (✉)
Department of Molecular Biology, 210 Lewis Thomas Laboratory,
Princeton University, Princeton, NJ 08544, USA
e-mail: icristea@princeton.edu

A.G. Woods and C.C. Darie (eds.), *Advancements of Mass Spectrometry in Biomedical Research*, Advances in Experimental Medicine and Biology 806,
DOI 10.1007/978-3-319-06068-2_11, © Springer International Publishing Switzerland 2014

11.1 Overview

Proteins carry out the majority of functions in a cell, and their regulation at a core of health or disease processes. While cells contain a multitude of proteins, of various sizes and abundances, it is almost never the case for one protein to have only one function. Any given one protein usually has the remarkable ability to perform numerous functions. The main tactic through which proteins carry out these diverse functions is through formation of numerous interactions. These interactions can be dynamic, spatially and temporally defined, and stable or transient in nature. For example, one enzyme can have many substrates, and their regulation can have different downstream impacts on cellular pathways. Similarly, one protein can be part of multiple protein complexes that can have distinct functions. Given their fundamental contribution to cellular processes, the study of protein–protein interactions has become an essential part of biological discovery. In this chapter, we discuss one of the most commonly utilized approaches for studying protein interactions—immunoaffinity purification coupled with mass spectrometry analysis (IP-MS). We start by describing the types of optimizations that need to be considered when designing an IP-MS experiment to ensure efficient isolation and accurate characterization of protein complexes. Next, we discuss what controls should we performed and how mass spectrometry data can be used to distinguish specific versus background interactions. Within this context, we cover some of the most frequently implemented label-free and metabolic labeling approaches. Lastly, we describe some of the recent developments in capturing transient associations and measuring the relative stability of interactions. Application of cross-linking approaches for studying protein complex structures and transient interactions is also discussed.

11.2 Methods for Isolating Protein Complexes

Common workflows for characterization of protein complexes. Immunoaffinity purification (IP) of proteins is a powerful approach for characterizing proteins of interest, their direct and indirect interactions required for formation of complexes, as well as their posttranslational modifications (PTMs). This information provides critical insights into the functions of proteins in different pathways, as well as regulation of their functions by various mechanisms (e.g., inhibiting or activating PTMs) [1, 2]. A standard workflow for isolating protein complexes is illustrated in Fig. 11.1. This workflow starts with the selection of an appropriate cell line or tissue sample, effective lysis of the sample, isolation and elution of the target protein with its interactions, followed by mass spectrometry analysis and identification and quantification of the co-isolated proteins. Each step of this process can be modified and optimized based on the nature of the protein of interest, its subcellular localization and abundance, and the overall goal of the study. Several important considerations for these optimization experiments are detailed below.

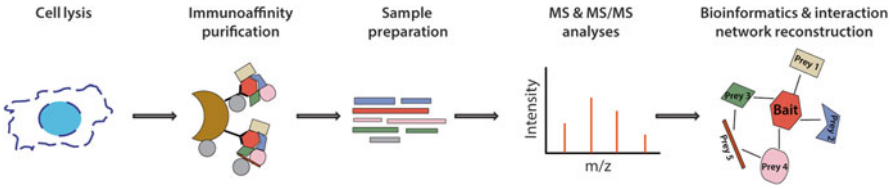


Fig. 11.1 Common workflow for immunoaffinity purification mass spectrometry experiments. Cells expressing the protein of interest are lysed and protein complexes are isolated by immunoaffinity purification. Eluted proteins are processed for MS analysis. MS spectra are analyzed to identify proteins within isolated complex(es) and bioinformatics tools are implemented to generate protein interaction networks

Optimizing conditions for immunoaffinity purifications. One of the first questions that needs to be addressed prior to IP is whether the endogenous protein or its tagged version would be a suitable candidate for the study. Endogenous proteins can be purified from tissue or cells at their biological levels, providing the best representation of their functional states. Isolation of an endogenous protein requires availability of antibody with high specificity and affinity to ensure efficient and clean purification. Endogenous protein isolation is extensively utilized in small- and large-scale studies [3–6]. For example, Li et al. isolated the nuclear DNA sensor IFI16 to define its localization-dependent antiviral functions [5]. Large-scale isolation of endogenous transcriptional and signaling proteins and their complexes was used by Malovannaya et al. to provide insight into the composition of human coregulator protein complexes [3]. However, antibody production can be costly and not all antibodies are commercially available or stored in a buffer compatible with IP conditions (e.g., large amounts of glycerol or storage in Tris buffer can interfere with coupling of an antibody to some resins), requiring additional purification steps. Therefore, these antibodies are routinely used in smaller-scale isolations for confirmatory studies. For example, Tsai et al. reported co-isolation of endogenous sirtuin 7 with B-WICH components and Pol I, supporting its role in regulation of Pol I transcription [4]. As an alternative to the use of antibodies for immunoaffinity purification studies, recent reports have proposed the use of small molecules, such as activity-based chemical probes or inhibitors, covalently linked to a resin for isolation of enzymes and their complexes [7–10]. For example, a large-scale study used histone deacetylase inhibitors to assess their affinity to different complexes [9]. Other approaches for isolation of endogenous targets include the use of nucleic acids and engineered binding proteins (reviewed by Ruigrok et al. [11]), which are actively incorporated in various biomedical studies [12, 13]. With their lower cost of production and higher stability than antibodies, these molecules have become valuable tools for isolating and characterizing protein complexes [14].

A more commonly utilized approach in studies of protein–protein interactions involves the tagging of the protein of interest, followed by its isolation using a tag-specific antibody. This method can be customized for the use of different tags (e.g., FLAG, EGFP, HA) and expression of the fusion protein from endogenous

(genomic) or exogenous (e.g., tetracycline-inducible) promoters. For example, Quantitative BAC InteraCtomics (QUBIC) approach utilizes expression of tagged proteins from native promoters, followed by IP and quantitative MS analysis, which aids the identification of specific interacting partners [15]. The use of a fluorescent tag, such as green fluorescent protein (GFP), allows combining information regarding protein localization and interactions [16] and provides a complementary validation of protein–protein interactions, as shown for virus–host protein interactions [17, 18]. Tandem affinity purification (TAP) strategies, which use multiple isolation steps via different tags, are useful for achieving cleaner purifications, leading to the isolation of fewer nonspecific interactions, however, at the expense of weaker interacting partners [19–21]. A variation of TAP tagging can be used, where two proteins that are known to be present in a complex are tagged and purified simultaneously (bimolecular affinity purification), allowing for the specific isolation of a homogeneous population of protein complexes [22]. For all these approaches, one major concern that has to be addressed when using tags for protein IP is whether tagging alters protein function. To verify its functional state, the localization and function (e.g., enzymatic activity) of the tagged protein can be compared against endogenous control (e.g., [2]).

The choice of the affinity resin also has an impact on the success of the IP experiment, influencing the efficiency of the isolation and level of nonspecific interactions. Common choices include sepharose and agarose beads, as well as the steady growth in popularity magnetic beads [22–24]. The surface area of the bead determines not only its capacity for the number of antibody molecules that it can bind to, but also the nonspecific associations to the resin itself. Resins with various chemistries are available for antibody binding (i.e., antibody-binding proteins, primary amines reactive groups, cross-linking to affinity ligands) and can help reduce the amount of eluting immunoglobulin molecules, limiting the interference in MS analysis. In the "Determining specificity of interactions" section in this chapter, the impact of the resin choice on the amount of nonspecific interactions is discussed in detail.

The lysis of the selected cells or tissue is the first step of an IP experiment, and it can impact the preservation of protein–protein interactions. Therefore, the procedure selected for lysis and the composition of the lysis buffers require careful considerations. Mechanical disruption can be performed on wet or frozen samples. For example, cryogenic lysis was shown to provide an effective and reproducible disruption of cellular organelles and membranes, while helping to maintain protein complexes and PTMs [25, 26]. This method is appropriate for different cell types and was successfully applied in studies in bacteria, yeast, mammalian cells, and tissues, as well as following viral infection (as reviewed in [27]). If it is necessary to preserve intact intracellular structures, fractionation steps can be added to the protocol. For example, nuclear-cytoplasmic fractionation was used for assessment of localization-dependent protein–protein interactions of HDAC5 mutated at different phosphorylation sites that regulate its nuclear-cytoplasmic shuttling [1]. Importantly, the stringency of the lysis buffer in an IP experiment determines the nature of isolated interactions. Low salt concentrations and mild detergents may allow preservation of weak interactions, while isolation in a more stringent buffer will enrich for strongly bound interacting partners [28]. Miteva et al. compared the presence of

distinct SIRT6 interactions under mild or stringent lysis conditions, which allowed determining their relative stability, as well as validation of specificity [6]. Addition of sonication step and RN/DNases to the lysis buffer can help remove interactions dependent on nucleic acid binding. It is also important to consider the compatibility of the lysis buffer detergents with the downstream sample analysis by MALDI or ESI/LC-MS/MS. For example, analysis of membrane-bound proteins can be hindered by the necessity to use harsh detergents that are detrimental for MS analysis, although it was shown that *n*-octylglucoside detergent is compatible with MALDI-MS [29]. Another solution is to use cleavable detergents that can be removed from the sample prior to the analysis [30, 31]. Of course, the composition of lysis buffer and duration of lysis has a profound impact on the level of observed nonspecific interacting partners, as discussed in detail in the "Determining specificity of interactions" section.

Following the isolation of the protein of interest, the elution conditions from the beads can also be optimized for different purposes, such as to reduce immunoglobulin contamination, to preserve native protein folding, or to assess the stability of interactions. Most commonly utilized elution buffers are sodium dodecyl sulfate- or lithium dodecyl sulfate-based, which denature the isolated proteins and are suitable for in-solution digest prior to MS analysis [32]. Basic (e.g., ammonium hydroxide and ethylenediaminetetraacetic acid) or acidic (e.g., trichloroacetic acid) elutions are also denaturing, but reduce the amount of background protein contamination [26]. For analysis of native proteins and their complexes, non-denaturing conditions can be used, such as competitive binders [33, 34].

Isolated protein complexes can be analyzed by bottom-up, middle-down, or top-down MS approaches [35–37]. For reduction of sample complexity, protein mixtures can be resolved by gel electrophoresis prior to digestion. Combination of different proteases can be used to improve sequence coverage and identification of PTMs [5]. Separation by liquid chromatography, performed either offline or online with the mass spectrometer, is also used to decrease the sample complexity and provide an in-depth analysis. Different types of peptide fragmentation, such as collision-induced dissociation (CID), electron transfer dissociation (ETD), and higher energy C-trap dissociation (HCD), have also significantly enhanced the current ability to characterize proteins (as reviewed in [38]). Targeted mass spectrometry-based approaches, such as selective reaction monitoring (SRM), further help the identification and quantification of low levels of proteins [39].

11.3 Determining Specificity of Interactions

The complex and dynamic nature of protein–protein interactions coexisting in a cell presents challenges for any IP study. During cell lysis, proteins lose their intracellular localizations, triggering an opportunity for numerous nonspecific interactions to occur. Additionally, nonspecific associations can occur with the resin, tags, or antibodies used for the study. In this section, we discuss some of the sources of

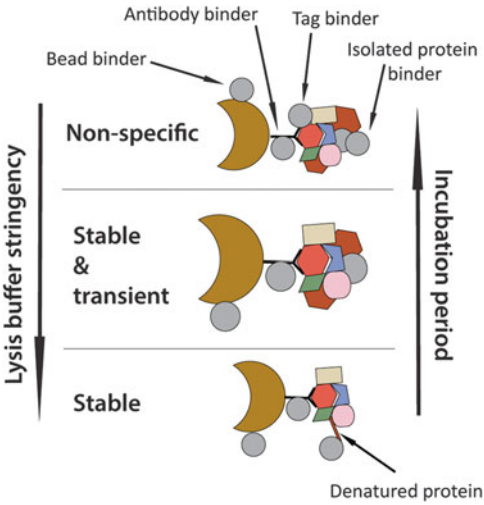


Fig. 11.2 Dependence of interaction specificity on IP conditions. Optimized lysis and incubation conditions, such as stringency of lysis buffer and incubation time with beads and antibodies, allow retention of specific stable and transient interactions, while reducing the number of nonspecific associations. The latter include proteins that bind to beads, tag, immunoglobulin molecules, or to specific isolated proteins

nonspecific interactions and how these can be minimized when designing the IP workflow, as well as approaches used for determining the specificity of observed interactions following data analysis.

Sources of nonspecific interactions. The presence of background proteins in affinity-purified protein mixtures is determined by multiple factors (Fig. 11.2). Nonspecific interactions can include proteins that bind to resin (e.g., magnetic beads), to immunoglobulin molecules, to the tag, and to other isolated proteins. For instance, although polyclonal antibodies have higher affinities and provide more efficient isolations, they also tend to accumulate more nonspecific associations than monoclonal antibodies. As for the choice of the IP resin, it was observed that it can also differentially impact the level and type of nonspecific binders [24, 40]. Sepharose beads seemed to preferentially isolate nonspecific nucleic acid-binding factors, while magnetic beads are prone to association with cytoskeletal and structural proteins [24]. However, magnetic beads can be collected on a magnet, removing the need for a centrifugation step, which reduces sample loss and effectively removes flow-through. Additionally, magnetic beads were preferred for isolating organelles or larger structures or macromolecules, given their feature of surface binding [41, 42].

Proteins that nonspecifically associate with isolated protein complexes are a significant source of background contamination that can be reduced by optimizing IP conditions (Fig. 11.2). For instance, the composition of the lysis buffer and the incubation period with the beads/resin and the antibody greatly influence nonspecific binding. More stringent buffers that contain higher concentrations of salts and detergents can be used to focus on the isolation of strong interactions, helping to reduce nonspecific interactions that tend to be weaker. On the other hand, very stringent buffers that are used to improve extraction of proteins from membranes and intracellular vesicles can lead to protein denaturation, which introduces additional surfaces for nonspecific binding (e.g., to heat shock proteins [43]). In addition, Cristea et al. demonstrated that the length of time used for the incubation of cell lysates with the

beads affects the number and abundance of nonspecific interactions [16]. It was suggested to keep the incubation periods as short as possible, ranging from several minutes up to 1 h, depending on the abundance of the targeted protein and the affinity of the antibody. Therefore, optimization of lysis buffer composition and immunoaffinity purification period allows achieving cleaner isolations, while preserving specific interactions (Fig. 11.2). Additionally, nonspecific interactions can occur upon cell disruption and mixing of proteins from different intracellular compartments. This type of background can be avoided by performing a fractionation step prior to the IP workflow and the lysis/time optimizations mentioned above.

Designing appropriate control experiments and validating interactions. Suitable controls have to accompany every IP experiment because, even with optimized conditions, low levels of nonspecific binders will still be present in the final isolation. Control experiments have to be carefully designed to follow the same conditions as the IP of the protein of interest. For instance, when isolating a tagged protein, a suitable control would be a cell line that expresses only the tag. If the targeted protein is localized within a particular cellular compartment, the control cell line should be designed to afford localization of the tag alone within the same compartment (e.g., by addition of a nuclear localization signal for expression within the nucleus [18]). Similarly, isolation of an endogenous protein requires a control incubation done in parallel using beads coupled to immunoglobulin molecules, which will capture proteins that associate nonspecifically to the antibody. If isolations are performed at different stages of a biological process, such as cell cycle or viral infection, it is necessary to introduce a control experiment for each time point [17]. This will account for variations in the type and abundance of nonspecific interactions throughout the process. Altogether, these control isolations will help differentiate nonspecific associations that occur via the tag, the antibody, and the resin type.

In a collaborative effort, several proteomic laboratories have provided their data from numerous control isolations performed in different cell types and using various resins, tags, and antibodies to generate a repository freely available to the scientific community [40]. This resource, termed the CRAPome database, allows users to determine the frequency of appearance of a protein of interest in control IPs or to analyze their own datasets in comparison to controls available online and assess the presence of common background contaminants. This repository is continuously expanding with new data submitted and processed according to the established workflow.

Validation of isolated interactions is another critical step in all studies aimed at characterizing protein–protein interactions. Such experiments include reciprocal isolations, where an identified prey protein of interest is used as bait in a follow-up IP experiment to confirm the co-isolation of the initial targeted protein. However, it is necessary to keep in mind that reciprocal IPs might prove challenging in identifying a target protein that has low levels of expression. As a consequence, its signal in the prey IP might be suppressed by the presence of more abundant interactions. On the other hand, the prey IP might not be feasible if an antibody is not available for it, but the use of tagging can overcome this problem. Co-localization studies using confocal microscopy or simple binary approaches (e.g., yeast two-hybrid) are also successfully applied for validating interactions [4, 6, 44].

Label-free methods for determining interaction specificity. Qualitative and quantitative data derived from MS analysis of co-isolated proteins contain valuable information regarding the specificity of identified interactions. Therefore, several algorithms were developed for this purpose [49].

Various scoring systems, such as the socio-affinity index and purification enrichment score, were utilized for the analysis of interactions isolated from large-scale IP studies in yeast [45–48]. Their purpose was to decrease the number of identified false-positive interactions, while retaining abundant interactions that are frequently assigned as false-negatives. For example, V-ATPase, an abundant and common contaminant in numerous studies, can be selectively rescued if assigned as a possible specific interaction [45]. Computational approaches were also applied in studies of mammalian interactomes. Among them is the interaction reliability score that was derived for the study of transcription and RNA processing complexes to assign high-confidence interactions [49].

Quantitative data generated by MS analysis in the form of spectral counts (total number of spectra observed per protein) are becoming increasingly utilized for predicting the specificity of interactions [2, 50–52]. In this approach, the number of spectra observed for a particular protein in the bait versus control IP indicates whether this interaction is likely specific for the targeted protein. For further analysis, normalized spectrum abundance factor (NSAF) was introduced by Paoletti et al. to account for the number of amino acids that a protein contains, with larger proteins expected to generate a higher number of spectral counts in MS analysis [53]. When combined with the protein abundance factor (PAX) [54] that reflects total protein abundance in a cell, resulting NSAF/PAX ratio becomes a good indicator of the enrichment of a particular interaction among co-isolated proteins [4]. However, it should be kept in mind that PAX values are not yet derived for all cell types and can change drastically under different environmental conditions or under stress (e.g., during viral infection).

The SAINT (significance analysis of interactome) algorithm was developed by Nesvizhskii et al. to generate a probability model for distributions of false-positives and false-negatives in IP data and to assign confidence scores to identified interactions [52]. For example, SAINT was utilized in a large-scale interactome study of the insulin receptor/target of rapamycin pathway in *Drosophila*, helping the identification of interactions important in controlling cell growth upon stimulation with insulin [55]. More recently, the SAINT algorithm was further optimized to account for the large dynamic range of spectral counts that is frequently observed in human interactomes, such as in the global interaction network of all eleven human histone deacetylases [2]. This study led to the identification of numerous HDAC-containing protein complexes, as well as a previously unrecognized function for HDAC11 in mRNA splicing. The CRAPome database mentioned above also utilized the SAINT algorithm for analyzing the collection of control IPs derived from different cell types and performed in various laboratories [40]. Another spectral counting-based program, called CompPASS, uses several scoring systems to derive confidence scores for interactions found in multiple parallel nonreciprocal IPs, without the use of control IPs [51]. Algorithms that utilize other MS data, such as MS1 signals (MasterMap)

and peak intensities (MiST), are also being developed with the goal of overcoming some of the limitations of spectral counting approaches, such as dependence of interaction abundances on bait and prey levels, efficiency of IP, and detection by MS [56, 57]. A more complete summary of available algorithms is reported in [27].

The label-free approaches mentioned above have several advantages over labeling approaches that will be discussed in the next section. They do not require expensive reagents, can be used for the analysis of tissue samples, and can be applied in both small- and large-scale studies. However, these approaches have certain limitations. For example, protein abundance has a significant impact on the assessment of specificity (i.e., cannot reliably quantify changes for proteins with low spectral counts). Additionally, as is the case for most methods, these approaches cannot fully segregate background proteins from specific interactions within the multitude of co-isolated proteins.

Labeling methods for determining interaction specificity. Metabolic and chemical labeling approaches were introduced into MS analysis workflows to provide absolute or relative protein quantification. During the last decade, the application of these approaches within targeted or global studies has revolutionized the field of proteomics and its ability to contribute to critical biological discoveries. These labeling methods have certain limitations, such as their challenging application to tissue samples and variations in sample processing prior to chemical labeling. Nevertheless, in recent years, the application of these approaches was expanded to include their incorporation into IP workflows for the downstream analysis of interaction specificity.

Labeling with stable isotopes, such as ^{15}N, and later with heavy amino acids in cell culture (SILAC) were among the first metabolic labeling methods to be introduced within mass spectrometry-based workflows [58, 59]. For identification of interaction specificity using metabolic labeling, Chait and colleagues developed the I-DIRT (Isotopic Differentiation of Interactions as Random or Targeted) approach and applied it to the study of DNA polymerase ∈ complex in yeast [60]. In their workflow, cells expressing the affinity-tagged protein are grown in medium containing naturally occurring amino acids, termed isotopically light medium. In contrast, wild-type cells are grown in isotopically heavy medium that contains amino acids labeled with heavy isotopes (e.g., ^{13}C). Proteins are immunoaffinity purified from a 1:1 mixture of these light and heavy cell lysates, and the specific interacting partners can be recognized as having only or predominantly light isotopic peaks (Fig. 11.3a). An "SRM-like" I-DIRT approach, in which the specificity of interaction for selected proteins of interest can be assessed using targeted MS/MS, may be utilized to analyze low abundance interactions [4]. To assess interaction specificity in studies of endogenous protein complexes, QUICK (quantitative immunoprecipitation combined with knockdown) strategy was developed [61–64]. In this workflow, light-labeled cell cultures are treated with RNAi against the protein of interest, while heavy-labeled cells serve as nontargeted controls. In subsequent MS analysis, light and heavy peptide intensities are compared to assign nonspecific (1:1 heavy to light rations) and specific (heavy isotopic peaks with higher intensity than light) interactions.

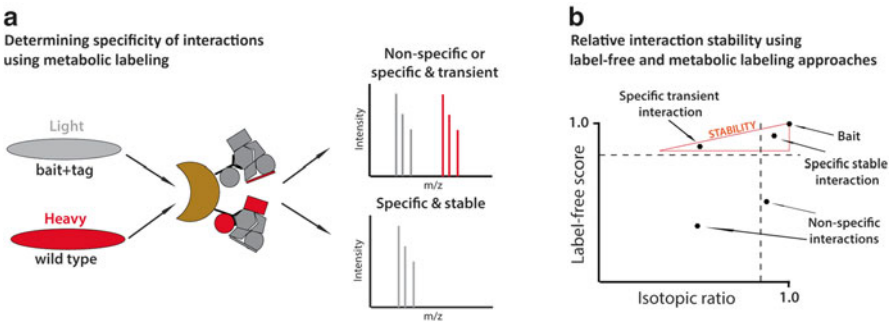


Fig. 11.3 Determining specificity and relative stability of interactions using label-free and metabolic labeling approaches. (**a**) In the I-DIRT approach, wild-type cells grown in "heavy" medium and cells expressing the tagged protein of interest grown in "light" medium are mixed prior to IP. Isolated complexes are analyzed by MS and isotopic ratios for each protein are indicative of the specificity and stability of the interaction. (**b**) When label-free quantification (e.g., SAINT) is combined with a metabolic labeling approach (e.g., I-DIRT), the relative stability of interactions can be assessed. Specific transient interactions can be observed with high SAINT scores and low I-DIRT ratios

QUICK can also be combined with cross-linking to stabilize protein complexes in cell extracts prior to IP, as demonstrated for VIPP1 complex functioning in chloroplast biogenesis [65]. Some of the disadvantages of the QUICK approach include SILAC-associated costs, arginine-to-proline conversions that produce difficulties in data interpretation, and uncontrollable alterations in protein expression due to RNAi knockdown. These concerns were addressed in an alternative QUICK method utilizing ^{15}N labeling and affinity modulation of protein–protein interactions [66]. SILAC approaches in the form of PAM (purification after mixing) SILAC and MAP (mixing after purification) SILAC were also used to assess the interaction specificity in several studies [67, 68]. These methods allow distinguishing between stable and transient interactions, which will be further discussed in the next section. Affinity purification of integral membrane proteins using nanodiscs, which circumvent the need for stringent detergents, was combined with SILAC in the analysis of interacting partners of bacterial channel, transporter, and integrase proteins [69]. In the study of phosphatase and tensin homologue (PTEN) interactions, IPs from two different cell lines using three different approaches (two tags and one endogenous) were combined in parallel affinity purification (PAP) SILAC approach to assign specific interactions with minimum number of false positives [70].

Chemical labeling was also successfully applied for the analysis of protein interactions and is typically done after IP, at either the protein or peptide level. In ICAT (isotope-coded affinity tag) approach, cysteine residues of intact proteins in bait and control experiments are labeled with heavy or light ICAT reagents, respectively [71, 72]. Therefore, specific interaction partners would produce peptide spectra with higher intensities for heavy isotope-containing peaks. Inherent to this method is the quantification of only cysteine-containing peptides. The iTRAQ multiplex labeling method [73], which tags peptides at N-terminal and lysine amines, was also applied

to distinguish true interactions within co-isolated protein mixtures when comparing differentially labeled bait and control IPs, as shown for grb2 [74]. Isotope-coded protein labeling (ICPL) was also recently combined with IP-MS to analyze native β-tubulin complexes in bovine retinal tissue [75]. Besides being compatible with analysis of tissue samples, this method allows for simultaneous analysis of several control samples, which can account for nonspecific binders to the beads and immunoglobulin molecules. Currently, chemical labeling is not as widely applied as metabolic labeling to the analysis of interaction specificity. However, its compatibility with tissue samples and its multiplexing feature that allows direct comparison of multiple samples and diverse controls makes it a useful tool, expected to continue to aid the discovery of protein complex compositions relevant in biomedical research.

11.4 Cross-Linking Methods

In MS studies of protein complexes, cross-linking is used for covalent joining of two or more molecules. Chemical cross-linking reagents have various chemistries (e.g., amino-, sulfhydryl-, carboxyl-reactive), sizes that determine the distance between cross-linked peptides, and add-on features for easier detection and identification (e.g., reversible cross-linkers). In this section, the use of cross-linking in studies of protein complex structures is discussed, while its application to capturing transient protein interactions is described in the section on "Determining stability of interactions".

Solving protein complex structures. Continuous advances in cross-linking methodologies, the improved sensitivity of MS instrumentation, and the development of automated algorithms for database searching have significantly expanded the use of cross-linking for characterization of protein complexes. Examples of elegant structures resolved using cross-linking methodologies include the yeast 19S proteasome lid, RNA Pol II, phage DNA packaging machinery, human protein phosphatase 2A (PP2A) complexes, INO80 nucleosome complex, TRiC/CCT chaperonin to name a few [76–80]. These cross-linking strategies are powerful at defining the exact points of contacts between proteins of interest. One caveat to keep in mind is that proteins can be part of multiple protein complexes that can have common components. For example, the histone deacetylases HDAC1 and HDAC2 are known to form the core of several distinct complexes, such as NuRD and Sin3A complexes [2]. Therefore, even after identifying the point of contact between two proteins, one may not know where the interaction takes place, i.e., which complex or which conformational state of the complex is represented. To partly address this issue, cross-linking is frequently combined with knowledge from X-ray crystallography studies. Nevertheless, as crystallography results usually reflect a more static or stable conformation of a protein, these issues should still be kept in mind.

Cross-linking was also integrated with other biochemical or mass spectrometry tools to help define protein structures and interactions within macromolecular

assemblies. For example, a combination of cross-linking with hydrogen/deuterium exchange was used to decipher inter-subunit interactions critical in the assembly of HIV-1 capsid protein, which complemented studies performed using X-ray crystallography and cryo-electron microscopy [81]. In another study, the folding of the immune receptor NKR-P1C was resolved using cross-linking, molecular modeling, and ion mobility mass spectrometry [82].

One common limitation of cross-linking when combined with MS analysis is the possible low abundance of resulting cross-linked peptides, which can make spectra interpretation challenging. To solve this problem for the analysis of the 12-subunit Pol II complex structure, Chen et al. utilized a strong cation exchange chromatography to enrich for cross-linked peptides which have multiple charges [76]. For isotopic labeling, Zelter et al. digested cross-linked peptides in the presence of $H_2{}^{18}O$, which were then identified by their characteristic isotopic peak distribution in MS spectra [83]. Several strategies were proposed that utilize modifications of the cross-linker itself to provide easier enrichment, detection, and identification. These include affinity tags, reporter tags, isotopic and fluorescence labeling, and cleavable cross-linking, all of which can be used in combination [84]. For example, Chowdhury et al. combined an alkyne enrichment tag and NO_2 detection tag when constructing a CLIP (click-enabled linker for interacting proteins) cross-linking reagent [85]. For the study of the 20S proteasome complex in yeast cells, Kao et al. designed a disuccinimidyl sulfoxide (DSSO) cross-linker that is cleavable by collision-induced dissociation and can be identified at the MS^3 level [86].

The expansion in types of cross-linking strategies and applications places demands on the developments of streamlined procedures for MS data interpretation. To make this process more automated, Herzog et al. utilized specialized xQuest search engine [79, 87]. In this workflow, isotopic pairs of cross-linked peptide ions were matched against a database of candidate peptides, upon which their sequences were assigned. This algorithm was used to solve the structures of human PP2A, INO80 nucleosome, TRiC/CCT eukaryotic chaperonin, and other complexes [79, 80, 88, 89]. However, this process has a laborious scoring procedure that requires manual verification of the cross-linked peptide spectra. In addition, reliable identification of isotopically labeled cross-linked peptides in this method can suffer from incomplete labeling. To overcome this limitation, Goodlett and coworkers developed an alternative cross-linking strategy that uses Popitam search engine [90] and can identify unlabeled cross-linkers [91]. In their workflow, cross-linked peptides are considered as complementary pairs of peptides modified by an unknown mass. The spectra are interpreted by matching to theoretical spectra of single linear peptides, and further analyzed against the masses of precursor tryptic peptides and manually validated. SEQUEST [92] searching was also further optimized for the identification of cross-linked peptides from a database containing all possible products of cross-linking, which allowed matching complex spectra of cross-linked peptide pairs more efficiently [93]. For automatic validation of database search results, Walzthoeni et al. introduced the xProphet software that uses a target-decoy strategy to estimate false discovery rates in large datasets derived from cross-linking studies [94]. Many other database processing algorithms and bioinformatics tools are being continuously

developed and released to address challenges associated with deciphering complex cross-linked peptide spectra [95–100]. Overall, further improvements in cross-linking strategies for easier detection and identification, as well as in the software for analysis of MS spectra of cross-linked peptides are required. However, the combination of cross-linking strategies with crystallography studies, computational modeling, and other quantitative mass spectrometry methods provides powerful approaches in proteomics that will continue to shed light on the structures of heterogeneous protein complexes.

11.5 Studying Transient and Fast-Exchanging Interactions

Coupling of efficient IP strategies with highly sensitive MS analysis leads to identification of numerous interactions, direct and indirect, transient and stable, which provide valuable information about protein function. Several methods can be used to distinguish between direct and indirect interactions, such as yeast two-hybrid and protein arrays [101]. These methods do not utilize mass spectrometry analysis and are not discussed in this review. The identification of transient and fast-exchanging interactions, such as enzyme-substrate interactions, presents challenges in mass-spectrometry proteomic workflows. These interactions can be either lost during the IP process or can be falsely assigned as nonspecific in metabolic labeling experiments. Therefore, methods are continuously being developed to help the capture and identification of these interactions in MS-based experiments.

Determining interaction stability using metabolic labeling. As mentioned in the earlier section, time-controlled PAM SILAC and MAP SILAC approaches have also been used for identification of specific interactions that are dynamic in nature [102]. In PAM SILAC approach, samples are mixed prior to purification, which allows for transient interactions to exchange quickly between light and heavy forms, resulting in equivalent levels of heavy and light ions. However, with decreased incubation time during purification, the level of heavy ions will increase. Moreover, if mixing of heavy and light-labeled samples is done after purification (MAP SILAC), same interactions will have predominant heavy ions because there will be no "light" labeled proteins present in the isolated sample. The MAP SILAC approach also allows for identification of fast-exchanging interactions, which would require short incubation times in order to be accurately assigned as specific when using the PAM SILAC approach. This approach was applied in studies of 26S proteasome and COP9 signalosome complexes [67, 68].

A combination of spectral counting (SAINT) and metabolic labeling (I-DIRT) approaches was utilized by Joshi et al. to measure relative interaction stabilities within HDAC-containing protein complexes (Fig. 11.3) [2]. In their workflow, unlabeled bait samples were analyzed against control IPs to generate a list of interactions using SAINT scores, with scores >0.8 reflective of likely specific interactions. In parallel experiments, metabolic labeling of cells expressing tagged HDACs was performed, and an I-DIRT approach was used as described above. By integrating SAINT

and I-DIRT scores for each isolated protein, stability profiles could be assigned to specific interacting partners. Proteins with >0.8 SAINT scores and ~0.5 I-DIRT scores corresponded to fast-exchanging proteins, while proteins with >0.8 SAINT scores and I-DIRT scores closer to 1.0 indicated stable interactions. For example, HDAC5 and HDAC7 interact transiently with the NCoR complex due to their nucleocytoplasmic shuttling, while HDAC3 is a stable component of this complex. Similarly, HDAC1 was shown to be an integral component of several chromatin remodeling complexes, while transiently associating with transcription factors and DNA-binding proteins. Therefore, this approach allows confident identification of novel specific interacting partners and assignment to transient or stable associations.

Detecting transient interactions using cross-linking. Several cross-linking methods were incorporated into IP-MS workflows to study stable and transient interactions in cell culture. The main requirement for a cross-linking reagent to be used for such analysis is its cell permeability. One of the most widely utilized reagents in these studies is formaldehyde. For example, TAP of formaldehyde cross-linked SCF ubiquitin ligase complex under denaturing conditions was utilized by Tagwerker et al. to preserve and characterize novel ubiquitination targets, as well as identify transient or weak interacting partners [103]. For a more quantitative analysis, Guerrero et al. combined TAP, formaldehyde cross-linking, and SILAC approaches to characterize 26S proteasome interactions in yeast [104]. Zero distance cross-linking using photoinducible amino acids [105] introduced into growing mammalian cells allowed identification of a direct interaction between endoplasmic reticulum stress protein MANF and GRFP78 that regulates stress-induced cell death [106]. One of the disadvantages of irreversible cross-linking is that, following the immunoaffinity purification of a protein complex, there is a low accessibility for trypsin at the core of the isolated complex, hindering the identification of selected proteins by MS. To resolve this issue, as well as other challenges connected to irreversible cross-linking, numerous studies employ reversible cross-linking. For example, reversion of formaldehyde cross-links was used in the SPINE method for detection of interacting partners of Strep-tagged membrane proteins in bacteria [107]. Another reversible cross-linking methodology—ReCLIP—utilizes thiol-cleavable cross-links and was used in the study of p120-catenin and E-cadherin complex [108, 109].

To capture transient interactions and address specificity of interactions within the same experiment, a transient I-DIRT approach was reported. This approach used cross-linking combined with isotopic labeling in yeast culture and was applied to the study of NuA3 multi-subunit complex [110]. In this workflow, cells expressing the tagged protein are grown in light media, while wild-type cells are grown in heavy media. Upon mixing cross-linked heavy and light cell cultures and purifying the target protein, MS analysis is performed and used to assign stable specific (~100 % light peptides), nonspecific (1:1 light:heavy peptide ratio), and transient (intermediate ratios) interacting partners [111].

Recent years have also seen significant developments in cross-linking reagents. To overcome the challenge of identifying cross-linked peptides in the mixed spectra generated from a complex mixture of cross-linked proteins with various

intermediate products, Bruce laboratory developed Protein Interaction Reporter (PIR) technology [112]. In their methodology, the cross-linking reagent is designed to contain two labile bonds that can be cleaved during MS/MS analysis. Upon cleavage, a reporter ion is released to mark the presence of a cross-linked peptide, while cleavage of the second labile bond generates single peptides from the cross-linked pair for further fragmentation and sequencing. This technology was utilized in the study of the Potato leafroll virus capsid structure and in defining interactions of bacterial chaperones and membrane proteins [113–115].

Cross-linking with formaldehyde has also been utilized in mouse models, where the reagent was introduced via transcardiac perfusion in a time-controlled manner [116]. This method was applied in combination with isotopic labeling with iTRAQ to assign interaction specificity in studies of cellular prion protein (PrP^c), oxidative stress sensor DJ-1, and amyloid precursor protein interactomes [116–119]. In addition to identifying specific interactions of PrP^c, Watts et al. also suggested that information derived from the MS/MS analysis of cross-linked proteins could be used to distinguish direct and indirect interactions [118]. For instance, proteins that have high sequence coverage and share similar domain structures were most likely to represent direct interacting partners.

The task of identifying transient, stable, direct, and indirect interactions is not trivial. However, metabolic labeling and cross-linking approaches incorporated into IP-MS workflows discussed above have significantly aided these studies. There is no doubt that studies of protein–protein interactions and resulting macromolecular complexes will continue to expand our understanding of critical biological processes. Further methodological developments are needed. Approaches that specifically capture one moment in a cellular process, temporally and spatially defined, are continuously being developed and improved. While studies in cell systems provide simple models with extraordinary specificity and insight into concrete cellular pathways, expansion of interaction studies to animal models allows for in vivo validation and a systems view of the changes caused by perturbations in a single protein functions. As protein interactions are at the core of cellular, tissue, and organ functions, their study will continue to shed light onto fundamental questions in both basic science and clinical research.

Acknowledgments We are grateful for funding from NIH grants DP1DA026192, R21AI102187, and R21 HD073044-01A1, an HFSPO award RGY0079/2009-C to IMC, and an NSF graduate fellowship to HGB.

References

1. Greco TM et al (2011) Nuclear import of histone deacetylase 5 by requisite nuclear localization signal phosphorylation. Mol Cell Proteomics 10(2):M110 004317
2. Joshi P et al (2013) The functional interactome landscape of the human histone deacetylase family. Mol Syst Biol 9:672

3. Malovannaya A et al (2011) Analysis of the human endogenous coregulator complexome. Cell 145(5):787–799
4. Tsai YC et al (2012) Functional proteomics establishes the interaction of SIRT7 with chromatin remodeling complexes and expands its role in regulation of RNA polymerase I transcription. Mol Cell Proteomics 11(5):60–76
5. Li T et al (2012) Acetylation modulates cellular distribution and DNA sensing ability of interferon-inducible protein IFI16. Proc Natl Acad Sci U S A 109(26):10558–10563
6. Miteva YV, Cristea IM (2014) A proteomic perspective of SIRT6 phosphorylation and interactions, and their dependence on its catalytic activity. Mol Cell Proteomics 13(1):168–183
7. Adam GC, Sorensen EJ, Cravatt BF (2002) Chemical strategies for functional proteomics. Mol Cell Proteomics 1(10):781–790
8. Salisbury CM, Cravatt BF (2007) Activity-based probes for proteomic profiling of histone deacetylase complexes. Proc Natl Acad Sci U S A 104(4):1171–1176
9. Bantscheff M et al (2011) Chemoproteomics profiling of HDAC inhibitors reveals selective targeting of HDAC complexes. Nat Biotechnol 29(3):255–265
10. Cen Y et al (2011) Mechanism-based affinity capture of sirtuins. Org Biomol Chem 9(4):987–993
11. Ruigrok VJ et al (2011) Alternative affinity tools: more attractive than antibodies? Biochem J 436(1):1–13
12. Gronwall C, Stahl S (2009) Engineered affinity proteins—generation and applications. J Biotechnol 140(3–4):254–269
13. Brody E et al (2012) Life's simple measures: unlocking the proteome. J Mol Biol 422(5):595–606
14. Wiens M et al (2011) Isolation of the silicatein-alpha interactor silintaphin-2 by a novel solid-phase pull-down assay. Biochemistry 50(12):1981–1990
15. Hubner NC, Mann M (2011) Extracting gene function from protein-protein interactions using Quantitative BAC InteraCtomics (QUBIC). Methods 53(4):453–459
16. Cristea IM et al (2005) Fluorescent proteins as proteomic probes. Mol Cell Proteomics 4(12):1933–1941
17. Cristea IM et al (2006) Tracking and elucidating alphavirus-host protein interactions. J Biol Chem 281(40):30269–30278
18. Li T, Chen J, Cristea IM (2013) Human cytomegalovirus tegument protein pUL83 inhibits IFI16-mediated DNA sensing for immune evasion. Cell Host Microbe 14(5):591–599
19. Li Y (2010) Commonly used tag combinations for tandem affinity purification. Biotechnol Appl Biochem 55(2):73–83
20. Gavin AC, Maeda K, Kuhner S (2011) Recent advances in charting protein-protein interaction: mass spectrometry-based approaches. Curr Opin Biotechnol 22(1):42–49
21. Rees JS et al (2011) In vivo analysis of proteomes and interactomes using parallel affinity capture (iPAC) coupled to mass spectrometry. Mol Cell Proteomics 10(6):M110 002386
22. Maine GN et al (2009) Bimolecular affinity purification (BAP): tandem affinity purification using two protein baits. Cold Spring Harb Protoc 2009(11):pdb prot5318
23. Ayyar BV et al (2012) Affinity chromatography as a tool for antibody purification. Methods 56(2):116–129
24. Trinkle-Mulcahy L et al (2008) Identifying specific protein interaction partners using quantitative mass spectrometry and bead proteomes. J Cell Biol 183(2):223–239
25. Archambault V et al (2003) Genetic and biochemical evaluation of the importance of Cdc6 in regulating mitotic exit. Mol Biol Cell 14(11):4592–4604
26. Cristea IM, Chait BT (2011) Affinity purification of protein complexes. Cold Spring Harb Protoc 2011(5):pdb prot5611
27. Miteva YV, Budayeva HG, Cristea IM (2013) Proteomics-based methods for discovery, quantification, and validation of protein-protein interactions. Anal Chem 85(2):749–768
28. Conlon FL et al (2012) Immunoisolation of protein complexes from Xenopus. Methods Mol Biol 917:369–390

29. Cadene M, Chait BT (2000) A robust, detergent-friendly method for mass spectrometric analysis of integral membrane proteins. Anal Chem 72(22):5655–5658
30. Norris JL, Porter NA, Caprioli RM (2003) Mass spectrometry of intracellular and membrane proteins using cleavable detergents. Anal Chem 75(23):6642–6647
31. Ye X et al (2009) Optimization of protein solubilization for the analysis of the CD14 human monocyte membrane proteome using LC-MS/MS. J Proteomics 73(1):112–122
32. Wisniewski JR et al (2009) Universal sample preparation method for proteome analysis. Nat Methods 6(5):359–362
33. Darie CC et al (2011) Identifying transient protein-protein interactions in EphB2 signaling by blue native PAGE and mass spectrometry. Proteomics 11(23):4514–4528
34. Dubois F et al (2009) Differential 14-3-3 affinity capture reveals new downstream targets of phosphatidylinositol 3-kinase signaling. Mol Cell Proteomics 8(11):2487–2499
35. Chait BT (2011) Mass spectrometry in the postgenomic era. Annu Rev Biochem 80:239–246
36. Tipton JD et al (2011) Analysis of intact protein isoforms by mass spectrometry. J Biol Chem 286(29):25451–25458
37. Aebersold R, Mann M (2003) Mass spectrometry-based proteomics. Nature 422(6928):198–207
38. Yates JR, Ruse CI, Nakorchevsky A (2009) Proteomics by mass spectrometry: approaches, advances, and applications. Annu Rev Biomed Eng 11:49–79
39. Picotti P, Aebersold R (2012) Selected reaction monitoring-based proteomics: workflows, potential, pitfalls and future directions. Nat Methods 9(6):555–566
40. Mellacheruvu D et al (2013) The CRAPome: a contaminant repository for affinity purification-mass spectrometry data. Nat Methods 10(8):730–736
41. Alber F et al (2007) Determining the architectures of macromolecular assemblies. Nature 450(7170):683–694
42. Selimi F et al (2009) Proteomic studies of a single CNS synapse type: the parallel fiber/purkinje cell synapse. PLoS Biol 7(4):e83
43. Hendrick JP, Hartl FU (1993) Molecular chaperone functions of heat-shock proteins. Annu Rev Biochem 62:349–384
44. Bard-Chapeau EA et al (2013) EVI1 oncoprotein interacts with a large and complex network of proteins and integrates signals through protein phosphorylation. Proc Natl Acad Sci U S A 110(31):E2885–E2894
45. Gavin AC et al (2006) Proteome survey reveals modularity of the yeast cell machinery. Nature 440(7084):631–636
46. Krogan NJ et al (2006) Global landscape of protein complexes in the yeast Saccharomyces cerevisiae. Nature 440(7084):637–643
47. Collins SR et al (2007) Toward a comprehensive atlas of the physical interactome of Saccharomyces cerevisiae. Mol Cell Proteomics 6(3):439–450
48. Babu M et al (2012) Interaction landscape of membrane-protein complexes in Saccharomyces cerevisiae. Nature 489(7417):585–589
49. Jeronimo C et al (2007) Systematic analysis of the protein interaction network for the human transcription machinery reveals the identity of the 7SK capping enzyme. Mol Cell 27(2):262–274
50. Sardiu ME et al (2008) Probabilistic assembly of human protein interaction networks from label-free quantitative proteomics. Proc Natl Acad Sci U S A 105(5):1454–1459
51. Sowa ME et al (2009) Defining the human deubiquitinating enzyme interaction landscape. Cell 138(2):389–403
52. Choi H et al (2011) SAINT: probabilistic scoring of affinity purification-mass spectrometry data. Nat Methods 8(1):70–73
53. Paoletti AC et al (2006) Quantitative proteomic analysis of distinct mammalian mediator complexes using normalized spectral abundance factors. Proc Natl Acad Sci U S A 103(50): 18928–18933

54. Wang M et al (2012) PaxDb, a database of protein abundance averages across all three domains of life. Mol Cell Proteomics 11(8):492–500
55. Glatter T et al (2011) Modularity and hormone sensitivity of the Drosophila melanogaster insulin receptor/target of rapamycin interaction proteome. Mol Syst Biol 7:547
56. Rinner O et al (2007) An integrated mass spectrometric and computational framework for the analysis of protein interaction networks. Nat Biotechnol 25(3):345–352
57. Jager S et al (2012) Global landscape of HIV-human protein complexes. Nature 481(7381): 365–370
58. Oda Y et al (1999) Accurate quantitation of protein expression and site-specific phosphorylation. Proc Natl Acad Sci U S A 96(12):6591–6596
59. Ong SE et al (2002) Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. Mol Cell Proteomics 1(5):376–386
60. Tackett AJ et al (2005) I-DIRT, a general method for distinguishing between specific and nonspecific protein interactions. J Proteome Res 4(5):1752–1756
61. Selbach M, Mann M (2006) Protein interaction screening by quantitative immunoprecipitation combined with knockdown (QUICK). Nat Methods 3(12):981–983
62. Ge F et al (2010) Identification of novel 14-3-3zeta interacting proteins by quantitative immunoprecipitation combined with knockdown (QUICK). J Proteome Res 9(11): 5848–5858
63. Meixner A et al (2011) A QUICK screen for Lrrk2 interaction partners—leucine-rich repeat kinase 2 is involved in actin cytoskeleton dynamics. Mol Cell Proteomics 10(1):M110 001172
64. Zheng P et al (2012) QUICK identification and SPR validation of signal transducers and activators of transcription 3 (Stat3) interacting proteins. J Proteomics 75(3):1055–1066
65. Heide H et al (2009) Application of quantitative immunoprecipitation combined with knockdown and cross-linking to Chlamydomonas reveals the presence of vesicle-inducing protein in plastids 1 in a common complex with chloroplast HSP90C. Proteomics 9(11):3079–3089
66. Schmollinger S et al (2012) A protocol for the identification of protein-protein interactions based on 15 N metabolic labeling, immunoprecipitation, quantitative mass spectrometry and affinity modulation. J Vis Exp 67(2):4083
67. Wang X, Huang L (2008) Identifying dynamic interactors of protein complexes by quantitative mass spectrometry. Mol Cell Proteomics 7(1):46–57
68. Fang L et al (2008) Characterization of the human COP9 signalosome complex using affinity purification and mass spectrometry. J Proteome Res 7(11):4914–4925
69. Zhang XX et al (2012) Nanodiscs and SILAC-based mass spectrometry to identify a membrane protein interactome. J Proteome Res 11(2):1454–1459
70. Gunaratne J et al (2011) Protein interactions of phosphatase and tensin homologue (PTEN) and its cancer-associated G20E mutant compared by using stable isotope labeling by amino acids in cell culture-based parallel affinity purification. J Biol Chem 286(20):18093–18103
71. Gygi SP et al (1999) Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. Nat Biotechnol 17(10):994–999
72. Ranish JA, Brand M, Aebersold R (2007) Using stable isotope tagging and mass spectrometry to characterize protein complexes and to detect changes in their composition. Methods Mol Biol 359:17–35
73. Ross PL et al (2004) Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. Mol Cell Proteomics 3(12):1154–1169
74. Zieske LR (2006) A perspective on the use of iTRAQ reagent technology for protein complex and profiling studies. J Exp Bot 57(7):1501–1508
75. Vogt A et al (2013) Isotope coded protein labeling coupled immunoprecipitation (ICPL-IP): a novel approach for quantitative protein complex analysis from native tissue. Mol Cell Proteomics 12(5):1395–1406
76. Chen ZA et al (2010) Architecture of the RNA polymerase II-TFIIF complex revealed by cross-linking and mass spectrometry. EMBO J 29(4):717–726
77. Sharon M et al (2006) Structural organization of the 19S proteasome lid: insights from MS of intact complexes. PLoS Biol 4(8):e267

78. Fu CY et al (2010) A docking model based on mass spectrometric and biochemical data describes phage packaging motor incorporation. Mol Cell Proteomics 9(8):1764–1773
79. Herzog F et al (2012) Structural probing of a protein phosphatase 2A network by chemical cross-linking and mass spectrometry. Science 337(6100):1348–1352
80. Leitner A et al (2012) The molecular architecture of the eukaryotic chaperonin TRiC/CCT. Structure 20(5):814–825
81. Lanman J et al (2003) Identification of novel interactions in HIV-1 capsid protein assembly by high-resolution mass spectrometry. J Mol Biol 325(4):759–772
82. Rozbesky D et al (2013) Structural model of lymphocyte receptor NKR-P1C revealed by mass spectrometry and molecular modeling. Anal Chem 85(3):1597–1604
83. Zelter A et al (2010) Isotope signatures allow identification of chemically cross-linked peptides by mass spectrometry: a novel method to determine inter-residue distances in protein structures through cross-linking. J Proteome Res 9(7):3583–3589
84. Paramelle D et al (2013) Chemical cross-linkers for protein structure studies by mass spectrometry. Proteomics 13(3–4):438–456
85. Chowdhury SM et al (2009) Identification of cross-linked peptides after click-based enrichment using sequential collision-induced dissociation and electron transfer dissociation tandem mass spectrometry. Anal Chem 81(13):5524–5532
86. Kao A et al (2011) Development of a novel cross-linking strategy for fast and accurate identification of cross-linked peptides of protein complexes. Mol Cell Proteomics 10(1):M110 002212
87. Rinner O et al (2008) Identification of cross-linked peptides from large sequence databases. Nat Methods 5(4):315–318
88. Tosi A et al (2013) Structure and subunit topology of the INO80 chromatin remodeler and its nucleosome complex. Cell 154(6):1207–1219
89. Jennebach S et al (2012) Crosslinking-MS analysis reveals RNA polymerase I domain architecture and basis of rRNA cleavage. Nucleic Acids Res 40(12):5591–5601
90. Hernandez P et al (2003) Popitam: towards new heuristic strategies to improve protein identification from tandem mass spectrometry data. Proteomics 3(6):870–878
91. Singh P et al (2008) Characterization of protein cross-links via mass spectrometry and an open-modification search strategy. Anal Chem 80(22):8799–8806
92. Eng JK, McCormack AL, Yates JR (1994) An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. J Am Soc Mass Spectrom 5(11):976–989
93. McIlwain S et al (2010) Detecting cross-linked peptides by searching against a database of cross-linked peptide pairs. J Proteome Res 9(5):2488–2495
94. Walzthoeni T et al (2012) False discovery rate estimation for cross-linked peptides identified by mass spectrometry. Nat Methods 9(9):901–903
95. Panchaud A et al (2010) xComb: a cross-linked peptide database approach to protein-protein interaction analysis. J Proteome Res 9(5):2508–2515
96. Maiolica A et al (2007) Structural analysis of multiprotein complexes by cross-linking, mass spectrometry, and database searching. Mol Cell Proteomics 6(12):2200–2211
97. Chu F et al (2010) Finding chimeras: a bioinformatics strategy for identification of cross-linked peptides. Mol Cell Proteomics 9(1):25–31
98. Lee YJ et al (2007) Shotgun cross-linking analysis for studying quaternary and tertiary protein structures. J Proteome Res 6(10):3908–3917
99. Nadeau OW et al (2008) CrossSearch, a user-friendly search engine for detecting chemically cross-linked peptides in conjugated proteins. Mol Cell Proteomics 7(4):739–749
100. Yang B et al (2012) Identification of cross-linked peptides from complex samples. Nat Methods 9(9):904–906
101. Braun P (2012) Interactome mapping for analysis of complex phenotypes: insights from benchmarking binary interaction assays. Proteomics 12(10):1499–1518
102. Kaake RM, Wang X, Huang L (2010) Profiling of protein interaction networks of protein complexes using affinity purification and quantitative mass spectrometry. Mol Cell Proteomics 9(8):1650–1665

103. Tagwerker C et al (2006) A tandem affinity tag for two-step purification under fully denaturing conditions: application in ubiquitin profiling and protein complex identification combined with in vivo cross-linking. Mol Cell Proteomics 5(4):737–748
104. Guerrero C et al (2006) An integrated mass spectrometry-based proteomic approach: quantitative analysis of tandem affinity-purified in vivo cross-linked protein complexes (QTAX) to decipher the 26 S proteasome-interacting network. Mol Cell Proteomics 5(2):366–378
105. Suchanek M, Radzikowska A, Thiele C (2005) Photo-leucine and photo-methionine allow identification of protein-protein interactions in living cells. Nat Methods 2(4):261–267
106. Glembotski CC et al (2012) Mesencephalic astrocyte-derived neurotrophic factor protects the heart from ischemic damage and is selectively secreted upon sarco/endoplasmic reticulum calcium depletion. J Biol Chem 287(31):25893–25904
107. Muller VS et al (2011) Membrane-SPINE: an improved method to identify protein-protein interaction partners of membrane proteins in vivo. Proteomics 11(10):2124–2128
108. Smith AL et al (2012) Association of Rho-associated protein kinase 1 with E-cadherin complexes is mediated by p120-catenin. Mol Biol Cell 23(1):99–110
109. Smith AL et al (2011) ReCLIP (reversible cross-link immuno-precipitation): an efficient method for interrogation of labile protein complexes. PLoS One 6(1):e16206
110. Smart SK et al (2009) Mapping the local protein interactome of the NuA3 histone acetyltransferase. Protein Sci 18(9):1987–1997
111. Byrum S et al (2012) Analysis of stable and transient protein-protein interactions. Methods Mol Biol 833:143–152
112. Tang X, Bruce JE (2010) A new cross-linking strategy: protein interaction reporter (PIR) technology for protein-protein interaction studies. Mol Biosyst 6(6):939–947
113. Chavez JD et al (2012) Cross-linking measurements of the Potato leafroll virus reveal protein interaction topologies required for virion stability, aphid transmission, and virus-plant interactions. J Proteome Res 11(5):2968–2981
114. Zheng C et al (2011) Cross-linking measurements of in vivo protein complex topologies. Mol Cell Proteomics 10(10):M110 006841
115. Zhang H et al (2009) Identification of protein-protein interactions and topologies in living cells with chemical cross-linking and mass spectrometry. Mol Cell Proteomics 8(3):409–420
116. Schmitt-Ulms G et al (2004) Time-controlled transcardiac perfusion cross-linking for the study of protein interactions in complex tissues. Nat Biotechnol 22(6):724–731
117. Bai Y et al (2008) The in vivo brain interactome of the amyloid precursor protein. Mol Cell Proteomics 7(1):15–34
118. Watts JC et al (2009) Interactome analyses identify ties of PrP and its mammalian paralogs to oligomannosidic N-glycans and endoplasmic reticulum-derived chaperones. PLoS Pathog 5(10):e1000608
119. Knobbe CB et al (2011) Choice of biological source material supersedes oxidative stress in its influence on DJ-1 in vivo interactions with Hsp90. J Proteome Res 10(10):4388–4404

Chapter 12
Mass Spectrometry-Based Tissue Imaging of Small Molecules

Carly N. Ferguson, Joseph W.M. Fowler*, Jonathan F. Waxer*, Richard A. Gatti, and Joseph A. Loo

Abstract Mass spectrometry imaging (MSI) of tissue samples is a promising analytical tool that has quickly become associated with biomedical and pharmacokinetic studies. It eliminates several labor-intensive protocols associated with more classical imaging techniques and provides accurate histological data at a rapid pace. Because mass spectrometry is used as the readout, MSI can be applied to almost any molecule, especially those that are biologically relevant. Many examples of its utility in the study of peptides and proteins have been reported; here we discuss its value in the mass range of small molecules. We explore its success and potential in the analysis of lipids, medicinals, and metal-based compounds by featuring representative studies from MSI laboratories around the globe.

*Authors contributed equally.

C.N. Ferguson • J.W.M. Fowler • J.F. Waxer
Department of Chemistry and Biochemistry, University of California-Los Angeles, Los Angeles, CA 90095, USA

R.A. Gatti
Department of Pathology and Laboratory Medicine, David Geffen School of Medicine, University of California-Los Angeles, Los Angeles, CA 90095, USA

Department of Human Genetics, David Geffen School of Medicine, University of California-Los Angeles, Los Angeles, CA 90095, USA

J.A. Loo (✉)
Department of Chemistry and Biochemistry, University of California-Los Angeles, Los Angeles, CA 90095, USA

Department of Biological Chemistry, David Geffen School of Medicine, University of California-Los Angeles, Los Angeles, CA 90095, USA
e-mail: JLoo@chem.ucla.edu

A.G. Woods and C.C. Darie (eds.), *Advancements of Mass Spectrometry in Biomedical Research*, Advances in Experimental Medicine and Biology 806,
DOI 10.1007/978-3-319-06068-2_12, © Springer International Publishing Switzerland 2014

12.1 Introduction

Mass spectrometry-based imaging of tissue samples and surfaces is a relatively new technology with a significant function in medicinal development [33]. Mass spectrometry imaging (MSI) was developed in the laboratory of Dr. Richard Caprioli and has continued to garner interest in labs around the world as its applications expand [33, 38]. The methodology essentially provides a chemical map of a surface by incorporating mass spectrometry instrumentation and assigning spatially relevant coordinates [33]. The specificity of analyte detection is provided by the mass spectrometry readout.

MSI is most often coupled with matrix-assisted laser desorption/ionization (MALDI), although we will briefly discuss an alternative ionization method. MALDI utilizes a matrix compound that is mixed and crystallized with the analyte of interest followed by desorption/ionization with a laser beam. MALDI matrices absorb efficiently at the wavelength of the specified laser. In MSI, the fundamentals of MALDI are utilized in a spatially relevant manner. A thin surface (for our purposes, a thin tissue section) is evenly coated in a MALDI matrix, and this surface is subsequently analyzed by rastering across it with laser irradiation. The spatial resolution is controlled by the operator and is limited by the laser spot size (e.g., typically 20–100 μm), and thousands of spectra are collected at specific *X*–*Y* coordinates covering the surface of interest. After collection, all spectra are compiled into one average spectrum, wherein masses of interest can be selectively highlighted and their spatial distribution depicted in a virtual image [33]. This workflow is shown in Fig. 12.1. Tissue samples (whole organs) are typically sliced on a cryotome at 10–20 μm thickness and thaw-mounted onto glass slides coated with conductive material. Matrix is applied via any one of a number of automated mechanisms, typically consisting of a nebulizer that evenly coats the entire slide with a thin layer of matrix crystals [33, 38].

There are several benefits of implementing MSI. Traditional protein and peptide imaging methods, such as immunohistochemistry (IHC) staining, require the use of a specific antibody [14]. Additionally, these staining methods only allow for a small

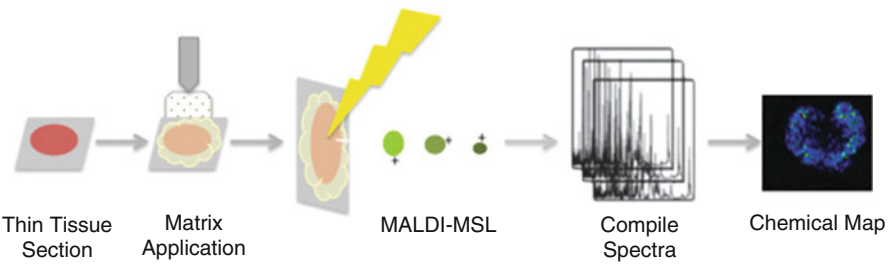


Fig. 12.1 Schematic of a MALDI-MSI experiment showing a typical tissue sample preparation protocol. A laser then rasters across the tissue surface. At each point, or pixel, a mass spectrum is collected. These spectra are averaged, and individual masses are selected to create maps, such as the one shown on the right (kidney tissue image of a small molecule metabolite)

number of molecules of interest to be analyzed on any given tissue section [1, 42]. MSI does not require antibodies, and the use of a mass spectrometer's analyzer and detector also allow for the visualization of thousands of molecules of interest simultaneously as all ionized molecules are nonspecifically detected. Similar benefits are true of MSI experiments for detection of small molecules. The mass accuracy of the MS analyzer, e.g., time-of-flight (TOF) analyzers most often applied for MSI, often allows for the identification of the molecule of interest, and advanced MS methods, such as tandem mass spectrometry (MS/MS), yield additional molecular information to derive identification.

A recent study by the Caprioli group demonstrates the functionality of MSI in a pathological sense. This work probed proteomic differences between two dermatological conditions: Spitz nevus (SN) and Spitzoid malignant melanoma (SMM). These two conditions are virtually indistinguishable by traditional pathology methods; however, an MSI study was able to identify a key proteomic difference that distinguishes the benign SN from the malignant SMM with 97 % accuracy. This study showcases the ability of MSI to complement existing histopathology methods for more confident diagnoses [29].

In addition to interest in proteins and peptides, small molecules of all types are compounds of interest in many clinically relevant studies. Herein we discuss the advances of MSI in three classes of small molecules: lipids, drug compounds, and nanoparticles. Combined, these studies broadly indicate the bright future MSI holds in the biomedical field.

12.2 Lipids

Lipids have considerable diversity in chemical structure and biological function and comprise the external lipid bilayer as well as subcellular organelles, including the mitochondria and surrounding nuclear membrane [21]. Due to this diversity in structure and localization, techniques that pinpoint specific lipids within tissues, while not delocalizing the compound of interest, are of great interest. The techniques generally used to identify lipids, however, involve the extraction of the lipids prior to analysis, which destroys relevant information regarding overall spatial distribution [36]. These traditional procedures for lipid analysis utilize destructive assays, which are subsequently coupled with mass spectrometry and/or liquid chromatography [5, 13, 24, 47, 51]. Fluorescence imaging could theoretically overcome these challenges, as fluorescent tagging procedures preserve location, reveal the exact position of lipids, and allow researchers to observe the rapid and dynamic changes in the location and structure of lipids [43]. However, most fluorescent tags are as large as the lipid molecules themselves, which would likely cause changes to the overall location and metabolism of the lipid. Therefore, in order to successfully study lipids within a biological system, analytical techniques must be able to overcome two major challenges: (i) preservation of the structural and locational information of individual lipids, and (ii) provide a high level of molecular specificity to

differentiate the diversity in lipid structure [20]. Modern mass spectrometry, due to its excellent sensitivity and molecular specificity, is arguably a method of choice for lipid analysis. The mass spectrometric analysis of lipids is traditionally achieved through lipid extraction from a sample and subsequent analysis by electrospray ionization (ESI), with or without prior chromatographic separation [20]. Although this method provides detailed structural information, it is not useful for measuring the spatial distribution of specific lipids. Therefore, new mass spectrometry approaches have been developed that provide direct surface analysis capabilities.

MALDI-MSI of Lipids. Recently, advances in MALDI-MSI have allowed for in-depth qualitative, quantitative, and spatial analyses of several types of lipid species. Although the preparation and specific parameters needed to map lipids varies between specific species, MALDI-MSI has been shown to be a more efficient and convenient technique to detect lipids compared to traditional methods. Techniques, such as fluorescence confocal microscopy and fluorescent tagging of molecules, are poor in lipid characterization because of the relatively small size of lipids as well as their ubiquity throughout many cell types [55]. The abundance of lipids in various cell structures, such as internal and external membranes, is beneficial to MALDI IMS because it allows for the direct visualization of these various compounds on a tissue section. Furthermore, the sensitivity of MS for molecular weight ranges under 1,000 Da is especially high [3].

MALDI generates two-dimensional molecular maps by ionizing molecules on a tissue sample through direct exposure to a laser. This is ideal for analysis of lipid compounds because of their inherent ability to ionize and produce positive or negative ions. Lipid molecules are amphipathic molecules comprised of either phosphate anions or nitrogen-centered cations that readily ionize during MALDI [9, 21].

Several techniques exist to prepare tissue sections to examine various types of lipid species in order to obtain good quality mass spectra. Generally, washing the mounted tissue slides with aqueous, volatile salt solutions simplifies the spectra to lipid compounds, such as phosphotidylcholine, that generate an abundance of positive ions through MALDI [49]. This will remove interfering salts and generate well-resolved mass spectra by taking out interfering chemical noise. Lipid species that form a high abundance of negative ions during the MALDI process, such as phosphatidylenolamine, naturally produce simpler spectra because they do not form alkali metal attachment ions [55].

A final advantage of MALDI-MSI of lipids is that it can be used in combination with other techniques. For example, MALDI can be used after thin layer chromatography (TLC) allowing for precise isolation of lipid compounds directly on the TLC plate, versus a more labor-intensive extraction process [55]. MALDI-MSI has also been used in conjunction with structural information obtained from ESI experiments, immunoblotting, and histological information obtained from staining to provide a multimodal model of lipid localization in the brain [12]. MALDI-MSI analysis of lipids has also allowed advancements in forensic investigation as well as diagnosis of diseases such as breast cancer and traumatic brain injury [51, 55]. MALDI-MSI has proven invaluable in the growth of lipid research and continues to spread into other avenues of biochemistry and analytical chemistry.

DESI of Lipid Compounds. MALDI-MSI is often used for peptide and protein analysis and has been used for lipid analysis as well [50]. MALDI-MSI utilizes a preparative step that evenly covers the tissue with an organic matrix, assisting with ionization of the sample [22, 44]. The type of matrix used, however, greatly influences the efficiency of ionization for different classes of molecules, and experimental conditions used for lipid imaging are different from those used for proteins in terms of matrix used, mass-to-charge (*m/z*) range, and the acquisition mode when using TOF analyzers [17, 22, 23, 32, 44].

Desorption electro spray ionization (DESI) is an ambient desorption/ionization technique that is based on the direct examination of unprepared, unmodified samples in the open environment and is commonly used for drug, metabolite, and lipid imaging. Recently, the development of dimethylformamide (DMF)-based solvent combinations has minimized the destructive nature of the technique and has enabled DESI-MSI to be performed with preservation of morphological features [18, 50]. By preserving morphological features, histological and immunohistochemical analysis can be performed after MSI [17]. The ability to preserve tissue morphology for histological examination, while assessing the lipid and protein profiles on the same section of tissue, provides an opportunity to correlate findings [18]. Eberlin et al. demonstrated that a single tissue section initially used for DESI-MSI of lipids can then be used for protein MSI, and subsequently hemotoxylin and eosin (H&E) staining to acquire morphological information [18].

By combining the techniques of DESI-MSI, MALDI-MSI, and H&E staining, Eberlin et al. [18] unambiguously matched the morphological and chemical features of mouse brain and human brain cancer tissue samples. They found that, prior to DESI-MSI of lipids, using either acetonitrile (ACN):DMF (1:1) or ethanol:DMF (1:1) solvent systems did not disturb the native protein localization, as ion images of control tissue sections showed similar protein spatial distributions, as well as similar co-localization of lipids and proteins [18]. Furthermore, results were able to be obtained from a sample of human glioma grade III, having first undergone DESI-MSI, followed by MALDI-MSI, and finally optical imaging after H&E staining [18]. In addition, their analysis of a control tissue section, which had not been subjected to DESI-MSI, revealed similar spatial distributions of their select proteins of interest [18]. This new approach combines the unique strengths of DESI and MALDI for lipid and protein MSI and allows the unambiguous matching of morphological and chemical features. Eberlin et al. [18] conclude that the combined methods of DESI, MALDI, and H&E on the same tissue section enable a more complete evaluation and are expected to not only enhance diagnostic capabilities, but also allow insights into the pathophysiology of disease.

These capabilities were further demonstrated by the same group in a study that employed ambient ionization-based MSI for characterization of tumor borders. Several brain tumors were analyzed via DESI-MSI to develop lipid-based classifiers to distinguish between tumor and healthy tissue. This technique provides fast classification and remains promising for intraoperative use to differentiate visibly indistinguishable tumor borders [19].

MSI with FT-ICR MS for Lipid Analysis. MALDI-MSI with TOF analyzers does not provide the ultra-high resolution and mass accuracy necessary for direct identification of low molecular weight compounds. However, MALDI can be coupled with Fourier transform ion cyclotron resonance (FT-ICR) MS in order to obtain high mass accuracy and identify compounds of interest with a high degree of confidence in tissue [31].

Vidova et al. [48] characterized the major lipid components of the ocular lens using MALDI-TOF-MSI. They found that the major lipid components were mainly long chain phosphatidylcholines (PC) and sphingomyelins (SM) [41]. However, MALDI-TOF-MSI did not provide the proper degree of resolution and mass accuracy to determine the exact masses of the lipid components in order to produce accurate spatial distribution results of compounds close in molecular weight. Therefore, Vidova et al. [48] used MALDI-FT-ICR-MSI in order to determine the masses of the specific PC and SM species within the ocular lens. Sections of porcine eyes were prepared and analyzed with MALDI-FT-ICR-MSI in positive ion mode. An example of the usefulness of FT-ICR MS in this specific study was in the analysis of a lipid compound with a mass of 787 Da. The study found the lipid aggregated in its protonated form at *m/z* 787.6685 and also found an unknown molecule at *m/z* 787.6042. Ultra-high resolution MS allowed for the discrimination between these two peaks, whereas a lower resolution MS instrument would not be able to resolve the two and would disrupt quantitative, qualitative, and spatial distribution analyses. By utilizing the high mass accuracy of FT-ICR MS, Vidova et al. [48] was able to find spatial and concentration-dependent distributions of various species of lipids within the ocular lens.

Lipids in Traumatic Brain Injury. MALDI-MSI is also an important tool that can be used to develop molecular biomarkers of disease. Recently, Woods et al. [52] used MALDI-MSI to characterize and quantify the spatial distribution of ganglioside species in mouse brains after being subjected to low-level explosive detonations. This study set out to find a biochemical connection between exposure to explosive blasts and traumatic brain imagery (TBI). Gangliosides account for 6 % of total brain mass and are included in several biochemical and metabolic pathways [26]. Ceramides are the scaffold molecules that form gangliosides, and the effects of proximity blasts from explosions may cause a disruption in the various pathways of these two lipids [4]. Woods et al. [52] studied the changes in the amount and spatial distribution of gangliosides and ceramides in mouse model brains exposed to blast explosions. Ceramides were found to be present at higher concentrations in TBI brains; however, this was determined via ESI due to insufficient MALDI ionization efficiency. It was found that exposure to explosive blasts increases the amounts of a ganglioside GM2 in several portions of the brain and causes a subsequent decrease in the concentration of ceramide species (Fig. 12.2). This was the first successful study to display an increase in GM2 from a nongenetic cause [52]. The findings of Woods et al. [52] point towards the possibility of using gangliosides and ceramides as biomarkers in the detection and analysis traumatic brain injury in patients, and furthermore, implicate MALDI-MSI as a useful strategy for biomarker discovery.

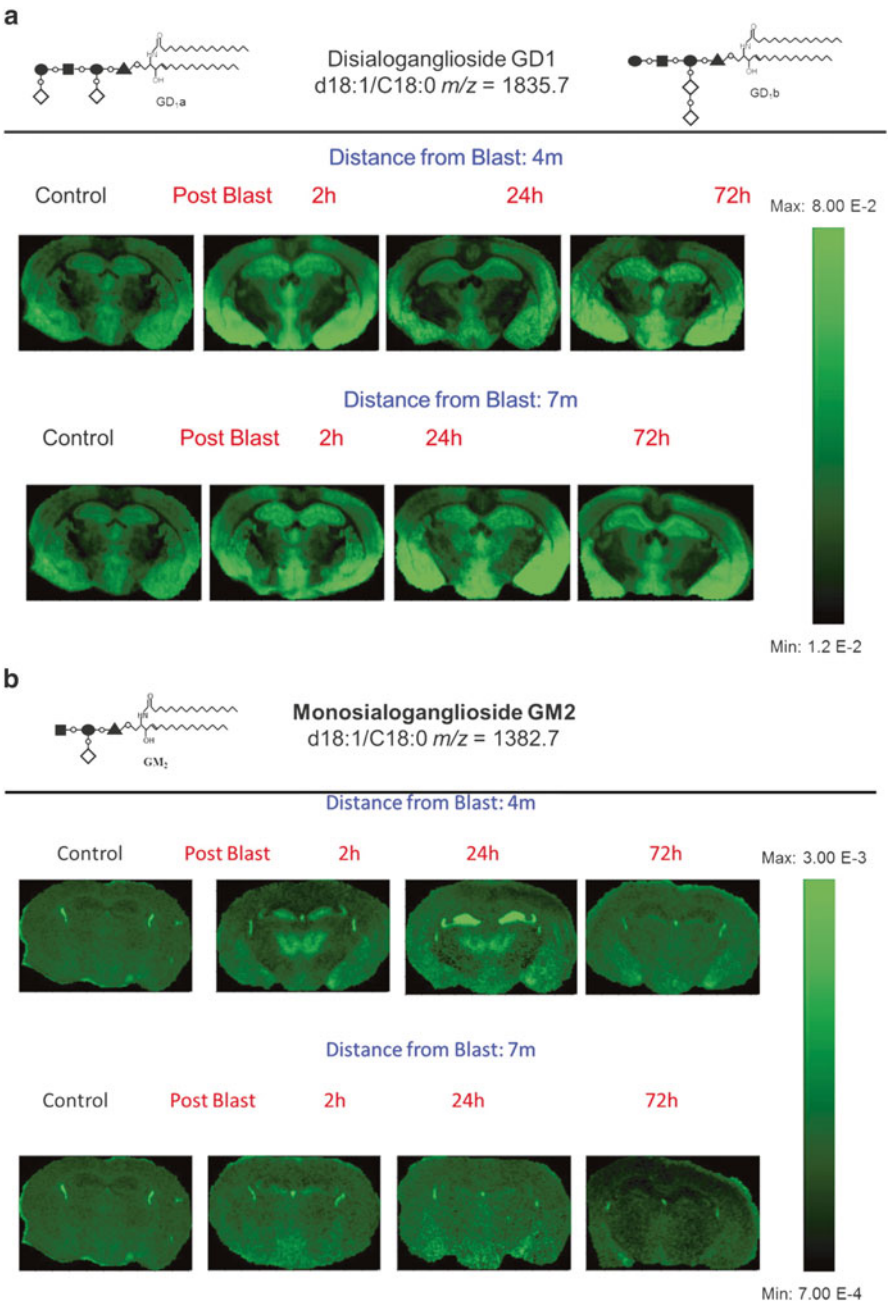


Fig. 12.2 MALDI-MSI of two gangliosides: (**a**) GD1d18:1/C18:0 and (**b**) GM2 d18:1/C18:0 in brains of control mice and mice 2, 24, and 72 h after open field blast exposure at 4 and 7 m. In panel a, GD1 is present mainly in gray matter areas (hippocampus, cortex, and hypothalamus), and there is no observable change in the distribution of GD1 between control and blast groups. In panel b, for the controls, the GM2 peak was highly localized in the lateral and the dorsal third ventricles, while increases were observed in the hippocampus and thalamus for blast exposure, especially for 4 m 2 and 24 h post-blast samples. Reprinted with permission from the *Journal of the American Chemical Society*. Copyright © 2013 American Chemical Society [52]

12.3 Medicinal Development

For the discovery and development of new drug entities, knowledge of a given compound's biodistribution is a critical factor. To where does a new drug compound distribute in the body? What are the metabolites of this new drug and where do these metabolites accumulate? What are the quantitative measures of their distribution? These questions must be answered in order to move from an in vivo animal model to human clinical trials. Because the approval process of a novel therapeutic is long, arduous, and very costly, areas in which technology can speed the acquisition of this information are open to modification [6]. The majority of these technology-driven areas throughout the drug development timeline occur within the preclinical stage. Traditionally, these early questions are answered via methods such as autoradiography and fluorescent tagging, each with its own set of limitations [37, 45]. Both methods require labor-intensive synthesis efforts to apply the necessary tags. Autoradiography involves the use of radioactive materials, and thus comes with its own set of regulatory issues. When radiolabeled versions of novel compounds are delivered to an animal, an autoradiograph is produced; however, there is little to no indication of whether the observed image is the result of an intact compound or a metabolic fragment [3, 45]. Conversely, in a metabolite study, only metabolites containing radiolabeled atoms will be detected and observed. Fluorescence tags are significantly larger than radioactive labels and require a significant amount of testing due to the possibility that the tag may interfere with the compound's membrane penetrance and efficacy [40]. Compounds are often less than 500 Da, meaning even the smallest dye molecule can have a large effect on the uptake and subsequent biodistribution of the compounds of interest. Images are further complicated by molecules exhibiting autofluorescence, and issues similar to radiolabeling, such as difficulty attributing a signal to intact or fragmented compound [40]. The synthesis of a pure compound itself provides a challenge.

These issues can be circumvented by MSI, as its non-targeted nature means all molecules on a given surface are chemically mapped, and no tag is required. With proper controls in place, previously unknown metabolites can be uncovered by utilizing tandem mass spectrometry (MS/MS) capabilities, essentially determining the structure of an unknown peak. Below we discuss a few case studies that have utilized MSI in drug development [6, 8].

Fosdevirine. MALDI-MSI has already found success in the pharmaceutical industry, as evidenced by the recent study by Castellino et al. [7]. Human immunodeficiency virus (HIV) is a widespread disease with limited treatments available. It is an autoimmune disorder that develops when viral strains display varying resistance to the available treatments. It remains an active area of research in the pharmaceutical industry [7].

The group of Castellino conducted experiments on a non-nucleoside reverse transcriptase inhibitor that had passed all preclinical testing and entered into Phase IIb clinical trials. Fosdevirine was developed by GlaxoSmithKline to treat a wide range of HIV-1 strains, the most common subtype as well as the most lethal.

Many of these strains have become resistant to other forms of existing treatment, such as Efavirenz [7]. Rigorous preclinical testing identified Fosdevirine as a promising drug due to its efficacy in low doses, and its effectiveness against both single and double mutants of reverse transcriptase that confer remarkable resistance against other drugs. The drug moved into a human model, with Phase I testing conducted on healthy individuals indicating no observable toxicity or side effects. Further testing on HIV-1-infected individuals that had not previously received treatment confirmed earlier findings from healthy individuals. However, Phase IIb trials involving HIV-1-infected individuals that had undergone alternative treatments resulted in 25 % of all subjects experiencing seizures, with no link to neurological problems in medical histories [7].

MSI was employed to study its biodistribution and metabolism in rabbit, minipig, and monkey brain tissue in order to evaluate potential differences associated with adverse neurological symptoms. This was done in conjunction with liquid chromatography-MS (LC-MS) analyses of cerebral spinal fluid (CSF) from seizure patients as well as animal models. LC-MS analyses of CSF found two cysteine adduct metabolites, M22 and M16, to be present in seizure patients as well as rabbits and minipigs that exhibited central nervous system (CNS) toxicity due to Fosdevirine treatment. While these results indicate M22 and M16 as potential effectors of CNS toxicity, they do not provide spatial distribution information within the tissues themselves. MSI studies further indicated a potential mechanism for this CNS toxicity, providing important localization information. M22 was observed to localize to the white matter portion of the brain in minipig and rabbit, both of which exhibited Fosdevirine-induced CNS toxicity. In monkeys, where no such toxicity was observed, Fosdevirine was found to localize to the gray matter portion of the brain. It is hypothesized that binding of M22 and M16 to the GABAa receptor could be the cause of this CNS toxicity. With MSI's capability to multiplex with other modes of imaging, it is possible to utilize a technique such as immunohistochemical staining in order to observe that M22 co-localizes with the GABAa receptor in rabbit and minipig brain, but not monkey brain [7]. Together, this localization to the white matter and co-localization with the GABAa receptor in model organisms exhibiting Fosdevirine CNS toxicity show the advantages of combining MSI with traditional staining protocols [6–8].

Paclitaxel. Optimized sample preparation protocols are important in the field of MSI. One such step is the matrix selection, both type and application method can be varied in order to optimize ionization of specific molecules [38, 39]. This optimization is especially crucial in small drug molecule studies due to the ion suppression effects that are observed in the low *m/z* range of the spectrum. Because traditional MALDI matrices are, themselves, small molecules, they contribute a number of interfering, intense ion peaks in a mass spectrum. This can often suppress signal from target drug molecules that are present in much lower levels than the saturated matrix solutions. Additionally, the spatial resolution of any given imaging run can be limited by the size of the crystallized matrix particles themselves, often exceeding the size of the laser spot and thus decreasing the resolution of the data [6, 8, 38, 39].

One common solution to these problems is the use of nanoparticle-assisted laser desorption/ionization (NALDI). Nanoparticles are typically composed of an inorganic material and exhibit such desirable characteristics as low heat capacity and efficient photo-adsorption. They are also smaller in size than matrix crystals, and thus do not interfere with spatial resolution as readily; additionally, they have a large surface area that allows for maximum adsorption of the desired analytes [35].

Morosi et al. [35] employed NALDI for analysis of the biodistribution of a small drug molecule used for cancer treatments, Paclitaxel, in solid tumors. This group utilized MSI in organs excised from mice injected with Paclitaxel, and further utilized MSI in the study of human xenograft tumors excised from mice injected with Paclitaxel. MSI served two major advantages in the detection of Paclitaxel: negative ion mode and tandem mass spectrometry. Paclitaxel ionizes more efficiently in negative ion mode and is typically suppressed by abundant lipids in positive ion mode over the *m/z* 800–1,000 range (Paclitaxel MW 853.9). Thus, MSI allows for the option of collecting data for observation of a negatively charged ion of Paclitaxel, specifically a fragment ion at *m/z* 284.2. MS/MS coupled to MSI experiments allowed for the confirmation of this peak by observing the transition from *m/z* 284.2 to *m/z* 72.6 [35].

This group also employed quantitative MSI by utilizing a deuterated internal standard, D5-Paclitaxel. The internal standard was spotted onto control tissues, a calibration curve was constructed, and Paclitaxel signal was measured and compared to the internal standard signal in a specific region of interest. The success of this quantitative method was further shown by the differences in normalized intensities in liver when mice were treated with different levels of the drug. This group indicates that their success with quantitative MSI was likely due to the homogeneity of the nanoparticles used, versus the traditional heterogeneous MALDI matrix crystals [35].

Read-Through Compounds. The treatment and management of genetic disorders are major target areas in the drug development field since no such drugs are yet of proven efficacy. One common genotype is caused by nonsense mutations, resulting in a malformed or unstable protein. Such nonsense mutations create a premature termination codon (PTC). Instead of translating a gene sequence in its entirety, the mRNA falls off the ribosome resulting in a truncated form of the protein. This protein fragment is subsequently degraded, and thus no full length, functional protein is produced by the patients' cells [25].

Examples of PTC disorders include Duchenne muscular dystrophy (MD), cystic fibrosis, and ataxia-telangiectasia (A-T). A-T is often caused by a single point mutation in the ataxia-telangiectasia mutated (ATM) gene encoding for ATM protein. This serine/threonine kinase is active in DNA repair, mainly by phosphorylation of a large number of proteins involved in DNA repair and cancer. Individuals deficient in ATM exhibit such symptoms as impaired cerebellum development, increased risk of cancer (one third of all A-T patients develop cancer), and enhanced susceptibility to radiation [10, 11, 46].

Aminoglycosides, such as Gentamicin, could potentially be used to treat PTC disorders by binding to the small ribosomal subunit at the decoding site and

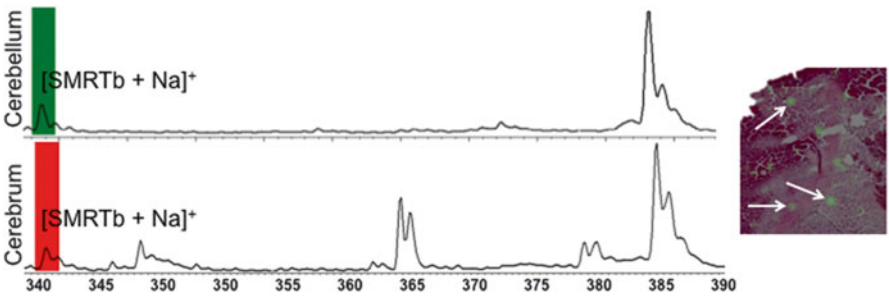


Fig. 12.3 MALDI-MSI of brains from SMRT-treated mice. The top panel shows a sodiated adduct found in cerebellum, image shown on right: green spots (*white arrows*) correspond to signal. Tissue was H&E stained following MSI. The bottom panel shows a sodiated adduct found in the cerebrum

inducing translational read-through of the stop codon [54]. While this has been successful in laboratory studies in "proof of concept," it would engender adverse side effects such as toxicity and potential deafness [2, 28, 56]. Treating neurological diseases like A-T would also require that the aminoglycoside cross the blood–brain barrier. Due to their large and rigid structure, aminoglycosides do not penetrate a healthy blood–brain barrier. A new class of small molecule read-through (SMRT) compounds has recently been developed that overcomes most of these issues [2, 10, 11, 15, 25, 28, 46, 54, 56].

Although SMRT compounds have shown promise in vitro, there has been little information concerning the biodistribution of these compounds. In a mouse model, we determined the biodistribution of a derivative of RTC13 utilizing MSI [15, 16]. The compound was administered intraperitoneally to mice, and its biodistribution was studied in all major organs at various time points. The SMRT compound was found to cross the blood–brain barrier without any off-target build-up. These results were confirmed by MS/MS, and tissue structure was elucidated by H&E staining, as shown in Fig. 12.3 (in preparation). These results highlight the utility of MSI as an initial biodistribution screening technique due to its speed and direct analysis of unmodified, novel compounds.

12.4 Metals and Nanoparticles

A newer area of MSI being investigated is the imaging of nanoparticles and other metal elements. Metal analysis is typically done using inductively coupled plasma MS (ICP-MS), which uses a plasma flame to ionize metal particles wherein they are detected by a mass spectrometer [57, 58]. While ICP-MS is often coupled with liquid samples in an electrospray-like set-up, recent studies have utilized laser ablation ICP-MS (LA-ICP-MS) to study dried samples, such as droplets, tissues, and other

surfaces. The coupling of ICP-MS to laser ablation has allowed for spatially relevant images to be collected, and thus MSI of metal elements is possible. We will discuss two examples of LA-ICP-MSI analyses of metals in tissue [57, 58].

Aside from the study of innate metal composition of tissues, there is growing interest in metal- and non-metal-based nanoparticle biodistribution. This stems from the recent development of nanoparticles as drug delivery vehicles [34]. Nanoparticles are promising due to their small, biocompatible nature and their ability to package many varieties of molecules, from proteins to small drug compounds, and deliver them to specific areas of a cell or organism. We will discuss a recent study that further stretched the capabilities of MALDI-MSI for the purposes of imaging nanoparticles in tissue [53].

Endogenous Metal Ions. LA-ICP-MS has long been used for the analysis of trace metals in samples; however, LA-ICP-MSI is a more recent application of this useful technique [57, 58]. Many metal ions play key roles in biological processes, thus it is of interest to study their distribution in tissue samples. Lear et al. [30] further improved the LA-ICP-MSI protocols in order to allow for the detection of traditionally undetectable metal ions by altering the reaction gas contents in a typical experiment.

Tissue surfaces are extremely complex, with thousands of different compounds present in a 1 μm^2 pixel, thus it is advantageous to develop techniques that will allow for enhanced signal intensity of particular molecules of interest. Specifically in LA-ICP-MSI, argon gas is employed to provide electrons as well as assist in the nebulizing process of ions of interest; however, this use of Ar significantly interferes with a number of signal regions, most notably that of iron (Fe). Additional impurities in the Ar gas source can lead to further eclipsing of signals of interest. This group employed the use of an H_2 reaction gas to dissociate ArO clusters and other interfering compounds while maintaining sensitivity in the detection of several metal ions, including: Mn, Fe, Cu, and Zn [30].

Mouse brain was used to study the distribution of each of the above mentioned metal ions. For Cu and Zn ions, specifically ^{63}Cu and ^{66}Zn, there are no known interferences and thus it was shown that use of H_2 reaction gas did not improve signal intensity during an LA-ICP-MSI experiment; however, both ions were able to be mapped at a 30 μm spatial resolution. The success of the H_2 reaction gas was evident in the analysis of the ^{56}Fe and ^{57}Fe ions, as well as the ^{55}Mn ion, all of which experience heavy interference with Ar gas clusters and impurities. Background signal was reduced after introduction of the reaction gas, and thus a much clearer ^{56}Fe image was produced at a spatial resolution of 6 μm. This improvement was even more pronounced in the analysis of ^{57}Fe, which is more heavily suppressed by interfering signals due to its lower abundance. The adaptability of this technique for analysis of important metal ions indicates its promise as an analytical tool for analysis of all biologically relevant molecules, from metal ions to proteins [30].

Platinum from cis*-platin drug*. Drug toxicity studies are imperative in advancing novel compounds into clinical trials. *Cis*-platin is a commonly used drug complex in the treatment of many types of cancers; however, it is known to exhibit nephrotoxicity when given in larger doses. It is of great interest to observe the

biodistribution of platinum (Pt) in *cis*-platin-treated individuals, as well as the differential distributions of copper and zinc (Cu and Zn), two important metal ions involved in biological processes [57].

Traditionally, Pt has been tracked in tissue via neutron activation analysis and autoradiography. These techniques, while sensitive, require labor-intensive techniques further complicated by the need for a nuclear reactor to produce neutrons. Less expensive alternatives to nuclear reactors often result in lower sensitivity and still involve the use of radioactive materials. Zoriy et al. [57] turned to LA-ICP-MSI to study the biodistribution of Pt, a technique that does not require the use of radioactive materials and is capable of whole-tissue analysis. This technique had been previously employed to show a correlation between the Alzheimer's disease-related amyloid beta protein (Aβ) and trace element concentration [27].

This group was able to image the biodistribution of Pt, Cu, and Zn in mouse kidney derived from *cis*-platin-dosed animals. Hemotoxylin and eosin staining was used to distinguish specific kidney structures, with LA-ICP-MSI experiments utilizing a spatial resolution of 50 μm. Zoriy et al. [57] were also able to employ internal standards for quantitative purposes. They observed a higher concentration of Cu in the capsule and external cortex (glomeruli), and a higher concentration of Zn in the inner cortex (tubules). Pt concentration was highest in the medulla, decreased in the inner cortex, and was lowest along the periphery of the kidney. The Pt gradient reflected the typical primary urine gradient found in the kidney structures. This is thought to be the first alternative imaging method for Pt that did not require the use of radioactive materials and thus shows the versatility of MSI [57].

Nanoparticles. A growing problem in medicine is the treatment of drug-resistant diseases, especially drug-resistant cancers. These cancers often require a combination of drugs working synergistically to effectively kill the fast-growing cells. Determining the perfect treatment cocktail can be a daunting task for physicians and researchers, and further prescribing of a large number of medications can create a variety of burdens for the patient in question. This also leads to complications in FDA investigational new drug (IND) studies. Nanoparticle drug delivery vehicles are a promising solution to these woes for a number of reasons [34].

Nanoparticles are typically small, biocompatible molecules on the scale of nanometers. Their surfaces are easily modified via various synthesis routes, and this allows for specific targeting of nanoparticles to desired cell types [34]. Additionally, their mostly hollow cores allow for packaging of a variety of molecules. Recent studies have indicated nanoparticles as effective delivery vehicles to carry molecules to specific drug-resistant cancer cells and subsequently release them; in this particular case these molecules included siRNA and a small drug compound [34].

Traditional nanoparticle imaging methods require the use of a large dye molecule, which can interfere with the activity and resulting biodistribution of the nanoparticles in question. Nanoparticles, such as gold or silica, are often used as MALDI matrices due to their limited background noise contributions in a mass spectrum [35]. However, they are not commonly detected using LDI-based methods due to their variation in size, the fact that they often contain fixed charges, and the lack of a predictive spectral profile.

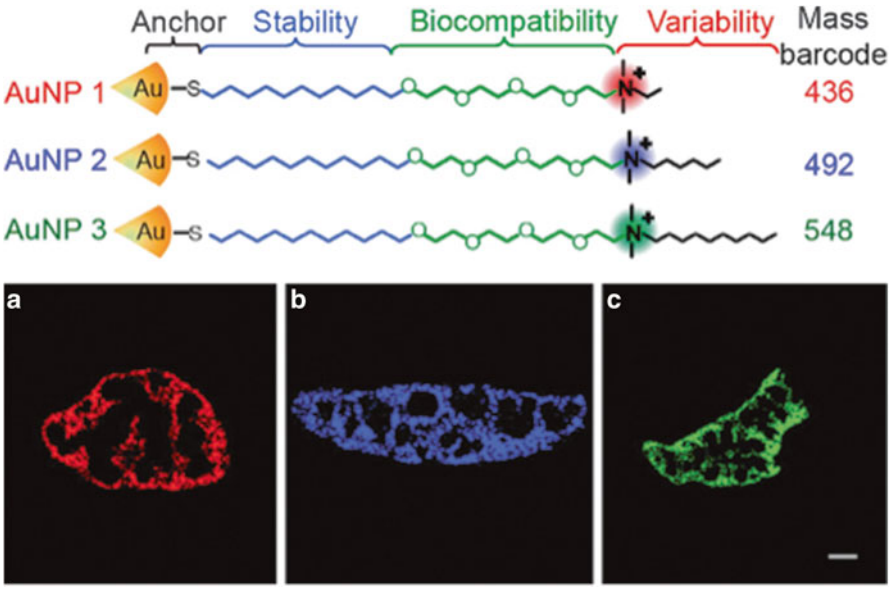


Fig. 12.4 Structures of the surface monolayers on the AuNPs: the "mass barcode" is the *m*/*z* of the AuNP surface ligand. LDI-MS images of AuNPs in mouse spleens. The biodistributions of AuNPs are shown in panel (**a**) (AuNP 1), (**b**) (AuNP 2), and (**c**) (AuNP 3). Reprinted and adapted with permission from the *Journal of the American Chemical Society*. Copyright © 2013 American Chemical Society [53]

Yan et al. [53] have circumvented these limitations by establishing a spectral profile, or "barcode," of the nanoparticles involved via attachment of a ligand molecule to the surface of the nanoparticle. These ligands are smaller in size than traditional dyes and serve the added function of modifying the surface properties of the nanoparticles. This study was able to identify a molecular ion attributed to the gold nanoparticle core, as well as three distinct molecular ions attributed to each of the ligands attached, as shown in Fig. 12.4. This illustrates a promising future for the use of LDI-MS-based imaging methods for the detection of such nontraditional molecules as nanoparticles [53].

12.5 Conclusion

Since its inception nearly two decades ago, the field of MSI has grown rapidly. New developments in sample preparation protocols, desorption/ionization methods, instrumentation (e.g., mass spectrometry analyzers and detectors), and data processing and data display have aided in its incorporation into bioanalytical chemistry and even medicine. We have discussed a number of applications in the realm of small molecules, and many more examples of MSI's utility can be found in the

study of peptides and proteins. Going forward, it is likely that MSI will become a permanent fixture in biomedical research. Further advancements in software and instrumentation are already aiding its incorporation into clinical settings, and the continued, collaborative goal towards standardized protocols will aid in its development as a pharmaceutical and clinical tool.

Acknowledgments This work was supported by the Ruth L. Kirschstein National Research Service Award (Grant No. GM007185, UCLA Cellular and Molecular Biology Training Grant, for C.N.F.) and the US National Institutes of Health Shared Instrumentation Program (Grant No. S10 RR025600 to J.A.L.).

References

1. Andersson M, Groseclose MR, Deutch AY, Caprioli RM (2008) Imaging mass spectrometry of proteins and peptides: 3D volume reconstruction. Nat Methods 5:101–108
2. Barton-Davis ER, Shoturma DI, Sweeney HL (1999) Contribution of satellite cells to IGF-I induced hypertrophy of skeletal muscle. Acta Physiol Scand 167:301–305
3. Börner K, Malmberg P, Månsson J-E, Nygren H (2007) Molecular imaging of lipids in cells and tissues. Int J Mass Spectrom 260:128–136
4. Buccoliero R, Futerman AH (2003) The roles of ceramide and complex sphingolipids in neuronal cell function. Pharmacol Res 47:409–419
5. Bulley NR, Fattori M, Meisen A, Moyls L (1984) Supercritical fluid extraction of vegetable oil seeds. J Am Oil Chem Soc 61:1362–1365
6. Castellino S (2012) MALDI imaging MS analysis of drug distribution in tissue: the right time!(?). Bioanalysis 4:2549–2551
7. Castellino S, Groseclose MR, Sigafoos J, Wagner D, De Serres M, Polli JW, Romach E, Myer J, Hamilton B (2013) Central nervous system disposition and metabolism of fosdevirine (GSK2248761), a non-nucleoside reverse transcriptase inhibitor: an LC-MS and matrix-assisted laser desorption/ionization imaging MS investigation into central nervous system toxicity. Chem Res Toxicol 26:241–251
8. Castellino S, Groseclose MR, Wagner D (2011) MALDI imaging mass spectrometry: bridging biology and chemistry in drug development. Bioanalysis 3:2427–2441
9. Chaurand P, Cornett DS, Caprioli RM (2006) Molecular imaging of thin mammalian tissue sections by mass spectrometry. Curr Opin Biotechnol 17:431–436
10. Chun HH, Gatti RA (2004) Ataxia-telangiectasia, an evolving phenotype. DNA Repair (Amst) 3:1187–1196
11. Chun HH, Sun X, Nahas SA, Teraoka S, Lai C-H, Concannon P, Gatti RA (2003) Improved diagnostic testing for ataxia–telangiectasia by immunoblotting of nuclear lysates for ATM protein expression. Mol Genet Metab 80:437–443
12. Colsch B, Jackson SN, Dutta S, Woods AS (2011) Molecular microscopy of brain gangliosides: illustrating their distribution in hippocampal cell layers. ACS Chem Neurosci 2:213–222
13. Conte E, Milani R, Morali G, Abballe F (1997) Comparison between accelerated solvent extraction and traditional extraction methods for the analysis of the herbicide diflufenican in soil. J Chromatogr A 765:121–125
14. Coons AH, Creech HJ, Jones RN, Berliner E (1942) The demonstration of pneumococcal antigen in tissues by the use of fluorescent antibody. J Immunol 45:159–170
15. Du L, Damoiseaux R, Nahas S, Gao K, Hu H, Pollard JM, Goldstine J, Jung ME, Henning SM, Bertoni C, Gatti RA (2009) Nonaminoglycoside compounds induce read through of nonsense mutations. J Exp Med 206:2285–2297

16. Du L, Jung ME, Damoiseaux R, Completo G, Fike F, Ku JM, Nahas S, Piao C, Hu H, Gatti RA (2013) A new series of small molecular weight compounds induce read through of all three types of nonsense mutations in the ATM gene. Mol Ther 21(9):1653–1660
17. Eberlin LS, Ferreira CR, Dill AL, Ifa DR, Cheng L, Cooks RG (2011) Nondestructive, histologically compatible tissue imaging by desorption electrospray ionization mass spectrometry. Chembiochem 12:2129–2132
18. Eberlin LS, Liu X, Ferreira CR, Santagata S, Agar NYR, Cooks RG (2011) Desorption electrospray ionization then MALDI mass spectrometry imaging of lipid and protein distributions in single tissue sections. Anal Chem 83:8366–8371
19. Eberlin LS, Norton I, Orringer D, Dunn IF, Liu X, Ide JL, Jarmusch AK, Ligon KL, Jolesz FA, Golby AJ, Santagata S, Agar NYR, Cooks RG (2013) Ambient mass spectrometry for the intraoperative molecular diagnosis of human brain tumors. Proc Natl Acad Sci 110:1611–1616
20. Ellis SR, Brown SH, In Het Panhuis M, Blanksby SJ, Mitchell TW (2013) Surface analysis of lipids by mass spectrometry: more than just imaging. Prog Lipid Res 52:329–353
21. Fahy E, Subramaniam S, Brown HA, Glass CK, Merrill AH, Murphy RC, Raetz CRH, Russell DW, Seyama Y, Shaw W, Shimizu T, Spener F, Van Meer G, Vannieuwenhze MS, White SH, Witztum JL, Dennis EA (2005) A comprehensive classification system for lipids. J Lipid Res 46:839–862
22. Ferguson L, Bradshaw R, Wolstenholme R, Clench M, Francese S (2011) Two-step matrix application for the enhancement and imaging of latent fingermarks. Anal Chem 83:5585–5591
23. Fuchs B, Süs R, Schiller J (2010) An update of MALDI-TOF mass spectrometry in lipid research. Prog Lipid Res 49:450–475
24. García-Ayuso LE, Velasco J, Dobarganes MC, Luque de Castro MD (2000) Determination of the oil content of seeds by focused microwave-assisted soxhlet extraction. Chromatographia 52:103–108
25. Gatti RA (2012) SMRT compounds correct nonsense mutations in primary immunodeficiency and other genetic models. Ann N Y Acad Sci 1250:33–40
26. Holthuis JCM, Pomorski T, Raggers RJ, Sprong H, Van Meer G (2001) The organizing potential of sphingolipids in intracellular membrane transport. Physiol Rev 81:1689–1723
27. Hutchinson RW, Cox AG, McLeod CW, Marshall PS, Harper A, Dawson EL, Howlett DR (2005) Imaging and spatial distribution of β-amyloid peptide and metal ions in Alzheimer's plaques by laser ablation–inductively coupled plasma–mass spectrometry. Anal Biochem 346:225–233
28. Lai CH, Chun HH, Nahas SA, Mitui M, Gamo KM, Du L, Gatti RA (2004) Correction of ATM gene function by aminoglycoside-induced read-through of premature termination codons. Proc Natl Acad Sci U S A 101:15676–15681
29. Lazova R, Seeley EH, Keenan M, Gueorguieva R, Caprioli RM (2012) Imaging mass spectrometry—a new and promising method to differentiate spitz nevi from spitzoid malignant melanomas. Am J Dermatopathol 34:82–90
30. Lear J, Hare DJ, Fryer F, Adlard PA, Finkelstein DI, Doble PA (2012) High-resolution elemental bioimaging of Ca, Mn, Fe, Co, Cu, and Zn employing LA-ICP-MS and hydrogen reaction gas. Anal Chem 84:6707–6714
31. Marshall AG, Hendrickson CL, Jackson GS (1998) Fourier transform ion cyclotron resonance mass spectrometry: a primer. Mass Spectrom Rev 17:1–35
32. McCombie G, Knochenmuss R (2004) Small-molecule MALDI using the matrix suppression effect to reduce or eliminate matrix background interferences. Anal Chem 76:4990–4997
33. McDonnell LA, Heeren RM (2007) Imaging mass spectrometry. Mass Spectrom Rev 26:606–643
34. Meng H, Liong M, Xia T, Li Z, Ji Z, Zink JI, Nel AE (2010) Engineered design of mesoporous silica nanoparticles to deliver doxorubicin and P-glycoprotein siRNA to overcome drug resistance in a cancer cell line. ACS Nano 4:4539–4550
35. Morosi L, Spinelli P, Zucchetti M, Pretto F, Carrà A, D'incalci M, Giavazzi R, Davoli E (2013) Determination of paclitaxel distribution in solid tumors by nano-particle assisted laser desorption ionization mass spectrometry imaging. PLoS One 8:e72532

36. Murphy RC, Hankin JA, Barkley RM (2009) Imaging of lipid species by MALDI mass spectrometry. J Lipid Res 50:S317–S322
37. Rasey JS, Krohn KA, Grunbaum Z, Spence AM, Menard TW, Wade RA (1986) Synthesis, biodistribution, and autoradiography of radiolabeled S-2-(3-methylaminopropylamino)ethylphosphorothioic acid (WR-3689). Radiat Res 106:366–379
38. Reyzer M, Caprioli R (2011) Imaging Mass Spectrometry. In: Banoub J (ed) Detection of biological agents for the prevention of bioterrorism. Springer, The Netherlands
39. Rubakhin SS, Jurchen JC, Monroe EB, Sweedler JV (2005) Imaging mass spectrometry: fundamentals and applications to drug discovery. Drug Discov Today 10:823–837
40. Rudin M, Weissleder R (2003) Molecular imaging in drug discovery and development. Nat Rev Drug Discov 2:123–131
41. Rujoi M, Estrada R, Yappert MC (2004) In situ MALDI-TOF MS regional analysis of neutral phospholipids in lens tissue. Anal Chem 76:1657–1663
42. Sanchez-Carbayo M (2006) Antibody arrays: technical considerations and clinical applications in cancer. Clin Chem 52:1651–1659
43. Schultz C, Neef AB, Gadella TW, Goedhart J (2010) Imaging lipids in living cells. Cold Spring Harb Protoc doi:10.1101/pdb.top83
44. Schwamborn K, Caprioli RM (2010) MALDI imaging mass spectrometry—painting molecular pictures. Mol Oncol 4:529–538
45. Solon EG, Schweitzer A, Stoeckli M, Prideaux B (2009) Autoradiography, MALDI-MS, and SIMS-MS imaging in pharmaceutical discovery and development. AAPS J 12:11–26
46. Swift M, Morrell D, Cromartie E, Chamberlin AR, Skolnick MH, Bishop DT (1986) The incidence and gene frequency of ataxia-telangiectasia in the United States. Am J Hum Genet 39:573–583
47. Taylor S, King J, List G (1993) Determination of oil content in oilseeds by analytical supercritical fluid extraction. J Am Oil Chem Soc 70:437–439
48. Vidova V, Pol J, Volny M, Novak P, Havlicek V, Wiedmer SK, Holopainen JM (2010) Visualizing spatial lipid distribution in porcine lens by MALDI imaging high-resolution mass spectrometry. J Lipid Res 51:2295–2302
49. Wang H-YJ, Liu CB, Wu H-W (2011) A simple desalting method for direct MALDI mass spectrometry profiling of tissue lipids. J Lipid Res 52:840–849
50. Watrous JD, Alexandrov T, Dorrestein PC (2011) The evolving field of imaging mass spectrometry and its impact on future biological research. J Mass Spectrom 46:209–222
51. Wenk MR (2005) The emerging field of lipidomics. Nat Rev Drug Discov 4:594–610
52. Woods AS, Colsch B, Jackson SN, Post J, Baldwin K, Roux A, Hoffer B, Cox BM, Hoffer M, Rubovitch V, Pick CG, Schultz JA, Balaban C (2013) Gangliosides and ceramides change in a mouse model of blast induced traumatic brain injury. ACS Chem Neurosci 4:594–600
53. Yan B, Kim ST, Kim CS, Saha K, Moyano DF, Xing Y, Jiang Y, Roberts AL, Alfonso FS, Rotello VM, Vachet RW (2013) Multiplexed imaging of nanoparticles in tissues using laser desorption/ionization mass spectrometry. J Am Chem Soc 135:12564–12567
54. Yoshizawa S, Fourmy D, Puglisi JD (1998) Structural origins of gentamicin antibiotic action. EMBO J 17:6437–6448
55. Zemski Berry KA, Hankin JA, Barkley RM, Spraggins JM, Caprioli RM, Murphy RC (2011) MALDI imaging of lipid biochemistry in tissues by mass spectrometry. Chem Rev 111:6491–6512
56. Zingman LV, Park S, Olson TM, Alekseev AE, Terzic A (2007) Aminoglycoside-induced translational read-through in disease: overcoming nonsense mutations by pharmacogenetic therapy. Clin Pharmacol Ther 81:99–103
57. Zoriy M, Matusch A, Spruss T, Becker JS (2007) Laser ablation inductively coupled plasma mass spectrometry for imaging of copper, zinc, and platinum in thin sections of a kidney from a mouse treated with cis-platin. Int J Mass Spectrom 260:102–106
58. Zoriy MV, Dehnhardt M, Matusch A, Becker JS (2008) Comparative imaging of P, S, Fe, Cu, Zn and C in thin sections of rat brain tumor as well as control tissues by laser ablation inductively coupled plasma mass spectrometry. Spectrochim Acta Part B 63:375–382

Chapter 13
Redox Proteomics: From Bench to Bedside

Karina Ckless

Abstract In general protein posttranslation modifications (PTMs) involve the covalent addition of functional groups or molecules to specific amino acid residues in proteins. These modifications include phosphorylation, glycosylation, S-nitrosylation, acetylation, lipidation, among others (Angew Chem Int Ed Engl 44(45):7342–7372, 2005). Although other amino acids can undergo different kinds of oxidative posttranslational modifications (oxPTMs) (Exp Gerontol 36(9):1495–1502, 2001), in this chapter oxPTM will be considered specifically related to Cysteine oxidation, and redox proteomics here is translated as a comprehensive investigation of oxPTMs, in biological systems, using diverse technical approaches. Protein Cysteine residues are not the only amino acid that can be target for oxidative modifications in proteins (Exp Gerontol 36(9):1495–1502, 2001; Biochim Biophys Acta 1814(12):1785–1795, 2011), but certainly it is among the most reactive amino acid (Nature 468(7325):790–795, 2010). Interestingly, it is one of the least abundant amino acid, but it often occurs in the functional sites of proteins (J Mol Biol 404(5):902–916, 2010). In addition, the majority of the Cysteine oxidations are reversible, indicating potential regulatory mechanism of proteins. The global analysis of oxPTMs has been increasingly recognized as an important area of proteomics, because not only maps protein caused by reactive oxygen species (ROS) and reactive nitrogen species (RNS), but also explores protein modulation involving ROS/RNS. Furthermore, the tools and strategies to study this type oxidation are also very abundant and developed, offering high degree of accuracy on the results. As a consequence, the redox proteomics field focuses very much on analyzing Cysteine oxidation in proteins under several experimental conditions and diseases states. Therefore, the identification and localization of oxPTMs within cellular milieu became critical to understand redox regulation of proteins in physiological

K. Ckless, Ph.D. (✉)
Department of Chemistry, State University of New York at Plattsburgh (SUNY Plattsburgh), Hudson Hall 231, 101 Broad Street, Plattsburgh, NY 12901, USA
e-mail: kckle001@plattsburgh.edu

A.G. Woods and C.C. Darie (eds.), *Advancements of Mass Spectrometry in Biomedical Research*, Advances in Experimental Medicine and Biology 806,
DOI 10.1007/978-3-319-06068-2_13, © Springer International Publishing Switzerland 2014

and pathological conditions, and consequently an important information to develop better strategies for treatment and prevention of diseases associated with oxidative stress.

There is a wide range of techniques available to investigate oxPTMs, including gel-based and non-gel-based separation approaches to be combined with sophisticated methods of detection, identification, and quantification of these modifications. The strategies and approaches to study oxPTMs and the respective applications related to physiological and pathological conditions will be discussed in more detail in this chapter.

13.1 Oxidative Posttranslational Modifications (oxPTM) on Proteins

13.1.1 Cysteine Reactivity and Oxidation

Cysteine (Cys) is the most nucleophilic amino acid in proteins and, therefore, sensitive to oxidative modification. This intrinsic reactivity is designed to perform diverse biochemical functions. The p*K*a of the free cys thiol is around 8.5. However, small changes in the microenvironment within or adjacent to the protein can lower Cys p*K*a and cause ionization of the thiol group. As a consequence, Cys residues will increase its reactivity. Because of its reactivity at physiological pH, Cys residues are often found in active site of several enzymes, such as proteases, oxidoreductases, and acetyl transferases. Besides to its function in catalysis, Cys residues in proteins are susceptible to several oxidative posttranslational oxidations (oxPTM) including formation of sulfenic (SOH), sulfinic (SO_2H), sulfonic (SO_3H) acids, as well as the covalent binding of the tripeptide glutathione (GSH) and nitric oxide (NO), respectively named S-glutathionylation (PSSG) and S-nitrosylation (PSNO). PSSG and PSNO have been considered redox "footprint" in proteins and an important mechanism of protein regulation [1, 2], similar to phosphorylation [3]. Whereas the appropriated chemical denomination for NO binding to nucleophilic centers in proteins is generally called nitrosation, in this chapter the NO binding to Cys will be referred as S-nitrosylation, in analogy to phosphorylation and other posttranslational modifications associated to cell signaling (Fig. 13.1) [4].

13.1.1.1 S-glutathionylation

Protein S-glutathionylation (PSSG) has been considered as a mechanism of redox-signal transduction as well as a cellular response to oxidative and/or nitrosative stress [5]. Interrelations between thiol oxidation and S-nitrosylation (PSNO) leading to the formation of mixed disulfides between Cys-protein thiols and the

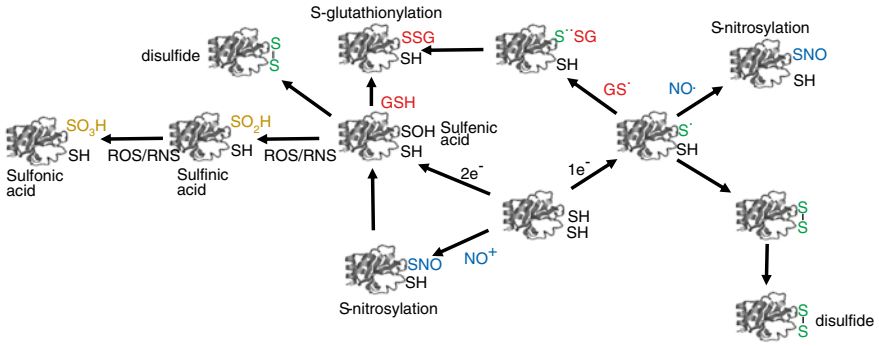


Fig. 13.1 Potential Cys-protein oxidation leading to different oxPTM. Thiol in proteins can undergo one electron oxidation and generate a thyil radical that can lead to S-glutathionylation or S-nitrosylation in presence of glutathione or NO, respectively. The thiyl radical can also facilitate interprotein or intraprotein disulfide. The two electron oxidation of thiol can generate sulfenic acid leading to S-glutathionylation in presence of glutathione. Alternatively, sulfenic acid can promote interprotein or intraprotein disulfide

tripeptide glutathione (GSH) may serve to transduce oxidative and nitrosative stimuli into a functional response at various levels of cellular signaling [6]. Stress-induced PSSG is poorly understood and several stimuli have been proposed to induce this oxPTM. PSSG can be induced in cells by mild oxidative stress or stimulated by cytokines [7]. In addition to understand what causes the formation of this oxPTM, it is also the subject of several studies to understand the chemical mechanisms by which PSSG occurs in biological systems. There are a number of in vitro studies using purified proteins and low molecular mass thiols proposing different mechanisms by which both PSSG and PSNO can occur in Cys residues within proteins. These in vitro studies intent to provide detailed information regarding these modifications and the chemical mechanisms of formation in a complex cell environment. The formation of PSSG in proteins does not occur directly by the addition of GSH to Cys thiol. As a redox reaction, one electron from protein-Cys (PSH) and another from GSH-Cys are needed to be lost at the sulfur atoms, producing respective thiyl radicals (S•) [6].

$$PS^{\bullet} + GS^{\bullet} \rightarrow PSSG$$

The ratio between reduced glutathione (GSH) and its respective oxidized pair (GSSG) has been considered as a central parameter of the cells redox state. Early study indicates that an increase in GSSG levels might be the main mechanism for PSSG formation [8]. In this case the mechanism involved would be the transfer of the modification between two different thiols (a process called "transglutathionylation").

$$PSH + GSSG \rightarrow PSSG + GSH$$

However, there are other publications arguing against it, since that in many situations GSSG levels do not increase and the GSH/GSSG levels are maintained. In fact, the intracellular concentration of reduced GSH in mammalian cells is in the millimolar range and the concentration of GSSG is less than 5 % of GSH. In addition to the arguments that intracellular GSSG does not increase substantially, the disulfide exchange between GSSG and another thiol (including PSH) is thermodynamically unfavored and considered a slow reaction due to the typical redox potentials for Cys residues [9]. Therefore S-glutathionylation of proteins by GSSG in vivo is not likely an efficient mechanism. Alternatively, peroxynitrite ($ONOO^-$), nitric oxide (NO), and endogenous low molecular weight S-nitrosothiols, such as S-nitrosoglutathione (GSNO), are thought to be the main agents leading to PSSG [10]. The nucleophilic attack of a protein thiolate on GSNO to produce a mixed disulfide is one of among several mechanisms in which protein S-glutathionylation can occur in vivo [11].

Regardless of the mechanism of formation, PSSG has been shown to be associated with regulation of the activity of several enzymes including carbonic anhydrase III [12], protein kinase C (PKC) [13], human aldose reductase [14], and human immunodeficiency virus, type I protease [15], glyceraldehyde-3-phosphate dehydrogenase (GAPDH) [5], papain [16], and the transcription factors AP-1 [17] and NF-κB [18] and its major regulator IKKβ [19]. In each case the effects of PSSG can be reversed by reducing agents. In biological systems, the degradation of PSSG is specifically catalyzed by a well-known enzymatic system, the glutaredoxins (GRX). The GRX enzymes, also known as thioltransferases, belong to the thioredoxin superfamily of enzymes and have been demonstrated to regulate the protein PSSG [20] GRXs in mammalian cells are oxidoreductases that catalyze the reduction of protein-glutathione mixed disulfides or deglutathionylation, using GSH as a cofactor [21]. Two mammalian GRX proteins have been identified to date. GRX1 is a cytosolic protein, whereas GRX2 contains a mitochondrial leader sequence but can also occur in the nucleus following alternative splicing [22, 23].

13.1.1.2 S-nitrosylation

Nitric oxide (NO) is an important signaling molecule, playing significant roles in physiology and pathophysiology [24]. By chemical definition, NO is a free radical, i.e., it has one unpaired electron in its outmost orbital, which makes it very unstable compared to many other chemical species (although it is relatively stable compared to other free radicals). As a free radical, NO does not only rapidly decompose but also reacts with other species, the reactivity of NO is due to its small size, high diffusion rate, and hydrophobicity. NO also undergoes reactions with oxygen (O_2), superoxide ions ($O_2^{-\bullet}$), and reducing agents to generate a variety of radical and non-radical derivatives, collectively named "reactive nitrogen species" (RNS). RNS include nitroxyl (HNO), N_2O_3, peroxynitrite, and S-nitrosothiols, such as *S*-nitrosoglutathione (GSNO). All these species show distinctive reactivity toward particular targets. RNS can modify lipids and DNA, as well as different protein

amino acid residues. The cellular activities of NO are transduced via multiple chemical reactions including the "classical" and direct interaction with heme and other metal centers in metalloproteins. RNS, on the other hand can also react with non-metal portions of the proteins, generating for the most part reversible posttranslational modifications. Of the particular interest are the reactions of NO with thiols on Cys residues that result in protein S-nitrosylation (PSNO) [25, 26]. PSNO has been thoroughly studied in the last two decades, because it represents an important oxPTM that can transduce NO-dependent signals and therefore modulate several signaling pathways [26, 27]. A large group of proteins have been characterized as targets for S-nitrosylation, including metabolic and structural proteins, ion channels, transcription factors, proteases, among others. S-nitrosylation in many cases is believed to regulate protein activity and function (reviewed in [26]). For example, S-nitrosylation has been shown to inhibit the activity of caspases [28]. Similarly, S-nitrosylation of inhibitory kappa B kinase (IKK) [29] and the p50 subunit of the transcription factor nuclear kappa B (NF-κB) are believed to be responsible for the NO-induced inhibition of NF-κB [30]. The mechanisms by which S-nitrosylation occurs on proteins are still the subject of investigation. Oxidized nitrogen species are the most important candidates for modifying a reduced thiol. N_2O_3 is the higher oxidation state for RNS, and it is generated by the reaction of NO with O_2, in particular in hydrophobic environments [31]. This RNS has been considered the typical S-nitrosylating species [32, 33].

$$\mathrm{PSH} + \mathrm{N_2O_3} \rightarrow \mathrm{PSNO} + \mathrm{H^+} + \mathrm{NO_2^-}$$

Another possible mechanism that can contribute PSNO formation in biological systems is transnitrosation. Transnitrosation reaction is the transfer of NO moiety between an S-nitrosothiol, such as GSNO to a thiol small molecule, and protein thiols [34].

$$\mathrm{PSH} + \mathrm{GSNO} \rightarrow \mathrm{PSNO} + \mathrm{SH}$$

Since GSNO can participate in transnitrosation reactions with proteins, modulation of GSNO levels should directly affect levels of protein S-nitrosylation. Several enzymes have been reported to degrade GSNO in vitro including protein disulfide isomerase, xanthine dehydrogenase/oxidase, superoxide dismutase (SOD), and glutathione peroxidase (GPRx) [35]. However, two enzymes have been shown to degrade GSNO and indirectly affect protein S-nitrosylation in vivo: GSNO reductase (also known as alcohol dehydrogenase class III or glutathione-dependent formaldehyde dehydrogenase) and carbonyl reductase 1. The only enzymatic system known so far that can directly denitrosylate protein thiols is the thioredoxin system (thioredoxin 1/thioredoxin reductase 1, Trx1/TrxR1). There is evidence that Trx1/TrxR1 catalyzes denitrosylation of proteins such as endothelial nitric oxide synthase (eNOS), caspase 3, caspase 8, caspase 9, protein tyrosine phosphatase 1B, glyceraldehyde 3-phosphate dehydrogenase (GAPD), and *N-ethylmaleimide-sensitive* factor [36–39]. More recent proteomics studies identified over a 100 putative substrates of Trx1 denitrosylation [34].

13.2 Detection and Quantification of oxPTM

In the past decade several approaches have been used to detect and identify oxPTM. Some approaches include techniques that specifically identify individual modifications such as PSSG, PSNO, sulfinic acid formation, among others. Other approaches are more global, measuring the overall "redox status" of Cys moieties in protein, and in this particular situation it is also possible to quantify the overall Cys oxidation (Fig. 13.2). In the more specific methods an antibody targeting the modification or a selective chemical label (or tag) can be used to identify a particular oxPTM. The advantage of this method is that the information can be helpful to evaluate the biological impact of a definite oxPTM, which can be associated with physiological or pathological condition. The drawback of this targeted approach is that other important modifications may be unnoticed. In addition, there are concerns about the specificity of antibodies toward very small epitopes, such as PSNO. To overcome the potential lack of specificity of antibodies a complementary approach utilizing chemical derivatization (labeling) of the oxPTM may be a complementary strategy [40, 41].

13.2.1 Cysteine-Labeling Strategies

Most of the Cys oxidations are fully reversible in vivo, because highly efficient enzymatic and nonenzymatic systems exist to control the redox status of proteins and maintain the cellular redox homeostasis. In addition, many of these modifications are relatively labile, and thiols themselves are prone to artifactual modification during protein isolation and labeling [42]. Therefore, capturing oxPTM in biological systems is very challenging, and both indirect and direct methods involving

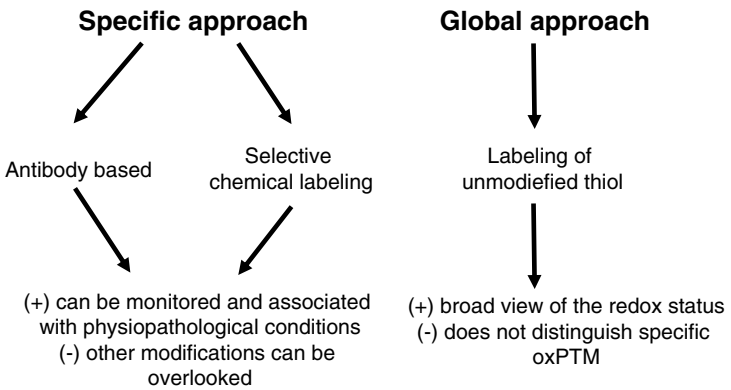


Fig. 13.2 Overview of the approaches utilized to detect and quantify oxPTM

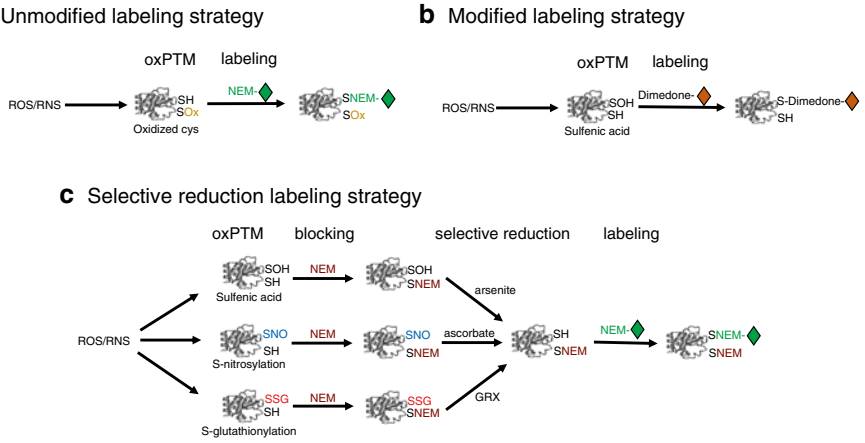


Fig. 13.3 General Cys-labeling strategies for detection of oxPTM. (**a**) Unmodified (reduced) Cys is labeled with a detectable probe and the modified (oxidized) Cys is not labeled. The decrease in labeling (and signal) indicates increases in the modification. (**b**) Chemical-specific detectable probe is used to detect a particular modification, such as sulfenic acid. (**c**) In the selective reduction labeling strategy, the free thiols are blocked with thiol-reactive NEM, prior to specific reduction, and labeling with a distinctive detectable probe

Cys chemical labeling are available. In general, the strategies to study oxPTM are complementary, followed by appropriated techniques of protein separation and identification. The reduced Cys in proteins can be chemically labeled utilizing a number of thiol-reactive compounds, including *N*-ethylmalemide (NEM), iodoacetamide (IAM), iodoacetic acid (IAA), and thiosulfates. These chemicals can have limited accessibility to some Cys moieties in proteins, and to improve the accessibility, the labeling can be performed in presence of sodium dodecyl sulfate (SDS) [43]. The reactivity of Cys is largely dependent on ionization to thiolate anion. Although the average p*K*a of thiol group in proteins is approximately 8.5, this value can range from 2.5 to 12 [4]. The majority of methods to evaluate changes in the redox state of Cys rely on the "loss of signal" due to Cys oxidation, or restoration of labeling by reducing agents, "gain of signal." In the "loss of signal" method, ROS or RNS oxidizes reactive thiols in proteins and a thiol-reactive compound is utilized to label the reduced "unmodified" thiols (Fig. 13.3a). As a result, increases in overall Cys oxidation (SOx) will cause a decrease in labeling and therefore in the signal. In the simplest and direct "gain of signal" approach, a selective probe can be used to label certain type of Cys oxidation, and therefore it will be a gain in signal when respective Cys oxidation increases. This direct approach is very popular to detect and identify proteins that show increases in sulfenic acid formation. For instance, a fluorescent tagged-dimedone (5,5-dimethyl-1,3-cyclohexanedione) or a dimedone derivative can be used to specifically react with sulfenic acid within proteins and label proteins that can be further purified and identified [44] (Fig. 13.3b).

Finally, there is the "gain signal" approach that is based on the restoration of oxidized Cys by specific reducing system, before thiol labeling. Consequently the increases in signal are proportional to Cys oxidation. In this particular case a "blocking" step with thiol-reactive compound prior to the specific reduction is necessary to distinguish between the endogenous thiols from the newly reduced ones. In addition, the blocking and labeling thiol-reactive compounds must differ to allow the distinction between the two pools of labeled thiols. Typically, the blocking agent is an unconjugated thiol-reactive compound (NEM) and the labeling is a thiol-reactive compound conjugated with biotin or a fluorophore, or isotopically modified. This method is typically used to identify PSNO and PSSG, as well as sulfenic acid [4, 40, 41]. Depending on the type of Cys oxidation, the reduction step can be performed using arsenite, ascorbate, or glutaredoxin to reduce specifically sulfenic acid [45], PSNO [40, 46], and PSSG [41], respectively (Fig. 13.3c).

13.2.2 Protein Separation and oxPTM Identification

Following one of these labeling strategies, the Cys-labeled proteins can now be assessed by a range of analytical procedures based on the nature of the label used as well as on the general objective of the study. The most low cost methods for Cys-labeled protein separation and identification are the gel-based methods followed by antibody detection. The simplest version among gel-based methods is the global analysis of Cys-labeled proteins. Typically a mixture of Cys-labeled proteins is separated by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and then blotted against the conjugated label. More sophisticated versions of this approach utilizes regular two-dimension gel electrophoresis (2-DE) or difference gel electrophoresis (DIGE), in which multiple samples can be labeled with fluorescent dyes prior to 2-DE separation [47]. The resulting blot or gel can be further submitted to imaging analysis for Cys-labeling quantification. In this manner, it is possible to analyze a global redox profile (redox proteomics) of a specific or nonspecific oxPTM, using modified and selected reduced labeling or unmodified labeling strategies, respectively. This global approach is commonly used for comparison between control and a pathological state or cell treatment that induces ROS and RNS. For a more specific analysis, a Cys-labeled protein can be initially immunoprecipitated from cell lysates or tissue homogenate followed by separation using SDS-PAGE, and blotted against biotin or any other conjugate, using streptavidin or antibodies. Thus it is possible to detect specific (modified labeling or selective reduced labeling) or nonspecific oxPTM (unmodified labeling) of a target protein. Alternatively, after immunoprecipitation and SDS-PAGE separation, Cys-labeled proteins can be submitted to mass spectrometry (MS) or liquid chromatography–mass spectrometry (LC/MS) separation and identification of labeled peptides for protein identification as well as determination of Cys moieties that are target for oxPTM. And finally, purified Cys-labeled proteins can be directly submitted to mass spec for identification of modified Cys residues (Fig. 13.4).

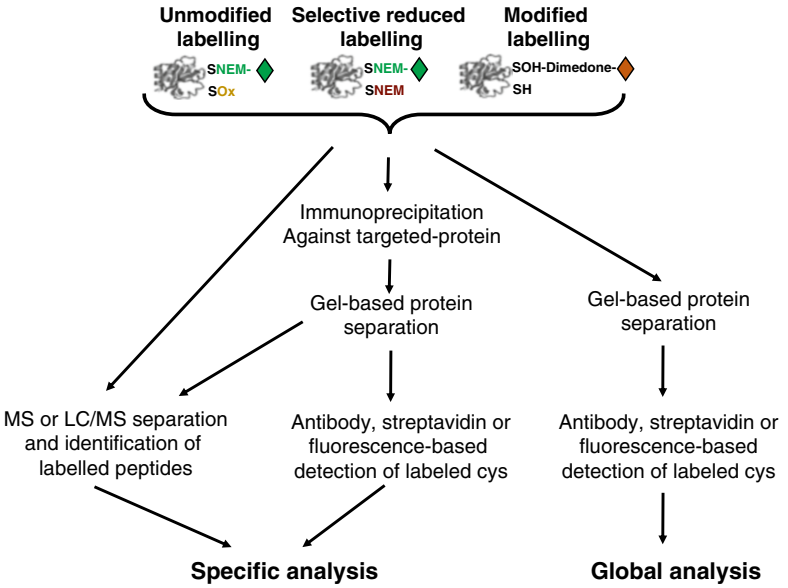


Fig. 13.4 Flow chart displaying the different methods utilized in protein separation and oxPTM identification, following Cys labeling

13.2.3 "In Situ" oxPTM Detection

Although valuable information can be gained from the detection, identification, and quantification of oxPTM in proteins using cell lysates or tissue homogenates, the ability to localize them in intact cells or tissues under physiological or pathological conditions will provide additional insights into the spatial relationship between oxPTM patterns and pathophysiological alterations. Although elegant studies have been published utilized antibodies to detect in situ protein PSNO, PSSG, and protein sulfenylation, antibodies directed against S-nitrosocysteine or any other small epitope may suffer from a lack of specificity. In addition, these oxPTM are very labile or rapidly can be regenerated by endogenous recovering systems, in particular during cell lysates or tissue homogenate preparations. Therefore, in situ oxPTM labeling can offer an alternative to overcome these difficulties as well as to provide valuable information regarding intracellular localization of oxPTM. So far, the approaches utilized to investigate oxPTM are based on Cys-labeling strategies described previously. The selective reduction of oxPTM with ascorbate or GRX system can be used to detect PSNO (Fig. 13.5) [40] or PSSG (Fig. 13.6) [41, 48] proteins, respectively in intact lung epithelial cells or in lung tissues. Given the limitations associated with some methods for detection of PSSG proteins, the catalytic activity of GRX1 to specifically reduce PSSG proteins in intact cells followed by

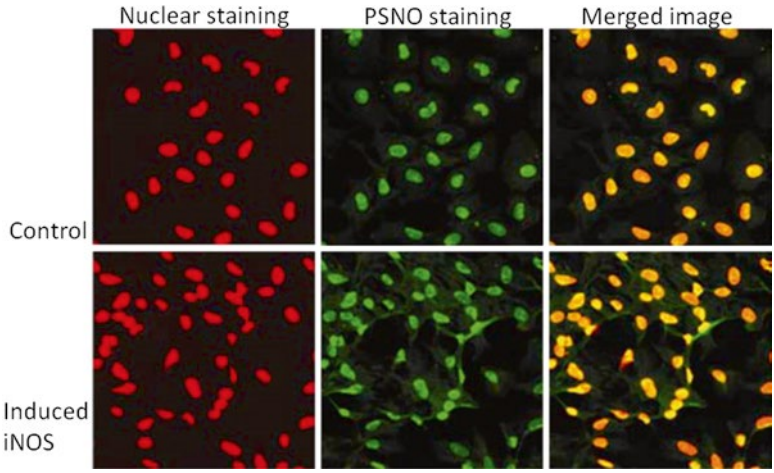


Fig. 13.5 Detection of PSNO in intact cells following induction of inducible nitric oxide synthase (iNOS). Cells were stimulated with tumor necrosis factor alpha (TNFα) and interferon gamma (IFNγ) to induce iNOS and increase RNS content. To visualize PSNO in intact cells, a "gain of signal" approach was utilized, as illustrated in Fig. 13.3c (*middle*). NEM-biotin labeling was used to detect PSNO. Streptavin-FITC (green fluorescence, *middle column*) was used to detect NEM-biotin and propidium iodide (PI) was used for nuclear staining (red fluorescence, *left column*). Confocal laser scanning analysis was utilized to visualize the images. Figure modified from previously published [3]

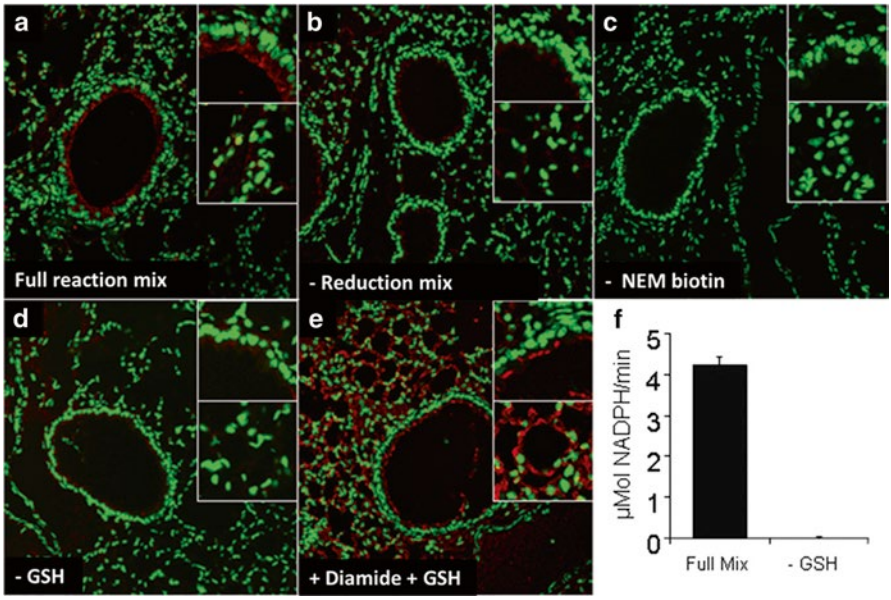


Fig. 13.6 Detection of PSSG in mouse lung tissue using Grx1-based Cysteine-specific reduction. To evaluate the localization of PSSG in intact cells, a "gain of signal" approach was utilized, as illustrated in Fig. 13.3c (*bottom*). NEM-biotin labeling was used to detect PSSG. Streptavin-Alexa

Cys labeling with NEM conjugated with fluorophore has been successfully used to specifically detect this oxPTM. Through the comprehensive implementation of diverse reagent controls, mouse GRX1 overexpression and knockdown approach the specificity of detection of S-glutathionylated proteins in intact cells and reveal their unique cellular localization in response to increases in ROS [41, 48].

As mentioned previously, dimedone and its derivatives can selectively react with sulfenic acid in Cys moieties, forming a stable derivatization product. A highly cell membrane permeable propyl azide dimedone derivative has been used to label sulfenic acid in intact cells [49]. This cell permeable dimedone and other dimedone derivatives can also allow to localized sulfenylated Cys in proteins [50], similar to in situ detection of PSNO and PSSG.

13.2.4 oxPTM Quantification

Depending on the overall objective of study, the global analysis with potential quantification of oxPTM in a particular pathological condition is more relevant than the identification of a targeted protein or Cys residues in proteins. The OxiCAT method combines isotope-coded affinity tag (ICAT) with differential thiol labeling to quantify oxPTM. Basically in this method, cells are lysed in presence of trichloroacetic acid (TCA) and other denaturants to fully expose Cys residues and prevent thiol-disulfide exchange. A lighter isotope (^{12}C) ICAT reagent can be used to label the unmodified Cys followed by reduction of all reversible Cys oxidation (PSNO, PSSG, sulfenylation, disulfides) with Tris (2-carboxyethyl) phosphine (TCPE). The resulting newly reduced thiols are labeled with the heavier isotope (^{13}C) ICAT reagent. Then all proteins are trypsinized and purified and the ratio of light to heavy isotope-labeled Cys can be quantified by LC–MS (Fig. 13.7) [51].

13.3 Clinical Applications of Redox Proteomics

Reactive oxygen and nitrogen species (ROS/RNS) are implicated in the regulation of several proteins involved in a wide range of biological processes, including gene expression and protein translation, basic metabolism, cell signaling, and apoptosis.

Fig. 13.6 (continued) fluor 569 (red fluorescence) was used to detect NEM-biotin and Sytox green was used for nuclear staining (green fluorescence). Confocal laser scanning analysis was utilized to visualize the images. (**a**) Basal level of PSSG in control mouse lung tissue. (**b**) Omission of reduction mix (negative control). (**c**) Omission of NEM-biotin-labeling agent (negative control). (**d**) Omission of GSH from the reduction mix (negative control). (**e**) Pretreatment of tissue with diamide and GSH to induce PSS (positive control). (**f**) Assessment of Grx1 activity in vitro using a full enzymatic mix or omitting GSH as a negative control. Figure modified from previously published [2]

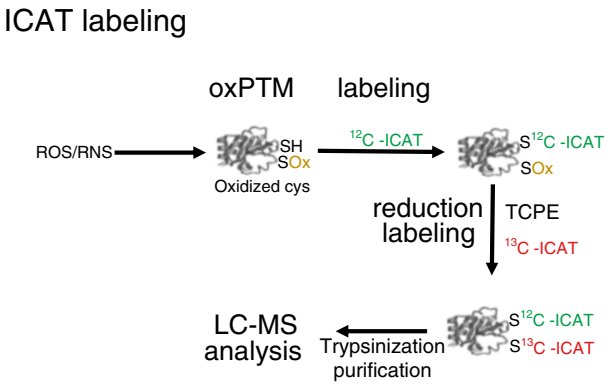


Fig. 13.7 Differential Cys labeling for oxPTM quantification. A lighter isotope (^{12}C) ICAT reagent is used to label the unmodified Cys. A reducing reagent (TCPE) reduces all reversible Cys oxidation (PSNO, PSSG, sulfenylation, disulfides). Then the newly reduced thiols are labeled with the heavier isotope (^{13}C) ICAT reagent for further analysis by LC–MS

Therefore, it is expected that changes in cellular redox homeostasis are associated with many different pathological conditions, including neurodegenerative illnesses, and cardiovascular diseases [52]. In recent years, a role for redox proteomics analysis has emerged with the intent of exploring alterations of redox status which target protein species in several diseases.

The brain tissue is very susceptible to oxidative damage because its high rate of oxygen consumption, high content of unsaturated lipids and relative lack of antioxidant enzymes. Therefore it is not a surprise that high ROS input is associated with several prominent central nervous system disorders [53], such as amyotrophic lateral sclerosis (ALS, also referred to as Lou Gehrig's disease), Alzheimer's disease (AD), Parkinson's disease (PD), and Huntington's disease (HD). ALS is a neurodegenerative disorder that results in loss of motor neurons, leading to a rapidly progressive form of muscle paralysis and subsequent death [54]. There is no available cure, and current therapies only provide minimal relief of the symptoms. In addition, to corroborate with this devastating scenario, the molecular mechanism of development of ALS is still unknown. Only 20 % of ALS patients are afflicted by familial ALS, which is associated with mutation of Cu,Zn-Superoxide dismutase 1 (SOD1), indicating that perturbations in redox homeostasis may play a role in the mechanism of these disease. The first evidence that ROS play a role in the etiology of this disease came from early studies demonstrating that protein carbonyls were increased as much as 119 % in spinal cord tissue of sporadic ALS patients [55]. Several independent studies also demonstrated that there is a consistent increase in protein nitration (nitrotyrosine) in ALS patients, which are classical biomarkers for oxidative stress, consistent with high input of ROS and/or RNS (reviewed in [56]). In addition, in vivo studies using mutant superoxide dismutase (G93A-SOD1) transgenic mice that develop ALS-like motor neuron disease demonstrated that four different proteins in the spinal cord of transgenic mice have higher specific proteins carbonyl compared to those of non-transgenic mice. These proteins are SOD1, translationally controlled tumor protein (TCTP), ubiquitin carboxyl-terminal hydrolase-L1 (UCH-L1), and, possibly, αB-crystallin, suggesting that oxidative

modification of these proteins may play an important role in the neurodegeneration of ALS [57].

Another common neurodegenerative disease that is consistently associated with augmentation of ROS production in the brain is Alzheimer's disease (AD). AD affects over 30 million people all around the globe and is marked by behavioral and cognitive impairment. Clinically, AD is a progressive neurodegenerative dementia characterized by the presence of senile plaques, consisting of deposits of β-amyloid peptide (Aβ) [58]. In addition to the accumulation of extracellular beta-amyloid plaques, AD is also assumed to be caused by neurofibrillary tangles of abnormally insoluble hyperphosphorylated Tau protein [59–61]. As a result of increased ROS and RNS production in AD brain, a significant higher levels of protein carbonyls and 3-nitrotyrosine is observed [62], similar to what was found in ALS patients. More recently publication from the same group discusses on the proteins that are identified, by redox proteomics, as selective targets of 4-hydroxy-2-nonenal (HNE) modification during the progression and pathogenesis of AD. Among the oxidatively modified proteins identified in AD are proteins that play important roles in regulating energy metabolism, cellular signaling, pH regulation, neuronal communication, antioxidant and detoxification, neurotransmitter regulation, and tau hyperphosphorylation [63].

Parkinson's disease (PD)-affected patients suffer from progressive loss of dopaminergic neurons in basal ganglia and the substantia nigra, accumulation of Lewy bodies. Clinically, this disease is manifested by a complex and variable motor and non-motor symptoms. ROS and RNS play important role in PD, indicated by dopamine cell degeneration and mitochondrial dysfunction and excitotoxicity. The role of proteomics in detecting abnormal proteins in PD has been recently reviewed by Licker and collaborators [64]. Using 2-DE analysis Basso and collaborators [65] found several different proteins expressed between PD and control group. Out of a total of 37 identified proteins, they found 16 differentially regulated, including elements of iron metabolism (ferritin H), glutathione-related redox metabolism (glutathione *S*-transferase M3, P1, O1), or novel redox proteins such as SH3BGRL [66].

Although the central nervous system represents the most susceptible target to oxidative stress, production of high levels of ROS and RNS in peripheral blood is known to have a negative impact to the overall cardiovascular system. A recent study using animal model of ischemia and reperfusion and an ICAT-based proteomic approach revealed that of the nearly 200 Cys-containing peptides that we detected in all treatment groups, about a quarter became reversibly oxidized when blood flow was first reestablished. However, oxidation was no longer evident after longer reperfusion, most probably because of recovery of oxidized Cys. In contrast, protein carbonyls accumulated during reperfusion. A majority of the thiol changes were in mitochondrial proteins and the study provides support for mitochondrial oxidative stress occurring primarily during the early stage of reperfusion [67].

Although, the technologies available for proteomic studies related to diseases or pathological conditions have improved considerably in the past decade, the search for consistent differences in protein profile and reliable biomarkers between patients and normal controls is still an ongoing battle. However, solid progress has been made in the proteomics field, in particular regarding neurodegenerative diseases

despite the challenges in obtaining biological material to perform the analyses. Typically to study neurodegenerative illnesses in humans, most studies have relied on samples of cerebrospinal fluid (CSF) obtained from living patients or brain tissue obtained at autopsy. In contrast, in animal models of diseases sampling can be available as needed regarding type, number, quantity, and quality of samples. Another limiting factor is the sample preparation or stability, in particular for studies where the main purpose is to detect oxidation in proteins, such Cys oxidation (oxPTM). The main limitation for this kind of study is the sample preparation due to the reversible nature of the oxidative modifications on the Cysteine thiols, relying on Cysteine labeling. Because of the advances in the technology, redox proteomics is an emerging branch of proteomics that can identify different types of oxidation in target proteins with potential use as biomarkers. This global analysis of oxPTM can provide knowledge to the scientists to develop alternative therapies for illnesses related to oxidative stress.

References

1. Weerapana E et al (2010) Quantitative reactivity profiling predicts functional cysteines in proteomes. Nature 468(7325):790–795
2. Reddie KG, Carroll KS (2008) Expanding the functional diversity of proteins through cysteine oxidation. Curr Opin Chem Biol 12(6):746–754
3. Mannick JB, Schonhoff CM (2002) Nitrosylation: the next phosphorylation? Arch Biochem Biophys 408(1):1–6
4. Leonard SE, Carroll KS (2011) Chemical 'omics' approaches for understanding protein cysteine oxidation in biology. Curr Opin Chem Biol 15(1):88–102
5. Mohr S et al (1999) Nitric oxide-induced S-glutathionylation and inactivation of glyceraldehyde-3-phosphate dehydrogenase. J Biol Chem 274(14):9427–9430
6. Martinez-Ruiz A, Lamas S (2007) Signalling by NO-induced protein S-nitrosylation and S-glutathionylation: convergences and divergences. Cardiovasc Res 75(2):220–228
7. Thomas JA, Poland B, Honzatko R (1995) Protein sulfhydryls and their role in the antioxidant function of protein S-thiolation. Arch Biochem Biophys 319(1):1–9
8. Cotgreave IA, Gerdes RG (1998) Recent trends in glutathione biochemistry—glutathione-protein interactions: a molecular link between oxidative stress and cell proliferation? Biochem Biophys Res Commun 242(1):1–9
9. Dalle-Donne I et al (2005) S-glutathionylation in human platelets by a thiol-disulfide exchange-independent mechanism. Free Radic Biol Med 38(11):1501–1510
10. Percival MD et al (1999) Inhibition of cathepsin K by nitric oxide donors: evidence for the formation of mixed disulfides and a sulfenic acid. Biochemistry 38(41):13574–13583
11. Ji Y et al (1999) S-nitrosylation and S-glutathiolation of protein sulfhydryls by S-nitroso glutathione. Arch Biochem Biophys 362(1):67–78
12. Cabiscol E, Levine RL (1995) Carbonic anhydrase III. Oxidative modification in vivo and loss of phosphatase activity during aging. J Biol Chem 270(24):14742–14747
13. Ward NE et al (1998) Irreversible inactivation of protein kinase C by glutathione. J Biol Chem 273(20):12558–12566
14. Cappiello M et al (1996) Specifically targeted modification of human aldose reductase by physiological disulfides. J Biol Chem 271(52):33539–33544
15. Davis DA et al (1996) Regulation of HIV-1 protease activity through cysteine modification. Biochemistry 35(7):2482–2488

16. Xian M et al (2000) Inhibition of papain by S-nitrosothiols. Formation of mixed disulfides. J Biol Chem 275(27):20467–20473
17. Klatt P, Lamas S (2002) c-Jun regulation by S-glutathionylation. Methods Enzymol 348:157–174
18. Pineda-Molina E, Lamas S (2002) S-glutathionylation of NF-kappa B subunit p50. Methods Enzymol 359:268–279
19. Reynaert NL et al (2006) Dynamic redox control of NF-kappaB through glutaredoxin-regulated S-glutathionylation of inhibitory kappaB kinase beta. Proc Natl Acad Sci U S A 103(35):13086–13091
20. Shelton MD, Kern TS, Mieyal JJ (2007) Glutaredoxin regulates nuclear factor {kappa}-B and intercellular adhesion molecule in Muller cells: model of diabetic retinopathy. J Biol Chem 282(17):12467–12474
21. Gravina SA, Mieyal JJ (1993) Thioltransferase is a specific glutathionyl mixed disulfide oxidoreductase. Biochemistry 32(13):3368–3376
22. Fernando MR et al (2006) Mitochondrial thioltransferase (glutaredoxin 2) has GSH-dependent and thioredoxin reductase-dependent peroxidase activities in vitro and in lens epithelial cells. FASEB J 20(14):2645–2647
23. Su D, Gladyshev VN (2004) Alternative splicing involving the thioredoxin reductase module in mammals: a glutaredoxin-containing thioredoxin reductase 1. Biochemistry 43(38):12177–12188
24. Kerwin JF Jr, Lancaster JR Jr, Feldman PL (1995) Nitric oxide: a new paradigm for second messengers. J Med Chem 38(22):4343–4362
25. Eu JP et al (2000) An apoptotic model for nitrosative stress. Biochemistry 39(5):1040–1047
26. Stamler JS, Lamas S, Fang FC (2001) Nitrosylation. The prototypic redox-based signaling mechanism. Cell 106(6):675–683
27. Hess DT et al (2001) S-nitrosylation: spectrum and specificity. Nat Cell Biol 3(2):E46–E49
28. Mannick JB et al (1999) Fas-induced caspase denitrosylation. Science 284(5414):651–654
29. Reynaert NL et al (2004) Nitric oxide represses inhibitory kappaB kinase through S-nitrosylation. Proc Natl Acad Sci U S A 101(24):8945–8950
30. Marshall HE, Stamler JS (2001) Inhibition of NF-kappa B by S-nitrosylation. Biochemistry 40(6):1688–1693
31. Moller MN et al (2007) Acceleration of nitric oxide autoxidation and nitrosation by membranes. IUBMB Life 59(4–5):243–248
32. Wink DA et al (1993) Reactions of the bioregulatory agent nitric oxide in oxygenated aqueous media: determination of the kinetics for oxidation and nitrosation by intermediates generated in the NO/O2 reaction. Chem Res Toxicol 6(1):23–27
33. Hogg N (2002) The biochemistry and physiology of S-nitrosothiols. Annu Rev Pharmacol Toxicol 42:585–600
34. Smith BC, Marletta MA (2012) Mechanisms of S-nitrosothiol formation and selectivity in nitric oxide signaling. Curr Opin Chem Biol 16(5–6):498–506
35. Benhar M, Forrester MT, Stamler JS (2009) Protein denitrosylation: enzymatic mechanisms and cellular functions. Nat Rev Mol Cell Biol 10(10):721–732
36. Forrester MT et al (2009) Detection of protein S-nitrosylation with the biotin-switch technique. Free Radic Biol Med 46(2):119–126
37. Benhar M et al (2008) Regulated protein denitrosylation by cytosolic and mitochondrial thioredoxins. Science 320(5879):1050–1054
38. Sengupta R et al (2009) Nitric oxide and dihydrolipoic acid modulate the activity of caspase 3 in HepG2 cells. FEBS Lett 583(21):3525–3530
39. Ito T, Yamakuchi M, Lowenstein CJ (2011) Thioredoxin increases exocytosis by denitrosylating N-ethylmaleimide-sensitive factor. J Biol Chem 286(13):11179–11184
40. Ckless K et al (2004) In situ detection and visualization of S-nitrosylated proteins following chemical derivatization: identification of Ran GTPase as a target for S-nitrosylation. Nitric Oxide 11(3):216–227
41. Reynaert NL et al (2006) In situ detection of S-glutathionylated proteins following glutaredoxin-1 catalyzed cysteine derivatization. Biochim Biophys Acta 1760(3):380–387

42. Chouchani ET et al (2011) Proteomic approaches to the characterization of protein thiol modification. Curr Opin Chem Biol 15(1):120–128
43. Charles R, Jayawardhana T, Eaton P (2014) Gel-based methods in redox proteomics. Biochim Biophys Acta 1840(2):830–837
44. Thamsen M, Jakob U (2011) The redoxome: proteomic analysis of cellular redox networks. Curr Opin Chem Biol 15(1):113–119
45. Saurin AT et al (2004) Widespread sulfenic acid formation in tissues in response to hydrogen peroxide. Proc Natl Acad Sci U S A 101(52):17982–17987
46. Jaffrey SR, Snyder SH (2001) The biotin switch method for the detection of S-nitrosylated proteins. Sci STKE 2001(86):pl1
47. Chan HL, Sinclair J, Timms JF (2012) Proteomic analysis of redox-dependent changes using cysteine-labeling 2D DIGE. Methods Mol Biol 854:113–128
48. Aesif SW et al (2009) In situ analysis of protein S-glutathionylation in lung tissue using glutaredoxin-1-catalyzed cysteine derivatization. Am J Pathol 175(1):36–45
49. Leonard SE, Reddie KG, Carroll KS (2009) Mining the thiol proteome for sulfenic acid modifications reveals new targets for oxidation in cells. ACS Chem Biol 4(9):783–799
50. Charles RL et al (2007) Protein sulfenation as a redox sensor: proteomics studies using a novel biotinylated dimedone analogue. Mol Cell Proteomics 6(9):1473–1484
51. Leichert LI et al (2008) Quantifying changes in the thiol redox proteome upon oxidative stress in vivo. Proc Natl Acad Sci U S A 105(24):8197–8202
52. Valko M et al (2007) Free radicals and antioxidants in normal physiological functions and human disease. Int J Biochem Cell Biol 39(1):44–84
53. Lehtinen MK, Bonni A (2006) Modeling oxidative stress in the central nervous system. Curr Mol Med 6(8):871–881
54. Perry JJ, Shin DS, Tainer JA (2010) Amyotrophic lateral sclerosis. Adv Exp Med Biol 685:9–20
55. Yan LJ, Orr WC, Sohal RS (1998) Identification of oxidized proteins based on sodium dodecyl sulfate-polyacrylamide gel electrophoresis, immunochemical detection, isoelectric focusing, and microsequencing. Anal Biochem 263(1):67–71
56. Beal MF (2002) Oxidatively modified proteins in aging and disease. Free Radic Biol Med 32(9):797–803
57. Poon HF et al (2005) Redox proteomics analysis of oxidatively modified proteins in G93A-SOD1 transgenic mice—a model of familial amyotrophic lateral sclerosis. Free Radic Biol Med 39(4):453–462
58. Fraser PE, Levesque L, McLachlan DR (1993) Biochemistry of Alzheimer's disease amyloid plaques. Clin Biochem 26(5):339–349
59. Hardy J et al (2006) Tangle diseases and the tau haplotypes. Alzheimer Dis Assoc Disord 20(1):60–62
60. Binder LI et al (2005) Tau, tangles, and Alzheimer's disease. Biochim Biophys Acta 1739(2–3):216–223
61. Williams DR (2006) Tauopathies: classification and clinical update on neurodegenerative diseases associated with microtubule-associated protein tau. Intern Med J 36(10):652–660
62. Butterfield DA, Perluigi M, Sultana R (2006) Oxidative stress in Alzheimer's disease brain: new insights from redox proteomics. Eur J Pharmacol 545(1):39–50
63. Sultana R, Perluigi M, Allan Butterfield D (2013) Lipid peroxidation triggers neurodegeneration: a redox proteomics view into the Alzheimer disease brain. Free Radic Biol Med 62:157–169
64. Licker V et al (2009) Proteomics in human Parkinson's disease research. J Proteomics 73(1):10–29
65. Basso M et al (2004) Proteome analysis of human substantia nigra in Parkinson's disease. Proteomics 4(12):3943–3952
66. Werner CJ et al (2008) Proteome analysis of human substantia nigra in Parkinson's disease. Proteome Sci 6:8

67. Kumar V et al (2013) Redox proteomics of thiol proteins in mouse heart during ischemia/reperfusion using ICAT reagents and mass spectrometry. Free Radic Biol Med 58:109–117
68. Walsh CT, Garneau-Tsodikova S, Gatto GJ Jr (2005) Protein posttranslational modifications: the chemistry of proteome diversifications. Angew Chem Int Ed Engl 44(45):7342–7372
69. Levine RL, Stadtman ER (2001) Oxidative modification of proteins during aging. Exp Gerontol 36(9):1495–1502
70. Di Domenico F et al (2011) Circulating biomarkers of protein oxidation for Alzheimer disease: expectations within limits. Biochim Biophys Acta 1814(12):1785–1795
71. Marino SM, Gladyshev VN (2010) Cysteine function governs its conservation and degeneration and restricts its utilization on protein surfaces. J Mol Biol 404(5):902–916

Chapter 14
Analysis of Fluorinated Proteins by Mass Spectrometry

Linda A. Luck

Abstract ^{19}F NMR has been used as a probe for investigating bioorganic and biological systems for three decades. Recent reviews have touted this nucleus for its unique characteristics that allow probing in vivo biological systems without endogenous signals. ^{19}F nucleus is exceptionally sensitive to molecular and microenvironmental changes and thus can be exploited to explore structure, dynamics, and changes in a protein or molecule in the cellular environment. We show how mass spectrometry can be used to assess and characterize the incorporation of fluorine into proteins. This methodology can be applied to a number of systems where ^{19}F NMR is used.

14.1 Background

^{19}F NMR has become a useful tool in the study of protein structure and dynamics for the last three decades [1, 2]. Although high-resolution X-ray and multidimensional NMR analysis provide the best structural information, all proteins are not amenable to these techniques. ^{19}F NMR provides data for these proteins and compliments the 3D structural data by monitoring specific locations within a known structure. Fluorine labels can be incorporated into proteins biosynthetically or chemically into specific sites. New frontiers in the area of ^{19}F NMR have been progressing in the area of pharmaceuticals to reveal drug uptake, biodistribution, and metabolism in vivo [3]. The fact that there are not endogenous biological fluorines allows good detection of the probes. The nucleus also has much promise for application in molecular biology in preclinical trial in animals that can be ultimately applied to humans.

L.A. Luck, Ph.D. (✉)
Department of Chemistry, State University of New York
at Plattsburgh, 101 Broad Street, Plattsburgh, NY 12901, USA
e-mail: luckla@plattsburgh.edu

A.G. Woods and C.C. Darie (eds.), *Advancements of Mass Spectrometry in Biomedical Research*, Advances in Experimental Medicine and Biology 806,
DOI 10.1007/978-3-319-06068-2_14, © Springer International Publishing Switzerland 2014

^{19}F nucleus has a spin of ½ and occurs at 100 % natural abundance which gives nearly the same sensitivity of the ^{1}H nucleus. The most notable of its features is the fact that there is a lack of fluorine in biological samples, thus there are no background signals to contend with as there are in ^{1}H NMR samples which may be nearly 90 % water. The spectra are generally simple in terms of acquisition times and usually require lower protein concentrations than those needed for high resolution multidimensional NMR studies. An additional bonus for the ^{19}F NMR studies is that the chemical shift range for the ^{19}F nucleus is 30 times larger than ^{1}H, which also adds to the simplicity of the spectra since most resonances are not overlapped. An aromatic fluorine nucleus such as one replacing a ^{1}H on a tryptophan, phenylalanine, or tyrosine residue is particularly susceptible to electrostatic effects and shows the highest dispersion of the ^{19}F NMR shifts. In addition the ^{19}F nucleus has lone pair electrons which allow exquisite sensitivity to local microenvironment. Paramagnetic probes have large affects on the fluorine nucleus and have been used for solvent accessibility studies [4–6]. These studies provide insight not obtained by the other structural techniques.

The ^{19}F nucleus is small and can be exchanged for a proton or a hydroxyl on an amino acid. A host of fluorinated amino acids (e.g., F-tryptophan, F-tyrosine, difluoromethionine, trifluoroleucine, pentafluoroleucine, F-proline, and F-serine) are commercially available. Fluorinated amino acids are usually biosynthetically incorporated into proteins by the expression of the cloned gene with the ^{19}F-labeled amino acid in the media. In addition, new methodology has surfaced in the past year where fluoroindole can be utilized in the media for biosynthesis of fluorotryptophans [7]. In this case and in other biosynthetic methods the extent of incorporation is not 100 % because unlabeled amino acids are in the media or are required during growth of the bacteria. Although the percent of incorporation of fluorine may be less than 100, all sites incorporate equally well, thus the NMR spectra are useful [5, 8]. A number of methodologies have been used to increase the percent incorporation of the ^{19}F label in the protein. These include using an auxotrophic bacterial strain that limits the synthesis of amino acid endogenously, adding label after induction of protein synthesis, and addition of glyphosate which inhibits aromatic protein synthesis. Knowing the percentage of fluorine labeling is crucial to many experiments (e.g., 2D ^{19}F Heteronuclear and Homonuclear NMR) and the effects of ^{19}F labeling on protein structure and activity must be studied in order to draw conclusions regarding the native structure when using ^{19}F NMR studies.

If a protein is 100 % labeled with fluorine then all sites contain the specific fluorinated amino acid. If the fluorine labeling is less than 100 % the proteins are a heterogeneous mixture. If for example the protein is being labeled with F-serine and there are six serine residues; one protein may have six sites labeled with fluorine and another protein may have only one site labeled. Assessment of the detrimental effects of the fluorine labeling may be misleading.

The immediate consequences for a switch of a proton atom to the more electronegative fluorine atom in proteins are the C–F bond is polarized in the opposite direction to the C–H bond and the C–F bond is more stable. In addition the C–F bond is less polarizable than the C–H bond. In an aliphatic chain the C–F bond may be a weak hydrogen bond acceptor but in a ring structure the C–F is less likely to have

this characteristic. The switch from an H to F is considered isosteric with F slightly larger than H. Fluorine substitutions have been made to increase the hydrophobicity of pharmaceuticals and improve their bioavailability [9]. Thus changes with the fluorine substitution have some bearing on the immediate environment of the ^{19}F nucleus and possibly the structure and dynamics of the protein. Each protein has its own unique character and the different nature of the fluorine label employed in the study can result in perturbations or fluorinated proteins without large deleterious effects.

There have been a number of ways to analyze the incorporation of labeled amino acids into proteins. In the case of tryptophan analogues spectroscopic methods based on the absorption spectrum and chromatographic separations can give insight into the percentage of analogue incorporation [10]. Another means to quantitate the extent of fluorine incorporation into a protein is to compare integrals of a well-resolved peak in the ^{19}F NMR spectrum to the integral of an internal standard [8]. This methodology can be used with the NMR standard in the sample or in an insert in the NMR tube during the experiment. We show a study where the internal standard actually bound to the protein of interest (vide supra). Thus this methodology is not always useful. Mass spectrum analysis has been used for analysis of fluorinated tryptophan-labeled proteins and presents itself as a method of choice for other fluorinated amino acid analogues not amenable to fluorescence studies. In light of the recent interest in designing highly fluorinated hydrophobic proteins with the specific goal of enhancing protein stability and other chemical properties, a variety of fluorinated analogues of leucine, valine, isoleucine, glycine, and other aliphatic amino acids have been utilized [11, 12]. Mass spectrum analysis hails as the best methodology by providing a general strategy for quantitative data for analogue incorporation for any of the amino acids. Substitution of fluorine for proton which is a change in mass of 18 Da is easily recognizable in the mass spectrum of proteins <70 kDa. The purpose of this chapter is to report how we have used mass spectral analysis to determine fluorine analogue incorporation into proteins and show how mass spectrometry was used to identify specific mutations in the proteins that identified peaks in the ^{19}F NMR spectrum.

14.2 Proteins

Two proteins will be discussed in this chapter to illustrate the utility of mass spectrometry as a means of analyzing fluorine incorporation into proteins and mutants produced to identify peaks in the ^{19}F NMR spectrum. The *Escherichia coli* D-galactose and D-glucose receptor (GGR) is an aqueous sugar-binding protein and the first component in the chemosensory and transport pathways for these sugars. The protein which is very stable, well characterized by structural studies with mutations easily made and tolerated by the protein structure [8]. Glucose measurements especially for blood sugar in diabetes are crucial, thus GGR has been extensively investigated as a platform for a variety of electrochemical biosensors for the detection of blood glucose levels [13–17]. This utilitarian protein has been used as our model for investigation of a number of fluorinated analogues including 4, 5, and 6 fluorotryptophan 2, 3 and 4 fluorophenylalinine, tetradeutero-5-fluorotryptophan and recently

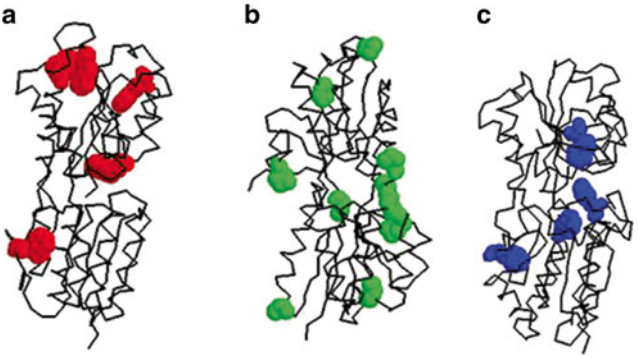


Fig. 14.1 Structure of the *E. coli* D-galactose and D-glucose receptor with the Trp residues highlighted (**a**) and the Pro residues highlighted (**b**). Structure of the *E. coli* L-leucine-specific receptor with the Trp residues highlighted (**c**) [23, 25]

H-*cis*-4-fluoro-proline (4F-*cis*-Pro) and H-*trans*-4-fluoro-proline (4F-*trans*-Pro) [18]. Figure 14.1a identifies the positions of the five tryptophan residues in GGR. Figure 14.1b illustrates the position of the ten proline residue sites in GGR.

The *Escherichia coli* leucine-specific protein (LBP) is a component of the transport pathway for hydrophobic amino acids. This protein as with the GGR is stable and easily manipulated for a number of biophysical studies. The LBP protein has been by ^{19}F NMR and fluorescence along with its counterpart the L-leucine, isoleucine, valine protein (LIV) [19–21]. LBP and LIV are models for the ionotropic glutamate receptors that control a wide variety of normal neuronal processes including learning and memory. Activation of these neurotransmitter receptors is involved in a number of neurodegenerative diseases, notably stroke and epilepsy. Study of the model proteins will allow us to understand the basic structural changes that contributing to the pathological effects of stroke and epilepsy, and shed light on some factors that will be beneficial in increasing cognition in our aging population. Thus we are pursuing studies on these proteins as platforms for biosensors to detect a variety of substances that may be toxic to our neurosystem [22]. Our laboratory was the first to report the affinity of LBP for phenylalanine and solved the closed structure by X-ray with phenylalanine in the binding site [23, 24]. The discovery of the phenylalanine affinity for LBP was made when 4-F-phenylalanine (4F-Phe) was used as an internal standard in the ^{19}F NMR spectrum to determine incorporation of 5-F-tryptophan (5F-Trp) in LBP (vide supra). Figure 14.1c illustrates the four positions of the tryptophan residues in closed form of LBP with leucine in the binding pocket.

14.3 Production of the Fluorine-Labeled Proteins

The production of the 4F-Trp-labeled GGR was accomplished in the strain *E. coli* NM303, a non auxotroph strain with 50 μg/mL 4F-Trp. The production of 4F-c*is*-Pro-labeled GGR utilized the compound H-*cis*-4-fluoro-proline, purchased from

BACHEM. Fifty micrograms per milliliter 4F-*cis*-Pro was used in the media and isolation of both of the fluorinated proteins was performed in the same manner as the wild-type protein [8]. The techniques and materials to generate the mutant LBP proteins and 5F-Trp-labeled LBP mutants have been described previously [19, 20, 23, 24].

14.4 NMR Measurements

NMR sample contained 0.6 mL of 1.9 mM LBP in 10 mM phosphate pH 6.9 with 10 % D_2O as lock. Spectra were obtained on a 500 MHz NMR with at 25 °C. Processing used 25 Hz linebroadening.

14.5 Mass Spectrometry Measurements

Samples were analyzed using an AB Sciex 4000 QTrap (AB Sciex, Framingham, MA) hybrid triple quadrupole/linear ion trap liquid chromatography–mass spectrometer (LCMS). Positive electrospray ionization (ESI) was used as the ionization source. Source temperature was maintained at 300 °C. Sheath gas (GS1) flow was set at 30, auxiliary gas flow (GS2) at 40, curtain gas flow at 25, and the declustering potential was set to 80. The mass spectrometer was operated in single quadrupole mode, scanning from *m*/*z* 700–1400. AB Sciex Biotools Beyesian Protein Reconstruct software was used for spectral deconvolution.

Analytes were separated using a Shimadzu Prominence high performance liquid chromatography (HPLC) system (Shimadzu Scientific Instruments, Columbia, MD) across water to acetonitrile (ACN) gradient, using 0.1 % formic acid and 0.025 % trifluoroacetic acid as ion pairing reagents. At the beginning of each run, the mobile phase was held at 30 % ACN for 3 min, increased to 70 % over 20 min, increased to 95 % over 2 min, and held at 95 % for 5 min. The LC column was then requilibrated to the initial conditions for 15 min. Flow was maintained at 80 μL/min. One microliter of each sample was injected onto a Phenomenex (Torrance, CA) Jupiter C4 300A reversed phase HPLC column (150 mm × 1 mm × 5 μm).

14.6 Mass Spectrometry of 4-Fluorotryptophan and 4F-*cis*-Pro-Labeled GGR

The ^{19}F NMR studies of 5F-Trp-labeled GGR showed great utility for the conformational changes that occurred during ligand binding [4, 8, 26]. The protein is easily manipulated and we have further explored the incorporation of a number of fluorine and carbon-13 labels. One question that plagued us during these studies was the multiplicity of the ^{19}F NMR peaks near the calcium binding site [26]. To address

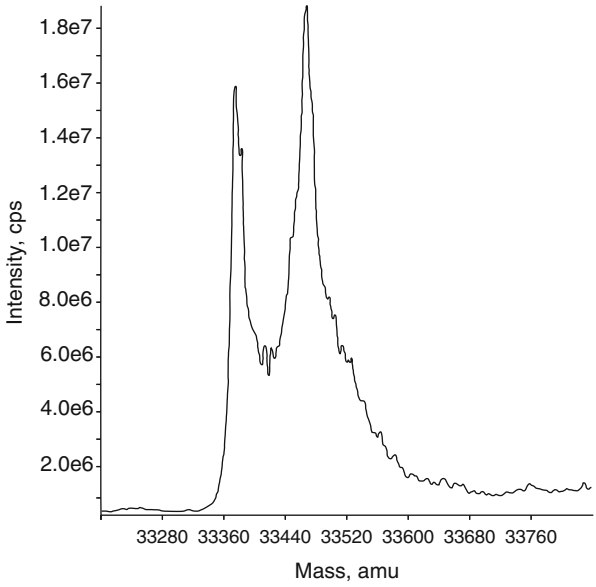


Fig. 14.2 The mass spectrometry of the 4F-Trp-labeled GGR

this question we have started experiments to determine whether these peaks are due to the 5F-Trp analogue or the neighboring proline residues that may show *cis* or *trans* orientation. To accomplish this we are exploring these questions by ^{19}F NMR studies and X-ray crystallography of the fluorinated proteins. We have commenced our investigation by incorporating 4F-Trp, 4F-*cis*-Pro, and 4F-*trans*-Pro into GGR. The 4F-Trp analogue is essentially nonfluorescent at ambient temperature; making it a "silent" analogue for absorption spectroscopy, therefore fluorine incorporation studies using an absorption technique referred to as LINCS analysis was not possible [27]. Figure 14.2 shows the initial mass spectrometry data for the incorporation of 4F-Trp into GGR in a non auxotrophic strain of *E. coli*. The spectrum shows a peak for native GGR at a mass of 33,370 and a broad peak for the incorporation of five fluorine atoms at 33,460. The percent incorporation is over 50 % which is sufficient for ^{19}F NMR studies but not for a number of physical biochemical methods such as differential scanning calorimetry for protein integrity or titrating calorimetry for determining the binding of sugar and calcium ligands to the fluorinated protein. Our efforts continue to increase the incorporation to nearly 100 % using an auxotroph strain of *E. coli*.

The 4F-*cis*-Pro incorporation into the ten sites on the GGR protein was the most complex project we have attempted to date. The mass spectral analysis of our third attempt to incorporate this fluorine label is shown in Fig. 14.3. The spectrum shows a distinct resonance for native GGR at 33,373 mass units. Since it is not fully labeled GGR shows a heterogeneous population of masses which corresponds to fluorine in any of the ten proline sites. Thus the observed masses for the fluorinated GGR at 33,388, 33,406, 33,424, 33,442, 33,460, 33,478, 33,496, 33,514, 33,532 and 33,550

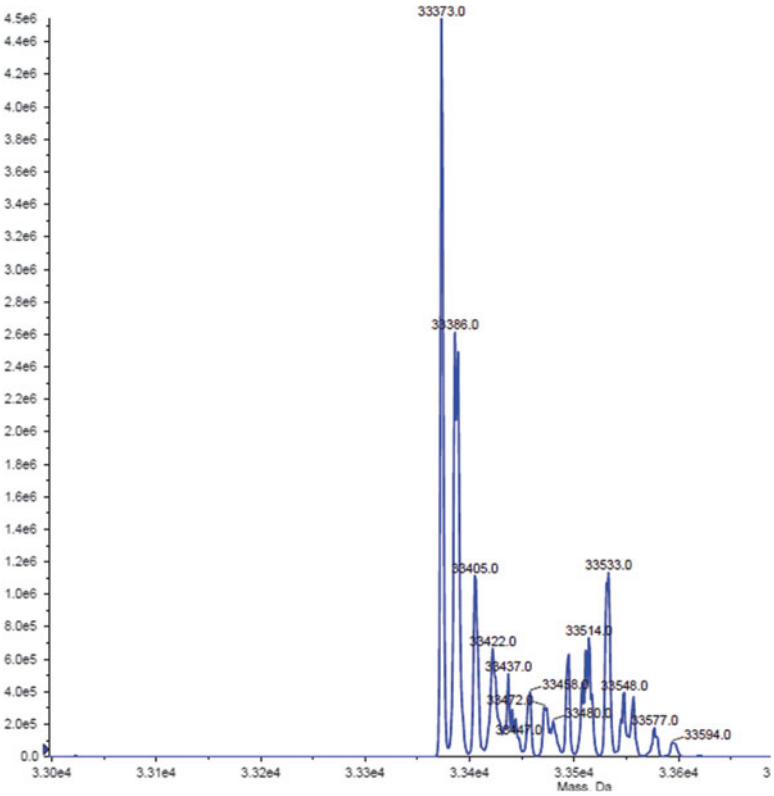


Fig. 14.3 The mass spectrum of 4F-*cis*-Proline-labeled GGR. The largest peak is the native GGR with no label

correspond to 1–10 fluorine atoms incorporated per protein. The peaks for these masses are detected in the spectrum, therefore we can conclude that the 4F-*cis*-Pro is tolerated by the biosynthetic system even though there have been reports where proteins are compromised with this analogue [28]. We will continue our work by transforming our plasmid into the proline auxotroph cell line JM83 with increased amounts of 4F-*cis*-Pro in the media to enhance the incorporation percentage of this fluorine analogue.

14.7 Mass Spectrometry of 5-Fluorotryptophan LBP and LBP Mutants

Initial studies of the 5F-Trp-labeled LBP protein used an internal standard of 4F-Phe as an internal standard for the ^{19}F NMR spectrum and for calculation of the percent incorporation of fluorine into the protein as was done for the GGR [4]. Figure 14.4

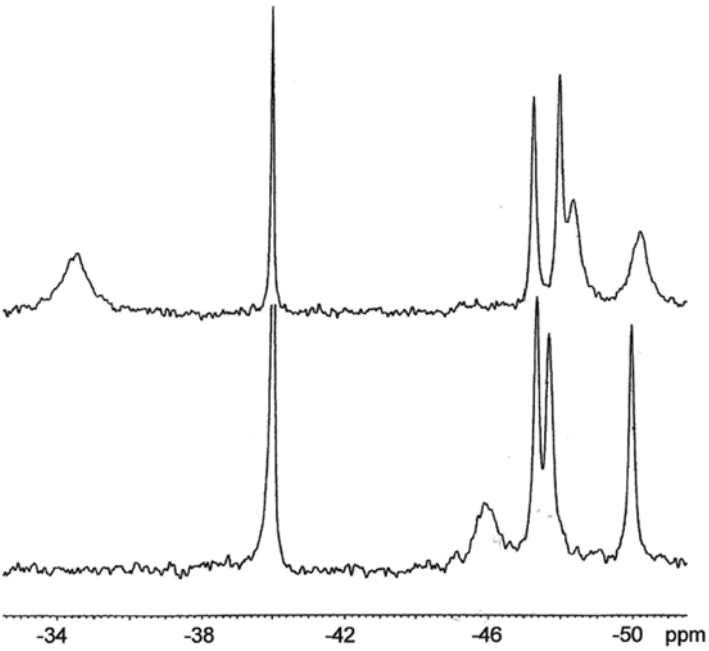


Fig. 14.4 The ^{19}F NMR spectra of 5F-Trp-labeled LBP with 4F-Phe added as an internal standard. Top spectrum is without ligand L-leucine and the bottom spectrum is with excess L-leucine. The internal standard resonance is present in the top spectrum because LBP does not bind to the D form of 4F-Phe. The standard 4F-Phe was a racemic mixture of both D and L forms

shows the ^{19}F NMR spectra of the LBP protein without L-leucine (top) and with excess L-leucine (bottom). The resonance at −39 ppm is the internal standard 4F-Phe. Contrary to the studies with ^{19}F NMR GGR which used the 3F-Phe as an internal standard without compromising the experiments, the internal standard 4F-Phe bounds to the LBP. This consequently led to further studies identifying the new ligands for LBP. The resonance to the far left of the top spectrum in Fig. 14.4 corresponds to the L-4F-Phe in the bound state in the LBP protein. This broad resonance disappears when excess L-leucine is added to the NMR tube. The ^{19}F NMR of the LBP without 4F-Phe in the sample shows four distinct sharp resonances with L-leucine [19]. The spectrum in Fig. 14.4 (bottom trace) shows a protein in slow exchange between two ligands in the NMR tube. This is a good example why an internal standard should not be used directly in a sample and why mass spectrometry is a better choice for calculating the percent incorporation of a fluorine label.

LBP has an inducible promoter and thus with minimal media and enhanced amounts of fluoro amino acids added prior to induction the percentage of fluorinated label is greatly increased. Figure 14.5 shows the mass spectrum of 5F-Trp-labeled LBP with 100 % incorporation. Native LBP has a 36,983 mass and 5F-Trp-labeled protein shows a mass of 37,054.

To identify the resonances in the NMR spectrum we changed the tryptophan residues to phenylalanine or tyrosine. Mass spectrometry can also be used to quickly

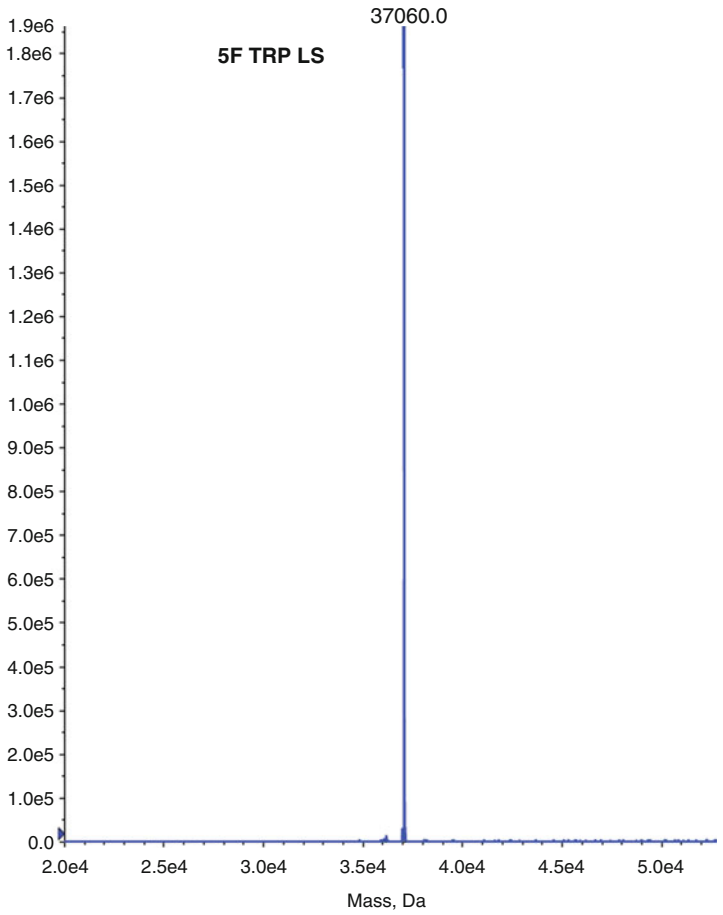


Fig. 14.5 The mass spectrum of 5F-Trp-labeled LBP

Table 14.1 The calculated masses and masses obtained by mass spectrometry of LBP proteins and mutants

Protein	Calculated weight	Mass spec weight
LBP WILD TYPE	36,983	36,986
5F-Trp LBP	37,054	37,060
LBP W336F	36,944	36,943
LBP W18, 278, 320F	36,869	36,868
LBP W18F	36,944	36,948
LBP W18Y	36,960	36,966
LBP W278F	36,944	36,942
LBP D1C	36,972	36,988

identify if the correct mutation is made. Table 14.1 shows the mass spectrometric analysis of a number of LBP mutants that were used in our NMR, biosensor, and fluorescent studies.

14.8 Conclusions

The studies shown in the chapter demonstrates the applicability of mass spectrometry for studying the percent incorporation of fluorine labels in proteins prepared for ^{19}F NMR studies. For 4F-Trp incorporation this method is crucial since absorption methods are not amenable. We clearly make a case that this is the method of choice for most fluorine analogues. In addition the results from the mass spectrometry strongly support the site directed mutagenesis data and confirm the mutant protein production for analysis of the resonances in the ^{19}F NMR spectra. The effectiveness of the ^{19}F NMR data is greatly facilitated by careful quantification and protein characterization by liquid chromatography electrospray mass spectrometry.

Acknowledgments The author would like to thank Bruce O'Roarke for analyzing the proteins by mass spectrometry. This work has been supported by the NSF for the LCMS instrumentation in the Chemistry Department at the University of Vermont (CHE MRI-0821501, LAL Co-PI), SUNY at Plattsburgh President's Award and Mini Grant.

References

1. Danielson MA, Falke JJ (1996) Annu Rev Biophys Biomol Struct 25:163–195
2. Yu JX, Hallac RR, Chiguru S, Mason RP (2013) Prog Nucl Magn Reson Spectrosc 70:25–49
3. Reid DG, Murphy PS (2008) Drug Discov Today 13:473–480
4. Luck LA, Falke JJ (1991) Biochemistry 30:6484–6490
5. Zemsky J, Rusinova E, Nemerson Y, Luck LA, Ross JB (1999) Proteins 37:709–716
6. Crabb JW, Carlson A, Chen Y, Goldflam S, Intres R et al (1998) Protein Sci 7:746–757
7. Crowley PB, Kyne C, Monteith WB (2012) Chem Commun (Camb) 48:10681–10683
8. Luck LA, Falke JJ (1991) Biochemistry 30:4248–4256
9. Yu J, Cui D, Zhao D, Mason RP (2008). In: GH A. Tressaud (ed) Fluorine and health. Elsevier B.V., Amsterdam, pp 198–276
10. Senear DF, Mendelson RA, Stone DB, Luck LA, Rusinova E, Ross JB (2002) Anal Biochem 300:77–86
11. Buer BC, Marsh EN (2012) Protein Sci 21:453–462
12. Salwiczek M, Nyakatura EK, Gerling UI, Ye S, Koksch B (2012) Chem Soc Rev 41:2135–2171
13. Luck LA, Moravan MJ, Garland JE, Salopek-Sondi B, Roy D (2003) Biosens Bioelectron 19:249–259
14. Carmon KS, Baltus RE, Luck LA (2004) Biochemistry 43:14249–14256
15. Andreescu S, Luck LA (2008) Anal Biochem 375:282–290
16. Tripathi A, Wang J, Luck LA, Suni II (2007) Anal Chem 79:1266–1270
17. Sokolov I, Subba-Rao V, Luck LA (2006) Biophys J 90:1055–1063
18. Luck LA, Vance JE, O'Connell TM, London RE (1996) J Biomol NMR 7:261–272
19. Salopek-Sondi B, Luck LA (2002) Protein Eng 15:855–859
20. Salopek-Sondi B, Skeels MC, Swartz D, Luck LA (2003) Proteins 53:273–281
21. Salopek-Sondi B, Swartz D, Adams PS, Luck LA (2002) J Biomol Struct Dyn 20:381–387
22. Ahmed AH, Loh AP, Jane DE, Oswald RE (2007) J Biol Chem 282:12773–12784
23. Magnusson U, Salopek-Sondi B, Luck LA, Mowbray SL (2004) J Biol Chem 279:8747–8752

24. Luck LA, Johnson C (2000) Protein Sci 9:2573–2576
25. Vyas MN, Jacobson BL, Quiocho FA (1989) J Biol Chem 264:20817–20821
26. Luck LA, Falke JJ (1991) Biochemistry 30:4257–4261
27. Waxman E, Rusinova E, Hasselbacher CA, Schwartz GP, Laws WR, Ross JB (1993) Anal Biochem 210:425–428
28. Holzberger B, Obeid S, Welte W, Diederichs K, Marx A (2012) Chem Sci 3:2924–2931

Chapter 15
Mass Spectrometry for Proteomics-Based Investigation Using the Zebrafish Vertebrate Model System

Reshica Baral*, Armand G. Ngounou Wetie*, Costel C. Darie, and Kenneth N. Wallace

Abstract The zebrafish (*Danio rerio*) is frequently being used to investigate the genetics of human diseases as well as resulting pathologies. Ease of both forward and reverse genetic manipulation along with conservation of vertebrate organ systems and disease causing genes has made this system a popular model. Many techniques have been developed to manipulate the genome of zebrafish producing mutants in a vast array of genes. While genetic manipulation of zebrafish has progressed, proteomics have been under-utilized. This review highlights studies that have already been performed using proteomic techniques and as well as our initial proteomic work comparing changes to the proteome between the *ascl1a*−/− and WT intestine.

Abbreviations

CID	Collision-induced dissociation
ESI-MS	Electrospray ionization mass spectrometry
LC-MS/MS	Liquid chromatography mass spectrometry
m/*z*	Mass/charge
MALDI-MS	MALDI mass spectrometry
MS	Mass spectrometry

*Author contributed equally with all other contributors.

R. Baral • K.N. Wallace (✉)
Department of Biology, Clarkson University, 8 Clarkson Avenue, Potsdam, NY 13699-5810, USA
e-mail: kwallace@clarkson.edu

A.G. Ngounou Wetie • C.C. Darie
Biochemistry and Proteomics Group, Department of Chemistry and Biomolecular Science, Clarkson University, 8 Clarkson Avenue, Potsdam, NY 13699-5810, USA

A.G. Woods and C.C. Darie (eds.), *Advancements of Mass Spectrometry in Biomedical Research*, Advances in Experimental Medicine and Biology 806,
DOI 10.1007/978-3-319-06068-2_15, © Springer International Publishing Switzerland 2014

Mw	Molecular weight
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
TIC	Total ion current/chromatogram

15.1 Introduction

Initially, the zebrafish model system was used extensively for investigations in early vertebrate development due to the ease of genetically manipulation [1]. In recent years, investigations utilizing zebrafish have expanded into organ development [2–6], disease modeling [7, 8], and toxicology [9]. Similarity of zebrafish in genetic identity to humans combined with homologous organ systems has demonstrated the usefulness of zebrafish as a model system for using the model system for investigation of numerous human diseases and disorders [10, 11]. Other features such as small size, production of high number of externally developing embryos, optical clarity of embryos, and fast embryogenesis make zebrafish attractive for studying these diseases.

During this time, the number of genetic tools has increased dramatically which is complemented by the sequencing and annotation of the zebrafish genome [12]. Currently, there are techniques involving both forward and reverse genetics that have increased the usefulness of the zebrafish system. One of the now classic approaches is random generation of point mutations utilizing *N*-ethyl-*N*-nitrosourea (ENU). In the mid 1990s, two large-scale ENU screens were performed which highlighted the ability of this vertebrate to be screened for novel mutations, which replicate defects in human diseases and disorders [13]. Since then there have been many additional forward genetic ENU screens including novel screens such as looking for genes that have a maternal effect on development of the embryo [14, 15]. These forward genetic screens are now being complemented by the Zebrafish Mutation Project that aims to make mutations in each gene within the zebrafish genome using TILLING (targeting induced local lesions in genomes) methods [16, 17], which have currently generated mutants in 46 % of the zebrafish protein coding genes [10]. In addition to TILLING mutations in the zebrafish genome are also created by transcription activator-like effector nucleases (TALENs) [18] and the CRISPR-Cas system [19], which can be targeted to individual genes to create double strand breaks in the DNA. Double strand breaks are then repaired inaccurately by the non-homologous end joining repair (NHEJ) system to create a variety of mutations at the break site.

Compared to development of genetic tools, proteomic analysis has not been an extensively utilized technique in the zebrafish system. Proteomic analysis can identify changes in structural and enzymatic content of organs modified by disease states, genetic changes, and differences over a developmental time course [20–29]. Thus proteomics can complement current methods of genetic modification to the zebrafish system and add another tool kit to the analysis. Already, a variety of methods have been applied to analyze the changes in the zebrafish proteome. Here we will review some of the ways that mass spectrometry (MS) is being used to investigate zebrafish proteomics and then provide an example of preliminary MS analysis that we have begun to compare differences in protein expression between wild type and mutant zebrafish embryos.

15.2 Experimental Design

ascl1a−/− embryos and WT siblings were grown to 5dpf. Intestines from mutants and WT were dissected into T-PERTissue Protein Extraction Reagent (Pierce) with Halt Protease Inhibitor Cocktail and EDTA (Pierce) with double the recommended concentration in order to prevent protein degradation. Whole protein extract was run on an 8 % SDS-PAGE gel and stained with Coomassie. In-gel trypsin digestion was performed on excised bands according to published procedures [30–32]. Samples were prepared for MS analysis and run on a nanoliquid chromatography tandem MS (nanoLC-MS/MS) using a NanoAcquity UPLC coupled to a QTOF Micro mass spectrometer (both from Waters Corp.). Raw data was processed using a ProteinLynx Global Server (PLGS version 2.5 from Waters Corp.) and the resulting pkl files were submitted to Mascot database search and proteins were identified using the UniProt database. This procedure is extensively described elsewhere [30–32].

15.3 Modeling Disease

A number of proteomic studies have investigated changes between the healthy conditions compared to the diseased state. In a study by Lu et al., proteomic alterations of zebrafish gills to *Aeromonas hydrophila* bacterial infection are observed [33]. Here as would be expected, proteins regulating bacterial infection and host immunity were unregulated. Also, proteins modulating increases in energy metabolism were up regulated along with cytoskeletal proteins. The latter may be needed due to increased phagosomes to begin clearing the infection.

In another study, Chen et al. determined the proteomic response to hypoxia [34]. In this study, adult zebrafish were exposed to hypoxic conditions for 48 h followed by extraction and preparation of only white muscle. The hypoxic conditions resulted in an increase in glycolytic enzymes in skeletal muscle with down-regulation of aerobic ATP production. This study also found increases in glycolytic muscle fibers and hemoglobin alpha variants suggesting a relatively rapid shift to low oxygen conditions.

Hogl et al. used proteomic analysis of a zebrafish knockout BACE1 to identify additional membrane targets of the enzyme [35]. BACE1 cleaves the amyloid precursor protein and is therefore a drug target for Alzheimer's disease. In the BACE1 knockout proteomic analysis, the majority of proteins were unaltered but a group of 24 intrinsic membrane proteins accumulated suggesting that they were no longer cleaved. Further analysis of identified substrates can determine whether reducing the activity of BACE1 into combat Alzheimer's disease will result in additional complications.

Fleming et al. utilized proteomics to help determine the timing of when the blood–brain barrier develops [36]. Exclusion of the drug from the head but not the body suggests development of a functional barrier between the blood and brain. Lack of drug accumulation in the head coupled with functional studies suggests that the blood–brain barrier matures between 8 and 10 days post fertilization (dpf).

While this investigation primarily identifies the timing of a developmental event, the same assay might be used to characterize breakdown of the blood–brain barrier during a disease process.

15.4 Comparative Proteomics

Proteomics can also take advantage of the changes between zebrafish developmental time points and even between different species to identify proteins that may be critical for initiating these changes. For example, sex determination within zebrafish is not strictly due to chromosomal inheritance. Groh et al. has used proteomic analysis to compare differences in protein composition between the testis and the ovary. Initially a global shotgun approach to generate an overall list of differential proteins was produced [37]. Once differentially expressed proteins were identified, these investigators utilized the selected reaction monitoring (SRM) technique coupled with representational difference analysis (RDA) to more reliably detect low abundance proteins that are likely to be important in determination of either the testis or the ovary [38].

Lin et al. have used proteomics to identify evolutionary changes in the crystallins between sighted fish such as zebrafish and compare loss of diversity in these proteins to loss of diversity in nocturnal rice eel and walking catfish [39]. Both rice eel and walking catfish have degenerative eyes. Here the number of overall proteins is limited by only including the lens in this shotgun proteomic approach. While there are a number of lens proteins, a number of other high abundance proteins are excluded by selective tissue acquisition. This approach identifies the loss of alpha-crystallin in the rice eel, which may function as a chaperone protein preventing aggregation of the other crytallins.

15.5 Changes to Protein Function

As in other systems antibodies to specific proteins can be immunoprecipitated and MS can be performed on the resulting pull-down proteins [40]. With this method, interacting proteins can be identified as well as modifications to the target and associated proteins. Within the zebrafish system there has been a lack of good and reliable antibodies to specific proteins. Many of the antibodies used in the zebrafish system are specific to proteins from other systems and cross-reacting with their homologous zebrafish counterparts.

The lack of zebrafish-specific antibodies has severely limited the identification of interacting proteins by immunoprecipitation. Deflorian et al. however, has generated a host of antibodies that were then screened for unique non-overlapping patterns [41]. From the numerous antibodies generated, 9 monoclonal antibodies were selected due to interesting expression patterns while one was used for immunoprecipitation. The protein was then identified by mass spectroscopy and found to be vitellogenin 1. Even though this screening method is labor intensive, it would produce a number of good antibodies to identify interacting proteins by immunoprecipitation.

Even when an appropriate zebrafish antibody is not present, the protein can be expressed with a FLAG tag in order to immunoprecipitate the protein [42]. Alternatively, Taskinen et al. expressed and purified a new avidin protein in *E. coli* to identify their native state. Here, MS demonstrated that the protein strongly binds biotin but the oligomeric state is unstable at low ionic strength [43].

15.6 Comparison of *ascl1a*−/− to WT Embryos

We have previously shown that *ascl1a*−/− embryos fail to develop intestinal epithelial secretory cells and only differentiate about 10 % of the normal number of enteric neurons by the end of embryogenesis [44]. While the Notch signaling pathway appears to play a role in formation of the intestinal secretory cells, the reason for failure of enteric neural differentiation is less clear. Mis-expressed proteins within the *ascl1a*−/− mutants may provide information about what cellular processes are affected. We have therefore begun to use a shotgun proteomic approach to identify mis-expressed proteins.

In order to limit the tissue involved in the analysis, we have dissected and pooled intestines from both *ascl1a*−/− and WT embryos at 5 dpf (workflow summary Fig. 15.1). Analysis of these extracts by SDS-PAGE demonstrates striking

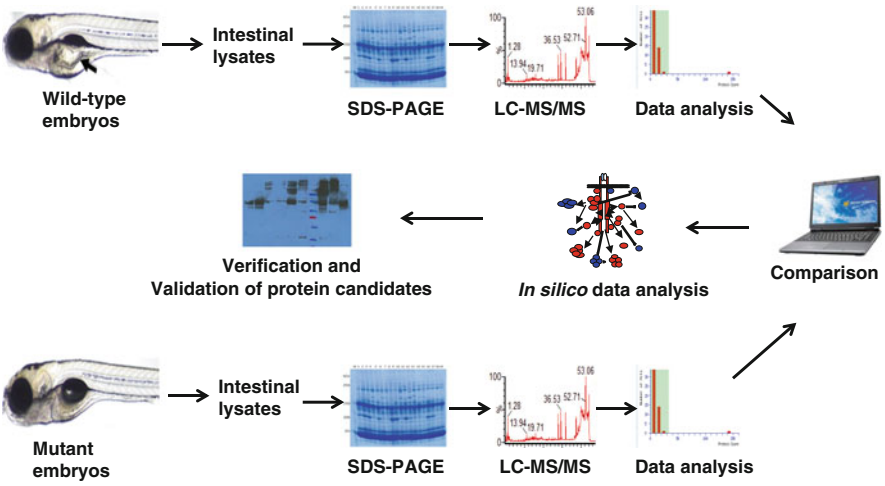


Fig. 15.1 Workflow for comparison of WT and mutant embryos. Intestinal lysates for both WT and mutant embryos 5 dpf embryos are prepared and run on SDS-PAGE gels followed by nanoliquid chromatography tandem MS (nanoLC-MS/MS) using a NanoAcquity UPLC coupled to a QTOF Micro mass spectrometer (both from Waters Corp.). Raw data is processed using a ProteinLynx Global Server (PLGS version 2.5 from Waters Corp.) and the resulting pkl files are submitted to Mascot database search and proteins are identified using the UniProt database. Proteins mis-regulated in mutant embryos are then verified by whole mount immunohistochemistry, western, or quantitative PCR

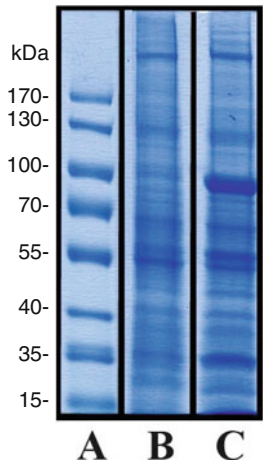


Fig. 15.2 Comparison of WT and *ascl1a*–/– 5 dpf intestine protein extract. (*A*) Molecular mass standards (in kDa). (*B*) WT intestinal protein extract. (*C*) *ascl1a*–/– protein extract

differences between WT and *ascl1a*–/– embryos (Fig. 15.2). In-gel tryptic digestion and nanoLC-MS/MS analysis identified a series of proteins that were specific to (or more abundant in) WT or *ascl1a*–/– embryos. An example of such a differentially expressed protein is Flotilin-2a (Fig. 15.3), a protein found in WT but not in the *ascl1a*–/– embryos.

Further comparison of alteration, increase or decrease of the levels of proteins and their post-translational modifications such as phosphorylation of acetylation between *ascl1a*–/– and WT embryos will likely suggest mechanisms for the lack of differentiated enteric neurons.

15.7 Conclusions and Perspectives

While proteomic analysis in the zebrafish system has still seen limited applications, there have been a number of promising investigations. As proteomics becomes more frequently utilized in the zebrafish system, it is likely to confirm predictions based on mRNA expression. These approaches may also reveal surprising differences in protein expression that would not be predicted at the mRNA expression level. Proteomics can give more of a global picture of structural and enzymatic changes due to alterations in either the genome or the disease state that is not possible by mRNA expression alone.

While proteomic techniques will not replace genetic approaches, these techniques promise to be a valuable complement. After shotgun proteomic analysis are completed, refined predictions for what proteins are important can be made and these and other less abundant proteins can be targeted in future experiments. One approach might be to refine the input cells using a genetic approach. For example, the relative ease with which transgenic fish can be produced has increased dramatically in the past few years. This coupled with increased annotation of the genome increases the ability to create fish with specific cells labeled by fluorescent proteins.

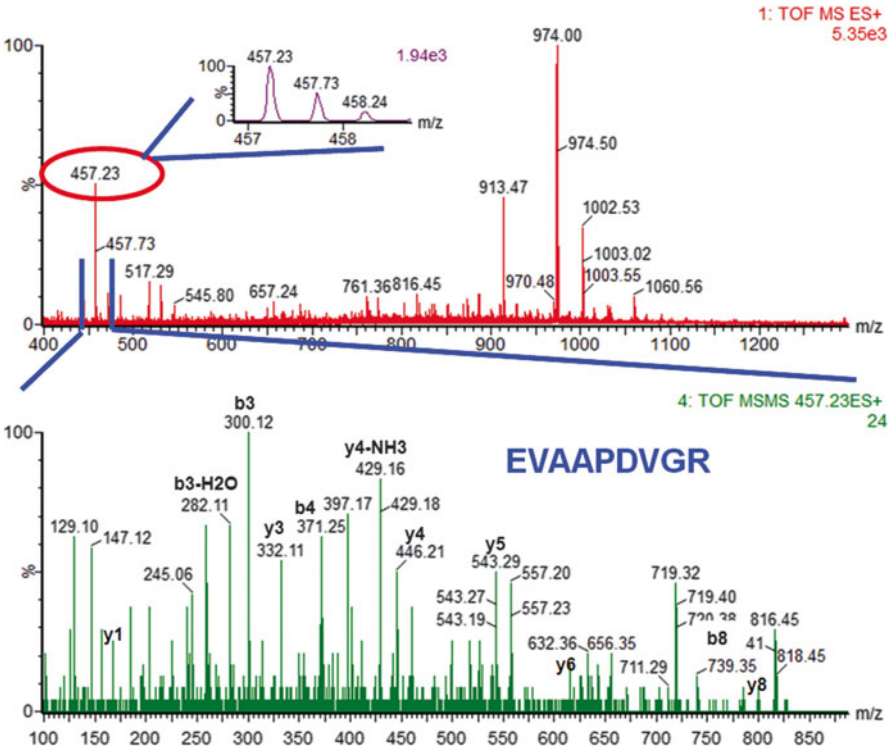


Fig. 15.3 Identification of a protein unique to *ascl1a*−/− mutants. NanoLC-MS/MS analysis led to identification of a doubly charged precursor peak with mass to charge ratio (*m*/*z*) of 457.23 (MS mode). Selection of this precursor peak for fragmentation by MS/MS (MS/MS mode) produced a series of fragment b and y ions, whose analysis led to identification of a peptide with the amino acid sequence EVAAPDVGR, which was part of the protein Flotillin-2a. This protein was found in the protein extract from WT but absent in *ascl1a*−/− 5 dpf intestine

Specific cell types from fluorescently labeled lines can then be isolated by Fluorescent Activated Cell sorting (FACs). Homogenous pools of cells can then be obtained for use in proteomic analysis. This will lessen the content of extraneous protein to add in detection of more low abundance proteins. Lessening unwanted proteins and other MS techniques to increase detection of low abundance proteins will continue to make proteomics a valuable tool for the zebrafish system.

Acknowledgements We would like to thank Ms. Laura Mulderig, Scott Nichols, and their colleagues (Waters Corporation) for their generous support in setting up the Proteomics Center at Clarkson University. C.C.D thanks Drs. Thomas A. Neubert (New York University), Belinda Willard (Cleveland Clinic), and Gregory Wolber and David Mclaughin (Eastman Kodak Company) for donation of a TofSpec2E MALDI-MS (each). This work was supported in part by the Keep a Breast Foundation (KEABF-375-35054), the Redcay Foundation (SUNY Plattsburgh), the Alexander von Humboldt Foundation, SciFund Challenge, private donations (Ms. Mary Stewart Joyce and Mr. Kenneth Sandler), David A. Walsh fellowship, and by the U.S. Army Research Office (DURIP grant #W911NF-11-1-0304).

References

1. Nusslein-Volhard C (1994) Of flies and fishes. Science 266:572–574
2. Schmidt R, Strahle U, Scholpp S (2013) Neurogenesis in zebrafish—from embryo to adult. Neural Dev 8:3
3. Wallace KN, Akhter S, Smith EM, Lorent K, Pack M (2005) Intestinal growth and differentiation in zebrafish. Mech Dev 122:157–173
4. Wallace KN, Dolan AC, Seiler C, Smith EM, Yusuff S, Chaille-Arnold L, Judson B, Sierk R, Yengo C, Sweeney HL, Pack M (2005) Mutation of smooth muscle myosin causes epithelial invasion and cystic expansion of the zebrafish intestine. Dev Cell 8:717–726
5. Wallace KN, Pack M (2003) Unique and conserved aspects of gut development in zebrafish. Dev Biol 255:12–29
6. Olden T, Akhtar T, Beckman SA, Wallace KN (2008) Differentiation of the zebrafish enteric nervous system and intestinal smooth muscle. Genesis 46:484–498
7. Bakkers J (2011) Zebrafish as a model to study cardiac development and human cardiac disease. Cardiovasc Res 91:279–288
8. Santoriello C, Zon LI (2012) Hooked! Modeling human disease in zebrafish. J Clin Invest 122:2337–2343
9. Silvana A, Mihaela G, Özel RE, Kenneth NW (2011) Methodologies for toxicity monitoring and nanotechnology risk assessment. In: Biotechnology and nanotechnology risk assessment: minding and managing the potential threats around us. American Chemical Society, Washington, pp 141–180
10. Seth A, Stemple DL, Barroso I (2013) The emerging use of zebrafish to model metabolic disease. Dis Model Mech 6:1080–1088
11. Sadler KC, Rawls JF, Farber SA (2013) Getting the inside tract: new frontiers in zebrafish digestive system biology. Zebrafish 10:129–131
12. Harrow JL, Steward CA, Frankish A, Gilbert JG, Gonzalez JM, Loveland JE, Mudge J, Sheppard D, Thomas M, Trevanion S, Wilming LG (2014) The Vertebrate Genome Annotation browser 10 years on. Nucleic Acids Res 42:D771–D779
13. Driever W, Solnica-Krezel L, Schier AF, Neuhauss SC, Malicki J, Stemple DL, Stainier DY, Zwartkruis F, Abdelilah S, Rangini Z, Belak J, Boggs C (1996) A genetic screen for mutations affecting embryogenesis in zebrafish. Development 123:37–46
14. Dosch R, Wagner DS, Mintzer KA, Runke G, Wiemelt AP, Mullins MC (2004) Maternal control of vertebrate development before the midblastula transition: mutants from the zebrafish I. Dev Cell 6:771–780
15. Wagner DS, Dosch R, Mintzer KA, Wiemelt AP, Mullins MC (2004) Maternal control of development at the midblastula transition and beyond: mutants from the zebrafish II. Dev Cell 6:781–790
16. Kettleborough RN, Bruijn E, Eeden F, Cuppen E, Stemple DL (2011) High-throughput target-selected gene inactivation in zebrafish. Methods Cell Biol 104:121–127
17. Kettleborough RN, Busch-Nentwich EM, Harvey SA, Dooley CM, de Bruijn E, van Eeden F, Sealy I, White RJ, Herd C, Nijman IJ, Fenyes F, Mehroke S, Scahill C, Gibbons R, Wali N, Carruthers S, Hall A, Yen J, Cuppen E, Stemple DL (2013) A systematic genome-wide analysis of zebrafish protein-coding gene function. Nature 496:494–497
18. Sander JD, Cade L, Khayter C, Reyon D, Peterson RT, Joung JK, Yeh JR (2011) Targeted gene disruption in somatic zebrafish cells using engineered TALENs. Nat Biotechnol 29: 697–698
19. Hwang WY, Fu Y, Reyon D, Maeder ML, Tsai SQ, Sander JD, Peterson RT, Yeh JR, Joung JK (2013) Efficient genome editing in zebrafish using a CRISPR-Cas system. Nat Biotechnol 31:227–229
20. Darie CC, Shetty V, Spellman DS, Zhang G, Xu C, Cardasis HL, Blais S, Fenyo D, Neubert TA (2008) Blue Native PAGE and mass spectrometry analysis of the ephrin stimulation-dependent protein-protein interactions in NG108-EphB2 cells. Springer, Düsseldorf

21. Darie CC, Litscher ES, Wassarman PM (2008) Structure, processing, and polymerization of rainbow trout egg vitelline envelope proteins. Springer, Düsseldorf
22. Darie CC (2013) Mass spectrometry and its application in life sciences. Aust J Chem 66:1–2
23. Ngounou Wetie AG, Sokolowska I, Woods AG, Darie CC (2013) Identification of post-translational modifications by mass spectrometry. Aust J Chem 66:734–748
24. Ngounou Wetie AG, Sokolowska I, Woods AG, Roy U, Deinhardt K, Darie CC (2014) Protein-protein interactions: switch from classical methods to proteomics and bioinformatics-based approaches. Cell Mol Life Sci 71(2):205–228
25. Ngounou Wetie AG, Sokolowska I, Woods AG, Roy U, Loo JA, Darie CC (2013) Investigation of stable and transient protein-protein interactions: Past, present, and future. Proteomics 13(3–4):538–557
26. Ngounou Wetie AG, Sokolowska I, Woods AG, Wormwood KL, Dao S, Patel S, Clarkson BD, Darie CC (2013) Automated mass spectrometry-based functional assay for the routine analysis of the secretome. J Lab Autom 18:19–29
27. Sokolowska I, Ngounou Wetie AG, Woods AG, Darie CC (2012) Automatic determination of disulfide bridges in proteins. J Lab Autom 17:408–416
28. Sokolowska I, Ngounou Wetie AG, Woods AG, Darie CC (2013) Applications of mass spectrometry in proteomics. Aust J Chem 66:721–733
29. Woods AG, Sokolowska I, Yakubu R, Butkiewicz M, LaFleur M, Talbot C, Darie CC (2011) Blue native page and mass spectrometry as an approach for the investigation of stable and transient protein-protein interactions. In: Andreescu S, Hepel M (eds) Oxidative stress: diagnostics, prevention, and therapy. American Chemical Society, Washington
30. Petrareanu C, Macovei A, Sokolowska I, Woods AG, Lazar C, Radu GL, Darie CC, Branza-Nichita N (2013) Comparative proteomics reveals novel components at the plasma membrane of differentiated HepaRG cells and different distribution in hepatocyte-and biliary-like cells. PLoS One 8:e71859
31. Sokolowska I, Gawinowicz MA, Ngounou Wetie AG, Darie CC (2012) Disulfide proteomics for identification of extracellular or secreted proteins. Electrophoresis 33:2527–2536
32. Sokolowska I, Ngounou Wetie AG, Roy U, Woods AG, Darie CC (2013) Mass spectrometry investigation of glycosylation on the NXS/T sites in recombinant glycoproteins. Biochim Biophys Acta 1834:1474–1483
33. Lu A, Hu X, Wang Y, Shen X, Li X, Zhu A, Tian J, Ming Q, Feng Z (2014) iTRAQ analysis of gill proteins from the zebrafish (Danio rerio) infected with Aeromonas hydrophila. Fish Shellfish Immunol 36:229–239
34. Chen K, Cole RB, Rees BB (2013) Hypoxia-induced changes in the zebrafish (Danio rerio) skeletal muscle proteome. J Proteomics 78:477–485
35. Hogl S, van Bebber F, Dislich B, Kuhn PH, Haass C, Schmid B, Lichtenthaler SF (2013) Label-free quantitative analysis of the membrane proteome of Bace1 protease knock-out zebrafish brains. Proteomics 13:1519–1527
36. Fleming A, Diekmann H, Goldsmith P (2013) Functional characterisation of the maturation of the blood-brain barrier in larval zebrafish. PLoS One 8:e77548
37. Groh KJ, Nesatyy VJ, Segner H, Eggen RI, Suter MJ (2011) Global proteomics analysis of testis and ovary in adult zebrafish (Danio rerio). Fish Physiol Biochem 37:619–647
38. Groh KJ, Schonenberger R, Eggen RI, Segner H, Suter MJ (2013) Analysis of protein expression in zebrafish during gonad differentiation by targeted proteomics. Gen Comp Endocrinol 193:210–220
39. Lin YR, Mok HK, Wu YH, Liang SS, Hsiao CC, Huang CH, Chiou SH (2013) Comparative proteomics analysis of degenerative eye lenses of nocturnal rice eel and catfish as compared to diurnal zebrafish. Mol Vis 19:623–637
40. ten Have S, Boulon S, Ahmad Y, Lamond AI (2011) Mass spectrometry-based immunoprecipitation proteomics—the user's guide. Proteomics 11:1153–1159
41. Deflorian G, Cinquanta M, Beretta C, Venuto A, Santoriello C, Baldessari D, Pezzimenti F, Aliprandi M, Mione M, de Marco A (2009) Monoclonal antibodies isolated by large-scale screening are suitable for labeling adult zebrafish (Danio rerio) tissues and cell structures. J Immunol Methods 346:9–17

42. Jurynec MJ, Xia R, Mackrill JJ, Gunther D, Crawford T, Flanigan KM, Abramson JJ, Howard MT, Grunwald DJ (2008) Selenoprotein N is required for ryanodine receptor calcium release channel activity in human and zebrafish muscle. Proc Natl Acad Sci U S A 105: 12485–12490
43. Taskinen B, Zmurko J, Ojanen M, Kukkurainen S, Parthiban M, Maatta JA, Leppiniemi J, Janis J, Parikka M, Turpeinen H, Ramet M, Pesu M, Johnson MS, Kulomaa MS, Airenne TT, Hytonen VP (2013) Zebavidin—an avidin-like protein from zebrafish. PLoS One 8:e77207
44. Roach G, Heath Wallace R, Cameron A, Emrah Ozel R, Hongay CF, Baral R, Andreescu S, Wallace KN (2013) Loss of *ascl1a* prevents secretory cell differentiation within the zebrafish intestinal epithelium resulting in a loss of distal intestinal motility. Dev Biol 376:171–186

Chapter 16
Mass Spectrometry-Based Biomarkers in Drug Development

Ronald A. Miller and Daniel S. Spellman

Abstract Advances in mass spectrometry, proteomics, protein bioanalytical approaches, and biochemistry have led to a rapid evolution and expansion in the area of mass spectrometry-based biomarker discovery and development. The last decade has also seen significant progress in establishing accepted definitions, guidelines, and criteria for the analytical validation, acceptance, and qualification of biomarkers. These advances have coincided with a decreased return on investment for pharmaceutical research and development and an increasing need for better early decision making tools. Empowering development teams with tools to measure a therapeutic interventions impact on disease state and progression, measure target engagement, and to confirm predicted pharmacodynamic effects is critical to efficient data-driven decision making. Appropriate implementation of a biomarker or a combination of biomarkers can enhance understanding of a drugs mechanism, facilitate effective translation from the preclinical to clinical space, enable early proof of concept and dose selection, and increase the efficiency of drug development. Here we will provide descriptions of the different classes of biomarkers that have utility in the drug development process as well as review specific, protein-centric, mass spectrometry-based approaches for the discovery of biomarkers and development of targeted assays to measure these markers in a selective and analytically precise manner.

16.1 Biomarkers

A "biomarker" has been defined by The Biomarkers Definitions Working Group as "a characteristic that is objectively measured and evaluated as an indicator of normal biologic processes, pathogenic processes, or pharmacologic responses to a

R.A. Miller • D.S. Spellman (✉)
Molecular Biomarkers Department, Merck Research Laboratories,
770 Sumneytown Pike, West Point, PA 19486, USA
e-mail: daniel_spellman@merck.com

A.G. Woods and C.C. Darie (eds.), *Advancements of Mass Spectrometry in Biomedical Research*, Advances in Experimental Medicine and Biology 806,
DOI 10.1007/978-3-319-06068-2_16, © Springer International Publishing Switzerland 2014

therapeutic intervention" [7, 30]. Biomarkers can be a single measured entity such as an mRNA, a protein, lipid, or monoamine, a specific cell type, an enzymatic activity, or something more complex such as a panel of proteins [55] or mRNA [12] that define a biologically significant signature [43, 64], organ functions, or characteristic changes in biological structures. Although Biomarkers as an independent field of research is relatively new, they have been used in clinical practice for hundreds if not thousands of years and in basic research for decades [32]. There are many examples of common biomarkers used everyday in medical practice; body temperature as a marker for fever and likely viral or bacterial infection, complete blood count (CBC) to monitor for the presence of many forms of disease, vital markers such as blood pressure and heart rate as an indicator of cardiovascular and renal function, cholesterol values as a risk indicator for coronary and vascular disease, and urine sugar and blood glucose levels for diabetes [20]. Some more modern examples that may align better with our current concepts of a biomarker are; C-reactive protein as a marker for inflammation [13], prostate specific antigen (PSA) as a marker of prostate cancer [16], or BRCA1 gene mutations as risk factors for breast and ovarian cancer [10].

Within the biomedical community what are typically described as "biomarkers" are most commonly molecular biomarkers. These markers are often discovered using genomic technologies such as mRNA profiling [8], genome-wide association studies (GWAS) [68], and whole genome sequencing [47], or proteomic technologies such as mass spectrometry-based discovery approaches [1], multiplexed immunoassays [38], aptamer arrays [23], or protein arrays [11]. In addition to genomic and proteomic platforms, metabolomics, lipidomics, glycomics, and imaging biomarkers such as volumetric magnetic resonance imaging (MRI), and positron emission tomography (PET) tracers are commonly used approaches in identification of biomarkers [48]. All of these approaches may prove useful in delivering a tool to measure a therapeutic interventions impact on disease state and progression, measure target engagement, measure a pharmacodynamic effect, or drug safety.

16.2 Biomarkers in Drug Development

There is a considerable investment involved in drug development often exceeding 800 million dollars spent to bring a drug to market [17, 54]. Overall, the pharmaceutical industry suffers regular setbacks (financially as well as efficient use of time and other resources) due to the inherent complications involved in navigating the unknown parameters involved in human physiology. In light of these challenges, significant focus is placed on generating data more efficiently, in a cost-effective manner, for earlier and more informed decision making. The "right" biomarkers can be used throughout the developmental process, by illuminating the mechanism of action, target engagement (preclinical and clinical), pharmacodynamic effect, efficacy, and safety. In the best of circumstances (when a biomarker proves useful and a program advances to the clinic) these markers translate into the clinical setting

providing quicker diagnosis, proof of concept, and potentially higher successful treatment rates with corresponding less expense and more timely development [20, 37, 69]. Key biomarker types in drug development include target engagement, pharmacodynamic, disease-related, and safety biomarkers. There are many different uses for these biomarkers, from exploratory hypothesis generation to definitive go/no-go decision making in late-phase drug development.

Disease-related biomarkers are often used to diagnose the presence or risk of disease in and individual (Diagnostic), determine the development of a disease over time (Progression), or determine the likely course an individual will take in disease progression (Prognostic). Disease-related markers can also have utility in further segregating a population and evaluating the appropriate therapeutic treatment and probable therapeutic response (Segmentation). The accuracy and utility of disease-related biomarkers can vary from individual to individual throughout the population. Therefore it is important to understand how that accuracy may vary in order to truly understand a biomarker's utility. A robust disease-specific biomarker could hold the key to establishing a preventative therapy and driving proof-of-concept results [20, 69].

Target engagement biomarkers measure drug-target interactions in preclinical and clinical settings [24, 35, 40, 72]. These biomarkers are critical in the drug development process and enable the establishment of correlations between target engagement and measurements of drug efficacy, pharmacodynamics, pharmacokinetics, and safety. Measuring target occupancy in vivo is a critical step in validation of a drug target. Without measurements of target engagement, it can be very difficult to determine the basis for a drug's efficacy or lack thereof [69]. If a drug fails to produce an expected therapeutic effect with full target occupancy then the proposed target and mechanism should be invalidated for the intended clinical indication. Conversely, if a drug produces an effect without occupancy of the intended target, the efficacy of the therapeutic is likely produced though an "off-target" effect. Understanding target engagement can help to determine drug doses that produce efficacy or side effects as well as potentially elucidating drug interactions with off-target proteins. Accurately determining target engagement allows efficacy and toxicity data to be correlated with drug selectivity in vivo [58].

Pharmacodynamics (PD) is the evaluation of what the body, biological system, organ, cell, or drug target does in response to a therapeutic intervention (i.e., a biochemical and/or physiological effect related to the mechanisms of drug action). This is in contrast to what the body or system does to the drug (Pharmacokinetics or PK). A useful PD biomarker should change (either decrease or increase) in a dose-dependent manner allowing for the measurement of a dynamic range of effect by a drug on a target or system. Establishing the relationship between drug concentration and effect (PKPD) is an essential element of the drug development process [28]. PD studies are often carried out in cells or isolated tissues (in vitro pharmacology) in early preclinical development or tissues and biological fluids isolated from treated animals or subjects (in vivo pharmacology) in the later stages of preclinical development and clinical development.

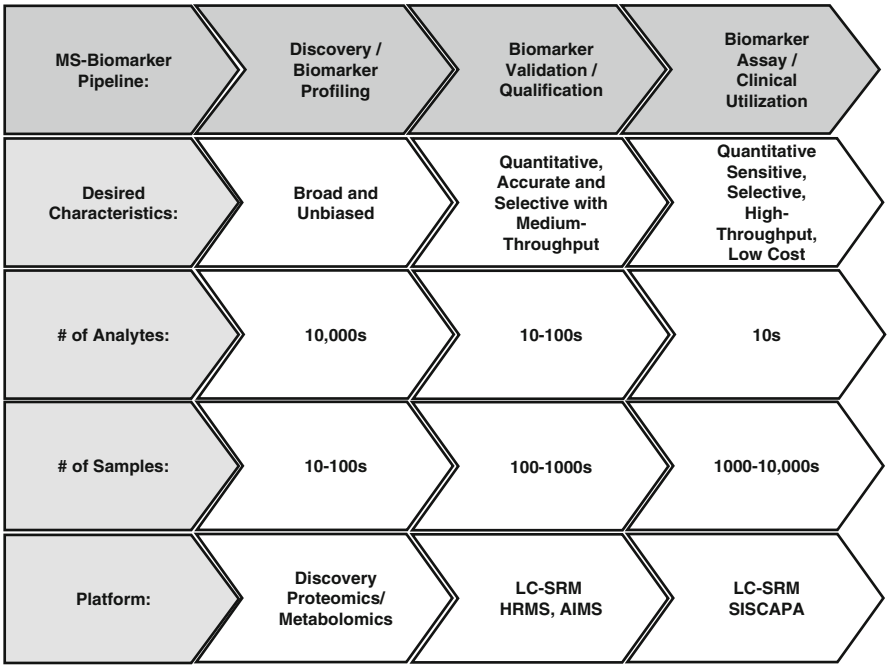


Fig. 16.1 A pipeline for the discovery and of protein-based biomarkers and development of assays for targeted validation and translation to the clinic

Safety biomarkers have been used for decades both in preclinical and clinical research and are often used to characterize specific safety issues during drug development. These biomarkers are another critical component of the drug development process and focus on understanding molecular basis of toxicity, its interplay with disease, and defining the maximal tolerated dose [39]. Characterizing safety issues typically includes the monitoring of certain systematic proteins found in the liver and kidneys [69]. These biomarkers are critical in determining the best balance of safety and efficacy, and correlation with PK/PD data allows for refinement of drug formulation, administration, and dosing/schedule schemes [39]. Safety biomarkers can also play an important role in describing the potential therapeutic advantages of new molecules over "standard of care" or competitor molecules (i.e., differentiation) [69].

Much like candidate pharmaceuticals, biomarkers can progress through a pipeline of development from discovery to the clinic (Fig. 16.1). The necessary characteristics and analytical platform of choice for a biomarker assay can vary significantly from broad, unbiased profiling to highly targeted and analytically precise. As a marker progresses through this pipeline, it may have many "fit-for-purpose" incarnations. For example, a differential proteomics experiment may identify a large list of candidate biomarkers that provide insight into a drugs mechanism of action in early preclinical development. That list may be further interrogated with a multiplexed, targeted, relative quantitation assay to further qualify which candidate markers have

utility as PD endpoints. This multiplexed assay may be employed across multiple preclinical species and reference human samples as the drug progresses through the later stages of preclinical development. Results from these experiments may then guide further refinement, optimization, development, and analytical validation of a targeted absolute quantitation assay for implementation in early clinical studies. Others have described a comprehensive process for protein biomarker discovery and validation separated into several essential steps including; discovery, qualification, verification, research assay optimization, clinical validation, and commercialization [55]. For the purposes of our discussion, we will focus on approaches to biomarker discovery, and targeted approaches for biomarker qualification and validation.

16.3 Biomarker Discovery

Biomarker discovery is most simply described as the broad and unbiased analysis of biomolecules across sample sets by employing techniques used to measure differential expression between conditions. A wide net is cast to analyze tens-to-hundreds of samples for tens of thousands of analytes in order to (hopefully) catch a few candidates with the most attractive properties to allocate resources towards pursuing. Mass spectrometry (MS) has emerged as the primary technology utilized in the discovery of biomarkers. Here we focus on protein analysis, but it should be mentioned that mass spectrometry is the dominant platform for biomarker discovery across many types of biomolecules such as lipids (lipidomics) and endogenous small molecules (metabolomics). Mass spectrometers generate mass spectra which plot the mass-to-charge ratio (m/z, where m is mass and z is charge state of the analyte) of the ions observed (x-axis) versus detected ion abundance (y axis). The characteristics of ions present in mass spectra such as accurate mass, charge, and signal intensity can provide a wealth of information including analyte identity and changes in abundance across samples. The details of some of the most promising MS-based protein biomarker discovery platforms will be discussed later.

When one initiates a biomarker discovery project it is important to have a clear vision of the end goal and sample type. Typical sample types with clinical utility would be primarily plasma (i.e., blood) or other accessible biological fluids such as urine, cerebral spinal fluid (CSF), synovial fluid, or saliva. However, if a preclinical assay is the end goal, cell culture or tissue homogenate directly from the likely model system would be preferred. Tissue sample or interstitial fluid around the tissue, from the most closely affected region is advantageous. This "proximal sample" approach isn't abundantly obtainable in human subjects, but is generally worth considering during the Discovery phase [55]. Discovery experiments in biofluids are fraught with challenges due to their immense complexity. Considering that a "healthy" adult human varies its proteome normally in various circadian stages (hourly, daily, monthly, seasonally), age-based progression/regression, and gender-based differences (menstruation, pregnancy, adolescent maturity, etc.), it is a daunting task to consider which population(s) to consider as standards to compare to the

other states described in the study (i.e., disease or treatment). Consider, too, how many complications can be compounded by certain diseases which brings layers of complexity to one disease-based target (i.e., various manifestations, states, and subtypes of a given disease). Adding to these challenging considerations, tens of thousands of proteins present in a daunting dynamic range of concentrations 10 to 11 orders of magnitude in protein abundance [2]. Cell line homogenates and tissue lysates, although often less variable within a treatment or sample group (biological replicates), pose similar problems for proteomic analysis. The Biomarker field faces yet another challenge from the influence of pre-analytical variables such as sample collection methods, sample handling and aliquoting, and sample storage. One must take great care to minimize the impact of these factors through careful planning and experimental design to either prevent or correct for any bias pre-analytical variables may introduce.

A key to addressing the complexity inherent in biological matrices is enrichment or fractionation of a targeted sub-population of proteins or peptides. Even the best discovery MS platforms are limited to 4 to 5 orders of magnitude in dynamic range of detection. Several such approaches have emerged in the last decade including; immunodepletion of abundant proteins and immunoaffinity capture of targeted proteins and peptides [19, 29], chromatographic fractionation such as strong cation exchange (SCX) [44], enrichment of modified subproteomes such as phosphoenrichment via metal affinity (TiO_2, IMAC) [6] (Fig. 16.2). There is no foreseeable "black box" method that can address all of the respective needs for biomarker discovery. The appropriate choice of sample preparation must be first be approached by critical analysis of a number of study design factors including; potential clinical needs (i.e., must be blood based, or proximal fluids are available), biology of interest (i.e., kinase substrates vs. protein complexes vs. secreted proteins…) sample source, the degree of purity required (the relative concentration of proteins or peptides of interest as compared to abundant proteins of your sample), expected modulation from the therapy (i.e., increase or decrease abundance), complexity of workflow (with a clinical method the desired goal) and available sample set for proper statistical power. Biological sample volumes vary depending on the source (tissue homogenate vs. CSF draws), and sometimes individual samples are not practical. The trade-off to small sample volume availability is to pool the sample and neglect biological variability (although certain criteria can be pooled in sets relating to sex, age, disease progression, and phenotype as necessary).

Like the sample preparation stage, the analysis, quantitation, and statistical treatment of the data have criteria that dictate the relative utility of the platform. A clear understanding of the question being asked is critical in designing a biomarker discovery experiment. Are you designing a hypothesis-generating experiment (i.e., what dose or time-dependent changes can be observed after treatment X?) or a hypothesis-testing experiment (i.e., do analytes a, b, c, and d demonstrate a significant change after drug treatment?). These two hypothetical experiments would require very different experimental designs likely with different sample numbers per treatment group and different analytical platforms.

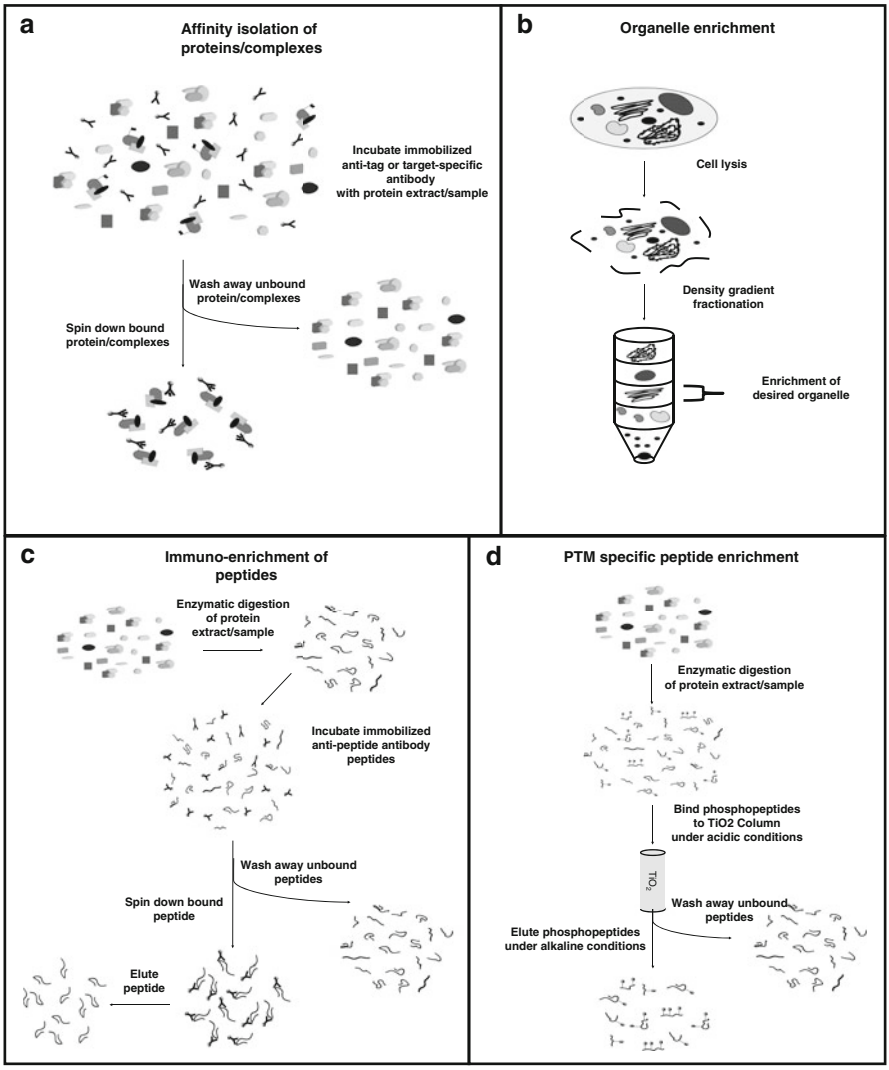


Fig. 16.2 Presented are several strategies for the targeted isolation and characterization of proteins and peptides for biomarker discovery: (**a**) Affinity isolation of proteins and their binding partners with immobilized antibodies, (**b**) enrichment of subcellular compartments and organelles through density-gradient fractionation, (**c**) immuno-enrichment of peptides with immobilized anti-peptide antibodies (**d**) enrichment of post-translationally modified peptides such as phosphopeptides with TiO_2

Once the samples are enriched, purified, and analyzed, one can quickly observe that the magnitude of information contained within datasets can be daunting. Before an attempt to recognize the best candidate(s) for follow-up, several stages of bioinformatic and biostatistical analyses are typically necessary to obtain what may

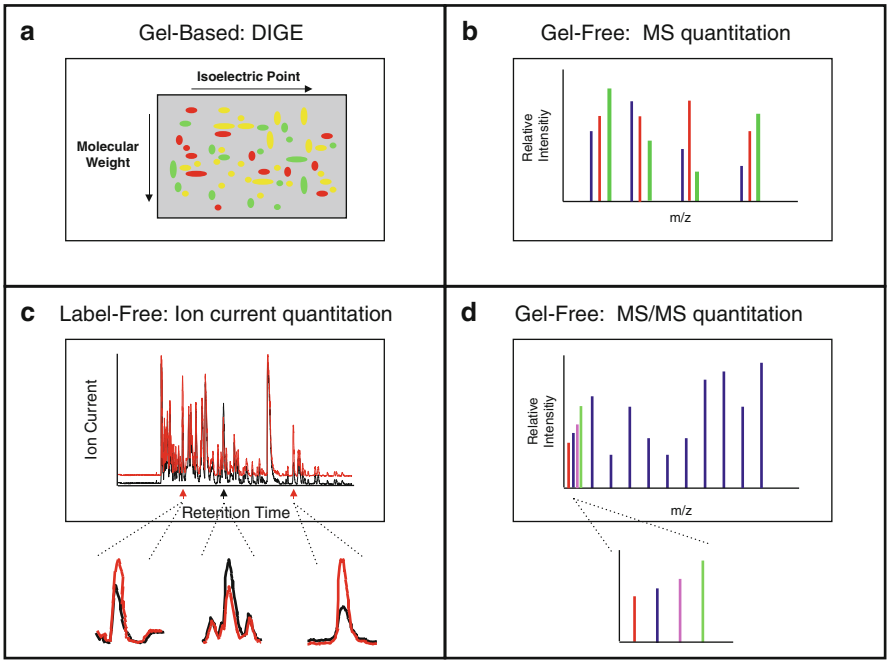


Fig. 16.3 There are several strategies for quantifying differences in abundance between protein samples: (**a**) Gel-based methods such as differential fluorescent labeling of samples followed by 2DE and image analysis for determination of relative intensities, (**b**) gel-free, chemical, enzymatic, or metabolic labeling of samples for quantitation of relative abundance of proteins based on MS signal intensities of the different labeled versions of their corresponding peptides, (**c**) quantitation of chemically labeled, isobaric peptides from protein digests from reporter ion intensities in MS/MS spectra, and (**d**) label-free quantitation where differential abundance is determined by comparing ion currents between multiple LC-MS runs

become the shortlist for the "analytes of interest." The focus typically begins wide and then narrows with each iteration of the data verified by mechanism of action studies, proof-of-concept experiments and exogenous data from collaborators, and the available literature. The initial studies are data-driven, usually obtained by differential analysis of a control (e.g., unperturbed state) compared to a perturbed state of conditions (verified disease state, treatment cohort, transgenic models, etc.). Quantitative tools such as metabolic or chemical labeling or deferential ion intensity analysis (described in more detail below and Fig. 16.3) were developed for this purpose. For comparative (differential) analysis between two states (control and disease state, for example), results can be compared from intra-run analysis to retain the fidelity of the comparison. Resulting lists of proteins should then be interrogated by a well-defined consensus of acceptance criteria. Certain questions must be entertained to avoid bias. For example, what constitutes a true change in protein expression levels? Much discussion has been debated about this topic [36], however the true answers lie within the experimental expectations. Proteins or peptides that show

differences in expression between two or more states can result in tens, hundreds, or thousands of candidates. The list that emerges from the Discovery experiment should proceed to be further evaluated using targeted approaches such as those described below to weed out the false positives inherent with the current platform.

16.4 Proteomic Biomarker Discovery Workflows

In most protein-based workflows, proteins are enzymatically converted (usually by trypsin digestion) into their constituent peptides. This is commonly referred to as the "bottom–up" approach. Once digested, the complex mixture of peptides can be resolved by methods such as ultra-high pressure liquid chromatography (UPLC) in line with the mass spectrometer (LC-MS). Sequence information can then be obtained by tandem MS (MS/MS) in which precursor ions are fragmented by collision with gas (e.g., N_2, Ar) to produce fragment ions. The resulting data is provided as fragmentation lists containing both fragment mass and precursor mass data. These raw data reports can be uploaded into dedicated software or web-based programs (examples include SEAQUEST (http://fields.scripps.edu/sequest) and MASCOT (http://www.matrixscience.com)) and searched against in silico digests of appropriate protein and nucleotide databases. Peptide identities are assigned based on the degree of correlation between the experimentally observed and the theoretical spectra. Protein identities are reported back by concatenating identified peptides and scores (often probability based) which are assigned to each peptide and subsequently assigned in the prediction of the progenitor protein [1]. Several factors can affect the outcome's integrity including, but not limited to, the mass accuracy of the peptide precursor mass, and methods to assess the statistical validity of the results and obtain a high-confidence list of peptide and protein identities are available [4, 9, 45, 46, 57, 62]. In order to maximize instrument sensitivity, chromatographic separation is often performed on columns ~75 μm in diameter and at flow rates of ~200 nL/min in order to allow detection of peptides at the low nanogram per milliliter level [21]. An alternative approach that has emerged over the last several years is the so called "top–down" approach wherein proteins are measured at the intact (i.e., non-cleaved) level. This approach enables the detection of protein isoforms and/or post-translationally modified forms as independent species and can be directly interpreted, resulting in accurate, relevant biological information [33, 41, 59].

Despite the efforts made to isolate protein or protein subsets from the original matrix (enrichment, depletion, fractionation, see Fig. 16.2) and digestion products from other peptides (chromatography), the mass spectrometer is still subjected to an immensely complex sample with components which work against the ionization and detection of analytes of interest [55]. Ideally, a true representation of the proteome would be accessible from the sample preparation though LC/MS platform, but this is hardly the case. Even with fractionation and other protein purification techniques, the limitations of typical HPLC analysis are manifested in the knowledge that only soluble peptides are being injected and separated.

Other chromatographic techniques are available to exploit less soluble peptides and proteins (membrane-based, for example) which will not be covered here [70]. The detection aspect of decoding the complex mixture of peptides, even when chromatographically separated is subjected to the instrument's ability to resolve peptide masses to parts-per-billion mass differentials. In addition, as can be expected, database search results can only be as reliable as the instrument accuracy of the input masses. During the Discovery phase, ion trap mass spectrometers lend their strength to these requirements, such as Thermo Fisher's (Waltham, MA, USA) Orbitrap or Waters' (Milford, MA, USA) SYNAPT lines of mass spectrometers which deliver the requisite high performance necessary in the search for unbiased database matches.

Approaches incorporating MS for proteomic biomarker discovery include full scan (precursor scan) ion intensity analysis such as differential mass spectrometry (dMS) [42, 51], pattern-based methods that produce MS-derived protein or peptide patterns via matrix-assisted laser desorption–ionization (MALDI) [67], analysis of selected spots from differential protein displays, such as two-dimensional electrophoresis (2DE) [53], and LC-MS/MS identity-based methods that yield lists of peptides in combination with relative quantitation information from metabolic, enzymatic, or chemical labeling approaches [27, 49, 50, 56, 61, 65, 74] (Fig. 16.3).

Historically, the most widely used of quantitative proteomic approaches has been 2DE. The most common form of 2DE involves isoelectric focusing of proteins using immobilized pH gradient (IPG) strips followed by a dimension of SDS-PAGE to separate proteins by molecular weight. This approach has been shown to have reasonable sensitivity and resolution with the numbers of spots identified reaching into the thousands for a single gel [53]. However, necessary expertise, difficult optimization for specific samples, limited reproducibility, quantitative sensitivity (often due to the limits of dyes), loss of protein at the extreme ends of the pH scale as well as hydrophobic proteins, and the need for several replicates limit the practical uses and throughput of 2DE. A more recent improvement of 2DE analysis has been the introduction of fluorescent dyes for use in difference gel electrophoresis (DIGE) analysis [63]. DIGE systems use size and charge-matched, spectrally resolvable fluorescent dyes that allow for loading of multiple samples into a single gel, eliminating problems of gel-to-gel variations (Fig. 16.3a). These methods rely on image analysis for quantitation, limiting those proteins quantified to those that can be visualized by a specific stain. Even those which limit gel-to-gel variation such as DIGE still require separate labeling of samples after the chosen isolation or enrichment protocol and introduce sample handling variations for even the most experienced users.

Perhaps even more powerful are gel-free quantitative methods including chemical, enzymatic, and metabolic labeling of specific protein populations and subsequent quantitation with the mass spectrometer. One of the first gel-free quantitation methods developed was ICAT (isotopically labeled chemical affinity tags) [27]. The approach involves chemically (i.e., covalently) labeling protein peptides with isotopic tags. ICAT experiments utilize heavy (i.e., isotopically labeled) and light (i.e., unlabeled) versions of the ICAT reagent that are incorporated into protein

isolates from multiple conditions, these populations are combined for MS analysis, and their relative intensities determined (Fig. 16.3b). A similar concept has been employed in iTRAQ (isobaric tags for relative and absolute quantification) and TMT (Tandem Mass Tags) with the distinction that chemical labels are isobaric and yield unique fragment ions that can be quantified in MS/MS spectra (Fig. 16.3c) [56, 65]. Enzymatic labeling with protein hydrolysis in the presence of heavy water (H_2O_{18}) is another widely used approach to distinguish between two samples and quantify relative intensities in the mass spectrometer [61, 74]. These approaches allow multiple samples to be mixed after digestion and labeling, and enable simultaneous analysis on the mass spectrometer. The benefits are self-evident. Inherent errors caused by attempting to align serial injections are eliminated, and the mixed sample is self-normalizing by the described "standard" sample (usually the control) enhancing sample-to-sample accuracy regardless of processing inconsistencies or errors. Both chemical and enzymatic labeling will continue to be useful in their application to samples which cannot be metabolically labeled such as tissue or clinical samples.

An approach that eliminates most sample-to-sample variation is the use of metabolic labeling. Incorporation of deuterium and heavy isotopes of nitrogen and carbon into proteins in cell culture [49, 50] or in vivo (*C. elegans*, *D. melonogaster* [34], *R. norvegicus* [73], *S. solfataricus* [60], and *M. musculus* [76]) allows for relative quantitation of differentially labeled versions of peptides in the mass spectrometer. The most popular of metabolic labeling approaches is SILAC (stable isotopic labeling of amino acids in culture) [49, 50]. With such approaches, after protein extraction, samples are combined and all treatments they are subject to are identical. This is in contradiction to all of the above mentioned approaches which require parallel handling of samples. Another measurement which exploits metabolic labeling techniques is the determination of protein half-life and kinetics. This technique enables assessment of protein, synthesis, turnover and degradation, and the effects of a specific stimulus on those processes [18, 25, 52].

Label-free quantitation with the mass spectrometer is one technique that has shown great promise in the field of biomarkers. One of the most prominent of these approaches is full scan (precursor scan) ion current-based quantitation, often referred to as dMS [42, 51]. These methods compare ion intensities of specific accurate *m/z* profiles between multiple LC-MS runs. The results are typically subjected to chromatographic processing such as baseline correction and normalization, and differential signals are extracted (Fig. 16.3d). Data acquired using such approaches on state or the art high resolution mass spectrometers (e.g., Thermo Orbitrap and Waters SYNAPT) and with high mass accuracy (single digit ppm) enables detection and measurement of thousands of analytes in the same mass spectrum, and resolution of seemingly isobaric co-eluting species [22, 42]. Powerful software tools are required to analyze such data and several algorithms have been developed (Elucidator, Rosetta Biosoftware, MaxQuant (www.maxquant.org)) Once differential signals are apparent the focus turns to defining what specific *m/z* signals are responsible, often through more targeted MS experiments.

16.5 Biomarker Qualification and Validation

There are two primary components of biomarker validation, namely (1) validation of the assay or method and (2) clinical validation (e.g., qualification) [69]. Qualification involves correlating biomarker dynamics to the observed state, and ascribing a level of "significance" for describing the state under observable phenotypical or behavioral perturbation. Qualification also often includes correlating multiple analytical methods (e.g., ELISA, Western Blotting, etc.) to describe the same end-point [22]. Validation verifies that the method used shows rugged, repeatable, and reliable evidence that the findings are accurate within the experimental conditions [26]. Validation of biomarkers is not trivial especially in light of population variance when attempting to describe the normal from the abnormal levels en route to disease diagnosis [55]. Generally, the goals of the method strive for undisputed biomarker selectivity and a sensitivity that can be used to quantitatively monitor increases or decreases in endogenous concentration deemed as "significant" during qualification. Such correlations can require large population data, translating into a need for more samples and even replicates of such samples. The results are used to provide high-powered statistically significant data that can be used to infer an individual's result to an "expected" range and the biomarker's respective action thresholds. The higher the population, the more powerful the results and the more sensitive the biomarker assay becomes for diagnosis and monitoring [22].

The general MS-based approach to biomarker qualification and validation employs targeted LC-MS/MS, multiple reaction monitoring (MRM) experiments and stable isotope dilution (SID) [14, 15] (Fig. 16.4). The SID workflow incorporates isotopically labeled peptides as internal standards for trypsin-cleaved endogenous (unlabeled) peptides. The internal standard is processed along with the sample to minimize issues inherent with workflow loss. Finally, the mixture is analyzed by LC-MS/MS MRM. The advantages to this approach are, simultaneous qualitative identification and quantitative analysis, by comparing the ratio of endogenous-to-standard peptides [5] (described below in more detail). In addition to the targeted quantitative approaches described below, an intermediate and targeted relative quantitation approach, such as a multiplexed LC-MS/MS MRM or data-independent acquisition on a discovery platform, can be useful in narrowing a list of candidate biomarkers in follow-up qualification experiments [31].

16.6 Targeted Mass Spectrometry Assays

Once a biomarker is selectively identified, verification and quantitative validation can be attempted through the use of mass spectrometry. Triple quadrupole instruments have a distinct advantage in the building of targeted biomarker assay platforms due to the selectivity, the sensitivity, and the ability to monitor multiple reactions simultaneously throughout the duration of a chromatographic elution gradient. The evolution of computational processing speeds has become a boon to attaining

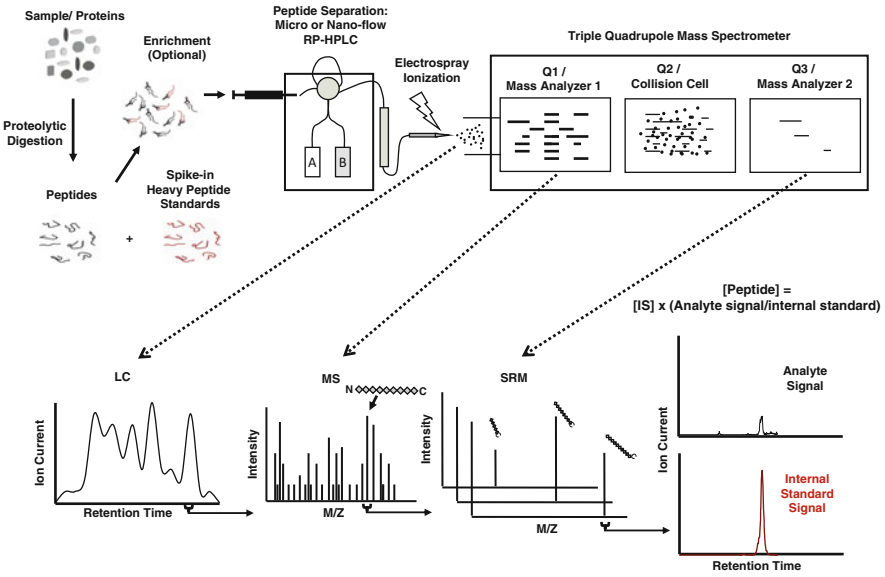


Fig. 16.4 The mass spectrometry approach to quantitation of protein markers. Selected/multiple reaction monitoring provides the high sensitivity and amino acid selectivity needed to detect peptides in biological samples. Peptide concentrations are obtained by multiplying the [IS] concentration by the signal ratio of the analyte/internal standard

millisecond rapid duty cycles. As a result, triple quadrupole mass spectrometers have the ability to "simultaneously" monitor extensive lists of single reaction transitions (e.g., product or fragment ions) allowing for rapid serial iterations of sequence analysis for multiple peptides. The "simultaneous" detection is a misnomer, since it is the duty cycle which is iteratively monitoring discreet, verified transitions collectively in microsecond time windows. These windows are the key to the accumulation of selective, sensitive and accurate, qualitative and quantitative scans. Confidence in the accuracy of biomarker identification and quantitation assays is improved as a function of the number of transitions monitored per sequence. Reporting data based on a single transition is not as strong as monitoring several. However, the combination of precursor mass accuracy, novel, predictable and verifiable fragmentation patterns, and sequence-specific retention times contribute confidence in interpreting even single transition data [14, 15, 66, 71, 75].

MRM of a single sequence provides confidence by increasing the number of subsequent transitions which act as internal controls for verifying the precursor's molecular weight and retention time. As computational power advances, so does the ability to reduce MS scan rates and the ability for a mass spectrometer to change scanning focus in real-time (especially for triple quadrupole instruments). Faster in-line chromatography options have also contributed to overall throughput advances in time and consumable usage. Ultra-high pressure liquid chromatography (as previously noted, UPLC) delivers peptide separations on an abbreviated timeline.

Previous to UPLC, its progenitor (high performance liquid chromatography, HPLC) would typically require an hour or more over an hour of gradient time to adequately separate even a fractionated peptide mixture. The historic trade-off to peptide peak selectivity was peak broadening that inadvertently affects sensitivity by lowering the peak height and challenging peak integration (used to calculate overall area which is correlated to sample concentration). These run times have been decreased considerably with little sacrifice to resolution. Column bore sizes have also been reduced to a point where nanoliter/minute flow times are possible simultaneously allowing sub-microliter injection volumes thereby reducing sample volume requirements which can also translate into higher observed concentration sensitivity options. As described previously, the use of isotopically labeled peptide standards and SID can allow this targeted approach to infer absolute quantitation. The mass spectrometer can exploit the co-elution of labeled and endogenous analytes for peptide identification purposes while measuring the distinct molecular weights of the precursor and resulting product ions [5, 14, 15] (Fig. 16.4).

A list of peptide biomarkers can be analyzed in a single chromatographic run, with several transitions per peptide within an order of μs/scan. The list can be further expanded by scheduling ranges of detection within the chromatographic run itself allowing an increase in targeted analysis for one injection. The development of such methods are considerably time consuming, but powerful once validated. Synthetic standard peptides or purified, digested protein should be considered in order to ensure the expected accuracy in elucidating transition state conditions. Several vendor-specific software, as well as publically available programs such as Skyline (MacCoss Laboratories, https://skyline.gs.washington.edu/), provide ways to decrease the overall time of method development by calculating the estimated collision energy (CE) of any peptide sequence entered using [CE = slope × (precursor m/z) + intercept], which then can be verified extrinsically. An entire protein sequence, a list of proteins, or a complete proteome can be uploaded into the program and, once the digesting enzyme is selected, an algorithm will automatically calculate the cleavage sites (with or without post-translational modifications). Target peptides can then be injected for MRM analysis providing verification of the digested peptides in the sample. If the option is selected to calculate collision energy values, a graph is generated showing whether or not the calculation is accurate. Despite the time-consuming process, the resulting method proves to be highly selective, sensitive, robust, and repeatable.

An intriguing targeted analysis platform has become recently available which involves the exploitation of anti-peptide antibodies for fractionation and purification. The method has been described as “Stable Isotope Standards and Capture By Anti-Peptide Antibodies” or “SISCAPA” (SISCAPA Assay Technologies, Inc., Washington, DC; (http://www.siscapa.com/) [3]). The utility of this approach is derived from the selectivity of the anti-peptide antibodies, which allow for low abundance peptides to be concentrated from samples without over-processing and risking loss. Due to the specificity of the immune-enrichment, an initial pre-fractionation step can be eliminated. The enriched peptide sample can then be analyzed by MRM-based targeted analysis resulting in sensitive characterization, and due to

the added advantage of introducing a sequence-complimentary stable isotope, the method is quantitative. The available anti-peptide antibody library is already considerable and expanding. Advantages include relative affordability and availability of the anti-peptide antibodies.

16.7 Conclusions

Appropriate implementation of a biomarker or a combination of biomarkers can enhance understanding of a drugs mechanism, facilitate effective translation from the preclinical to clinical space, enable early proof of concept and dose selection, and increase the efficiency of drug development. Here we have provided descriptions of the different classes of biomarkers that have utility in the drug development process and reviewed specific, protein-centric, mass spectrometry-based approaches for the discovery of biomarkers and development of targeted assays to measure these markers in a selective and analytically precise manner. Biomarkers, when properly validated and employed, have the potential to enhance disease detection, accelerate the drug development process, and improve drug safety.

References

1. Aebersold R, Mann M (2003) Mass spectrometry-based proteomics. Nature 422(6928): 198–207
2. Anderson NL, Anderson NG (2002) The human plasma proteome: history, character, and diagnostic prospects. Mol Cell Proteomics 1(11):845–867
3. Anderson NL, Anderson NG, Haines LR, Hardie DB, Olafson RW, Pearson TW (2004) Mass spectrometric quantitation of peptides and proteins using stable isotope standards and capture by anti-peptide antibodies (SISCAPA). J Proteome Res 3(2):235–244
4. Baldwin MA (2004) Protein identification by mass spectrometry: issues to be considered. Mol Cell Proteomics 3(1):1–9
5. Barr JR, Maggio VL, Patterson DG Jr, Cooper GR, Henderson LO, Turner WE, Smith SJ, Hannon WH, Needham LL, Sampson EJ (1996) Isotope dilution—mass spectrometric quantification of specific proteins: model application with apolipoprotein A-I. Clin Chem 42(10): 1676–1682
6. Batalha IL, Lowe CR, Roque AC (2012) Platforms for enrichment of phosphorylated proteins and peptides in proteomics. Trends Biotechnol 30(2):100–110
7. Biomarkers Definitions Working Group (2001) Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. Clin Pharmacol Ther 69(3):89–95
8. Breitling R (2006) Biological microarray interpretation: the rules of engagement. Biochim Biophys Acta 1759(7):319–327
9. Carr S, Aebersold R, Baldwin M, Burlingame A, Clauser K, Nesvizhskii A (2004) The need for guidelines in publication of peptide and protein identification data: working group on publication guidelines for peptide and protein identification data. Mol Cell Proteomics 3(6):531–533
10. Carser JE, Quinn JE, Michie CO, O'Brien EJ, Mccluggage WG, Maxwell P, Lamers E, Lioe TF, Williams AR, Kennedy RD, Gourley C, Harkin DP (2011) BRCA1 is both a prognostic

and predictive biomarker of response to chemotherapy in sporadic epithelial ovarian cancer. Gynecol Oncol 123(3):492–498
11. Chandra H, Reddy PJ, Srivastava S (2011) Protein microarrays and novel detection platforms. Expert Rev Proteomics 8(1):61–79
12. Chung C, Christianson M (2014) Predictive and prognostic biomarkers with therapeutic targets in breast, colorectal, and non-small cell lung cancers: a systemic review of current development, evidence, and recommendation. J Oncol Pharm Pract 20(1):11–28
13. Danesh J, Wheeler JG, Hirschfield GM, Eda S, Eiriksdottir G, Rumley A, Lowe GD, Pepys MB, Gudnason V (2004) C-reactive protein and other circulating markers of inflammation in the prediction of coronary heart disease. N Engl J Med 350(14):1387–1397
14. Desiderio DM, Kai M (1983) Preparation of stable isotope-incorporated peptide internal standards for field desorption mass spectrometry quantification of peptides in biologic tissue. Biomed Mass Spectrom 10(8):471–479
15. Desiderio DM, Kai M, Tanzer FS, Trimble J, Wakelyn C (1984) Measurement of enkephalin peptides in canine brain regions, teeth, and cerebrospinal fluid with high-performance liquid chromatography and mass spectrometry. J Chromatogr 297:245–260
16. Dhanasekaran SM, Barrette TR, Ghosh D, Shah R, Varambally S, Kurachi K, Pienta KJ, Rubin MA, Chinnaiyan AM (2001) Delineation of prognostic biomarkers in prostate cancer. Nature 412(6849):822–826
17. Dimasi JA, Hansen RW, Grabowski HG (2003) The price of innovation: new estimates of drug development costs. J Health Econ 22(2):151–185
18. Elbert DL, Mawuenyega KG, Scott EA, Wildsmith KR, Bateman RJ (2008) Stable isotope labeling tandem mass spectrometry (SILT): integration with peptide identification and extension to data-dependent scans. J Proteome Res 7(10):4546–4556
19. Fang X, Zhang WW (2008) Affinity separation and enrichment methods in proteomic analysis. J Proteomics 71(3):284–303
20. Frank R, Hargreaves R (2003) Clinical biomarkers in drug discovery and development. Nat Rev Drug Discov 2(7):566–580
21. Gatlin CL, Kleemann GR, Hays LG, Link AJ, Yates JR III (1998) Protein identification at the low femtomole level from silver-stained gels using a new fritless electrospray interface for liquid chromatography-microspray and nanospray mass spectrometry. Anal Biochem 263(1): 93–101
22. Gillette MA, Mani DR, Carr SA (2005) Place of pattern in proteomic biomarker discovery. J Proteome Res 4(4):1143–1154
23. Gold L, Ayers D, Bertino J, Bock C, Bock A, Brody EN, Carter J, Dalby AB, Eaton BE, Fitzwater T, Flather D, Forbes A, Foreman T, Fowler C, Gawande B, Goss M, Gunn M, Gupta S, Halladay D, Heil J, Heilig J, Hicke B, Husar G, Janjic N, Jarvis T, Jennings S, Katilius E, Keeney TR, Kim N, Koch TH, Kraemer S, Kroiss L, Le N, Levine D, Lindsey W, Lollo B, Mayfield W, Mehan M, Mehler R, Nelson SK, Nelson M, Nieuwlandt D, Nikrad M, Ochsner U, Ostroff RM, Otis M, Parker T, Pietrasiewicz S, Resnicow DI, Rohloff J, Sanders G, Sattin S, Schneider D, Singer B, Stanton M, Sterkel A, Stewart A, Stratford S, Vaught JD, Vrkljan M, Walker JJ, Watrobka M, Waugh S, Weiss A, Wilcox SK, Wolfson A, Wolk SK, Zhang C, Zichi D (2010) Aptamer-based multiplexed proteomic technology for biomarker discovery. PLoS One 5(12):e15004
24. Grimwood S, Hartig PR (2009) Target site occupancy: emerging generalizations from clinical and preclinical studies. Pharmacol Ther 122(3):281–301
25. Gustavsson N, Greber B, Kreitler T, Himmelbauer H, Lehrach H, Gobom J (2005) A proteomic method for the analysis of changes in protein concentrations in response to systemic perturbations using metabolic incorporation of stable isotopes and mass spectrometry. Proteomics 5(14):3563–3570
26. Gutman S, Kessler LG (2006) The US Food and Drug Administration perspective on cancer biomarker development. Nat Rev Cancer 6(7):565–571
27. Gygi SP, Rist B, Gerber SA, Turecek F, Gelb MH, Aebersold R (1999) Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. Nat Biotechnol 17(10):994–999

28. Rang HP (2006) Pharmacology: its role in drug discovery. In: Rang HP (ed) Drug discovery and development: technology in transition. Elsevier, Philadelphia
29. Hakimi A, Auluck J, Jones GD, Ng LL, Jones DJ (2014) Assessment of reproducibility in depletion and enrichment workflows for plasma proteomics using label-free quantitative data-independent LC-MS. Proteomics 14(1):4–13
30. Institute Of Medicine (US) Forum On Drug Discovery, D. A. T (2009) Accelerating the development of biomarkers for drug safety: workshop summary. The National Academies Collection: reports funded by National Institutes of Health. National Academies Press (US), Washington
31. Jaffe JD, Keshishian H, Chang B, Addona TA, Gillette MA, Carr SA (2008) Accurate inclusion mass screening: a bridge from unbiased discovery to targeted assay development for biomarker verification. Mol Cell Proteomics 7(10):1952–1962
32. Peters KE, Walters CC, Moldowan JM (2007) The biomarker guide: biomarkers and isotopes in the environment and human history, 2nd edn. Cambridge University Press, Cambridge
33. Kelleher NL (2004) Top-down proteomics. Anal Chem 76(11):197A–203A
34. Krijgsveld J, Ketting RF, Mahmoudi T, Johansen J, Artal-Sanz M, Verrijzer CP, Plasterk RH, Heck AJ (2003) Metabolic labeling of C. elegans and D. melanogaster for quantitative proteomics. Nat Biotechnol 21(8):927–931
35. Krishna R, Herman G, Wagner JA (2008) Accelerating drug development using biomarkers: a case study with sitagliptin, a novel DPP4 inhibitor for type 2 diabetes. AAPS J 10(2):401–409
36. Li Q (2010) Assigning significance in label-free quantitative proteomics to include single-peptide-hit proteins with low replicates. Int J Proteomics 2010
37. Mahajan R, Gupta K (2010) Food and drug administration's critical path initiative and innovations in drug development paradigm: challenges, progress, and controversies. J Pharm Bioallied Sci 2(4):307–313
38. Marquette CA, Corgier BP, Blum LJ (2012) Recent advances in multiplex immunoassays. Bioanalysis 4(8):927–936
39. Marrer E, Dieterle F (2010) Impact of biomarker development on drug safety assessment. Toxicol Appl Pharmacol 243(2):167–179
40. Matthews PM, Rabiner EA, Passchier J, Gunn RN (2012) Positron emission tomography molecular imaging for drug development. Br J Clin Pharmacol 73(2):175–186
41. Mazur MT, Cardasis HL, Spellman DS, Liaw A, Yates NA, Hendrickson RC (2010) Quantitative analysis of intact apolipoproteins in human HDL by top-down differential mass spectrometry. Proc Natl Acad Sci U S A 107(17):7728–7733
42. Meng F, Wiener MC, Sachs JR, Burns C, Verma P, Paweletz CP, Mazur MT, Deyanova EG, Yates NA, Hendrickson RC (2007) Quantitative analysis of complex peptide mixtures using FTMS and differential mass spectrometry. J Am Soc Mass Spectrom 18(2):226–233
43. Mor G, Visintin I, Lai Y, Zhao H, Schwartz P, Rutherford T, Yue L, Bray-Ward P, Ward DC (2005) Serum protein markers for early detection of ovarian cancer. Proc Natl Acad Sci U S A 102(21):7677–7682
44. Motoyama A, Yates JR III (2008) Multidimensional LC separations in shotgun proteomics. Anal Chem 80(19):7187–7193
45. Nesvizhskii AI (2010) A survey of computational methods and error rate estimation procedures for peptide and protein identification in shotgun proteomics. J Proteomics 73(11): 2092–2123
46. Nesvizhskii AI, Aebersold R (2004) Analysis, statistical validation and dissemination of large-scale proteomics datasets generated by tandem MS. Drug Discov Today 9(4):173–181
47. Ng PC, Kirkness EF (2010) Whole genome sequencing. Methods Mol Biol 628:215–226
48. O'Connor JP, Jackson A, Asselin MC, Buckley DL, Parker GJ, Jayson GC (2008) Quantitative imaging biomarkers in the clinical development of targeted therapeutics: current and future perspectives. Lancet Oncol 9(8):766–776
49. Ong SE, Blagoev B, Kratchmarova I, Kristensen DB, Steen H, Pandey A, Mann M (2002) Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. Mol Cell Proteomics 1(5):376–386

50. Ong SE, Foster LJ, Mann M (2003) Mass spectrometric-based approaches in quantitative proteomics. Methods 29(2):124–130
51. Paweletz CP, Wiener MC, Bondarenko AY, Yates NA, Song Q, Liaw A, Lee AY, Hunt BT, Henle ES, Meng F, Sleph HF, Holahan M, Sankaranarayanan S, Simon AJ, Settlage RE, Sachs JR, Shearman M, Sachs AB, Cook JJ, Hendrickson RC (2010) Application of an end-to-end biomarker discovery platform to identify target engagement markers in cerebrospinal fluid by high resolution differential mass spectrometry. J Proteome Res 9(3):1392–1401
52. Pratt JM, Petty J, Riba-Garcia I, Robertson DH, Gaskell SJ, Oliver SG, Beynon RJ (2002) Dynamics of protein turnover, a missing dimension in proteomics. Mol Cell Proteomics 1(8): 579–591
53. Rabilloud T (2002) Two-dimensional gel electrophoresis in proteomics: old, old fashioned, but it still climbs up the mountains. Proteomics 2(1):3–10
54. Rawlins MD (2004) Cutting the cost of drug development? Nat Rev Drug Discov 3(4): 360–364
55. Rifai N, Gillette MA, Carr SA (2006) Protein biomarker discovery and validation: the long and uncertain path to clinical utility. Nat Biotechnol 24(8):971–983
56. Ross PL, Huang YN, Marchese JN, Williamson B, Parker K, Hattan S, Khainovski N, Pillai S, Dey S, Daniels S, Purkayastha S, Juhasz P, Martin S, Bartlet-Jones M, He F, Jacobson A, Pappin DJ (2004) Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. Mol Cell Proteomics 3(12):1154–1169
57. Sadygov RG, Liu H, Yates JR (2004) Statistical models for protein validation using tandem mass spectral data and protein amino acid sequence databases. Anal Chem 76(6):1664–1671
58. Simon GM, Niphakis MJ, Cravatt BF (2013) Determining target engagement in living systems. Nat Chem Biol 9(4):200–205
59. Siuti N, Kelleher NL (2007) Decoding protein modifications using top-down mass spectrometry. Nat Methods 4(10):817–821
60. Snijders AP, De Vos MG, Wright PC (2005) Novel approach for peptide quantitation and sequencing based on 15N and 13C metabolic labeling. J Proteome Res 4(2):578–585
61. Staes A, Demol H, Van Damme J, Martens L, Vandekerckhove J, Gevaert K (2004) Global differential non-gel proteomics by quantitative and stable labeling of tryptic peptides with oxygen-18. J Proteome Res 3(4):786–791
62. States DJ, Omenn GS, Blackwell TW, Fermin D, Eng J, Speicher DW, Hanash SM (2006) Challenges in deriving high-confidence protein identifications from data gathered by a HUPO plasma proteome collaborative study. Nat Biotechnol 24(3):333–338
63. Swatton JE, Prabakaran S, Karp NA, Lilley KS, Bahn S (2004) Protein profiling of human postmortem brain using 2-dimensional fluorescence difference gel electrophoresis (2-D DIGE). Mol Psychiatry 9(2):128–143
64. Tello-Montoliu A, Marin F, Roldan V, Mainar L, Lopez MT, Sogorb F, Vicente V, Lip GY (2007) A multimarker risk stratification approach to non-ST elevation acute coronary syndrome: implications of troponin T, CRP, NT pro-BNP and fibrin D-dimer levels. J Intern Med 262(6):651–658
65. Thompson A, Schafer J, Kuhn K, Kienle S, Schwarz J, Schmidt G, Neumann T, Johnstone R, Mohammed AK, Hamon C (2003) Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. Anal Chem 75(8):1895–1904
66. Tiller PR, Cunniff J, Land AP, Schwartz J, Jardine I, Wakefield M, Lopez L, Newton JF, Burton RD, Folk BM, Buhrman DL, Price P, Wu D (1997) Drug quantitation on a benchtop liquid chromatography-tandem mass spectrometry system. J Chromatogr A 771(1–2):119–125
67. Villanueva J, Philip J, Entenberg D, Chaparro CA, Tanwar MK, Holland EC, Tempst P (2004) Serum peptide profiling by magnetic particle-assisted, automated sample processing and MALDI-TOF mass spectrometry. Anal Chem 76(6):1560–1570
68. Visscher PM, Brown MA, Mccarthy MI, Yang J (2012) Five years of GWAS discovery. Am J Hum Genet 90(1):7–24
69. Wagner JA (2008) Strategic approach to fit-for-purpose biomarkers in drug development. Annu Rev Pharmacol Toxicol 48:631–651

70. Whitelegge JP (2013) Integral membrane proteins and bilayer proteomics. Anal Chem 85(5): 2558–2568
71. Wieboldt R, Campbell DA, Henion J (1998) Quantitative liquid chromatographic-tandem mass spectrometric determination of orlistat in plasma with a quadrupole ion trap. J Chromatogr B Biomed Sci Appl 708(1–2):121–129
72. Wong DF, Tauscher J, Grunder G (2009) The role of imaging in proof of concept for CNS drug discovery and development. Neuropsychopharmacology 34(1):187–203
73. Wu CC, Maccoss MJ, Howell KE, Matthews DE, Yates JR III (2004) Metabolic labeling of mammalian organisms with stable isotopes for quantitative proteomic analysis. Anal Chem 76(17):4951–4959
74. Yao X, Freas A, Ramirez J, Demirev PA, Fenselau C (2001) Proteolytic 18O labeling for comparative proteomics: model studies with two serotypes of adenovirus. Anal Chem 73(13): 2836–2842
75. Yost RA, Perchalski RJ, Brotherton HO, Johnson JV, Budd MB (1984) Pharmaceutical and clinical analysis by tandem mass spectrometry. Talanta 31(10 Pt 2):929–935
76. Zanivan S, Krueger M, Mann M (2012) In vivo quantitative proteomics: the SILAC mouse. Methods Mol Biol 757:435–450

Chapter 17
Detection of Biomedically Relevant Stilbenes from Wines by Mass Spectrometry

Veronica Andrei *, Armand G. Ngounou Wetie *, Iuliana Mihai, Costel C. Darie, and Alina Vasilescu

Abstract Stilbenes represent a class of compounds with a common 1,2-diphenylethylene backbone that have shown extraordinary potential in the biomedical field. As the most well-known example, resveratrol proved to have anti-aging effects and significant potential in the fight against cardiovascular diseases and some types of cancer. Mass spectrometry is an analytical method of critical importance in all studies related to stilbenes that are important in the biomedical field. From the discovery of new natural compounds and mapping the grape metabolome up to advanced investigations of stilbenes' potential for the protection of human health in clinical studies, mass spectrometry has provided critical analytical information. In this review we focus on various approaches related to mass spectrometry for the detection of stilbenes—such as coupling with chromatographic separation methods and direct infusion—with presentation of some illustrative applications. Clearly, the potential of mass spectrometry for assisting in the discovery of new stilbenes of biomedical importance, elucidating their mechanisms of action, and quantifying minute quantities in complex matrices is far from being exhausted.

*Author contributed equally with all other contributors.

V. Andrei • I. Mihai • A. Vasilescu (✉)
International Centre of Biodynamics, 1B Intrarea Portocalelor, Sector 6, Bucharest 060101, Romania
e-mail: avasilescu@biodyn.ro

A.G. Ngounou Wetie • C.C. Darie
Biochemistry and Proteomics Group, Department of Chemistry and Biomolecular Science, Clarkson University, 8 Clarkson Avenue, Potsdam, NY 13699-5810, USA

A.G. Woods and C.C. Darie (eds.), *Advancements of Mass Spectrometry in Biomedical Research*, Advances in Experimental Medicine and Biology 806,
DOI 10.1007/978-3-319-06068-2_17, © Springer International Publishing Switzerland 2014

17.1 Introduction

Consumption of wines in moderation can have a positive influence on the human body due to their high content in antioxidants. Antioxidants have the ability to scavenge-free radicals species which are formed during metabolic processes in human cells which may inhibit normal functioning of the cells [1]. Ever since the French Paradox relating wine consumption to the low prevalence of cardiovascular diseases in spite of the unhealthy diet in France [2], wine antioxidants have been the focus of intensive research, with one compound, resveratrol, becoming particularly famous.

Resveratrol (3,5,4-trihydroxystilbene) is a stilbene, part of the phytoalexins group, an important class of de novo synthesized substances as response to pathogenic infections in plants [3, 4]. Resveratrol is found as a mixture of *cis* and *trans* isomers, synthesized in plants by stilbene-synthase [5]. The oxidative dimerization of resveratrol leads to oligomers called viniferins [6] (Fig. 17.1). Stilbenes (from stilbos- "shining" in Greek) are a class of substances with a common 1,2-diphenylethylene backbone. The *cis*- and *trans*-isomers of stilbenes have different pharmacological activities; the *trans*-isomer of resveratrol, for example, performs better in antioxidant and anticancer assays. Hydroxylated derivatives of stilbenes are called stilbenoids.

Resveratrol and resveratrol derivatives have been shown to be powerful antioxidants [3]. As many diseases develop due to the production of radical oxygen species in the human body, in the quest for new strategies against diseases, resveratrol and related stilbenes have been used in many in vivo and in vitro studies investigating their potential beneficial effect on human health [7–11]. A range of protective and preventive effects are currently widely attributed to resveratrol and its derivatives, including anti-aging [8], antioxidant [3], as an enhancer of NO production in endothelial cells [12], cardio protection [13], as a reducer of the invasion of breast cancer cells [14], etc.

This comprehensive area of well-known pharmacological properties of resveratrol exists due to a large research activity directed towards understanding the pharmacokinetics, the pharmacodynamics, and the metabolism of this compound. Resveratrol was at the center of scientific world's attention in 2003 when a study published in the journal *Nature* showed that this molecule can increase the lifespan of *Saccharomyces cerevisiae* [16]. Later, in 2006, Professor Sinclair's group found also that resveratrol can prolong the lifespan of obese mice, although it didn't have any effect on normal mice [17]. Clinical trials were launched, first based on resveratrol, than on some of its derivatives, examining the clinical efficiency of these compounds in managing diseases like diabetes, obesity, Alzheimer's disease, or even cancer [18, 19]. This year it has been confirmed that resveratrol has anti-aging effects, being demonstrated that it activates the protein SIRT 1 through an allosteric mechanism [20]. Some clinical trials are currently in phase II, showing promise for the use of stilbenes in the prevention of cardiovascular disease and for delaying aging effects. On the other hand, in a recent study intended to evaluate the safety and pharmacokinetics of resveratrol administered to humans in a single dose to healthy patients [21], it has been demonstrated that this compound is rapidly metabolized and it has a low bioavailability. Resveratrol is rapidly adsorbed after oral administration, with maximum levels in the human body being reached in approximately 30–60 min [22]. New stilbenoids have recently been identified [23, 24],

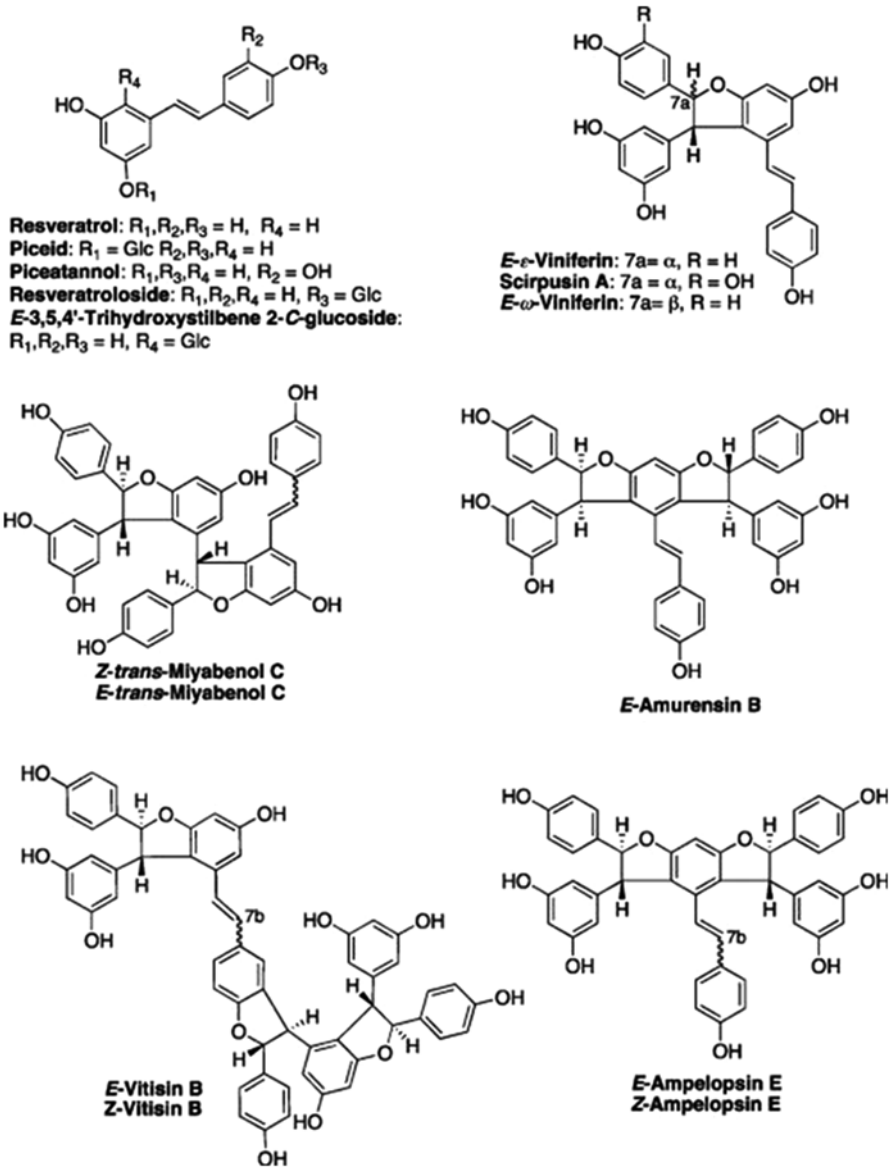


Fig. 17.1 Structures of compounds isolated from *V. amurensis* or *M. rotundifolia*. Reprinted with permission from [15]. Copyright (2013) American Chemical Society

hence the rising interest in developing new methods and strategies for quantification of stilbenoid compounds. Also, new challenges posed by the necessity to sometimes analyze very small volumes of samples gave rise to different analytical approaches and more complex techniques and instrumentation.

Grapes and wines can be considered a natural source of nutraceutical compounds, including stilbenes. The content of stilbenes in these natural matrices is

influenced by a set of factors such as the wine making process, variety, climate, etc. Different strategies are currently taken into consideration in the wine-making process for increasing not only the content of stilbenes in grapes, but also the efficiency of their transfer to the obtained wines. For this purpose, a rapid screening of complex samples for the desired stilbenes is necessary and the analytical information has to be available in a timely manner. In this work we report on the application of mass spectrometry (MS), one of the most used analytical methods for the detection of biomedically important stilbenes from wines.

17.2 Grape and Wine Stilbenes

Increased efforts have been directed in the biomedical field towards developing new strategies to prevent, treat, or identify the causes of major diseases. The detection and quantification of biochemical compounds found in plants, which may act against various types of diseases, is therefore very important. A tremendous amount of work has been done over the years to evaluate the chemical composition of grapes and wines. According to the literature, more than 1,000 components have been identified in wine [25].

The antioxidant properties of grapes and wine are mainly due to the fact that they contain a large amount of polyphenols with the ability to scavenge reactive oxygen species. Resveratrol and its derivatives can be found in different parts of the plant such as grape canes [26, 27] or grape skin [28] and the most grape product—wine [5, 15, 23]. The evolution and synthesis of stilbenes in plant organisms depends on the pathogenic infection [29]. One of the most common fungal infections which may affect the grapes and all the plant organs during growth is *Botrytis cinerea*. The pathways of stilbene formation in plants as response to fungal infections are not completely understood [4]. To gain more knowledge, *Vitis vinifera* cells were inoculated with methyl-β-cyclodextrin and methyl jasmonate, which lead to an increase in the concentration of resveratrol [30]. Both isomers of resveratrol, *cis*- and *trans*-, were found inside and outside of the cells, in contrast to another stilbene, piceid, which was found only inside the cells. Using the same elicitor (methyl jasmonate) in grape culture cells leads to the formation of *trans*-ε-viniferin, *trans*-δ-viniferin, *trans*-3-methylviniferin, and *trans*-piceatannol [31].

The amount of stilbenes formed as response to fungal infections depends on the grape variety [32]. Postharvest irradiation with UV light can also stimulate the synthesis of stilbenoids in grapes [24, 33, 34]. Up to 15.25 mg/kg of total stilbenes (of which 10.89 mg/kg was resveratrol) were found in grapes subjected to UV irradiation, compared to control where only 0.61 mg/kg have been found. In addition to resveratrol, p-viniferin (1.06 mg/kg), α-viniferin (0.34 mg/kg), and piceatannol (2.95 mg/kg) have been quantified in the UV-irradiated grapes samples. Similar results have been observed with white grapes, although the amounts of stilbenes were lower [35]. Besides increasing the concentration of stilbenes in wines, postharvest exposure of grapes to ultraviolet C was also found to result in the formation of new compounds (e.g., isorhapontigenin [24]).

The concentration of resveratrol in wines is influenced not only by grape variety [27], geographical region, climate [26], the chemical treatments applied to the vine to prevent pathogenic infections, the occurrence of fungal infections in grapes, postharvest treatments, etc., but also by the vinification procedures [36, 37]. Evolution and stability of resveratrol and its derivatives during wine making and maturation is of great interest [36, 38]. It has been observed that, during wine maturation, the amount of both isomers of resveratrol decrease, from 0.37 mg/L of resveratrol found in the grape juice, to less than the detection limit in the final product, the same trend being also observed for piceid. The decrease of stilbenoids in wines exposed to UV irradiation [33] was proportionally much lower, from 6.52 mg/L resveratrol and 1.90 mg/L picetannol initially to 4.12 mg/L and 1.05 mg/L, respectively, in bottled wines.

Wines from all over the world have been shown to contain stilbenes. Red grapes have a higher content in resveratrol than white and rose grapes [36, 39]. A review published in 2007 [40] has shown that Canada produces wines with some of the highest resveratrol contents (3.2 ± 1.5 mg/L), calculated as mean between lowest and highest concentration in cited articles. In the same study the highest level of resveratrol, 3.6 ± 2.9 mg/L was found in Pinot Noir wines from France, while wines of Agiorgitiko variety grown in Greece displayed the lowest contents (0.6 ± 0.2 mg/L). In Italy, a total of 1.31 mg/L of *trans*-resveratrol have been found in Nero d'Avola, the main red wine variety from grapes grown in Sicily [41]. North African red wines also contain stilbenoids. A quantity of 1.20 mg/L of *trans*-viniferin has been found in Merlot wine, 0.69 mg/L in Cabernet Sauvignon from Algeria [42]. In Brazil, 2.27 mg/L of *cis*-resveratrol were determined in Cabernet Sauvignon; however, the highest level was observed in a local variety, Tannat, which contained 5.49 mg/L. A content of 7.81 mg/L *trans*-resveratrol was found in an Australian Pinot Noir [43]. Stilbene levels in wines from the Idaho Valley region in the US, expressed as *trans-resveratrol*, ranged from 0.97 mg/L (Riesling) to 12.88 mg/L (Cabernet Sauvignon) [44].

17.3 Mass Spectrometry Analysis of Grapes and Wines Samples

Identification and detection of stilbenes are typically performed directly by electrochemistry [45], nuclear magnetic resonance (NMR), mass-spectrometry (MS [46]), and capillary electrophoresis [47] or, most of the time, by coupling chromatographic separation with sensitive detection procedures. High-performance liquid chromatography (HPLC) was coupled with DAD, fluorescence, electrochemical, or mass spectrometry detectors [48–50]. The coupling of HPLC or GC with MS allows assessing in detail complex matrices with high sensitivity, being successfully used for both analytical applications and basic research [51]. The mass fragmentation patterns acquired by mass spectrometry can help to identify a wide range of compounds, either by comparison with an external standard solution or based on mass spectral data available in MS libraries.

The mass spectrometer is more than just another analytical instrument. Currently, it is a very important analytical tool in many fields such as chemistry, biochemistry,

and medicine and the complexity of this instrument has increased tremendously. A significant number of ionization methods and types of mass analyzers have been developed and used in different ways for multiple applications [52–54]. Intensive research is being conducted in the field of biomedicine and the interest in discovering new compounds with improved bioavailability that can be used to fight major diseases continues to rise.

Mass spectrometry has been shown to be invaluable in the analysis of stilbenes from grapes, grape cell cultures, and wines, as well as for food supplements. For example, although more than 400 stilbenes from various plants are known in the present [55], other 23 new stilbenes were only discovered in 2013 with the help of MS [23]. Some cosmetics and numerous dietary supplements that contain stilbenoids and especially resveratrol can be found today on the market under many presentations: crèmes, capsules, beverages, chocolate bars, etc. (e.g., Resveratrol WINETIME bar™) [8].

The metabolic pathways and the rate of adsorption of active stilbenes from such products are not well investigated. Pharmacokinetics of stilbenes and the effect of various compounds from this class on various cell lines are two important research domains where the use of MS has a critical importance [56].

A variety of approaches have been used with mass spectrometry for discovering and quantifying stilbenoids in different types of matrices including: HPLC-ESI-MS for the detection of two important stilbenes, *trans*-resveratrol and *trans*-piceid in chocolate [57], the identification of a new resveratrol hexoside by reversed phase-HPLC coupled with atmospheric-pressure chemical ionization (APCI) tandem mass spectrometry (MS/MS) in cocoa liquor [58], and the use of high resolution electrospray ionization mass spectrometry (HR-ESI-MS) in defining the chemical structure of novel stilbenoids from *Polygonum cuspidatum* [59] or from *Rumex bucephalophorus* roots [60] (compounds with a very important biological activity against α-glucosidase, important in controlling hyperglycemia). Mass spectrometry was either used in conjunction with separation techniques—typically GC or HPLC—or was directly applied to the samples (direct infusion ESI-MS).

17.3.1 HPLC-MS for the Analysis of Stilbenes

HPLC is the most commonly used separation technique coupled with mass spectrometry for the quantification and detection of stilbenoids. This coupling combines the strength of chromatographic separation with the specificity and resolution achieved with mass spectrometry. Ions produced from the various sample components are identified and quantitated by different approaches. Various types of ionization methods (electrospray ionization ESI [61, 62]; atmospheric pressure chemical ionization APCI [63]; atmospheric pressure photoionisation (APPI–MSn) [64], and mass analyzers (TOF [65], QTOF, triple quadrupole (QQQ) [66], ion trap [43])) were used for the analysis of stilbenes. The MS data was acquired either in the positive or negative mode and compared with authentic markers for identification. Some

reports claimed that the negative ionization mode offered more information related to the chemical structure of the compounds that could be used for confirming peak identity, as compared to the positive ionization mode [43]. Most authors preferred selected ion monitoring (SIM) over multiple ion monitoring as MS analyzing mode, in order to reach the best sensitivity and reproducibility [67].

Several examples of the practical application of mass spectrometry are detailed further below, to underline the power of this technique, pertaining in particular to the identification of stilbenoids in complex matrices or to explain the pharmacokinetics of these compounds.

In vitro adsorption of resveratrol and its derivatives from the dietary source of roasted and boiled peanuts has been studied in details using HPLC-UV and HPLC coupled with a quadrupole MS [56]. The assay has been done on a human adenocarcinoma cell line (Caco-2). The separation of the compounds was made by reversed-phase LC on a classical C18 column using gradient elution and the spectra acquisition in the range of 200–600 nm. The LC eluent was transferred in the MS interface without stream splitting. This method uses an APCI source (positive ionization) and the ion abundance was acquired in the range of 50–600 amu. MS was used to distinguish the aglycone and glycosidic forms of resveratrol after gastro-intestinal digestion. The MS spectra analysis revealed indeed the presence of resveratrol diglycosides in peanuts. Further the transepithelial transport of resveratrol has been investigated and the results of this study have shown that the hydrolytic products of resveratrol glycosides are transported at a higher rate than the glycosidic forms. They can be found in a higher amount in roasted peanuts in comparison with the boiled sample. The same type of ionization source but operating in the negative mode has been successfully used also for stilbene detection in wines with great sensitivity [63].

A study aiming at the characterization of some newly discovered stilbenes from downy mildew-infected grapevine leaves found that APPI lead to cleaner MS spectra and allowed to determine resveratrol oligomers with higher sensitivity compared to ESI ionization. MS^n spectra obtained were used to propose chemical structures for the unknown stilbenes [64].

The relationship between stilbene composition of grape skins or stems and that of the corresponding wine has been studied by HPLC-ESI-MS [61]. The MS spectra were recorded in both positive and negative modes. The MS data allowed to conclude that although both grape stems and skin contain high amounts of stilbenes, their transfer rates to wine are very low: only 4 % of resveratrol and 14 % of piceid in stem-contact wine were contributed by stem tissue while the transfer rate for grape skin is lower that 11 %. The HPLC-ESI-MS technique has been successfully used to determine compounds like *trans*-resveratrol and δ-viniferin [32], piceid metabolites in rats [68], isomers of resveratrol dimer, and their analogues [69] or for the analysis of polyphenols (including resveratrol) in order to classify wines according to their geographic origin, grape variety, and vintage [65].

Applications based on mass spectrometry detection kept pace with modern chromatographic separation or with improved extraction procedures, for a rapid and sensitive analysis. Following this trend, 41 stilbenes (from which 23 new ones) were determined in the same run in a 2013 study by researchers at the University of

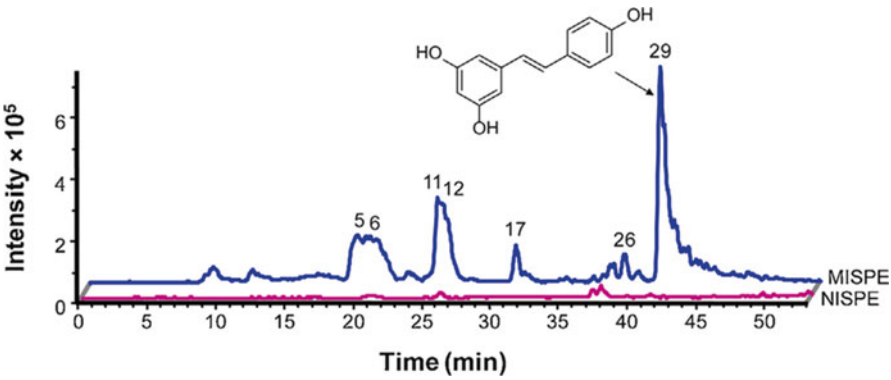


Fig. 17.2 Base peak chromatograms (BPC) of a Pinot noir red wine sample after treatment using either an (E)-resveratrol-templated MISPE (back chromatogram) cartridge or the corresponding NISPE (front chromatogram) cartridge. The samples were analyzed in the *m/z* range from 100 to 1,200, using LC–ESI-MS/MS in the negative ionization mode. MISPE: molecularly imprinted polymer solid-phase extraction. NISPE: non-imprinted polymer solid-phase extraction. Reprinted from [43]. Copyright (2013), with permission from Elsevier

British Columbia, Canada. The new discoveries were possible by coupling the fast separation by UHPLC with ESI-Q-TOF MS detection [23].

In another example, resveratrol and several secondary metabolites were identified by RP-HPLC and μLC–ESI ion trap MS/MS following selective solid-phase extraction (SPE) with reusable molecularly imprinted polymers (MIPs, Fig. **17.2**). The coupling between selective pre-concentration using MIPs and the sensitivity of MS detection allowed to achieve a detection limit of 8.87×10^{-3} mg/L for *trans-resveratrol* [43].

The application of mass spectrometry in mechanistic studies of resveratrol against radical oxygen species can give insight into the pathways involved. Based on HPLC-ESI-MS data [70], a mechanism has been proposed for the interaction of resveratrol with $^1O_2$ (Fig. 17.3). The authors have shown that an endoperoxide intermediate is formed, followed by a hydrolyzed intermediate and quinine as final product. This suggests that resveratrol can be useful as a drug for treating $^1O_2$-mediated diseases.

17.3.2 GC-MS for the Analysis of Stilbenes

Stilbenoids have been analyzed also by coupling gas chromatography with mass spectrometry (GC-MS). A chemical derivatization step has to precede analysis by gas chromatography in order to transform the nonvolatile stilbenes into volatile, thermostable compounds. Various extraction methods for stilbenes have been developed to facilitate their sensitive detection by GC-MS such as: SPE [71] or solid-phase microextraction (SPME) [71], dispersive liquid–liquid microextraction (DLLME) [5], stir bar sorbtive extraction [72, 73], directly suspended droplet microextraction (DSDME) [74], etc.

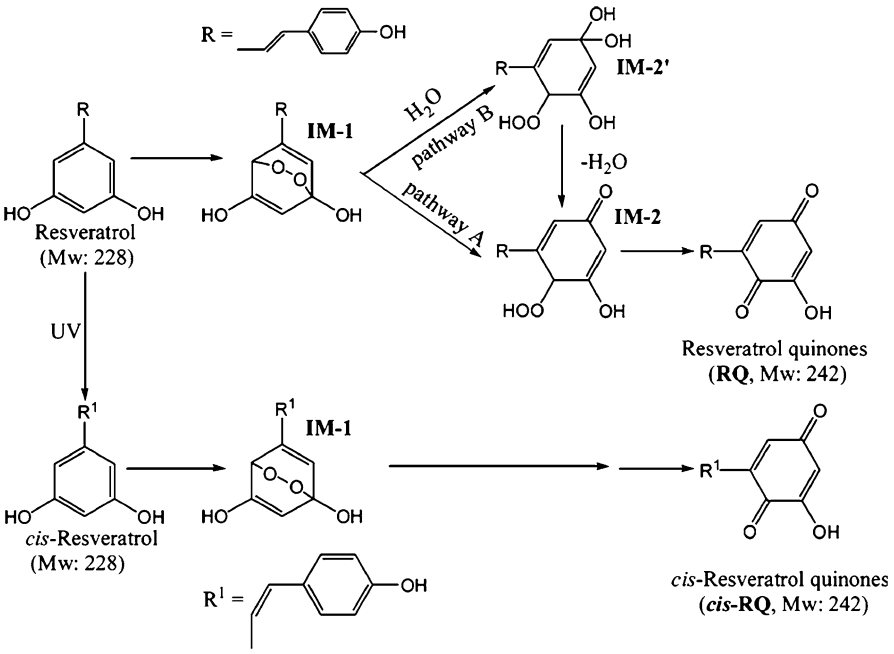


Fig. 17.3 Proposed mechanism for resveratrol against $^1O2$. Reprinted with permission from [70]. Copyright (2010) American Chemical Society

Such an example is the detection of five polyphenols (*trans*- and *cis*-resveratrol, piceatannol, catechin, and epicatechin) in wine and grapes using a SPME-GC coupled to a quadrupole mass selective spectrometer equipped with an inert ion source that operates in electron-impact (EI) mode at 70 eV [75]. The derivatization reaction used in the study is silylation, the most commonly used for GC analysis of polyphenols. An alternative derivatization approach relies on acetylation [71].

A DLLME technique has been proposed for the first time in a 2012 study for the determination of three hydroxylated stilbenes (*trans*-pterostilbene, resveratrol, and piceatannol) in wine samples. By coupling this with GC-EI-MS analysis, it was possible to quantify the above stilbenes in the range from 0.6 and 5 ng/mL [5].

Resveratrol and its stilbenes analogues were suggested to act as chemopreventive agents in colon cancer. Studies have been conducted both in vitro on colon cancer cell lines and in vivo on immunodeficient mice to check the efficiency of 24 stilbenes against this disease. An important part in this work was the GC-MS analysis of the serum from mice treated with the investigated compounds [76].

A method based on stir bar sorptive extraction coupled to gas chromatography–mass spectrometry by means of a thermal desorption unit (SBSE-TD-GC–MS) has been optimized for the determination of *cis*/*trans* isomers of resveratrol, piceatannol, and oxyresveratrol in wines [73]. This study uses a GC coupled with a quadrupole mass selective spectrometer and an inert ion source. The compounds were quantified in the SIM mode in order to improve the sensitivity (Fig. 17.4).

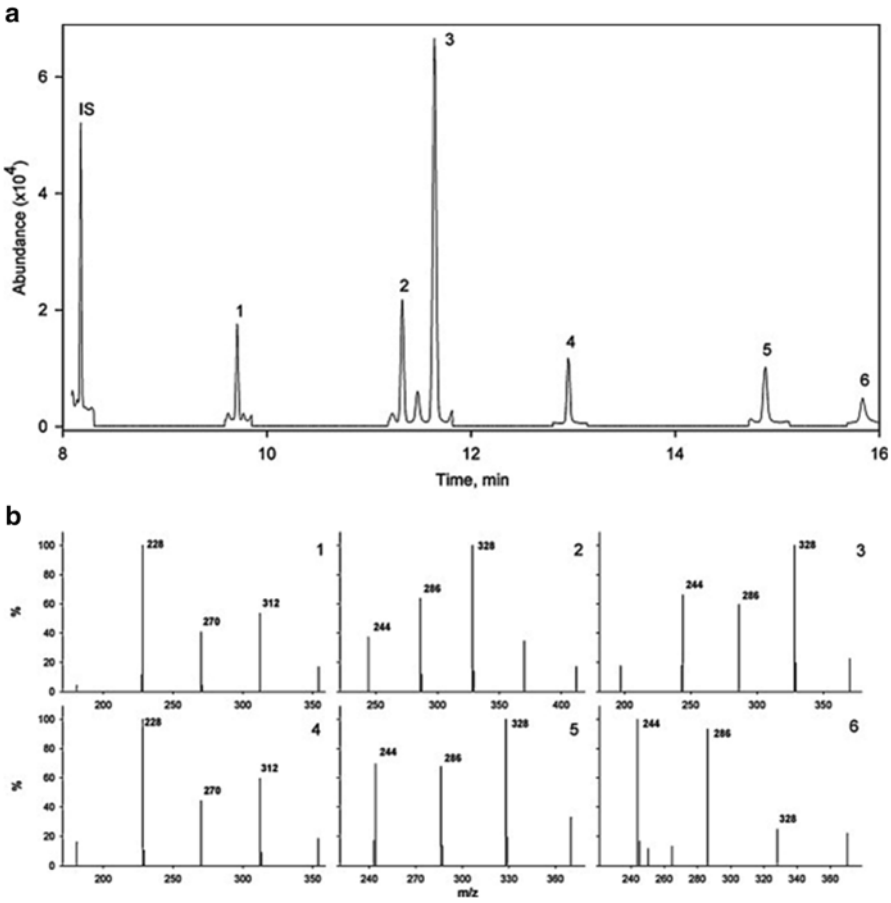


Fig. 17.4 (**a**) SBSE-TD-GC–MS chromatogram obtained for a spiked white wine fortified at 2 μg L^{-1} under SIM mode. Peaks correspond to: (1) *cis*-resveratrol, (2) *cis*-oxyresveratrol, (3) *cis-piceatannol*, (4) *trans*-resveratrol, (5) *trans*-oxyresveratrol, and (6) *trans*-piceatannol. (**b**) Mass spectra of each compound. Reprinted from [73] with permission from Elsevier

The major compound determined was *trans*-resveratrol, with concentrations in the range of 3–230 μg/L, depending on the type of wine.

In a few studies, although stilbenes were separated by HPLC, detection by mass spectrometry was done off-line. For example, a new resveratrol dimer (*cis-ε-viniferin*) was isolated and identified in an Algerian red wine using HPLC and MALDI-TOF MS [42]. Spectra were recorded in the positive-ion mode using the reflectron and with an accelerating voltage of 20 kV. One advantage of this technique was that it allowed the simultaneous detection of *trans*-ε-viniferin. Consequently, based on MS data it has been shown that both viniferin isomers exhibited marked cytotoxic activity breast human cancer cell lines and antimutagenic activity [65].

17.3.3 Mass Spectrometry Analysis of Stilbenes Without Separation

17.3.3.1 Direct Infusion ESI-MS

In direct infusion mass spectrometry, the samples are analyzed without prior extraction or separation. This eliminates the bias due to sample pretreatment and allows a fast screening of samples for the compounds of interest. Identification of these compounds is made by comparing their ESI-MS/MS fragmentation pattern with that of standard compounds from spectral libraries. A high number of substances from different chemical classes such as organic acids, inorganic acids, and phenolic compounds can thus be identified in the same run in a matter of minutes [46].

The MS spectra of commercial tannin and of several types of red wine (FN, PN, CS, NM), from the same vineyard and harvest (2012) collected by ESI-MS direct infusion in positive mode, are shown in Fig. 17.5 (*m*/*z* range 100–850). Also included is the profile acquired for an oenological tannin used in the same vineyard for tannin correction of wine in some vinification procedures. The ESI-MS and ESI-MS/MS were acquired using a QTOF Micro mass spectrometer in positive mode and a micro ESI source with the capillary voltage at 3,200 V, at a flow rate of 5 μl/min, according to published procedures [77–79].

As observed, the wines analyzed have common peaks either in all wines and tannin, such as the peak with *m*/*z* of 381.27, peaks common to a particular wine and tannin (i.e., peaks with *m*/*z* of 397.25 and 719.42, found in tannin and in FN wine or peak with *m*/*z* of 274.25 found in tannin and PN wine), peaks found in some wines, but not in others or in tannins (i.e., peak with *m*/*z* of 809.17 found in PN and NM and peak with *m*/*z* of 513.07, found in PN and CS; none of these peaks were observed in tannin), or peaks specific to one type of wine (i.e., peak with *m*/*z* of 459.12, specific to tannins or peak with *m*/*z* of 493.12, specific to CS wine). The peaks with the highest intensity that were either common or specific to tannin or to wines were selected for fragmentation and are currently investigated for identification. MS/MS fragmentation of the peaks observed in the MS in tannin and various wines produced a series of spectra shown in Figs. 17.6 and 17.7 (they are currently under investigation).

All these MS spectra of tannins and various wines were recorded under ESI positive ionization, under slightly acidic conditions (red wines are acidic, with a pH ~3.5). However, the full composition of a particular wine (and tannin) is usually not revealed under a particular set of experimental conditions. For example, analyzing the same wines under very acidic conditions (i.e., in acetonitrile (ACN) containing 0.1 % formic acid (FA)) can increase the number of molecules that one can identify. Furthermore, analyzing the same samples by ESI under negative ionization at neutral or alkaline conditions will surely lead to identification of additional peaks that correspond to molecules which are part of and perhaps specific to some particular wines. Examples of MS spectra of wines analyzed in ACN or in ACN containing 0.1 % FA are shown in Fig. 17.8 (FN wine) and Fig. 17.9 (PN wine). In FN wine

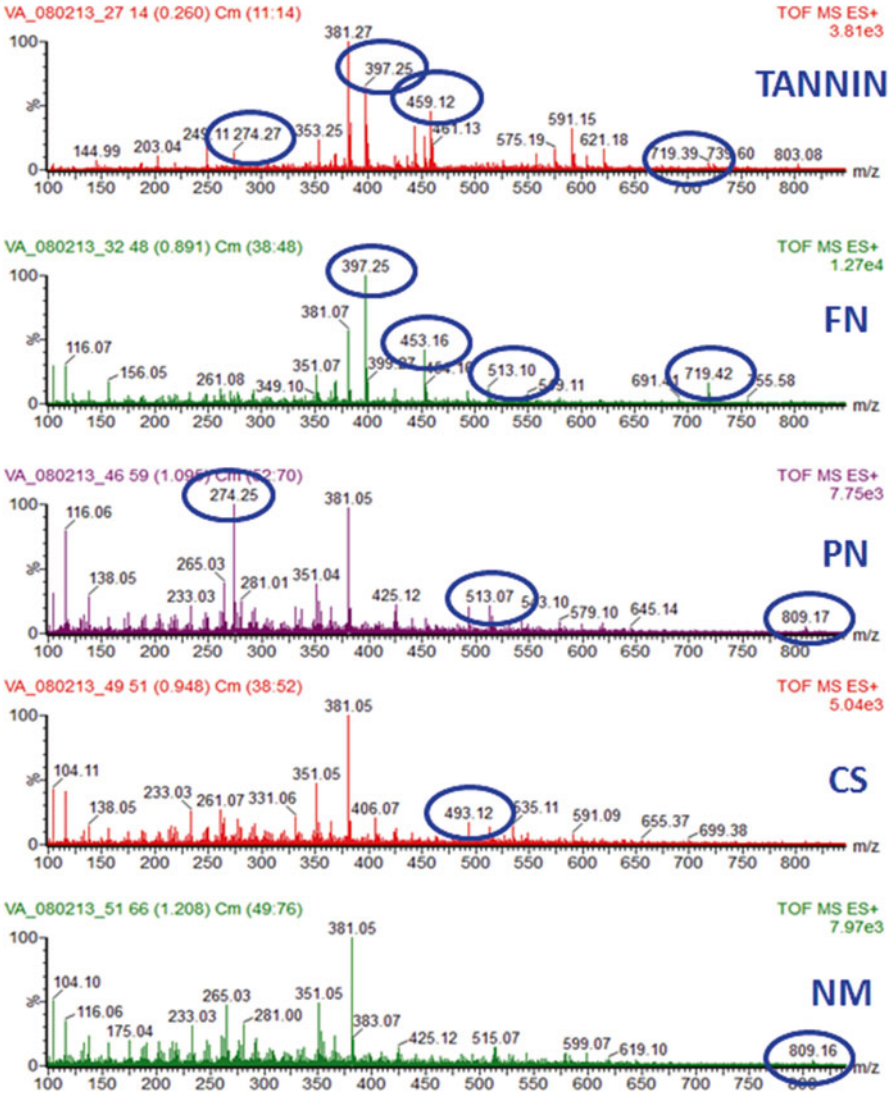


Fig. 17.5 ESI-MS analysis (direct infusion) of tannin and various wines (FN, PN, CS, NM). Shown are the peaks with the *m/z* ranging from 100 to 850. Circled are the peaks either common to all wines or specific to some wines and tannin or specific to one particular wine. These peaks are discussed in the text

(Fig. 17.8), the peaks with *m/z* of 397.24, or 719.42 are specific to FN analyzed in ACN with no FA, while the peaks with *m/z* of 331.02, 535.02, or 639.04 are specific to FN analyzed in ACN with 0.1 % FA. Common peaks were also observed (i.e., peaks with *m/z* of 453.16 or 493.02). In PN wine (Fig. 17.9), peaks specific to PN analyzed in ACN without FA (*m/z* of 116.06, 513.08 or 809.19) or peaks specific to

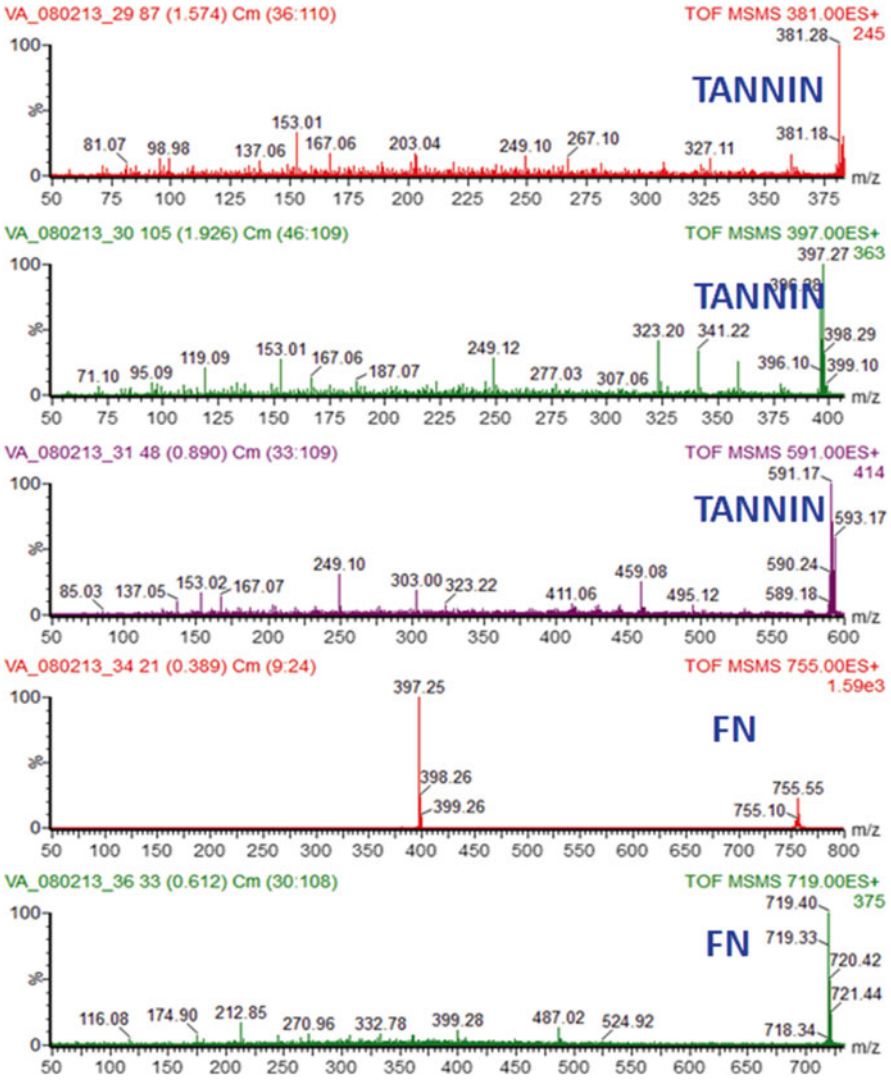


Fig. 17.6 ESI-MS/MS analysis (direct infusion) of precursor ions detected in ESI-MS in tannin (*m/z* of 381.28, 397.27, and 591.17) and FN wine (*m/z* of 755.55, 719.00). The collision energy was optimized for each precursor and varied from 5 to 50 V

PN analyzed in AC with 0.1 % FA (*m/z* of 365.08 or 599.04), as well as peaks common to both conditions (*m/z* of 381.05 or 493.12) were also observed.

The MS spectra are a snapshot capturing unique characteristics of each type of wine and can be used by applying chemometrics for data interpretation in order to differentiate wine varieties. The high throughput of direct infusion ESI-MS allowed, among others, to follow the evolution of samples of must and wines obtained from

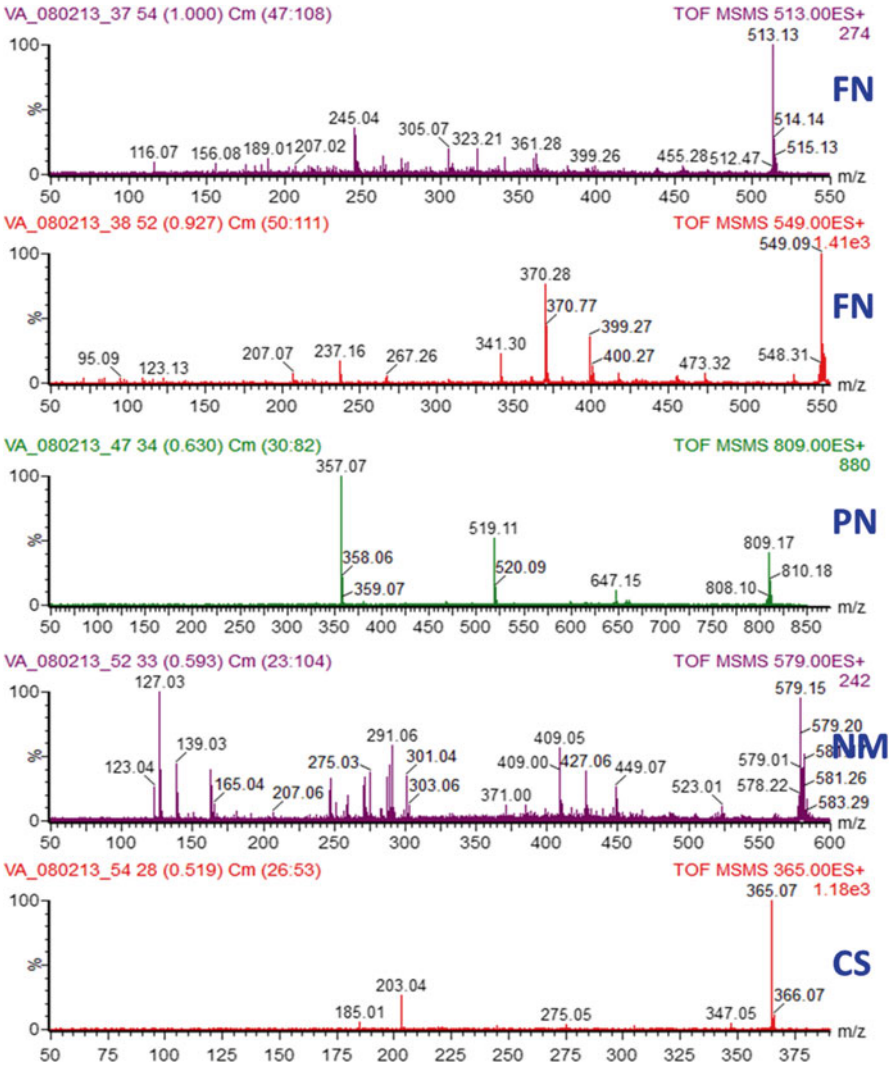


Fig. 17.7 ESI-MS/MS analysis (direct infusion) of precursor ions detected in ESI-MS in FN wine (*m/z* of 513.13 and 549.09), PN wine (*m/z* of 809.00), NM wine (*m/z* of 579.00), and CS wine (*m/z* of 365.00). The collision energy was optimized for each precursor and varied from 5 to 50 V

different grape varieties before and after fermentation by acquiring specific fingerprints and identifying marker ions for wine and must. Principal Component Analysis was used to group samples according to these markers [80, 81]. As some authors noted, ESI-MS in negative mode can provide more analytical information compared to the positive mode as it leads to lower amounts of salt adducts and to a higher number and variety of ions [82].

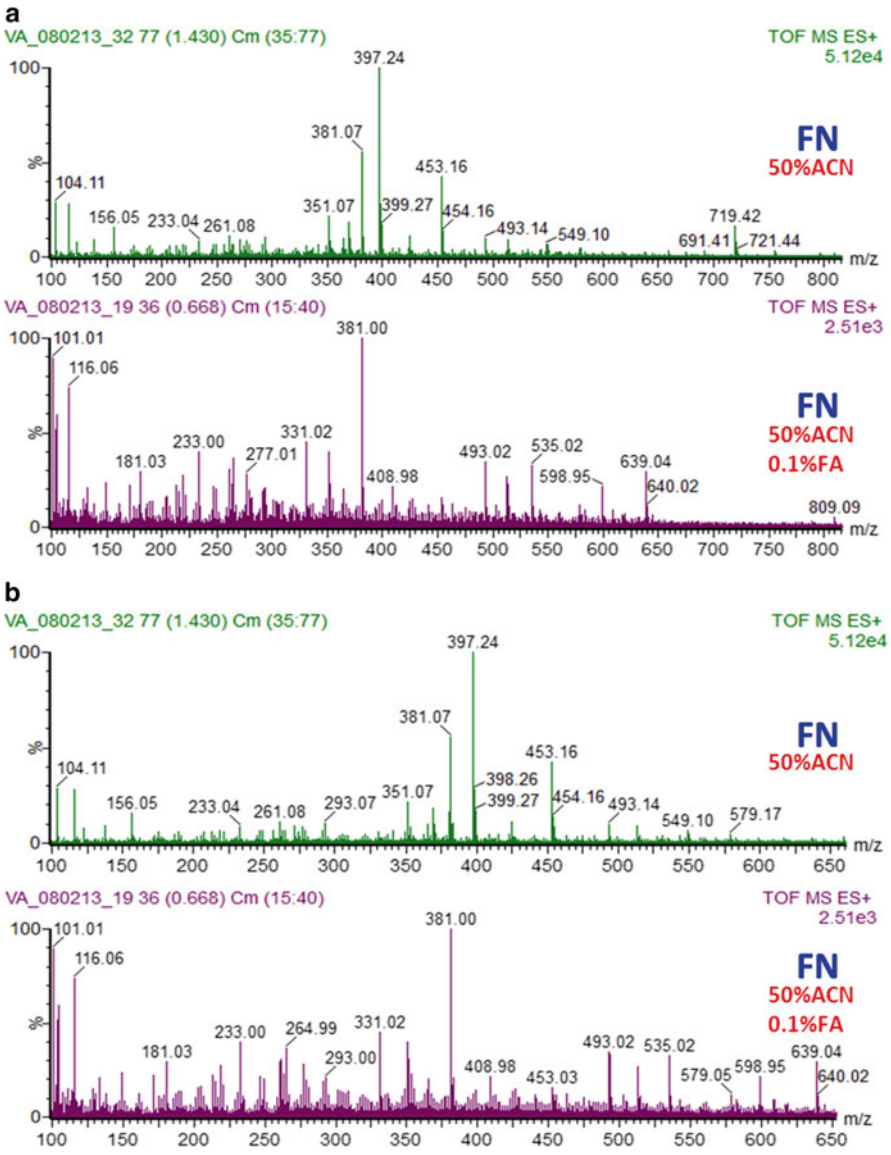


Fig. 17.8 ESI-MS analysis (direct infusion) of FN wine analyzed in ACN and ACN with 0.1 % FA. The *m/z* range is 100–850 (**a**) and 100–650 (**b**)

The increasing evolution of analytical instrumentation providing higher mass resolution, robustness, and accuracy has raised the ESI-MS technique from the rank of "fingerprinting" tool to one of metabolite profiling [83]. ESI-MS combined with single or hybrid quadrupole, time-of-flight, ion trap, Orbitrap, and FT-ICR-MS technology are time-consuming analytical approaches that offer the possibility to obtain

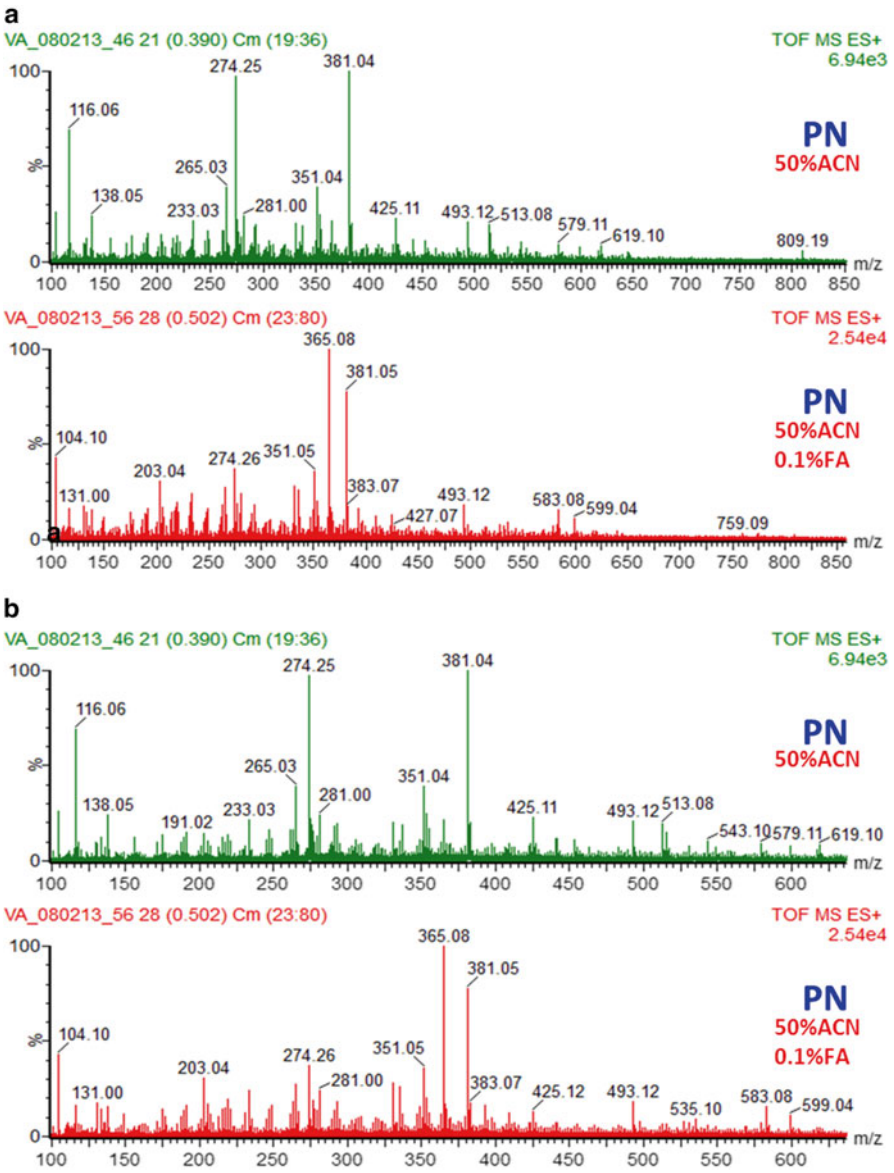


Fig. 17.9 ESI-MS analysis (direct infusion) of PN wine analyzed in ACN and ACN with 0.1 % FA. The *m/z* range is 100–850 (**a**) and 100–650 (**b**)

a complete metabolome analysis in complex samples [84]. For example, the ion cyclotron resonance-Fourier transform mass spectrometry (ICR-FT/MS) has been used to develop a nontargeted method for metabolite profiling in oenology [85] or to differentiate grapes and corresponding wines from distinct vineyards, according to complex chemical fingerprints [86].

More towards biomedical applications, direct infusion ESI coupled with an ion trap mass spectrometer has been used to study the interaction between e-viniferin glucoside (VG), a resveratrol-derived dimer, and amyloid β-peptides responsible for triggering neuronal degeneration [87]. Using this technique it was possible to observe a non-covalent complex between VG and Aβ. The formation of this complex leads to decreased cytotoxicity of Aβ against PC12 cells lines in vitro.

17.4 Conclusions

Resveratrol and related stilbenes represent a class of compounds with extraordinary potential for the biomedical field, due to their aging-delaying effects and their potential against cardiovascular diseases and cancer. New chemical compounds from this class, and even more importantly, their interactions in the human body remain yet to be discovered. Mass spectrometry proved to be of critical importance for the study of stilbenes that are important in the biomedical field. Its use is expected to increase in the future, from the simple screening of samples in search of those containing the desired stilbenes up to complex investigations aiming to unravel new chemical structures and understanding their interactions.

Acknowledgments VA and AV acknowledge financial support by a grant of the Romanian National Authority for Scientific Research, CNDI—UEFISCDI, project number PN-II-PT-PCCA-2011-3.1-1809.

References

1. Sen S, Chakraborty R (2011) The role of antioxidants in human health. In: Silvana A, Hepel M (eds) Oxidative stress: diagnostics, prevention, and therapy, vol 1083. American Chemical Society, Washington, DC, pp 1–37
2. Renaud S, de Lorgeril M (1992) Wine, alcohol, platelets, and the French paradox for coronary heart disease. Lancet 339(8808):1523–1526
3. Nopo-Olazabal C, Hubstenberger J, Nopo-Olazabal L, Medina-Bolivar F (2013) Antioxidant activity of selected stilbenoids and their bioproduction in hairy root cultures of muscadine grape (Vitis rotundifolia Michx.). J Agric Food Chem 61(48):11744–11758
4. Jeandet P, Delaunois B, Conreux A, Donnez D, Nuzzo V, Cordelier S, Clément C, Courot E (2010) Biosynthesis, metabolism, molecular engineering, and biological functions of stilbene phytoalexins in plants. Biofactors 36(5):331–341
5. Rodríguez-Cabo T, Rodríguez I, Cela R (2012) Determination of hydroxylated stilbenes in wine by dispersive liquid–liquid microextraction followed by gas chromatography mass spectrometry. J Chromatogr A 1258:21–29
6. Langcake P, Pryce RJ (1977) A new class of phytoalexins from grapevines. Experientia 33(2):151–152
7. Karuppagounder SS, Pinto JT, Xu H, Chen H-L, Beal MF, Gibson GE (2009) Dietary supplementation with resveratrol reduces plaque pathology in a transgenic model of Alzheimer's disease. Neurochem Int 54(2):111–118
8. Valenzano DR, Terzibasi E, Genade T, Cattaneo A, Domenici L, Cellerino A (2006) Resveratrol prolongs lifespan and retards the onset of age-related markers in a short-lived vertebrate. Curr Biol 16(3):296–300

9. Bishayee A, Politis T, Darvesh AS (2010) Resveratrol in the chemoprevention and treatment of hepatocellular carcinoma. Cancer Treat Rev 36(1):43–53
10. González-Sarrías A, Gromek S, Niesen D, Seeram NP, Henry GE (2011) Resveratrol oligomers isolated from carex species inhibit growth of human colon tumorigenic cells mediated by cell cycle arrest. J Agric Food Chem 59(16):8632–8638
11. Wang K-T, Chen L-G, Tseng S-H, Huang J-S, Hsieh M-S, Wang C-C (2011) Anti-inflammatory effects of resveratrol and oligostilbenes from Vitis thunbergii var. taiwaniana against lipopolysaccharide-induced arthritis. J Agric Food Chem 59(8):3649–3656
12. Appeldoorn MM, Venema DP, Peters THF, Koenen ME, Arts ICW, Vincken J-P, Gruppen H, Keijer J, Hollman PCH (2009) Some phenolic compounds increase the nitric oxide level in endothelial cells in vitro. J Agric Food Chem 57(17):7693–7699
13. Tomé-Carneiro J, Larrosa M, González-Sarrías A, Tomás-Barberán F, García-Conesa M, Espín J (2013) Resveratrol and clinical trials: the crossroad from in vitro studies to human evidence. Curr Pharm Des 19(34):6064–6093
14. Ko HS, Lee H-J, Kim S-H, Lee E-O (2012) Piceatannol suppresses breast cancer cell invasion through the inhibition of MMP-9: involvement of PI3K/AKT and NF-κB pathways. J Agric Food Chem 60(16):4083–4089
15. Pawlus AD, Sahli R, Bisson J, Rivière C, Delaunay J-C, Richard T, Gomès E, Bordenave L, Waffo-Téguo P, Mérillon J-M (2013) Stilbenoid profiles of canes from Vitis and Muscadinia species. J Agric Food Chem 61(3):501–511
16. Howitz KT, Bitterman KJ, Cohen HY, Lamming DW, Lavu S, Wood JG, Zipkin RE, Chung P, Kisielewski A, Zhang L-L, Scherer B, Sinclair DA (2003) Small molecule activators of sirtuins extend Saccharomyces cerevisiae lifespan. Nature 425(6954):191–196
17. Baur JA, Pearson KJ, Price NL, Jamieson HA, Lerin C, Kalra A, Prabhu VV, Allard JS, Lopez-Lluch G, Lewis K, Pistell PJ, Poosala S, Becker KG, Boss O, Gwinn D, Wang M, Ramaswamy S, Fishbein KW, Spencer RG, Lakatta EG, Le Couteur D, Shaw RJ, Navas P, Puigserver P, Ingram DK, de Cabo R, Sinclair DA (2006) Resveratrol improves health and survival of mice on a high-calorie diet. Nature 444(7117):337–342
18. Udenigwe CC, Ramprasath VR, Aluko RE, Jones PJ (2008) Potential of resveratrol in anticancer and anti-inflammatory therapy. Nutr Rev 66(8):445–454
19. Kumar A, Kaundal RK, Iyer S, Sharma SS (2007) Effects of resveratrol on nerve functions, oxidative stress and DNA fragmentation in experimental diabetic neuropathy. Life Sci 80(13):1236–1244
20. Hubbard BP, Gomes AP, Dai H, Li J, Case AW, Considine T, Riera TV, Lee JE, Sook Yen E, Lamming DW, Pentelute BL, Schuman ER, Stevens LA, Ling AJY, Armour SM, Michan S, Zhao H, Jiang Y, Sweitzer SM, Blum CA, Disch JS, Ng PY, Howitz KT, Rolo AP, Hamuro Y, Moss J, Perni RB, Ellis JL, Vlasuk GP, Sinclair DA (2013) Evidence for a common mechanism of SIRT1 regulation by allosteric activators. Science 339(6124):1216–1219
21. Walle T, Hsieh F, DeLegge MH, Oatis JE Jr, Walle UK (2004) High absorption but very low bioavailability of oral resveratrol in humans. Drug Metab Dispos 32(12):1377–1382
22. Bishayee A (2009) Cancer prevention and treatment with resveratrol: from rodent studies to clinical trials. Cancer Prev Res (Phila) 2(5):409–418
23. Moss R, Mao Q, Taylor D, Saucier C (2013) Investigation of monomeric and oligomeric wine stilbenoids in red wines by ultra-high-performance liquid chromatography/electrospray ionization quadrupole time-of-flight mass spectrometry. Rapid Commun Mass Spectrom 27(16):1815–1827
24. Fernández-Marín MI, Guerrero RF, García-Parrilla MC, Puertas B, Richard T, Rodriguez-Werner MA, Winterhalter P, Monti J-P, Cantos-Villar E (2012) Isorhapontigenin: a novel bioactive stilbene from wine grapes. Food Chem 135(3):1353–1359
25. Gómez-Míguez MJ, Cacho JF, Ferreira V, Vicario IM, Heredia FJ (2007) Volatile components of Zalema white wines. Food Chem 100(4):1464–1473
26. Vergara C, von Baer D, Mardones C, Wilkens A, Wernekinck K, Damm A, Macke S, Gorena T, Winterhalter P (2012) Stilbene levels in grape cane of different cultivars in southern Chile: determination by HPLC-DAD-MS/MS method. J Agric Food Chem 60(4):929–933

27. Lambert C, Richard T, Renouf E, Bisson J, Waffo-Teguo P, Bordenave L, Ollat N, Merillon J-M, Cluzet S (2013) Comparative analyses of stilbenoids in canes of major Vitis vinifera L. cultivars. J Agric Food Chem 61(47):11392–11399
28. Lago-Vanzela ES, Da-Silva R, Gomes E, García-Romero E, Hermosín-Gutiérrez I (2011) Phenolic composition of the edible parts (flesh and skin) of Bordô grape (Vitis labrusca) using HPLC-DAD-ESI-MS/MS. J Agric Food Chem 59(24):13136–13146
29. Mattivi F, Vrhovsek U, Malacarne G, Masuero D, Zulini L, Stefanini M, Moser C, Velasco R, Guella G (2011) Profiling of resveratrol oligomers, important stress metabolites, accumulating in the leaves of hybrid Vitis vinifera (Merzling×Teroldego) genotypes infected with Plasmopara viticola. J Agric Food Chem 59(10):5364–5375
30. Martinez-Esteso MJ, Sellés-Marchart S, Vera-Urbina JC, Pedreño MA, Bru-Martinez R (2011) DIGE analysis of proteome changes accompanying large resveratrol production by grapevine (Vitis vinifera cv. Gamay) cell cultures in response to methyl-β-cyclodextrin and methyl jasmonate elicitors. J Proteomics 74(8):1421–1436
31. Donnez D, Kim K-H, Antoine S, Conreux A, De Luca V, Jeandet P, Clément C, Courot E (2011) Bioproduction of resveratrol and viniferins by an elicited grapevine cell culture in a 2 L stirred bioreactor. Process Biochem 46(5):1056–1062
32. Timperio AM, D'Alessandro A, Fagioni M, Magro P, Zolla L (2012) Production of the phytoalexins trans-resveratrol and delta-viniferin in two economy-relevant grape cultivars upon infection with Botrytis cinerea in field conditions. Plant Physiol Biochem 50:65–71
33. Guerrero RF, Puertas B, Jiménez MJ, Cacho J, Cantos-Villar E (2010) Monitoring the process to obtain red wine enriched in resveratrol and piceatannol without quality loss. Food Chem 122(1):195–202
34. Fernández-Marín MI, Guerrero RF, García-Parrilla MC, Puertas B, Ramírez P, Cantos-Villar E (2013) Terroir and variety: two key factors for obtaining stilbene-enriched grapes. J Food Compos Anal 31(2):191–198
35. Guerrero RF, Puertas B, Fernández MI, Piñeiro Z, Cantos-Villar E (2010) UVC-treated skin-contact effect on both white wine quality and resveratrol content. Food Res Int 43(8): 2179–2185
36. Roldán A, Palacios V, Caro I, Pérez L (2010) Evolution of resveratrol and piceid contents during the industrial winemaking process of sherry wine. J Agric Food Chem 58(7):4268–4273
37. Atanacković M, Petrović A, Jović S, Bukarica LG, Bursać M, Cvejić J (2012) Influence of winemaking techniques on the resveratrol content, total phenolic content and antioxidant potential of red wines. Food Chem 131(2):513–518
38. Sawaya ACHF, Catharino RR, Facco EMP, Fogaça A, Godoy HT, Daudt CE, Eberlin MN (2011) Monitoring of wine aging process by electrospray ionization mass spectrometry. Food Sci Tech (Campinas) 31(3):730–734
39. Gatto P, Vrhovsek U, Muth J, Segala C, Romualdi C, Fontana P, Pruefer D, Stefanini M, Moser C, Mattivi F, Velasco R (2008) Ripening and genotype control stilbene accumulation in healthy grapes. J Agric Food Chem 56(24):11773–11785
40. Stervbo U, Vang O, Bonnesen C (2007) A review of the content of the putative chemopreventive phytoalexin resveratrol in red wine. Food Chem 101(2):449–457
41. Todaro A, Palmeri R, Barbagallo RN, Pifferi PG, Spagna G (2008) Increase of trans-resveratrol in typical Sicilian wine using β-Glucosidase from various sources. Food Chem 107(4):1570–1575
42. Amira-Guebailia H, Valls J, Richard T, Vitrac X, Monti J-P, Delaunay J-C, Mérillon J-M (2009) Centrifugal partition chromatography followed by HPLC for the isolation of cis-ε-viniferin, a resveratrol dimer newly extracted from a red Algerian wine. Food Chem 113(1):320–324
43. Hashim SNNS, Schwarz LJ, Boysen RI, Yang Y, Danylec B, Hearn MTW (2013) Rapid solid-phase extraction and analysis of resveratrol and other polyphenols in red wine. J Chromatogr A 1313:284–290
44. Lee J, Rennaker C (2007) Antioxidant capacity and stilbene contents of wines produced in the Snake River Valley of Idaho. Food Chem 105(1):195–203

45. Corduneanu O, Janeiro P, Brett AMO (2006) On the electrochemical oxidation of resveratrol. Electroanalysis 18(8):757–762
46. Biasoto ACT, Catharino RR, Sanvido GB, Eberlin MN, da Silva MAAP (2010) Flavour characterization of red wines by descriptive analysis and ESI mass spectrometry. Food Qual Prefer 21(7):755–762
47. Chu Q, O'Dwye M, Zeece MG (1998) Direct analysis of resveratrol in wine by micellar electrokinetic capillary electrophoresis. J Agric Food Chem 46(2):509–513
48. Boutegrabet L, Fekete A, Hertkorn N, Papastamoulis Y, Waffo-Téguo P, Mérillon JM, Jeandet P, Gougeon RD, Schmitt-Kopplin P (2011) Determination of stilbene derivatives in Burgundy red wines by ultra-high-pressure liquid chromatography. Anal Bioanal Chem 401(5):1513–1521
49. Jiménez Sánchez JB, Crespo Corral E, Santos Delgado MJ, Orea JM, Ureña AG (2005) Analysis of trans-resveratrol by laser ionization mass spectrometry and HPLC with fluorescence detection. J Chromatogr A 1074(1–2):133–138
50. Beňová B, Adam M, Onderková K, Královský J, Krajíček M (2008) Analysis of selected stilbenes in Polygonum cuspidatum by HPLC coupled with CoulArray detection. J Sep Sci 31(13):2404–2409
51. Gross J (2011) Electrospray ionization. Mass spectrometry. Springer, Berlin, pp 561–620. ISBN 978-3-642-10709-2
52. Mirsaleh-Kohan N, Robertson WD, Compton RN (2008) Electron ionization time-of-flight mass spectrometry: historical review and current applications. Mass Spectrom Rev 27(3):237–285
53. Monge ME, Harris GA, Dwivedi P, Fernández FM (2013) Mass spectrometry: recent advances in direct open air surface sampling/ionization. Chem Rev 113(4):2269–2308
54. Wasinger VC, Zeng M, Yau Y (2013) Current status and advances in quantitative proteomic mass spectrometry. Int J Proteomics 2013:180605
55. Shen T, Wang X-N, Lou H-X (2009) Natural stilbenes: an overview. Nat Prod Rep 26(7):916–935
56. Chukwumah Y, Walker L, Vogler B, Verghese M (2011) In vitro absorption of dietary trans-resveratrol from boiled and roasted peanuts in Caco-2 cells. J Agric Food Chem 59(23):12323–12329
57. Counet C, Callemien D, Collin S (2006) Chocolate and cocoa: new sources of trans-resveratrol and trans-piceid. Food Chem 98(4):649–657
58. Jerkovic V, Bröhan M, Monnart E, Nguyen F, Nizet S, Collin S (2010) Stilbenic profile of cocoa liquors from different origins determined by RP-HPLC-APCI(+)-MS/MS. Detection of a new resveratrol hexoside. J Agric Food Chem 58(11):7067–7074
59. Li F, Zhan Z, Liu F, Yang Y, Li L, Feng Z, Jiang J, Zhang P (2013) Polyflavanostilbene A, a new flavanol-fused stilbene glycoside from Polygonum cuspidatum. Org Lett 15(3):674–677
60. Kerem Z, Bilkis I, Flaishman MA, Sivan L (2006) Antioxidant activity and inhibition of α-glucosidase by trans-resveratrol, piceid, and a novel trans-stilbene from the roots of Israeli Rumex bucephalophorus L. J Agric Food Chem 54(4):1243–1247
61. Sun B, Ribes AM, Leandro MC, Belchior AP, Spranger MI (2006) Stilbenes: quantitative extraction from grape skins, contribution of grape solids to wine and variation during wine maturation. Anal Chim Acta 563(1–2):382–390
62. Ivanova V, Dörnyei Á, Márk L, Vojnoski B, Stafilov T, Stefova M, Kilár F (2011) Polyphenolic content of Vranec wines produced by different vinification conditions. Food Chem 124(1):316–325
63. Bravo MN, Silva S, Coelho AV, Boas LV, Bronze MR (2006) Analysis of phenolic compounds in Muscatel wines produced in Portugal. Anal Chim Acta 563(1–2):84–92
64. Jean-Denis JB, Pezet R, Tabacchi R (2006) Rapid analysis of stilbenes and derivatives from downy mildew-infected grapevine leaves by liquid chromatography–atmospheric pressure photoionisation mass spectrometry. J Chromatogr A 1112(1–2):263–268
65. Jaitz L, Siegl K, Eder R, Rak G, Abranko L, Koellensperger G, Hann S (2010) LC–MS/MS analysis of phenols for classification of red wine according to geographic origin, grape variety and vintage. Food Chem 122(1):366–372

66. Di Lecce G, Arranz S, Jáuregui O, Tresserra-Rimbau A, Quifer-Rada P, Lamuela-Raventós RM (2014) Phenolic profiling of the skin, pulp and seeds of Albariño grapes using hybrid quadrupole time-of-flight and triple-quadrupole mass spectrometry. Food Chem 145:874–882
67. Rodriguez-Naranjo MI, Gil-Izquierdo A, Troncoso AM, Cantos E, Garcia-Parrilla MC (2011) Melatonin: a new bioactive compound in wine. J Food Compos Anal 24(4–5):603–608
68. Wang D, Zhang Z, Ju J, Wang X, Qiu W (2011) Investigation of piceid metabolites in rat by liquid chromatography tandem mass spectrometry. J Chromatogr B 879(1):69–74
69. Kong QJ, Ren XY, Hu N, Sun CR, Pan YJ (2011) Identification of isomers of resveratrol dimer and their analogues from wine grapes by HPLC/MSn and HPLC/DAD-UV. Food Chem 127(2):727–734
70. Jiang L-Y, He S, Jiang K-Z, Sun C-R, Pan Y-J (2010) Resveratrol and its oligomers from wine grapes are selective 1O2 quenchers: mechanistic implication by high-performance liquid chromatography–electrospray ionization–tandem mass spectrometry and theoretical calculation. J Agric Food Chem 58(16):9020–9027
71. Montes R, García-López M, Rodríguez I, Cela R (2010) Mixed-mode solid-phase extraction followed by acetylation and gas chromatography mass spectrometry for the reliable determination of trans-resveratrol in wine samples. Anal Chim Acta 673(1):47–53
72. Arbulu M, Sampedro MC, Sanchez-Ortega A, Gómez-Caballero A, Unceta N, Goicolea MA, Barrio RJ (2013) Characterisation of the flavour profile from Graciano Vitis vinifera wine variety by a novel dual stir bar sorptive extraction methodology coupled to thermal desorption and gas chromatography–mass spectrometry. Anal Chim Acta 777:41–48
73. Cacho JI, Campillo N, Viñas P, Hernández-Córdoba M (2013) Stir bar sorptive extraction with gas chromatography–mass spectrometry for the determination of resveratrol, piceatannol and oxyresveratrol isomers in wines. J Chromatogr A 1315:21–27
74. Viñas P, Martínez-Castillo N, Campillo N, Hernández-Córdoba M (2011) Directly suspended droplet microextraction with in injection-port derivatization coupled to gas chromatography–mass spectrometry for the analysis of polyphenols in herbal infusions, fruits and functional foods. J Chromatogr A 1218(5):639–646
75. Viñas P, Campillo N, Martínez-Castillo N, Hernández-Córdoba M (2009) Solid-phase microextraction on-fiber derivatization for the analysis of some polyphenols in wine and grapes using gas chromatography–mass spectrometry. J Chromatogr A 1216(9):1279–1284
76. Paul S, Mizuno CS, Lee HJ, Zheng X, Chajkowisk S, Rimoldi JM, Conney A, Suh N, Rimando AM (2010) In vitro and in vivo studies on stilbene analogs as potential treatment agents for colon cancer. Eur J Med Chem 45(9):3702–3708
77. Sokolowska I, Ngounou Wetie AG, Woods AG, Darie CC (2012) Automatic determination of disulfide bridges in proteins. J Lab Autom 17(6):408–416
78. Sokolowska I, Woods AG, Wagner J, Dorler J, Wormwood K, Thome J, Darie CC (2011) Mass spectrometry for proteomics-based investigation of oxidative stress and heat shock proteins. In: Andreescu S, Hepel M (eds) Oxidative stress: diagnostics, prevention, and therapy, vol 1083. American Chemical Society, Washington, DC, pp 369–411
79. Ngounou Wetie AG, Sokolowska I, Woods AG, Wormwood KL, Dao S, Patel S, Clarkson BD, Darie CC (2013) Automated mass spectrometry-based functional assay for the routine analysis of the secretome. J Lab Autom 18(1):19–29
80. Villagra E, Santos LS, Vaz BG, Eberlin MN, Felipe Laurie V (2012) Varietal discrimination of Chilean wines by direct injection mass spectrometry analysis combined with multivariate statistics. Food Chem 131(2):692–697
81. de Souza PP, Resende AMM, Augusti DV, Badotti F, Gomes Fde C, Catharino RR, Eberlin MN, Augusti R (2014) Artificially-aged cachaça samples characterised by direct infusion electrospray ionisation mass spectrometry. Food Chem 143:77–81
82. Cooper HJ, Marshall AG (2001) Electrospray ionization Fourier transform mass spectrometric analysis of wine. J Agric Food Chem 49(12):5710–5718
83. Draper J, Lloyd AJ, Goodacre R, Beckmann M (2013) Flow infusion electrospray ionisation mass spectrometry for high throughput, non-targeted metabolite fingerprinting: a review. Metabolomics 9(1):4–29

84. Gross J (2004) Electron ionization. Mass spectrometry. Springer, Berlin, pp 193–222. ISBN 978-3-642-07388-5
85. Gougeon RD, Lucio M, Frommberger M, Peyron D, Chassagne D, Alexandre H, Feuillat F, Voilley A, Cayot P, Gebefügi I, Hertkorn N, Schmitt-Kopplin P (2009) The chemodiversity of wines can reveal a metabologeography expression of cooperage oak wood. Proc Natl Acad Sci U S A 106(23):9174–9179
86. Roullier-Gall C, Boutegrabet L, Gougeon RD, Schmitt-Kopplin P (2014) A grape and wine chemodiversity comparison of different appellations in burgundy: vintage vs terroir effects. Food Chem 152:100–107
87. Richard T, Poupard P, Nassra M, Papastamoulis Y, Iglésias M-L, Krisa S, Waffo-Teguo P, Mérillon J-M, Monti J-P (2011) Protective effect of ε-viniferin on β-amyloid peptide aggregation investigated by electrospray ionization mass spectrometry. Bioorg Med Chem 19(10): 3152–3155

Chapter 18
Mass Spectrometric DNA Adduct Quantification by Multiple Reaction Monitoring and Its Future Use for the Molecular Epidemiology of Cancer

Bernhard H. Monien

Abstract The formation of DNA adducts is considered essential for tumor initiation. Quantification of DNA adducts may be achieved by various techniques of which LC-MS/MS-based multiple reaction monitoring has become the most prominent in the past decade. Adducts of single nucleosides are analyzed following enzymatic break-down of a DNA sample following adduct enrichment usually by solid-phase extraction. LC-MS/MS quantification is carried out using stable isotope-labeled internal reference substances. An upcoming challenge is the use of DNA adducts as biomarkers either for internal exposure to electrophilic genotoxins or for the approximation of cancer risk. Here we review recent studies in which DNA adducts were quantified by LC-MS/MS in DNA samples from human matrices. We outline a possible way for future research to aim at the development of an "adductome" approach for the characterization of DNA adduct spectra in human tissues. The DNA adduct spectrum reflects the inner exposure of an individual's tissue to electrophilic metabolites and, therefore, should replace the conventional and inaccurate external exposure in epidemiological studies in the future.

Abbreviations

2-AAF	2-Acetylaminofluorene
4-ABP	4-Aminobiphenyl
BaP	Benzo[*a*]pyrene
CYP	Cytochrome P450 monooxygenase

B.H. Monien (✉)
Lead, Genotoxic Food Contaminants Research Group, German Institute of Human Nutrition (DIfE) Potsdam-Rehbrücke, Arthur-Scheunert-Allee 114-116, 14558 Nuthetal, Germany
e-mail: monien@dife.de

A.G. Woods and C.C. Darie (eds.), *Advancements of Mass Spectrometry in Biomedical Research*, Advances in Experimental Medicine and Biology 806,
DOI 10.1007/978-3-319-06068-2_18, © Springer International Publishing Switzerland 2014

dA	2′-Deoxyadenosine
dC	2′-Deoxycytidine
dG	2′-Deoxyguanosine
HAA	Heterocyclic aromatic amine
LC-MS/MS	Liquid chromatography—tandem mass spectrometry
MP	1-Methylpyrene
MRM	Multiple reaction monitoring
PAH	Polycyclic aromatic hydrocarbon
PhIP	2-Amino-1-methyl-6-phenylimidazo[4,5-*b*]pyridine
TLC	Thin-layer chromatography

18.1 Introduction

Chemical carcinogens cause different DNA lesions, such as adducts, strand breaks, and cross-linked DNA strands. If not fixed by DNA repair mechanisms the damage can lead to mutation events. If this occurs in genes involved in the regulation of the cell cycle, differentiation, or cell–cell interaction, autonomous growth and metastatic potential could be achieved [21]. The vast majority of chemical carcinogens, e.g., heterocyclic aromatic amines (HAA) [24], polycyclic aromatic hydrocarbons (PAHs) [25], *N*-nitrosamines [23], and halogenated hydrocarbons [19, 22] are relatively inert. Enzyme-catalyzed oxidation and conjugation reactions can convert procarcinogens into electrophilic metabolites that can react via nucleophilic substitutions with proteins, RNA or DNA. Cytochrome P450 (CYP) is primarily involved in oxidation reactions. For example, the epoxidation of aflatoxin B1 to aflatoxin B1 *exo*-8,9-oxide is catalyzed primarily by CYP3A4 [20]. CYP1A1 and 1B1 substantially contribute to the bioactivation of benzo[*a*]pyrene (BaP) generating two enantiomers of *trans*-7,8-diols that can be further epoxidized to the ultimate carcinogen BaP-7*R*,8*S*-diol-9*S*,10*R*-epoxide by CYP1A1, 1A2, 1B1, and 2C9 [26, 50]. But also conjugation reactions catalyzed by sulfotransferases, glutathione *S*-transferases, and *N*-acetyl transferases play critical roles in metabolic activation of procarcinogens resulting in covalent modifications of the DNA [48].

Formation of DNA adducts is considered the initial event in cancer development. Studies of animals exposed to common carcinogenic compounds showed that increasing concentrations of DNA adducts were usually associated with growing tumor numbers; however, the correlations were not necessarily linear [36, 42]. Also, DNA adducts can be found in organs that do not develop tumors indicating that other factors, e.g., the tissue-specific capacity of cell-proliferation, co-determine the risk for tumor induction. As tumors do not form in the absence of DNA adducts in animal models, DNA adduct formation is considered a “necessary but not sufficient” requirement for cancer development.

Concentrations of DNA adducts in a particular tissue of an animal treated with a test substance or of human origin may be quantified by different techniques, such as ^{32}P-postlabeling [46] or LC-MS/MS multiple reaction monitoring (MRM) [12, 27].

Since its introduction in 1981 ^{32}P-postlabeling was the gold standard of DNA adduct quantification at the end of the last century due to its superior sensitivity. The applicability of the method resulted in a multitude of studies intending to establish correlations between concentrations of DNA adducts in particular tissues of tumor patients vs. control subjects. Various reviews summarize exemplary data in the field [1, 15, 39, 56]. However, many attempts to correlate DNA adducts and tumor incidence were inconclusive and the value of the data recorded by ^{32}P-postlabeling is questioned today due to the insufficient specificity of the method. Technical details are explained in the next paragraph. In the last decade LC-MS/MS-based techniques were used increasingly for the highly sensitive and specific quantification of DNA adducts via monitoring of analyte-specific molecular fragmentation reactions [27]. We present the results of several recent studies describing the application of LC-MS/MS analytical techniques for the quantification of DNA adducts from human biological matrices. Future research will be directed towards monitoring of multiple DNA adducts in order to characterize the inner exposure of the human genome to electrophilic substances. This "adductome" approach may greatly improve the interpretation of epidemiological data related to cancer development.

18.2 Quantification of DNA Adducts: Technical Details

A variety of analytical methods for the quantification of DNA adducts are available. Until 1981, radioactively labeled carcinogens were used to calculate DNA adduct levels. Later, alternative methods have been used, such as ^{32}P-postlabeling [46], immunofluorescent detection [43], a competitive radioimmunoassay [54], and LC-MS/MS MRM. Table 18.1 provides a brief overview of the techniques. The choice of methods depends on various factors such as the amount of DNA available, the chemical nature of DNA adducts (hydrophobicity), and the scientific question (for example, the search for genotoxins of yet unknown identity in a complex mixture of compounds or quantification of well-defined DNA adducts). ^{32}P-postlabeling was introduced in 1981 and is still attractive because of the sensitivity in the detection of adducts formed from large hydrophobic substances such as PAHs. The method combines the insertion of a radioactive [^{32}P]phosphoryl group at the 5′-hydroxy position of the 3′-mononucleotide adducts after DNA cleavage with subsequent separation of labeled adducts by thin-layer chromatography (TLC). Sensitive autoradiography is used to visualize the chromatographic pattern of DNA adducts on the TLC plate. ^{32}P-postlabeling is still a valuable tool in DNA adduct analyses, especially when molecular adduct structures are unidentified. Thus, ^{32}P-postlabeling is advantageous for the detection of DNA adducts of genotoxic substances in mixtures of environmental xenobiotics of unknown composition, for example, in food plants [2]. The sensitive technique enables the detection of adduct levels in the range of 1 adduct/10^{10} nucleotides using only 10 micrograms DNA [11]. However, the pattern on the TLC plate often shows a number of spots of unknown origin, whose identities can only be conjectured by co-chromatography of standard substances.

Table 18.1 Common techniques for the quantification of DNA adducts

Method	Procedure	Advantages	Drawbacks
^{32}P-postlabeling	– Enzymatic hydrolysis of DNA into 3′-mononucleotides – Nucleotide adducts are labeled with ^{32}P-phosphate at the 5′-end – Separation of adducts over TLC and detection by scintillation counting	– Requires only 1–10 µg DNA – Sensitive – Knowledge of DNA-adduct structures is not required	– Radioactive labeling – Unspecific detection – Underestimation of adduct levels
LC-MS/MS	– Enzymatic hydrolysis of DNA into nucleosides – Enrichment of adducts by extraction methods – LC-MS/MS	– Specific detection – Sensitive quantification	– Requires >10 µg DNA – Isotope-labeled reference substances are desirable
Immunoassays	– Use of antiserum for a DNA adduct in competitive immunoassay, endpoints, e.g., radioactivity or fluorescence	– Inexpensive – Immunohisto-chemistry allows studying localization of adducts	– Unspecific detection (crossreactions) – Requires 50–100 µg DNA – Overestimation of adduct levels

In the past decade, LC-MS/MS techniques have become more important for the quantification of DNA adducts, even though standard substances are indispensable. The LC-MS/MS techniques are characterized by several advantages, including high specificity for the detected DNA adducts, straightforward quantification by isotope-labeled internal reference compounds, and high-throughput capability. The specific detection of nucleoside adducts is based on collision-induced fragmentation of recurring structural motifs. Usually, the adducts are formed by nucleophilic substitution of reactive metabolites and atoms of 2′-deoxyadenosine (dA), 2′-deoxyguanosine (dG), and 2′-deoxycytidine (dC). Figure 18.1a shows the dG adduct of 1-methylpyrene (MP), N^2-((pyren-1-yl)methyl)-dG (N^2-MP-dG). MP is a common carcinogenic food contaminant [33]. It can be bioactivated by CYP-catalyzed hydroxylation at the exocyclic methyl group and subsequent sulfo conjugation resulting in a highly reactive sulfate ester, which undergoes nucleophilic substitutions with exocyclic nitrogens of dA, dG, or dC [33].

Figure 18.1b shows the dG adduct of 2-amino-1-methyl-6- phenylimidazo[4,5-*b*] pyridine (PhIP), a highly carcinogenic HAA isolated from well-done meat [13]. Similar to MP, PhIP is bioactivated by CYP-catalyzed hydroxylation at the exocyclic nitrogen and subsequent sulfo conjugation [14]. The sulfate ester of N^2-hydroxy-PhIP causes the formation of an adduct of dG on *C*8, *C*8-(2-amino-1-methyl-6-phenylimidazo[4,5-*b*]pyridine-N^2-yl)-dG (*C*8-PhIP-dG) (Fig. 18.1b).

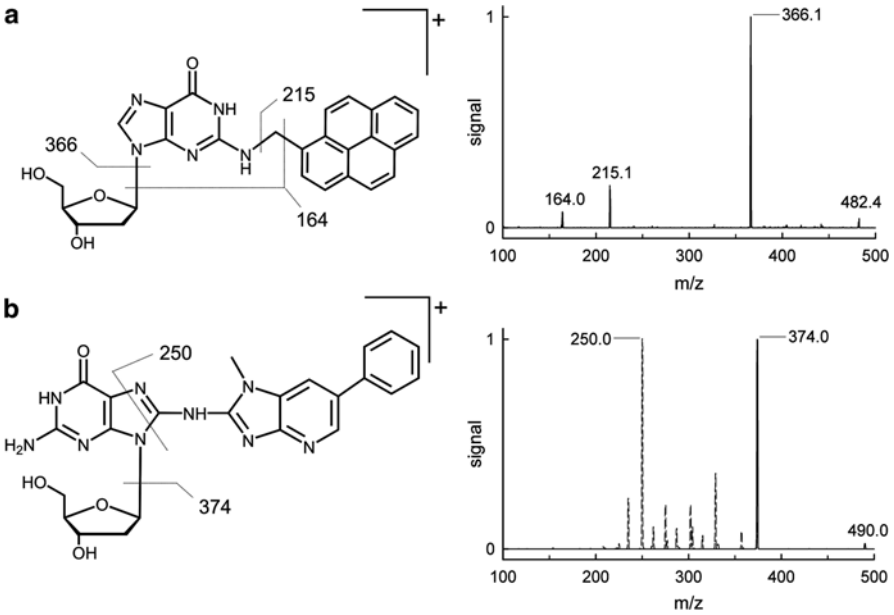


Fig. 18.1 Molecular fragmentation of DNA adducts in the collision cell of a triple quadrupole mass spectrometer. Fragmentation patterns of (**a**) N^2-((pyren-1-yl)methyl)-dG (N^2-MP-dG) and (**b**) *C*8-(2-amino-1-methyl-6-phenylimidazo[4,5-*b*]pyridine-N^2-yl)-dG (*C*8-PhIP-dG). The fragment spectrum of N^2-MP-dG (**a**) recorded by collision-induced dissociation showed ions at $m/z = 366.1$ (the aglycone of N^2-MP-dG), $m/z = 215.1$ (the MP-cation), and $m/z = 164.0$ (protonated N^2-methylguanine). The collision-induced dissociation of *C*8-PhIP-dG (**b**) generated fragments of $m/z = 374.1$ (the aglycone of *C*8-PhIP-dG) with a collision energy of 20 eV (*solid line* in the mass spectrum) and a group of fragments at a collision energy of 50 eV (*dashed line* in the mass spectrum) with a dominating signal at $m/z = 250.0$

The molecular structures of the adducts determine the fragmentation reactions in the collision cell of the mass spectrometer, which allows their specific detection and quantification. The breakage of the glycosidic bond in the protonated nucleoside adduct $[M+H]^+$ leads to the neutral loss of 2′-deoxyribose with a mass of 116 Da and the formation of the base adduct $[B+H]^+$ (Fig. 18.1). Another route of collision-induced dissociation of the precursor ion $[M+H]^+$ leads to the release of a positively charged fragment of adduct molecules, which can be observed for numerous different N^6-adducts of dA and N^2-adducts of dG [30–32, 44]. However, the adduct *C*8-PhIP-dG breaks at higher collision energies in many different fragments that cannot be assigned unambiguously to particular molecular structures (see mass spectrum in Fig. 18.1b). Figure 18.2 shows four chromatograms resulting from the collision-induced fragmentation analysis of another common exemplary DNA adduct formed after the uptake of the rodent carcinogen furfuryl alcohol which is a food contaminant present at high levels in the human diet [31]. The upper panel shows the neutral loss of the 2′-deoxyribose ($332.1 \rightarrow 216.1$), which is used as a *quantifier signal*. The chromatogram of the second panel results from the cleavage of the positively

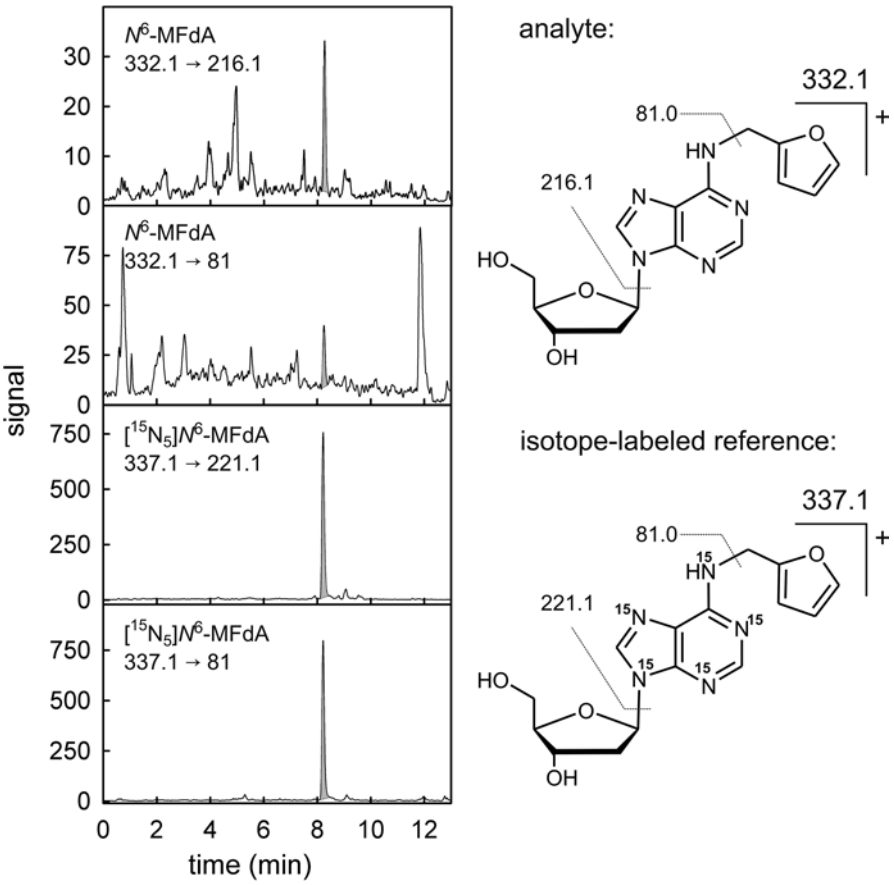


Fig. 18.2 LC-MS/MS analytical quantification of N^6-((furan-2-yl)methyl)-2′-deoxyadenosine (N^6-MFdA) formed by the rodent carcinogen furfuryl alcohol. The chromatograms are LC-MS/MS MRM traces of N^6-MFdA in a digest of DNA isolated from the liver of furfuryl alcohol-treated mice. The fragmentations 332.1 → 216.1 (*first panel*) and 332.1 → 81 (*second panel*) allowed detecting N^6-MFdA and were monitored together with the transitions 337.1 → 221.1 (*third panel*) and 337.1 → 81 (*fourth panel*) of the internal isotope-labeled standard [$^{15}N_5$]N^6-MFdA (38.0 fmol/injection). The ratio of peak areas for the transition 332.1 → 216.1 (N^6-MFdA) and for the transition 337.1 → 221.1 ([$^{15}N_5$]N^6-MFdA) was used to calculate the N^6-MFdA content of the DNA. Details of the method are outlined in [31]

charged methylfuran fragment (332.1 → 81). These traces are monitored together with two additional MRM signals from the transitions 337.1 → 221.1 (third panel) and 337.1 → 81 (fourth panel) of the internal isotope-labeled standard [$^{15}N_5$]N^6-MFdA. The ratio of peak areas for the transition 332.1 → 216.1 (N^6-MFdA) and for the transition 337.1 → 221.1 ([$^{15}N_5$]N^6-MFdA) is used for the quantification of the N^6-MFdA in the sample.

The main advantage of this procedure is the specificity of detection. The identity of a particular nucleoside adduct is confirmed by a specific retention time of the chromatography and the MS/MS-monitoring of several selected fragmentation reactions.

In return, the scientist accepts that other possible adducts remain unobserved. Further, the sensitivity of LC-MS/MS detection with detection limits in the range 1–10 adducts/10^9 nucleotides using 100 μg DNA is somewhat lower compared to that of ^{32}P-postlabeling [12, 27]. However, the quantification of the adducts by LC-MS/MS provides more accurate results in comparison with ^{32}P-postlabeling and immunoassays: Beland et al. determined adduct levels by LC-MS/MS, ^{32}P-postlabeling, and a fluoroimmunoassay in liver DNA of mice treated with [2,2′-^{3}H]-4-aminobiphenyl (4-ABP) and compared the results with the quantification by the ^{3}H-labeled DNA. The result of the LC-MS/MS method agreed best to the adduct concentration determined by scintillation counting, while ^{32}P-postlabeling and the fluoroimmunoassay grossly under- or overestimated the correct adduct concentration [4]. This is consistent with other comparative studies. For a methodological comparison between LC-MS/MS and ^{32}P-postlabeling, we determined the adduct concentrations in hepatic DNA of rats treated with the active metabolite of MP. The concentrations of N^2-MP-dG as determined by LC-MS/MS were on average 3.4 times higher compared to the amounts determined by ^{32}P-postlabeling. Also in the case of adducts of BaP [51] and PhIP [16] LC-MS/MS methods reported 3.7- and 20-fold higher adduct levels compared to ^{32}P-postlabeling, respectively. Factors that may contribute to the underestimation of DNA adduct concentration by ^{32}P-postlabeling are the incomplete digestion of the sample DNA as well as a partial phosphorylation of the modified nucleotides [4, 40, 47]. We showed that an over-digestion of the MP adducts leading to an unintentional loss of the 3′-phosphate also contributes to the adduct loss. The N^2-MP-dG-3′-phosphate proved not to be entirely resistant to dephosphorylation in the presence of micrococcus nuclease, spleen phosphodiesterase, and nuclease P1. About 20 % of the N^2-MP-dG-3′-phosphate was hydrolyzed to N^2-MP-dG and was thus lost to the ^{32}P-postlabeling by T4 polynucleotide kinase [33].

In summary, the advantages of the LC-MS/MS MRM prevail. There are no difficulties resulting from nonspecific enzymatic reactions as in ^{32}P-postlabeling because the DNA samples are digested completely to nucleosides in preparation for LC-MS/MS analysis. An effective solid-phase extraction for enrichment of the adducts allows for almost total isolation of modified nucleosides. The use of stable isotope-labeled standard substances ensures compensation of the analyte losses during the workup for a highly specific detection and ultimately for a convenient quantification of the analytes. In addition, the use of radioactivity is avoided and the time for sample preparation is shorter compared to ^{32}P-postlabeling, which allows the daily processing of 100 samples and the future application of LC-MS/MS techniques in routine analyses of DNA adducts.

18.3 The Scope of DNA Adducts as Human Biomarkers

DNA adducts are considered a prerequisite for the development of tumors. Otteneder et al. established a tentative correlation of hepatic DNA adduct concentrations and the incidence of liver tumors after chronic exposure of common carcinogens in mice and rats [36]. The calculated adduct concentrations at the TD_{50} (dose that caused a

50 % increase of tumors in the treated animals over the controls) ranged from 53 adducts/10^8 nucleosides for aflatoxin B1 to 2,083 adducts/10^8 nucleosides for *N*-nitrosodimethylamine in rats. In mice, the adduct concentrations at the TD_{50} ranged from 812 adducts/10^8 nucleosides for ethylene up to 5,543 adducts/10^8 nucleosides for 2-AAF. This vague correlations between adduct levels and tumorigenic effect of the substances pointed out that the carcinogenic potency of individual DNA adducts may differ greatly [36]. Nevertheless, due to the application of LC-MS/MS for the specific quantification of single DNA adducts it should be theoretically possible to predict the hepatic cancer risk. However, there are several reasons arguing against the prediction of tumor incidences from DNA adduct levels [37]: (1) Usually, the carcinogenicity of a particular genotoxin does not increase in a linear fashion with increasing DNA adduct levels but depends on the species, the gender, and the tissue [42]. This is in part due to tissue-specific differences in DNA repair and cell proliferation, both of which influence the effects of DNA adducts [37]. For example, male mice are twice as sensitive with respect to the hepatocarcinogenic effect of the adducts of *N*-nitrosodimethylamine compared to male rats [36]. (2) Tissue samples from healthy persons are usually not accessible. Non-invasive studies are restricted to the adduct analysis of DNA samples from leukocytes, cells of the sputum, breast milk, and from urine. This restricts the prospect of the possible future DNA adduct analyses for the prediction of a tissue-specific cancer risk in humans.

Although DNA adducts may not be used as biomarkers of effect, they may offer a superior tool for the characterization of the inner exposure to electrophilic compounds. Epidemiological studies investigating the association between exposure to complex mixtures of compounds, e.g., food, and tumor incidences, are greatly hampered by the modeling of the exposure. The subject's exposure, for example, food uptake, is usually deduced from questionnaires which are recognized as a source of inaccuracy ("recall bias"). Moreover, the "external" exposure of a subject to a compound does not necessarily reflect the internal effect of a bioactive metabolite to the individual's genome. The sequence of events between uptake of genotoxic carcinogens and a mutation includes the following steps: (1) absorption and bioactivation of the genotoxin, (2) possible detoxification of the reactive metabolite, (3) reaction with proteins, RNA or DNA (only the latter case is of importance for tumor initiation), (4) persistence of the DNA adduct or removal by repair mechanisms, and (5) proliferation of the cell containing the DNA adduct. This sequence of steps varies greatly between individuals. Further, factors of life style, e.g., alcohol consumption and permanent drug medication, were shown to exert considerable effects on tissue-specific concentrations of DNA adducts [28, 32]. And finally, there are carcinogens in complex mixtures for which the actual exposure is very difficult to calculate because accurate concentrations of the compounds cannot be determined. For example, the accurate intake of methyleugenol remains elusive due to variations of its content in food plants and spices [53]. Therefore, characterization of the "adductome," the pattern of an individual's inner exposure towards electrophilic compounds, would incorporate all interindividual differences in absorption, bioactivation, detoxification, and other parameters, such as life style and medication, that would influence the formation of DNA adducts. The replacement of the external exposure

with the characterization of the adduct-load of human DNA samples (the "adductome") may increase the scope of interpretation in future epidemiological studies targeted at the origins of tumor development.

18.4 DNA Adduct Analyses in Human Biological Matrices: Current Research

Numerous studies were published recently, in which LC-MS/MS MRM was used for the analysis of DNA adducts from different environmental carcinogens in human samples, e.g., from 4-ABP in pancreas tissue from smokers [49], from acrolein [57] or acetaldehyde [8] in leukocytes of smokers, from estrogen in breast tumor tissue [10], from PhIP and 4-ABP in saliva samples [6], or from tamoxifen in colon tissue [7]. All reports describe pilot studies with less than 50 persons using different kinds of biological matrices including saliva, pancreatic tissue, leukocytes, or breast tissue. The outcomes of the studies varied. Chen et al. showed that acetaldehyde adducts in DNA of leukocytes of 25 smokers decreased within several weeks of smoking abstinence [8]. The concentration of *C*8-PhIP-dG in the DNA of epithelial buccal cells did not correlate to consumption of grilled meat or smoking in 37 persons [6], and there was also no association between the 4-ABP adduct of dG in pancreatic tissue samples and smoker status of twelve participants [49]. The emphases of these works were on the description of the LC-MS/MS techniques demonstrating the feasibility of the studies [6–8, 10, 49, 57].

More recently, levels of *C*8-PhIP-dG were determined in adjacent tissue of mammary tumors from 70 patients using a sensitive LC-MS/MS method. The adduct was detectable in merely one sample, at a level of three molecules *C*8-PhIP-dG/10^9 nucleotides [18]. This result is in conflict with previous studies using immunohistochemistry and ^{32}P-postlabeling analytical methods. Zhu et al. reported elevated dG-PhIP adduct concentrations in normal breast tissue of 87 from 106 mammary tumor patients using immunohistochemistry (limit of detection ~1 adduct/10^7 nucleotides) [58]. Gorlewska-Roberts described the detection of dG-PhIP in 30 DNA samples from exfoliated ductal epithelial cells isolated from milk samples of 64 lactating women (mean value 4.7 adducts/10^7 nucleotides, no limit of detection reported) [17]. The discrepancies between results from studies using either highly specific LC-MS/MS MRM or the less selective immunohistochemistry and ^{32}P-postlabeling suggest critical revisions of many older biomarker studies that found correlations between tumor incidence and occurrence of DNA adducts [38, 39, 55]. The elaborate method of accelerator mass spectrometry was used by Brown et al. to detect dG-N^2-tamoxifen in colon DNA of women who were treated with a single dose of 20 mg [^{14}C]-labeled tamoxifen [7]. This supported the hypothesis of a causal relationship between tamoxifen therapy and the increased risk for the incidence of colorectal tumors in tamoxifen-treated women [35]. Taken together, these studies show that progressing development of LC-MS/MS technical equipment allows adduct analyses at sensitivities that were reported previously only for ^{32}P-postlabeling.

There are only few studies focusing on the correlation of increased DNA adduct levels and cancer risk in which specific LC-MS/MS was used to assess DNA damage. A Chinese cohort of 18,000 men were enrolled in a study in order to clarify the consequences of aflatoxin B1 intake and hepatitis B viral infection on the development of hepatic cancer. Samples of urine were analyzed for the adduct aflatoxin B1-N^7-guanine, which originates from hydrolysis of the *N*-glycosidic bond in the DNA adduct aflatoxin B1-N^7-dG. Men without hepatitis B infection but with measurable urine concentrations of aflatoxin B1-N^7-guanine faced a three-fold higher risk for the development of hepatic tumors compared to subjects of the control group. The relative risk was even increased in individuals infected with hepatitis (RR = 59.4, CI = 16.6, 212.0) [45].

Various studies presented correlations between concentrations of DNA adducts from the lipid peroxidation products malon dialdehyde and 4-hydroxy-2-nonenal as biomarkers for oxidative stress. The latter was shown to generate etheno-adducts, e.g., 1,N^6-etheno-dA (εdA) and N^2,3-etheno-dG (εdG). Increases of hepatic etheno-adducts were found in patients with either Wilson's disease or primary hemochromatosis, both of which induce hepatic oxidative stress [34]. Elevated urinary εdA concentrations were found in patients with alcoholic liver disease, chronic hepatitis, and liver cirrhosis, all of which are precancerous illnesses [3]. Bartsch et al. suggested that etheno-adducts in urine and needle liver biopsies may be explored as putative risk markers and to evaluate chemopreventive and therapeutic intervention strategies. However, laborious validation is required for the application of DNA adducts as biomarkers of cancer risk, in case a specific cancer incidence can be attributed to a particular DNA adduct. The validation of a tumor-DNA adduct correlation requires a prospective nested case-control study, in which a large group of participants have to be monitored over many years until cancer develops [41].

18.5 "Adductomics": Monitoring of Multiple DNA Adducts

Recently, many reports were published about DNA adduct quantification using LC-MS/MS-based techniques. Usually, the scientists focused on DNA adducts derived from single carcinogens. However, the association between the exposure to a single carcinogen and the development of a specific tumor as observed for aflatoxin B1 and hepatic cancer is not a common observation. Humans are exposed to complex mixtures of carcinogens. For example, food uptake confronts the organism with a plethora of mutagenic and carcinogenic compounds including mycotoxins, HAA, PAH, heavy metals, *N*-nitrosoalkylamines, substituted furans, etc. Most reports presented in the preceding paragraph were proof-of-concept studies showing the applicability of a novel analytical method for the quantification of DNA adducts of single genotoxins. However, single genotoxins only contribute to the overall cancer risk, which should be better described by the sum of all DNA lesions. Apart from the efforts to further increase the sensitivities of LC-MS/MS-based methods future research will be aimed at simultaneous analyses of different DNA

adducts reflecting the exposure to many of carcinogens. Singh et al. described a method for the quantification of multiple DNA adducts by LC-MS/MS neutral loss of 2′-deoxyribose in a digest of calf thymus DNA incubated with a mixture of dihydrodiol-epoxides of different PAHs [52]. More recently, human autopsy tissue samples were analyzed for 16 different DNA adducts originating from lipid peroxidation demonstrating that the "adductomic" strategy is transferable to studies with human samples for the assessment of internal exposure to electrophilic substances [9]. This approach was further used to analyze samples of gastric mucosa from Japanese and Chinese cancer patients who underwent gastrectomy [29]. These studies were initial steps on the way to develop techniques for the quantification of multiple DNA adducts in human biomatrices. Two further hurdles should be mentioned that determine the progress in this field. Since tissue samples of living human subjects are usually not available (except from cancer patients undergoing surgery), the future use of the methods requires adaption to the analysis of (small amounts of) DNA samples obtained from noninvasive procedures, e.g., from leukocytes. Thus, the progress depends in part on the continuous instrumental advance yielding optimized chromatographic and mass spectrometric equipment. Second, future prediction of cancer risk from assessment of multiple DNA adduct levels, *i.e.*, the application of the DNA adduct spectrum as human biomarker, requires a validation in prospective molecular epidemiology studies. This will be a time-consuming endeavor. However, future analyses of DNA adduct spectra may greatly amplify the significance of human biomonitoring and may extend the scope of interpretations in epidemiological studies.

18.6 Conclusions

The number of reports on sensitive LC-MS/MS analytical techniques for quantification of adducts in DNA samples of human origin is constantly increasing. This is, in part, due to the continuous instrumental advance yielding optimized chromatographic and mass spectrometric equipment, which allows compensating the sensitivity advantage of ^{32}P-postlabeling. As a result, LC-MS/MS MRM is currently the method of choice for DNA adduct analyses. ^{32}P-postlabeling is still attractive for screening purposes if target DNA adducts are not yet characterized or if facilities for organic synthesis of standard substances are lacking.

The assessment of DNA adducts as biomarkers for the definition of internal exposure to environmental carcinogens has several advantages over traditional exposure estimation. Most importantly, DNA adducts account for interindividual differences in uptake, elimination, distribution, metabolism, and repair among exposed individuals as well as for different uptake routes. Consequently, DNA adducts may be helpful tools for the establishment of biologically plausible associations between exposure and disease in epidemiological studies. For example, application of DNA adducts as biomarkers of exposure, e.g., to reactive metabolites in common food carcinogens, such as furfuryl alcohol or methyleugenol, could be

very helpful because the external exposure of these substances resulting from many sources is very difficult to determine [5, 31]. Further, DNA adducts may serve as valuable endpoints in intervention studies. However, the association between tissue concentrations of single adducts and the outcome of a cancer study does not seem plausible. We believe that future analyses of multiple DNA adducts providing an overview of genomic damage due to reactive electrophiles of exogenous or endogenous origin will serve as a valid parameter for molecular epidemiology of cancer.

Acknowledgments The author gratefully acknowledges financial support from the German Research Foundation (MO 2520/1-1) and the German Institute of Human Nutrition (DIfE).

References

1. Angerer J, Ewers U, Wilhelm M (2007) Human biomonitoring: state of the art. Int J Hyg Environ Health 210:201–228
2. Baasanjav-Gerber C, Monien BH, Mewis I, Schreiner M, Barillari J, Iori R et al (2011) Identification of glucosinolate congeners able to form DNA adducts and to induce mutations upon activation by myrosinase. Mol Nutr Food Res 55:783–792
3. Bartsch H, Arab K, Nair J (2011) Biomarkers for hazard identification in humans. Environ Health 10 suppl 1:S11
4. Beland FA, Doerge DR, Churchwell MI, Poirier MC, Schoket B, Marques MM (1999) Synthesis, characterization, and quantitation of a 4-aminobiphenyl-DNA adduct standard. Chem Res Toxicol 12:68–77
5. Benford D, Bolger PM, Carthew P, Coulet M, DiNovi M, Leblanc JC et al (2010) Application of the margin of exposure (MOE) approach to substances in food that are genotoxic and carcinogenic. Food Chem Toxicol 48(suppl 1):S2–S24
6. Bessette EE, Spivack SD, Goodenough AK, Wang T, Pinto S, Kadlubar FF et al (2010) Identification of carcinogen DNA adducts in human saliva by linear quadrupole ion trap/multistage tandem mass spectrometry. Chem Res Toxicol 23:1234–1244
7. Brown K, Tompkins EM, Boocock DJ, Martin EA, Farmer PB, Turteltaub KW et al (2007) Tamoxifen forms DNA adducts in human colon after administration of a single [^{14}C]-labeled therapeutic dose. Cancer Res 67:6995–7002
8. Chen L, Wang M, Villalta PW, Luo X, Feuer R, Jensen J et al (2007) Quantitation of an acetaldehyde adduct in human leukocyte DNA and the effect of smoking cessation. Chem Res Toxicol 20:108–113
9. Chou PH, Kageyama S, Matsuda S, Kanemoto K, Sasada Y, Oka M et al (2010) Detection of lipid peroxidation-induced DNA adducts caused by 4-oxo-2(E)-nonenal and 4-oxo-2(E)-hexenal in human autopsy tissues. Chem Res Toxicol 23:1442–1448
10. Embrechts J, Lemiere F, Van Dongen W, Esmans EL, Buytaert P, Van Marck E et al (2003) Detection of estrogen DNA-adducts in human breast tumor tissue and healthy tissue by combined nano LC-nano ES tandem mass spectrometry. J Am Soc Mass Spectrom 14:482–491
11. Farmer PB, Brown K, Tompkins E, Emms VL, Jones DJ, Singh R et al (2005) DNA adducts: mass spectrometry methods and future prospects. Toxicol Appl Pharmacol 207:293–301
12. Farmer PB, Singh R (2008) Use of DNA adducts to identify human health risk from exposure to hazardous environmental pollutants: the increasing role of mass spectrometry in assessing biologically effective doses of genotoxic carcinogens. Mutat Res 659:68–76
13. Felton JS, Knize MG, Shen NH, Lewis PR, Andresen BD, Happe J et al (1986) The isolation and identification of a new mutagen from fried ground beef: 2-amino-1-methyl-6-phenylimidazo[4,5-b]pyridine (PhIP). Carcinogenesis 7:1081–1086

14. Glatt HR (2006) Metabolic factors affecting the mutagenicity of heterocyclic amines. In: Skog K, Alexander J (eds) Acrylamide and Other Health Hazardous Compounds in Heat-treated Foods. Woodhead Publishing, Cambridge, pp 358–404
15. Godschalk RW, Van Schooten FJ, Bartsch H (2003) A critical evaluation of DNA adducts as biological markers for human exposure to polycyclic aromatic compounds. J Biochem Mol Biol 36:1–11
16. Goodenough AK, Schut HA, Turesky RJ (2007) Novel LC-ESI/MS/MS(n) method for the characterization and quantification of 2′-deoxyguanosine adducts of the dietary carcinogen 2-amino-1-methyl-6-phenylimidazo[4,5-*b*]pyridine by 2-D linear quadrupole ion trap mass spectrometry. Chem Res Toxicol 20:263–276
17. Gorlewska-Roberts K, Green B, Fares M, Ambrosone CB, Kadlubar FF (2002) Carcinogen-DNA adducts in human breast epithelial cells. Environ Mol Mutagen 39:184–192
18. Gu D, Turesky RJ, Tao Y, Langouet SA, Nauwelaers GC, Yuan JM et al (2012) DNA adducts of 2-amino-1-methyl-6-phenylimidazo[4,5-*b*]pyridine and 4-aminobiphenyl are infrequently detected in human mammary tissue by liquid chromatography/tandem mass spectrometry. Carcinogenesis 33:124–130
19. Guengerich FP (2003) Activation of dihaloalkanes by thiol-dependent mechanisms. J Biochem Mol Biol 36:20–27
20. Guengerich FP, Johnson WW, Shimada T, Ueng YF, Yamazaki H, Langouet S (1998) Activation and detoxification of aflatoxin B1. Mutat Res 402:121–128
21. Hanahan D, Weinberg RA (2000) The hallmarks of cancer. Cell 100:57–70
22. International Agency for Research on Cancer (2008) 1,3-butadiene, ethylene oxide and vinyl halides (vinyl fluoride, vinyl chloride and vinyl bromide), vol 97. International Agency for Research on Cancer, Lyon
23. International Agency for Research on Cancer (2007) Smokeless tobacco and some tobacco-specific N-nitrosamines, vol 89. International Agency for Research on Cancer, Lyon
24. International Agency for Research on Cancer (2010) Some aromatic amines, organic dyes, and related exposures, vol 99. International Agency for Research on Cancer, Lyon
25. International Agency for Research on Cancer (2010) Some non-heterocyclic polycyclic aromatic hydrocarbons and some related exposures, vol 92. International Agency for Research on Cancer, Lyon
26. Kim JH, Stansbury KH, Walker NJ, Trush MA, Strickland PT, Sutter TR (1998) Metabolism of benzo[*a*]pyrene and benzo[*a*]pyrene-7,8-diol by human cytochrome P450 1B1. Carcinogenesis 19:1847–1853
27. Klaene JJ, Sharma VK, Glick J, Vouros P (2012) The analysis of DNA adducts: the transition from ^{32}P-postlabeling to mass spectrometry. Cancer Lett 334:10–19
28. Ma L, Kuhlow A, Glatt HR (2002) Ethanol enhances the activation of 1-hydroxymethylpyrene to DNA adduct-forming species in the rat. Polycycl Aromat Comp 22:933–946
29. Matsuda T, Tao H, Goto M, Yamada H, Suzuki M, Wu Y et al (2013) Lipid peroxidation-induced DNA adducts in human gastric mucosa. Carcinogenesis 34:121–127
30. Monien BH, Engst W, Barknowitz G, Seidel A, Glatt HR (2012) Mutagenicity of 5-hydroxymethylfurfural in V79 cells expressing human SULT1A1: identification and mass spectrometric quantification of DNA adducts formed. Chem Res Toxicol 25: 1484–1492
31. Monien BH, Herrmann K, Florian S, Glatt HR (2011) Metabolic activation of furfuryl alcohol: formation of 2-methylfuranyl DNA adducts in *Salmonella typhimurium* strains expressing human sulfotransferase 1A1 and in FVB/N mice. Carcinogenesis 32:1533–1539
32. Monien BH, Müller C, Bakhiya N, Donath C, Frank H, Seidel A et al (2009) Probenecid, an inhibitor of transmembrane organic anion transporters, alters tissue distribution of DNA adducts in 1-hydroxymethylpyrene-treated rats. Toxicology 262:80–85
33. Monien BH, Müller C, Engst W, Frank H, Seidel A, Glatt HR (2008) Time course of hepatic 1-methylpyrene DNA adducts in rats determined by isotope dilution LC-MS/MS and ^{32}P-postlabeling. Chem Res Toxicol 21:2017–2025

34. Nair J, Carmichael PL, Fernando RC, Phillips DH, Strain AJ, Bartsch H (1998) Lipid peroxidation-induced etheno-DNA adducts in the liver of patients with the genetic metal storage disorders Wilson's disease and primary hemochromatosis. Cancer Epidemiol Biomarkers Prev 7:435–440
35. Newcomb PA, Solomon C, White E (1999) Tamoxifen and risk of large bowel cancer in women with breast cancer. Breast Cancer Res Treat 53:271–277
36. Otteneder M, Lutz WK (1999) Correlation of DNA adduct levels with tumor incidence: carcinogenic potency of DNA adducts. Mutat Res 424:237–247
37. Paini A, Scholz G, Marin-Kuan M, Schilter B, O'Brien J, van Bladeren PJ et al (2011) Quantitative comparison between in vivo DNA adduct formation from exposure to selected DNA-reactive carcinogens, natural background levels of DNA adduct formation and tumour incidence in rodent bioassays. Mutagenesis 26:605–618
38. Perera F, Mayer J, Jaretzki A, Hearne S, Brenner D, Young TL et al (1989) Comparison of DNA adducts and sister chromatid exchange in lung cancer cases and controls. Cancer Res 49:4446–4451
39. Phillips DH (2002) Smoking-related DNA and protein adducts in human tissues. Carcinogenesis 23:1979–2004
40. Phillips DH, Castegnaro M (1999) Standardization and validation of DNA adduct postlabelling methods: report of interlaboratory trials and production of recommended protocols. Mutagenesis 14:301–315
41. Poirier MC (2004) Chemical-induced DNA damage and human cancer risk. Nat Rev Cancer 4:630–637
42. Poirier MC, Beland FA (1992) DNA adduct measurements and tumor incidence during chronic carcinogen exposure in animal models: implications for DNA adduct-based human cancer risk assessment. Chem Res Toxicol 5:749–755
43. Poirier MC, Stanley JR, Beckwith JB, Weinstein IB, Yuspa SH (1982) Indirect immunofluorescent localization of benzo[a]pyrene adducted to nucleic acids in cultured mouse keratinocyte nuclei. Carcinogenesis 3:345–348
44. Punt A, Delatour T, Scholz G, Schilter B, van Bladeren PJ, Rietjens IM (2007) Tandem mass spectrometry analysis of N^2-(*trans*-isoestragol-3′-yl)-2′-deoxyguanosine as a strategy to study species differences in sulfotransferase conversion of the proximate carcinogen 1′-hydroxyestragole. Chem Res Toxicol 20:991–998
45. Qian GS, Ross RK, Yu MC, Yuan JM, Gao YT, Henderson BE et al (1994) A follow-up study of urinary markers of aflatoxin exposure and liver cancer risk in Shanghai, People's Republic of China. Cancer Epidemiol Biomarkers Prev 3:3–10
46. Randerath K, Reddy MV, Gupta RC (1981) ^{32}P-labeling test for DNA damage. Proc Natl Acad Sci U S A 78:6126–6129
47. Reddy MV (2000) Methods for testing compounds for DNA adduct formation. Regul Toxicol Pharmacol 32:256–263
48. Rendic S, Guengerich FP (2012) Contributions of human enzymes in carcinogen metabolism. Chem Res Toxicol 25:1316–1383
49. Ricicki EM, Soglia JR, Teitel C, Kane R, Kadlubar F, Vouros P (2005) Detection and quantification of N-(deoxyguanosin-8-yl)-4-aminobiphenyl adducts in human pancreas tissue using capillary liquid chromatography-microelectrospray mass spectrometry. Chem Res Toxicol 18:692–699
50. Shou M, Korzekwa KR, Crespi CL, Gonzalez FJ, Gelboin HV (1994) The role of 12 cDNA-expressed human, rodent, and rabbit cytochromes P450 in the metabolism of benzo[*a*]pyrene and benzo[*a*]pyrene *trans*-7,8-dihydrodiol. Mol Carcinog 10:159–168
51. Singh R, Gaskell M, Le Pla RC, Kaur B, Azim-Araghi A, Roach J et al (2006) Detection and quantitation of benzo[a]pyrene-derived DNA adducts in mouse liver by liquid chromatography-tandem mass spectrometry: comparison with ^{32}P-postlabeling. Chem Res Toxicol 19:868–878
52. Singh R, Teichert F, Seidel A, Roach J, Cordell R, Cheng MK et al (2010) Development of a targeted adductomic method for the determination of polycyclic aromatic hydrocarbon DNA adducts using online column-switching liquid chromatography/tandem mass spectrometry. Rapid Commun Mass Spectrom 24:2329–2340

53. Smith B, Cadby P, Leblanc JC, Setzer RW (2010) Application of the margin of exposure (MoE) approach to substances in food that are genotoxic and carcinogenic: example: methyleugenol, CASRN: 93-15-2. Food Chem Toxicol 48(suppl 1):S89–S97
54. Spodheim-Maurizot M, Saint-Ruf G, Leng M (1980) Antibodies to N-hydroxy-2-aminofluorene modified DNA as probes in the study of DNA reacted with derivatives of 2-acetylaminofluorene. Carcinogenesis 1:807–812
55. Tang D, Phillips DH, Stampfer M, Mooney LA, Hsu Y, Cho S et al (2001) Association between carcinogen-DNA adducts in white blood cells and lung cancer risk in the physicians health study. Cancer Res 61:6708–6712
56. Vineis P, Husgafvel-Pursiainen K (2005) Air pollution and cancer: biomarker studies in human populations. Carcinogenesis 26:1846–1855
57. Zhang S, Balbo S, Wang M, Hecht SS (2010) Analysis of acrolein-derived 1, N2-propanodeoxyguanosine adducts in human leukocyte DNA from smokers and nonsmokers. Chem Res Toxicol 24:119–124
58. Zhu J, Chang P, Bondy ML, Sahin AA, Singletary SE, Takahashi S et al (2003) Detection of 2-amino-1-methyl-6-phenylimidazo[4,5-*b*]-pyridine-DNA adducts in normal breast tissues and risk of breast cancer. Cancer Epidemiol Biomarkers Prev 12:830–837

Chapter 19
Using Breast Milk to Assess Breast Cancer Risk: The Role of Mass Spectrometry-Based Proteomics

Sallie S. Schneider, Roshanak Aslebagh, Armand G. Ngounou Wetie, Susan R. Sturgeon, Costel C. Darie, and Kathleen F. Arcaro

Abstract Although mammography and treatment advances have led to declines in breast cancer mortality in the United States, breast cancer remains a major cause of morbidity and mortality. Breast cancer in young women is associated with increased mortality and current methods of detecting breast cancers in this group of women have known limitations. Tools for accurately assessing personal breast cancer risk in young women are needed to identify those women who would benefit the most from earlier intervention. Proteomic analysis of breast milk could identify biomarkers of breast cancer risk and provide a tool for identifying women at increased risk. A preliminary analysis of milk from four women provides a proof of concept for using breast milk to assess breast cancer risk.

S.S. Schneider
Pioneer Valley Life Sciences Institute, Springfield, MA 01107, USA

R. Aslebagh • A.G. Ngounou Wetie • C.C. Darie
Biochemistry and Proteomics Group, Department of Chemistry and Biomolecular Science, Clarkson University, 8 Clarkson Avenue, Potsdam, NY 13699-5810, USA

S.R. Sturgeon
Division of Biostatistics and Epidemiology, School of Public Health and Health Sciences, University of Massachusetts Amherst, Amherst, MA 01003, USA

K.F. Arcaro (✉)
Department of Veterinary and Animal Sciences, University of Massachusetts Amherst, Amherst, MA 01003, USA
e-mail: karcaro@vasci.umass.edu

A.G. Woods and C.C. Darie (eds.), *Advancements of Mass Spectrometry in Biomedical Research*, Advances in Experimental Medicine and Biology 806,
DOI 10.1007/978-3-319-06068-2_19, © Springer International Publishing Switzerland 2014

19.1 Introduction

19.1.1 Overview

Although mammography and treatment advances have led to declines in breast cancer mortality in the United States, breast cancer remains a major cause of morbidity and mortality, and racial disparities in the burden of this disease exist. Mammography has greatly increased the early detection of breast cancer, but has known limitations, especially in younger women with dense breasts. Breast milk provides access to breast tissue, in the form of exfoliated epithelial cells, that when combined with mass spectrometry-based proteomics may offer a novel way to identify young women at increased risk of developing breast cancer leading to earlier detection and improved outcomes.

19.1.2 Breast Cancer Incidence and Mortality

Nearly 1.4 million women across the world were diagnosed with breast cancer in 2008 and it is predicted that the number of cases will rise to 2.1 million by 2030. The highest age-adjusted incidence rates occur in westernized countries and are attributed to industrialization, urbanization, and access to mammography [1]. In the United States, about 232,000 new cases of invasive breast cancer were diagnosed in 2013. In the same year, approximately 39,620 US women were expected to die from breast cancer [2]. The current lifetime risk is 12.3 %, with 1 in 8 women expected to develop breast cancer in their lifetime.

In the United States, breast cancer mortality rates decreased by 34 % between 1990 and 2010 [2]. This decline in mortality is attributed to improvements in treatment and early detection. The benefits of these medical advances have not been observed equally among all groups of women. In 2010, breast cancer mortality among African American women was 41 % higher than in White women, despite the fact that the overall incidence of breast cancer is lower among Black women [2]. Lack of access to both mammographic screening and high quality care explain much of the racial differences in death rates, but differences in breast cancer tumor characteristics may also contribute to the higher mortality among Black women.

19.1.3 Breast Cancer in Young Women and Pregnancy-Associated Breast Cancer

Approximately 12 % of all breast cancers are diagnosed in women under 45 years of age [3]. Because of the density of the breast tissue in young women, mammography is less effective in detecting tumors in younger women. The breast tumors that develop

in young women tend to be aggressive, lack expression of the estrogen receptor, and are more likely to be associated with poorer survival compared to tumors that develop in older women. In addition, incidence rates among young Black women (less than 40 years old) are higher than those among similarly-aged White women [2]. Epidemiology studies show that giving birth at a young age, having multiple live births and prolonged breastfeeding provide protection against developing breast cancer [4, 5]. However, paradoxically, the hormonal changes associated with pregnancy result in a transient (about 5 years) increase in breast cancer risk [6, 7]. In summary, breast cancer in younger women remains a significant challenge for detection and treatment.

19.1.4 Limits of Mammography and Risk Assessment

Current guidelines of the U.S. Preventive Services Task Force recommend a mammogram every 2 years for all women between the ages of 50 and 74. Routine mammography screening is not recommended for younger women [8]. Despite its recognized value, mammography has limitations including false positives, false negatives, and over diagnosis. There is also a need for improved risk assessment. Personalized risk assessment would provide women the opportunity to make informed decisions regarding various interventions, such as Tamoxifen. In recent years, considerable effort has been directed towards identifying molecular markers of breast cancer risk and early disease; however, the collection of breast tissue for analysis remains a major obstacle. Methods of obtaining breast tissue from non-symptomatic women include collection of nipple aspirate, ductal lavage, and fine needle aspirate. While these methods have provided valuable information regarding breast biology, they have several limitations. Ductal lavage and fine needle aspirate are invasive and do not survey all of the ducts and lobes of a breast and therefore may not provide a sufficient screen of the breast limiting their value for predicting risk. Collection of nipple aspirate fluid is minimally invasive but about half the women sampled are determined to be non-secretors and half of the samples collected may not contain epithelial cells [9, 10] Access to the breast tissue present in breast milk of young women could likely provide a means of identifying molecular markers associated with breast cancer risk.

19.1.5 Breast Milk

To date most of the research on the composition of human milk has been aimed at understanding the effects of breast milk on the growth and development of the infant, the biology of lactation, and the effects of mastitis [11]. Prior proteome analyses have focused on the proteins secreted into the milk, including growth factors, cytokines, and cleaved membrane proteins, and have documented changes in the

content of immune-related proteins with length of lactation; from colostrum through mature milk to weaning [12, 13]. A few studies have used ELISA methods to examine the relationship of a limited number of secreted proteins with breast cancer risk factors [14, 15]. However, the biomarkers most likely to be important for assessing breast cancer risk may be missed if only secreted proteins are examined, as the primary function of secreted proteins is to aid in the development of the infant. Therefore, it will be important to include intracellular proteins of the breast epithelial cells in proteome analyses of breast milk aimed at assessing breast cancer risk.

Milk droplets are produced in epithelial cells in the lobules of the gland and travel down the ductal system to the nipple, collecting exfoliated epithelial cells along the way. The intracellular proteins of the exfoliated epithelial cells could provide information on changes in protein expression related to risk. In addition to the relatively small percentage of epithelial cells, breast milk contains numerous leukocytes [16] with their own array of intracellular proteins. Thus breast milk provides an excellent sampling of intracellular, secreted, and vesicle cell-derived proteins from the entire mammary gland.

19.1.6 Epigenetic Markers of Risk in Breast Milk

To date, the use of breast milk for risk assessment and early diagnosis is limited to a few studies. For example we recently examined exfoliated epithelial cells isolated from breast milk for epigenetic signals associated with increased breast cancer risk [16, 17]. Comparisons of DNA methylation between women at average and increased risk revealed greater DNA methylation of several tumor suppressor genes in DNA from milk cells of women at increased risk [17]. These results provide proof of concept for the use of breast milk in an assay to assess risk, and are therefore encouraging, however, DNA methylation is only one way in which transcription and hence protein production is controlled. Therefore, an assay based directly on protein biomarkers could provide improved risk assessment.

19.1.7 Mass Spectrometry (MS)-Based Proteomics Analysis of Breast Milk for Assessing Breast Cancer Risk

Proteomics is an emerging approach that analyzes the proteome, i.e., the complete protein array or complement in a specific protein sample, such as the intracellular or extracellular protein complement, soluble, or membrane protein complement, or the cellular, tissue, organ or organism protein complement [18–22]. A variety of tissues such as biopsies of tumor and surrounding tissue or biological fluids such as blood, saliva, urine, tears, bronchial secretions, or breast milk can be used to conduct discovery-based proteomic research. Protein samples analyzed by proteomic approaches from different groups of individuals (for example blood from

Table 19.1 Demographics of the four African American women who provided breast milk samples

Sample	Current age (years)	Age at first birth	Pregnancies	Live births	Baby's age (days)	Family history of BC
1	32	32	1	1	44	No
2	34	34	1	1	150	No
3	36	22	6	5	136	No
4	29	27	2	2	356	No

healthy controls and breast cancer cases), can be compared in quantitative terms through identification of the proteomes in each sample, and through protein quantification, protein post-translational modifications, and protein–protein interactions [21, 23–25]. Such comparisons can identify a single protein or a protein signature that is specific to one condition, which can be used to assess future risk of developing a disease. Ideally, this type of case–control analysis should be nested in a large prospective cohort study to allow discrimination of a proteome risk signature from a signature of the disease itself. The general proteomic approach could also be used for the assessment of treatment outcomes, as well as for the prognosis of disease progression.

In a pilot study we explored proteomic analysis of whole milk using MS with a focus on identifying proteins that could be associated with breast cancer risk. For this preliminary study we selected milk from four young African American women (see Table 19.1). Participants were between the ages of 29–34 years at the time of milk collection. Age at first birth ranged from 22 to 34 years old. Two women were nursing their first child, one woman was nursing her second, and the fourth woman was nursing her fifth child. Baby's age, a surrogate for length of lactation, ranged from 44 to 356 days. None of the women had a family history of breast cancer.

Details of the biochemical fractionation of the breast milk samples used in the present study are described elsewhere [13]. Briefly, 2 mL of whole milk from each participant was directly subjected to ultracentrifugation, and 150 μL of supernatant was run on an SDS-PAGE gel. Because the ultracentrifugation step was not preceded by a slow centrifugation step, which is often used to deplete milk samples of cellular content, it is likely that the current analysis included a large number of intracellular proteins released from both the leukocytes and epithelial cells during the ultracentrifugation, as well as the more abundant secreted proteins.

Figure 19.1 presents the SDS-PAGE gel for the four samples. As can be seen in the gel, the overall pattern of most of the major bands is similar among all four samples and bands that include some of the high abundance proteins (e.g., lactotransferrin at 78 kDa) are clearly present. Sample specific-bands are also present. The red arrows in Fig. 19.1 mark bands that differ among the four samples. For example, samples 1 and 2 have a high molecular weight band (arrow a) that is missing from the other two samples, and sample 3 is lacking or has reduced levels at several bands present in the other samples (arrows e, f, and g).

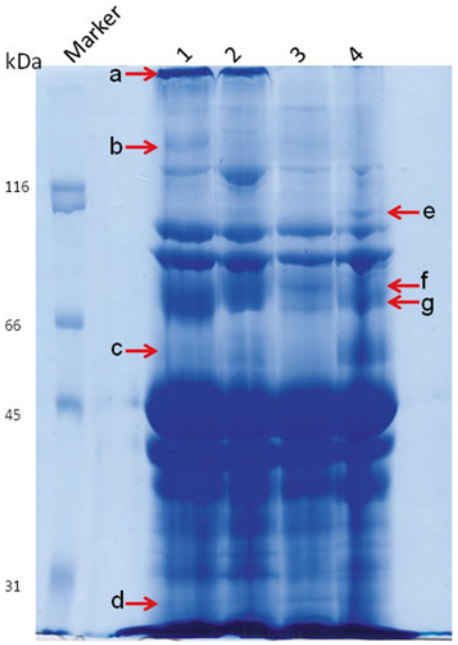


Fig. 19.1 SDS-PAGE gel showing protein band profiles from four human milk samples demonstrate differences among women. *Red arrows* indicate differences in protein bands among the samples

To identify proteins, the bands for each sample were cut into ten pieces, and enzymatically digested and run on nanoliquid chromatography-tandem mass spectrometry according to a published procedure [24]. Scaffold 4 was used to identify proteins using the stringent setting of 99 % probability of correctly identifying a protein from a minimum of two peptides. A relatively small number of proteins, about 100, were identified in each sample and there was considerable similarity among samples.

To gain an initial understanding of the role of the identified proteins, we merged the protein datasets and used pathway software to provide a descriptive analysis. Thus, the data in Figs. 19.2 and 19.3 are based on a compilation of proteins from all four samples and are presented to demonstrate the range of biological processes represented by the identified proteins (Fig. 19.2), and the cellular component to which the proteins belong (Fig. 19.3). With respect to biological processes, the general categories of proteins in Fig. 19.2 are very similar to the proteome signature of milk from five Italian women that was previously reported by Picariello and colleagues [13] suggesting that this method produces similar results and in general the milk proteome is qualitatively similar. However, the inclusion of the cells resulted in an increased intracellular and nuclear component. Twenty-three of the identified proteins locate to the extracellular region, 23 to the intracellular organelle, and 14 to the nucleus (Fig. 19.3), suggesting that both proteins secreted into the milk and proteins from the cells in the milk were detected with the mass spectrometry analysis.

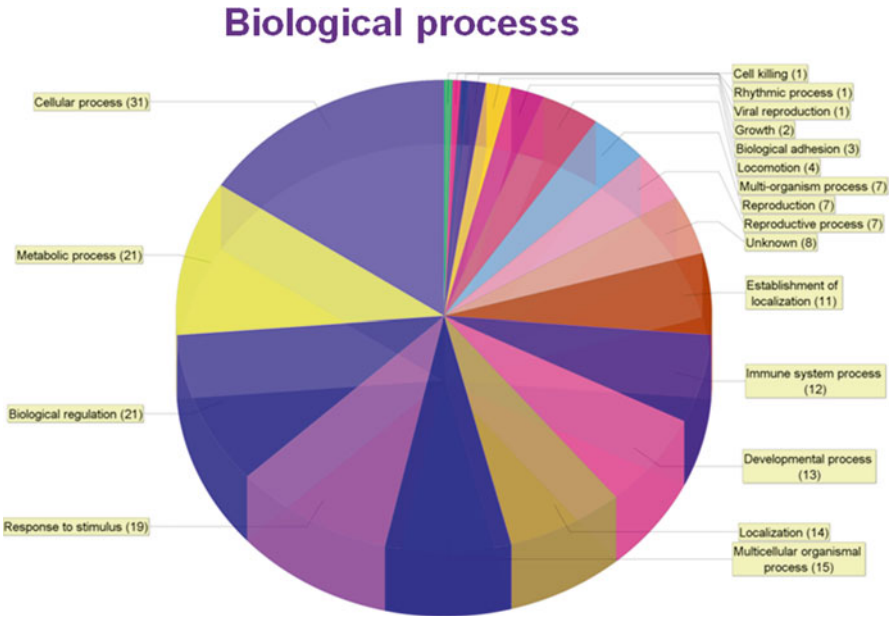


Fig. 19.2 Pie chart showing the distribution of proteins detected in the four milk samples based on biological process

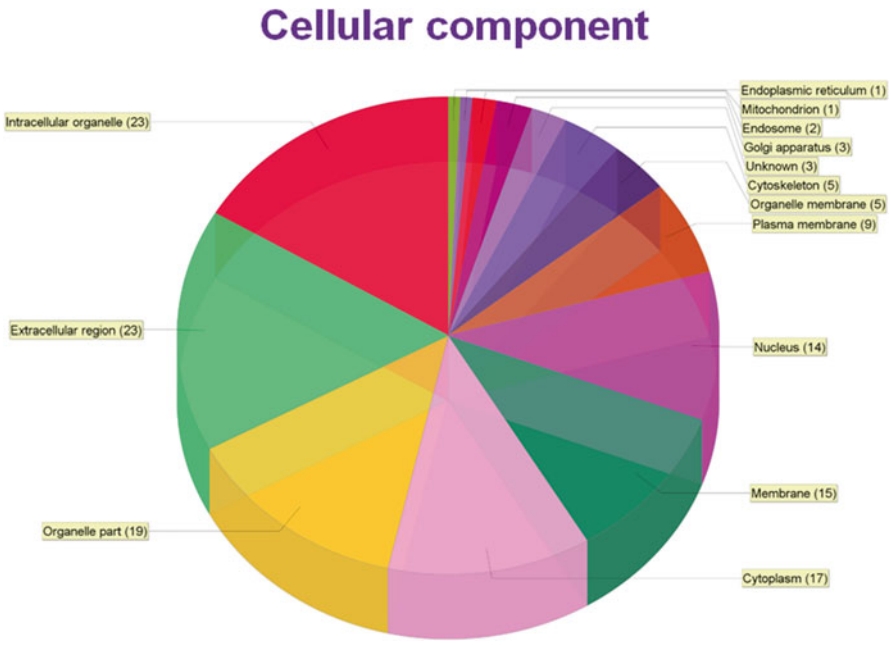


Fig. 19.3 Pie chart showing the distribution of proteins detected in the four milk samples based on cellular component

Comparison of the lower abundance proteins from the individual MS analyses for each of the four milk samples revealed some interesting preliminary differences that require further verification. In particular, the protein spectra of milk from participant 2 included proteins absent from the other milk samples and not previously identified in other milk proteome studies [13] that could be biomarkers of risk. Mucin 16 (previously known as CA125), a membrane-associated protein and known cancer biomarker [26] was identified by two peptides in sample 2 and not in the other three samples. Mucin 16 is involved in regulating cell adhesion which has led to its use as a biomarker for resectability in epithelial ovarian cancer [27]. The milk from subject 2 also contained detectable Axin2, an intracellular protein, and a marker of activated Wnt/β-catenin signaling. This finding is of particular interest because our previous analysis of epithelial cells from breast milk revealed DNA promoter methylation of a critical Wnt pathway antagonist, SFRP1, in women at increased risk of developing breast cancer [17]. Promoter methylation of SFRP1 leads to decreased SFRP1 protein levels which we have observed to result in enhanced Axin2 expression (S. Schneider, personal communication), increased proliferation, and decreased sensitivity to cell death in an immortalized human mammary epithelial cell line [28]. While these results are preliminary they support the hypothesis that analysis of proteins in breast milk may provide information regarding breast cancer risk.

19.2 Conclusion

This preliminary proteomic analysis of breast milk focused on detecting proteins associated with breast cancer risk and therefore both the secreted proteins and cellular proteins were included, as we did not remove the cellular fraction. Proteins previously identified with breast cancer were detected in one sample supporting the potential of a proteomic analysis of breast milk for detecting increased breast cancer risk. While encouraging, significant improvements in methodology are required as the number of proteins detected was low. Methods that deplete the sample of high abundance proteins without removing the low abundance proteins must be optimized. Other challenges for exploiting breast milk and proteomics for early detection and risk estimate relate to the establishment of a large prospective cohort of young women with stored milk specimens for a nested case–control study to robustly examine the usefulness of the proteomic approach. Special attention should be given to Black women who are at increased risk of developing breast cancer at a young age and to those with a family history of the disease.

Acknowledgements The authors thank Eva P. Browne and Elizabeth C. Punska for their help with milk collection and proofreading the manuscript. This work was supported by an award from the Avon Foundation for Women to KFA and awards from the Keep a Breast Foundation (KEABF-375-35054), the David A. Walsh fellowship, and the U.S. Army research office (DURIP grant #W911NF-11-1-0304) to CCD.

References

1. Breast Cancer Worldwide. http://www.wcrf.org/cancer_statistics/cancer_facts/women-breast-cancer.php
2. American Cancer Society (2013) Breast cancer facts and figures 2013–2014. American Cancer Society, Atlanta. http://www.cancer.org/research/cancerfactsfigures/breastcancerfactsfigures/
3. SEER cancer statistics factsheets: breast cancer. National Cancer Institute, Bethesda. http://seer.cancer.gov/statfacts/html/breast.html
4. NCI—reproductive history and breast cancer risk. http://www.cancer.gov/cancertopics/factsheet/Risk/reproductive-history
5. González-Jiménez E, García PA, Aguilar MJ, Padilla CA, Alvarez J (2013) Breastfeeding and the prevention of breast cancer: a retrospective review of clinical histories. http://www.ncbi.nlm.nih.gov/pubmed/23937211 "J Clin Nurs. 2013 Aug 13. doi: 10.1111/jocn.12368. [Epub ahead of print]
6. Borges VF, Schedin PJ (2012) Pregnancy-associated breast cancer: an entity needing refinement of the definition. Cancer 118(13):3226–3228
7. Callihan EB, Gao D, Jindal S, Lyons TR, Manthey E, Edgerton S, Urquhart A, Schedin P, Borges VF (2013) Postpartum diagnosis demonstrates a high risk for metastasis and merits an expanded definition of pregnancy-associated breast cancer. Breast Cancer Res Treat 138(2):549–559
8. Screening for breast cancer, topic page (2010) U.S. Preventive Services Task Force. http://www.uspreventiveservicestaskforce.org/uspstf/uspsbrca.htm
9. Paweletz CP, Trock B, Pennanen M, Tsangaris T, Magnant C, Liotta LA, Petricoin EF (2001) Proteomic patterns of nipple aspirate fluids obtained by SELDI-TOF: potential for new biomarkers to aid in the diagnosis of breast cancer. Dis Markers 17(4):301–307
10. Filassi JR, Zonta MA, Trinconi A, Calvagno D, Velame de Oliveira F, Ricci MD, Baracat E, Longatto-Filho A (2013) Can breast nipple fluid collected with automated aspiration and preserved in based-liquid solution improve the cytological samples? Acta Cytol 57(3):276–280
11. Ballard O, Morrow AL (2013) Human milk composition: nutrients and bioactive factors. Pediatr Clin North Am 60(1):49–74
12. Roncada P, Stipetic LH, Bonizzi L, Burchmore RJ, Kennedy MW (2013) Proteomics as a tool to explore human milk in health and disease. J Proteomics 88:47–57
13. Picariello G, Ferranti P, Mamone G, Klouckova I, Mechref Y, Novotny MV, Addeo F (2012) Gel-free shotgun proteomic analysis of human milk. J Chromatogr A 1227:219–233
14. Arcaro KF, Browne EP, Qin W, Zhang K, Anderton DL, Sauter ER (2012) Differential expression of cancer-related proteins in paired breast milk samples from women with breast cancer. J Hum Lact 28(4):543–546
15. Qin W, Zhang K, Kliethermes B, Ruhlen R, Browne E, Arcaro K, Sauter E (2012) Differential expression of cancer associated proteins in breast milk based on age at first full term pregnancy. BMC Cancer 12(1):100
16. Wong C, Anderton D, Smith Schneider S, Wing M, Greven M, Arcaro K (2010) Quantitative analysis of promoter methylation in exfoliated epithelial cells isolated from breast milk of healthy women. Epigenetics 5(7):645–655
17. Browne E, Punska E, Lenington S, Otis C, Anderton D, Arcaro K (2011) Increased promoter methylation in exfoliated breast epithelial cells in women with a previous breast biopsy. Epigenetics 6(12):1425–1435
18. Darie C (2013) Mass spectrometry and its applications in life sciences. Aust J Chem 66(7):719–720
19. Ngounou Wetie AG, Sokolowska IJ, Woods AG, Darie CC (2013) Identification of post-translational modifications by mass spectrometry. Aust J Chem 66(7):734–748
20. Sokolowska I, Ngounou Wetie AG, Woods AG, Darie CC (2013) Applications of mass spectrometry in proteomics. Aust J Chem 66(7):721–733

21. Sokolowska I, Ngounou Wetie AG, Roy U, Woods AG, Darie CC (2013) Mass spectrometry investigation of glycosylation on the NXS/T sites in recombinant glycoproteins. Biochim Biophys Acta 1834(8):1474–1483
22. Woods AG, Sokolowska I, Yakubu R, Butkiewicz M, LaFleur M (2011) Blue native PAGE and mass spectrometry as an approach for the investigation of stable and transient protein-protein interactions. ACS Symp Ser 1083:341–368
23. Ngounou Wetie AG, Sokolowska I, Woods AG, Roy U, Deinhardt K, Darie CC (2014) Protein-protein interactions: switch from classical methods to proteomics and bioinformatics-based approaches. Cell Mol Life Sci 71(2):205–228
24. Ngounou Wetie AG, Sokolowska I, Woods AG, Wormwood KL, Dao S, Patel S, Clarkson BD, Darie CC (2013) Automated mass spectrometry-based functional assay for the routine analysis of the secretome. J Assoc Lab Autom 18(1):19–29
25. Woods AG, Sokolowska I, Darie CC (2012) Identification of consistent alkylation of cysteine-less peptides in a proteomics experiment. Biochem Biophys Res Commun 419(2):305–308
26. Mukhopadhyay P, Chakraborty S, Ponnusamy MP, Lakshmanan I, Jain M, Batra SK (2011) Mucins in the pathogenesis of breast cancer: implications in diagnosis, prognosis and therapy. Biochim Biophys Acta 1815(2):224–240
27. Fleming ND, Cass I, Walsh CS, Karlan BY, Li AJ (2011) CA125 surveillance increases optimal resectability at secondary cytoreductive surgery for recurrent epithelial ovarian cancer. Gynecol Oncol 121(2):249–252
28. Gauger KJ, Hugh JM, Troester MA, Schneider SS (2009) Down-regulation of sfrp1 in a mammary epithelial cell line promotes the development of a cd44high/cd24low population which is invasive and resistant to anoikis. Cancer Cell Int 9:11

Chapter 20
Cancer Secretomes and Their Place in Supplementing Other Hallmarks of Cancer

Sapan Patel, Armand G. Ngounou Wetie, Costel C. Darie, and Bayard D. Clarkson

Abstract The secretome includes all macromolecules secreted by cells, in particular conditions at defined times, allowing cell–cell communication. Cancer cell secretomes that are altered compared to normal cells have shown significant potential for elucidating cancer biology. Proteins of secretomes are secreted by various secretory pathways and can be studied using different methods. Cancer secretomes seem to play an important role in known hallmarks of cancers such as excessive proliferation, reduced apoptosis, immune invasion, angioneogenesis, alteration in energy metabolism, and development of resistance against anti-cancer therapy [1, 2]. If a significant role of an altered secretome can be identified in cancer cells, using advanced mass spectrometry-based techniques, this may allow researchers to screen and characterize the secretome proteins involved in cancer progression and open up new opportunities to develop new therapies. We aim to elaborate upon recent advances in cancer cell secretome analysis using different proteomics techniques. In this review, we highlight the role of the altered secretome in contributing to already recognized and emerging hallmarks of cancer and we discuss new challenges in the field of secretome analysis.

S. Patel
Memorial Sloan Kettering Cancer Center, Molecular Pharmacology and Chemistry Program, 415 East 68th Street, New York, NY 10065, USA

Biochemistry and Proteomics Group, Department of Chemistry and Biomolecular Science, Clarkson University, 8 Clarkson Avenue, Postdam, NY 13699-5810, USA

A.G. Ngounou Wetie • C.C. Darie
Biochemistry and Proteomics Group, Department of Chemistry and Biomolecular Science, Clarkson University, 8 Clarkson Avenue, Postdam, NY 13699-5810, USA

B.D. Clarkson (✉)
Memorial Sloan Kettering Cancer Center, Molecular Pharmacology and Chemistry Program, 415 East 68th Street, New York, NY 10065, USA

Molecular Pharmacology and Chemistry Program, Sloan-Kettering Institute for Cancer Research, 1275 York Avenue, Box 96, New York, NY 10021, USA
e-mail: clarksob@mskcc.org

A.G. Woods and C.C. Darie (eds.), *Advancements of Mass Spectrometry in Biomedical Research*, Advances in Experimental Medicine and Biology 806,
DOI 10.1007/978-3-319-06068-2_20, © Springer International Publishing Switzerland 2014

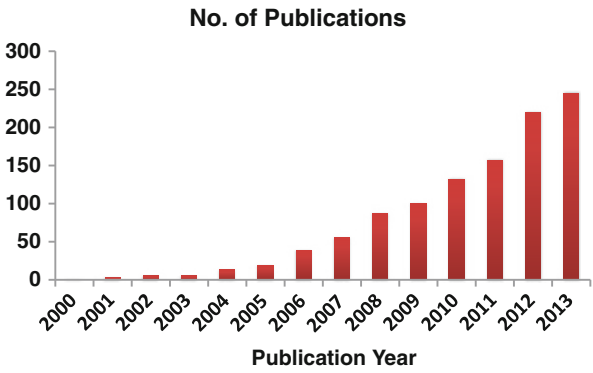


Fig. 20.1 Number of publications describing the secretome. *Bar graph* showing the number of articles describing secretome studies from 2000, when it was first described, to 2013. NCBI Pubmed Central, with a query of "secretome" and publication year, was used to obtain data

20.1 The Secretome

Secretome is the term used to describe proteins or other molecules secreted by cells into the extracellular compartment. Tjalsma et al. first used the term "secretome" in a genome-based study of secretory proteins in *Bacillus subtilis* bacteria [3]. The original intention to use the term secretome was to address secretory pathway constituents and secreted factors. Approximately 10 % of the human genome constitutes secretome proteins. The definition proposed by Agrawal et al. is "A global group of proteins secreted into the extracellular space by a cell, tissue, organ, or organism at any given time and condition through known and unknown constitutive mechanisms regulating secretory organelles [4]." These secreted proteins are an important class of molecules involving cell motility factors, growth factors, extracellular proteases, cytokines, anti-survival proteins, angiogenic factors, and other bioactive molecules. They are involved in various pathophysiological functions such as cell differentiation, invasion, metastasis, autophagy, apoptosis, tissue organization, immune surveillance, angiogenesis, and cell-to-cell communication of cancer cells. There are already many reviews focusing on identification of various secretome constituents in many diseases including obesity, diabetes, inflammatory diseases, AIDS, and caners. Since the year 2000, when the secretome term was first used, to the year 2013, the number of articles describing secretome studies has been progressively rising as shown in Fig. 20.1.

20.2 Protein Secretion Pathways

It is first necessary to understand the protein secretion pathway before understanding its role in physiological functions and the methods used for secretome analysis. These are the pathways responsible not only for secreting an ample amount of protein in the extracellular microenvironment of normal cells but also in cancer cells. In eukaryotic cells, the proteins released by cells and/or tissues are released by both classical and non-classical secretion mechanisms.

20.2.1 Classical Secretory Pathway

In the classical secretion mechanism, also known as the conventional secretion pathway, upon stimulation and activation of intracellular signaling pathways the storage granules or vesicles are released into the extracellular space. This process is known as exocytosis in which newly synthesized proteins are translocated into the lumen of endoplasmic reticulum (ER), transported through the Golgi complex and released into the extracellular space. The signaling peptide sequence inherently present in the messenger RNAs encoding these proteins plays an important role by guiding newly synthesized proteins from ribosomes to the ER membrane during translation and initiates the transport of the growing polypeptide across the ER membrane into ER lumen. A typical ER signal sequence contains an amino terminal with one or more positively charged amino acids followed by 6–12 hydrophobic residues. Usually the secreted proteins are synthesized as protein precursors and the N-terminal signal peptide sequence is cleaved from the protein when the polypeptide chain is continuing to grow on the ribosome [5, 6].

20.2.2 Non-classical Secretory Pathway

Some of the proteins are also secreted into the extracellular space using an unconventional method. Many soluble proteins lack a typical signal peptide and they are transported from the cytoplasm or nucleus using unconventional secretion mechanisms. Proteins secreted by non-classical secretory pathways lack the N-terminal signal peptide and are secreted into the ECM (extra cellular matrix) by ER-to-Golgi independent secretion pathways. These pathways are categorized into two major categories that are further divided into four categories. We will describe the principle mechanism of four unconventional protein secretion pathways in detail.

20.2.2.1 Non-vesicular Transport

Direct Translocation of Cytoplasmic Proteins Across the Plasma Membrane

The type-1 unconventional secretions translocate the proteins without using any vesicular intermediates. The secretion of fibroblast growth factor 2 (FGF2), fibroblast growth factor 1 (FGF1), and Annexin A2 (ANXA2) are some of the examples by which proteins are secreted into the extracellular space by this mechanism.

Release of FGF2 involves interaction of FGF2 with the phosphoinositide phosphatidylinositol (4,5)-biophosphate (PIP4, 5) at the inner cell membrane and with heparin sulfate at the outer plasma membrane. This interaction results in the formation of PIP4,5-induced oligomerization at the inner membrane. This process is dependent on the active confirmation of FGF2 and also stimulated further by

phosphorylation of FGF2 at tyrosine residue 82. Oligomerization of FGF2 by this mechanism prevents the loss of membrane potential and loss of lipid asymmetry that allows translocation of protein to the cell surface [7–11].

ABC Transporter-Based Secretion

This is also a non-vesicular mode of protein translocation. These pathways translocate lapidated peptide and proteins across the plasma membrane. There are several examples by which farnesylated, geranylated, and acylated form of proteins are exported by the ABC transporters.

20.2.2.2 Vesicular Transport (Organelle Carriers)

Vesicular intermediates like autophagic membrane, the endosome, and secretory lysosomes are the organelles carriers involved in unconventional secretion pathways.

Autophagy-Based Secretion (Secretion Through Autophagasome-Like Vesicles)

This type of secretion uses intracellular vesicles such as secretory lysosomes, microvesicles, and multivesicular bodies to transport the cytoplasmic protein into the extracellular space. Secretion of IL-1β relies on this mechanism, which is processed into active IL-1β by Caspase-1 and released as a mature form of IL-1β.

Golgi Bypass

Secretion of CFTR, CD45, Connexin 26 and 30, Pannexins 1 and 3, Serglycin and Smoothened depends on this mechanism. It is an anterograde mode of transport of protein from ER to the plasma membrane without any involvement of the Golgi complex.

20.3 Release of Extracellular Vesicles

In addition to proteins, cells also release membranous vesicles termed as Extracellular vesicles (EVs) into the extracellular space. According to their mode of biogenesis, they are categorized into three major groups, which include ectosomes (shedding microvesicles), exosomes, and apoptotic bodies. The vesicles are secreted into body fluids including breast milk, blood plasma, urine, saliva, amniotic fluids, and pleural ascites. It is important to study the role of these vesicles in diverse physiological and pathophysiological conditions as these vesicles contain proteins, lipids, and RNA and thus act as great reservoirs of disease biomarkers [12, 13].

20.3.1 Shedding Microvesicles

Ectosomes that were considered an experimental artifact seem to be involved in important biological processes. Their sizes range between 50 and 1,000 nm in diameter. Expression of phosphatidylserine (PS) on the membrane is a characteristic feature of shedding microvesicles. They shed from cells by budding of the plasma membrane followed by fission. Proliferating tumor cells and the cells forming the tumor microenvironment release them containing proteases that can degrade ECM and help to promote angiogenesis. These vesicles also transfer mRNA and miRNAs between tumor cells and endothelial cells creating a favorable environment for neoangiogenesis followed by angiogenesis. In addition, anti-cancer drugs like doxorubicin and other chemotherapeutic agents are entrapped into the membrane of these shedding vesicles that decreases the availability of drug to cancer cells and are thus involved in drug resistance [14].

20.3.2 Exosomes

Fusion of multivesicular endosomes with the cell surface results in the formation of exosomes having a size of 40–100 nm in diameter [15]. Release of exosomes into the tumor microenvironment helps to transport various proteins, transcription factors, miRNAs and affect the cell signaling at distant organs as well as modulate a niche to promote favorable tumor environments. In cancer, changes in the cargo of exosomes depend on the level of p53, PTEN, and APC as shown by various studies. Small noncoding RNAs, miRNAs, which can influence the functions of various genes, are also one of the members exported via exosomes. Export of miRNA causes exchange of genetic information between the cells affecting various processes like modulation of immune response and promoting angiogenesis. Thus, exosomes have a potential role in tumor progression that affects various processes like changes in the tumor microenvironment, cancer metastasis, and angiogenesis and drug resistance [16].

20.3.3 Apoptotic Bodies

Fragmented apoptotic cells release 50–5,000 nm in diameter sized vesicles known as apoptotic bodies. Programmed cell death, otherwise known as apoptosis is a major mechanism of cell death in normal cells as well as in tumor cells. Apoptosis is a well orchestrated mechanism involving various steps like condensation of nuclear chromatin, membrane blebbing, disintegration of cellular content, and formation of membrane enclosed vesicles known as apoptosomes or apoptotic bodies [17]. Several studies suggest that release of apoptotic bodies mediate transfer of genetic information allowing intercellular communication that leads to changes in the tumorigenic phenotype of cancer cells.

20.4 Inhibitors to Study Secretory Pathways

20.4.1 Classical Secretion Pathway Inhibitors

Brefeldin A and Exo1 are inhibitors of classical secretory pathways, inhibiting Arf1 which is a GTP-binding protein involved in initiation of bud formation. To test whether a particular protein is released by a conventional or unconventional mechanism, these inhibitors are used to compare the level of the protein in the supernatant of treated and untreated cells using western blots or an assay system developed for that particular protein. After treatment with inhibitors, if there is significant change in the level of bait protein then there is a probability that the classical secretory pathway is responsible [18].

20.4.2 Non-classical Secretion Pathway Inhibitors

There are at least four major non-classical pathways as discussed earlier. To test the mechanism responsible for secretion of a particular protein by other than the classical secretion pathway multiple inhibitors are available. ATP transporters contribute to the release of several proteins. Glyburide, an inhibitor of ABC transporter, is used by several researchers to determine the contribution of ABC transporters in the secretion of particular proteins. In a similar manner, Cytochalasin D, an inhibitor of actin polymerization, and nocodazole, an inhibitor of microtubule polymerization, can be used to determine the role of pathways dependent on the cytoskeleton components, actin and microtubule respectively. The effects of other inhibitors like methylamine (inhibitor of endosomal and lysosomal function), ouabain (inhibitor of the Na^+, K^+-ATPase pump), calcimycin (membrane permeable Ca^{2+} ionophore), and EGTA (membrane impermeable Ca^{2+} chelator) are used to examine the role of this inhibitor in the non-classical secretory pathway [18].

20.5 Sources of Cancer Secretome

Secretomes of microorganisms like bacteria, yeast, and fungi are easiest to study since they do not require serum or other exogenous protein additives in their culture medium and the supernatant collected from the culture of these microorganisms contains only the proteins secreted by them. In mammalian cell secretome studies, it is absolutely essential to obtain the secretome as pure as feasible with the least amount of non-secretome serum proteins possible before analysis. There are two major sources of concentrated secretome proteins utilized to study cancer cell secretomes that are described below.

20.5.1 Cancer Cell Line Supernatant

20.5.1.1 In Vitro Secretome Sample

Major secretome studies in mammalian cells, especially in freshly obtained cancer cells or frozen-thawed cells, are performed using the cultured cancer cells because they are most similar to the physiological state of the cells in the tumors in vivo. The routine methods to collect the secreted proteins by tumor cell lines are used by the majority of the researchers. The cells of interest are cultured in serum-supplemented media until they reach the desired confluence (approximately 60–70 %). Totally serum-free or protein-free media are unsuitable because the cells die quickly in the complete absence of any serum proteins. Normally, the cancer cells are cultured in serum-supplemented media, but it is important to reduce the concentration of the serum proteins as much as possible to assure successful analysis of the cancer secretome. Cells reaching sufficient confluences are washed using PBS or serum-free medium to remove the bovine serum proteins. Before collecting the cell supernatant, the cells are incubated in the serum-free medium for another 12–24 h. With serum depletion, the growth of cells is reduced and they autolyse releasing cytosolic proteins [19] The main reason to grow the cells in serum-free medium is because the serum is highly enriched in albumin and many other serum proteins that mask the identification of low abundance proteins in secretome. After 12–24 h, serum-depleted supernates containing the secreted proteins of cells of interest are collected. It is necessary to optimize the incubation time, the cell confluence and to perform morphological and cell viability assays using trypan blue to obtain supernates containing a minimal amount of cytosolic proteins and also reducing the detrimental effects of serum protein depletion on the cells' health. To monitor cell lysis, two proteins, beta-actin and beta-tubulin, are used. Once the supernatant is collected it is necessary to remove any floating cells and cell debris by centrifugation followed by sterile filtration. Because the concentrations of some proteins are too low to measure in the range of ng/ml, it is necessary to concentrate the large amounts of supernate collected from the cells of interest. There are various methods, which can be used, for concentrating the supernates containing the secreted proteins but not all methods are suitable since the volume of media is so large. There are several methods like acetone precipitation, which requires a fivefold access of acetone, and dye precipitation, which excludes an important class of a secreted protein, namely proglycoproteins and ultrafiltration. Ultrafiltration is the method of choice of all the methods described, despite leakage of some low molecular proteins. Despite these refinements, substantial amounts of intracellular proteins are detected which are believed to be due to cell death and lysis or leakage. Evidence also suggests that the non-classical pathway using vesicles and exosomes can also secrete these proteins. Alpha-enolase is a glycolytic enzyme that is detected in significant amount in the secretome of all types of cultured cell medium. Discrimination of proteins secreted and release of intracellular protein due to non-physiological stress and mechanical injury are the biggest challenge. One of the few techniques

that enable high-density cell culture with minimal cell lysis is the use of hollow-fiber culture systems. It is also observed that intracellular proteins present in the condition medium of the cells are believed to be secreted due to leakage or cell lysis also contribute towards extracellular functions that are totally different than when the proteins are present in the intracellular compartment.

20.5.2 Proximal Biological Fluid

Most of the studies in the cancer secretomics field use cancer cell lines to collect the secreted proteins from the supernatant in vitro. Nevertheless, the in vitro culture system of cell line is far from recapitulating the heterogeneity of human tumors and also lacks the host-tumor microenvironment that plays a significant role in the progression and survival of any human tumors. It is also important to note that the genomic and phenotypic characteristics change over long passage time and contribute towards evolution of totally new populations of cell lines that differ substantially from the original line. To date only a few studies have investigated the cancer secretome under in vivo conditions.

20.5.3 In Vivo Secretome Profiling

Analysis of body fluids including serum, cerebrospinal fluid, nipple aspirate fluid, urine, saliva, tears, amniotic fluid, pleural effusion, bronchial washes, ear fluids, tears, interstitial fluids, synovial fluid, and tumor interstitial fluids have been used to investigate the proteome of these body fluids in individuals in healthy and diseased conditions. The major problems in analysis are the detection of low abundance proteins, the contribution of a specific tissue to the proteome of a specific body fluid, and the contamination by proteins in the blood. Interstitial fluids contain proteins that are secreted by cells and tissues in their local environment. The major challenge is the flow of interstitial fluids and proteins from the tissue of interest that cannot be studied in isolation using the circulatory fluid. But few studies such as micro-dialysis devices, capillary ultrafiltration, or ultraslow novel micro-dialysis devices have been used to collect the secretory proteins [20–22].

20.5.4 Ex Vivo Secretome Profiling

Using this method, the freshly obtained tissue specimen is incubated in serum-free basal medium for 12–24 h. Following incubation the conditioned medium is collected and processed for proteomic profiling.

Another strategy that can be useful is to incubate tissues in the medium containing a heavy isotope of lysine for several hours that allows identification of de novo synthesized proteins secreted into the extracellular medium.

20.6 Application of Mass Spectrometry in Cancer Secretome Analysis

To understand the complexities of neoplastic disease, it is important to consider the biological capabilities that human tumors have acquired during their development. The six major hallmarks noted earlier to be considered, include sustaining proliferative signaling, evading growth suppressors, resisting cell death, enabling replicative immortality, inducing neo-angiogenesis, and activating invasion and metastasis and also the two emerging hallmarks: reprogramming energy metabolism and evading immune destruction [2]. The most important characteristic of cancer cells, which leads to a fatal outcome, is their proclivity to continue proliferating while disregarding pre-existing homeostatic regulatory molecular pathways in each organ and tissue. Cancer secretomes also have a major impact on cancer progression, and studies describing the secretomes of various types of cancer have been rising with the hope to find differentially-regulated proteins (Fig. 20.2).

20.6.1 Identification of Growth Factors Involved in Sustaining Proliferative Signaling

Transmission of signals is important for normal stem and progenitor cells (S/P cells) to activate them from the quiescent state into an active proliferative state. Normal cells require diffusible growth factors (GFs), hormones, extracellular

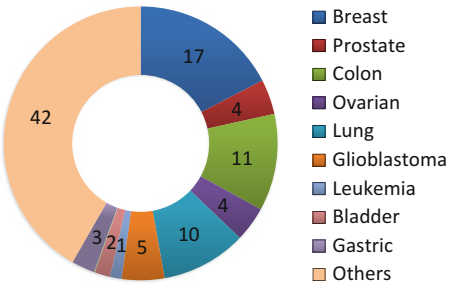


Fig. 20.2 Percentage distribution of publications describing the cancer secretome. Pie chart showing the percentage of articles describing the cancer secretome studies from 2000, when it was first described, to 2013. NCBI Pubmed Central, with a query of "secretome" and cancer type, was used to obtain data

matrix elements, and cell–cell communication that transmit the proliferative signals from the cell surface receptors to the nucleus. Binding of GFs to receptors and their interactions are highly specific. Thus, the growth of normal cells is dependent on these exogenous stimulations. In the case of cancer cells, they do not necessarily rely on the external growth factors or other normal exogenous signals, but may produce autocrine or paracrine growth factors or other stimulatory molecules, which reduces their dependence on stimulation from the surrounding tumor microenvironment. The secretion and production of PDGF is one example showing the role of a GF in activating an autocrine stimulatory loop and causing excessive proliferation [1, 23–26].

20.6.2 Identification of Factors Involved in Acquiring Insensitivity to Antigrowth Signals

Tight regulation of quiescence and tissue homeostasis is vital in maintaining normal tissue growth. There are two main signals including soluble growth factor inhibitors and surface receptor inhibitors that are present on the surrounding cells present in the tumor microenvironment.

Two mechanisms are important in the process of antigrowth signals. Actively proliferating S/P cells can enter a quiescent state (Go) when stimulated by these signals and can reenter the cell cycle when extracellular signals are available. Another process is commitment to differentiate whereby the cells undergo a strictly limited number of transit amplifying or maturation divisions before reaching a state of terminal differentiation when they can no longer divide.

During the cell cycle, normal S/P cells sense the outside signals to make the decision whether to continue proliferating, to resume quiescence, or to proceed to differentiate. Cancer cells develop multiple complex mechanisms to evade these antigrowth signals and to continue to proliferate and survive.

The soluble-signaling molecule TGFβ is the most well studied antigrowth factor. It decreases the expression of the c-myc gene, which regulates cells in the G1 portion of the cell cycle.

20.6.3 Identification of Anti-apoptotic Molecules

Tumor cells are not only capable of expanding in number by proliferation but are also sometimes able to develop resistance to dying, another important normal mechanism for limiting excessive cell growth, namely apoptosis or programmed cell death. Upon activation of apoptosis, signals are activated wherein apoptotic machinery composed of upstream regulators and downstream effectors play important roles.

Some of the major morphological changes that occur during apoptosis include cell shrinkage, membrane blebbing, the cytoplasm becomes dense, the organelles becomes tightly packed, and pyknosis occurs when the chromatin condenses and there is nuclear and DNA fragmentation. A major difference between apoptosis and necrosis is that apoptotic cells do not release their cellular content into the microenvironment but rather are phagocytosed by surrounding macrophages. Survival Signals conveyed by IGF-1/IGF-2, IL-3 and death signals conveyed by Fas-ligands and TNFα monitor normality and abnormality in the outer environment and decide whether cell should survive or die [2, 27].

20.6.4 Identification of Angiogenesis-Inducing Proteins

Like normal cells, cancer cells of course also require nutrients and oxygen to sustain their growth and the ability to get rid of metabolic waste and carbon dioxide. All cells in tissue reside within 100 μm of a capillary blood vessel and this process is highly regulated by the closely coordinated growth of vessels and parenchyma. This coordination is made possible by both positive and negative signals produced by the surrounding tissue in stimulating or inhibiting angiogenesis. These signals are soluble and diffusible factors and interact with surface receptors present on the endothelial cells. The importance of inducing and sustaining angiogenesis in xenograft tumors has been well studied. Vascular endothelial growth factor (VEGF) and Fibroblast growth factors bind to tyrosine kinase receptors and promote angiogenesis. Thrombospondin-1 is an angiogenesis inhibitor involved in inhibition of blood vessel formation [2, 28].

20.6.5 Identification of Tissue Invasion and Metastasis Promoting Protein

Growth of the initial cancer is often not the only cause of cancer-related death, but secondary tumors or metastases also can play a major role. Invasion and metastasis can be divided into five stages. During the invasion and migration stages the cells detach from primary tumor and invade adjacent tissues. The intravasation stage allows them to adhere to blood and lymphatic vessels and make the blood vessel leaky by secreting proteolytic enzymes, which allow the cancer cells to infiltrate the blood vessels while extravasation permits them to leave the blood stream and colonize in different tissues. Once the cancer cells colonize at distant tissue sites, the cells proliferate and induce neo-angiogenesis by secreting growth factors, which makes the surrounding environment affable to supporting their survival and growth.

20.6.6 Identification of Proteins Involved in Reprogramming Energy Metabolism

Normal cells utilize glucose to produce pyruvate in the presence of oxygen, a process known as glycolysis. Pyruvate is generated in the cytosol and transported to mitochondria generating carbon dioxide and releasing ATPs. During anaerobic conditions when limited oxygen is available, the glucose is converted into lactatic acid. Cancer cells are often able to reprogram their energy metabolism pathways and sustain their growth by "aerobic glycolysis" despite the presence of oxygen, a distinguishing feature was first observed by Otto Warburg, the great German biochemist and physiologist in 1930 [29], a mechanism that links metabolism and cancer through increased aerobic glycolysis. Cancer cells are able to adjust the energy metabolism requirement by reprogramming glucose metabolism and thus their energy production. Surprisingly, two subpopulations of cancer cells are found in some tumors, which differ, in their energy-generating and energy-utilizing pathways. Glucose-dependent cells in such tumors secrete lactate that is imported by the other subpopulation by employing the citric acid cycle mechanism [30].

20.6.7 Identification of Proteins Involved in Evading Immune Destruction

The immune system plays an important role in repressing, resisting, and eradicating tumor formation. In 1909 Paul Ehrlich postulated the involvement of the immune system in cancer [31], but its possible role in repressing cancer was at first not widely accepted due to the lack of sufficient experimental evidences. Recently, several studies have shown that the immune system can indeed sometimes recognize and eliminate the majority of tumor cells in established cancers and eliminate nascent tumors. This work led to the theory of immune surveillance, and subsequent experimental studies have strongly supported the existence and physiological relevance of cancer immune surveillance in both mouse and human studies. Due to inherent genetic instability of tumors, mechanisms have been developed to eliminate tumor cells of high immunogenicity, leaving tumor variants of low immunogenicity behind. This process is known as immune sculpting. It was believed that immune sculpting occurs in the early development of tumors when the cancer is not clinically detectable. Thus, cancer cells are sometimes capable of disabling an important component of the immune system that could lead to their destruction [32]. As an example cancer cells can secrete factors like TGF-β or other immune suppressive factors to disable CD8+ cytotoxic T lymphocytes (CTL) and Natural Killer (NK) cells [33, 34] Secretion of phosphatidylserine triggers activation of macrophages and dendritic cells causing phagocytosis of apoptotic cells that subsequently leads to secretion of anti-inflammatory mediators like IL-10, TGF-β, and

PGE_2, thus supporting the malignant progression of tumor cells. Tumor cells also secrete soluble Fas, soluble FasL, and MHC-class 1 related chain A proteins thereby helping tumor cells to avoid cytolysis by CTL and NK cells [35]. Secretion of IL-10, VEGF, and TGF-β inhibits the maturation impairing inflammatory responses of dendritic cells [36].

20.6.8 Identification of Proteins Involved in Supporting Resistance Against Anti-cancer Therapy

Resistance to anti-cancer drugs may be attributed to multiple gene mutations, gene amplification, or epigenetic changes that modify uptake, metabolism, or export of drugs from cancer cells. However, the tumor microenvironment and the heterogeneity of the cancer cell population in tumors are also involved in mediating resistance against anti-cancer therapies [37] Several oncoprotein-targeted therapies have been developed and tested in clinical trials with partial clinical responses, but the tumors often recur within a few months or days of treatment indicating that a substantial proportion of the cells are resistant to these therapies. Several investigators have proposed that factors secreted by the tumor microenvironment not only contribute to the growth and metastatic potential of tumor but are also involved in the development of innate drug resistance [38–40]. Using a co-culture system, Straussman et al. have shown how 23 stromal cell types influence the resistance of 45 cancer cells lines to 35 anti-cancer drugs. They have shown how secretion of HGF leads to reactivation of MAPK and PI3K signaling pathways and finally inhibition to RAF inhibitors in BRAF mutant melanoma cell lines [41] In a similar study using different receptor tyrosine kinase (RTK) ligands, other researchers have shown that drug-induced growth inhibition of 41 human-tumor-derived cell lines can be rescued by exposing them to one or more RTK ligands [42].

20.6.9 The Role of the Cancer Stem Cell Secretome in Cancer Progression

The cancer cells within tumors were once considered to be a fairly homogeneous population of cells. However, more recently there have been numerous studies showing that most tumors are extremely heterogeneous, and it is now widely accepted that a clonal subpopulation of cells exists with distinct tumor-initiating properties known as tumor-initiating cells (TICs). This clonal heterogeneity contributes to different degrees of proliferation, differentiation, and metastatic invasiveness, and is also characterized by differences in their cellular proteome and release of secretome proteins. Within these clonal subpopulations, there are special types of rare cells known as cancer stem cells (CSCs). CSCs exhibit similar

properties to normal stem cells like self-renewal and the ability to differentiate. Those CSCs might originate as a mutated normal stem cell or early progenitor cell and be capable of initiating a fatal malignance by uniclonal expansion was proposed to be the case in CML and the acute myelogenous leukemias by Fialkow [43] and several other investigators over 50 years ago (reviewed in Clarkson and Rubinow [44]). Using a severe combined immunodeficiency mouse model (SCID/NOD), it was found that CD34+ CD38- cells are able to differentiate in vivo into leukemic blasts, which were designated as leukemia initiating cells (LICs) [45]. Differential expression of cell-surface markers present on the neoplastic cell populations allows their separation using flow cytometry and the resulting fractions can be tested to evaluate their cancer-initiating properties.

The role of CSCs was also established in solid tumors including breast [46], prostate [47], colon [48], ovarian [49], and skin cancers [50]. How the CSCs arise and are maintained at stem cell stage is poorly understood. Various investigators have proposed that the CSCs niche plays an important role in promoting their survival, invasiveness, metastatic properties, and resistance to anti-cancer therapies. Using co-culture systems to investigate how the secretome of astrocytes contributes to the invasive phenotype of glioblastomas (GBMs) from GBM-stem-like cells (GSCs), Rath et al. identified that secreted factors from the normal microenvironment contribute significantly to the invasive/metastatic potential of GSCs [51].

It is well-established that CSCs are not only resistant to anti-cancer therapies that lead to tumor recurrence, but that a CSC-intrinsic drug resistance mechanism might have an equal contribution. To address this question, Emmink et al. investigated secretomes of CSCs and isogenic-differentiated tumor cells isolated from colon cancer patients with metastatic tumors. Using this proteomic approach they found that in addition to secreting proteins related to cell survival, the cells secrete at least two drug-metabolizing enzymes, aldehyde dehydrogenase 1 (ALDH1A1) and bleomycin hydrolase (BLMH) [52, 53].

20.6.10 Identification of Proteins Secreted from the Tumor Microenvironment

The tumor microenvironment has been recognized as a major factor influencing the biology of tumors. It is composed of non-cancer cells including cancer-associated fibroblasts (CAFs) and immune inflammatory cells such as monocytes/macrophages, neutrophils, and lymphocytes which are recruited to and reside in tumor stromas and also vascular cells like endothelial cells and pericytes as well as local and bone marrow-derived stromal stem and progenitor cells [54]. The interaction of tumor and microenvironment cells can promote the proliferation, metastasis, drug resistance, and immune invasion of cancer cells. Cancer cells are able to modify their protein expression profile and thus lead to alteration of secretory proteins [55].

Interaction of cancer cell with CAFs allows specific communication in assisting cancer growth. Proteins in the CAFs secretome such as HGF and TGF-β have been shown to initiate cancer, and once a tumor is initiated, the CAFs also assist in promoting proliferation and progression of cancer cells in their niches by releasing proteinases like MMP-2 and MMP-9, growth factors like IGF-1, chemokines like TNF-α, and angiogenesis factors like VEGF. Fat tissue is abundant in the breast and the breast cancer cell-adipocyte interaction seems to have an important role in breast cancer progression. A study conducted by Celis et al. to investigate secreted proteins in the interstitial fluid from breast cancer patients identified alterations in several inflammatory cytokines, growth factors, angiogenic factors, and proteinases, suggesting specific interactions and communications within the complex cancer microenvironment [56–59].

20.6.11 Identification of Biomarkers

Further development of new biomarkers for detection, diagnosis, prognosis, and disease monitoring for cancer is necessary. Biomarkers are important tools to evaluating the physiological state of disease at specific times. Transformation of normal cells to a cancerous state is a complex process causing alteration at genetic and epigenetic levels leading to changes in respective protein product levels. The carcinoembryogenic antigen (CEA), prostate specific antigen (PSA), alpha-feto-protein (AFP), CA 125, CA15-3, and CA19-9 are established as clinically useful cancer biomarkers regularly employed in clinics. Studies to discover promising new specific and sensitive biomarkers have been undertaken by many researchers, but most of the proposed biomarkers have failed to gain acceptance in general usage due to lack of specificity, sensitivity, and performance in the clinical setting.

Cancer secretome analysis allows investigation of novel biomarkers using advanced proteomic tools. An example of a proteomic strategy on how to investigate two different secretomes of treated and non-treated cancer cell populations is given in Fig. 20.3a as well as a high throughput method consisting in bypassing the time consuming gel electrophoresis step (Fig. 20.3b) [60]. Additional proteomics strategies are described elsewhere [61–75]. Studies employing primary lung cells and normal bronchial epithelial cells from six lung cancer patients identified CD98, fascin, 14-3-3 η, and polymeric immunoglobulin receptor/secretory component as potential biomarkers [76]. Aiming to discover a biomarker in melanoma, Pardo et al. studied the proteomic composition of patient serum and identified cathepsin D and gp100 as potential biomarkers [77]. The majority of biomarkers identified in major studies appear promising until the validation phase but then fail to enter the critical phase of biomarker development raising concern about the future value of these biomarker studies. However, some studies have suggested that secretome analysis may be a feasible and efficient method for identification of potential biomarkers [78–82].

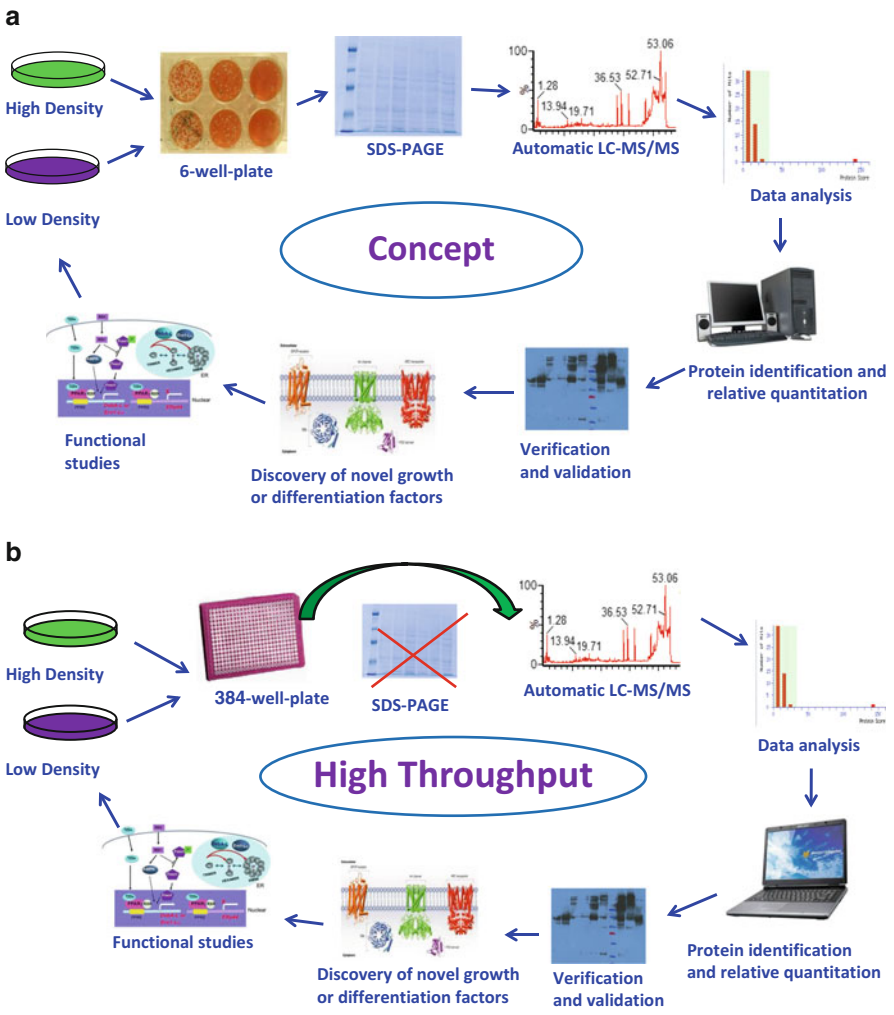


Fig. 20.3 (**a**) General principle for analysis of secretomes. The secreted proteins obtained from the cells growing in different conditions are collected from supernates and concentrated proteins were separated using sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and digested proteins are analyzed by LC-MS/MS. (**b**) The modified method allowing automated analysis of secretome proteins using high throughput methods like 384-well plates. Reprinted with permission from Ngounou Wetie et al., 2013 [60]

20.7 Investigation of the Cell Secretome

20.7.1 Genome-Based Computational Prediction

Computational methods have a significant role in predicting the composition of the secretomes of many types of organisms. The availability of whole genome sequencing provides an opportunity to examine the secretome using analytical and

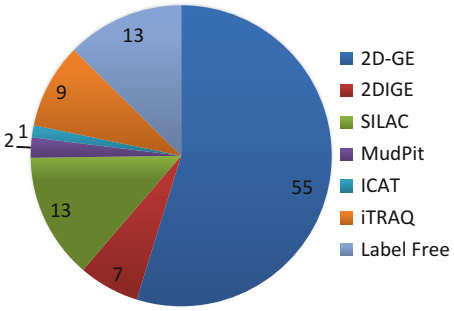


Fig. 20.4 Percentage distribution of publications describing the mass spectrometry-based method to study the cancer secretome. Pie chart showing the percentage of articles describing the mass spectrometry method used to study the cancer secretome studies from 2000, when it was first described, to 2013. NCBI Pubmed Central, with query keywords of "secretome," cancer type and mass spectrometry, was used to obtain data

computational tools. Prediction of specific secretory proteins using bioinformatics relies on N-terminal signal sequence present in the proteins. Proteins having a signal peptide are secreted via the classical ER-Golgi pathways and can be predicated using SignalP software. To determine the identity of proteins secreted via non-classical secretion pathways SecretomeP is widely used [59, 83–87].

20.7.2 Proteomic Approaches

An improvement in proteomic methods, and increased sensitivity of mass spectrometry instruments has led to an increase in cell secretome studies [88]. Figure 20.4 shows the pie chart describing the percentage of publications, from 2000 to 2013, utilizing different proteomic methods for analyzing secretomes obtained from different types of cancer cells. In this section, we have described the different methods that are being used by researchers to evaluate secretomes. Table 20.1 summarizes the findings from some of the cancer secretome studies where different proteomic approaches were used and their outcomes that led to identification of possible candidates that might have important roles in cancer progression.

20.7.2.1 Gel-Based Proteomic Technologies

Two-Dimensional Gel Electrophoresis

Two-dimensional gel electrophoresis (2-DE) is the method of choice for the separation of complex protein samples utilized by many laboratories. In this method the proteins are separated on the basis of their net charge using Isoelectric focusing in first dimension. Then in second dimension, the proteins are separated on the basis of their molecular masses on SDS-PAGE (SDS Polyacrylamide Gel Electrophoresis).

Table 20.1 Outcomes of cancer secretome studies using different mass spectrometry-based techniques

Two-dimensional gel electrophoresis				
Cancer	Candidate	Property	Notes	Reference
Nasopharyngeal carcinoma	Galectin-1 (identified 11 protein which were differentially expressed)	ELISA	Study comparing secreted proteins from nasopharyngeal carcinoma-associated stromal fibroblasts and normal fibroblasts	[90]
Glioblastoma	HDGF	Western blotting, siRNA silencing	Identification of novel candidate supporting neo-angiogenesis in gliomas	[103]
Nasopharyngeal carcinoma	APP, cystatin C	Western blotting, pathway inhibition	Identification of TGF-α stimulated secretory proteins	[104]
Bladder cancer cell	SPARC, tPA, clusterin	Western blot, antibody inhibition, functional studies	Identification of secreted proteins involved in aggressive bladder cancer cell phenotype	[105]
Melanoma	Lactate dehydrogenase B, M2 pyruvate kinase, cathepsin D, and galectin 1, syntenin	Role of syntenin in invasion process using functional studies	Characterization of secreted proteins comparing melanoma cell line and melanoma cell line with profound invasive capacity	[106]
Breast cancer cell microenvironment	Multiple proteins	Validated calreticulin, cellular retinoic acid-binding protein II, chloride intracellular channel protein 1, EF-1-beta, galectin 1, peroxiredoxin-2, platelet-derived endothelial cell growth factor, protein disulfide isomerase, and ubiquitin carboxyl-terminal hydrolase 5 using a tissue microarray	Established upregulated proteins in tumor interstitial fluid compared to normal interstitial fluids. This study used a large set of patient sample (68 tumor samples)	[107]
Colorectal tumor explants	32 differentially expressed proteins	Potential biomarker	Study comparing secretome of colorectal cancer liver metastatic and normal colon mucosa specimen	[108]
Colon carcinoma	"Secreted protein, acidic and rich in cysteins" (SPARC)	Functional studies	Differential Secretome between Smad4-negative and Smad4-positive colon carcinoma cells	[109]

Differential gel electrophoresis				
Cancer	Influence	Confirmation	Notes	Reference
Pancreatic	14-3-3 σ	Western blotting	Effect of gemcitabine on pancreatic cancer cell secretion	[40]
	Lactotransferrin	Western blotting		
Colon adenocarcinoma	TCTP, TGFβIp, NPC2		Secretome study comparing differentiated and undifferentiated Caco-2 cell line	[110]
Large cell lung cancer	IL-25, stathmin, vimentin		Compared two cell lines obtained from patients	[111]
Colon cancer	25 Differentially expressed secretome proteins		Identification of secretome comparing Smad4-deficient and Smad4-re-expressing human colon carcinoma cells.	[112]
SILAC				
Cancer	Influence	Confirmation	Notes	Reference
Colon			Alteration of 45 proteins. Shown importance of interaction between epithelial cells and colon cancer cells	[113]
Gastric	Activation MMP-1, MMP2 and MMP-3	In-vivo imaging	Activation of proteolytic processing by gastric cancer-associated myofibroblasts	[114]
Colorectal cancer-associated liver metastasis	GDF15, S100A8/A9, SERPINI1	ELISA	Allows discrimination from healthy individual	[115]
	NEO1, SERPINI1, PODXL	siRNA silencing and Antibody blocking	Effect on cellular adhesion	
	SOSTDC1, CTSS, EFNA3, CD137L/TNFSF9, ZG16B and Midkine	siRNA silencing and Antibody blocking	Decrease in migration and invasion of highly metastatic cells	
	SOSTDC1, EFNA3, CD137L/TNFSF9	siRNA silencing	Reduced liver colonization	
Gastric	PCSK9, LMAN2, PDAP1	Immunohistochemistry	Overexpression of 263 proteins in gastric cancer cell lines as compared to normal gastric epithelial cells	[116]
Breast cancer		Super-SILAC mix	Identification of N-glycosylated secretome proteins in different stages of breast cancer	[117]

(continued)

Table 20.1 (continued)

SILAC				
Cancer	Influence	Confirmation	Notes	Reference
Hepatocellular carcinoma cells	61 Upregulated and 11 down-regulated proteins (twofold) PAI-1	Western blot, RT-PCR, overexpression	Identification of thyroid hormone (T3)-regulated secretome in hepatoma cells	[118]
Metastatic glioblastoma multiforme	ADAM9, ADAM10, cathepsin B, cathepsin L1, osteopontin, neuropilin-1, semaphoring-7A, suprabasin, CH3L1	Role of CH3L1 using Antibody inhibition and functional studies	Identification of secreted proteins exclusively secreted in highly invasive cell line	[119]
Esophageal squamous cell carcinoma (ESCC)	PDIA3, GDI2, LGALS3BP	Western blot, immunohistochemistry	Identification of secretome comparing non-neoplastic and ESCC	[120]
Mammary fibroblast	More than 1,000 proteins identified		Investigation of secretome comparing TGF-beta knockout and WT mammary fibroblast	[121]
Pancreatic cancer	145 Differentially secreted proteins	Subset of proteins validate using Western blotting	Evaluate differential secretome between pancreatic cancer-derived cells and non-neoplastic pancreatic ductal cells for biomarker discovery	[122]
Label-free methods				
Cancer	Candidate	Property	Notes	Reference
Lung adenocarcinoma	PARK7	Overexpression	Used hollow-fiber system to study secretome	[123]
Plexiform neurofibroma	RARRES1	ELISA, mRNA levels, Western blot	Comparison between human plexiform neurofibroma derived Schwann Cells and normal nerve tissue. 22 proteins were present at high levels and 9 were present at low levels in NF1 cells	[124]

Breast cancer bone metastasis	128 Proteins up/down-regulated CST6	Functional studies like proliferation, colony formation, migration, and invasion of breast cancer cells	Establishment of CST6 as a bonafide breast cancer osteolytic metastatic suppressor	[94]
Colorectal	EFEMP2	ELISA, immunohistochemistry	In this study lectin affinity-based approach was used to capture secreted protein and quantified using label-free spectral counting comparing fresh CRC tissue and paired normal colorectal tissue	[79]
Colorectal	Release of extracellular vesicles (EV)		Comparison of proteomes between EVs derived from primary colorectal cancer cells and their metastatic derivatives	[125]
Breast cancer	IL-18	Western blotting, ELISA, functional studies,	Study identifying change in secretome of the breast cancer cells and doxorubicin-resistant breast cancer cells	[39]
Human pancreatic ductal adenocarcinoma	CD9, vimentin, PLOD3, SH3L3, PCBP1, SFRS1	Proposed to be involved in pancreatic cancer	Identification of functional networks using functional enrichment analysis of secreted proteins from pancreatic cancer cell lines	[126]
Colorectal cancer metastasis	GDF15, TFF3, AGR2, TGM2, LCN2, IGFBP7	Western blot analysis, ELISA	Characterization of differential secretomes present in primary colorectal cell line and its lymph node metastatic cell line	[127]

The proteins resolved using 2-DE are subjected to either coomassie blue or silver staining and then the protein spots are excised from the gel and the spots are destained, reduced, alkylated, and subjected to enzymatic digestion. Then the digested protein samples are subjected to mass spectrometry to identify specific proteins.

The advantage of using this method is that it is relatively inexpensive, can be applied to a wide range of samples and is able to detect any post-translational modification of specific protein isoforms. The disadvantages are poor separation of membrane and hydrophobic proteins that might result in the precipitation of these proteins. The method has been used in the vast majority of cancer secretome studies [89, 90].

Differential In-Gel Electrophoresis

Two-dimensional differential gel electrophoresis (DIGE) is a modification of 2-DE [91–93]. It is designed to detect differential protein expression that is reproducible and allows quantification. It is used to measure quantitative changes in the proteomes of two different samples. In the DIGE method, before running the gel the proteins are mixed with different fluorescent cyanine dyes (e.g., Cy2, Cy3, Cy5). Two samples are mixed and run on the gel in both dimensions that allows the first separation on the basis of net charge and the second on the basis of molecular masses. Once the run is complete, the gel is scanned with the excitation wavelength of each dye. Two different gel images are then superimposed to measure the differences in the protein expression levels. After image analysis, the gel spots are excised and subjected to the same procedure for protein sample preparation that reduces the variation by using different gel and a difference in the washing. This procedure is highly reproducible and the method allows detection of unknown secretory proteins with a high resolving power. It also allows identification and quantification of proteins differentially or uniquely expressed in any two methods under study [94]. Using this technique, we analyzed the secretome in two supernate samples of ALL3 cells grown in QBSF-60 media. The ALL3 cells were originally obtained from the pleural fluid of an adult patient with Ph+ALL. After growing large number of these Bcr–Abl expressing cells in QBSF-60, multiple aliquots were frozen in liquid nitrogen to preserve the cells as closely as possible to their stage in the pleural effusion. Periodically aliquots were thawed for further studies and then a portion of the cells was refrozen. A series of studies have shown that the ALL3 cells only grow at a high cell density (HD) but die at a low cell density (LD) and studies using transwells or cell-free supernates showed that the HD cells secrete diffusible factors that rescue the LD cells, preventing some of them from dying and stimulating them to resume growing rapidly (Fig. 20.5).

A HD and an unstimulated LD QBSF supernates were separated by DIGE for the investigation of the differences in their proteomes. DIGE, with its much higher resolution than 1D-SDS-PAGE should improve the identification of extremely low-abundant growth-stimulating factors present in the supernatant of high-density cells.

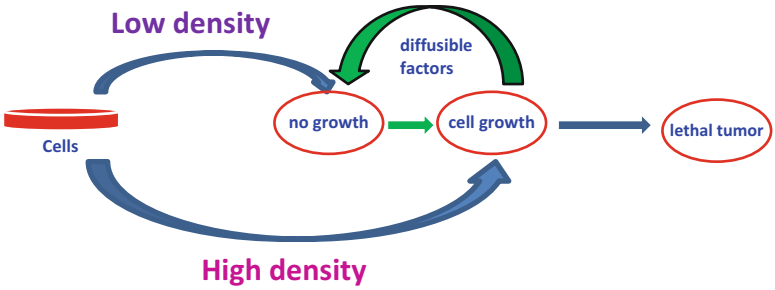


Fig. 20.5 Schematic for identification of a diffusible factor(s) secreted from high-density ALL3 cells that can stimulate growth of dying LD ALL3 cells. Reprinted with permission from Ngounou Wetie et al., 2013 [60]

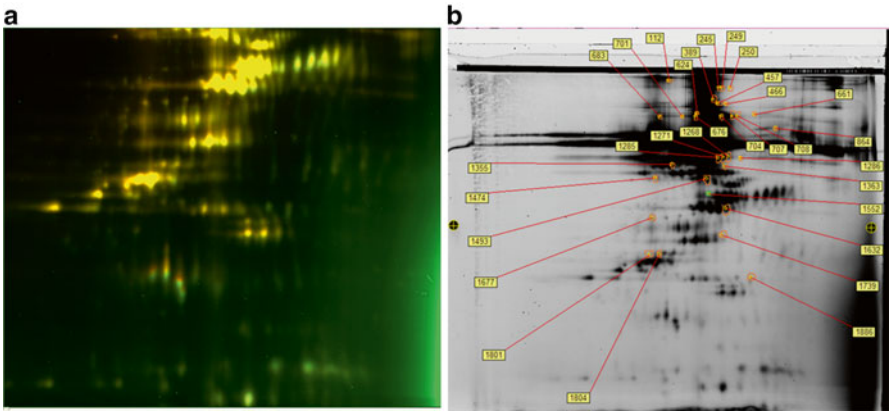


Fig. 20.6 Analysis of secretome proteins, obtained from ALL3 cells growing at high and low density, using DIGE method. The supernates from both samples are collected and treated with cyanin3 and cyanin5 dyes, then they are mixed together and separated using two-dimensional gel electrophoresis method. (**a**) Differentially expressed proteins are visualized by fluorescence-based method (**b**) The proteins are designated with numbers and processed for MS/MS analysis

The supernatants were concentrated and then mixed with Cy3 (green) and Cy5 (red) for visualization of the differences in protein expression (Fig. 20.6a). The DIGE gel shows that many of the high-abundant proteins are common to both conditions (yellow spots), whereas low-abundant proteins are more likely present in one condition and absent in the other. Figure 20.6b shows the spots that were depicted from these runs for further analysis. We performed a trypsin digestion of these spots followed by nano-liquid chromatography tandem mass spectrometry (nano-LC-MS/MS). In Fig. 20.7, we show the MS/MS of peptides from spot 250 (Fig. 20.6b) that led to the identification of three proteins (heat shock protein 90, elongation factor 2 and heat shock 70 kDa protein 6).

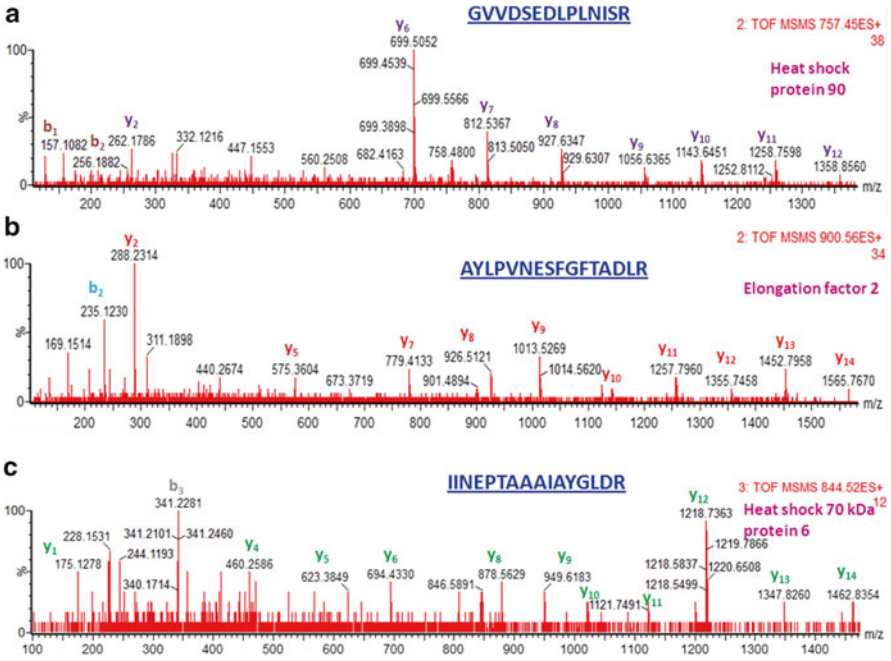


Fig. 20.7 Identification of three proteins, from spot 250 (Fig. 20.6b), using Nano-LC-MS/MS. The precursor ions with *m/z* of 757.45 (**a**), 900.56 (**b**), and 844.52 (**c**) were fragmented by MS/MS and produced a series of b and y ions whose analysis led to identification of peptides with amino acid sequence GVVDSEDLPLNISR (**a**), AYLPVNESFGFTADLR (**b**) and IINEPTAAAIAYGLDR (**c**), that were part of Heat shock protein 90 (**a**), elongation factor 2 (**b**), and Heat shock 70 kDa protein 6 (**c**)

20.7.2.2 Gel-Free MS-Based Technologies

Multidimensional Protein Identification Technology

In multidimensional protein identification technology (MudPIT), the separation in the liquid chromatography column is based on the difference in the retention of particular proteins or peptide using nonpolar stationary phase or polar solvent. Using different percentages of organic solvent separation and stationary phase allows separation based on an increasing degree of hydrophobicity of proteins or peptides. Strongly retained proteins or peptides can be resolved nicely using gradient elution [95].

Isotope-Coded Affinity Tag

In isotope-coded affinity tag (ICAT), two different samples are reacted to cysteine of proteins using light C12 ICAT and heavy C13 ICAT. Then treated samples are mixed together and the mixture is subjected to tryptic digestion followed by cation exchange

chromatography and multiple fractions are collected. All fractions collected are subjected to avidin chromatography isolation of cysteine-containing tryptic peptides. The tryptic peptides are subjected to LC-tandem mass spectrometry (MS/MS) analysis allowing identification of ICAT peptide pairs and quantifying the relative [^{12}C]/[^{13}C] ratios. ProICAT software from Applied Biosystem is used to identify and quantify the ICAT-derivatized peptide pairs that differ by exactly 9 Da [96].

Cleavable Isotope-Coded Affinity Tag

The cleavable isotope-coded affinity tag (cICAT) reagent is made up of four parts (a) A protein reactive group (b) An affinity tag (c) An isotopically labeled liner (d) An acid cleavage site.

(a) Protein reactive group (iodoacetamide): The reactive group covalently links the ICAT to proteins by alkylation of free cysteines.
(b) An affinity tag (Biotin): Simplifies analysis of ICAT reagent-labeled peptides thus reducing complexity of the peptide mixture.
(c) An isotopically labeled linker ($C_{10}H_{17}N_3O3$): Allows comparison of peptides labeled with heavy and light reagents, providing ratios of the concentration of the proteins in the original samples.
(d) An acid cleavage site: Allows removal of the biotin portion of the label and linker by addition of TFA, thus reducing the overall mass of a tag on peptide and increasing peptide fragmentation efficiency in the MS.

In this method, control and experimental samples are dissolved with ^{12}C (control sample) or ^{13}C (experimental sample) in the solution and then combined and the mixture is digested with trypsin. The mixture is separated in SCX cartridge and peptides are collected using an avidin column; the biotin tag is removed by treating the peptides with TFA and subjected to LC-MS/MS. The ^{12}C- or ^{13}C- ICAT moiety allows quantification of the control and experimental samples [97, 98].

Stable Isotope Labeling by Amino Acids in Cell Culture (SILAC)

Mammalian cells are supplied with minimum essential amino acids to support cell growth since they cannot synthesize these amino acids. These amino acids can be isotopically labeled and if they are given to cells in the culture the isotopic labeled amino acids will be incorporated into newly synthesized protein chains. Thus, the particular amino acids will be replaced by isotopically labeled analog after a certain number of cell doublings. It allows the comparison of control and experimental samples when one is labeled and the other is unlabeled. Proteins are collected from the two different samples and mixed together and the proteins and peptides can be analyzed following quantification of differentially expressed proteins in control and test samples [99].

Isobaric Tag for Relative and Absolute Quantization

In isobaric tag for relative and absolute quantization (iTRAQ) methods, the protein samples are derivatized with one of the many tags. All tags have similar overall mass but they vary in terms of the distribution of heavy isotopes around their structure. The tag compromises three portions: (a) A peptide reactive group (b) the reporter group (c) a balance group. The reporter group allows relative quantification of peptide upon MS/MS. Recent advances allow analysis of eight samples in a single experiment. The protein samples from different samples are collected, digested with trypsin, treated with one of the several tags, and mixed together. The peptide mixture is subjected to LC-MS/MS and a ratio of different reporter ions allows relative quantification of that peptide in each sample [100].

Surface Enhanced Laser Desorption Ionization-Time of Flight-Mass Spectrometry

Hutchens et al. developed SELDI technology in 1993 ([101], 576–580). Surface-enhanced laser desorption ionization-time of flight-mass spectrometry (SELDI-TOF-MS) is a modification of matrix-assisted laser desorption ionization (MALDI). Proteins or peptide samples are spotted on a surface having different chemical functionalities like metal-binding, hydrophilic, hydrophobic, ion-exchange and activated surfaces as well as biological functionality like antibodies, other proteins, and DNA. Use of the commercially available ChipBiology system uses the above mechanism to retain proteins on a solid phase chromatographic surface that follows ionization and detection by time-of-flight (TOF) MS [102].

20.7.2.3 Label-Free Proteomics

Label-free protein profiling is a simple and feasible method to support complex experimental designs. Increased sensitivity and resolution of mass spectrometers and reproducibility of peptide ion counts has attracted many researchers to utilize a label-free protein profiling approach. Samples to be quantified are prepared using identical reagents and proper internal controls are added to each sample to evaluate experiment-to-experiment variation. One of the approaches termed as MS^E spectra are recorded at alternate low (precursor) and high (product) fragmentation voltages. Spectra generated are analyzed using database searching an algorithm and peptide ratios are compared to observe the difference in relative protein levels between different samples.

20.8 Conclusion

Various factors are important in cancer progression and metastasis. Secretory proteins released by cancer cells and by stromal cells surrounding the cancer cells are both important in mediating the hallmarks of cancer. In this review we have

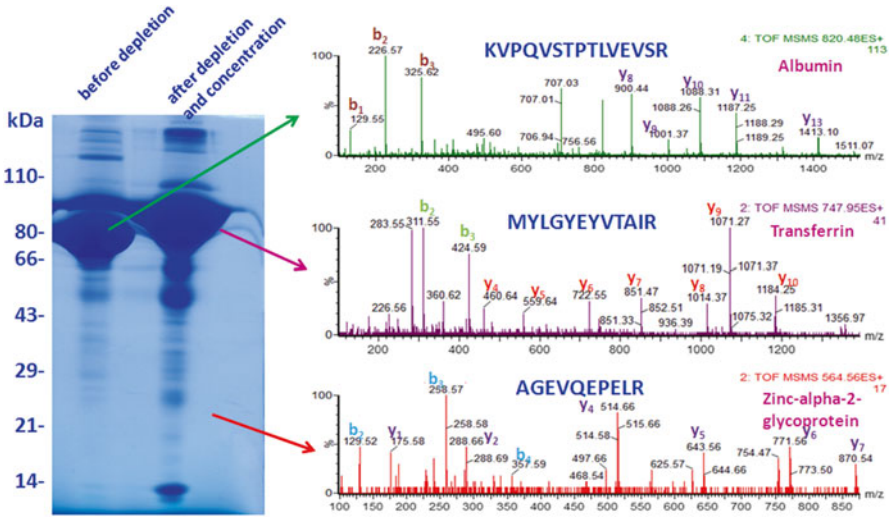


Fig. 20.8 Analysis of proteins before and after depletion: The supernates were albumin-depleted. Non-depleted and depleted samples are separated using SDS-PAGE. Before depletion, albumin is the dominant protein and after depletion transferrin is the dominant protein hampering identification of low-abundant proteins. MS/MS spectra of peptides showing identification of three proteins, the two most abundant proteins, albumin and transferrin, and a low-abundant protein zinc-alpha-2-glycoprotein. Reprinted with permission from Ngounou Wetie et al., 2013 [60]

attempted to update the contribution of secreted proteins to the pathophysiology of cancer and address challenges in identifying the specific low abundance proteins responsible for the abnormalities. The secretome is released from cells via both classical and non-classical pathways and it is altered in various diseases including cancer. These pathways can be studied in detail using various inhibitors. Changes in the cancer secretome and also in the secretome of normal cells within tumor microenvironment can be greatly affected by cancer cells.

Proteomic methods are powerful means to investigate and identify the secretory proteins. However, there are still many challenges that need to be addressed for proteomics to display its full potential in the analysis of the secretome. For instance, the presence of high-abundant serum proteins (e.g., albumin, transferrin) can prevent the detection of the disease-relevant molecules. This problem is being addressed by using serum-free culturing media (e.g., CellGro). However, the viability of cells in these media is limited and thus hampers the conduct of long-term studies. Figure 20.8 shows an albumin-depleted secretome that was concentrated after depletion. Though removal of most of the albumin was achieved, the other high abundant protein (transferrin) is now dominant after concentration. This shows the broad dynamic range represented in secretome samples and constitutes one of the major barriers to overcome in proteomics in general.

The classical gel-based proteomic approach is the method of choice allowing identification of key proteins involved in cancer progression, although a shift is occurring to the use of gel-free proteomic methods by many investigators due to its

sensitivity, simple processing, and automation. The advancing technology will permit a more detailed examination of the secretomes of tumor cells and provide new valuable information about the biology of cancer, which hopefully will be important in developing new approaches to early detection, more effective therapies, and prevention of malignant progression and metastases.

Acknowledgements This work was supported in part by The Enid A Haupt Charitable Trust, and The E./S. Sindina Lymphoma Research Fund, the MeadWestvaco Corp (to BDC). This work was also supported in part by the Keep a Breast Foundation (KEABF-375-35054), the David A. Walsh fellowship, and the U.S. Army research office (DURIP grant #W911NF-11-1-0304) (to CCD). CCD & AGNW would also like to thank Dr. Oscar Alzate, Dr. Robert M. Dekroon, and Ms. Mihaela Mocanu (Lineberger Comprehensive Cancer Center, University of North Carolina, Chapel Hill, NC) for the initial steps in DIGE analysis.

References

1. Hanahan D, Weinberg RA (2000) The hallmarks of cancer. Cell 100:57–70
2. Hanahan D, Weinberg RA (2011) Hallmarks of cancer: the next generation. Cell 144:646–674
3. Tjalsma H, Bolhuis A, Jongbloed JD, Bron S, van Dijl JM (2000) Signal peptide-dependent protein transport in Bacillus subtilis: a genome-based survey of the secretome. Microbiol Mol Biol Rev 64:515–547
4. Agrawal GK, Jwa NS, Lebrun MH, Job D, Rakwal R (2010) Plant secretome: unlocking secrets of the secreted proteins. Proteomics 10:799–827
5. Walter P, Gilmore R, Blobel G (1984) Protein translocation across the endoplasmic reticulum. Cell 38:5–8
6. Mellman I, Warren G (2000) The road taken: past and future foundations of membrane traffic. Cell 100:99–112
7. van Vliet C, Thomas EC, Merino-Trigo A, Teasdale RD, Gleeson PA (2003) Intracellular sorting and transport of proteins. Prog Biophys Mol Biol 83:1–45
8. Rabouille C, Malhotra V, Nickel W (2012) Diversity in unconventional protein secretion. J Cell Sci 125:5251–5255
9. Nickel W (2003) The mystery of nonclassical protein secretion. A current view on cargo proteins and potential export routes. Eur J Biochem 270:2109–2119
10. Nickel W, Rabouille C (2009) Mechanisms of regulated unconventional protein secretion. Nat Rev Mol Cell Biol 10:148–155
11. Zhang M, Schekman R (2013) Cell biology. Unconventional secretion, unconventional solutions. Science 340:559–561
12. Kalra H, Adda CG, Liem M, Ang CS, Mechler A, Simpson RJ, Hulett MD, Mathivanan S (2013) Comparative proteomics evaluation of plasma exosome isolation techniques and assessment of the stability of exosomes in normal human blood plasma. Proteomics 13:3354–3364
13. Ji H, Greening DW, Barnes TW, Lim JW, Tauro BJ, Rai A, Xu R, Adda C, Mathivanan S, Zhao W, Xue Y, Xu T, Zhu HJ, Simpson RJ (2013) Proteome profiling of exosomes derived from human primary and metastatic colorectal cancer cells reveal differential expression of key metastatic factors and signal transduction components. Proteomics 13:1672–1686
14. Cocucci E, Racchetti G, Meldolesi J (2009) Shedding microvesicles: artefacts no more. Trends Cell Biol 19:43–51
15. Fevrier B, Vilette D, Archer F, Loew D, Faigle W, Vidal M, Laude H, Raposo G (2004) Cells release prions in association with exosomes. Proc Natl Acad Sci U S A 101:9683–9688
16. Azmi AS, Bao B, Sarkar FH (2013) Exosomes in cancer development, metastasis, and drug resistance: a comprehensive review. Cancer Metastasis Rev 32(3–4):623–642

17. Akers JC, Gonda D, Kim R, Carter BS, Chen CC (2013) Biogenesis of extracellular vesicles (EV): exosomes, microvesicles, retrovirus-like vesicles, and apoptotic bodies. J Neurooncol 113:1–11
18. Dinter A, Berger EG (1998) Golgi-disturbing agents. Histochem Cell Biol 109:571–590
19. Chirico WJ (2011) Protein release through nonlethal oncotic pores as an alternative nonclassical secretory pathway. BMC Cell Biol 12:46
20. Huang CM, Nakatsuji T, Liu YT, Shi Y (2008) In vivo tumor secretion probing via ultrafiltration and tissue chamber: implication for anti-cancer drugs targeting secretome. Recent Pat Anticancer Drug Discov 3:48–54
21. Chiu KH, Chang YH, Liao PC (2013) Secretome analysis using a hollow fiber culture system for cancer biomarker discovery. Biochim Biophys Acta 1834:2285–2292
22. Pavlou MP, Kulasingam V, Sauter ER, Kliethermes B, Diamandis EP (2010) Nipple aspirate fluid proteome of healthy females and patients with breast cancer. Clin Chem 56:848–855
23. Hermanson M, Funa K, Hartman M, Claesson-Welsh L, Heldin CH, Westermark B, Nister M (1992) Platelet-derived growth factor and its receptors in human glioma tissue: expression of messenger RNA and protein suggests the presence of autocrine and paracrine loops. Cancer Res 52:3213–3219
24. Engebraaten O, Bjerkvig R, Pedersen PH, Laerum OD (1993) Effects of EGF, bFGF, NGF and PDGF(bb) on cell proliferative, migratory and invasive capacities of human brain-tumour biopsies in vitro. Int J Cancer 53:209–214
25. Pantazis P, Pelicci PG, Dalla-Favera R, Antoniades HN (1985) Synthesis and secretion of proteins resembling platelet-derived growth factor by human glioblastoma and fibrosarcoma cells in culture. Proc Natl Acad Sci U S A 82:2404–2408
26. Plate KH, Breier G, Farrell CL, Risau W (1992) Platelet-derived growth factor receptor-beta is induced during tumor development and upregulated during tumor progression in endothelial cells in human gliomas. Lab Invest 67:529–534
27. Lowe SW, Lin AW (2000) Apoptosis in cancer. Carcinogenesis 21:485–495
28. Falcon BL, Pietras K, Chou J, Chen D, Sennino B, Hanahan D, McDonald DM (2011) Increased vascular delivery and efficacy of chemotherapy after inhibition of platelet-derived growth factor-B. Am J Pathol 178:2920–2930
29. Warburg O, Posener K, Negelein E (1930) Ueber den Stoffwechsel der Tumoren. Biochem Z 152:319–344
30. Dang CV (2012) Links between metabolism and cancer. Genes Dev 26:877–890
31. Ehrlich P (1909) Über den jetzigen stand der karzinomforschung. Ned Tijdschr Geneeskd 5:273–290
32. Dunn GP, Bruce AT, Ikeda H, Old LJ, Schreiber RD (2002) Cancer immunoediting: from immunosurveillance to tumor escape. Nat Immunol 3:991–998
33. Shields JD, Kourtis IC, Tomei AA, Roberts JM, Swartz MA (2010) Induction of lymphoidlike stroma and immune escape by tumors that express the chemokine CCL21. Science 328:749–752
34. Yang L, Pang Y, Moses HL (2010) TGF-beta and immune cells: an important regulatory axis in the tumor microenvironment and progression. Trends Immunol 31:220–227
35. Kim R, Emi M, Tanabe K (2006) Cancer immunosuppression and autoimmune disease: beyond immunosuppressive networks for tumour immunity. Immunology 119:254–264
36. Rabinovich GA, Gabrilovich D, Sotomayor EM (2007) Immunosuppressive strategies that are mediated by tumor cells. Annu Rev Immunol 25:267–296
37. Tredan O, Galmarini CM, Patel K, Tannock IF (2007) Drug resistance and the solid tumor microenvironment. J Natl Cancer Inst 99:1441–1454
38. Harbinski F, Craig VJ, Sanghavi S, Jeffery D, Liu L, Sheppard KA, Wagner S, Stamm C, Buness A, Chatenay-Rivauday C, Yao Y, He F, Lu CX, Guagnano V, Metz T, Finan PM, Hofmann F, Sellers WR, Porter JA, Myer VE, Graus-Porta D, Wilson CJ, Buckler A, Tiedt R (2012) Rescue screens with secreted proteins reveal compensatory potential of receptor tyrosine kinases in driving cancer growth. Cancer Discov 2:948–959
39. Yao L, Zhang Y, Chen K, Hu X, Xu LX (2011) Discovery of IL-18 as a novel secreted protein contributing to doxorubicin resistance by comparative secretome analysis of MCF-7 and MCF-7/Dox. PLoS One 6:e24684

40. Takata T, Ishigaki Y, Shimasaki T, Tsuchida H, Motoo Y, Hayashi A, Tomosugi N (2012) Characterization of proteins secreted by pancreatic cancer cells with anticancer drug treatment in vitro. Oncol Rep 28:1968–1976
41. Straussman R, Morikawa T, Shee K, Barzily-Rokni M, Qian ZR, Du J, Davis A, Mongare MM, Gould J, Frederick DT, Cooper ZA, Chapman PB, Solit DB, Ribas A, Lo RS, Flaherty KT, Ogino S, Wargo JA, Golub TR (2012) Tumour micro-environment elicits innate resistance to RAF inhibitors through HGF secretion. Nature 487:500–504
42. Wilson TR, Fridlyand J, Yan Y, Penuel E, Burton L, Chan E, Peng J, Lin E, Wang Y, Sosman J, Ribas A, Li J, Moffat J, Sutherlin DP, Koeppen H, Merchant M, Neve R, Settleman J (2012) Widespread potential for growth-factor-driven resistance to anticancer kinase inhibitors. Nature 487:505–509
43. Fialkow PJ, Gartler SM, Yoshida A (1967) Clonal origin of chronic myelocytic leukemia in man. Proc Natl Acad Sci U S A 58:1468–1471
44. Clarkson B, Rubinow SI (1977) Growth kinetics in human leukemia. In: The University of Texas System Cancer Center M.D. Anderson Hospital and Tumor Institute 29th Annual Symposium on Fundamental Cancer Research, pp 592–627
45. Bonnet D, Dick JE (1997) Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. Nat Med 3:730–737
46. Al-Hajj M, Wicha MS, Benito-Hernandez A, Morrison SJ, Clarke MF (2003) Prospective identification of tumorigenic breast cancer cells. Proc Natl Acad Sci U S A 100:3983–3988
47. Lang SH, Frame FM, Collins AT (2009) Prostate cancer stem cells. J Pathol 217:299–306
48. O'Brien CA, Pollett A, Gallinger S, Dick JE (2007) A human colon cancer cell capable of initiating tumour growth in immunodeficient mice. Nature 445:106–110
49. Zhang S, Balch C, Chan MW, Lai HC, Matei D, Schilder JM, Yan PS, Huang TH, Nephew KP (2008) Identification and characterization of ovarian cancer-initiating cells from primary human tumors. Cancer Res 68:4311–4320
50. Schatton T, Murphy GF, Frank NY, Yamaura K, Waaga-Gasser AM, Gasser M, Zhan Q, Jordan S, Duncan LM, Weishaupt C, Fuhlbrigge RC, Kupper TS, Sayegh MH, Frank MH (2008) Identification of cells initiating human melanomas. Nature 451:345–349
51. Rath BH, Fair JM, Jamal M, Camphausen K, Tofilon PJ (2013) Astrocytes enhance the invasion potential of glioblastoma stem-like cells. PLoS One 8:e54752
52. Emmink BL, Verheem A, Van Houdt WJ, Steller EJ, Govaert KM, Pham TV, Piersma SR, Borel Rinkes IH, Jimenez CR, Kranenburg O (2013) The secretome of colon cancer stem cells contains drug-metabolizing enzymes. J Proteomics 91:84–96
53. Skalnikova H, Motlik J, Gadher SJ, Kovarova H (2011) Mapping of the secretome of primary isolates of mammalian cells, stem cells and derived cell lines. Proteomics 11:691–708
54. Li H, Fan X, Houghton J (2007) Tumor microenvironment: the role of the tumor stroma in cancer. J Cell Biochem 101:805–815
55. Hanahan D, Coussens LM (2012) Accessories to the crime: functions of cells recruited to the tumor microenvironment. Cancer Cell 21:309–322
56. Celis JE, Gromov P, Cabezon T, Moreira JM, Ambartsumian N, Sandelin K, Rank F, Gromova I (2004) Proteomic characterization of the interstitial fluid perfusing the breast tumor microenvironment: a novel resource for biomarker and therapeutic target discovery. Mol Cell Proteomics 3:327–344
57. Celis JE, Gromov P, Moreira JM, Cabezon T, Friis E, Vejborg IM, Proess G, Rank F, Gromova I (2006) Apocrine cysts of the breast: biomarkers, origin, enlargement, and relation with cancer phenotype. Mol Cell Proteomics 5:462–483
58. Celis JE, Moreira JM, Cabezon T, Gromov P, Friis E, Rank F, Gromova I (2005) Identification of extracellular and intracellular signaling components of the mammary adipose tissue and its interstitial fluid in high risk breast cancer patients: toward dissecting the molecular circuitry of epithelial-adipocyte stromal cell interactions. Mol Cell Proteomics 4:492–522
59. Karagiannis GS, Pavlou MP, Diamandis EP (2010) Cancer secretomics reveal pathophysiological pathways in cancer molecular oncology. Mol Oncol 4:496–510

60. Ngounou Wetie AG, Sokolowska I, Woods AG, Wormwood KL, Dao S, Patel S, Clarkson BD, Darie CC (2013) Automated mass spectrometry-based functional assay for the routine analysis of the secretome. J Lab Autom 18:19–29
61. Darie CC (2013) Mass spectrometry and its applications in life sciences. Aust J Chem 66:719–720
62. Ngounou Wetie AG, Sokolowska I, Woods AG, Roy U, Loo JA, Darie CC (2013) Investigation of stable and transient protein–protein interactions: past, present, and future. Proteomics 13:538–557
63. Sokolowska I, Wetie AGN, Woods AG, Darie CC (2013) Applications of mass spectrometry in proteomics. Aust J Chem 66:721–733
64. Ngounou Wetie AG, Sokolowska I, Woods AG, Darie CC (2013) Identification of post-translational modifications by mass spectrometry. Aust J Chem 66:734–748
65. Ngounou Wetie AG, Sokolowska I, Woods AG, Roy U, Deinhardt K, Darie CC (2014) Protein-protein interactions: switch from classical methods to proteomics and bioinformatics-based approaches. Cell Mol Life Sci 71(2):205–228
66. Sokolowska I, Gawinowicz MA, Ngounou Wetie AG, Darie CC (2012) Disulfide proteomics for identification of extracellular or secreted proteins. Electrophoresis 33:2527–2536
67. Sokolowska I, Ngounou Wetie AG, Roy U, Woods AG, Darie CC (2013) Mass spectrometry investigation of glycosylation on the NXS/T sites in recombinant glycoproteins. Biochim Biophys Acta 1834:1474–1483
68. Sokolowska I, Ngounou Wetie AG, Woods AG, Darie CC (2012) Automatic determination of disulfide bridges in proteins. J Lab Autom 17:408–416
69. Sokolowska I, Woods AG, Wagner J, Dorler J, Wormwood K, Thome J, Darie CC (2011) Mass spectrometry for proteomics-based investigation of oxidative stress and heat shock proteins. In: Andreescu S, Hepel M (eds) Oxidative stress: diagnostics, prevention, and therapy. American Chemical Society, Washington, DC
70. Woods AG, Ngounou Wetie AG, Sokolowska I, Russell S, Ryan JP, Michel TM, Thome J, Darie CC (2013) Mass spectrometry as a tool for studying autism spectrum disorder. J Mol Psychiatry 1:6
71. Woods AG, Sokolowska I, Darie CC (2012) Identification of consistent alkylation of cysteine-less peptides in a proteomics experiment. Biochem Biophys Res Commun 419:305–308
72. Woods AG, Sokolowska I, Yakubu R, Butkiewicz M, LaFleur M, Talbot C, Darie CC (2011) Blue native page and mass spectrometry as an approach for the investigation of stable and transient protein-protein interactions. In: Andreescu S, Hepel M (eds) Oxidative stress: diagnostics, prevention, and therapy. American Chemical Society, Washington, DC
73. Petrareanu C, Macovei A, Sokolowska I, Woods AG, Lazar C, Radu GL, Darie CC, Branza-Nichita N (2013) Comparative proteomics reveals novel components at the plasma membrane of differentiated HepaRG cells and different distribution in hepatocyte-and biliary-like cells. PLoS One 8:e71859
74. Sokolowska I, Dorobantu C, Woods AG, Macovei A, Branza-Nichita N, Darie CC (2012) Proteomic analysis of plasma membranes isolated from undifferentiated and differentiated HepaRG cells. Proteome Sci 10:47
75. Sokolowska I, Woods AG, Gawinowicz MA, Roy U, Darie CC (2012) Identification of a potential tumor differentiation factor receptor candidate in prostate cancer cells. FEBS J 279:2579–2594
76. Xiao T, Ying W, Li L, Hu Z, Ma Y, Jiao L, Ma J, Cai Y, Lin D, Guo S, Han N, Di X, Li M, Zhang D, Su K, Yuan J, Zheng H, Gao M, He J, Shi S, Li W, Xu N, Zhang H, Liu Y, Zhang K, Gao Y, Qian X, Cheng S (2005) An approach to studying lung cancer-related proteins in human blood. Mol Cell Proteomics 4:1480–1486
77. Pardo M, Garcia A, Antrobus R, Blanco MJ, Dwek RA, Zitzmann N (2007) Biomarker discovery from uveal melanoma secretomes: identification of gp100 and cathepsin D in patient serum. J Proteome Res 6:2802–2811
78. Xue H, Lu B, Lai M (2008) The cancer secretome: a reservoir of biomarkers. J Transl Med 6:52

79. Yao L, Lao W, Zhang Y, Tang X, Hu X, He C, Hu X, Xu LX (2012) Identification of EFEMP2 as a serum biomarker for the early detection of colorectal cancer with lectin affinity capture assisted secretome analysis of cultured fresh tissues. J Proteome Res 11(6):3281–3294
80. Stastna M, Van Eyk JE (2012) Investigating the secretome: lessons about the cells that comprise the heart. Circ Cardiovasc Genet 5:o8–o18
81. Stastna M, Van Eyk JE (2012) Secreted proteins as a fundamental source for biomarker discovery. Proteomics 12:722–735
82. Pavlou MP, Diamandis EP (2010) The cancer cell secretome: a good source for discovering biomarkers? J Proteomics 73:1896–1906
83. Caccia D, Dugo M, Callari M, Bongarzone I (2013) Bioinformatics tools for secretome analysis. Biochim Biophys Acta 1834:2442–2453
84. Emanuelsson O, Brunak S, von Heijne G, Nielsen H (2007) Locating proteins in the cell using TargetP, SignalP and related tools. Nat Protoc 2:953–971
85. Lawlor K, Nazarian A, Lacomis L, Tempst P, Villanueva J (2009) Pathway-based biomarker search by high-throughput proteomics profiling of secretomes. J Proteome Res 8:1489–1503
86. Mukherjee P, Mani S (2013) Methodologies to decipher the cell secretome. Biochim Biophys Acta 1834:2226–2232
87. Clark HF, Gurney AL, Abaya E, Baker K, Baldwin D, Brush J, Chen J, Chow B, Chui C, Crowley C, Currell B, Deuel B, Dowd P, Eaton D, Foster J, Grimaldi C, Gu Q, Hass PE, Heldens S, Huang A, Kim HS, Klimowski L, Jin Y, Johnson S, Lee J, Lewis L, Liao D, Mark M, Robbie E, Sanchez C, Schoenfeld J, Seshagiri S, Simmons L, Singh J, Smith V, Stinson J, Vagts A, Vandlen R, Watanabe C, Wieand D, Woods K, Xie MH, Yansura D, Yi S, Yu G, Yuan J, Zhang M, Zhang Z, Goddard A, Wood WI, Godowski P, Gray A (2003) The secreted protein discovery initiative (SPDI), a large-scale effort to identify novel human secreted and transmembrane proteins: a bioinformatics assessment. Genome Res 13:2265–2270
88. Brown KJ, Formolo CA, Seol H, Marathi RL, Duguez S, An E, Pillai D, Nazarian J, Rood BR, Hathout Y (2012) Advances in the proteomic investigation of the cell secretome. Expert Rev Proteomics 9:337–345
89. Mbeunkui F, Fodstad O, Pannell LK (2006) Secretory protein enrichment and analysis: an optimized approach applied on cancer cell lines using 2D LC-MS/MS. J Proteome Res 5:899–906
90. Ge S, Mao Y, Yi Y, Xie D, Chen Z, Xiao Z (2012) Comparative proteomic analysis of secreted proteins from nasopharyngeal carcinoma-associated stromal fibroblasts and normal fibroblasts. Exp Ther Med 3:857–860
91. DeKroon RM, Osorio C, Robinette JB, Mocanu M, Winnik WM, Alzate O (2011) Simultaneous detection of changes in protein expression and oxidative modification as a function of age and APOE genotype. J Proteome Res 10:1632–1644
92. DeKroon RM, Robinette JB, Osorio C, Jeong JS, Hamlett E, Mocanu M, Alzate O (2012) Analysis of protein posttranslational modifications using DIGE-based proteomics. Methods Mol Biol 854:129–143
93. Winnik WM, Dekroon RM, Jeong JS, Mocanu M, Robinette JB, Osorio C, Dicheva NN, Hamlett E, Alzate O (2012) Analysis of proteins using DIGE and MALDI mass spectrometry. Methods Mol Biol 854:47–66
94. Jin L, Zhang Y, Li H, Yao L, Fu D, Yao X, Xu LX, Hu X, Hu G (2012) Differential secretome analysis reveals CST6 as a suppressor of breast cancer bone metastasis. Cell Res 22:1356–1373
95. Mauri P, Scarpa A, Nascimbeni AC, Benazzi L, Parmagnani E, Mafficini A, Della Peruta M, Bassi C, Miyazaki K, Sorio C (2005) Identification of proteins released by pancreatic cancer cells by multidimensional protein identification technology: a strategy for identification of novel cancer markers. FASEB J 19:1125–1127
96. Colangelo CM, Williams KR (2006) Isotope-coded affinity tags for protein quantification. Methods Mol Biol 328:151–158
97. Qu J, Jusko WJ, Straubinger RM (2006) Utility of cleavable isotope-coded affinity-tagged reagents for quantification of low-copy proteins induced by methylprednisolone using liquid chromatography/tandem mass spectrometry. Anal Chem 78:4543–4552

98. Yi EC, Li XJ, Cooke K, Lee H, Raught B, Page A, Aneliunas V, Hieter P, Goodlett DR, Aebersold R (2005) Increased quantitative proteome coverage with (13)C/(12)C-based, acid-cleavable isotope-coded affinity tag reagent and modified data acquisition scheme. Proteomics 5:380–387
99. Ong SE, Blagoev B, Kratchmarova I, Kristensen DB, Steen H, Pandey A, Mann M (2002) Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. Mol Cell Proteomics 1:376–386
100. Unwin RD, Griffiths JR, Whetton AD (2010) Simultaneous analysis of relative protein expression levels across multiple samples using iTRAQ isobaric tags with 2D nano LC-MS/MS. Nat Protoc 5:1574–1582
101. Hutchens TW, Yip TT (1993) New desorption strategies for the mass spectrometric analysis of macromolecules. Rapid Commun Mass Spectrom 7:576–580
102. Issaq HJ, Veenstra TD, Conrads TP, Felschow D (2002) The SELDI-TOF MS approach to proteomics: protein profiling and biomarker identification. Biochem Biophys Res Commun 292:587–592
103. Thirant C, Galan-Moya EM, Dubois LG, Pinte S, Chafey P, Broussard C, Varlet P, Devaux B, Soncin F, Gavard J, Junier MP, Chneiweiss H (2012) Differential proteomic analysis of human glioblastoma and neural stem cells reveals HDGF as a novel angiogenic secreted factor. Stem Cells 30:845–853
104. Tang CE, Guan YJ, Yi B, Li XH, Liang K, Zou HY, Yi H, Li MY, Zhang PF, Li C, Peng F, Chen ZC, Yao KT, Xiao ZQ (2010) Identification of the amyloid beta-protein precursor and cystatin C as novel epidermal growth factor receptor regulated secretory proteins in nasopharyngeal carcinoma by proteomics. J Proteome Res 9:6101–6111
105. Makridakis M, Roubelakis MG, Bitsika V, Dimuccio V, Samiotaki M, Kossida S, Panayotou G, Coleman J, Candiano G, Anagnou NP, Vlahou A (2010) Analysis of secreted proteins for the study of bladder cancer cell aggressiveness. J Proteome Res 9:3243–3259
106. Rondepierre F, Bouchon B, Bonnet M, Moins N, Chezal JM, D'Incan M, Degoul F (2010) B16 melanoma secretomes and in vitro invasiveness: syntenin as an invasion modulator. Melanoma Res 20:77–84
107. Gromov P, Gromova I, Bunkenborg J, Cabezon T, Moreira JM, Timmermans-Wielenga V, Roepstorff P, Rank F, Celis JE (2010) Up-regulated proteins in the fluid bathing the tumour cell microenvironment as potential serological markers for early detection of cancer of the breast. Mol Oncol 4:65–89
108. Shi HJ, Stubbs R, Hood K (2009) Characterization of de novo synthesized proteins released from human colorectal tumour explants. Electrophoresis 30:2442–2453
109. Volmer MW, Radacz Y, Hahn SA, Klein-Scory S, Stuhler K, Zapatka M, Schmiegel W, Meyer HE, Schwarte-Waldhoff I (2004) Tumor suppressor Smad4 mediates downregulation of the anti-adhesive invasion-promoting matricellular protein SPARC: Landscaping activity of Smad4 as revealed by a "secretome" analysis. Proteomics 4:1324–1334
110. Garcia-Lorenzo A, Rodriguez-Pineiro AM, Rodriguez-Berrocal FJ, Cadena MP, Martinez-Zorzano VS (2012) Changes on the Caco-2 secretome through differentiation analyzed by 2-D differential in-gel electrophoresis (DIGE). Int J Mol Sci 13:14401–14420
111. Yousefi Z, Sarvari J, Nakamura K, Kuramitsu Y, Ghaderi A, Mojtahedi Z (2012) Secretomic analysis of large cell lung cancer cell lines using two-dimensional gel electrophoresis coupled to mass spectrometry. Folia Histochem Cytobiol 50:368–374
112. Volmer MW, Stuhler K, Zapatka M, Schoneck A, Klein-Scory S, Schmiegel W, Meyer HE, Schwarte-Waldhoff I (2005) Differential proteome analysis of conditioned media to detect Smad4 regulated secreted biomarkers in colon cancer. Proteomics 5:2587–2601
113. Zeng X, Yang P, Chen B, Jin X, Liu Y, Zhao X, Liang S (2013) Quantitative secretome analysis reveals the interactions between epithelia and tumor cells by in vitro modulating colon cancer microenvironment. J Proteomics 89:51–70
114. Holmberg C, Ghesquiere B, Impens F, Gevaert K, Kumar JD, Cash N, Kandola S, Hegyi P, Wang TC, Dockray GJ, Varro A (2013) Mapping proteolytic processing in the secretome of

gastric cancer-associated myofibroblasts reveals activation of MMP-1, MMP-2, and MMP-3. J Proteome Res 12:3413–3422
115. Barderas R, Mendes M, Torres S, Bartolome RA, Lopez-Lucendo M, Villar-Vazquez R, Pelaez-Garcia A, Fuente E, Bonilla F, Casal JI (2013) In-depth characterization of the secretome of colorectal cancer metastatic cells identifies key proteins in cell adhesion, migration, and invasion. Mol Cell Proteomics 12:1602–1620
116. Marimuthu A, Subbannayya Y, Sahasrabuddhe NA, Balakrishnan L, Syed N, Sekhar NR, Katte TV, Pinto SM, Srikanth SM, Kumar P, Pawar H, Kashyap MK, Maharudraiah J, Ashktorab H, Smoot DT, Ramaswamy G, Kumar RV, Cheng Y, Meltzer SJ, Roa JC, Chaerkady R, Prasad TS, Harsha HC, Chatterjee A, Pandey A (2013) SILAC-based quantitative proteomic analysis of gastric cancer secretome. Proteomics Clin Appl 7:355–366
117. Boersema PJ, Geiger T, Wisniewski JR, Mann M (2013) Quantification of the N-glycosylated secretome by super-SILAC during breast cancer progression and in human blood samples. Mol Cell Proteomics 12:158–171
118. Chen CY, Chi LM, Chi HC, Tsai MM, Tsai CY, Tseng YH, Lin YH, Chen WJ, Huang YH, Lin KH (2012) Stable isotope labeling with amino acids in cell culture (SILAC)-based quantitative proteomics study of a thyroid hormone-regulated secretome in human hepatoma cells. Mol Cell Proteomics 11(M111):011270
119. Formolo CA, Williams R, Gordish-Dressman H, MacDonald TJ, Lee NH, Hathout Y (2011) Secretome signature of invasive glioblastoma multiforme. J Proteome Res 10:3149–3159
120. Kashyap MK, Harsha HC, Renuse S, Pawar H, Sahasrabuddhe NA, Kim MS, Marimuthu A, Keerthikumar S, Muthusamy B, Kandasamy K, Subbannayya Y, Prasad TS, Mahmood R, Chaerkady R, Meltzer SJ, Kumar RV, Rustgi AK, Pandey A (2010) SILAC-based quantitative proteomic approach to identify potential biomarkers from the esophageal squamous cell carcinoma secretome. Cancer Biol Ther 10:796–810
121. Xu BJ, Yan W, Jovanovic B, An AQ, Cheng N, Aakre ME, Yi Y, Eng J, Link AJ, Moses HL (2010) Quantitative analysis of the secretome of TGF-beta signaling-deficient mammary fibroblasts. Proteomics 10:2458–2470
122. Gronborg M, Kristiansen TZ, Iwahori A, Chang R, Reddy R, Sato N, Molina H, Jensen ON, Hruban RH, Goggins MG, Maitra A, Pandey A (2006) Biomarker discovery from pancreatic cancer secretome using a differential proteomic approach. Mol Cell Proteomics 5:157–171
123. Chang YH, Lee SH, Chang HC, Tseng YL, Lai WW, Liao CC, Tsay YG, Liao PC (2012) Comparative secretome analyses using a hollow fiber culture system with label-free quantitative proteomics indicates the influence of PARK7 on cell proliferation and migration/invasion in lung adenocarcinoma. J Proteome Res 11:5167–5185
124. Chen HL, Seol H, Brown KJ, Gordish-Dressman H, Hill A, Gallo V, Packer R, Hathout Y (2012) Secretome survey of human plexiform neurofibroma derived Schwann cells reveals a secreted form of the RARRES1 protein. Int J Mol Sci 13:9380–9399
125. Choi DS, Choi DY, Hong BS, Jang SC, Kim DK, Lee J, Kim YK, Kim KP, Gho YS (2012) Quantitative proteomics of extracellular vesicles derived from human primary and metastatic colorectal cancer cells. J Extracell Vesicles:1–15, 18704
126. Schiarea S, Solinas G, Allavena P, Scigliuolo GM, Bagnati R, Fanelli R, Chiabrando C (2010) Secretome analysis of multiple pancreatic cancer cell lines reveals perturbations of key functional networks. J Proteome Res 9:4376–4392
127. Xue H, Lu B, Zhang J, Wu M, Huang Q, Wu Q, Sheng H, Wu D, Hu J, Lai M (2010) Identification of serum biomarkers for colorectal cancer metastasis using a differential secretome approach. J Proteome Res 9:545–555

Chapter 21
Thiostrepton, a Natural Compound That Triggers Heat Shock Response and Apoptosis in Human Cancer Cells: A Proteomics Investigation

Cristinel Sandu, Armand G. Ngounou Wetie, Costel C. Darie, and Hermann Steller

Abstract Thiostrepton is a natural antibiotic produced by bacteria of *Streptomyces* genus. We identified Thiostrepton as a strong hit in a cell-based small molecule screen for DIAP1 stability modulators. It was shown previously that Thiostrepton induces upregulation of several gene products in *Streptomyces lividans*, including the TipAS and TipAL isoforms, and that it can induce apoptotic cell death in human cancer cells. Furthermore, it was suggested that thiostrepton induces oxidative and proteotoxic stress, as inferred from the transcriptional upregulation of stress-related genes and endoplasmic reticulum (ER) stress genes. We used a combination of biochemical and proteomics approaches to investigate the effect of Thiostrepton and other compounds in human cells. Our mass-spectrometry data and subsequent biochemical validation shows that Thiostrepton (and MG-132 proteasome inhibitor) trigger upregulation of heat shock proteins HspA1A, Hsp70, Hsp90α, or Hsp105 in various human cancer cells. We propose a model where Thiostrepton-induced proteasome inhibition leads to accumulation of protein aggregates that trigger a heat shock response and apoptosis in human cancer cells.

C. Sandu • H. Steller (✉)
Howard Hughes Medical Institute, Strang Laboratory of Apoptosis and Cancer Biology, The Rockefeller University, New York, NY 10065, USA
e-mail: steller@rockefeller.edu

A.G. Ngounou Wetie • C.C. Darie
Biochemistry and Proteomics Group, Department of Chemistry and Biomolecular Science, Clarkson University, Potsdam, NY 13699, USA

A.G. Woods and C.C. Darie (eds.), *Advancements of Mass Spectrometry in Biomedical Research*, Advances in Experimental Medicine and Biology 806,
DOI 10.1007/978-3-319-06068-2_21, © Springer International Publishing Switzerland 2014

Abbreviations

DIAP1	*Drosophila* inhibitor of apoptosis protein 1
DMSO	Dimethyl sulphoxide
m/z	Mass/charge
MS	Mass spectrometry
NAC	*N*-acetylcysteine
nanoLC-MS/MS	Nanoliquid chromatography tandem mass spectrometry
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
WB	Western blotting

21.1 Introduction

Thiostrepton is a natural antibiotic produced by bacteria of *Streptomyces* genus [1–3]. In our laboratory, Thiostrepton was identified as a strong hit in a cell-based small molecule screen for *Drosophila* inhibitor of apoptosis protein 1 (DIAP1) stability modulators. Thiostrepton was previously identified as a hit in a cell-based screen for compounds that stabilize a mutated von Hippel-Lindau (VHL) protein [4]. Its antibiotic function is explained by the ability to block protein synthesis in bacteria, through binding to and inhibiting the activity of the ribosome (reviewed in [5]). It was shown that Thiostrepton treatment induces the upregulation of several gene products in *Streptomyces lividans*, including the TipAS and TipAL isoforms. TipAS binds thiostrepton covalently as part of *Streptomyces* defense mechanism against this antibiotic. Adduct formation implicates one or more dehydroalanine side chains in Thiostrepton and a cysteine residue in TipAS [6]. In mammals, Thiostrepton is used as a topical medication in veterinary medicine for treatment of mastitis caused by gram-negative bacteria. Thiostrepton was shown more recently to induce cell death in various cancer cell types, including breast [7, 8], melanoma [9, 10], leukemia [11], liver cancer [11], or malignant mesothelioma [12]. It was suggested that the mechanism of Thiostrepton toxicity in human cancer cells is caused by specific depletion of oncogenic transcription factor Forkhead Box Protein M1 (FoxM1), which is overexpressed in human cancer cells [8, 13]. Recently, it was shown that thiostrepton interacts with FoxM1 in vitro, albeit with micromolar affinity [14]. Thiostrepton was reported also as an inhibitor of the 20S proteasome chemotrypsin activity [10, 15]. Furthermore, it was suggested that Thiostrepton induces oxidative and proteotoxic stress, as inferred from the transcriptional upregulation of heat shock, oxidative stress, and endoplasmic reticulum (ER) stress genes [10]. Oxidative and proteotoxic stress effects were relieved by treatment with anti-oxidant *N*-acetylcysteine (NAC). Finally, it was suggested that the mechanism of oxidative stress induced by Thiostrepton is caused by its adduct formation with mitochondrial peroxiredoxin-3 (PRX3), as inferred from a protein mobility shift change and was proposed to act by disabling an important anti-oxidative network [12].

21.2 Materials and Methods

All compounds in this work, unless otherwise specified were dissolved in dimethyl sulphoxide (DMSO). Compounds were purchased from commercial inventories as follows: MG-132 (Calbiochem, USA), Thiostrepton (Tocris Biosciences, MO, USA). IAP-antagonist (Compound-3) was kindly provided by Dr. Patrick G. Harran, UCLA or synthesized by Ouathek Ouerfelli and Barney at the Organic Synthesis Core Facility (OSCF) of the MSKCC. The antibodies used in this work were purchased as follows: rabbit anti-cleaved PARP (New England Biolabs, Ipswich, MA, USA), mouse anti-β–Actin-HRP (Sigma-Aldrich, St. Louis, Mo, USA), mouse anti-HspA1A (LifeSpan Biosciences, Inc., Seattle, WA, USA), mouse anti-Hsp70 (Stressgen, MI, USA), mouse anti-p53 (Santa Cruz Biotechnology, Inc., Dallas, TX, USA), mouse anti-XIAP (BD Transduction Laboratories, San Jose, CA, USA), mouse anti-Usp9X (Novus Biologicals, LLC., Littleton, Co, USA), mouse anti-FoxM1 (Abcam, Cambridge, MA, USA).

HEK293 cells and melanoma cell lines A375, UACC-257, MeWo, SK-Mel-5 and Malme-3M were grown in DMEM medium supplemented with 10 % FBS, and antibiotics. Human epidermal melanocytes HEMa-LP were purchased from Cascade Biologics (Carlsbad, CA, USA) and cultures as suggested by the provider in Medium 254 supplemented with human melanocytes growth supplement (HMGS). To test the effect of compounds on the cell lines mentioned above, 7.5 million cells were seeded in a 10 cm dish and treated with DMSO, Thiostrepton, MG-132 or IAP-antagonist resuspended in DMSO. After 21 h treatment, the cells were harvested and lysed in Lysis buffer (10 mM Tris, pH 8.0, 100 mM NaCl, 0.5 % NP-40). 15 µg cell extract was used for SDS-PAGE and WB detection. SDS-PAGE and WB were performed as previously described [16, 17]. NanoLC-MS/MS analysis was performed as in [18–20]. MS-based relative quantitation was performed as in [21].

21.3 Results and Discussion

In order to understand its mechanisms of action in human cells, we investigated Thiostrepton in numerous assays pertaining to different molecular pathways. In one such assay we treated human embryonic kidney fibroblasts (HEK293 cells) overnight with DMSO (control), Thiostrepton, a peptidomimetic IAP-antagonist as well as a proteasome inhibitor (MG-132) and inspected by SDS-PAGE eventual changes in the protein pattern. Interestingly, in Thiostrepton-treated sample we observed the appearance of a distinctly over-expressed band at approximately 70 kDa, as well as two other bands less pronounced at approximately 90 and 100 kDa, when compared to DMSO treatment (Fig. 21.1). In addition, these bands were also increased in the MG-132 treated samples. Because MG-132 is a proteasome inhibitor and Thiostrepton has been previously linked to proteasome inhibition, an explanation for these observations is that these upregulated bands are labile proteasome

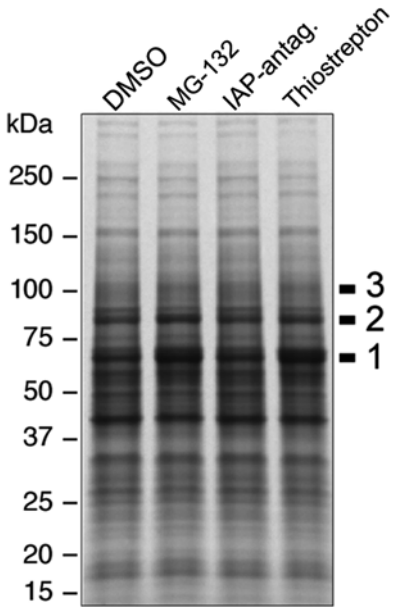


Fig. 21.1 Thiostrepton and MG-132 proteasome inhibitor trigger upregulation of certain proteins in HEK293 cells. Coomassie-stained SDS-PAGE gel showing the pattern of proteins in extracts of HEK293 cells treated with DMSO, MG-132 (5 μM), IAP-antagonist (20 μM), and Thiostrepton (10 μM) for 21 h. Certain bands, labeled as 1–3 appear upregulated in MG-132 and Thiostrepton-treated samples

substrates that are stabilized following proteasome inhibition. To determine the nature of these proteins, we excised the bands 1–3 from the Coomassie-stained SDS-PAGE gel and analyzed them by nanoliquid chromatography-tandem mass spectrometry (nanoLC-MS/MS). Data analysis revealed that the protein in Band 1 is heat shock protein HspA1A, the protein in band 2 is heat shock protein HSP90AA1, and the protein in band 3 is heat shock protein HSP105. Examples of MS/MS spectra whose analysis led to identification of peptides which were part of HSPA1A, HSP90AA1, or HSP105 proteins are shown in Fig. 21.2. Quantitative analysis of HSPA1A, HSP90AA1, and HSP105 revealed an increase in the levels of all these proteins in the MG-132- and Thiostrepton-treated cells, as compared with the DMSO-treated cells, in agreement with the data shown in Fig. 21.1. However, upregulation of HSPs was not triggered by the peptidomimetic IAP-antagonist, as its profile is very similar to that of DMSO samples (Fig. 21.3). Taken together, these data suggest that MG-132 and Thiostrepton, but not IAP-antagonist, stimulate the heat shock response, manifested through increased levels of HSPs.

To further validate the MS findings, we employed antibodies for HspA1A and Hsp70 to probe by Western blot (WB) for upregulation of these proteins under the compounds' treatment. Indeed, extracts of HEK293 cells treated with Thiostrepton or MG-131 showed a significant accumulation of HspA1A and Hsp70 proteins (Fig. 21.4), thus confirming the data in Figs. 21.1, 21.2, and 21.3. Besides the effect on heat shock protein response, Thiostrepton and MG-132 trigger apoptosis in HEK293 cells, as detected by the appearance of cleaved PARP (cPARP) [22]. In addition, we observed a dramatic decrease in p53 and XIAP levels in HEK293 cells

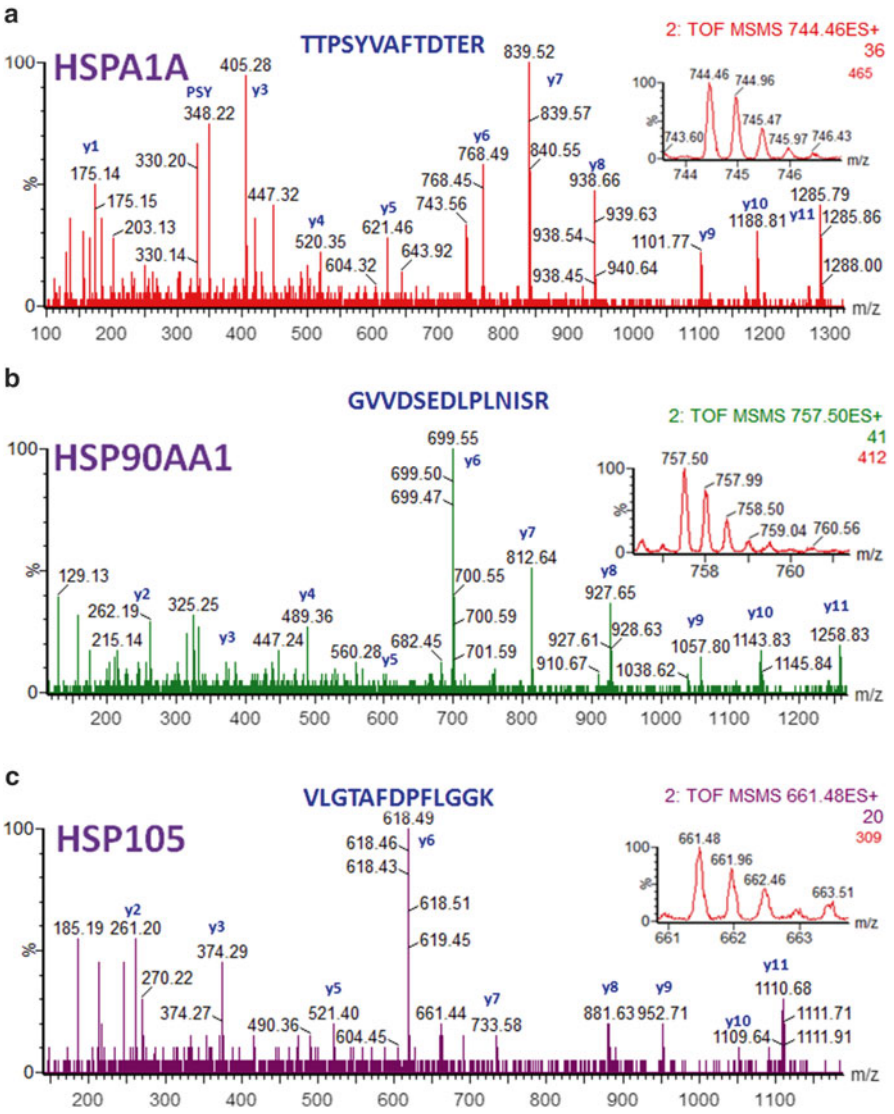


Fig. 21.2 NanoLC-MS/MS analysis of the SDS-PAGE gel bands 1, 2, and 3. The Coomassie-stained gel bands were excised, digested with trypsin, and analyzed by nanoLC-MS/MS. The MS spectra that contain doubly charged precursor ions with mass/charge ratio (*m*/*z*) of 744.46 (**a**), 757.50 (**b**), and 661.48 (**c**) were fragmented by MS/MS and produced a series of y ions, whose analysis led to identification of peptides with the amino acid sequence TTPSYVAFTDTER (**a**), GVVDSEDLPLNISR (**b**), and VLGTAFDPFLGGK (**c**), which were parts of HSPA9A (**a**), HSP90AA1 (**b**), and HSP105 (**c**)

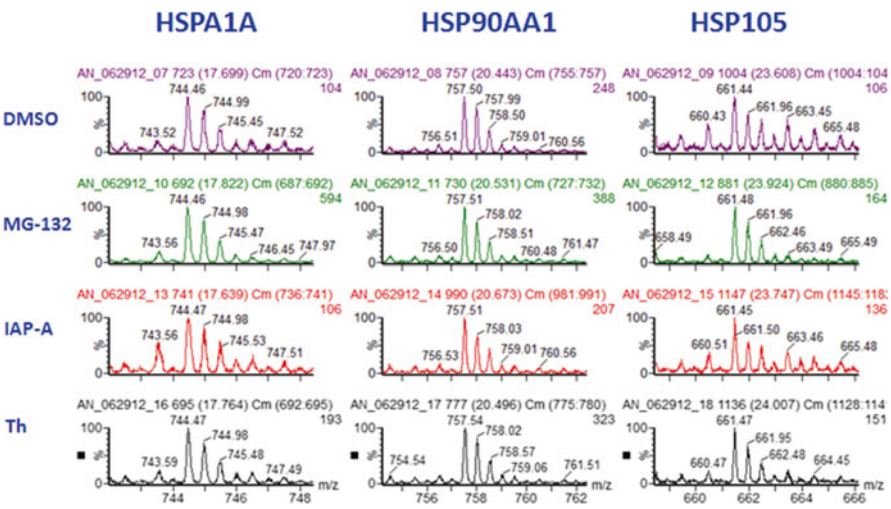


Fig. 21.3 Relative quantitation of the HSPs. Relative quantitation of the intensity of the precursor ions shown in Fig. 21.2, which were from the gel bands treated with DMSO, MG-132, IAP-antagonist (IAP-A), and Thiostrepton (Th). The number of counts for each precursor ion is indicated

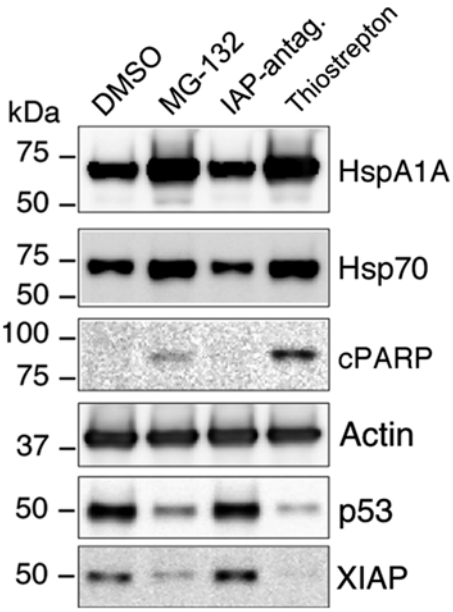


Fig. 21.4 Thiostrepton and MG-132 trigger heat shock response and apoptosis in HEK293 cells. Western blot detection of HspA1A, Hsp70, cleaved PARP (cPARP), Actin, p53, and XIAP in extracts of HEK293 cells treated with DMSO, MG-132 (5 μM), IAP-antagonist (20 μM), and Thiostrepton (10 μM) for 21 h

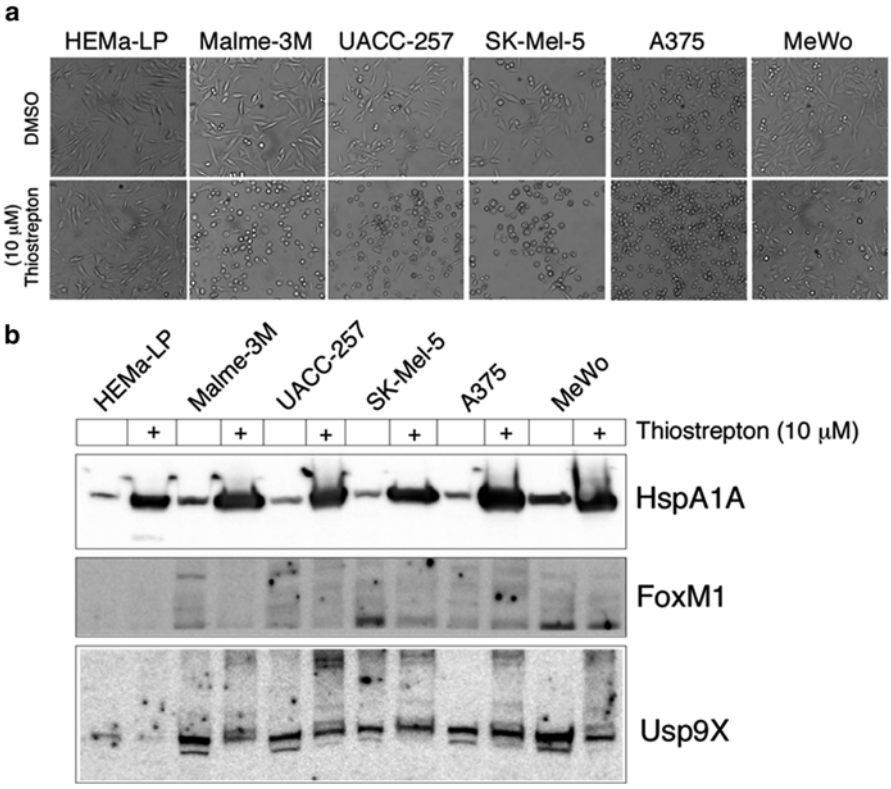


Fig. 21.5 Thiostrepton induce cell death, heat shock response, and downregulation of proteins in melanoma. (**a**) Light microscopy micrographs of wt melanocytes (HEMa-LP) and melanoma cell lines (MALME-3M, UACC-257, SK-Mel-5, A375 and MeWo) treated for 21 h with either DMSO (control) or 10 μM Thiostrepton. (**b**) Western blot detection of HspA1A, FoxM1, and Usp9X in extract of melanocytes and melanoma cells treated as in panel (**a**)

treated with Thiostrepton or MG-132. During apoptosis there are major changes in the cell in terms of not only protein profile but also transcriptional profile. It is a common fact, that under apoptotic conditions certain proteins in the cells are degraded, as is the case here with p53 and XIAP.

Next we wanted to establish whether the observations in HEK293 can be extended to cancer cells. For this purpose we analyzed a panel of melanoma cell lines (Malme-3, UACC-257, SK-Mel5, A375, and MeWo). As a control for the melanoma cells, we used Normal Epidermal Fibroblasts (HEMa-LP cells). Interestingly treatment with 10 μM Thiostrepton for 21 h led to cell death in all cancer lines with the exception of normal cells HEMa-LP, as can be visualized by changes in cell morphology (Fig. 21.5a). Next, we investigated the effect of Thiostrepton on heat shock response in normal melanocytes and melanoma cells by WB. Interestingly, Thiostrepton triggers a heat shock response in melanoma cells

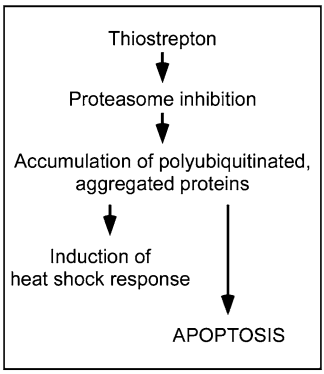


Fig. 21.6 Model of Thiostrepton-induced heat shock protein response and apoptosis in human cells

as well, as can be visualized by the significant increase in HspA1A (Fig. 21.5b). Besides the heat shock response, we investigated the Thiostrepton effects on other proteins in the cell such as FoxM1 and Usp9X. Both proteins were found to be upregulated in cancer cells when compared to melanocytes HEMa-LP and treatment of Thiostrepton leads, beside cell death, to a decrease in both proteins. This is also consistent with the observed decrease in p53 and XIAP in HEK293 cells (Fig. 21.3). Thus we have shown that Thiostrepton induces cell death in melanoma cells, triggers heat shock response and loss of many proteins from the cells (i.e., p53, XIAP, FoxM1, Usp9X).

Taken together, our data suggests that Thiostrepton (like MG-132) inhibits protein degradation and thereby causes heat shock induction (Fig. 21.6). It is well documented that proteasome inhibition leads to accumulation of protein aggregates and activation of chaperones, and that prolonged proteotoxic stress promotes apoptosis. Although Thiostrepton triggers a heat shock response in both normal and cancer cells (Fig. 21.5b), it appears that normal cells have a greater capacity to survive under these conditions whereas cancer cells undergo apoptosis (Fig. 21.5a).

Acknowledgements H. Steller is an Investigator of the Howard Hughes Medical Institute. Part of this work was supported by NIH grant R01GM60124 to H.S. and a grant from Melanoma Research Alliance. Part of this work was supported by the U.S. Army Research Office through the Defense University Research Instrumentation Program (DURIP grant #W911NF-11-1-0304) to CCD. C.S. thanks Jerry Chipuk (The Mount Sinai Hospital, NY) for the kind gift of MeWo cell line. CS and CCD thank Dr. Alisa G. Woods (Clarkson University) for discussions regarding the manuscript.

References

1. Donovick R, Pagano JF, Stout HA, Weinstein MJ (1955) Thiostrepton, a new antibiotic. I. In vitro studies. Antibiot Annu 3:554–559
2. Dutcher JD, Vandeputte J (1955) Thiostrepton, a new antibiotic. II. Isolation and chemical characterization. Antibiot Annu 3:560–561
3. Jambor WP, Steinberg BA, Suydam LO (1955) Thiostrepton, a new antibiotic. III. In vivo studies. Antibiot Annu 3:562–565

4. Ding Z, German P, Bai S, Feng Z, Gao M, Si W, Sobieski MM, Stephan CC, Mills GB, Jonasch E (2012) Agents that stabilize mutated von Hippel-Lindau (VHL) protein: results of a high-throughput screen to identify compounds that modulate VHL proteostasis. J Biomol Screen 17:572–580
5. Walsh CT, Acker MG, Bowers AA (2010) Thiazolyl peptide antibiotic biosynthesis: a cascade of post-translational modifications on ribosomal nascent proteins. J Biol Chem 285: 27525–27531
6. Chiu ML, Folcher M, Griffin P, Holt T, Klatt T, Thompson CJ (1996) Characterization of the covalent binding of thiostrepton to a thiostrepton-induced protein from Streptomyces lividans. Biochemistry 35:2332–2341
7. Halasi M, Zhao H, Dahari H, Bhat UG, Gonzalez EB, Lyubimo AV, Tonetti DA, Gartel AL (2010) Thiazole antibiotics against breast cancer. Cell Cycle 9:1214–1217
8. Kwok JM, Myatt SS, Marson CM, Coombes RC, Constantinidou D, Lam EW (2008) Thiostrepton selectively targets breast cancer cells through inhibition of forkhead box M1 expression. Mol Cancer Ther 7:2022–2032
9. Bhat UG, Zipfel PA, Tyler DS, Gartel AL (2008) Novel anticancer compounds induce apoptosis in melanoma cells. Cell Cycle 7:1851–1855
10. Qiao S, Lamore SD, Cabello CM, Lesson JL, Munoz-Rodriguez JL, Wondrak GT (2012) Thiostrepton is an inducer of oxidative and proteotoxic stress that impairs viability of human melanoma cells but not primary melanocytes. Biochem Pharmacol 83:1229–1240
11. Bhat UG, Halasi M, Gartel AL (2009) Thiazole antibiotics target FoxM1 and induce apoptosis in human cancer cells. PLoS One 4:e5592
12. Newick K, Cunniff B, Preston K, Held P, Arbiser J, Pass H, Mossman B, Shukla A, Heintz N (2012) Peroxiredoxin 3 is a redox-dependent target of thiostrepton in malignant mesothelioma cells. PLoS One 7:e39404
13. Radhakrishnan SK, Bhat UG, Hughes DE, Wang IC, Costa RH, Gartel AL (2006) Identification of a chemical inhibitor of the oncogenic transcription factor forkhead box M1. Cancer Res 66:9731–9735
14. Hegde NS, Sanders DA, Rodriguez R, Balasubramanian S (2011) The transcription factor FOXM1 is a cellular target of the natural product thiostrepton. Nat Chem 3:725–731
15. Bhat UG, Halasi M, Gartel AL (2009) FoxM1 is a general target for proteasome inhibitors. PLoS One 4:e6593
16. Sandu C, Ryoo HD, Steller H (2010) Drosophila IAP antagonists form multimeric complexes to promote cell death. J Cell Biol 190:1039–1052
17. Carrington PE, Sandu C, Wei Y, Hill JM, Morisawa G, Huang T, Gavathiotis E, Werner MH (2006) The structure of FADD and its mode of interaction with procaspase-8. Mol Cell 22:599–610
18. Woods AG, Sokolowska I, Darie CC (2012) Identification of consistent alkylation of cysteine-less peptides in a proteomics experiment. Biochem Biophys Res Commun 419:305–308
19. Sokolowska I, Gawinowicz MA, Ngounou Wetie AG, Darie CC (2012) Disulfide proteomics for identification of extracellular or secreted proteins. Electrophoresis 33:2527–2536
20. Sokolowska I, Woods AG, Gawinowicz MA, Roy U, Darie CC (2012) Identification of a potential tumor differentiation factor receptor candidate in prostate cancer cells. FEBS J 279:2579–2594
21. Darie CC, Deinhardt K, Zhang G, Cardasis HS, Chao MV, Neubert TA (2011) Identifying transient protein-protein interactions in EphB2 signaling by blue native PAGE and mass spectrometry. Proteomics 11:4514–4528
22. Lazebnik YA, Kaufmann SH, Desnoyers S, Poirier GG, Earnshaw WC (1994) Cleavage of poly(ADP-ribose) polymerase by a proteinase with properties like ICE. Nature 371:346–347

Chapter 22
Using Proteomics to Unravel the Mysterious Steps of the HBV-Life-Cycle

Norica Branza-Nichita, Catalina Petrareanu, Catalin Lazar, Izabela Sokolowska, and Costel C. Darie

Abstract Infection with Hepatitis B virus (HBV) is the most common cause of liver disease in the world. Infection becomes chronic in up to 10 % of adults, with severe consequences on liver function, including inflammation, fibrosis, cirrhosis, and eventually hepatocellular carcinoma (HCC). HCC is a fast progressing disease causing the death of approximately one million patients annually; current treatment has very limited success, mainly due to late-stage diagnosis and poor screening methodologies. Therefore, unraveling the complex HBV-host cell interactions during progression of the disease is of crucial importance, not only to understand the mechanisms underlying carcinogenesis, but importantly, for the development of new biomarkers for prognostic and early diagnosis. This is an area of research strongly influenced by proteomic studies, which have benefited in the last decade from major technical improvements in accuracy of quantification and sensitivity, large-scale analysis of low-abundant proteins, such as those from clinical samples being now possible and widely applied. This work is a critical review of the impact of the proteomic studies on our current understanding of HBV-associated pathogenesis, diagnostics, and treatment.

N. Branza-Nichita, Ph.D. (✉) • C. Petrareanu
Department of Viral Glycoproteins, Institute of Biochemistry of the Romanian Academy, Splaiul Independentei, 296, Sector 6, Bucharest 060031, Romania

Department of Analytical Chemistry and Environmental Engineering, Faculty of Applied Chemistry and Materials Science, Politehnica University of Bucharest, Calea Grivitei, 132, Sector 1, Bucharest 010707, Romania
e-mail: nichita@biochim.ro

C. Lazar
Department of Viral Glycoproteins, Institute of Biochemistry of the Romanian Academy, Splaiul Independentei, 296, Sector 6, Bucharest 060031, Romania

I. Sokolowska • C.C. Darie
Biochemistry and Proteomics Group, Department of Chemistry and Biomolecular Science, Clarkson University, 8 Clarkson Avenue, Potsdam, NY 13699-5810, USA

A.G. Woods and C.C. Darie (eds.), *Advancements of Mass Spectrometry in Biomedical Research*, Advances in Experimental Medicine and Biology 806,
DOI 10.1007/978-3-319-06068-2_22, © Springer International Publishing Switzerland 2014

Abbreviations

CID	Collision-induced dissociation
HBV	Hepatitis B virus
HCC	Hepatocellular carcinoma
HCV	Hepatitis C virus
m/*z*	Mass/charge
MS	Mass spectrometry
nanoLC-MS/MS	Nanoliquid chromatography tandem mass spectrometry
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
TIC	Total ion current

22.1 Introduction

Viruses depend entirely on a host cell for their multiplication and propagation and have evolved complex mechanisms to exploit and manipulate the cell environment to ensure them continuous survival. In turn, the cell responds to the pathogen aggression by activating specific or non-specific defense pathways. Understanding the molecular details of the virus-host cell relationship is thus a difficult task, requiring a large variety of cell factors engaged in the viral life cycle or altered upon virus infection to be identified and characterized. The information gained from such studies will not only help the development of new antiviral compounds but importantly, may also assist in the early diagnosis of specific infectious diseases.

There have been many approaches taken in the last decades to investigate viral pathogenesis, including drug inhibition of signaling pathways, expression of functionally inactive dominant-negative variants of target proteins or small interfering (si) RNA-mediated knockdown. Among these techniques, the mass spectrometry (MS)-based proteomics has the advantage to provide a large-scale analysis of protein expression. Using this method, not only individual viral and cellular proteins that are functionally modified by infection are identified, but also whole signal transduction pathways and potential protein–protein interactions, thus, a comprehensive understanding of the mechanism supporting a particular condition is provided. The development of new protein purification technologies and subcellular fractionation strategies has led to a significant improvement of the MS sensitivity in detecting low-abundant or regulatory proteins trafficked between different cellular compartments. Despite these advances, only a small number of viruses and associated pathologies have been investigated in both animal and plants, using this approach.

More than 30 years after its discovery, the Hepatitis B virus (HBV) still causes one of the most frequent and serious viral disease, an estimated 350 million people being currently chronic carriers worldwide, representing a large reservoir for new infections. Chronically infected patients are at high risk to develop severe liver diseases, such as liver fibrosis, cirrhosis, and eventually hepatocellular carcinoma (HCC), one

of the most prevalent forms of cancer in humans [1, 2]. Treatment of these patients comes at a significant cost and is often inefficient, about 500,000 deaths being registered yearly as a consequence of liver complications following HBV infection.

HBV is the prototype member of the *Hepadnaviridae* family, a group of enveloped DNA viruses with a specific tropism for liver cells. While hepatocytes are the only confirmed host site for hepadnaviruses infection, other cells such as biliary epithelial or lymphocytic cells have been proposed to be a target [3]. However, as these reported extrahepatic infections are incompletely characterized or even controversial, potential implications for pathogenesis are still to be evaluated [4, 5].

The HBV genome is organized in a partially double-stranded DNA molecule of about 3.2-kb adopting a relaxed circular conformation [6]. The rather small encoding capacity of the HBV genome is compensated by the existence of four partially overlapping open reading frames (ORF) and further increased by the initiation of translation from several in-frame ATG codons, leading to the synthesis of at least seven viral polypeptides. Of these, the core and polymerase are essential for nucleocapsid assembly and genome replication, while the small (S), medium (M), and large (L) envelope proteins form the surface layer of mature virions and play different roles in infection. Establishing the exact function of the other two HBV gene products, the X protein (HBx) and e-antigen (HBeAg) in the viral life-cycle has been a subject of intense scrutiny. Clearly, HBx has pleiotropic roles in both virus replication and pathogenesis [7], by transactivating viral and cellular promoters [8], interfering with various cellular signaling pathways [9, 10] or inhibiting the nuclear DNA repair [11]. While dispensable for infection in vivo [12] HBeAg appears to play a role in inducing immune tolerance [13].

The mature HBV virion (or the Dane particle) is a spherical structure consisting of an icosahedral nucleocapsid surrounded by the lipid bilayer, which anchors the envelope proteins. The virus is highly infectious in vivo and has the unique property to assemble spherical or filamentous subviral envelope particles (SVPs) devoid of nucleocapsid, which are non-infectious and secreted in great excess over the virions [14].

There are currently two classes of approved agents against HBV infection: viral polymerase inhibitors based on nucleoside/nucleotide analogs and alpha interferon (IFN-α) that modulates both the host immune response and viral replication [15]. These therapies not only have limited success in eradicating HBV, due to the persistent type of viral replication, but are also associated with major side-effects, selection of escape mutants, and recurrence following drug withdrawal [16, 17]. Interestingly, efficient anti- HBV vaccines are available on market, however, recent studies have shown that between 5 and 10 % of the vaccinated individuals failed to develop a protective immune response and remain vulnerable to infection [18]. In addition, the vaccine is not effective in patients already infected. In this context, the development of new antivirals is necessary to fight HBV infections and future strategies may involve combination therapies targeting viral proteins other than polymerase, or cellular functions required at different steps of the viral life cycle.

The molecular basis of HBV organization and replication, as well as the mechanisms leading to chronic HBV infections are now well defined. Identification of cellular factors contributing to HBV productive infection and the virus–host

interactions that induce severe liver damage is essential for the prevention of new infections and the establishment of the best therapy to manage chronically infected patients.

The proteomics studies employed so far to investigate HBV infection have focused on several major directions: (1) comprehensive analysis of host cell proteome modifications as a consequence of HBV infection; (2) proteomic analysis of HBV-associated HCC; (3) identification of differentially expressed proteins in Hepatitis C Virus (HCV)—and HBV-originating HCC; (4) proteomics of cells permissive for HBV infection.

22.2 Analysis of Proteome Modifications in HBV-Replicating Cells

Despite being highly infectious in human patients, HBV infection occurs at very low rates in vitro. Before the discovery of the replicating HepaRG cells in 2002, primary hepatocytes were the only cells susceptible for HBV infection. However, hepatocytes are difficult to procure and have limited stability in tissue culture, which have a negative impact on the reproducibility and reliability of the experimental results. Therefore, studies of HBV pathogenesis and antiviral screenings have been performed in alternative cellular models that support high levels of viral replication.

HepG2.2.15 and HepAD38, both derived from HepG2 cells, containing copies of the HBV genome of more than 1-unit-length are accepted cellular models and have been largely used in the last two decades to study HBV-life-cycle at post-entry levels [19, 20]. In addition to supporting full HBV replication, the HepG2.2.15 cell line is also able to secrete a large amount of SVPs, HBeAg, and infectious virions [21], which indicate it as a reliable model to investigate host–virus interactions.

To reveal cellular alterations induced by HBV, most proteomics studies involved a comparative analysis of HepG2.2.15 and HepG2 proteome profiles, using two-dimensional (2DE) gel electrophoresis-based separation of proteins coupled with MS. Such an investigation revealed 61 proteins differentially expressed between HepG2.2.15 and HepG2 cells, which were clustered in several functional groups, according to their biological function [22]. The main group contained significantly upregulated proteolysis factors, which were either components (PSMB6) or activators (PSME1, PSME2) of the proteasome complex, or related to ubiquitination (SKP1A, UBE2N, UBC). Interestingly, elevated levels of both proteasome subunits and lysosomal proteases were also found in hepatic tumors isolated from an HBx transgenic mouse [23]. In addition, a direct, HBx-dependent correlation between the inhibition of cellular proteasome activities and hepadnavirus replication was shown in an independent study [24], strongly supporting the hypothesis advanced by the authors that the increased synthesis of these proteins in HBV-replicating cells might be a consequence of the HBx protein expression. The implication of the cellular proteolytic activities into the HBV-life-cycle was further confirmed in a functional study, showing that the ER degradation-enhancing mannosidase-like proteins

(EDEMs) are significantly upregulated in HepG2.2.15, or other cells expressing high amounts of HBV envelope proteins. In turn, EDEM1 specifically interacts with (at least) the S protein and delivers folding-competent oligomers to degradation reducing both the amounts of secreted SVPs and virions [25].

The second group of modified proteins was calcium ion-binding members of the Annexin and S100 families with a role in cell protection against an excess of intracellular calcium concentration [26]. Previous studies have demonstrated that cytosolic calcium is a crucial regulator of HBV replication and nucleocapsid assembly functions which are also mediated by HBx [10, 27]. Unlike the proteolytic factors, these calcium-binding proteins were down-regulated in HepG2.2.15 cells; however, the significance of this attenuated expression and its relationship with the HBV life-cycle remained to be established in future functional studies.

The third group of differentially expressed proteins was involved in retinol (vitamin A) metabolism. Specifically, the levels of retinal dehydrogenase 1 (ALDH1), plasma retinol-binding protein precursor (RBP), and cellular retinol-binding protein 1 (CRBP1) were significantly increased in HepG2.2.15 cells. Following binding by CRBP, retinol is either stored within cells, as esters, or transformed into retinoic acid, the last reaction of this metabolic pathway being catalyzed by ALDH [28]. The retinoic acid binds to and stimulates several nuclear receptors, which further act as transcription factors regulating the expression of many genes, including those of HBV, through activation of its enhancer elements [29, 30]. The plasma transport of retinol is ensured by RBP [31], which also facilitates its cellular uptake upon binding to the STRA6 receptor [32]. Interestingly, high expression of ALDH1 was also found in liver biopsies from HCV-infected patients, by an independent proteomic study [33]. Thus, the retinol metabolism may be an attractive target for the development of new antiviral therapies against not only HBV [34], but other viruses as well.

Another highly upregulated protein in HepG2.2.15 cells reported by this study was complexin-2. This protein is involved in calcium-dependent exocytosis, primarily in the nervous system [35]; however, recent studies associated complexin-2 with regulated exocytosis in other cell types [36], which raises the question of a potential role in liver cells, possibly related to HBV morphogenesis. As for many other targets identified in comparative proteomics, this hypothesis too awaits confirmation by future functional studies.

Another comprehensive proteomic study investigated the network of protein–protein interactions in HepG2 and HepG2.2.15 cells aiming to identify the protein complexes modulating HBV-life-cycle [37]. In this work, the authors used 2D blue native (BN)/SDS-PAGE, which employs a nondenaturing compound (Coomassie Blue G-250) to negatively charge the protein complexes and enable them to migrate on SDS-PAGE while preserving their association [38, 39]. This analysis lead to the identification of two multi-protein complexes in HepG2.2.15 cells, that were absent in the parental HepG2 cell line. The MS analysis revealed that the most abundant proteins present in these complexes were the members of the heat-shock (HS) family of molecular chaperones, HSP60, HSP70, and HSP90, known to assist different steps of protein synthesis, folding, and oligomerisation [40]. Interestingly, while the HSP90 and HSP70 association in response to other viral infections had been documented

previously [41, 42], the presence of HSP60 in this network was novelty. The complex was further demonstrated by co-immunoprecipitation experiments, which showed that HSP60, HSP70, and HSP90 interacted with each other exclusively in HepG2.2.15 cells, indicating a potential function in HBV replication and/or assembly. Such a role has been assigned to either HSP60 or HSP90 in independent studies showing that HSP60 is required for proper folding of the HBV polymerase [43], while HSP90 facilitates the polymerase interaction with the pregenomic RNA during nucleocapsid assembly [44].

The proteomic results were validated in functional studies by silencing HSP70 and HSP90 in HepG2.2.15 cells. Moreover, HSP90 was specifically inhibited by 17-Allylamino-17-demethoxygeldanamycin (17-AAG), a small compound initially developed as cancer cells inhibitor [45]. These additional investigations showed that secretion of HBV virions and SVPs from HepG2.2.15 cells was inhibited by HSP90 or HSP70 down-regulation, while no significant cellular toxicity was observed. Similar effects on viral genome and protein secretion were obtained when 17-AAG was used to treat the HepG2.2.15 cells. Together, these results not only validate the proteomics data, but also add valuable insights into the mechanisms of HBV-host cell interactions and provide alternative targets for the development of new viral inhibitors.

Lipid rafts play an important role in the life cycle of different viruses, including HBV [46, 47]. These structures, defined as cholesterol- and sphingolipid-rich, detergent-resistant membrane microdomains (DRMs) function as a platform for anchorage of glycosylphosphatidylinositol (GPI)-tagged proteins, cytoskeleton elements, and signaling molecules as well as vesicles involved in cargo sorting and trafficking [48–50]. Because of their insolubility in detergents, DRMs components are not usually extracted during cell lysis and fail to be identified in whole cell proteomics analysis, requiring special protocols for purification and extraction. Such a procedure was very recently applied to characterize the host cell lipid rafts modified by HBV infection, using HepG2.2.15 as a cellular model [51]. In this study the authors used 2DE-MS/MS in combination with stable isotope labeling with amino acids in cell culture (SILAC)-based quantitative proteomics, to reveal differentially expressed lipid-raft proteins in HepG2.2.15 and the parental HepG2 cells. Together, 97 proteins with changed profiles were identified by these approaches that were grouped according to their biological function in signal transduction and metabolic molecules, transport and vesicle trafficking regulators, and ion channels. Interestingly, of the 44 proteins identified by the 2DE approach, 34 of which were also confirmed by SILAC. Several proteins identified in this proteomic analysis have been previously associated with the life cycle of different viruses. Of these, the decay accelerating factor for complement (DAF), a GPI-anchored protein, has been shown to directly interact with HBsAg using an alternative approach; these results are in support of a role played by DAF in HBV infection, which, however, needs further functional confirmation.

Interestingly, CHMP4B, a traffic regulator which form a complex with ESCRT-III was upregulated in HepG2.2.15, which may provide new insights into the complex mechanism of HBV secretion. Unlike nucleocapsids and SVPs, fully enveloped virions appear to use a ESCRT/Vps4B-dependent secretion pathway, therefore elevated levels of CHMP4B may imply that HBV uses lipid rafts as exit sites [14].

This study also identified a significant number of proteins functioning as receptors in signaling pathways such as PI3K/Akt or the Ras/Raf/ERK that are known to be modulated by HBx, sustaining a role for this protein in HCC development. In addition, proteins involved in immune response, cell adhesion, or electron transport, normally localized in different intracellular compartments, were found in DRMs. This unusual relocation may reflect novel functions of these proteins in HBV-infected cells that could be beneficial for the virus life-cycle, or components of a defense mechanism elicited by the cell. It would be very interesting to identify these new roles and learn to exploit them in targeted therapies against infected cells.

The consequences of HBV replication on the host cell protein expression profile were also investigated in HepAD38 cells, which produce HBV transcripts under the control of a tetracycline-responsive promoter [19]. Viral replication in these cells occurs in a controllable and inducible manner allowing a direct comparison of cellular modifications in the absence or presence of HBV-encoded factors. Moreover, this system avoids a potential influence of a different cell genetic background on the experimental outcome that cannot be totally ruled out when HepG2 and HepG2.2.15 cells are compared.

In this study, the authors performed a combination of 2DE and matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF)-MS to identify differentially expressed proteins when HBV replication is induced in HepAD38 cells [52]. Of the 23 cellular proteins that were modulated by HBV, 70 % were significantly upregulated. Further functional studies were carried out for GRP78 (also known as BiP or HSPA5), which was more than fivefold upregulated in HBV-replicating HepAD38 cells. Interestingly, an influence of HBV infection on the GRP78 levels had been demonstrated before and it was suggested to depend on the viral genotype/subgenotype [53]. GRP78/BiP is a molecular chaperone with crucial role in the maintenance of the endoplasmic reticulum (ER) homeostasis and key stress sensors in the ER. The ER stress signaled by the GRP78/BiP pathway is a major prosurvival branch of the unfolded protein response (UPR) [54].

Unexpectedly, GRP78 knockdown in HepAD38 resulted in increased efficiency of HBV replication and viral antigen secretion, while overexpression of the molecular chaperone had opposite effects, suggesting that GRP78 may act as an antiviral cellular protein. Moreover, kinetics experiments demonstrated that GRP78 mRNA begins to accumulate at the onset of HBV replication. Investigations of the interferon (IFN)—inducible signaling pathway showed that GRP78 inhibition of HBV replication occurs through stimulation of IFN-β1 expression, which, in turn, activates GRP78. The correlation between GRP78 expression and HBV replication was further confirmed in vivo, by immunohistochemical staining of HCC tumor biopsies resected from HBV-infected patients, pre- or post-treatment with the replication inhibitor, lamivudine. The authors observed that GRP78 expression was significantly increased in pre-treatment biopsies, while an opposite effect was revealed in post-treatment samples. Further studies are required to fully understand the mechanism of HBV inhibition by GRP78, and how this unusual effect can be exploited from a therapeutic point of view.

It is worth to note that despite its abundance in HepAD38 cells, previous proteomic studies failed to find altered levels of GRP78 in HepG2.2.15 cells, while

chaperones beneficial for HBV life-cycle, such as HSP60, HSP70, or HSP90 were readily identified. This suggests that GRP78 upregulation may occur at the beginning of HBV infection as an early antiviral mechanism developed by the host cell. If this line of cellular defense is overcome and the virus continues to replicate it will eventually modulate the cell response to attenuate the antiviral effects.

22.3 Proteomic Analysis of HBV-Associated HCC

Despite chronic HBV infection being associated with HCC development for more than 30 years and the compelling evidence of a role of HBx protein in this pathogenesis gathered in the last decade, the exact mechanisms of tumorigenesis and HCC progression are still to be defined. HCC is a fast progressing disease, with a very high mortality rate mostly due to late-stage diagnosis which makes successful treatment impossible [55]. It is estimated that less than 5 % of diagnosed patients survive for up to 5-years, most cases being registered in developing countries. Early diagnostics is of crucial importance in the management of the disease, therefore, HCC biomarkers are valuable tools not only for the screening of the high-risk population, but also, to understand the mechanisms triggering carcinogenesis and develop treatment strategies against it. This analysis was performed both, in vitro, using the available cellular models supporting HBV infection and replication, or in vivo, using clinical samples from HCC-diagnosed patients.

The first approach ensures an unlimited source of samples, with homogenous origins, thus, high reproducibility and low costs. Moreover, secreted proteins with a role in infection and pathogenesis can be readily identified in culture medium and further validated in vivo, as biomarkers with potential clinical use.

Given the properties of the HBx protein and its function in tumorigenesis, most in vitro studies have focused on the consequences of this protein expression on the global proteomic profile of the host cell.

Based on the observation that clinical manifestations and progression of HBV-infected HCC patients depend on the virus genotype (B or C), it of interest to determine whether different HBx genotypes can account for the outcome of the disease. To this aim, genotype-specific HBxs were expressed in non-HCC Chang cells to identify cellular proteins associated with specific HBx [56]. The results have shown that several calcium-binding proteins and degradation proteins are differently expressed in transfected cells in HBx genotype-specific manner. Significantly modified profiles were found for cytoskeletal proteins involved in cell migration. These differences were later associated with the binding properties of the proline-rich domain of the HBx protein, which is the only variable region among the well-conserved HBV genotypes. This domain consisting of a PXXP motif is involved in many interactions responsible for the HBx functions, including those with cytoskeletal proteins, which promote cell motility and invasion [57]. Interestingly, this motif had been implicated previously in intracellular signal transduction mediated by other viral proteins [58].

In a study published recently, HepG2 cells were transfected with HBx cDNA and a stable cell line (HepG2-HBx) was selected [59]. Using 2DE and MS analysis, the authors further investigated how the HBx protein expression modified the proteome of the HepG2 cells. The analysis revealed that expression of 60 proteins varied more than twofold between HepG2-HBx and the parental cell line and the identity of 54 of them was established by MS/MS analysis. Previous studies had demonstrated that HBV infection results in epigenetic modifications of important tumor suppressor genes (TSGs), a process mediated by the HBx protein and strongly linked to HCC development [60, 61]. The viral protein was shown to activate the DNA methyltransferase 1 (DNMT1) by a mechanism involving the positive regulatory p16 (INK4a)-cyclin D1-cyclin-dependent kinase (CDK) 4/6-retinoblastoma protein (pRb)-E2F1 pathway, leading to DNA hypermethylation and repression of TSGs [62]. In this context, it was interesting to determine whether the expression of the most altered proteins was also regulated by epigenetic modifications. The analysis showed that two down-regulated proteins involved in calcium homeostasis, S100A6, and S100A4, were hypermethylated in HBx-expressing cells; moreover, cell treatment with a DNA methylase inhibitor 5-aza-2′-deoxycytidine (AZA) was able to restore their expression. Surprisingly, the promoter of aldehyde dehydrogenase 1 (ALDH1) was found hypomethylated in HepG2-HBx cells, suggesting that in addition to DNA hypermethylation, HBx can also induce specific hypomethylation of host cell genes. Moreover, these results validated and provided a mechanism for the altered proteome profile observed previously in HBx expressing HepG2.2.15 cells, showing that cytosolic calcium-binding proteins of the S100 family were significantly down-regulated, while ALDH1 expression was increased compared to the parental HepG2 cell line [22]. The physiological relevance of these findings was discussed in the chapter above.

The role of HBV in epigenetic regulation of host cell genes was addressed in a comprehensive screening study, by coupling iTRAQ and liquid chromatography (LC)-MS/MS to identify proteins down-regulated in HepG2.2.15 compared to HepG2 cells, which regained expression following treatment with the DNMT inhibitor, 5-aza-2′-deoxycytidine [63]. The differentially expressed proteins reactivated by this inhibitor were further grouped according to their biological function including signaling regulation (five proteins), metabolic enzymes (six proteins), and structural proteins (four proteins). Interestingly, synthesis of the calcium-binding proteins S100A6 and S100A4 was significantly lower in HepG2.2.15 than HepG2 cells, in agreement with the proteomics studies published previously. In addition, expression of Annexin A2, another calcium-binding protein was decreased in HepG2.2.15 cells and significantly restored following Aza treatment. The response of S100A6 and Annexin A2 genes to the methylation inhibitor was further confirmed by real-time RT-PCR and Western blotting. Moreover, analysis of promoter methylation by bisulfite genome sequencing (BGS) showed increased values in HepG2.2.15 compared to HepG2 cells for both, S100A6 (43.18 and 7.89 %) and Annexin2 (97.22 and 2.78 %), strongly supporting the hypothesis that down-regulation of these genes is a consequence of the epigenetic regulation induced by HBV. It is important to note that while proteins of the S100 family were consistently found down-regulated

in HBV-replicating cells, inhibition of Annexin2 expression is more controversial, as previous studies had reported upregulated levels of this protein in HCC [64]. It may well be that the negative regulation of Annexin2 expression through epigenetic modifications depends strictly on the level of HBx and HBV replication, which occurs at high rates in HepG2.2.15. In the absence of Annexin2 and other calcium-binding proteins down-regulated during infection, higher amounts of cytoplasmic calcium are available to sustain HBV replication in these cells.

The proteomics studies performed in vitro using HCC-derived cell lines have certainly provided insights on the global changes induced by HBV on host factors, which could be further exploited as potential markers in HCC diagnosis. Development of early diagnostic biomarkers is very important in HCC management as specific symptoms are absent at the onset of the disease. However, to gain clinical relevance, these results must be validated in a large number of samples from HCC patients.

The first in vivo study employed 2DE coupled with MALDI-TOF-MS to analyze non-tumor, cirrhotic, and tumor tissues isolated from livers of HCC patients. Of the 21 proteins differentially expressed in these samples, the authors proposed that at least lamin B1 could be considered as a novel marker for cirrhosis [65].

A second proteomics study addressed the tissue heterogeneity issue when using biological samples from patients for proteomic studies, by selecting pairs of normal and tumor tissues from the same liver [66]. Moreover, to identify potential markers in HCC progression, samples from well and poorly differentiated tumors were also selected for analysis. By employing two-dimensional difference gel electrophoresis (2D-DIGE) coupled with MS, the authors found metabolic and oxidative-stress-related proteins were down-regulated in the poorly differentiated samples. Of these proteins, methionine adenosyltransferase (MAT) is particularly significant, being involved in the methylation cycle of the liver and the formation of *S*-adenosylmethionine (AdoMet). AdoMet is an important metabolite supporting mitochondrial function and the management of oxidative stress in the liver, its deficiency being linked to steatohepatitis and HCC development [67], therefore, MAT might be a good candidate marker to monitor HCC progression.

A major issue encountered when analyzing tissue samples is the contamination with proteins from other sources, such as blood. To overcome this problem, a number of proteomic studies employed laser capture microdissection (LCM), which allows selection of a cell population from a heterogeneous tissue under the microscope. LCM coupled with (^{13}C) isotope-coded affinity tag (cICAT) protein labeling 2D-LC-MS/MS was used to investigate the protein profiles in HCC and non-HCC tissues from a patient with HBV infection [68]. In addition to the large number of total and differentially expressed proteins in HCC and non-HCC tissues revealed by the qualitative and quantitative analysis, this approach allowed identification of lower-abundant proteins and enrichment of basic and transmembrane proteins, in a rapid and accurate manner. Although many of the proteins identified in this study are involved in signal transduction pathways linked to carcinogenesis, future studies in a larger number of clinical samples are required to validate them as potential HCC diagnostic markers.

To increase LCM sensitivity, the following study coupled this technology with electrospray ionization mass spectrometry (ESI-MS), which allows the identification of very small amounts of protein isolated from a sample [69]. Of the proteins isolated from HCC patients, showing a significantly altered expression in tumor and healthy tissues, peroxiredoxin 2 (Prx II), Apolipoprotein A-I precursor, (ApoA-I), and 3-hydroxyacyl-CoA dehydrogenase type II appeared relevant for HBV-related HCC. Interestingly, unlike in other types of tumor cells such as breast, thyroid or lung, which express high amounts of Prx II, this antioxidant enzyme was found down-regulated in HCC tissues by 2DE and ESI-MS/MS. This behavior was further confirmed by immunohistochemical and Western blotting analysis. This shows the complexity of the pathways involved in the development and progression of tumors of different origins, and suggests that while being a good candidate as tumor progression marker in other types of cancer, Prx II may function as a tumor suppressor in HCC.

It is important to note that despite the important amount of data that resulted from the proteomics studies of HCC tissues and the variety of approaches used including technical improvements to increase their efficiency and accuracy, most of the proteins were not validated by complementary analysis or in larger populations. Moreover, there was little overlap between the differentially expressed proteins proposed by these studies, strongly suggesting that a more comprehensive analysis using a wider range of clinical samples is needed. Such an investigation was more recently proposed by Sun and colleagues [70] who employed the 2D-DIGE technique to analyze the protein expression profiles in tumor and healthy liver tissues from 12 HCC patients. In addition to proteins previously linked to HCC, such as members of the HSP70 and HSP90 families, the proliferating cell nuclear antigen, the hepatoma-derived growth factor, which were overexpressed and metabolic enzymes involved in the methylation cycle in the liver, which were down-regulated, other proteins were reported for the first time. Two of the proteins from this category were the HSP70/HSP90-organizing protein (HOP), a chaperone molecule involved in proliferation and myogenesis of cardiac tissue [71] and the heterogeneous nuclear ribonucleoproteins C1/C2 (hnRNP C1/C2), a telomerase-binding factor, were strongly upregulated in cancer samples. To validate these proteins as potential HCC markers, their expression was further confirmed by Western blotting and immunohistochemistry staining in 12 and 70 pairs respectively, of tumor and non-tumor HCC tissues.

Very relevant for biomarkers discovery are those tissue proteins that are released into the blood at the onset or during progression of the disease, since their levels can be easily monitored in serum samples. Therefore, comparative proteomic studies have focused on the analysis of serum from patients at different stages of HCC.

A global serological profile of proteins with markedly changed patterns in HCC was performed by 2D-MALDI-TOF MS using serum from 18 patients with chronic HBV infection and 5 healthy individuals, with particular focus on potential alterations occurring among protein isoforms [72]. The patients were grouped according to their necroinflammatory score (NS) in low (L) and high (H) NS, which correlated with their serum alanine aminotransferase (ALT) levels. One of the most altered proteins identified in this study was haptoglobin and its multiple isoforms, which

increased slightly in the LNS serum samples compared to control while hardly detectable in HNS patients. This suggests that in the absence of significant inflammation, HBV infection does not lead to serious liver injury, which may correspond to the immune-tolerant phase observed during the evolution of the disease. In contrast, severe inflammation of the liver may have dramatic functional consequences leading to reduced haptoglobin secretion from the liver.

A very interesting behavior was also observed for the ApoA-I isoforms: while the overall levels were reduced in both LNS and HNS serum, the individual profile of the three isoforms followed opposite trends in HBV-infected patients. Previous reports have shown that ApoA-I biosynthesis is post-translationally regulated [73]; thus it is plausible that the observed variation of the isoforms is the result of this control and may have functional consequences in HCC progression. Further characterization of these post-translational modifications may contribute to the current understanding of HBV infection and diagnosis.

A qualitative modification was confirmed in this study for transthyretin, a protein that is significantly down-regulated in the acute-phase of liver diseases [74]. Expression of one of the transthyretin isoforms was reduced in both LNS and HNS serum samples, indicating that serious liver dysfunctions may occur at early stages of HBV infection.

Analysis of both quantity and quality of serum proteins in patients with liver cancer that could be correlated with the progression of the disease was the subject of another proteomic investigation performed by 2D-MALDI-TOF-MS [75]. Here too, particular attention was paid to the clinical status of patients who were grouped in several categories: normal, chronically infected with HBV, either without symptoms or with active disease and HCC diagnostic. The samples obtained from patients in each group were pooled in order to prevent a high variability of the results between individuals. Using this strategy, two proteins abundantly expressed in the healthy group and consistently down-regulated in the HBV-infected group were reported, and further identified as a fragment of the complement C3 and an isoform of the ApoA-I, respectively. The exact relationship between the reduced amount of these proteins and liver pathogenesis due to HBV infection remained to be established, as well as their potential relevance as HCC biomarkers. Nevertheless, these early proteomic studies showed that appropriate technical strategies can overcome the problematic features of serum proteome profiling, such as sample variability and high concentration of non-relevant proteins. Other similar strategies addressing these issues included protein deglycosylation [76] or albumin and immunoglobulin G depletion [77] before proteome analysis. Using this last approach, the proteomic profiles of 60 sera grouped in normal, HBV-infected and HCC patients were investigated by 2D coupled with MALDI-TOF-MS/MS. The study reported eight proteins differentially expressed between the three groups. Expression of the heat-shock protein 27 (HSP27), one the most upregulated proteins in HCC sera, was validated by Western blotting. Interestingly, the protein was absent in sera from healthy patients while overexpressed in two HBV-positive and 90 % of HCC sera, which makes it a promising candidate as an HCC biomarker [77]. Very recent validation studies employing an immunoenzymatic assay to detect the serum level of HSP27

in a larger population, including 151 patients at different stages of chronic liver disease, confirmed the upregulation of the chaperone in HCC; however, the disease was rather related to HCV infection and a direct relationship between the Hsp27 levels and tumor prognostic risk factors could not be clearly established [78].

Based on previous reports showing aberrant glycosylation as one of the most important post-translational modification of proteins in liver disease, an interesting study was set-up to analyze the global profile of the oligosaccharides released from serum glycoproteins in healthy and HCC-diagnosed woodchucks and further identify those proteins significantly modified by 2DE-MS [79]. Woodchucks can develop HCC following infection with Woodchuck Hepatitis Virus (WHV), a hepadnavirus sharing many similarities with HBV which has been long used as a model for the human variant [80]. Very high levels of alpha-1, 6-linked fucose was found in serum from HCC-woodchucks compared to healthy subjects. As expected, a larger number of proteins were found to be fucosylated in the HCC compared with normal sera, some of them, such as alpha-fetoprotein and α-1-antitrypsin being reported previously in human HCC [81]. Of the novel fucosylated proteins, expression of the Golgi Protein 73 (GP73) was further validated in the sera of human patients with HBV infection and HCC. Although the exact role of GP73 in HCC development was not clear, it was suggested that its secretion from liver cells may indicate a serious dysfunction of Golgi-residing glycosylation pathways during HCC. It is worth mentioning that the relevance of GP73 hyperfucosylation in HCC is still a pending question, however, several following studies performed in significantly larger populations have confirmed the upregulation of this protein in serum of patients with chronic HBV infection. Importantly, the concentration of the protein was shown to vary in earlier phases of the disease making GP73 a promising biomarker candidate for liver fibrosis and cirrhosis diagnosis [82].

Serum proteomic analysis was also performed by surface enhanced laser desorption/ionization, (SELDI)-TOF-MS analysis was applied to generate models for diagnosis of different stages of chronic liver disease and HBV infection by assigning significantly modified proteins into optimized classification trees [83]. The clinical sensitivity and specificity of the models were investigated using double blind test samples and very good prediction profiles were obtained for the diagnosis of chronic HBV carriers, cirrhotic, and HCC patients.

SELDI-TOF was also combined with ProteinChip followed by 2DE analysis and sequential protein identification through peptide-mass fingerprinting (PMF) to determine the serum proteome profiling of HBV-infected patients including those at early stages of chronic liver disease [84]. The protein signals were statistically interpreted and a predictive model was built up based on sequential algorithmic analyses of the proteins most changed between the clinically differentiated patients. One of the proteins that gradually increased in concentration from normal to HBV-infected and the HCC groups was serum amyloid A (SAA), further confirmed by immunoprecipitation and Western blotting. Given that SAA has been previously associated with acute phases of inflammation in other diseases, additional studies are required to validate the relevance of this discovery; nevertheless it is hoped that SAAs could be used in combination with other markers in early HCC prediction and monitoring.

22.4 Identification of Differentially Expressed Proteins in HCV- and HBV-Originating HCC

HCC is one of the most frequent malignant tumors worldwide causing the death of about one million people annually [85]. Chronic HBV and HCV infections are responsible for 50 and 30 % of HCC cases, respectively; however, significant differences between the mechanisms of HBV and HCV-induced malignancy have been revealed in recent years [86]. In addition to encoding the oncogenic HBx protein discussed above, HBV is able to integrate its genome into susceptible genes of the host DNA, inducing genomic instability and insertional mutagenesis which play an important role in tumor pathogenesis. Targeted proteomic studies have also shown significant epigenetic changes induced by HBV infection resulting in down-regulation of TSGs at early stages of HCC development. Unlike HBV, HCV is an RNA virus and unable to integrate into the host genome. The predominant role in HCV-induced carcinogenesis is played by the viral proteins and their interactions with the host factors responsible for the cell cycle regulation and control [87].

Currently, many transcriptomic and proteomic studies address the comparison between HBV- and HCV-associated HCC to understand in more detail the etiology-specific oncogenic pathways and help the development of new therapeutic strategies according to the causative pathogen. A first comprehensive proteomic research investigated the protein patterns in HBV- and HCV-related HCC using 2DE together with MALDI-TOF-MS [88]. The study included 21 pairs of tumor and surrounding liver tissues clinically classified in three groups, HBV-positive, HCV-positive and HBV/HCV-negative. The authors identified 60 proteins varying significantly between the three groups. Interestingly, the pool of proteins highly upregulated in HCC samples, regardless of their etiology, was dominated by members of the HSP family. In addition, nucleophosmin and elongation factor 2 were also increased in all HCC subjects, while carbamoyl phosphate synthase I (CPSASE I) and argininosuccinate synthase were down-regulated. These results confirm previous reports showing more abundant expression of nucleophosmin in cancer and highly proliferating rather than the normal cells [89] and low levels of CPSASE I mRNA in HCC tissues [90].

Of the proteins disregulated in HBV-dependent HCC, the Cu/Zn superoxide dismutase (SOD Cu/Zn) was found to decrease, this tendency being confirmed by Western blotting analysis. This is an interesting observation as SOD is a metalloenzyme protecting the cell from oxygen-free radicals which have been rather associated with HCV infections [91]. In agreement with this proteomic study, decreased SOD levels had also been reported in HCC and a positive correlation with patients survival rate was established [92]. In contrast, 4-aminobutyrate aminotransferase (GABA-AT), HSP27, and enoyl-CoA hydratase were upregulated in HCV-dependent HCC. Although the molecular bases of the differences observed between the protein profiles in HBV- and HCV-related HCC are still to be defined, this proteomic investigation strongly suggests that HCC is a “different” disease depending on the etiologic factors and therefore specific approaches must be taken to increase the treatment efficiency.

Interestingly, a detailed transcriptomic investigation in HBV-related HCC has shown a significant upregulation of the Wnt/β-catenin pathway in tumor sections [93],

a process which was later associated with expression of the HCV core protein, too [94]. Wnt genes encode a family of secreted cysteine-rich proteins located near the cell or in association with extracellular matrix proteins, which constitute the evolutionarily conserved Wnt-signaling pathway. This pathway has crucial roles in organ development and has been found disregulated in an increasing number of cancers; moreover, mutations of Wnt ligands or receptor genes have been discovered in colorectal cancer, desmoid tumor, or hepatoblastoma [95, 96]. These data implied that HBV and HCV share common mechanisms leading to carcinogenesis as well, and that the Wnt-controlled transduction pathway is one of them. Indeed, using a functional proteomics approach, enhanced levels of Wnt-1 protein were found associated with nuclear factor-kappa B (NF-kB)- signaling complexes HCC tissues, regardless of the HBV or HCV etiology [97]. Activation of the NF-kB pathway had been previously shown in liver cancer tissues. However, a molecular link to hepatocarcinogenesis was missing [98].

A very recent proteomic study employed 2DE-LC-MS/MS to compare the protein profiles in three progressive stages of HCV and HBV infection, including chronic active hepatitis, cirrhosis, and HCC [99]. The study was performed using blood serum from 40 diagnosed patients and 7 healthy individuals. Of the 62 protein spots with significantly different expression between the three groups, 42 were identified by MS and assigned to 15 proteins. Interestingly, a similar proteome pattern was observed in chronic active hepatitis patients infected with either HBV or HCV, while differences became obvious with progression of the disease. Of the differentially expressed proteins, α1-acid glycoprotein (AGP), leucine-rich α-2-glycoprotein (LRG) and clusterin (CLU) were upregulated in HBV- compared to HCV-cirrhotic patients and Ig gamma chain C followed an opposite trend. AGP is a glycoprotein synthesized and secreted by hepatocytes and its serum level is associated with liver pathogenesis. Increased concentrations of AGP were shown in acute hepatitis and HCC, while the protein is reduced in chronic hepatitis and liver cirrhosis [100]. Therefore, low expression of AGP in HCV-cirrhotic patients may reflect more advanced liver tissue damage in HCV compared to HBV cirrhotic patients. LRG was one of the most upregulated proteins in HBV- but not in HCV–HCC patients, but the reason for this difference is not known. LRG is mostly synthesized by liver cells and highly increased in mice livers following lipopolysaccharide treatment. Although its physiological functions are still to be established it is hoped that this protein may find clinical applicability in discriminating between HCC of different etiologies.

22.5 Proteomics of Cells Permissive for HBV Infection

The proteomic studies addressing the HBV-host cell relationship described above have been performed in cells replicating the virus either stably or transiently, or expressing different virus-encoded proteins. These data have certainly contributed to our understanding of the HBV-induced pathogenesis and the cellular adaptive response. However, none of these cell lines is able to support HBV infection, which

hampered considerably the investigation of the early steps of the viral life cycle. Owing to its highly specific tissue tropism, HBV infection in vitro was restricted for years to human and Tupaia primary hepatocytes and it is notoriously inefficient. Moreover, the heterogeneity of these cells and the difficulty to maintain them viable in culture for a reasonable amount of time have made their use in vitro very complicated and unreliable. The development of the HepaRG cell line, isolated from the liver tumor of an HCV-infected patients and the discovery of its unique properties have provided a ground-breaking tool to investigate HBV infection in a more robust and reproducible system [101]. These proliferating cells can be infected with HBV when differentiated in the presence of DMSO, insulin, and hydrocortisone, while maintained quiescent at confluence. Interestingly, HepaRG differentiation results in the development of two cell populations with either hepatocyte-like or biliary-like epithelial phenotypes, suggesting that these cells have bipotent progenitor features. When optimally differentiated, the HepaRG cells express high levels of adult hepatocyte-specific proteins, such as albumin and aldolase B, membrane transporters and nuclear transcription factors and chaperones involved in oxidative-stress relief. In addition, phase I and II biotransformation enzymes involved in drug metabolism are highly expressed during differentiation, which has promoted these cells as an excellent in vitro model to study xenobiotic metabolism and genotoxic compounds [102].

In this context, a thorough characterization of the proteomic profile of the HepaRG cells in the absence and presence of HBV was very interesting, not only to understand the unique properties of these cells, but importantly, to reveal new host cell factors with potential role in the early steps of infection, which was not possible in previous cellular models. Using 2DE coupled with LC-MS/MS, the first proteomic analysis on HepaRG revealed 44 proteins differentially expressed between naive and HBV-infected cells [103]. The proteins were grouped according to their biological function, such as calcium and apoptosis regulators, DNA/RNA processing enzymes, or scaffold proteins. Many of these factors had been reported previously in HBV-replicating cells (copper chaperone for superoxide dismutase, glutathione *S*-transferase, annexin A1, heterogeneous nuclear ribonucleoproteins), or in other viral cellular systems (heterogeneous nuclear ribonucleoproteins A/B, Ras GTPase activating protein-binding protein 1, PDZ domain containing-protein).

It is important to note that this study compared the HepaRG proteome profile in the absence or presence of HBV infection and in both cases the cells were first differentiated. However, this experimental setting makes difficult to evidence the host cell factors rendering the HepaRG cells permissive to HBV infection, since most likely, upregulation of HBV receptors and other proteins involved in virus entry takes place during the complex differentiation process. To gain more insights into the mechanisms underlying HepaRG differentiation and susceptibility to HBV infection, our group performed two independent proteomic studies on non-differentiated (ND) and differentiated (D) cells. As plasma membrane is the first barrier encountered by a pathogen during cell infection, we focused our analysis on this compartment, aiming to identify differentially regulated proteins with potential role in HBV entry, or functional signaling networks that are activated at early stages of infection.

Cellular fractionation and purification of the plasma membrane fraction also provided the means to enrich low-abundant proteins or factors trafficked between different compartments and the cell surface. Using LC-MS/MS, the first study revealed 111 and 98 proteins in (D) cells exclusively, in duplicated experiments [104]. Of these proteins, three members of the Annexin family A1, A2, and A5 were identified with very high Mascot scores. This is an interesting result which may shed some light on the controversial function of these calcium-binding proteins, which were found either upregulated or inhibited in independent proteomics studies, as discussed above. This behavior was correlated with the rate of HBV replication, which varies greatly between different cell lines and tumor tissues and strongly depends on cytoplasmic calcium. The identification of annexins at the plasma membrane of (D) HepaRG cells suggest additional functions played by these proteins depending on their intracellular location. Notably, annexin A2 was shown to bind to negatively charged phospholipid of cholesterol-rich membranes [105], which are highly upregulated in (D) cells [47]. This suggests a potential role of this protein in the organization of lipid raft microdomains membrane, which might be relevant for HBV infection.

The second proteomic analysis used the same strategy, intending to improve the list of differentially expressed proteins at the plasma membrane of (ND) and (D) cells [106]. In this study, components of the plasma membrane fraction were significantly enriched compared to the total cell lysates, as demonstrated following identification of relevant markers by Western blotting. This was a crucial advantage, which permitted detection of novel proteins, unreported before. Two independent analyses revealed the identity of 246 and 177 proteins respectively that were only expressed in (D) cells, while 51 and 45 proteins respectively were exclusively found in (ND) samples. Notably, more trafficking and signal transduction regulators, as well as proteins previously associated with viral infections were reproducibly found (D) compared to (ND) cells. Among the differentially expressed proteins, strong Mascot scores were obtained for cathepsin D, carboxylesterase, cyclophilins A and B, members of the HSP 90, and protein disulfide-isomerases (PDI) families of proteins, annexins (A1, A2, and A5). For example, the MS and MS/MS spectra leading to identification of the VDATEESDLAQQYGVR and SNFAEALAAHK peptides, which belong to PDI, are shown in Fig. 22.1. Similarly, the RFDVSGYPTLK and VDATAETDLAK peptides, which are part of PDI-A4, are shown in Fig. 22.2.

Upregulation of some of these proteins was further confirmed by Western blotting or by comparison of the relative intensity of precursor ions. An example of this last analysis is shown for the PDI and PDI-A4-derived peptides (Fig. 22.3). To confirm the quantitative results, two peptides per protein were involved in quantification, which was performed as in Darie et al., 2011 [107].

As observed, the relative quantitation showed a clear increase of the amounts of the peptides derived from either PDI or PDI-A4, strongly suggesting that the PDI proteins are localized at the plasma membrane of (D) cells.

In addition, the proteins with significantly altered levels of expression in the two HepaRG cell populations resulting from the differentiation treatment are grouped according to their biological role, cellular location, and molecular function (Fig. 22.4).

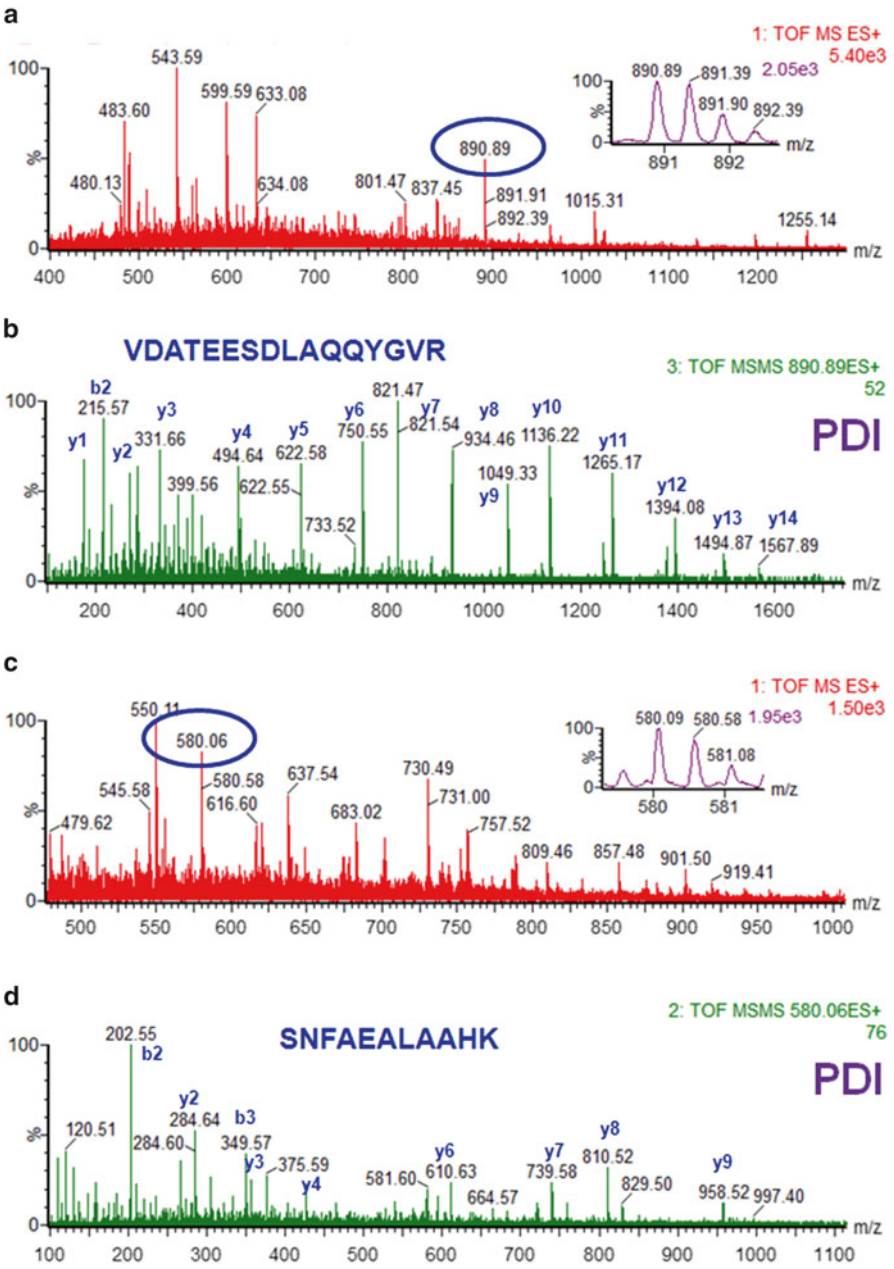


Fig. 22.1 NanoLC-MS/MS of plasma membrane samples from (ND) and (D) HepaRG cells and PDI identification. Panels (**a**) and (**c**) show the MS spectra and the precursor ions with *m/z* of 890.89 (2+) (**a**) and 580.06 (2+) (**c**) selected for fragmentation. Fragmentation of the precursor ion with *m/z* of 890.89 (2+) produced the MS/MS spectrum shown in panel (**b**) and fragmentation of the precursor ion with *m/z* of 580.06 (2+) produced the MS/MS spectrum shown in (**d**). Data analysis of the MS and MS/MS spectra resulted in identification of peptides with the amino acid sequence VDATEESDLAQQYGVR (**a**, **b**) and SNFAEALAAHK (**c**, **d**) which belong to PDI

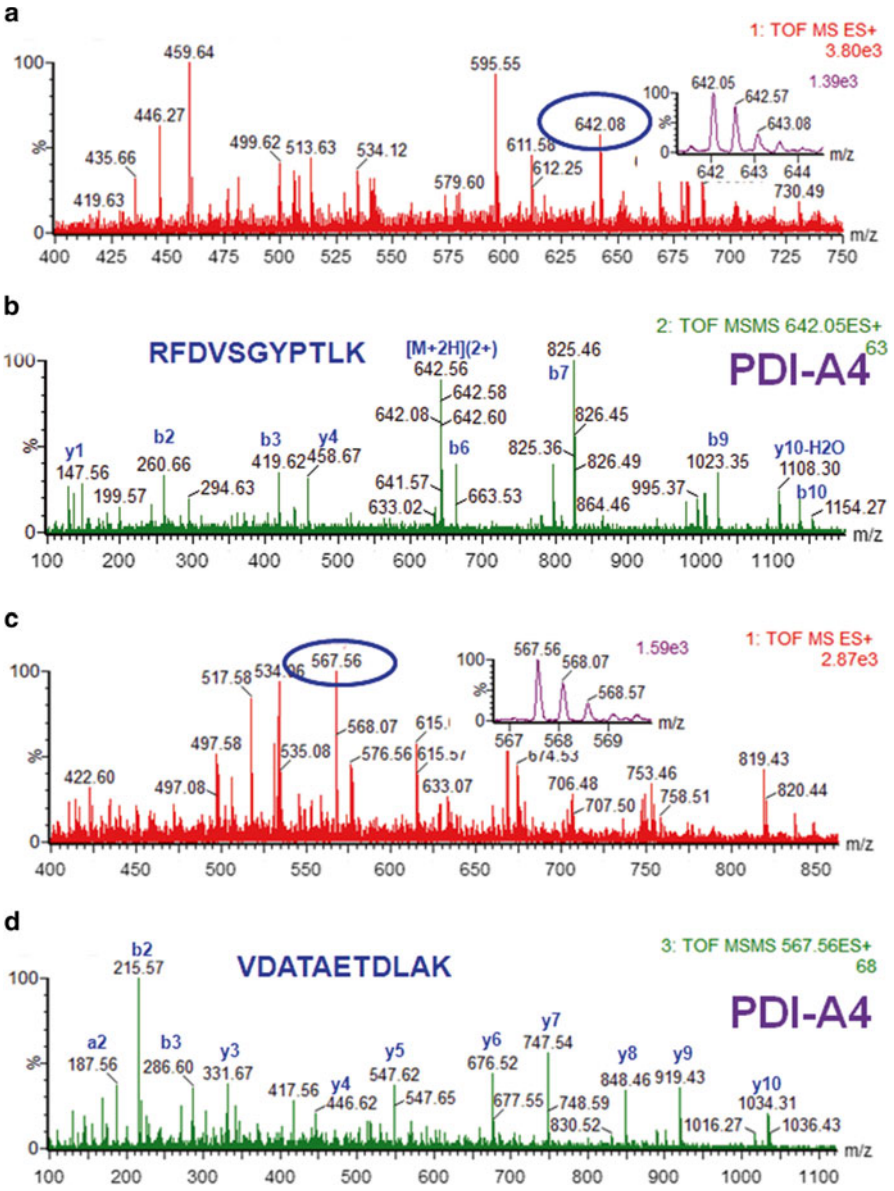


Fig. 22.2 NanoLC-MS/MS of plasma membrane samples from (ND) and (D) HepaRG cells and PDI-A4 identification. Panels (**a**) and (**c**) show the MS spectra and the precursor ions with *m/z* of 642.08 (2+) (**a**) and 567.56 (2+) (**c**) selected for fragmentation. Fragmentation of the precursor ion with *m/z* of 642.08 (2+) produced the MS/MS spectrum shown in (**b**) and fragmentation of the precursor ion with *m/z* of 567.56 (2+) produced the MS/MS spectrum shown in (**d**). Data analysis of the MS and MS/MS spectra resulted in identification of peptides with the amino acid sequence RFDVSGYPTLK (**a**, **b**) and VDATAETDLAK (**c**, **d**) which belong to PDI-A4

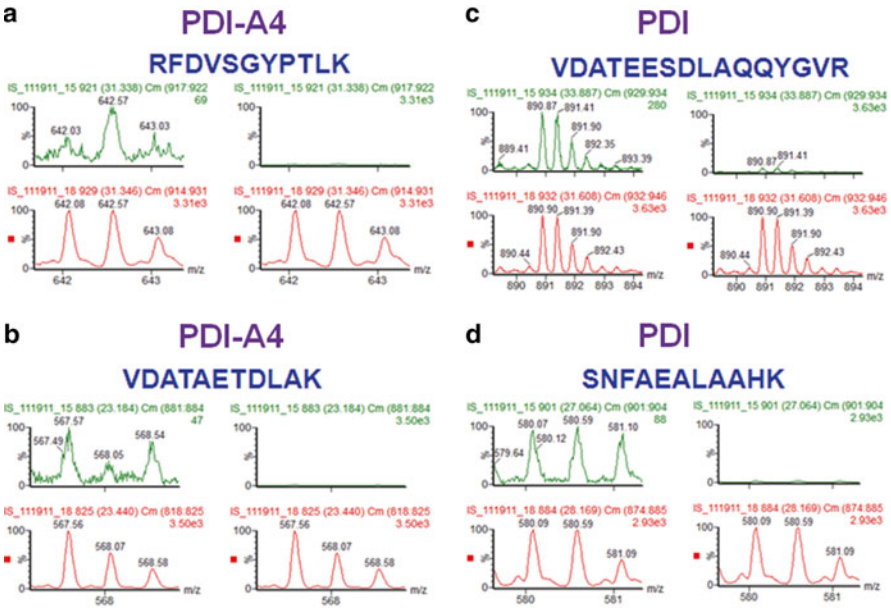


Fig. 22.3 Relative quantitation of the precursor ions with *m/z* of 642.08 (2+) (**a**), 567.56 (2+) (**b**), 890.89 (2+) (**c**), and 580.06 (2+) (**d**). These ions led to identification of the peptides with the indicated amino acid sequence, which derive from PDI-A4 or PDI. For the relative quantitation, the spectra from the extracted ion chromatogram was summed for one minute and either directly compared for the number of counts (indicated on the spectrum on the left side of each precursor (**a**–**d**) or compared using the vertical axes linked (the precursor peak with the highest number of counts was considered to have a 100 % intensity)

Interestingly, further investigation of these proteins by confocal microscopy revealed that annexin 1, cyclophilin A, and cathepsin D overexpression was restricted to hepatocyte-like cells, while PDIs appeared increased in both, biliary- and hepatocyte-like cells, as exemplified in Fig. 22.5.

Identification of proteins such as cathepsin D, or PDIs, at the plasma membrane of (D) cells appears surprising, as traditionally, these proteins have been localized at different intracellular sites, such as late endosomes or lysosomes and the ER, respectively, where they fulfill specific functions. Future functional studies will investigate the significance of these changes and the potential role that these enzymes may play in the early steps of HBV entry, such as nucleocapsid uncoating and genome release. In a recently published report we have shown that HBV infection in HepaRG cells is favored by the reducing milieu of the endocytic pathway, possibly by destabilization of the disulfide-linked envelope proteins. Although this process remains to be characterized at molecular level, it is highly plausible to involve the upregulated PDIs.

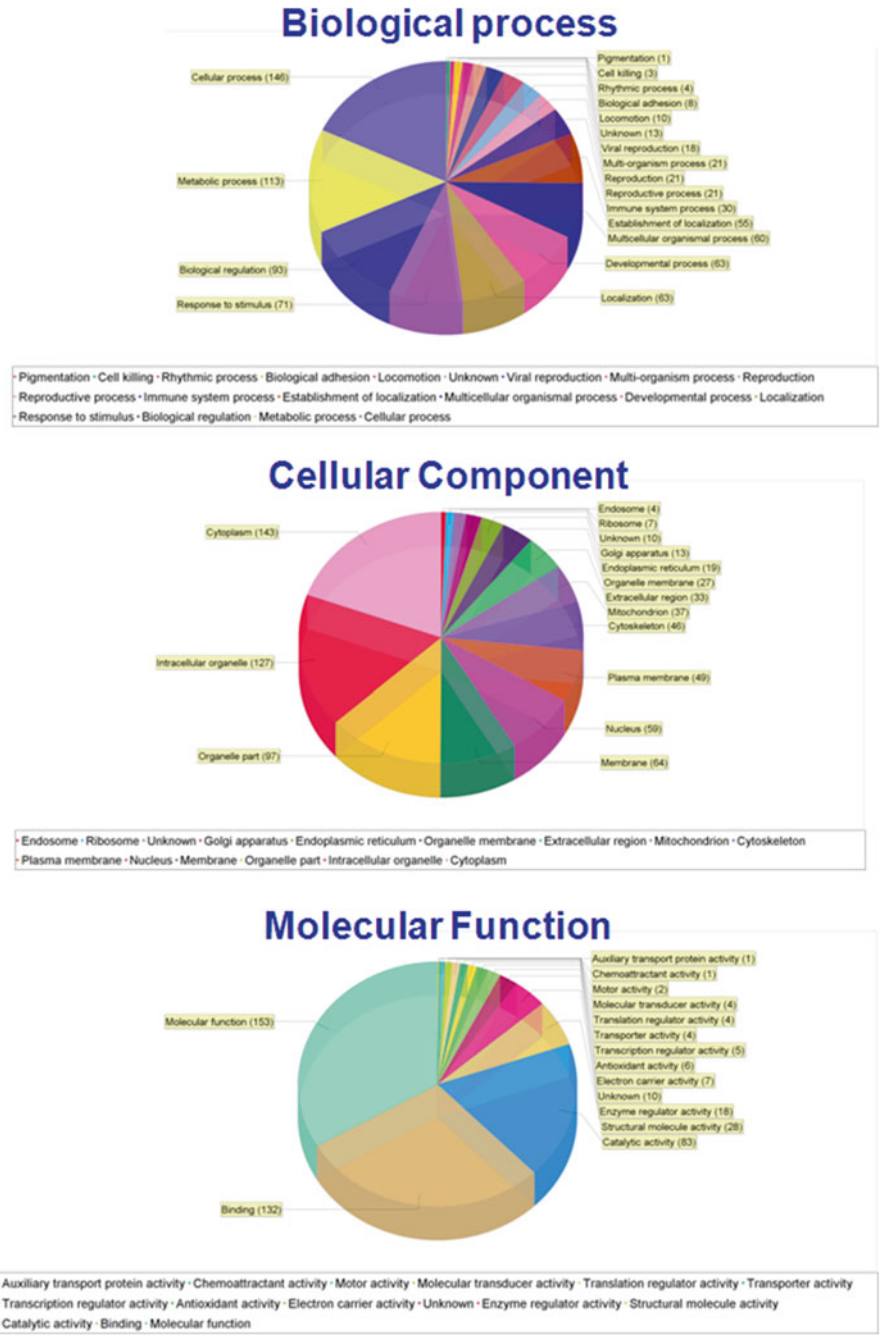


Fig. 22.4 Classification of plasma membrane proteins isolated from (ND) and (D) HepaRG cells. The identified proteins were grouped according to their biological role, cellular localization, and molecular function as in reference [106]

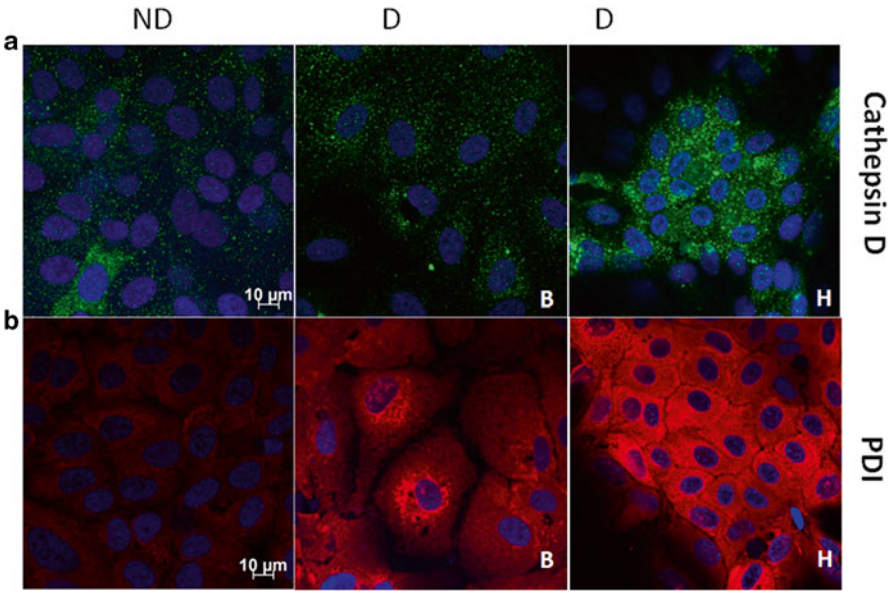


Fig. 22.5 Analysis of differentially expressed proteins in (ND) and (D) HepaRG cells. Expression of Cathepsin D (panel **a**) and PDI (panel **b**) in (ND), hepatic (H) and biliary (B) cells was evidenced using specific antibodies. DAPI was used to visualize the nuclei (*blue*). Samples were analyzed with a Zeiss LSM 710 confocal microscope. Images were taken with a 63× objective and processed with the ZEN software. Scale bar is 10 µm

22.6 Conclusion Remarks

A large number of proteins differentially expressed in HBV pathogenesis have been discovered using proteomics (summarized in Fig. 22.6) and now doubt, many of these findings have contributed to our current knowledge of HBV infection. The continuous improvement of this technology has provided a powerful tool to identify very low-abundant proteins in samples from patients at various stages of livers disease, with relevance for biomarkers development. However, it is important to note that little overlap was found between the proteins reported in different studies, which is most probably related to the variety of technologies employed and the sample origin. In addition, functional characterization of key proteins is scarce in most proteomics studies. Despite the impressive amount of data available, there is still a long road for potential new biomarkers to meet the requirements for clinical approval and unfortunately, none of the candidate proteins has been able to reach this point, so far. Clearly, further validation using a higher number of patients and more standardized methodologies are required. An integrative approach combining this technology with other global analysis such as transcriptomics and metabolomics may also increase the significance and impact of proteomics data and ultimately help understand the HBV-induced pathology in a more comprehensive and systematic manner.

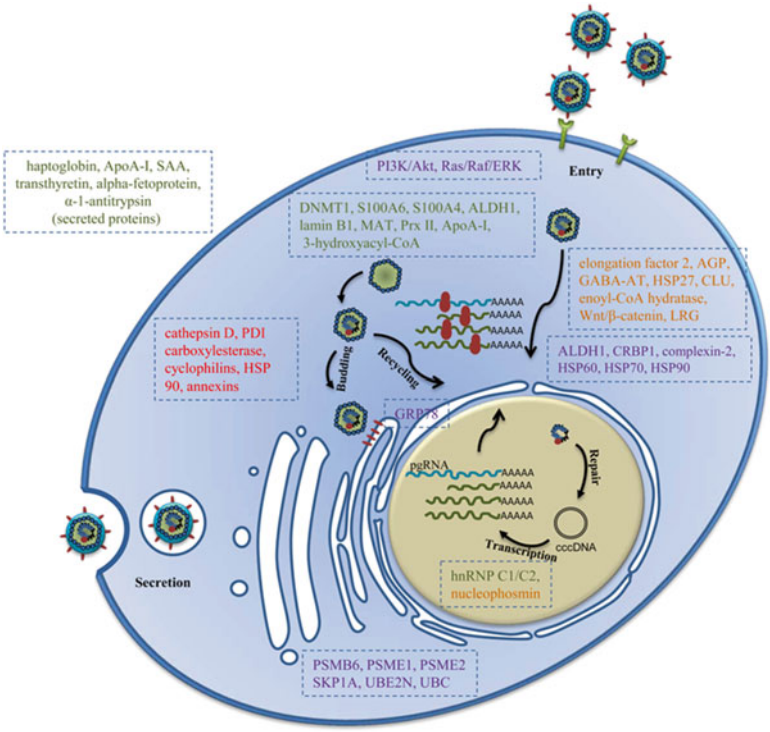


Fig. 22.6 Schematic representation of HBV-induced host cell protein modifications. The HBV life-cycle is depicted and differentially expressed proteins in HBV-replicating cells (*purple*); HBV-associated HCC (*green*); HCV- and HBV-originating HCC (*orange*); cells permissive for HBV infection (*red*)

Acknowledgements The work was supported by the Romanian Academy project 3 of the Institute of Biochemistry and POSDRU/89/1.5/S/60746; Catalina Petrareanu was supported by the Sectoral Operational Programme Human Resources Development 2007–2013 of the Romanian Ministry of Labour, Family, and Social Protection through the Financial Agreement POSDRU/107/1.5/S/76903 grant.

References

1. Beck J, Nassal M (2007) Hepatitis B virus replication. World J Gastroenterol 13:48–64
2. Lee WM (1997) Hepatitis B virus infection. N Engl J Med 337:1733–1745
3. Yoffe B, Noonan CA, Melnick JL, Hollinger FB (1986) Hepatitis B virus DNA in mononuclear cells and analysis of cell subsets for the presence of replicative intermediates of viral DNA. J Infect Dis 153:471–477
4. Kock J, Theilmann L, Galle P, Schlicht HJ (1996) Hepatitis B virus nucleic acids associated with human peripheral blood mononuclear cells do not originate from replicating virus. Hepatology 23:405–413

5. Stoll-Becker S, Repp R, Glebe D, Schaefer S, Kreuder J, Kann M, Lampert F, Gerlich WH (1997) Transcription of hepatitis B virus in peripheral blood mononuclear cells from persistently infected patients. J Virol 71:5399–5407
6. Hoofnagle JH (1992) Molecular biology of the hepatitis B virus. In: McLachlan A. CRC Press, Boca Raton, 1991, p 330. $179.95. Hepatology 15: 976–977
7. Bouchard MJ, Schneider RJ (2004) The enigmatic X gene of hepatitis B virus. J Virol 78:12725–12734
8. Twu JS, Schloemer RH (1987) Transcriptional trans-activating function of hepatitis B virus. J Virol 61:3448–3453
9. Benn J, Schneider RJ (1994) Hepatitis B virus HBx protein activates Ras-GTP complex formation and establishes a Ras, Raf, MAP kinase signaling cascade. Proc Natl Acad Sci U S A 91:10350–10354
10. Bouchard MJ, Wang LH, Schneider RJ (2001) Calcium signaling by HBx protein in hepatitis B virus DNA replication. Science 294:2376–2378
11. Becker SA, Lee TH, Butel JS, Slagle BL (1998) Hepatitis B virus X protein interferes with cellular DNA repair. J Virol 72:266–272
12. Chang C, Enders G, Sprengel R, Peters N, Varmus HE, Ganem D (1987) Expression of the precore region of an avian hepatitis B virus is not required for viral replication. J Virol 61:3322–3325
13. Milich DR, Jones JE, Hughes JL, Price J, Raney AK, McLachlan A (1990) Is a function of the secreted hepatitis B e antigen to induce immunologic tolerance in utero? Proc Natl Acad Sci U S A 87:6599–6603
14. Patient R, Hourioux C, Roingeard P (2009) Morphogenesis of hepatitis B virus and its subviral envelope particles. Cell Microbiol 11:1561–1570
15. Papatheodoridis GV, Manolakopoulos S, Dusheiko G, Archimandritis AJ (2008) Therapeutic strategies in the management of patients with chronic hepatitis B virus infection. Lancet Infect Dis 8:167–178
16. Aloman C, Wands JR (2003) Resistance of HBV to adefovir dipivoxil: a case for combination antiviral therapy? Hepatology 38:1584–1587
17. Delaney WE, Locarnini S, Shaw T (2001) Resistance of hepatitis B virus to antiviral drugs: current aspects and directions for future investigation. Antivir Chem Chemother 12:1–35
18. Cardell K, Akerlind B, Sallberg M, Fryden A (2008) Excellent response rate to a double dose of the combined hepatitis A and B vaccine in previous nonresponders to hepatitis B vaccine. J Infect Dis 198:299–304
19. Ladner SK, Otto MJ, Barker CS, Zaifert K, Wang GH, Guo JT, Seeger C, King RW (1997) Inducible expression of human hepatitis B virus (HBV) in stably transfected hepatoblastoma cells: a novel system for screening potential inhibitors of HBV replication. Antimicrob Agents Chemother 41:1715–1720
20. Sells MA, Chen ML, Acs G (1987) Production of hepatitis B virus particles in Hep G2 cells transfected with cloned hepatitis B virus DNA. Proc Natl Acad Sci U S A 84:1005–1009
21. Acs G, Sells MA, Purcell RH, Price P, Engle R, Shapiro M, Popper H (1987) Hepatitis B virus produced by transfected Hep G2 cells causes hepatitis in chimpanzees. Proc Natl Acad Sci U S A 84:4641–4644
22. Tong A, Wu L, Lin Q, Lau QC, Zhao X, Li J, Chen P, Chen L, Tang H, Huang C, Wei YQ (2008) Proteomic analysis of cellular protein alterations using a hepatitis B virus-producing cellular model. Proteomics 8:2012–2023
23. Cui F, Wang Y, Wang J, Wei K, Hu J, Liu F, Wang H, Zhao X, Zhang X, Yang X (2006) The up-regulation of proteasome subunits and lysosomal proteases in hepatocellular carcinomas of the HBx gene knockin transgenic mice. Proteomics 6:498–504
24. Zhang Z, Protzer U, Hu Z, Jacob J, Liang TJ (2004) Inhibition of cellular proteasome activities enhances hepadnavirus replication in an HBX-dependent manner. J Virol 78:4566–4572
25. Lazar C, Macovei A, Petrescu S, Branza-Nichita N (2012) Activation of ERAD pathway by human hepatitis B virus modulates viral and subviral particle production. PLoS One 7:e34169

26. Rogers J (1989) Calcium-binding proteins: the search for functions. Nature 339:661–662
27. Choi Y, Gyoo Park S, Yoo JH, Jung G (2005) Calcium ions affect the hepatitis B virus core assembly. Virology 332:454–463
28. Napoli JL (1999) Retinoic acid: its biosynthesis and metabolism. Prog Nucleic Acid Res Mol Biol 63:139–188
29. Balmer JE, Blomhoff R (2002) Gene expression regulation by retinoic acid. J Lipid Res 43:1773–1808
30. Huan B, Siddiqui A (1992) Retinoid X receptor RXR alpha binds to and trans-activates the hepatitis B virus enhancer. Proc Natl Acad Sci U S A 89:9059–9063
31. Quadro L, Blaner WS, Salchow DJ, Vogel S, Piantedosi R, Gouras P, Freeman S, Cosma MP, Colantuoni V, Gottesman ME (1999) Impaired retinal function and vitamin A availability in mice lacking retinol-binding protein. EMBO J 18:4633–4644
32. Kawaguchi R, Yu J, Honda J, Hu J, Whitelegge J, Ping P, Wiita P, Bok D, Sun H (2007) A membrane receptor for retinol binding protein mediates cellular uptake of vitamin A. Science 315:820–825
33. Diamond DL, Jacobs JM, Paeper B, Proll SC, Gritsenko MA, Carithers RL Jr, Larson AM, Yeh MM, Camp DG 2nd, Smith RD, Katze MG (2007) Proteomic profiling of human liver biopsies: hepatitis C virus-induced fibrosis and mitochondrial dysfunction. Hepatology 46:649–657
34. Mawson AR, Steele TA (2001) Possible role of retinoids in hepatitis B virus-associated liver damage. Exp Biol Med 226:734–739
35. Melia TJ Jr (2007) Putting the clamps on membrane fusion: how complexin sets the stage for calcium-mediated exocytosis. FEBS Lett 581:2131–2139
36. Falkowski MA, Thomas DD, Groblewski GE (2010) Complexin 2 modulates vesicle-associated membrane protein (VAMP) 2-regulated zymogen granule exocytosis in pancreatic acini. J Biol Chem 285:35558–35566
37. Liu K, Qian L, Wang J, Li W, Deng X, Chen X, Sun W, Wei H, Qian X, Jiang Y, He F (2009) Two-dimensional blue native/SDS-PAGE analysis reveals heat shock protein chaperone machinery involved in hepatitis B virus production in HepG2.2.15 cells. Mol Cell Proteomics 8:495–505
38. Camacho-Carvajal MM, Wollscheid B, Aebersold R, Steimle V, Schamel WW (2004) Two-dimensional Blue native/SDS gel electrophoresis of multi-protein complexes from whole cellular lysates: a proteomics approach. Mol Cell Proteomics 3:176–182
39. Wittig I, Braun HP, Schagger H (2006) Blue native PAGE. Nat Protoc 1:418–428
40. Craig EA, Gambill BD, Nelson RJ (1993) Heat shock proteins: molecular chaperones of protein biogenesis. Microbiol Rev 57:402–414
41. Jindal S, Young RA (1992) Vaccinia virus infection induces a stress response that leads to association of Hsp70 with viral proteins. J Virol 66:5357–5362
42. Nakagawa S, Umehara T, Matsuda C, Kuge S, Sudoh M, Kohara M (2007) Hsp90 inhibitors suppress HCV replication in replicon cells and humanized liver mice. Biochem Biophys Res Commun 353:882–888
43. Park SG, Jung G (2001) Human hepatitis B virus polymerase interacts with the molecular chaperonin Hsp60. J Virol 75:6962–6968
44. Hu J, Toft DO, Seeger C (1997) Hepadnavirus assembly and reverse transcription require a multi-component chaperone complex which is incorporated into nucleocapsids. EMBO J 16:59–68
45. Sydor JR, Normant E, Pien CS, Porter JR, Ge J, Grenier L, Pak RH, Ali JA, Dembski MS, Hudak J, Patterson J, Penders C, Pink M, Read MA, Sang J, Woodward C, Zhang Y, Grayzel DS, Wright J, Barrett JA, Palombella VJ, Adams J, Tong JK (2006) Development of 17-allylamino-17-demethoxygeldanamycin hydroquinone hydrochloride (IPI-504), an anti-cancer agent directed against Hsp90. Proc Natl Acad Sci U S A 103:17408–17413
46. Bremer CM, Bung C, Kott N, Hardt M, Glebe D (2009) Hepatitis B virus infection is dependent on cholesterol in the viral envelope. Cell Microbiol 11:249–260

47. Macovei A, Radulescu C, Lazar C, Petrescu S, Durantel D, Dwek RA, Zitzmann N, Nichita NB (2010) Hepatitis B virus requires intact caveolin-1 function for productive infection in HepaRG cells. J Virol 84:243–253
48. Allen JA, Halverson-Tamboli RA, Rasenick MM (2007) Lipid raft microdomains and neurotransmitter signalling. Nat Rev Neurosci 8:128–140
49. Balasubramanian N, Scott DW, Castle JD, Casanova JE, Schwartz MA (2007) Arf6 and microtubules in adhesion-dependent trafficking of lipid rafts. Nat Cell Biol 9:1381–1391
50. Helms JB, Zurzolo C (2004) Lipids as targeting signals: lipid rafts and intracellular trafficking. Traffic 5:247–254
51. Xie N, Huang K, Zhang T, Lei Y, Liu R, Wang K, Zhou S, Li J, Wu J, Wu H, Deng C, Zhao X, Nice EC, Huang C (2012) Comprehensive proteomic analysis of host cell lipid rafts modified by HBV infection. J Proteomics 75:725–739
52. Ma Y, Yu J, Chan HL, Chen YC, Wang H, Chen Y, Chan CY, Go MY, Tsai SN, Ngai SM, To KF, Tong JH, He QY, Sung JJ, Kung HF, Cheng CH, He ML (2009) Glucose-regulated protein 78 is an intracellular antiviral factor against hepatitis B virus. Mol Cell Proteomics 8:2582–2594
53. Chua PK, Wang RY, Lin MH, Masuda T, Suk FM, Shih C (2005) Reduced secretion of virions and hepatitis B virus (HBV) surface antigen of a naturally occurring HBV variant correlates with the accumulation of the small S envelope protein in the endoplasmic reticulum and Golgi apparatus. J Virol 79:13483–13496
54. Dong D, Ni M, Li J, Xiong S, Ye W, Virrey JJ, Mao C, Ye R, Wang M, Pen L, Dubeau L, Groshen S, Hofman FM, Lee AS (2008) Critical role of the stress chaperone GRP78/BiP in tumor proliferation, survival, and tumor angiogenesis in transgene-induced mammary tumor development. Cancer Res 68:498–505
55. Parkin DM, Bray F, Ferlay J, Pisani P (2001) Estimating the world cancer burden: Globocan 2000. Int J Cancer 94:153–156
56. Tan TL, Chen WN (2005) A proteomics analysis of cellular proteins associated with HBV genotype-specific HBX: potential in identification of early diagnostic markers for HCC. J Clin Virol 33:293–298
57. Tan TL, Feng Z, Lu YW, Chan V, Chen WN (2006) Adhesion contact kinetics of HepG2 cells during Hepatitis B virus replication: involvement of SH3-binding motif in HBX. Biochim Biophys Acta 1762:755–766
58. Macdonald A, Crowder K, Street A, McCormick C, Saksela K, Harris M (2003) The hepatitis C virus non-structural NS5A protein inhibits activating protein-1 function by perturbing ras-ERK pathway signaling. J Biol Chem 278:17775–17784
59. Tong A, Gou L, Lau QC, Chen B, Zhao X, Li J, Tang H, Chen L, Tang M, Huang C, Wei YQ (2009) Proteomic profiling identifies aberrant epigenetic modifications induced by hepatitis B virus X protein. J Proteome Res 8:1037–1046
60. Shim YH, Yoon GS, Choi HJ, Chung YH, Yu E (2003) p16 Hypermethylation in the early stage of hepatitis B virus-associated hepatocarcinogenesis. Cancer Lett 190:213–219
61. Zhong S, Tang MW, Yeo W, Liu C, Lo YM, Johnson PJ (2002) Silencing of GSTP1 gene by CpG island DNA hypermethylation in HBV-associated hepatocellular carcinomas. Clin Cancer Res 8:1087–1092
62. Jung JK, Arora P, Pagano JS, Jang KL (2007) Expression of DNA methyltransferase 1 is activated by hepatitis B virus X protein via a regulatory circuit involving the p16INK4a-cyclin D1-CDK 4/6-pRb-E2F1 pathway. Cancer Res 67:5771–5778
63. Niu D, Sui J, Zhang J, Feng H, Chen WN (2009) iTRAQ-coupled 2-D LC-MS/MS analysis of protein profile associated with HBV-modulated DNA methylation. Proteomics 9:3856–3868
64. Mohammad HS, Kurokohchi K, Yoneyama H, Tokuda M, Morishita A, Jian G, Shi L, Murota M, Tani J, Kato K, Miyoshi H, Deguchi A, Himoto T, Usuki H, Wakabayashi H, Izuishi K, Suzuki Y, Iwama H, Deguchi K, Uchida N, Sabet EA, Arafa UA, Hassan AT, El-Sayed AA, Masaki T (2008) Annexin A2 expression and phosphorylation are up-regulated in hepatocellular carcinoma. Int J Oncol 33:1157–1163

65. Lim SO, Park SJ, Kim W, Park SG, Kim HJ, Kim YI, Sohn TS, Noh JH, Jung G (2002) Proteome analysis of hepatocellular carcinoma. Biochem Biophys Res Commun 291: 1031–1037
66. Liang CR, Leow CK, Neo JC, Tan GS, Lo SL, Lim JW, Seow TK, Lai PB, Chung MC (2005) Proteome analysis of human hepatocellular carcinoma tissues by two-dimensional difference gel electrophoresis and mass spectrometry. Proteomics 5:2258–2271
67. Martinez-Chantar ML, Corrales FJ, Martinez-Cruz LA, Garcia-Trevijano ER, Huang ZZ, Chen L, Kanel G, Avila MA, Mato JM, Lu SC (2002) Spontaneous oxidative stress and liver tumors in mice lacking methionine adenosyltransferase 1A. FASEB J 16:1292–1294
68. Li C, Hong Y, Tan YX, Zhou H, Ai JH, Li SJ, Zhang L, Xia QC, Wu JR, Wang HY, Zeng R (2004) Accurate qualitative and quantitative proteomic analysis of clinical hepatocellular carcinoma using laser capture microdissection coupled with isotope-coded affinity tag and two-dimensional liquid chromatography mass spectrometry. Mol Cell Proteomics 3:399–409
69. Ai J, Tan Y, Ying W, Hong Y, Liu S, Wu M, Qian X, Wang H (2006) Proteome analysis of hepatocellular carcinoma by laser capture microdissection. Proteomics 6:538–546
70. Sun W, Xing B, Sun Y, Du X, Lu M, Hao C, Lu Z, Mi W, Wu S, Wei H, Gao X, Zhu Y, Jiang Y, Qian X, He F (2007) Proteome analysis of hepatocellular carcinoma by two-dimensional difference gel electrophoresis: novel protein markers in hepatocellular carcinoma tissues. Mol Cell Proteomics 6:1798–1808
71. Shin CH, Liu ZP, Passier R, Zhang CL, Wang DZ, Harris TM, Yamagishi H, Richardson JA, Childs G, Olson EN (2002) Modulation of cardiac growth and development by HOP, an unusual homeodomain protein. Cell 110:725–735
72. He QY, Lau GK, Zhou Y, Yuen ST, Lin MC, Kung HF, Chiu JF (2003) Serum biomarkers of hepatitis B virus infected liver inflammation: a proteomic study. Proteomics 3:666–674
73. Panduro A, Lin-Lee YC, Chan L, Shafritz DA (1990) Transcriptional and posttranscriptional regulation of apolipoprotein E, A-I, and A-II gene expression in normal rat liver and during several pathophysiologic states. Biochemistry 29:8430–8435
74. Qian X, Samadani U, Porcella A, Costa RH (1995) Decreased expression of hepatocyte nuclear factor 3 alpha during the acute-phase response influences transthyretin gene transcription. Mol Cell Biol 15:1364–1376
75. Steel LF, Shumpert D, Trotter M, Seeholzer SH, Evans AA, London WT, Dwek R, Block TM (2003) A strategy for the comparative analysis of serum proteomes for the discovery of biomarkers for hepatocellular carcinoma. Proteomics 3:601–609
76. Comunale MA, Mattu TS, Lowman MA, Evans AA, London WT, Semmes OJ, Ward M, Drake R, Romano PR, Steel LF, Block TM, Mehta A (2004) Comparative proteomic analysis of de-N-glycosylated serum from hepatitis B carriers reveals polypeptides that correlate with disease status. Proteomics 4:826–838
77. Feng JT, Liu YK, Song HY, Dai Z, Qin LX, Almofti MR, Fang CY, Lu HJ, Yang PY, Tang ZY (2005) Heat-shock protein 27: a potential biomarker for hepatocellular carcinoma identified by serum proteome analysis. Proteomics 5:4581–4588
78. Gruden G, Carucci P, Lolli V, Cosso L, Dellavalle E, Rolle E, Cantamessa A, Pinach S, Abate ML, Campra D, Brunello F, Bruno G, Rizzetto M, Perin PC (2013) Serum heat shock protein 27 levels in patients with hepatocellular carcinoma. Cell Stress Chaperones 18:235–241
79. Block TM, Comunale MA, Lowman M, Steel LF, Romano PR, Fimmel C, Tennant BC, London WT, Evans AA, Blumberg BS, Dwek RA, Mattu TS, Mehta AS (2005) Use of targeted glycoproteomics to identify serum glycoproteins that correlate with liver cancer in woodchucks and humans. Proc Natl Acad Sci U S A 102:779–784
80. D'Ugo E, Argentini C, Giuseppetti R, Canitano A, Catone S, Rapicetta M (2010) The woodchuck hepatitis B virus infection model for the evaluation of HBV therapies and vaccine therapies. Expert Opin Drug Discov 5:1153–1162
81. Naitoh A, Aoyagi Y, Asakura H (1999) Highly enhanced fucosylation of serum glycoproteins in patients with hepatocellular carcinoma. J Gastroenterol Hepatol 14:436–445
82. Wei H, Li B, Zhang R, Hao X, Huang Y, Qiao Y, Hou J, Li X, Li X (2013) Serum GP73, a marker for evaluating progression in patients with chronic HBV infections. PLoS One 8:e53862

83. Cui J, Kang X, Dai Z, Huang C, Zhou H, Guo K, Li Y, Zhang Y, Sun R, Chen J, Li Y, Tang Z, Uemura T, Liu Y (2007) Prediction of chronic hepatitis B, liver cirrhosis and hepatocellular carcinoma by SELDI-based serum decision tree classification. J Cancer Res Clin Oncol 133:825–834
84. He QY, Zhu R, Lei T, Ng MY, Luk JM, Sham P, Lau GK, Chiu JF (2008) Toward the proteomic identification of biomarkers for the prediction of HBV related hepatocellular carcinoma. J Cell Biochem 103:740–752
85. Yang JD, Roberts LR (2010) Hepatocellular carcinoma: A global view. Nat Rev Gastroenterol Hepatol 7:448–458
86. Honda M, Kaneko S, Kawai H, Shirota Y, Kobayashi K (2001) Differential gene expression between chronic hepatitis B and C hepatic lesion. Gastroenterology 120:955–966
87. Nguyen H, Mudryj M, Guadalupe M, Dandekar S (2003) Hepatitis C virus core protein expression leads to biphasic regulation of the p21 cdk inhibitor and modulation of hepatocyte cell cycle. Virology 312:245–253
88. Kim W, Oe Lim S, Kim JS, Ryu YH, Byeon JY, Kim HJ, Kim YI, Heo JS, Park YM, Jung G (2003) Comparison of proteome between hepatitis B virus- and hepatitis C virus-associated hepatocellular carcinoma. Clin Cancer Res 9:5493–5500
89. Chan WY, Liu QR, Borjigin J, Busch H, Rennert OM, Tease LA, Chan PK (1989) Characterization of the cDNA encoding human nucleophosmin and studies of its role in normal and abnormal growth. Biochemistry 28:1033–1039
90. Kinoshita M, Miyata M (2002) Underexpression of mRNA in human hepatocellular carcinoma focusing on eight loci. Hepatology 36:433–438
91. Selimovic D, El-Khattouti A, Ghozlan H, Haikel Y, Abdelkader O, Hassan M (2012) Hepatitis C virus-related hepatocellular carcinoma: an insight into molecular mechanisms and therapeutic strategies. World J Hepatol 4:342–355
92. Lin MT, Wang MY, Liaw KY, Lee PH, Chien SF, Tsai JS, Lin-Shiau SY (2001) Superoxide dismutase in hepatocellular carcinoma affects patient prognosis. Hepatogastroenterology 48:1102–1105
93. Xu XR, Huang J, Xu ZG, Qian BZ, Zhu ZD, Yan Q, Cai T, Zhang X, Xiao HS, Qu J, Liu F, Huang QH, Cheng ZH, Li NG, Du JJ, Hu W, Shen KT, Lu G, Fu G, Zhong M, Xu SH, Gu WY, Huang W, Zhao XT, Hu GX, Gu JR, Chen Z, Han ZG (2001) Insight into hepatocellular carcinogenesis at transcriptome level by comparing gene expression profiles of hepatocellular carcinoma with those of corresponding noncancerous liver. Proc Natl Acad Sci U S A 98:15089–15094
94. Fukutomi T, Zhou Y, Kawai S, Eguchi H, Wands JR, Li J (2005) Hepatitis C virus core protein stimulates hepatocyte growth: correlation with upregulation of wnt-1 expression. Hepatology 41:1096–1105
95. Miller JR (2002) The Wnts. Genome Biol 3:REVIEWS3001
96. Reya T, Clevers H (2005) Wnt signalling in stem cells and cancer. Nature 434:843–850
97. Lee TH, Tai DI, Cheng CJ, Sun CS, Lin CY, Sheu MJ, Lee WP, Peng CY, Wang AH, Tsai SL (2006) Enhanced nuclear factor-kappa B-associated Wnt-1 expression in hepatitis B- and C-related hepatocarcinogenesis: identification by functional proteomics. J Biomed Sci 13:27–39
98. Tai DI, Tsai SL, Chang YH, Huang SN, Chen TC, Chang KS, Liaw YF (2000) Constitutive activation of nuclear factor kappaB in hepatocellular carcinoma. Cancer 89:2274–2281
99. Sarvari J, Mojtahedi Z, Taghavi SA, Kuramitsu Y, Shamsi Shahrabadi M, Ghaderi A, Nakamura K (2013) Differentially expressed proteins in chronic active hepatitis, cirrhosis, and HCC related to HCV infection in comparison with HBV infection: a proteomics study. Hepat Mon 13:e8351
100. Mondal G, Chatterjee U, Das HR, Chatterjee BP (2009) Enhanced expression of alpha1-acid glycoprotein and fucosylation in hepatitis B patients provides an insight into pathogenesis. Glycoconj J 26:1225–1234
101. Gripon P, Rumin S, Urban S, Le Seyec J, Glaise D, Cannie I, Guyomard C, Lucas J, Trepo C, Guguen-Guillouzo C (2002) Infection of a human hepatoma cell line by hepatitis B virus. Proc Natl Acad Sci U S A 99:15655–15660

102. Guillouzo A, Guguen-Guillouzo C (2008) Evolving concepts in liver tissue modeling and implications for in vitro toxicology. Expert Opin Drug Metab Toxicol 4:1279–1294
103. Narayan R, Gangadharan B, Hantz O, Antrobus R, Garcia A, Dwek RA, Zitzmann N (2009) Proteomic analysis of HepaRG cells: a novel cell line that supports hepatitis B virus infection. J Proteome Res 8:118–122
104. Sokolowska I, Dorobantu C, Woods AG, Macovei A, Branza-Nichita N, Darie CC (2012) Proteomic analysis of plasma membranes isolated from undifferentiated and differentiated HepaRG cells. Proteome Sci 10:47
105. Rescher U, Gerke V (2004) Annexins—unique membrane binding proteins with diverse functions. J Cell Sci 117:2631–2639
106. Petrareanu C, Macovei A, Sokolowska I, Woods AG, Lazar C, Radu GL, Darie CC, Branza-Nichita N (2013) Comparative proteomics reveals novel components at the plasma membrane of differentiated HepaRG cells and different distribution in hepatocyte- and biliary-like cells. PLoS One 8:e71859
107. Darie CC, Deinhardt K, Zhang G, Cardasis HS, Chao MV, Neubert TA (2011) Identifying transient protein-protein interactions in EphB2 signaling by blue native PAGE and mass spectrometry. Proteomics 11:4514–4528

Chapter 23
Oxidative Stress and Antibiotic Resistance in Bacterial Pathogens: State of the Art, Methodologies, and Future Trends

Mouna Marrakchi, Xiaobo Liu, and Silvana Andreescu

Abstract Despite the significant advances of modern medicine and the availability of a wide variety of antibiotics for the treatment of microbial infections, there is an alarming increase of multiresistant bacterial pathogens. This chapter discusses the status of bacterial resistance mechanisms and the relationship with oxidative stress and provides an overview of the methods used to assess oxidative conditions and their contribution to the antibiotic resistance.

23.1 Introduction

Humans, animals, and plants are in continuous contact with beneficial, harmless, or pathogenic bacteria. Diagnosis of bacterial infections and efficient treatment of infectious diseases are crucial for human health [1]. Eighty years ago, Alexander Fleming's discovery of penicillin had changed the world of modern medicine by introducing the age of useful antibiotic [2]. A discovery rewarded in 1945 by the Nobel Prize in Medicine. Yet, within 15 years of his findings, Fleming presciently

M. Marrakchi (✉)
Department of Chemistry and Biomolecular Science, Clarkson University, 8 Clarkson Ave, Potsdam, NY 13699-5810, USA

Laboratoire d'Ecologie et de Technologie Microbienne, Carthage University, INSAT, Centre Urbain Nord, BP 676, 1080 Charguia Cedex, Tunisia
e-mail: mounamarrakchi@yahoo.fr

X. Liu • S. Andreescu
Department of Chemistry and Biomolecular Science, Clarkson University, 8 Clarkson Ave, Potsdam, NY 13699-5810, USA
e-mail: eandrees@clarkson.edu

A.G. Woods and C.C. Darie (eds.), *Advancements of Mass Spectrometry in Biomedical Research*, Advances in Experimental Medicine and Biology 806,
DOI 10.1007/978-3-319-06068-2_23, © Springer International Publishing Switzerland 2014

hypothesized that bacteria would likely attain resistance to any antibiotic treatment given the right circumstance. The continued emergence of single and multiple antibiotic-resistant bacterial strains is one of the more important societal issues today. Justifiably, the focus of antibiotic resistance research in the last half century has been on the elucidation of the mechanisms by which microbes can physically alter a drug's structure, disrupt the interaction between a drug and its cellular target, or alter the behavior and efficiency of its own transport machinery to reduce access to a drug's cellular target [3].

Despite the significant advances in antibacterial therapy today, death of bacteria in response to antibiotic exposure remains largely unknown. Although it is possible to measure the physiological effects of antibiotics, such as loss of membrane permeability, changes in cell morphology, and molecular effects (e.g., inhibition or cellular pathways) [4], the causes of the ever-increasing prevalence of antibiotic-resistant strains remain obscure. Indeed, this developed resistance has many consequences. In many cases, the infected person with a resistant microorganism is more likely to require hospitalization, to double the duration of their hospital stay, and/or to experience increased risk of death and morbidity. Understanding the mechanisms of antibiotic resistance can advance novel therapeutic approaches and serve as a foundation for the development of new antibiotics.

This chapter discusses the status of bacterial resistance mechanisms and the relationship with oxidative stress and provides an overview of the methods used to assess oxidative stress and mechanisms of antibiotic resistance. A general survey of conventional biological methodologies and the role of proteomics in assessing bacterial resistance and oxidative stress are provided.

23.2 Antibiotics and Bacterial Resistance Mechanisms

Antibiotics (compounds that are by definition "against life") are typically antibacterial drugs, interfering with processes that are essential to bacterial growth or survival without harm to the eukaryotic host harboring the infecting bacteria [5]. Antibiotics can have bactericidal effect (resulting in cell death) or bacteriostatic effect (stop bacterial growth). In order to understand the mechanism by which bacteria develop antibiotic resistance, it is important to study the different targets and reaction pathways for the main classes of antibacterial drugs and bacterial pathogens.

23.2.1 How Antibiotics Work

It is not clear whether there is a single major mechanism of bacterial cell death from antibiotics or many. Most antibiotics function through inhibition of essential cellular processes, an intervention to which there are likely to be many consequences, and as a result, there would be equally numerous ways to achieve bacterial killing [4].

There are three proven targets for the main antibacterial drugs:

1. *Bacterial cell wall biosynthesis*: Most bacteria produce a cell wall that is composed partly of a macromolecule called peptidoglycan, itself made up of amino sugars and short peptides. Penicillin, one of the first antibiotics to be used widely (and other β-Lactam antibiotics), targets cell wall synthesis by inhibiting the formation of peptidoglycan cross-links in the bacterial cell wall, a process also known as transpeptidation. The result is a very fragile cell wall that bursts, killing the bacterium.
2. *Bacterial protein synthesis*: Drugs that target and inhibit protein synthesis can be divided in two classes: 50S ribosome inhibitors and the 30S inhibitors. Indeed the ribosome (serving as the primary site of biological protein synthesis (translation)) is composed of two ribonucleprotein subunits, the 50S and 30S, which assemble (during the initiation phase) following the formation of a complex between an mRNA transcript, N-formylmethionine-charged aminoacyl tRNA, several initiation factors, and a free 30S subunit [6]. Tetracycline, for example, can cross the membranes of bacteria and accumulate in high concentrations in the cytoplasm. Tetracycline then binds to a single site on the 30S ribosomal subunit and blocks that key RNA interaction, which shuts off the lengthening protein chain.
3. *Bacterial DNA replication and repair*: Bacterial chromosomal topology is maintained by the activities of topoisomerase I, topoisomerase IV, and DNA gyrase (topoisomerase II) [7]. These reactions are exploited by the synthetic quinolone class of antimicrobials which target DNA–topoisomerase complexes. The quinolone class of antimicrobials interferes with the maintenance of chromosomal topology by targeting topoisomerase II and topoisomerase IV, trapping these enzymes at the DNA cleavage stage and preventing strand rejoining [6]. The process leads to the complete inhibition of cell division and results to bacteriostatic effects and ultimately cell death.

While the antibiotic drug–target interactions and their respective direct effects are well known, as discussed above, the bacterial responses to antibiotic drug treatments that contribute to cell death are complex and not as well understood. It was reported that the three major classes of bactericidal drugs, regardless of drug–target interaction, all utilize a common mechanism of inactivation whereby they stimulate the production of lethal doses of hydroxyl radicals via the Fenton reaction [8]. The generation of these highly destructive hydroxyl radicals is the result of the iron misregulation by the superoxide-mediated oxidation of iron–sulfur clusters, a process that promotes a breakdown of iron regulatory dynamics [3]. This oxidative stress contributes to bactericidal antibiotic-mediated cell death. Kohanski and colleagues have studied the antibiotic-induced stress response networks to determine how the primary effect of a given bactericidal drug triggers aspects of cell death that are common to all bactericidal drugs. They showed that an aminoglycoside-antibiotic (which is known to be a protein synthesis inhibitor) also induces oxidative stress and cell death. These studies indicate that oxidative stress is involved in the antibiotic resistance effects in pathogenic bacteria and that exposure of bacteria to antibiotics may alter the antioxidant defense system and redox mechanisms in cells.

The discovery of the existence of a common oxidative damage cellular death pathway can be helpful for the development of more effective antibacterial therapies. ROS (reactive oxygen species), such as superoxide ($O_2^{\bullet -}$), and hydroxyl radicals ($HO^{\bullet}$), as well as RNS (reactive nitrogen species) such as nitric oxide (NO) and peroxynitrite ($ONOO^-$), are highly toxic, as a result of their actions as oxidizing and nitrating agents, and can have damaging effects on bacterial physiology. There is still much to be learned about how oxidative stress related changes in bacterial physiology, affect antibiotic-mediated cell death and the emergence of resistance [6]. Here, we provide an overview of the current state of the art in this field and the methods that can be used to study such effects.

23.2.2 How Bacteria Fight Antibiotics Effects

Bacteria can develop resistance to virtually any antimicrobial agent at varietal stages [9, 10]. The evaluation of a new antimicrobial agent typically involves the study of organisms that are either naturally resistant or susceptible to that agent, thus defining a broad spectrum of activity for that agent [11]. The following section discusses the origin, evolution, and current understanding of antibiotic resistance and the processes that make antibiotic resistance inevitable.

In bacteria, the front line of this resistance system is the cell envelope. In Gram-negative bacteria this includes the outer membrane, which is composed of an asymmetric lipopolysaccharide–phospholipid bilayer, and provides an effective physical barrier to the entry of molecules (including many antibiotics) into the cell. Outer membrane-spanning porins that facilitate the entry of small molecules into cells also passively excludes many antibiotics. In Gram-positive bacteria the absence of an outer membrane results in increased sensitivity to many antibiotics. Nonetheless, many Gram-positive bacteria, such as *Mycobacteria* species, can fight the cytotoxic effects of antibiotics using physiological defenses [12]. In addition to the protective cell envelope, there are different other mechanisms of acquired antimicrobial resistance. These include possible changes in the drug target (e.g., reduction of receptor affinity and the substitution of an alternative pathway), the production of a detoxifying enzyme, or decreased antibiotic uptake (through diminished permeability or an active efflux system) [11]. Genetic modifications can also be used to increase bacteria resistance against antibiotics [13]. These modifications are performed via plasmid conjugation, phage-based transduction, or lateral gene transfer [14, 15], activation of latent mobile genetic elements, and the mutagenesis of its own DNA [12, 13]. The presence of antibiotic resistance elements in pathogenic bacteria is, mostly the result of the horizontal gene transfer, a process by which bacteria acquire resistance genes form environmental bacteria [16, 17]. A principal mechanism for the fast spread of antibiotic-resistance genes through bacterial populations is that such genes get collected on plasmids that are independently replicated within and passed between bacterial cells and species [5]. This fast acquisition of resistance is facilitated by the environmental antibiotic pressure.

Under antibiotic stress, a few spontaneous drug-resistant mutants can enhance the survival capacity of the overall population in that stressful environment. This protective effect is resultant from the production and sharing of the metabolite indole (produced by antibiotic-resistant mutants), a signaling molecule, that could turn on drug-efflux pumps and activate oxidative stress protective mechanisms [18]. In any event, the survival of a bacterium amidst oxidative stress depends on the evolution of a series of defense mechanisms, which include:

1. Detoxifying enzymes (enzymatic antibiotic inactivation) [19] and free radical-scavenging substrates.
2. DNA and protein repair systems.
3. Competition by substrates favoring bacterial survival.

In many cases, these defenses may be coordinately regulated [20].

23.3 The Role of Oxidative Stress, Antibiotic Function, and Emerging Bacterial Resistance Against It

In addition to the general antibiotic effects mentioned above, which include inhibition of cell wall assembly, protein synthesis, and DNA replication, a string of several other mechanisms have been correlated to cell death meditated by antibiotics. Prominently, oxidative stress has been suggested as a possible pathway involved in antibiotic effects and also in the development of antibiotic resistance in bacteria [21]. Studies suggested oxidative stress as a secondary mechanism to the primary modes of action of antibiotics [22].

Oxidative stress is a redox disequilibrium state, in which the generation of ROS overwhelms the antioxidant defense mechanisms [23]. According to the hierarchical oxidative stress model, a minor level of oxidative stress merely induces protective effects. However, at high level, excessive ROS may cause severe damage to cells, including necrosis and apoptosis [24]. Therefore, quantitative study of ROS release in cells is of particular interest in assessing the relationship between antibiotics and oxidative stress. Examples of techniques that have been used for this purpose are summarized in Table 23.1 and discussed below. Figure 23.1 provides an overview of the various parameters and methods involved in such studies.

Methods for monitoring ROS are a critical first step in unraveling the role of these species and their contribution to bacterial resistance and antibiotics susceptibility. Continuous monitoring of these species in biological systems is a significant challenge due to their high reactivity and short lifetime. Moreover, study of their kinetic characteristics is difficult due to the many interrelated redox reactions and low concentrations that change dynamically over time. Commonly used techniques involve indirect absorbance or fluorescence measurements [25, 26]. However, fluorescence probes are relatively nonspecific [27] due to light sensitivity, photobleaching, and poor selectivity. Many biological studies infer ROS levels from indirect

Table 23.1 Methods and strategies used to assess oxidative stress in antibiotic-treated bacteria

Methods and strategies	Microorganisms	Stress source	Results	References
Chemiluminescence method	*Staphylococcus aureus*, *E. coli*, *Pseudomonas aeruginosa*, *Enterococcus faecalis*	Ciprofloxacin, ampicillin, chloramphenicol, ceftazidime, piperacillin,	Fluorescent dye was oxidized by ROS, resulting increasing fluorescence intensity, demonstrated the cumulative ROS in antibiotic-treated bacteria.	[22, 39–41]
Investigation of Fenton reaction	*E. coli, Staphylococcus aureus*	Norfloxacin, ampicillin, kanamycin, chloramphenicol	Iron level regulation was disrupted by superoxide radicals, resulted an increased hydroxyl radical formation.	[3, 8, 40]
Addition of ROS scavenger	*E. coli, M. tuberculosis, Mycobacterium smegmatis*	Oxolinic acid, moxifloxacin, quinolone, ciprofloxacin, isoniazid, rifampin, streptomycin, clofazimine	Addition of thiourea and/or 2,2′-bipyridyl to cell culture increased survival rate, showed the existence of ROS in the culture and the killing effect was from ROS.	[3, 40, 42]
Mutant strain validation	*Staphylococcus aureus*, *E. coli*, *Pseudomonas aeruginosa, Enterococcus faecalis*	Ciprofloxacin, ampicillin, chloramphenicol, ceftazidime, piperacillin	Specific antibiotics sensitive mutant or knock out strain mutant were used to show the presence and augment of ROS in bacteria treated by these drugs.	[39, 41]
Gene expression	*E. coli, P. putida, P. aeruginosa*	Norfloxacin, ampicillin, kanamycin, chloramphenicol	Specific gene can be expressed and used as hallmark to verify cells are undergoing oxidative stress.	[3, 22]
MALDI-mass spectrometry	*Paracoccidioides* yeast, *Staphylococcus aureus, Bacillus anthracis, Leishmania donovani*	Hydrogen peroxide, photodynamic therapy (PDT) treatment	Oxidative stress-responsive proteins/isoforms were identified and grouped by function.	[50, 51, 54, 56]
ESI-mass spectrometry	*Fusobacterium nucleatum, Helicobacter pylori*	Atmospheric oxygen exposure, hydrogen peroxide	Oxidative stress-responsive proteins/isoforms were identified and grouped by function.	[52, 53]
Gel electrophoresis	*Paracoccidioides* yeast, *Staphylococcus aureus, Bacillus anthracis, Leishmania donovani, Fusobacterium nucleatum, Helicobacter pylori*	Hydrogen peroxide, PDT treatment, atmospheric oxygen exposure	Oxidative stress-responsive proteins/isoforms were isolated and predetermined according to their molecular weight.	[50–53]
PCR	*Paracoccidioides* yeast, *Staphylococcus aureus, Bacillus anthracis, Leishmania donovani, Fusobacterium nucleatum, Helicobacter pylori*	Hydrogen peroxide, PDT treatment, atmospheric oxygen exposure	DNA sequences responsible for oxidative stress were identified and expressed to validate proposed hypothesis.	[50, 53–56]

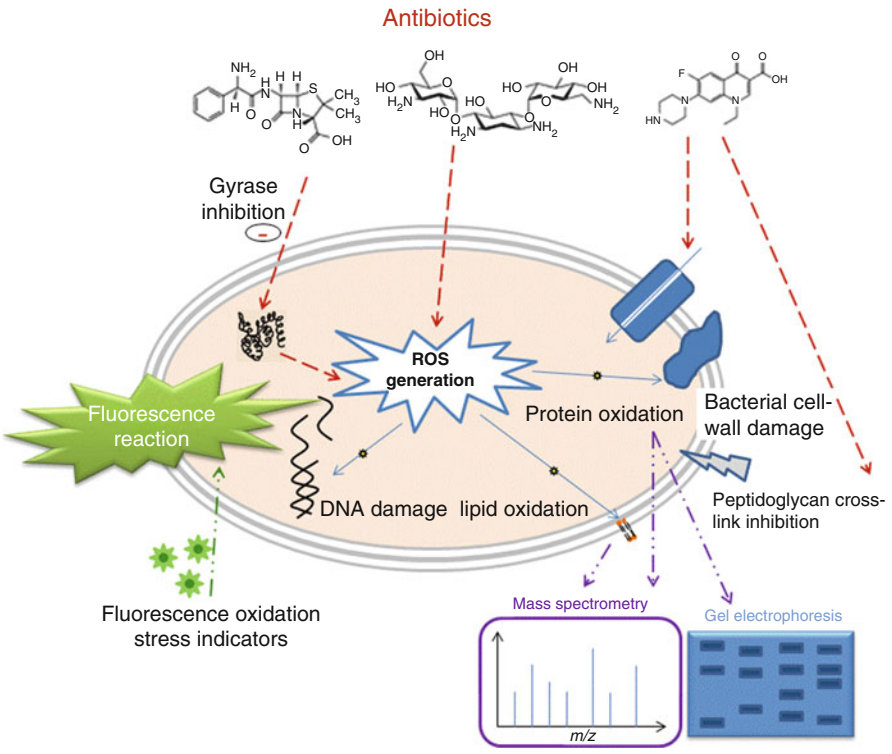


Fig. 23.1 Antibiotics effects on bacteria and assessment methods

measurements of products of protein nitrosylation (3-nitrotyrosine) or lipid peroxidation (4-hydroxynonenal). There are few methods that can measure ROS directly. Electron paramagnetic resonance (EPR) [28, 29] can be used to assess production of free radicals directly but the method has relatively low sensitivity and measurements can be challenging in complex biological environments. Other methods include chemiluminescence (CL), electrospray ionization-mass spectrometry (ESI-MS), gel electrophoresis, polymerase chain reaction (PCR), gene expression, and matrix-assisted laser desorption ionization mass spectrometry (MALDI-MS).

Electrochemistry is another method that can provide real-time measurements of ROS in complex biological environments [30]. Electrochemical sensors are relatively inexpensive and easy-to-use. Most work has been done with individual microelectrodes by Amatore's group to study cell secretion [31–34]. Practical problems that impede the broad implementation of electrochemical microsensors in the study of oxidative stress are mainly related to the difficulties of calibration and operation in complex biological samples due to interferences and instability of the radicals. Additionally and unfortunately, commercial microelectrochemical probes and high throughput electrochemical instrumentation are of limited availability and are not widely accessible to life scientists. Sensors that could simultaneously and

continuously assess the evolution of multiple ROS with high selectivity can provide quantitative measurements of free radicals in real time with high spatial and temporal resolution directly in bacterial cultures. Examples of applications of this technology were demonstrated in several biological environments [35, 36]. However, the use of these sensors to study bacterial pathogens and antibiotic susceptibility are limited [37, 38].

Several studies have used chemiluminescence (CL) to probe the presence of ROS in bacteria. Albesa et al. used CL measurements to assess the involvement of oxidative stress, particularly the role of superoxide in the action of antibiotics against different bacteria including *Staphylococcus aureus*, *Escherichia coli*, *Pseudomonas aeruginosa*, and *Enterococcus faecalis* [39]. Using CL, Albesa found that diverse antibiotics can increase superoxide release in various strains but only those strains that are sensitive to antibiotics show oxidative stress response [39]. In addition to superoxide, several studies have demonstrated the role of hydroxyl radicals ($HO^{\bullet}$) as an essential contributor to the oxidative stress response, and their involvement in the mode of action of antibiotics. Kohanski et al. studied *E. coli* treated with norfloxacin, ampicillin, and kanamycin and *Staphylococcus aureus* treated with norfloxacin and chloramphenicol and demonstrated release of hydroxyl radicals via the Fenton reaction [8]. A hydroxyl radical-specific dye, hydroxyphenyl fluorescein (HPF), was used in this study to assess the formation of hydroxyl radicals. In the Fenton reaction, the ferrous iron is driving formation of hydroxyl radicals; therefore iron chelators will inhibit hydroxyl radical generation allowing direct quantification of oxidative effects due to the presence of these radicals. Similar results were also observed by Grant et al. in norfloxacin-treated *E. coli* after addition of HPF [40]. Dwyer et al. hypothesized that generation of hydroxyl radicals might also involve superoxide-mediated oxidation of the iron–sulfur clusters, and showed that the generation of superoxide radicals disrupted iron regulatory dynamics, inducing iron misregulation in cells [3]. Yeom et al. also showed that antibiotics could accelerate cell death by promoting the Fenton reaction leading to an oxidative stress response in ampicillin-treated *Pseudomonas aeruginosa* [22]. Results of this study demonstrated that the antibiotic action is affected by modulation of reduced nicotinamide-adenine dinucleotide (NADH) levels and iron chelation. In addition to direct or indirect measurements of ROS, the involvement of ROS in cell death and antibiotic resistance can also be assessed by using ROS scavengers. Theoretically, an ROS scavenger will neutralize excessive ROS and increase the percentage of surviving bacteria after antibiotics exposure. Therefore, the addition of antioxidants to bacteria exposed to an oxidative stress environment would protect cells from the damaging effects of ROS. Goswami et al. investigated the protective effect of antioxidants in *Escherichia coli* exposed to ciprofloxacin. Both glutathione and ascorbic acid antioxidants have shown substantial protective effects, which demonstrated the involvement of ROS in the antibiotics-mediated cell death [41]. Reduced killing effects were observed when thiourea, a hydroxyl radical scavenger, was added to an antibiotic-treated cell culture. These results validate the hypothesis that cell death is mediated by hydroxyl radicals. Wang et al. reported similar effects after adding thiourea and/or 2,2′-bipyridyl to oxolinic acid, moxifloxacin, and quinolone-treated

E. coli [42]. Grant et al. used thiourea to remove hydroxyl radicals in *Mycobacterium smegmatis* and *M. tuberculosis* treated with ciprofloxacin, isoniazid, rifampin, streptomycin, and clofazimine, respectively, in order to establish the relationship between dissolved oxygen and ROS [40].

Another strategy to assess ROS release in antibiotic-treated bacteria is to use mutant strains that have an altered antioxidant defense mechanism. The antioxidant defensive system in bacteria comprises specific antioxidative enzymes including superoxide dismutase (SOD), catalase, and peroxidase. Through manipulation of *E. coli* strains, knockout strain can be generated to create an artificial imbalance state that allows selective study of oxidant/antioxidant mechanisms. Goswami et al. studied the effect of mutations in oxidative stress defense genes in SOD and alkyl hydroperoxide reductase on the sensitivity of *E. coli* to antibiotics and found that both superoxide and H_2O_2 may be involved in the antibacterial action of ciproflaxin [41]. Additional information about oxidative stress in bacteria treated by antibiotics can be obtained by phenotypic and gene expression analysis. Dwyer et al. performed several phenotypic and genetic analyses on gyrase-inhibited *E. coli*. and demonstrated that both superoxide and hydroxyl radical oxidative species are generated following inactivation by gyrase inhibitors. This oxidative response can amplify the inhibitor effect by oxidatively damaging DNA, proteins, and lipids [3]. The antimicrobial gyrase-catalyzed effects included DNA strand breakage and damage of the replication machinery that may ultimately result in cell death.

Investigation of oxidative stress-related processes and the role of ROS in bacteria exposed to antibiotics provide valuable information regarding the mechanism of bactericidal antibiotic-mediated cell death and the involvement of ROS in this process. These studies can potentially reveal novel aspects related to the mode of action of antibiotics against bacteria, information that can be useful in the future for the development of new antimicrobial drugs. Moreover, since ROS can cause damage to cell membrane and most importantly, proteins, it is equally important to assess changes in the cell proteome to establish the effect of the antibiotic-induced oxidative stress on the cell's proteome. Proteomic approaches such as mass spectrometry (MS) and gel electrophoresis, e.g., sodium dodecyl-sulfate polyacrylamide gel electrophoresis (SDS-PAGE) can be used to study these effects. Ongoing efforts in this direction are summarized in the following section.

23.4 Proteomic Investigation of Oxidative Stress in Bacteria

Under conditions that can cause oxidative stress, bacterial cells are exposed to excessive ROS that can oxidize membrane fatty acids, initiating lipid peroxidation [43], oxidize proteins [44], and cause DNA damage [45, 46]. The immediate effects on proteins include tyrosine hydroxylation, methionine or cysteine oxidation, and formation of carbonyl group on side chain amino acids [47]. As a result, modified proteins could be used as potential markers for oxidative stress. Several advanced instrumentations are available for the analysis of the proteome (e.g., proteins, peptides, glycans,

protein interactions, and post-translational modifications). These include ESI-MS, MALDI-MS, and chromatography-coupled and tandem techniques such as HPLC-ESI-MS, LC-MS, and LC-MS/MS [47, 48]. In general, these methods are used in conjunction with specific biochemical techniques such as SDS-PAGE, PCR, and cell labeling. As compared to genomic analysis, proteomics studies of bacterial cells are more challenging and difficult due to many variable physicochemical parameters, wide dynamic ranges, and relative protein abundances that differ among different cells. Proteomic tools can reveal information of bacterial surface exposed and cell envelope proteins [49] as well as bacterial secretome. Studies of cell surface proteins could reveal the interaction of cells with the environment and could predict an oxidative stress response. Changes in the secretome of various bacteria can show differences in secreted markers that are related to toxicity and protein alterations. While such studies are relevant for the investigation of bacterial pathogens and their interaction with the environment, few reports have been published on the use of proteomics to assess the relationship between oxidative stress and bacterial pathogenicity.

Western blotting, two-dimensional gel electrophoresis, gel imaging, and mass spectrometry were used to perform a full proteomic analysis of *Paracoccidioides* exposed to H_2O_2, used as an example of ROS, for the purpose of mimicking oxidative stress. Furthermore, intracellular NADPH/$NADP^+$ ratio were determined. One hundred and seventy-nine oxidative stress-responsive proteins/isoforms were identified and grouped. *Paracoccidioides* yeast response was characterized by up-regulated proteins/isoforms that represented a total of 64.8 % of all proteins/isoforms identified in this study [50]. Dosselli et al. used a proteomic approach to examine the oxidative response of *Staphylococcus aureus* undergoing photodynamic therapy (PDT), an antimicrobial method of killing bacteria by ROS. Several functional classes of proteins appeared to be selectively affected by PDT treatment. Moreover, cell growth and nutrition uptake were also inhibited by this treatment [51].

In another proteomic study, NAD-specific glutamate dehydrogenase, phosphoglycerate kinase, and Acyl-CoA dehydrogenase in *Fusobacterium nucleatum* were found to be up-regulated by oxidative stress after 72 h of atmospheric oxygen exposure and four additional exposure cycles [52]. Huang et al. showed that the treatment of *Helicobacter pylori* with 10 mM H_2O_2 induced overexpression of the following: cytotoxin-associated protein A (CagA), vacuolating cytotoxin (VacA), adherence-associated protein (AlpA), alkylhydroperoxide reductase (AhpC), catalase (KatA), serine protease (HtrA), aconitate hydratase, and fumarate reductase [53]. Combined results from 2D gel electrophoresis and MALDI-TOF MS analysis showed expression of 60 different proteins in *Bacillus anthracis* treated with 0.3 mM H_2O_2; 17 of these proteins are differently expressed over time. Time-dependent changes in generation of metabolic and repair/protection proteins were also studied [54]. Shu et al. reported a proteomic study of the oxidative stress response induced by low-dose H_2O_2 in *Bacillus anthracis* targeting activity and efficacy of Dps-like proteins (Dps1, Dps2, and Dps3) encoded by *Bacillus cereus*. Electrophoretic mobility shift assay (EMSA) and real-time reverse transcription-PCR were used to determine transcription level of the dps genes. The deletion of dps1 and dps2 caused a dramatic decrease in the survival rate. Since Dps1 and Dps2

were induced by oxidative stress, the authors concluded that bacteria developed certain defensive strategy when facing oxidative stress, while Dps3 only responded to general stress [55]. Sardar et al. investigated the changes in the proteome of *Leishmania donovani* promastigote during oxidative and nitrosative stresses using real-time PCR, MALDI-TOF/TOF mass spectrometry, Western blot, and iTRAQ labeling. There were nine proteins that were involved in redox homeostasis and were up-regulated while three others were down-regulated. For proteins involved in β-oxidation, TCA cycle, mitochondrial respiration, and oxidative phosphorylation, up-regulation was observed. The heat shock proteins (HSPs) and chaperone also were up-regulated. Antioxidant levels in *L. donovani* promastigotes decrease when the cells were treated with menadione, SNAP, and the combination of these two reagents. Possibly, antioxidants were consumed to maintain the normal physiological condition [56]. Through proteomic analysis, damage by ROS and/or RNS can be comprehensively investigated and fully revealed. Furthermore, a recent proteomic study indicates that pathogenic bacteria exhibit a complex response to ROS that includes the rapid adaption of metabolic pathways in response to oxidative-stress challenges [57]. A better understanding about the adaptation and the protection mechanisms of bacteria against oxidative stress are slowly being obtained using the above-mentioned methods.

23.5 Future Perspectives and Emerging Trends

This chapter reviewed the main mechanisms of antibiotic action, as well as mechanisms allowing bacteria to develop antibiotic resistance. We also introduced the role of oxidative stress in antibiotic-induced cell death as well as the how pathogenic bacteria have developed antibiotic resistance to oxidative stress. Various methods commonly used to assess oxidative stress were summarized along with their advantages and limitations, and their contribution to the study of antibiotic resistance. Since ROS are difficult to measure, the major challenge in this field remains the identification of a molecular connection between antibiotics and oxidative stress in bacteria [22]. Development of quantitative analytical methods to allow real-time quantitative measurements of ROS and antioxidant status could facilitate fundamental future investigations in this field. Improving the understanding of the intracellular communication and molecular mechanism used by bacteria is important in future research developments to be able to rationally design effective clinical interventions to respond to the growing threat of resistant bacterial infections [18].

The acquisition of multidrug resistance is a serious problem for the modern medicine. Resistance to antibiotics is facilitated by the presence of antibiotic-resistance genes on transferable genetic elements and also by the use of antibiotics in a way that allows them to act as selective agents. The use of antimicrobials for the promotion of animal growth and its link to the increased resistance has been a topic of heated debate [58]. As in humans, subtherapeutic doses in animals can select for resistant strains; if the bacteria cross from animal hosts to human hosts, then reservoirs of resistance may markedly reduce the effective lifetime of human antibiotics [5].

As a possible solution to this growing problem of resistance to conventional antibiotics, some authors suggested the use of antimicrobial peptides to partially substitute low effective antibiotics [59–61]. The interest in antimicrobial peptides began with the work of Fleming in 1922, with his discovery of antimicrobial activities in different secretions (saliva, nasal mucus, and tears…), blood, leukocytes, and lymphatic tissues called lysozyme [62]. Antimicrobial peptides are an abundant and diverse group of molecules. Their amino acid composition, amphipathicity, cationic charge, and size allow them to attach to and insert into membrane bilayers to form pores by "barrel-stave," "carpet," or "toroidal-pore" mechanisms [59]. One of the challenges to the use of these peptides as antimicrobial human therapy is their potential for toxicity. All clinical trials to date have used topical applications to address surface infections, rather than the more effective systemic administration (parenteral and oral). Another disadvantage of natural peptides is the potential sensitivity to proteases, creating potentially unfavorable pharmacokinetics. And, finally, the high cost of manufacturing peptides has limited both the testing and development of large numbers of variants and clinical targets to which these molecules can be applied [60].

Another emerging field expected to open new avenues in the fight against bacterial infection is nanotechnology. Nanotechnology is emerging as a new interdisciplinary field of chemistry, physics, and material science with broad applicability to biology and medicine [63]. A wide number of engineered nanoparticles (NPs) have shown excellent antibacterial activity on several Gram-positive and Gram-negative bacteria. The scientific debate concerning the mechanism of the antibacterial effect of NPs is still open. High surface area to volume ratios and unique chemico-physical properties of various nanomaterials are believed to contribute to the observed antimicrobial activities [64–73]. Among the different antimicrobial agents, silver has been the most extensively studied and used since ancient times to fight infections and prevent food spoilage. The antibacterial, antifungal, and antiviral properties of silver ions, silver compounds, and silver NPs have been extensively studied [67, 74–77]. In addition to silver, nitric oxide-releasing NPs (NO-NPs) are effective candidates in the inhibition of growth of many resistant bacteria (e.g., methicillin-resistant bacteria, Gram-negative bacteria resistant to commonly used antibiotics) [78]. The authors of this study suggested that as NO provides multiple mechanisms of bactericidal and immunological activity, the risk of pathogen resistance to NO-NPs is limited. Although many scientists are presenting nanoparticles (or "Nanoantibiotics") as a new promising paradigm for treating multiresistant bacteria [63, 79, 80], other works indicate that the development of NP resistance is also possible [81]. While NPs have demonstrated potential as effective antimicrobial agents against multidrug-resistant bacteria, future fundamental studies are needed to evaluate the specific toxicity mechanisms and assess the risks of a possible development of resistance, before these can be fully implemented in real world applications.

Acknowledgments Mouna Marrakchi is grateful to the Fulbright Foundation for the research fellowship as visiting professor in Clarkson University from September 2013 to June 2014. This material is based upon work supported by the National Science Foundation under Grant Nos. 0954919 and 1336493. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

References

1. Peeling RW, Smith PG, Bossuyt PMM (2006) A guide for diagnostic evaluations. Nat Rev Microbiol 4(12 Suppl):S2–S6
2. Fleming A (2001) On the antibacterial action of cultures of a penicillium, with special reference to their use in the isolation of B. influenzae. 1929. Bull World Health Organ 79(8):780–790
3. Dwyer DJ et al (2007) Gyrase inhibitors induce an oxidative damage cellular death pathway in Escherichia coli. Mol Syst Biol 3:91
4. Wright GD (2007) On the road to bacterial cell death. Cell 130(5):781–783
5. Walsh C (2000) Molecular mechanisms that confer antibacterial drug resistance. Nature 406(6797):775–781
6. Kohanski MA, Dwyer DJ, Collins JJ (2010) How antibiotics kill bacteria: from targets to networks. Nat Rev Microbiol 8(6):423–435
7. Champoux JJ (2001) DNA topoisomerases: structure, function, and mechanism. Annu Rev Biochem 70:369–413
8. Kohanski MA et al (2007) A common mechanism of cellular death induced by bactericidal antibiotics. Cell 130(5):797–810
9. Davies J (1994) Inactivation of antibiotics and the dissemination of resistance genes. Science 264(5157):375–382
10. Levy SB, Marshall B (2004) Antibacterial resistance worldwide: causes, challenges and responses. Nat Med 10(12):S122–S129
11. Jacoby GA, Archer GL (1991) New mechanisms of bacterial resistance to antimicrobial agents. N Engl J Med 324(9):601–612
12. Wright GD (2007) The antibiotic resistome: the nexus of chemical and genetic diversity. Nat Rev Microbiol 5(3):175–186
13. Dwyer DJ, Kohanski MA, Collins JJ (2009) Role of reactive oxygen species in antibiotic action and resistance. Curr Opin Microbiol 12(5):482–489
14. Ochman H, Lawrence JG, Groisman EA (2000) Lateral gene transfer and the nature of bacterial innovation. Nature 405(6784):299–304
15. Baquero F, Martinez JL, Canton R (2008) Antibiotics and antibiotic resistance in water environments. Curr Opin Biotechnol 19(3):260–265
16. Allen HK et al (2010) Call of the wild: antibiotic resistance genes in natural environments. Nat Rev Microbiol 8(4):251–259
17. Thomas CM, Nielsen KM (2005) Mechanisms of, and barriers to, horizontal gene transfer between bacteria. Nat Rev Microbiol 3(9):711–721
18. Lee HH et al (2010) Bacterial charity work leads to population-wide resistance. Nature 467(7311):82–85
19. D'Costa VM et al (2006) Sampling the antibiotic resistome. Science 311(5759):374–377
20. Hassett DJ, Cohen MS (1989) Bacterial adaptation to oxidative stress: implications for pathogenesis and interaction with phagocytic cells. FASEB J 3(14):2574–2582
21. Imlay JA (2013) The molecular mechanisms and physiological consequences of oxidative stress: lessons from a model bacterium. Nat Rev Microbiol 11(7):443–454
22. Yeom J, Imlay JA, Park W (2010) Iron homeostasis affects antibiotic-mediated cell death in pseudomonas species. J Biol Chem 285(29):22689–22695
23. Halliwell B (1994) Free radicals, antioxidants, and human disease: curiosity, cause, or consequence? Lancet 344(8924):721–724
24. Xia T et al (2006) Comparison of the abilities of ambient and manufactured nanoparticles to induce cellular toxicity according to an oxidative stress paradigm. Nano Lett 6(8):1794–1807
25. Kalyanaraman B et al (2012) Measuring reactive oxygen and nitrogen species with fluorescent probes: challenges and limitations. Free Radic Biol Med 52(1):1–6
26. Murrant CL, Reid MB (2001) Detection of reactive oxygen and reactive nitrogen species in skeletal muscle. Microsc Res Tech 55(4):236–248

27. Murphy MP et al (2011) Unraveling the biological roles of reactive oxygen species. Cell Metab 13(4):361–366
28. Villamena FA, Zweier JL (2004) Detection of reactive oxygen and nitrogen species by EPR spin trapping. Antioxid Redox Signal 6(3):619–629
29. White JR, Dearman HH (1965) Generation of free radicals from phenazine methosulfate, streptonigrin, and riboflavin in bacterial suspensions. Proc Natl Acad Sci U S A 54(3): 887–891
30. Borgmann S (2009) Electrochemical quantification of reactive oxygen and nitrogen: challenges and opportunities. Anal Bioanal Chem 394(1):95–105
31. Amatore C et al (2006) Monitoring in real time with a microelectrode the release of reactive oxygen and nitrogen species by a single macrophage stimulated by its membrane mechanical depolarization. Chembiochem 7(4):653–661
32. Amatore C et al (2008) Real-time amperometric analysis of reactive oxygen and nitrogen species released by single immunostimulated macrophages. Chembiochem 9(9):1472–1480
33. Amatore C, Arbault S, Koh AC (2010) Simultaneous detection of reactive oxygen and nitrogen species released by a single macrophage by triple potential-step chronoamperometry. Anal Chem 82(4):1411–1419
34. Amatore C et al (2008) Electrochemical monitoring of single cell secretion: vesicular exocytosis and oxidative stress. Chem Rev 108(7):2585–2621
35. Ganesana M, Erlichman JS, Andreescu S (2012) Real-time monitoring of superoxide accumulation and antioxidant activity in a brain slice model using an electrochemical cytochrome c biosensor. Free Radic Biol Med 53(12):2240–2249
36. Njagi J et al (2010) A sensitive electrochemical sensor based on chitosan and electropolymerized Meldola blue for monitoring NO in brain slices. Sens Actuators B Chem 143(2):673–680
37. Karasinski J et al (2005) Multiarray sensors with pattern recognition for the detection, classification, and differentiation of bacteria at subspecies and strain levels. Anal Chem 77(24): 7941–7949
38. Karasinski J et al (2007) Detection and identification of bacteria using antibiotic susceptibility and a multi-array electrochemical sensor with pattern recognition. Biosens Bioelectron 22(11):2643–2649
39. Albesa I et al (2004) Oxidative stress involved in the antibacterial action of different antibiotics. Biochem Biophys Res Commun 317(2):605–609
40. Grant SS et al (2012) Eradication of bacterial persisters with antibiotic-generated hydroxyl radicals. Proc Natl Acad Sci 109(30):12147–12152
41. Goswami M, Mangoli SH, Jawali N (2006) Involvement of reactive oxygen species in the action of ciprofloxacin against Escherichia coli. Antimicrob Agents Chemother 50(3):949–954
42. Wang X et al (2010) Contribution of reactive oxygen species to pathways of quinolone-mediated bacterial cell death. J Antimicrob Chemother 65(3):520–524
43. Mead J, Pryor W (1976) Free radicals in biology. Academic, New York, pp 51–68
44. Brot N et al (1981) Enzymatic reduction of protein-bound methionine sulfoxide. Proc Natl Acad Sci U S A 78(4):2155
45. Demple B, Linn S (1982) 5,6-Saturated thymine lesions in DNA: production by ultraviolet light or hydrogen peroxide. Nucleic Acids Res 10(12):3781–3789
46. Levin DE et al (1982) A new Salmonella tester strain (TA102) with A X T base pairs at the site of mutation detects oxidative mutagens. Proc Natl Acad Sci 79(23):7445–7449
47. Sokolowska I et al (2011) Mass spectrometry for proteomics-based investigation of oxidative stress and heat shock proteins. Oxidative Stress: Diagnostics, Prevention, and Therapy 1083: 369–411
48. James P (1997) Protein identification in the post-genome era: the rapid rise of proteomics. Q Rev Biophys 30(4):279–331
49. Solis N, Cordwell SJ (2011) Current methodologies for proteomics of bacterial surface-exposed and cell envelope proteins. Proteomics 11(15):3169–3189

50. de Arruda Grossklaus D et al (2013) Response to oxidative stress in Paracoccidioides yeast cells as determined by proteomic analysis. Microbes Infect 15(5):347–364
51. Dosselli R et al (2012) Molecular targets of antimicrobial photodynamic therapy identified by a proteomic approach. J Proteomics 77:329–343
52. Silva VL et al (2010) Use of 2-D electrophoresis and ESI mass spectrometry techniques to characterize *Fusobacterium nucleatum* proteins up-regulated after oxidative stress. Anaerobe 16(2):179–182
53. Huang C-H, Chiou S-H (2011) Proteomic analysis of upregulated proteins in Helicobacter pylori under oxidative stress induced by hydrogen peroxide. Kaohsiung J Med Sci 27(12): 544–553
54. Kim SH et al (2013) Proteomic analysis of the oxidative stress response induced by low-dose hydrogen peroxide in Bacillus anthracis. J Microbiol Biotechnol 23(6):750–758
55. Shu J-C et al (2013) Differential regulation and activity against oxidative stress of Dps proteins in Bacillus cereus. Int J Med Microbiol 303(8):662–673
56. Sardar AH et al (2013) Proteome changes associated with Leishmania donovani promastigote adaptation to oxidative and nitrosative stresses. J Proteomics 81:185–199
57. Deng X et al (2013) Proteome-wide quantification and characterization of oxidation-sensitive cysteines in pathogenic bacteria. Cell Host Microbe 13(3):358–370
58. Witte W (1998) Medical consequences of antibiotic use in agriculture. Science 279(5353):996–997
59. Brogden KA (2005) Antimicrobial peptides: pore formers or metabolic inhibitors in bacteria? Nat Rev Microbiol 3(3):238–250
60. Hancock REW, Sahl HG (2006) Antimicrobial and host-defense peptides as new anti-infective therapeutic strategies. Nat Biotechnol 24(12):1551–1557
61. Zasloff M (2002) Antimicrobial peptides of multicellular organisms. Nature 415(6870):389–395
62. Fleming A (1922) On a remarkable bacteriolytic element found in tissues and secretions. Proc R Soc Lond B 93:306–317
63. Moritz M, Geszke-Moritz M (2013) The newest achievements in synthesis, immobilization and practical applications of antibacterial nanoparticles. Chem Eng J 228:596–613
64. Amato E et al (2011) Synthesis, characterization and antibacterial activity against Gram positive and Gram negative bacteria of biomimetically coated silver nanoparticles. Langmuir 27(15):9165–9173
65. Baek YW, An YJ (2011) Microbial toxicity of metal oxide nanoparticles (CuO, NiO, ZnO, and Sb2O3) to Escherichia coli, Bacillus subtilis, and Streptococcus aureus. Sci Total Environ 409(8):1603–1608
66. Bandyopadhyay S et al (2012) Comparative toxicity assessment of CeO2 and ZnO nanoparticles towards Sinorhizobium meliloti, a symbiotic alfalfa associated bacterium: use of advanced microscopic and spectroscopic techniques. J Hazard Mater 241–242:379–386
67. El Badawy AM et al (2011) Surface charge-dependent toxicity of silver nanoparticles. Environ Sci Technol 45(1):283–287
68. Heinlaan M et al (2008) Toxicity of nanosized and bulk ZnO, CuO and TiO2 to bacteria Vibrio fischeri and crustaceans Daphnia magna and Thamnocephalus platyurus. Chemosphere 71(7):1308–1316
69. Kumar A et al (2011) Cellular uptake and mutagenic potential of metal oxide nanoparticles in bacterial cells. Chemosphere 83(8):1124–1132
70. Li M, Zhu L, Lin D (2011) Toxicity of ZnO nanoparticles to Escherichia coli: mechanism and the influence of medium components. Environ Sci Technol 45(5):1977–1983
71. Raghupathi KR, Koodali RT, Manna AC (2011) Size-dependent bacterial growth inhibition and mechanism of antibacterial activity of zinc oxide nanoparticles. Langmuir 27(7): 4020–4028
72. Thill A et al (2006) Cytotoxicity of CeO2 nanoparticles for Escherichia coli. Physico-chemical insight of the cytotoxicity mechanism. Environ Sci Technol 40(19):6151–6156

73. Tong TZ et al (2013) Cytotoxicity of commercial nano-TiO2 to Escherichia coli assessed by high-throughput screening: effects of environmental factors. Water Res 47(7):2352–2362
74. Kvitek L et al (2008) Effect of surfactants and polymers on stability and antibacterial activity of silver nanoparticles (NPs). J Phys Chem C 112(15):5825–5834
75. Mohanty S et al (2012) An investigation on the antibacterial, cytotoxic, and antibiofilm efficacy of starch-stabilized silver nanoparticles. Nanomedicine 8(6):916–924
76. Pal S, Tak YK, Song JM (2007) Does the antibacterial activity of silver nanoparticles depend on the shape of the nanoparticle? A study of the gram-negative bacterium Escherichia coli. Appl Environ Microbiol 73(6):1712–1720
77. Rai M, Yadav A, Gade A (2009) Silver nanoparticles as a new generation of antimicrobials. Biotechnol Adv 27(1):76–83
78. Friedman A et al (2011) Susceptibility of Gram-positive and -negative bacteria to novel nitric oxide-releasing nanoparticle technology. Virulence 2(3):217–221
79. Huh AJ, Kwon YJ (2011) "Nanoantibiotics": a new paradigm for treating infectious diseases using nanomaterials in the antibiotics resistant era. J Control Release 156(2):128–145
80. Pelgrift RY, Friedman AJ (2013) Nanotechnology as a therapeutic tool to combat microbial resistance. Adv Drug Deliv Rev 65(13–14):1803–1815
81. Aruguete DM et al (2013) Antimicrobial nanotechnology: its potential for the effective management of microbial drug resistance and implications for research needs in microbial nanotoxicology. Environ Sci Process Impacts 15(1):93–102

Chapter 24
Proteomic Approaches to Dissect Neuronal Signaling Pathways

Heather L. Bowling and Katrin Deinhardt

Abstract With an increasing awareness of mental health issues and neurological disorders, "understanding the brain" is one of the biggest current challenges in biological research. This has been recognized by both governments and funding agencies, and includes the need to understand connectivity of brain regions and coordinated network activity, as well as cellular and molecular mechanisms at play. In this chapter, we will describe how we have taken advantage of different proteomic techniques to unravel molecular mechanisms underlying two modulators of neuronal function: Neurotrophins and antipsychotics.

24.1 Introduction

Over the past years, proteomic studies have significantly advanced our understanding of the mechanisms of brain function. They have highlighted differences between normal and diseased brains (e.g. [1, 10]), identified the molecular composition of synapses [39], and in a landmark study, identified the entire and stoichiometric protein composition of a synaptic vesicle as the first ever fully characterized organelle [35]. It is well beyond the scope of this chapter to describe all the advances in neuroscience

H.L. Bowling
Skirball Institute of Biomolecular Medicine, New York University School of Medicine, New York, NY 10016, USA
e-mail: Heather.Bowling@med.nyu.edu

K. Deinhardt (✉)
Institute for Life Sciences and Centre for Biological Sciences, University of Southampton, Life Sciences Building 85, Southampton, SO17 1BJ, UK
e-mail: K.Deinhardt@soton.ac.uk

A.G. Woods and C.C. Darie (eds.), *Advancements of Mass Spectrometry in Biomedical Research*, Advances in Experimental Medicine and Biology 806,
DOI 10.1007/978-3-319-06068-2_24, © Springer International Publishing Switzerland 2014

that were aided by proteomic approaches. Instead, here we focus on our recent studies that took advantage of proteomics to unravel molecular mechanisms of action, two distinct modulators of neuronal function: Neurotrophins and antipsychotics.

Neurotrophins are a family of growth factors comprising the first ever described growth factor, nerve growth factor (NGF) [8], as well as brain-derived neurotrophic factor (BDNF), and neurotrophins (NT) 3 and 4 [27]. Neurotrophins are best characterized for their role in mediating neuronal outgrowth, arborization, and survival, as well as synaptic plasticity. They signal through their cognate Trk receptors, the receptor tyrosine kinases (RTKs) TrkA, B, and C, and all neurotrophins also bind to and signal through the tumor necrosis family-type receptor, p75 pan-neurotrophin receptor [12, 13, 27, 37]. Similar to other secreted molecules, neurotrophins are initially synthesized as precursors, the proneurotrophins, which are then processed either intra- or extracellularly to generate the mature ligand [7]. However, proneurotrophins are not just precursors to the mature ligand, but can also act as active signaling molecules. They bind with high affinity to a complex of p75NTR and a sortilin family member, and interestingly, display largely opposing effects to their mature counterparts, inducing apoptosis and long-term depression of synapses and discouraging outgrowth [13, 22, 27, 36]. While the importance of neurotrophins for the formation, maintenance, and plasticity of the nervous system is undisputed, the intracellular mechanisms mediating their actions are not fully characterized. Below, we will describe our recent approaches to broaden our current knowledge of the intracellular pathways triggered by neurotrophins, and speculate how future approaches may further our understanding.

Given the significance of growth factors for the functioning of the central nervous system, it is not surprising, a large effort is put into finding small molecule mimetics to for example substitute for loss of BDNF in neurodegenerative conditions, and that drugs commonly used to treat mental health disorders interact with neurotrophic pathways. For example, the serotonin reuptake inhibitor fluoxetine, which is commonly used to treat major depression, also raises BDNF levels, and this BDNF increase is essential for the beneficial effects of the drug [2]. In this chapter we will discuss how proteomics have advanced our understanding of the intracellular mechanisms of neurotrophin action as well as of the trophic actions of antipsychotics.

24.2 Immediate Activation of Signaling Cascades: Isolating Tyrosine Phosphorylated Substrates

Neurotrophins signal through their cognate RTKs, and accordingly, many of the canonical pathways downstream of all RTKs, such as MEK-Erk1/2, PI3 kinase-Akt, and PLCγ signaling, are also activated by neurotrophins [12, 27, 37]. To broaden our understanding of neurotrophin signaling in primary neurons, we took an unbiased approach by isolating tyrosine-phosphorylated proteins following BDNF and NT-3 treatment, followed by mass spectrometry analysis. While many meaningful

studies have been completed through the addition of unlabelled samples to the mass spectrometer and their comparison following separate mass spectrometry analyses, the problem of natural technical variability persists. To address this technical variability that stems from comparing two groups in individual mass spectrometry analyses, [26] created stable isotope labeling of amino acids in culture (SILAC). In this method, cells in culture are fed amino acids (commonly Arginine and Lysine, as trypsin cleaves the C-terminal to these amino acids) that have different carbon or nitrogen weights (e.g. they have Carbon-13 instead of Carbon-12) that will in turn be incorporated into nascent proteins. Typically, cells are grown for ten divisions in SILAC media before being subjected to differential treatment and mass spectrometry. In order to adapt this technology to post-mitotic neurons, we analyzed the labeling efficiency over time and found that neurons grown in SILAC media for 10 days from the time of plating show >80 % labeling of their proteome [34]. We then used the SILAC approach to identify proteins that are tyrosine phosphorylated downstream of BDNF in primary cortical neurons. To this end, treated and untreated cells were lysed, lysates were combined and tyrosine phosphorylated proteins and their interactors were immunoprecipitated. Since the difference in weight will correspond to a specific higher or lower weight when measured by the mass spectrometer, the two treatment groups can be separated by the mass spectrometer and assigned to their original treatment groups, limiting experiment-by-experiment technical variability, including which proteins are measured in each group. This label (heavy or light) is reversed in subsequent experiments to avoid label bias in peptide detection or integration [34, 41]. Using this technique, we found a number of proteins that had not been identified as activated downstream of BDNF-TrkB signaling before, including a number of cytoskeleton-associated proteins. In addition, we identified and characterized the ESCRT machinery-associated protein, Hrs, as a novel regulator of TrkB turnover in cortical neurons [34].

While we observed significant labeling of the proteome after 10 days in culture, individual proteins of low turnover rate will show less label incorporation and thus may be assigned to the wrong experimental group. To avoid misinterpretation, label incorporation rates for individual proteins were measured and results were adjusted accordingly [34]. This labor-intensive correction step can be circumvented by using multiplex SILAC, i.e. labeling all treatment groups, one with “medium” and one with “heavy” amino acids and thus excluding “light” peptides from the analysis. Using this modified approach, we investigated which proteins are tyrosine phosphorylated downstream of NT-3 treatment in cortical neurons [40]. We expected to see a large overlap with the BDNF experiment since both neurotrophins activate canonical RTK signaling pathways. Interestingly, despite activating both TrkC and also TrkB robustly, many of the non-canonical downstream targets of NT-3 were quite different from those identified in response to BDNF. In fact, many of the candidates coming out of this screen were associated with synaptic vesicles and exocytosis [40]. This striking difference may be due to differential receptor engagement or receptor–ligand complex stability. It is well characterized that both signal localization and duration may critically influence the cellular outcome [14, 19, 25], which may in part underlie these observed differences.

24.3 Immediate Activation of Signaling Pathways: Characterizing Protein–Protein Interactions

In addition to phosphorylation events, signals are transduced through recruitment of adaptor proteins into effector complexes and formation of protein complexes. This is not only important downstream of RTK signaling, but also allows to dissect signaling events downstream of receptors lacking a kinase domain, such as p75NTR. One approach to isolate protein complexes is through affinity purification of the protein of interest and then submitting them to mass spec analysis, ideally from SILAC labelled cells [29]. These approaches have not only considerably broadened our understanding of cellular processes, but more recently have also highlighted how the cellular context may critically determine the composition of interaction networks [33]. We have used classical co-immunoprecipitation followed by mass spectrometry to identify a novel interactor of p75NTR, Trio. We further showed that Trio dissociates from p75NTR in response to proneurotrophin treatment and thus leads to an acute growth cone collapse response in primary neurons, discouraging neuronal growth [15]. This identified an intracellular pathway of proneurotrophin action on neuronal morphology, which opposes the well-established function of mature neurotrophins in promoting outgrowth.

Another unbiased approach to study protein complexes forming in response to a stimulus is through separating intact complexes from a lysate by blue native polyacrylamide gel electrophoresis, followed either directly by mass spectrometry of the entire complexes, or by a separation of the complexes into their subunits by SDS-PAGE. The resulting 2D gel may then be analyzed by Coomassie or silver staining to look for shifts between conditions, by mass spectrometry to identify components or by Western blot to probe for the presence of candidates. Using this technique, we isolated protein complexes containing tyrosine-phosphorylated proteins in response to ephrinB1 stimulation in NG108 cells. We detected many established and a number of new interactions, both leading to assembly or disassembly of complexes in response to ephrinB1 and implicating focal adhesion kinase and WAVE complexes in ephrinB1 signaling [11]. It will be of interest to use these approaches to further characterize differential p75NTR signaling in response to mature proneurotrophins, as well as to characterized to protein interaction networks downstream of Trk signaling in primary neurons.

24.4 Watching the Cell Change Its Proteome: Monitoring the Activation of Translation in Response to a Signal

Activation of signaling pathways often leads to translational activation and adjustment of the cellular makeup. Proteome changes have been identified by mass spectrometry following BDNF treatment of hippocampal neurons and radiolabeling of newly synthesized proteins [24], and in cortical synaptoneurosomes using MudPIT

(multidimensional protein identification technology) and relative quantification of all protein spectra [21]. Interestingly, both screens revealed that BDNF stimulation leads to upregulation of the translational machinery itself, thus increasing capability of the cells to synthesize new proteins. In addition, upregulation of a large variety of proteins including cytoskeletal and synaptic components were detected [21, 24].

Another way of measuring newly synthesized proteins is by labeling them with SILAC media. We have taken advantage of this approach to analyze proteomic changes in response to antipsychotics [4]. Previous studies have shown that the kinase Akt can be acutely activated after antipsychotic treatment [3], and that chronic antipsychotic treatment can lead to alterations in the proteomic content of neurons [6, 23]. We demonstrated that the antipsychotic haloperidol activated protein synthesis within 30 min of treatment in neurons in culture via the Akt—mammalian target of rapamycin complex 1 (mTORC1) signaling cascade that is known to influence both translation initiation and elongation in neurons [4, 31].

To measure which proteins were being synthesized in response to the antipsychotic haloperidol, we performed proteomic analysis at an early stage (5 h) and later stage (48 h) of treatment. Striatal neurons were grown in culture for 7 days. They were then treated with either an antipsychotic or vehicle for 48 h in the presence of SILAC Multiplex media (Medium: [$^{13}C_6$] Arginine [4,4,5,5-D_4] Lysine, Heavy: [$^{13}C_6$; $^{15}N_4$] Arginine [$^{13}C_6$; $^{15}N_2$] Lysine). This way we were able to detect nascent proteins that were synthesized and not degraded during the 48 h treatment period that were relevant to antipsychotic action. The advantage of this method is the ability to analyze nascent proteins in the mass spectrometry analysis. A potential caveat to this approach is that while it is very useful, SILAC label incorporation into primary neurons occurs at approximately 1 % per hour on average (based on [40]). This makes it difficult or impossible to quantify changes in total protein amounts by mass spectrometry after only a few hours of treatment without enrichment for newly synthesized proteins.

However, it is of keen interest to the translation community to study short timepoints to measure acute responses to stimuli. In order to achieve this, several new techniques were introduced. One such technique was developed by Schmidt et al. [32]—surface sensing of translation (SUnSET) (Fig. 24.1). In SUnSET, cells are fed a sublethal dose of the antibiotic puromycin, then lysed and analyzed with a monoclonal anti-puromycin antibody by Western blot. This allows identification of nascent proteins, as puromycin is structurally similar to tRNA, and the ribosome will stochastically incorporate it in place of a tRNA onto nascent peptide chains, forming a covalent bond. Puromycin differs from a tRNA in an oxygen-to-nitrogen substitution, which prevents the ribosome from hydrolyzing it as it would a tRNA and this causes premature truncation and release of the peptide from the ribosome. We combined SUnSET with Multiplex SILAC to allow for the labeling, isolation, and quantitation of nascent proteins following the short labeling period (5 h). Using this approach, we were able to enrich for bonafide proteins regulated by antipsychotic treatment in a relatively early time period of 5 h. SUnSEt allowed for the isolation of nascent proteins via an anti-puromycin immunoprecipitation, and SILAC enabled us to combine the lysates from the two treatment groups before

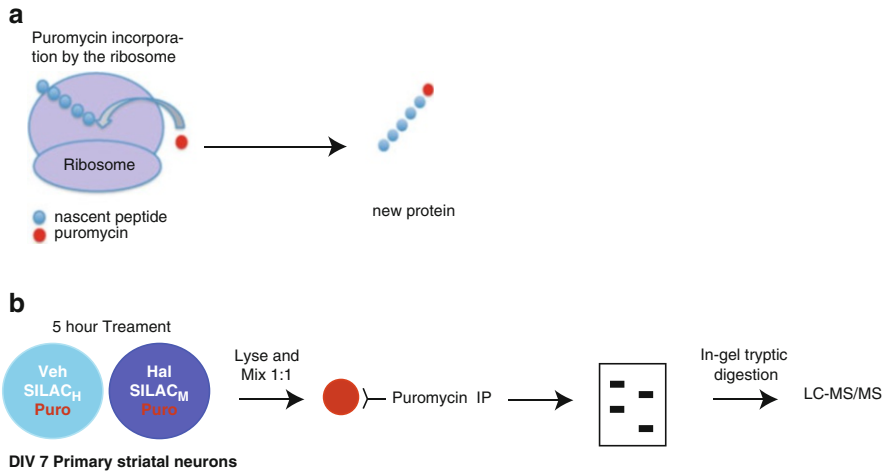


Fig. 24.1 Schematic of puromycin labelling and SUnSET/multiplex SILAC experiments. (**a**) Puromycin mimics a tRNA and is stochastically incorporated by the ribosome onto nascent peptide chains, causing release from the ribosome, and allowing for the labelling of nascent peptides. (**b**) Schematic of SUnSET/multiplex SILAC experiment. DIV7 primary striatal neurons were labelled for 5 h with puromycin and multiplex SILAC media upon treatment with vehicle or the antipsychotic, haloperidol. Following treatment, the neurons were lysed, mixed at a 1:1 ratio, and puromycin labelled proteins were immunoprecipitated, run on an acrylamide gel, isolated using in-gel tryptic digestion, and subjected to LC-MS/MS

immunoprecipitation and mass spectrometry analysis to reduce technical variability. With the enrichment of the nascent proteins with the puromycin immunoprecipitation, proteins could be measured and quantified by mass spectrometry, which was not possible without the enrichment. Using this method, we were able to identify several hundred proteins that were unique to antipsychotic treatment in the first 5 h that provided insight into the biological mechanisms of early antipsychotic response, specifically an upregulation of ribosomal proteins, including ribosomal protein S6 (Rps6). An increase in these proteins suggests "priming" for later stage changes in cellular protein content, such as the increase in specific cytoskeletal proteins identified after 48 h of treatment that were not detected at 5 h, which included ankyrin repeat-rich membrane-spanning protein (ARMS/Kidins220). Interestingly, there were several different types of proteins identified from ribosomal proteins, to mitochondrial proteins and cytoskeletal proteins, and nuclear proteins, suggesting that this technique is not limited to labeling specific subsets of only one type of protein. These data, especially the detection of ARMS/Kidins220, allowed us to predict an acute increase in morphological complexity, which was dependent on mTORC1 signaling effectors associated with protein synthesis [4].

Despite the fact that this method was helpful in identifying proteins synthesized in a short time period, there are drawbacks to its use. Though the cells are presented with a sublethal dose of puromycin, puromycin itself interrupts the normal function of the ribosome by causing premature truncation and release of nascent peptides. Over time, this disrupts cellular function—both by truncating important proteins

and by ultimately clogging the ribosome with puromycin instead of tRNAs. In our experiments, we analyzed striatal neurons treated with puromycin over time. While we measured SILAC incorporation at 5 h of treatment, we found that SILAC incorporation was markedly reduced in the presence of puromycin at 8, 12, and 24 h (unpublished observation).

An alternative approach to enrich newly synthesized proteins is BONCAT (biorthogonal non-canonical amino acid tagging), which is achieved by feeding cells a non-toxic amino acid analog (such as AHA or L-Azidohomoalanine) containing a heavily charged azide group or an alkyne group. Cells are amino acid starved for 30 min to encourage them to take up the tagged analog and other amino acids from the surrounding media. The tag is an amino acid analog, so it mimics an amino acid (methionine in the case of AHA) and is naturally incorporated by the ribosome onto newly synthesized proteins [17]. After labeling for the desired period of time, cells are lysed. Using an alkyne linked to biotin, a bead or fluorophore, a copper-catalyzed Huisgen-azide alkyne reaction is performed to form a covalent bond between the tag (biotin or fluorophore) and the nascent proteins [4, 17, #342], [38, #343]. Using Western blot or immunohistochemistry, general protein synthesis may be visualized. To target specific proteins, biotin can be removed using a filtration process and using the specific streptavidin-biotin interaction, nascent proteins can be precipitated from the lysate of the cell and examined by Western blot. Since this technique involves separating only the nascent proteins, it is possible to examine shorter timepoints to examine total proteomic changes, thus allowing to measure acute treatment responses. However, because there is no difference in label between treatment groups, separate mass spectrometry analyses must be performed and analyzed separately. Thus, while the new population can be analyzed under different conditions, the problem of technical variability persists.

To solve this problem, SILAC and BONCAT were combined to create QuaNCAT (quantitative noncanonical amino acid tagging). In QuaNCAT, cells are treated with the methionine analog AHA and SILAC media, and the BONCAT method is used to isolate nascent proteins while SILAC is used to measure the proteins by mass spectrometry with less variability so direct quantitative comparisons can be made. This method was pioneered by Howden et al. [20]. In their study, they plated freshly isolated human CD4+ T cells. These cells were starved of arginine, methionine, and lysine for 1 h then stimulated them for 2 h in methionine, arginine, and lysine-free media supplemented with AHA and Medium: [$^{13}C_6$] Arginine; [$^{2}H_4$] Lysine or Heavy: [$^{13}C_6$,$^{15}N_4$] Arginine; [$^{13}C_6$,$^{15}N_2$]-Lysine. Lysates from stimulated and unstimulated cells were combined and AHA-tagged proteins underwent cycloaddition via the Huisgen azide-Alkyne cycloaddition reaction, and were isolated by NeutrAvidin pull-down, digested with trypsin, and analyzed by LC-MS/MS. Since AHA has been shown to be non-toxic to cells [16], one could use this technique in place of the SUnSET/SILAC combination to label proteins made in a specific short timeframe and isolate them for quantitative analysis. This technique would allow for longer windows of time while analyzing only nascent proteins that respond to a given stimulus such as drug treatment. One possible application of this technique is in further drug studies, where one could examine changes in the 12–14 h treatment range that are still challenging with SILAC alone, but beyond the scope of the SUnSET technique.

24.5 Special Challenges Posed by Non-dividing Cells: Post-mitotic Systems and Data Analysis

Though there are many new and exciting methods to use in the field of neuroscience for examining proteomic changes, one problem that continues to plague the field is analyzing subtler proteomic changes. Cells undergoing mitosis, such as commonly used cell lines, have an overall higher rate of translation compared to quiescent cells (reviewed in [18, 30]). The need to identify significant smaller scale shifts is acute in quiescent and post mitotic systems such as neurons where few proteins have been reported to change above twofold or 200 % with various treatments [4, #336], [5], [34, #9]; however, changes as low as 20 % can be statistically significant in neurons and validated by Western blotting [5]. This begs the question of whether even smaller changes that are biologically relevant may be significant. Unfortunately, this point remains controversial in the field as the reliability and sensitivity of measuring proteomic changes can be variable, thus, any protein of interest in this range must be independently validated using a candidate approach. While there is no standard method of analyzing data in the proteomics field, many researchers agree that proteomic changes should be presented in ratio form (e.g. treated vs. untreated, transgenic vs. wildtype) and normalized to a peptide that falls in the middle of the calculated range of peptides in the dataset [9]. An alternative method suggested by the Yates lab aims to further refine this ratio ("fold change") calculation. Using an automated program called Census that they created, they remove statistical outliers from datasets by identifying and removing poor quality peptide measurements [5, 28]. They then use QuantCompare to generate protein ratios for peptides that meet the criterion and identify statistically significant changes in individual proteins. This new approach allows for a greater refinement of signal by purging outliers, which in turn facilitates the identification of significant shifts between treatment groups that are smaller. It will be exciting to see how this new analytical approach will aid to further advance our understanding of the molecular mechanisms underlying brain function.

Acknowledgment This work was supported by the Human Frontiers Science Program and the FNES start-up award (KD).

References

1. Andreev VP, Petyuk VA, Brewer HM, Karpievitch YV, Xie F, Clarke J, Camp D, Smith RD, Lieberman AP, Albin RL, Nawaz Z, El Hokayem J, Myers AJ (2012) Label-free quantitative LC-MS proteomics of Alzheimer's disease and normally aged human brains. J Proteome Res 11(6):3053–3067
2. Bath KG, Jing DQ, Dincheva I, Neeb CC, Pattwell SS, Chao MV, Lee FS, Ninan I (2012) BDNF Val66Met impairs fluoxetine-induced enhancement of adult hippocampus plasticity. Neuropsychopharmacology 37:1297–1304

3. Beaulieu JM, Gainetdinov RR, Caron MG (2007) The Akt-GSK-3 signaling cascade in the actions of dopamine. Trends Pharmacol Sci 28:166–172
4. Bowling H, Zhang G, Bhattacharya A, Pérez-Cuesta LM, Deinhardt K, Hoeffer CA, Neubert TA, Gan WB, Klann E, Chao MV (2014) Antipsychotics activate mTORC1-dependent translation to enhance neuronal morphological complexity. Sci Signal 7(308):ra4
5. Butko MT, Savas JN, Friedman B, Delahunty C, Ebner F, Yates JR III, Tsien RY (2013) In vivo quantitative proteomics of somatosensory cortical synapses shows which protein levels are modulated by sensory deprivation. Proc Natl Acad Sci U S A 110:E726–E735
6. Chan MK, Tsang TM, Harris LW, Guest PC, Holmes E, Bahn S (2011) Evidence for disease and antipsychotic medication effects in post-mortem brain from schizophrenia patients. Mol Psychiatry 16:1189–1202
7. Chao MV, Bothwell M (2002) Neurotrophins: to cleave or not to cleave. Neuron 33:9–12
8. Cohen S, Levi-Montalcini R (1957) Purification and properties of a nerve growth-promoting factor isolated from mouse sarcoma 180. Cancer Res 17:15–20
9. Cox J, Mann M (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat Biotechnol 26:1367–1372
10. Culver BP, Savas JN, Park SK, Choi JH, Zheng S, Zeitlin SO, Yates JR III, Tanese N (2012) Proteomic analysis of wild-type and mutant huntingtin-associated proteins in mouse brains identifies unique interactions and involvement in protein synthesis. J Biol Chem 287: 21599–21614
11. Darie CC, Deinhardt K, Zhang G, Cardasis HS, Chao MV, Neubert TA (2011) Identifying transient protein-protein interactions in EphB2 signaling by blue native PAGE and mass spectrometry. Proteomics 11:4514–4528
12. Deinhardt K, Chao MV (2009) Neurotrophin signaling in development. Handbook of cell signaling. Academic press, Oxford, pp 1913–1918
13. Deinhardt K, Chao MV (2014) Shaping neurons: long and short range effects of mature and proBDNF signalling upon neuronal structure. Neuropharmacology 76:603–609
14. Deinhardt K, Jeanneteau F (2012) More than just an off-switch: the essential role of protein dephosphorylation in the modulation of BDNF signaling events. In: Huang C (ed) Protein phosphorylation in human health. Rijeka, InTech, pp 217–232
15. Deinhardt K, Kim T, Spellman DS, Mains RE, Eipper BA, Neubert TA, Chao MV, Hempstead BL (2011) Neuronal growth cone retraction relies on proneurotrophin receptor signaling through Rac. Sci Signal 4:ra82
16. Dieterich DC, Hodas JJ, Gouzer G, Shadrin IY, Ngo JT, Triller A, Tirrell DA, Schuman EM (2010) In situ visualization and dynamics of newly synthesized proteins in rat hippocampal neurons. Nat Neurosci 13:897–905
17. Dieterich DC, Link AJ, Graumann J, Tirrell DA, Schuman EM (2006) Selective identification of newly synthesized proteins in mammalian cells using bioorthogonal noncanonical amino acid tagging (BONCAT). Proc Natl Acad Sci U S A 103:9482–9487
18. Fuge EK, Braun EL, Werner-Washburne M (1994) Protein synthesis in long-term stationary-phase cultures of Saccharomyces cerevisiae. J Bacteriol 176:5802–5813
19. Harrington AW, St Hillaire C, Zweifel LS, Glebova NO, Philippidou P, Halegoua S, Ginty DD (2011) Recruitment of actin modifiers to TrkA endosomes governs retrograde NGF signaling and survival. Cell 146:421–434
20. Howden AJ, Geoghegan V, Katsch K, Efstathiou G, Bhushan B, Boutureira O, Thomas B, Trudgian DC, Kessler BM, Dieterich DC, Davis BG, Acuto O (2013) QuaNCAT: quantitating proteome dynamics in primary cells. Nat Methods 10(4):343–346
21. Liao L, Pilotte J, Xu T, Wong CC, Edelman GM, Vanderklish P, Yates JR III (2007) BDNF induces widespread changes in synaptic protein content and up-regulates components of the translation machinery: an analysis using high-throughput proteomics. J Proteome Res 6:1059–1071
22. Lu B, Pang PT, Woo NH (2005) The yin and yang of neurotrophin action. Nat Rev Neurosci 6:603–614

23. Ma D, Chan MK, Lockstone HE, Pietsch SR, Jones DN, Cilia J, Hill MD, Robbins MJ, Benzel IM, Umrania Y, Guest PC, Levin Y, Maycox PR, Bahn S (2009) Antipsychotic treatment alters protein expression associated with presynaptic function and nervous system development in rat frontal cortex. J Proteome Res 8:3284–3297
24. Manadas B, Santos AR, Szabadfi K, Gomes JR, Garbis SD, Fountoulakis M, Duarte CB (2009) BDNF-induced changes in the expression of the translation machinery in hippocampal neurons: protein levels and dendritic mRNA. J Proteome Res 8:4536–4552
25. Marshall CJ (1995) Specificity of receptor tyrosine kinase signaling: transient versus sustained extracellular signal-regulated kinase activation. Cell 80:179–185
26. Ong SE, Blagoev B, Kratchmarova I, Kristensen DB, Steen H, Pandey A, Mann M (2002) Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. Mole Cell Proteomics 1(5):376–386
27. Park H, Poo MM (2013) Neurotrophin regulation of neural circuit development and function. Nat Rev Neurosci 14:7–23
28. Park SK, Venable JD, Xu T, Yates JR III (2008) A quantitative analysis software tool for mass spectrometry-based proteomics. Nat Methods 5:319–322
29. Piechura H, Oeljeklaus S, Warscheid B (2012) SILAC for the study of mammalian cell lines and yeast protein complexes. Methods Mol Biol 893:201–221
30. Pyronnet S, Sonenberg N (2001) Cell-cycle-dependent translational control. Curr Opin Genet Dev 11:13–18
31. Santini E, Klann E (2011) Dysregulated mTORC1-dependent translational control: from brain disorders to psychoactive drugs. Front Behav Neurosci 5:76
32. Schmidt EK, Clavarino G, Ceppi M, Pierre P (2009) SUnSET, a nonradioactive method to monitor protein synthesis. Nat Methods 6:275–277
33. Song J, Wang Z, Ewing RM (2013) Integrated analysis of the Wnt responsive proteome in human cells reveals diverse and cell-type specific networks. Mol Biosyst 10:45–53
34. Spellman DS, Deinhardt K, Darie CC, Chao MV, Neubert TA (2008) Stable isotopic labeling by amino acids in cultured primary neurons: application to brain-derived neurotrophic factor-dependent phosphotyrosine-associated signaling. Mol Cell Proteomics 7:1067–1076
35. Takamori S, Holt M, Stenius K, Lemke EA, Gronborg M, Riedel D, Urlaub H, Schenck S, Brugger B, Ringler P, Muller SA, Rammner B, Grater F, Hub JS, De Groot BL, Mieskes G, Moriyama Y, Klingauf J, Grubmuller H, Heuser J, Wieland F, Jahn R (2006) Molecular anatomy of a trafficking organelle. Cell 127:831–846
36. Teng KK, Felice S, Kim T, Hempstead BL (2010) Understanding proneurotrophin actions: recent advances and challenges. Dev Neurobiol 70:350–359
37. Teng KK, Hempstead BL (2004) Neurotrophins and their receptors: signaling trios in complex biological systems. Cell Mol Life Sci 61:35–48
38. Wang Q, Chan TR, Hilgraf R, Fokin VV, Sharpless KB, Finn MG (2003) Bioconjugation by copper(I)-catalyzed azide-alkyne [3 + 2] cycloaddition. J Am Chem Soc 125:3192–3193
39. Witzmann FA, Arnold RJ, Bai F, Hrncirova P, Kimpel MW, Mechref YS, McBride WJ, Novotny MV, Pedrick NM, Ringham HN, Simon JR (2005) A proteomic survey of rat cerebral cortical synaptosomes. Proteomics 5:2177–2201
40. Zhang G, Deinhardt K, Chao MV, Neubert TA (2011) Study of neurotrophin-3 signaling in primary cultured neurons using multiplex stable isotope labeling with amino acids in cell culture. J Proteome Res 10:2546–2554
41. Zhang G, Neubert TA (2009) Use of stable isotope labeling by amino acids in cell culture (SILAC) for phosphotyrosine protein identification and quantitation. Methods Mol Biol 527:79–92, xi

Chapter 25
Investigating a Novel Protein Using Mass Spectrometry: The Example of Tumor Differentiation Factor (TDF)

Alisa G. Woods, Izabela Sokolowska, Katrin Deinhardt, and Costel C. Darie

Abstract Better understanding of central nervous system (CNS) molecules can include the identification of new molecules and their receptor systems. Discovery of novel proteins and elucidation of receptor targets can be accomplished using mass spectrometry (MS). We describe a case study of such a molecule, which our lab has studied using MS in combination with other protein identification techniques, such as immunohistochemistry (IHC) and Western blotting. This molecule is known as tumor differentiation factor (TDF), a recently-found protein secreted by the pituitary into the blood. TDF mRNA has been detected in brain; not heart, placenta, lung, liver, skeletal muscle, or pancreas. Currently TDF has an unclear function, and prior to our studies, its localization was only minimally understood, with no understanding of receptor targets. We investigated the distribution of TDF in the rat brain using IHC and immunofluorescence (IF). TDF protein was detected in pituitary and most other brain regions, in specific neurons but not astrocytes. We found TDF immunoreactivity in cultured neuroblastoma, not astrocytoma. These data suggest that TDF is localized to neurons, not to astrocytes. Our group also conducted studies to identify the TDF receptor (TDF-R). Using LC-MS/MS and Western blotting, we identified the members of the Heat Shock 70-kDa family of proteins (HSP70) as potential TDF-R candidates in both MCF7 and BT-549 human breast cancer cells (HBCC)

*Author contributed equally with all other contributors.

A.G. Woods • I. Sokolowska • C.C. Darie (✉)
Biochemistry & Proteomics Group, Department of Chemistry & Biomolecular Science, Clarkson University, 8 Clarkson Avenue, Potsdam, NY 13699-5810, USA
e-mail: cdarie@clarkson.edu

K. Deinhardt
Centre for Biological Sciences, University of Southampton, Life Sciences Building 85, Southampton SO17 1BJ, UK

Institute for Life Sciences, University of Southampton, Life Sciences Building 85, Southampton SO17 1BJ, UK

A.G. Woods and C.C. Darie (eds.), *Advancements of Mass Spectrometry in Biomedical Research*, Advances in Experimental Medicine and Biology 806,
DOI 10.1007/978-3-319-06068-2_25, © Springer International Publishing Switzerland 2014

and PC3, DU145, and LNCaP human prostate cancer cells (HPCC), but not in HeLa cells, NG108 neuroblastoma, or HDF-a and BLK CL.4 cell fibroblasts or fibroblast-like cells. These studies have combined directed protein identification techniques with mass spectrometry to increase our understanding of a novel protein that may have distinct actions as a hormone in the body and as a growth factor in the brain.

Abbreviations

CNS	Central nervous system
GFAP	Glial fibrillary acidic protein
GS9L	Astrocytoma cell line
IF	Immunofluorescence
IHC	Immunohistochemistry
NeuN	Neuron-specific DNA-binding nuclear protein
NG108-15	Neuroblastoma × glioma cell line
TDF	Tumor differentiation factor
TDF-R	TDF receptor
WB	Western blotting

25.1 Introduction

New proteins identified in the brain can be crucial for understanding normal, pathological, and developmental central nervous system (CNS) processes. Specific proteins may have numerous roles, including regulating hormonal systems mediating neuroprotection, mitosis, process growth, and other functions [1–4]. Many proteins are locally distributed and produced within the CNS, but may also have functions outside of the CNS and may be transported via the bloodstream.

Tumor differentiation factor (TDF) is a novel, understudied protein. TDF was initially isolated from a human pituitary cDNA library and is a108-amino acid polypeptide protein. It was identified in GH3 pituitary tumor cell line and in human whole brain extract. TDF and TDF-P1 (a 20-amino acid peptide selected from the open reading frame of TDF), can differentiate human prostate cancer and breast cancer cells. TDF induces polarization and formation of cell junctions and basement membrane, synthesis of milk protein, and over-expression of E-cadherin. TDF has no known differentiation effect on kidney, hepatoma, fibroblasts, and leukemic lymphocytic cells. It may be a hormone, based on localization to the pituitary and human blood serum [5–9]. TDF has a molecular weight that is consistent with an anterior pituitary-derived hormone, not a posterior pituitary-derived hormone [9–11]. TDF may have a CNS function as well as a function outside of the pituitary gland. As an example of such a role, anterior-pituitary-derived hormones are produced outside of the anterior pituitary in the nervous, muscular, reproductive, or immune systems [12].

25.1.1 Identification of TDF in the Pituitary and in the Brain

Because TDF was known to be present in the brain extract and its cellular pattern was uncharacterized, we initiated a study to examine whether TDF is produced by the pituitary and to further investigate whether TDF is present in parts of the brain. We examined TDF protein in rat brain via immunohistochemistry (IHC) and immunofluorescence (IF) (Fig. 25.1). We found TDF protein in pituitary and several other brain regions. TDF immunoreactivity was present in neurons and their processes. Cells appearing to express TDF included cerebellar Purkinje cells, large-soma neurons in the hilus and sparse, large-soma cells in the granule cell region and pyramidal cell region of the hippocampus, individual cells in the striatal fundus and caudoputamen as well as pituitary gland cells, and apparent pyramidal neurons in the cerebral cortex.

Double-staining for detection of TDF and glial fibrillary acidic protein (GFAP), a marker for astroglia, showed no co-localization, indicating that TDF is not produced by astrocytes. Double-staining for TDF with the neuronal marker NeuN showed co-localization, indicating that TDF is made by neurons (Figs. 25.2 and 25.3). However, not all NeuN positive cells were TDF immunoreactive, suggesting that TDF is made by a subset of neurons. Supporting this idea, a subset of isolated primary hippocampal neurons grown in cell culture were also TDF-immunopositive (Fig. 25.4a), as shown in our published work [13]. Western blotting (WB) of NG108 neuroblastoma and GS9L astrocytoma cell lysate revealed TDF immunoreactivity in cultured neuroblastoma, not astrocytoma (Fig. 25.4b). Our experiments suggested that TDF is found in neurons, not astrocytes, and was the first report showing TDF in any cell type. This was the first report showing cellular localization of TDF in the brain and the first cellular localization for TDF protein demonstrated in any organ, confirming RNA data showing that TDF is produced by the pituitary gland and in the brain [6]. TDF may act as a pituitary-derived hormone peripherally and in the CNS, where it is made by distinct CNS neurons.

We tested the specificity of our antibody with controls, including incubation with the TDFP1 peptide versus random peptide, comparison with an antibody generated against a different TDF portion and omission of primary and secondary antibodies. These methods have been established as valid antibody specificity controls [14]. Incubation with citrate treatment did not change TDF staining; this is known to reduce non-specific staining [15]. Our antibody appeared to be specific to TDF.

25.1.2 Isolation and Identification of TDF Receptor Candidates from Androgen-Resistant DU145 Cells by Affinity Purification and Nanoliquid-Chromatography Tandem Mass Spectrometry (nanoLC-MS/MS)

Complementary to this study showing the localization of TDF in the brain, we also investigated the putative TDF receptor (TDF-R) candidates using MS and other directed methods. We used Affinity Purification (AP) chromatography to purify

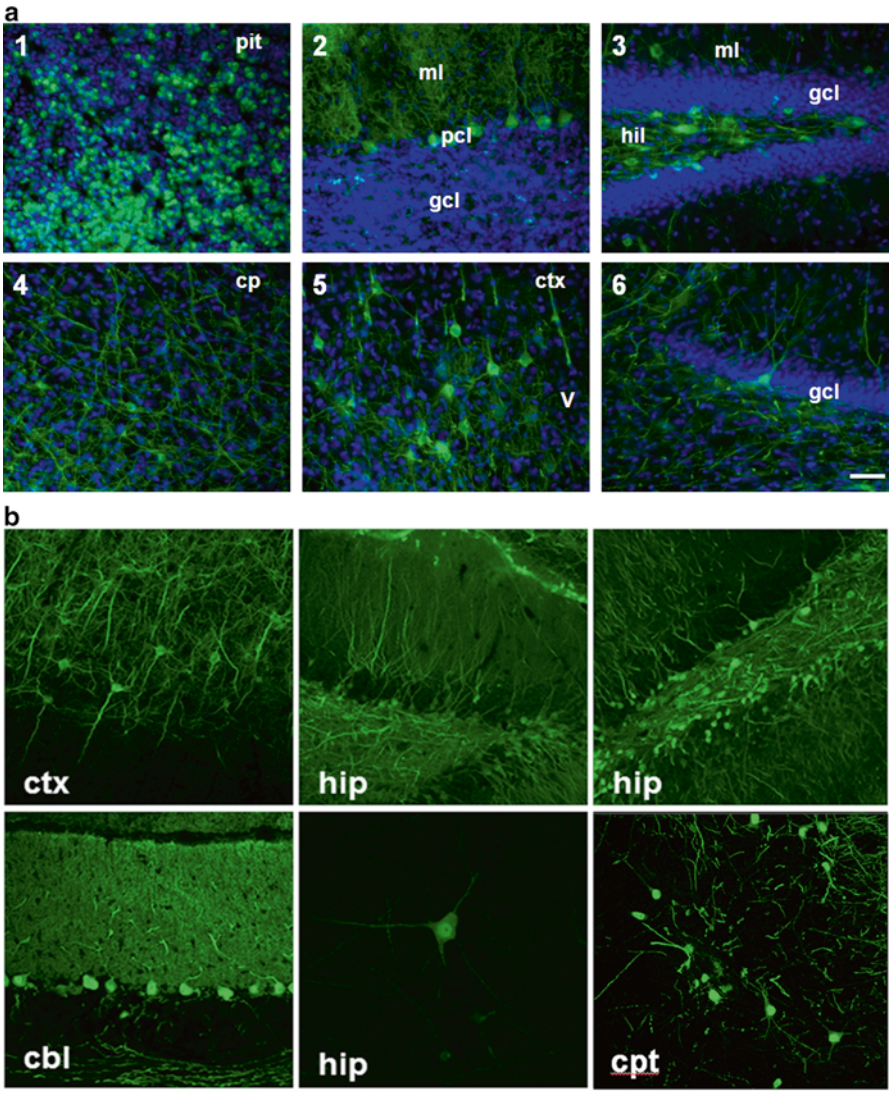


Fig. 25.1 TDF is found in pituitary and in other brain regions. (**a**) IHC of brain and pituitary tissues using anti-TDF-Ab (*green*) counter-stained with DAPI (*blue*). *Panel 1*: staining in pituitary. Some DAPI+ cells were TDF immunopositive, others were not. *Panel 2*: cerebellum. TDF immunoreactivity in Purkinje cell bodies and dendrites. TDF immunoreactivity is only occasionally seen in the granule cell layer. *Panel 3*: Hippocampus. Large cell bodies and processes stain in hilus, traversing into the molecular layer. Molecular layer cell bodies were also TDF immunopositive. *Panel 4*: in caudoputamen occasional cell bodies but primarily processes are TDF immunopositive. *Panel 5*: cortex, large cell bodies with pyramidal cell morphology and processes are TDF immunopositive. Many DAPI+ cells do not possess TDF immunoreactivity. *Panel 6*: Within hippocampal granule cell layer, a single large neuron and its processes stain. TDF immunoreactivity is primarily in process in this region. Scale bar = 50 μm. *pit* pituitary, *ml* molecular layer, *pcl* Purkinje cell layer, *gcl* granule cell layer, *hil* hilus, *cp* caudoputamen, *ctx* cortex. Figure adapted from [13] with permission from the publisher. (**b**) Additional labeling of TDF with anti-TDF Ab, showing that TDF is localized in both cell body and processes. Magnification: ×40. *Ctx* cortex, *hip* hippocampus, *clb* cerebellum, *cpt* caudoputamen

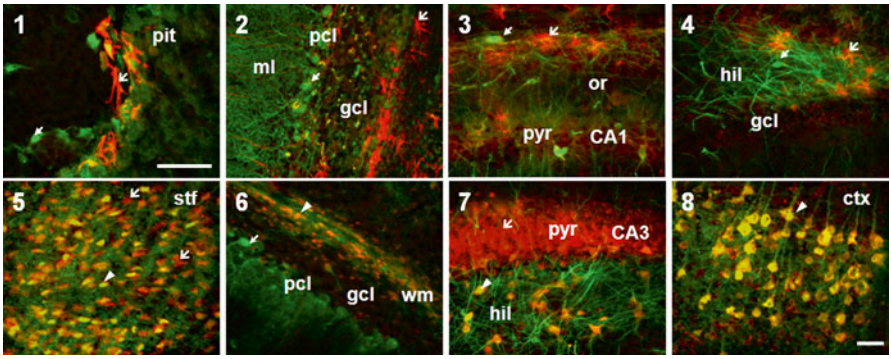


Fig. 25.2 TDF is found in neurons but not in astroglia. Immunostaining of the brain and pituitary using double-staining TDF and GFAP colocalization (1–4) or of TDF and NeuN (5–8). *Panel 1*: GFAP-immunopositive (+) cells (*red*) do not colocalize with TDF-immunopositive (+) cells (*green*) in pituitary (*small arrowhead* = anti-TDF+ alone, *large arrowhead* = anti-GFAP+ alone, *Panels 1–4*). GFAP-immunopositive astrocytes (*red*) do not colocalize with TDF-immunopositive cells (*green*) in cerebellum (*Panel 2*), hippocampal dentate gyrus (*Panel 3*) or hippocampal hilus (*Panel 4*). *Panel 5*: TDF immunoreactive cell bodies co-localize with NeuN, however all anti-NeuN+ neurons are not anti-TDF+ (*large arrowheads* = NeuN alone, *triangular arrowheads* = NeuN colocalized with TDF, *Panels 5–8*). *Panel 6*: In cerebellum, TDF alone appears in the Purkinje cell layer (*small arrowhead*). Some small cells in the granule cell layer and white matter are anti-TDF+ and co-localize with NeuN. White matter processes are also anti-TDF+. *Panel 7*: Pyramidal cells in hippocampal CA3 are anti-NeuN+ but not anti-TDF+. In surrounding hilus cells are anti-NeuN+ and anti-TDF+. *Panel 8*: Neurons with pyramidal cell morphology are both anti-NeuN+ and anti-TDF+, processes for these cells have only TDF immunoreactivity. Some cells are only anti-NeuN+. *pit* pituitary, *ml* molecular layer, *pcl* Purkinje cell layer, *gcl* granule cell layer, *or* stratum oriens of hippocampus, *pyr* pyramidal cells, *CA3* Cornus Ammonus 3 of hippocampus, *hil* hilus, *stf* striatal fundus, *wm* white matter, *ctx* cortex. Scale bar = 50 μM. Figure adapted from [13] with permission from the publisher

putative TDF-R candidates and LC-MS/MS to identify and characterize receptor candidates. We analyzed steroid-responsive MCF7 human breast cancer cells (HBCC) and steroid-resistant DU145 human prostate cancer cells (HPCC). We grew HPCC and HBCC in vitro, lysed and purified potential TDF-R candidates via TDF-P1 coupled to agarose beads. After purification of putative TDF-R candidates, we separated eluates by SDS-PAGE. Gel pieces were digested using trypsin. Peptide mixtures were extracted and analyzed using LC-MS/MS after which a Mascot database search was conducted to identify proteins. We isolated, identified, and characterized potential TDF-R candidates according to published procedures [16–20]. The results of these studies have been published [7–9, 11, 21–23].

In the AP and LC-MS/MS experiments using DU145 cells as starting material, we identified potential TDF receptor candidates with high confidence: GRP78 precursor or BiP (gi6470150), heat shock 70 kDa protein 8 isoform 1 (HSP8, gi62897129), heat shock 70 kDa protein (HSP1, gi386785), Heat shock 90Bb protein (HSP90Bb, gi20149594), Heat shock protein 90 (HSP90, gi306891), Sequestosome 1 (gi119574171), and Valosin-containing protein (gi11305821). HSP90Bb and HSP90 were not considered as they were only identified by one peptide.

The protein we identified with the highest probability to be the potential TDF receptor was Glucose regulated protein (GRP78), a 78 kDa protein and a heat shock

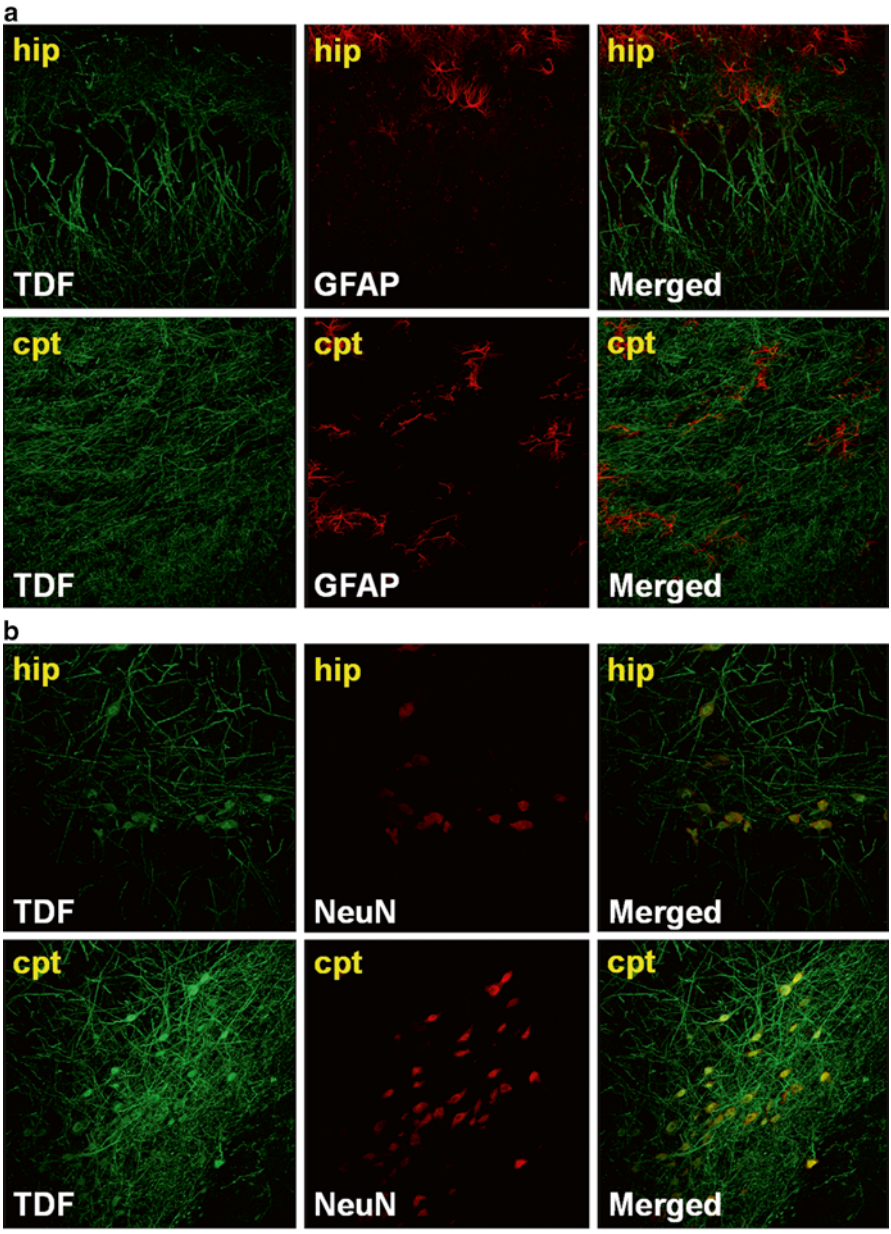


Fig. 25.3 Higher resolution immunolabeling of hippocampus (hip) and caudaputamen (cpt) with anti-TDF Ab and anti-GFAP Ab (**a**) and anti-TDF Ab and NeuN Ab (**b**) showing that TDF is found in neurons but not in astroglia. IHC was visualized by confocal microscopy. Magnification: ×40

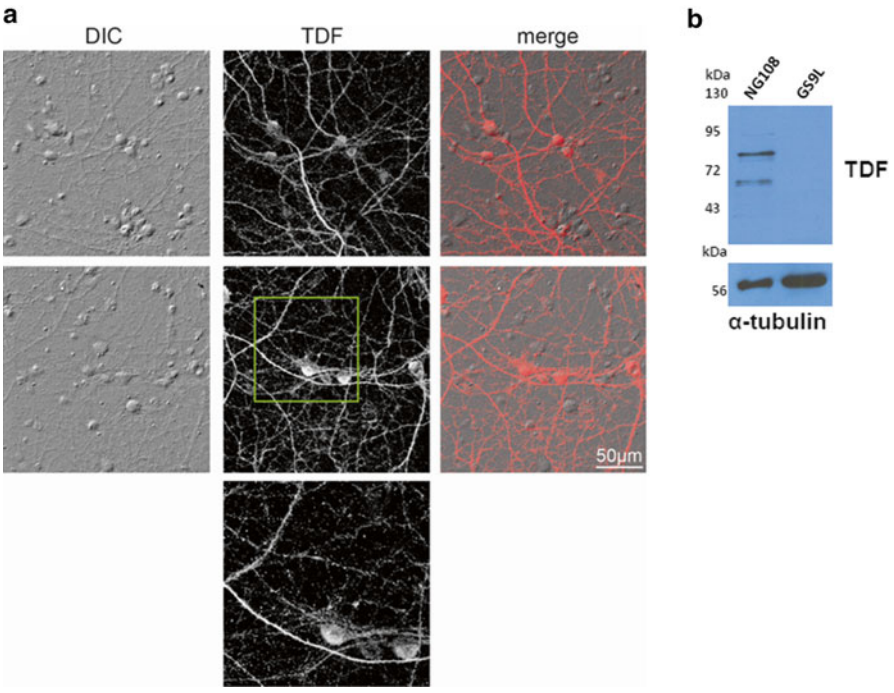


Fig. 25.4 (**a**) TDF is found in cultured hippocampal neurons. Shown are two examples of mature primary hippocampal cultures stained with anti-TDF-Abs. *Bottom panel* shows a higher magnification of the boxed region. Scale bar, 50 μm. (**b**) TDF is found in NG108 neuroblastoma but not in GS9L astrocytoma cell lysate. WB for TDF protein in NG108 neuroblastoma and GS9L astrocytoma cell lysate. Two bands appear (possibly monomer and dimer). The blot was stripped and reprobed with anti-α-tubulin. Molecular mass markers indicated in kDa. Figure adapted from [13] with permission from the publisher

protein (HSP) family member, also called Heat shock 70 kDa protein 5 (HSP70 or HSP5) or immunoglobulin heavy chain-binding protein (BiP). This protein is involved in the folding and assembly of endoplasmic reticulum proteins but may also be found in cytosol and cell membranes [24, 25]. Cancer cells have high HSP expression [24, 26–33], which are essential to cancer cell survival [34]. Therefore, HSP inhibitors may act for anticancer treatments [35, 36]. HSPs (HSP70 and HSP90) are associated with estrogen and androgen receptors [37–41]. HSPs regulate cell proliferation. Inhibition of one HSP (HSP90) can lead to dysregulation of a different HSPs (HPS70) as well as inhibition of cell proliferation [42]. HSP70 and HSP90 block apoptotic pathways, promoting cell differentiation [42]. HSPs may determine whether the cells undergo apoptosis or differentiation [43]. GRP78 complexes at the cell surface with Cripto, an oncoprotein that signals via MAPK/ERK, PI3K/Akt and Smad2/3 pathways, mediating signaling in human tumor, mammary epithelial, and embryonic stem cells [24]. Active Cripto from Cripto-GRP78 complex increases cell proliferation, decreases cell adhesion and down-regulates of

E-Cadherin. But Cripto alone cannot accomplish these cellular events and must complex withGRP78, acting as an oncogene [24]. When Cripto is not complexed with GRP78, it may promote cellular differentiation [43]. Two other HSPs, heat shock 70 kDa protein 8 isoform 1 (HSP8, gi62897129), heat shock 70 kDa protein (HSP70, gi386785) are also currently being investigated by our group.

Using androgen-resistant DU145 cells and AP we also identified Sequestosome 1 as a putative TDF-R. This is a 47 kDa cytoplasmatic protein, also known as Ubiquitin-binding protein p62. This protein is likely involved in cell signaling, receptor internalization, protein turnover, and protein–protein interactions. We did not identify this protein in the AP experiments using MCF7 cells, so it may have been a contaminant. However, we still consider this protein on our list of TDF receptor candidates, pending further experiments. The final protein identified using AP and LC-MS/MS with DU145 cells was Valosin-containing protein, also known as Transitional endoplasmic reticulum ATPase, an 89 kDa protein. This is usually found as a homohexamer [44] with a molecular mass of 540 kDa [45–47]. The MS/MS spectra that led to identification of the TDF-R candidates (HSP70, GRP78, and HSP90) are shown in Fig. 25.5.

25.1.3 Isolation and Identification of TDF Receptor Candidates from Estrogen-Responsive MCF7 Cells by AP and LC-MS/MS

We identified four proteins with high confidence via LC-MS/MS experiments performed with the material isolated from estrogen-responsive MCF7 cells. These putative receptor proteins included: GRP78 precursor (gi386758); heat shock 70 kDa protein 8 isoform 1 (HSP8, gi5729877 heat shock 70 kDa protein 1 (HSP1, gi4529893); and Heat shock 70 kDa protein 9 (HSPA9, gi12653415). The only strong TDF receptor candidates were GRP78 and HSP70, identified with high confidence experiments using both cancer cell types. Signalosome 1 and Valosin-containing protein were found to be unlikely TDF-R candidates. The other potential TDF-R candidates (heat shock 70 kDa protein 8 isoform 1, heat shock-induced protein, and heat shock 70 kDa protein 9 isoform) are all HSPs from the same family with GRP78, currently being investigated.

25.1.4 Validation of the AP and LC-MS/MS by AP and Western Blotting

Validation of the TDF-R was investigated using AP and WB in several cell lines, shown in Fig. 25.6 and summarized in Table 25.1. Based on these studies, we found that GRP78 is the primary TDF-R candidate, but HRP70 may also be a

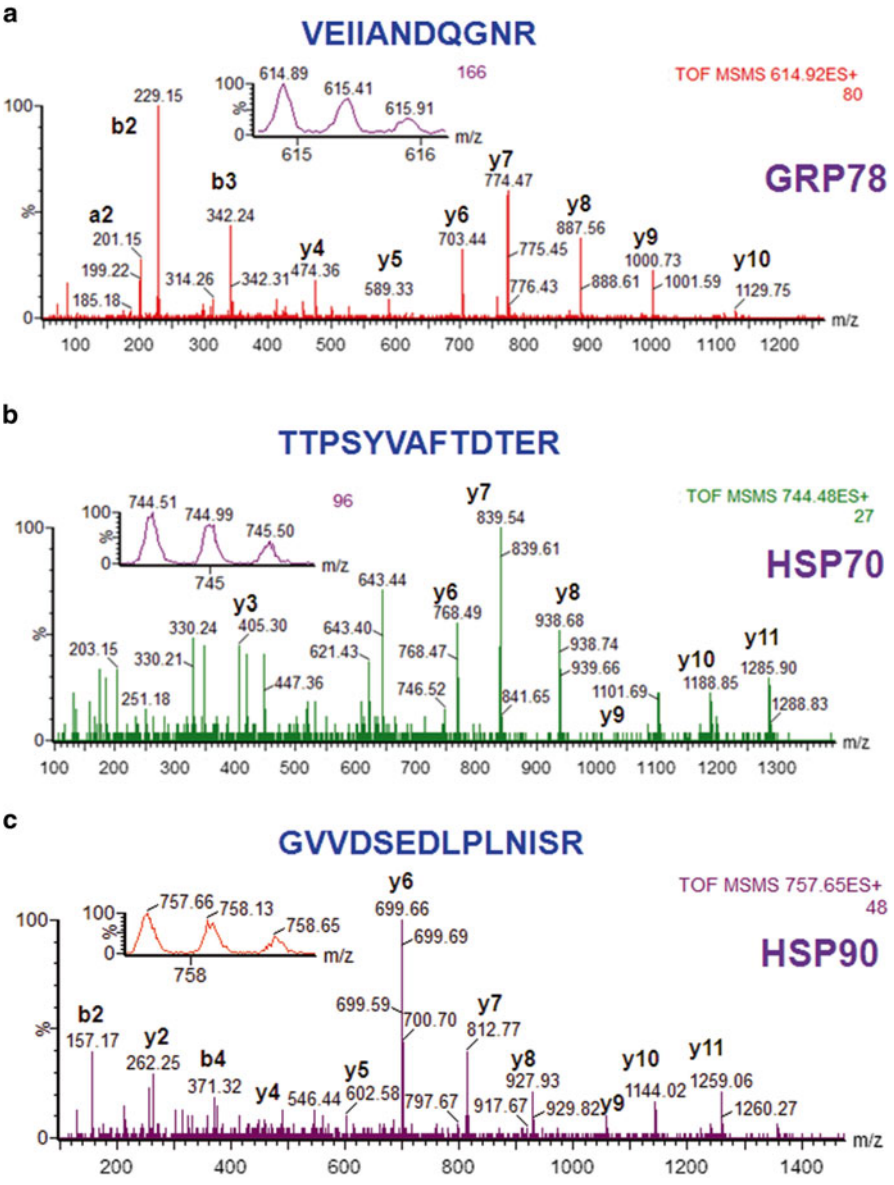


Fig. 25.5 Identification of TDF-R candidates in DU145 cells using AP and LC-MS/MS (AP-MS). The potential receptors for TDF protein were purified from cell lysate using AP, resulting samples were separated by SDS-PAGE and the gel bands were excised and digested by trypsin. The peptides mixture was analyzed by LC-MS/MS to identify the purified proteins. (**a**) MS/MS spectrum of peptide VEIIANDQGNR that led to identification of GRP78 as TDF-R candidate. (**b**) MS/MS spectrum of peptide TTPSYVAFTDTER that led to identification of HSP70 as TDF-R candidate. (**c**) MS/MS spectrum of peptide GVVDSEDLPLNISR that led to identification of HSP90 as TDF-R candidate. Figure adapted from [21] with permission from the publisher

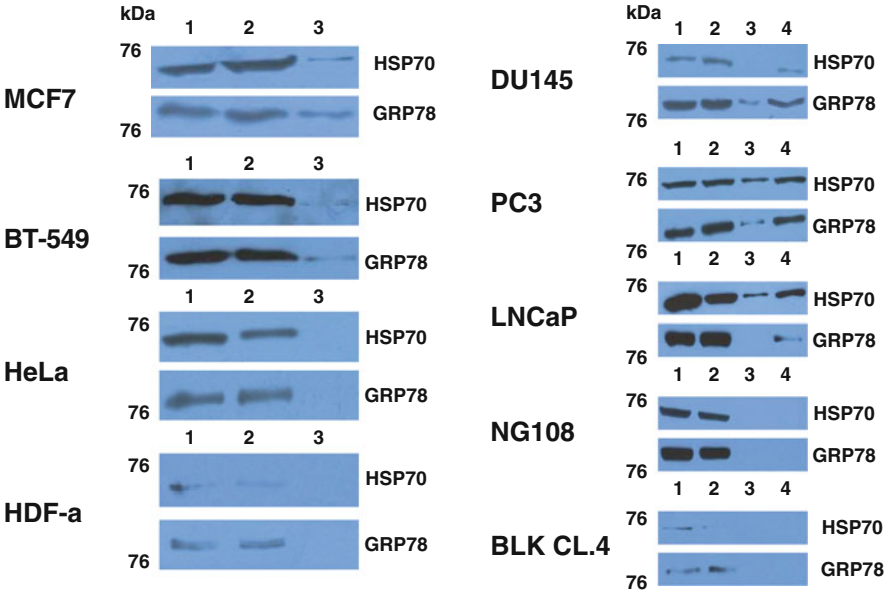


Fig. 25.6 AP and WB analyses of various cell lines using anti-GRP78 and HSP70 antibodies. The TDF-R candidates were purified from the cell lines indicated by AP using TDF-P1 peptide and then the eluates were investigated by WB. The cell lysates were prepared from HBCC (MCF7 & BT-549), HPCC (DU145, PC3 & LNCaP), cancerous cells (HeLa & NG108 neuroblastoma), and normal cells (HDF-a fibroblasts & BLK CL.4 fibroblast-like cells). In WBs: (1) input cell lysate, (2) flow through, (3) eluate, and (4) 5× concentrated eluate. The molecular mass markers are indicated (kDa). Figure adapted from [9] with permission from the publisher

Table 25.1 Summary of the cell lines where GRP78 and HSP70 proteins were identified by AP and LC-MS/MS or by AP and WB

Cell lines		GRP78	HSP70
Breast cancer cell lines	MCF7	+	+
	BT-549	+	+
Prostate cancer cell lines	DU145	+	+
	PC3	+	+
	LNCaP	+	+
Other cancer cell lines	HeLa	–	–
	NG108	–	–
Normal cell lines	HDF-a	–	–
	BLK CL.4	–	–

Table reproduced from [9] with permission from the publisher

TDF-R candidate, representing a family of proteins including HSP8, HSP1, and HSP70. Both GRP78 and HSP70 are TDF-R may also form a GRP78-HSP70 core protein complex.

25.1.5 GRP78 and HSP70 May Share a Steroid-Independent Mechanism of TDF-Induced Cell Differentiation in Breast and Prostate Cells But Not in Other Cells

We initially tested HBCC to validate TDF-R candidates, finding GRP78 and HSP70 by AP and WB in both MCF7 steroid-responsive HBCC and in steroid resistant BT-549 HBCC (cells that do not express estrogen receptors) [48]. This suggests that TDF-R candidates may activate the TDF pathway via a steroid-independent pathway possibly parallel to the steroid pathway. We found a similar trend in steroid-responsive LNCaP HPCC and steroid-resistant DU145 and PC3 HPCC, suggesting that the TDF pathway may be common to breast and prostate cancer cells. We did not find TDF-R candidates in AP and WB experiments in non-breast, non-prostate HeLa cancerous cells, NG108 neuroblastoma x glioma cells, as well as in non-breast, non-prostate, non-cancerous human dermal fibroblasts (HDF-a) or normal embryonal fibroblast-like cells (BLK CL.4). TDF-R candidates appear to be restricted to breast and prostate cells, and are not found in non-breast, non-prostate cancerous cells or normal non-breast, non-prostate cells. Therefore, the TDF pathway may be restricted to steroid-responsive and steroid-resistant breast and prostate cells.

25.1.6 Investigation of GRP78 and HSP70 as Potential TDF-R Candidates Using Immunofluorescence and Confocal Microscopy

Using IF and confocal microscopy we observed that GRP78 Ab interacts with their antigens outside the plasma membrane, compared with CM-Dil, a plasma membrane marker [7]. This pattern was observed in both steroid-responsive MCF7 and steroid-resistant BT-549 cells, as well as in HeLa cancerous cells, not in HDF-a cells. We also found that HSP70 antibodies identified their antigens outside of plasma membrane in all cell lines studied [7]. This led us to propose that 1) the HSP70 is not a specific TDF-R, 2) staining was not specific, or 3) HSP70 exists outside the plasma membrane in all cells studied, but the TDF-R is a multi-subunit receptor with specificity and activity that depends on the HSP70 with GRP78 interaction to form a TDF-R complex, consisting of at minimum GRP78 and HSP70. Other subunits could fit to our theory, including HSP90Bb or HSP90 beta. These were identified in our experiments and are known to interact with HSP70 proteins.

In a follow-up study, GRP78 protein was detected by antibodies outside the plasma membrane of steroid-resistant DU145 (Fig. 25.7) and PC3 HPCC, steroid-responsive LNCaP HPCC, and in the non-prostate NG108 neuroblastoma cells, but not in the normal fibroblast-like BLK CL.4 cells [8]. IF experiments examining HSP70 showed similar staining in all cell lines studied, including DU145, PC3, LNCaP HPCC, NG108 neuroblastoma, and normal fibroblast-like BLK CL.4 cells.

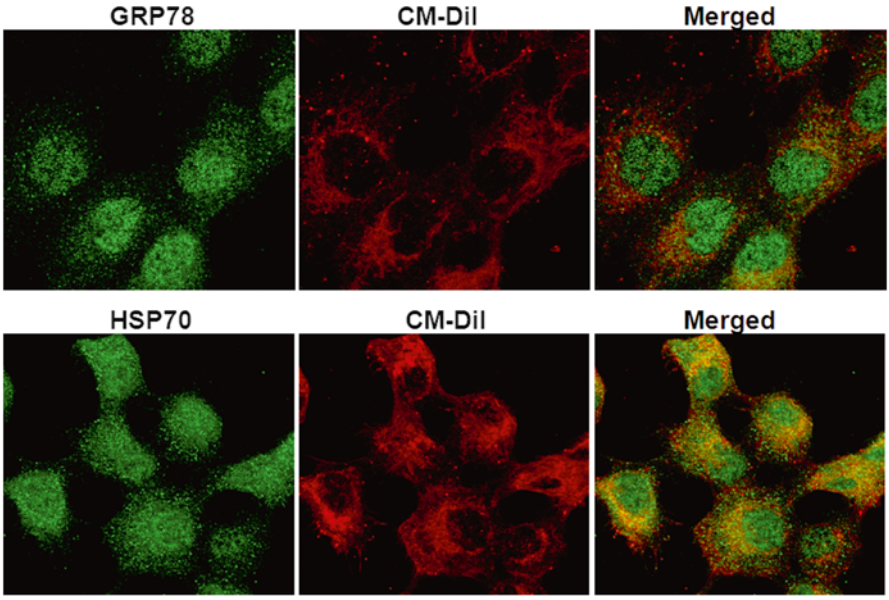


Fig. 25.7 Immunolocalization of GRP78 and HSP70 proteins in DU145 cells by confocal microscopy. GRP78 and HSP70 proteins were detected by anti-GRP78 and anti-HSP70 antibodies and then visualized with AlexaFluor 488 antibodies (*green*). Plasma membrane was stained with CM-DiI (*red*). The merged images are also shown. Figure adapted from [9] with permission from the publisher

While GRP78 is detected outside the cells in HeLa and NG108 cancerous cells, detection of this protein is likely not sufficient to be named a TDF-R. Therefore, the proposal that TDF-R may contain more than one subunit seems even more likely. We propose that the interaction of GRP78 with HSP70 and possibly other proteins may lead to a functional TDF-R, which may also be a TDF-induced-formation of TDF-R complex or a constitutive TDF-R complex (Fig. 25.8).

Overall, these studies have combined directed protein identification techniques with MS to increase our understanding of a novel protein that may have distinct actions as a hormone and in the brain. We found that TDF protein is localized to distinct cells, specifically neurons and pituitary cells, and may further be found in other regions of the body. Preliminary in situ *hybridization* studies confirm these findings, and the concept that TDF may be synthesized locally by neurons. Our preliminary unpublished studies have also identified TDF immunoreactivity in specific breast and prostate cells in healthy adult rat tissue. Further studies will continue to explore TDF localization. We have additionally identified the TDF-R candidates using AP, LC-MS/MS, and confocal IF. GRP78 and HSP70 currently appear to be the most promising candidates. These studies underscore how MS can be combined with other techniques to understand a novel protein.

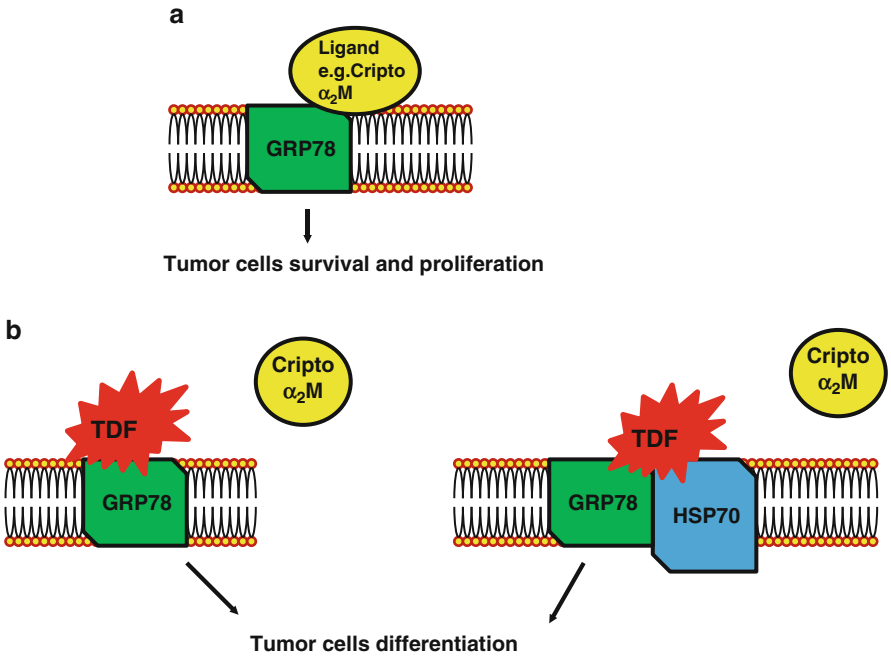


Fig. 25.8 Hypothetical mechanism of action of TDF which considers TDF-R as a protein complex. (**a**) This model considers more interaction partners for GRP78 (e.g., Cripto, alpha2 macroglobulin (a2M)). These GRP78 partners, either soluble (a2M) or membrane-bound (Cripto) are competitively replaced by TDF. (**b**) TDF-R is considered as one molecule (GRP78) or a protein complex formed from at least GRP78 and HSP70. Figure adapted from [9] with permission from the publisher

Acknowledgments We would like to thank Ms. Laura Mulderig, Scott Nichols, and their colleagues (Waters Corporation) for their generous support in setting up the Proteomics Center at Clarkson University. C.C.D. thanks Drs. Thomas A. Neubert (New York University, Belinda Willard (Cleveland Clinic), and Gregory Wolber & David Mclaughin (Eastman Kodak Company) for donation of a TofSpec2E MALDI-MS (each). C.C.D. thanks his advisors, Vlad Artenie, Wolfgang Haehnel, Paul M. Wassarman & Thomas A. Neubert advice and support. This work was supported in part by private donations (Spring Beck Richardson, Guillaume Mimoun, Bob and Karen Brown), the David A. Walsh fellowship, and by the U.S. Army research office (DURIP grant #W911NF-11-1-0304).

References

1. Woods AG, Poulsen FR, Gall CM (1999) Dexamethasone selectively suppresses microglial trophic responses to hippocampal deafferentation. Neuroscience 91(4):1277–1289
2. Woods AG et al (1998) Deafferentation-induced increases in hippocampal insulin-like growth factor-1 messenger RNA expression are severely attenuated in middle aged and aged rats. Neuroscience 83(3):663–668

3. Guthrie KM et al (1997) Astroglial ciliary neurotrophic factor mRNA expression is increased in fields of axonal sprouting in deafferented hippocampus. J Comp Neurol 386(1):137–148
4. Dallner C et al (2002) CNTF and CNTF receptor alpha are constitutively expressed by astrocytes in the mouse brain. Glia 37(4):374–378
5. Platica M et al (1992) Pituitary extract causes aggregation and differentiation of rat mammary tumor MTW9/Pl cells. Endocrinology 131(6):2573–2580
6. Platica M et al (2004) A pituitary gene encodes a protein that produces differentiation of breast and prostate cancer cells. Proc Natl Acad Sci U S A 101(6):1560–1565
7. Sokolowska I et al (2012) Identification of potential tumor differentiation factor (TDF) receptor from steroid-responsive and steroid-resistant breast cancer cells. J Biol Chem 287(3):1719–1733
8. Sokolowska I et al (2012) Identification of a potential tumor differentiation factor receptor candidate in prostate cancer cells. FEBS J 279(14):2579–2594
9. Sokolowska I et al (2013) Characterization of tumor differentiation factor (TDF) and its receptor (TDF-R). Cell Mol Life Sci 70:2835–2848
10. Caldwell HK, Young III WS (2006) Oxytocin and vasopressin: genetics and behavioral implications. In: Lim R, Lajtha A (eds) Handbook of neurochemistry and molecular neurobiology: neuroactive proteins and peptides (3rd edn), vol 40. Springer, Berlin, pp 573–607
11. Roy U et al (2012) Structural investigation of tumor differentiation factor (TDF). Biotechnol Appl Biochem 59:445–450
12. Harvey S (2010) Extrapituitary growth hormone. Endocrine 38(3):335–359
13. Woods AG et al (2013) Identification of tumor differentiation factor (TDF) in select CNS neurons. Brain Struct Funct. [Epub ahead of print]
14. Burry RW (2011) Controls for immunocytochemistry: an update. J Histochem Cytochem 59(1):6–12
15. Shi SR et al (1993) Antigen retrieval technique utilizing citrate buffer or urea solution for immunohistochemical demonstration of androgen receptor in formalin-fixed paraffin sections. J Histochem Cytochem 41(11):1599–1604
16. Ngounou Wetie AG et al (2013) Automated mass spectrometry-based functional assay for the routine analysis of the secretome. J Lab Autom 18(1):19–29
17. Sokolowska I et al (2012) Proteomic analysis of plasma membranes isolated from undifferentiated and differentiated HepaRG cells. Proteome Sci 10(1):47
18. Sokolowska I et al (2012) Disulfide proteomics for identification of extracellular or secreted proteins. Electrophoresis 33(16):2527–2536
19. Sokolowska I et al (2013) Mass spectrometry investigation of glycosylation on the NXS/T sites in recombinant glycoproteins. Biochim Biophys Acta 1834(8):1474–1483
20. Sokolowska I et al (2012) Automatic determination of disulfide bridges in proteins. J Lab Autom 17(6):408–416
21. Roy U et al (2013) Tumor differentiation factor (TDF) and its receptor (TDF-R): is TDF-R an inducible complex with multiple docking sites? Mod Chem Appl 1(108)
22. Roy U et al (2013) Structural evaluation and analyses of tumor differentiation factor. Protein J 32(7):512–518
23. Sokolowska I et al (2011) Mass spectrometry for proteomics-based investigation of oxidative stress and heat shock proteins. In: Andreescu S, Hepel M (eds) Oxidative stress: diagnostics, prevention, and therapy. American Chemical Society, Washington, DC
24. Kelber JA et al (2009) Blockade of Cripto binding to cell surface GRP78 inhibits oncogenic Cripto signaling via MAPK/PI3K and Smad2/3 pathways. Oncogene 28(24):2324–2336
25. Ni M et al (2009) Regulation of PERK signaling and leukemic cell survival by a novel cytosolic isoform of the UPR regulator GRP78/BiP. PLoS One 4(8):e6868
26. Graner MW et al (2009) Heat shock protein 70-binding protein 1 is highly expressed in high-grade gliomas, interacts with multiple heat shock protein 70 family members, and specifically binds brain tumor cell surfaces. Cancer Sci 100(10):1870–1879
27. Wu B, Wilmouth RC (2008) Proteomics analysis of immunoprecipitated proteins associated with the oncogenic kinase cot. Mol Cells 25(1):43–49

28. Fu Y, Lee AS (2006) Glucose regulated proteins in cancer progression, drug resistance and immunotherapy. Cancer Biol Ther 5(7):741–744
29. Kuwabara H et al (2006) Glucose regulated proteins 78 and 75 bind to the receptor for hyaluronan mediated motility in interphase microtubules. Biochem Biophys Res Commun 339(3):971–976
30. Lim SO et al (2005) Expression of heat shock proteins (HSP27, HSP60, HSP70, HSP90, GRP78, GRP94) in hepatitis B virus-related hepatocellular carcinomas and dysplastic nodules. World J Gastroenterol 11(14):2072–2079
31. Fernandez PM et al (2000) Overexpression of the glucose-regulated stress gene GRP78 in malignant but not benign human breast lesions. Breast Cancer Res Treat 59(1):15–26
32. Menoret A, Bell G (2000) Purification of multiple heat shock proteins from a single tumor sample. J Immunol Methods 237(1–2):119–130
33. Stoeckle MY et al (1988) 78-kilodalton glucose-regulated protein is induced in Rous sarcoma virus-transformed cells independently of glucose deprivation. Mol Cell Biol 8(7):2675–2680
34. Daneshmand S et al (2007) Glucose-regulated protein GRP78 is up-regulated in prostate cancer and correlates with recurrence and survival. Hum Pathol 38(10):1547–1552
35. Didelot C et al (2007) Anti-cancer therapeutic approaches based on intracellular and extracellular heat shock proteins. Curr Med Chem 14(27):2839–2847
36. Solit DB, Rosen N (2006) Hsp90: a novel target for cancer therapy. Curr Top Med Chem 6(11):1205–1214
37. Nair SC et al (1996) A pathway of multi-chaperone interactions common to diverse regulatory proteins: estrogen receptor, Fes tyrosine kinase, heat shock transcription factor Hsf1, and the aryl hydrocarbon receptor. Cell Stress Chaperones 1(4):237–250
38. Nemoto T, Ohara-Nemoto Y, Ota M (1992) Association of the 90-kDa heat shock protein does not affect the ligand-binding ability of androgen receptor. J Steroid Biochem Mol Biol 42(8):803–812
39. Sanchez ER et al (1990) The 56-59-kilodalton protein identified in untransformed steroid receptor complexes is a unique protein that exists in cytosol in a complex with both the 70- and 90-kilodalton heat shock proteins. Biochemistry 29(21):5145–5152
40. Veldscholte J et al (1992) Anti-androgens and the mutated androgen receptor of LNCaP cells: differential effects on binding affinity, heat-shock protein interaction, and transcription activation. Biochemistry 31(8):2393–2399
41. Veldscholte J et al (1992) Hormone-induced dissociation of the androgen receptor-heat-shock protein complex: use of a new monoclonal antibody to distinguish transformed from nontransformed receptors. Biochemistry 31(32):7422–7430
42. Jensen MR et al (2008) NVP-AUY922: a small molecule HSP90 inhibitor with potent antitumor activity in preclinical breast cancer models. Breast Cancer Res 10(2):R33
43. Lanneau D et al (2007) Apoptosis versus cell differentiation: role of heat shock proteins HSP90, HSP70 and HSP27. Prion 1(1):53–60
44. DeLaBarre B, Brunger AT (2003) Complete structure of p97/valosin-containing protein reveals communication between nucleotide domains. Nat Struct Biol 10(10):856–863
45. Darie CC, Shetty V, Spellman DS, Zhang G, Xu C, Cardasis HL, Blais S, Fenyo D, Neubert TA (2008) Blue native PAGE and mass spectrometry analysis of the ephrin stimulation-dependent protein-protein interactions in NG108-EphB2 cells. Applications of mass spectrometry in life safety, NATO Science for Peace and Security Series. Springer, Düsseldorf, Germany, pp 3–22
46. Darie CC et al (2011) Identifying transient protein-protein interactions in EphB2 signaling by blue native PAGE and mass spectrometry. Proteomics 11(23):4514–4528
47. Woods AG (2011) Blue native PAGE and mass spectrometry as an approach for the investigation of stable and transient protein-protein interactions. In: Oxidative stress: diagnostics and therapy. Chapter 12, pp 341–367
48. Vladusic EA et al (2000) Expression and regulation of estrogen receptor beta in human breast tumors and cell lines. Oncol Rep 7(1):157–167

Chapter 26
Mass Spectrometry for the Study of Autism and Neurodevelopmental Disorders

Armand G. Ngounou Wetie, Robert M. Dekroon, Mihaela Mocanu, Jeanne P. Ryan, Costel C. Darie, and Alisa G. Woods

Abstract Mass spectrometry (MS) has been increasingly used to study central nervous system disorders, including autism spectrum disorders (ASDs). The first studies of ASD using MS focused on the identification of external toxins, but current research is more directed at understanding endogenous protein changes that occur in ASD (ASD proteomics). This chapter focuses on how MS has been used to study ASDs, with particular focus on proteomic analysis. Other neurodevelopmental disorders have been investigated using this technique, including genetic syndromes associated with autism such as fragile X syndrome and Smith–Lemli–Opitz syndrome.

A.G. Ngounou Wetie • C.C. Darie
Biochemistry & Proteomics Group, Department of Chemistry & Biomolecular Science, Clarkson University, 8 Clarkson Avenue, Potsdam, NY 13699-5810, USA

R.M. Dekroon • M. Mocanu
Lineberger Comprehensive Cancer Center, University of North Carolina, Chapel Hill, NC 27599, USA

J.P. Ryan
SUNY Plattsburgh Neuropsychology Clinic and Psychoeducation Services, 101 Broad Street, Plattsburgh, NY 12901, USA

A.G. Woods (✉)
Biochemistry & Proteomics Group, Department of Chemistry & Biomolecular Science, Clarkson University, 8 Clarkson Avenue, Potsdam, NY 13699-5810, USA

SUNY Plattsburgh Neuropsychology Clinic and Psychoeducation Services, 101 Broad Street, Plattsburgh, NY 12901, USA
e-mail: awoods@clarkson.edu

A.G. Woods and C.C. Darie (eds.), *Advancements of Mass Spectrometry in Biomedical Research*, Advances in Experimental Medicine and Biology 806,
DOI 10.1007/978-3-319-06068-2_26, © Springer International Publishing Switzerland 2014

26.1 Introduction

Mass spectrometry (MS)-based proteomics are a new methodology that is increasingly being used to study autism and neurodevelopmental disorders. Initial research into these syndromes focused on identification of genetic information, and the number of studies involving genetic research are highly disproportionate relative to those involving proteomic research [1]. Much of this focus may have to do with historical factors (such as the cloning of the human genome), training of biomedical scientists, and available technology.

Proteomics can complement genetic information about neurodevelopmental disorders such as autism spectrum disorder (ASD) by elucidating disorder subtypes, by identifying convergent protein abnormalities caused by different genes, or by expanding on posttranslational modifications (PTMs) or protein perturbations caused by environmental factors [2]. Proteomics measure active protein levels [3] or PTMs such as glycosylation, phosphorylation, and formation/destruction of disulfide bridges that regulate protein three-dimensional structure [2, 4–16]. Proteomics is defined as the study of the total complement of proteins at the time of collection in a specific cell, organelle, extracellular fluid, or tissue [5, 7, 17–23]. MS is considered the workhorse of proteomics that can be used to identify endogenous compounds, such as proteins, or exogenously introduced compounds, such as toxins [24].

Cerebrospinal fluid, urine, blood sera, saliva, hair, and tissue biopsies can all be analyzed using MS. For investigations in central nervous system (CNS) disorders, proteomic analysis of brain tissue, including synaptic proteins, using MS is also possible, an approach that tends to be restricted to animal models [25, 26]. Different bodily materials for analysis may confer different advantages and challenges. Blood serum is a complex, containing proteins, nucleic acids, lipids, and other metabolic products [27], making it a potential source of biomarkers. Blood, however, may require an additional depletion step to remove abundant proteins [28], which can complicate the analysis. Urine analysis is noninvasive and convenient, and urine is easy to obtain in large quantities. Urine contains mainly plasma proteins and proteins secreted from the urinary tract and kidney, primarily containing lower molecular weight proteins below 30 kDa [29]. The urinary proteome is less complex compared to the serum proteome making low-abundant molecules much easier to detect. This is also true of the salivary proteome, with the additional advantage that saliva contains less salt than urine [30]. Salts can complicate proteomic experiments [31]. Compared to blood plasma, 2,290 proteins have been found in saliva versus 2,698 proteins found in plasma [32]. Therefore, several biological fluids should be potentially used for analysis in neurodevelopmental disorders such as ASD. Studies using these materials have already been conducted, but due to the relatively small number of studies and new/rapidly evolving technology, this field is still in its infancy [31].

26.2 Mass Spectrometry in Autism Research

26.2.1 Autism Overview

ASDs have been defined by the presence of three categories of impairments: (1) social deficits, (2) repetitive/stereotypical behaviors and interests, and (3) communication difficulties [33]. Until recently ASDs were separated into three primary categories classified separately on the DSM-IV-TR: Asperger's syndrome (AS), autism, and pervasive developmental disorder not otherwise specified (PDD-NOS) [34, 35]. In May 2013 the DSM-5 was introduced, collapsing all categories into one umbrella term, ASD. Clinical aspects are divided into an impairment dyad: (1) persistent deficits in social communication issues and (2) restricted, repetitive behavior, interests, and activities [33, 34].

The incidence of ASDs is on the rise, with 1 in 88 US children estimated as having an ASD; one in 54 boys and one in 252 girls are affected [36]. A United States government survey (Centers for Disease Control and Prevention; CDC) of parents has reported that ASD prevalence may be as high as 1 in 50 [37]. Based on the rapid recent rise in the numbers of children diagnosed, ASD prevalence likely cannot be accounted for by improved diagnosis alone. Current etiological theories support a genetic and environmental interaction [38, 39].

Many genetic studies of ASD have been conducted [40], although this has not made the etiology of the disorder clear. ASD is more than 20 times increased in first-degree relatives [41], based on an analysis in twins. From 60 to 92 % ASD concordance occurs in identical twins and from 0 to 10 % in fraternal twins [41, 42]. Although genetic influences seem clear, numerous associated genes have been suggested, without clear consistency in many instances. Over 100 genes and 40 genomic *loci* have been associated with ASD [43, 44]. A recent study of de novo mutations in ASD suggested that several mutations identified in ASDs are not associated with the disorder [45]. A probable heterogeneity of ASD causality [46] therefore exists, and genetic information alone does not provide a comprehensive picture for understanding ASD [47]. Genes involved in ASD may share molecular mechanisms, despite the large number of candidates [48], suggesting functional convergence at the protein level. Many genes likely collaborate with environmental influences to cause ASDs [46, 49, 50]. This genetic complexity and the fact that genetics alone do not provide clear predictive markers limit the utility of genetic testing for ASDs [49]. In fact, clinical geneticists rule out known genetic syndromes when evaluating individuals for ASD, leaving 80 % of families without a definitive ASD diagnosis [51].

Many studies have pointed towards biological disturbances in ASDs [23, 52–54]; however, a clear diagnostic biomarker for ASD is not currently available. ASDs are therefore diagnosed based on behavioral symptoms, and although gold-standard behavioral tests exist [55, 56], biomarkers could improve screening and diagnosis further. ASDs still often go unrecognized in children and even in adults, and current instruments used to screen for ASDs can produce false positive or false negative diagnoses [57]. For this reason, moving beyond genetic analysis in the search for

ASD biomarkers may provide new inroads to diagnosis and understanding of ASD causes. MS-based proteomics provides one such option for exploration. Before reviewing MS-based proteomics, the current focus of MS analysis in ASDs, it is of interest to review initial MS studies of ASD, which focused on exogenous factors, rather than endogenous proteins.

26.2.2 Analysis of Exogenous Toxins in ASD Using MS

Urine is the most frequently studied biomaterial in ASD research, and has generally been used to examine changes in exogenously introduced compounds proposed to have a causal role in ASD. The two major theories studied in this regard are the opioid and heavy metal theories of ASD etiology.

One group of researchers used high pressure liquid chromatography (HPLC) and matrix-assisted laser desorption/ionization-time-of-flight mass spectrometry (MALDI-TOF MS) to study the urinary contents of boys with autism ($n=65$) versus controls ($n=158$). The investigation was directed at testing the hypothesis that exogenously derived opioid peptides are found in urine of children with ASD. This idea has not had strong empirical support [58] and indeed, no significant differences were found between the urine contents of the two groups [59]. A separate study tested this idea using LC–UV-MS (liquid chromatography–ultraviolet-MS), and opioid peptides were not detected in ten children with autism or their ten non-autistic siblings [60].

A separate study tested for opioids in individuals with ASDs using highly sensitive liquid chromatography–electrospray ionization-tandem mass spectrometry (LC–ESI-MS/MS or LC–MS/MS). The researchers developed this method to be selective for opioid peptides in urine of children with ASD ($n=54$) versus age-matched controls ($n=15$) using a method called solid-phase extraction (SPE). Targeted peptides were gliadinomorphin, beta-casomorphin, deltorphin 1, and deltorphin 2; however, no evidence for an excess of these were found in ASDs versus controls [61]. These data again refute the endogenous opioid hypothesis of ASD.

Inductively coupled plasma-mass spectrometry (ICP-MS) was used to examine whether children with autism possess excessive heavy metals, another theory of ASD that has been generally unsubstantiated [62, 63]. This type of MS is good for analyzing metals [64]. Urine from children with ASD ($n=15$) and controls ($n=4$) was tested for arsenic, cadmium, lead, and mercury during a "provoked excretion" induced by the oral chelator meso-2,3-dimercaptosuccinic acid (DMSA). Urinary levels of metals were compared to baseline. Three children with ASD excreted more metal than baseline during the provoked excretion, although two were near the limit of detection. A third child excreted mercury, which was corrected by a change in diet. This MS study overall did not support that heavy metals are found in children with ASD [65]. An MS study that analyzed the heavy metal content of hair did provide some support for the heavy metal theory. Mercury, lead, arsenic, and cadmium were measured in hair from children with ASD, 1–6 years old ($n=45$), compared with gender-, age-, and race-matched non-ASD children ($n=45$). Arsenic,

cadmium, and lead were detected at significantly lower levels in hair of children with ASD versus controls. Mercury was lower in hair of children with ASD, but this effect was not significant [66]. The authors speculated that children with ASD have more difficulty excreting toxic metals compared to non-ASD children [66].

The Childhood Autism Risk from Genetics and the Environment (CHARGE) Study is being conducted to address numerous susceptibility factors for autism, and their interactions, including the presence of heavy metals in people with ASD [67]. Inductively coupled MS was used to determine that mercury levels were unaltered in children 2–5 years old with ASD relative to controls in children [68] ($n=309$ and $n=143$, respectively). Plasma polybrominated diphenyl ethers (PBDEs) were measured in children with autism and developmental delay using GC–MS; however, no differences were observed [69]. ICP-MS was used to detect increased plasma levels of copper in 152 individuals with ASD relative to controls [70]. Taken together, MS studies have not demonstrated strong and consistent evidence for the introduction of either heavy metals or dietary opioids as having a causal influence in the etiology of ASD.

26.2.3 MS Analysis of Endogenous Proteins in ASDs

Current analyses using MS to study ASDs have focused on the search for endogenously produced biomarkers that could aid in diagnosis or understanding of ASDs. For example, Taurines et al. (2010) found protein differences in individuals with ASD and attention-deficit hyperactivity disorder (ADHD) ($n=9$ ASD+ADHD, ASD alone $n=7$) compared to age-matched controls ($n=12$). Three peaks with differences were observed; however, the actual proteins were not identified [71], although the investigators speculated that the protein may be an apolipoprotein [72]. Confirmation of elevations in apolipoprotein (apo) B-100 and apo A-IV was observed in a second study conducted in children with high functioning versus low-functioning autism. Upregulation in complement factor H-related protein, complement C1q and fibronectin 1, and apo B-100 was also measured in children with ASD compared to matched controls in the same study [73]. Confirming the possible presence of dysregulated complement proteins in ASD, a more recent MS analysis of blood plasma used surface-enhanced laser desorption/ionization-time-of-flight (SELDI-TOF) MS to examine peptides in plasma of children with ASD versus control subjects. Increases in the peaks of three peptides were measured, corresponding to fragments of C3 complement protein [74, 75]. Further proteomic studies will hopefully help confirm the existence of complement protein and apolipoprotein disturbances in ASDs.

Our group has recently identified two-dimensional differential in-gel electrophoresis (2D-DIGE) as a powerful technique for the differential investigation of the salivary proteome of ASD subjects and matched controls. DIGE is mostly used in protein expression profiling experiments of at least two samples or conditions allowing the determination of the relative abundance of proteins. In 2D-DIGE, protein samples are primarily labeled with spectrally resolvable fluorescent cyanine dyes (Cy2, 3, and 5) prior to 2D electrophoresis [76]. Thereafter, the gels are scanned for visualization of the protein bands. This method is very accurate and

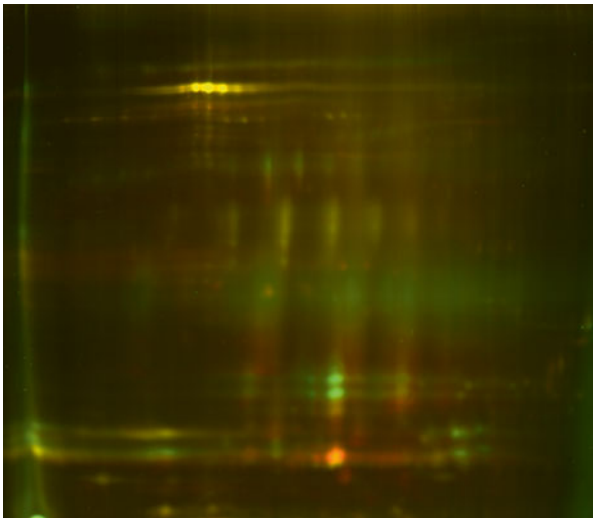


Fig. 26.1 Difference gel electrophoresis (DIGE). In DIGE (analytical gel), protein samples are primarily labeled with spectrally resolvable fluorescent cyanine dyes (Cy2, 3, and 5) prior to 2D electrophoresis. Thereafter, the gels are scanned for visualization of the protein bands. Many of the proteins are differentially expressed between the two conditions as many of the spots on the gel are either primarily *green* or *red* and very few are *yellow*. Here were analyzed saliva from children with ASD (Cy3 or *green*) and matched controls (Cy5 or *red*)

sensitive, reproducible, and has a broad dynamic range (over five orders of magnitude) [77, 78]. In our 2D-DIGE experiments, minimal labeling chemistry (labeling of a single lysine per protein using *N*-hydroxysuccinimidyl ester) was undertaken with Cy3 and Cy5 CyDye DIGE fluors (Fig. 26.1) and Cy2 CyDye DIGE fluor was used to label a pooled sample of equal amounts of each sample (internal standard, data not shown), according to published procedures [79–81]. The gel in Fig. 26.1 shows that many of the proteins are differentially expressed between the two conditions as many of the spots on the gel are either primarily green or red and very few are yellow. Based on the statistical calculations of the analysis of variance (ANOVA) spots with significant differences between the ASD sample and control subject were picked for nano-liquid chromatography–tandem mass spectrometry (nano-LC–MS/MS) and MALDI-MS/MS (Fig. 26.2) [11, 82–84]. One protein that was found to be elevated in ASD patients is cystatin-S which was identified in spot 394 (Fig. 26.2). An example of an MS/MS spectrum of the peptide that led to the identification of this protein in our nano-LC–MS/MS is shown in Fig. 26.3a. Another protein that was found to be upregulated in our MALDI-MS/MS experiments was protein S100-A9 also called migration inhibitory factor-related protein 14 (MRP-14) or calgranulin-B. The y- and b-ion series that led to the identification of this protein are shown (Fig. 26.3b). These findings are novel as these proteins have not been associated before with ASD and should therefore be further investigated.

When one looks to identify meaningful differences between ASD and controls (or any other disease or disorder and controls), one must also take into

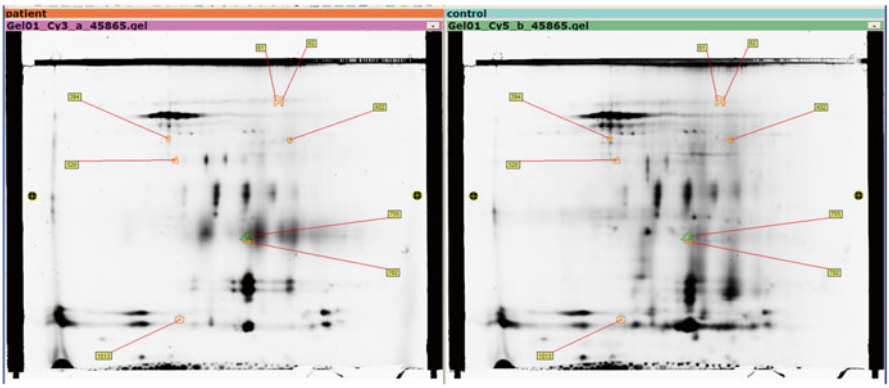


Fig. 26.2 DIGE (preparative gel). The preparative gels allow protein spots to be picked and then digested by trypsin and the peptides extracted and further analyzed by MALDI-MS or nano-LC–MS/MS. Shown are spots marked to be extracted that are specific to ASD (*left*) or control (*right*) samples

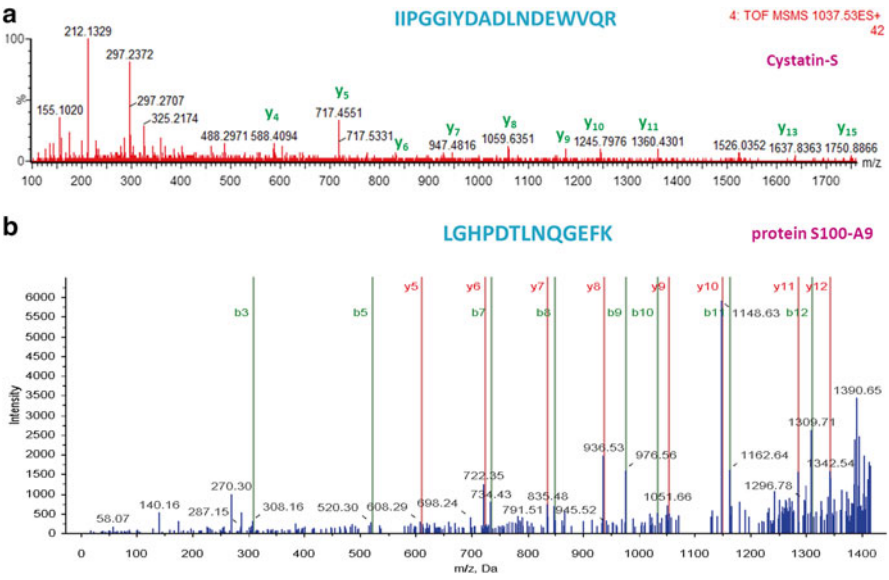


Fig. 26.3 Nano-LC–MS/MS of DIGE spots. (**a**) One protein that was found to be elevated in ASD patients is cystatin-S which was identified in spot 394 (shown in Fig. 26.2). Analysis of the MS/MS spectrum of a precursor peak with *m/z* of 1,037.53 (2+) led to identification of a peptide with the sequence IIPGGIYDADLNDEWVQR, that led to the identification of cystatin-S. (**b**) Another protein that was found to be upregulated in our DIGE and identified by MALDI-MS/MS experiments was protein S100-A9 also called migration inhibitory factor-related protein 14 (MRP-14) or calgranulin-B. Analysis of the y- and b-ion series shown in this MS/MS spectrum led to the identification of a peptide with the sequence LGHPDTLNQGEFK which is part of this protein

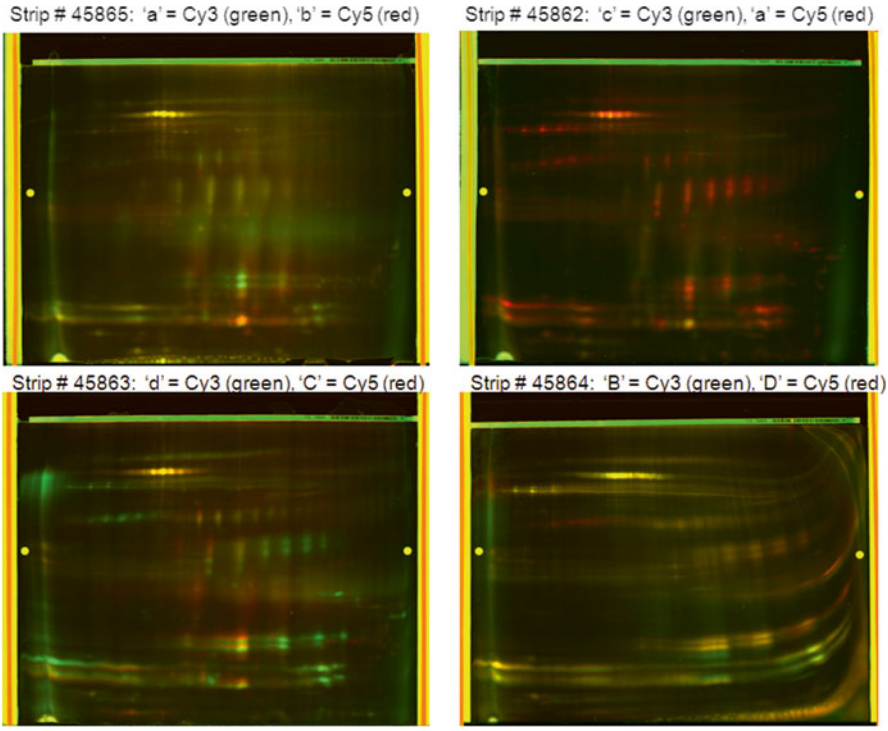


Fig. 26.4 DIGE of saliva from children with ASD and matched controls to investigate the circadian rhythmicity; a = ASD AM, b = controls AM, c = ASD PM, d = ASD PM. *AM* morning, *PM* afternoon

account the time when a sample is collected and analyzed. Historically, the blood and urine samples are collected in the hospitals in the morning. However, not much is known regarding the time of the day for when saliva should be collected. Therefore, such investigations must be performed prior choosing the time of collection during the day. An example of such comparison is where saliva from children with ASD and matched controls were compared to identify the circadian rhythmicity. For this comparison, the saliva was collected in the morning (AM) or afternoon (PM). These saliva, as analyzed by DIGE, are shown in Figs. 26.4 and 26.5.

26.2.4 MS-Based Metabolomic and Small Molecule Analysis in ASD

Urinary analysis of endogenous metabolites has been conducted in individuals with ASD using MS. LC–MS/MS was used to analyze the urine of young adults with severe ASD and with schizophrenia versus controls ($n = 15$, 18, and 18, respectively).

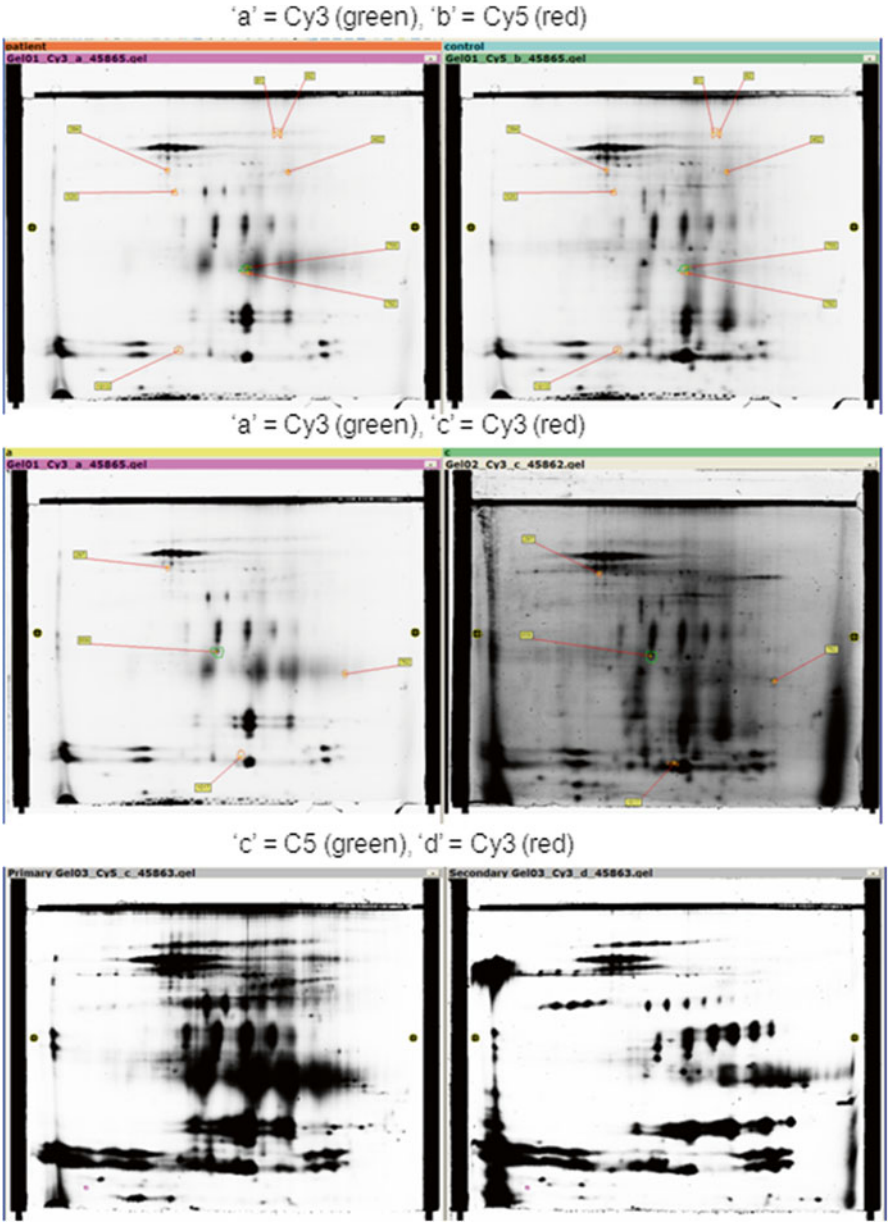


Fig. 26.5 DIGE of saliva from children with ASD and matched controls to investigate the circadian rhythmicity; a = ASD AM, b = controls AM, c = ASD PM, d = ASD PM. *AM* morning, *PM* afternoon

Butofenine, a molecule related to serotonin, was increased relative to controls in individuals with ASD and schizophrenia and correlated with scores of hyperactivity in these subjects [85]. This finding is consistent with serotonin abnormalities reported in ASD in several MS studies [58, 86, 87].

Oxidative stress, the generation of reactive oxygen species that cause cellular and tissue damage [88], may contribute to ASD etiology [89], and this has been supported by studies using MS. GC–MS identified higher urinary markers for lipid peroxidation and endothelium activation in individuals with ASDs ($n=26$) versus controls ($n=12$). Lipid peroxidation correlated with platelet and endothelium activation [90]. MS analyses have supported that oxidative stress and dysregulation of serotonin may be present in ASD.

26.2.5 Posttranslational Modification Analysis in ASD

In addition to providing information about protein levels, identity, and differences in protein signatures, MS-based proteomics may also provide information about PTMs. One PTM that has been associated with disease processes is phosphorylation. For example, hyper-phosphorylated tau has been measured in the saliva of subjects with Alzheimer's disease relative to controls [91]. HPLC–ESI-ion trap-MS or high performance liquid chromatography–electrospray ionization-ion trap-mass spectrometry was used to measure salivary peptides in individuals with ASDs ($n=27$) compared to age-matched controls. Hypo-phosphorylation of statherin, histatin 1, and acidic proline-rich proteins (entire and truncated isoforms) was significantly lower in subjects with ASDs than in controls. Lower phosphorylation of at least one peptide occurred in 66 % of ASD subjects [92].

Overall, proteomic evidence uncovered by MS supports a role for altered complement proteins and apolipoproteins in ASDs. These proteins may form the basis for a biomarker signature, aiding in future biological diagnosis of ASD. PTMs could be an additional piece of information to compose such a biomarker signature, although more MS studies of human ASD biomaterials are needed.

26.3 MS Analysis in Fragile X Syndrome

Fragile X syndrome (FXS) is caused by a mutation in the fragile X mental retardation (FMR1) gene and reduction or the absence of fragile X mental retardation protein (FMRP). This causes many CNS proteins to function abnormally [93, 94]. People with FXS can have cognitive deficits, seizures, autism, and ADHD [95]. FXS is a cause of 5 % of autism cases [96]. Individuals with FXS do not always develop ASDs, only between 15 and 60 % for reasons that are not well understood [96–99]. No medications are approved for FXS but clinical trials have been conducted for arbaclofen [100] and minocycline [101, 102] to treat FXS symptoms, and more drugs for FXS are in the pipeline [103]. Biological endpoints for FXS clinical trials are greatly needed [104], and would supplement behavioral endpoints. MS-based proteomics could aid in the search for such biological endpoints; however, MS-based proteomic studies to date have concentrated largely on animal models, rather than on human biomaterials [1].

Proteomic studies in FXS animal models have provided a plethora of proteins potentially altered in [26, 105–109]. For example, in mouse cerebellar cells transfected to produce an FXS model, 56 proteins were differentially expressed including 16 proteins with decreased levels and 40 with increased levels, relative to non-transfected cells. SILAC was used to examine primary neuronal culture in an FXS knock-out mouse. More than 100 proteins associated with nervous system function were found to have altered levels [106]. SILAC was also used to analyze an FXS drosophila model, in which alterations in actin-binding protein profilin and microtubulin-associated protein futsch were observed [110]. FXS drosophila models utilizing MS protein identification have also identified changes in monoamine-associated enzymes phenylalanine hydroxylase (Henna) and GTP cyclohydrolase (Punch) [108], and changes in Hsp60B-, Hsp68-, and Hsp90-related protein TRAP1 [109]. In an embryonic FXS drosophila model, 28 proteins were altered, including CCT subunits [107]. In FXS knock-out mice iTRAQ was used to measure changes in 23 hippocampal synapse proteins, including increase in Basp1, Gap43, and Cend1 [26]. Although these protein changes are interesting, they seem to vary depending on the model used, and would be most useful if confirmed in humans.

Human studies using MS to analyze FXS biomaterials have been somewhat limited, but could be useful, particularly in light of the current need for biological endpoints for FXS clinical trials [104]. Acetylation defects of annexin-A1 were found via MS in leukocytes taken from individuals with FXS [111]. MALDI-TOF MS was used to analyze methylation in blood of individuals for FXS and was found to be superior to Southern blotting, the current gold standard for FXS diagnosis. Differential methylation of two fragile X-related epigenetic elements within the fragile X mental retardation gene intron 1 (FMR1) was related to depletion of FMRP expression [112]. MALDI-TOF MS was used to examine FMR1 intron 1 methylation in blood of 62 premutation (between 55 and 200 CGG repeats in FMR1 gene) or 18 full mutation (more than 200 CGG repeats) females versus controls ($n=74$). FMR1 intron 1 hyper-methylation predicted verbal cognitive impairment [113]. This study is of particular interest since an MS-based biochemical measurement was paired with a behavioral test. This interdisciplinary approach may increase for assessment of neurodevelopmental disorders in the future, particularly as MS-based proteomics become more widely used.

26.4 MS Analysis in Smith–Lemli–Opitz Syndrome

Smith–Lemli–Opitz syndrome (SLOS) is genetically inherited decrease in cholesterol synthesis-associated symptoms of ASD, intellectual/learning problems, and numerous physical problems [114–117]. Total or partial deficiency of *Dhcr7* is the cause of SLOS [118]. This causes tissue cholesterol and total sterol levels to be substantially depleted, and 7-dehydrocholesterol levels (7DHC) to increase. High levels of 7DHC inhibit *Hmgcr* further causing deficits in cellular cholesterol [119]. Prenatally, SLOS is predicted using ultrasonography based on malformations and intrauterine lack of

growth as well as reduced maternal free estriol in serum. Sterol analysis is used to confirm the diagnosis in amniotic fluid or chorionic villus biopsy [118]. SLOS symptoms improve incompletely with cholesterol supplementation. In addition, there are individuals with ASD without SLOS with low total cholesterol [117].

MS studies focusing on SLOS and SLOS animal models have focused on the analysis of sterols, specifically 7DHC and cholesterol [120–124]. LC–MS/MS was used to analyze amniotic fluid in 6 SLOS subjects versus controls, to measure differences in 7DHC to cholesterol ratios between the two groups, in a proof-of-concept study focusing on prenatal screening [121]. Atmospheric pressure thermal desorption chemical ionization mass spectrometry (APTDCI-MS) has been used to differentiate normal from SLOS dried blood spots, specifically by analyzing cholesterol, 7DHC, and cholesteryl stearate, for the purpose of SLOS screening [123]. Nanostructure-initiator mass spectrometry (NIMS) has been utilized to image brain sterol metabolites in a mouse model of SLOS, which identified both 7DHC and cholesterol localized to abnormally developing brain structures (cerebellum and brainstem) in a non-diffuse pattern that was distinct from normally developing mice [124]. Gas chromatography–MS was used to determine sterol profiles (7DHC/cholesterol ratio) in plasma and erythrocyte membranes of 276 unaffected (reference values) and 16 SLOS subjects. The authors suggested that this method could be used to determine cholesterol biosynthesis defects [120].

Recently the first MS proteomic study in an SLOS model (rat retina) was reported [125]. This rat model exhibits some of the characteristics of SLOS, specifically retinal degeneration and visual impairment, and is generated by treatment with AY9944, a drug that inhibits DHCR7 (3beta-hydroxysterol-delta7-reductase), the same enzyme that is defective in SLOS [126]. Retinas from the SLOS rat model ($n=5$) were compared to retinas from age- and sex-matched control rats ($n=5$), via nano-RPLC (reverse-phase liquid chromatography). Statistically significant differences in 101 proteins were measured, including proteins that regulate lipid metabolism, oxidative stress, vision, proteolysis, cell death, and vesicular/membrane transport. Western blot and immunohistochemistry were used to validate specific protein targets. Cathepsin D, glial fibrillary acidic protein, Stat3, and μ-crystallin were elevated in the SLOS model retina, while ApoE was decreased.

Further studies focusing on proteomics using MS may supplement the reports describing sterol levels. For example, it may be of interest to analyze apolipoprotein levels in individuals with SLOS. Protein targets may also be utilized in therapeutic approaches for SLOS, which is currently inadequately treated.

26.5 Conclusions

In conclusion, MS is a useful tool for the analysis of human biomaterials, and may be employed for further study of neurodevelopmental disorders. For ASD, MS analysis directed at proteomic biomarker discovery may provide tools for ASD identification. This could provide the basis for a biological diagnostic test for ASD, since

no such test currently exists. Biomarker discovery in ASD may also reveal methods for treatment monitoring, and can help further elucidate the causes and biological consequences of this disorder. FXS and SLOS, in contrast to ASD, can be diagnosed via genotyping. As reviewed, MS-based proteomics methods have also been tested for potential diagnosis in these syndromes. For these disorders, MS-based analyses can be employed in biomarker identification as well, directed towards protein target discovery, for understanding the consequences of these disorders and for monitoring the effects of medications. MS-based research in neurodevelopmental disorders is currently a field in its infancy. MS as a tool holds great potential to further our understanding of these disorders.

Acknowledgments The authors would like to thank Dr. Witold Winnik (Environment Protection Agency, Research Triangle Park, Durham, NC) for the MALDI-MS analysis of the DIGE spots. The authors would like to thank Ms. Laura Mulderig, Scott Nichols, and their colleagues (Waters Corporation) for their generous support in setting up the Proteomics Center at Clarkson University. CCD thanks Drs. Thomas A. Neubert (New York University), Belinda Willard (Cleveland Clinic), and Gregory Wolber and David Mclaughin (Eastman Kodak Company) for donation of a TofSpec2E MALDI-MS (each). CCD thanks his advisors, Vlad Artenie, Wolfgang Haehnel, Paul M. Wassarman, and Thomas A. Neubert, for their advice and support. This work was supported in part by the Keep a Breast Foundation (KEABF-375-35054), the Redcay Foundation (SUNY Plattsburgh), the Alexander von Humboldt Foundation, SciFund Challenge, private donations (Spring Beck Richardson and Gillaume Mimoun), the David A. Walsh fellowship, and by the U.S. Army research office (DURIP grant #W911NF-11-1-0304).

References

1. Wormwood K, Sokolowska I, Ryan JP, Russell S, Darie CC, Woods AG (2013) The potential for proteomics in understanding neurodevelopmental disorders. J Proteom Bioinform S5
2. Junaid MA, Pullarkat RK (2001) Proteomic approach for the elucidation of biological defects in autism. J Autism Dev Disord 31:557–560
3. Anderson L, Seilhamer J (1997) A comparison of selected mRNA and protein abundances in human liver. Electrophoresis 18:533–537
4. Darie C (2013) Post-translational modification (PTM) proteomics: challenges and perspectives. Mod Chem Appl 1:e114
5. Darie CC, Shetty V, Spellman DS, Zhang G, Xu C, Cardasis HL, Blais S, Fenyo D, Neubert TA (2008) Blue native PAGE and mass spectrometry analysis of the ephrin stimulation-dependent protein-protein interactions in NG108-EphB2 cells. Springer, Düsseldorf, Germany
6. Darie CC, Deinhardt K, Zhang G, Cardasis HS, Chao MV, Neubert TA (2011) Identifying transient protein-protein interactions in EphB2 signaling by blue native PAGE and mass spectrometry. Proteomics 11:4514–4528
7. Ngounou Wetie AG, Sokolowska I, Woods AG, Darie CC (2013) Identification of post-translational modifications by mass spectrometry. Aust J Chem 66:734–748
8. Ngounou Wetie AG, Sokolowska I, Woods AG, Roy U, Deinhardt K, Darie CC (2014) Protein-protein interactions: switch from classical methods to proteomics and bioinformatics-based approaches. Cell Mol Life Sci 71:205–228
9. Ngounou Wetie AG, Sokolowska I, Woods AG, Roy U, Loo JA, Darie CC (2012) Investigation of stable and transient protein-protein interactions: past, present, and future. Proteomics 13:538–557

10. Ngounou Wetie AG, Sokolowska I, Woods AG, Wormwood KL, Dao S, Patel S, Clarkson BD, Darie CC (2013) Automated mass spectrometry-based functional assay for the routine analysis of the secretome. J Lab Autom 18:19–29
11. Sokolowska I, Dorobantu C, Woods AG, Macovei A, Branza-Nichita N, Darie CC (2012) Proteomic analysis of plasma membranes isolated from undifferentiated and differentiated HepaRG cells. Proteome Sci 10:47
12. Sokolowska I, Gawinowicz MA, Ngounou Wetie AG, Darie CC (2012) Disulfide proteomics for identification of extracellular or secreted proteins. Electrophoresis 33:2527–2536
13. Sokolowska I, Ngounou Wetie AG, Roy U, Woods AG, Darie CC (2013) Mass spectrometry investigation of glycosylation on the NXS/T sites in recombinant glycoproteins. Biochim Biophys Acta 1834:1474–1483
14. Sokolowska I, Ngounou Wetie AG, Woods AG, Darie CC (2012) Automatic determination of disulfide bridges in proteins. J Lab Autom 17:408–416
15. Sokolowska I, Ngounou Wetie AG, Wormwood K, Michel TM, Thome J, Darie CC, Woods AG (2013) The potential of biomarkers in psychiatry: focus on proteomics. J Neural Transm 1–10
16. Woods AG, Sokolowska I, Darie CC (2012) Identification of consistent alkylation of cysteine-less peptides in a proteomics experiment. Biochem Biophys Res Commun 419:305–308
17. Woods AG, Sokolowska I, Yakubu R, Butkiewicz M, LaFleur M, Talbot C, Darie CC (2011) Blue native page and mass spectrometry as an approach for the investigation of stable and transient protein-protein interactions. In: Andreescu S, Hepel M (eds) Oxidative stress: diagnostics, prevention, and therapy. American Chemical Society, Washington, DC
18. Darie C (2013) Mass spectrometry and proteomics: principle, workflow, challenges and perspectives. Mod Chem Appl 1:e105
19. Darie CC, Litscher ES, Wassarman PM (2008) Structure, processing, and polymerization of rainbow trout egg vitelline envelope proteins. Springer, Düsseldorf, Germany
20. Darie CC (2013) Mass spectrometry and its application in life sciences. Aust J Chem 66:1–2
21. Ngounou Wetie AG, Sokolowska I, Wormwood K, Michel TM, Thome J, Darie CC, Woods AG (2013) Mass spectrometry for the detection of potential psychiatric biomarkers. J Mol Psychiatry 1:8
22. Sokolowska I, Ngounou Wetie AG, Woods AG, Darie CC (2013) Applications of mass spectrometry in proteomics. Aust J Chem 66:721–733
23. Woods AG, Sokolowska I, Taurines R, Gerlach M, Dudley E, Thome J, Darie CC (2012) Potential biomarkers in psychiatry: focus on the cholesterol system. J Cell Mol Med 16:1184–1195
24. Woods AG, Ngounou Wetie AG, Sokolowska I, Russell S, Ryan JP, Michel TM, Thome J, Darie CC (2013) Mass spectrometry as a tool for studying autism spectrum disorder. J Mol Psychiatry 1:6
25. Thomas RH, Foley KA, Mepham JR, Tichenoff LJ, Possmayer F, MacFabe DF (2010) Altered brain phospholipid and acylcarnitine profiles in propionic acid infused rodents: further development of a potential model of autism spectrum disorders. J Neurochem 113:515–529
26. Klemmer P, Meredith RM, Holmgren CD, Klychnikov OI, Stahl-Zeng J, Loos M, van der Schors RC, Wortel J, de Wit H, Spijker S, Rotaru DC, Mansvelder HD, Smit AB, Li KW (2011) Proteomics, ultrastructure, and physiology of hippocampal synapses in a fragile X syndrome mouse model reveal presynaptic phenotype. J Biol Chem 286:25495–25504
27. Lista S, Faltraco F, Prvulovic D, Hampel H (2013) Blood and plasma-based proteomic biomarker research in Alzheimer's disease. Prog Neurobiol 101–102:1–17
28. Shi T, Su D, Liu T, Tang K, Camp DG II, Qian WJ, Smith RD (2012) Advancing the sensitivity of selected reaction monitoring-based targeted quantitative proteomics. Proteomics 12:1074–1092
29. Stalmach A, Albalat A, Mullen W, Mischak H (2013) Recent advances in capillary electrophoresis coupled to mass spectrometry for clinical proteomic applications. Electrophoresis 34:1452–1464

30. Amado FM, Ferreira RP, Vitorino R (2013) One decade of salivary proteomics: current approaches and outstanding challenges. Clin Biochem 46:506–517
31. Righetti PG, Boschetti E, Candiano G (2012) Mark Twain: how to fathom the depth of your pet proteome. J Proteome 75:4783–4791
32. Hu S, Li Y, Wang J, Xie Y, Tjon K, Wolinsky L, Loo RR, Loo JA, Wong DT (2006) Human saliva proteome and transcriptome. J Dent Res 85:1129–1133
33. APA (2013) Diagnostic and statistical manual of mental disorders, 5th edn. American Psychiatric Association, Arlington, VA
34. Woods AG, Mahdavi E, Ryan JP (2013) Treating clients with Asperger's syndrome and autism. Child Adolesc Psychiatry Ment Health 7:32
35. APA (2000) Diagnostic and statistical manual of mental disorders (4th text revision). American Psychiatric Association, Arlington, VA
36. Autism and Developmental Disabilities Monitoring Network Surveillance Year Principal Investigators, Centers for Disease Control, and Prevention (2012) Prevalence of autism spectrum disorders–Autism and Developmental Disabilities Monitoring Network, 14 sites, United States, 2008. MMWR Surveill Summ 61:1–19
37. Blumberg SJ, Bramlett MD, Kogan MD, Schieve LA, Jones JR, Lu MC (2013) Changes in prevalence of parent-reported autism spectrum disorder in school-aged U.S. children: 2007 to 2011–2012. National Health Statistics reports; no 65. (Department of Health and Human Services). Centers for Disease Control and Prevention, National Center for Health Statistics, Hyattsville, MD
38. Zhu L, Wang X, Li XL, Towers A, Cao X, Wang P, Bowman R, Yang H, Goldstein J, Li YJ, Jiang YH (2014) Epigenetic dysregulation of SHANK3 in brain tissues from individuals with autism spectrum disorders. Hum Mol Genet 23:1563–1578
39. Volk HE, Kerin T, Lurmann F, Hertz-Picciotto I, McConnell R, Campbell DB (2013) Autism spectrum disorder: interaction of air pollution with the MET receptor tyrosine kinase gene. Epidemiology 25:44–47
40. State MW, Levitt P (2011) The conundrums of understanding genetic risks for autism spectrum disorders. Nat Neurosci 14:1499–1506
41. Bailey A, Le Couteur A, Gottesman I, Bolton P, Simonoff E, Yuzda E, Rutter M (1995) Autism as a strongly genetic disorder: evidence from a British twin study. Psychol Med 25:63–77
42. Le Couteur A, Bailey A, Goode S, Pickles A, Robertson S, Gottesman I, Rutter M (1996) A broader phenotype of autism: the clinical spectrum in twins. J Child Psychol Psychiatry 37:785–801
43. Betancur C (2011) Etiological heterogeneity in autism spectrum disorders: more than 100 genetic and genomic disorders and still counting. Brain Res 1380:42–77
44. Nickl-Jockschat T, Michel TM (2011) Genetic and brain structure anomalies in autism spectrum disorders. Towards an understanding of the aetiopathogenesis? Nervenarzt 82:618–627
45. Neale BM, Kou Y, Liu L, Ma'ayan A, Samocha KE, Sabo A, Lin CF, Stevens C, Wang LS, Makarov V, Polak P, Yoon S, Maguire J, Crawford EL, Campbell NG, Geller ET, Valladares O, Schafer C, Liu H, Zhao T, Cai G, Lihm J, Dannenfelser R, Jabado O, Peralta Z, Nagaswamy U, Muzny D, Reid JG, Newsham I, Wu Y, Lewis L, Han Y, Voight BF, Lim E, Rossin E, Kirby A, Flannick J, Fromer M, Shakir K, Fennell T, Garimella K, Banks E, Poplin R, Gabriel S, DePristo M, Wimbish JR, Boone BE, Levy SE, Betancur C, Sunyaev S, Boerwinkle E, Buxbaum JD, Cook EH Jr, Devlin B, Gibbs RA, Roeder K, Schellenberg GD, Sutcliffe JS, Daly MJ (2012) Patterns and rates of exonic de novo mutations in autism spectrum disorders. Nature 485:242–245
46. Leblond CS, Heinrich J, Delorme R, Proepper C, Betancur C, Huguet G, Konyukh M, Chaste P, Ey E, Rastam M, Anckarsater H, Nygren G, Gillberg IC, Melke J, Toro R, Regnault B, Fauchereau F, Mercati O, Lemiere N, Skuse D, Poot M, Holt R, Monaco AP, Jarvela I, Kantojarvi K, Vanhala R, Curran S, Collier DA, Bolton P, Chiocchetti A, Klauck SM, Poustka F, Freitag CM, Waltes R, Kopp M, Duketis E, Bacchelli E, Minopoli F, Ruta L, Battaglia A, Mazzone L, Maestrini E, Sequeira AF, Oliveira B, Vicente A, Oliveira G, Pinto D, Scherer SW, Zelenika D, Delepine M, Lathrop M, Bonneau D, Guinchat V, Devillard F, Assouline B, Mouren MC, Leboyer M, Gillberg C, Boeckers TM, Bourgeron T (2012) Genetic and func-

tional analyses of SHANK2 mutations suggest a multiple hit model of autism spectrum disorders. PLoS Genet 8:e1002521
47. Eapen V, Crncec R, Walter A (2013) Exploring links between genotypes, phenotypes, and clinical predictors of response to early intensive behavioral intervention in autism spectrum disorder. Front Hum Neurosci 7:567
48. Lanz TA, Guilmette E, Gosink MM, Fischer JE, Fitzgerald LW, Stephenson DT, Pletcher MT (2013) Transcriptomic analysis of genetically defined autism candidate genes reveals common mechanisms of action. Mol Autism 4:45
49. Gurrieri F (2012) Working up autism: the practical role of medical genetics. Am J Med Genet C Semin Med Genet 160C:104–110
50. Lee TL, Raygada MJ, Rennert OM (2012) Integrative gene network analysis provides novel regulatory relationships, genetic contributions and susceptible targets in autism spectrum disorders. Gene 496:88–96
51. Carter MT, Scherer SW (2013) Autism spectrum disorder in the genetics clinic: a review. Clin Genet 83:399–407
52. Nickl-Jockschat T, Michel TM (2011) The role of neurotrophic factors in autism. Mol Psychiatry 16:478–490
53. Theoharides TC, Kempuraj D, Redwood L (2009) Autism: an emerging 'neuroimmune disorder' in search of therapy. Expert Opin Pharmacother 10:2127–2143
54. Veenstra-VanderWeele J, Blakely RD (2012) Networking in autism: leveraging genetic, biomarker and model system findings in the search for new treatments. Neuropsychopharmacology 37:196–212
55. Lord C, Risi S, Lambrecht L, Cook EH, Leventhal BL, DiLavore PC, Pickles A, Rutter M (2000) The autism diagnostic observation schedule–generic: a standard measure of social and communication deficits associated with the spectrum of autism. J Autism Dev Disord 30:205–223
56. Lord C, Risi S, DiLavore PS, Shulman C, Thurm A, Pickles A (2006) Autism from 2 to 9 years of age. Arch Gen Psychiatry 63:694–701
57. Dereu M, Roeyers H, Raymaekers R, Meirsschaut M, Warreyn P (2012) How useful are screening instruments for toddlers to predict outcome at age 4? General development, language skills, and symptom severity in children with a false positive screen for autism spectrum disorder. Eur Child Adolesc Psychiatry 21:541–551
58. Lam KS, Aman MG, Arnold LE (2006) Neurochemical correlates of autistic disorder: a review of the literature. Res Dev Disabil 27:254–289
59. Cass H, Gringras P, March J, McKendrick I, O'Hare AE, Owen L, Pollin C (2008) Absence of urinary opioid peptides in children with autism. Arch Dis Child 93:745–750
60. Hunter LC, O'Hare A, Herron WJ, Fisher LA, Jones GE (2003) Opioid peptides and dipeptidyl peptidase in autism. Dev Med Child Neurol 45:121–128
61. Dettmer K, Hanna D, Whetstone P, Hansen R, Hammock BD (2007) Autism and urinary exogenous neuropeptides: development of an on-line SPE-HPLC-tandem mass spectrometry method to test the opioid excess theory. Anal Bioanal Chem 388:1643–1651
62. Albizzati A, More L, Di Candia D, Saccani M, Lenti C (2012) Normal concentrations of heavy metals in autistic spectrum disorders. Minerva Pediatr 64:27–31
63. Owens SE, Summar ML, Ryckman KK, Haines JL, Reiss S, Summar SR, Aschner M (2011) Lack of association between autism and four heavy metal regulatory genes. Neurotoxicology 32:769–775
64. Montaser A, McLean JA, Liu H, Mermet J-M (1998) An introduction to ICP spectrometries for elemental analysis. In: Montaser A (ed) Plasma mass spectrometry. Wiley, Hoboken, NJ
65. Soden SE, Lowry JA, Garrison CB, Wasserman GS (2007) 24-hour provoked urine excretion test for heavy metals in children with autism and typically developing controls, a pilot study. Clin Toxicol 45:476–481
66. Kern JK, Grannemann BD, Trivedi MH, Adams JB (2007) Sulfhydryl-reactive metals in autism. J Toxic Environ Health A 70:715–721

67. Hertz-Picciotto I, Croen LA, Hansen R, Jones CR, van de Water J, Pessah IN (2006) The CHARGE study: an epidemiologic investigation of genetic and environmental factors contributing to autism. Environ Health Perspect 114:1119–1125
68. Hertz-Picciotto I, Green PG, Delwiche L, Hansen R, Walker C, Pessah IN (2010) Blood mercury concentrations in CHARGE Study children with and without autism. Environ Health Perspect 118:161–166
69. Hertz-Picciotto I, Bergman A, Fangstrom B, Rose M, Krakowiak P, Pessah I, Hansen R, Bennett DH (2011) Polybrominated diphenyl ethers in relation to autism and developmental delay: a case–control study. Environ Heal 10:1
70. Russo AJ, Devito R (2011) Analysis of copper and zinc plasma concentration and the efficacy of zinc therapy in individuals with Asperger's syndrome, pervasive developmental disorder not otherwise specified (PDD-NOS) and autism. Biomark Insights 6:127–133
71. Taurines R, Dudley E, Conner AC, Grassl J, Jans T, Guderian F, Mehler-Wex C, Warnke A, Gerlach M, Thome J (2010) Serum protein profiling and proteomics in autistic spectrum disorder using magnetic bead-assisted mass spectrometry. Eur Arch Psychiatry Clin Neurosci 260:249–255
72. Taurines R, Dudley E, Grassl J, Warnke A, Gerlach M, Coogan AN, Thome J (2011) Proteomic research in psychiatry. J Psychopharmacol 25:151–196
73. Corbett BA, Kantor AB, Schulman H, Walker WL, Lit L, Ashwood P, Rocke DM, Sharp FR (2007) A proteomic study of serum from children with autism showing differential expression of apolipoproteins and complement proteins. Mol Psychiatry 12:292–306
74. Momeni N, Brudin L, Behnia F, Nordstrom B, Yosefi-Oudarji A, Sivberg B, Joghataei MT, Persson BL (2012) High complement factor I activity in the plasma of children with autism spectrum disorders. Autism Res Treat 2012:868576
75. Momeni N, Bergquist J, Brudin L, Behnia F, Sivberg B, Joghataei MT, Persson BL (2012) A novel blood-based biomarker for detection of autism spectrum disorders. Transl Psychiatry 2:e91
76. Unlu M, Morgan ME, Minden JS (1997) Difference gel electrophoresis: a single gel method for detecting changes in protein extracts. Electrophoresis 18:2071–2077
77. Lilley KS, Razzaq A, Dupree P (2002) Two-dimensional gel electrophoresis: recent advances in sample preparation, detection and quantitation. Curr Opin Chem Biol 6:46–50
78. Tonge R, Shaw J, Middleton B, Rowlinson R, Rayner S, Young J, Pognan F, Hawkins E, Currie I, Davison M (2001) Validation and development of fluorescence two-dimensional differential gel electrophoresis proteomics technology. Proteomics 1:377–396
79. DeKroon RM, Osorio C, Robinette JB, Mocanu M, Winnik WM, Alzate O (2011) Simultaneous detection of changes in protein expression and oxidative modification as a function of age and APOE genotype. J Proteome Res 10:1632–1644
80. DeKroon RM, Robinette JB, Osorio C, Jeong JS, Hamlett E, Mocanu M, Alzate O (2012) Analysis of protein posttranslational modifications using DIGE-based proteomics. Methods Mol Biol 854:129–143
81. Londono C, DeKroon RM, Mocanu M, Booe J, Winnik WM, Swank A, Osorio C, Hamlett ED, Alzate O (2012) Proteomic analysis of mice expressing human ApoE demonstrates no differences in global protein solubility between APOE 3 and APOE 4 young mice. Electrophoresis 33:3745–3755
82. Petrareanu C, Macovei A, Sokolowska I, Woods AG, Lazar C, Radu GL, Darie CC, Branza-Nichita N (2013) Comparative proteomics reveals novel components at the plasma membrane of differentiated HepaRG cells and different distribution in hepatocyte-and biliary-like cells. PLoS One 8:e71859
83. Roy U, Sokolowska I, Woods AG, Darie CC (2012) Structural investigation of tumor differentiation factor. Biotechnol Appl Biochem 59:445–450
84. Winnik WM, Dekroon RM, Jeong JS, Mocanu M, Robinette JB, Osorio C, Dicheva NN, Hamlett E, Alzate O (2012) Analysis of proteins using DIGE and MALDI mass spectrometry. Methods Mol Biol 854:47–66
85. Emanuele E, Colombo R, Martinelli V, Brondino N, Marini M, Boso M, Barale F, Politi P (2010) Elevated urine levels of bufotenine in patients with autistic spectrum disorders and schizophrenia. Neuro Endocrinol Lett 31:117–121

86. Kałuzna-Czaplinska J, Michalska M, Rynkowski J (2010) Determination of tryptophan in urine of autistic and healthy children by gas chromatography/mass spectrometry. Med Sci Monit 16:CR488–CR492
87. Pedersen OS, Liu Y, Reichelt KL (1999) Serotonin uptake stimulating peptide found in plasma of normal individuals and in some autistic urines. J Pept Res 53:641–646
88. Sokolowska I, Woods AG, Wagner J, Dorler J, Wormwood K, Thome J, Darie CC (2011) Mass spectrometry for proteomics-based investigation of oxidative stress and heat shock proteins. In: Andreescu S, Hepel M (eds) Oxidative stress: diagnostics, prevention, and therapy. American Chemical Society, Washington, DC
89. Ghanizadeh A, Akhondzadeh S, Hormozi M, Makarem A, Abotorabi-Zarchi M, Firoozabadi A (2012) Glutathione-related factors and oxidative stress in autism, a review. Curr Med Chem 19:4000–4005
90. Yao Y, Walsh WJ, McGinnis WR, Pratico D (2006) Altered vascular phenotype in autism: correlation with oxidative stress. Arch Neurol 63:1161–1164
91. Shi M, Sui YT, Peskind ER, Li G, Hwang H, Devic I, Ginghina C, Edgar JS, Pan C, Goodlett DR, Furay AR, Gonzalez-Cuyar LF, Zhang J (2011) Salivary tau species are potential biomarkers of Alzheimer's disease. J Alzheimers Dis 27:299–305
92. Castagnola M, Messana I, Inzitari R, Fanali C, Cabras T, Morelli A, Pecoraro AM, Neri G, Torrioli MG, Gurrieri F (2008) Hypo-phosphorylation of salivary peptidome as a clue to the molecular pathogenesis of autism spectrum disorders. J Proteome Res 7:5327–5332
93. Brown W (2002) The molecular biology of fragile X mutation. In: Hagerman R, Hagerman PJ (eds) Fragile X syndrome: diagnosis, treatment, and research. Johns Hopkins University Press, Baltimore, MD, pp 110–135
94. De Rubeis S, Fernandez E, Buzzi A, Di Marino D, Bagni C (2012) Molecular and cellular aspects of mental retardation in the Fragile X syndrome: from gene mutation/s to spine dysmorphogenesis. Adv Exp Med Biol 970:517–551
95. Hagerman RJ, Berry-Kravis E, Kaufmann WE, Ono MY, Tartaglia N, Lachiewicz A, Kronk R, Delahunty C, Hessl D, Visootsak J, Picker J, Gane L, Tranfaglia M (2009) Advances in the treatment of fragile X syndrome. Pediatrics 123:378–390
96. McCary LM, Roberts JE (2012) Early identification of autism in fragile X syndrome: a review. J Intellect Disabil Res 57:803–814
97. Boyle L, Kaufmann WE (2010) The behavioral phenotype of FMR1 mutations. Am J Med Genet C: Semin Med Genet 154C:469–476
98. Hagerman R, Lauterborn J, Au J, Berry-Kravis E (2012) Fragile X syndrome and targeted treatment trials. Results Probl Cell Differ 54:297–335
99. Smith LE, Barker ET, Seltzer MM, Abbeduto L, Greenberg JS (2012) Behavioral phenotype of fragile X syndrome in adolescence and adulthood. Am J Intellect Dev Disabil 117:1–17
100. Berry-Kravis EM, Hessl D, Rathmell B, Zarevics P, Cherubini M, Walton-Bowen K, Mu Y, Nguyen DV, Gonzalez-Heydrich J, Wang PP, Carpenter RL, Bear MF, Hagerman RJ (2012) Effects of STX209 (arbaclofen) on neurobehavioral function in children and adults with fragile X syndrome: a randomized, controlled, phase 2 trial. Sci Transl Med 4:152ra127
101. Dziembowska M, Pretto DI, Janusz A, Kaczmarek L, Leigh MJ, Gabriel N, Durbin-Johnson B, Hagerman RJ, Tassone F (2013) High MMP-9 activity levels in fragile X syndrome are lowered by minocycline. Am J Med Genet A 161A:1897–1903
102. Leigh MJ, Nguyen DV, Mu Y, Winarni TI, Schneider A, Chechi T, Polussa J, Doucet P, Tassone F, Rivera SM, Hessl D, Hagerman RJ (2013) A randomized double-blind, placebo-controlled trial of minocycline in children and adolescents with fragile X syndrome. J Dev Behav Pediatr 34:147–155
103. Jacquemont S, Berry-Kravis E, Hagerman R, von Raison F, Gasparini F, Apostol G, Ufer M, Des Portes V, Gomez-Mancilla B (2014) The challenges of clinical trials in fragile X syndrome. Psychopharmacology 231:1237–1250
104. Berry-Kravis E, Hessl D, Abbeduto L, Reiss AL, Beckel-Mitchener A, Urv TK, Outcome Measures Working G (2013) Outcome measures for clinical trials in fragile X syndrome. J Dev Behav Pediatr 34:508–522

105. Darie CC, Biniossek ML, Gawinowicz MA, Milgrom Y, Thumfart JO, Jovine L, Litscher ES, Wassarman PM (2005) Mass spectrometric evidence that proteolytic processing of rainbow trout egg vitelline envelope proteins takes place on the egg. J Biol Chem 280:37585–37598
106. Liao L, Park SK, Xu T, Vanderklish P, Yates JR III (2008) Quantitative proteomic analysis of primary neurons reveals diverse changes in synaptic protein content in fmr1 knockout mice. Proc Natl Acad Sci U S A 105:15281–15286
107. Monzo K, Dowd SR, Minden JS, Sisson JC (2010) Proteomic analysis reveals CCT is a target of Fragile X mental retardation protein regulation in Drosophila. Dev Biol 340: 408–418
108. Zhang YQ, Friedman DB, Wang Z, Woodruff E III, Pan L, O'Donnell J, Broadie K (2005) Protein expression profiling of the drosophila fragile X mutant brain reveals up-regulation of monoamine synthesis. Mol Cell Proteomics 4:278–290
109. Zhang YQ, Matthies HJ, Mancuso J, Andrews HK, Woodruff E III, Friedman D, Broadie K (2004) The Drosophila fragile X-related gene regulates axoneme differentiation during spermatogenesis. Dev Biol 270:290–307
110. Xu P, Tan H, Duong DM, Yang Y, Kupsco J, Moberg KH, Li H, Jin P, Peng J (2012) Stable isotope labeling with amino acids in Drosophila for quantifying proteins and modifications. J Proteome Res 11:4403–4412
111. Kaufmann WE, Cohen S, Sun HT, Ho G (2002) Molecular phenotype of Fragile X syndrome: FMRP, FXRPs, and protein targets. Microsc Res Tech 57:135–144
112. Godler DE, Slater HR, Bui QM, Ono M, Gehling F, Francis D, Amor DJ, Hopper JL, Hagerman R, Loesch DZ (2011) FMR1 intron 1 methylation predicts FMRP expression in blood of female carriers of expanded FMR1 alleles. J Mol Diagn 13:528–536
113. Godler DE, Slater HR, Bui QM, Storey E, Ono MY, Gehling F, Inaba Y, Francis D, Hopper JL, Kinsella G, Amor DJ, Hagerman RJ, Loesch DZ (2012) Fragile X mental retardation 1 (FMR1) intron 1 methylation in blood predicts verbal cognitive impairment in female carriers of expanded FMR1 alleles: evidence from a pilot study. Clin Chem 58:590–598
114. Aneja A, Tierney E (2008) Autism: the role of cholesterol in treatment. Int Rev Psychiatry 20:165–170
115. Bukelis I, Porter FD, Zimmerman AW, Tierney E (2007) Smith-Lemli-Opitz syndrome and autism spectrum disorder. Am J Psychiatry 164:1655–1661
116. Diaz-Stransky A, Tierney E (2012) Cognitive and behavioral aspects of Smith-Lemli-Opitz syndrome. Am J Med Genet C: Semin Med Genet 160C:295–300
117. Tierney E, Bukelis I, Thompson RE, Ahmed K, Aneja A, Kratz L, Kelley RI (2006) Abnormalities of cholesterol metabolism in autism spectrum disorders. Am J Med Genet B Neuropsychiatr Genet 141B:666–668
118. DeBarber AE, Eroglu Y, Merkens LS, Pappu AS, Steiner RD (2011) Smith-Lemli-Opitz syndrome. Exp Rev Mol Med 13:e24
119. Fitzky BU, Moebius FF, Asaoka H, Waage-Baudet H, Xu L, Xu G, Maeda N, Kluckman K, Hiller S, Yu H, Batta AK, Shefer S, Chen T, Salen G, Sulik K, Simoni RD, Ness GC, Glossmann H, Patel SB, Tint GS (2001) 7-Dehydrocholesterol-dependent proteolysis of HMG-CoA reductase suppresses sterol biosynthesis in a mouse model of Smith-Lemli-Opitz/RSH syndrome. J Clin Invest 108:905–915
120. Corso G, Gelzo M, Barone R, Clericuzio S, Pianese P, Nappi A, Dello Russo A (2011) Sterol profiles in plasma and erythrocyte membranes in patients with Smith-Lemli-Opitz syndrome: a six-year experience. Clin Chem Lab Med 49:2039–2046
121. Griffiths WJ, Wang Y, Karu K, Samuel E, McDonnell S, Hornshaw M, Shackleton C (2008) Potential of sterol analysis by liquid chromatography-tandem mass spectrometry for the prenatal diagnosis of Smith-Lemli-Opitz syndrome. Clin Chem 54:1317–1324
122. Meljon A, Watson GL, Wang Y, Shackleton CH, Griffiths WJ (2013) Analysis by liquid chromatography-mass spectrometry of sterols and oxysterols in brain of the newborn Dhcr7(Delta3-5/T93M) mouse: a model of Smith-Lemli-Opitz syndrome. Biochem Pharmacol 86:43–55

123. Paglia G, D'Apolito O, Gelzo M, Dello Russo A, Corso G (2010) Direct analysis of sterols from dried plasma/blood spots by an atmospheric pressure thermal desorption chemical ionization mass spectrometry (APTDCI-MS) method for a rapid screening of Smith-Lemli-Opitz syndrome. Analyst 135:789–796
124. Patti GJ, Shriver LP, Wassif CA, Woo HK, Uritboonthai W, Apon J, Manchester M, Porter FD, Siuzdak G (2010) Nanostructure-initiator mass spectrometry (NIMS) imaging of brain cholesterol metabolites in Smith-Lemli-Opitz syndrome. Neuroscience 170:858–864
125. Tu C, Li J, Jiang X, Sheflin LG, Pfeffer BA, Behringer M, Fliesler SJ, Qu J (2013) Ion current based proteomic profiling of the retina in a rat model of Smith-Lemli-Opitz syndrome. Mol Cell Proteomics 12:3583–3598
126. Fliesler SJ (2010) Retinal degeneration in a rat model of Smith-Lemli-Opitz Syndrome: thinking beyond cholesterol deficiency. Adv Exp Med Biol 664:481–489

Chapter 27
Biomarkers in Major Depressive Disorder: The Role of Mass Spectrometry

Alisa G. Woods, Dan V. Iosifescu, and Costel C. Darie

Abstract Major depressive disorder (MDD) is common. Despite numerous available treatments, many individuals fail to improve clinically. MDD continues to be diagnosed exclusively via behavioral rather than biological methods. Biomarkers—which include measurements of genes, proteins, and patterns of brain activity—may provide an important objective tool for the diagnosis of MDD or in the rational selection of treatments. Proteomic analysis and validation of its results as biomarkers is less explored than other areas of biomarker research in MDD. Mass spectrometry (MS) is a comprehensive, unbiased means of proteomic analysis, which can be complemented by directed protein measurements, such as Western Blotting. Prior studies have focused on MS analysis of several human biomaterials in MDD, including human post-mortem brain, cerebrospinal fluid (CSF), blood components, and urine. Further studies utilizing MS and proteomic analysis in MDD may help solidify and establish biomarkers for use in diagnosis, identification of new treatment targets, and understanding of the disorder. The ultimate goal is the validation of a biomarker or a biomarker signature that facilitates a convenient and inexpensive predictive test for depression treatment response and helps clinicians in the rational selection of next-step treatments.

A.G. Woods (✉)
Biochemistry & Proteomics Group, Department of Chemistry & Biomolecular Science, Clarkson University, 8 Clarkson Avenue, Potsdam, NY 13699-5810, USA

SUNY Plattsburgh Neuropsychology Clinic and Psychoeducation Services, 101 Broad Street, Plattsburgh, NY 12901, USA
e-mail: awoods@clarkson.edu

D.V. Iosifescu
Mood Disorders and Anxiety Program, Mount Sinai Hospital, 1428 Madison Avenue, New York, NY 10029, USA

C.C. Darie
Biochemistry & Proteomics Group, Department of Chemistry & Biomolecular Science, Clarkson University, 8 Clarkson Avenue, Potsdam, NY 13699-5810, USA

A.G. Woods and C.C. Darie (eds.), *Advancements of Mass Spectrometry in Biomedical Research*, Advances in Experimental Medicine and Biology 806,
DOI 10.1007/978-3-319-06068-2_27, © Springer International Publishing Switzerland 2014

27.1 Introduction

Major depressive disorder (MDD) occurs frequently, affecting 7.1 % people each year and 14.4 % of people over the course of a lifetime [1]. According to the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5), the diagnosis of MDD requires at least five symptoms, present for at least 2 weeks; among the five symptoms at least one should be depressed mood and/or loss of interest or pleasure [2]. Other symptoms include change in sleep, appetite, fatigue/energy loss, feelings of worthlessness or guilt, diminished concentration, and suicidal thoughts [2, 3]. In addition to suffering from impairment in normal functioning, patients with MDD experience increased rates of other comorbid medical illnesses, including diabetes, arthritis, cardiovascular disease as well as comorbid psychiatric disorders [4–10].

Effective treatments for MDD are greatly needed, since MDD is very common and associated with high costs of health care as well as high morbidity and mortality [8, 9, 11–13]. Unfortunately, partial or suboptimal response to treatment occurs commonly in MDD, and only 30–40 % of patients receiving treatment will fully remit (achieving near-complete resolution of symptoms) [14–17]. The Sequenced Treatment Alternatives to Relieve Depression (STAR*D) study is one of the largest and more comprehensive studies of MDD treatments ever conducted. Responses to up to four successive treatment steps were studied in individuals with MDD. Remission was achieved in less than one third of patients receiving first- and second-line treatments; rates of remission were significantly lower in patients with treatment-resistant depression (TRD, i.e., those who have had failed two or more consecutive treatments) [14, 15]. The cumulative rate of remission was 67 % after all four levels. These sobering results indicate that the current treatment options are not very effective; even when effective treatment effects are very slow. Each level of treatment requires 12 weeks; a full year of treatments may be required in order to achieve significant positive outcomes (remission) in two thirds of the patients. This problem of treatment-resistant depression has led to the continued search for novel, more effective antidepressant treatments [18–25].

Given the current only partially effective antidepressant treatments and the imperfect, trial-and-error methods for treatment selection in MDD, predictive biomarkers could be particularly useful. Ideally, biomarkers predictive of treatment response to specific therapies could eliminate multiple and lengthy treatment steps by selecting the right treatment from the start.

The use of biomarkers to predict or identify a disease or disorder is an essential strategy in many areas of medicine, and is increasingly studied in psychiatric research. The United States Food and Drug Administration (FDA) defines a biomarker as an objective measurement of a normal biological process, a pathological, biological process or an objective measurement that indicates response to a therapeutic [26]. Biomarkers are commonly used for diseases such as cancer [27–29], however, no biomarker is routinely used clinically in psychiatry, despite the assumption that biological changes underlie or contribute to many psychiatric problems [30, 31].

Biomarkers could have several functions in psychiatry, including clarification of the etiology of psychiatric problems, and could be used as predictors of therapeutic response [32].

Many different biological parameters may potentially serve as biomarkers, such as specific genetic or epigenetic abnormalities, metabolites, proteins, or brain activity patterns. For example, specific changes in theta power on prefrontal leads in electroencephalograms (EEG) from baseline to week 1 of treatment have been used to predict antidepressant treatment outcomes at week 8 [33]. Moreover, EEG data has been shown to be specific enough to guide clinical treatment in patients with MDD [34]. Both resting-state EEGs as well as evoked potentials have been explored as biomarkers predictive of antidepressant treatment response [35]. EEGs may provide a useful tool for aiding in antidepressant therapy however access to the technology and imperfect predictive ability have limited its development so far. It is likely that, similar to other areas of medicine, biomarkers useful in psychiatry will involve a composite of clinical and biological factors [35]; proteins found in human biomaterials could play a significant role.

Brain neuroanatomical changes, as measured by MR morphometry, diffusion tensor imaging (DTI), or even functional MRI may also provide depression biomarkers. For example, DTI studies have suggested that white matter abnormalities may be present in treatment-resistant depression [36]. DTI has also been used to indicate that MDD may be associated with abnormal microstructure in brain reward/aversion regions, specifically the ventral tegmental area and dorsolateral prefrontal white matter [37]. While neuroimaging provides the potential to understand the impairment in neurocircuitry underlying TRD, it is unclear whether such results are specific enough to become useful biomarkers for diagnosis or treatment selection. Brain imaging techniques are also not easily accessible to all clinicians and expensive. Such instruments tend to be limited to major medical centers. Therefore, additional biomarkers, more easily deployed in the general population and based on blood or saliva samples that can be easily analyzed in remote labs, warrant exploration.

Genomics has witnessed a major search for biomarkers in psychiatry and candidate genomic biomarkers have been found for several psychiatric conditions, most notably, serotonin transport protein polymorphisms as predictors of depression and anxiety [38–41]. Specific gene mutations may interact with environmental factors to ultimately increase the risk of psychiatric problems [42–45]. However, despite the initial promise of genetic information, a recent mega-analysis of genome-association studies in MDD failed to find consistent genetic associations [46]. The search for gene mutations as psychiatric biomarkers may be flawed, since genomic information does not necessarily reflect active protein levels [47] or tell the researcher about possible important posttranslational modifications (PTMs) such as glycosylation, phosphorylation, or formation/destruction of disulfide bridges to keep/disturb the protein's three-dimensional structure [48]. Epigenetics, studying the transcription status of genes, is probably a better reflection of temporal changes in gene expression; but the study of epigenetics with respect to biomarkers in MDD is still in its infancy. Additional proteomic information may expand possibilities for a disorder's identification in clinical tests. Proteins represent the functional molecules in a

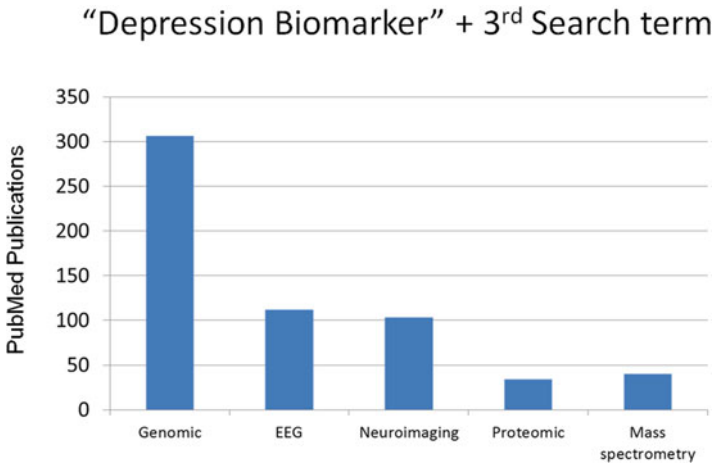


Fig. 27.1 PubMed search results for "depression biomarker" plus genomic, electroencephalogram (EEG), neuroimaging, proteomics, or mass spectrometry

biological system; therefore, study of proteins may take a researcher closer to identifying the cause of a disorder and could also suggest targets for therapeutics. Protein profiling of candidate biomarkers in psychiatry is therefore an area of/with great potential [49, 50]. Mass spectrometry (MS) is currently the technique of choice for proteomic studies [32, 49–73].

A recent PubMed search of depression and biomarkers reveals the extent to which different biomarker approaches are emphasized in the field of depression research. Using the search terms "depression biomarker" and adding a variable third term, yielded the following results: genomic = 306 articles, EEG = 112 articles, neuroimaging = 103 articles, proteomic = 34 articles, mass spectrometry = 40 articles (Fig. 27.1). This search underscores the need for expansion into the proteomic realm in the field of depression research.

MS-based proteomics is very useful for protein biomarker identification for several reasons. Using a variety of proteomics approaches, a researcher can detect and determine multiple properties of a specific sample [55, 56, 61, 63, 65–67]. Virtually all information about a protein/peptide such as mass, sequence information, and the charge state can be measured and identified using MS-based proteomic methods [66, 69, 70, 72, 74, 75]. Knowledge-based methods for protein biomarker discovery can also be employed, such as enzyme-linked immunoabsorbant assay (ELISA) [76], Western Blotting [72, 77], or immunohistochemistry [72, 78, 79]. These techniques can be used independently, with the downside being that the "discovery" aspect is negated by the need to identify a target protein. They may also be used to validate and expand upon the more comprehensive screening provided by MS. A general proteomics approach is presented in Fig. 27.2. Here the protein samples from various sources are either fractionated, digested, and then analyzed by MS (liquid chromatography-tandem mass spectrometry), followed by data analysis, validation, and follow up, or analyzed without a priori digestion (i.e., by matrix-assisted laser desorption

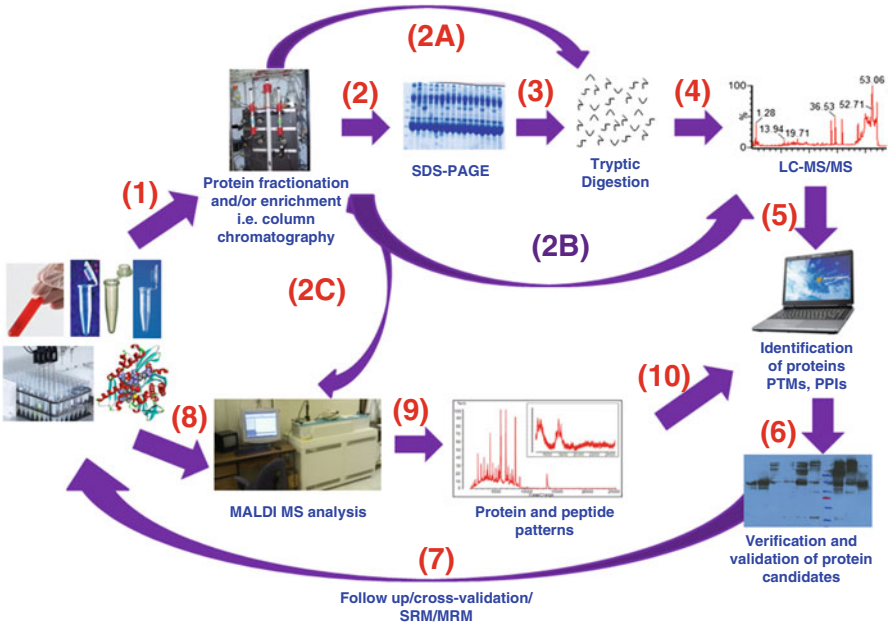


Fig. 27.2 Schematic of a full proteomics experiment. The protein sample is separated/fractionated under non-denaturing (*1*) and denaturing (*2*) conditions, digested (*3*) and analyzed by LC-MS/MS (*4*) (*bottom-up* proteomics). Database search and data analysis (*5*) led to identification of proteins, post-translational modifications (PTMs) in proteins and protein–protein interactions (PPIs), that were verified and then validated by Western blotting (*6*), and then monitored by single reaction monitoring (SRM), multiple reaction monitoring (MRM), and/or cross-validation. The original protein samples can also be analyzed by MALDI-MS (8) that leads to identification of specific protein/peptide patterns whose data analysis (*10*) leads to identification of proteins, PTMs, and PPIs. The fractionated samples in (*2*) can also be investigated without using electrophoretic denaturing conditions (*2A*) in a process called in solution digestion, or analyzed without any digestion (*2B*) in a process called *top-down* proteomics. In a third alternative, the fractionated samples from (*2*) are analyzed without digestion directly by MALDI-MS (*2C*). © 2013 Wormwood KL, et al. Reproduced from Wormwood et al., J. Proteomics Bioinform 2013, S5, http://dx.doi.org/10.4172/jpb.S5-001

ionization MS or MALDI-MS) followed by data analysis, verification, and follow up. Many times, when the expertise and instrumentation are available, both methods are used simultaneously and their outcome complements each other.

27.1.1 Protein Biomarkers in MDD

Depression has indeed been associated with several protein biomarkers, which have been identified using both directed methods as well as mass spectrometry [49, 80]. For example, serum brain-derived neurotrophic factor (BDNF) appears to be decreased relative to healthy controls, based on immunological measurement techniques [81–85]. A recent study using ELISA confirmed that ketamine treatment

indeed results in rapid upregulation of BDNF in those patients who are responders, measured in blood sera. Ketamine has been recently identified as producing rapid antidepressant response [86], which is in contrast to many antidepressants currently used, which take several weeks to produce a response [87]. Greater understanding of the factors producing BDNF upregulation may help in predicting responders to ketamine or possibly other antidepressants. MS and comprehensive proteomics may further expand upon directed approaches to yield more biomarker candidates.

27.1.2 Use of Animal Models and MS

MS and proteomic approaches have been utilized to identify putative biomarkers in depression animal models. An animal study examined the effects of chronic stress on proteins found in the hippocampus of rats. MALDI-TOF-MS was utilized to identify found 27 potential protein markers that were dysregulated relative to non-stressed control animals. These proteins were known to have roles in neurogenesis, oxidative metabolism, transcription, and signal transduction [88]. This is particularly relevant to depression based on the observation that the hippocampus may decrease in volume in individuals who are depressed [89]. A separate study Using iTRAQ labeling, a label-based method used to identify and quantify proteins, coupled with MS, quantified 2,000 proteins from rat hippocampus in a stress model of depression. Seventy-three proteins were found to be differentially expressed. Deficits in vesicular release proteins (SNCA, SYN-1, and AP-3) were found with a relationship to stress susceptibility. Increased expression of a sodium channel protein (SCN9A) also predicted susceptibility to stress [90]. Such susceptibility could provide a model for biomarkers predictive of depression, although naturally these results would need to be replicated in humans.

27.1.3 Protein Biomarkers in MDD: Human Studies

In addition to animal models, MS has indeed also yielded biomarker insights in humans. Alawam et al., used MALDI-MS with a C18 magnetic beads protocol to analyze serum samples from 39 individuals with depression and 30 controls. C18 magnetic silica beads are used for protein and peptide sample concentration, desalting, and fractionation. C18 beads are coated with C18 alkyl groups, which allow them to provide a reversed phase surface chemistry and, through fractionation of proteins and peptides, reduce sample complexity. No protein signals distinguished participants with depression from controls, however, analysis of individual peptide signals identified three signals that were distinct in the depression group. The specific identities of these peptides were however, not reported [91]. Protein identification is the next step to establish biomarkers, although comprehensive protein fingerprints via spectral differences could also serve as a potential diagnostic.

Proteomic analysis of brain tissue is an approach that can be used to identify MDD biomarkers, although naturally the approach is limited to post-mortem tissue. A proteomic study using difference gel electrophoresis (DIGE) identified 59 potential biomarkers in cerebral cortex and 11 in amygdala in post-mortem brain tissue from suicide victims. Proteins identified were those that control biological functions and structures such as metabolism, the redox system, the cytoskeleton, synaptic function, and proteolysis [92]. Another group used shotgun proteomics to analyze post-mortem dorsolateral prefrontal cortex brain tissue from 24 MDD patients and 12 matched controls [93]. Shotgun proteomics refers to the use of bottom-up proteomics techniques in identifying proteins in complex mixtures. This method usually takes advantage of combining HPLC with MS. Distinct protein fingerprints resulted from significantly, differentially expressed proteins between subgroups of MDD patients (with and without psychosis) as well as between MDD subjects and healthy controls. Differentially expressed proteins between MDD patients and healthy controls were those involved in metabolism, transport, cell communication and signalling, cell growth and maintenance, protein metabolism, and regulation of nucleic acid metabolism [93].

In addition to the analysis of brain tissue, MS and proteomic approaches have been used to analyze bodily fluids. The advantage of this approach is naturally the possibility of identifying biomarkers in live subjects, with the hope that the proteins identified reflect brain processes or have the potential for diagnosis and/or prediction of response. Cerebrospinal fluid (CSF) is naturally the closest fluid to brain tissue, and therefore may be most suitable for reflecting brain content. The downside of CSF use is the invasive nature of collection and need for a spinal tap. CNS has been analyzed using MALDI-TOF-MS to compare individuals with MDD versus controls. This study reported 11 proteins and 144 peptide features that were significantly different in the CSF from depressed patients versus controls. The investigators also found differences in the phosphorylation pattern of several CSF proteins, underscoring the additional utility of MS for identifying PTMs as well as protein levels. Identified dysregulated proteins in this study were those involved with neuroprotection, neuronal development, sleep regulation, and amyloid plaque deposition in aging brain [94], reflecting possible disruptions of these systems by depression, involvement of these systems in the etiology of depression, or both.

Blood plasma or serum is more accessible than CSF, and may also be used for proteomic identification in MDD. Depression has been associated with disturbances in cholesterol transport and metabolism, which is mediated by apolipoproteins [32]. Indeed, a study utilizing multidimensional liquid chromatography-tandem mass spectrometry with validation by immunoblotting or ELISA to analyze blood plasma, several protein biomarkers that are altered in depressed individuals. The primary function of the identified proteins was in lipid metabolism and immunoregulation [95]. This is consistent with prior observations of increased disease incidence with depression [4–10] as well as lipid perturbations that have been observed in depression [32].

Urine analysis could be most convenient and amenable to an MDD diagnostic or predictive test. Liquid chromatography-tandem mass spectrometry (LC-MS/MS)

was used to corroborate levels of urinary serotonin in individuals with depression taking antidepressants. A strong correlation was measured between urinary serotonin levels measured by ELISA. Using ELISA, serotonin levels detected in depressed patients were significantly lower than in control subjects and 5-hydroxytryptophane and/or selective serotonin re-uptake inhibitors significantly increased urinary serotonin levels [96]. Biological monitoring of depression and response to antidepressants may therefore be possible via urinary analysis.

Saliva is an accessible biofluid that has not been fully exploited in depression research. Saliva collection is noninvasive and convenient. In all, 2,290 proteins have been found in saliva compared with 2,698 proteins found in plasma. Nearly 40 % of proteins believed to be candidate markers for diseases such as cancer, cardiovascular disease, and stroke can be found in whole saliva, Therefore, saliva presents a convenient, noninvasive, accessible bodily fluid for analysis [97]. Potential salivary biomarkers have been identified for other disorders, specifically Alzheimer's disease [98] and autism spectrum disorder [99]. Although we have not been able to identify analyses of MDD salivary proteome, salivary cortisol in MDD has been analyzed using directed methods (competitive radioimmunoassay for cortisol) rather than the comprehensive screening provided by MS. High salivary cortisol was not found to be a risk factor for MDD, however, low mean salivary cortisol concentration and a small difference between morning and evening cortisol concentrations were identified as possible risk factors [100]. Saliva may be an unexploited biofluid for MS analysis in MDD that warrants further exploration.

27.1.4 Metabolomics for MS-Based Biomarker Discovery in MDD

Based on prevalent hypotheses that small molecules, such as monoamine neurotransmitters, are dysregulated in MDD [2], metabolomics is an approach that may yield biomarkers in MDD. Metabolomics is the investigation of the metabolites, or small molecules that exist on a cell or a bodily fluid at a particular timepoint, and can be performed using specialized MS approaches. In a study comprising 26 subjects with MDD versus controls, blood sera content was measured using a 3200 QTRAP MS system. The investigators found depletions in tryptophan, lysine, and gamma amino butyric acid (GABA) in the subjects with MDD [101]. Proton nuclear magnetic resonance ((1)H NMR) spectrometry has also been used to distinguish metabolomic differences in MDD plasma relative to control plasma. Investigators have reported use of this method to accurately diagnose MDD in 26 subjects with sensitivity and specificity of 92.8 % and 83.3 %, respectively [102]. The prevalent monoamine theory of depression as well as the known mechanism of action of many antidepressant medications on neurotransmitters (small molecules) suggests that metabolomics may be particularly useful for MS-based exploration of biomarkers in MDD. The ability to measure metabolites in blood, urine, or saliva could provide a useful diagnostic and predictive test for MDD.

27.1.5 Protein Preparation Methods

Preparing samples for MS analysis can be a crucial component of biomarker discovery. In fact, it may be considered the most important component of the MS analysis. Depending on how the sample is prepared, one can get either useless results or significant results with high clinical relevance. This is well reflected in the NCI Clinical Proteomic Tumor Analysis Consortium (CPTAC) study, in which different proteomics labs from the United States, with different instrumentation and instrument settings obtained different results [103]. For example, the methods by which samples are collected are crucial, including how they are processed, frozen (and how many times they are thawed). Instrument settings need to be identical, for high reproducibility. Furthermore, It also matters whether the samples to be analyzed (i.e., sera) are depleted or not and if so, whether the depletion is performed under identical conditions in different labs. At its best, due to the secretome and/or functional degradome (truncated proteins with physiological significance), as well as posttranslational proteomics (labile or transient/reversible phosphorylations, acetylations, methylations, glycosylations, etc.), the samples have to be kept as intact as possible and then processed. For example, one study identified transferrin and fibrinogen as potential biomarkers, based on 6 M HCL hydrolyzation of serum proteins followed by MALDI-TOF MS [104]. However, HCl incubation caused rapid protein decomposition. To distinguish between samples from individuals with MDD versus control samples, the same investigators reported that protein hydrolysis using 20 % TFA and sinapinic acid were the optimal reagents for protein hydrolysis and found that the same markers, fibrinogen and transferrin, could be distinguished as potential biomarkers for MDD, with higher peaks identified for fibrinogen relative to controls and lower peaks for transferrin (relative intensity) [105]. However, analysis of the forced, uncontrolled fragmentation of the proteins from the samples may be dangerous, as it may be difficult to reproduce and translate into clinical settings. Therefore, sample preparation is and will always be the most crucial part of the proteomics that will allow one to obtain significant results with clinical relevance.

27.1.6 Effects of Antidepressant Treatment on Protein Biomarkers

In addition to understanding depression, naturally MS-based proteomics may ultimately elucidate markers that correspond with outcomes of antidepressant treatment. Initial proteomic analysis in animal models indicates possible protein changes associated with the use of antidepressant medications. A proteomic study analyzed monoamine reuptake inhibitors (MAOIs) and protein changes that occurred in rat hippocampal cytosolic extract. Either venlafaxine (primarily noradrenergic) or fluoxetine (primarily serotonergic) was systemically administered to adult rats for

2 weeks, after which 2D gel electrophoresis identified 31 proteins that were upregulated and two that were down-regulated. MALDI TOF MS was used to identify these proteins. Treatment with both antidepressants led to the upregulation of several factors associated with central nervous system function. These include those associated with neurogenesis (IGF-1, glia maturation factor) neuronal process outgrowth, and maintenance (hippocampal cholinergic neurostimulating peptide [HCNP]-precursor; PCTAIRE-3; serine protease inhibitor 2.1), and anti-apoptotic activity (dimethylargininase-1 L-*N,N*-dimethylarginine dimethylaminohydrolase-1 [DDAH1], antioxidant protein-2 [AOP-2], and pyruvate dehydrogenase-E1 [PDH-E1]) [106]. Whether similar markers can be measured in human bodily fluids remains to be seen.

Using directed protein measurements (immunoblot and electrophoretic mobility shift assays) of total cyclic AMP response element binding protein (tCREB) in T-lymphocytes, it has been found that subjects with MDD and low baseline tCREB were significantly more likely to respond (78 %) to selective serotonin reuptake inhibitor (SSRI) treatment than individuals with high baseline tCREB (36 %) [107]. This study illustrates that biomarkers that are predictive of antidepressant response could be very useful to directing and selecting treatments. The question remains as to whether tCREB is a specific marker for SSRI response, or antidepressant response in general. MS and proteomic have the potential to reveal further protein biomarkers and expand upon this therapeutic area, which will become increasingly important as more treatment options with different mechanisms of action become available.

27.2 Concluding Remarks

Although various biomarkers for MDD have been identified using an assortment of methods, the true test of a biomarker will be its validity, reproducibility, and prediction of therapeutic response. In addition, biomarkers for use in clinical tests will need to be identified in an inexpensive, quick, and convenient manner. For this reason, protein tests may be particularly useful. Proteins can be derived from easily obtainable bodily fluids, such as blood and even saliva. MS has the potential to identify new protein or even metabolite targets that may ultimately be used in tests. Such tests may be conducted via MS, or using directed protein-identification methods that are quick, cheap, and easy. The hope is to develop a form of personalized medicine for MDD that can direct the choice of treatment, particularly in light of the high rate of depression relapse and the high rate of treatment-resistance.

Acknowledgments This work was supported in part by the U.S. Army research office through the Defense University Research Instrumentation Program (DURIP grant #W911NF-11-1-0304). We thank supporters of the Biochemistry & Proteomics Group, specifically Bob and Karen Brown, Bonhomie Imports, Wolverine, and the SciFund challenge contributors.

References

1. Kessler RC, Petukhova M, Sampson NA, Zaslavsky AM, Wittchen HU (2012) Twelve-month and lifetime prevalence and lifetime morbid risk of anxiety and mood disorders in the United States. Int J Methods Psychiatr Res 21:169–184
2. To SE, Zepf RA, Woods AG (2005) The symptoms, neurobiology, and current pharmacological treatment of depression. J Neurosci Nurs 37:102–107
3. APA (2013) Diagnostic and statistical manual of mental disorders, 5th edn. American Psychiatric Association, Arlington, VA
4. Woods AG (2004) Understanding depression and diabetes. Diabetes Self Manag 21:6, 8, 11–12
5. Strine TW, Mokdad AH, Balluz LS, Gonzalez O, Crider R, Berry JT et al (2008) Depression and anxiety in the United States: findings from the 2006 Behavioral Risk Factor Surveillance System. Psychiatr Serv 59:1383–1390
6. Niranjan A, Corujo A, Ziegelstein RC, Nwulia E (2012) Depression and heart disease in US adults. Gen Hosp Psychiatry 34:254–261
7. Mathew CS, Dominic M, Isaac R, Jacob JJ (2012) Prevalence of depression in consecutive patients with type 2 diabetes mellitus of 5-year duration and its impact on glycemic control. Indian J Endocrinol Metab 16:764–768
8. Kessler RC, Chiu WT, Demler O, Merikangas KR, Walters EE (2005) Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. Arch Gen Psychiatry 62:617–627
9. Kessler RC, Berglund P, Demler O, Jin R, Koretz D, Merikangas KR et al (2003) The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R). JAMA 289:3095–3105
10. Chapman D, Perry G, Strine T (2005) The vital link between chronic disease and depressive disorders. Prev Chronic Dis 2:A14
11. Shim RS, Baltrus P, Ye J, Rust G (2011) Prevalence, treatment, and control of depressive symptoms in the United States: results from the National Health and Nutrition Examination Survey (NHANES), 2005–2008. J Am Board Fam Med 24:33–38
12. Stewart WF, Ricci JA, Chee E, Hahn SR, Morganstein D (2003) Cost of lost productive work time among US workers with depression. JAMA 289:3135–3144
13. Wang PS, Lane M, Olfson M, Pincus HA, Wells KB, Kessler RC (2005) Twelve-month use of mental health services in the United States: results from the National Comorbidity Survey Replication. Arch Gen Psychiatry 62:629–640
14. Rush AJ, Trivedi MH, Wisniewski SR, Nierenberg AA, Stewart JW, Warden D et al (2006) Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: a STAR*D report. Am J Psychiatry 163:1905–1917
15. Trivedi MH, Rush AJ, Wisniewski SR, Warden D, McKinney W, Downing M et al (2006) Factors associated with health-related quality of life among outpatients with major depressive disorder: a STAR*D report. J Clin Psychiatry 67:185–195
16. Trivedi MH, Fava M, Wisniewski SR, Thase ME, Quitkin F, Warden D et al (2006) Medication augmentation after the failure of SSRIs for depression. N Engl J Med 354:1243–1252
17. Rush AJ, Trivedi MH, Wisniewski SR, Stewart JW, Nierenberg AA, Thase ME et al (2006) Bupropion-SR, sertraline, or venlafaxine-XR after failure of SSRIs for depression. N Engl J Med 354:1231–1242
18. Woods AG (2008) Give a man a fish. Essential fatty acids in health and disease. Diabetes Self Manag 25:8, 11–12, 14
19. Roth R (2012) Lisdexamfetamine dimesylate augmentation for executive dysfunction in adults with fully or partially remitted major depressive disorder. In: The 165th annual meeting of the American Psychiatric Association, Philadelphia, PA
20. Piet J, Hougaard E (2011) The effect of mindfulness-based cognitive therapy for prevention of relapse in recurrent major depressive disorder: a systematic review and meta-analysis. Clin Psychol Rev 31:1032–1040

21. Pehrson AL, Sanchez C (2013) Serotonergic modulation of glutamate neurotransmission as a strategy for treating depression and cognitive dysfunction. CNS Spectr 19(2):121–133
22. Murrough JW, Iosifescu DV, Chang LC, Al Jurdi RK, Green CM, Perez AM et al (2013) Antidepressant efficacy of ketamine in treatment-resistant major depression: a two-site randomized controlled trial. Am J Psychiatry 170:1134–1142
23. Mantione E, Micheloni S, Alcaino C, New K, Mazzaferro S, Bermudez I (2012) Allosteric modulators of alpha4beta2 nicotinic acetylcholine receptors: a new direction for antidepressant drug discovery. Future Med Chem 4:2217–2230
24. Keefe R (2012) Lisdexamfetamine dimesylate in the treatment of cognitive dysfunction in patients with partially or fully remitted major depressive disorder. In: 165th annual meeting of the American Psychiatric Association, Philadelphia, PA
25. Celada P, Bortolozzi A, Artigas F (2013) Serotonin 5-HT1A receptors as targets for agents to treat psychiatric disorders: rationale and current status of research. CNS Drugs 27:703–716
26. Martins-De-Souza D, Wobrock T, Zerr I, Schmitt A, Gawinecka J, Schneider-Axmann T et al (2010) Different apolipoprotein E, apolipoprotein A1 and prostaglandin-H2 D-isomerase levels in cerebrospinal fluid of schizophrenia patients and healthy controls. World J Biol Psychiatry 11:719–728
27. Pallis AG, Fennell DA, Szutowicz E, Leighl NB, Greillier L, Dziadziuszko R (2011) Biomarkers of clinical benefit for anti-epidermal growth factor receptor agents in patients with non-small-cell lung cancer. Br J Cancer 105:1–8
28. Phillips KA, Marshall DA, Haas JS, Elkin EB, Liang SY, Hassett MJ et al (2009) Clinical practice patterns and cost effectiveness of human epidermal growth receptor 2 testing strategies in breast cancer patients. Cancer 115:5166–5174
29. Ross JS (2011) Biomarker-based selection of therapy for colorectal cancer. Biomark Med 5:319–332
30. Lakhan SE, Vieira K, Hamlat E (2010) Biomarkers in psychiatry: drawbacks and potential for misuse. Int Arch Med 3:1
31. Singh I, Rose N (2009) Biomarkers in psychiatry. Nature 460:202–207
32. Woods AG, Sokolowska I, Taurines R, Gerlach M, Dudley E, Thome J et al (2012) Potential biomarkers in psychiatry: focus on the cholesterol system. J Cell Mol Med 16:1184–1195
33. Cook IA, Hunter AM, Gilmer WS, Iosifescu DV, Zisook S, Burgoyne KS et al (2013) Quantitative electroencephalogram biomarkers for predicting likelihood and speed of achieving sustained remission in major depression: a report from the biomarkers for rapid identification of treatment effectiveness in major depression (BRITE-MD) trial. J Clin Psychiatry 74:51–56
34. Greenblatt JM, Sussman C, Jameson M, Yuan L, Hoffman DA, Iosifescu DV (2011) Retrospective chart review of a referenced EEG database in assisting medication selection for treatment of depression in patients with eating disorders. Neuropsychiatr Dis Treat 7:529–541
35. Iosifescu DV (2011) Electroencephalography-derived biomarkers of antidepressant response. Harv Rev Psychiatry 19:144–154
36. Hoogenboom WS, Perlis RH, Smoller JW, Zeng-Treitler Q, Gainer VS, Murphy SN et al (2014) Limbic system white matter microstructure and long-term treatment outcome in major depressive disorder: a diffusion tensor imaging study using legacy data. World J Biol Psychiatry 15:122–134
37. Blood AJ, Iosifescu DV, Makris N, Perlis RH, Kennedy DN, Dougherty DD et al (2010) Microstructural abnormalities in subcortical reward circuitry of subjects with major depressive disorder. PLoS One 5:e13945
38. Barnett JH, Smoller JW (2009) The genetics of bipolar disorder. Neuroscience 164: 331–343
39. Kvajo M, McKellar H, Gogos JA (2010) Molecules, signaling, and schizophrenia. Curr Top Behav Neurosci 4:629–656
40. Poelmans G, Pauls DL, Buitelaar JK, Franke B (2011) Integrated genome-wide association study findings: identification of a neurodevelopmental network for attention deficit hyperactivity disorder. Am J Psychiatry 168:365–377

41. Weber H, Kittel-Schneider S, Gessner A, Domschke K, Neuner M, Jacob CP et al (2011) Cross-disorder analysis of bipolar risk genes: further evidence of DGKH as a risk gene for bipolar disorder, but also unipolar depression and adult ADHD. Neuropsychopharmacology 36:2076–2085
42. Caspi A, Hariri AR, Holmes A, Uher R, Moffitt TE (2010) Genetic sensitivity to the environment: the case of the serotonin transporter gene and its implications for studying complex diseases and traits. Am J Psychiatry 167:509–527
43. Kolassa IT, Kolassa S, Ertl V, Papassotiropoulos A, De Quervain DJ (2010) The risk of posttraumatic stress disorder after trauma depends on traumatic load and the catechol-o-methyltransferase Val(158)Met polymorphism. Biol Psychiatry 67:304–308
44. Lahey BB, Rathouz PJ, Lee SS, Chronis-Tuscano A, Pelham WE, Waldman ID et al (2011) Interactions between early parenting and a polymorphism of the child's dopamine transporter gene in predicting future child conduct disorder symptoms. J Abnorm Psychol 120:33–45
45. Roy A, Sarchiopone M, Carli V (2009) Gene-environment interaction and suicidal behavior. J Psychiatr Pract 15:282–288
46. Major Depressive Disorder Working Group of the Psychiatric GWAS Consortium, Ripke S, Wray NR, Lewis CM, Hamilton SP, Weissman MM et al (2013) A mega-analysis of genome-wide association studies for major depressive disorder. Mol Psychiatry 18:497–511
47. Anderson L, Seilhamer J (1997) A comparison of selected mRNA and protein abundances in human liver. Electrophoresis 18:533–537
48. Junaid MA, Pullarkat RK (2001) Proteomic approach for the elucidation of biological defects in autism. J Autism Dev Disord 31:557–560
49. Ngounou Wetie AG, Sokolowska I, Wormwood K, Michel TM, Thome J, Darie CC et al (2013) Mass spectrometry for the detection of potential psychiatric biomarkers. J Mol Psychiatry 1:8
50. Woods AG, Ngounou Wetie AG, Sokolowska I, Russell S, Ryan JP, Michel TM et al (2013) Mass spectrometry as a tool for studying autism spectrum disorder. J Mol Psychiatry 1:6
51. Darie C (2013) Investigation of protein-protein interactions by blue native-PAGE & mass spectrometry. Mod Chem Appl 1:e111
52. Darie CC, Biniossek ML, Gawinowicz MA, Milgrom Y, Thumfart JO, Jovine L et al (2005) Mass spectrometric evidence that proteolytic processing of rainbow trout egg vitelline envelope proteins takes place on the egg. J Biol Chem 280:37585–37598
53. Darie CC, Biniossek ML, Jovine L, Litscher ES, Wassarman PM (2004) Structural characterization of fish egg vitelline envelope proteins by mass spectrometry. Biochemistry 43:7459–7478
54. Darie CC, Biniossek ML, Winter V, Mutschler B, Haehnel W (2005) Isolation and structural characterization of the Ndh complex from mesophyll and bundle sheath chloroplasts of Zea mays. FEBS J 272:2705–2716
55. Darie CC, Deinhardt K, Zhang G, Cardasis HS, Chao MV, Neubert TA (2011) Identifying transient protein-protein interactions in EphB2 signaling by blue native PAGE and mass spectrometry. Proteomics 11:4514–4528
56. Darie CC, Janssen WG, Litscher ES, Wassarman PM (2008) Purified trout egg vitelline envelope proteins VEbeta and VEgamma polymerize into homomeric fibrils from dimers in vitro. Biochim Biophys Acta 1784:385–392
57. Darie CC, Litscher ES, Wassarman PM (2008) Structure, processing, and polymerization of rainbow trout egg vitelline envelope proteins. Nato Scie Peace Secu 23–36
58. Darie CC, Shetty V, Spellman DS, Zhang GJ, Xu CF, Cardasis HL et al (2008) Blue native page and mass spectrometry analysis of ephrin stimulation-dependent protein-protein interactions in Ng108-Ephb2 cells. Nato Scie Peace Secu 3–22
59. Ngounou Wetie AG, Sokolowska I, Woods AG, Darie CC (2013) Identification of post-translational modifications by mass spectrometry. Aust J Chem 66:734–748
60. Ngounou Wetie AG, Sokolowska I, Woods AG, Roy U, Deinhardt K, Darie CC (2014) Protein-protein interactions: switch from classical methods to proteomics and bioinformatics-based approaches. Cell Mol Life Sci 71:205–228

61. Ngounou Wetie AG, Sokolowska I, Woods AG, Roy U, Loo JA, Darie CC (2013) Investigation of stable and transient protein-protein interactions: past, present, and future. Proteomics 13:538–557
62. Ngounou Wetie AG, Sokolowska I, Woods AG, Wormwood KL, Dao S, Patel S et al (2013) Automated mass spectrometry-based functional assay for the routine analysis of the secretome. J Lab Autom 18:19–29
63. Sokolowska I, Dorobantu C, Woods AG, Macovei A, Branza-Nichita N, Darie CC (2012) Proteomic analysis of plasma membranes isolated from undifferentiated and differentiated HepaRG cells. Proteome Sci 10:47
64. Sokolowska I, Gawinowicz MA, Ngounou Wetie AG, Darie CC (2012) Disulfide proteomics for identification of extracellular or secreted proteins. Electrophoresis 33:2527–2536
65. Sokolowska I, Ngounou Wetie AG, Roy U, Woods AG, Darie CC (1834) Mass spectrometry investigation of glycosylation on the NXS/T sites in recombinant glycoproteins. Biochim Biophys Acta 2013:1474–1483
66. Sokolowska I, Ngounou Wetie AG, Woods AG, Darie CC (2012) Automatic determination of disulfide bridges in proteins. J Lab Autom 17:408–416
67. Sokolowska I, Ngounou Wetie AG, Woods AG, Darie CC (2013) Applications of mass spectrometry in proteomics. Aust J Chem 66:721–733
68. Sokolowska I, Ngounou Wetie AG, Wormwood K, Michel TM, Thome J, Darie CC et al (2013) The potential of biomarkers in psychiatry: focus on proteomics. J Neural Transm. [Epub ahead of print]
69. Sokolowska I, Woods AG, Gawinowicz MA, Roy U, Darie CC (2012) Identification of potential tumor differentiation factor (TDF) receptor from steroid-responsive and steroid-resistant breast cancer cells. J Biol Chem 287:1719–1733
70. Sokolowska I, Woods AG, Gawinowicz MA, Roy U, Darie CC (2013) Characterization of tumor differentiation factor (TDF) and its receptor (TDF-R). Cell Mol Life Sci 70:2835–2848
71. Sokolowska I, Woods AG, Wagner J, Dorler J, Wormwood K, Thome J et al (2011) Mass spectrometry for proteomics-based investigation of oxidative stress and heat shock proteins. In: Andreescu S, Hepel M (eds) Oxidative stress: diagnostics, prevention, and therapy. American Chemical Society, Washington, DC
72. Woods AG, Sokolowska I, Deinhardt K, Sandu C, Darie CC (2013) Identification of tumor differentiation factor (TDF) in select CNS neurons. Brain Struct Funct. [Epub ahead of print]
73. Woods AG, Sokolowska I, Yakubu R, Butkiewicz M, LaFleur M, Talbot C et al (2011) Blue native page and mass spectrometry as an approach for the investigation of stable and transient protein-protein interactions. In: Andreescu S, Hepel M (eds) Oxidative stress: diagnostics, prevention, and therapy. American Chemical Society, Washington, DC
74. Roy U, Sokolowska I, Woods AG, Darie CC (2012) Structural investigation of tumor differentiation factor. Biotechnol Appl Biochem 59:445–450
75. Taurines R, Dudley E, Grassl J, Warnke A, Gerlach M, Coogan AN et al (2011) Proteomic research in psychiatry. J Psychopharmacol 25:151–196
76. Haile CN, Murrough JW, Iosifescu DV, Chang LC, Al Jurdi RK, Foulkes A et al (2014) Plasma brain derived neurotrophic factor (BDNF) and response to ketamine in treatment-resistant depression. Int J Neuropsychopharmacol 17(2):331–336
77. Kazuno AA, Ohtawa K, Otsuki K, Usui M, Sugawara H, Okazaki Y et al (2013) Proteomic analysis of lymphoblastoid cells derived from monozygotic twins discordant for bipolar disorder: a preliminary study. PLoS One 8:e53855
78. Woods AG, Poulsen FR, Gall CM (1999) Dexamethasone selectively suppresses microglial trophic responses to hippocampal deafferentation. Neuroscience 91:1277–1289
79. Woods AG, Guthrie KM, Kurlawalla MA, Gall CM (1998) Deafferentation-induced increases in hippocampal insulin-like growth factor-1 messenger RNA expression are severely attenuated in middle aged and aged rats. Neuroscience 83:663–668
80. Schneider B, Prvulovic D (2013) Novel biomarkers in major depression. Curr Opin Psychiatry 26:47–53

81. Yoshida T, Ishikawa M, Niitsu T, Nakazato M, Watanabe H, Shiraishi T et al (2012) Decreased serum levels of mature brain-derived neurotrophic factor (BDNF), but not its precursor proBDNF, in patients with major depressive disorder. PLoS One 7:e42676
82. Kotan Z, Sarandol E, Kirhan E, Ozkaya G, Kirli S (2012) Serum brain-derived neurotrophic factor, vascular endothelial growth factor and leptin levels in patients with a diagnosis of severe major depressive disorder with melancholic features. Ther Adv Psychopharmacol 2:65–74
83. Molendijk ML, Bus BA, Spinhoven P, Penninx BW, Kenis G, Prickaerts J et al (2011) Serum levels of brain-derived neurotrophic factor in major depressive disorder: state-trait issues, clinical features and pharmacological treatment. Mol Psychiatry 16:1088–1095
84. Huang TL, Lee CT, Liu YL (2008) Serum brain-derived neurotrophic factor levels in patients with major depression: effects of antidepressants. J Psychiatr Res 42:521–525
85. Karege F, Perret G, Bondolfi G, Schwald M, Bertschy G, Aubry JM (2002) Decreased serum brain-derived neurotrophic factor levels in major depressed patients. Psychiatry Res 109:143–148
86. Carlson PJ, Diazgranados N, Nugent AC, Ibrahim L, Luckenbaugh DA, Brutsche N et al (2013) Neural correlates of rapid antidepressant response to ketamine in treatment-resistant unipolar depression: a preliminary positron emission tomography study. Biol Psychiatry 73:1213–1221
87. Harmer CJ, Cowen PJ (2013) 'It's the way that you look at it'—a cognitive neuropsychological account of SSRI action in depression. Philos Trans R Soc Lond B Biol Sci 368:20120407
88. Mu J, Xie P, Yang ZS, Yang DL, Lv FJ, Luo TY et al (2007) Neurogenesis and major depression: implications from proteomic analyses of hippocampal proteins in a rat depression model. Neurosci Lett 416:252–256
89. Posener JA, Wang L, Price JL, Gado MH, Province MA, Miller MI et al (2003) High-dimensional mapping of the hippocampus in depression. Am J Psychiatry 160:83–89
90. Henningsen K, Palmfeldt J, Christiansen S, Baiges I, Bak S, Jensen ON et al (2012) Candidate hippocampal biomarkers of susceptibility and resilience to stress in a rat model of depression. Mol Cell Proteomics 11:M111.016428
91. Alawam K, Dudley E, Donev R, Thome J (2012) Protein and peptide profiling as a tool for biomarker discovery in depression. Electrophoresis 33:3830–3834
92. Kekesi KA, Juhasz G, Simor A, Gulyassy P, Szego EM, Hunyadi-Gulyas E et al (2012) Altered functional protein networks in the prefrontal cortex and amygdala of victims of suicide. PLoS One 7:e50532
93. Martins-de-Souza D, Guest PC, Harris LW, Vanattou-Saifoudine N, Webster MJ, Rahmoune H et al (2012) Identification of proteomic signatures associated with depression and psychotic depression in post-mortem brains from major depression patients. Transl Psychiatry 2:e87
94. Ditzen C, Tang N, Jastorff AM, Teplytska L, Yassouridis A, Maccarrone G et al (2012) Cerebrospinal fluid biomarkers for major depression confirm relevance of associated pathophysiology. Neuropsychopharmacology 37:1013–1025
95. Xu HB, Zhang RF, Luo D, Zhou Y, Wang Y, Fang L et al (2012) Comparative proteomic analysis of plasma from major depressive patients: identification of proteins associated with lipid metabolism and immunoregulation. Int J Neuropsychopharmacol 15:1413–1425
96. Nichkova MI, Huisman H, Wynveen PM, Marc DT, Olson KL, Kellermann GH (2012) Evaluation of a novel ELISA for serotonin: urinary serotonin as a potential biomarker for depression. Anal Bioanal Chem 402:1593–1600
97. Hu S, Li Y, Wang J, Xie Y, Tjon K, Wolinsky L et al (2006) Human saliva proteome and transcriptome. J Dent Res 85:1129–1133
98. Shi M, Sui YT, Peskind ER, Li G, Hwang H, Devic I et al (2011) Salivary tau species are potential biomarkers of Alzheimer's disease. J Alzheimers Dis 27:299–305
99. Castagnola M, Messana I, Inzitari R, Fanali C, Cabras T, Morelli A et al (2008) Hypophosphorylation of salivary peptidome as a clue to the molecular pathogenesis of autism spectrum disorders. J Proteome Res 7:5327–5332
100. Grynderup MB, Kolstad HA, Mikkelsen S, Andersen JH, Bonde JP, Buttenschon HN et al (2013) A two-year follow-up study of salivary cortisol concentration and the risk of depression. Psychoneuroendocrinology 38:2042–2050

101. Xu HB, Fang L, Hu ZC, Chen YC, Chen JJ, Li FF et al (2012) Potential clinical utility of plasma amino acid profiling in the detection of major depressive disorder. Psychiatry Res 200:1054–1057
102. Zheng P, Gao HC, Li Q, Shao WH, Zhang ML, Cheng K et al (2012) Plasma metabonomics as a novel diagnostic approach for major depressive disorder. J Proteome Res 11:1741–1748
103. Ellis MJ, Gillette M, Carr SA, Paulovich AG, Smith RD, Rodland KK et al (2013) Connecting genomic alterations to cancer biology with proteomics: the NCI Clinical Proteomic Tumor Analysis Consortium. Cancer Discov 3:1108–1112
104. Lo LH, Huang TL, Shiea J (2009) Acid hydrolysis followed by matrix-assisted laser desorption/ionization mass spectrometry for the rapid diagnosis of serum protein biomarkers in patients with major depression. Rapid Commun Mass Spectrom 23:589–598
105. Huang TL, Cho YT, Su H, Shiea J (2013) Principle component analysis combined with matrix-assisted laser desorption ionization mass spectrometry for rapid diagnosing the sera of patients with major depression. Clin Chim Acta 424:175–181
106. Khawaja X, Xu J, Liang JJ, Barrett JE (2004) Proteomic analysis of protein changes developing in rat hippocampus after chronic antidepressant treatment: implications for depressive disorders and future therapies. J Neurosci Res 75:451–460
107. Lim SW, Kim S, Carroll BJ, Kim DK (2013) T-lymphocyte CREB as a potential biomarker of response to antidepressant drugs. Int J Neuropsychopharmacol 16:967–974

Chapter 28
Application of Mass Spectrometry to Characterize Localization and Efficacy of Nanoceria In Vivo

Karin L. Heckman, Joseph Erlichman, Ken Reed, and Matthew Skeels

Abstract In vivo study of nanomaterials is complicated by the physical and chemical changes induced in the nanomaterial by exposure to biological compartments. A diverse array of proteins can bind to the nanomaterial, forming a protein corona which may alter the dispersion, surface charge, distribution, and biological activity of the material. Evidence suggests that unique synthesis and stabilization strategies can greatly affect the composition of the corona, and thus, the in vivo properties of the nanomaterial. Protein and elemental analyses techniques are critical to characterizing the nature of the protein corona in order to best predict the in vivo behavior of the nanomaterial. Further, as described here, inductively coupled mass spectroscopy (ICP-MS) can also be used to quantify nanomaterial deposition in tissues harvested from exposed animals. Elemental analysis of ceria content demonstrated deposition of cerium oxide nanoparticles (CeNPs) in various tissues of healthy mice and in the brains of mice with a model of multiple sclerosis. Thus, ICP-MS analysis of nanomaterial tissue distribution can complement data illustrating the biological, and in this case, therapeutic efficacy of nanoparticles delivered in vivo.

K.L. Heckman (✉) • J. Erlichman
Departments of Biology, St. Lawrence University, Canton, NY, USA
e-mail: kheckman@stlawu.edu

K. Reed
Cerion NR x LLC, 1 Blossom Road, Rochester, NY 14610, USA

M. Skeels
Departments of Chemistry, St. Lawrence University, Canton, NY, USA

A.G. Woods and C.C. Darie (eds.), *Advancements of Mass Spectrometry in Biomedical Research*, Advances in Experimental Medicine and Biology 806,
DOI 10.1007/978-3-319-06068-2_28, © Springer International Publishing Switzerland 2014

Abbreviations

AC	Analytical centrifugation
AUC	Area under the curve
BBB	Blood–brain barrier
CA/EDTA	Citric acid/EDTA
CeNPs	Cerium oxide nanoparticles
DLS	Dynamic light scattering
EAE	Experimental autoimmune encephalomyelitis
ICP-MS	Inductively coupled mass spectrometry
LC-MS/MS	Liquid chromatography-mass spectrometry
MS	Multiple sclerosis
PEGylation	Polyethylene glycol addition
ROS	Reactive oxygen species
SDS-PAGE	SDS-polyacrylamide gel electrophoresis
SEC	Size exclusion chromatography

28.1 Introduction

The extravasation of drugs and nanomedicines through the blood–brain barrier (BBB) represents a significant obstacle for the development of effective therapeutics for neurological diseases. Of the over 40,000 current medicinal formulations, less than 1 % of the drugs gain entry and are active in the central nervous system [1]. Consequently, for many therapeutics, the BBB is a principle obstruction in the development of pipeline drugs to treat central nervous system disorders. Brain microvascular endothelial cells and glial end-feet, which constitute the anatomical basis of the BBB, form tight junctions due to a lack of fenestration, thereby reducing the diffusion of molecules across the vascular bed. However, cell and molecular biology studies have provided insights into the function of the BBB and the transport systems, enzymes, and receptors that regulate the penetration of molecules into the brain. Further, the field of nanotechnology has opened the door to the delivery of drugs and medicines through the BBB; several nanoparticle-bound drugs including doxorubicin [2–4], loperamide [5, 6], a novel anticonvulsive agent [7], and others [8–10] can be successfully delivered across the BBB, often times utilizing the brain's endogenous transport mechanisms.

Nanomaterials are systems (1–100 nm) that are potentially useful tools in medicine; by virtue of their small size, they readily interact with endogenous biomacromolecules (DNA, protein, lipids) and, thus, are able to act at the cellular, molecular, and genetic level. Both targeting nanoparticles to specific organs and cell types and understanding the interactions of nanomaterials with endogenous biomolecules remain significant challenges in nanomedicine development. Regardless of the route of administration, all nanomaterials encounter various cell types, complement system

enzymes, lipids, and a myriad of proteins both in the blood and the extravascular space. The complexity of the mammalian serum proteome is immense and estimated to contain over 5,000 different proteins at concentrations spanning over 10 orders of magnitude [11, 12]. While changes in the physical and chemical properties of nanomaterials can alter their affinity for biomolecules, predicting these interactions, a priori, is difficult at best. In vitro test beds have not shown good translational efficacy compared to the effects observed in intact animals, except for the most toxic of materials. This discordance in findings may occur, in part, as a result of the dynamic change in the protein corona in vivo which can evolve with changes in protein concentration and affinity to the nanomaterial as well as the ionic strength of the biological compartment the particle is resident in. For example, it is not unusual for a nanoparticle delivered to the lung to cross ~15 different physiological barriers (and hence a number of unique biological compartments) to reach a target in the brain.

Characterizing the interactions between nanomaterials and the proteins with which they associate is critical for the prediction and understanding of biological outcomes [13–18]. The protein corona can change the intrinsic physical–chemical properties of the nanoparticle (size, degree of subsequent particle aggregation, and surface properties) which in turn affect the biological activity of the nanomaterial. Understanding the biomolecular relationship between the protein corona and tissues may permit more selective targeting to cell receptors, cellular compartments or organelles, and specific biomolecules. In particular, the corona has been shown to affect nanoparticle uptake [19, 20], blood coagulation [21], targeting [15, 16], protein aggregation [22], and particle distribution [23].

In biosystems, the corona may evolve and take on an evolving identity during its lifetime as the nanomaterial traverses through biological systems [24]. Once a nanomaterial enters a biological system, the most abundant proteins are adsorbed on the surface. However, over time they are replaced by higher affinity proteins in a process termed the *Vroman Effect* [25], which may continually alter the properties and behavior of nanomaterials in biological systems. Thus, understanding the dynamic nature of the corona of nanomedicines would greatly aid in predicting the potential biological interactions and the evolution of nanomaterial efficacy and distribution within intact organisms over time (Fig. 28.1).

28.2 Structure and Composition of the Corona

The majority of adsorbed biomacromolecules on the surface of nanomaterials in blood plasma are proteins, and recently some minor traces of lipids have also been reported [26]. The protein corona is thought to be divided into two parts distinguished by binding affinities. Proteins with high affinity for the nanomaterial are known to comprise the “hard” corona, consisting primarily of proteins that bind tightly to the nanoparticle and have very slow exchange rates. Proteins of lower affinity comprise the “soft” corona and consist of proteins that interact loosely with the proteins that make up the “hard” corona and have more rapid exchange rates [27].

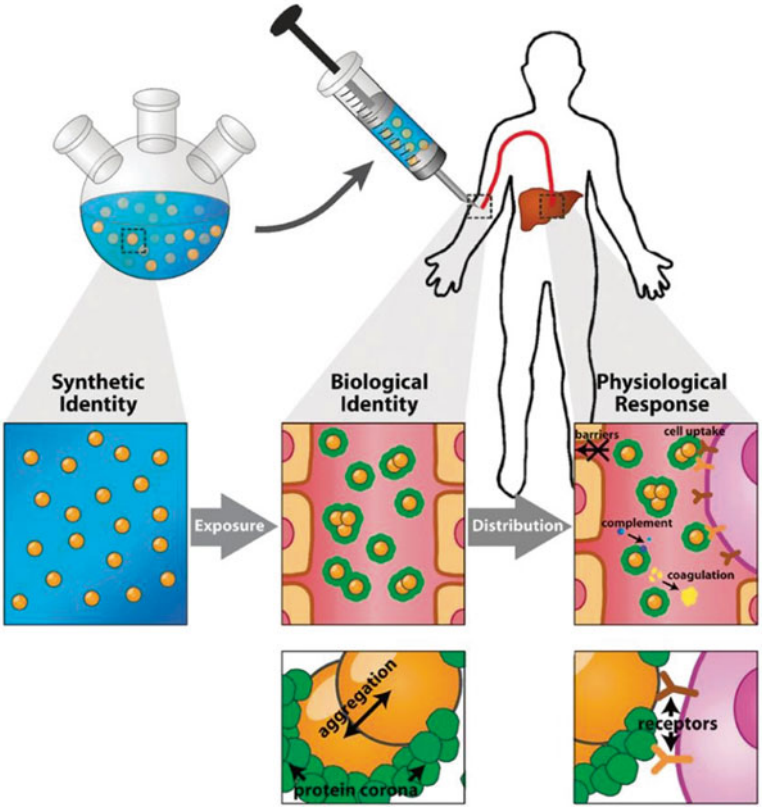


Fig. 28.1 The relationship between the identity of a nanoparticle and its physiological response. The synthetic identity of the nanoparticle includes the many physical and chemical properties of the synthetic nanoparticle. Upon exposure to biosystems, the nanoparticle becomes encapsulated by a corona of adsorbed proteins and begins to take on its biological identity. The biological identity determines the physiological response, including the distribution and activity of the nanoparticle. Reproduced from [27] with permission of The Royal Society of Chemistry

In their recent review of 63 nanomaterials from 26 studies, Walkey and Chan [27] noted 125 unique plasma proteins, termed the "adsorbome," in the protein corona of one standardized nanomaterial. This list is relatively new and will certainly be expanded as more studies are performed on the corona of nanomaterials. In this study, the most common plasma protein corona consists of approximately 2–6 proteins that are adsorbed in high abundance with many more proteins adsorbed in lower abundance. Interestingly, the adsorbed proteins in high abundance were not conserved across materials, and only a small subset of the adsorbome proteins bind to most nanoparticles [27]. However, some trends in protein binding have been observed. Dextran coated nanoparticles are typically bound by antibodies, while cationic and hydrophobic nanoparticle surfaces show strong affinity for albumin [28]. Hydrophobic surfaces tend to attract apolipoproteins, while surfaces presenting hydroxyl groups and positive charges promote binding of proteins from the complement system (C3) [28].

28.3 Hard vs. Soft Corona

The influence of differing nanoparticle surface modification on the protein corona has been reviewed [23], and it was observed that albumin, IgG, fibrinogen, and apolipoproteins are present in the corona of all of the studied nanomaterials. This is not surprising, given the relative abundance of these molecules in the blood, but it is also likely that in time these molecules are desorbed and replaced by higher affinity proteins as they traverse different biological compartments. Such association of proteins with the surface of the nanomaterial is primarily entropy driven [17]. Many water molecules associated with the nanomaterial are liberated when one protein, such as fibrinogen, apolipoprotein, or albumin, binds to the surface. The increase in entropy of the released water molecules is larger than the decrease in entropy of the corona-bound proteins. Proteins bound to the surface in this manner generally do not lose their active conformation [17].

Because of this type of interaction, the hard corona can evolve over time [29]. In the study by Lundqvist (2008), nanoparticles were first incubated in plasma followed by a secondary incubation in salt solutions that mimicked cytosol (i.e., intracellular fluid). Following this second incubation, the composition of the corona was characterized and compared to control samples that were incubated in each fluid separately. The results suggested that there is a distinct evolution of the corona, and that the final corona displays a roadmap or blueprint of its history. The authors of this work suggest that the identity of the corona may be used to potentially predict nanoparticle behavior in biosystems [29].

Though it is apparent that the protein corona has significant biological implications, it remains difficult to predict the interactions between biomacromolecules and nanomaterials. The coating and surface characteristics of nanoparticles appear to provide some indication of the nature of their corona interactions but predicting these interactions has proven difficult and will require intensive study if these materials are to be developed for biological applications.

28.4 Minimizing the Extent of Protein Corona Formation

Several polymers, such as polysaccharides, have been used to coat nanoparticles to efficiently increase circulation time [30]. Among the various approaches, addition of polyethylene glycol (PEG) (abbreviated here PEGylation), is the most widely used strategy to create a steric barrier on the surface of nanoparticles and has been a mainstay in functionalizing a wide variety of nanoparticles [31, 32]. PEGylation can change the physical and chemical properties of the nanomaterial, including its conformation, charge, and hydrophobicity. For many nano-formulations, PEGylation increases drug stability, decreases protein adsorption, improves solubility, and retention time of the conjugates in blood, and reduces proteolysis, renal excretion, and uptake by reticuloendothelial organs (i.e., liver and spleen). These factors result in a prolonged blood residence time, thereby allowing a reduced dosing frequency

[33]. For example, Walkey et al. [34] demonstrated that the surface density of PEG influences serum protein adsorption and the subsequent biological fate of gold nanoparticles, including their uptake by macrophages.

Unfortunately, PEGylation does have drawbacks. PEGylation strongly inhibits cellular uptake and endosomal escape, which results in significant loss of activity of the delivery system [35]. If the biological impact of the particle involves surface chemistry, high density PEGylation can dampen catalytic activity, and the desorption of PEG can lead to immune responses directed against the PEG molecule [36]. Despite its limitations, however, PEGylation continues to be an important strategy in improving the biocompatibility and performance of many nanomaterials.

Recently, we have shown that a stabilizer package of citric acid/EDTA (CA/EDTA) confers many of the benefits of PEGylation without the drawbacks. Citric acid stabilization by itself is easily washed off the particle surface, however the addition of EDTA generates a shell layer that is very durable even when exposed to high ionic strength solutions. Interestingly, the unique CA/EDTA stabilizer decreases nanoceria aggregation, lengthens circulation time, and limits reticuloendothelial organ deposition, without impairing the catalytic properties of this nanomaterial as it participates in redox reactions [37]. Although this approach to stabilization may not be amenable to all nanomaterials, it is readily applied to metal oxides. The addition of CA/EDTA also allows the particle to retain a smaller hydrodynamic radius compared to a similar particle that has been PEGylated. Importantly, intracellular localization and translocation of the particles across the BBB are improved by the addition of CA/EDTA; removal of the stabilizer prevents extravasation of the particle into the brain parenchyma (unpublished data).

28.5 Techniques to Study Protein Corona Composition

Thus, native nanoparticles can certainly be changed by decoration with proteins. The corona thickness and density determine the overall size of the nanoparticle as well as the underlying nanoparticle surface area exposure. The identity and quantity of adsorbed proteins must also be considered, as these factors regulate biological distribution and route of metabolism. Observation of such characteristics can be accomplished by a few widely employed techniques, including dynamic light scattering (DLS), analytical centrifugation (AC), and size exclusion chromatography (SEC) [38, 39].

Identification of the proteins decorating the corona is generally performed after isolation of the adsorbed protein from the nanomaterial surface. Subsequently, the protein corona is removed from the nanomaterial with high temperatures, high salt concentrations, detergents, enzymes, thiols, or a combination of the aforementioned treatments. Depending on the strength of the nanomaterial–protein interaction, the isolation may result in only partial removal of proteins, which could certainly bias the results. Upon isolation of proteins from the corona, the proteins are generally

separated by size using SDS-polyacrylamide gel electrophoresis (SDS-PAGE), followed by band extraction from the gel, digestion with trypsin [40], and identification via tandem liquid chromatography-mass spectrometry (LC-MS/MS) [13, 41].

Since it is not possible to recover nanoparticles once they are injected into an animal, protein–nanomaterial interactions are often modeled by incubating the nanomaterials in a physiological medium containing plasma or serum to mimic blood protein adsorption. With this technique, the corona forms, the nanomaterial is washed, and the proteins comprising the "hard corona" are identified. However, a drawback is that neither serum nor plasma fully replicates the in vivo environment; serum lacks coagulation factors and plasma lacks blood enzymes. In addition, this strategy does not address the changes in protein composition as the particle traverses through tissues, cells, and organelles.

28.6 In Vivo Properties of Nanomaterials

The goal of nanomedicine is to administer nanomaterials to whole organisms for diagnostic or therapeutic purposes [42]. Since the physical and functional characteristics of nanomaterials are so heavily impacted by formation of the corona, the in vivo behavior of these materials may not directly correlate with observed in vitro effects. Aside from altered biological activity, the nanomaterials may have unexpected or undesired distribution to organ systems. For example, the reticuloendothelial organs, liver, and spleen are common sites of nanoparticle deposition when materials are delivered intravenously or subcutaneously [43]. This distribution pattern is likely due in part to the uptake of nanoparticles by circulating or tissue-resident immune cells (i.e., Kupfer cells) [43], a process that is enhanced by the binding of blood complement proteins to the surface. Though strategies have been implemented in attempt to specifically target nanomaterials to desired locations by the addition of moieties during synthesis [44], more general nanoparticle stabilization (or certainly no stabilization) allows organismal distribution to be driven more by the biochemistry of serum protein interactions.

Deposition of nanomaterials to the liver and spleen can result in tissue toxicity and organ dysfunction [45, 46], in particular if levels of the nanoparticles in these tissues remain elevated over time. Unless the nanomaterial is injected directly into a target tissue or organ, there are relatively low levels detected in other organs following nanomaterial administration [47–49], which can limit the range of applications of such substances. As previously noted, delivery of materials into the brain through the BBB is a challenge regardless of the size and composition. The exception to this may be nanomaterials that are small enough to pass through the tight junctions between endothelial cells making up the BBB. However, to reach the brain in the first place, the nanoparticles must avoid uptake by phagocytic immune cells [43, 50], aggregation of particles themselves [51], and binding to other large proteins that would significantly increase their size [50].

Recently, we developed a method to synthesize very small, monodispersed cerium dioxide nanoparticles (2 nm) that are stabilized with CA/EDTA [37]. Cerium oxide nanoparticles (CeNPs) that are unstabilized or minimally stabilized with sodium citrate have illustrated limited deposition in brain tissues [46, 52, 53], though high levels were detected in liver and spleen tissues. A predominant functional property of CeNPs is their ability to fluctuate between +3 and +4 states, which provides unparalleled antioxidant activity that has been demonstrated in both acellular [54–57] and cellular [58–62] in vitro systems. As noted, however, these in vitro models have limitations in their translation to outcomes in intact biological systems, since the nanomaterials are applied directly to the cells, instead of having to traverse biological compartments in order to reach potential target sites. Thus, though the in vitro capabilities of CeNPs are attractive as a therapeutic mechanism, there is little guarantee that this antoxidant activity would be preserved and observed in vivo regardless of the organ target.

28.7 Use of Murine Multiple Sclerosis Model to Demonstrate Biological Efficacy of Cerium Oxide Nanoparticles

A number of neurodegenerative and autoimmune diseases involve high levels of reactive oxygen species (ROS) and other free radicals as at least part of their pathogenesis. The mechanism of such molecules is to steal electrons from other macromolecules such as DNA and lipids, which results in stabilization of the free radical, but causes damage to cells and cellular structures as a result [63]. Thus, antioxidants that can donate or accept electrons from other molecules, generally free radicals, limit the reactivity of these molecules and theoretically block disease development or progression [64]. Previously we have shown that CeNPs are capable of reducing the accumulation of many biologically relevant free radicals in the brain including superoxide anion, nitric oxide, and peroxynitrite [58, 65]. The disease multiple sclerosis (MS) is a neurodegenerative, autoimmune disease involving the targeting of antigens in the myelin sheath surrounding neurons [66]. Immune cells mediate damage to the myelin, including macrophages that are recruited by activated autoreactive T cells. Macrophages play a critical role in protection from pathogens by producing high amounts of ROS (among other killing mechanisms) [67], but in the context of an autoimmune disease, the ROS are inappropriately targeting a “self” tissue instead of providing protection against infection. Destruction of the myelin sheaths of neurons in MS patients causes poor nerve impulse conduction that can result in cognitive and motor function deficits [66].

The murine model of MS, experimental autoimmune encephalomyelitis (EAE) [68], provides an excellent opportunity to test the efficacy of nanomaterials for several reasons. First, neutralization of free radicals can theoretically provide alleviation of disease pathogenesis, so the efficacy of antioxidant CeNPs could be measured by tracking disease symptom severity. Second, the site of disease pathogenesis in the brain and spinal cord allows assessment of how well nanoceria can be delivered

to these organs without direct administration into these sites. Since EAE development involves disruption of the BBB, additional healthy animals must be similarly treated with nanoparticles to confirm similar brain deposition of nanoceria in organisms with intact BBBs.

Given limited brain deposition of other nanoceria formulations and the understanding that physical and chemical characteristics of nanoparticles affect protein adsorption and delivery, development of a novel nanoceria formulation was undertaken as previously noted. Synthesis of these CeNPs involved stabilization with a unique combination of citrate and EDTA [37]. The zeta potential of these custom-synthesized CeNPs is considerably less negative than other formulations (−23.5 mV), and the particles are smaller (hydrodynamic radius = 2.9 nm). Further, the particles remain highly dispersed in physiological salt solutions and resist pelleting even at very high centrifugation rates (100,000 × *g*) [37]. The protein–nanomaterial interactions of our CeNPs were modeled by incubating the nanomaterial in a physiological medium containing mouse serum to mimic blood protein adsorption. After incubation in the serum, the protein-decorated CeNPs were collected by centrifugation and analyzed. Isolated hard corona proteins were separated via SDS-PAGE and the gel bands were excised from the gel and digested with trypsin, and the resulting peptide mixtures were then extracted from the gel pieces and concentrated on a SpeedVac concentrator and then cleaned with a C18 Micro ZipTip (Millipore). The clean peptide mixtures were then analyzed by nanoliquid chromatography tandem mass spectrometry (nanoLC-MS/MS) using a NanoAcquity UPLC coupled with a QTOF Micro mass spectrometer (both from Waters Corporation). The ions were analyzed in data-dependent mode. The resulting raw data were converted into pkl files using ProteinLynx Global Server (version 2.4. from Waters Corporation) and the pkl files were used to perform database searches using the web-based Mascot database search engine. The Mascot search parameters included: Swissprot database, peptide tolerance of ±0.9 Da, MS/MS tolerance of ±0.5 Da, enzyme used: trypsin, one missed cleavage, and cysteines modified to carbamydomethyl as fixed modification.

The full procedure for protein digestion and peptide extraction, as well as the parameters used for nanoLC-MS/MS and data processing are described in detail elsewhere [69–73]. An example of an outcome of such a proteomics experiment is shown in Table 28.1 (note: the numbers on the right are the Mascot scores; the higher the score, the higher the probability that a protein is identified with high confidence). Far fewer proteins were observed in the corona than predicted [74], likely due to the presence and characteristics of the CA/EDTA stabilizer.

The relatively minimal protein corona generated as a result of the addition of highly durable stabilizers also appears to affect nanoceria biodistribution, as we observed high quantities of CeNPs in the brain [37], relative to those reported by other studies with differing nanoceria formulations [46, 52, 53]. Though ceria was detected in the spleens and livers of healthy mice injected with CeNPs, the levels were nearly 100-fold lower than those measured in animals that received citrate stabilized or unstabilized nanoceria [46, 52, 53]. This unique distribution pattern may be due to high concentrations of ApoA1 and serum albumin (relative to immunoglobulins) bound to the CeNPs. The near impossibility of isolating untagged

Table 28.1 The protein composition of the hard corona of monodispersed cerium dioxide nanoparticles (2 nm) stabilized with CA/EDTA[a]

Proteins found in corona	Score
Alpha-2-macroglobulin	331
Serum albumin	288
Ig alpha chain C region	271
Ig mu chain C region secreted form	249
Hemoglobin subunit beta-1	207
Alpha-1-antitrypsin 1–2	180
Complement C3	172
Apolipoprotein A-I	171
Ig kappa chain V–III region PC 2880/PC 1229	110
Ig gamma-2B chain C region	96
Haptoglobin	82
Ig gamma-2A chain C region, A allele	76
Hemoglobin subunit alpha	61
Alpha-2-macroglobulin	57
Hemopexin	46
Ig heavy chain V region 441	44
Ig gamma-3 chain C region	38
Ig gamma-1 chain C region, membrane-bound form	35
Kininogen-1	33
Serine protease inhibitor A3K	30

[a]Previously unpublished results

CeNPs from such biological tissues necessitates the use of ICP-MS to most efficiently quantify the ceria content of these organs as a proxy for better, direct assessment of nanoparticle content. Transmission electron microscopy (TEM) could be used, but it is time consuming and only qualitative. STEM-EELS could add elemental analysis capability to the TEM technique, but often the deposition of the element of interest is so low that it cannot be detected in thin sections by this method.

Administration of the CeNPs to animals induced with a chronic progressive form of EAE illustrated similar brain deposition to that observed in healthy animals (Fig. 28.2) [37]. TEM imaging of cerebellum tissue harvested from a CeNP-treated animal revealed the presence of the nanoparticles within cells including the mitochondria, neuronal myelin sheaths, and cell membranes (Fig. 28.2). However, simple target tissue deposition was insufficient to demonstrate the retention of the CeNPs' potential biological activity. EAE-induced mice develop progressive loss of motor function, first in the hind limbs and later in the front limbs [75]. A disease assessment paradigm exists to assign a clinical score of disease severity to each animal twice daily, based upon its limb mobility. The scale for these scores is 0–5, with higher scores reflecting more severe symptoms. Intravenous CeNP administration at varying doses (10, 20, or 30 mg/kg) followed a preventative or therapeutic regimen relative to disease induction: preventative: days −1, 0, 1, 3, 7, 14, 21, 28, 35; 3-day therapeutic: days 3, 7, 14, 21, 28, 35; or 7-day therapeutic: days 7, 14, 21, 28, 35 [37]. Using clinical scores as an indication of disease severity, the

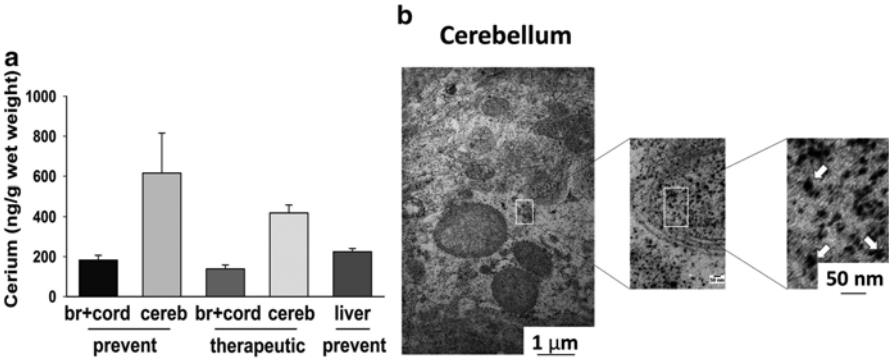


Fig. 28.2 ICP-MS analysis demonstrated the presence of CeNPs in the brains of mice with EAE. Ceria was detected in the cerebellum of CeNP-treated EAE animals (**a**) with intracellular distribution shown by transmission electron microscopy (**b**). *Arrows* indicate nanoparticle location. Reprinted with permission from Heckman, K. L., W. Decoteau, A. Estevez, K. J. Reed, W. Costanzo, D. Sanford, J. C. Leiter, J. Clauss, K. Knapp, C. Gomez, P. Mullen, E. Rathbun, K. Prime, J. Marini, J. Patchefsky, A. S. Patchefsky, R. K. Hailstone, and J. S. Erlichman, 2013, Custom Cerium Oxide Nanoparticles Protect against a Free Radical Mediated Autoimmune Degenerative Disease in the Brain: ACS Nano. Copyright 2013 American Chemical Society

preventative and therapeutic treatment regimens were effective at alleviating EAE symptoms, in particular at higher doses (Fig. 28.3) [37]. Area under the curve (AUC) analysis allows for assessment of the cumulative disease severity throughout the entire disease course (Fig. 28.4).) In fact, the efficacy of the CeNPs was similar to that of the currently used MS drug fingolimod (Figs. 28.3 and 28.4) [37].

The cumulative amount of CeNPs delivered to the EAE mice correlated with the improvements in disease severity (measured by clinical scores) and also with the amount of ceria detected in the brains of treated animals (detected by ICP-MS) (Fig. 28.4) [37]. Given these correlations, it seems logical to conclude that the CeNPs reach the brain and protect against this ROS-mediated disease. However, these metrics do not specifically demonstrate the mechanism utilized by the CeNPs to interfere with disease pathogenesis and thus protect against symptom progression. To test whether levels of ROS in the brain may have been neutralized by the antioxidant properties of the CeNPs, brains were harvested from EAE animals late in the disease course, and sections were stained with CM-DCFDA, a total ROS indicator dye (Fig. 28.5) [37]. Levels of ROS were reduced approximately 30 % in brain sections of CeNP-treated animals, relative to those of control or fingolimod-treated animals, indicating that antioxidant activity was indeed observed in those animals treated with CeNPs and is likely a mechanism of CeNP-mediated disease protection.

Complete analysis of the biological properties and efficacy of nanoceria is not only dependent upon the standard biological metrics of disease severity, but also upon the biochemical techniques that enable both analysis of in vivo ceria deposition and more fundamental in vitro CeNP characterization. However, in context of the in vivo

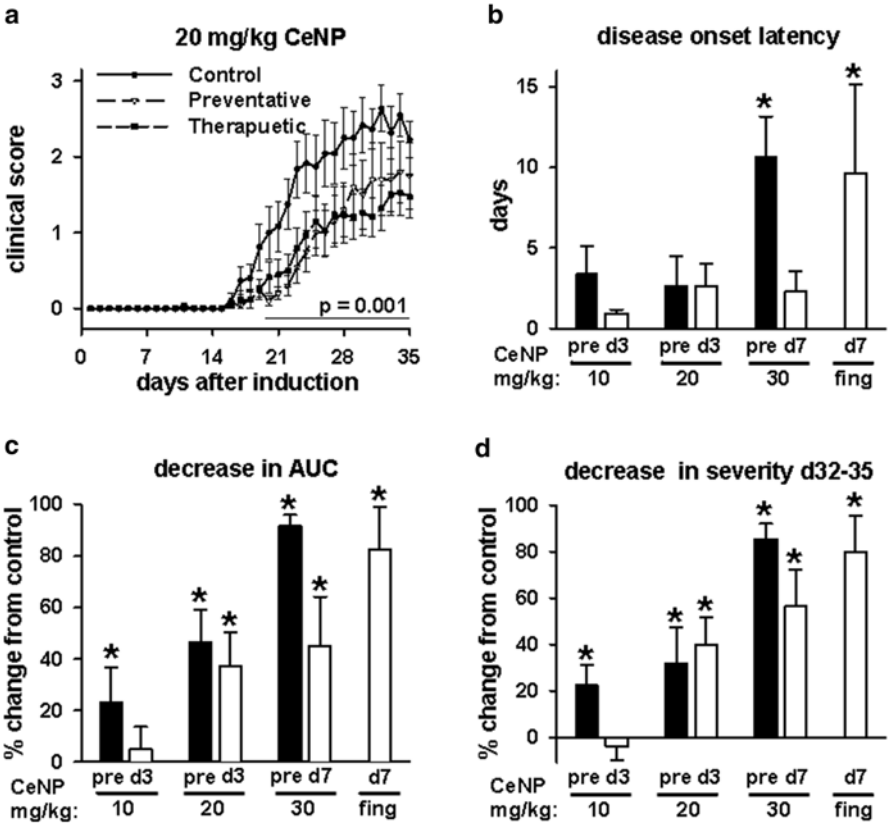


Fig. 28.3 EAE disease severity was alleviated by CeNP treatment in preventative and therapeutic treatment regimens. Higher clinical scores are indicative of worse disease symptoms; (**a**) depicts the kinetics of disease progression in mice treated with 20 mg/kg CeNPs compared to control. Disease onset (**b**) and area under the curve (AUC) analysis (**c**) illustrate the improved efficacy of CeNPs delivered at higher doses, in particular compared to the currently used drug, fingolimod. The longevity of the CeNP treatment is illustrated by decreased clinical scores at late stages of disease (**d**). Reprinted with permission from Heckman, K. L., W. Decoteau, A. Estevez, K. J. Reed, W. Costanzo, D. Sanford, J. C. Leiter, J. Clauss, K. Knapp, C. Gomez, P. Mullen, E. Rathbun, K. Prime, J. Marini, J. Patchefsky, A. S. Patchefsky, R. K. Hailstone, and J. S. Erlichman, 2013, Custom Cerium Oxide Nanoparticles Protect against a Free Radical Mediated Autoimmune Degenerative Disease in the Brain: ACS Nano. Copyright 2013 American Chemical Society

EAE studies, there is still an incomplete understanding of what unique protein corona enables the delivery of the CeNPs into the brain in a configuration that preserves their antioxidant activity. The ideal experiment would be to somehow retrieve the CeNPs from the brain tissue for characterization of bound proteins. Given the small size of the nanoparticles, this is a formidable task, especially given the intracellular compartmentalization that has likely occurred. In the absence of an efficient retrieval

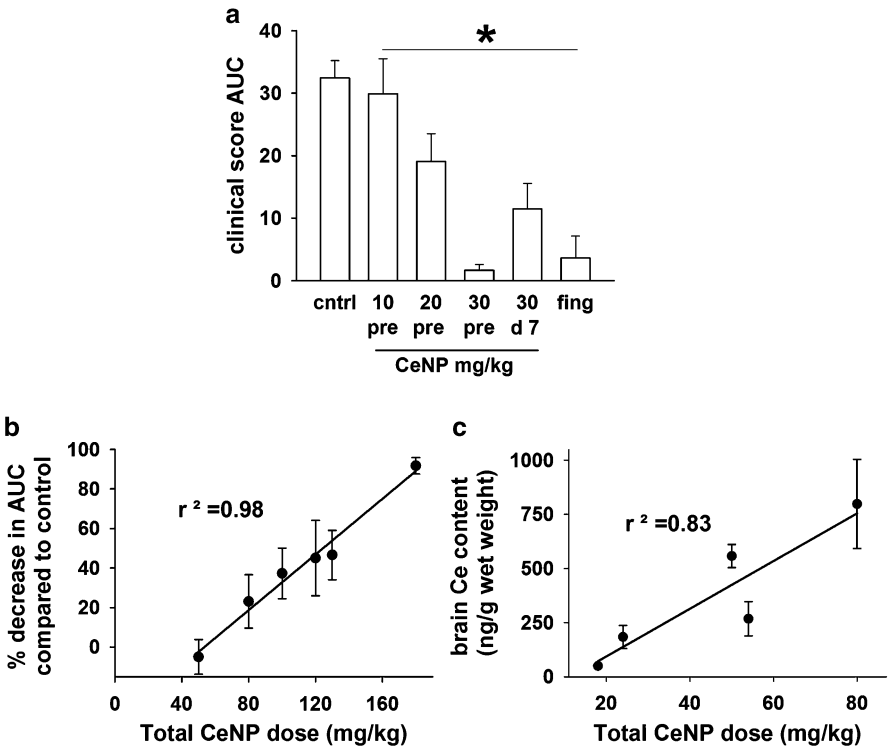


Fig. 28.4 Cumulative disease severity was lessened by CeNP treatment, as dosage correlated with biological activity and tissue deposition. AUC analysis indicates that overall EAE disease severity was worse in control animals and those treated with lower doses of CeNPs (**a**). However, as CeNP dose increased, so did the reduction in clinical disease severity (**b**). Further, the increased deposition of ceria in the brains of CeNP-treated animals (**c**) suggests that the presence of the nanoparticles provided this protection against disease pathogenesis. Reprinted with permission from Heckman, K. L., W. Decoteau, A. Estevez, K. J. Reed, W. Costanzo, D. Sanford, J. C. Leiter, J. Clauss, K. Knapp, C. Gomez, P. Mullen, E. Rathbun, K. Prime, J. Marini, J. Patchefsky, A. S. Patchefsky, R. K. Hailstone, and J. S. Erlichman, 2013, Custom Cerium Oxide Nanoparticles Protect against a Free Radical Mediated Autoimmune Degenerative Disease in the Brain: ACS Nano. Copyright 2013 American Chemical Society

technique, immunohistochemistry analysis of brain sections of CeNP-treated animals could be useful. If the CeNPs could be molecularly tagged in a manner that would not drastically increase size or disrupt protein corona interactions, the nanoparticles could be detected along with staining to detect the co-localization of the nanomaterial with a bound, probable serum protein. Further description of the protein corona that allowed for the successful deposition of these custom CeNPs in the brain (with simultaneous limited deposition in reticuloendothelial organs) will allow for the synthesis of new nanomaterials specifically targeted for brain delivery.

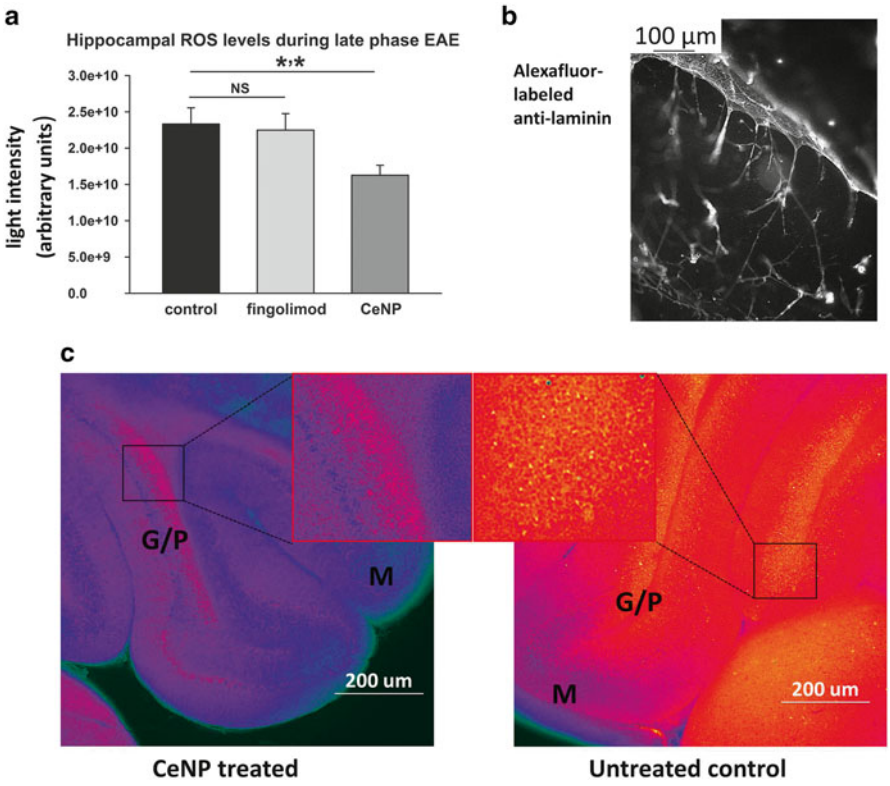


Fig. 28.5 Lower levels of ROS were detected in the hippocampus of CeNP-treated EAE animals. Late in the disease course, hippocampus brain slices were stained with a total ROS indicator dye (**c**); quantitative analysis (**a**) demonstrated that the CeNPs retained their antioxidant activity upon reaching the brain. Microvasculature staining (**b**) and lower magnification analysis of the ROS indicator dye staining in the brain (**c**) demonstrated that reduced ROS levels are found outside the brain vasculature. In particular, ROS levels in the granular cells (typically a site of neurodegeneration in MS) were lower in the CeNP-treated mice compared to controls. Reprinted with permission from Heckman, K. L., W. Decoteau, A. Estevez, K. J. Reed, W. Costanzo, D. Sanford, J. C. Leiter, J. Clauss, K. Knapp, C. Gomez, P. Mullen, E. Rathbun, K. Prime, J. Marini, J. Patchefsky, A. S. Patchefsky, R. K. Hailstone, and J. S. Erlichman, 2013, Custom Cerium Oxide Nanoparticles Protect against a Free Radical Mediated Autoimmune Degenerative Disease in the Brain: ACS Nano. Copyright 2013 American Chemical Society

28.8 Summary

Nanotechnology is a "toolbox" that provides nanosized building blocks for new materials, devices, and systems for broad scientific applications. Engineered nanomaterials have unique physical–chemical properties that make them promising candidates for many biomedical applications, including tissue regeneration, drug,

and gene delivery, and in vivo monitoring of disease processes. The burgeoning possibilities of nano-scale structure manipulations will likely increase. Furthermore, because cellular molecules and man-made nanoparticles have similar dimensions, the latter may directly interact with vital cellular processes. Understanding the nano-biological interface requires a thorough understanding of factors that influence protein adsorption which are central in dictating biodistribution, fate and, to varying degrees, biological action. Consequently, new biological testing paradigms must be developed to examine the physiological effects of these man-made materials to assess both their biomedical potential as well as toxicology. Mass spectroscopy in its various forms (ICP-MS, LCMS) is a critical tool in the development of this technology in biological settings.

Acknowledgments The authors acknowledge Armand G. Ngounou Wetie and Dr. Costel C. Darie (Biochemistry & Proteomics Group, Clarkson University) for the mass spectrometry work.

References

1. Pardridge W (2003) Blood–brain barrier drug targeting: the future of brain drug development. Mol Interv 3:90–105
2. Gulyaev AE, Gelperina SE, Skidan IN, Antropov AS, Kivman GY, Kreuter J (1999) Significant transport of doxorubicin into the brain with polysorbate 80-coated nanoparticles. Pharm Res 16:1564–1569
3. Hekmatara T, Bernreuther C, Khalansky AS, Theisen A, Weissenberger J, Matschke J, Gelperina S, Kreuter J, Glatzel M (2009) Efficient systemic therapy of rat glioblastoma by nanoparticle-bound doxorubicin is due to antiangiogenic effects. Clin Neuropathol 28:153–164
4. Steiniger SC, Kreuter J, Khalansky AS, Skidan IN, Bobruskin AI, Smirnova ZS, Severin SE, Uhl R, Kock M, Geiger KD, Gelperina SE (2004) Chemotherapy of glioblastoma in rats using doxorubicin-loaded nanoparticles. Int J Cancer 109:759–767
5. Alyautdin RN, Petrov VE, Langer K, Berthold A, Kharkevich DA, Kreuter J (1997) Delivery of loperamide across the blood–brain barrier with polysorbate 80-coated polybutylcyanoacrylate nanoparticles. Pharm Res 14:325–328
6. Ulbrich K, Hekmatara T, Herbert E, Kreuter J (2009) Transferrin- and transferrin-receptor-antibody-modified nanoparticles enable drug delivery across the blood–brain barrier (BBB). Eur J Pharm Biopharm 71:251–256
7. Friese A, Seiller E, Quack G, Lorenz B, Kreuter J (2000) Increase of the duration of the anticonvulsive activity of a novel NMDA receptor antagonist using poly(butylcyanoacrylate) nanoparticles as a parenteral controlled release system. Eur J Pharm Biopharm 49:103–109
8. Alyautdin RN, Tezikov EB, Ramge P, Kharkevich DA, Begley DJ, Kreuter J (1998) Significant entry of tubocurarine into the brain of rats by adsorption to polysorbate 80-coated polybutylcyanoacrylate nanoparticles: an in situ brain perfusion study. J Microencapsul 15:67–74
9. Gao K, Jiang X (2006) Influence of particle size on transport of methotrexate across blood brain barrier by polysorbate 80-coated polybutylcyanoacrylate nanoparticles. Int J Pharm 310:213–219
10. Wilson B, Samanta MK, Santhi K, Kumar KP, Paramakrishnan N, Suresh B (2008) Targeted delivery of tacrine into the brain with polysorbate 80-coated poly(n-butylcyanoacrylate) nanoparticles. Eur J Pharm Biopharm 70:75–84
11. Lai KK, Kolippakkam D, Beretta L (2008) Comprehensive and quantitative proteome profiling of the mouse liver and plasma. Hepatology 47:1043–1051

12. Monopoli MP, Walczyk D, Campbell A, Elia G, Lynch I, Bombelli FB, Dawson KA (2011) Physical-chemical aspects of protein corona: relevance to in vitro and in vivo biological impacts of nanoparticles. J Am Chem Soc 133:2525–2534
13. Cedervall T, Lynch I, Foy M, Berggård T, Donnelly SC, Cagney G, Linse S, Dawson KA (2007) Detailed identification of plasma proteins adsorbed on copolymer nanoparticles. Angew Chem Int Ed Engl 46:5754–5756
14. Kratz F, Elsadek B (2012) Clinical impact of serum proteins on drug delivery. J Control Release 161:429–445
15. Mirshafiee V, Mahmoudi M, Lou K, Cheng J, Kraft ML (2013) Protein corona significantly reduces active targeting yield. Chem Commun (Camb) 49:2557–2559
16. Salvati A, Pitek AS, Monopoli MP, Prapainop K, Bombelli FB, Hristov DR, Kelly PM, Åberg C, Mahon E, Dawson KA (2013) Transferrin-functionalized nanoparticles lose their targeting capabilities when a biomolecule corona adsorbs on the surface. Nat Nanotechnol 8:137–143
17. Lynch I, Dawson KA (2008) Protein-nanoparticle interactions. NanoToday 3:40–47
18. Sahoo B, Goswami M, Nag S, Maiti S (2007) Spontaneous formation of a protein corona prevents the loss of quantum dot fluorescence in physiological buffers. Chem Phys Lett 445:217–220
19. Dutta D, Sundaram SK, Teeguarden JG, Riley BJ, Fifield LS, Jacobs JM, Addleman SR, Kaysen GA, Moudgil BM, Weber TJ (2007) Adsorbed proteins influence the biological activity and molecular targeting of nanomaterials. Toxicol Sci 100:303–315
20. Tenzer S, Docter D, Kuharev J, Musyanovych A, Fetz V, Hecht R, Schlenk F, Fischer D, Kiouptsi K, Reinhardt C, Landfester K, Schild H, Maskos M, Knauer SK, Stauber RH (2013) Rapid formation of plasma protein corona critically affects nanoparticle pathophysiology. Nat Nanotechnol 8:772–781
21. Dobrovolskaia MA, Aggarwal P, Hall JB, McNeil SE (2008) Preclinical studies to understand nanoparticle interaction with the immune system and its potential effects on nanoparticle biodistribution. Mol Pharm 5:487–495
22. Linse S, Cabaleiro-Lago C, Xue WF, Lynch I, Lindman S, Thulin E, Radford SE, Dawson KA (2007) Nucleation of protein fibrillation by nanoparticles. Proc Natl Acad Sci U S A 104:8691–8696
23. Aggarwal P, Hall JB, McLeland CB, Dobrovolskaia MA, McNeil SE (2009) Nanoparticle interaction with plasma proteins as it relates to particle biodistribution, biocompatibility and therapeutic efficacy. Adv Drug Deliv Rev 61:428–437
24. Dell'Orco D, Lundqvist M, Oslakovic C, Cedervall T, Linse S (2010) Modeling the time evolution of the nanoparticle-protein corona in a body fluid. PLoS One 5:e10949
25. Vroman L, Adams AL, Fischer GC, Munoz PC (1980) Interaction of high molecular weight kininogen, factor XII, and fibrinogen in plasma at interfaces. Blood 55:156–159
26. Hellstrand E, Lynch I, Andersson A, Drakenberg T, Dahlbäck B, Dawson KA, Linse S, Cedervall T (2009) Complete high-density lipoproteins in nanoparticle corona. FEBS J 276: 3372–3381
27. Walkey CD, Chan WC (2012) Understanding and controlling the interaction of nanomaterials with proteins in a physiological environment. Chem Soc Rev 41:2780–2799
28. Karmali PP, Simberg D (2011) Interactions of nanoparticles with plasma proteins: implication on clearance and toxicity of drug delivery systems. Expert Opin Drug Deliv 8:343–357
29. Lundqvist M, Stigler J, Elia G, Lynch I, Cedervall T, Dawson KA (2008) Nanoparticle size and surface properties determine the protein corona with possible implications for biological impacts. Proc Natl Acad Sci U S A 105:14265–14270
30. Sheng Y, Liu C, Yuan Y, Tao X, Yang F, Shan X, Zhou H, Xu F (2009) Long-circulating polymeric nanoparticles bearing a combinatorial coating of PEG and water-soluble chitosan. Biomaterials 30:2340–2348
31. Knop K, Hoogenboom R, Fischer D, Schubert US (2010) Poly(ethylene glycol) in drug delivery: pros and cons as well as potential alternatives. Angew Chem Int Ed Engl 49:6288–6308
32. Kim YK, Minai-Tehrani A, Lee JH, Cho CS, Cho MH, Jiang HL (2013) Therapeutic efficiency of folated poly(ethylene glycol)-chitosan-graft-polyethylenimine-Pdcd4 complexes in H-ras12V mice with liver cancer. Int J Nanomedicine 8:1489–1498

33. Lundqvist M, Stigler J, Cedervall T, Berggård T, Flanagan MB, Lynch I, Elia G, Dawson K (2011) The evolution of the protein corona around nanoparticles: a test study. ACS Nano 5:7503–7509
34. Walkey CD, Olsen JB, Guo H, Emili A, Chan WC (2012) Nanoparticle size and surface chemistry determine serum protein adsorption and macrophage uptake. J Am Chem Soc 134:2139–2147
35. Hatakeyama H, Akita H, Harashima H (2013) The polyethyleneglycol dilemma: advantage and disadvantage of PEGylation of liposomes for systemic genes and nucleic acids delivery to tumors. Biol Pharm Bull 36:892–899
36. Ishida T, Wang X, Shimizu T, Nawata K, Kiwada H (2007) PEGylated liposomes elicit an anti-PEG IgM response in a T cell-independent manner. J Control Release 122:349–355
37. Heckman KL, Decoteau W, Estevez A, Reed KJ, Costanzo W, Sanford D, Leiter JC, Clauss J, Knapp K, Gomez C, Mullen P, Rathbun E, Prime K, Marini J, Patchefsky J, Patchefsky AS, Hailstone RK, Erlichman JS (2013) Custom cerium oxide nanoparticles protect against a free radical mediated autoimmune degenerative disease in the brain. ACS Nano 7:10582–10596
38. Cedervall T, Lynch I, Lindman S, Berggård T, Thulin E, Nilsson H, Dawson KA, Linse S (2007) Understanding the nanoparticle-protein corona using methods to quantify exchange rates and affinities of proteins for nanoparticles. Proc Natl Acad Sci U S A 104:2050–2055
39. Walczyk D, Bombelli FB, Monopoli MP, Lynch I, Dawson KA (2010) What the cell "sees" in bionanoscience. J Am Chem Soc 132:5761–5768
40. Shevchenko A, Wilm M, Vorm O, Mann M (1996) Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. Anal Chem 68:850–858
41. Capriotti AL, Caracciolo G, Caruso G, Cavaliere C, Pozzi D, Samperi R, Laganà A (2010) Analysis of plasma protein adsorption onto DC-Chol-DOPE cationic liposomes by HPLC-CHIP coupled to a Q-TOF mass spectrometer. Anal Bioanal Chem 398:2895–2903
42. Palombo M, Deshmukh M, Myers D, Gao J, Szekely Z, Sinko PJ (2014) Pharmaceutical and toxicological properties of engineered nanomaterials for drug delivery. Annu Rev Pharmacol Toxicol 54:581–598
43. Almeida JP, Chen AL, Foster A, Drezek R (2011) In vivo biodistribution of nanoparticles. Nanomedicine (Lond) 6:815–835
44. Nam J, Won N, Bang J, Jin H, Park J, Jung S, Park Y, Kim S (2013) Surface engineering of inorganic nanoparticles for imaging and therapy. Adv Drug Deliv Rev 65:622–648
45. Nalabotu SK, Kolli MB, Triest WE, Ma JY, Manne ND, Katta A, Addagarla HS, Rice KM, Blough ER (2011) Intratracheal instillation of cerium oxide nanoparticles induces hepatic toxicity in male Sprague–Dawley rats. Int J Nanomedicine 6:2327–2335
46. Tseng MT, Lu X, Duan X, Hardas SS, Sultana R, Wu P, Unrine JM, Graham U, Butterfield DA, Grulke EA, Yokel RA (2012) Alteration of hepatic structure and oxidative stress induced by intravenous nanoceria. Toxicol Appl Pharmacol 260:173–182
47. Hirst SM, Karakoti A, Singh S, Self W, Tyler R, Seal S, Reilly CM (2013) Bio-distribution and in vivo antioxidant effects of cerium oxide nanoparticles in mice. Environ Toxicol 28:107–118
48. Yokel RA, Au TC, MacPhail R, Hardas SS, Butterfield DA, Sultana R, Goodman M, Tseng MT, Dan M, Haghnazar H, Unrine JM, Graham UM, Wu P, Grulke EA (2012) Distribution, elimination, and biopersistence to 90 days of a systemically introduced 30 nm ceria-engineered nanomaterial in rats. Toxicol Sci 127:256–268
49. Yokel RA, Florence RL, Unrine JM, Tseng MT, Graham UM, Wu P, Grulke EA, Sultana R, Hardas SS, Butterfield DA (2009) Biodistribution and oxidative stress effects of a systemically-introduced commercial ceria engineered nanomaterial. Nanotoxicology 3:234–248
50. Dobrovolskaia MA, McNeil SE (2007) Immunological properties of engineered nanomaterials. Nat Nanotechnol 2:469–478
51. Uskoković V (2013) Entering the era of nanoscience: time to be so small. J Biomed Nanotechnol 9:1441–1470
52. Hardas SS, Butterfield DA, Sultana R, Tseng MT, Dan M, Florence RL, Unrine JM, Graham UM, Wu P, Grulke EA, Yokel RA (2010) Brain distribution and toxicological evaluation of a systemically delivered engineered nanoscale ceria. Toxicol Sci 116:562–576

53. Hardas SS, Sultana R, Warrier G, Dan M, Florence RL, Wu P, Grulke EA, Tseng MT, Unrine JM, Graham UM, Yokel RA, Butterfield DA (2012) Rat brain pro-oxidant effects of peripherally administered 5 nm ceria 30 days after exposure. Neurotoxicology 33:1147–1155
54. Campbell CT, Peden CH (2005) Chemistry. Oxygen vacancies and catalysis on ceria surfaces. Science 309:713–714
55. Schalow T, Laurin M, Brandt B, Schauermann S, Guimond S, Kuhlenbeck H, Starr DE, Shaikhutdinov SK, Libuda J, Freund HJ (2005) Oxygen storage at the metal/oxide interface of catalyst nanoparticles. Angew Chem Int Ed Engl 44:7601–7605
56. Sayle TX, Molinari M, Das S, Bhatta UM, Möbus G, Parker SC, Seal S, Sayle DC (2013) Environment-mediated structure, surface redox activity and reactivity of ceria nanoparticles. Nanoscale 5:6063–6073
57. Andreescu S, Ornatska M, Erlichman JE, Estevez A, Leiter JC (2012) Biomedical applications of metal oxide nanoparticles. In: Matijevic E (ed) Fine particles in medicine and pharmacy. Springer, New York, pp 57–100
58. Estevez AY, Pritchard S, Harper K, Aston JW, Lynch A, Lucky JJ, Ludington JS, Chatani P, Mosenthal WP, Leiter JC, Andreescu S, Erlichman JS (2011) Neuroprotective mechanisms of cerium oxide nanoparticles in a mouse hippocampal brain slice model of ischemia. Free Radic Biol Med 51:1155–1163
59. Tarnuzzer RW, Colon J, Patil S, Seal S (2005) Vacancy engineered ceria nanostructures for protection from radiation-induced cellular damage. Nano Lett 5:2573–2577
60. Schubert D, Dargusch R, Raitano J, Chan SW (2006) Cerium and yttrium oxide nanoparticles are neuroprotective. Biochem Biophys Res Commun 342:86–91
61. Colon J, Herrera L, Smith J, Patil S, Komanski C, Kupelian P, Seal S, Jenkins DW, Baker CH (2009) Protection from radiation-induced pneumonitis using cerium oxide nanoparticles. Nanomedicine 5:225–231
62. Estevez AY, Erlichman JS (2011) Cerium oxide nanoparticles for the treatment of neurological oxidative stress diseases. In: Andreescu S (ed) Oxidative stress: diagnostics, prevention, and therapy. American Chemical Society, Washington, DC, pp 255–288
63. Circu ML, Aw TY (2010) Reactive oxygen species, cellular redox systems, and apoptosis. Free Radic Biol Med 48:749–762
64. Kancheva VD, Kasaikina OT (2013) Bio-antioxidants—a chemical base of their antioxidant activity and beneficial effect on human health. Curr Med Chem 20:4784–4805
65. Ganesana M, Erlichman JS, Andreescu S (2012) Real-time monitoring of superoxide accumulation and antioxidant activity in a brain slice model using an electrochemical cytochrome c biosensor. Free Radic Biol Med 53:2240–2249
66. Hemmer B, Nessler S, Zhou D, Kieseier B, Hartung HP (2006) Immunopathogenesis and immunotherapy of multiple sclerosis. Nat Clin Pract Neurol 2:201–211
67. Raivich G, Banati R (2004) Brain microglia and blood-derived macrophages: molecular profiles and functional roles in multiple sclerosis and animal models of autoimmune demyelinating disease. Brain Res Brain Res Rev 46:261–281
68. Mix E, Meyer-Rienecker H, Hartung HP, Zettl UK (2010) Animal models of multiple sclerosis–potentials and limitations. Prog Neurobiol 92:386–404
69. Ngounou Wetie AG, Sokolowska I, Woods AG, Wormwood KL, Dao S, Patel S, Clarkson BD, Darie CC (2013) Automated mass spectrometry-based functional assay for the routine analysis of the secretome. J Lab Autom 18:19–29
70. Ngounou Wetie AG, Sokolowska I, Woods AG, Darie CC (2013) Identification of post-translational modifications by mass spectrometry. Aust J Chem 66:734–748
71. Petrareanu C, Macovei A, Sokolowska I, Woods AG, Lazar C, Radu GL, Darie CC, Branza-Nichita N (2013) Comparative proteomics reveals novel components at the plasma membrane of differentiated HepaRG cells and different distribution in hepatocyte-and biliary-like cells. PLoS One 8:e71859
72. Sokolowska I, Ngounou Wetie AG, Woods AG, Darie CC (2013) Applications of mass spectrometry in proteomics. Aust J Chem 66:721–733

73. Sokolowska I, Ngounou Wetie AG, Roy U, Woods AG, Darie CC (2013) Mass spectrometry investigation of glycosylation on the NXS/T sites in recombinant glycoproteins. Biochim Biophys Acta 1834:1474–1483
74. Monopoli MP, Aberg C, Salvati A, Dawson KA (2012) Biomolecular coronas provide the biological identity of nanosized materials. Nat Nanotechnol 7:779–786
75. Ledeboer A, Wierinckx A, Bol JG, Floris S, Renardel de Lavalette C, De Vries HE, van den Berg TK, Dijkstra CD, Tilders FJ, van dam AM (2003) Regional and temporal expression patterns of interleukin-10, interleukin-10 receptor and adhesion molecules in the rat spinal cord during chronic relapsing EAE. J Neuroimmunol 136:94–103

Chapter 29
Bottlenecks in Proteomics

Armand G. Ngounou Wetie, Devon A. Shipp, and Costel C. Darie

Abstract Mass spectrometry (MS) is the core for advanced methods in proteomic experiments. When effectively used, proteomics may provide extensive information about proteins and their post-translational modifications, as well as their interaction partners. However, there are also many problems that one can encounter during a proteomic experiment, including, but not limited to sample preparation, sample fractionation, sample analysis, data analysis & interpretation, and biological significance. Here we discuss some of the problems that researchers should be aware of when performing a proteomic experiment.

29.1 Introduction

Mass spectrometry (MS) is the core for advanced methods in proteomic experiments. Integration of MS with a variety of other analytical methods has made it possible to examine virtually all types of samples derived from tissues, organs, and organisms and has led to the identification and quantification of thousands of proteins and peptides from complex biological samples. Unfortunately, due to expense and time requirements, MS-based proteomics remain prohibitive for many labs that could greatly benefit from this technology.

A.G. Ngounou Wetie • C.C. Darie (✉)
Biochemistry & Proteomics Group, Department of Chemistry & Biomolecular Science, Clarkson University, 8 Clarkson Avenue, Potsdam, NY 13699-5810, USA
e-mail: cdarie@clarkson.edu

D.A. Shipp
Polymer Science & Materials Group, Department of Chemistry & Biomolecular Science, Clarkson University, 8 Clarkson Avenue, Potsdam, NY 13699-5810, USA

A.G. Woods and C.C. Darie (eds.), *Advancements of Mass Spectrometry in Biomedical Research*, Advances in Experimental Medicine and Biology 806,
DOI 10.1007/978-3-319-06068-2_29, © Springer International Publishing Switzerland 2014

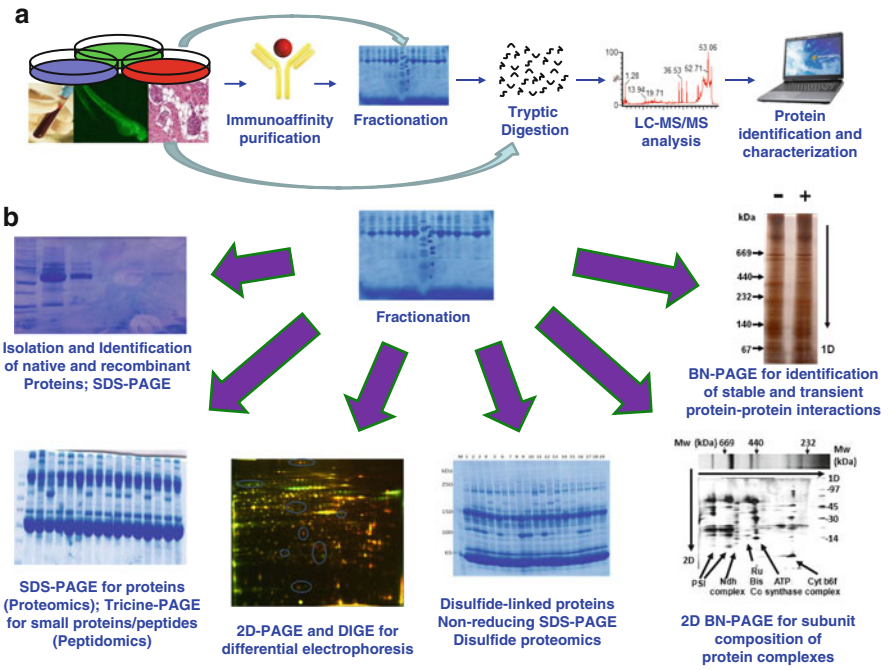


Fig. 29.1 (**a**) Common steps of a proteomic experiment. In many experiments, the samples are fractionated by electrophoresis and then digested in gel. When the biological materials are fluids (e.g., blood, saliva, etc), the electrophoresis-based fractionation step may be bypassed. (**b**) The types of protein samples and the types of electrophoresis that could benefit from a degradable gel

Proteomics is the study of proteome which is the whole protein complement in a cell or organism at any given time [1–8]. The proteome of an organism or even a single type of cell is much more complex than its corresponding genome. This is mostly due to alternation that can be introduced via alternate splicing and post-translational modifications (PTMs) which affect virtually all proteins. A proteome differs from cell to cell and from time to time with composition depending on the physiological or pathological state of cells or organisms. Understanding that proteins are the actual effectors of biological function has been an essential part of biochemistry for over a 100 years [9]. Due to complexity, the analysis of a proteome is extremely challenging. Therefore, modern biochemical technologies with improved separation and identification methods need to be introduced.

The workflow in a proteomic experiment involves sample fractionation by a biochemical approach, followed by enzymatic digestion (usually trypsin), peptide extraction, and MS analysis (Fig. 29.1a) [10–17]. When the peptide mixture is analyzed by MALDI-MS, the proteins of interest are identified using a procedure named peptide mass fingerprinting. Alternatively, the peptide mixture is further fractionated by HPLC on different columns (usually reverse-phase-HPLC or -LC), followed by ESI-MS analysis. The combination of LC and ESI-MS is usually named LC-MS/MS, and analysis of a protein using this approach provides not only the

protein identity, but also sequence information for that particular protein. In addition to qualitative information provided by MALDI-MS or LC-MS/MS analysis, MS may also provide quantitative information about a particular protein. Methods such as differential gel electrophoresis (DIGE) [18], isotope-coded affinity tag (ICAT) [19], stable isotope labeling by amino acids in cell culture (SILAC) [20], absolute quantitation (AQUA) [21], multiple reaction monitoring (MRM) [22], or spectral counting [23] allow detection, identification, and quantification of proteins or peptides. These methods are currently known as functional proteomics and are widely used in basic research.

Another dimension to identification and characterization of proteins is added due to the intensive PTMs of proteins. It is a great challenge to fully identify a PTM pattern at any given time in cells, tissues, and organisms and MS-based proteomics became the method of choice for their detection and characterization [24]. PTMs are important to virtually all biological processes. For example, glycosylation enhances many biological processes such as cell–cell recognition and influences the protein's biological activity [25, 26], while phosphorylation is a reversible and common PTM that plays a role in controlling and modifying the majority of cellular processes [27]. Another important modification of proteins is the formation of disulfide bridges, which plays a significant role in maintaining correct protein function [28, 29]. Additional stable or transient modifications in proteins include acetylation, ubiquitination, and methylation, which when combined with stable or transient protein–protein interactions (PPIs) [30], add an additional level of complexity to proteomic approaches [31].

29.2 Bottleneck # 1: Sample Preparation and Separation

Purification and fractionation steps are to be carefully selected to avoid loss of samples and introduction of errors, which would be reflected later in the MS analysis. Usually, the first step in proteomic workflows is the extraction of proteins from tissues, cells, or biological fluids. At this stage of sample preparation, a loss of material is usually encountered as the homogenization or alternative extraction procedures (e.g., centrifugation) are not very effective. Coevally, one should attempt to reduce sample complexity and matrix effects by isolating and enriching for the fractions of interest (subproteomes, organelle proteomes). To this end, subcellular fractionation is sometimes undertaken. However, any additional sample preparation steps that are performed increase the risk of introducing technical variability and contamination in the analysis. Also, the physiological state/time point of protein extraction from the source plays an important role in the outcome of the study and the biological significance that can be inferred from it. Therefore, the design of proteomic studies should be thoroughly thought through in order to draw meaningful conclusions from the experiments. Further, it should be noted that most sample preparation and purification techniques discriminate against "extreme" proteins (very hydrophobic/hydrophilic proteins, too low/high pI, low abundant, membrane proteins).

29.3 Bottleneck # 2: Gel Fractionation

Key parameters of MS-based proteomic experiments are sensitivity, resolution, dynamic range, and mass accuracy. Therefore, various elements need to be taken into consideration during a typical MS-based proteomic experiment (Fig. 29.1). One of the most important bottlenecks in these experiments is always the sample fractionation and processing, which happens prior to MS analysis. Sample fractionation is primarily performed by some kind of electrophoresis. Sample processing involves in-gel protein digestion, followed by peptide extraction and concentration, and MS analysis. This is a very time-consuming procedure. An alternative to electrophoresis and in-gel digestion steps is the in-solution digestion of proteins. The difference between the two methods is time and labor, as well as the scientific outcome. For example, protein fractionation by electrophoresis and in-gel digestion of proteins, up to the MS analysis step can take up to 3 days, while the in-solution digestion of proteins up to the MS analysis can take only about 5 h. The labor required when the electrophoresis step and in-gel digestion is used can be as much as 24 h of labor, while the in-solution digestion step typically requires a maximum of 3 h. The in-solution digestion step is advantageous over the electrophoresis and in-gel digestion step because of the time and labor savings. The reason for which many scientists do not use in-solution digestion is because of the scientific outcome. For example, without the SDS-PAGE electrophoretic fractionation step, one cannot create molecular mass constraints to monitor particular proteins (i.e., the focus is only on low abundant 50 kDa protein), simply because the samples without fractionation are analyzed as a whole and the low abundant proteins may not be identified by MS analysis. In addition, when the proteins are investigated for identification of their amino acid sequence information or for identification of their PTMs such as phosphorylation, acetylation, or glycosylation, and the proteins that bear these PTMs are not abundant, the SDS-PAGE step is almost mandatory. Furthermore, when the stable and transient PPIs are investigated, native gel electrophoresis step is a must for biochemical fractionation.

Native gel electrophoresis (Blue Native-PAGE or BN-PAGE and Colorless Native-PAGE or CN-PAGE) allows protein complexes from various sources to be separated according to their molecular weight (MW) and external charge (BN-PAGE) or based on the internal charge of subunits in protein complexes (CN-PAGE) [32, 33]. Therefore, to solve one of the most critical bottlenecks in proteomics, it would be ideal to combine the power of protein fractionation (electrophoresis) with the speed (low time and labor) of in-solution digestion, and to obtain the optimal scientific outcome, similar to the electrophoresis and in-gel digestion procedure. In other words, to use electrophoresis step for protein fractionation but not use the in-gel digestion; but to rather dissolve the gel followed by the low-time, low-labor in-solution digestion step and MS analysis.

One solution to this could be to use degradable gels, thus addressing the bottleneck issue and bypassing the in-gel protein digestion and peptide fragmentation by simply replacing the current electrophoresis gel crosslinker (*N*,*N*′-methylenebisacrylamide (MBA)) with a cleavable one. There are various types of

gel monomers used to build electrophoresis gels. The most frequently used are an acrylamide-MBA mixture, with MBA used as a crosslinker. Alternatives to MBA as a crosslinker are cleavable crosslinkers, which can be used to generate reversible polyacrylamide gels. They can be divided into those cleavable by oxidation (e.g., *N*,*N*′-1,2-dihydroxyethylenebisacrylamide or *N*,*N*′-diallyl-tartar-diamide) [34, 35] and those which undergo reductive cleavage (e.g., *N*,*N*′-bisacryloyl-cystamine) [36]. Recently, a modified gel system based on co-polymerization of acrylamide with MBA and ethylene-glycol-diacrylate to offer controllable pore size was also communicated [37]. However, there are numerous problems with these crosslinkers, which are far from being optimal for proteomic studies. For example, one of these crosslinkers (*N*,*N*′-bis(acryloyl)cystamine) is a disulfide-linked (commercially available) crosslinker. This crosslinker can be reduced by incubating the gel or gel piece in a reducing agent such as dithiothreitol or beta-mercaptoethanol. The end-effect is disintegration of the gel piece and release of the protein. However, experiments in our lab identified several problems with this crosslinker: (1) since this is a disulfide linker, the gel has to be run under non-reducing conditions, thus preventing one to investigate the cysteine-containing proteins, (2) the protein recovery efficiency is very low (we recovered about 10–20 % of the initially loaded protein), (3) the proteins that are extracted from the gel are over-alkylated, thus complicating the protein identification due to artificial alkylation, (4) the method is environmentally unfriendly, since it creates a large amount of reducing agent (e.g., beta-mercaptoethanol), (5) the protein is very diluted and requires concentration (the protein within the gel piece is usually incubated in 100 mM beta-mercaptoethanol). Therefore, designing new crosslinkers that are compatible with proteomic experiments, does not artificially modify proteins, allows the protein to be recovered with a high efficiency, and is both time- and labor efficient is important and would solve one of the most critical bottlenecks in proteomics.

29.4 Bottleneck # 3: Ionization

Another important bottleneck in MS-based proteomics is related to the ionization method. It is imperative to choose the appropriate matrices (MALDI), solvents, and salt for a specific analysis under consideration (ESI). Of the two mostly used soft ionization techniques (MALDI and ESI), it has been shown that ESI is the more appropriate for fragile and/or labile biomacromolecules. With ESI, in-source fragmentation is limited and PTMs as well as non-covalent attachments are preserved, whereas MALDI is more prone to loss of information. Another issue with MALDI is the high background generated by matrix ions that increases the noise and could hamper identification of low-abundant biomolecules. It has been suggested that the organic matrix could be replaced with an inorganic nanomaterial surface to increase the signal-to-noise ratio [38].

29.5 Bottleneck # 4: Data Analysis and Interpretation

Post-MS analysis, a huge amount of data generated has to be processed and converted into biological meaning. This requires understanding of the fragmentation pattern of the biomacromolecules at hand so that useful structural information can be obtained from the data. This task can be undertaken manually (this is time-consuming and requires a lot of expertise and skills) or can be done with the help of a multitude of available software. The goal here will be to extract biological information from mass spectrometric data through a completely simple automated approach that requests little time and effort from the user. One particular issue encountered during data analysis is the different outcome resulting from the analysis of MS data using different software which makes the interpretation of results somehow uncertain. Though these software programs are mostly very sophisticated, they are not always easy to use and to understand [10, 39]. In the case of MS/MS, the availability of reference spectra is of critical importance. Identification of novel proteins or copolymers is rendered difficult or could even be missed if corresponding reference spectra are absent. In this case, one must manually investigate the data in order to identify these novel entities. This is, for example, observable during Mascot searches, where peptides with very good scores are not assigned by the software for reasons mentioned above or some other unknown factors (organism with incomplete or unavailable genome sequence, novel isoforms of a protein). Therefore, researchers have to perform at least two searches (e.g., Mascot and Sequest) for confident identification [40, 41]. This holds true also for software used for visualization/presentation or quantitation of MS raw data such as Scaffold or DTAselect [42, 43].

As for the analysis of synthetic or derivatives of biomacromolecules, tandem MS libraries should be created and made accessible to the whole scientific community. Moreover, an effort should be undertaken to gather and organize all the published MS-based proteomic data in an openly accessible database so that researchers can gain access to them for analysis, evaluation, reproducibility, interpretation, and extraction of information. The idea is to store proteomic data obtained with different MS instruments in a single global database in a format that is compatible with a free online tool. An example of such a format is "mzXML," an open, generic XML (extensible markup language) for MS, MS/MS, and MS^n data output. We believe that this should be a requirement for all manuscript submissions to proteomic journals. The idea of a centralized, organized, structured, and openly accessible MS-data storage center will advance proteomic research tremendously [44–46].

29.6 Bottleneck # 5: False Positives, False Negatives, and Unassigned Spectra

Although this is really part of data analysis, this topic deserves special consideration. It is a great challenge to do a database search (i.e., using Mascot database search engine) and to realize that many, sometimes too many MS/MS spectra are unassigned to any peptide. In addition, some MS/MS spectra that correspond to a

peptide are not in the Mascot database search, due to the search parameters, i.e., an oxidized peptide will not be identified if we do not choose to search for it, or a peptide with two missed tryptic cleavage sites will not be identified if we use only one missed cleavage site during the search. Furthermore, searching in Mascot database with or without the decoy function eliminates some true positive spectra and only inspection of the raw data elucidates whether a MS/MS spectrum indeed corresponds to a peptide with a particular amino acid sequence or not. While all these "little bottlenecks" are not a problem and for most of them there is already a solution such as using the variable modification option for methionine oxidation or using two missed cleavage sites for trypsin, or submitting the MS/MS spectra as supplementary data for a manuscript submitted for publication, the largest problem for which there is still no solution is the natural and artificial modification of the peptides, which leads to two major problems: (1) a perfectly fine MS/MS has no peptide assigned because the peptide contains an artificial modification such as iodoacetamide modification of the N-terminal amino acids within peptides, leading to an unassigned MS/MS spectrum and an unidentifiable peptide and (2) artificial alteration of the peptides by experiment-induced peptide modification, which leads to errors in data quantitation. Again, in this case, the same example is given: when a peptide is alkylated at the non-cysteine amino acids, the amount of the precursor that corresponds to the unmodified peptide is sometimes very low. Such an example of artificial peptide modification was published by our lab with one example given in Fig. 29.2.

29.7 Bottleneck # 6: Instrumentation

MS instruments are very costly, delicate, and require a lot of maintenance and troubleshooting. This is one of the main reason MS is usually compared to NMR, which is thought to have the upper hand with regard to high-throughput, robustness, and the relative simplicity of the sample preparation steps. However, NMR cannot match the sensitivity and structural ability of MS [47, 48]. Additionally, one needs to possess a diverse range of instrument types in order to cover all possible applications, as one specific instrument cannot perform all kinds of experiments. Furthermore, there is still a need for mass spectrometers with better sensitivity, accuracy, and resolution. For example, dynamic ranges obtained by current MS instruments oscillates between 10^4 and 10^5, while the concentration range of proteins in human blood plasma is about 10^{12} [49, 50]. The aforementioned limitation is also related to another challenge in proteomics, which is the bias towards high-abundant and soluble proteins leading to the non-detection of disease-relevant low-abundant proteins [51]. This is also true for low-abundant PTMs or PPIs in complex samples. To partly remedy this situation, it is often recommended to operate one or more fractionation steps (usually chromatography) or enrichment of selected species (antibody-based, IMAC, chemical methods) prior to MS as a means of reducing sample complexity.

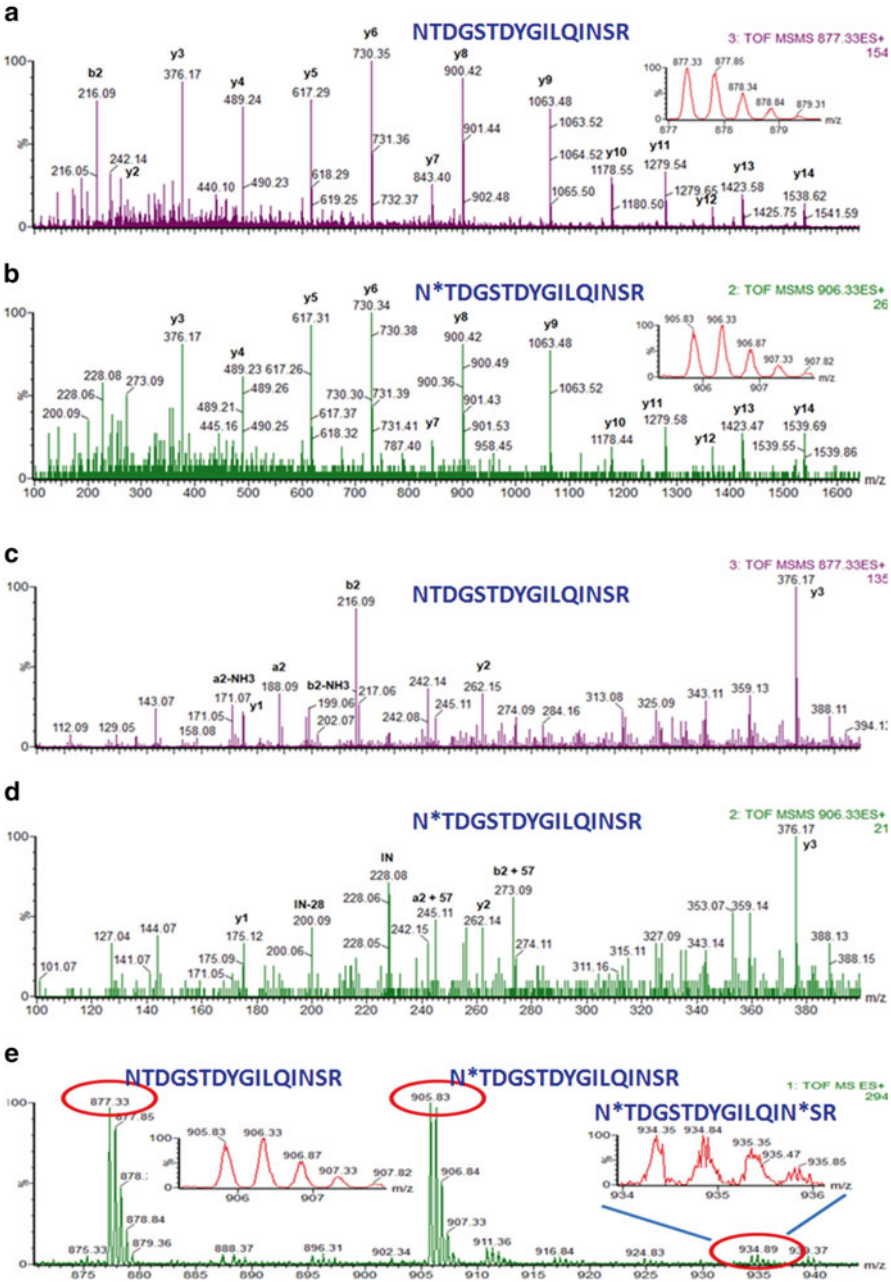


Fig. 29.2 NanoLC-MS/MS analysis of an SDS-PAGE gel band that was reduced by dithiothreitol (DTT), alkylated by iodoacetamide (IAA) and digested by trypsin. (**a**) MS/MS of the doubly-charged precursor ion with *m/z* of 877.33 produced a series of b and y product ions that led to the identification of the peptide with sequence NTDGSTDYGILQINSR that was part of the lysozyme protein. This peptide was identified during the Mascot database search. (**b**) MS/MS of the doubly-charged

29.8 Bottleneck # 7: Biological Significance

Perhaps one of the biggest bottlenecks in proteomics is the assignment of a biological role and of a biological, physiological, pathological, or clinical significance to a proteomic experiment. For example, when one performs a proteomics experiment, the outcome is a list of identified proteins or a list of ratios for a number of proteins. However, these proteins are not deeply investigated; having a list of proteins in a proteomic experiment is far from reflecting the physiological phenomenon that happens in an organelle, cell, tissue, organ, or organism. Information about the PTMs or interaction partners of these proteins is often not investigated. Instead, researchers focus on the investigation of several proteins (usually 3–5) by one (relative quantitation using the precursor ions or Western blotting) or more (Western blotting and immunofluorescence) methods, sometimes associated with functional experiments such as down-/up-regulation of these proteins, mutation of a protein, or transient/stable transfection of the DNA that encodes that protein. In addition, researchers rarely investigate more than 4–5 proteins in one proteomic experiment; in the proteomic world, these proteins are called the "validated proteins." While it is good to focus on a few proteins that are identified as crucial in a proteomic experiment, the other proteins are almost never investigated. The outcome of a proteomic dataset is published and almost never re-investigated for other proteins. Therefore, the biological significance in a dataset is explored only with regard to a few proteins; the rest of the information, even if it is deposited in a proteomic database (i.e., PRIDE), is lost or is almost never investigated. As such, a large amount of proteomic data is produced, but a very little amount is truly investigated.

Although not a bottleneck, another problem with assignment of a biological significance is the enzymatic activity of a protein and the position of a modification such as phosphorylation, acetylation, or methylation within the three-dimensional structure of that protein. One of these modifications, phosphorylation, is not only an important functional part of proteomics, but also an established proteomic subfield (phosphoproteomics). Phosphoproteomics is an important functional part of proteomics, and a method that quantitatively compares the phosphorylation of proteins from two different conditions, e.g., unstimulated and stimulated in vitro-grown cells. However, it is now clear that phosphorylation of a protein does not have

Fig. 29.2 (continued) precursor ion with *m*/*z* of 905.83 produced a series of b and y product ions that led to the identification of the mono-alkylated peptide with sequence N*TDGSTDYGILQINSR. This modification led to a 57 Da increase in the *m*/*z* of the precursor ion, compared with the precursor ion shown in (**a**). (**c**) The low mass range of the MS/MS spectrum shown in (**a**). (**d**) The low mass range of the MS/MS spectrum shown in (**b**). The *star* on the amino acids such as N*TDGSTDYGILQINSR indicates that the marked amino acid is alkylated. (**e**) MS spectrum showing the doubly-charged precursor ions with *m*/*z* of 877.33, *m*/*z* of 905.83, and *m*/*z* of 934.35 that corresponds to the unmodified peptide NTDGSTDYGILQINSR, mono-alkylated peptide N*TDGSTDYGILQINSR, and di-alkylated peptide N*TDGSTDYGILQIN*SR. Reproduced with permission from [52]

biological significance in itself, but rather the amino acid residues that are phosphorylated, the number of amino acids phosphorylated in one protein at a particular physiological condition, and the role in inserting a negative charge within a particular region(s) of the protein, as well as the biological effect that is observed or monitored. This information is almost never taken into account by any proteomic researcher. Yet, the role of protein phosphorylation and the number of phosphorylations in a protein, as well as the position of the phosphate group in a three dimensional structure of a protein seriously affects its enzymatic activity, as reflected in the biochemistry textbooks for glycogen synthase. Additional information associated with protein phosphorylation is also not considered. For example, protein truncation is almost never investigated in a regular proteomic experiment. Similarly, association of the protein phosphorylation with PPIs (particularly transient ones), is also not taken into account. As such, protein phosphorylation may also be considered a bottleneck in proteomics with regard to the interpretation of the biological significance of a phosphorylation of a specific protein, at a specific amino acid residue within a protein, at a particular time.

29.9 Conclusions

MS-based techniques have been employed for the comprehensive analysis of proteins. These methods have enabled multiple discoveries and significant advances in the biomedical field. However, it is only in combination with other omics disciplines (lipidomics, glycomics, transcriptomics, metabonomics) that proteomics will be able to provide complete answers to pressing biomedical questions of our time. We even envision that MS-based techniques could be expanded to the analysis of synthetic biomolecules, polymers, and other chemical entities. This has been mainly possible due to the recent development of better performing instrumentation. Nonetheless, there is still plenty of room for improvements in terms of instrumentation. An example is the fact that currently available instruments can only sequence precursor ions up to 5,000 m/z. Instruments that could go beyond this present limit are highly desirable for better identifications and sequence coverage. Further, there are not globally acceptable standards for the submission of proteomics data. For instance, the number of required technical and biological replicates varies from journal to journal. In some cases, two is enough and in other cases three is the minimum requested. One thing that has been shown in numerous studies is the much larger biological variability compared to the rather less significant technical variability. Another controversial point is the mention of false discovery rates (FDR) and where the threshold for this parameter should lie. Moreover, MS results should always be validated, preferentially with a different method such as immunoblotting or ELISA. However, MS methods like single/multiple reaction monitoring (SRM/MRM) are becoming more and more employed for validation purposes.

Though not mentioned as a bottleneck, the lack of adequate training for graduate students in this field is an important aspect hampering the large-scale use and

understanding of MS data analysis. Many students that work with mass spectrometry are good users, but do not understand in depth the complexities and challenges associated with mass spectrometry. There should be a transfer of knowledge from established MS experts towards interested graduate students.

Acknowledgements We would like to thank Ms. Laura Mulderig, Scott Nichols, and their colleagues (Waters Corporation) for their generous support in setting up the Proteomics Center at Clarkson University. CCD thanks Drs. Thomas A. Neubert (New York University), Belinda Willard (Cleveland Clinic), and Gregory Wolber & David Mclaughin (Eastman Kodak Company) for donation of a TofSpec2E MALDI-MS (each). C.C.D. thanks his advisors, Vlad Artenie, Wolfgang Haehnel, Paul M. Wassarman & Thomas A. Neubert for advice and support. This work was supported in part by the Keep a Breast Foundation (KEABF-375-35054), the Redcay Foundation (SUNY Plattsburgh), the Alexander von Humboldt Foundation, SciFund Challenge, the David A. Walsh fellowship, and by the U.S. Army research office (DURIP grant #W911NF-11-1-0304).

References

1. Darie CC, Shetty V, Spellman DS, Zhang G, Xu C, Cardasis HL, Blais S, Fenyo D, Neubert TA (2008) Blue native PAGE and mass spectrometry analysis of the ephrin stimulation-dependent protein-protein interactions in NG108-EphB2 cells. Applications of Mass Spectrometry in Life Safety, NATO Science for Peace and Security Series. Springer, Düsseldorf, Germany
2. Ngounou Wetie AG et al (2012) Investigation of stable and transient protein-protein interactions: past, present, and future. Proteomics 13(3–4):538–557
3. Darie CC (2013) Mass spectrometry and its application in life sciences. Aust J Chem 66:1–2
4. Ngounou Wetie AG et al (2013) Identification of post-translational modifications by mass spectrometry. Aust J Chem 66:734–748
5. Ngounou Wetie AG et al (2014) Protein-protein interactions: switch from classical methods to proteomics and bioinformatics-based approaches. Cell Mol Life Sci 71:205–228
6. Petrareanu C et al (2013) Comparative proteomics reveals novel components at the plasma membrane of differentiated HepaRG cells and different distribution in hepatocyte-and biliary-like cells. PLoS One 8(8):e71859
7. Sokolowska I et al (2012) Proteomic analysis of plasma membranes isolated from undifferentiated and differentiated HepaRG cells. Proteome Sci 10(1):47
8. Sokolowska I et al (2013) Applications of mass spectrometry in proteomics. Aust J Chem 66:721–733
9. James P (1997) Protein identification in the post-genome era: the rapid rise of proteomics. Q Rev Biophys 30(4):279–331
10. Aebersold R, Mann M (2003) Mass spectrometry-based proteomics. Nature 422(6928):198–207
11. Ngounou Wetie AG et al (2013) Automated mass spectrometry-based functional assay for the routine analysis of the secretome. J Lab Autom 18(1):19–29
12. Sokolowska I et al (2012) Disulfide proteomics for identification of extracellular or secreted proteins. Electrophoresis 33(16):2527–2536
13. Sokolowska I et al (2013) Mass spectrometry investigation of glycosylation on the NXS/T sites in recombinant glycoproteins. Biochim Biophys Acta 1834(8):1474–1483
14. Sokolowska I et al (2012) Automatic determination of disulfide bridges in proteins. J Lab Autom 17(6):408–416
15. Sokolowska I et al (2011) Mass spectrometry for proteomics-based investigation of oxidative stress and heat shock proteins. In: Andreescu S, Hepel M (eds) Oxidative stress: diagnostics, prevention, and therapy. American Chemical Society, Washington, DC

16. Woods AG et al (2012) Potential biomarkers in psychiatry: focus on the cholesterol system. J Cell Mol Med 16(6):1184–1195
17. Woods AG et al (2011) Blue native page and mass spectrometry as an approach for the investigation of stable and transient protein-protein interactions. In: Andreescu S, Hepel M (eds) Oxidative stress: diagnostics, prevention, and therapy. American Chemical Society, Washington, DC
18. Viswanathan S, Unlu M, Minden JS (2006) Two-dimensional difference gel electrophoresis. Nat Protoc 1(3):1351–1358
19. Gygi SP et al (1999) Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. Nat Biotechnol 17(10):994–999
20. Ong SE et al (2002) Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. Mol Cell Proteomics 1(5):376–386
21. Stemmann O et al (2001) Dual inhibition of sister chromatid separation at metaphase. Cell 107(6):715–726
22. Anderson L, Hunter CL (2006) Quantitative mass spectrometric multiple reaction monitoring assays for major plasma proteins. Mol Cell Proteomics 5(4):573–588
23. Liu H, Sadygov RG, Yates JR III (2004) A model for random sampling and estimation of relative protein abundance in shotgun proteomics. Anal Chem 76(14):4193–4201
24. Savitski MF, Savitski MM (2010) Unbiased detection of posttranslational modifications using mass spectrometry. Methods Mol Biol 673:203–210
25. Spiro RG (2002) Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds. Glycobiology 12(4):43R–56R
26. Marino K et al (2010) A systematic approach to protein glycosylation analysis: a path through the maze. Nat Chem Biol 6(10):713–723
27. Tarrant MK, Cole PA (2009) The chemical biology of protein phosphorylation. Annu Rev Biochem 78:797–825
28. Gorman JJ, Wallis TP, Pitt JJ (2002) Protein disulfide bond determination by mass spectrometry. Mass Spectrom Rev 21(3):183–216
29. McAuley A et al (2008) Contributions of a disulfide bond to the structure, stability, and dimerization of human IgG1 antibody CH3 domain. Protein Sci 17(1):95–106
30. Koh GC et al (2012) Analyzing protein-protein interaction networks. J Proteome Res 11(4):2014–2031
31. De Las Rivas J, Fontanillo C (2010) Protein-protein interactions essentials: key concepts to building and analyzing interactome networks. PLoS Comput Biol 6(6):e1000807
32. Krause F (2006) Detection and analysis of protein-protein interactions in organellar and prokaryotic proteomes by native gel electrophoresis: (membrane) protein complexes and supercomplexes. Electrophoresis 27(13):2759–2781
33. Schagger H, von Jagow G (1991) Blue native electrophoresis for isolation of membrane protein complexes in enzymatically active form. Anal Biochem 199(2):223–231
34. Anderson LE, McClure WO (1973) An improved scintillation cocktail of high-solubilizing power. Anal Biochem 51(1):173–179
35. O'Connell PB, Brady CJ (1976) Polyacrylamide gels with modified cross-linkages. Anal Biochem 76(l):63–73
36. Hansen JN (1976) Electrophoresis of ribonucleic acid on a polyacrylamide gel which contains disulfide cross-linkages. Anal Biochem 76(l):37–44
37. Bornemann S et al (2010) A novel polyacrylamide gel system for proteomic use offering controllable pore expansion by crosslinker cleavage. Electrophoresis 31(4):585–592
38. Wei J, Buriak JM, Siuzdak G (1999) Desorption-ionization mass spectrometry on porous silicon. Nature 399(6733):243–246
39. Yates JR, Ruse CI, Nakorchevsky A (2009) Proteomics by mass spectrometry: approaches, advances, and applications. Annu Rev Biomed Eng 11:49–79
40. Perkins DN et al (1999) Probability-based protein identification by searching sequence databases using mass spectrometry data. Electrophoresis 20(18):3551–3567

41. Eng JK, McCormack AL, Yates JR (1994) An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. J Am Soc Mass Spectrom 5(11):976–989
42. Tabb DL, McDonald WH, Yates JR III (2002) DTASelect and Contrast: tools for assembling and comparing protein identifications from shotgun proteomics. J Proteome Res 1(1):21–26
43. Craig R, Beavis RC (2003) A method for reducing the time required to match protein sequences with tandem mass spectra. Rapid Commun Mass Spectrom 17(20):2310–2316
44. Pedrioli PG et al (2004) A common open representation of mass spectrometry data and its application to proteomics research. Nat Biotechnol 22(11):1459–1466
45. Deutsch EW, Lam H, Aebersold R (2008) Data analysis and bioinformatics tools for tandem mass spectrometry in proteomics. Physiol Genomics 33(1):18–25
46. Deutsch E (2008) mzML: a single, unifying data format for mass spectrometer output. Proteomics 8(14):2776–2777
47. Lenz EM et al (2005) Metabonomics with ^{1}H NMR spectroscopy and liquid chromatography-mass spectrometry applied to the investigation of metabolic changes caused by gentamicin-induced nephrotoxicity in the rat. Biomarkers 10(2–3):173–187
48. Kussmann M, Rezzi S, Daniel H (2008) Profiling techniques in nutrition and health research. Curr Opin Biotechnol 19(2):83–99
49. Lescuyer P, Hochstrasser D, Rabilloud T (2007) How shall we use the proteomics toolbox for biomarker discovery? J Proteome Res 6(9):3371–3376
50. Surinova S et al (2011) On the development of plasma protein biomarkers. J Proteome Res 10(1):5–16
51. Drake RR et al (2011) Challenges to developing proteomic-based breast cancer diagnostics. OMICS 15(5):251–259
52. Woods AG, Sokolowska I, Darie CC (2012) Identification of consistent alkylation of cysteine-less peptides in a proteomics experiment. Biochem Biophys Res Commun 419(2):305–308

About the Editors

Dr. Alisa G. Woods received her B.S. in psychology from the State University of New York (SUNY) at Plattsburgh and her Ph.D. in biology from the University of California, Irvine, with Dr. Christine Gall. She continued with post-doctoral training at the University of Freiburg, Germany, with Dr. Michael Frotscher and completed a Master's degree in Mental Health Counseling at the University of Massachusetts, Boston. She is currently a Research Assistant Professor of Chemistry and Biomolecular Science at Clarkson University, where she works in the Laboratory for Biochemistry and Proteomics. She is also a counselor and neuropsychology researcher at the Neuropsychology Clinic at SUNY Plattsburgh.

Dr. Costel C. Darie started his biomedical career as a nurse at the Murgeni Hospital, Romania. He received his B.S. and M.S. in biochemistry at the Alexandru Ioan Cuza University from Iasi, Romania, and performed research under Dr. Vlad Artenie. He continued his Ph.D. studies with Dr. Wolfgang Haehnel at the Albert-Ludwigs Universitaet, Freiburg, Germany, after which he worked in the labs of Drs. Paul M. Wassarman and Micsunica Platica (Mount Sinai School of Medicine) and Thomas A. Neubert (New York University Medical Center). He is currently an Assistant Professor of Chemistry and Biomolecular Science at Clarkson University, where he is the lead of the Biochemistry and Proteomics Group.

A.G. Woods and C.C. Darie (eds.), *Advancements of Mass Spectrometry in Biomedical Research*, Advances in Experimental Medicine and Biology 806,
DOI 10.1007/978-3-319-06068-2, © Springer International Publishing Switzerland 2014

Index

A.G. Woods and C.C. Darie (eds.), *Advancements of Mass Spectrometry in Biomedical Research*, Advances in Experimental Medicine and Biology 806,
DOI 10.1007/978-3-319-06068-2,

© Springer International Publishing Switzerland 2014