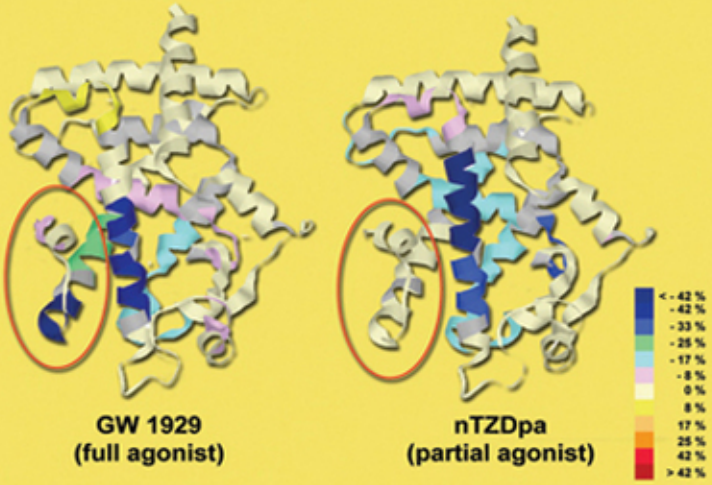
Protein and Peptide Mass Spectrometry in Drug Discovery



Edited by
Michael L. Gross
Guodong Chen
Birendra N. Pramanik

WILEY

PROTEIN AND PEPTIDE MASS SPECTROMETRY IN DRUG DISCOVERY

PROTEIN AND PEPTIDE MASS SPECTROMETRY IN DRUG DISCOVERY

Edited By

Michael L. Gross
Washington University

Guodong Chen
Bristol-Myers Squibb

Birendra N. Pramanik
Merck Research Laboratories

A JOHN WILEY & SONS, INC., PUBLICATION

Published by John Wiley & Sons, Inc., Hoboken, New Jersey
Published simultaneously in Canada

Library of Congress Cataloging-in-Publication Data

Protein and peptide mass spectrometry in drug discovery / edited By Michael L. Gross, Guodong Chen, Birendra N. Pramanik.
p. ; cm.
Includes bibliographical references.
ISBN 978-0-470-25817-0 (cloth)
1. Drug development. 2. Peptides--Spectra. 3. Proteins--Spectra. I. Gross, Michael L. II. Chen, Guodong, 1966- III. Pramanik, Birendra N., 1944-
[DNLM: 1. Drug Discovery. 2. Mass Spectrometry. 3. Peptides--analysis. 4. Proteins--analysis. QV 744]
RM301.25.P757 2011
615'.19--dc23

2011015204

Printed in the United States of America

oBook ISBN: 9781118116555
ePDF ISBN: 9781118116531
ePub ISBN: 9781118116548
eMobi ISBN: 9781118116524

10 9 8 7 6 5 4 3 2 1

CONTENTS

PREFACE

For over a decade mass spectrometry (MS) has been one of the most highly utilized analytical technique for analysis of proteins and peptides. This is largely due to continuous refinement of ionization methods, including electrospray ionization (ESI) and matrix-assisted laser desorption ionization (MALDI), the improvement of MS instrumentation, and the growth in the data processing. Various niche applications in neuroproteomics and antigenic peptides could have important implications in drug discovery, and these developments are described in two chapters. Furthermore there has been considerable research activity focused on the development of new methodologies for the analysis of proteins and peptides; these methods have exploited ongoing instrumentation improvements and include both bottom-up and top-down protein sequencing. New approaches include those in imaging, ion mobility, and the use of microwave radiation to speed proteolysis, and these new ideas are covered in three chapters in this volume. Accompanying the analytical developments are new techniques for the determinations of protein structure, of their interactions with peptides, proteins, and ligands including drugs, and their folding and unfolding. These techniques are described in detail in this volume.

One of the important and immediate applications in protein and peptide MS is for pharmaceutical analysis throughout each stage of drug development process, ranging from drug discovery to manufacturing. MS-based technologies play critical roles in providing qualitative and quantitative information to characterization of target proteins and protein products for therapeutic use, as illustrated in this volume.

We are delighted to bring together the work of contributors from academe and industry in highlighting current analytical approaches, industry practices, and modern strategies for the characterization of proteins and peptides in drug discovery. Our goal is to present a compilation of the latest methodologies and applications to practitioners of protein and peptide MS with the focus on drug discovery efforts.

We would like to acknowledge the special efforts and patience of all the authors, who have made significant contributions to this book.

Michael L. Gross
Guodong Chen
Birendra N. Pramanik

CONTRIBUTORS

Christine L. Andrews, Merck Research Laboratories, Cambridge, MA

Catherine A. Srebalus Barnes, Eli Lilly and Company, Indianapolis, IN

Fotini N. Bazoti, The Goulandris Natural History Museum, Kifissia, Greece

Adriana Bora, University of Illinois, Urbana, IL

Ajay K. Bose (deceased), Stevens Institute of Technology, Hoboken, NJ

Ron Bose, Washington University, St. Louis, MO

Michael Boyne, Washington University, St. Louis, MO

Richard M. Caprioli, Vanderbilt University, Nashville, TN

Guodong Chen, Bristol-Myers Squibb, Princeton, NJ

Hao Chen, Ohio University, Athens, OH

Ismael Cotte-Rodriguez, Procter & Gamble, Loveland, OH

Peter O'Connor, University of Warwick, Coventry, UK

Surendra Dasari, Vanderbilt University Medical Center, Nashville, TN

Jeffrey R. Enders, Vanderbilt University, Nashville, TN

Jacob Galan, Purdue University, West Lafayette, IN

Michael L. Gross, Washington University, St. Louis, MO

Anton Iliuk, Purdue University, West Lafayette, IN

Sandra A. Kerfoot, Seattle Childrens Research Institute, Seattle, WA

Michal Kliman, Vanderbilt University, Nashville, TN

Amareth Lim, Eli Lilly and Company, Indianapolis, IN

Cheng Lin, Boston University, Boston, MA

John A. McLean, Vanderbilt University, Nashville, TN

Zhixin Miao, Ohio University, Athens, OH

Urooj A. Mirza, Merck Research Laboratories, Kenilworth, NJ

Elliott Nickbarg, Merck Research Laboratories, Cambridge, MA

Birendra N. Pramanik, Merck Research Laboratories, Kenilworth, NJ

Michelle L. Reyzer, Vanderbilt University, Nashville, TN

Henry Rohrs, Washington University, St. Louis, MO

Patrick J. Rudewicz, Elan Pharmaceuticals, South San Francisco, CA

Sevugarajan Sundarapandian, Vanderbilt University, Nashville, TN

Jonathan V. Sweedler, University of Illinois, Urbana, IL

David L. Tabb, Vanderbilt University Medical Center, Nashville, TN

W. Andy Tao, Purdue University, West Lafayette, IN

Anthony Tsarbopoulos, University of Patras, Greece

Fang Xie, Pacific Northwest National Laboratory, Richland, WA

Xianshu Yang, Merck Research Laboratories, Cambridge, MA

Yun Zhang, Ohio University, Athens, OH

Michael R. Ziebell, Merck Research Laboratories, Cambridge, MA

PART I

METHODOLOGY

CHAPTER 1

Ionization Methods in Protein Mass Spectrometry

ISMAEL COTTE-RODRIGUEZ, YUN ZHANG, ZHIXIN MIAO, and HAO CHEN

Mass spectrometry (MS) has become one of the most powerful and popular modern physical-chemical methods to study the complexities of elemental and molecular processes in nature. The advent of new methods of ion generation, novel mass analyzers, and new tools for data processing has made it possible to analyze almost all chemical entities by MS, ranging from small organic compounds, large biological molecules, to whole living cells/tissues. As proteins fulfill a plethora of biochemical functions within every living organism, equally spectacular efforts and advances have been seen for protein ionization methods. In particular, the invention of matrix-assisted laser desorption ionization (MALDI) [1] and electrospray ionization (ESI) technologies [2,3] allow one to measure protein molecular weights, to determine sequences, and to probe conformations and post-translational modifications of proteins. In addition the mass range of species amenable for MS analysis has been increased immensely, enabling the transfer into the gas phase of ionized noncovalent species with masses well over one million (e.g., a 100 MDa single DNA ion [4]). These advances move MS into the range of intact protein oligomers and functional machineries.

This chapter is an introduction to various ionization methods for proteins. As this is a broad topic with an immense literature coverage including many excellent books [5,6] and reviews [7–17], we will emphasize some types of spray or laser-based protein ionization techniques, including atmospheric pressure MALDI, surface-enhanced laser desorption/ionization (SELDI), nanostructure-initiator MS (NIMS), sonic spray ionization (SSI), electrosonic spray ionization (ESSI), desorption electrospray ionization (DESI), fused-droplet electrospray ionization (FD-ESI), electrospray-assisted laser desorption ionization (ELDI), and matrix-assisted laser desorption electrospray ionization (MALDESI). We begin with the introduction of some historic facts for the

Protein and Peptide Mass Spectrometry in Drug Discovery, Edited by Michael L. Gross, Guodong Chen, and Birendra N. Pramanik.

development of protein ionization methods, followed with the description of each method including the ionization principles, strengths, and analytical applications.

1.1 HISTORY OF THE DEVELOPMENT OF PROTEIN MASS SPECTROMETRY

MS originates from ninteen-century physics. The first known mass spectrometer was built by J. J. Thomson in the early 1900s to study and measure the mass (m)-to-charge (z) (m/z) values of the "corpusules" that make up "positive rays" [18], a type of radiation initially observed by German physicist Eugen Goldstein. Following the seminal work of Thomson, MS underwent countless improvements in instrumentation, ionization methods, and applications. The classical ionization method, electron ionization (EI), was devised by Dempster and improved later by Bleakney [19] and Nier [20], and became a widely used standard for ionization of volatile organic compounds. This ionization technique requires extensive derivatization and evaporation of a nonvolatile analyte into the ion source, and it involves numerous fragmentation and rearrangement reactions.

Applications of MS to peptides (derivatized via acylation) begun in the late 1950s by Biemann [21] and McLafferty [22] The first methods that allowed analysis of nonderivatized peptides were field desorption (FD) and chemical ionization (CI) developed in the 1960s [23,24]. Ionization by CI is achieved by interaction of its volatile molecules with reagent ions. CI allows ionization without significant degree of ion fragmentation but still requires gas-phase samples. Field desorption was reported by Beckey in 1969 [25], in which electron tunneling triggered by a very high electric field results in ionization of gaseous analyte molecules.

It was plasma desorption (PD) [26] and fast atom bombardment (FAB) [27] that opened the way to protein analysis. PD ionization, invented by R. D. Macfarlane in 1976 [28], a breakthrough in the analysis of solid samples, involves ionization of materials in the solid state by bombardment with ions or neutral atoms formed as a result of the nuclear fission of the Californium isotope ^{252}Cf. In 1982 Sundqvist and coworkers obtained the first spectrum of a protein, insulin (Figure 1.1), using bombardment with a beam of 90 MeV $^{127}I^{20+}$ ions from a tandem accelerator [26]. Later, FAB involving focusing the sample in liquid matrix with a beam of neutral atoms or molecules, was implemented for the ionization of proteins up to 24 kDa [29]. In 1983 Blakely and Vestal [30] introduced thermospray ionization (TSI) to produce ions from an aqueous solution sprayed directly into a mass spectrometer. Thermospray is a form of atmospheric pressure ionization in MS, transferring ions from the liquid phase to the gas phase for analysis. It was particularly useful in coupling liquid chromatography with mass spectrometry [31].

The breakthrough for large molecule laser desorption ionization came in 1987 when Tanaka combined 30-nm cobalt particles in glycerol with a 337-nm nitrogen laser for ionization and showed that singly charged protein molecular ions up to about 35 kDa can be introduced to a mass spectrometer [32]. During that time, MALDI [15,33], first reported in 1985 by Hillenkamp, Karas, and their colleagues,

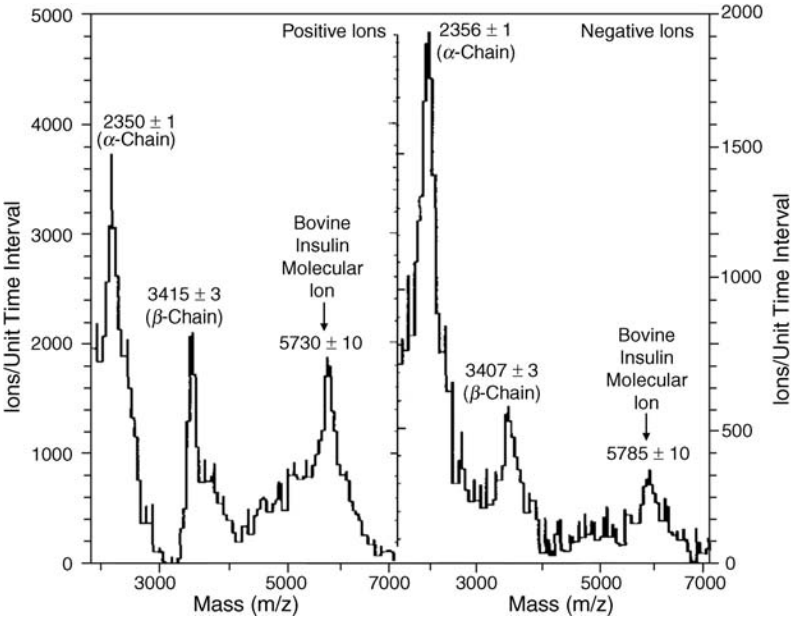


FIGURE 1.1 ^{127}I-PDMS spectra of bovine insulin recorded over a 1.5-h period with a 90-MeV ^{127}I (+20) beam current of 2000 s^{-1}. From [26]. Copyright permission was obtained from ACS.

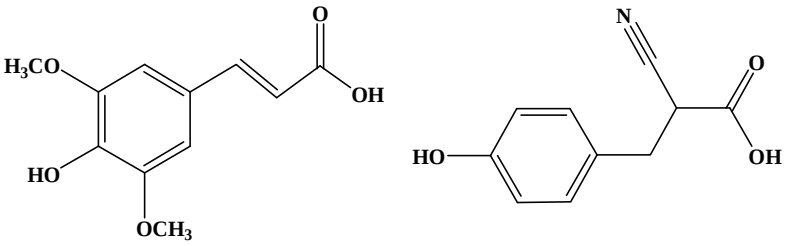
emerged as the culmination of a long series of experiments using desorption ionization (DI). MALDI is a soft ionization technique for the analysis of biomolecules and large organic molecules and has gained wide success in protein analysis, particularly when coupled with time-of-flight (TOF) instruments [34,35].

Another breakthrough occurred in 1984 when Fenn and coworkers used electrospray to ionize biomolecules [2]; the first ESI analyses of biopolymers including proteins were published in 1989 [3]. MALDI and ESI have revolutionized protein mass spectrometry since their invention in 1980s, and they have triggered the explosion in application of mass spectrometry for protein studies [36].

1.2 LASER-BASED IONIZATION METHODS FOR PROTEINS

1.2.1 Matrix-Assisted Laser Desorption/Ionization (MALDI)

Investigations of the wavelength influence in ultraviolet-laser desorption [33] led to invention of ultraviolet-laser matrix-assisted laser desorption ionization (UV-MALDI) between 1984 and 1986 and summarized in a 1987 paper [37]. In 1988 Karas and Hillenkamp reported ultraviolet-laser desorption (UVLD) of bioorganic compounds in the mass range above 10 kDa [1]. As a soft desorption ionization method, MALDI handles thermolabile, nonvolatile organic compounds, especially those with



sinapinic acid ***alpha*-cyano-4-hydroxycinnamic acid**

SCHEME 1.1 Structures of two common MALDI matrices.

high molecular weight and can be successfully used for the analysis of proteins, peptides, glycoproteins, oligosaccharides, and oligonucleotides. Its operation is relatively straightforward, although matrix preparation requires experience and perhaps some artistry.

MALDI is based on the bombardment of sample molecules with laser light, process that allows sample ionization [38]. It requires a specific matrix consisting of small organic compounds (e.g., nicotinic acid) that exhibit a strong resonance absorption at the laser wavelength used. The sample is premixed and diluted with the highly absorbing matrix and allowed to dry on a sample target. A range of compounds is suitable as matrices: sinapinic acid is a common one for protein analysis while *alpha*-cyano-4-hydroxycinnamic acid is often used for peptide analysis (the structures of matrices are shown in Scheme 1.1). This kind of acid serves well as a matrix for MALDI owing to the acid's ability to absorb laser radiation and also to donate protons (H^+) to the analyte of interest. Upon laser irradiation, energy is absorbed by the matrix in a localized region of the surface. As a result an explosive break up of the cocrystallized analyte/matrix sample occurs. The rapid expansion of the vaporized matrix in MALDI leads to the translational excitation of analyte molecules and the release of the analyte molecules from the surface of the condensed phase sample into vacuum. The analyte may be precharged (e.g., exist as a salt), and the intact analyte ion may simply be transferred as an ion from the solid to the vapor state upon laser irradiation of the matrix. Alternatively, a neutral analyte may be ionized through ion–molecule reactions (e.g., proton transfer reaction) occurring in the energized selvedge or interfacial region between the solid and gas phases.

MALDI has remarkable efficiency in producing intact molecular ions (often $[M + H]^+$, $[M + Na]^+$) of large biological compounds. MALDI ionization sensitivity is also extraordinary, and total amounts of sample loaded onto the target surface often are in the picomole to femtomole range. The method has tolerance to buffers and other additives and gives predominantly singly charged ions for large biomolecules [35]. TOF mass analyzers are ideal for use with this ionization technique because they are compatible with high-mass ions and pulsed-ion production [34,35]. TOF analyzers separate ions according to their *m/z* ratios by measuring the time it takes for ions, accelerated to the same kinetic energy, to travel through a field-free region known as the flight or drift tube. The heavier ions move slower than the lighter ones [6].

An important application of MALDI is chemical imaging, using a technique called matrix-assisted laser desorption/ionization imaging mass spectrometry (MALDI-IMS) [16]. Imaging combines parallel, high-throughput molecular analysis with location-specific information for the characterization of protein distributions directly from thin sections of intact biological tissue [39,40] and offers complementary information to two-dimensional (2D) gel electrophoresis and to shotgun proteomics for investigating proteomic differences. It is covered in Chapter 5 by Reyzer and Caprioli in this volume. Figure 1.2 illustrates the typical experimental process for MALDI imaging. Frozen tissue specimens are sectioned on a cryostat into about 5- to 20-mm thick sections. The sections are thaw-mounted onto conductive MALDI target plates. Matrix is applied to the sections, depending on the experiment to be performed: droplets (nL to pL) can be deposited in arrays or on discrete morphological areas, or a uniform coating of matrix can be applied to the entire tissue section.

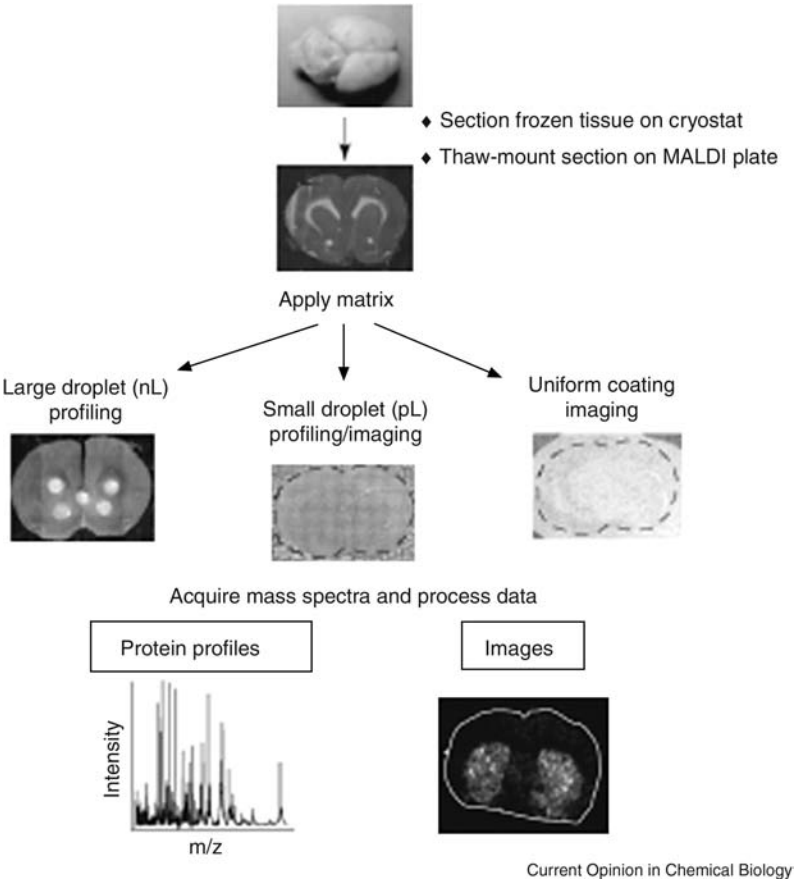


FIGURE 1.2 Typical experimental design for IMS. From [16]. Copyright permission was obtained from Elsevier. (*See the color version of this figure in Color Plates section.*)

Mass spectra are obtained from each spot or from across the entire tissue section in a defined raster pattern. The acquired spectra can then be examined and processed to form 2D molecule-specific ion images [16]. The power of MALDI-IMS technology is its capability to link reliably protein data with specific cellular regions within the tissue.

MALDI-IMS has been employed as an imaging technology in a wide variety of applications from the analysis of small molecules such as drugs and endogenous metabolites to high molecular weight proteins (e.g., MALDI-IMS of a mouse model of Parkinson's disease revealed a significant decrease in PEP-19 expression levels in the striatum after administration of the drug MPTP [41]).

1.2.2 Atmospheric Pressure Matrix-Assisted Laser Desorption/Ionization (AP-MALDI)

Atmospheric pressure matrix-assisted laser desorption/ionization (AP-MALDI) was first described by Laiko et al. [42]. In contrast to conventional vacuum MALDI, AP-MALDI can be operated at atmospheric pressure instead of high vacuum where ions are typically produced at 10 mTorr or less. During the ionization process, the solid-phase target material containing analyte sample and matrix is irradiated with a pulsed laser beam. The matrix absorbs the photon energy and undergoes fast heating and evaporation, which results in the formation of gaseous analyte ions [43]. Because the ionization of AP-MALDI occurs at atmospheric pressure, thermalization of the resulting ions takes place owing to collisions with the ambient gases used in AP-MALDI, accounting for the soft ionization nature of AP-MALDI [42].

The AP-MALDI source makes use of a high voltage potential that is applied between the target tip and the heated inlet transport capillary. The laser is focused onto the surface of the target plate. Ions are desorbed from the angled replaceable target tip and carried by the dry carrier nitrogen gas into a mass spectrometer [44]. The sensitivity of detection for AP-MALDI can be affected by the geometry of the target tip, and its position relative to the inlet orifice, the nitrogen gas flow rate, gas nozzle position, etc. [42]. Furthermore, when a Nd: YAG laser with high laser power rather than a nitrogen laser is used, the signal intensity can be improved [45].

Given that AP-MALDI and conventional vacuum MALDI share common ionization mechanisms, they have many similar features including simplicity of sample preparation and tolerance to interference from salts [42], which can be detrimental for biomolecule analysis. AP-MALDI is an extension of conventional MALDI, but it has some unique characteristics. First, samples are handled at atmospheric pressure. Second, AP-MALDI is a softer ionization technique than vacuum MALDI, which is favored for protein analysis. For example, the heavier peptides/glycopeptides from protein digestion are less likely to fragment by AP rather than vacuum MALDI; thus, more peptides can be detected (Figure 1.3) [42]. Third, AP-MALDI can employ liquid matrices to improve the ionization reproducibility, which otherwise results in source contamination in vacuum MALDI. Fourth, the AP-MALDI ion source is easy to exchange with other atmospheric ionization sources, allowing it to be easily coupled to different mass analyzers (e.g., quadrupole ion trap (QIT) [46], TOF [45],

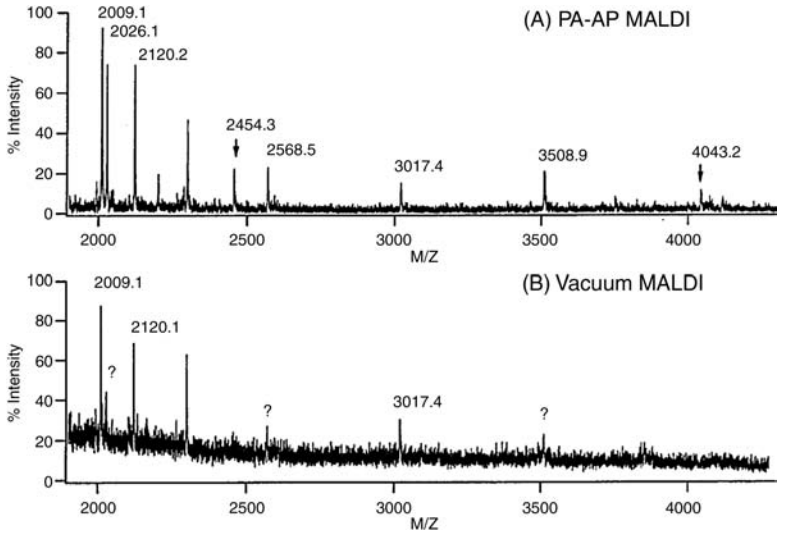


FIGURE 1.3 Partial mass spectra for tryptic digest of bovine fetuin. (*A*) AP-MALDI spectrum from 1.5-pmol deposition, showing eight identified peaks. (*B*) Vacuum MALDI spectrum for 1-pmol deposition, showing three identified peaks. From [42]. Copyright permission was obtained from ACS.

and Fourier transform ion cyclotron resonance (FT-ICR) [47]. Fifth, the analyte-matrix cluster ions in AP-MALDI caused by collisional cooling can be observed [43]. Nevertheless, the major disadvantage of AP-MALDI is the ion loss in atmospheric pressure interfaces, giving it lower sensitivity than conventional MALDI [46].

AP-MALDI MS has seen a variety of applications similar to those of conventional vacuum MALDI, including analysis in proteomics and determinations of oligosaccharides, DNA/RNA/PNA, lipids, bacteria, phosphopeptides, small molecules, and synthetic polymers [48]. The convenient and rapid exchange of the AP-MALDI source with other ionization sources and high throughput are attractive features. The major expected application for AP-MALDI is for the analysis of vacuum-incompatible samples like profiling of biological tissue samples, which requires the use of a wide range of liquid matrices at atmospheric pressure [43].

1.2.3 Surface-Enhanced Laser Desorption/Ionization (SELDI)

Surface-enhanced laser desorption/ionization (SELDI) as a prominent form of laser desorption/ionization (LDI) mass spectrometry was first described in 1993 by Hutchens and Yip [49]. It can be classified in three groups: surface-enhanced neat desorption (SEND), surface-enhanced affinity capture (SEAC), and surface-enhanced photolabile attachment and release (SEPAR) [50]. In SEND, analytes even for large molecules can be desorbed and ionized without adding matrix. This occurs because a compound with a chromophore to absorb laser energy is attached to the probe surface via physical adsorption or covalent modification [51]. In SEAC, the

probe surface plays an active role in the extraction, fractionation, cleanup, and/or amplification of the sample of interest. Common are chemical surfaces such as H50 (hydrophobic surface, similar to C_6–C_{12} reverse-phase chromatography materials), CM10 (weak-positive ion exchanger), Q10 (strong anion exchanger), IMAC30 (metal-binding surface), and biochemical surfaces containing antibodies, receptors, enzymes or DNA [50,52]. In SEPAR, an energy-absorbing molecule promotes analyte desorption and ionization, making this approach a hybrid of SECA and SEND [50]. Furthermore SELDI is commercially embodied in Ciphergen's ProteinChip® Array System (Ciphergen Biosystemes, Palo Alto, CA, USA), which simplifies the sample preparation with on-chip binding and detection [50]. SELDI is typically coupled with TOF mass spectrometers and is applied to detect proteins in tissue, urine, blood, and other clinical samples.

SELDI-TOF-MS is the extended form of MALDI-TOF-MS. The differences are the sample preparation and the software tools for interpreting the acquired data. In executing the SELDI process (Figure 1.4) [53], the first step is to select a chromatographic and preactivated ProteinChip array. Next, the protein sample solution is applied and incubated on the spots of the ProteinChip array. Third, by allowing the proteins to interact with the chromatographic array surface, on-spot contaminants and salts of the sample can be washed away to ensure efficient sample cleanup. This binding step to the SELDI surface can be viewed as a separation step, purifying the proteins bound to the surface. Fourth, matrices are added for the formation of a homogeneous layer of cocrystallized target proteins. After that, a laser beam is used to irradiate the spot, causing desorption and ionization of the proteins. The laser beam raster can be applied to cover selectively the entire spot surface, affording an output of the entire spot. Finally, multiple spectra are averaged to yield a final spectrum that displays the protein ions. Protein quantification is achieved by the correlation between the signal intensities and analyte concentrations of proteins in the sample.

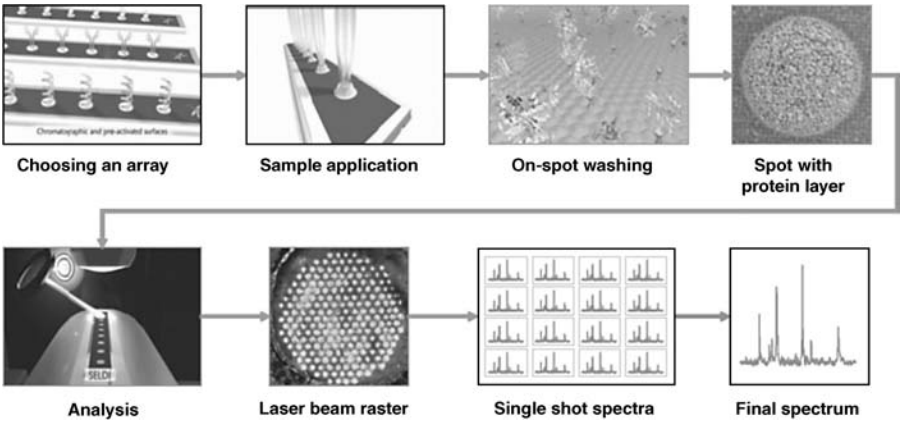


FIGURE 1.4 Experimental steps of SELDI-TOF-MS-based ProteinChip® System. From [53]. Copyright permission was obtained from Nature Publishing Group. (*See the color version of this figure in Color Plates section.*)