Advances in Experimental Medicine and Biology 1140

Alisa G. Woods
Costel C. Darie *Editors*

Advancements of Mass Spectrometry in Biomedical Research

Second Edition

Springer

Advances in Experimental Medicine and Biology

Volume 1140

More information about this series at http://www.springer.com/series/5584

Alisa G. Woods • Costel C. Darie
Editors

Advancements of Mass Spectrometry in Biomedical Research

Second Edition

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)
Advances in Experimental Medicine and Biology
ISBN 978-3-030-15949-8 ISBN 978-3-030-15950-4 (eBook)
https://doi.org/10.1007/978-3-030-15950-4

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

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

Preface

The use of mass spectrometry (MS) is becoming increasingly important for biomedical research and for clinical applications. Detection of exogenous substances, such as toxins, can be performed by MS 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 MS 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. Aspects of molecular analysis are considered (i.e. whether to use label-based or label-free analysis, or 1D-PAGE versus 2D-PAGE analysis), and then options are explored on how MS can be applied to understanding and specific diseases and disorders.

Numerous MS-based methodologies are now available to researchers, and this text reviews many cutting-edge and relevant technologies. Dudley, Ju, and Neagu focus on the application of matrix-assisted laser desorption/ionization (MALDI) in biomedicine. Xu's, Wood's, Melman's, Darie's, and Luque-Garcia's teams explore methods for full sequence analysis in proteins, mutated proteins, conjugated/derivatized proteins, and membrane proteins, while Gethings and Connolly examine how peak capacity can be best maximized and Spellman's group is focused on biomarker discovery. Brown's lab describes lipidomics, Mihasan's team focuses on bacterial proteomics, Arcaro's team describes breast milk proteomics, and Kendrick's team focuses on two-dimensional electrophoresis coupled with proteomics. While Neubert's and Haley's groups explore the use of stable isotope labeling by amino acids in cell culture (SILAC) for protein quantification using MS, Mihasan et al. discuss the complementary use of computational structural biology in MS, and Dupree et al. focus on environmental proteomics and proteogenomics.

Different aspects of proteins and other molecules can be studied using MS. Petre explores protein structures and interactions, Drochioiu's team explores metal-peptides interactions, Furdui's team explores the use of MS in redox precision medicine, and Munteanu et al. explore protein complexes, while Zamfir has tackled the use of MS to understand gangliosides. Ngounou Wetie et al., Yakubu et al., and Aslebagh et al. examine the analysis of protein post-translational modifications and protein-protein interactions and Samways focuses on MS-based analysis of ion channel structure and function. Spellman describes the workflow for biomarker discovery in a pharmaceutical company, while Ferguson et al. focus on its use for understanding small molecules. Small molecules are also explored by members of Mesaros, Dudley, Darie, and Boolani teams.

Focusing on health promotion, Andrei and colleagues examine how MS can be used to analyze biomedically relevant stilbenes from wine. With regard to increased understanding of diseases and disorders and of the environment that we live in, chapters by Monien, Arcaro,

Channaveerappa, Boolani, Cristea, or Crimmins describe the use of MS for analysis in cancer adductomics, breast cancer, sleep apnea, mental fatigue, development, or environmental exposure. Topics covered include quantifying DNA adducts, analysis of breast milk, tissue proteomics in sleep apnea, salivary proteomics in mental fatigue subjects, developmental proteomics, or nontargeted screening of legacy contaminants in environmental monitoring programs.

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

In the realm of diagnostics and therapeutics, it is worth noting the contributions from authors that are part of biotechnology and pharmaceutical industry. Companies like Waters Corporation, Bruker Daltonics, or Mobilion are focusing on making better mass spectrometers with more applications; others like Merck & Co and Biogen are focused on basic research and drug development; and other companies like Kendrick Labs focus on custom proteomics analysis for protein identification, characterization, and quantification.

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. Since the editors' university is primarily an undergraduate institution and MS is usually not a focus on teaching undergraduate students, we decided to show some examples of the MS-based applications in protein and small molecule analysis designed for teaching undergraduate students (Jayathirtha et al.). We do envision that very soon in the near future MS will be a required course at the undergraduate level at almost every university that wants to be competitive.

We thus present to the reader a comprehensive text, examining the many uses of MS 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 MS 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, USA Alisa G. Woods
Potsdam, NY, USA Costel C. Darie

The original version of this book was revised. The correction is available at
https://doi.org/10.1007/978-3-030-15950-4_47

Contents

Contributors

Veronica Andrei International Centre of Biodynamics, Bucharest, Romania

Guillermo Aragoneses-Cazorla Department of Analytical Chemistry, Faculty of Chemical Sciences, Complutense University of Madrid, Madrid, Spain

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

Cornelia Babii BioActive Group, Department of Biology, Alexandru Ioan Cuza University of Iasi, Iaşi, Romania

Ian A. Blair Penn SRP Center and Center of Excellence in Environmental Toxicology Center, Department of Systems Pharmacology and Translational Therapeutics Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

Daniel F. Bogenhagen Stony Brook Cancer Center, Stony Brook University School of Medicine, Stony Brook, NY, USA

Department of Pharmacology, Stony Brook University School of Medicine, Stony Brook, NY, USA

Ali Boolani Applied Physiology and Psychology Lab, Department of Physical Therapy, Clarkson University, Potsdam, NY, USA

Heather L. Bowling Center for Neural Science, New York University, New York, NY, USA

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

Pete Bryngelson Analytical Development, Biogen, Cambridge, MA, USA

Devika Channaveerappa Biochemistry and Proteomics Group, Department of Chemistry and Biomolecular Science, Clarkson University, Potsdam, NY, USA

Xiaofei Chen Department of Internal Medicine, Section on Molecular Medicine, Wake Forest School of Medicine, Winston-Salem, NC, USA

Gabriela N. Chirițoiu Department of Molecular Cell Biology, Institute of Biochemistry, Bucharest, Romania

Joanne B. Connolly Waters Corporation, Stamford Avenue, Wilmslow, SK9 4AX, Cheshire, UK

Dana Copolovici Faculty of Food Engineering, Tourism and Environmental Protection, "Aurel Vlaicu" University of Arad, Arad, Romania

Bernard S. Crimmins Department of Civil and Environmental Engineering, Clarkson University, Potsdam, NY, USA

AEACS, LLC, New Kensington, PA, USA

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

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

Daniel DeBord MOBILion Systems Inc., Exton, PA, USA

Liana Dehelean University of Medicine and Pharmacy, "Victor Babes", Timisoara, Romania

Katrin Deinhardt Biological Sciences, Faculty of Environmental and Life Sciences, University of Southampton, Southampton, UK

Liulin Deng MOBILion Systems Inc., Exton, PA, USA

Gabi Drochioiu Faculty of Chemistry, Alexandru Ioan Cuza University of Iasi, Iaşi, Romania

Ed Dudley Swansea University Medical School, Swansea University, Swansea, UK

Emmalyn J. Dupree Biochemistry and Proteomics Group, Department of Chemistry and Biomolecular Science, Clarkson University, Potsdam, NY, USA

Hector Estevez Department of Analytical Chemistry, Faculty of Chemical Sciences, Complutense University of Madrid, Madrid, Spain

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

Adrian Fifere Advanced Research Centre for Bionanoconjugates and Biopolymers, "Petru Poni" Institute of Macromolecular Chemistry of Romanian Academy, Iasi, Romania

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

Cristina M. Furdui Department of Internal Medicine, Section on Molecular Medicine, Wake Forest School of Medicine, Winston-Salem, NC, USA

Wake Forest Baptist Comprehensive Cancer Center, Winston-Salem, NC, USA

Center for Redox Biology and Medicine, Wake Forest School of Medicine, Winston-Salem, NC, USA

Estefania Garcia-Calvo Department of Analytical Chemistry, Faculty of Chemical Sciences, Complutense University of Madrid, Madrid, Spain

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, Stamford Avenue, Wilmslow, SK9 4AX, Cheshire, UK

Aurora Goodwin Applied Physiology and Psychology Lab, Department of Physical Therapy, Clarkson University, Potsdam, NY, USA

Todd M. Greco Department of Molecular Biology, Lewis Thomas Laboratory, Princeton University, Princeton, NJ, USA

Stephanie Grobe Applied Physiology and Psychology Lab, Department of Physical Therapy, Clarkson University, Potsdam, NY, USA

John D. Haley Biological Mass Spectrometry Center, Stony Brook University School of Medicine, Stony Brook, NY, USA

Stony Brook Cancer Center, Stony Brook University School of Medicine, Stony Brook, NY, USA

Department of Pathology, Stony Brook University School of Medicine, Stony Brook, NY, USA

Ahmed M. Hamid MOBILion Systems Inc., Exton, PA, USA

Yutaka Hashimoto Department of Molecular Biology, Lewis Thomas Laboratory, Princeton University, Princeton, NJ, USA

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

Matt Hoelter Kendrick Laboratories, Inc., Madison, WI, USA

Thomas M. Holsen Department of Civil and Environmental Engineering, Clarkson University, Potsdam, NY, USA

Sorin Alexandru Ibanescu Advanced Research Centre for Bionanoconjugates and Biopolymers, "Petru Poni" Institute of Macromolecular Chemistry of Romanian Academy, Iasi, Romania

Laura Ion Laboratory of Biochemistry, Department of Chemistry, Al. I. Cuza University of Iasi, Iaşi, Romania

Faculty of Chemistry, Alexandru Ioan Cuza University of Iasi, Iaşi, Romania

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

Nathan S. Kline Institute for Psychiatric Research, New York University School of Medicine, Orangeburg, NY, USA

Madhuri Jayathirtha Biochemistry and Proteomics Group, Department of Chemistry and Biomolecular Science, Clarkson University, Potsdam, NY, USA

Jon Johansen Kendrick Laboratories, Inc., Madison, WI, USA

Yue Ju Bruker Daltonics Inc., San Jose, CA, USA

Monica Jureschi Faculty of Chemistry, Alexandru Ioan Cuza University of Iasi, Iaşi, Romania

Vasantha Krishna Kadambar Department of Chemistry and Biomolecular Science, Clarkson University, Potsdam, NY, USA

Nancy Kendrick Kendrick Laboratories, Inc., Madison, WI, USA

Jingyun Lee Department of Internal Medicine, Section on Molecular Medicine, Wake Forest School of Medicine, Winston-Salem, NC, USA

Wake Forest Baptist Comprehensive Cancer Center, Winston-Salem, NC, USA

Joseph A. Loo Department 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

Ana Lacramioara Lungoci Advanced Research Centre for Bionanoconjugates and Biopolymers, "Petru Poni" Institute of Macromolecular Chemistry of Romanian Academy, Iasi, Romania

Ancuta Veronica Lupaescu Faculty of Chemistry, Alexandru Ioan Cuza University of Iasi, Iaşi, Romania

Jose L. Luque-Garcia Department of Analytical Chemistry, Faculty of Chemical Sciences, Complutense University of Madrid, Madrid, Spain

Andres Machuca Department of Analytical Chemistry, Faculty of Chemical Sciences, Complutense University of Madrid, Madrid, Spain

Narcisa Laura Marangoci Advanced Research Centre for Bionanoconjugates and Biopolymers, "Petru Poni" Institute of Macromolecular Chemistry of Romanian Academy, Iasi, Romania

Laura Maxon MOBILion Systems Inc., Exton, PA, USA

Artem Melman Department of Chemistry and Biomolecular Science, Clarkson University, Potsdam, NY, USA

Clementina Mesaros Penn SRP Center and Center of Excellence in Environmental Toxicology Center, Department of Systems Pharmacology and Translational Therapeutics Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

Iuliana Mihai International Centre of Biodynamics, Bucharest, Romania

Marius Mihăşan Biochemistry & Proteomics Group, Structural Biology & Molecular Modeling Unit, Department of Chemistry & Biomolecular Science, Clarkson University, Potsdam, NY, USA

Department of Biology, Alexandru Ioan Cuza University of Iasi, Iaşi, Romania

Ronald A. Miller Impact Analytical, Midland, MI, USA

Bernhard H. Monien German Institute of Human Nutrition (DIfE) Potsdam Rehbrücke, Nuthetal, Germany

Department of Food Safety, Federal Institute for Risk Assessment, Unit 54, Berlin, Germany

Cristian V. A. Munteanu Department of Bioinformatics and Structural Biochemistry, Institute of Biochemistry, Bucharest, Romania

Florentina-Daniela Munteanu Faculty of Food Engineering, Tourism and Environmental Protection, "Aurel Vlaicu" University of Arad, Arad, Romania

James Murrough Mood and Anxiety Disorders Program, Mount Sinai School of Medicine, New York, NY, USA

Shruti Nayak Proteomics Laboratory, Alexandria Center for Life Science, NYU Langone, New York, NY, USA

Anca-Narcisa Neagu Laboratory of Animal Histology, Faculty of Biology, "Alexandru Ioan Cuza" University of Iasi, Iasi, Romania

Thomas A. Neubert Kimmel Center for Biology and Medicine at the Skirball Institute and Department of Cell Biology, 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

Edward Nieves Department of Biochemistry, Albert Einstein College of Medicine, Bronx, NY, USA

Department of Developmental and Molecular Biology, Albert Einstein College of Medicine, Bronx, NY, USA

Brian K. Panama Department of Physiology and Biophysics, Jacobs School of Medicine and Biomedical Sciences, SUNY Buffalo, Buffalo, NY, USA

Brian T. Pentecost Department of Veterinary and Animal Sciences, University of Massachusetts Amherst, Amherst, MA, USA

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

Center for Fundamental Research and Experimental Development in Translation Medicine—TRANSCEND, Regional Institute of Oncology, Iaşi, Romania

Andrei-Jose Petrescu Department of Bioinformatics and Structural Biochemistry, Institute of Biochemistry, Bucharest, Romania

Ștefana M. Petrescu Department of Molecular Cell Biology, Institute of Biochemistry, Bucharest, Romania

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

Mariana Pinteală Advanced Research Centre for Bionanoconjugates and Biopolymers, "Petru Poni" Institute of Macromolecular Chemistry of Romanian Academy, Iasi, Romania

Ginny Powers Kendrick Laboratories, Inc., Madison, WI, USA

Rafael C. Prados-Rosales Department of Analytical Chemistry, Faculty of Chemical Sciences, Complutense University of Madrid, Madrid, Spain

CIC bioGUNE, Derio, Bizkaia, Spain

Michelle R. Robinson Pharmacokinetics, Pharmacodynamics and Drug Metabolism, Merck Research Laboratories, West Point, PA, USA

Alexandru Rotaru Advanced Research Centre for Bionanoconjugates and Biopolymers, "Petru Poni" Institute of Macromolecular Chemistry of Romanian Academy, Iasi, Romania

Urmi Roy Biochemistry and Proteomics Group, Structural Biology and Molecular Modeling Unit, 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

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

Emily R. Sekera Department of Chemistry, Natural Sciences Complex, University at Buffalo, Buffalo, NY, USA

Mihaela Silion Advanced Research Centre for Bionanoconjugates and Biopolymers, "Petru Poni" Institute of Macromolecular Chemistry of Romanian Academy, Iasi, Romania

Kathryn Sinclair Swansea University Medical School, Swansea University, Swansea, Wales, UK

Justin Snider Biological Mass Spectrometry Center, Stony Brook University School of Medicine, Stony Brook, NY, USA

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

Zoran Sosic Analytical Development, Biogen, Cambridge, MA, USA

Daniel S. Spellman Pharmacokinetics, Pharmacodynamics and Drug Metabolism, Merck Research Laboratories, West Point, PA, USA

Sydney Stradtman Biochemistry and Proteomics Group, Department of Chemistry and Biomolecular Science, Clarkson University, Potsdam, NY, USA

Allen W. Tsang Department of Internal Medicine, Section on Molecular Medicine, Wake Forest School of Medicine, Winston-Salem, NC, USA

Wake Forest Baptist Comprehensive Cancer Center, Winston-Salem, NC, USA

Center for Redox Biology and Medicine, Wake Forest School of Medicine, Winston-Salem, NC, USA

Alina Vasilescu International Centre of Biodynamics, Bucharest, Romania

Daifeng Wang Department of Biomedical Informatics, Stony Brook University School of Medicine, Stony Brook, NY, USA

Stony Brook Cancer Center, Stony Brook University School of Medicine, Stony Brook, NY, USA

Siyu Wang Penn SRP Center and Center of Excellence in Environmental Toxicology Center, Department of Systems Pharmacology and Translational Therapeutics Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

Yan Wang Analytical Development, Biogen, Cambridge, MA, USA

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

Louis M. Weiss Department of Pathology, Albert Einstein College of Medicine, Bronx, NY, USA

Department of Medicine, Division of Infectious Diseases, Albert Einstein College of Medicine, Bronx, NY, USA

Danielle Whitham Biochemistry and Proteomics Group, Department of Chemistry and Biomolecular Science, Clarkson University, Potsdam, NY, USA

Tyler Wilkinson Biochemistry and Proteomics Group, Department of Chemistry and Biomolecular Science, Clarkson University, Potsdam, NY, USA

Troy D. Wood Department of Chemistry, Natural Sciences Complex, University at Buffalo, Buffalo, NY, USA

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

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

MOBILion Systems Inc., Exton, PA, USA

Hanzhi Wu Department of Internal Medicine, Section on Molecular Medicine, Wake Forest School of Medicine, Winston-Salem, NC, USA

Chong-Feng Xu Analytical Development, Biogen, Cambridge, MA, USA

Rama R. Yakubu Department of Pathology, Albert Einstein College of Medicine, Bronx, NY, USA

Alina D. Zamfir National Institute for Research and Development in Electrochemistry and Condensed Matter, Timisoara, Romania

"Aurel Vlaicu" University of Arad, Arad, Romania

Fereshteh Zandkarimi Department of Biological Sciences, Quantitative Proteomics and Metabolomics Center, Columbia University, New York, NY, USA

Li Zang Analytical Development, Biogen, Cambridge, MA, USA

Guoan Zhang Proteomics and Metabolomics Core Facility, Weill Cornell Medicine, New York, NY, USA

Department of Biochemistry and Pharmacology, New York University School of Medicine, New York, NY, USA

Radu Zonda Advanced Research Centre for Bionanoconjugates and Biopolymers, "Petru Poni" Institute of Macromolecular Chemistry of Romanian Academy, Iasi, Romania

About the Editors

Costel C. Darie, PhD started his biomedical career as a nurse at the Murgeni Hospital, Romania. He received his BS and MS in biochemistry at the Al. I. 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 currently is an Associate Professor of Chemistry and Biomolecular Science at Clarkson University, where he is the Lead of the Biochemistry and Proteomics Group.

Alisa G. Woods, PhD, MS received her BS in psychology from the State University of New York (SUNY) at Plattsburgh, her PhD in biology from the University of California Irvine with Dr. Christine Gall, postdoctoral 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 Associate 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.

Mass Spectrometry for Proteomics-Based Investigation

1

Alisa G. Woods, Izabela Sokolowska, Armand G. Ngounou Wetie, Devika Channaveerappa, Emmalyn J. Dupree, Madhuri Jayathirtha, Roshanak Aslebagh, Kelly L. Wormwood, and Costel C. Darie

Abstract

Within the past years, we have witnessed a great improvement is 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 the proteins, but also to identify the protein's post-translational modifications (PTMs), protein isoforms, protein truncation, protein-protein interactions (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 the 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.

Keywords

Mass spectrometry · Proteomics · MALDI-MS · LC-MS/MS

Abbreviations

BN-PAGE	Blue native PAGE
CI	Chemical ionization
CID	Collision-induced dissociation
CN-PAGE	Colorless native PAGE
DIGE	Differential gel electrophoresis
EI	Electric ionization
ESI	Electrospray ionization
ESI-MS	Electrospray ionization mass spectrometry
FAB	Fast atom bombardment
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
PD	Plasma desorption
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

A. G. Woods · I. Sokolowska · A. G. Ngounou Wetie
D. Channaveerappa · E. J. Dupree · M. Jayathirtha · R. Aslebagh
K. L. Wormwood · C. C. Darie (✉)
Biochemistry & Proteomics Group, Department of Chemistry & Biomolecular Science, Clarkson University, Potsdam, NY, USA
e-mail: cdarie@clarkson.edu

A. G. Woods, C. C. Darie (eds.), *Advancements of Mass Spectrometry in Biomedical Research*, Advances in Experimental Medicine and Biology 1140, https://doi.org/10.1007/978-3-030-15950-4_1

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 sub-fields, such as proteomics, lipidomics, glycomics, and metabolomics. The focus of proteomics is to analyze proteins and protein derivatives (such as glycoproteins), peptides, post-translational modifications within proteins (PTMs), 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 SDS-PAGE, and a two dimensional separation can be performed by two-dimensional electrophoresis or by affinity purification followed by 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 post-translational modifications and interaction partners of some proteins (protein-protein interactions) [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), 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.

Proteomic analysis also can 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

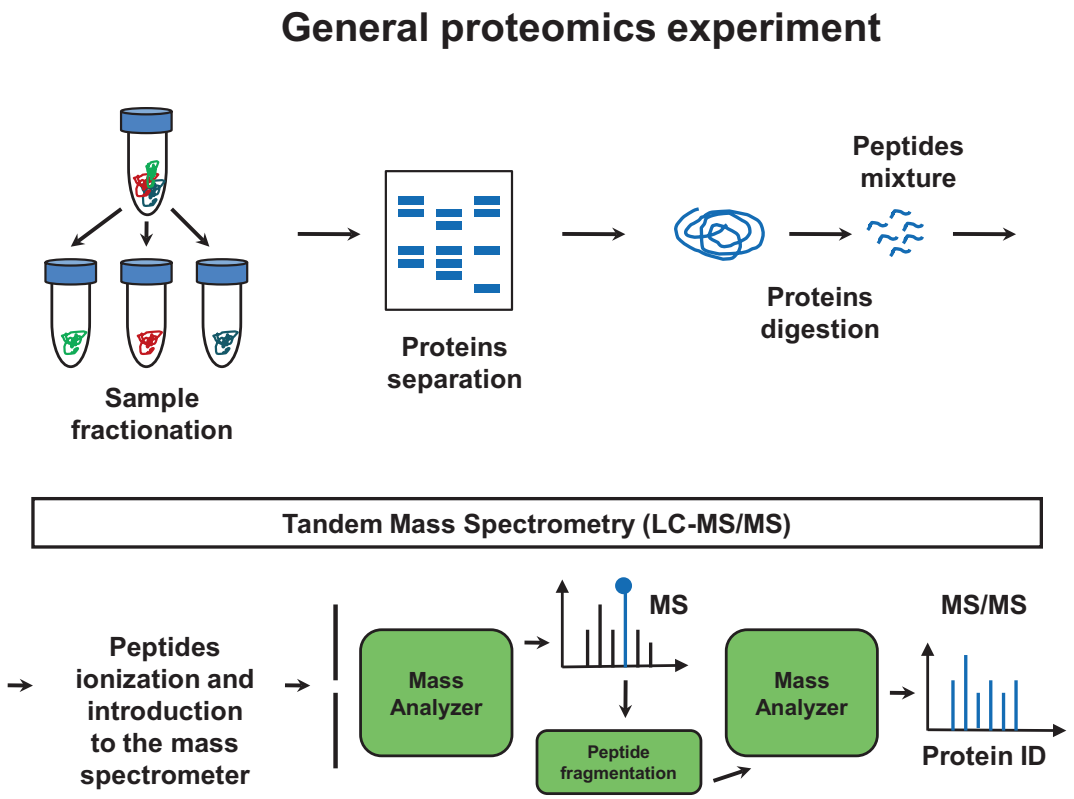

Fig. 1.1 General proteomic experiment workflow schematic. Reprinted and adapted with permission from the *Australian Journal of Chemistry CSIRO Publishing* http://www.publish.csiro.au/?paper=CH13137 [15]

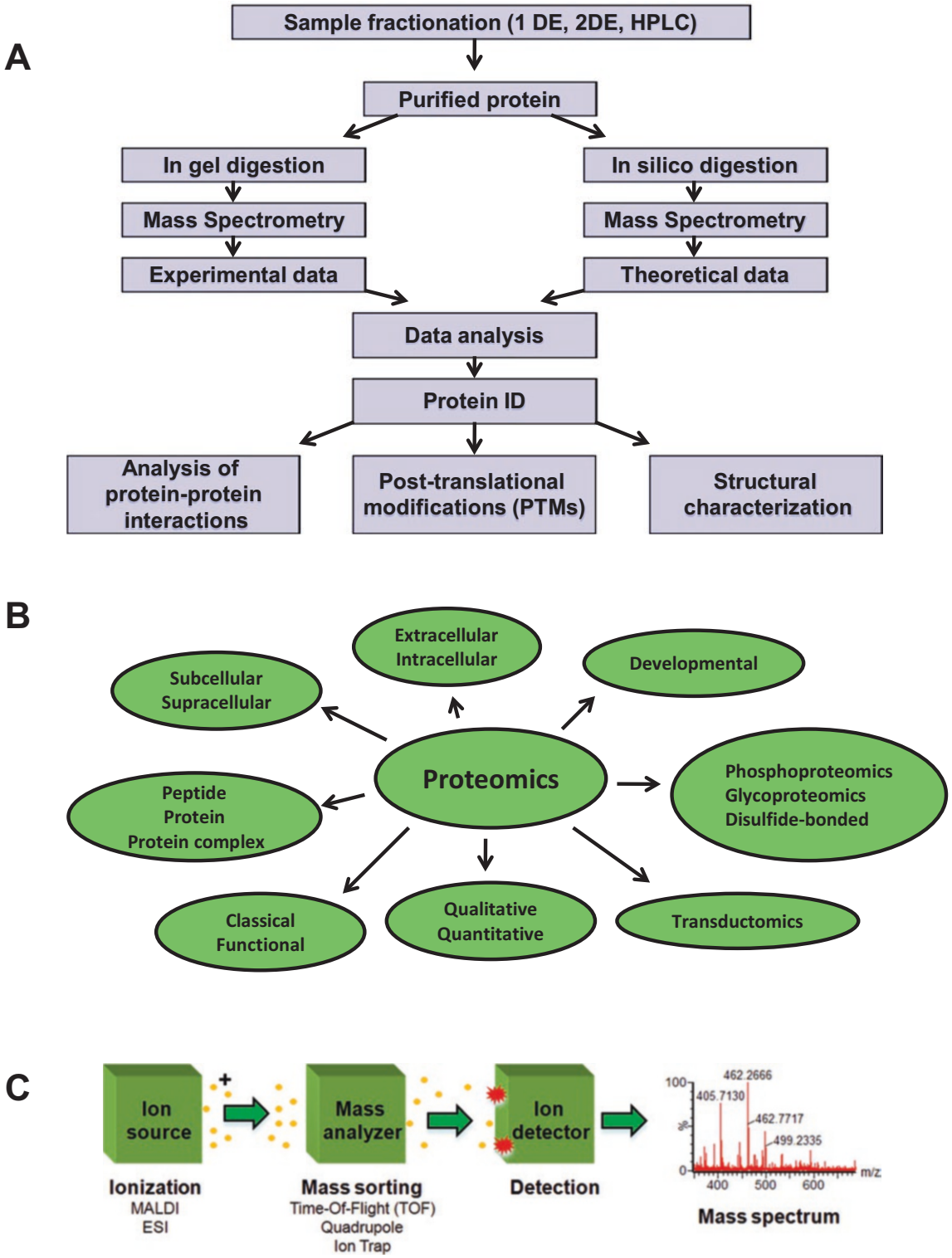

Fig. 1.2 General proteomics experiment. (**a**) Proteomics experiment workflow schematic. (**b**) Proteomics and applications schematic. (**c**) Mass spectrometer schematic. Reprinted and adapted with permission from the *Oxidative Stress: Diagnostics, Prevention, and Therapy*, S. Andreescu and M. Hepel,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

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-MS and 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 sodium dodecyl-sulfate-polyacrylamide gel electrophoresis (SDS-PAGE), reduced, 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 non-reducing 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 proteins and peptides (2–20 kDa), where SDS-PAGE has poor or 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] which are 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], is 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 (Fig. 1.2c). Initially, the sample is ionized and the ions produced by MALDI or ESI ionization 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.

1.3.1 Ionization Sources

Ionization of peptides is dependent on the electrical potential at the ion source and on the pH at which it is 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 1000 Da. EI provides good structural information derived from fragmentation. However, molecular mass determination is rather poor (poor signal or 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 1000 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 1500 Da [54].

1.3.2 Mass Analyzers

There are three main types of mass analyzers used for proteomics experiments: trapping type instruments (quadrupole ion trap—QIT, linear ion trap—IT, 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 an electrostatic gate pulses to inject ions into the ion trap. The ion-trap-based analyzers are 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 a 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 and 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 toward one rod. These instruments provide good reproducibility and low cost, but their resolution and accuracy is 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 ions with different masses to travel from the ion source to detector.

Mass spectrometers can have stand-alone analyzers or have a 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, and QQ-LIT; these instruments are also called hybrid mass spectrometers, and are highly sensitive and also have a high resolution [1, 57–59].

1.3.3 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, or 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 the peptides, which then become ionized, fly through the time of flight (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 × 1]/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 searches of the MALDI-MS spectra usually identifies that single protein or those proteins through a process named peptide mass fingerprinting (Fig. 1.3c).

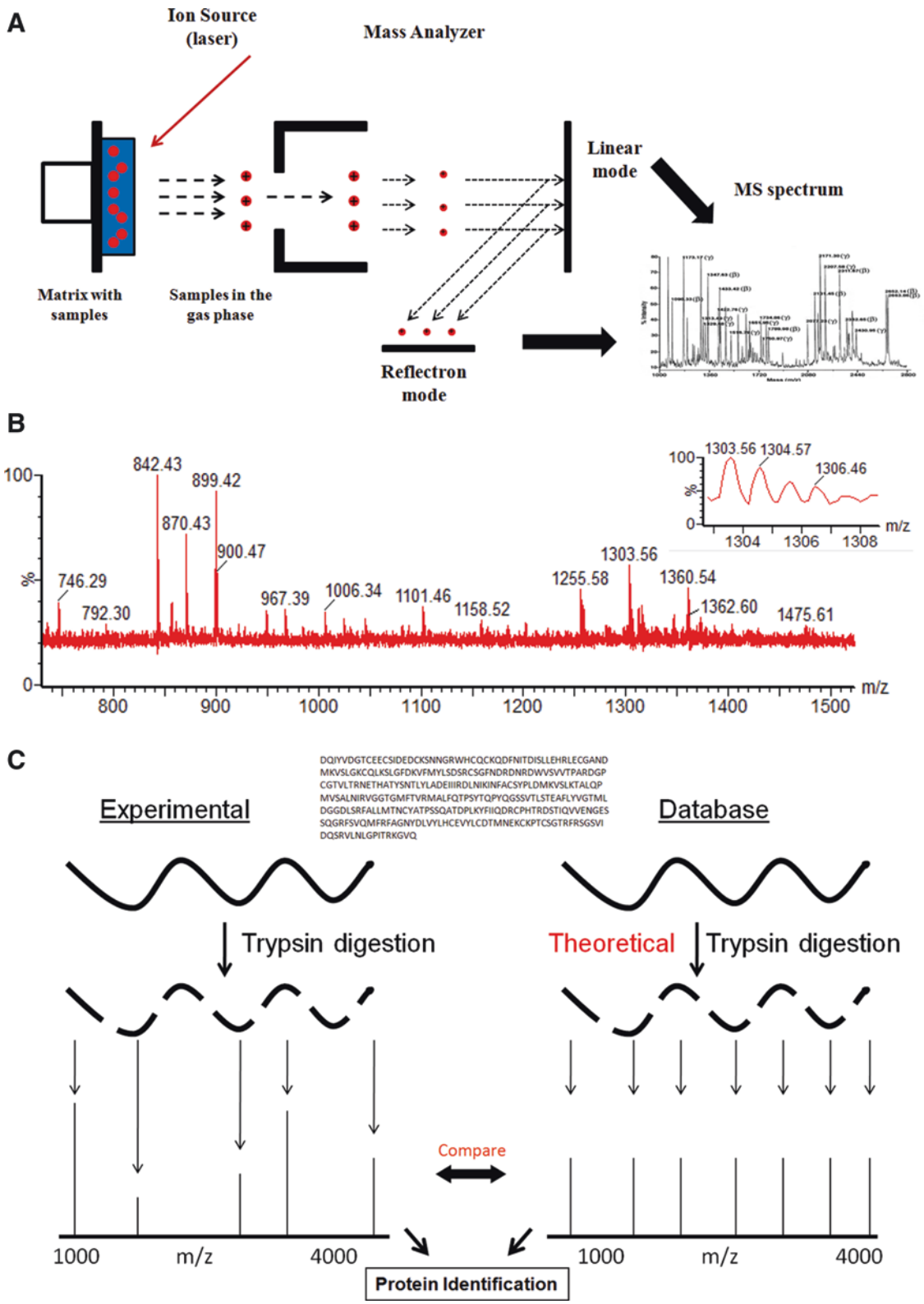

Fig. 1.3 MALDI-TOF MS. (**a**) MALDI TOF mass spectrometer principle. An ion source, a mass analyzer and detector are present on the instrument. At the detector the mass spectrum is detected/recorded. The mass analyzer is a TOF and can be used in linear mode or reflective mode. (**b**) A MALDI MS spectrum 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 mixture of peptides is analyzed by MALDI-MS. Peaks recorded in a MS spectrum are then converted to a peak list file, used for database search. An *in silico* experiment is similarly performed and the difference between the theoretical and experimental outcome during database search generally identifies one or more proteins through PMF. 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]

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 multiply 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 a MS spectrum in a process called direct infusion (ESI-MS mode). For example, if one has ten peptides in an Eppendorf tube, one can identify all ten 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 ten 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 (2500–3000 Da) than the regular peptides analyzed by MS (800–2500 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 not only be used for peptides, but also for investigation of proteins. This information is particularly useful for determining the molecular mass of those proteins, their potential post-translational modifications, and their conformation. In addition, the high molecular mass proteins can also be analyzed by ESI-MS in either positive mode (protonated) or negative mode (deprotonated), thus providing distinct, yet complementary information regarding the distribution of charges on the surface of the protein investigated. Examples of MS spectra for 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 peptides 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, since fragmentation of the ions that correspond to peptides happens manually, one peptide at a time. For example, if one has four 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 a 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 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 peptides 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 peptides mixtures,

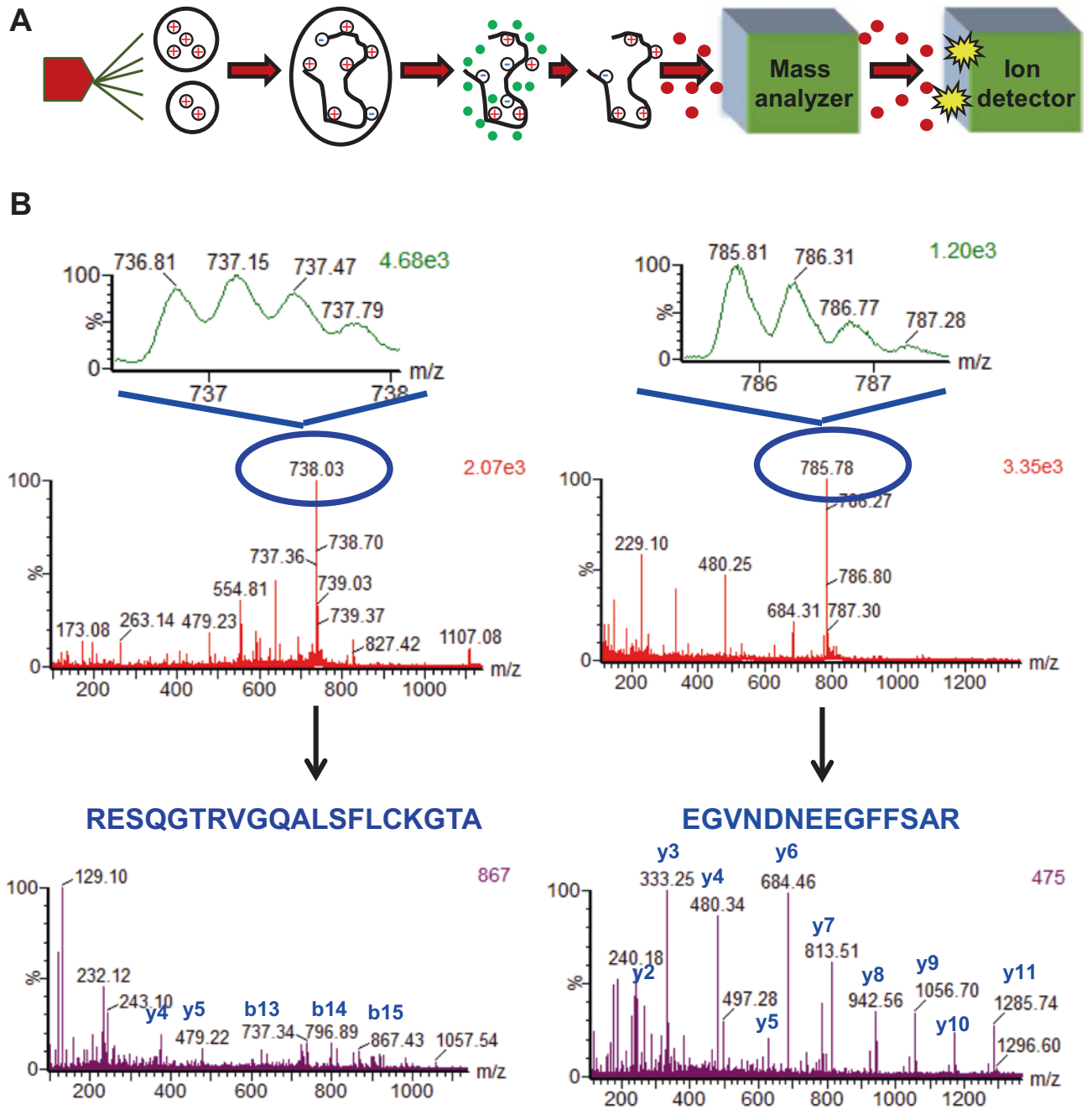

Fig. 1.4 ESI-MS of peptides. (**a**) An ESI-MS mass spectrometer. The ESI-MS has an ion source, in which the ions are ionized, a mass analyzer that ions travel through as well as an ion detector, which records the mass spectrum. In ESI-MS, the sample is liquid, under high temperature and high electric 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]

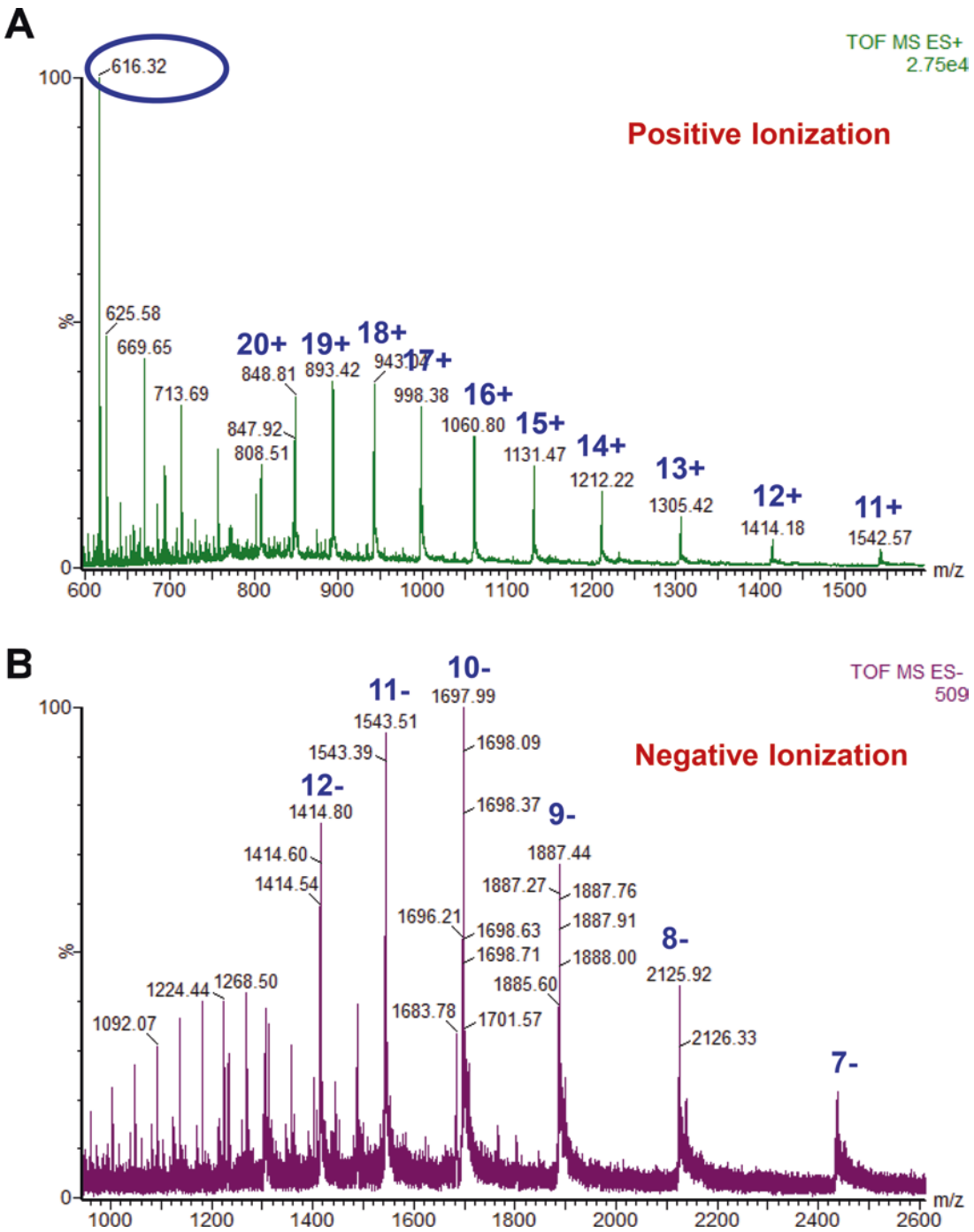

Fig. 1.5 ESI-MS proteins: ESI-MS spectra of intact 17 kDa protein, myoglobin, analyzed under acidic conditions (pH ~2). (**a**) MS spectrum in positive ionization; (**b**) MS spectrum analyzed in negative ionization. The positive (**a**) and negative (−) charges are indicated. The peak with m/z of 616.32 (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]

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 shortened 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 have 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: 1 s MS survey, 1 s MS/MS (Peak 1), 1 s MS/MS (Peak 2) and then again 1 s MS survey (Fig. 1.6a).

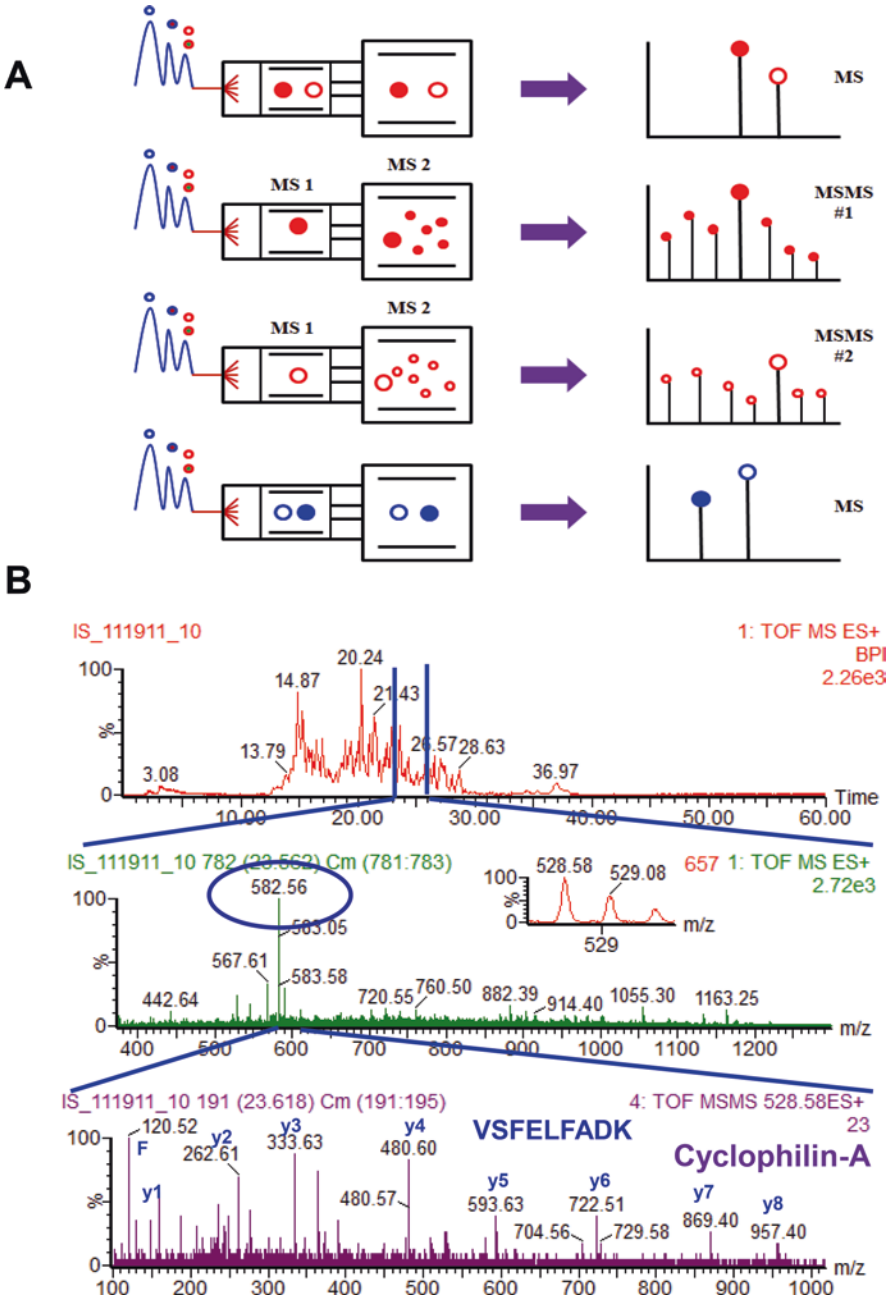

Fig. 1.6 LC-MS/MS experiment. (**a**) In each LC-MS/MS experiment, with elution of peptides from the HPLC gradually, the mass spectrometer analyzes corresponding ions via MS survey (recorded in an MS spectrum). Ions with highest intensity (typically one to eight ions; two ions in this example) are selected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

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 ~15 × 120 = 1800 proteins. However, 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 a MS run can identify a maximum of 15 × 90 = 1350 proteins. If the length of a 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 (~1350 × 6 = 8100 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. 8100 × 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 not the case. For example, the type and length of the gradient in HPLC (for example, sharp or shallow) does play an important role in peptide fractionation. An optimized versus an unoptimized 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. In practice, a good LC-MS/MS run usually leads to identification of about 500–1000 proteins. An example of a TIC, MS and MS/MS is shown in Fig. 1.6b.

1.7 Data Analysis

The raw data collected by mass spectrometers 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 for protein identification. 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 differential gel electrophoresis (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], and 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 are the first steps 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 intact proteins are investigated and a bottom-up approach when proteins are digested and the peptides mixture is analyzed (Fig. 1.7).

A top down approach allows for the identification of protein isoforms or any potential post-translational modifications within proteins [65]. In the 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].

Characterization of proteins is not easy, but it becomes even more complicated due to the extensive 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 on the proteins 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]. Characterization of glycoproteins is difficult because the glycosylation is not uniform and usually multiple glycoforms are simultaneously produced by the